Targeted inhibition of the *Plasmodium falciparum* Vitamin B6 producing enzyme Pdx1 and the biochemical and functional consequences thereof

Shaun Bernard Reeksting

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor

in the Faculty of Natural and Agricultural Sciences

Department of Biochemistry

University of Pretoria

South Africa

June 2013

Submission declaration

I, Shaun Bernard Reeksting, declare that this thesis, which I hereby submit for the degree *Philosophiae Doctor* in the Department of Biochemistry, at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: ……

Date: ……

Plagiarism declaration

I understand what plagiarism is and am aware of the University's policy in this regard.

I declare that this thesis is my own original work. Where other people's work has been used (either from printed source, internet or any other source), this has been properly acknowledged and referenced in accordance with departmental regulations.

I have not used work previously produced by another student or any other person to hand in as my own.

I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Student's signature Date

Summary

Malaria is caused by the parasite *Plasmodium falciparum* and still plagues many parts of the world. To date, efforts to control the spread of the parasites have been largely ineffective. Due to development of resistance by the parasites to current therapeutics there is an urgent need for new classes of therapeutics. The vitamin B_6 biosynthetic pathway consists of a PLP synthase which produces pyridoxal 5'-phosphate (PLP) within the parasite. The absence of this pathway in humans makes it attractive for selective targeting using small chemical molecules. The PLP synthase condenses D-ribose 5-phosphate (R5P) and DL-glyceraldehyde 3-phosphate (G3P) with ammonia to form PLP. Two proteins make up this PLP synthase – *Pf*Pdx1 and *Pf*Pdx2. Computational modelling of *Pf*Pdx1, and mapping of the R5P-binding site pharmacophore facilitated the identification of several ligands with predicted favourable binding interactions. Confirmatory testing of these on the purified *Pf*Pdx1 *in vitro* revealed D-erythrose 4-phosphate (E4P) and an analogue 4-phospho-D-erythronhydrazide (4PEHz) were capable of dose-dependently inhibiting the enzyme. The acyclic tetrose scaffold of E4P, with both aldehyde and phosphate group moieties, was thought to affect R5P imine bond formation in *Pf*Pdx1, possibly allowing the molecule to enter the R5P-binding site of *Pf*Pdx1. This hypothesis was supported by molecular docking simulations, and suggested that 4PEHz could similarly enter the R5P-binding site. 4PEHz was detrimental to the proliferation of cultured *P. falciparum* intraerythrocytic parasites and had an inhibitory concentration (IC_{50}) of 10 µM. The selectivity of 4PEHz in targeting *Pf*Pdx1 was investigated using transgenic cell lines over-expressing *Pf*Pdx1 and *Pf*Pdx2, revealing that complementation of PLP biosynthesis rescued the parasites from the detrimental effects of 4PEHz. Functional transcriptomic and proteomic characterisation of 4PEHz-treated parasites revealed that the expression of *Pf*Pdx2 increased during 4PEHz treatment, moreover showed that other PLP-related processes were affected. These results supported that *Pf*Pdx1 is targeted by 4PEHz, and affected PLP biosynthesis *de novo*. Results from this study allude to alternative regulation of *de novo* PLP biosynthesis within the parasites by E4P. Moreover, contributions from this work showed that the *de novo* vitamin B_6 pathway of *P. falciparum* is chemically targetable, and a potential strategy for the development of newer antimalarials.

Abbreviations

4PEHz	4-phospho-D-erythronhydrazide
ACT	artemisinin combinational therapeutics
ALAS	delta-aminolevulinic acid synthetase (ALAS)
ANCOVA	analysis of covariance
AspAT	aspartate aminotransferase or aspartate transaminase
ATP	adenosine triphosphate
BNI	Bernhard-Nocht Institute for Tropical Medicine
*Bs*Pdx1	*Bacillus subtilis* Pdx1
CDC	Centers for Disease Control
CPS	carbamoyl phosphate synthetases
CQ	chloroquine
CS	circumsporozoite
Cyc	cyclophillin
DDT	dichlorodiphenyltrichloroethane
DHA	dihydroartemisinin
DHAP	dihydroxyacetone phosphate
DHFR-TS	dihydrofolate reductase-thymidylate synthase
DHPS	dihydropteroate synthase
DOPE	discrete optimized protein energy
DS	Discovery Studio
DXP	deoxy-D-xylulose 5-phosphate
E4P	D-erythrose 4-phosphate
FC	fold change
F6P	D-fructose 6-phosphate
G3P	DL-glyceraldehyde 3-phosphate
GMAP	Global Malaria Action Plan
GO	Gene ontology
GSEA	Gene set enrichment analysis
Hb	haemoglobin
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HDAC	histone deactelyase
HPI	hours post invasion

HTS	high-throughput screening
IDC	intra-erythrocytic development cycle
IRS	indoor residual spraying
ITN	insecticide treated bed nets
KEGG	Kyoto encyclopaedia of genes and genomes
LDH	Lactate dehydrogenase
LLIN	long lasting impregnated nets
LPM	ligand pharmacophore mapping
MMV	The Medicines for Malaria Venture
MSA	multiple protein sequence alignment
MPD	(4*S*)-2-methyl-2,4-pentanediol
OAT	L-ornithine aminotransferase
ODC	L-ornithine decarboxylase
PABA	*p*-aminobenzoic acid
PAGE	polyacrylamide gel electrophoresis
*Pb*Pdx1	*Plasmodium berghei* Pdx1
*Pf*Pdx1	*Plasmodium falciparum* Pdx1
PfPdx2	*Plasmodium falciparum* Pdx2
PCR	polymerase chain reactions
PDF	probability density function
Pdx1	SNO glutamine amidotransferase family protein
Pdx2	SNO glutamine amidotransferase family protein
PdxK	pyridoxine kinases
PdxP	pyridoxal phosphatases
PfCRT	chloroquine resistance transporter
*Pf*MDR	multidrug resistance transporter 1
PL	pyridoxal
PLP	pyridoxal 5'-phosphate
PM	pyridoxamine
PN	pyridoxine
PNPase	4-nitrophenyl phosphatase
PMSF	phenylmethylsulphonyl fluoride
PPP	pentose phosphate pathway
PQ	primaquine

PS	phosphatidylserine
PSDC	phosphatidylserine decarboxylase
RBM	Roll Back Malaria Partnership
RMSD	root mean square deviation
RQI	RNA quality indicator
rRNA	ribosomal RNAs
R5P	D-ribose 5-phosphate
Ru5P	D-ribulose-5-phosphate
SAR	structure activity relationship
SCD	sickle cell disease
*Sc*Pdx1	*Sacchromyces cerevisiae* Pdx1
SDS	Sodium dodecyl sulphate
SHMT	serine hydroxymethyltransferase
SP	sulpha doxine-pyrimethamin
SPT	serine C-palmitoyltransferase
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
StSyn	seryl-tRNA synthetase
T	Treated
THF	tetrahydrofolate
TIM	triosephosphate isomerases
*Tm*Pdx2	*Thermotoga maritima* Pdx1
TPP	thiamine pyrophosphate
*Tt*Pdx1	*Thermus thermophilus* Pdx1
UNICEF	United Nations Children's Fund
UT	Untreated
WHO	World Health Organization
WT	wildtype

Table of Contents

List of Figures

List of Tables

Acknowledgements

My wife, Bianca Reeksting, for being so supportive, encouraging and understanding. Thank you for sacrificing so much, and enduring through the tough times.

My parents, Pearl and Bernard Reeksting, for their encouragement and support throughout my degree.

My supervisors, Prof. Carsten Wrenger and Prof. Lyn-Marie Birkholtz, as well as my co-supervisor Prof. Abraham Louw. You have been an inspiration and I am thankful for the guidance, patience and support you have provided to me.

Dr. Pieter B. Burger for help with computational approaches and drug libraries.

Dr. J. Knöckel for providing the *Pf*Pdx1 and *Pf*Pdx2 constructs. As well as the *Pf*Pdx1/*Pf*Pdx2 overexpressing cell lines. Without your expertise this work would not have been possible.

Dr. Ingrid B. Müller at the Bernhard-Nocht Institute for Tropical Medicine (BNI) for providing the *Pf*Pdx1 R164A mutant. Your insights into experimental design were really valuable to me, and I learnt a lot.

Mrs. Bärbel Bergmann for excellent experimental and technical assistance in the laboratory at the BNI.

The Biochemical Parasitology group at BNI for hosting and accommodating me in the laboratory, and always being keen for discussions over coffee.

Prof. Laurent Salmon for kindly providing R5P analogues.

Mrs. Janette Reader for her help in microarray experiments and data processing.

The DAAD for providing funding and giving me the opportunity to study in Germany.

The University of Pretoria for a postgraduate bursary.

Chapter 1
Literature background

The burden of malaria around the world, and the associated mortality rate is shocking. This parasitic infection results in more than 200 million clinical episodes of malaria which, in turn, results in more than 600 000 deaths per annum. Due to poor living conditions and the high socio-economic burden, sub-Saharan Africa is the most affected by malaria with > 90% of all malaria-related deaths occurring on this continent. Current statistics show that a child dies from malaria every 40 seconds, equating to an estimated 2000 children per day [1]. People living in malaria endemic areas have poor quality of life, and typically socio-economic development is poor in comparison to the rest of the world [1]. Renewed efforts have been focussing on more effective drug treatments, the development of a malaria vaccine, and alternative mosquito vector control measures with the aim of eradicating malaria. The current collection of antimalarial therapeutics is under threat of being rendered useless due to decreasing parasite susceptibility - also termed resistance development. There is still an urgent need for improved therapeutics, such as single-dose cure treatments that would improve patient compliance, as well as drugs that block transmission of the parasite [2, 3]. It has therefore become paramount to understand the parasite biology and biochemistry in order to improve prophylaxis, treatment, elimination and to stem parasite drug resistance development.

1.1 Epidemiology of malaria

1.1.1 Emergence and history

The word malaria is derived from Latin “malus” and “aria”, or Italian “mala” and “aria”, and means bad air as the disease was associated as coming from wetland or marsh conditions. Malaria is caused by an infection from parasites of the *Plasmodium* spp. and, according to earliest recorded history this parasite has plagued the world for more than 4000 years. In 2700 BC, Nei-Ching described the symptoms of malaria infections such as periodic or undulating fevers [4]. In the fourth century BC Hippocrates similarly described such symptoms and noticed that incidences occurred in wetland areas, the habitat of mosquitoes which transmit malaria [5]. In the early part of the first century AD, the Romans also recorded fevers amongst people living in wetland areas [5]. Bone tissue samples of an Egyptian mummy dating back to 1500 BC have been found to contain DNA from the malaria parasite [6]. These reported incidences suggest that the parasite has persisted for millennia. In 1880 Alphonse Laveran described *P. falciparum*

sexual blood stage forms of malaria termed gametocytes in a French soldier in Algeria [7]. During that time it was not known how these parasites were transmitted. In 1897, Ronald Ross discovered parasite oocysts in the gut of mosquitoes, establishing the developmental stages of the malaria parasite in the mosquito gut and salivary glands [8]. In 1948, Shortt and Garnham used sporozoites to infect monkeys to reveal exo-erythrocytic schizonts forms of *P. cynomolgi* and later *P. falciparum* parasites in the liver – the missing piece in the parasite's elusive life cycle [9, 10]. Evidence also suggests that malaria parasites from the genus *Plasmodium* are much older than 4 centuries and were believed to have changed hosts going from gorillas into early humans [11]. Estimations based on mutation rates of a cytochrome *b* gene have placed this host switching event to around 365 000 years ago [12]. These early ancestors of the human malaria parasites may still hold clues into pathogenesis and parasite–host interactions, and how to address parasite resistance development.

1.1.2 Transmission

Plasmodium sp. belongs to the phylum Apicomplexa, where the term 'apicomplexa' refers to the complex array of organelles located on the apical ends of these parasite invasive forms [13]. Five species of *Plasmodium* spp. can infect humans; *P. malariae*, *P. ovale*, *P. vivax*, *P. falciparum* and *P. knowlesi*. Almost 90% of deaths caused by *Plasmodium* sp. are as a result of *P. falciparum* infections [14]. *P. falciparum* infections have characteristic 48 h fever paroxysms, and cause severe and complicated malaria. The large number of fatalities and high recrudescence rate, or the re-emergence of parasites following unsuccessful clinical treatment and/or loss of protective immunity, makes *P. falciparum* malaria the most serious [15].

Infections by other *Plasmodium* species are becoming increasingly prevalent. *P. knowlesi* was previously thought to only infect macaque monkeys, however increased incidence of near-fatal human infections from Malaysia and Thailand highlighted the host switching capability of the parasite [16]. In part, poor diagnosis of *P. knowlesi* in humans was attributed to the inability of rapid diagnostic tests detecting this parasite, retarding diagnosis and complicating treatment [16]. Infections with *P. malariae* characteristically cause 72 h fever paroxysms [17]. Such *P. malariae* infections are relatively mild compared to *P. falciparum* infections, and rarely result in death [17-19]. P. vivax malaria has 48 h recurring paroxysms associated with fever, sweating and headaches, reflecting the asexual reproductive life cycle within the human hosts. *P. vivax* is not restricted to tropical climates and is more prevalent in South and Central America as well as in Asia [19-21]. Infections from this parasite account for more than 50% of infections outside of

Africa [19-21]. After successful treatment of *P. vivax* malaria a relapse can occur, which arises when dormant liver stage forms of the parasite or hypnozoites reactivate and reinvade [22, 23]. For this reason *P. vivax* is more resilient and more difficult to control [23]. There are a number of factors that determine the latency period of the hypnozoites, which can range between a few weeks up to 9 months, and these include geographic location, climate, and the initial levels of parasite inoculation [22, 23]. An increase in mortality rates due to *P. vivax* infections and mixed co-infections with *P. falciparum* has been observed [24]. Moreover, increased spread of drug resistant *P. vivax* parasites has highlighted the increasing burden of malaria caused by this parasite specie [24].

Malaria is transmitted by female *Anopheles* mosquitoes, although the genus contains more than 500 species only a select few can transmit *P. falciparum* parasites. These include *An. gambiae*, *An. funestus* and *An. arabiensis*, and are predominant in Africa [25]. These mosquitoes are habitually anthropophilic – they prefer human blood above other sources. These mosquitoes often seek shelter inside houses and have nocturnal feeding habits [25]. Mosquitoes that are infected with *P. falciparum* parasites have increased feeding intervals and consume greater volumes of blood during feeding [26]. The period of blood feeding followed by maturation of eggs in mosquito ovaries and then oviposition, known as the gonotrophic cycle, can be as short as 2 days in tropical climates [20]. The short gonotrophic cycle and increasing feeding behaviour in tropical climates determine the rate of transmission of malaria. Not surprisingly, events of feeding are also determined by the "attractiveness" of the host; *An. gambiae* mosquitoes have olfactory receptors that respond to 1-octen-3-ol, excreted from human skin, and are also capable of detecting 2,3-butanedione, a by-product produced by skin bacteria [27]. Some individuals may therefore be more likely to get bitten by mosquitoes. Skin microorganism diversity, and the ability of the microbiota to produce volatile odorants, is also believed to play a role in host attractiveness, where individuals with lower diversity are more readily targeted by the mosquitoes [28].

1.1.3 Life cycle of *P. falciparum* parasites

As outlined by Figure 1.1, *P. falciparum* parasites are introduced into the human host whilst an infected mosquito takes a blood meal. Specialised invasive forms termed sporozoites migrate towards the liver, where a single sporozoite is all that is required to establish an infection within the host [29, 30]. After successful invasion of hepatic tissues, sporozoites undergo exoerythrocytic schizogony and multiply to form merozoites [29].

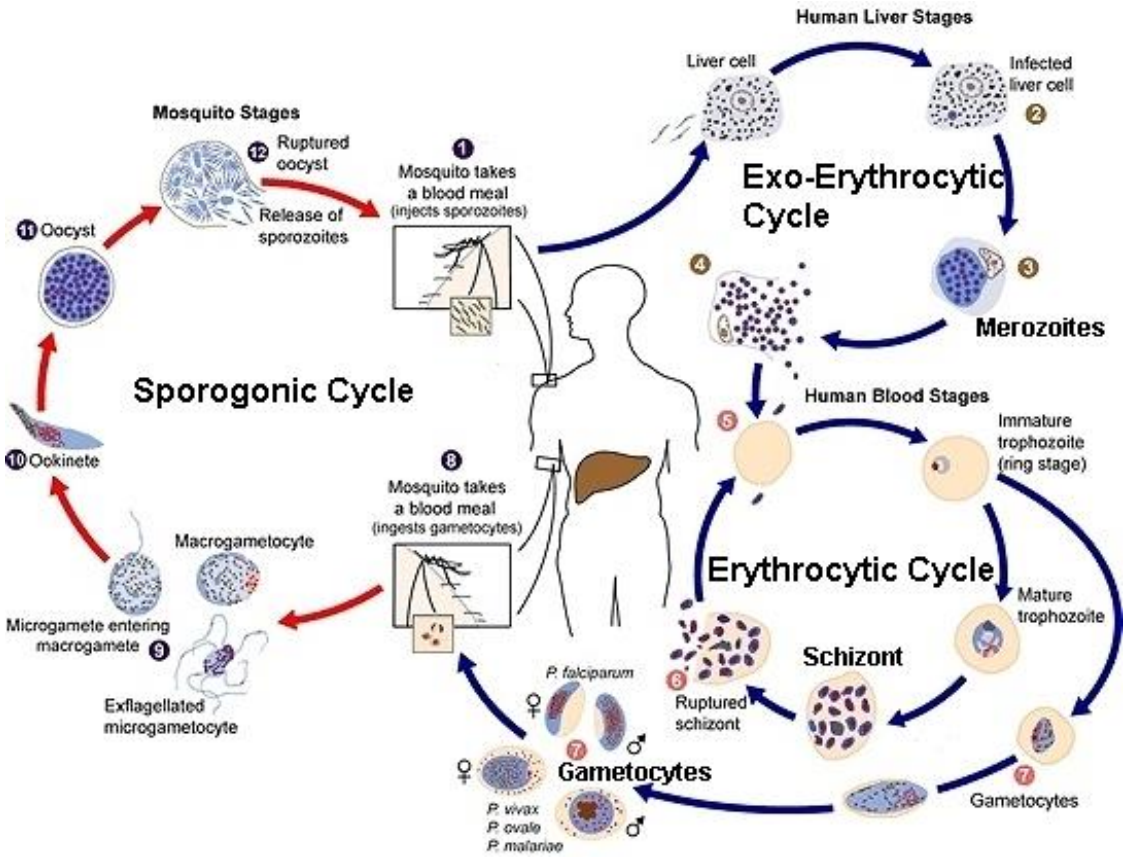


Figure 1.1: The life cycle of *P. falciparum*. Sporozoites, from an infected mosquito, invade the human skin (1), and migrate toward the liver to infect hepatic tissues (2). These parasites mature in the liver (3), and rupture to release merozoites (4). Merozoites invade erythrocytes (5) and mature to become schizonts. Mature schizonts rupture to release merozoites, which reinvade erythrocytes (5). Sexually-destined schizonts become gametocytes (7), which can be consumed by a mosquito during a blood-meal (8). Gametocytes mature in the midgut of the mosquito, and eventually rupture to release sporozoites, which invade the salivary glands of the infected mosquito. Figure adapted from the Centers for Disease Control (CDC) [31].

These merozoites invade human erythrocytes thereby initiating asexual erythrocytic schizogony, also termed the intra-erythrocytic development cycle (IDC). Merozoites are pear-shaped invasive forms of the parasite, which contain apically located organelles that facilitate attachment and invasion of the erythrocyte [29, 32]. Erythrocyte entry involves membrane protein interactions between parasites and erythrocytes, followed by invasion and the formation of the protective parasitophorous vacuole [29]. The IDC is composed of ring, trophozoite and schizont forms, in which the parasites survive by proteolytically digesting one of the most abundant host proteins within the erythrocyte – haemoglobin (Hb) [33]. The schizont stage marks the development of several daughter merozoites that explosively egress from the erythrocyte to invade new erythrocytes. Sexual development or gametocytogenesis involves the invasion and maturation of sexually committed merozoites [34]. This results in the formation of gametocytes that are accessible to mosquitoes during feeding, completing the cycle of disease transmission. Gametocytes are then ingested by the female mosquitoes during a blood meal taken from a *P.*

falciparum infected individual, thereby initiating the sporogonic stages of the parasite's life cycle. After recombination of gametocytes (male and female micro- and macro-gametes) within the mosquito's midgut lumen, the parasite matures into ookinetes, which migrate into the midgut epithelium and develops into oocysts [35]. These maturing oocysts, which start to bulge from the midgut epithelium into the mosquito haemocoelic cavity, eventually ruptures giving rise to millions of sporozoites that go on to invade the salivary glands in the mosquito and are available to reinitiate the invasion cycle [35].

1.1.4 Pathogenesis and clinical features

During the IDC, malaria disease symptoms become visible as characterised by periodic paroxysms, anaemia leading to hypoxia, hypoglycaemia, hyperlactaemia or metabolic acidosis, impaired consciousness and respiratory distress [36]. Septicaemia, other bacterial infections and febrile diseases can make accurate diagnosis of malaria difficult. The onset of severe malaria is characterised as severe anaemia and obstruction of brain microvasculature by parasites, and eventually results in death [37]. Symptoms of severe malaria are compounding, for instance the reduced oxygen carrying capacity (hypoxia) due to anaemia that leads to anaerobic metabolism and the accumulation of lactic acid, metabolic acidosis or hyperlactaemia [37]. Acidosis stimulates respiratory distress in which the human body tries to rid itself of excess CO_2 through rapid breathing to increase blood pH [36]. Cerebral malaria is caused by modification of the erythrocyte membrane by adhesion proteins including *P. falciparum* erythrocyte membrane protein-1 (*Pf*EMP1) and enables parasites to adhere to epithelial and endothelial vascular tissues, eventually leading to microvascular obstructions, impaired consciousness, coma and death [33].

1.1.5 Distribution

Globally, more than 36% of the world's population live in areas of malaria areas and in 2011 the World Health Organization (WHO) reported that malaria transmission occurs in more than 100 countries worldwide [38, 39]. Malaria is endemic to tropical and sub-Saharan regions of the world and in 2010 it was estimated that 91% of all malaria infections occurred in Africa [39]. Geographic elevation as well as annual rainfall patterns affects mosquito breeding habitats, which in turn determines the intensity of malaria transmission. There are therefore large differences in the risk of malaria infection throughout Africa [38]. Interestingly, red blood cell protein polymorphisms confer a degree of resistance against malaria infections and correlate to

the distribution of the parasite. The most striking of these genetically maintained resistance markers are haemoglobinopathies such as thalassemia, HbC, HbE, and sickle cell disease (SCD) or HbS [40]. The mechanism of SCD resistance to malaria involves diminished parasite invasion and poor parasite growth within the erythrocyte, with increased phagocytosis of sickle cell infected erythrocytes [40]. Although SCD protects from severe malaria, this polymorphism results in anaemia, which reduces oxygen delivery to tissues leading to fatigue and increased risk to bacterial infections.

1.1.6 Incidence rates (frequency)

The WHO has reported that during 2010 there were more than 216 million cases of malaria, with around 655 000 deaths of which 91% occurred in Africa [39]. However, malaria mortality rates are believed to be underestimated, and instead Murray *et al.* suggested that there were 1.2 million recorded fatalities in 2010 [41]. It is estimated that 714 000 (58%) of these deaths were children under the age of 5 [41], which is in contrast to the WHO reports that estimate around 86% of all malaria deaths to be children aged 5 and below [39]. Some estimations on the cumulative probability of dying from malaria have been proposed by Murray *et al.* [41]. The mortality risk is the highest in central as well as western sub-Saharan Africa, as outlined in endemicity maps of *P. falciparum* in 2010 (Figure 1.2) [41]. Dramatic reductions in malaria-related deaths have been noted in Zambia, Kenya, Tanzania and Ethiopia. However these regions are still associated with high malaria mortality rates [41].

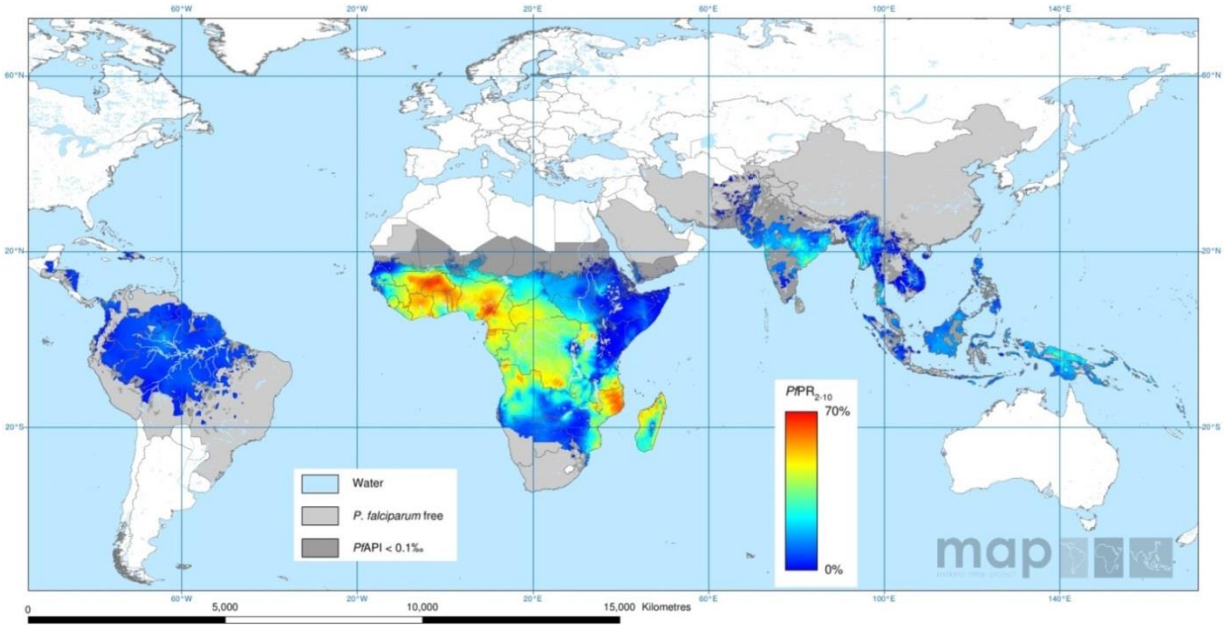


Figure 1.2: The spatial distribution of *P. falciparum* endemicity in 2010. The burden of malaria is shown as an intensity value of age-standardised *P. falciparum* parasite rate ($PfPR_{2-10}$). This value describes the estimated proportion of children between the ages of 2 and 10 that were infected with *P. falciparum* malaria during 2012. The figure was adapted as obtained from the Malaria Altas Project free resource database (http://www.map.ox.ac.uk). Maps were generated as described by Gething *et al.* [42].

1.2 Control and treatment strategies

Several control measures have been implemented with the goal of reducing and eliminating malaria transmission [3]. These include indoor residual spraying (IRS) of insecticides, the distribution of insecticide treated bed nets (ITN) and the use of antimalarial chemotherapeutics. Other means of control involve the development of a malaria vaccine. Two types of vaccine are feasible for malaria control and eradication; vaccines that block initial sporozoite invasion and transmission blocking vaccines which target sexual gametocyte forms. Global frameworks such as the Roll Back Malaria Partnership (RBM), launched in 1998 by the WHO and the United Nations Children's Fund (UNICEF), amongst others, aim to reduce malaria mortality by implementing preventative measures together with effective mosquito control programs. The Bill and Melinda Gates foundation Global Health Program also focus on developing malaria vaccines and more effective therapeutics and insecticides to reduce the spread of the parasites. The RBM initiated The Global Malaria Action Plan (GMAP, http://www.rbm.who.int/gmap/index.html) which aims to reduce the number of deaths caused by malaria to 0 in 2015. GMAP outlines three main goals for defeating malaria; control malaria to reduce the burden, eliminate malaria systematically, and develop tools to support global control and elimination. These ambitious goals are challenged by several factors including management of parasite resurgence and transmission from endemic areas. Moreover, the cost of elimination and eradication exceeds that

of treatment, and implementation of eradication schemes needs to be tailored for resource-limited areas.

1.2.1 Vaccines

Several stages of the parasite life cycle can be targeted using vaccines. Vaccines against the pre-erythrocytic sporozoites parasites are the most attractive. This is due to the fact that a small number of sporozoites (5-50) invade the human, compared to the large number present during the IDC, and therefore does not overwhelm immune responses. Humans do not exhibit any clinical symptoms of infection during sporozoite invasion and are not immune compromised, increasing the likelihood of parasite elimination during this stage of infection. Maturation of sporozoites typically takes around 6 days, allowing time for facilitated elimination of the parasites from the host [43-45]. Immunization with attenuated *P. falciparum* sporozoites has been effective in conferring protection, however, immunity is short lived (36-42 weeks), requires repetitive immunizations and is species specific [46, 47].

With the aim of developing a sporozoite vaccine the RTS,S/AS vaccine was designed to target circumsporozoite (CS) protein, a dominant membrane protein of the sporozoite. The antigen candidate consists of C-terminal conserved repeat regions in CS protein (termed RT) linked to a hepatitis B surface protein (termed S), which assemble into virus-like particles once expressed [47]. In Phase II clinical trials the vaccine had between 30% - 50% efficacy, protecting against malaria infections and the onset of severe malaria for up to 18 months. This is the most clinically advanced malaria vaccine to date [47, 48]. Phase III clinical trials are underway [49]. Preliminary results from Phase III trials are disappointing and suggested that the vaccine has lowered efficacy than observed previously [50, 51]. The administration of the vaccine to infants between the ages of 6 and 12 weeks conferred 31% reduced risk of developing clinical malaria, and 26% reduced risk of developing severe malaria [51]. The reduced efficacy in infants suggested that early vaccination, which is logistically easier, may not be as feasible as hoped for [50].

Vaccines designed against parasite stages of the IDC could, apart from preventing and combating initial emergence, reduce the symptoms of these proliferative stages [3]. Efforts to design such vaccines are hampered due to highly effective parasite immune evasion mechanisms including the antigenic variation of the parasite surface proteins [33, 52]. Additionally transmission-blocking vaccines are also needed to target gametocyte forms of the parasite [53].

This, in combination with antimalarial therapeutics, could prevent transmission of parasites [53]. This would aid in ensuring effective epidemic management and subsequently aid in elimination.

1.2.2 Vector control

ITN offer both physical and chemical protection, and together with IRS has dramatically reduced malaria incidence in areas of high transmission [14, 54-56]. ITN's have largely been replaced with long lasting impregnated nets (LLIN) which contain an embedded insecticide, rather than residually applied insecticide as found on ITN's [57]. Commonly used insecticides include pyrethroids such as permethrin and deltamethrin that are incorporated into the polymer before extrusion, or is applied as a resin coating over the polymeric fibers of the nets [57].

The insecticide bis[4-chlorophenyl]-1,1,1-trichloroethane, also known as dichlorodiphenyltrichloroethane (DDT) was introduced and extensively used during the 1940's as it was inexpensive ($3/kg in 1990) and extremely effective [58-60]. DDT was classified as a persistent organic pollutants (POP) by the Stockholm convention in May 2004 [61]. It is carcinogenic in animals, resulting in liver tumors, however, there is no substantial evidence that DDT causes cancer in humans [62]. DDT is still used in Africa and was reintroduced in South Africa due to increased malaria incidence rates and the rapid appearance of pyrethroid-resistant *An. funestus* mosquitoes [63]. The safer alternative to DDT includes various pyrethroids that are also the active ingredients in ITN. Pyrethroids, as well as DDT, target the sodium channels of the mosquito's neurons, resulting in paralysis leading to death [64]. Mosquito resistance against pyrethroids arises through either polymorphisms in the sodium channels, or through altered cytochrome P450-mediated metabolism and detoxification of these compounds [54, 64]. The increased spread of DDT-resistant mosquitoes, as well as the widespread inefficacy of pyrethroids is alarming, establishing an urgent need for better control measures to stem the spread of malaria [54, 65].

1.2.3 Therapeutics

Historically the administration of chemotherapeutics has been the most successful in treating and preventing malaria [46, 66]. Various effective antimalarials are listed in Table 1.1 and can be divided into several groups; 4-aminoquinolines, sesquiterpene endoperoxides, antifolates, and 8-aminoquinolines. These therapeutics are extensively discussed elsewhere, and here only selected

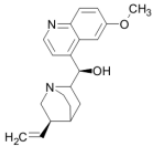
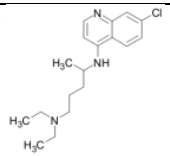
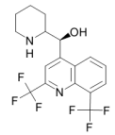
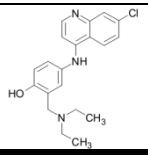
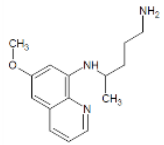
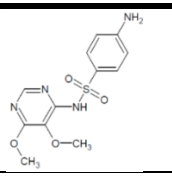
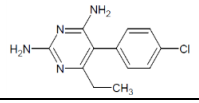
aspects concerning efficacy and drug resistance are mentioned [66, 67]. Antimalarials that have been widely used between the 1950's and 1980's belong to the 4-aminoquinoline group (Table 1.1) including quinine, chloroquine (CQ), mefloquine and amodiaquine all of which contain the quinoline chemical scaffold. CQ is considered as one of the most successful antimalarial drugs due to the fact that it was safe and relatively inexpensive [68]. Broadly, the 4-aminoquinolines are capable of interfering with the parasite's leading to the accumulation of haeme [69, 70]. More specifically, free-haeme from Hb digestion is polymerised into haemozoin, which is deposited in the acidic food vacuoles of the parasites in crystalline form. CQ is capable of binding to free haeme, prevent haemozoin formation and leading to the accumulation of toxic complexes and eventual parasite death [68].

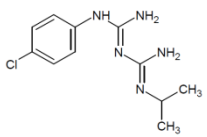
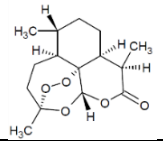
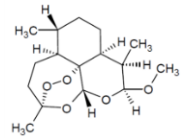
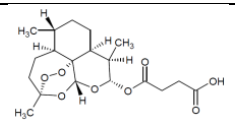
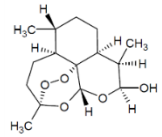
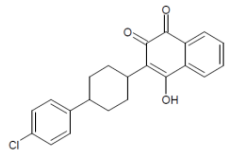
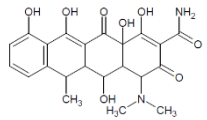
Widespread parasite resistance against the 4-aminoquinolines, especially against CQ, has limited their effectiveness in almost 90% of all endemic areas and led to discontinued use [70]. In the 1980's no affordable alternative was available and CQ resistance led to a 3-fold increase in malaria mortality [67]. Single nucleotide polymorphisms (SNP's) in the genes encoding *Pf*CRT (chloroquine resistance transporter) and *Pf*MDR (Multidrug resistance transporter 1) are linked to reduced parasite susceptibility to 4-aminoquinolines [69]. Both *Pf*CRT and *Pf*MDR are components of the parasite vacuolar membrane and reduce the accumulation of 4-aminoquinolines in the vacuole [70]. Related compounds to the 4-aminoquinolines include the 8-aminoquinoline, primaquine (PQ), is also an effective gametocidal capable of reducing mature gametocyte development and subsequently transmission of the parasites [71-73]. PQ is also effective against exoerythrocytic parasites [71, 72]. Other 8-aminoquinolines such as aablaquine (also known as bulaquine or elubaquine) and tafenoquine are currently in clinical trials [72, 74, 75].

Transmission-blocking therapeutics are essential to reduce the spread of resistant parasites and malaria by targeting gametocyte transmission [74, 76]. However, semi-immune populations carrying parasites are often asymptomatic, do not receive treatment, resulting in the continual transmission of parasites - and potentially resistant forms - for extended periods of time [74]. Several strategies aimed at targeting these asymptomatic carriers have been proposed, including mass drug administration, however these are not readily implemented [74]. Chemotherapeutics form part of the toolkit for eliminating malaria, but for this reason need to be used in combination with other control measures [76].Chemotherapeutics form part of the toolkit for eliminating malaria, but for this reason need to be used in combination with other control measures [76].

Antifolates are blood schizontocides that selectively target asexual forms of the parasites [77]. These drugs have been very successful antimalarials and interfere with the formation of folates, leading to decreased pyrimidine synthesis as DNA nucleobase precursors and hence affecting DNA formation [77, 78]. Progaunil, a triazine chloroguanide, was discovered in the 1940s and used to prevent and treat malaria [79]. The active metabolite of proguanil is cycloguanil which targets the dihydrofolate reductase (*Pf*DHFR) enzyme of the parasite [79]. Some antifolates such as sulphadoxine, which is a sulphonamide, inhibit dihydropteroate synthase (*Pf*DHPS) by mimicking *p*-aminobenzoic acid (PABA), whereas other antifolates such as pyrimethamine also target *Pf*DHFR (Table 1.1) [78]. Extensive amino acid mutations in the respective target proteins have led to resistance development and use of these drugs have become restricted. The introduction of the sulphadoxine-pyrimethamine (SP) drug combination in 1967 was rapidly followed by parasite resistance and decreased clinical effectiveness in the same year, despite this SP was utilised to effectively treat CQ-resistant parasites until the early 1990s [77, 80]. SP is now only used to treat malaria during pregnancy [79]. Multiple factors determine the rate and spread of resistance; SP has a longer half-life than other antifolate combinations (such as LapDap – chloroproguanil with dapsone) which was believed to have contributed to rapid resistance development [80]. Additionally, poor drug compliance leads to sub-optimal therapeutic dosing and can result in resistance development [80]. Newer antifolates are being developed to overcome problems associated with resistance, one such compound - P218 – which inhibits both wildtype (WT) and mutated DHFR has shown promise as a potential therapeutic candidate [81]. Antifolates remain good drugs against *P. falciparum* and alternative strategies include using different antifolate combinations with other effective antimalarials. Malarone®, a fixed-dose combination of the antifolate proguanil with atovaquone, is used to prevent and treat multi-drug resistance *P. falciparum*, and has relatively mild side effects [82]. Reports of treatment failure in a patient treated with Malarone® have however also raised concerns whether resistant parasites might emerge [82].

Table 1.1: Most widely used antimalarial therapeutics and resistance markers.

Name	Structure	Trade name	Target	Benefits	Disadvantages	Resistance marker (SNP's)	Ref.
4-aminoquinolines							
Quinine		Quininmax Aflukin	Haem detoxification mechanisms	Used for treatment of severe malaria. Can be used during pregnancy.	Short half-life (8h). Cannot be used as prophylactic.	*Pf*CRT (chloroquine resistance transporter) *Pf*MDR (Multidrug resistance transporter 1) *Pf*NHE1 (sodium proton exchanger 1)	[13, 67, 83-85]
Chloroquine		Resochin, Dawaquin, Daramal	haemozoin polymerisation Heme detoxification	Cheap, long half-life (60h). Suitable for children and pregnant woman.	Wide-spread resistance. Overdose risk.	*Pf*CRT, *Pf*MDR	
Mefloquine		Larium, Mephaquine	Haem detoxification or *Pf*MDR	Long half-life (14d) requiring weekly administration, ensures patient compliance.	Neuro-psychiatric side effects, not recommended during pregnancy. Vomiting in children.	*Pf*MDR	
Amodiaquine		Camoquin Flavoquine	Haem detoxification processes	-	Short half life	*Pf*CRT, *Pf*MDR	
8-aminoquinolines							
Primaquine (PQ)		Primaquine	*Pf*CRT transporter inhibitor	Gametocytocidal. Targets *P. vivax* hypnozoites	Not compatible with G6PD deficiency. Not recommended during pregnancy.	-	[13, 67, 71, 83]
Antifolates							
Sulphadoxine		Fansidar (combined with pyrimethamine)	Dihydropteroate synthase (*Pf*DHPS)	Inexpensive, long half-life (100 – 200h)	Extensive parasite resistance, severe allergic reactions.	SNP in *Pf*DHPS	[78]
Pyrimethamine		Daraprim	Dihydrofolate reductase (*Pf*DHFR)	Long half-life (80h)	Widespread resistance	SNP in *Pf*DHFR	[78, 84, 86]

Proguanil (cycloguanil *in vivo*)		Lapdap (combined with dapsone) Malarone® (combined with atovaquone)	*Pf*DHFR, may include additional targets.	-	Parasite resistance, short half-life	SNP in *Pf*DHFR	[67, 78, 86]
Sesquiterpene endoperoxides							
Artemisinin		Not used as monotherapy	Haemoglobin digestive processes (falcipain 4) or Mitochondrion electron transport processes.	Early ring and gametocidal activity	Short half-life.	No definitive resistance markers	[67, 87, 88]
Artemether		Coartem® (in combination with lumefantrine)		Effective against SP and chloroquine resistant parasites. Cheap.	Not recommend for pregnant woman.		
Artesunate		Larimal (in combination with amodiaquine)		More hydrophilic	Short half-life.		
Dihydro-artemisinin		Artenimol Eurartesim® (combination with piperaquine)		Active form	Short half-life.		
Other classes							
Atovaquone (hydroxynaptho quinone)		Malarone® (combination with proguanil) Mepron	Targets cytochrome bc_1 complex, arrests mitochondrial respiration	Well tolerated	Expensive	SNP in cytochrome b gene	[67, 86]
Doxycycline (tetracycline)		Vibramycin	Protein synthesis inhibitor, apicoplast targeting	Well tolerated, inexpensive.	Slow acting, used only as prophylactic. Cannot be used in children or during pregnancy	Increased copy numbers of *pfmdt* and *pftetQ*	[83, 89]

Compounds with a sesquiterpene endoperoxide functionality include artemisinin (Table 1.1) and its more hydrophilic derivatives artemether, artesunate and dihydroartemisinin (DHA) [88]. These have become the current mainstay antimalarial therapeutics. Artemisinin and derivatives are converted into to the clinically active DHA form [88]. Early studies showed that treatment of malaria-infected individuals with artesunate and artemether had a significant gametocidal effect, and are therefore capable of reducing transmission of malaria [90]. The compounds do not target mature gametocytes, rather the early gametocytes (stages I – III) as well as IDC-ring forms of the parasite [88, 90, 91]. Artemisinin inhibits chloroplast electron transport and is phytotoxic leading to impaired plant growth [92]. In the malaria parasite the endoperoxide ring of the compounds create reactive oxygen species (ROS), and are believed to affect mitochondrial electron transport functions [93]. Increased parasite recrudescence rates against artemisinin combinational therapeutics (ACT) was first reported in 2004, later in 2006, and most recently in 2009 [94]. A study in western Cambodia showed reduced clearance times in artesunate-treated patients compared to individuals in western Thailand [94]. Postulated polymorphisms in *Pf*ATPases or *Pf*MDR were not linked to ACT resistance development and there are still no definitive resistance markers for ACT-resistant parasites [87, 94].

The Medicines for Malaria Venture (MMV) is a non-profit organization jointly established by the Government of Switzerland, the WHO (RBM partnership), and private philanthropists in 1999. Through collaborating with industrial and academic partners the MMV aims to discover and develop newer and effective antimalarial drugs which would aid in reducing the burden of malaria (http://www.mmv.org). The MMV has developed antimalarials of which Eurartesim® (DHA and piperaquine combination) [95] and Coartem® *Dispersible* are successful commercially available products. The MMV is involved in lead generation and optimisation, and has several compounds in preclinical developmental stages. The MMV also boasts several compounds in Phase I, IIa, IIb and III clinical trials of which tafenoquine is one. A recent venture of the MMV in collaboration with, amongst others the University of Cape Town, has resulted in a compound which was identified from phenotypic whole cell high-throughput screens (HTS) [2]. The compound was highly effective against two drug resistant malaria strains and boasts a single dose cure capability against *Plasmodium berghei*-infected mice [2].

The discovery of new antimalarials is intensive and expensive. Typically from the original hit (molecule) to approved medicine can take up to 12 years and cost more than $1 billion [96, 97]. There are several phases of a typical drug discovery scheme program; generally a target is selected, followed by identification of hits (usually chemical antagonists or agonists – modulators of target function) [98]. The hits are generally confirmed using dose-response curves, and tested in

functional cellular assays to evaluate efficacy and membrane permeability. Hit-to-lead transition further establishes the hit cytotoxicity and biocompatibility. Lead optimisation aims to improve lead efficacy and pharmacokinetic properties. Large phenotypic or whole cell HTS are used to identify hits – its major drawback is high cost and time expenditure. Some strategies involve rigorous target validation, which is then followed by target-based HTS, which results in the identification of more specific molecules compared to whole cell HTS approaches. Attrition rates of target-based HTS hits during the hit-to-lead transition are expected to be less (considering that these are more specific), but these hits often have poorer permeability characteristics, requiring additional lead optimisation. An overview of these methodologies, as well as the benefits of computational screening, is given in Chapter 2.

Over the last 30 years only a few antimalarial therapeutics with unique chemical scaffolds have been identified or developed, and strategies have largely focussed on the modification of effective chemotypes [99]. From the collection of antimalarial therapeutics, it is estimated that there are only around 30 different chemical scaffolds, of which only around 10 are clinically useful [99]. More alarming is the fact that parasite resistance to almost all of these scaffolds have been reported, necessitating the identification and expansion of newer chemotypes. Different drug discovery approaches, such as targeting of validated proteins, optimisation of known effective chemotypes, and chemical screening, have resulted in the successful identification of antimalarials.

In drug discovery chemotypes can be expanded by modifications to known scaffolds, or by designing and identifying newer classes of compounds against cellular target [99]. Validated chemotypes can be considered as classes of compounds which have known efficacy, of which the mode of action (MOA) is not necessarily always known, and often of interest. However, conservation and optimisation of a chemotype does not render newer chemical scaffolds. Furthermore the occurrence of drug-resistance can render whole chemotypes useless. In the case of malaria this has occurred numerous times, as demonstrated for aminoquinolines [67]. This severely negates highly effective treatment and possible elimination of the parasites.

A question that arises in target-based drug discovery is whether a single protein is truly tractable or "druggable" *in vivo* and whether chemical attenuation does achieve the desired effect? For this reason target-based drug discovery essentially focuses on validated protein targets. A validated protein target can be considered as a target which has been successfully attenuated using chemotherapeutics and kills the parasite. Target validation includes demonstrating that chemical attenuation is achievable in an *in vitro* as well as *in vivo* setting [100]. Chemical knock-down as a tool for validation of the target is not always possible due to target-specific compounds having poor

bioavailability, low efficacy *in vivo*, high toxicity and off-target effects [99, 100]. Moreover, protein target validation using forward genetic approaches such as antisense- or RNAi-knockdown is based on the premise that the implementation of chemical compounds will have the same effect [100]. This is not always the case. The RNA product could have numerous biological functions apart from providing the ribosomes with a protein assembly sequence, and could obscure the interpretation of the knock-down [101]. Gene deletions or knock-outs and loss–of-function mutations are also used to validate proteins for their druggability. The increasing appreciation of the involvement of coding as well as non-coding regions of DNA in cellular homeostasis questions whether this approach is truly comparable to chemical attenuation of protein targets [101]. In the human genome it has become apparent that a large number of genes are interlaced and transcripts can be synthesized from both DNA strands [101]. Therefore these approaches can complement the validation of protein targets but should not be relied on for absolute validation of druggability.

From validated targets additional compounds can be generated based on chemical analogy or designed from structure-based mining approaches. However, currently there are only a few validated antimalarial drug targets. Moreover, the identification of newer chemical scaffolds to these targets depends on the finite availability of ligands that may have complementary fit with the target, and such approaches often do not yield more effective compounds than what is currently available [99, 102]. For this reasons there is an urgent necessity to exploit non-validated protein targets using chemotherapeutic intervention. This would aid in expanding the antimalarial armamentarium. Vitamin B_6 biosynthesis in the parasite, and the proteins involved in the pathway, are attractive candidates for drug design and identification.

1.3 The significance of vitamin B_6

Vitamin B_6 is one of the most versatile co-factors found in nature. The B_6 family is composed of several vitamers and their respective phosphorylated forms; pyridoxal (PL), pyridoxamine (PM) and pyridoxine (PN) as shown in Figure 1.3 [103-105]. The term vitamin B_6 is given to all of the vitamers [106]. Pyridoxal 5'-phosphate (PLP) is the most prominent vitamer acting as the active cofactor in PLP-dependent enzymes, and is also involved in non-enzymatic reactions [107-109]. PLP is involved in more than 140 different biochemical reactions ranging from transamination, deamination, racemisation and decarboxylation, and is most notably utilised as co-factor during amino acid metabolism [103, 110-112].

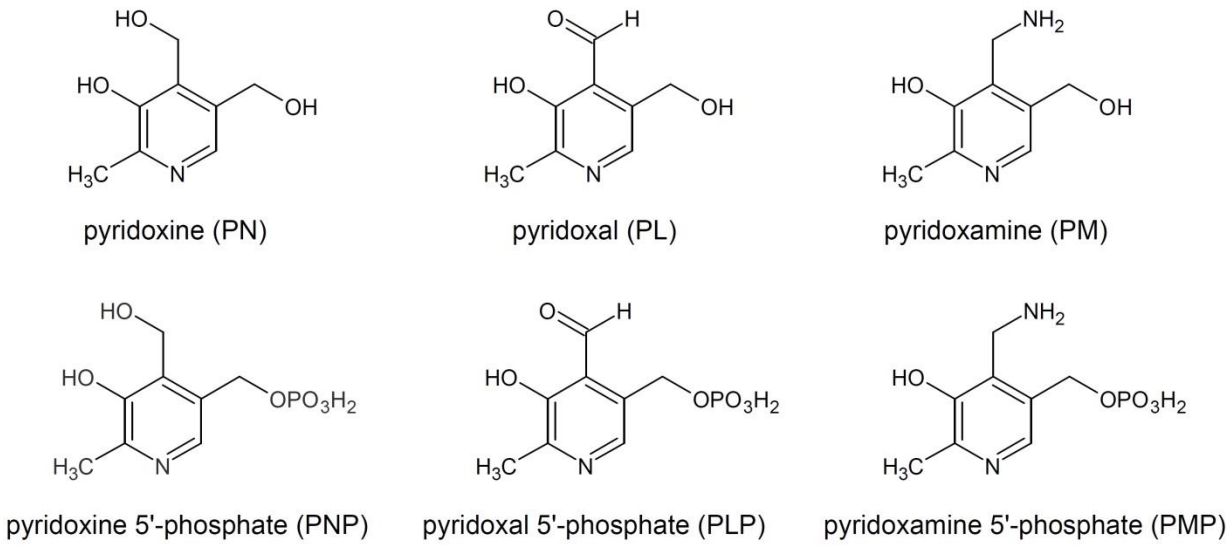


Figure 1.3: Different physiological forms of vitamin B_6. The basic pyridine scaffold of B_6 vitamers containing various functional chemical moieties attached to the 4' position of the pyridine ring [111]. Both pyridoxal 5'-phosphate and pyridoxal 5-phosphate are accepted nomenclature according to the International Union of Pure and Applied Chemistry (IUPAC) [113, 114].

The physiological significance of vitamin B_6 is further highlighted due to its anticancer properties. Recently, it has been demonstrated that a dietary supplement of vitamin B_6 effectively decreased colon tumorigenesis in mice, partly attributed to the antioxidative activity of the molecule [115]. Vitamin B_6, particularly PLP, can effectively scavenge ROS by quenching singlet molecular oxygen ($^1O_2$) [116, 117].

1.3.1 Vitamin B_6 metabolism

Vitamin B_6 biosynthesis is limited to most *Archaea*, bacteria, fungi and plants, whereas humans and mammals acquire this essential metabolite exclusively from their diet [106]. Two major routes of vitamin B_6 biosynthesis have been established, the deoxy-D-xylulose 5-phosphate (DXP) dependent and independent routes. The DXP-dependent pathway is found in *Escherichia coli* and the γ-division of proteobacteria. The pathway is composed of enzymes PdxA (E.C. 1.1.1.262) and PdxJ (E.C. 2.6.99.2), and produces PNP from DXP and 4-phosphohydroxy-L-threonine (Figure 1.4) [106]. The final PNP is oxidized to PLP by PdxH to complete the biosynthesis of the cofactor [106]. The DXP-independent PLP synthesis is more prevalent in bacteria (other than *E. coli*), archaea and eukaryotes [107]. The DXP-independent pathway consists of a PLP synthase which uses D-ribose 5-phosphate (R5P), DL-glyceraldehyde 3-phosphate (G3P), and L-glutamine as substrates (Figure 1.4) [106]. The PLP synthase is composed of two proteins Pdx1 and Pdx2. These proteins are present in organisms lacking proteins from the DXP-dependent pathway, and a clear divergence is evident

from the DXP-dependent route of PLP synthesis [116, 118]. Pdx1 and Pdx2 proteins are named YaaD and YaaE in *Bacillus subtilis*, SNZ and SNO in *Sacchromyces cerevisiae*, and PdxS and PdxT in *Geobacillus stearothermophilus*, respectively [107, 119]. Other names of Pdx1 proteins are PyroA in *Aspergillus nidulans* and SOR1 in the filamentous fungi *Cercospora nicotianae* [116, 120].

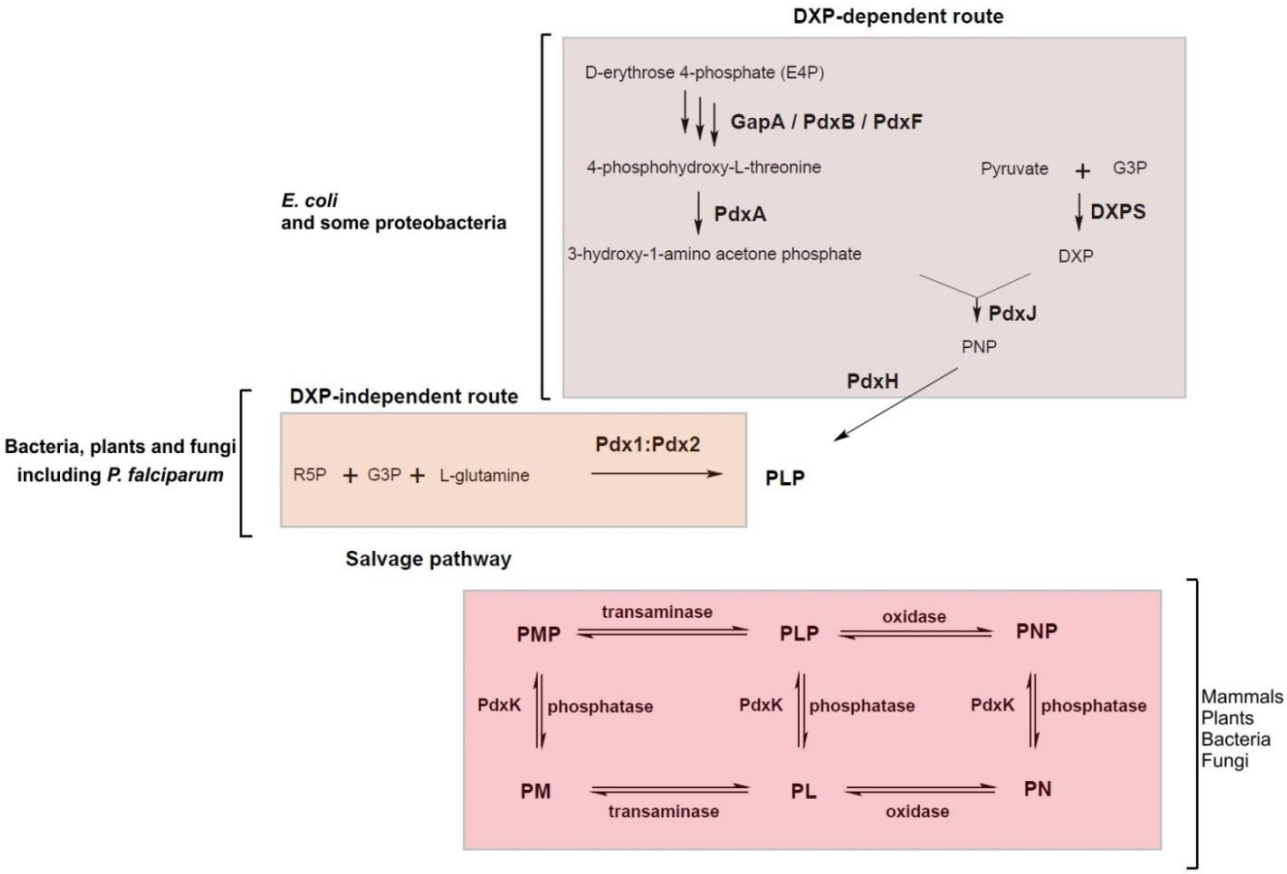


Figure 1.4: Summary of vitamin B_6 biosynthesis and salvage. There are two biosynthetic pathways for the formation of PLP. The DXP-dependent route utilises E4P, which undergoes enzymatic conversion,together with DXP to form PNP. This reaction is catalysed by PdxJ. DXP is derived from pyruvate and G3P by DXP synthase. The final PNP molecule is oxidised to PLP by PdxH. The DXP-independent route involves a single PLP synthase, composed of proteins Pdx1 and Pdx2. These proteins condense R5P, G3P and L-glutamine to form PLP. Salvage of the B_6 vitamers also constitute towards the PLP pool. Pyridoxine kinase (PdxK) are responsible for phosphorylation of PM, PL and PN. Phosphatases are involved in the dephosphorylated of PLP, PNP and PMP. For a comprehensive summary of biochemical entities in the pathways see Mukherjee *et al.* [121].

Mammals do not have the required biochemical machinery for *de novo* PLP synthesis, and instead salvage B_6 vitamers from their diet (Figure 1.4) [104]. Unphosphorylated vitamers are absorbed through diffusion from the intestine and pyridoxine kinases (PdxK, E.C. 2.71.35) are involved in phosphorylation to form PLP, PMP and PNP [106]. This effectively traps the different B_6 vitamers. Flavin mononucleotide (FMN)-dependent PNP oxidases are responsible for conversion of PNP and PMP to PLP (Figure 1.4) [122]. Pyridoxal phosphatases (PdxP, E.C. 3.1.3.74) can dephosphorylate

PLP, PNP and PMP in order to allow transport across cellular membranes [122]. Transaminases facilitate interconversion of PMP and PM to PLP and PL, respectively (Figure 1.4) whereas pyridoxine 5-phosphate oxidases (E.C. 1.4.3.5) are able to convert PNP and PN to PLP and PL respectively [106, 121, 123, 124].

1.3.2 Vitamin B_6 in the erythrocyte

Erythrocytes transport B_6 vitamers within the body. Dietary acquired PN, PL and PM enters the erythrocytes through diffusion, and are converted to their phosphorylated forms by PdxK [123-126]. Molecular trapping through phosphorylation, rather than active transport, drives the accumulation of B_6 within the erythrocytes (Figure 1.5) [127]. The phosphorylated vitamers PNP and PMP are converted to PLP by pyridoxine 5-phosphate oxidase (E.C. 1.4.3.5) and [123, 124] or via transamination reactions [121]. Within the erythrocyte, phosphorylated B_6 vitamers can be de-phosphorylated by PdxP whereas tissue non-specific alkaline phosphatases (E.C. 3.1.3.1) and liver acid phosphatases (E.C. 3.1.3.2) with broad substrate specificity performs this function in other tissues [123, 128]. In the erythrocyte a large proportion of PLP is bound to Hb and this both protects free PLP from hydrolysis and aids in establishing the concentration gradient to facilitate accumulation [127, 129]. PLP forms a Schiff base with a N-terminal valine of the β chains and reduces the oxygen affinity of Hb [130]. Hb constitutes 95% of the total erythrocyte protein content and is present at a intracellular concentration of 5mM [46]. The average erythrocyte PLP concentration is 410 pmol/g Hb, equivalent to 0.247 ng PLP/g Hb, and this low physiological concentration compared to Hb suggests that a very small proportion of Hb is PLP-bound (PLP binding to Hb is therefore not saturated at physiological conditions) [131]. Human erythrocytes possess the full complement enzymes required to interconvert B_6, contain greater oxidase activity compared to other tissues, and high levels of Hb, therefore are believed to be the main courier of B_6 within the body [128].

1.3.3 Vitamin B_6 metabolism in the malaria parasite

The malaria parasite occupies erythrocytes during asexual proliferation in the human, consuming host cell nutrients and proteolytically digests host proteins like Hb. This is complemented in part by *de novo* production of other essential metabolites, vitamins and amino acids. PLP in the parasite could be derived from two different routes – *de novo* biosynthesis through the action of PLP synthases, or through vitamin B_6 salvage pathways. In *P. falciparum* salvage of B_6 vitamers

minimally contributes toward the total vitamin B_6 pool, with recent isotopic tracer experimentation suggesting that *de novo* production through PLP synthase is the major route in which parasites obtain this essential cofactor (unpublished experiments, C. Wrenger). The necessary enzymes are present, as discussed below, however it has not yet definitively been established whether these promote salvage of B_6 vitamers or are involved in interconversion of different B_6 vitamers within the parasite.

Two potential salvage routes are possible - the salvaging unphosphorylated B_6 vitamers through diffusion from the erythrocyte or alternatively as Hb-bound forms (Figure 1.5). Pyridoxal kinases (PdxK) have been identified in *P. falciparum* and was shown to be involved in the phosphorylation of the various inactive B_6 vitamers, utilising adenosine triphosphate (ATP) as a substrate [132]. Through phosphorylation, PdxK effectively traps PLP within the parasites [133]. Prodrugs, consisting of PN piggybacked onto tryptophan methyl esters, were shown to be phosphorylated by parasitic PdxK resulting in trapping of toxic adducts capable of inhibiting other PLP-dependent enzymes [105]. Uptake of these was presumed to be through an amino acid transporter, which may recognise the tryptophan moiety.

Alternative sources of PLP which form part of the salvage pathway could include release of PLP during proteolysis of Hb in parasite acidic digestive food vacuoles (Figure 1.5). The recent discovery of acid phosphatases suggests that unphosphorylated B_6 vitamers could potentially be salvaged via this route [134]. The acid phosphatases are secreted via the plasma membrane to the parasite food vacuoles and have broad substrate acceptance, including the ability to accept PLP. This implies that PLP from Hb in the food vacuoles could be dephosphorylated, and capable of entering the parasites through diffusion. PdxK could re-activate PL into the active cofactor.

Enzymes involved in interconversion of PNP, PMP and PLP such putative pyridoxine oxidases (PF14_0570) have not experimentally been verified, and raises the question whether the parasites are capable of, or have need for, interconversion between PNP, PMP and PLP [135]. Other components of the salvage pathway include PLP phosphatase (PdxPP, PF07_0059) also termed 4-nitrophenyl phosphatase (PNPase), which dephosphorylate PLP to produce PL [133, 136]. *P. falciparum* parasites have a cytosolic expressed PNPase protein with broad substrate acceptance [136]. The exact contribution of these cytosolic PNPases and secreted acid phosphatases in B_6 metabolism remains to be determined.

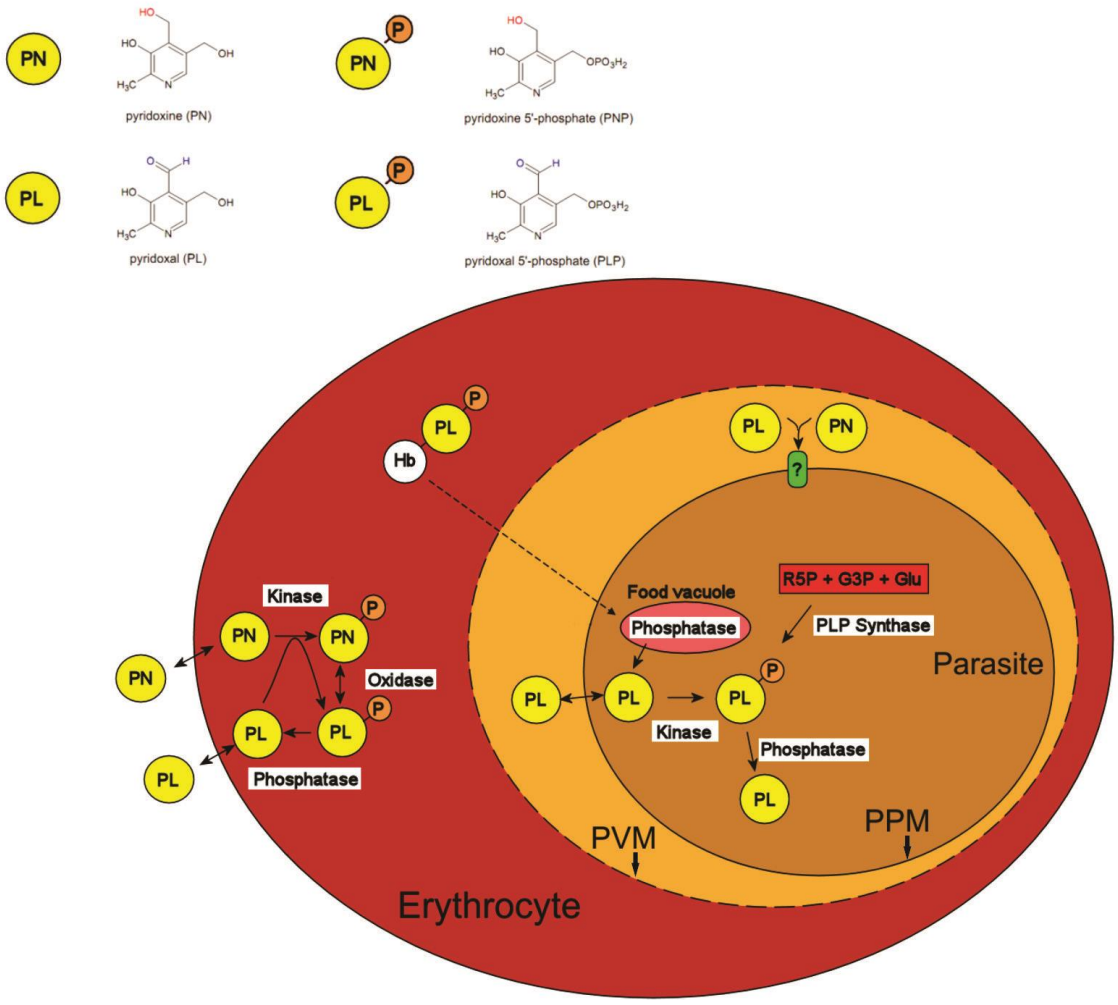


Figure 1.5: Erythrocyte and *P. falciparum* B_6 homeostasis. Unphosphorylated PN, PM and PL can diffuse into the erythrocyte, and are trapped within via phosphorylation by PdxK. The resulting PNP and PMP can be converted into PLP by oxidases and transaminases. PdxP phosphatases catalyse dephosphorylation of PLP and PNP to allow diffusion from the erythrocyte back to peripheral tissues. Unphosphorylated B_6 vitamers can diffuse into the malaria parasite. PLP can enter when attached to Hb, and is released in the acid food vacuoles in which phosphatases are involved. The parasite PdxK can convert salvaged PL to PLP. The parasites possess PLP synthases which can produce PLP from R5P, G3P and L-glutamine. PVM - parasitophorous vacuolar membrane and PPM - parasite plasma membrane.

In *P. falciparum* parasites, PLP is produced *de novo* by PLP synthase. The *P. falciparum* PLP synthase consists of *Pf*Pdx1 and *Pf*Pdx2 which assembles into a dimer, which arranges into a large dodecameric complex [133]. *Pf*Pdx1 and *Pf*Pdx2 form an obligate heterodimeric complex, and one cannot function without the other in the cellular environment (*in vitro Pf*Pdx1 can function without *Pf*Pdx2 when supplied with ammonia). *Pf*Pdx2 are glutaminases (or glutamine amidotransferases) which produce ammonia through the hydrolysis of L-glutamine which is channelled to Pdx1 subunits [104]. The *P. falciparum* Pdx2 protein was first annotated as the SNO-type pyridoxine biosynthesis protein in a proteomic study by Lasonder *et al.* [137]. Wrenger *et al.* confirmed the single gene copies of *Pf*Pdx1 and *Pf*Pdx2 were present in *P. falciparum* parasites, [132]. The

*Pf*Pdx1 and *Pf*Pdx2 proteins are translated from single gene copy mRNA's, with maximal mRNA transcript expression at 26 hours post invasion (HPI) corresponding to trophozoite parasite stages [132]. Gengenbacher *et al.* later revealed that corresponding protein levels only peaked at the later 34 HPI schizont parasite stages and the proteins are localized to the cytosol of the parasite [104]. Compared to the energy demanding PdxK ATP-catalysed reactions for salvage of PLP, *de novo* biosynthesis of PLP is a low cellular-energy cost route for maintaining the PLP levels.

1.3.4 Vitamin B_6 biosynthesis in *P. falciparum* as drug target

Free haem iron and haem moiety release during Hb digestion is associated with oxidative stress and generates superoxide anions (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radicals (OH·) [46]. The parasites combat oxidative damage through release of superoxide dismutases and catalases, and also inactivate free haem through polymerisation to form inert hemozoin [46]. PLP and other B_6 vitamers are potent antioxidants capable of quenching $^1O_2$ [117, 138, 139]. PLP can reduce superoxide anions (O_2^-), and is a potent antioxidant comparable to vitamin C [116, 117]. During the IDC of the malaria parasite $^1O_2$ was shown to peak during the trophozoite developmental stage and has been correlated with Hb digestion [140]. The expression of *Pf*Pdx1 and *Pf*Pdx2 during trophozoite and schizont parasite stages coincides with degradation of large amounts of Hb [132]. Moreover, *Pf*Pdx1 and *Pf*Pdx2 transcripts were overexpressed during methylene blue and cercosporin stress-treatment [132, 141]. This implied that *Pf*Pdx1 and *Pf*Pdx2 could aid during oxidate stress caused by ROS. Episomal over-expression of both *Pf*Pdx1 and *Pf*Pdx2 in the parasites resulted in increased levels of PLP and was shown to protect parasites from a sub-lethal dose of cercosporin, a singlet oxygen producing toxin from the pathogenic fungus *Cercospora nicotinae* [141]. Moreover, parasites over-expressing *Pf*Pdx1 and *Pf*Pdx2 proteins were less sensitive to cercosporin, emphasising the role PLP plays in neutralization of ROS [141]. The PLP synthase is therefore an important and indispensable protein that produces PLP which limits the cellular stress caused by ROS. These studies support *Pf*Pdx1 and *Pf*Pdx2 as good candidates for drug design; however there are no known inhibitors that target the PLP synthase. This therefore establishes the incentive for drug discovery on the PLP synthase.

Table 1.2: PLP-dependent enzymes in *P. falciparum*. Table was adapted from Müller *et al.* [108].

Enzyme name	EC number	PlasmoDB ID	Involvement in metabolism
Serine hydroxymethyltransferase (SHMT)	2.1.2.1	PFL1720w PF14_0534 (putative)	Folate metabolism
p-aminobenzoic acid synthetase	4.1.3.38	PFI1100w	Folate biosynthesis
Cysteine desulfurase	2.8.1.7	MAL7P1.150 (putative) PF07_0068 (putative)	Iron-sulphur cluster synthesis
Aspartate aminotransferase, aspartate transaminase (AspAT)	2.6.1.1 or 2.6.1.57	PFB0200c	Amino acid metabolism
Lysine decarboxylase	4.1.1.18	PFD0285c (putative) PFD0670c (putative)	Polyamine metabolism
Ornithine aminotransferase (OAT)	2.6.1.13	PFF0435w	Arginine and proline metabolism
Ornithine decarboxylase (ODC) in bifunctional S-adenosylmethionine decarboxylase-ODC	4.1.1.17	PF10_0322	Polyamine metabolism
Delta-aminolevulinic acid synthetase	2.3.1.37	PFL2210w	Porphyrin metabolism, amino acid metabolism
Serine C-palmitoyltransferase (SPT)	2.3.1.50	PF14_0155 (putative)	Sphingolipid metabolism
Phosphatidylserine decarboxylase (PSDC)	4.1.1.65	PFI1370c	Glycerophospholipid metabolism
Branched-chain amino acid aminotransferase	2.6.1.42	PF14_0557 (putative)	Pantothenate and CoA biosynthesis; amino acid metabolism

The malaria parasite possesses several PLP-dependent enzymes, and these have diverse involvement in amino acid, folate and polyamine-metabolism (Table 1.2). Several PLP-dependent enzymes are established targets for commercially available drugs; such as L-ornithine decarboxylase (ODC) for the treatment of African sleeping sickness caused by *Trypanosoma brucei gambiense* with α-difluoromethylornithine (Eflornithine) [142]), γ-aminobutyric acid (GABA)-aminotransferases for treatment of epilepsy using Vigabatrin [143], and L-3,4-dihydroxyphenylalanine (DOPA) decarboxylase for the treatment of Parkinson's disease and hypertension using carbidopa or Lodosyn [105, 144]. Some of these PLP-dependent enzymes have not been identified in *P. falciparum*, however the select few which are present, as listed in Table 1.2, are considered as good drug targets [145]. By targeting PLP biosynthesis and disrupting PLP homeostasis the multitude of PLP-dependent processes could be affected which should be lethal to the parasites. The diverse metabolic involvement of the PLP cofactor in these processes underscores the importance of the *de novo* PLP biosynthetic pathway, and makes *Pf*Pdx1 and *Pf*Pdx2 promising drug targets.

1.4 Hypothesis, research objective and aims

Hypothesis: Inhibition of *de novo* PLP biosynthesis in the malaria parasite *P. falciparum* is detrimental to parasite growth *in vitro* and results in compensatory responses functionally linked to PLP metabolism.

The objective of this study was the *in silico* identification of *Pf*Pdx1 inhibitors, analyses of their inhibitory capacity on parasite proliferation and the functional consequences thereof.

The aims of this study were to:

a) Use *in silico* / computational approaches to aid in the identification of potential inhibitory compounds against *Pf*Pdx1.
b) Determine the inhibitory capacity of *in silico*-identified and rationally selected compounds on *Pf*Pdx1 activity.
c) Analyse parasite proliferation when treated with *Pf*Pdx1 inhibitors.
d) Determine the global functional consequences of inhibiting *Pf*Pdx1 by analysis of its transcriptome and proteome of the parasites.

Chapter 2, describes homology modelling and structure-based virtual screening of *Pf*Pdx1 pharmacophores, of which hits were evaluated based on their chemical profiles and ability to dock into *Pf*Pdx1 homology models.

Subsequently, in Chapter 3, we asked if the *in silico* identified compounds were able to elicit inhibition of recombinantly expressed *Pf*Pdx1. Additionally, we also evaluated the inhibitory activity of rationally identified compounds on *Pf*Pdx1. Moreover, growth assays were used to show that an successful *Pf*Pdx1 inhibitor affected the proliferation of the parasites. Mutagenesis studies also helped identify essential residues required for a R5P reaction intermediate formation, contributing to the biochemical characterisation of this enzymatic activity.

Lastly, we asked if the identified *Pf*Pdx1 inhibitors were able to elicit inhibition of parasite proliferation and if so, what the functional consequences thereof were on the parasite. Chapter 4 describes an investigation of the inhibitory effect by monitoring global transcriptional and proteome-wide changes.

In a concluding discussion, Chapter 5 evaluates the scientific contributions made in this study and attempts to evaluate targeting of *de novo* PLP biosynthesis in the parasite as a viable strategy for novel antimalarial drug development.

1.5 Outputs

Research findings originating from this thesis has been presented at the following instances:

Conference proceedings:

1. Reeksting, S. B., Müller, I. B., Salmon, L., Louw, A. I., Birkholtz, L. M. and Wrenger, C. Identification of structure-based leads from the *Plasmodium falciparum* PLP synthase enzyme Pdx1. Polyamines in Parasites 2010, 6th Biennial Symposium, Phalaborwa, South Africa. (**Oral presentation**).
2. Reeksting, S. B., Müller, I. B., Salmon, L., Louw, A. I., Birkholtz, L. M. and Wrenger, C. Identification of inhibitors against the *Plasmodium falciparum* PLP synthase enzyme Pdx1. SASBMB 2012, Champagne castle, South Africa. (**Oral presentation**).
3. Reeksting, S. B., Müller, I. B., Salmon, L., Louw, A. I., Birkholtz, L. M. and Wrenger, C. Identification of inhibitors against the *Plasmodium falciparum* PLP synthase enzyme Pdx1. Molecular approaches to malaria (MAM) 2012. Melbourne, Australia. (**Poster presentation**).

Manuscripts:

1. Reeksting, S. B., Müller, I. B., Burger, P. B., Burgos, E. S., Salmon, L., Louw, A. I., Birkholtz, L. M. and Wrenger, C. (2013), Exploring inhibition of Pdx1, a component of the PLP synthase complex of the human malaria parasite *Plasmodium falciparum*. *Biochemical Journal***. 449**, p. 175-187.

2. Reeksting, S. B., Olivier, N., Burgos, E. S., Salmon, L., Louw, A. I., Birkholtz, L. M. and Wrenger, C. Functional consequences of targeting PLP synthesis as novel drug target in malaria parasites. S.B. Reeksting, J. Reader, N Olivier, C Wrenger and L Birkholtz. Manuscript in preparation.

Chapter 2

In silico modelling of *Pf*Pdx1 to facilitate selection of potential R5P-active site inhibitors

2.1 Introduction

2.1.1 Drug discovery

There are several different screening methods for the identification of novel antimalarial compounds and these include whole cell or phenotypic screening, target-based screening and virtual-based screening. Phenotypic screening or *in vitro* whole cell assays have been more successful in identification of chemical leads since 1) usually no additional modifications are required to improve pharmacokinetics or permeability of identified compounds and 2) an unbiased approach is employed [99, 146]. Another advantage of whole cell screens, which could lend support to its success, is the fact that the complete parasite biology is represented and therefore interrogated during screens. Antimalarials such as CQ and other 4-aminoquinolines affect haemozoin polymerisation, a cellular process, rather than a specific protein. One of the disadvantages of this approach is the fact that the molecular target is not always known. Resistance surveillance therefore becomes more intensive and tailoring the compound for improved efficacy is harder. Fragment-based libraries compliment phenotypic screens and uses small molecules (100 – 250 Da) at relatively high concentrations to provide an initial skeleton for additional chemical optimisations [147]. This drug design methodology argues that as the complexity of a ligand increases, the chance match to a potential drug target decreases [148]. A library of complex molecules is therefore not as likely to succeed in having the desired effect on a drug target compared to smaller and simpler molecule types.

Target-based screening involves the pre-selection and validation of essential parasites proteins, which are then either subjected to HTS or structure-based screening methods (if the target structure is available). In the laboratory target-based screening requires the establishment of sensitive *in vitro* assay procedures and is often hampered by poor quality and low quantities of recombinantly expressed proteins from malaria parasites. Moreover, compounds identified in such screens may not always have the required potency *in vivo*, necessitating chemical optimisation to either improve efficacy and/or pharmacokinetic and pharmacodynamic properties [99]. Pursuing compounds identified from target-based inhibition screens are often not successful as these fail in whole cell activity assays, even after successive compound optimisations [99, 146, 149]. It has been proposed that chemical optimisation to improve cell permeability and pharmacokinetics results in fewer leads

compared to whole-cell phenotypic screening approaches [99, 149]. Newer drug-delivery methods may, however, complement such shortcomings [99, 149]. One advantage of target-based drug identification is that these approaches can more effectively benefit from ligand-target binding information, such as NMR or crystallographic evidence of inhibitor interactions, resulting in rapidly optimised leads.

The advancement of computational techniques has dramatically aided in drug discovery and virtual-based screening approaches have become more prominent largely due to lower costs and being less time consuming compared to both whole cell and target-based screens [150, 151]. Virtual-based screening can either be used to extract useful information from ligands or the actual protein target structure or receptor. Ligand-based methods can be used to identify the 3D shape, chemical similarity or pharmacophore features of known active compounds and aid in the selection of analogous compounds [150]. [150]. When the target structure is known, receptor or structure-based computational or virtual screening methods can be used [150]. Structure-based drug design in turn aims to predict complementary interactions ligands have with a receptor. This enables the rationalization of binding modes and comparisons of the structure activity relationships (SAR) ligands possess. Structure-based methods include molecular docking simulations, which aim to identify low-energy binding partners, as well as pharmacophore approaches, which describe the necessary chemical features compounds should possess to bind effectively. An example of a successful structure-based study by Haque *et al.* identified potent inhibitors with K_i-values in the nanomolar range against two aspartyl protease plasmepsin proteins of the malaria parasite [152]. The computational technique was first utilised to derive building blocks for the library synthesis from a protein homologue cathepsin D, after which these compounds were tested on plasmepsin [152, 153]. Another study focussed on identifying more specific inhibitors against cysteine protease homologues from *Leishmania donovani*, *Trypanosoma cruzi* and *P. falciparum* (cysteine proteases termed falcipain-2 and falcipain-3) [154]. Computational libraries of around 335 000 compounds were filtered and docked into the *T. cruzi* structure as well as both falcipain-2 and falcipain-3 homology models. Several inhibitors with low micromolar efficacy were identified against falcipain-2 and the *L. donovani* proteins, and an additional pharmacophore representative of these proteins was created and deemed useful for additional interrogations of the active site [154].

A pharmacophore model is an ensemble of chemical features required by a compound, protein or cellular component for molecular recognition [155]. The pharmacophore includes chemical moieties on a molecule that are needed for interactions with a protein receptor and *vice versa* the amino acid residues of a protein that enable recognition of a molecule or cellular binding partner. Ligand-based pharmacophores can be created by overlaying known substrates, inhibitors or

products and extracting the shared chemical characteristics [155]. Structure-based pharmacophore screening approaches aim to map compounds to the chemical environment of the target active site. Structure-based pharmacophore mapping entails identifying complementary chemical features and their spatial arrangement in the active site. The receptor-ligand complex is often used to locate the active site, alternatively atoms in cavity regions and interaction maps are generated to probe potential ligand binding site. These pharmacophores are then screened using virtual screening and searching approaches. This approach helps to discover structurally diverse compounds which are capable of targeting the receptor protein and aims to provide a comprehensive repertoire of molecules which are effective. Structure-based pharmacophores are generated by identifying interactions between substrates or products and their respective protein receptors (Figure 2.1). A limitation to structure-based drug design could be the lack of crystallographic data of structures, and also validated targets, and this is often overcome by using homology modelling.

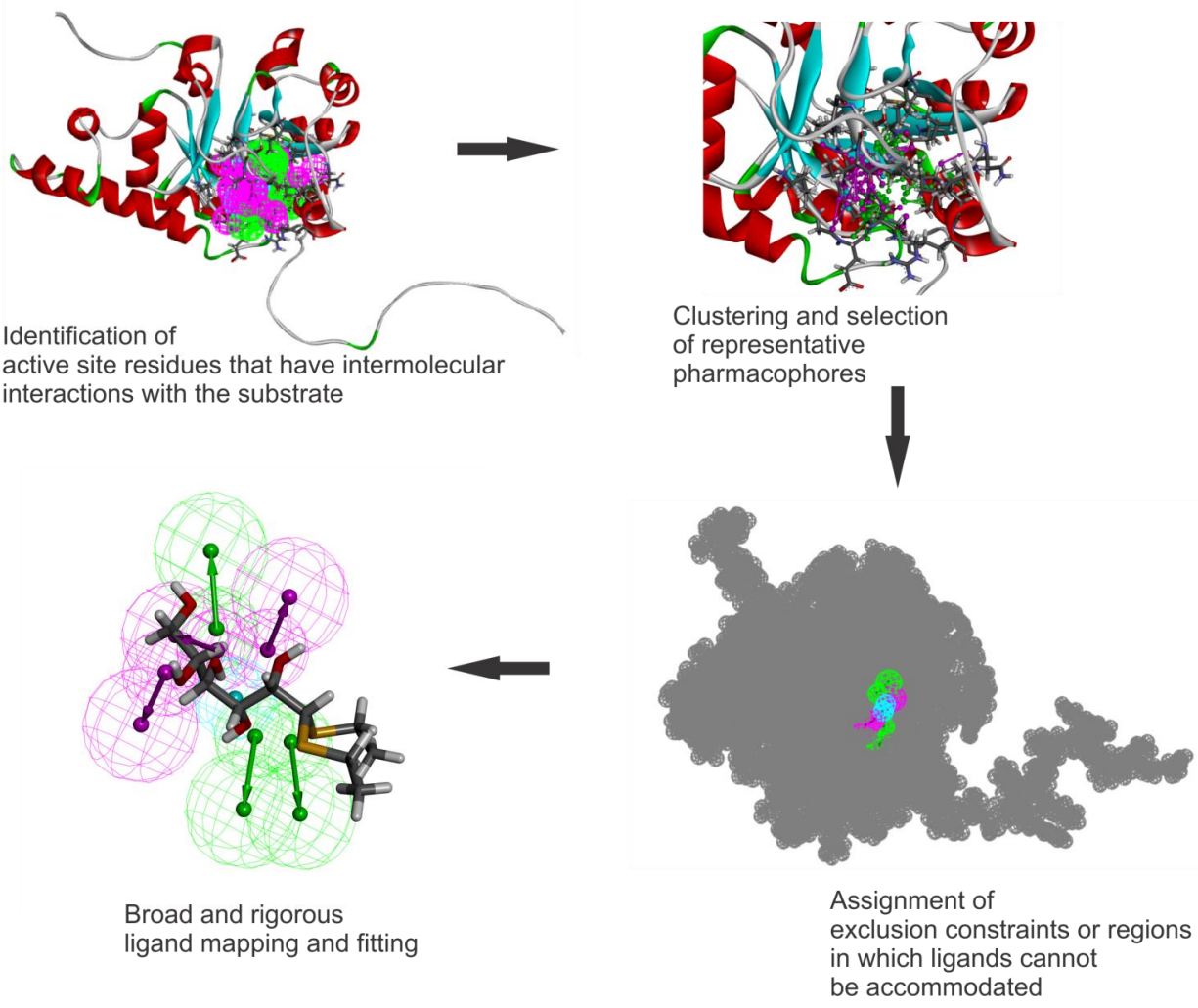


Figure 2.1: Workflow for the identification of ligands from a structure-based pharmacophore. Protein structures or homology models are used to extract pharmacophores from the active site. Disallowed regions are assigned in the hypothetical pharmacophore model based on the location of other amino acid residues that are not directly associated with the substrate. Ligands are then finally mapped to the pharmacophores and aligned with complementary chemical moieties.

2.1.2 The Pdx1 component of PLP synthase

The PLP synthase is composed of two proteins; Pdx1 and Pdx2 (Figure 2.2). In *B. subtilis*, *G. stearothermophilus* and *Thermotoga maritima* it was shown that the PLP synthase complex forms a large dodecameric ring structure, composed of two hexameric rings, each hexameric ring consisting of six Pdx1 subunits surrounded by six Pdx2 proteins [119, 156, 157]. Assembly of the *P. falciparum* PLP synthase is initiated when inactive *Pf*Pdx2 monomers decorate the two *Pf*Pdx1 hexameric rings to form a dodecamer of 424 kDa [133]. Structural characterisation of the plasmodial PLP synthase has highlighted important aspects of assembly and exploitable characteristics for the development of novel therapeutics [133]. During assembly of PLP synthases Pdx1 activates the Pdx2 glutaminase activity, resulting in the hydrolysis of L-glutamine, thereby releasing ammonia that is channelled to Pdx1 subunits.

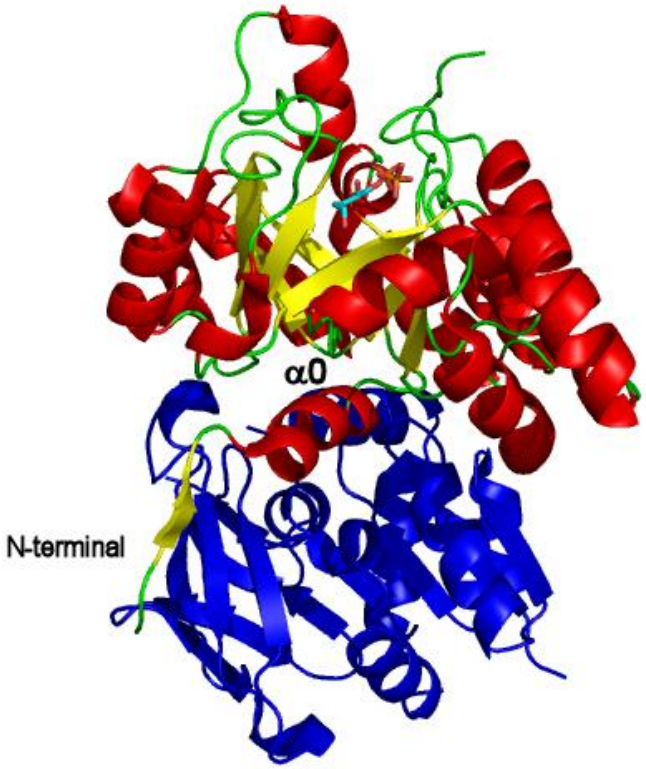


Figure 2.2: The heterodimeric structure of *T. maritima* PLP synthase (PDB: 2ISS). *Tm*Pdx1, coloured in red and yellow, associates with TmPdx2, coloured in blue [157]. *Tm*Pdx1 has TIM $(\alpha/\beta)_8$ barrel fold architecture with eight central parallel β-sheets (yellow) alternated with by eight α-helices (red), with some modifications. The N-terminal β-sheet and the α-helical region (α0 or αN) of *Tm*Pdx1 contacts the *Tm*Pdx2 monomer and have been shown to stimulate glutaminase activity in *Tm*Pdx2 [156, 157].

Pdx1 proteins are highly conserved, however the N-terminal region, spanning the first 20 amino acids, differs markedly between species and plays an important role in PLP production [158]. The Pdx1 N-terminal stimulates Pdx2 activity in the assembled heterodimeric complex, and this region in *Pf*Pdx1 is believed to contribute half of the total interactions between heterodimers [159]. The

crystal structure of *Tm*Pdx1 in complex with *Tm*Pdx2 revealed ordered interactions promoted by the Pdx1 N-terminal region, including residues 8 - 16 which form an α-helix, termed α0 or αN (Figure 2.2) [157]. The *B. subtilis* Pdx1 (*Bs*Pdx1) N-terminal region, composed of an additional short β-strand and α-helix, is essential for *Bs*Pdx2 binding and interaction [156]. Both *Tm*Pdx1 and *Bs*Pdx1 N-terminal regions have predominately polar interaction with Pdx2, involving both acidic and basic residues [157, 159]. (Figure 2.2) [157]. The *B. subtilis* Pdx1 (*Bs*Pdx1) N-terminal region, composed of an additional short β-strand and α-helix, is essential for *Bs*Pdx2 binding and interaction [156]. Both *Tm*Pdx1 and *Bs*Pdx1 N-terminal regions have predominately polar interaction with Pdx2, involving both acidic and basic residues [157, 159].

The *Pf*Pdx1 enzyme performs the bulk of the catalytic formation of PLP. The enzyme combines R5P, G3P and ammonia derived from active Pdx2 monomers, to synthesise PLP. The reaction mechanism was already comprehensively studied and elucidated at the initiation of this study [160-162]. These studies eluded to the complex imine chemistry involved in different active sites, however at the time only one site with the R5P substrate was definitively known from crystallographic evidence of *Tm*Pdx1, including evidence from mutational studies [157]. The R5P binding site is also termed the P1 site. Pdx1 enzymes can also bind D-ribulose-5-phosphate (Ru5P) in the P1 site, and this pentose sugar is also incorporated into the final PLP molecule, which is discussed in more detail in Chapter 3. Two additional regions in Pdx1 were later identified which were shown to accommodate PLP and G3P, and were termed the P2 and P3 sites respectively [156, 163, 164].

Contributions of amino acid side-chains in the *Tm*Pdx1 R5P-binding or P1 site are shown in Figure 2.3. The first part of the Pdx1 reaction mechanism is believed to involve imine formation of R5P with the internal K residue [160]. The involvement of a K82 residue was confirmed in *Tm*Pdx1, in which it was observed to be covalently bound, through the formation of an imine, to the C2 carbon of Ru5P (Figure 2.3) [157]. In the recently solved *Plasmodium berghei* structure of Pdx1 (*Pb*Pdx1) it was shown that unlike in the *Tm*Pdx1 structure, the K84 residue was covalently linked to the C1 carbon of R5P [164]. The *Pb*Pdx1 and *Tm*Pdx1 structures are therefore progressive snapshots of the initial R5P-binding stages, respectively. Other residues that are important in the R5P-binding site include a D25 in *Tm*Pdx1 (D27 in *Pb*Pdx1) involved in H-bond contacts with the hydroxyl groups located on the C2 or C3 carbon of R5P and Ru5P, respectively [157, 164]. Residues G215, G236 and S237 in *Tm*Pdx1 (Figure 2.3) are H-bonded to the phosphate group of Ru5P, and are believed to facilitate anchoring of the substrate [157]. Residues in *Pb*Pdx1 ascribed with similar roles include G217, S239 and G238 and additionally, G156 [164].

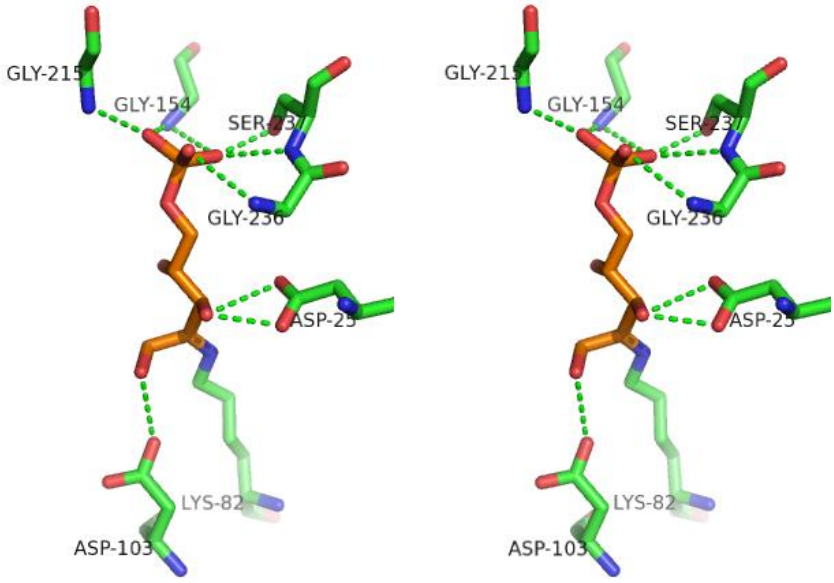


Figure 2.3: Stereoview of the *Tm*Pdx1 R5P-active site. The figure shows residues involved in hydrogen bonding to the Ru5P molecule, which include K82 forming a covalent imine bond with the substrate, as well as a series of glycine residues involved in stabilization of the phosphate moiety [157].

From a structure-based pharmacophore drug design approach, *Pf*Pdx1 is an attractive drug target as this protein is highly conserved and crystallographic information of Pdx1 homologues are available. The high degree of sequence identity shared amongst Pdx1 proteins enables quality homology modelling, and allows more truthful predictions regarding ligand-receptor binding interactions [165]. *Pf*Pdx1 homology models were generated and five representative structure-based pharmacophores were extracted from the R5P-binding site. These pharmacophores were used for the virtual screening of 2 million ligands to identify potential inhibitory compounds. Results from the virtual screening were confirmed with molecular docking to support pharmacophore-identified compounds.

2.2 Methods

2.2.1 Homology modelling

Alignment of templates and molecular modelling was performed using the Discovery Studio (DS) suite (v3.0, Accelrys Software Inc.) unless specified differently. Multiple sequence alignments (MSA) to compare the different templates were generated using the BLOSUM 62 pairwise alignment scoring matrix. The gap open penalty was set to 11, with an extension penalty of 1. *Pf*Pdx1 homology models were generated using *T. maritima* (*Tm*Pdx1, PDB accession: 2ISS) and *Thermus thermophilus* (*Tt*Pdx1, PDB accession: 2ZBT) Pdx1 as templates. At the time this study was initiated the *P. berghei* structure of Pdx1 (*Pb*Pdx1, PDB accession: 4ADU) was not available.

Models were generated using aligned templates in MODELER within the Build homology models protocol [166]. MODELER optimises spatial restraints to derive models from provided template structures. Two separate homology models were generated, one in which the Ru5P substrate from the 2ISS template was copied, the other model was built containing the (4*S*)-2-methyl-2,4-pentanediol (MPD) from the 2ZBT template. These structures were labelled *Pf*Pdx1_Ru5P and *Pf*Pdx1_MPD, respectively. The 2ZBT template crystal structure was solved with several water molecules in near vicinity of the MPD ligand. These were labelled as H_2O51, H_2O58, H_2O95, H_2O102, H_2O124 and H_2O139 in the 2ZBT structure, and were additionally included in the *Pf*Pdx1_MPD homology model, which was labelled *Pf*Pdx1_MPD_W. Five versions of the models were generated and the model with the lowest probability density function (PDF) total energy was used in further comparisons. Both the PDF total energy and PDF physical energy was calculated during this process. Additionally the discrete optimized protein energy (DOPE) score was computed for each homology model. Structural comparisons and superimpositions were made using both the DS (Accelrys Inc.) and the Java-based RCSB PDB comparison tool [167] with the Java implemented CE algorithm (jCE, version 1.1) [168]. Ramachandran plots were generated using PDBsum [169, 170] and the online resource PROCHECK was used to stereochemically assess the qualities and geometry of the homology models [171].

2.2.2 Protein minimisation

The *Pf*Pdx1 homology models were prepared for pharmacophore mapping by adding hydrogens using the CHARMm forcefield, after which the models were subjected to energy minimisation steps [172]. The CHARMm forcefield tool is a flexible molecular dynamics and modelling program which assigns atomic charges to atoms and residue types to residues in the models in preparation for

energy minimisation. In energy minimisations with the CHARMm minimisation algorithm 1000 steps of Steepest Descent were used with an RMS gradient tolerance of 3 kcal/mol.Å and a distance cutoff value of 14 Å for nonbonded interaction pairs. The maximum steps during minimisation were limited to 200 and the minimisation energy change was set to 0. This minimisation was followed by a conjugate gradient minimisation with the same parameters except for the RMS gradient which was set to 0.1 kcal/mol.Å. The substrate within the minimised structure was used to assign a binding site sphere with a radius of 8Å.

2.2.3 Interaction generation

The minimised homology models were subjected to the interaction generation protocol in the DS. This approach probes the receptor and identifies the interaction map or pharmacophore features in the assigned binding sphere based on amino acid properties. The position of the individual ligands in the models allowed the creation of an active site sphere. The sphere radius was, in each case, set to 8.0 Å. The water molecules in the *Pf*Pdx1_MPD_W model were included as receptor components. Water molecules form hydrogen bond (H-bond) networks which facilitate ligand-receptor interactions [173]. Although dynamic, inclusion of these waters provided additional features in the ensemble of pharmacophores not derived from active site residues. In the *Pf*Pdx1_MPD_W model these water were H_2O51 (O,H1,H2), H_2O58 (O3,H3,H4), H_2O95 (O1,H5,H6), H_2O102(O5,H7,H8), H_2O124(O2,H9,H10) and H_2O139 (O4,H11,H12). Ligands were deleted from the models before initiation of the protocol. Running parameters were altered with the density of lipophilic sites set to 5 and the density of polar sites was also set to 5.

2.2.4 Pharmacophore clustering

From interaction generation between 20 and 70 different pharmacophore features were identified from the respective homology models. These pharmacophore features were hierarchically clustered into a distance matrix allowing a pharmacophore cluster centre to be selected based upon the distance from another feature. A custom feature pharmacophore according to interactions displayed for the Ru5P ligand in the homology model, was also created. Exclusion pharmacophore features were assigned to residues located outside a 5Å radius from the identified pharmacophore features.

2.2.5 Database searching and ligand pharmacophore mapping (LPM)

The pharmacophores were screened against the drug-like subset of the Zinc database comparing 2 million ligands [174]. The database was generated as a multi-conformer composite database using catDB with a maximum of 250 representative conformations (Accelrys Software Inc.). In the database searching algorithm the best searching method was used which adds additional flexibility to the ligands during fitting. The drug-like database was screened against 5 pharmacophore models and the combined hits produced from the 5 pharmacophores were further submitted for scoring based-pharmacophore mapping. All of the 5100 ligands identified from database searching, were re-submitted for LPM in which all the conformations of the ligands were used. A rigid fitting method was employed with the energy threshold of the ligands set to 20 kcal/mol. During mapping the atoms of the ligands are optimally fitted to the pharmacophore features using an algorithm established by Kabsch *et al.* [175]. The fit value is computed using both the numerical weight assigned to individual pharmacophore features and the distance of the ligand atoms which are fitted to the centre of a corresponding pharmacophore feature. The individual pharmacophore features of *Pf*Pdx1 pharmacophores were equally weighted, with each feature assigned a numerical value of 1. For stringent selection of ligands at least n – 1 pharmacophore features had to match in order to map the ligands (where n represents the total number of pharmacophore features in the representative model).

Chemical diversity was determined using Chemmine tools (http://chemmine.ucr.edu), a freely available online tool used to clusters compounds based on pairwise similarities which are defined using Atom Pair and Tanimoto descriptors. Atom Pair descriptors refer to the atom organization in terms of the paths around non-hydrogen atoms in a molecule, whereas Tanimoto descriptors refer to the proportion of features shared amongst compounds, as retrieved from http://chemmine.ucr.edu/iframe/work/intro/tutorial. The 50-best ligands identified from LPM were clustered based on their single-linkage pairwise similarities.

2.2.6 Docking

To reinforce the selection of ligands against *Pf*Pdx1, 300 of the best-hit compounds, based on their pharmacophore fit-values, were further subjected to molecular docking simulations. Molecular docking simulations using LigandFit were conducted using a dreiding energy forcefield, [176, 177]. Molecular docking simulations using LigandFit were conducted using a dreiding energy forcefield, [176, 177]. During conformational searching of the R5P-active site 1500 Monte Carlo trials were

used to iteratively generate ligand conformations, with the dielectric constant set to 1.0, and a non-bonded cut-off distance of 10.0 Å. Docking was performed with a RMS threshold of 2.0 Å for ligand/binding site matching. Ligand poses were ranked according to DockScore functions based on forcefield approximations, taking into account the van der Waals and electrostatic interaction energy between ligand and receptor, as well as the ligand internal energy. Ligand poses were additionally evaluated using the LigScore empirical scoring functions [178]. This defines van der Waals interaction between ligands and the receptor, derived using polar attraction, and desolvation penalties incurred when ligands bind to a protein receptor. R5P was used as control in cross-docking to validate ligand docking and scoring. Following the release of the *Pb*Pdx1 structure in 2012 [164], additional docking was performed on the *Pf*Pdx1 homologue. The *Pb*Pdx1 structure was prepared by removing the covalent bond between K84 and R5P, followed by minimisation as described above. Post-minimisation, R5P was selected to assign a binding site and sphere radius, similar to the docking simulations described above. K84 was maintained in a protonated state during the docking simulations (predicted pK_a of 12.1).

2.3 Results

2.3.1 Alignments and template selection

A multiple protein sequence alignment (MSA) of the template Pdx1 crystal structures and the *Pf*Pdx1 homologue showed a 60% and 61% sequence identity with *T. maritima* Pdx1 (*Tm*Pdx1, PDB accession 2ISS) and *T. thermophilus* Pdx1 (*Tt*Pdx1, PDB accession 2ZBT), respectively (Figure 2.4). The sequence similarity was 79% and 77% compared to *Tm*Pdx1 and *Tt*Pdx1, respectively. After the release of the *P. berghei* Pdx1 crystal structure (*Pb*Pdx1, PDB accession 4ADU) an alignment was performed and showed a 86% sequence identity with *Pf*Pdx1, however at the time of homology modelling this structure was not available [164].

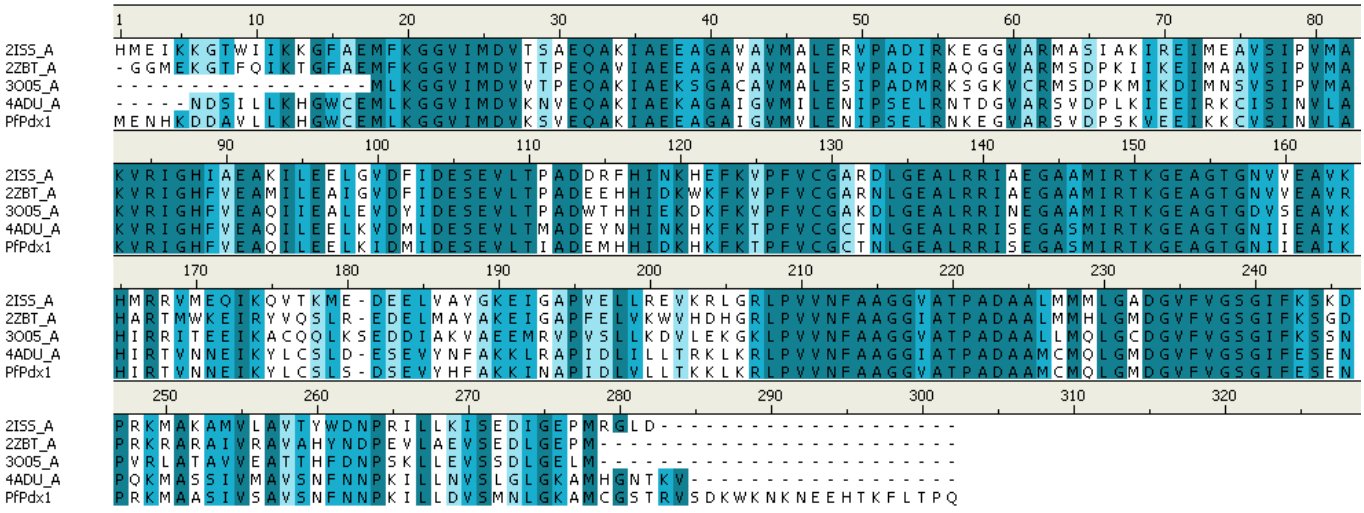


Figure 2.4: Multiple protein sequence alignment (MSA) generated for Pdx1 homologues. Aligned crystal structures of *T. maritima* (2ISS), *T. thermophilus* (2ZBT), *P. berghei* (4ADU), *S. cerevisae* Pdx1 (3O05) and *Pf*Pdx1 (pyridoxine biosynthetic enzyme pdx1 homologue, NCBI reference sequence XP_966196.1). Conserved residues indicated in dark blue, with residues with similar properties indicated in light blue.

The secondary structure topology profile of *Tm*Pdx1 and *Tt*Pdx1 are shown in Figure 2.5, together with the calculated residue conservation. These topology profiles were shown to be similar with some differences noted at the N-terminal region. *Tt*Pdx1 has more extensive β-sheet formation followed by a shorter α-helical region compared to *Tm*Pdx1 (Figure 2.5 A). Residues which form part of the R5P-binding site are indicated in the topology profiles (Figure 2.5). These included D25, M44, P50, K82, G154, G214 and G215 (as labelled in *Tm*Pdx1).

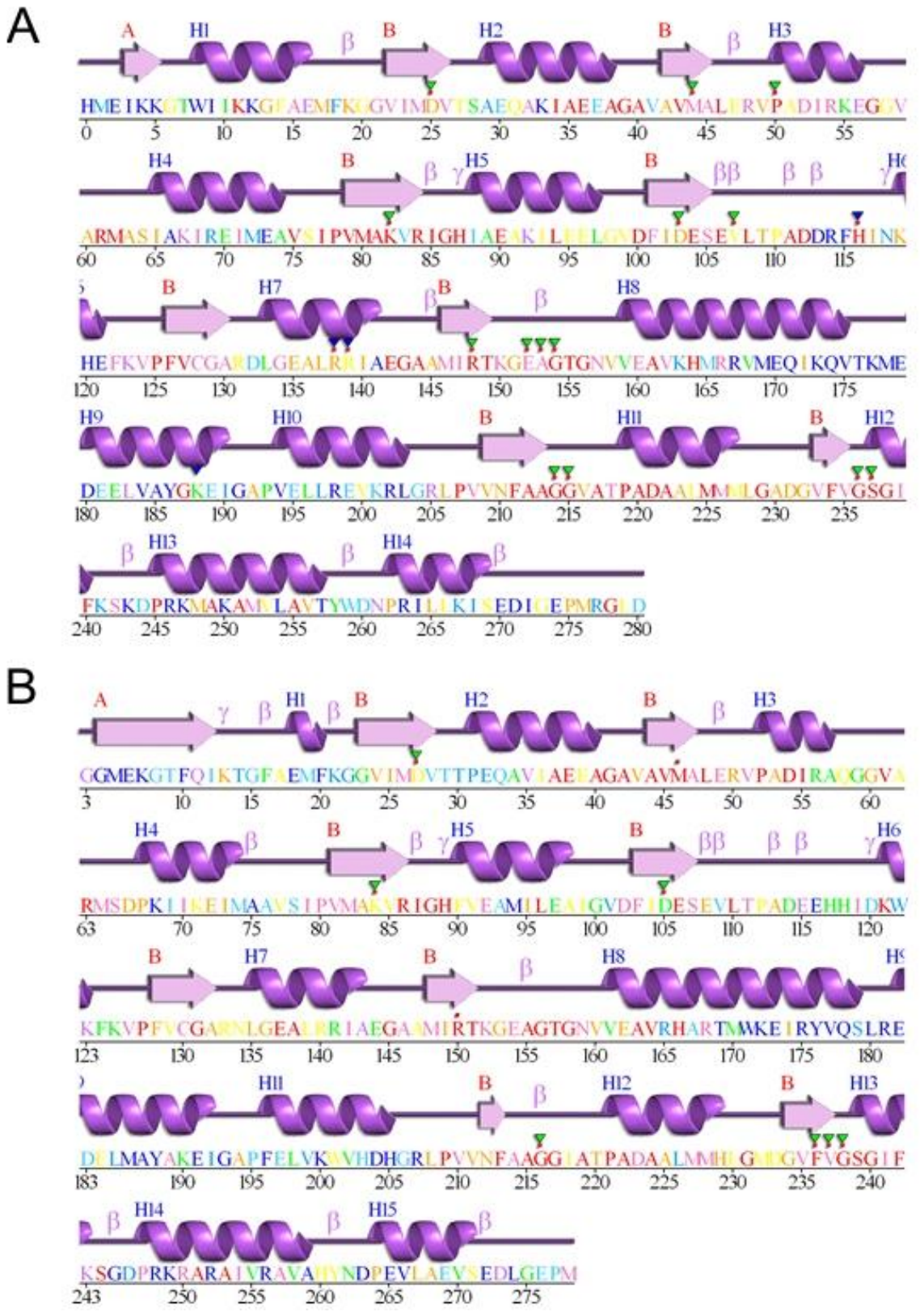


Figure 2.5: Secondary structure topology diagrams of A) *Tm*Pdx1 and B) *Tt*Pdx1. Topology diagrams were generated in PDBsum [169, 170]. Secondary structures such as α-helices are labeled with H, whereas β-sheet are labeled with B. β- and γ-turns are indicated by β and γ respectively. Active site residues with predicted interactions with ligands are indicated by red dots above the residue. Additional green triangles indicate that the substrate is part of the defined active site record in the PDB file. Residue conservation calculated using ConSurf-DB server is indicated by colour, in which least conserved residues are blue, somewhat conserved residues are light blue to green, and most conserved residues are shades of red.

A structural comparison between A chains of the templates, *Tm*Pdx1 and *Tt*Pdx1, revealed 74% sequence identity and 86% sequence with the C_α root mean square deviation (RMSD) of 1.0 Å (Figure 2.6 A). Residues surrounding the R5P substrate (5 Å radius) had a main chain RMSD (N, C_α, C, O) of 0.89 Å (Figure 2.6 B). Structural comparison of K82, D25, M44, R148 and G215 showed a good correlation with these residues adopting similar positioning (Figure 2.6 B). The mechanistically implicated K82 residue in *Tm*Pdx1, covalently bonded to the Ru5P ligand, was shown to have a similar orientation to K83 in *Tt*Pdx1 (Figure 2.6 B). The residue forms a H-bond with MPD in *Tt*Pdx1. Residue D25 of *Tm*Pdx1 was also shown to H-bond with the hydroxyl group of Ru5P, compared to the equivalent D27 residue in *Tt*Pdx1 which had two predicted H-bonds toward the hydroxyl groups of MPD. The S237 (S239 in TtPdx1) and P50 (P52 in *Tt*Pdx1) residues between the two templates differed significantly and is located more than 3.4 Å from each other (Figure 2.6 B).

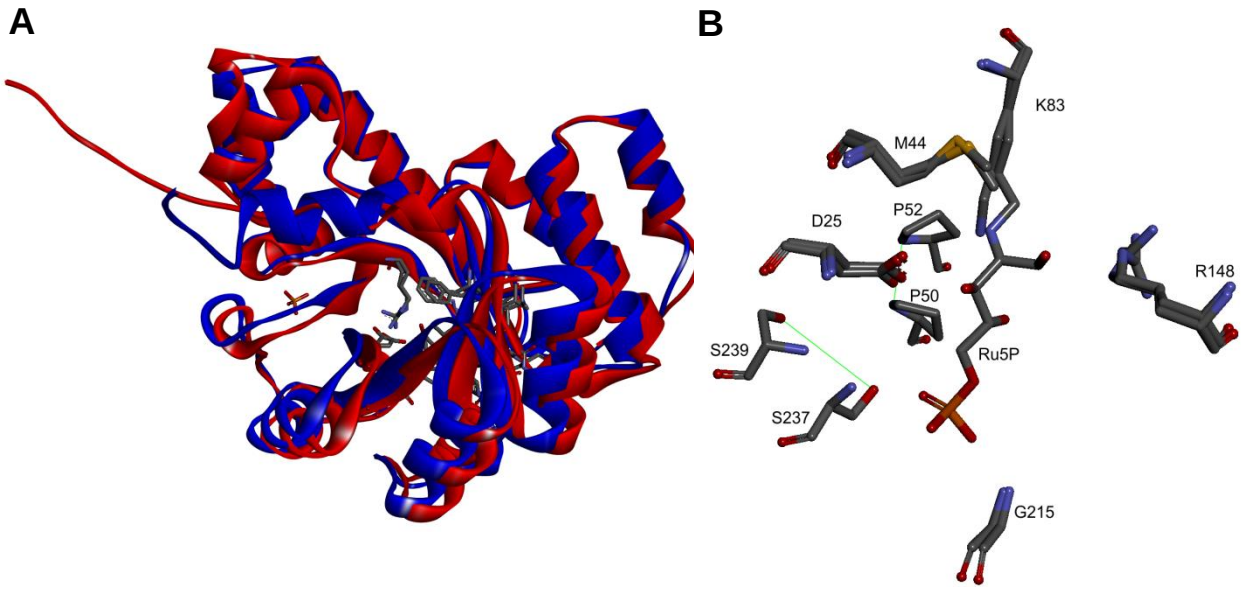


Figure 2.6: Comparison of *Tt*Pdx1 (PDB: 2ZBT) to *Tm*Pdx1 (PDB: 2ISS). A) Superimposition of the proteins revealed a C_α RMSD of 1.0 Å. **B)** Superimposition of R5P-binding site residues revealed conservation of residue locations, with a main chain RMSD of 0.89 Å. Markedly different residues were P50 and S237 which were located more than 3.4 Å from each other (*Tm*Pdx1 labelling).

2.3.2 Homology modelling of *Pf*Pdx1

Four homology models were generated based on the template sequence alignments. These models were generated each to include different substrate and water molecules from the different template structures. Three of the models contained the MPD ligand from the *Tt*Pdx1 structure

(*Pf*Pdx1_MPD, *Pf*Pdx1_MPDsco and *Pf*Pdx1_MPDw) whereas the other contained the Ru5P ligand from *Tm*Pdx1 (termed *Pf*Pdx1_Ru5P). The models were subjected to protein quality assessment in which the PDF total energy, physical PDF and discrete optimized protein energy (DOPE) scores were calculated before being used in further studies (Table 2.1). The PDF total energy was determined and gives the model-derived pseudo energy terms as wells as the stereochemical pseudo-energy terms. In addition the physical PDF energy was determined and describes the summed energies of stereochemical pseudo-energy terms, consisting of valence angles and torsion angles, valence bonds, improper torsion angles and soft sphere repulsions.

Table 2.1: Properties of homology models built using MODELER in DS.

	PDF total energy[a]	PDF physical energy[a]	DOPE score[b]
***Pf*Pdx1_MPD**	7355.68	761.53	-31510.11
***Pf*Pdx1_Ru5P**	6918.45	727.15	-31585.13
***Pf*Pdx1_MPDw**	7355.68	761.53	-31510.11
***Pf*Pdx1_MPDsco**	7355.68	761.53	-31510.11

[a] PDF total and physical energies are unit-less descriptors of the restraint violations in the models, in which a higher comparative value indicates greater violations
[b] DOPE score or relative stability of the model is used to compare different states of the same model, whereas the absolute unit-less score is not meaningful due to the algorithms not normalizing for protein size [179].

In addition, the individual PDF scores for individual residues of models reflecting their steric pseudo-energy environments were calculated and are shown in Figure 2.7. Residues such as P51 and S238 in the *Pf*Pdx1_MPD models had high PDF total scores indicating steric collisions which were further resolved using the CHARMm energy based minimisation procedure to resolve any steric clashes between side chain orientations. Resulting models such as *Pf*Pdx1_MPDsco had similar residue PDF total energy line plots as for *Pf*Pdx1_Ru5P (results not shown).

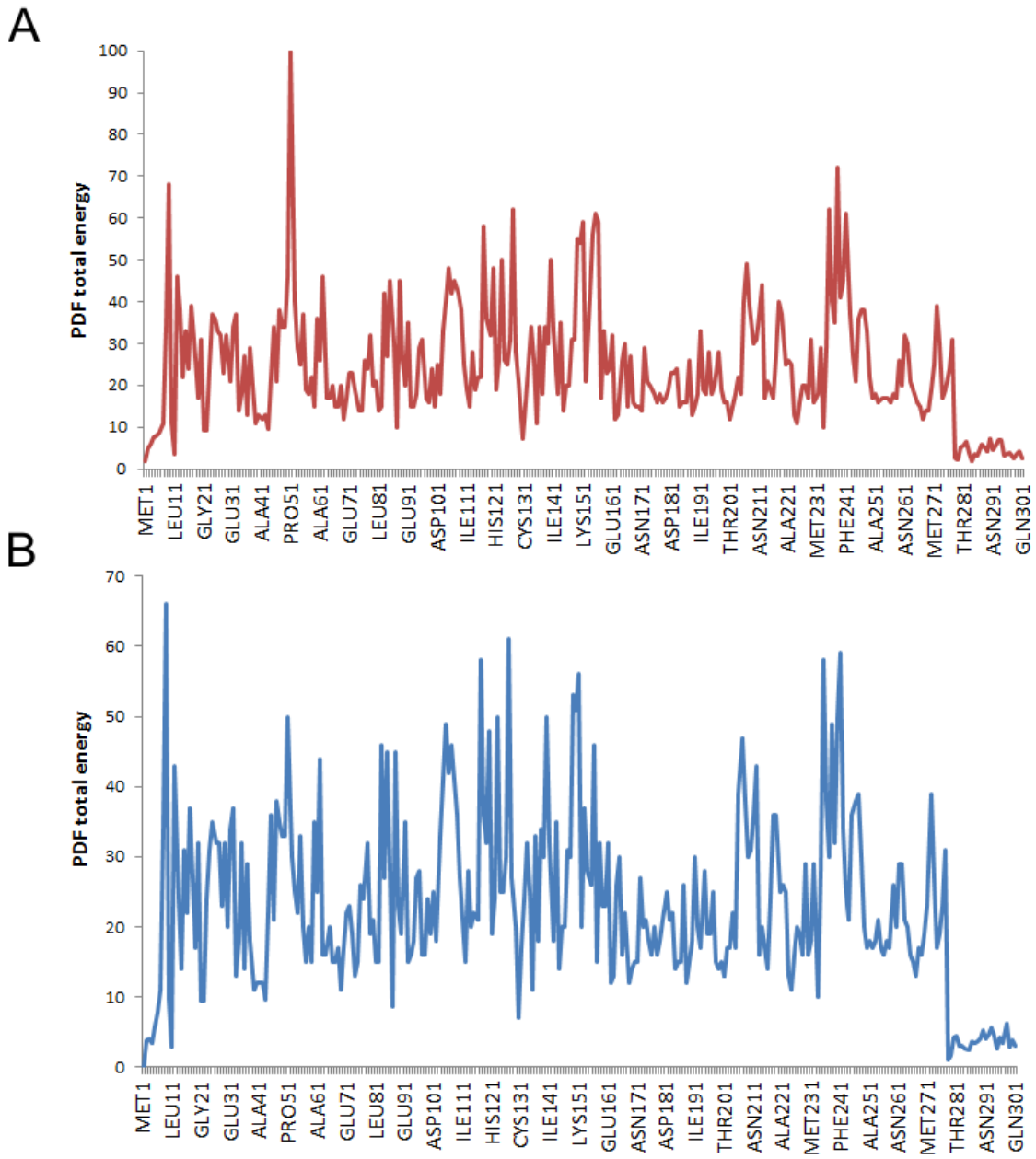
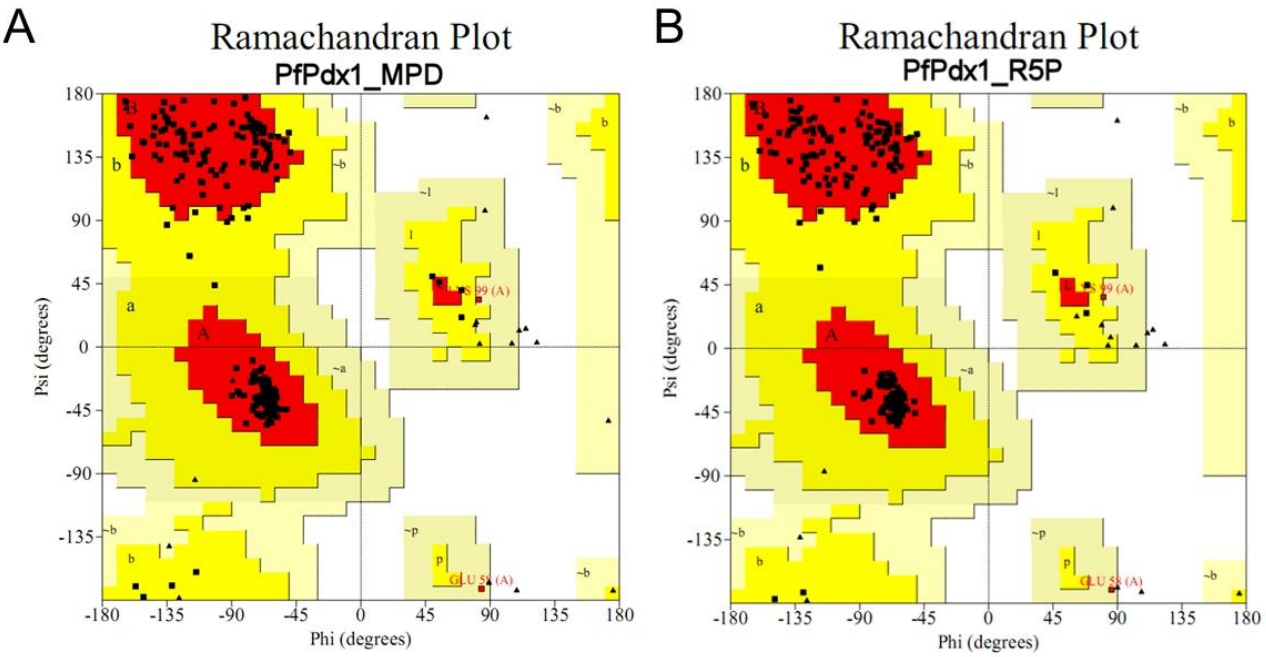


Figure 2.7: Line plot of the PDF total score in homology model (A) *Pf*Pdx1_MPD and (B) *Pf*Pdx1_Ru5P. Residues that had high PDF total scores include P51 and S238. The former P51 residue was calculated with high bond angle potential.

Moreover, Ramachandran plots were used to verify residue phi- (φ) and psi- (ψ) angle discrepancies. Ramachandran plots were generated for *Pf*Pdx1_MPD and *Pf*Pdx1_Ru5P (Figure 2.8), whereas the remaining models *Pf*Pdx1_MPDsco and *Pf*Pdx1_MPDw had similar plots and are not shown. The residue φ and ψ-angles were clustered into distinct regions in the plots, corresponding to secondary structural elements these occurred in [180]. There were 21 glycine residues and 9 proline residues in the *Pf*Pdx1 models. Glycine has greater backbone conformational flexibility due to the lack of side-chain, therefore allowed in additional regions of the plots, whereas

proline is rotationally constrained and limited to certain ψ-angles [180, 181]. No residues were found in disallowed regions with only two residues present in the generously allowed regions; E58 and K99 (Figure 2.8). The E58 and K99 residues were located on solvent exposed loop regions in *Pf*Pdx1, and were distant from the active site, and would not have influenced interpretation of the R5P-binding site pharmacophore.



	*Pf*Pdx1_MPD	*Pf*Pdx1_Ru5P
Residues in most favoured regions	92.9% (250)	94.4% (254)
Residues in additional allowed regions	6.3% (17)	4.8% (13)
Residues in generously allowed regions	0.7% (2)	0.7% (2)
Residues in disallowed regions	0.0% (0)	0.0% (0)

Figure 2.8: Ramachandran plots of *Pf*Pdx1_MPD (A) and *Pf*Pdx1_Ru5P (B). The plots represent the φ and ψ angle combinations of the residues in the *Pf*Pdx1 models. Areas coloured in red represent most favoured angle distribution regions, areas coloured in yellow and light yellow represent additionally and generously allowed angle distribution regions, respectively. Individual amino acid residues are indicated by black squares, or red squares when within generously allowed regions. Glycine residues are indicated as triangles. Values in the table represent the percentage of residues in respective regions with the number of residues in parenthesis.

The C_α RMSD's of *Pf*Pdx1 models prior to minimisation compared to the template structures revealed values of between 0.9 and 0.6. Å (Table 2.2). The comparative C_α RMSD value between our models and the later released *Pb*Pdx1 was 0.6 Å (Table 2.2).

Table 2.2: RMSD's of *Pf*Pdx1 homology models. The homology models were compared using the RCSB PDB comparison tool, using an established jCE algorithm [167]. The *Pf*Pdx1 models were compared to template and *P. berghei* Pdx1 structures.

	C_α RMSD compared to *Tm*PDx1 (2ISS) (Å)	C_α RMSD compared to *Tt*PDx1 (2ZBT) (Å)	C_α RMSD compared to *Pb*Pdx1 (4ADU) (Å)
***Pf*Pdx1_MPD**	0.38	0.87	0.68
***Pf*Pdx1_Ru5P**	0.27	0.92	0.64
***Pf*Pdx1_MPDw**	0.38	0.87	0.68
***Pf*Pdx1_MPDsco**	0.38	0.87	0.68

In addition the C_α and main chain atoms (N, C_α, C, O) RMSD values for the active site (within 5 Å radius of Ru5P/MPD) were calculated and compared to determine C_α backbone and side chain atom orientations (Table 2.3). The RMSD's of the *Pf*Pdx1 models revealed smaller deviations compared to *Tm*Pdx1 than to *Tt*Pdx1. The *Pf*Pdx1 active site residues were comparable to the *Pb*Pdx1 active site, with RMSD between 0.5 – 0.6 Å, indicating that good quality models were generated.

Table 2.3: RMSD's of residues surrounding the R5P-binding site ligand in *Pf*Pdx1 homology models. Un-minimised *Pf*Pdx1 models were compared to Pdx1 templates and *P. berghei* Pdx1 using residues surrounding the ligand in the active site. The C_α RMSD is reported and main chain atom RMSD is given in parenthesis.

	RMSD compared to *Tm*PDx1 (2ISS) (Å)	RMSD compared to *Tt*PDx1 (2ZBT) (Å)	RMSD compared to *Pb*Pdx1 (4ADU) (Å)
***Pf*Pdx1_MPD**	0.37 (0.36)	0.70 (0.73)	0.62 (0.66)
***Pf*Pdx1_Ru5P**	0.22 (0.24)	0.76 (0.76)	0.59 (0.64)
***Pf*Pdx1_MPDw**	0.37 (0.36)	0.70 (0.73)	0.62 (0.66)
***Pf*Pdx1_MPDsco**	0.37 (0.36)	0.70 (0.73)	0.62 (0.66)

R5P-binding site residues in the generated *Pf*Pdx1 homology models were spatially comparable to the *Tt*Pdx1 and *Tm*Pdx1 template structures (Table 2.3). In Figure 2.A1 of Appendix 2 residues aligned better with the *Tm*Pdx1 templates compared to *Tt*Pdx1 supporting the RMSD calculations in Table 2.3. The P52 residue of *Pf*Pdx1, which differed markedly between the two template structures, adopted a conformation similar to that of P51 in the *Tm*Pdx1 structure (Figure A2.2, Appendix 2). These results suggested that the high degree of sequence conservation between *Pf*Pdx1 and templates led to the creation of structurally similar homology models.

The overall topology of the *Pf*Pdx1 structure had a $(\alpha/\beta)_8$ barrel fold architecture with eight central parallel β-sheets alternated with by eight α-helices (Figure 2.8). This is in agreement with topologies of *Tm*Pdx1, *Bs*Pdx1 and *G. stearothermophilus* Pdx1 [119, 156, 157]. Additions to the $(\beta/\alpha)_8$ barrel structure include the N-terminal α-helix, αN, as well as α-helices α2a, α6a, α6b and

α8a located in between β/α pairs (Figure 2.8 A) [157]. A right handed 3_{10} helix, with characteristic 3-residue H-bonding pattern, was located between β8 and α8 (Figure 2.9 A).

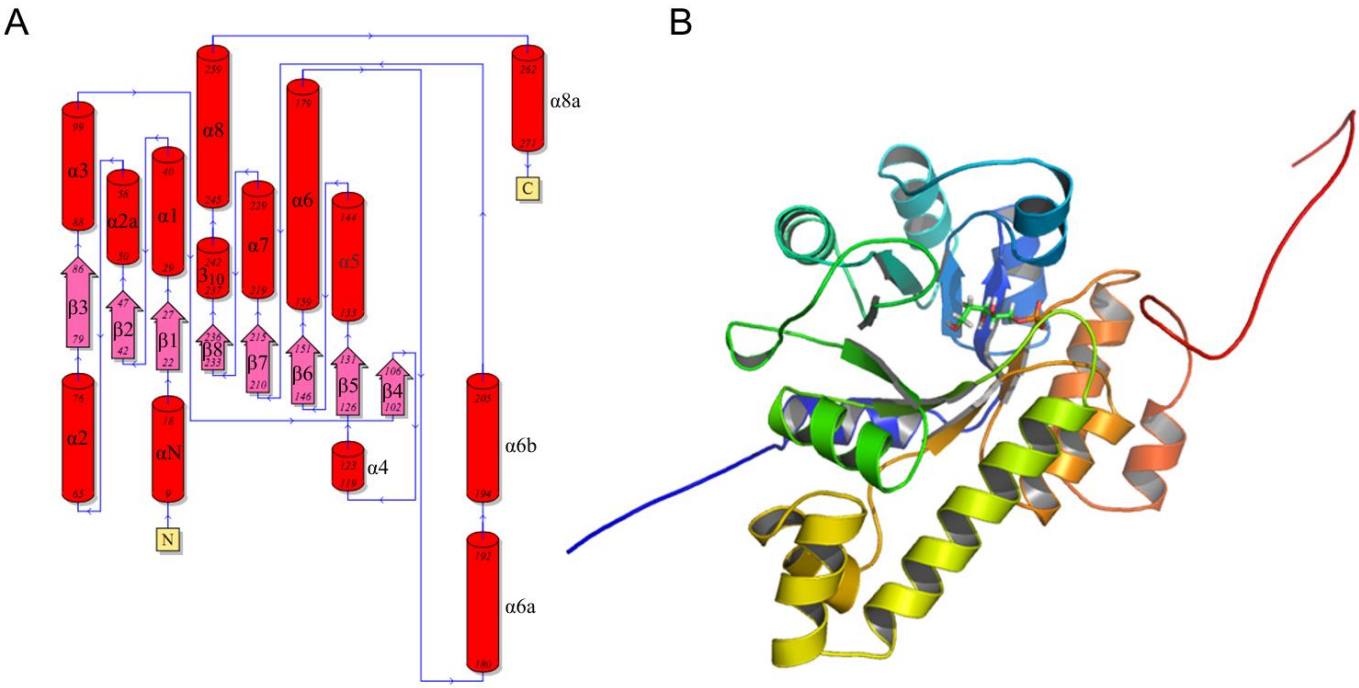


Figure 2.9: Topology and homology model of *Pf*Pdx1_Ru5P. A) The topology of *Pf*Pdx1 was comparable to *Tm*Pdx1 with eight central parallel β-sheets alternated with by eight α-helices. The topology diagram was generated using PDBsum [169, 170]. **B)** Homology model of *Pf*Pdx1 (model *Pf*Pdx1_Ru5P). The view is from the top down the β-barrel. The model is indicated with the N-terminal in blue and the C-terminal in red. The Ru5P ligand is shown as a stick line structure located in the β-barrel region.

The *Pb*Pdx1 was previously shown to complex with *Pf*Pdx2 resulting in the formation of chimeric PLP synthase dodecamers, which was shown to have comparable catalytic activities to *P. falciparum* PLP synthase [164]. The high degree of identity shared between the proteins (86%) therefore enabled *Pf*Pdx2 proteins to recognise similar structural elements in *Pb*Pdx1 required for dimerisation. The bacterial *Tm*Pdx1 has an additional β-strand on the N-terminal preceding αN which undergoes β completion with a central β-sheet in *Tm*Pdx2 upon association [157]. *Pb*Pdx1 was shown to lack this secondary structural element, and interactions between *Pb*Pdx1 and *Pf*Pdx2 predominated from αN region [164]. Superimposition of *Pf*Pdx1 homology models with *Pb*Pdx1 showed corresponding residues from αN in *Pf*Pdx1 could be involved with interactions *Pf*Pdx2 (Figure 2.10). The *Pf*Pdx1 homology models lacked the N-terminal β-strand. Residues from the *Pb*Pdx1 αN-region, identical to corresponding residues in *Pf*Pdx1, which form part of the interface with *Pf*Pdx2 include M19 (in *Pb*Pdx1) [164]. This residue in *Pf*Pdx1 (M18) is therefore similarly expected to be involved in contacts with *Pf*Pdx2 (Figure 2.10). *Pf*Pdx1 has 20 additional amino acids in the C-terminal region, which could not be aligned with any template structures, and this

region remained unstructured in the models (Figure 2.10). The *Pb*Pdx1 structure similarly had an extended C-terminal region, which did not adopt any secondary structural arrangement (Figure 2.10).

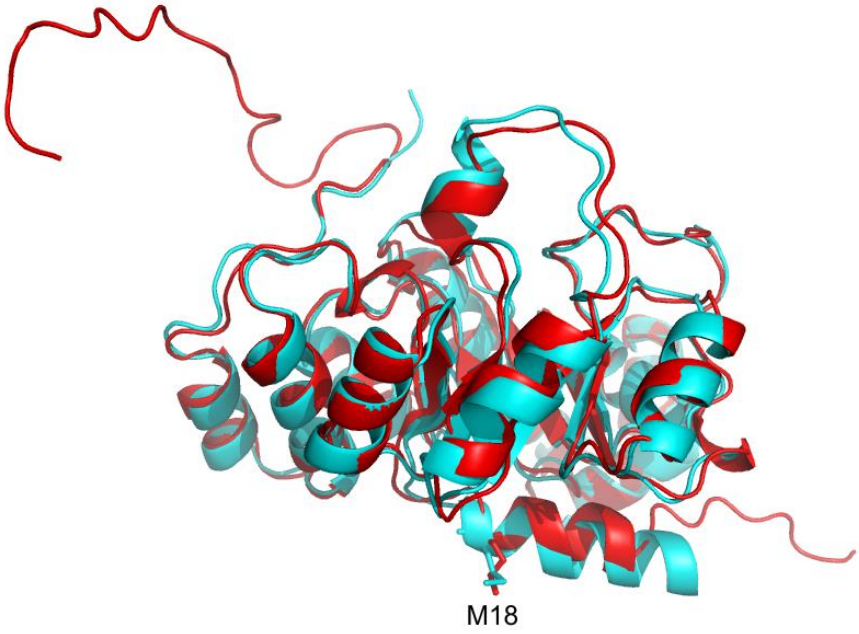


Figure 2.10: Overlay of *Pf*Pdx1 with *Pb*Pdx1. Of the 13 residues in the N-terminal α-helix of *Pf*Pdx1 (red), located in the lower part of the figure, were conserved and adopted similar orientations to *Pb*Pdx1 (cyan). The conserved methionine residue (M18 in *Pf*Pdx1**)** proposed to be involved in contacts with the *Pf*Pdx2 monomer in the assembled heterodimer, is indicated [164]. The unstructured C-terminal tails of *Pf*Pdx1 and *Pb*Pdx1 are shown on the left side of the figure**.**

2.3.3 Protein minimisation

Protein minimisation was performed to geometrically optimise the protein models, and remove steric clashes introduced by satisfaction of spatial restraints during superimposition model building by superimposition of φ and ψ angles of the template structures. The *Pf*Pdx1_MPDsco model was subjected to side chain optimisation to resolve steric clashes prior to protein minimisation, and had a lower initial potential energy compared to other models (Table 2.4). Comparing the final potential energies from the final structures with those of the initial structure indicated that some molecular geometric clashes and atomic overlaps were resolved during minimisation (Table 2.4). The final potential energy of the models was between -18800 and -19023 kcal/mol (Table 2.4).

Table 2.4: Properties of minimised *Pf*Pdx1 homology models.

	Initial potential energy (kcal/mol)	Final potential energy (kcal/mol)	Van der Waals energy (kcal/mol)	Electrostatic energy (kcal/mol)	Initial RMS gradient (kcal/(mol x Angstrom)	Final RMS gradient (kcal/(mol x Angstrom)
***Pf*Pdx1_MPD**	14682.3	-18824.6	-1941.5	-19347.3	1748.3	1.28
***Pf*Pdx1_Ru5P**	67073.7	-19358.4	-1923.7	-19822.6	8540.9	1.19
***Pf*Pdx1_MPDw**	148887.6	-19023.2	-1913.7	-19583.9	20231.9	1.23
***Pf*Pdx1_MPDsco**	-10065.7	-18816.6	-1884.9	-19420.8	12.0	1.28

Protein secondary structures of the minimised *Pf*Pdx1 homology model (Figure 2.11) were comparable to *Tm*Pdx1 (Figure 2.4). Secondary structures of other homology models are not shown, as these were identical to that of the *Pf*Pdx1_MPD model. In the assembled heterodimer, the N-terminal region of Pdx1 proteins spanning the first 20 amino acids interacts with Pdx2 and plays an important role in stimulating Pdx2 activity. This region, consisting of residues 8 – 16 form the αN α-helix, becomes ordered (adopts secondary structural arrangement) in *Tm*Pdx1 only in the presence of *Tm*Pdx2 monomers, [157]. The first 4 amino acids of the N-terminal region amongst different *Plasmodium* Pdx1 homologs have been shown to be variable, with structural evidence showing that the only ordered region in the N-terminal region of the *P. berghei* Pdx1 crystal structure consists of an α-helix spanning from residues D8 to M19 [158, 164]. Residues L10 – M18 in minimised *Pf*Pdx1 homology models adopted an α-helical arrangement (Figure 2.11 and Figure 2.12) similar to *Pb*Pdx1 structures.

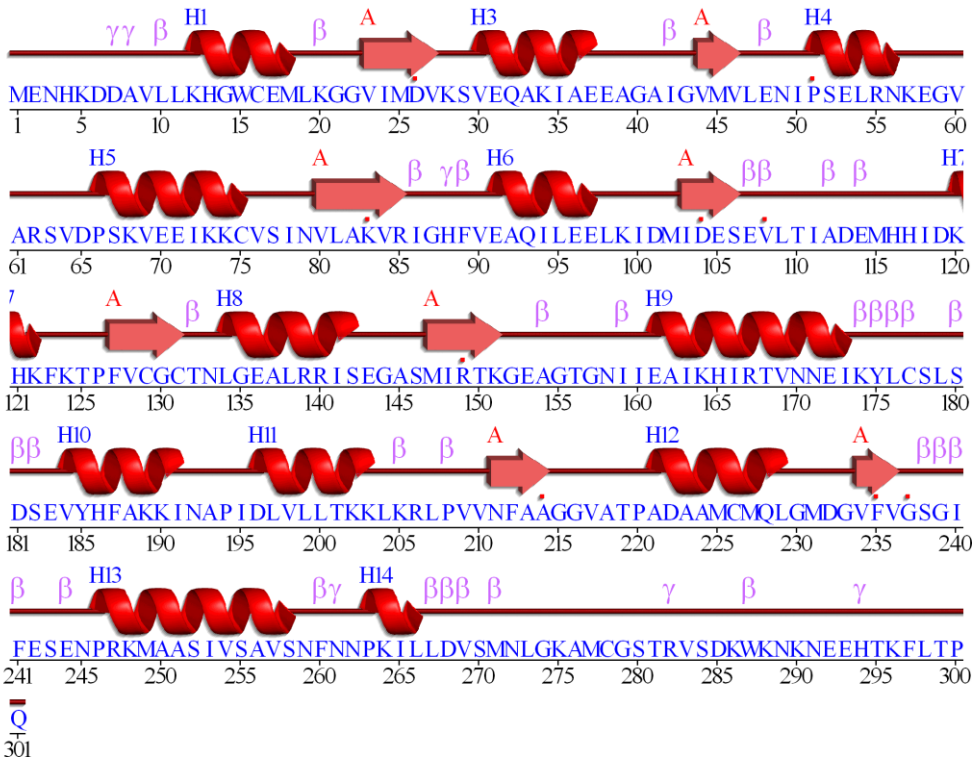


Figure 2.11: Secondary structure diagram of minimised *Pf*Pdx1_MPD. Predicted protein secondary structures were generated in PDBsum [169, 170]. Secondary structures such as α-helices are labeled with H, whereas β-sheet are labeled with A. β- and γ-turns are indicated by β and γ respectively.

The overall arrangement of secondary structural elements of the minimised *Pf*Pdx1 model was similar to *Tm*Pdx1 (Figure 2.12). Comparison to the *Tm*Pdx1 showed similar secondary structural motifs were conserved. Some difference in the *Pf*Pdx1_Ru5P model included shortening of the central β3 and β4 sheets (Figure 2.12), which was not observed in the remaining *Pf*Pdx1_MPD models (data not shown). The β3 regions of *Pf*Pdx1 is composed of residues identical to that of *Pb*Pdx1 which was shown to form a β-sheet from residue N80 to V85 [164]. These differences in secondary structure of this particular *Pf*Pdx1 model were related to the algorithm used to predict secondary structure during model creation. RMSD comparison (see below) suggested that there was a good correlation between minimised *Pf*Pdx1 models and the *Tm*Pdx1 and *Pb*Pdx1 structures.

Previous deletion mutagenesis of the C-terminal tail of *Pf*Pdx1 showed that residues 270-279 were essential for the catatlyitc activity of *Pf*Pdx1 and influenced R5P binding [182]. *Pf*Pdx1 models were aligned with the chimeric *Pb*Pdx1 structure in complex with *Pf*Pdx2 (Appendix Fig A2.3). The C-terminal tail of *Pf*Pdx1 models from residue M271 to T281 were shown to correlated with those from the *Pb*Pdx1 structure. Residues in close proximity to the Ru5P molecule included M277, which was located 5.27 Å from the phosphate groups of the molecule.

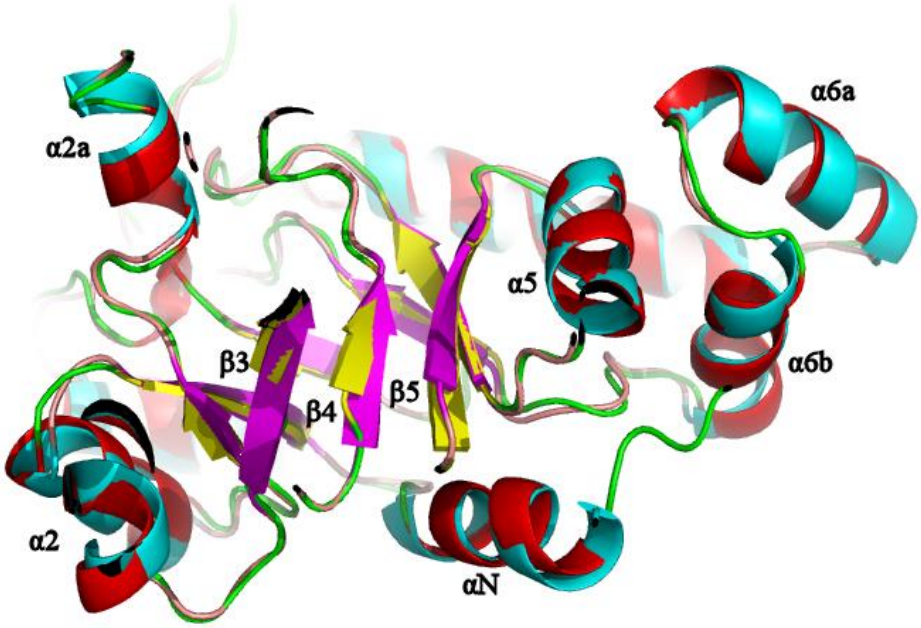


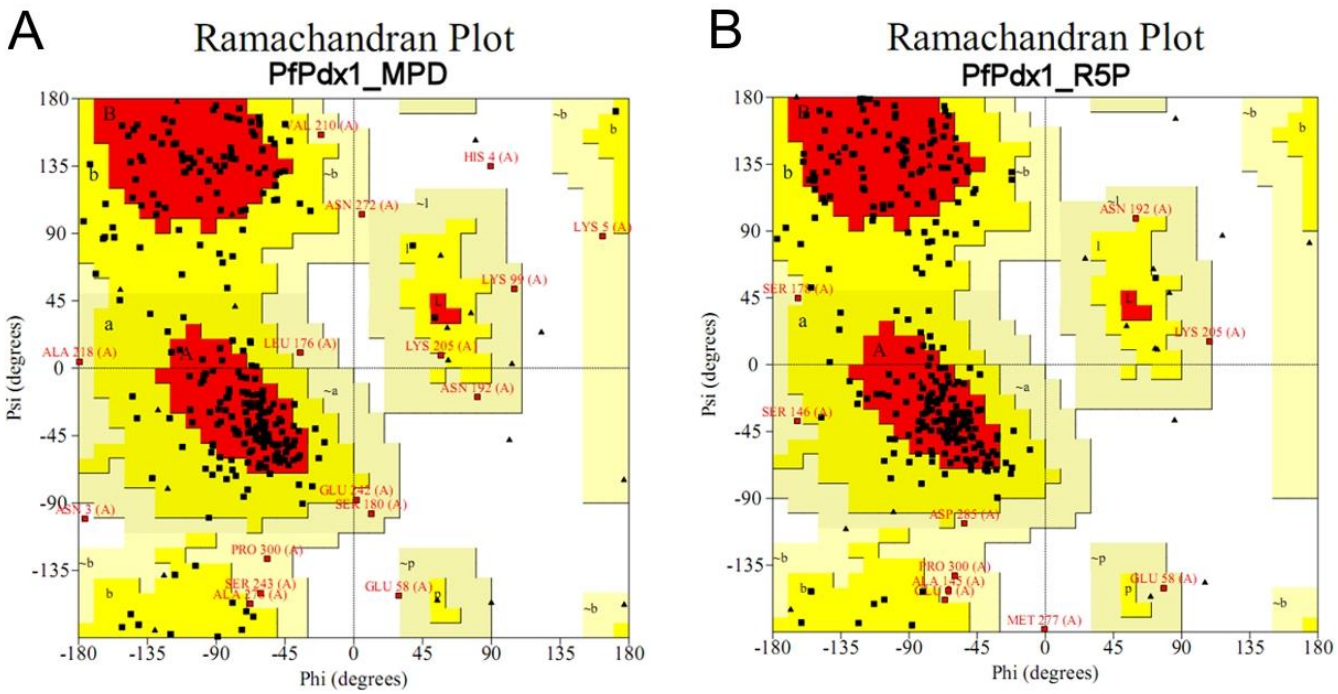
Figure 2.12: Side view of *Pf*Pdx1_Ru5P superimposed on the *Tm*Pdx1 template structure. *Pf*Pdx1_Ru5P model (α-helices indicated in red, β-sheets indicated in yellow) was overlaid on the *Tm*Pdx1 (α-helices indicated in cyan, β-sheets indicated in magenta). For ease of presentation the helix α3 was omitted from the image.

RMSD comparisons of the minimised models to reported structures and templates showed a good correlation (Table 2.5). The deviations of minimised models were slightly increased from unminimised models, however were still below 1.0 Å for the *Tm*Pdx1 template and *Pb*Pdx1 structure.

Table 2.5: RMSD's of minimised *Pf*Pdx1 models. The C_α RMSD's of *Pf*Pdx1 models were compared to template and *P. berghei* Pdx1 structures using the RCSB PDB comparison tool [167] with the jCE algorithm.

	C_α RMSD compared to *Tm*PDx1 (2ISS) (Å)	**C_α RMSD compared to *Tt*PDx1 (2ZBT) (Å)**	**C_α RMSD compared to *Pb*Pdx1 (4ADU) (Å)**
*Pf*Pdx1_MPD	0.79	1.09	0.96
*Pf*Pdx1_Ru5P	0.74	1.20	0.93
*Pf*Pdx1_MPDw	0.77	1.08	0.93
*Pf*Pdx1_MPDsco	0.80	1.14	0.98

Ramachandran plots after minimisation of *Pf*Pdx1 homology models revealed a greater percentage of residues within generously allowed regions, as well as three residues having φ and ψ angles in disallowed regions (Figure 2.13). Ramachandran plots of additional homology models are given in the appendix (Appendix 2, Figure 2.A4).



	*Pf*Pdx1_MPD	*Pf*Pdx1_Ru5P	*Pf*Pdx1_MPDsco	*Pf*Pdx1_MPDw
Residues in most favoured regions	66.9% (180)	69.5% (187)	65.8% (177)	67.7% (182)
Residues in additional allowed regions	27.5% (74)	27.1% (73)	28.6% (77)	28.3% (76)
Residues in generously allowed regions	4.5% (12)	3.0% (8)	4.5% (12)	3.0% (8)
Residues in disallowed regions	1.1% (3)	0.4% (1)	1.1% (3)	1.1% (3)

Figure 2.13: Ramachandran plots of *Pf*Pdx1 homology models after protein minimisation. Values represent percentages with the number of residues in parenthesis. The plots of **A)** *Pf*Pdx1_MPD and **B)** *Pf*Pdx1_Ru5P models represent the φ and ψ angle combinations of the residues in the *Pf*Pdx1 models. Areas coloured in red represent most favoured angle distribution regions, areas coloured in yellow and light yellow represent additionally and generously allowed angle distribution regions, respectively. Individual amino acid residues are indicated by black squares, or red squares when within generously allowed or disallowed regions. Glycine residues are indicated by triangles.

Some residues of *Pf*Pdx1 models in generously allowed regions of the Ramachandran plots were part of disordered N- and C-terminal tails. These were not removed from models and therefore biased the Ramachandran plots. Of the residues in disallowed regions most occurred on the disordered N-terminal surface exposed region, such as N3 and H4, or the C-terminal tail, such as M277 (Figure 2.13, also Appendix Figure A2.3). The E58 residue, positioned on a surface exposed loop region between α2a and α2, could not be minimised into alternative φ and ψ angles, and was also characterised into a disallowed region. Residues in surface exposed regions have higher propensities to exhibit conformations which result in characterization into disallowed

Ramachandran plot regions [183]. After energy minimisation, the orientation of residues in the active site was compared using RMSD calculations. The RMSD's of residues located in the R5P-binding active site of *Pf*Pdx1 were below 1.0 Å compared to the *Tm*Pdx1 and *Pb*Pdx1, suggesting conservation of the relative residue positioning (Table 2.6).

Table 2.6: RMSD's of the R5P-binding site of minimised *Pf*Pdx1 models. The *Pf*Pdx1 homology models were compared to templates and *Pb*Pdx1 using residues surrounding the ligand in the active site (5 Å radius from ligand). C_α RMSD is reported and main chain atoms RMSD given in parenthesis.

	RMSD compared to *Tm*PDx1 (2ISS) (Å)	**RMSD compared to *Tt*PDx1 (2ZBT) (Å)**	**RMSD compared to *Pb*Pdx1 (4ADU) (Å)**
*Pf*Pdx1_MPD	0.75 (0.83)	1.00 (1.08)	0.87 (0.96)
*Pf*Pdx1_Ru5P	0.70 (0.77)	1.06 (1.12)	0.87 (0.96)
*Pf*Pdx1_MPDw	0.76 (0.85)	1.00 (1.08)	0.86 (0.95)
*Pf*Pdx1_MPDsco	0.77 (0.87)	1.08 (1.14)	0.89 (0.99)

Residues in close proximity to the Ru5P ligand in *Pf*Pdx1 superimposed and aligned well with equivalent residues in the *Tm*Pdx1 structure (Figure 2.14 A). Spatial positioning of D26, K83, D104 and R149, located near the carbonyl component of the Ru5P ligand, were maintained (Figure 2.14 A). Residues involved in stabilization of the phosphate groups, such as G216, G237 and G155 (not shown) were comparable and aligned with each other (Figure 2.14 A). This implied that the R5P-binding site of the *Pf*Pdx1 homology models were comparable to that of *Tm*Pdx1, and that similar residues are involved in stabilization of the substrate in *Pf*Pdx1 [157]. The R148 residue from *Tm*Pdx1 H-bonds with the C1 hydroxyl group of Ru5P [157]. The equivalent residue in *Pf*Pdx1, R149, had a similar spatial positioning, and could therefore be involved in H-bonds with the R5P terminal carbonyl group (Figure 2.14 A).

The P51 residue in *Pf*Pdx1, which was shown to differ in spatial location between the *Tm*Pdx1 and *Tt*Pdx1 template structures, had a similar positioning to that of *Tm*Pdx1 (Figure 2.14 A). The equivalent P50 residue in *Tm*Pdx1 had non-bonded contacts with the ligand and was predicted to contribute in hydrophobic stabilisation to Ru5P [157]. Similar roles could be assigned to the P51 residue in the *Pf*Pdx1 model. Superimposition of *Pf*Pdx1 with *Pb*Pdx1 had comparable orientations of residues involved with contacts to the R5P substrate (Figure 2.14 B). The orientation of side chain atoms of the M45 residue in *Pf*Pdx1 was different compared to both the *Tm*Pdx1 template and *Pb*Pdx1 structures, in which the sulphur atom was rotated 180° towards the ligand. This residue has been predicted to be involved in hydrophobic contacts towards the Ru5P ligand in *Tm*Pdx1 [157]. The particular orientation the M45 residue in the *Pf*Pdx1 models, with methyl groups in

approximately similar positions compared to *Tm*Pdx1 and *Pb*Pdx1 structures, is still expected to promote hydrophobic interactions with the Ru5P substrate.

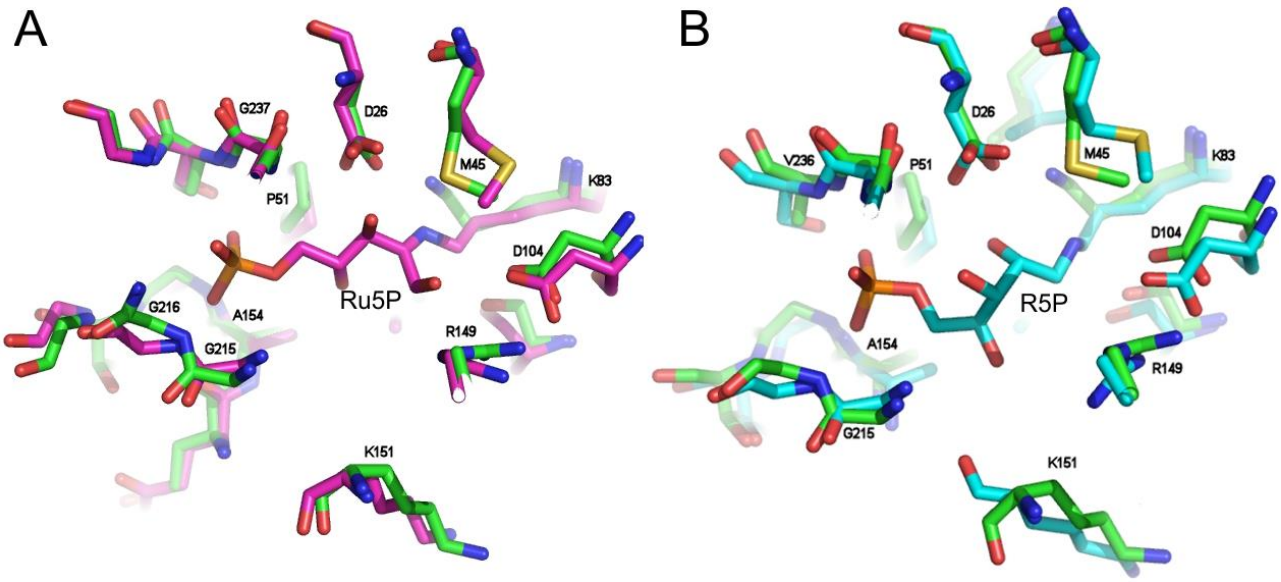


Figure 2.14: The R5P-binding site of *Pf*Pdx1 homology models. A) The *Pf*Pdx1_Ru5P model (green) was compared to *Tm*Pdx1 (magenta). The main chain RMSD of the active site residues was 0.77 Å suggesting good correlation of residue positioning. In contrast to *Tm*Pdx1 the M45 residue of *Pf*PDx1 was orientated with the sulphur atom rotated towards the Ru5P substrate, with the methyl group adopting similar positioning to that of the *Tm*Pdx1 structure. Residue labels are that of the *Pf*Pdx1 model. The Ru5P substrate from *Tm*Pdx1 is shown covalently bonded to K82 (*Tm*Pdx1 numbering) in the centre of the figure. **B)** The *Pf*Pdx1_Ru5P model (green) superimposed with *Pb*Pdx1 (cyan). Residues were labelled according to the *Pf*Pdx1 structure. The R5P ligand from *Pb*Pdx1 is shown in the centre of the figure.

2.3.4 Interaction generation and pharmacophore extraction

Pharmacophore chemical features were extracted from the R5P binding site of individual PfPdx1 homology models using the interaction generation protocol in DS. Ideally a pharmacophore should not consist of more than 7 features to make computational library searching practical and less resource intensive. The chemical features identified in the *Pf*Pdx1 active site were hierarchically clustered in order to generate a representative pharmacophore hypothesis. Of the chemical features identified, those in close proximity to the ligand (8 Å) were selectively retained during hierarchical clustering. The procedure identified pharmacophore H-bond donors, H-bond acceptors and hydrophobic features (Table 2.7). The individual pharmacophore features for each model are shown as stereo images (Figure 2.15, as well as Figure 2.A5 – Figure 2.A7 in the Appendix 2). The chemical features from the respective pharmacophores included residues known to have interactions with the Ru5P (in *Tm*Pdx1) and R5P substrates (in *Pb*Pdx1). Correlating chemical features were identified from the respective pharmacophore hypotheses, such as the side chain of K83 as H-bond donor in all except the custom assigned pharmacophore. A few pharmacophores contained chemical

features associated with residues which were not present in other pharmacophores, such as E107 in the *Pf*Pdx1_MPD pharmacophore which was flagged as H-bond donor feature, and is absent from other models. This suggested there were five different and diverse pharmacophore hypotheses generated, each representing a different set of chemical features of the R5P active site in *Pf*Pdx1.

From the *Pf*Pdx1_MPD structure (Figure 2.15 A) several H-bond donors were identified, and these include the NZ atom of K83, the main chain N-atoms of both E107 and G216. The side chain carbonyl O-atom of D26 and the main chain carbonyl O-atoms of A154 and A214 were identified as H-bond acceptors (Figure 2.15 A). Additionally, two hydrophobic features were identified in the central part of the cavity, which normally contains the ligand, located near E107. In the *Pf*Pdx1_Ru5P model the K83 as well as D26 was identified as H-bond donor and acceptor features, respectively, and were shared features amongst the models. Additionally side chain N-atoms from R149 as well as main chain N-atoms from G152 and A154 were recognised as H-bond donors.

Table 2.7: Summary of pharmacophore features identified from *Pf*Pdx1 homology models. Specific residues and atoms (in parentheses) involved in either H-bond acceptor or donor interactions as well as residues located near hydrophobic features are indicated. A custom pharmacophore was also built, in which pharmacophore features were manually assigned and aligned to residues in the *Pf*Pdx1_Ru5P model. Pharmacophores involving water atoms are labelled "(W)".

Model	Residues identified as H-bond donors (atoms)	Residues identified as H-bond acceptors (atoms)	Residues flagged near hydrophobic features
*Pf*Pdx1_MPD	K83 (NZ) E107 (N) G216 (N)	D26 (OD) A154 (O) A214 (O)	Central cavity region, located near E107.
*Pf*Pdx1_Ru5P	K83 (NZ) R149 (NE) G152 (N) A154 (N)	D26 (OD) A214 (O)	Central cavity region. Involving residues V108 and A154.
*Pf*Pdx1_MPDw	K83 (NZ) G216 (N) F235 - (W) - G216	D26 (OD) I24 - (W) - D26 V108 - (W) - R149	Central cavity region.
*Pf*Pdx1_MPDsco	K83 (NZ) A154 (N) G216 (N)	E107 (OE) F235 (O)	Close to V108, near to G152. Second feature near M45 and G237.
Custom assigned pharmacophore	R149 (NE) G155 (N) G237 (N) S238 (N)	D26 (OD) D104 (OD)	-

Consensus atoms involved in H-bond donor interactions identified from the four models include the side chain NZ-atom of K83, as well as main chain N-atoms from both R149 and G216. Consensus atoms predicted as H-bond acceptors include the side chain carboxyl O-atom of D26, and main chain carbonyl O-atoms of A154 and A214 (Table 2.7). Three water molecules in the *Pf*Pdx1_MPDw model were predicted with pharmacophore features. A water molecule predicted to

H-bond with both F235 and G216 was flagged as an H-bond donor, whereas water molecules located between I24 and D26, as well as between V108 and R149 were flagged with H-bond acceptor attributes (Figure 2.A5. In addition to pharmacophore features identified exclusion constraints were added to restrict mapping of ligands into areas containing amino acids which were not part of the pharmacophore. The exclusion constraints were assigned to all amino acids outside a 5 Å radius of the identified pharmacophore. These exclusion constraints impose geographical restraints which restrict ligands from binding outside the R5P-active site region.

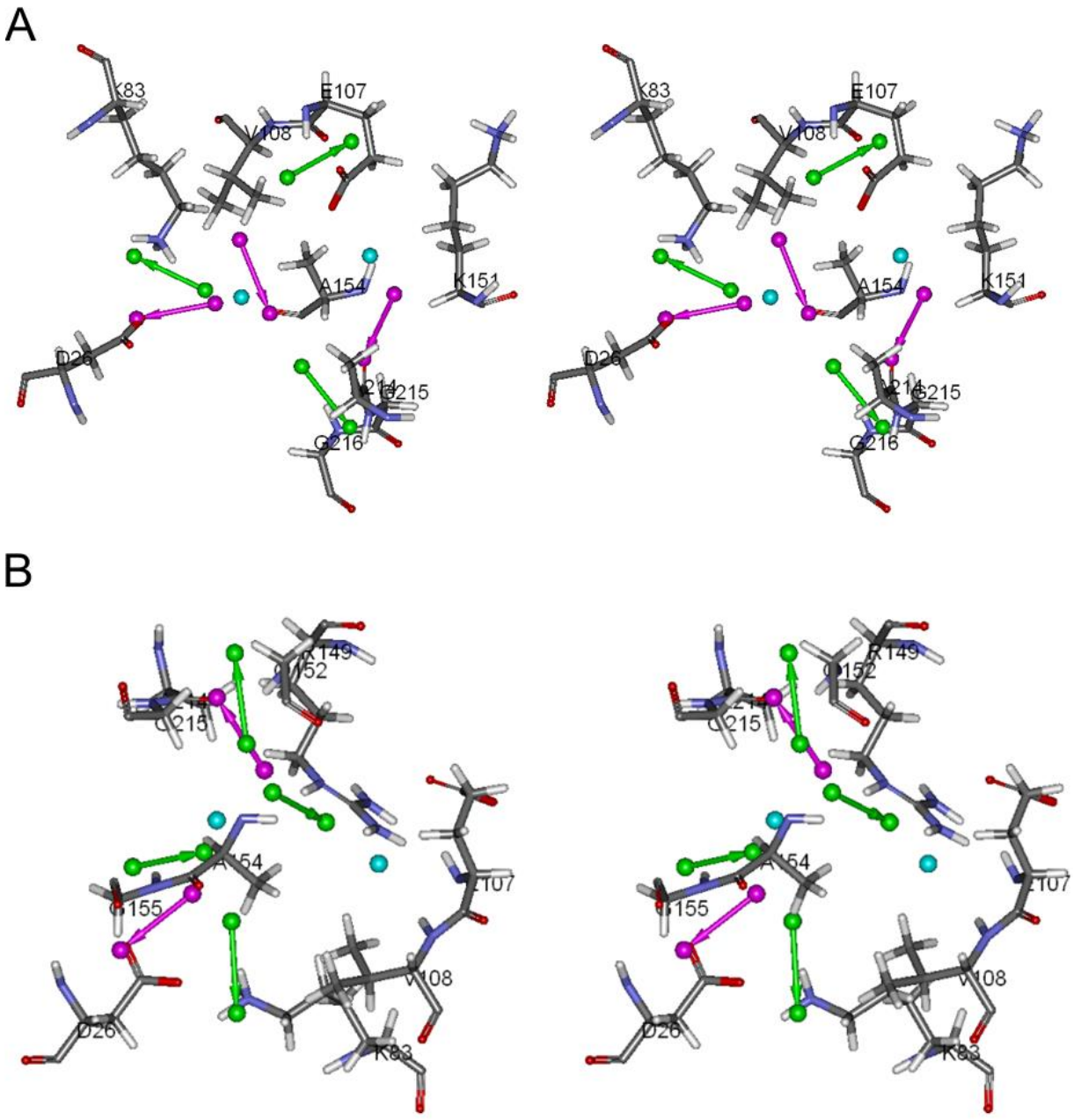


Figure 2.15: Stereo view of the pharmacophores extracted from the *Pf*Pdx1 homology models. Pharmacophores from the **A)** *Pf*Pdx1_MPD model and the **B)** *Pf*Pdx1_Ru5P model are shown with chemical features such as H-bond donors (green), H-bond acceptors (purple) and hydrophobic features (cyan) indicated.

2.3.5 Database searching and ligand pharmacophore mapping (LPM)

A database consisting of around 2.1 million commercially available compounds from the Zinc database drug-like subset [174] (zinc.docking.org) were searched using the individual pharmacophores generated from *Pf*Pdx1 models. The drug-like database consisted of ligands (molecules) pre-filtered using Lipinski's rule of five [184]. These rules or physiochemical properties such as; molecular weight ≤ 500, log *P* ≤ 5, H-bond donors ≤ 5 and H-bond acceptors ≤ 10, are associated with compounds which have good oral bioavailability or are "drug-like" [185]. The drug-like subset database was previously generated as a multi-conformer composite database using catDB with a maximum of 250 representative conformations (Accelrys Software Inc.) [186]. The drug-like subset database was screened against 5 pharmacophore models and 300 best-hit compounds were identified based on their pharmacophore fit-values. The database searching approach was employed as pre-selection method for compounds that match the pharmacophores. The 7028 ligands identified from the searching procedure (results not shown) were further refined by ranking the pharmacophore fit-values using ligand pharmacophore mapping (LPM). In this procedure, the ligands are strictly aligned to the respective pharmacophore, and a score is assigned for each ligand based upon corresponding distances from pharmacophore features. LPM identified a total of 3063 ligands which satisfied the pharmacophore match constraints from the five pharmacophores (Figure 2.16). The *Pf*Pdx1_MPDsco model pharmacophore had the most hit ligands, whereas the *Pf*Pdx1_MPD model pharmacophore had the least (Figure 2.16). A comparison of ligands identified from the models showed that the *Pf*Pdx1_MPDw and *Pf*Pdx1_MPDsco had around 217 shared hits (Figure 2.16). Only one molecule was shared in all of the pharmacophores (Figure 2.16).

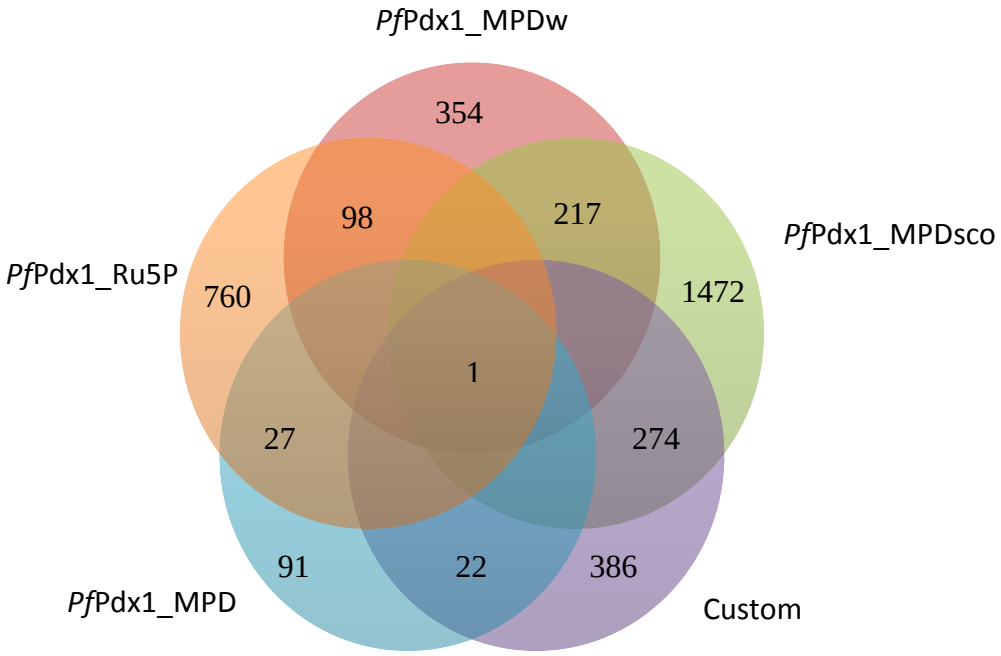


Figure 2.16: Venn diagram representing the number of ligands identified from the respective *Pf*Pdx1 pharmacophores. Common or shared hits are shown in overlapping regions and for simplicity, comparisons were made linearly with adjacent pharmacophores. The ZINC03957744 compound was identified as a hit in all five pharmacophore models.

Compounds were rigorously filtered based on their interaction with the pharmacophore in which at least n – 1 corresponding chemical groups had to align to a pharmacophore of n total features. This weighting during pharmacophore mapping ensured that chemical features of compounds that matched pharmacophore features were scored in the same manner. The sulphonamide ligand ZINC03560463 was found to satisfy corresponding pharmacophore features in the *Pf*Pdx1_Ru5P pharmacophore (Figure 2.17). The sulphonamide ligand ZINC03560463 was predicted to fit in four of the five pharmacophore models, and had the best fit-score in the *Pf*Pdx1_Ru5P model (Table 2.8). The top 10 best-scoring ligands identified during LPM based on their combined or summed fit values from the individual pharmacophores are shown in Table 2.8. The fit values of the top 100 ligands ranged from 8.45 to 3.66 (complete list not shown).

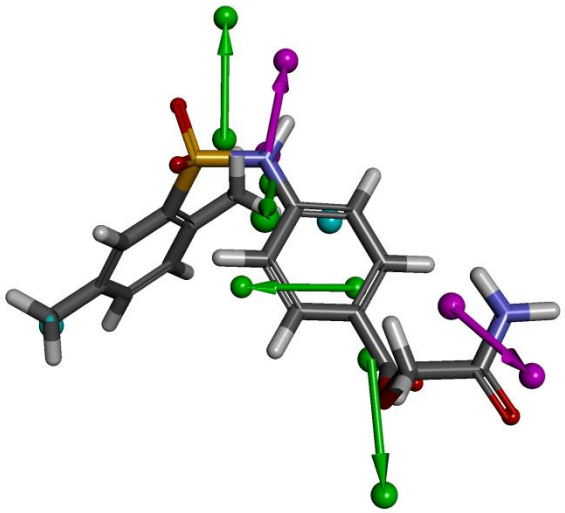
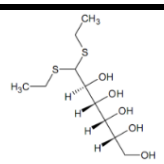
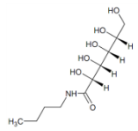
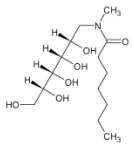
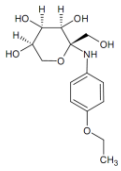
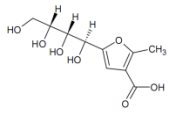
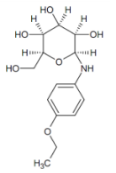
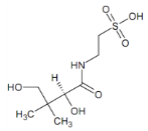
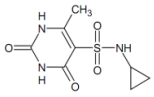
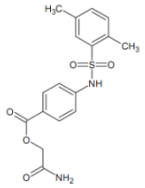
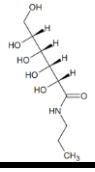


Figure 2.17: Complementary chemical features of ZINC03560463 that matched the *Pf*Pdx1_Ru5P pharmacophore. Only complementary features to the pharmacophore were matched, for example the amine substituent on the right side of the molecule, which is a H-bond donor, matched the H-bond acceptor feature (purple) of the pharmacophore. Similarly, the bottom H-bond donor feature (green) was matched by alignment of the O-atom of the ester group in the molecule.

Chemically diverse ligands were identified, however several of the best scoring ligands shared polyhydroxyl substituent characteristics or carbohydrate sugar backbones (Table 2.8). Compounds that did not match particular pharmacophore features in the respective models had a zero fit score. The representative summed fit-value accounted for mismatches in the pharmacophore fits, and enabled an average representation of ligand matching in the different PfPdx1 models. Other ligands with comparable fit-values had aromatic ring substituents, with sugar-like moieties (Table 2.8). Of the 3063 ligands ZINC01586791 had the best sum fit-value from all the pharmacophores.

Table 2.8: Best 10 ligands identified using LPM of *Pf*Pdx1 pharmacophores. The scores of the ligands from each pharmacophore were combined as a sum fit-value.

Name	Structure	Fit value in pharmacophore					Sum fit-value
		PfPdx1_MPDw	PfPdx1_Ru5P	PfPdx1_MPD	PfPdx1_MPDsco	Custom	
ZINC01586791		3.07	1.70	0.00	0.84	2.84	8.45
ZINC04533078		3.82	0.25	0.00	0.00	3.36	7.43
ZINC04546449		2.42	1.28	0.00	3.02	0.68	7.39
ZINC04773149		2.57	0.00	0.00	1.87	2.07	6.51
ZINC04206966		1.51	0.00	0.00	3.03	1.49	6.04
ZINC04353125		1.33	2.06	0.00	1.65	0.91	5.95
ZINC01666581		2.43	0.00	0.00	0.00	3.17	5.60
ZINC00093092		2.85	0.00	0.00	0.00	2.72	5.57
ZINC03560463		0.55	2.98	0.00	2.00	0.01	5.54
ZINC04533049		3.78	0.92	0.00	0.00	0.77	5.46

ZINC01586791 had the best combined fit score of all the pharmacophores (Figure 2.18 A). The hydroxyl substituents in the ligand were predicted to satisfy corresponding H-bond donor- and acceptor-interactions in the PfPdx1_MPDw (Figure 2.18 A). Similarly, hydroxyl substituents of the ZINC04206966 ligand mapped in the PfPdx1_MPDsco pharmacophore were identified based on both H-bond acceptor and donor properties (Figure 2.18 B). The molecule ranked fifth-best and had a combined fit score of 6.04 of all the pharmacophore models (Table 2.8).

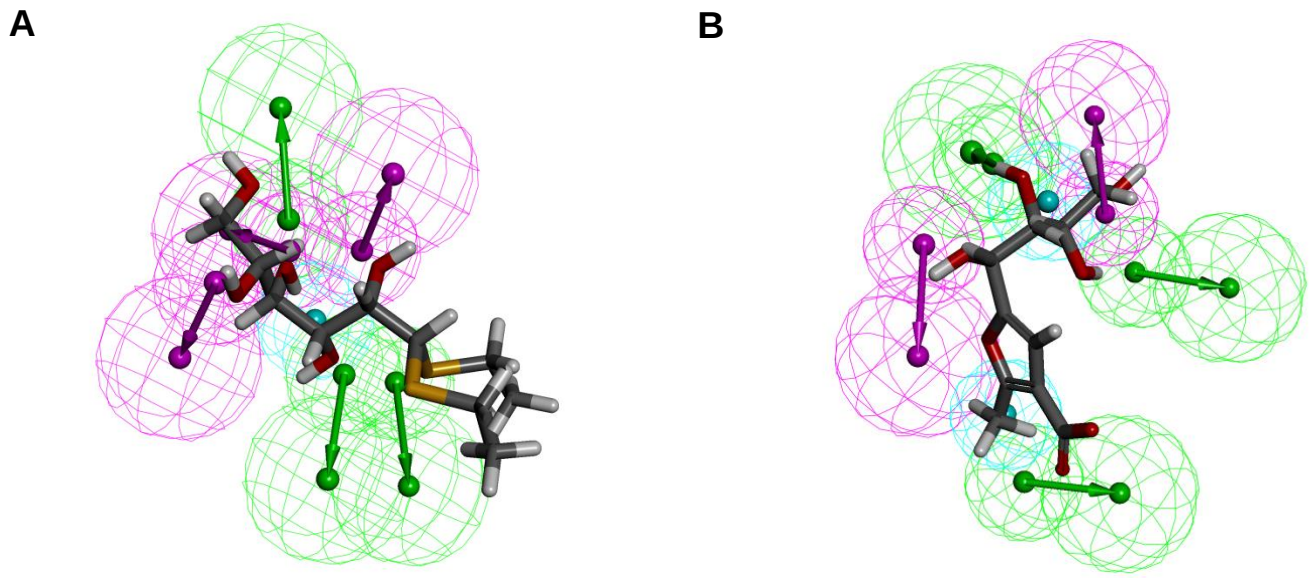


Figure 2.18: Ligands mapped into the *Pf*Pdx1 pharmacophore models using LPM. A) ZINC01586791 mapped in the *Pf*Pdx1_MPDw pharmacophore. The molecule was mapped into the pharmacophore in which 6 corresponding chemical features could be matched out of the 7 pharmacophore features present. A hydrophobic region matching the hydrophobic feature (cyan) was not satisfied. Hydroxyl substituent groups were found to match both H-bond donor (green) and acceptor (purple) features, in which the O-atom could act as an H-bond acceptor in turn matching the respective H-bond acceptor feature area in the pharmacophore, and vice versa for the H-atoms matching H-bond donor feature areas. **B)** ZINC04206966 mapped in the *Pf*Pdx1_MPDsco pharmacophore. The molecule satisfied 6 of the 7 pharmacophore features, of which the top hydrophobic feature (cyan) was not matched by a corresponding chemical moiety.

Chemical diversity of the 50 best-fitting compounds was assessed by pairwise similarity clustering using the online Chemmine tool. Clustering revealed the best 50 ligands from pharmacophore mapping could be grouped into distinctive groups based on their overall structure (Figure 2.19). Blocked regions of intense red represented molecules with highly similar chemical arrangements, whereas compounds with lesser degrees of similarity are indicated by shades of yellow to white for the least similar compounds (Figure 2.19). The overall distributions of chemical similarly suggested clusters of related compounds were identified from the *Pf*Pdx1 pharmacophores. Numerical values next to the ligand names in Figure 2.19 represent the best-scoring compounds as numerically listed in Table 2.8. The LPM was therefore not randomized and suggested systematic selection based on

the pharmacophore feature matching. A chemical cluster was found to contain compounds with polyhydroxyl characteristics, such as the best scoring compound ZINC01566791 (Table 2.8), and ZINC04533049 (Figure 2.19). This cluster also contained most of the best-scoring ligands from Table 2.8 (Figure 2.19).

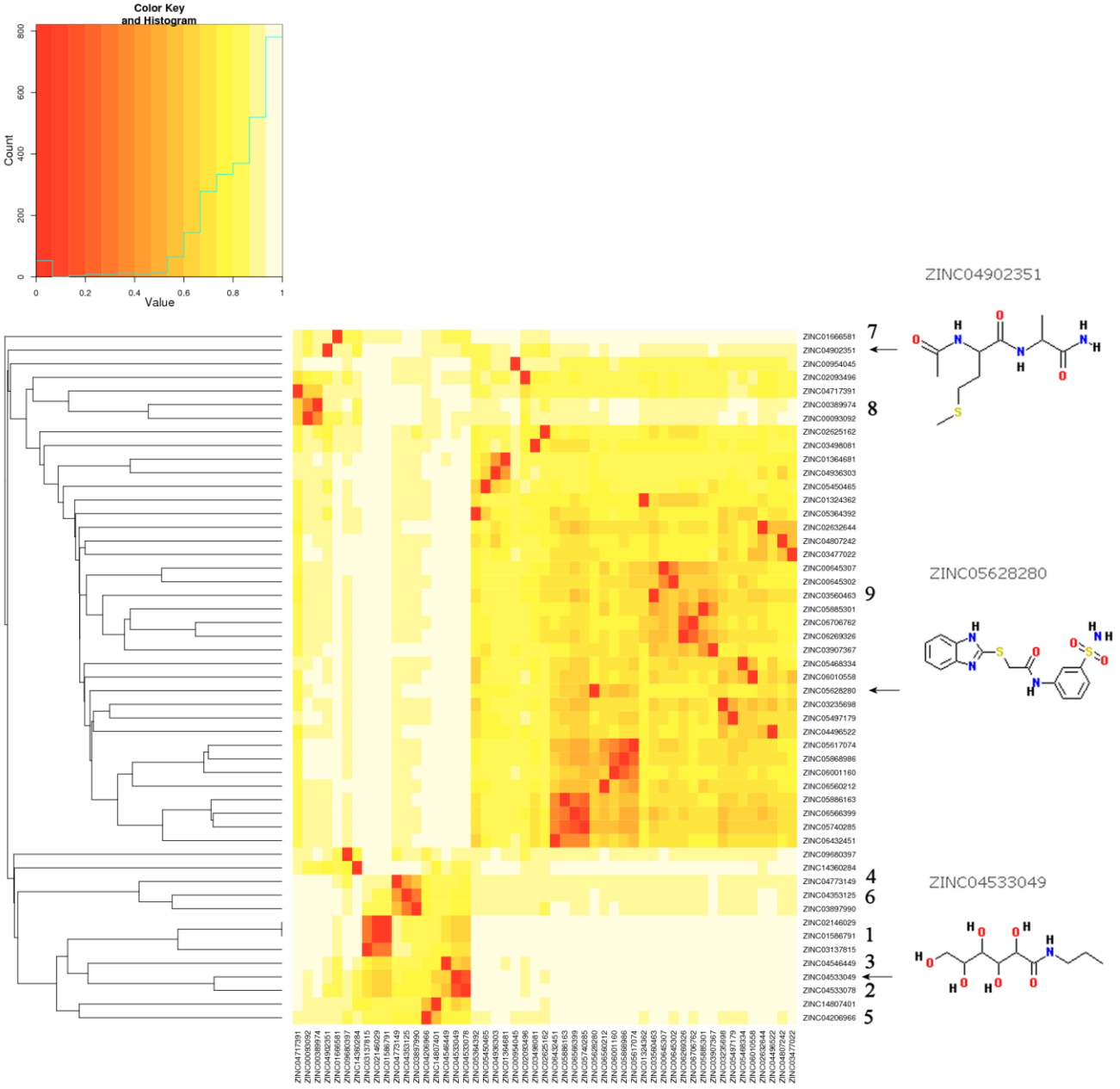


Figure 2.19: Structural pairwise similarity clustering of the best 50 pharmacophore-identified ligands by chemical similarity. Similar ligands are indicated in shades of red, in which the most intense red represents a high degree of chemical similarity or compound identity. Numerical values next to the ligand names, on the right side of the figure, represent the best-scoring in which 1 represent the best scoring ZINC01586791 ligand. The heat map was generated using Chemmine (http://chemmine.ucr.edu).

Distinctive scaffolds in the cluster include ZINC04773149 and ZINC04353125 (No. 4 and 6 respectively in Figure 2.19) both of which share aromatic rings connected to a hexose sugar moiety. Molecules such as ZINC01666581 and ZINC00093092 (No.7 and 8 in Figure 2.19, respectively) were clustered together due to similar number of sulphur atoms although still chemically diverse. ZINC00093092 is aromatic compared to the aliphatic peptide-like arrangement in ZINC0166581. Within the 50-best ligand cluster the molecule ZINC04902351 (indicated in Figure 2.19), which had a peptide backbone characteristic, was similar to ZINC01666581 (No. 7 in Figure 2.19), however, as indicated by the heat map, did not share a high degree of chemical similarity to other ligands. Another large cluster of ligands sharing similar aromatic arrangements included the molecule ZINC03560463, with a sulphonamide moiety (No. 9 in Figure 2.19, also previously shown in Figure 2.17), and ZINC05628280, a phenylated-benzoimidazol (Figure 2.18). These ligands had greater aromatic character and larger molecular weights compared to ligands from other clusters. There was also a greater distribution of similarity shared amongst ligands in this cluster as indicated by shades of red in the blocked regions (Figure 2.19).

The average log *P* values (partition coefficient) and molecular weight distribution compared to the summed LPM fit values of the best 277 ligands are shown in Figure 2.20 A and B. The best 277 ligands had summed fit values ranging from 8.5 – 2.6, and ligands with higher summed fit values tended to have smaller molecular weights (Figure 2.20 A) with more hydrophilic character with log *P* values smaller than -0.5 (Figure 2.20 B). This suggested that the LPM approach mostly ranked smaller molecules with greater hydrophilic character as better ligands for the R5P-active site pharmacophore. This correlated with the features in the respective *Pf*Pdx1 pharmacophores, which tended to contain more H-bond donor and acceptor features compared to the hydrophobic features. The higher ranked ligands had physiochemical properties which correlated with the substrate molecule R5P (228.09 g.mol^{-1} with a log *P* of -3.06). The distribution of ligand molecular weights from the drug-like subset from the Zinc database (before screening) ranged from 150 to 500 g.mol^{-1}, with the log *P* values ranging between 5 to -7 [174]. Within these ranges, the overall distribution of ligand hits from the *Pf*Pdx1 pharmacophore LPM tended to be smaller on average (357 g.mol^{-1}), with slightly hydrophobic character (Avg. log *P* = 1.97). The LPM approach was therefore not only selective for smaller hydrophilic ligands, but able to identify a diverse range of molecules with a wide spectrum of physiochemical properties.

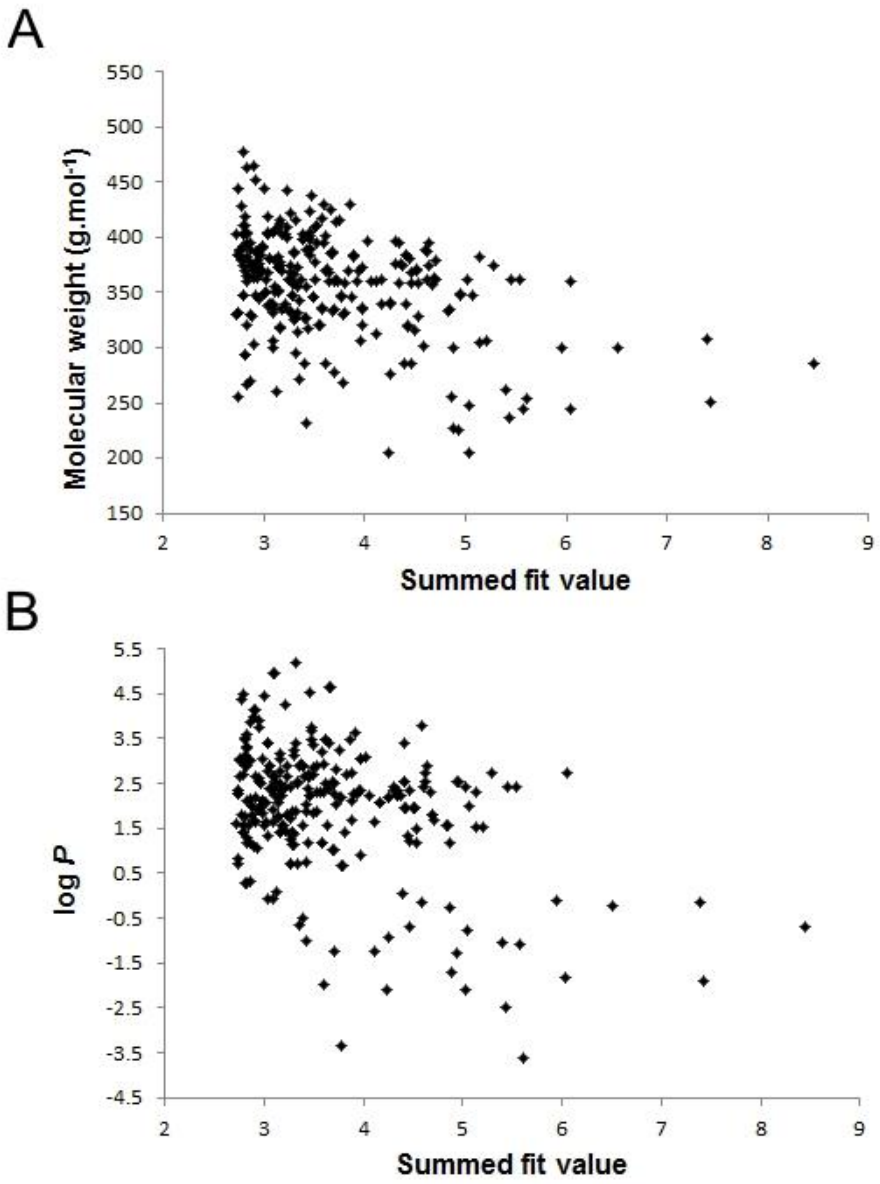


Figure 2.20: The physiochemical properties of the best 277 LPM ligand hits. The log *P* and molecular weight distribution of the best 277 LPM hit ligands were plotted against the summed fit values. The average molecular weight was 357 g.mol^{-1}, and the average log *P* value was 1.97. Chemical properties were obtained from the ZINC database.

2.3.6 Docking

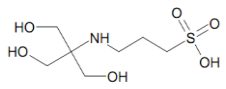
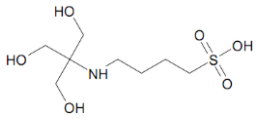
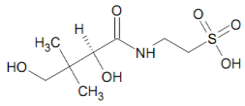
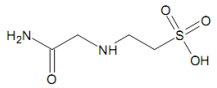
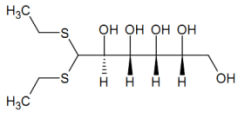
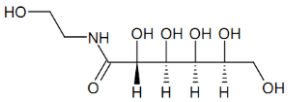
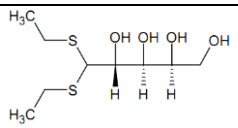
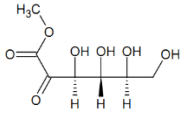
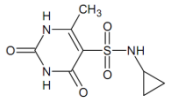
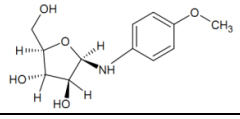
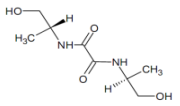
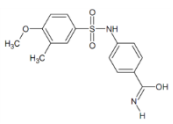
Molecular docking simulations were used to confirm the binding modes of the top 300 ligands, and aided in refining the LPM hits. Ligands were docked into the R5P active site of PfPdx1 homology models, and ligand orientations were scored using Ligscore scoring functions as described in the Methods section. The receptor-ligand binding affinities computed using the LigScore scoring functions have units of pKi (-log Ki) reflecting the predicted association between ligands and receptors.

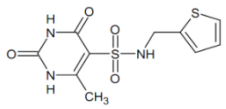
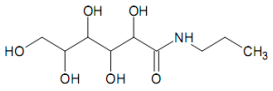
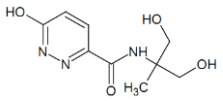
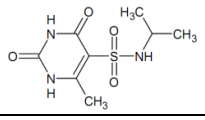
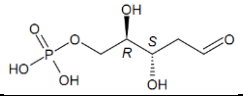
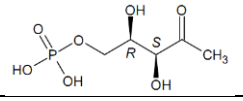
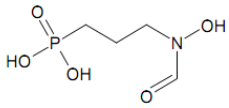
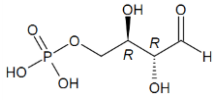
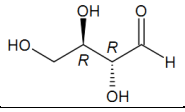
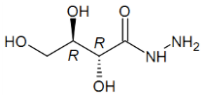
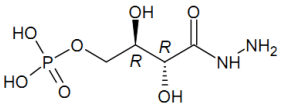
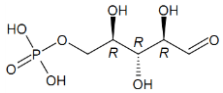
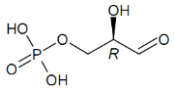
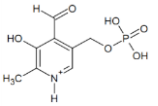
Following the release of the *Pb*Pdx1 structure in 2012 [164], additional docking was performed on the Pdx1 homologue for selected ligands. Docking into *Pf*Pdx1 supported pharmacophore mapping results, and some of the best scoring compounds from the former approach also had good Ligscores (Table 2.9). Some of the ZINC compounds were not commercially available (or affordable) and necessitated the inclusion of closely related analogues, and their docking values were also calculated in order to support selection for down-stream *in vitro* testing. Analogues of R5P were included and expected to have good active site binding capabilities.

Compounds which docked into the *Pf*Pdx1 R5P active site with large Ligscore values, or binding affinities, comparable to the control R5P ligand included ZINC01666581 (Figure 2.21 A). The sulphonated compound adopted an orientation similar to R5P (results not shown) with the sulphonated group H-bonding to corresponding residues known to stabilise the phosphate group of R5P in *Tm*Pdx1 [157]. This also corresponded with the LPM hit results in which this ligand was ranked 7th best. Several sulphonate analogues of ZINC01666581 were also included during docking to probe the binding effect of various substituted groups containing the sulphonate scaffold. The sulphonate groups of the ZINC05273895, TAPS and ACES were predicted with good ligand binding affinities (Table 2.1), and were orientated in similar positions to phosphate groups of R5P during molecular docking (results not shown). The ZINC05273895 ligand with shorter aliphatic chain had better predicted ligand binding affinities compared to TAPS. The sulphonated analogue ACES had the least substituted chemical groups, however was predicted with the poorest binding affinities compared to the other analogues. The two amino groups on ACES, of which the central secondary substituted group carries a positive charge, resulted in this ligand shifting away from the phosphate binding cavity region to H-bond with D104 and D26 (results not shown).

Other ligands which were predicted with good fit values in the pharmacophores such as ZINC01586791 were shown to have docking scores comparable to R5P (Table 2.9). The ZINC03897990 ligand was predicted with better binding affinities compared to R5P, whereas the LPM combined fit-value was 4.42 – within the top 100 hits (Table 2.9). This emphasised that docking was a useful complementary approach, capable of highlighting certain ligands with intermediate LPM scores. Docking confirmed the feasibility of the LPM approach in being able to identify ligands with good ligands binding affinities. Considerations brought into account during docking include active site shape matching, van der Waals interactions as well as the buried surface area between the ligand and the receptor representing attractive interactions. Some additionally included steric restraints introduced during docking could filter out ligands which are geometrically too large for the R5P active site.

Table 2.9: Ligands from the pharmacophore hits docked into *Pf*Pdx1 models. Rationally selected compounds, which shared similarity to pharmacophore-identified ligands, as well as LPM hits were docked into the *Pf*Pdx1 homology models. R5P was used as reference control during docking. LPM fit values of the ligands are given. LigScores as well as Dock scores for ligands docked into the *Pb*Pdx1 structure are given in parentheses.

Compound name or ZINC ID	LPM fit-value	Structure	LigScore 1[a]	LigScore2[a]	Dock score[b]
ZINC05273895	-		6.17	5.6	44.91
TAPS	-		5.75	5.39	61.35
ZINC01666581	5.60		5.97	5.51	54.76
ACES	-		5.22	4.94	82.18
ZINC01586791	8.45		4.76	5.36	24.01
ZINC04683167	-		6.14	5.71	40.09
ZINC03137815	4.88		5.4	6.03	33.31
NSC163924	-		6.21	5.87	45.38
ZINC00093092	5.57		4.55	4.95	44.91
ZINC03897990	4.42		6.87	6.24	49.51
ZINC14360284	5.04		4.97	4.93	32.52
ZINC06269326	4.46		5.6	5.9	34.71

ZINC04717391	4.61		5.27	6.02	42.14
S309737	-		5.65	5.05	41.51
ZINC09680397	4.94		4.85	5.79	42.94
ZINC00389974	5.07		4.5	5.28	41.49
DR5P	-		6.69	6.47	85.64
DXP	-		6.64	6.29	88.92
Fosmidomycin	-		6.27	5.99	91.33
E4P	-		6.40 (3.00)	6.09 (0.41)	85.17 (120.89)
D-erythrose	-		4.74 (3.78)	4.25 (2.15)	42.19 (48.20)
EHz	-		5.34 (4.07)	5.39 (3.46)	53.48 (46.63)
4PEHz	-		6.55 (3.73)	6.43 (1.46)	98.96 (117.01)
R5P (reference compound)	-		**6.55 (4.32)**	**6.02 (2.63)**	**90.15 (107.14)**
G3P	-		5.89	5.4	61.80
PLP	-		6.51	6.27	87.50

[a]LigScore1 and LigScore2, with units of pK_i (-log K_i), refers to predicted receptor-ligand binding affinities. [b]Dockscore refers to the unitless rigid body minimisation energy of the final ligand pose calculated during Monte Carlo trials.

Some rationally selected analogues of R5P which were included during docking include DR5P, DXP, E4P and 4PEHz. These ligands had Ligscore binding affinity values similar to R5P (Table 2.9). DR5P had Ligscore values marginally better than R5P, predicting that this molecule had good binding capabilities in the R5P active site. Other R5P analogues that lacked the phosphate group such as D-erythrose and EHz had lower score values compared to their phosphorylated counterparts E4P and 4PEHz in *Pf*Pdx1 models, respectively (Table 2.9). Another ligand which had comparable docking scores to R5P in the *Pf*Pdx1 models included 4PEHz (Figure 2.21 B). Molecular docking simulations predicted that the molecule is spatially orientated with phosphate groups in association with G237 and G216. Similar poses were shown for the R5P substrate in the *Tm*Pdx1 crystal structure [157].

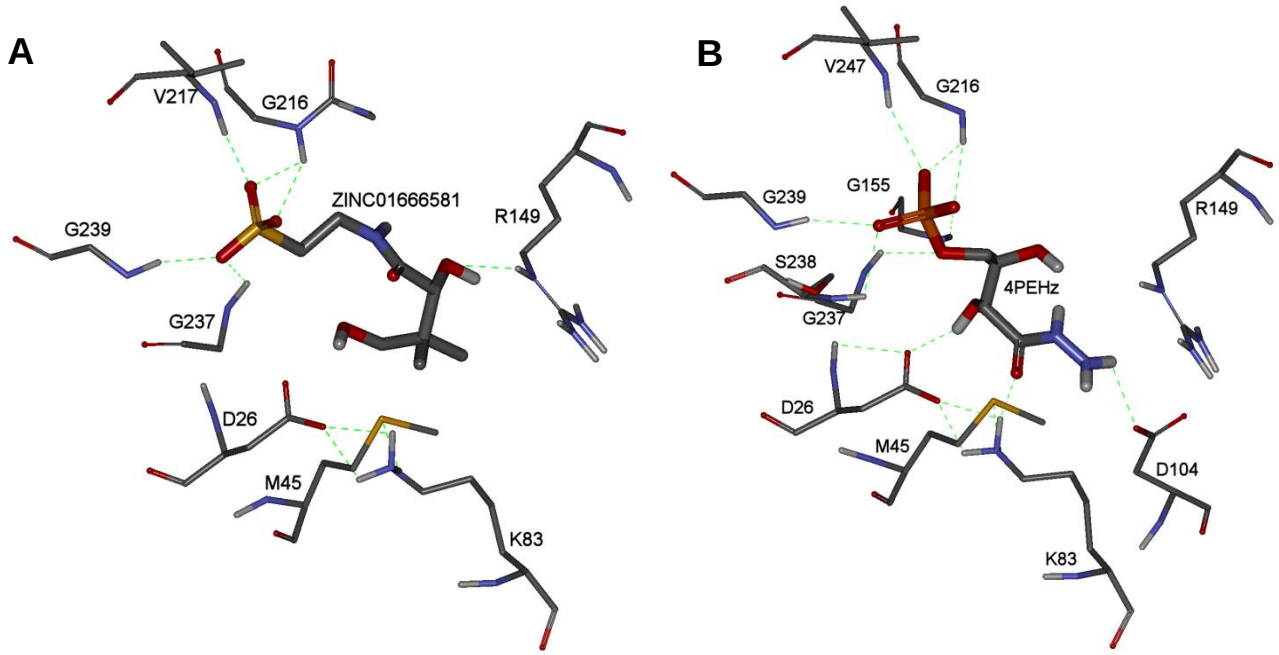


Figure 2.21: Ligands docked in the *Pf*Pdx1_Ru5P homology model. **A)** ZINC01666581 docked in *Pf*Pdx1_Ru5P. The molecule was orientated with sulphate groups H-bonded to G216, G239 and G237, and the C3 hydroxyl substituent was H-bonded to R149. **B)** 4PEHz docked into *Pf*Pdx1_Ru5P predicted that the phosphate group of the molecule to have H-bonds with G155, G216, V247, G239, S238 and G237, whereas the C3 carbonyl O-atom was H-bonded to K83, and the C4 hydroxyl substituent to D26. Atoms C, N, O, P, and H are coloured grey, blue, red, orange and white, respectively.

Additional docking of the ligands in the *Pb*Pdx1 structure was used to confirm predicted ligand binding affinities and orientations as observed in *Pf*Pdx1 models. The R5P-binding site in both *Pf*Pdx1 and *Pb*Pdx1 are highly comparable, and similar binding affinities are expected for ligands in the *Pb*Pdx1 structure. The Ligscore1 receptor-ligand binding affinity of R5P in the *Pb*Pdx1 structure was calcuated to be 4.32 (Table 2.9). The reference molecule had similar orientations compared to docking prediction for R5P in the *Pf*Pdx1 models (Figure 2.22 A). R5P analogues, such as E4P and D-erythrose were docked into the R5P-binding site of *Pb*Pdx1 and had marginally

lower Ligscores to compared R5P (Table 2.9). D-erythrose had better predicted ligand binding affinities compared to E4P (Table 2.9). The docked analogue E4P had similar orientations to R5P with the phosphate groups in close association with S239 and G238 (Figure 2.22 B).

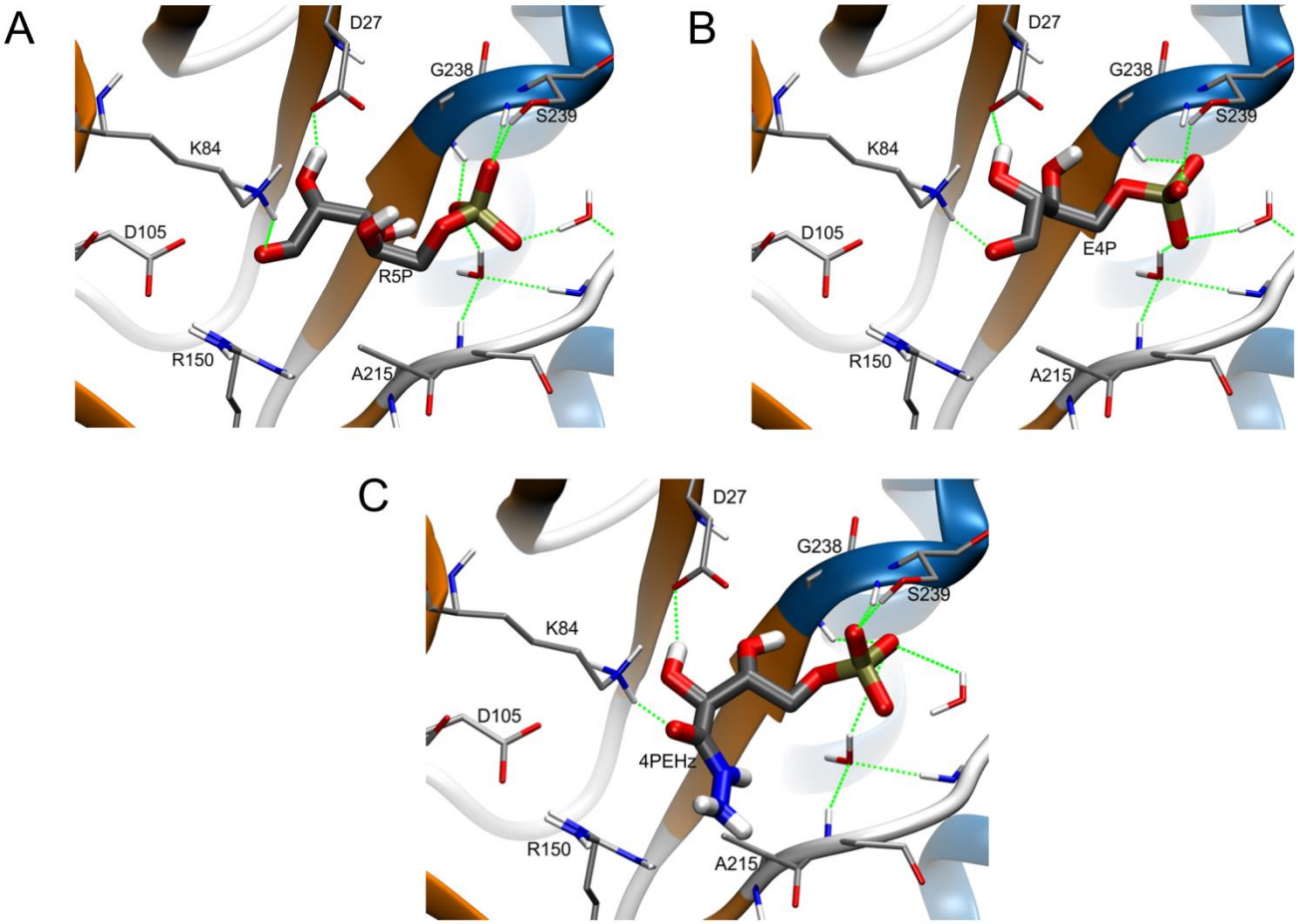


Figure 2.22: Compounds docked into *Pb*Pdx1. A) R5P docked into *Pb*Pdx1 using LigandFit. The phosphate groups of the molecule H-bonded with G238 and S239, as well as with K84. **B)** E4P docked into the R5P binding site of *Pb*Pdx1. The molecule was predicted to H-bond with K84, mainly involving the aldehyde and hydroxyl groups, similary interacting with residue D27. E4P was calculated to have a LigScore1 of 3.00. **C)** 4PEHz docked into *Pb*Pdx1. Similar to R5P, 4PEHz was predicted to adopt an orientation with the phosphate group H-bonded to G238 and S239. The hydrazide group of 4PEHz was orientated approxiately parralel to R150 in close proximity to D108 and both ketone and hydroxyl substituents of 4PEHz H-bonded with K84 and D27, respectively.

The 4PEHz ligand had somewhat less favourable predicted interactions in the *Pb*Pdx1 structure, resulting in poorer LigScores, compared to the unphosphorylated EHz ligand and R5P (Table 2.9). Poses of 4PEHz were scored using LigScore1 and resulted in a value of 3.73, compared to 4.32 for the R5P control. In the *Pb*Pdx1 structure 4PEHz adopted similar positioning compared to R5P, and phosphate groups of 4PEHz were spatially orientated in close proximity to G238 and S239 (Figure 2.22 C). Overall, docking confirmed that compounds from both the LPM and docking approaches

had comparable, if not better, binding affinities to R5P and were suitable candidates for further *in vitro* testing on *Pf*Pdx1.

2.4 Discussion

Homology modelling aims to predict protein conformation based on experimentally determined protein structures and sequence similarities [165]. MSA showed that *Pf*Pdx1 shared around 60% sequence identity with the *Tm*Pdx1 and *Tt*Pdx1 protein sequences. These alignments were used to infer structural information based on sequence conservation to create *Pf*Pdx1 homology models. The level of sequence identity in homology modelling influences the quality of the models. Sequence identities more than 50% are suitable for structure-based drug design and protein-ligand binding predictions [165]. The high degree of sequence identity of *Pf*Pdx1 with *Tm*Pdx1 and *Tt*Pdx1 suggested that accurate models could be generated. The *Tm*Pdx1 structure has a lower resolution (2.90 Å) compared to that of the *Tt*Pdx1 structure (1.65 Å) and at the time this study was initiated *Tm*Pdx1 was the only structure with information regarding the positioning and interactions of the Ru5P substrate. Several *Pf*Pdx1 homology models were built from *Tm*Pdx1 and *Tt*Pdx1 template structures, each incorporating different substrate molecules.

Previous MSA alignments of *Pf*Pdx1 also showed a high degree of identity between plasmodial homologues, all of which contain an additional 20 amino acids on the C-terminal tail end [158]. The Pdx1 homologue from *Toxoplasma gondii* had 58% identity compared to *Pf*Pdx1 [187]. The Pdx1 family of proteins share sequence conservation compared to Dna-K related molecular chaperones – one of the most-conserved protein families [188]. This high degree of conservation emphasises the important cellular and physiological role this protein plays.

The qualities of generated *Pf*Pdx1 homology models were evaluated based on energy violations and spatial restraints. The PDF total energy descriptor represents the summed homology-derived pseudo-energy terms and stereochemical pseudo-energy terms. This is a unit-less value which is useful to interpret the energy violations when individual models are compared. The PDF physical energy is the summed stereochemical pseudo-energy terms which comprise of valence bonds and angles, torsion angles, non-bonded potentials, soft-sphere repulsion energies and improper torsion angles. In both cases higher PDF energy terms indicate larger violations in predicted models. The PDF total energy was also useful to identify and resolve any particular residues with higher restraint violations. These values have no meaning by itself, and more useful to compare different conformers of homology models of the same protein. A lower total PDF energy value determined for *Pf*Pdx1_Ru5P suggested smaller restraint violations in the model incorporating the Ru5P ligand and showed less stereochemical pseudo-energy violations compared to the MPD-containing models.

The DOPE score, a statistical atomic distance-dependent potential descriptor, developed by Shen *et al.*, is a measure of protein stability or native state which can be correlated with model error [179].

The DOPE scores were computed for *Pf*Pdx1 homology models, and were an indication of the conformational energy as well as relative stability of the generated conformation. Considering that the DOPE score does not normalise for the size of the protein, the absolute score is not useful and instead the score is more useful in reporting the relative energies of multiple protein conformations generated during the homology modelling. Homology models with the lowest DOPE scores were selected for further down-stream predictions. The *Pf*Pdx1_MPD and *Pf*Pdx1_Ru5P models had similar DOPE scores and therefore comparable conformational energies.

Minimisation geometrically optimizes protein structures and removes any steric overlaps to relax the conformational state of the protein. Ramachadran plots revealed that minimised *Pf*Pdx1 models had less residues in most favoured regions and a transition into additionally allowed regions, as well as disallowed regions. RMSD comparisons average the deviation of atom positioning in compared macromolecules, and give an indication of structural convergence of model-to-template superimpositions. RMSD comparisons of minimised *Pf*Pdx1 homology models suggested a good correlation, or small deviations, to the template Pdx1 structures. The minimised *Pf*Pdx1 homology models also correlated to the *Pb*Pdx1 structure [164]. RMSD calculations of the C_{α} and main chain atoms in the active site regions of *Pf*Pdx1 homology models compared to the *Tm*Pdx1 and *Tt*Pdx1 templates, as well as compared to *Pb*Pdx1, were close to 1.0 Å. This suggested comparable spatial orientations of R5P-binding site residues and that accurate information could be extracted from the *Pf*Pdx1 R5P-binding site and residues to represent the pharmacophore.

Flicker *et al*. previously modelled the *Pf*Pdx1 using the *Bs*Pdx1 (PDB ID 2NV2) template structure [159]. The model was shown with two additional amino acid residues on the N-terminal. These were proposed to add additional H-bond character on the essential N-terminal component which interacts with the Pdx2 monomers. In this study *Pf*Pdx1 homology models could be created with similar N-terminal α-helical arrangement compared to *Tm*Pdx1 and *Pb*Pdx1. Residues in the αN region of *Pf*Pdx1, which shared approximately 70% identity to the *Pb*Pdx1 structure, could be superimposed with the *Pb*Pdx1 structure. Residues in the *Pf*Pdx1 N-terminal region are expected to have similar interactions with *Pf*Pdx2 compared to the interactions of *Pb*Pdx1 with *Pf*Pdx2 [164].

The topology of the final minimised protein homology model was comparable to the *Tm*Pdx1 crystal structure (PDB ID: 2ISS). The C-terminal tail of *Pf*Pdx1, consisting of 20 amino acids, included in homology models did not adopt any secondary structural conformations and remained unstructured. In the *Tm*Pdx1 structure only a short part (between residues 270 – 280) was shown to become ordered when Pdx1 dimerises with Pdx2, and the furthest terminal end residues remained disordered [157]. The *Pb*Pdx1 structure, in which the last 10 residues were not present, showed that

the last 30 amino acid residues of this C-terminal tail did not adopt any secondary conformation when dimerised with *Pf*Pdx2 (chimera with *P. falciparum* Pdx2 counterpart) [164]. The C-terminal region of *Pf*Pdx1 models was similarly disordered, and could be aligned with *Pb*Pdx1 residues up until residue T281.

It was shown that the *Pf*Pdx1$\Delta_{270\text{-}301}$ and *Pf*Pdx1$\Delta_{273\text{-}301}$ C-terminal deletion mutants were incapable of binding R5P and catalytically inactive, whereas a shorter mutant *Pf*Pdx1$\Delta_{279\text{-}301}$ was capable of binding R5P and forming PLP [182]. These studies suggested that residues preceding G279 were important, and could have communicated long range interactions to the R5P-binding site. C-terminal residues of the *Pf*Pdx1 models had spatially similar arrangement to those found in the *Pb*Pdx1 structure, which suggested accurate representation of residues that could be involved in long range contacts with residue of the R5P-binding site.

Consensus R5P-binding site residues in *Pf*Pdx1 models and templates included the R5P-binding residue K83 as well as D26, which were predicted to form H-bonds with the R5P substrate. Some noticeable differences between the *Tm*Pdx1 and *Pf*Pdx1 active site could include the hydrophobic contribution of the M44 residue. This residue was shown to have long range hydrophobic interactions with the C3 and C2 atoms of Ru5P in *Tm*Pdx1. In the *Pb*Pdx1 structure the equivalent M46 residue was as in close proximity to R5P, which was covalently bonded to the K84 residue [164]. In minimised *Pf*Pdx1 models the equivalent M45 was spatially orientated with sulphur groups closer towards the R5P substrate. The methyl group however occupied a similar positioning compared to other Pdx1 structures. The hydrophobic contributions of this residue are expected to be similar compared to *Tm*Pdx1 and *Pb*Pdx1. This residue was identified to form part of a hydrophobic pharmacophore feature shared with the G237 residue in the *Pf*Pdx1_MPDsco model. Other residues which contribute hydrophobic features in the *Pf*Pdx1 pharmacophore included V108, which is also implicated in *Tm*Pdx1 (V107 in *Tm*Pdx1) [157].

The pharmacophores extracted from the *Pf*Pdx1 homology models were diverse, with different residues in the R5P-binding site contributing to the final pharmacophore model. The R5P-binding site residue K83 plays a fundamental role in initial imine formation with either Ru5P or R5P [157, 164]. This residue was consistently identified as an H-bond donor in all pharmacophores, confirming that representative pharmacophores containing important features could be created. Considering that imine formation entails nucleophillic attack of the amine on the carbonyl group of R5P, K83 can become deprotonated during the course of the reaction coordinate. Nonetheless, in either the protonated or deprotonated states K83 remains a H-bond donor, as predicted in *Pf*Pdx1 pharmacophores. Other residues represented in the *Pf*Pdx1 pharmacophores include D26, of which

the side chain carbonyl group was predicted as a H-bond acceptor feature, capable of sharing the lone pair of electrons on the oxygen atom. This is consistent with the role attributed to this residue in *Tm*Pdx1 as well as *Pb*Pdx1 [157, 164]. Residues A154, G216, G237 and S238 were also identified as predominantly H-bond donor features. These residues are involved in R5P-phosphate group H-bond interactions in *Tm*Pdx1 and similarly in *Pf*Pdx1 can be considered as the phosphate-group pharmacophore region.

Pharmacophore-based screening and confirmatory docking of positive hits was performed using the ZINC database containing 2 million compounds and resulted in the identification of around 3000 ligands. Plotting of the physiochemical properties of the ligands suggested that a chemically diverse set of ligands were identified, but also emphasised that smaller molecules with a hydrophillic character were better matched to the *Pf*Pdx1 pharmacophore. The diversity of the ligands suggested that matching chemical features on the ligands were found, whereas the basic chemical framework these were situated on was different. This becomes useful when active molecules are identified, and the solubility or pharmacokinetic properties need to be optimised . The average molecular weight and log *P* values of the top scoring compounds from the LPM were in close agreement to properties of most approved drug libraries (MW of 375 Da, log *P* of 3.4) [99]. The distribution of physiochemical properties from the initial drug-like library suggested that pharmacophore screening effectively filtered compounds from the database. Clustering of the scaffold composition of the best hits further emphasised the identification of molecules from other diverse chemical scaffolds, and were grouped into distinct classes. The 4 best scoring ligands had similar chemical arrangements, with polyhydroxylated moieties, implying that from the large *in silico* library some selectivity was achieved, and shared scaffolds could reproducibly be identified. The top scoring molecules had modified polyhydroxylated or carbohydrate backbones similar to R5P. pharmacophoreP mapping therefore successfully identified molecules with shared characteristics to known substrates of *Pf*Pdx1.

In this study docking was used to reinforce pharmacophore hits and introduced additional filtering for selection of ligands for *in vitro* testing. Docking is more computationally intensive, and therefore only the top scoring compounds from LPM were docked. LPM spatially aligns flexible ligands in order to maximise the number of geometrically overlapping chemical features with that of the target or receptor [189]. Docking predicts and generates binding modes of ligands which is comparable to crystallographic ligand binding modes [190]. Docking considers the geometrical fit of ligands in the active site, by calculating the local energy minimum of the complex, whilst incorporating protein flexibility. Additional shape restraints during docking introduced more stringent fitting conditions for further refining of pharmacophore hits. The predicted ligand binding

affinities from docking simulations were used as criteria to evaluate ligand suitability, and helped in assessing ligand binding modes. R5P could be docked back into the *Pf*Pdx1 models and had binding modes comparable to that of Ru5P in *Tm*Pdx1 and R5P in *Pb*Pdx1 [157, 164]. The top scoring molecules from LPM scored were similarly predicted with the high ligand binding affinities during docking.

The LPM fit-values and data from molecular docking were as criteria to select molecules for further *in vitro* testing. Ligands with appreciable LPM fit-values such ZINC04546449 were excluded during docking. Docking in this case allowed the exclusion of ligands which did not match the shape restraints of the R5P active site. The ZINC04546449 ligand was in the top 10 LPM hits, suggesting a good fit in all of the pharmacophores, however contained a C_6 aliphatic portion which was not expected to be accommodated in the active site. The ligand failed to dock in any of the *Pf*Pdx1 models, and confirmed that the ligand should be excluded for further testing on the enzyme. This molecule additionally shared chemical similarity to ZINC04683167, which docked better, and was a better alternative to test.

A few compounds with good LPM fit-values such as ZINC04533078 were not commercially obtainable, and an alternative analogue ZINC04683167 was used. Docking in this case was used to confirm ligand binding affinities, which were within the comparable range of R5P for the selected analogue. Docking was useful to prioritise selection of certain ligands such as ZINC03897990 which had an intermediate LPM fit-score (4.42), ranking 50th on the LPM-hit list, whereas the ligand had predicted binding affinities better than R5P. Docking therefore confirmed the binding capabilities of the ligand. Docking confirmed that ligands with polyhydroxylated moieties such as ZINC01586791 and ZINC03137815 identified during LPM had potential for binding in the R5P active site of the *Pf*Pdx1 models. Analogues of these such as NSC163924 and ZINC04683167 (which replaced ZINC04546449), which were not screened in LPM, were included in the selection for *in vitro* testing to determine the effect of alternative side chain moieties from the polyhydroxylated scaffold.

Molecular docking simulations predicted R5P analogues, such as DR5P, DXP and E4P with better or comparable fit values to R5P in both *Pf*Pdx1 models and the *Pb*Pdx1 structure. The phosphate groups of most of these analogues were predicted to form H-bonds with residues A154, G216, G237 and S238 – the phosphate-binding region in the R5P active site. These corresponding residues in *Tm*Pdx1, which are highly conserved, form H-bonds with the phosphate group of R5P [157]. This emphasized that coordinated anchoring of the phosphate moiety could be an important requirement for binding in the R5P site of *Pf*Pdx1. These analogues were therefore included for

additional testing on *Pf*Pdx1. As linear forms (acyclic) these analogues had either terminal aldehyde or carbonyl groups, a feature shared with the R5P substrate. These groups are considered to be relatively reactive and have the potential to interact with other chemical groups [191]. For this reason these are generally undesirable to include in virtual libraries, and were manually included during docking studies. When considering the substrates of *Pf*Pdx1, G3P and R5P, these molecules bearing the same terminal aldehyde arrangement are also reactive, in particular the reactivity of R5P enables this molecule to bind to K83 in the active site [160]. These shared features on the R5P analogues, such as DR5P and E4P, were therefore appropriate and believed to have potential in affecting R5P binding in *Pf*Pdx1. Additional details regarding the reactivity of R5P, and the binding mechanism in Pdx1 proteins, is discussed in Chapter 3.

Several sulphonated ligands such as ZINC01666581 were identified during LPM, and the binding modes during docking suggested that the sulphonate groups could associate with the components of the phosphate-binding region of the R5P-active site. The LPM and docking therefore suggested that compounds carrying this sulphonate moiety may be accommodated within the R5P-active site, in particular with these groups corresponding to the phosphate group of R5P. It is not uncommon for phosphate groups in ligands to be substituted for sulphonate groups which reduces potential chemical and enzymatic degradation and thus increases stability of compounds [192]. Moreover sulphonate groups carry less negative charge and provide a neutral characteristic to the compound, which is more desirable in order to cross cellular membranes [192]. For this reason additional ligands such as ZINC05273895, TAPS and ACES were included during *in vitro* testing to probe whether sulphonated ligands may be accommodated.

2.5 Conclusions

The current rate at which the limited antimalarial chemotypes are becoming ineffective due to drug-resistance necessitates the discovery of better and more effective chemotherapeutics. Newer therapeutic targets in essential pathways need to be investigated and exploited. PLP biosynthesis in the parasites provides the PLP cofactor to myriad array of different parasite-essential processes, and for this reason has become an attractive pathway for the development of newer chemotherapeutics. Both *Pf*Pdx1 and *Pf*Pdx2 are integral parts of the PLP synthase complex and *Pf*Pdx1 forms the catalytic centre in which all the substrates condense to form PLP.

Homology models created of the *Pf*Pdx1 were comparable to a recently solved *Pb*Pdx1 structure, as well as *Tm*Pdx1. The high degree of sequence identity and structural homology shared amongst Pdx1 proteins is indicative of structural conservation owed to its specialised biochemical function. When this study was initiated the only definitive substrate binding site in the Pdx1 protein was the R5P-active site. Residues in the R5P-binding site of *Pf*Pdx1 were conserved and had similar spatially orientations compared to *Tm*Pdx1 and *Pb*Pdx1. Residues previously implicated in substrate molecular interactions were highlighted during extraction of the *Pf*Pdx1 pharmacophores, and were used to create a spatially and chemically representative site of the R5P cavity in *Pf*Pdx1. Ligands that matched the pharmacophore features were identified, and molecular docking confirmed favourable shape complementary and the potential to form H-bonds with R5P site residues. The *in silico* derived ligands had diverse chemical scaffolds, which suggested that the computational screening effectively searched the initial compound library, and spatially arranged complementary chemical features on the ligands could be found on more than one type of chemical framework. This suggests that the expansion into different lead molecules or scaffold hopping from actives could be possible, an approach often used to improve pharmacokinetic properties and solubility of chemical leads [193].

Docking reinforced LPM and showed that top scoring ligands from pharmacophore screening had favourable shape complementarity and the potential to form H-bonds in the R5P active site. This ultimately aided in selection of ligands for *in vitro* testing. A series of sulphonated analogues were selected for further testing on the protein based on predictions that the sulphonate groups interact with residues in the phosphate-binding region in the R5P-cavity. These may therefore interfere with R5P-binding in the protein, and could be potential inhibitors. Additionally, analogues of R5P such as DR5P and E4P were predicted with favourable binding affinities, and had binding modes comparable to R5P. Several polyhydroxylated sugar-like ligands were also highlighted during docking and were selected for additional testing.

Chapter 3

In vitro identification of inhibitors against *Pf*Pdx1 and their potential MOA

3.1. Introduction

3.1.1. On the inner workings of the PLP synthase complex

The Pdx1 protein combines R5P and G3P, including ammonia provided through the hydrolysis of L-glutamine by Pdx2, to form PLP. Merging of these small molecules occurs in the Pdx1 component of the PLP synthase complex, and how this is coordinated is not entirely understood yet. The Pdx1 reaction mechanism leading to the formation of PLP is fascinating and intricately complex. Some significant findings regarding the inner working of this enzyme are briefly introduced, which alludes to peculiar functions that make it an attractive drug target.

The Pdx1 enzyme perform a multitude of different complex chemical functions including triose- and pentose-phosphosugar isomerizations, condensation of two phosphosugars, imine formation, and closing of the pyridine ring after incorporation of ammonia [119, 162]. The *Pf*Pdx1 protein is highly conserved in organisms utilising the DXP-independent B_6 pathway, including *P. falciparum* parasites [117, 188]. From this high degree of structural conservation it can be deduced that the multiple Pdx1 active sites have an uncompromising nature, and point mutations are not readily tolerated in this gene. This could hold merit in terms of inhibitor design, since fewer incidences of rescue of resistant phenotypes due to point mutations in *Pf*Pdx1 could be expected.

Pdx1 proteins have the ability to utilise alternative substrates such as dihydroxyacetone phosphate (DHAP) and Ru5P as alternative sugars [194]. In particular *Pf*Pdx1 also has this ability [104]. DHAP and G3P are interchangeable triose sugar substrates, and it was shown that Pdx1 contain triosephosphate isomerase (TIM) activity, able to catalyze the conversion of G3P into DHAP [194]. Both R5P and Ru5P are accepted substrates in Pdx1 proteins, however, *B. subtilis* Pdx1 (*Bs*Pdx1) was demonstrated to have a lower K_m value for R5P, which suggests this to be the preferred substrate [195]. Both R5P and Ru5P are incorporated into the final PLP molecule, and the enzyme has been proposed to have R5P isomerase activity.

As mentioned previously, the R5P binding site or P1 site is known from crystallographic evidence [157, 164]. The P2 site or PLP-binding site is located on the surface, and was shown to be solvent accessible with four water molecules complexed with PLP [163]. However, where the R5P and G3P substrates combine, and whether this occurs in the P2 site is not known [163]. A third site or P3 site in which G3P was bound, has been identified, and the G3P molecule had H-bond contacts to E116

and R164 [163]. Equivalent residues of *Pf*Pdx1 that could be involved in G3P binding are R164 and D119.

It is widely accepted that the R5P / Ru5P substrates form imines with an active site lysine as one of the first steps of the reaction mechanism [157, 161, 162]. Hanes *et al.* identified initial reaction ^{13}C-nuclear magnetic resonance (NMR) peaks characteristic of an imine bond between the N atom of lysine and the C1 carbon of the R5P substrate (2 in Figure 3.1) [160]. Additional evidence to support this was shown in the *Pb*Pdx1 structure shown with bound R5P [164]. Earlier structural evidence of Ru5P bound in the *Tm*Pdx1 crystal structure has proven that the molecule is capable of forming an imine with the active site lysine, however substrate isomerisation and reshuffling of Ru5P could lead to reversion into the bound R5P state [157, 160, 194].

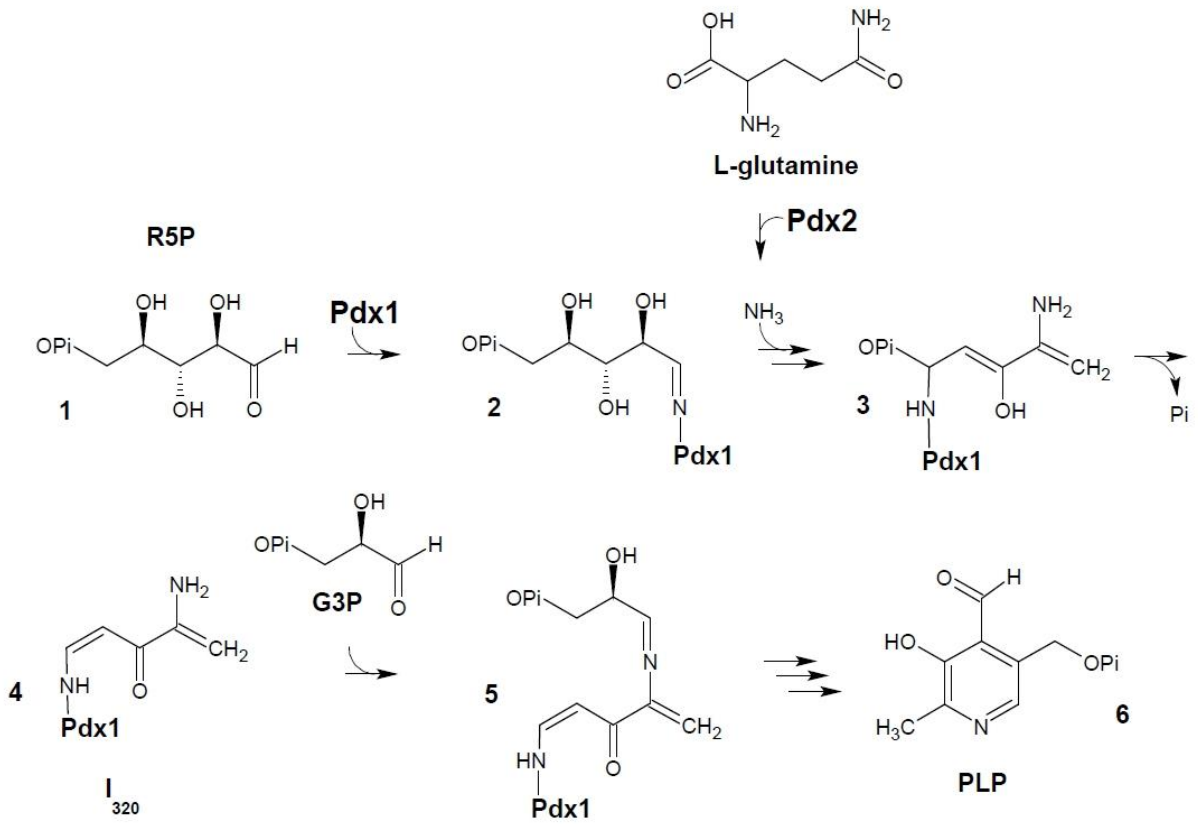


Figure 3.1: The Pdx1 reaction mechanism. Acyclic aldehyde forms of R5P (1) bind via an imine bond to K83 (2). The bound intermediate undergoes isomerization, followed by an additional imine formation with ammonia derived from Pdx2 monomers (3). A C1 to C5 lysine migration also occurs during this step (3). Loss of phosphate results in the formation of I_{320} (4). This bound species has a characteristic absorbance at 320 nm, and forms within the enzyme in the absence of G3P. The I_{320} chromophore undergoes imine formation with G3P (5), and leads to ring closure and formation of PLP (6). This figure was redrawn, as adapted from Hanes *et al.* [160].

Raschle *et al.* identified a novel chromophoric intermediate, called I_{320}, with absorbance maximum (λ_{max}) at 315 nm [162]. This species forms when Pdx1 was incubated with Pdx2 and glutamine with

only the pentose substrate R5P. Formation of the I_{320} chromophoric species was also found to be dependent on the R5P concentration in the presence of the nitrogen source, including ammonium sulphate [162]. Systematically, incubation of R5P with Pdx1 results first in the formation of the imine- or amine-intermediate, then into I_{320} once the ammonium sulphate source was introduced (4 in Figure 3.1) [161, 162]. Hanes *et al.* proved that the I_{320} chromophore was a ketone intermediate species of R5P which further undergoes imine formation with G3P leading to eventual formation of PLP (5 in Figure 3.1) [160, 196]. The I_{320} intermediate can therefore be considered as a marker for R5P binding and an indication that R5P is effectively converted into an intermediate receptive for conjugation of G3P.

Some reshuffling of both the R5P substrate and lysine migration in the R5P-active site has been proposed; Hanes *et al.* showed that there is an imine migration (also called lysine migration) from the C1 of R5P (2 in Figure 3.1) to the C5 (3 in Figure 3.1) after incorporation of ammonia (there are multiple step involved) [160]. This implies that there is a state during the reaction mechanism when an early intermediate form of R5P dissociates from the internal lysine, in order to form the new C5 imine bond with the same lysine. This step could have implications for inhibitor design, and compounds that interfere with imine formation initially as well as during this step could have potential to affect the enzyme activity. A second proposal by Zein *et al.* suggested the possibility of an imine shuffling in *Tm*Pdx1 involving both K81 and K149 residues in *Tm*Pdx1 [157]. The R5P was suggested to be shuffled or transferred to a new site in which the K149 residue forms the imine bond (in this case with the C2 carbon). Mutagenesis of K149 did disrupt PLP formation, consistent with the notion that this residue has direct catalytic involvement [157]. The K149 in *Tm*Pdx1 was suggested to be capable of rotating, coming in close proximity with the Ru5P-K82 imine adduct [157]. Both these proposal suggest that imine formation in the R5P-active site is essential for catalytic conversion of R5P, and compounds that interfere with these steps could be valuable during drug design. Many questions remain regarding the reaction mechanism of Pdx1. The site of G3P-binding in Pdx1 has only recently been proposed and how this substrate is brought into proximity to R5P is not yet known [163]. The active site residues such as D26, K83 as well as K151 are strictly conserved in *Pf*Pdx1 suggesting a similar reaction mechanism.

Recent mutagenesis studies on the *Pf*Pdx1 protein have also highlighted important residues in dodecamer assembly as well as enzyme activity. Two triple mutants of *Pf*Pdx1 were created the DKK (D26A, K83A and K151A) mutant involved mutation of the active site residues, and the RHE (R85A, H88A and E91A) mutant involved residues involved in Pdx1:Pdx1 monomer contacts [133]. Both of these *Pf*Pdx1 mutants were not capable of maintaining dodecameric quaternary structure of the PLP synthase and existed as inactive Pdx1 hexamer and monomer structures,

respectively. [133]. In the DKK mutant the residues that destabilise *Pf*Pdx1 assembly were shown to exclusively the K151A mutation [133]. The monomeric RHE *Pf*Pdx1 mutant, which also did not support Pdx2 activity, suggested that these highly conserved RHE Pdx1:Pdx1 interface residues coordinates Pdx1 assembly to form at least a hexameric Pdx1 crown [133]. The ability of these mutants to form the I_{320} intermediate – a marker of R5P binding - was not tested. The previously created *Pf*Pdx1 DKK and RHE mutants were therefore tested in order to establish the involvement of these residues in the catalytic process of forming PLP.

The C-terminal tail of the *Pf*Pdx1 also contributes to oligomeric assembly and substrate binding. A C-terminally truncated *Pf*Pdx1 (*Pf*Pdx1$\Delta_{273\text{-}301}$ and *Pf*Pdx1$\Delta_{270\text{-}301}$) was monomeric and this 30 amino acid unstructured region is suspected to be involved in initial substrate binding and cooperative assembly of the PLP synthase [182]. This mutant was also unable to form I_{320}, and through long-range interaction is believed to aid in organisation of the R5P active site. Here the *Pf*Pdx1$\Delta_{270\text{-}301}$ mutant was recreated and the catalytic capabilities in terms of PLP and I_{320} formation were confirmed in this study. This also served as a control during analysis of other *Pf*Pdx1 mutants and their ability to form I_{320}. Additionally, the residue located in the P3 G3P binding site was mutated to confirm this residues involvement in PLP formation as well as I_{320} formation. This residue was located more than 20 Å from the R5P binding site, and was not expected to affect I_{320} formation [163]. In the previous chapter, the identification and selection of potential inhibitors of *Pf*Pdx1 were reported. Here the effect of these compounds was tested on purified *Pf*Pdx1 *in vitro*. Some active inhibitors were identified and the potential mode-of-actions of these are discussed further.

3.2. Methods

3.2.1. Expression and purification of *Pf*Pdx1

Plasmids of wild-type *P. falciparum* Pdx1 and Pdx2, as well as K83A, K151A, RHE, ERR, DKK *Pf*Pdx1 mutants, cloned into the pASK-IBA3 vector, were obtained from Knöckel *et al.* [132]. Plasmids were transformed into the *E. coli* BLR (DE3) expression cell line using heat shock transformation. *Pf*Pdx1 plasmid DNA (1 µL or 10 ng) was added to 100 µL heat shock competent cells (prepared as described below), allowing incubation on ice for 30 min. Cells were subjected to a heat shock at 42°C for 60 sec, after which the cells were placed on ice for 2 min. Preheated LB medium (Luria-Berthani sterile medium containing 10% w/v tryptone, 5% w/v yeast extract, 10% w/v NaCl, pH 7.5; 800 µL, 37°C) was added to the transformed cells and the cells grown with shaking (200 rpm) at 37°C for 1 h. Transformed cells (100 uL) were plated out on LB-amp agar plates (LB-broth with 1% w/v agar and 50 µg/mL ampicillin) and grown overnight at 37°C.

For the preparation of heat shock competent cells 5 mL of LB-broth was inoculated with *E. coli* BLR (DE3) and grown overnight at 37°C in a shaking incubator at 200 rpm. The cells were diluted 1:100 into 250 mL fresh LB-broth and were incubated at 37°C with shaking (200 rpm) until an optical density (OD) at 600 nm of 0.5 was reached. The cells were then cooled on ice for 15 min and subsequently centrifuged at 3000*g* for 15 min at 4°C. Whilst working on ice, the remaining pellet was gently re-suspended with 50 mL 80 mM cold $CaCl_2$ and left on ice for 40 min. The cells were centrifuged at 3000*g* for 15 min at 4°C and the pellet was again re-suspended with 5 mL 80 mM cold $CaCl_2$ as well as 2 mL 50% w/v glycerol. Cells were divided into 100 µL aliquots and frozen at -70°C.

Selected single colonies were used to inoculate LB broth (10 mL) containing 50 µg/mL ampicillin (LB-amp) and grown overnight at 37°C. Cultures were diluted 1:100 into LB-amp media and grown to an OD_{600nm} of 0.5. Expression was induced using 200 ng/mL anhydrotetracycline (IBA, Germany) and grown for 4 h at 37°C with shaking. Cells were centrifuged at 2000*g* for 10 min (Avanti J-26 XP, Beckman Coultier) and the remaining pellets were re-suspended using 15 mL buffer W, consisting of 100 mM Tris-HCl, 150 mM NaCl adjusted to pH 8.0 using HCl. Phenylmethylsulphonyl fluoride (PMSF, 0.1 mM) and 5 mg lysozyme (Sigma Aldrich) was added for every 10 mL re-suspended bacterial cell pellet. The re-suspended pellets were sonicated at 40 kHz using pulsing at output 6, duty cycle 40 with 1 min resting intervals on ice (Sonifier 250, Branson). The lysed bacterial mixture was clarified by centrifugation at 50000*g* for 50 min at 4°C (Eppendorf). Strep-Tag fusion proteins were purified using strep-avidin affinity resin according to the manufacturer recommendations (IBA, Germany), as briefly described here. Strep-Tactin®

Sepharose resin (3 mL of 50% suspension) was packed in polypropylene (PP) cartridges containing sintered filters. The resin was equilibrated with 10 column bed volumes (15 mL) of buffer W. The cleared cell lysates (approximately 15 mL) was added to the resin and allowed to elute through at least three consecutive times. Non-specifically adhered proteins were washed off using 3 x 5 mL of buffer W. Elution buffer (2 mL), which consisted of 2.5 mM desthiobiotin and 1 mM ethylenediaminetetraacetate (EDTA) in buffer W at pH 8.0, was eluted through the resin three consecutive times. The resin was regenerated with 15 mL of buffer R, which consisted of 1 mM hydroxy-azophenyl-benzoic acid (HABA), in buffer W at pH 8.0. The columns were finally rinsed with 10 column bed volumes (15 mL) of buffer W.

Protein concentrations were determined using Bradford protein reagent [197], consisting of 0.01% w/v Coomassie Brilliant Blue G-250 (Merck, Germany) dissolved in 4.7 % v/v ethanol (EtOH) containing 8.5% v/v H_3PO_4. Protein containing solutions (10 µL) were diluted with 90 µL buffer W, and pippeted into 1 mL plastic cuvettes (Plastibrand, Brand). Bradford reagent (900 µL) was added and the reactions were incubated at room temperature (RT) for 5 min. Absorbance of the protein-dye solutions were determined at 595 nm and compared to a standard calibration curve of bovine serum albumin (BSA) to calculate the protein concentrations in the analysed samples.

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) was performed as previously established by Laemmli *et al.* [198], with modifications. Electrophoresis was accomplished in Biometra Minigel protein electrophoresis housings (Biometra, Göttingen, Germany) with plastic gel cassettes of sizes of 8.6 x 7.7 cm (Invitrogen, California, USA). Protein samples were denatured in loading buffer (final concentrations of each component; 50 mM Tris, pH 6.8, 2% w/v SDS, 10% v/v glycerol, 0.02% v/v β-mercaptoethanol, 0.05% w/v bromophenol blue). Samples were boiled for 5 min to denature the proteins. A 10% separating gel was prepared by adding 2.5 mL dH_2O, 1.5 mL of 3 M Tris-HCl pH 8.9, 2 mL acrylamide/bis-acrylamide solution (29.1% w/v acrylamide, 0.9% w/v bis-acrylamide in dH_2O), 60 µL of 10% SDS, 50 µL 10% w/v ammonium persulphate (APS), and a final concentration of 33.5 nM (5µL) *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED) to a 50 mL PP tube. This solution was poured into assembled plates, and allowed to polymerise at 4°C for 1 h. The stacking gel component at 5% consisted of 1 mL of 0.5 M Tris-HCl, pH 6.7, 667 µL acrylamide/bis-acrylamide solution, 40 µL 10% w/v SDS, 50 µL 10% w/v APS, and a final concentration of 33.5 nM (5µL) TEMED in a 50 mL PP tube. The solution was cast on top of the separating gel component together with an appropriate comb. Electrophoresis buffer consisted of 25 mM Tris-Cl, 192 mM Glycine, 0.1% w/v SDS. The protein gels were stacked at 80 V, followed by separating conditions at 120 V. The PageRuler unstained protein ladder (Fermentas) with standard reference proteins ranging from 10 –

200 kDa, was used as molecular weight marker. SDS-PAGE gels were stained with Coomassie solution, containing 0.1% w/v Coomassie Brilliant blue R-250 (Serva, Germany), 48% v/v EtOH and 10% v/v glacial acetic acid in dH_2O. Destaining was performed using a solution containing 20% v/v EtOH and 10% v/v glacial acetic acid.

3.2.2. *Pf*Pdx1 activity assays and inhibitor preparation

*Pf*Pdx1 activity was determined spectrophotometrically through detection of a PLP-Tris Schiff base at 414 nm [199]*Pf*Pdx1 activity was assayed in reaction buffer consisting of 100 mM Tris-HCl, 150 mM NaCl at pH 8.0 with purified *Pf*Pdx1 (between 100 – 250 µg, indicated below) including 0.5 mM R5P (Sigma Aldrich, prepared in $dddH_2O$), 0.5 mM G3P (Sigma Aldrich, diluted in $dddH_2O$) and 20 mM NH_4Cl (Merck), unless otherwise stated. Alternatively when the activity of *Pf*Pdx1 was assayed together with *Pf*Pdx2 (between 100 – 250 µg, indicated below), NH_4Cl was replaced with 20 mM L-glutamine (Merck). Blank reactions contained only G3P and NH_4Cl, or G3P and L-glutamine when *Pf*Pdx2 was present. Calibration curves with a PLP standard (Sigma Aldrich, prepared in $dddH_2O$) were made in 100 mM Tris-HCl, 150 mM NaCl pH 8.0, and used to determine the concentration of PLP in the *Pf*Pdx1 incubations. Spectrophotometric measurements at 414 nm, including spectrophotometric wavelength scans and I_{320} measurements, were made using an Uvikon 933B spectrophotometer (Bio-Tek Kontron) in 70µL cuvettes (Plastibrand, Brand). The ability to form PLP was determined by comparing the absorption at 414 nm with a standard calibration curve of PLP in the same buffer conditions. The enzyme activity was expressed in $nmol.min^{-1}.mg^{-1}$.

Both initial inhibitor screening tests as well as inhibitor IC_{50} assays on *Pf*Pdx1 were performed in reaction buffer (100 mM Tris-HCl, 150 mM NaCl at pH 8.0) together with 0.5 mM R5P, 0.5 mM G3P and 20 mM NH_4Cl with 100 µg purified *Pf*Pdx1 in a 250 µL reaction. Assays in which *Pf*Pdx2 was included contained 100 µg of both purified *Pf*Pdx1 and *Pf*Pdx2 in reaction buffer containing 0.5 mM R5P, 0.5 mM G3P and 20 mM L-glutamine. All the inhibitors were prepared in $dddH_2O$, unless otherwise stated. Initial inhibitor screening involved testing inhibitors between 0.5 and 3 mM concentrations. Reactions were set up with the addition of the inhibitor and purified *Pf*Pdx1 or PfPdx1 with *PfP*dx2 occurring last. The enzymes were not pre-incubated with the inhibitors. An uninhibited control reaction was included to ensure active protein was present, and was used as reference to express inhibited reactions as a percentage of the uninhibited control. Inhibitor screening experiments were performed in duplicate on three or more independent occasions.

3.2.3. Mutagenesis of *Pf*Pdx1

Previously created mutants of *Pf*Pdx1 DKK, RHE and ERR were kindly provided by Dr. I.B. Müller and Prof. C. Wrenger [133]. Two additional mutant *Pf*Pdx1 proteins were created in order to study their effects on the enzyme activity and I_{320} species formation. Two different *Pf*Pdx1 mutants were created, the one contained a 31 amino acid deletion on the C-terminal tail (*Pf*Pdx1$\Delta_{270\text{-}301}$) and the other a R164A. Primers designed are listed in Table 3.1. All primers were ordered from Sigma-Aldrich and dissolved in dddH_2O (Millipore) to a final concentration of 300 µM. A working stock solution was prepared by using a 1/10 dilution of the original primer stock solution, with a final concentration of 30 µM.

Table 3.1: **Primers used in mutagenesis polymerase chain reactions (PCR) to create ΔC- and R164A *Pf*Pdx1.** Regions in bold indicate the sites of the introduced modification. The pASK-IBA3 primers were used for automated sequencing.

Primer	Sequence (5' - 3')
ΔC sense	GCGCGCGGTCTCGAATGGAAAATCATAAAGATGATGC
ΔC antisense	GCGCGCGGTCTCAGCG**CTAACATCTAAAAGTATTTTAGGG**
R164 sense	GCTATTAAACATATA**GCA**ACTGTAAATAATGAA
R164 antisense	TTCATTATTTACAGT**TGC**TATATGTTTAATAGC
pASK-IBA3 sense	GAGTTATTTTACCACTCCCT
pASK-IBA3 antisense	CGCAGTAGCGGTAAACG

3.2.3.1. Deletion mutagenesis PCR of *Pf*Pdx1 (*Pf*Pdx1$\Delta_{270\text{-}301}$)

Deletion of the C-terminal tail residues 270 to 301 (*Pf*Pdx1$\Delta_{270\text{-}301}$) of *Pf*Pdx1 was performed using deletion mutagenesis PCR of *Pf*Pdx1 previously cloned in the expression plasmid pASK-IBA3 (IBA, Germany) [132]. The PCR contained 35 ng of expression plasmid with 30 pmol each of the ΔC sense and antisense primer using recombinant *Thermus aquaticus* DNA polymerase in PCR Supermix (Invitrogen). PCR conditions were as follows; 95°C for 3 min, then 35 cycles of 95°C for 30 s, 41°C for 90 s, 60°C for 2 min in a Thermocycler UNO II (Biometra). Parent template was removed from the deletion mutagenesis PCR product using *DpnI* (New England Biolabs). PCR products (20 - 40 µL) were digested with 1 µL of *DpnI* (20 U), whereby reactions were incubated for 2 h at 37°C. Following parent template removal general molecular cloning techniques as described in Section 3.2.4 were followed to ligate the fragments into the pASK-IBA3 expression vector. Fragments were purified according to methodology described in Section 3.2.4.1. The fragments were digested using *BsaI* to enable sticky-end ligation into pre-digested pASK-IBA3 vector as described in Section 3.2.4.3. Digested fragments were again purified, as previously described, and ligated into the vector and transformed into an *E. coli* cloning cell line as outlined in

Section 3.2.4.4. The insertion of the correctly-sized fragments was verified by isolating plasmids and performing RE digestion and automated nucleotide sequencing as described in Section 3.2.4.5

3.2.3.2. Site-directed point mutation of R164 in *Pf*Pdx1

The *Pf*Pdx1 R164A mutant was kindly provided by Dr. I.B. Müller. The mutation was created with R164 sense and antisense primers given in Table 3.1, as described previously [133].

3.2.4. General molecular cloning techniques

3.2.4.1. Purification of PCR and digested fragments

Purification of PCR products was performed using the Pure Link PCR purification kit (Invitrogen) according to manufacturer recommendations, which entailed the addition of 180µL binding buffer to 45 µL PCR samples, and adding the sample to provided spin columns. Columns were spun at 10 000*g* for 1 min, discarding the eluent. Adhered fragments on the columns were washed with 650 µL wash buffer, and centrifuged at 10 000*g* to remove unwanted wash products. The spin columns were additionally spun at 10 000*g* to remove any residual liquids. The PCR products were eluted using 50 µL dddH_2O pre-heated to 50°C.

3.2.4.2. Agarose gel electrophoresis

Gel electrophoresis was performed as previously described by Sambrook *et al.* [200]. A 1% w/v agarose gel was prepared by adding 1 g of agarose (Fermentas) to 100 mL of TE buffer (90 mM Tris, 2 mM EDTA at pH 8.0). The mixture was heated in order to dissolve the agarose, and allowed to cool, after which 5 µL EtBr (10 mg/mL) was added to a final concentration of 0.5 µg/mL. The agarose solution was cast into gel chambers (Biometra, 8.2 cm x 7.1 cm). Samples (5 µL) were mixed with 2 µL sample loading dye (Promega) and loaded into cast gels together with an appropriate DNA molecular weight marker, in this case GeneRuler™ 1000 BP ladder (Fermentas). Electrophoresis was performed for 60 min at 120V in a TE buffer. Bands were visualised under an UV-illuminator to confirm correct PCR products and digested fragments.

3.2.4.3. Restriction enzyme (RE) digestion

The PCR products were digested using *Bsa*I (New England Biolabs, 10 000 U/mL). A typical 100 µL digestion reaction contained 50 µL purified PCR product, 10 µL buffer 3 (New England Biolabs, 100 mM $NaCl_2$, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1mM dithiothreitol (DTT), pH 7.9), 3 µL *BsaI* (30 U) and 1 µL bovine serum albumin (BSA, 10 mg/mL) to a final concentration of 100 µg/mL. Water ($dddH_2O$) was added to a final reaction volume of 100 µL. Reactions were digested at 37°C for 3h. The pASK-IBA3 plasmids were digested to expose sticky-ends prior to ligation with inserts in reactions conditions containing 30 µL vector, 10 µL buffer 3, 3 µL *BsaI* (30 U) and 1 µL BSA with 56 µL $dddH_2O$ digested in conditions similar to above.

3.2.4.4. Ligation

*Pf*Pdx1$\Delta_{270\text{-}301}$ gene constructs were ligated into pASK-IBA3 expression plasmids. Ligation was performed using T4-DNA ligase (New England Biolabs) at 14°C for 24 h in a reaction containing 16 µL purified fragment, 1µL digested pASK-IBA3 plasmid, 2 µL ligation buffer and 1 µL T4-DNA ligase at a 3:1 insert to vector ratio. Plasmids from the ligation reactions were directly transformed into *E. coli* XL10 Gold competent cells (Stratagene) by heat shock transformation using 100 µL competent cells by adding 20 µL of ligation mixture. The cells were incubated on ice then subjected to heat shock as described in section 3.2.1. For plasmid isolation, colonies were selected from the plates and used to inoculate 2 mL LB-amp, and grown overnight at 37°C with shaking.

3.2.4.5. Plasmid isolation

Plasmids were isolated from transformed XL10 gold *E. coli* cells using Gold Plasmid miniprep kit (Peqlab) following manufactures recommendation. Briefly, the 2 mL LB-amp inoculum was pelleted at 10000*g* for 2 min at RT. Solution I (250 µL) was added to the pellets and these were disrupted using vortexing. Solution II (250 µL) was then added and the tubes gently inverted. Solution III (350 µL) was added and the tubes were inverted, after which the tubes were centrifuged at 10000*g* for 10 min. The supernatant was transferred to a PerfectBind DNA spin column and centrifuged at 10000*g* for 1 min, discarding the flow through material. The adhered material on the columns was washed with 500 µL PW plasmid buffer, and centrifuged at 10000*g* for 1 min, discarding the flow-through material. Residual wash buffer was removed by additionally centrifuging the spin columns for 5 min at 10000*g*. Plasmids were eluted from the columns by

adding 50 µL $dddH_2O$, preheated to 50°C, allowing a brief incubation period and then centrifuging the columns at 10000*g* for 3 min.

To verify the correct cloning of the *Pf*Pdx1$\Delta_{270\text{-}301}$ gene construct RE digestion was performed. Purified plasmids were digested using 10 U each of *XbaI* (NEB, 10 000 U/mL) and *HindIII* (NEB 10 000 U/mL), 10 µL of plasmid DNA, together with 2 µg BSA protein (New England Biolabs) and 2 µL Buffer 2 (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9) filled to a final volume of 20 µL using $dddH_2O$. Fragments were digested at 37°C for 3h. Samples were additionally submitted for automated sequencing (Seqlab, Germany). Reactions were prepared by adding 6 µL of plasmid DNA with 1 µL of IBA3 sense primer, and similarly 1 µL of antisense IBA 3 primer for reverse sequencing reactions. Correct plasmids were transformed into *E. coli* BLR DE3 and purified as described in Section 3.2.1.

3.2.5. Culturing of *P. falciparum* parasites

3.2.5.1. Culture media and culturing conditions

P. falciparum (3D7) parasites were maintained in continuous culture according to Trager *et al.* [201], as modified by Das Gupta *et al.*[201]. The RMPI 1640 culture medium contained 25 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), 10 mM glucose, 20 mM sodium bicarbonate, 25 mM hypoxanthine and 0.5% w/v AlbuMAX II (Invitrogen) at pH 7.4. Parasite cultures were maintained at 2% haematocrit using human O+ erythrocytes in flasks (Nunc) with shaking at 200 rpm at 37°C. The culture media was replaced every second day through centrifugation of the cultures at 3000*g* for 5 min, removing the supernatant, and carefully re-suspending the pellets in the appropriate volume of culture media. Flasks were deoxygenated or gassed using 5% CO_2, 5% O_2 and 90% N_2. The parasitaemia ranged between 1% - 4% depending on the experiments as indicated below.

3.2.5.2. Sorbitol synchronisation

Before IC_{50} assays the synchronicity of the parasite life-stages was controlled by D-sorbitol synchronisation to selectively retain ring-stage parasites, as previously described by Lambros *et al.* [202]. A solution containing 5% w/v D-sorbitol was prepared in PBS and was filter sterilized using 0.22 µM syringe filters (Sartorius). Cultures (50 mL) consisting of mainly ring stage parasites (between 1 to 5% parasitaemia) were centrifuged at 3000*g* for 5 min, after which the supernatants were discarded. The pellets were re-suspended in 5 volumes (10 mL) of the 5% D-sorbitol solution, and incubated at RT for 10 min. The cultures were centrifuged at 3000*g* for 5 min, and the supernatant was removed. The pellets were re-suspended in complete culture media, following gassing procedures as indicated previously.

3.2.5.3. IC_{50} determinations

^{3}H-hypoxanthine incorporation assays were performed as reported by Das Gupta *et al.* [203]. Cultured parasites were washed in RMPI media lacking hypoxanthine prior to IC_{50} determinations. Cultured parasites (250 µL) at initial 2% hemotocrit and 1% parasitaemia were tested in 96 well plate formats with serial dilutions of compound varying from 400 µM – 0.1 µM. After 24 h incubation at 37°C, 100 nCi ^{3}H-hypoxanthine (10 µL of 2.5 µCi/µL) was added. The cells were harvested after another 24 h incubation using an Inotech IH110 cell harvester. The parasite nucleic acid material was collect on pre-wetted (using dddH_2O) paper filter mats (Perkin Elmer) using the cell harvester, in which the dddH_2O enables lysing of the cells. The filter mats were subsequently rinsed in dddH_2O and dried under vacuum. The filter mats were dried at 80°C for 15 min then placed in plastic sachets together with 4 mL Ultima Gold™ scintillation fluid (Perkin Elmer) and counted using a Trilux liquid scintillation counter (Wallac). The counts per minute (CPM) generated for untreated parasites within the plates were used as the 100% uninhibited growth control. Assays were performed in triplicate within 96-well plates, and at least three independent biological experiments for each compound.

3.2.5.4. Growth inhibition assays of transfected parasites

Asynchronous parasites, co-transfected with *Pf*Pdx1 and *Pf*Pdx2 contained on pARL1a^{-} vectors with blasticidin and WR99210 antibiotic resistance (kind gift of Carsten Wrenger, Hamburg) together with a Mock cell line containing vector with only antibiotic resistance were maintained in culture, as described previously [141]. The parasite growth was monitored over a period of seven

days (168 h) in the presence of 1 µM 4PEHz, as well as in untreated cultures of the same transgenic cell line. Long term growth assays were performed using 1 mL parasite culture in 12 well format plates (Nunc, Denmark) at 4% haematocrit and a starting 1% parasitaemia. Blasticidin S-hydrochloride (Fluka) and WR99210 were maintained at final concentrations of 2.1 mM (1 µg/µl) and 5 nM, respectively. Both *Pf*Pdx1/*Pf*Pdx2-overexpressing and mock parasites were maintained in culture for 7 days. Spent media was removed daily from suspended erythrocytes by gently tilting the plates, and replaced with fresh media containing both antibiotics. Fresh 4PEHz was added to both treated-*Pf*Pdx1/*Pf*Pdx2 as well as treated-mock parasites every day at a final concentration of 1 µM. The growth rate or parasitaemia was assessed by morphological monitoring and counting of Giemsa (Merck, Giemsa's azur eosin methylene blue solution) stained thin smears using oil immersion light microscopy (Axioskop 20,Zeiss) at 100x magnification, counting ≥1000 total cells (uninfected together with infected) from triplicate slides.

Considering that parasite growth was exponential over time, log transformations were performed to linearize the growth rate data. The growth rate was expressed as the doubling time. The doubling time was a measure of effective doubling of parasite numbers, however considering that the parasite has a 48 hour life cycle and the multiplicity of *P. falciparum* invasions can range from 4 to 16, the term doubling time may be misleading. The doubling time should therefore in this case be interpreted as the effective growth rate of the parasites, and not the actual rate of doubling. Alternatively, the growth rate constant (K), was used to compare the growth of parasites. The effective doubling time has been used to relate different growth patterns in gene disrupted clones of *P. falciparum* by Liu *et al.* [204].

The doubling time or effective growth rate was calculated using the following formula;

$$Doubling\ time = \frac{\ln(2)}{K}$$

In which K represents the growth rate constant with units of h^{-1}. Data (or the cumulative parasitaemia over time) of the 7-day growth inhibition assays were log transformed, and the slope is used to calculate K, by using the following equation;

$$K = \ln(10^{slope})$$

In which slope refers to the slope value from the log graph.

Results were analysed using Graphpad Prism 5, in which the null hypothesis of equal regression lines slopes was tested in a 99% confidence interval (P = 0.01). Analysis of covariance (ANCOVA)

was conducted using JMP9 (JMP, Version 9, SAS Institute Inc.), testing whether the slopes significantly differed from each other with $\alpha = 0.05$.

3.3. Results

3.3.1. Expression of *Pf*Pdx1 and initial inhibitor screening

Protein expression and purification of *Pf*Pdx1 and *Pf*Pdx2 was performed as reported previously [158]. The Strep-tag on *Pf*Pdx1 facilitated the purification of the protein on affinity-based resin, and no visible contaminating proteins other than the 37.5 kDa *Pf*Pdx1 were noted on SDS-PAGE with Coomassie staining (Figure 3.2 A). As outlined in Brooks *et al.* during inhibitor screening the linearity of the enzyme assay needs to be to ensure that detection of the product is not compromised or becomes saturated in the product ranges measured [205]. The PLP product is measured as a Schiff-base adduct of Tris at 414 nm [199]. PLP standard calibration curves were linearly correlated with increasing concentrations of PLP, corresponding to the increased formation of a Schiff base between PLP and Tris (Figure 3.2 B). This showed that PLP could accurately be quantified from unknown samples within these ranges.

Ideally during inhibitor screening in the enzyme is maintained in initial velocity conditions, or the linear range where the velocity does not change with time [172]. The relatively low sensitivity for detection of PLP and the relatively slow turnover rate for PLP necessitated using substrate conditions, as well as enzyme concentrations, that led to sufficiently detectable levels of PLP. For this reason initial activity of the enzyme was verified using conditions of 1 mM R5P and G3P with 40 mM NH_4Cl. The formation of PLP by *Pf*Pdx1 was spectrophotometrically monitored as the formation of the PLP-Schiff base at 414 nm over time (Figure 3.2 C). PLP standard calibration curves, such as Figure 3.2 B, were used to calculate the amount of PLP formed by *Pf*Pdx1, and showed that the PLP formed during the reaction fell within the linear ranges for the quantitation of PLP. Linear responses in enzyme activity over time with fixed substrate concentrations showed that no substrate-limiting conditions were reached, and the formation of PLP did not result in product inhibition on *Pf*Pdx1 for the duration of monitoring (Figure 3.2 C). Product inhibition can potentially interfere with interpretation of screening results, and is often associated with plateauing of enzyme activity for extended incubation periods. With linear PLP-formation rates, equating to enzyme activity, this also suggested that the protein remained stable for at least 5 h at 37°C.

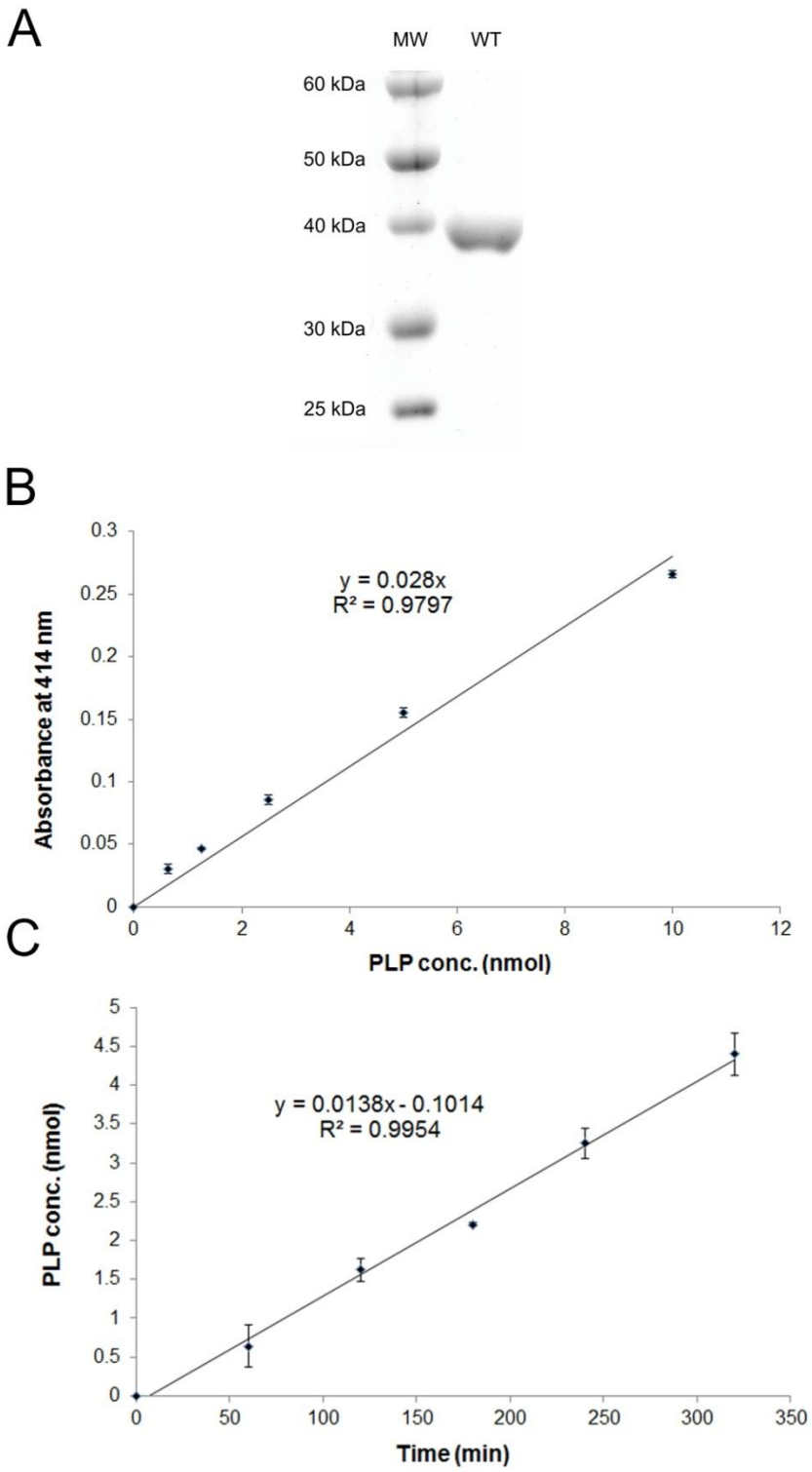
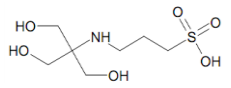
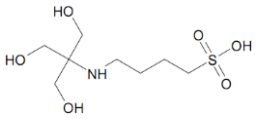
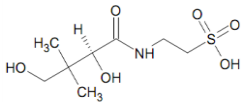
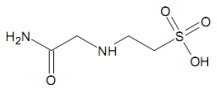
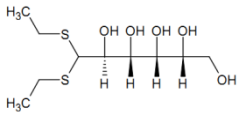
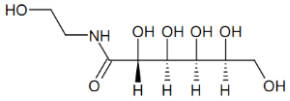
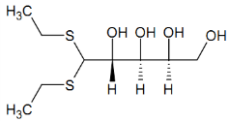
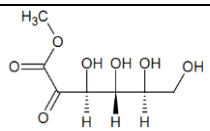
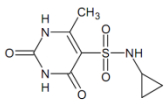
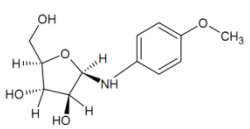
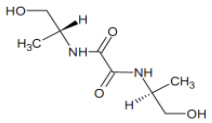


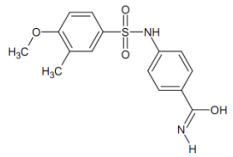
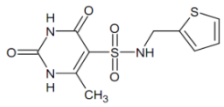
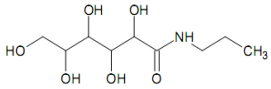
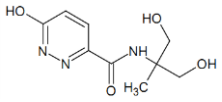
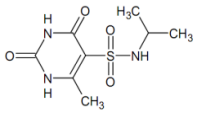
Figure 3.2: SDS-PAGE and enzyme activity of purified *Pf*Pdx1. **A**) Coomassie stained SDS-PAGE gel (black and white image) showing purified protein samples of wildtype (WT) *Pf*Pdx1. One hundred microgram of *Pf*Pdx1 was loaded together with a molecular weight marker (MW). **B)** Standard PLP calibration curve at 414 nm (A_{414}). Error bars indicate the standard deviation from a single experiment derived from triplicate determinations. There was a linear correlation between the PLP concentration and absorbance as reflected by R^2 of the fitted line with 0.97. **C)** *Pf*Pdx1 enzyme activity in reaction buffer (1 mM R5P, 1 mM G3P and 40 mM NH_4Cl) with 100 µg (final concentration 0.125 µg/mL). Aliquots from the incubated solution were removed at hourly intervals. Results represent data from a single experiment with triplicate determinations. Error bars indicate the standard deviation.

The reaction conditions during inhibitor screening were decreased to 0.5 mM R5P and G3P, as well as 20 mM NH_4Cl, with 200 µg of purified *Pf*Pdx1, which still resulted in sufficiently detectable levels of PLP during the reaction incubation times (between 4 – 10 nmol total PLP formed). The reaction times varied between 2 and 3 h at 37°C. The average specific activity of the protein under these conditions was 310 ± 25 $pmol.min^{-1}.mg^{-1}$ protein from more than 10 independently performed experiments. This was marginally lower than previously reported values of 746 ± 76 $pmol.min^{-1}.mg^{-1}$ using 0.5 mM R5P, 1.0 mM G3P and 10 mM NH_4Cl [133, 158], and 695 ± 71 $pmol.min^{-1}.mg^{-1}$ using 1 mM R5P and G3P with 10 mM NH_4Cl [182].

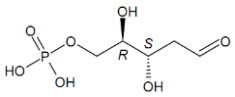
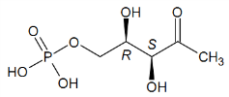
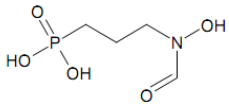
Ligands identified as potential inhibitors of *Pf*Pdx1 by pharmacophore-based drug design (Chapter 2), were tested for their ability to inhibit *Pf*Pdx1. Initial concentrations ranging from 0.5 mM to 3 mM was used for each ligand, in reaction conditions containing 200 µg purified *Pf*Pdx1, 0.5 mM R5P and G3P, as well as 20 mM NH_4Cl. The inhibitors were tested at concentrations equivalent or greater than that of the substrates to promote detection of inhibitory effects on the enzyme. Control reactions excluding inhibitors were included in each screening assay (Table 3.2). Several ligands that were identified contained sulphonate substituent groups (ZINC05273895, TAPS, ACES and ZINC01666581) but did not significantly inhibit the formation of PLP, even when tested at 3 mM (Table 3.2). Additionally, compounds ZINC01586791, ZINC04683167 and ZINC03137815 with carbohydrate backbones also did not show any significant inhibitory action (Table 3.2). NSC163924 with a terminal diol-ester arrangement was the most effective of the conjugated carbohydrates, with a 88% *Pf*Pdx1 activity remaining at 0.5 mM, however this was not statistically significant ($P \geq 0.05$, n = 5). Some compounds containing aromatic rings (ZINC00093092, ZINC04717391, ZINC09680397, ZINC00389974) were predicted to bind within the R5P active site, with the premise that these could potentially mimic PLP in this site. These compounds, however, did not show appreciable inhibitory activity (Table 3.2). The ZINC03897990 compound was coloured with absorption spectra overlapping with that of PLP at 414 nm, and was falsely reflected to increase the *Pf*Pdx1 activity. The pyrimidine compound, ZINC00093092, was shown to significantly (P = 0.046) increase *Pf*Pdx1 activity at 1 mM (Table 3.2).

Table 3.2: Inhibitory activity of *in silico* identified compounds against *Pf*Pdx1. Results are expressed as the average percentage inhibition compared to uninhibited control *Pf*Pdx1 activity calculated from three or more independent experiments performed in duplicate with the standard error of the mean (SEM) indicated. Statistical comparisons, of which the *P*-value is reported, were made using an unpaired two-tailed Student's t-test with $n \geq 3$ in which *P*-values smaller than 0.05 were considered significant.

Compound name or ZINC ID	**Structure**	**Conc. (mM)**	***P*-value**	***Pf*Pdx1 inhibition (% remaining activity)**
ZINC05273895		0.5	0.318	96 ± 3
TAPS		0.5	0.476	103 ± 4
ZINC01666581		3	0.565	96 ± 5
ACES		3	0.348	96 ± 3
ZINC01586791		0.5	0.893	99 ± 5
ZINC04683167		0.5	0.369	107 ± 6
ZINC03137815		1	0.928	99 ± 7
NSC163924		0.5	0.108	88 ± 6
ZINC00093092		1	0.046	113 ± 2
ZINC03897990		0.5	0.053	200 ± 20 *
ZINC14360284		0.5	0.453	110 ± 10

ZINC06269326		0.5	0.326	107 ± 5
ZINC04717391		0.5	0.774	99 ± 3
S309737		0.5	0.974	99 ± 5
ZINC09680397		0.5	0.341	95 ± 4
ZINC00389974		0.5	0.292	105 ± 3

Rationally selected compounds

Compound name	**Structure**	**Conc. (mM)**	***P*-value**	***Pf*Pdx1 inhibition (% remaining activity)**
DR5P		12	0.082	56 ± 4 (Fig 3.6A)
DXP		5	0.114	88 ± 4
Fosmidomycin		20	0.701 (n = 2)	103 ± 8

*absorbance of compound interfered with detection of PLP.

The overall lack of significant inhibition of *Pf*Pdx1 by *in silico*-identified ligands necessitated additional testing of rationally selected molecules. Analogues of R5P were therefore tested in order to establish the competitive binding interactions since it was anticipated that these analogues could potentially compete or interfere with initial R5P entry into the active site. Compounds tested included 2-deoxy-D-ribose 5 phosphate (DR5P), deoxy-D-xylulose 5-phosphate (DXP), and fosmidomycin (Table 3.2). DR5P showed some inhibitory effect on PLP production (Table 3.2) and was investigated further as discussed below. When tested at 5 mM the acyclic keto-sugar, DXP, marginally decreased *Pf*Pdx1 activity to 88% of that of the uninhibited control *Pf*Pdx1 reactions,

however this was not statistically significant (Table 3.2). The closely related analogue of DXP – fosmidomycin - had no effect on *Pf*Pdx1 activity at 20 mM (Table 3.2).

3.3.2. The inhibition of *Pf*Pdx1 by DR5P, and the effects of related analogues

Spectrophotometric wavelength scans, which allowed the simultaneous monitoring of PLP and I_{320} formation, were used to show that DR5P inhibited *Pf*Pdx1 (Figure 3.3 A). In the presence of DR5P *Pf*Pdx1 showed concomitant decreases in PLP formation as the concentrations of DR5P increased (Figure 3.3 A). The I_{320} formation with DR5P was not significantly altered compared to the uninhibited protein *Pf*Pdx1. The possibility that DR5P could mimic R5P and be incorporated into the final PLP structure prompted further assays with DR5P, as well as other analogues, in which R5P was excluded. However, incubation of *Pf*Pdx1 with DR5P, together with NH_4Cl and G3P did not result in the formation of visible chromophoric species at 414 nm (Figure 3.3 B). DR5P therefore was shown to inhibit *Pf*Pdx1, and could not serve as an alternative substrate in place of R5P. Results suggested that DR5P could mimic R5P and potentially enter the active site, thereby interfering with R5P processing or imine bond formation.

Remarkably the Pdx1 protein is able to isomerise Ru5P to R5P, and also DHAP to G3P. For this reason it was of interest to verify the possibility that other sugar analogues of R5P could serve as a substrate in the *Pf*Pdx1 reaction. The enzyme was incubated with several analogues, including the required G3P and NH_4Cl substrates (Figure 3.3 B and C). In incubations of *Pf*Pdx1 with D-erythrose 4-phosphate (E4P) the formation of a 320 nm chromophoric species similar to incubations with the substrate R5P was noted (Figure 3.3 B). Incubations containing R5P, E4P, DXP or DR5P in the absence of G3P confirmed that only E4P and R5P led to the formation of the chromophore (Figure 3.3 C). These alternative sugars were therefore not considered to support formation of PLP. The effect of E4P on *Pf*Pdx1 I_{320} formation and enzyme activity was investigated further.

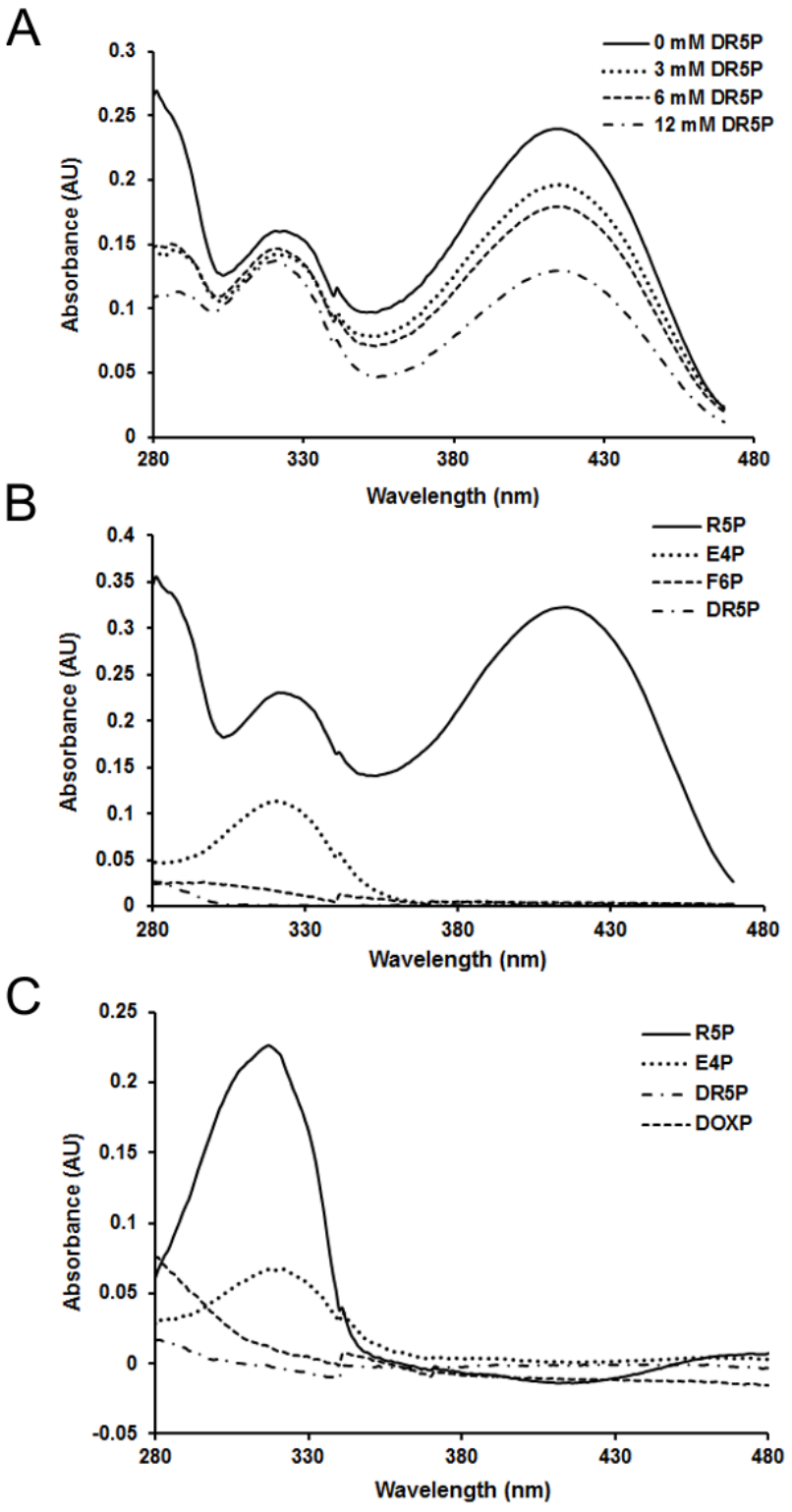


Figure 3.3: Effect of R5P-analogues on *Pf*Pdx1 and I_{320} formation. A) Progressive i**nhibition of *Pf*Pdx1 by increasing concentrations of DR5P in reactions containing 200 µg purified *Pf*Pdx1 with 0.5 mM of each R5P and G3P, using 20 mM NH_4Cl as nitrogen source. Decreased levels of PLP measured at 414 nm correlated with increased concentrations of DR5P.** B) **U**tilisation of alternative sugar **substrates using 200 ug *Pf*Pdx1 incubated with 3 mM R5P, 3 mM E4P, 12 mM D-fructose 6-phosphate (**F6P**) or 12 mM DR5P respectively, in the presence of 1 mM G3P and 20 mM NH_4Cl. Note the 320 nm maximum absorbance only for R5P and E4P.** E4P **was shown to promote formation of an I_{320} species. C) Alternative sugars, E4P (2 mM), DR5P (2 mM) and DXP (2 mM) incubated** with *Pf*Pdx1 in the presence of 0.5 mM G3P and 20 mM NH_4Cl. The positive control included 0.5 mM R5P lacking G3P to

demonstrate formation of the chromophoric intermediate - I_{320}. Solely E4P of the sugar analogues formed the I_{320} species.

3.3.3. Erythrose 4-phosphate inhibits *Pf*Pdx1

E4P was assayed with *Pf*Pdx1 in reaction conditions with R5P, G3P and NH_4Cl to establish whether the compound affected the enzyme activity. Increasing concentrations of E4P was shown to produce the 320 nm chromophoric species (Figure 3.4 A) and also dose-dependently affected PLP production, reflected as decreased 414 nm-absorbance values (Figure 3.4 A). To further elucidate the SAR the dose-response relationship of both E4P and its non-phosphorylated analogue D-erythrose was tested on *Pf*Pdx1 (Figure 3.4 B). The IC_{50} of E4P was calculated as 3.7 ± 0.9 mM and D-erythrose as 160 ± 46 mM (Figure 3.4 B. The calculated IC_{50} of D-erythrose was more than 40-fold weaker against *Pf*Pdx1.

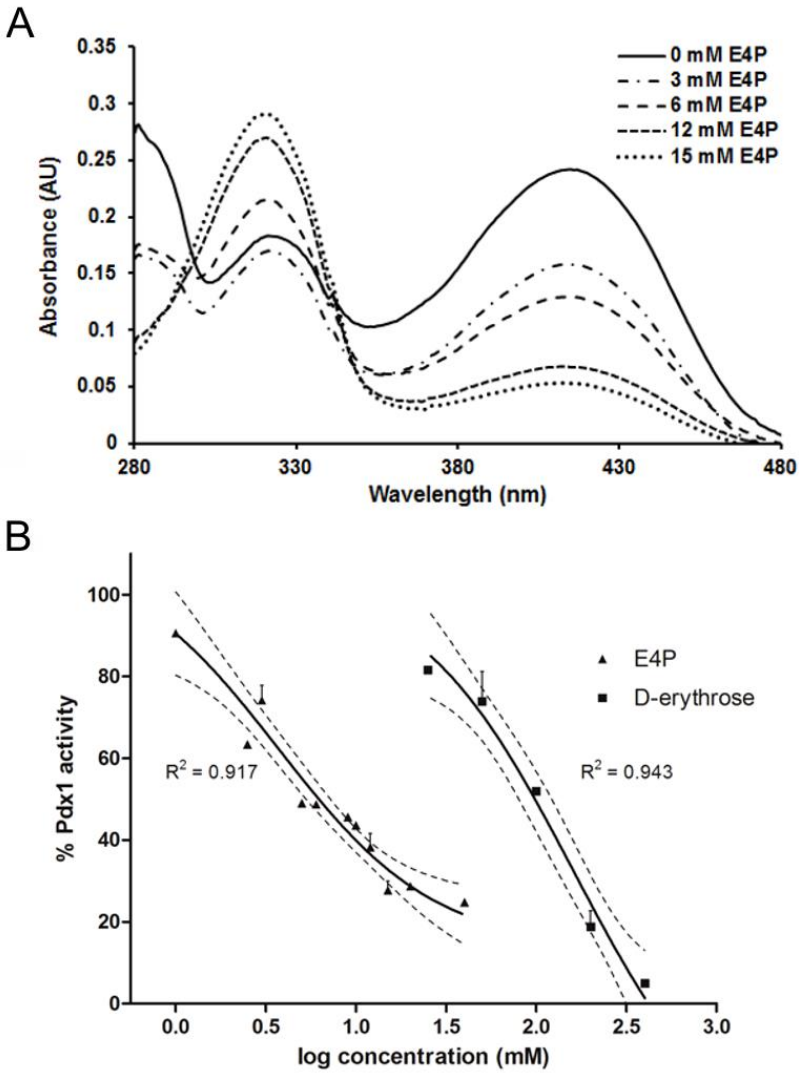


Figure 3.4: Inhibitory effect of E4P on *Pf*Pdx1. A) Representative UV-visible absorbance spectra of *Pf*Pdx1 incubated with increasing concentrations of E4P. Reactions assays contained 0.5 mM R5P, 0.5 mM G3P, 20 mM NH_4Cl with 100 µg of *Pf*Pdx1 and reaction incubation time was 2.5 h at 37°C. **B)** Dose response assays of E4P and D-erythrose on *Pf*Pdx1. Assays contained 100 µg of purified *Pf*Pdx1, with 0.5

mM G3P and R5P, with 20 mM NH_4Cl. Values represent data from three independent experiments performed in duplicate compared to an uninhibited control. Error bars represent the standard error of the mean (SEM) with the 95% confidence interval (CI) presented as dotted lines.

On average at approximately 3 mM E4P (or 375 nmol E4P) there was 50% *Pf*Pdx1 activity remaining compared to the uninhibited control (Figure 3.4 B, with log concentration of 0.477). The total amount of PLP formed in the uninhibited reactions ranged between 9 – 11 nmol (result not shown), and the E4P was therefore present at an excess molar ratio of approximately 37 to 1 compared to PLP. To demonstrate that excess E4P levels did not interfere with the quantitation of PLP incubations with only PLP and E4P in Tris-buffer conditions were set-up (Figure 3.5 A and B). No enzyme was present in these incubations, and served to verify if detection of PLP was affected by E4P.

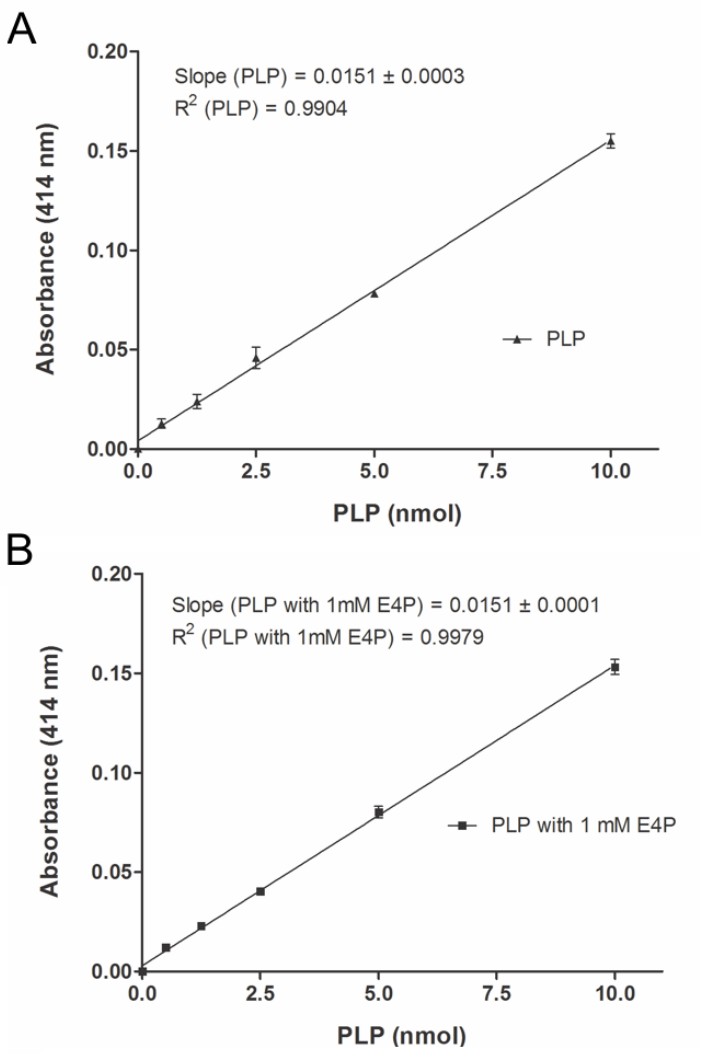


Figure 3.5: Detection of PLP is not affected by E4P. Standard calibration curves of PLP in Tris-Cl buffer conditions with concentrations ranging from 10 µM to 200 µM with **A)** No E4P, and **B)** 1 mM (50 nmol) E4P. Reaction conditions included 0.5 mM R5P, G3P and 20 mM NH_2Cl with incubations at 37 °C for 1.5 h. No *Pf*Pdx1 enzyme was present. Results represent data from four independent experiments performed in triplicate. The E4P-to-PLP molar ratio varied from 80-fold excess to 5-fold excess E4P.

Regression lines of increasing concentrations of PLP, in the absence and presence of molar excess E4P, were generated (Figure 3.5 A and B). Statistical analyses of the linear regression slope and intercepts revealed no significant differences in the slopes of the two different regression lines (P = 0.975). Similarly, ANCOVA showed no significant difference (P = 0.401) between the two regression lines. The lowest point of the PLP standard curve, which has a good fit with the regression line, represented reactions containing 0.5 nmol PLP with 50 nmol E4P, equating to 100-fold molar excess E4P-to-PLP. Even at these high excess E4P conditions the ability to detect PLP was not influenced. If E4P influenced PLP detection the effect would have been noticed as a tapering towards zero at the lower PLP concentration ranges, where conditions of excess E4P are high. This was however not the case. The enzymatic capability of *Pf*Pdx1 to produce PLP was therefore influenced by E4P, leading to further investigations of E4P-analogues and their effect on *Pf*Pdx1, as discussed below.

3.3.4. Characterisation of R5P-I_{320} forming residues

The formation of an I_{320} chromophore, an indication of R5P binding, in *Tm*Pdx1 was previously shown to involve the K83 residue [157]. Similar residues that could be involved in R5P-related I_{320} formation in *Pf*Pdx1 include mechanistically implicated residues such as K83 and D26. Moreover, it was shown that deletion of 31 amino acids from the C-terminal tail of *Pf*Pdx1 abolishes formation of the I_{320} chromophoric intermediate [182], which suggests that I_{320} formation, which follows R5P binding and isomerization, also involves long-range interactions in the protein. Several additional mutant forms of *Pf*Pdx1 were therefore tested to investigate possible involvement of alternative residues in I_{320} formation. These included previously created K83A, K151A, DKK, RHE and ERR *Pf*Pdx1 mutants, as well as the C-terminal tail truncated *Pf*Pdx1$\Delta_{270\text{-}301}$. Additionally, G3P was discovered to occupy a surface exposed site in *S. cerevisiae* Pdx1 (*Sc*Pdx1) [163]. The site for G3P-binding was postulated to include R164, as well as D110 and E116. The equivalent *Pf*Pdx1 residue that could have been involved in G3P-imine formation was R167. To confirm the importance of this residue the enzymatic activity and ability to form the I_{320} specie was tested on a R167A *Pf*Pdx1 mutant (kindly provided by Dr. I.B. Müller).

The *Pf*Pdx1 DKK mutant (D25A, K83A and K151A) did not form the I_{320} species in the presence of R5P, whereas the WT enzyme was catalytically active and in the absence of G3P capable of formatting the I_{320} specie (Figure 3.6 A). Due to the lowered yields from *Pf*Pdx1 DKK expression only 19.8 µg protein could be included in the assays. For this reason lack of I_{320} formation was

confirmed in the K83A and K151A *Pf*Pdx1 mutants, to establish which residue from the DKK mutation plays a role in I_{320} formation. Both the K83A and K151A mutations abolished the formation of I_{320} (Figure 3.6 B). This confirmed the observations in the DKK mutant, and established that both K83 and K151 are essential in initial R5P intermediate formation.

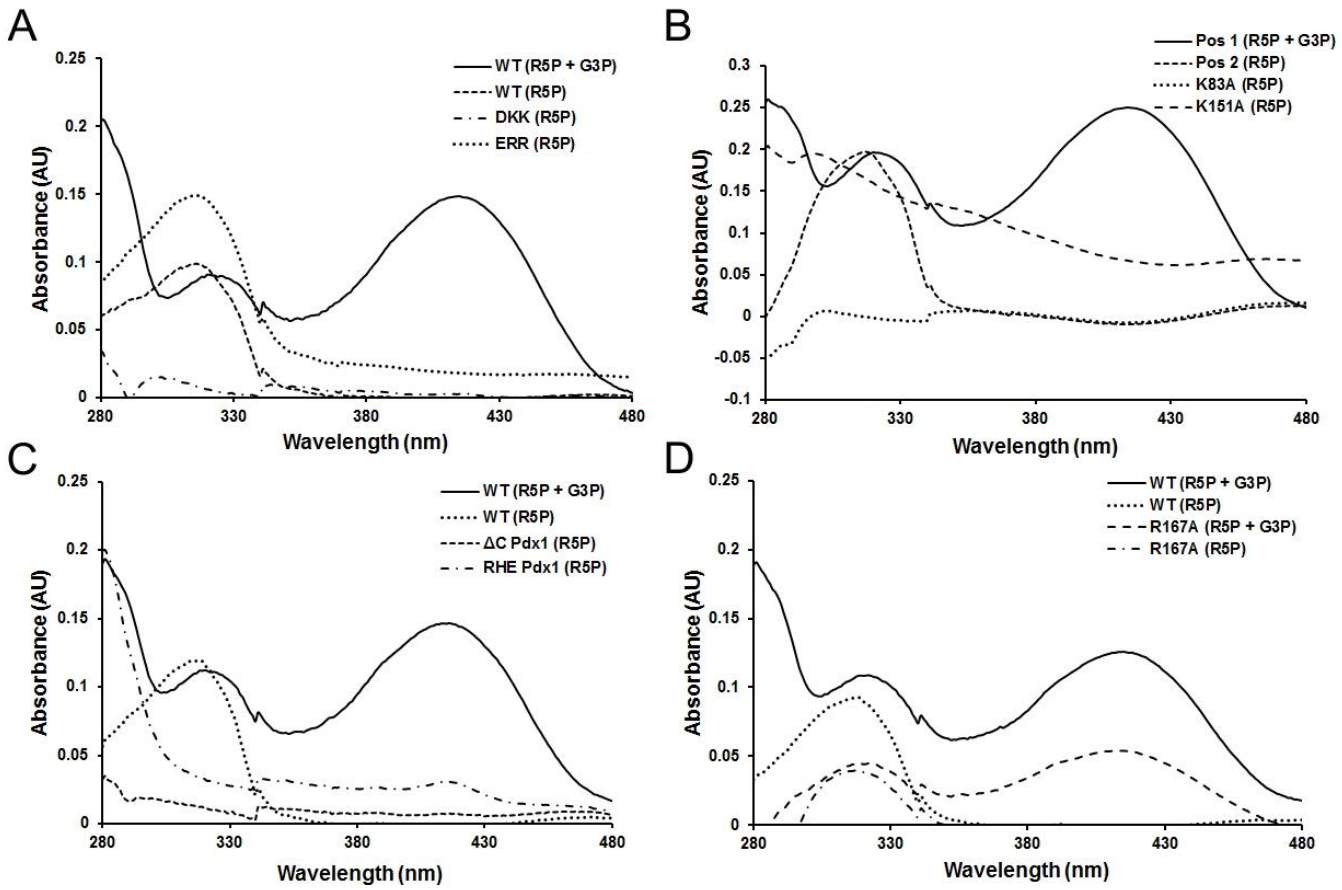


Figure 3.6: Formation of R5P-related I_{320} with *Pf*Pdx1 mutants. A) *Pf*Pdx1 DKK (19.8 µg) and ERR (100 µg) mutants of incubated with 0.5 mM R5P and 20 mM NH_4Cl. Reaction incubation time was 3 h at 37°C. **B)** *Pf*Pdx1 K83A and K151A. The protein mutants (200 µg) were incubated with 0.5 mM R5P and 20 mM NH_4Cl. Higher baseline values for the K151A mutant was attributed to background interferences from the protein purification steps. **C)** *Pf*Pdx1$\Delta_{270\text{-}301}$ and RHE mutants (100 µg each) incubated with 0.5 mM R5P. **D)** *Pf*Pdx1 R167A (100 ug each) incubated with 0.5 mM R5P. In each case positive control reactions included active *Pf*Pdx1 WT (R5P and G3P) with 0.5 mM G3P, 0.5 mM R5P and 20 mM NH_4Cl, as well as *Pf*Pdx1 WT (R5P), in which G3P was omitted to demonstrate active I_{320} formation. Results represent data from experiments performed in duplicate.

The ERR triple mutant *Pf*Pdx1 had comparable I_{320} to WT *Pf*Pdx1, which excluded their involvement in I_{320} formation (Figure 3.6 A). Higher base-line absorbance in the ERR mutant may be attributed to spectral interference which could be related to the properties of the mutant or ineffective removal of some contaminating proteins during protein purification. The relative percentage of the ERR mutant to form I_{320} as listed in Table 3.3 is therefore slightly higher. The truncated C-terminal *Pf*Pdx1$\Delta_{270\text{-}301}$, lacking 31 amino acids, was created, and this was confirmed

through nucleotide sequencing (results not shown) and SDS-PAGE of the recombinantly expressed mutant (Appendix 3, Figure A3.1). The *Pf*Pdx1$\Delta_{270\text{-}301}$ mutant was unable to form the I_{320}, and also not able to generate PLP (Figure 3.6 C). This finding corroborates with previous observations for the C-terminal region of *Pf*Pdx1 [182]. Unexpectedly the RHE *Pf*Pdx1 triple mutant was also unable to form I_{320} or PLP (Figure 3.6 C). Mutation of the R167 residue, believed to be part of the G3P-binding cavity in *Pf*Pdx1, resulted in decreased I_{320} formation (Figure 3.6 D). The R167A mutant was also catalytically impaired, and PLP formation in this variant was around 50% of that of WT *Pf*Pdx1. The properties of *Pf*Pdx1 variants in terms of I_{320} formation, *Pf*Pdx2 activation and PLP formation capabilities are summarised in Table 3.3.

Table 3.3: Relative catalytic capabilities of *Pf*Pdx1 variants. The capability of *Pf*Pdx1 variants to form I_{320} was determined spectrophotometrically and expressed relative to WT *Pf*Pdx1. Additionally, the ability of these mutants to stimulate *Pf*Pdx2 activity and the ability to form PLP is also summarized, with references in brackets. Results represent data from duplicate observations given with the standard deviation.

	R5P I_{320} formation (% compared to WT)	***Pf*Pdx2 activation (% compared to WT)**	**PLP formation (% relative to WT)**
*Pf*Pdx1 WT	100	100 [133]	100
*Pf*Pdx1 ERR	184.3 ± 21.4	104.5 [133]	0 [133]
*Pf*Pdx1 K83A	0.15 ± 0.1	64.6 [133]	0 [133]
*Pf*Pdx1 DKK	24.9 ± 25	103 [133]	0 [133]
*Pf*Pdx1 ΔC	10.1 ± 0.6	99 [182]	0 [182]
*Pf*Pdx1 RHE	28.9 ± 1.6	0 [133]	0 [133]
*Pf*Pdx1 R167A	41.6 ± 6.3	N.D	46.5 ± 3.3

Footnotes: Not determined (N.D.)

3.3.5. E4P analogues and their effect on *Pf*Pdx1

From the E4P lead structure a closely related analogue, 4PEHz [206], was tested on *Pf*Pdx1. The hydrazide derivative was an effective *Pf*Pdx1 inhibitor, and dose-dependently decreased the production of PLP by *Pf*Pdx1 (Figure 3.7 A). This was also accompanied by increases in 320 nm species (Figure 3.7 A). From several independent experiments the IC_{50} of 4PEHz against *Pf*Pdx1 was calculated to be 43 ± 8 µM (Figure 3.7 B). In assay conditions containing both *Pf*Pdx1 and *Pf*Pdx2 with L-glutamine as nitrogen source, the IC_{50} value increased to 16 ± 4 µM (Figure 3.7 C) for 4PEHz. An unphosphorylated analogue of 4PEHz - EHz – was synthesised and used to confirm the importance of the phosphate groups, similar to what was observed for E4P and D-erythrose. The unphosphorylated compound EHz had an IC_{50} of 902 ± 206 µM (Figure 3.7 B) and was more than 20 times weaker than 4PEHz.

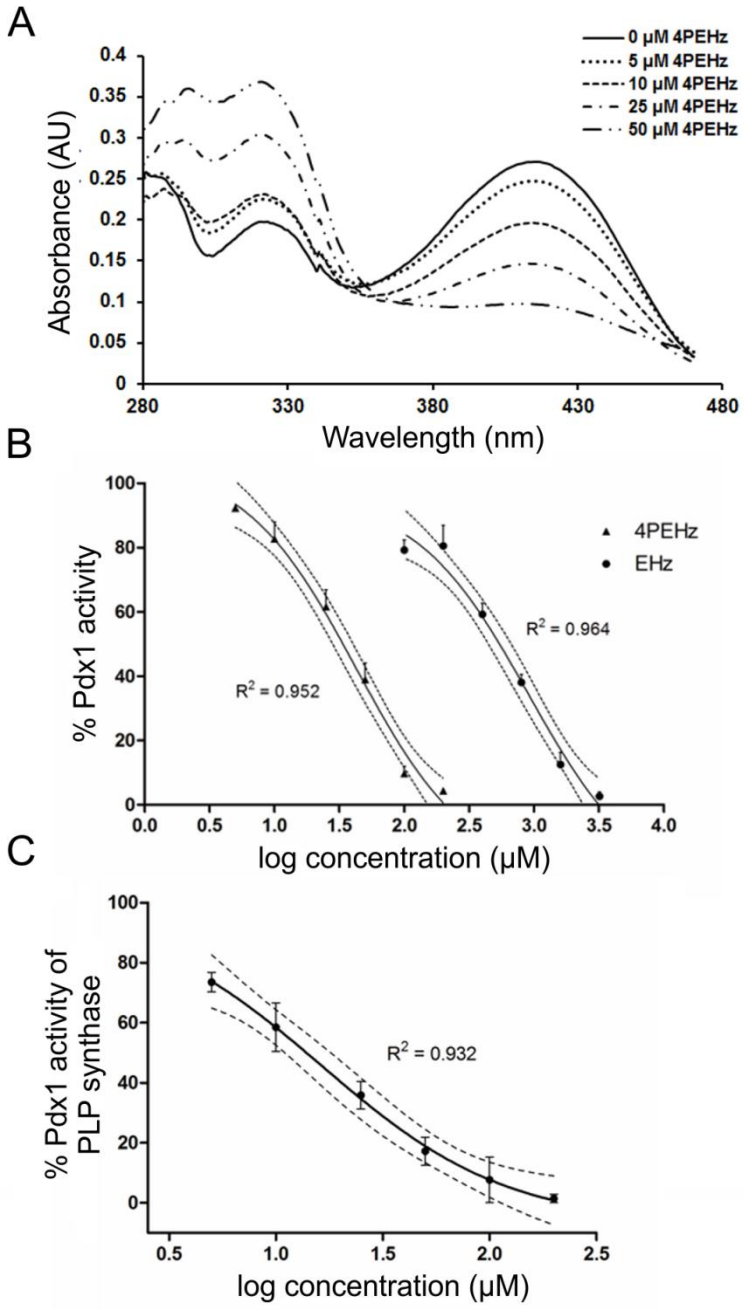
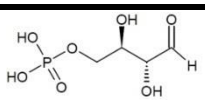
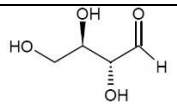
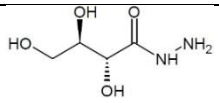
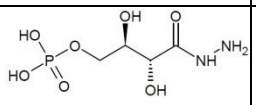


Figure 3.7: E4P analogues and their effect on *Pf*Pdx1. A) Incubation of *Pf*Pdx1 (200 µg) with increasing concentrations of 4PEHz and measured PLP production visible at 414 nm. Results shown represent a single observation from independent triplicate experiments. **B**) Dose-response curves of 4PEHz and EHz inhibition on *Pf*Pdx1 (200 ug). Results were generated from three independently performed experiments, performed in duplicate, with the SEM indicated. The 95% CI is shown as dotted lines. **C)** Effect of 4PEHz on oligomeric *Pf*Pdx1 and *Pf*Pdx2. Reaction conditions included 100 µg of each *Pf*Pdx1 and *Pf*Pdx2 with 0.5 mM R5P, G3P and 20 mM L-glutamine as only nitrogen source. Results were generated from three independent experiments performed in in duplicate, with the SEM indicated. The 95% CI is given in dotted lines.

The inhibitory effect of E4P and analogues against *Pf*Pdx1 are reported in Table 3.4. The trend from the IC_{50} values was that phosphorylated molecules had increased inhibitory capabilities compared to their unphosphorylated counterparts. As mentioned, the unphosphorylated counterparts of E4P and 4PEHz were predicted with lowered binding capability in the *Pf*Pdx1 R5P-binding site (Table 3.4,

see also Chapter 2). This was not so pronounced in the *Pb*Pdx1 model, and rather the unphosphorylated molecules were predicted with overall better LigScores (Table 3.4).

Table 3.4: E4P analogues inhibit *Pf*Pdx1. IC_{50} values were calculated from dose-response curves of the respective compounds incubated with purified *Pf*Pdx1, as reported in the Experimental section. IC_{50} values represent data from three of more independent experiments performed in duplicate, with the SEM indicated. LigScore and Dockscore values were generated from docking in *Pf*Pdx1 (see Chapter 2), with values in parenthesis representing data from docking in *Pb*Pdx1.

Compound	Structure	LigScore1[a]	LigScore2[a]	Dock score[b]	IC_{50} on *Pf*Pdx1
E4P		6.40 (3.00)	6.09 (0.41)	85.17 (120.89)	3.7 ± 0.9 mM
D-erythrose		4.74 (3.78)	4.25(2.15)	42.19 (48.20)	160 ± 46 mM
EHz		5.34 (4.07)	5.39 (3.46)	53.48 (46.63)	902 ± 206 µM
4PEHz		6.55 (3.73)	6.43 (1.46)	98.96 (117.01)	43 ± 8 µM

[a]LigScore1 and LigScore2, with units of pK_i (-log K_i), refers to predicted receptor-ligand binding affinities.
[b]Dockscore refers to the unitless rigid body minimisation energy of the final ligand pose calculated during Monte Carlo trials.

3.3.6. 4PEHz and EHz affect parasite growth

Considering the moderate efficacy of 4PEHz and EHz on *Pf*Pdx1, we tested these compounds *in vitro* against *P. falciparum* (3D7) parasites. Parasite growth was monitored as a function of ^{3}H-labelled hypoxanthine incorporation over a 48 h period, after which the cells were harvested and the amount of incorporated material used to describe parasite growth. The 4PEHz compound had appreciable activity against cultured parasites with an IC_{50} of 10.4 ± 1.2 µM, compared to 138 ± 9 µM as determined for EHz (Figure 3.8). Considering that E4P is a natural occurring molecule within the parasites, excess concentrations were not tested on the parasites.

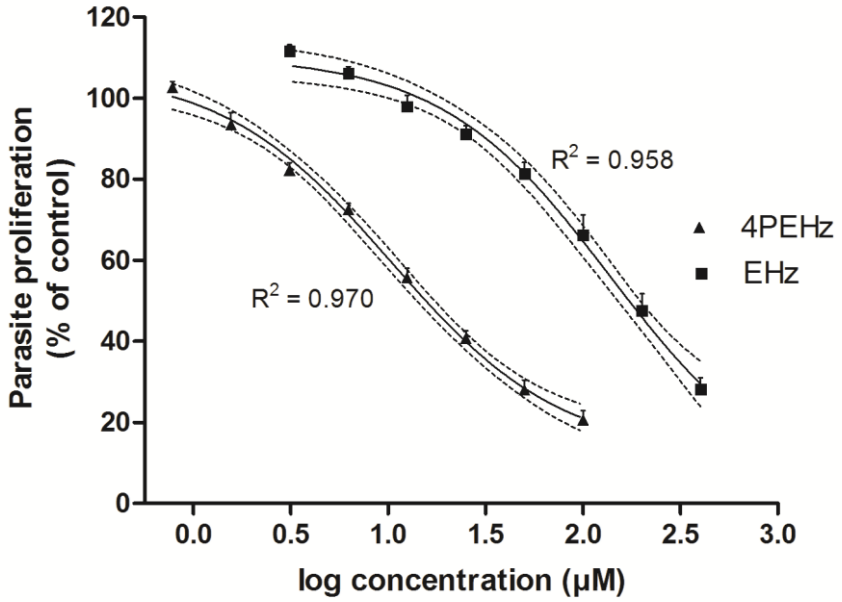


Figure 3.8: Efficacy of 4PEHz and EHz against cultured *P. falciparum* (3D7) parasites. Proliferation was assessed by monitoring the [^{3}H]-hypoxanthine incorporation after 48 h. Values represent data from three of more independent experiments performed in duplicate expressed as the percentage of uninhibited controls.

To further elucidate the effect of 4PEHz *in vitro* and determine the specificity of the compound transgenic *P. falciparum* parasites over-expressing *Pf*Pdx1 and *Pf*Pdx2 (*Pf*Pdx1/*Pf*Pdx2) as well as to control parasites containing a mock plasmid were treated with 4PEHz. The parasite proliferation relative to untreated mock parasites decreased significantly ($P < 0.05$) by 20, 65 and 80% over 48, 72 and 120 h, respectively (Figure 3.9). In contrast, 4PEHz-treated parasites that overexpressed *Pf*Pdx1/*Pf*Pdx2 had unchanged proliferation relative to untreated *Pf*Pdx1/*Pf*Pdx2 parasites (Figure 3.9). This extended period of monitoring was necessary to account for potential delayed drug effects, such as delayed death caused by doxycycline in *P. falciparum* [207], and ensured that the observed rescue effect of *Pf*Pdx1 and *Pf*Pdx2 complementation was sustained through at least three parasite replication cycles. On day 7 the proliferation of the parasites in 4PEHz-treated *Pf*Pdx1/*Pf*Pdx2 parasite appeared to be more diminished compared to untreated cells, however these differences were not statistically significant (Figure 3.9). These results indicate that complementation of PLP metabolism *in vitro* is able to rescue parasites from the effects of 4PEHz which is most likely specifically targeting *Pf*Pdx1 within the parasite.

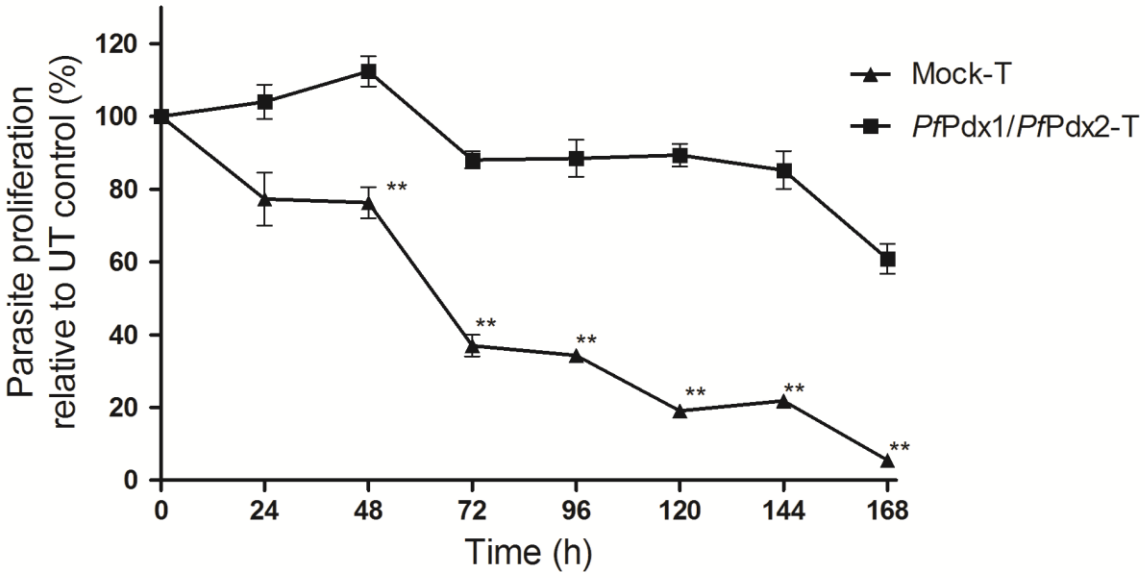


Figure 3.9: Effects of 4PEHz on the proliferation of parasites complemented with *Pf*Pdx1/*Pf*Pdx2- and mock plasmids. The proliferation of mock and *Pf*Pdx1/*Pf*Pdx2 overexpressing parasites treated with 4PEHz are expressed as a percentage relative to the respective UT control cell lines. Results represent averaged values from two independent experiments performed in triplicate with bars indicating SEM and ** indicating $P < 0.05$.

When comparing the slopes of exponential growth curves, and the growth rate constants (K), it was shown that mock-treated parasites (K = 0.0276 h^{-1}) were significantly affected (Figure 3.10 A) compared to untreated mock parasites (K = 0.0412 h^{-1}). The effective doubling time, calculated from the growth rate constant, was 25.10 h in mock-treated parasites compared to untreated parasites with 16.80 h (Figure 3.10 A). Statistical analyses comparing the slopes of the regression lines revealed a significant difference between the slopes (Figure 3.10 A) with $P < 0.0001$. Similarly ANCOVA analyses comparing mock-UT with mock-T revealed that the lines differed significantly (F-ratio probability < 0.0001). These differences in the slope therefore reflected that the growth rates of the mock-treated parasites were significantly diminished compared to untreated parasites. In contrast, the growth rate (K) of the *Pf*Pdx1/*Pf*Pdx2 overexpressing parasites treated with 1 µM 4PEHz (K = 0.0441 h^{-1}) did not differ significantly and remained unchanged compared to that of untreated *Pf*Pdx1/*Pf*Pdx2 overexpressing parasites (K = 0.0457 h^{-1}, Figure 3.10 B). The regression lines of *Pf*Pdx1/*Pf*Pdx2-UT and *Pf*Pdx1/*Pf*Pdx2-T were not significantly different (P = 0.0138). Using ANCOVA the regression lines of *Pf*Pdx1/*Pf*Pdx2-T were compared to *Pf*Pdx1/*Pf*Pdx2-UT and did not significantly differ from each other (F-ratio probability = 0.0986). This suggested that complementation of *Pf*Pdx1 and *Pf*Pdx2 rescued the parasites from the inhibitory effects of 4PEHz.

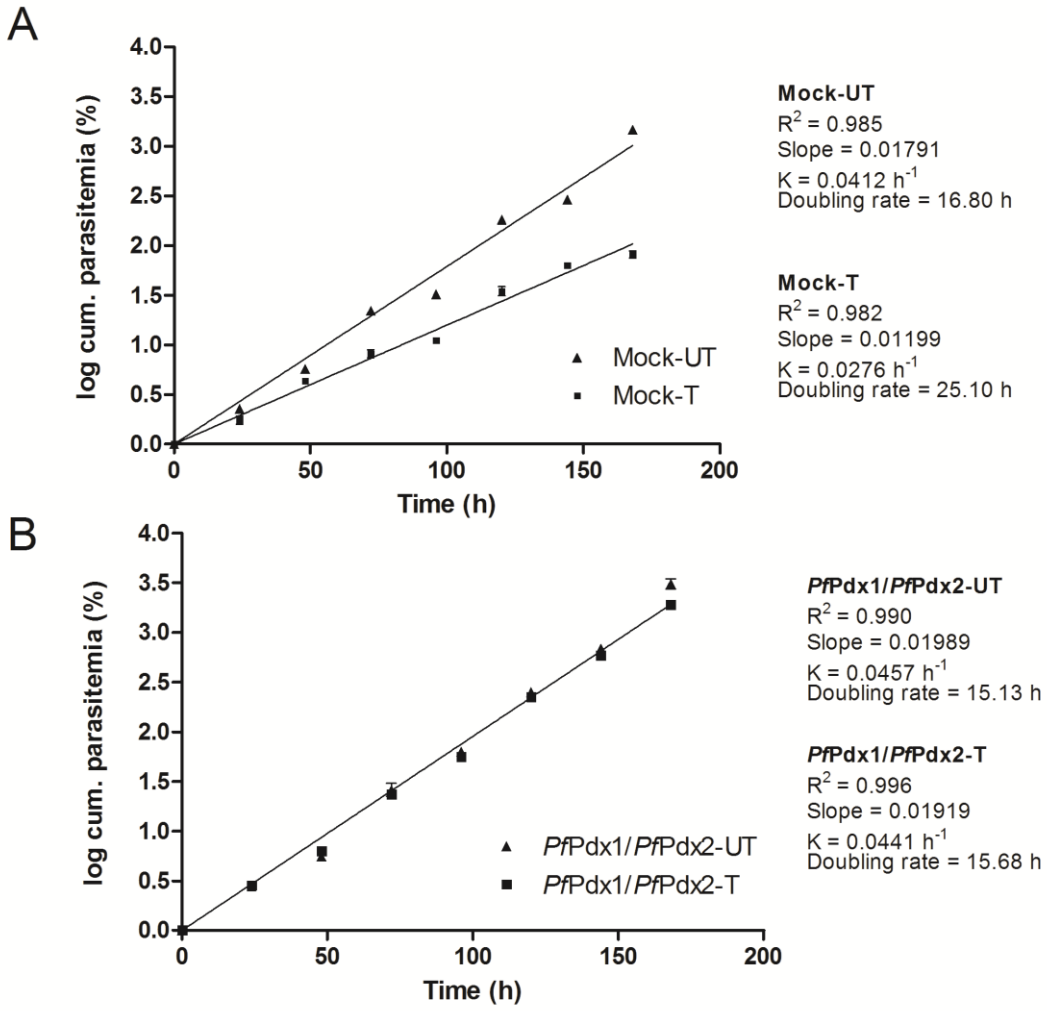


Figure 3.10: Effect of 4PEHz on the exponential growth of transgenic *P. falciparum*. A) Mock control and **B)** *Pf*Pdx1/*Pf*Pdx2 overexpressing cells were treated with 1 µM 4PEHz in continuous culture for seven days as described in Methods section. The cumulative parasitaemia was calculated from the observed parasitaemia by factoring in the dilution during culturing. The data represents average values from two independent experiments performed in triplicate, with the SEM indicated.

The morphological stage composition of the transgenic mock and *Pf*Pdx1/*Pf*Pdx2 parasites were monitored during growth inhibition assays with 4PEHz (Appendix Figure A3.2 and A3.3). The life-stage compositions (percentage of rings, trophozoites and schizonts) of mock-treated parasites was significantly different at 120 h compared to untreated parasites. This was however the only time point affected in the mock-treated parasites. Comparing morphological stages in *Pf*Pdx1/*Pf*Pdx2 parasites, the treatment with 4PEHz did not significantly affect the asexual life-stage compositions (Figure A3.3 A and B).

3.4. Discussion

3.4.1. *Pf*Pdx1 pharmacophore-based ligands

Ligands identified during structure-based pharmacophore screening were tested against purified *Pf*Pdx1. Several sulphonated ligands such as ZINC05273895, ZINC01666581, TAPS and ACES were predicted to bind in the R5P-binding site of *Pf*Pdx1, however did not show appreciable inhibitory activity against *Pf*Pdx1. The hydroxyl substituent in the sulphonate group generally has a pK_a of -1.3, and carries a negative charge at pH 7.0. This group was predicted to mimic the phosphate group of R5P in docking simulations. R5P has two negative charges attributed to the phosphate groups at pH 7.0 and is anchored within the active site with 4 H-bonds to the phosphate group [157]. These molecules, excluding ACES, had bulky terminal groups or tertiary substituted group which could have hindered effective entry into the R5P-active site. ACES contained two amino groups (predicted pK_a of 8.01 and 15.31) which could have influenced the charge of the molecule. At pH 7 the secondary substituted amino group in the aliphatic chain of ACES would be charged, and from inhibition data of *Pf*Pdx1 this is not accommodated in the R5P active site. The other sulphonated compounds, ZINC05273895, ZINC01666581 and TAPS similarly contained secondary amine substituents, and the overall charge of the molecules tended to be neutral or slightly negative. The lack of inhibition by these sulphonated analogues could suggest the charge of the molecule and the stereochemical bulkiness could be important determinants to consider for binding in the R5P site of *Pf*Pdx1.

The polyhydroxyl substituents in R5P, part of the carbohydrate characteristic, functions primarily as H-bond donors, as previously demonstrated in related Pdx1 structures [157]. *Pf*Pdx1 pharmacophore screening successfully identified comparable compounds with polyhydroxyl substituents and docking predicted that some of these could bind in the R5P-active site. These molecules included ZINC0158679, which was the highest scoring compound from pharmacophore screening, as well as ZINC04683167 and ZINC03137815. None of these compounds showed significant inhibitory activity against *Pf*Pdx1 during *in vitro* assays. Binding of R5P (or Ru5P) to K83 requires the formation of a Schiff base, necessitating a terminal aldehyde or ketone arrangement. This could be an important feature in Pdx1 proteins to selectively bind the substrate. The lack of similar groups in the polyhydroxylated analogues could have blocked these from entering the active site.

Several ligands identified from the structure-based pharmacophore screening had aromatic character, such as ZINC03897990, ZINC00093092, ZINC0471739, ZINC09680397 and ZINC00389974. During docking the ZINC03897990 ligand was predicted with better binding

scores compared to R5P. The lack of efficacy of these ligands suggested that these are not effectively accommodated in this site, or any *Pf*Pdx1 site. The pyrimidine compound ZINC00093092 marginally and significantly increased *Pf*Pdx1 activity; the compound could potentially act as an allosteric activator, however no tangible conclusions can be drawn without further investigation.

3.4.2. Substrate preference of *Pf*Pdx1

The Pdx1 reaction mechanism has been elucidated and is summarised in Figure 3.1. A recent mechanistic proposal suggested that initial imine formation with the terminal aldehyde group of R5P occurs with the Pdx1 K83 residues, followed by isomerisation into a Ru5P imine adduct [160]. This bound Ru5P-intermediate is converted into the characteristic I_{320} UV-absorbing chromophore. Hanes *et al.* used isotopically-labelled R5P and NMR to determine the identity of I_{320} [160]. This intermediate forms after the Ru5P-intermediate undergoes a C1 to C5 imine migration, as shown in the second to third transition in Figure 3.1. Loss of the phosphate group from this intermediate results in the formation of I_{320} (Figure 3.1). Only R5P and ammonia are required to form I_{320}, and incubation of Pdx1 with only these substrates produces the characteristic UV-absorbing species.

The fact that Pdx1 has R5P isomerase activity which leads to the formation of I_{320} suggests that the α-hydroxy aldehyde terminal arrangement on R5P, capable of undergoing conjugation, is essential for initial substrate binding and I_{320} formation. [104, 162]. We tested whether other sugar substrates were accepted by *Pf*Pdx1 in the presence of G3P and were expected to be capable of undergoing at least partial imine formation with internal K83 residues in *Pf*Pdx1. DR5P was able to inhibit *Pf*Pdx1 reflecting that some competitive interactions for internal imine formation could exist. The weak inhibition by DR5P could be related to the inability to undergo isomerisation to a Ru5P adduct or the fact that the internal I_{320} specie cannot be displaced by DR5P. With terminal aldehyde functional groups DR5P could compete for K83 Schiff base formation, however, the lack of the C2 hydroxyl substituent could inhibit further isomerisation. This is supported by the fact that DXP, equivalent to the Ru5P scaffold structure apart from the stereochemistry of a single hydroxyl substituent and a terminal ketone group, was not effective at inhibiting *Pf*Pdx1. The initial formation of the Schiff-base requires a reactive aldehyde, which is only present on DR5P, and not DXP. It was therefore surprising that fosmidomycin was ineffective against *Pf*Pdx1, as this compound contains a terminal aldehyde moiety required for initial binding except the one carbon length difference compared to R5P. The molecule however contained more aliphatic character,

potentially prohibiting active site entry, and it was hypothesised that hydrolysable phosphate groups are required for binding as is discussed in more detail later in this section.

E4P was an effective inhibitor of *Pf*Pdx1 with an IC_{50} of 3.7 mM, E4P is structurally similar to both R5P and G3P with a terminal α-hydroxy aldehyde arrangement. The terminal α-hydroxy aldehyde and stereochemical arrangement of the hydroxyl backbone together with hydrolysable phosphate groups in E4P was suggested to be important functional groups to effectively inhibit *Pf*Pdx1. This was supported by the fact that D-erythrose was almost 50-times weaker than E4P emphasising that the hydrolysable phosphate group is also essential.

E4P was also shown to promote the formation of a chromophoric species at 320 nm – this specie formed independently whether *Pf*Pdx1 was present or absent. E4P was previously shown to dimerise into a cyclic hemiacetal form [208]. Moreover, these 320 nm species were formed in the absence of PLP, excluding the possibility of PLP-E4P Schiff-base adducts. Whether E4P dimerises, or reacts with Tris to form these chromophoric species is therefore unknown. In order to demonstrate that the E4P chromophore does not interfere with the detection of PLP molar excess conditions of E4P were used with PLP standard curves. E4P was shown not to interfere with detection of PLP in Tris buffer conditions used during the PfPdx1 enzyme assays. This therefore discounted E4P or Tris-adducts from potentially reacting with PLP, which could have skewed the *Pf*Pdx1 enzyme assays. The results therefore suggested as a linear aldose E4P affected *Pf*Pdx1, possibly by mimicking R5P and could be able to bind to the K83 in the R5P-active site.

Parker *et al.* showed that E4P was capable of forming a Schiff base with a lysine (K186) residue in 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS) thereby irreversibly inactivating the enzyme [209]. Phosphoglucoisomerases are also inhibited by E4P [210]. The PGI active site are remarkably similar to that of triosephosphate isomerases (TIM), moreover the PGI active site contains lysine residues speculated to be involved in the reaction coordinate [211-213]. Considering the TPI activity of *Pf*Pdx1, capable of utilising both DHAP and G3P, this activity of *Pf*Pdx1 could be targeted by E4P. The definitive location for G3P isomerisation has not yet been established, however this could involve residues located in the P3 site. This therefore allows speculation into the MOA of E4P. There are three potential ways in which E4P can inhibit *Pf*Pdx1; either outcompeting R5P for K83 imine formation, and also undergoing isomerisation. This would affect binding of R5P and lead to disrupted PLP formation. An alternative route could involve E4P binding to the postulated G3P-binding site K120 (equivalent to *Sc*Pdx1 K117) residue, whereby E4P would mimic G3P, interfering with downstream PLP ring-closure steps [163]. The susceptibility of PGI, with TPI

likeness, provides a third alternative, suggesting an unknown site in which isomerisation of G3P occurs, where E4P can potentially bind.

Parasitic E4P is a vital precursor, for the production of chorismate, and downstream phenylalanine, tyrosine and p-aminobenzoic acid utilised for the production of folates [214]. Previous studies have failed to identify homologous sequences of transaldolases in *P. falciparum*, required for the production of E4P from sedoheptulose 7-phosphate [215]. Instead a transketolase (PFF0530w) was cloned and characterised from the parasites and were proposed to produce E4P from F6P and G3P [216]. Earlier studies have reported the presence of 3-deoxy-7-phosphoheptulonate synthase, which condenses E4P and phosphoenol pyruvate as part of the initial stages in the shikimate pathway [217]. Under physiological conditions E4P could therefore regulate PLP production through inhibition of *Pf*Pdx1.

The E4P analogue - 4PEHz - was shown to inhibit *Pf*Pdx1 within low micromolar efficacy. This was also supported by molecular docking simulations. Potentially 4PEHz binds to *Pf*Pdx1 and influences the passage of labile ammonia from the Pdx2 monomer, compared to relatively unrestricted mobility of ammonia in *Pf*Pdx1 only incubations. 4PEHz was previously reported as a weak inhibitor of spinach R5P isomerase with a K_i of 1.8 ± 0.2 mM [206, 218]. Our findings suggest that 4PEHz occupies the same site as E4P, and is supported by the fact that hydrolysable phosphate groups improved binding characteristics. EHz and 4PEHz can potentially form Schiff-bases with PLP through linkage with hydrazide groups. This could be the mechanism by which either compound inhibits the enzyme. These adducts may form in either the P2 and/or P3 sites of the enzyme, leading to PLP-conjugation, and thereby affect the enzyme activity. However, considering that both compounds contain hydrazide functional groups, the difference of inhibitory efficiency suggests some specificity towards *Pf*Pdx1 involving the phosphate group.

During oligomerization of Pdx1 and Pdx2 into the functional PLP synthase Pdx1 activates the glutaminase activity in Pdx2 monomers leading to the formation of a putative opening through which ammonia is channelled [156]. Oligomerization also leads to increased R5P binding affinity in Pdx1 [219]. We noted increased inhibitory efficacy of 4PEHz in the *Pf*Pdx1:*Pf*Pdx2 complex suggesting improved binding interactions of 4PEHz, similar to previous observations for R5P, and further supports that 4PEHz may bind in the R5P binding site.

4PEHz inhibited parasite growth at low micromolar concentrations, compared to the unphosphorylated EHz analogue which was at least 12 fold weaker. The IC_{50} of 4PEHz against cultured parasites was comparable to the IC_{50} against the oligomeric *Pf*Pdx1/*Pf*Pdx2 enzyme complex. Phosphorylated molecules are expected to have poor pharmacokinetic properties, and

passage through the parasite membranes may be hindered. For this reason it was thought that EHz might be a better growth inhibitor than 4PEHz, which was not the case. Some phosphorylated metabolites like AMP can be taken up by parasites, however the exact biological mechanism of transport or protein-mediated transport/uptake still remains unknown [220]. How 4PEHz enters the parasites remains unknown, and ideally isotopic labelling studies would be necessary to study the route of entry.

EHz affected the proliferation of the parasites with an IC_{50} of 138 µM, whereas the IC_{50} for inhibition of *Pf*Pdx1 was 902 µM. The efficacy of EHz *in vitro* was therefore more potent than expected. EHz was predicted to have improved pharmacokinetic properties compared to 4PEHz owed to the lack of the phosphate group, and results presented support the *in vitro* efficacy of EHz. It could however not be ruled out that EHz had alternative intraparasitic targets. 4PEHz was similarly more potent in growth inhibition assays compared to enzyme inhibition results. The potentiated effect of 4PEHz in the parasite could be related to the parasitic dependency on B_6, with inhibition of PLP biosynthesis having severe consequences, resulting in parasite death.

The complementation of PLP biosynthesis through overexpression of *Pf*Pdx1 and *Pf*Pdx2 in *P. falciparum* parasites *in vitro* was previously shown to increase tolerance to cercosporin-induced oxidative stress [140, 141]. Here we similarly observed that parasites which over-express *Pf*Pdx1 and *Pf*Pdx2 were not affected by the *Pf*Pdx1-inhibtor 4PEHz. In contrast, parasites harbouring the same expression plasmid, without *Pf*Pdx1 and *Pf*Pdx2, had reduced growth rates and were significantly attenuated by 4PEHz. The doubling times of the mock parasites were 25.1 h compared to untreated mock parasites with 16.8 h. The doubling times should not be misinterpreted as the time in which the parasite number doubles, rather as a mathematical term which allows comparison of proliferation. The life cycle of the parasite is 48 h, and these parasites cannot double in 16 h. Factored in with the multiplicity of parasite expansion, these doubling times refer to the effective doubling rate over the course of seven days. The overall effect of 4PEHz on the doubling time serves as a comparison between the same cell lines only. 4PEHz therefore affected the doubling times of parasites lacking additional *Pf*Pdx1 and *Pf*Pdx2. Almost 3-fold greater PLP levels were previously reported in *Pf*Pdx1/*Pf*Pdx2 complemented parasites [141]. This complementation of PLP synthesis *in vitro* counteracts the effects of 4PEHz, and suggests that 4PEHz interferes primarily with this part of the parasite metabolism. This suggested that 4PEHz is specific against the *de novo* PLP pathway, and with supporting information from *Pf*Pdx1 inhibitory studies, the compound was suggested to target *Pf*Pdx1 in the parasite. These results support 4PEHz as a novel lead compound for targeting vitamin B_6 biosynthesis within the malaria parasite. How 4PEHz affects functional pathways in the parasite was further analysed using transcriptomics and proteomics in Chapter 4.

3.4.3. Mutagenesis of *Pf*Pdx1

Recent elucidation of Pdx1 has suggested K149 is involved during intermediate transfer towards the P2 site, whereas D26 and K83 constitute the P1 binding site [157, 221]. Our results confirmed residues involved in I_{320} formation include DKK, and more specifically both K151 and K83, equivalent to active site residues *Tm*Pdx1. This corroborates with previous observations elsewhere [157, 162]. Residues located in the PLP-binding site of *Pf*Pdx1 include E136A, R139A and R140A and these were only recently shown to be involved in intermolecular contacts with PLP [157, 163, 221]. The ERR *Pf*Pdx1 triple mutant (E136A, R139A and R140A) was capable of forming the I_{320} intermediate and confirmed the lack of involvement of these residues in R5P isomerisation, however, underscored their catalytic importance more than likely during PLP binding and release.

The mutagenesis of *Pf*Pdx1$\Delta_{270\text{-}301}$ corroborated with previous results which showed that the C-terminal deletion results in abolished I_{320} formation [182]. Additionally, the R164 residue which is proposed to be part of the G3P binding site or P3 site in Pdx1 was also suggested to be important in I_{320} as well as PLP formation. This residue is part of the Pdx1:Pdx1 contact region which G3P is located, and in the *Sc*Pdx1 structure this residue is located more than 20 Å from the P1 or R5P binding site [163]. The effect on I_{320} formation of *Pf*Pdx1 could therefore suggest that some long range interactions could be involved from the P3 site that promote R5P binding and initial substrate modulation.

Surprisingly, the RHE triple mutant did not form I_{320}, and sheds new light on the importance of these residues in *Pf*Pdx1. The RHE residues, located on the loop between β3 and α3, was previously shown to confer dodecamer assembly, and subsequently dictated *Pf*Pdx2 binding and activation [133]. Lack of I_{320} formation in the RHE mutant indicates that these residues are intricately involved in R5P binding, additionally taking part in active site reorganisation during dodecamer assembly. Indeed a glycine residue (G84) located on the same loop region as the RHE residues was shown to have H-bonded contacts with D110 which forms part of the G3P P3 binding site in *Sc*Pdx1 [163]. These RHE residues in *Pf*Pdx1 may therefore participate in long-range interactions, communicated through the G3P binding site, which results in the formation of I_{320}.

3.5. Conclusions

Compounds were identified that could inhibit *Pf*Pdx1. These compounds were structurally similar to R5P and had reactive C1 groups, such as aldehyde and hydrazide groups. The MOA was proposed to involve binding in the R5P-binding site of *Pf*Pdx1. The most potent inhibitor of *Pf*Pdx1 attenuated the proliferation of *in vitro* cultured *P. falciparum* parasites, and was suggested to target PLP biosynthesis. Transcriptomics and proteomics were used to further verify whether the compound affected *Pf*Pdx1 and PLP-related processes and is discussed in Chapter 4.

Chapter 4
Functional consequences of perturbing vitamin B6 biosynthesis in *P. falciparum* parasites.

4.1 Introduction

4.1.1 Transcriptomics in the malaria parasite

Integrated systems biology is becoming increasingly essential in describing the functional organisation of drug-induced changes by integrating X-omics approaches; genomics, transcriptomics, proteomics, metabolomics and bioinformatics [222, 223]. Transcriptomics aims to describe the dynamics of transcript expression and co-expression, as well as quantities of the transcripts, which consists of mRNA, rRNA and non-coding RNA components [223]. Currently transcriptomics provides the most comprehensive view of changes within the cell, and provides insight into the functions and dynamics of the expressed genome [224]. Microarrays, serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS), quantitative reverse transcriptase (RT)-PCR (qRT-PCR) and whole transcriptome shotgun sequencing (WTSS or RNA-seq) are some transcript profiling methodologies [224]. SAGE, MPSS, as well as RNA-seq are methodologies that do not require prior knowledge of sequence information, and are useful approaches on non-model organisms which have not been sequenced. Microarrays are more affordable, high throughput and can interrogate transcriptional changes in the whole genome, however rely on prior sequence information to design oligos [224]. The genome of *P. falciparum* has been sequenced and revealed that it consisted of 5400 protein-encoding genes, of which 60% were found to be transcriptionally active during the IDC [225]. This has led to the implementation of a large number of microarray profiling studies to mine information regarding the *P. falciparum* transcriptome.

Microarray-based transcriptomic studies on *P. falciparum* parasites have aided in understanding differential transcript expression in diverse life-stage developmental forms [225-228], in distinct parasites strains [229], as well as transcript copy number variations and polymorphisms in different laboratory isolates [230]. Different pathogenicity factors in field isolates [230, 231], as well expression of transcripts potentially involved in drug transport in drug-resistant strains [232] have been elucidated using microarray studies. Microarrays have been used to correlate *in vivo* parasite proliferation to *in vitro* culturing conditions [233], and the effect of different environmental factors on *in vitro* cultured parasites [234, 235]. Some important features identified include monocistronic transcript expression patterns of most chromosomal genes during the IDC [225]. These transcripts

reach maximal expression only once during the 48 h cycle, whereas apicoplast gene expression patterns were shown to be polycistronic [225].

A large number of microarray studies on *P. falciparum* have focused on elucidating transcriptomic responses due to drug pressure as well as drug resistance mechanisms [207, 236-241]. With this information the compensatory transcriptome adaptations due to chemical knock-down of the target could be elucidated, which should aid in creating more effective drugs, as well as drug combinations. Decreased transcript expression during drug treatment could suggest that these processes potentiate the activity of a compound. A hypothetical example is decreased transcription of membrane transporters that facilitate import of the active compound. Increased expression of transcripts could facilitate removal of the drug, as has been demonstrated for mefloquine resistant parasites which have increased the *P. falciparum* MDR transporter gene copy numbers [77].

Studies have shown that the *P. falciparum* transcriptome is dynamic, and that different small molecules have profoundly different effects on transcript expression patterns. Hu *et al.* constructed interaction networks or gene-function prediction networks based on global expression patterns from a wide range of different small molecule growth inhibitors [242]. These different molecules stimulated distinctly different sub-populations of transcripts, and provided insights into the MOA of similar classes of inhibitors. Interesting small molecules such as apicidin and trichostatin A – both histone deactelyase inhibitors (HDAC inhibitors) – with different chemical scaffolds were shown to stimulate very similar transcript expression profiles. Other molecules such as quinine and CQ induced different transcription patterns compared to the former compounds, but were similar compared to one another. It was also elegantly shown with microarrays that the mechanism of drug resistance against DHA involves increased *pfmdr*1 transcript expression, confirming that this transporter is involved in conferring resistance to the drug [238]. The delayed cytotoxic responses of tetracyclines such as doxycycline on parasites were elucidated using microarrays and attributed to impaired apicoplast gene transcription [207]. A microarray study aimed at determining the MOA of a potent piperazine compound ACT-213615 showed that the drug induced differential expression of around 550 genes in late schizonts [237]. Even though the definitive MOA could not be determined, it was shown that the compound has a distinctive transcriptomic responses apart from artemisinin or CQ [237].

In some cases, it has been shown that the transcriptome of parasites is relatively unresponsive due to drug treatment. The antifolate WR99210 drug which specifically targets dihydrofolate reductase-thymidylate synthase (DHFR-TS) did not affect the target transcriptional expression [241]. The folate cofactor, tetrahydrofolate (THF), which is recycled by DHFR-TS, is essential for production

of nucleobases and inhibition of this process could have resulted in a paralysed state. In CQ-treated parasites 600 transcripts were differentially expressed, of which a only small percentage (~3%) had fold change (FC) values above 4 [240]. The authors failed to identify clear CQ-responsive networks linked to low level transcripts expression, however suggested that small changes in transcript abundances could have profound synergistic effects on cell physiology. The lack of differential transcript expression of the known targets in each case led the authors to speculate that the transcriptome is rigid and unresponsive. In contrast, this could provide clues into the mechanism of action of CQ; the drug might affect a master transcription factor and thereby rendering parasites unable to differentially express transcripts, or the toxic accumulation of heme induces a similar arrest in transcription.

In this chapter, a functional genomic study was undertaken to determine the global effects of treatment of the parasite with 4PEHz, and whether inhibition of *Pf*Pdx1 affects processes related to PLP metabolism. The compound was shown to have some specificity for *Pf*Pdx1, and inhibition of protein activity is expected to disrupt metabolic processes which could lead to alterations in transcript expression patterns. Additionally, proteins from 4PEHz-treated parasites were extracted to determine the extent of translational and post-translational effects caused by the anticipated PLP-depleted state in the parasites.

4.2 Methods

4.2.1 Parasite culturing and drug treatment

P. falciparum (3D7) parasites were maintained under aseptic conditions in continuous culture according to Trager *et al.* [201], as previously indicated by Smit *et al.*[243]. The RMPI 1640 culture medium contained 25 mM HEPES, 10 mM glucose, 20 mM sodium bicarbonate, 25 mM hypoxanthine and 0.5% w/v AlbuMAX II (Invitrogen) at pH 7.4. Parasite cultures were maintained at 5% haematocrit using human O^+ erythrocytes in 75 cm^3 flasks (Nunc) with shaking at 200 rpm at 37°C. Before drug treatments, the synchronicity of the parasite life-stages was controlled by three consecutive D-sorbitol (5% w/v in 1x PBS) treatments, each 8 h apart in early and late ring-stage parasite cultures. Briefly, parasite cultures (50 mL, 1-5% parasitaemia, ring-stage parasites) were collected by centrifugation at 3000*g* for 5 min, supernatants discarded and pellets re-suspended in 5 volumes (10 mL) of 5% D-sorbitol and incubated at RT for 10 min. The cultures were centrifuged as before and the pellets were re-suspended in complete culture media. Flasks were deoxygenated using 5% CO_2, 5% O_2 and 90% N_2 and kept in a shaking incubator (120 rpm) at 37°C.

The highly synchronous cultures (approx. 95% schizonts) were dived into treated (T) and untreated (UT) groups at 10% parasitaemia, 5% haematocrit. The treated group was treated with 20 µM 4PEHz (equivalent to 2 x IC_{50}) at the start of the ring stages (defined as 0 hours post invasion, HPI). After 24 h, the media was replaced with fresh media, maintaining drug pressure for the 4PEHz treatments. Sampling was done by harvesting infected erythrocytes at three different time points each 12 h apart (t_1 = 12 HPI, t_2 = 24 HPI and t_3 = 36 HPI) corresponding to rings, trophozoites and schizont forms of the parasite. This entailed removing 10 mL culture from the flasks and washing the harvested infected erythrocytes twice with 10 pellet volumes of cold 1x PBS as above to remove any remaining drug or culture media components. The washed infected erythrocyte pellets were flash frozen in liquid nitrogen and subsequently stored at -70°C. Additional 5 mL culture samples were harvested from both T and UT groups of each time point for proteomic analyses. Experiments were performed in duplicate, with two technical replicates for each time point.

4.2.2 Transcriptomics

4.2.2.1 RNA isolation

RNA was isolated from infected erythrocytes using Tri-reagent (Sigma-Aldrich) combined with RNeasy mini kit (Qiagen) according to the manufacturer's protocol for isolation of total RNA from animal tissues using on-column DNase-1 digestion (Qiagen), as reported previously [243, 244].

Frozen infected erythrocyte pellets were thawed at RT and gently agitated to facilitate thawing. The cells were lysed through the addition of 600 µL lysis buffer (RNAeasy kit, Qiagen) and agitated by vortexing. The lysed cell mixture was pipetted onto a QIA-shredder column, and the eluent collected after centrifugation at 13 000*g* for 2 min was separated into two equal parts in microcentrifuge tubes. Tri-reagent (Sigma-Aldrich), measuring to 600 µL, was added to each tube, followed by incubation at RT for 5 min. Tri-reagent is a RNA isolation agent that contains guanidinium thiocyanate and phenol that denatures proteins and inhibits RNAses, and following addition of chloroform allows separation of mostly RNA into upper aqueous phases, originally developed by Chomczynski *et al.* [245]. Chloroform (400 µL, Merck) was added to the tubes and these vigorously vortexed for 15 s. The chloroform-containing tubes were incubated at RT for 10 min, centrifuged at 13 000*g* for 15 min and the upper aqueous phases transferred to a clean microcentrifuge tubes. Ethanol (700 µL of 70% v/v, Merck) was added to each tube to precipitate RNA, and the contents of the previously split tubes combined into one tube. The solution (700 µL) was consecutively loaded onto a RNAeasy spin column and centrifuged for 8000*g* for 15 sec until all of the solution was passed through the column. The spin column was washed by adding 350 µL wash buffer RW1 (Qiagen) followed by centrifugation as before. On column DNase-1 digestion was performed with 80 µL DNase-1 solution in buffer RDD (Qiagen, 10 Kunitz units) at RT for 15 min. The digested genomic DNA was washed from the columns using 350 µL wash buffer RW1, followed by a final 2x washing in 500 µL buffer RPE (Qiagen) and drying of the columns by centrifugation at 8000*g* for 2 min to ensure complete removal of remaining liquid. RNA was collected by adding 30 µL of RNase-free water (RNAeasy mini kit, Qiagen) allowing incubation for 2 min at RT, followed by centrifugation at 8000*g* for 2 min, repeated twice. The collected RNA was stored at -70°C.

4.2.2.2 Assessment of RNA purity and quality

RNA was quantified using Nanodrop® ND-1000 (*NanoDrop* Technologies), which uses the UV absorption of the sample at 260 nm, together with the molecular extinction coefficient of RNA, to calculate the RNA concentration. The quality and purity of the RNA was assessed using the Experion automated electrophoresis system (Bio-Rad), which involves; separation, staining and destaining, RNA band detection, imaging and data analysis [246]. RNA (200 ng) was loaded onto the Experion™ RNA StdSens Lab-Chip microfluidic chip (Bio-Rad) according to the manufacturers recommendations. Loaded samples are labelled with a fluorescent dye enabling the detection of the RNA through laser-induced fluorescence. RNA quality is qualitatively and

quantitatively assessed, the former reported as the RNA quality indicator (RQI) which determined by analysing the integrity of the ribosomal RNA, the latter determined by comparison to an internal molecular weight standard.

4.2.2.3 cDNA synthesis

First strand cDNA synthesis was performed by mixing 4 µg RNA, 250 pmol OligodT primer ($dT_{23}VN$, Inqaba Biotechnical Industries) and 775 pmol random nonamer (N_9, Inqaba Biotechnical Industries) which was incubated at 70°C for 10 min, followed by incubation at 4°C for 10 min. A 10× aminoallyl dNTP stock was prepared with final concentrations of 5 mM 5-(3-aminoallyl)-2ˈ-deoxyuridine-5ˈ-triphosphate (aa-dUTP), 5 mM dTTP, 5 mM dGTP, 5 mM dCTP, 10 mM dATP. After incubation of the RNA-containing tubes 6 µL 5× Superscript first strand reaction buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM $MgCl_2$), 10 mM DTT, 40 U RNasin (40 U/µL, natural protein-based RNase inhibitor, Promega), 1.7 µL 10× aminoallyl dNTP stock (0.28 mM aa-dUTP, dTTP, dGTP, dCTP and 0.56 mM dATP) and 340 U Superscript III (Invitrogen) was added, to a final volume of 30 µL with RNase-free water. The reverse transcriptase reactions were incubated at 42°C for 20 h. Remaining RNA was hydrolysed by adding a final concentration of 142 mM NaOH (Fluka) and 71 mM EDTA (pH 8.0, Merck), followed by incubation at 65°C for 15 min. The synthesised cDNA was subjected to a clean-up procedure to remove hydrolysed RNA, using the PCR Clean-up kit (Macherey-Nagel). Briefly, 400 µL buffer NT was added to the cDNA-containing tubes, after which the sample was transferred to the Nucleospin extract II spin column and incubated at RT for 4 min. Under these chaotropic salt conditions the cDNA selectively binds to the silica matrix, allowing removal of any unbound material. The spin columns were centrifuged at 12 000*g* for 2 min, after which the membrane was washed using 500 µL buffer NT, followed by centrifugation as above. Columns were dried with a final centrifugation step at 12000*g* for 2 min. Pre-heated RNase-free water (30 µL) was pipetted onto the membranes, allowing incubation for 4 min at RT, followed by centrifugation at 12000*g* for 90 s. The cDNA concentration and purity was determined using a Nanodrop® ND-1000 (*NanoDrop* Technologies), which uses the UV absorption of the sample at 280 nm, together with the molecular extinction coefficient of DNA, to calculate the DNA concentration. A cDNA reference pool comprising of all the time points was created by combining 1.5 µg cDNA from each of the 24 samples, this was mixed and split into 24 aliquots.

4.2.2.4 cDNA labelling

cDNA (1.5 µg) of each sample, together with 1.5 µg reference pool samples, were dried *in vacuo*, after which the samples were resuspended in 2.5 µL water (SABAX, Adcock-Ingram). The dyes were prepared by adding 12 µL dimethyl sulphoxide (DMSO, 99%, Sigma-Aldrich) to each Cy3 and Cy5 (GE Healtcare) dyes. Reference pool samples were labelled with Cy3, whereas Cy5 dye was used to label the individual cDNA samples. The dissolved dyes (2.5 µL) were diluted in 100 mM freshly prepared Na_2CO_3 (pH 9.0) and added to resuspended cDNA samples. The reaction mixtures were gently mixed, and incubated at RT in a desiccator for 2 h. Excess dye was removed using the QIAquick PCR purification kit (Qiagen) by adding 100 µL of buffer PBI (Proprietary, Qiagen) to the labelling mixture, after which the sample was loaded onto the QIAquick spin column and allowed to incubate at RT for 4 min. The spin columns were centrifuged at 12 000*g* for 1 min followed by three washing steps comprised of the addition of 500 µL buffer PE (Proprietary, Qiagen), followed by centrifugation at 12000*g* for 2 min, discarding the eluent in each case. Spin column membranes were thoroughly dried by centrifugation at 12 000*g* for 2 min. Labelled cDNA was eluted from the column by applying 30 µL water, allowing incubation at RT for 4 min, followed by centrifugation at 12000*g* for 90 sec, collecting the eluted liquid. Dye incorporation efficiency was determined using a Nanodrop® ND-1000 (*NanoDrop* Technologies), where the efficiency was calculated using equation 4.1, in which more than 10 labelled nucleotide per 1000 nucleotides was required to proceed.

Eq. 4.1 $$\text{Efficiency} = \frac{\text{pmol dye} \times 324.5\ \text{pg/mol}^*}{\text{ng DNA}}$$

* average mass of one dNTP

4.2.2.5 Sample fragmentation and slide oligonucleotide hybridization

Individual Cy5 labelled cDNA samples (20 pmol) were combined with 20 pmol of the Cy3 labelled reference pool cDNA sample, together with 5 µL 10× blocking buffer (Proprietary, Agilent), 1 µL 25× fragmentation buffer (Proprietary, Agilent) to a final 25 µL volume. The tubes were incubated at 60°C for 30 min, after which 25 µL 2× GE hybridization buffer was gently added to ensure no unwanted foaming occurred, and stored on ice. The arrays consisted of 8 chambers into which a total 20 pmol of each Cy5 and Cy3-labelled cDNA was loaded into each chamber. Samples were loaded into assembled hybridization chambers in the gasket slide, after which the microarray oligonucleotide slide (Agilent, with 'Agilent' side down), was carefully placed on top. The slide

chamber assembly was tightened and placed in a hybridization oven at 65°C for 18 h on rotational speed 10, making sure of liquid spreading evenly over the slide.

4.2.2.6 Hybridization clean-up

Hybridized slides were removed from assembly chambers and carefully disassembled. The slides were washed twice in wash buffer 1 (Proprietary, Agilent) for 1 min, followed by washing in preheated (37°C) wash buffer 2 (Proprietary, Agilent) for 1 min. Slides were dried by centrifugation using an Agilent desktop centrifuge, after which slides were scanned on a Axon GenePix 4000B scanner (Molecular devices).

4.2.2.7 Data processing

An improved 60-mer Agilent platform was created, partly adapted from the 70-mer Operon Array containing 8089 *P. falciparum* genes as described by Smit *et al.* [243] At the time, the most recent annotated *P. falciparum* (strain 3D7) genome from PlasmoDB 5.4 (www.plasmodb.org) was used to design the 60-mer based array. After removal of unnecessary "NULL" Operon –specific controls the remaining 7004 oligonucleotides were shorted to 60 nucleotides by using a 10 nucleotide scanning window from either the 3'- or 5'-ends, whilst ensuring that the microarray annealing temperature (Tm) was maintained close to 65°C. Equation 4.2 was used to calculate the Tm.

Eq. 4.2 $$\mathrm{Tm} = 64.9°\mathrm{C} + 41°\mathrm{C}(\mathrm{GC}\text{-}16.4)/\mathrm{N}$$

in which GC corresponds to the number of G and C nucleotides in the target sequence, where N is the total length of the sequence.

The 60-mer sequences were validated using NBLAST (www.ncbi.nlm.nih.gov/BLAST) from which all the submitted sequences had target E-values below 10^{-6}. Additionally, the 60-mer sequences were submitted, as FASTA file, into the ArrayOligoSelector tool (http://arrayoligosel.sourceforge.net/) which optimizes oligo selection based on sequence complexity, uniqueness and lack of self-complementarity. The redesigned sequences were again validated using NBLAST and submitted for printing onto Agilent slides. Arrays were interpreted using the Axon GenePix 6.0 software (Molecular devices). Automated spot detection was performed to ensure correct spot quality, saturation and signal-to-noise ratios using criteria as listed in Table 4.1.

Table 4.1: Parameters used for spot flagging and assessment in GenePix.

Parameter [a]	Function	Flag	Average percentage of flagged spots
Saturation [b]	[F532 % Sat.] > 20 Or [F635 % Sat.] > 20	Bad	2.8 – 4.6 %
Signal-to-noise (S/N) [c]	[SNR 532] < 3 And [SNR 635] < 3	Bad	29 - 32 %

[a] Flagging of spots based on circularity and intensity was not performed due to the low number of positive flagging events. [b] Spots saturation was evaluated in both foreground red channels (532 nm) and green channels (635 nm) and flagged as bad, therefore were not used, if the saturation level was above 20. [c] Spots with S/N levels below 3 were removed from arrays by this function.

For simplicity, the Cy5-red and Cy3-green intensity values from an array spot is expressed as A and M values, where M-values represents difference in average log intensity of spots in the array (Eq. 4.3). For a particular spot, if the red channel value is greater than the green, this results in a positive M-value which equates to greater levels of red-labelled cDNA, or in this case Cy5 labelled sample, compared to green-labelled Cy3 reference cDNA. The $\log_2$ ratios of the individual colour channels are used to relate expression in terms of M, as listed in equation 4.3.

Eq. 4.3 $$M = \log_2 R - \log_2 G$$

$$= \log_2 \left(\frac{R}{G}\right)$$

Where R is the intensity of in the red channel at 532 nm for the Cy5 dye and G is the spot intensity in the green channel at 635 nm for the Cy3 dye. The M-value also represents the $\log_2$ fold change ($\log_2$FC).

The A-values represents the log intensity in the colour channel for a particular spot, and is calculated by the average log intensities of both red and green channel intensities [247]. The average intensity of a spot was interpreted as the A-value using equation 4.4, which

Eq. 4.4 $$A = \frac{(\log_2 R + \log_2 G)}{2}$$

In which R is the intensity of in the red channel and G is the spot intensity in the green channel.

The LIMMA (Linear models for microarray data, [248], www.bioconductor.com) and LIMMA-GUI packages were used to analyse the gene expression microarray data. The background from all the arrays was corrected using a subset of 50 [249]. For values of M the assumption is made that the majority of genes are unaffected by the drug treatment and therefore plots of M are expected to be centred around zero [247]. With both M- and A-values, MA plot can be generated, and represent the distribution of spot data. This is useful to compare between-array M-values and determines the

extent of normalisation which is required. Within-array and between-array comparisons of both A- and M-values are necessary for accurate estimations of expression levels and consider both intrinsic array variations as well as biological differences between samples from separate arrays. Within-array normalisation methods include Loess and Robust spline normalisation, the former method fits a normalisation curve using local linear regression, the latter employs Bayes statistics together with robust fitting of regression splines to normalise the data [250, 251]. Gquantile is a between-array normalisation technique which scales the quartiles of the data [250]. Robust-spline normalisation was performed within each array, after which Gquantile normalisation was performed between the arrays which incorporates the use of the Cy3-labelled cDNA reference pool [252].

Differently expressed transcripts were identified by firstly comparing UT arrays with T array using Pearson correlations, after which between-array replicate spots were correlated using linear modelling (lmFit) which makes use of empirical Bayes methods [253]. Transcripts with a *P*-value smaller than 0.05, with the $\log_2$ ratio greater or equal to 0.75 or smaller and equal to -0.75 were considered significant ($\log_2$ ratio $\geq$ 0.75 and $\log_2$ ratio $\leq$ -0.75). The $\log_2$FC for differential expression in T vs. UT comparisons was calculated by subtracting the $\log_2$ (Cy5/Cy3) in the treated sample from the $\log_2$ (Cy5/Cy3) of the untreated sample. The fold change (FC) could be calculated using the antilog$_2$ of the $\log_2$ ratio, according to Eq. 4.5. The FC significance level was therefore greater and equal to 1.68 or smaller and equal to -1.68.

Eq. 4.5 $$\mathrm{FC} = 2^{\log_2 \mathrm{ratio}} \; or \; 2^{M}$$

The log-odds value or B-statistic, a statistical descriptor generated during array data comparison, refers to the probability that the gene of interest was differentially expressed, where a B-statistic of 0 corresponds to a 50-50 chance that the gene is differently expressed. Rankings of expression data using the B-statistic results in the same order if the data are ranked according to *P*-value (or the adjusted *P*-value). Some transcripts were represented by two or more spots in the arrays, and accounts for alternatively spliced variants of the same mRNA. In the case were duplicate identifiers were present, expression values from spots with the greatest B-statistic, which corresponds to the smallest adjusted *P*-value, were used.

Gene ontology (GO) terms of the differentially expressed transcripts were obtained from PlasmoDB 8.1 (www.plasmodb.org) and the transcripts were arranged according to GO biological function term. Grouping was verified using Micro Array Data Interface for Biological Annotation (MADIBA) [254] (http://www.bi.up.ac.za/MADIBA/) and DAVID (http://david.abcc.ncifcrf.gov/). Hierarchical clustering, by average linkage, was performed using Cluster [255] (http://rana.lbl.gov/EisenSoftware.htm). Gene set enrichment analysis (GSEA) was performed

using gene set enrichment analysis, GSEA software, and Molecular Signature Database (MSigDB) (http://www.broad.mit.edu/gsea/) [256]. Normalised M-values from the Tt_3 and UTt_3 array data were compared in a gene set permutation type analysis. The number of permutation was 1000, with a weighted enrichment statistic and signal-to-noise metric for ranking genes.

4.2.2.8 Real-time quantitative RT-PCR (qRT-PCR)

Real-time qRT-PCR was used to verify gene expression levels in UT and 4PEHz-treated parasites from the microarray study. The cDNA from T and UT samples of t_3 were diluted to 1.5 ng/µL with DNase-free dddH_2O (Invitrogen). A dilution series was created from pooled UT samples corresponding to 1/10, 1/20, 1/50 and 1/100 dilutions. The housekeeping transcripts; lactate dehydrogenase (LDH,) cyclophillin and seryl-tRNA synthetase (StSyn) were used to create the standard curve. Additionally, standard curves of the six target transcripts were created using the same dilution series of UT pooled cDNA at 1.5 ng/µL starting concentration (Table 4.2). Reactions were set up in a 384-well plate (Roche) consisting of 5× KAPA SYBR Fast qPCR (Kapa Biosystems), 3.75 ng cDNA and 5 pmol of each forward and reverse primer, as listed in Table 4.2,in a final 10 µL reaction volume. The plates were amplified and analysed in a Lightcycler 480 (Roche). The plates were incubated at 95°C for 10 min, followed by 48 amplification cycles consisting of 95°C for 10 s, annealing at 53°C for 5s, and extension at 72°C for 7s. A melting curve analysis was performed to verify whether primer dimers were present. Results were analysed using QBase Plus (Biogazelle).

Table 4.2: **Real-time qRT-PCR primers used to validate microarray data**.

PlasmoDB ID	Description	5' - 3' sequence
PFE0505w Cyclophilin	CycloF	AATTCTTTGACCATCTTAATCATTC
	CycloR	CAAAACAATTTTACTTCCTTGGGTTA
PF13_0141 Lactate dehydrogenase (LDH)	LDHF	GATTTGGCTGGAGCAGATATGTA
	LDHR	CAACAATAATAAAAGCATTTGGACAA
PF07_0073 Seryl-tRNA synthetase (StSyn)	StSynF	TTCGGCACATTCTTCCATAA
	StSynR	AAGTAGCAGGTCATCGTGGT
PF14_0224 (serine/threonine protein phosphatase)	PF14_0224F	AGTGCAGATACAGCTAATTAT
	PF14_0224R	TTGCCTTGACAAAATCGTTAT
PFB0695c (acyl-CoA synthetase, PfACS8)	PFB0695cF	GTCAAAAGGAACGCACTAA
	PFB0695cR	TTTTCTTGAGAACGTGATT
PFI1370c phosphatidylserine decarboxylase	PFI1370cF	AGCCATCAGTGCATATAAT
	PFI1370cR	ATTATGGATGAGCCTACTTTA
PF11_0256 pyruvate dehydrogenase E1 alpha subunit	PF11_0256F	TTACAACAAGCAGAATTAGAT
	PF11_0256R	TCATTAGAAATGTCGTCAAAT
PF14_0155 serine C-palmitoyltransferase, putative	PF14_0155F	AATGCAATACAAGCTGCTATG
	PF14_0155R	CAAACAGCCAAAATAGAATCA
PF08_0066 lipoamide dehydrogenase	PF08_0066F	AAATAATTTGAGTAAACCCATA
	PF08_0066R	AATTTCAACACCGACATTATC

4.2.3 Proteomics

4.2.3.1 Protein extraction

Additional UTt_3 and Tt_3 parasite samples (5 mL) taken at the same timepoint as for the transcirptome analyses were used for proteomic analyses. The infected erythrocyte pellets (250 µL) were thawed at RT, and re-suspended in 5 mL PBS. Parasites were isolated from the erythrocytes using saponin lysis by adding 5 µL 10% w/v saponin (Merck) followed by incubation on ice for 5 min. The samples were centrifuged at 4500*g* for 15 min at 4°C and the red supernatant was discarded. The remaining pellets were washed two consecutive times using 1 mL PBS to re-suspend the pellet and centrifuging at 15 000*g* for 2 min at 4°C. The remaining pellets were lysed using 500 µL lysis buffer consisting of 8 M urea, 2 M thiourea, 2% CHAPS, 0.5% w/v freshly prepared DTT. Samples were sonified seven times using a sonifier (Branson 450) with microtip at 20% duty cycle for 1 min with pulsing, in each case immediately halting once sample frothing was observed, with 1 min resting intervals on ice. Samples were clarified by centrifugation at 16 000*g* for 30 min at 4°C, and the remaining supernatant containing the soluble protein fraction was used in downstream

processes. The 2-D quant kit (GE Healthcare) was used to quantify the parasite protein samples, according to manufacturer's instructions.

4.2.3.2 Polyacrylamide gel electrophoresis

SDS-PAGE was performed as previously established by Laemmli *et al.* with modifications [198]. Electrophoresis was performed in Biometra Minigel protein electrophoresis housings (Biometra) with glass assembly plates. Protein samples were denatured in sample loading buffer (200 mM Tris-Cl, pH 6.8, 8% SDS, 40% glycerine, 0.02% v/v β-mercaptoethanol, 0.05% w/v bromophenol blue) by adding 2 - 18 µL of sample loading dye and boiling samples for 5 min. A 10% separating gel was prepared by adding 1.5 M Tris-Cl pH 8.7, bis-acrylamide (29.1% acrylamide, 0.9% bis-acrylamide), 0.1% SDS, 10% APS, and 33.5 nM (5µL) TEMED. The stacking gel component at 5% consisted of Tris-Cl, pH 6.7, Bis-acrylamide, 0.1% SDS, 10% APS, 33.5 nM (5 µL) TEMED. Electrophoresis buffer consisted of 25 mM Tris-Cl, 192 mM Glycine, and 0.1% w/v SDS. Loaded protein samples were run at 120V. SDS-PAGE gels were stained with Coomassie, containing 0.3 mM Coomassie R-250, 40% v/v MeOH, 7% v/v glacial acetic acid. Destaining was performed using a solution containing 40% v/v MeOH and 7% v/v glacial acetic acid.

4.2.3.3 MS Proteomics

Excised lanes from the SDS-PAGE gel was submitted to Dr Salome Smit at the Proteomics laboratory, Central Analytical Facility, Tygerberg Campus, University of Stellenbosch. Briefly, the proteins were extracted from the acrylamide gel and digested with trypsin. The digested protein samples were subjected to mass spectrometric analysis using Thermo Scientific EASY-nLC II connected to a LTQ OrbitrapVelos mass spectrometer (Thermo Scientific) with nano-electropsray source. The Thermo Proteome Discoverer 1.3 software program (PD 1.3, Thermo Scientific) was used to identify proteins via automated database searching using Mascot (Matrix Science) and Sequest (Yates lab, University of Washington) of all tandem mass spectra against the PlasmoDB 7.2 and SwissProt_57.15 databases. Proteins were considered positive matches at least 2 tryptic peptides per protein was identified with a Mascot score threshold of 20 and Sequest score threshold of 1.5.

4.3 Results

4.3.1 Transcriptomics

4.3.1.1 RNA integrity

RNA is enzymatically degraded within the cell by RNases and suffers from chemical instability at higher temperatures. The integrity and the quality of RNA extracted from treated (T) and untreated (UT) infected erythrocytes was assessed using automated electrophoresis. Traditional methods for assessing RNA quality compare the quantity of smaller and larger ribosomal RNAs (rRNA), where the 28S/18S ratio greater than 1.8 is considered as relativity intact RNA [257]. The RNA quality indicator (RQI) was developed using a series of increasingly degraded RNA samples, which were used to construct an indicator value ranging from 1 to 10, where 1 is completely degraded RNA [258]. The RQI algorithm interprets both the rRNA 28S and 18S regions from electrophoregrams, and also takes into account the pre-18S regions in which degraded RNA can occur [258]. Electropherograms of RNA extracted from both treated and untreated infected erythrocytes indicated the various time points had narrow and resolved fluorescence peaks visible for 18S and 28S ribosomal RNA with little or no shorter RNA fragments reflected as low baseline visible at the early part of the electropherogram (Figure 4.1). The RQI values for all the RNA extracts were above 7 (results not shown), indicative of good quality RNA.

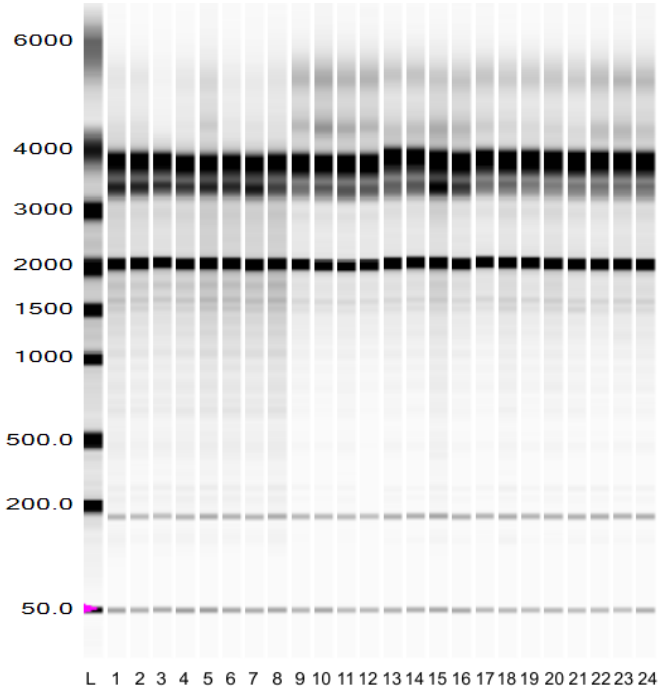


Figure 4.1: Electropherogram of RNA isolated from untreated and 4PEHz-treated parasites. A molecular weight marker, denominated in base-pairs (BP), is indicated on the left of the figure. Samples from 1 to 24 represent 200 ng RNA extracted from UTt_1 (1-4) and Tt_1 (5-8), UTt_2 (9-12) and Tt_2 (13-16) and UTt_3 (17-20) and Tt_3 (21-24), respectively. The 18S and 28S rRNA is visible at approximately 2000 and 3500 BP, respectively. Low levels of smearing visible between 2000 and 50 BP indicated that almost no degraded RNA was present.

4.3.1.2 Hybridization and raw image processing

The Agilent microarray platform, representing ~15K (15 744) oligonucleotides of which 7004 were unique identifiers, was used to determine the transcriptomic effects of 4PEHz treatment on the intraerythrocytic development of *P. falciparum* parasites. Positive control spots in each individual array included bright and dull spots at the corner of the array (Figure 4.2). Microarray spots representing Cy3 (green) labelled reference pool cDNA or Cy5 (red) labelled cDNA oligonucleotides were correlated to mapped grids containing the target information. A common reference-designed microarray was used in which the Cy3 reference pool consisted of equal amounts of pooled cDNA from all T- and UT-extracts.

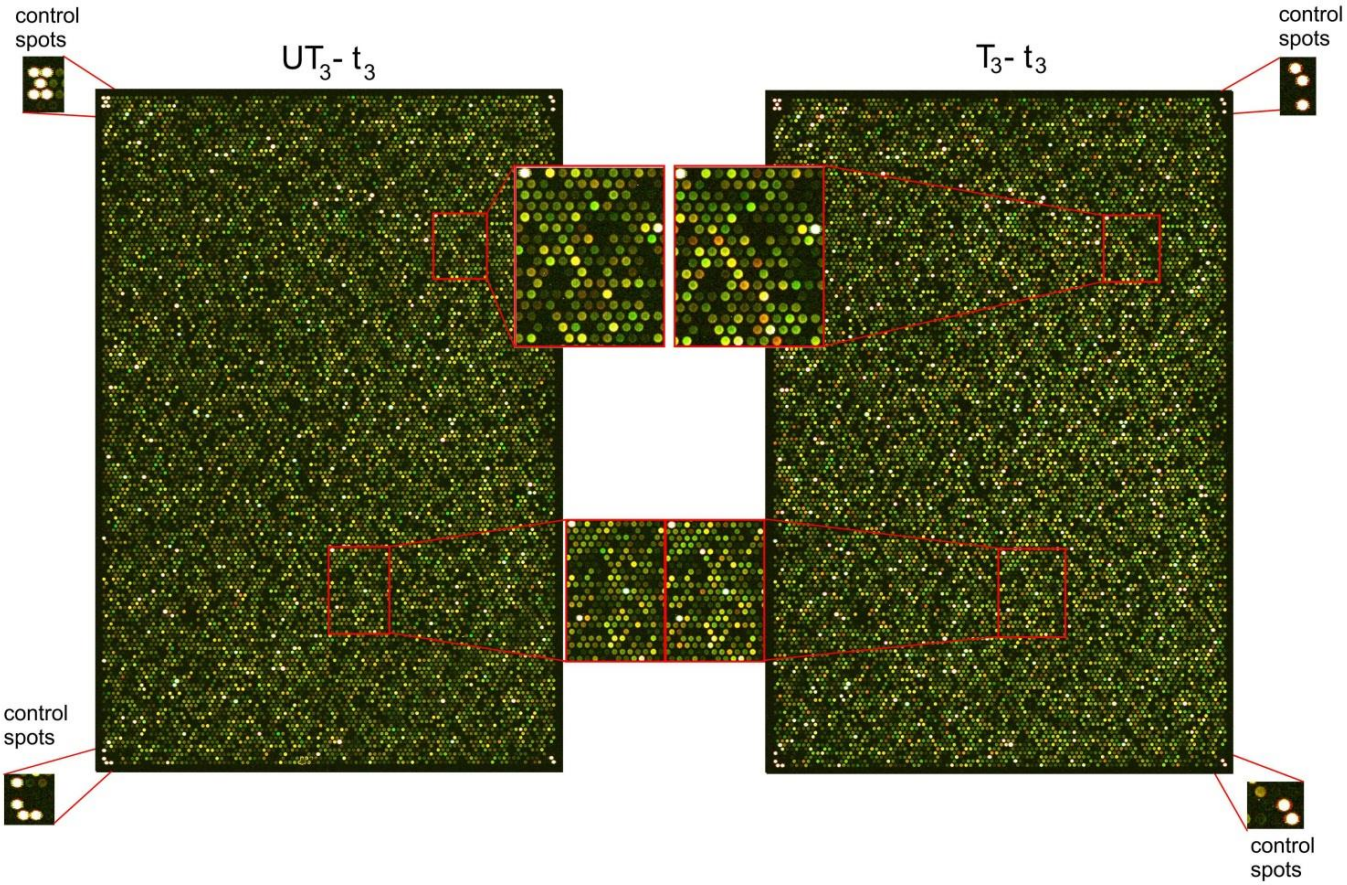


Figure 4.2: Individual arrays from a single UTt_3 and Tt_3 sample. The expanded regions in the centre compare the same region of spots from different arrays, and as can be seen the Tt_3 array some spots have higher red intensities, which indicate greater abundance of those particular transcripts in the array. The corner regions contained bright and dark control spots which are used to determine whether cDNA hydridization occurred successfully.

Spots from the arrays were subjected to evaluations based on signal intensity, spot saturation and signal to noise levels as indicated in the methods section. When comparing two individual arrays (UTt_3 and Tt_3), differences in spot brightness, intensity and colour reflects are noticeable (Figure

4.2). The spots, of either red or green colour channel, were interpreted with imaging. Spots from all of the slides were subjected to spot filtering procedures which entailed removing low signal-to-noise spots, as well as spots which were over saturated. On average between all the slides 2.8 – 4.6% of spots were removed based on over saturation. Between 29 – 32% of spots were removed from further interpretations based on the low (< 3) signal-to-noise ratios. On average more than 11K spots in each slide conformed to the quality filtering parameters and were further subjected to spot and array normalization procedures.

4.3.1.3 Normalisation of microarray image data

Background or ambient spot signal is caused by residual material left on the arrays after washing or by the non-specific binding of labelled samples [249]. Background correction of the individual red- and green-colour channels is necessary to remove biased interpretation and accounts for array-based variance including differences in labelling efficiency and spatial-associated differences with the array, ensuring correct interpretation of biological variation in terms of gene expression [249, 252]. In experiments performed the corrected adaptive background method (*normexp*) was used, with an offset of 50, thereby also removing unnecessary control spots thereby only retaining important information regarding target spots [248]. Background processing using the *normexp+offset* normalisation, instead of traditional subtraction of local background, gives lower false discovery rates [249].

Before normalisation, the log intensity (A) values of with the arrays had median values of between 3 and 4 with outlier values ranging as low as -1 (Figure 4.3 A). The $\log_2$ ratio of expression (M) values of the data had median values between -1 and 0 (Fig 4.5 B). MA plots, pre- and post-normalisation, generated from arrays of t_2 and t_3 in the 4PEHz study are shown in the Appendix section (Fig 4.A1- 4.A3). MA plots of array 7 – 12, as shown in Fig 4.A1, showed M values centred at 0. Loess normalisation aims to produce median centred data, and the quartiles are forced to have the same values using Gquantile between-array normalisation. Plots of Loess and Gquantile normalised data were median centred (Figure 4.3 C, D), however the differences were still visible in the quartile values of the data, suggesting ineffective Gquantile scaling of all the data. Array 4 and 14 had visibly smaller 25^{th} percentile values in A-plot (Figure 4.3 C), with some outliers as well, which is reflected in M-plots where array 4 and 14 can be seen with lowered median values (Figure 4.3 D). Robust spline normalisation together with Gquantile scaling resulted in improved median correction for A-values (Figure 4.3 E), in which the quartile values were effectively scaled to each other. The M-values appeared more evenly centred to 0 compared to Loess normalisation (Figure 4.3 F).

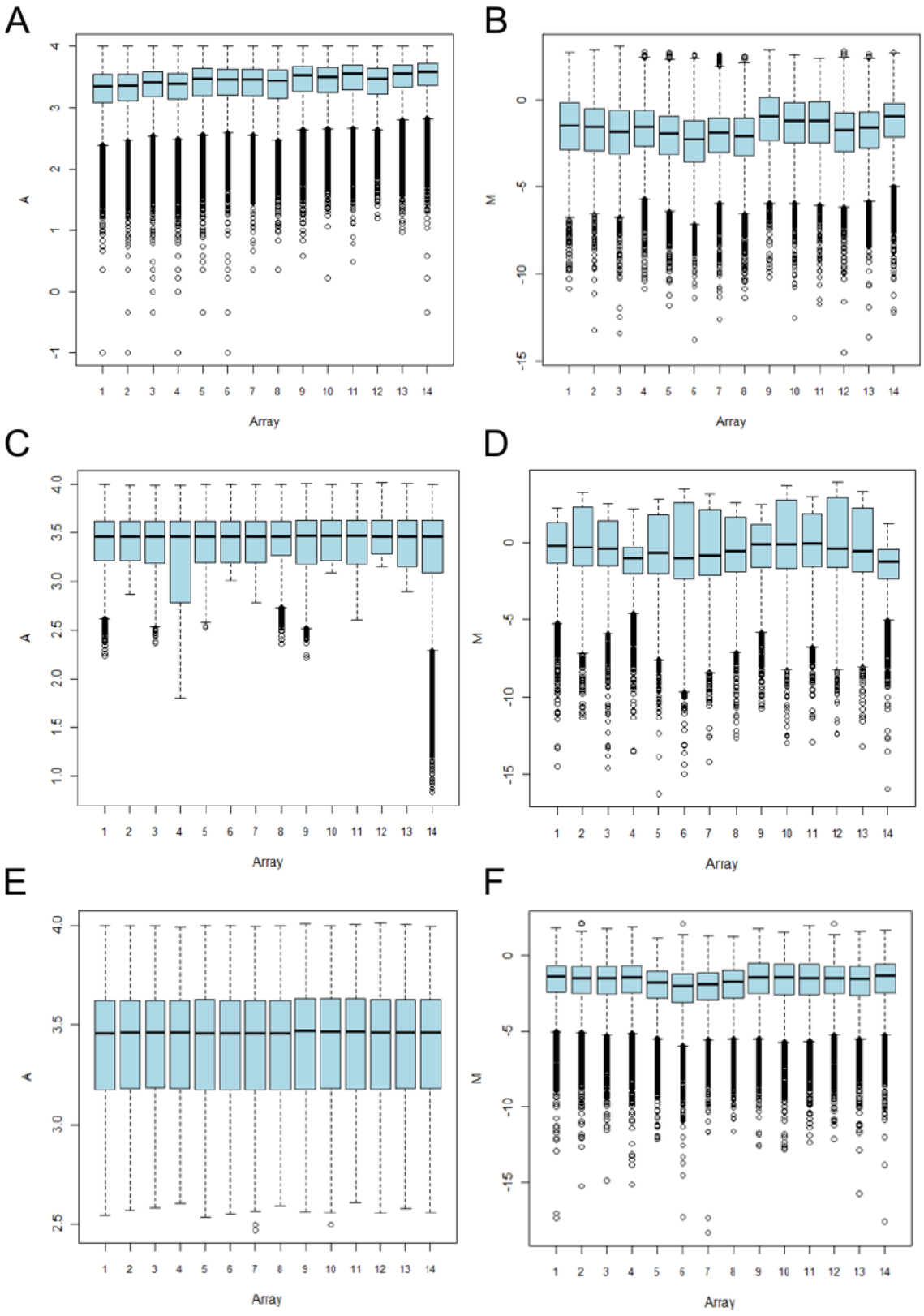


Figure 4.3: Boxplots of the log intensity (A) and log ratio expression (M) before and after array normalization. A and B) Plots of A and M before normalization, respectively. **C and D)** Plots of A and M after Loess and Gquantile normalisation**. E and F)** Boxplots of A and M after Robust spline and Gquantile normalisation. Boxplots represent the median of the data represented by a solid black line, boxed by quartiles with the top 75th percentile and bottom 25th percentile. Dotted lines represent the minimum and maximum values together with outliers marked in circles.

Density plots were used to visualise the distribution of the array data, and represents the frequency at which spot intensity values were observed, which is useful for determining array outliers [247]. Relating to the previous assumption that the M-values from an array remain unchanged or centred at 0, similarly when comparing one array to another it is expected that the overall frequency and intensity of the spots remain the same across the arrays. This is however hardly ever the case and another reason why normalisation is employed to account for between-array variation. By using a Cy3 cDNA reference pool the green channel data across all of the arrays can be normalised to one another (Figure 4.4). The green Cy3 non-normalised arrays data is depicted as green lines which have similar intensity and frequency distribution, however small between-array variations can be observed (Figure 4.4 A). After normalization, a single green representative line is observed (Figure 4.4 B and C). Loess normalisation, with Gquantile scaling, produced irregular red channel frequencies (Figure 4.4 B), compared to Robust spline normalisation with Gquantile showed markedly improved normalisation of the arrays (Figure 4.4 C). The Loess normalisation method therefore was not able to effectively normalise the data between different arrays and data from the Robust spline method was used determine differential gene expression.

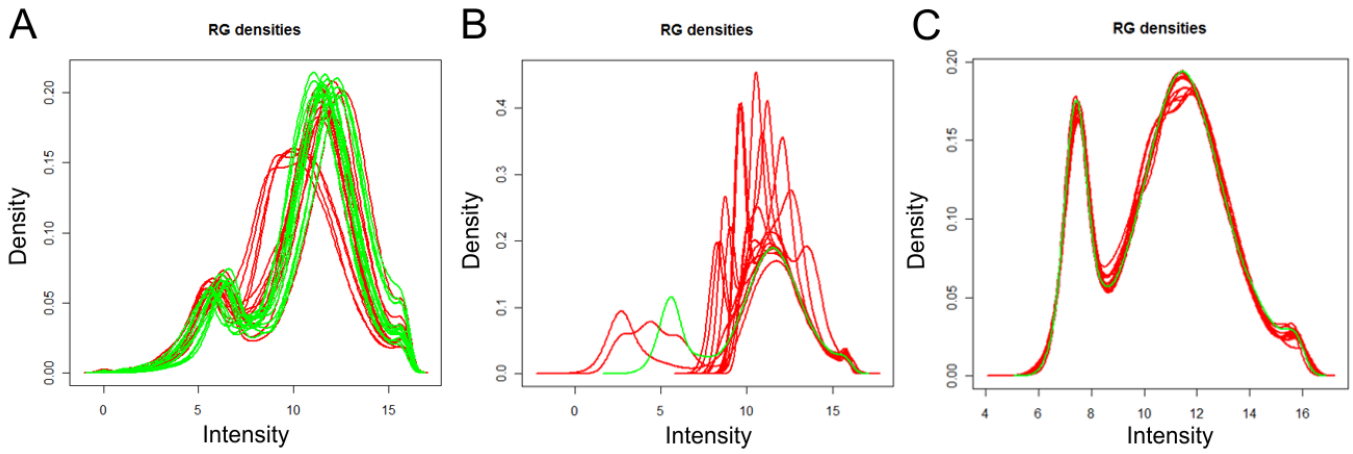


Figure 4.4: Density plots used to interpret array data before and after normalization. Density plots of array data **A)** before normalisation, **B)** after Loess and Gquantile normalisation and **C)** after Robust spline and Gquantile normalisation. The y-axis density represents the frequency of mRNA's in relationship to the x-axis intensity of the mRNA's.

After normalization all the Cy5 and Cy3 intensity values of the spots on the slides, apart from those omitted due to bad intensity or oversaturation, were used in calculations to determine the $\log_2$ FC. From these calculations only transcripts which had *P*-values smaller than 0.05, with the $\log_2$ ratio greater or equal to 0.75 or smaller and equal to -0.75 were considered significant ($\log_2$ ratio $\geq$ 0.75

and $\log_2$ ratio ≤ -0.75). Across all of these slides, on average, around 11K transcript spots in each slide were processed to derive individual spot $\log_2$ FC expression values.

4.3.2 Pearson correlations between UT and T samples

Previous microarray analyses with cytostatic therapeutics showed that Pearson correlations are necessary to ensure that, considering the 'just-in-time" nature of the transcriptome [225], correct phases of the parasite are compared to one another to observe drug-specific perturbation effects [259]. These correlations align global transcript expression patterns from T and UT parasites of different time points. As shown in the previous chapter, 4PEHz treatment of the malaria parasite did not affect the morphological composition of the parasite populations throughout a seven day period. Pearson correlations were necessary to ensure the correct T and UT-time point samples were compared to each other.

Table 4.3: Pearson correlations between UT and T array data of the 4PEHz treatment study.

Sample	Pearson Correlation (r)
$UTt_1 : Tt_1$	0.891
$UTt_2 : Tt_2$	0.847
$UTt_3 : Tt_3$	0.516
$UTt_1 : Tt_2$	0.093
$UTt_1 : Tt_3$	-0.521
$UTt_2 : Tt_1$	0.024
$UTt_2 : Tt_3$	0.288
$UTt_3 : Tt_1$	-0.486
$UTt_3 : Tt_2$	-0.280
$UTt_1 : UTt_2$	0.018
$UTt_1 : UTt_3$	-0.436
$UTt_2 : UTt_3$	-0.186

The highest correlation was found between the matching UT and T from each time point (Table 4.3). From the first two time points there was a high degree of correlation, which indicated preservation of global transcript expression patterns (Table 4.3). Pearson correlations between UTt_3 and Tt_3 suggested that there were greater variations in transcript expression profiles compared to the t_1 and t_2 comparisons (Table 4.3). Correlations of UT from any of the time points with any other T sample were either uncorrelated or anti-correlated (Table 4.3). Transcriptional progression is also marked by uncorrelated and anti-correlated correlations between the UTt_1 vs. UTt_2 and UTt_3 (Table 4.3), as expected due to the "just-in-time" nature of the transcriptome. From these results it was suggested that comparisons of gene expression could be made from the same time points. To

summarise, Pearson correlations suggested that the global transcript expression patterns from the first two time points were highly correlated, whereas the largest transcriptome perturbations occurred in t_3.

4.3.3 Differentially expressed transcripts

Comparison of the transcript expression ratios between UTt_1 vs. Tt_1 samples revealed only 25 unique transcripts were significantly expressed with $P < 0.05$ and had differential expression with the $\log_2$ FC > 0.75 or $\log_2$ FC < -0.75 (Figure 4.5). Of these transcripts 52% (13/25) of the transcripts had decreased expression and 48% (12/25) had increased abundance (Figure. 4.7). For the second time point (UTt_2 vs. Tt_2), 14 genes were found to be significantly expressed, all of which had decreased abundance.

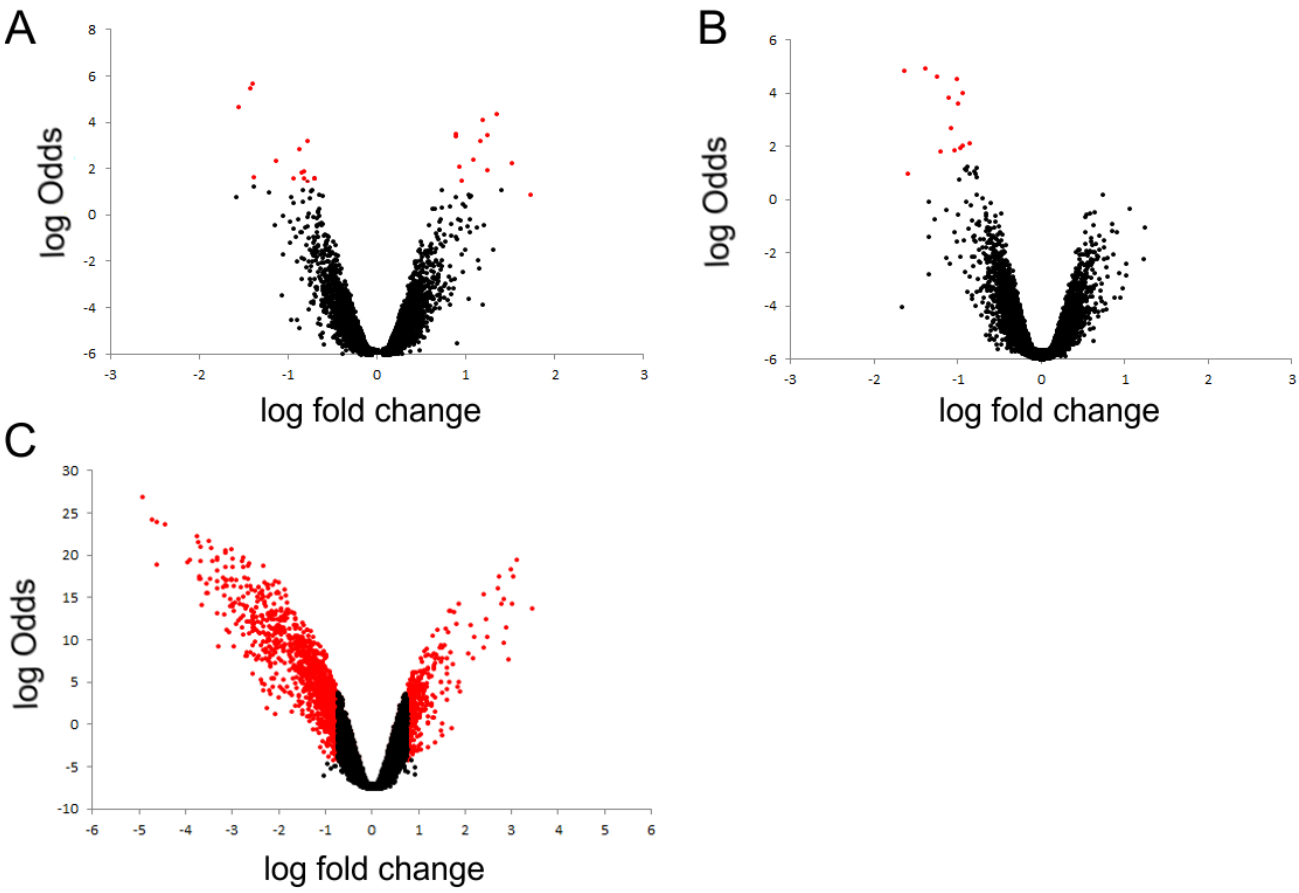


Figure 4.5: Volcanoplots generated for differently expressed transcripts affected by 4PEHz treatment. **A)** UTt_1 compared Tt_1 of which 13 unique transcripts had decreased abundance (red spots indicated $p < 0.05$, with $\log_2$FC > 0.75 or $\log_2$FC < -0.75) and 12 unique transcripts had increased expression. **B)** Volcanoplot of UTt_2 vs. Tt_2 in which 14 transcripts (from the 23 in total) were unique and that had significant decreased abundance. **C)** Volcanoplot of UTt_3 compared to Tt_3 in which 787 unique transcripts were differential expressed. Transcripts which had differential expression with the $\log_2$ FC > 0.75 or $\log_2$ FC < -0.75, and with $P < 0.05$ are marked in red.

Comparisons between the UTt_3 and Tt_3 samples showed 787 unique differently expressed transcripts. Of these 68% (535/787) of the transcripts had decreased transcript expression compared to 32% (252/787) which had increased expression. There were only a few differentially expressed transcripts in the first two time points in contrast the third time point, implicating that 4PEHz affects the parasites during later developmental stages of the IDC. The complete list of the differentially expressed transcripts from only the third t_3 time point is provided in the Appendix 4 (Table A4.1). The 20 most affected genes in time point t_3 (UTt_3 and Tt_3) are listed in Table 4.4. GO classification of transcripts with the lowest log_2 FC suggested these are involved in cell motility (GO term: 0006928), actin filament organisation (GO term: 0007015) and pathogenesis (GO term: 009405) (Table 4.4). Transcripts which had the highest abundance were involved in cyclic nucleotide metabolic processes (GO term: 0009187), protein polymerisation (GO:0051258) and components of the tricarboxylic acid cycle (GO:0006099).

Table 4.4: **Twenty transcripts that showed the greatest differential decreased and increased abundance during 4PEHz treatment in t_3.**

PlasmoDB ID	PlasmoDB description	log_2FC	Adj. P-value	Annotated GO biological process
Decreased abundance				
MAL7P1.176	erythrocyte binding antigen-175	-4.93	2.99E-12	regulation of immune response
MAL13P1.60	erythrocyte binding antigen-140	-4.74	2.89E-11	cell-cell adhesion
PFL2520w	reticulocyte binding protein homologue 3, pseudogene	-4.63	8.53E-10	-
PF08_0035	conserved *Plasmodium* protein, unknown function	-3.96	6.99E-10	-
PF13_0058	conserved *Plasmodium* protein, unknown function	-3.93	6.47E-10	entry into host cell, pathogenesis
PFA0110w	ring-infected erythrocyte surface antigen	-3.77	1.15E-10	-
PF10_0039	membrane skeletal protein IMC1-related	-3.75	1.71E-10	-
PFD0110w	normocyte binding protein 1,reticulocyte binding protein homologue 1	-3.72	2.53E-09	translation, translational initiation
PF10_0003	rifin	-3.70	6.50E-10	antigenic variation
PF10_0138	conserved *Plasmodium* protein, unknown function	-3.69	2.92E-10	cytoskeleton organization
PF07_0061	conserved *Plasmodium* protein, unknown function	-3.56	3.94E-09	-
PFL2460w	coronin	-3.52	1.71E-10	phosphate ion transport
MAL8P1.70	conserved *Plasmodium* protein, unknown function	-3.49	2.93E-09	entry into host cell
PFL1435c	myosin D	-3.46	2.99E-10	-
PF11_0268	kelch protein, putative	-3.44	6.50E-10	actin filament organization, actin filament-based movement
PF10_0262	conserved *Plasmodium* protein, unknown function	-3.35	6.38E-10	cytokinesis, microtubule-based process
PFI0540w	conserved *Plasmodium* protein, unknown function	-3.34	1.37E-09	-
PF10_0346	merozoite surface protein 6	-3.33	6.50E-10	cell-cell adhesion, pathogenesis
MAL13P1.176	reticulocyte binding protein 2 homologue b	-3.33	5.53E-09	-
PFE1285w	membrane skeletal protein IMC1-related	-3.22	2.73E-09	entry into host cell

Increased abundance				
PF11_0256	pyruvate dehydrogenase E1 alpha subunit	1.65	4.28E-08	acetyl-CoA biosynthetic process from pyruvate, fatty acid biosynthetic process, tricarboxylic acid cycle
PFL0795c	male development gene 1	1.65	4.18E-05	-
MAL13P1.129	conserved *Plasmodium* protein, unknown function	1.66	1.59E-06	-
PF10_0164	early transcribed membrane protein 10.3	1.67	4.17E-08	-
PFI0315c	dynein light intermediate chain 2, cytosolic	1.69	0.002833	-
PF14_0672	phosphodiesterase delta, putative	1.72	4.82E-06	cyclic nucleotide metabolic process, signal transduction
MAL7P1.162	dynein heavy chain, putative	1.81	1.29E-07	microtubule-based movement
PF14_0040	secreted ookinete adhesive protein	1.81	4.73E-05	-
PF11_0092	mechanosensitive ion channel protein	1.84	3.09E-05	-
PF14_0272	CPW-WPC family protein	1.84	2.31E-08	-
PF11_0479	conserved *Plasmodium* protein, unknown function	1.89	7.35E-05	-
PFL1910c	conserved *Plasmodium* protein, unknown function	2.05	1.96E-06	translation
PFL2385c	conserved *Plasmodium* protein, unknown function	2.09	1.44E-07	-
PF14_0467	conserved *Plasmodium* protein, unknown function	2.16	3.50E-06	-
PFB0355c	serine repeat antigen 2	2.40	9.86E-07	immunoglobulin production, proteolysis, regulation of immune response
PFE1205c	conserved *Plasmodium* membrane protein, unknown function	2.69	5.40E-09	-
PFD1050w	alpha tubulin 2	2.71	2.68E-09	microtubule-based movement, protein polymerization
PF14_0708	probable protein, unknown function	3.00	2.31E-08	-
PFI1520w	asparagine-rich antigen, putative	3.10	6.47E-10	-
PFI1230c	conserved *Plasmodium* protein, unknown function	3.43	4.15E-08	-

The adjusted P-value (adj. P-value) indicates the result of a moderated t-statistic test employed in LIMMA which normalises the standard errors from spot or probe results, and is interpreted as a normal t-statistic result. The $\log_2$FC value, or also the $\log_2$(Cy5/Cy3), indicates the relative expression of a transcript in Tt_3 in comparison with the UTt_3 stage, and is considered treatment-specific.

4.3.4 Gene ontology (GO) and functional annotation

GO charts of significantly expressed transcripts were constructed to determine processes affected by 4PEHz treatment (Figure 4.6). The predicted GO process term, a useful descriptive attribute assigned to genes according to the associated/predicted cellular function, cellular process or exact molecular function, was derived from PlasmoDB v7.0. Of the 251 transcripts which had increased abundance in t_3 a large proportion of transcripts (63%) represented conserved proteins with unknown function, and these were omitted from GO charts. Of the 535 transcripts with decreased abundance, the majority of transcripts (73%) similarly represented transcripts of which the product has no known function.

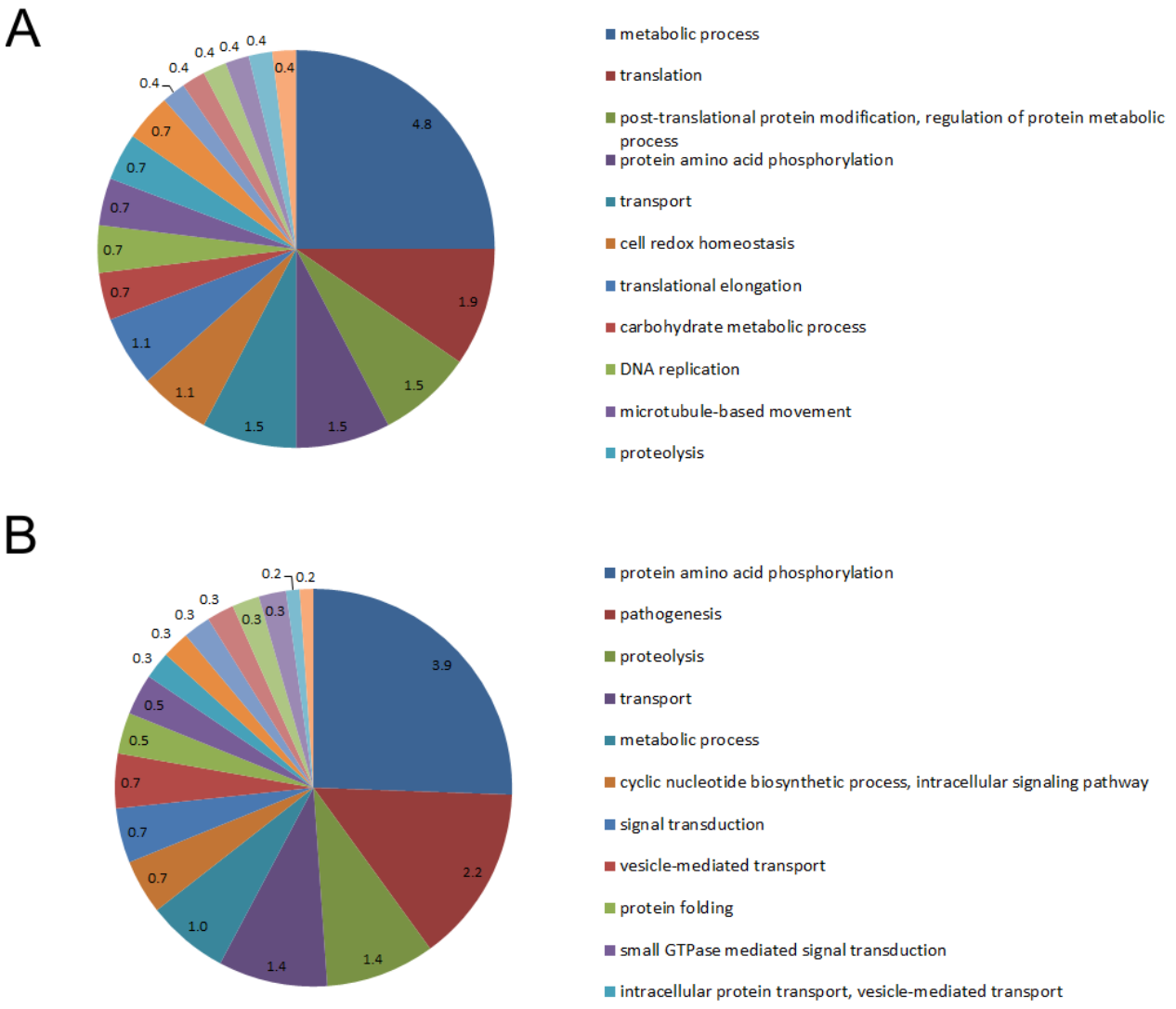


Figure 4.6: Differently expressed transcripts from UTt_3:Tt_3 classified according to PlasmoDB predicted GO processes. A) Transcripts which had increased abundance are associated with metabolic processes, as well as translation, and cell redox homeostasis. **B)** Transcripts which had decreased abundance were shown to be involved in protein amino acid phosphorylation, pathogenesis and proteolysis. Transcripts with unknown or hypothetical function, were not included in both instances. Values represent percentages of the total amount of transcripts.

Functional annotation of differentially expressed transcripts in DAVID was useful to determine additional processes which may have been excluded from the PlasmoDB database [260, 261]. Functional annotation revealed both increased and decreased transcript perturbations in apicoplast- and mitochondrial-related processes (Figure 4.7). Transcripts which had decreased abundance had involvement in phosphorylation, were associated with protein kinase activity, and had contributed in regulation of transcription (Figure 4.7 A). Transcripts with increased abundance were involved in cofactor metabolic processes, as well as acetyl-CoA metabolism (Figure 4.7 B). A few transcripts which had increased expression were also involved in DNA repair such as PF11_0087 (Rad51 homolog), and MAL13P1.346 (DNA repair endonuclease, putative) (Table A4.1).

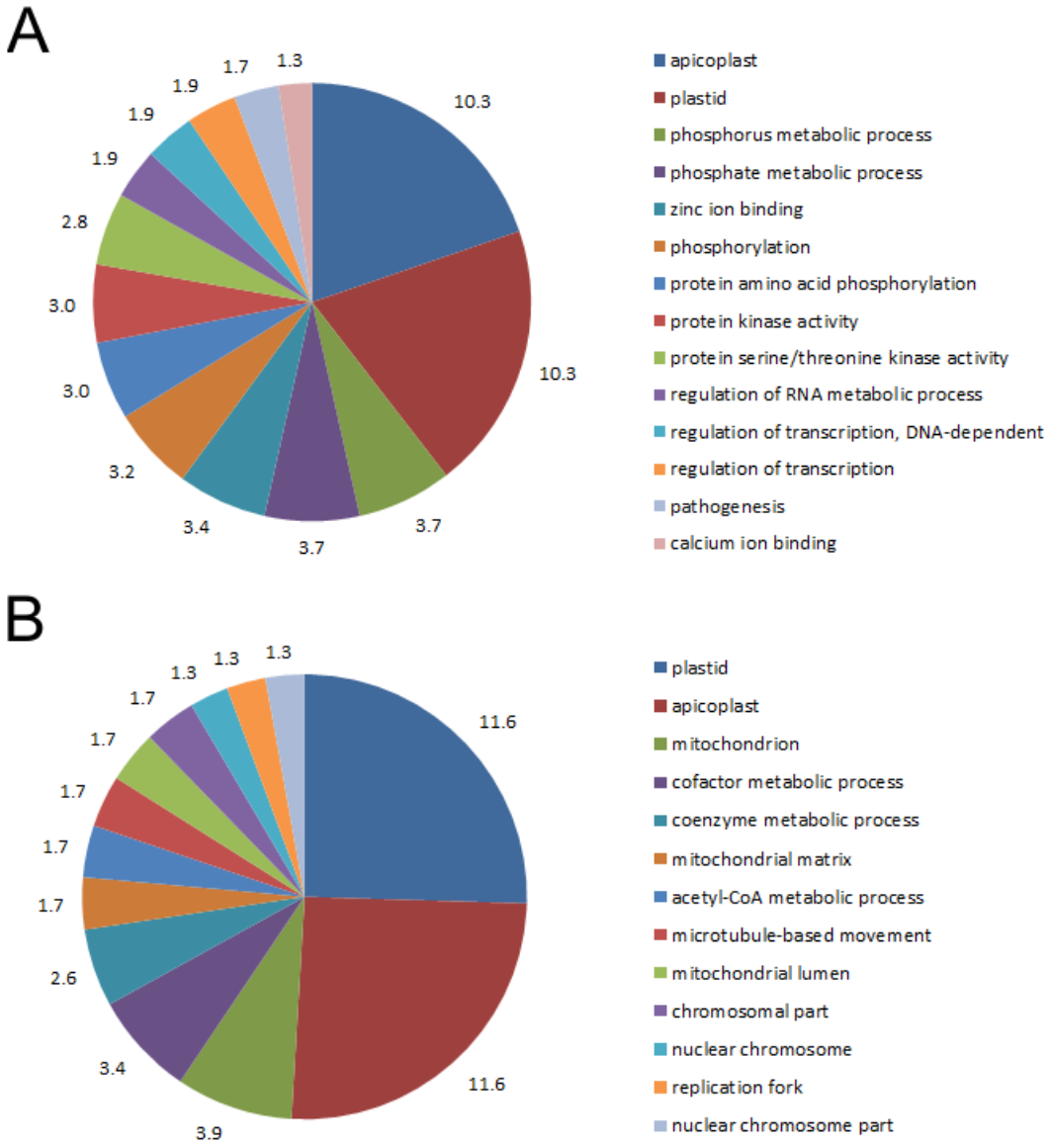


Figure 4.7: Functional annotation of differentially expressed transcripts using DAVID. Functional annotation was performed on **A)** transcripts which had decreased abundance and **B)** transcripts with increased abundance from UTt_3 vs. Tt_3.

There was some correlation in transcript expression across the different time points (Figure 4.8). In all three time points PFC0120w (cytoadherence linked asexual protein 3.1), PF14_0102 (rhoptry-associated protein 1) and PFI1445w (high molecular weight rhoptry protein 2) had decreased expression (Figure 4.8 A). These transcripts are associated with cytoadherence, and the rhoptries of the parasite. Corresponding transcripts in both t_1 and t_3 time points are shown in Figure 4.8 B. Transcripts MAL7P1.124 (conserved *Plasmodium* protein, unknown function) and PFL1565c (CG2-related protein, putative), had decreased and increased abundance at t_1, respectively, with opposite expression levels at the t_3 time point (Figure 4.8 B). When comparing transcript expression patterns from t_2 and t_3 time points (Figure 4.8 C), the PFD0955w transcript (apical merozoite protein) had a more pronounced suppression at t_3. The gene product of PFD0955w has a predicted role in attachment of GPI anchors to proteins, and plays a role in pathogenesis. Overall, transcript expressions across the different time points were shown to correlate, transcripts which had decreased expression during t_1 and t_2 had decreased expression in the t_3.

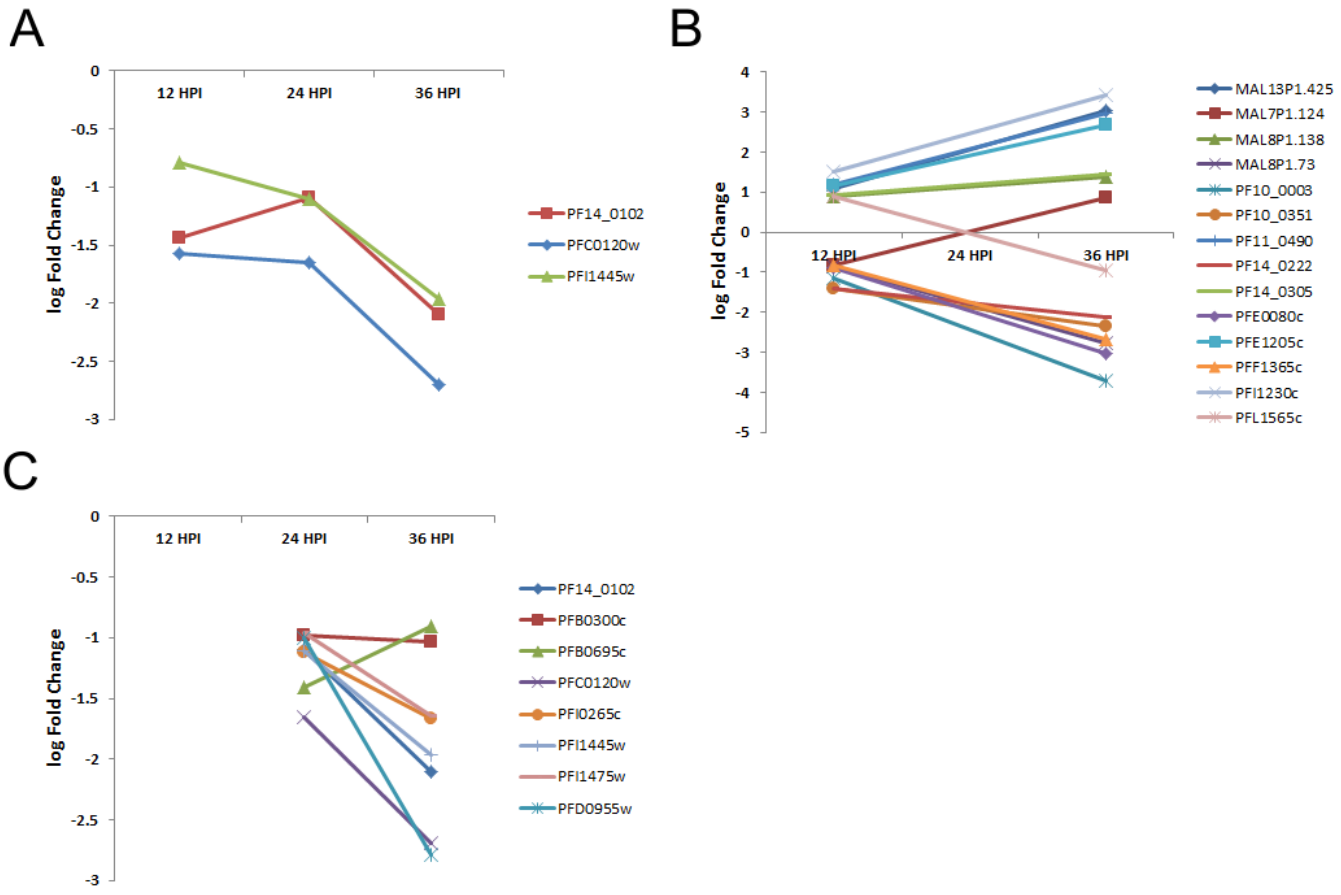


Figure 4.8: Transcripts which had significant expression and correlated between t_1 (12HPI), t_2 (24HPI) and t_3 (36HPI). A) PF14_0102, PFC0120w and PFI1445w had decreased abundance and this became more pronounced during the late t_3 phases. **B)** Transcripts PFL1565c and MAL7P1.124 were shown to have increased and decreased abundance, respectively, which changed to the opposite condition in the late t_3 time point. **C)** PFD0955w and PFC0120w had decreased expression levels during t_2 and greater decreases in expression at t_3.

4.3.5 PLP synthase and PLP-related perturbations

A degree of specificity was shown for 4PEHz in *Pf*Pdx1/*Pf*Pdx2 over-expressing cells, as discussed previously. PLP-related processes as well as processes affiliated with *Pf*Pdx1 and *Pf*Pdx2 substrates were therefore expected to be affected by 4PEHz treatment and it was postulated that these effects were visible in the transcriptome. The *Pf*Pdx2 transcript was significantly affected and had increased expression in t_3 (Table 4.5). Other transcripts of enzymes involved in PLP production such as pyridoxal and pyridoxine kinase (PFF0775w) were not significantly affected by 4PEHz treatment (Table 4.4). Of the PLP dependent enzymes in *P. falciparum*, a putative serine hydroxymethyltransferases (SHMT, PF14_0534) had increased abundance during t_1 and t_2. For these two time points the adjusted *P*-value was 0.049 and 0.319, suggesting that the t_2 expression level could not be considered as statistically significant (Table 4.5). Two other PLP-dependent enzyme transcripts were affected, a putative serine C-palmitoyltransferase (SPT, PF14_0155) and phosphatidylserine decarboxylase (PSDC, PFI1370c) (Table 4.5). The majority of other transcripts of PLP-dependent enzymes showed a general trend of increased expression during the t_3 stages, most of these were however not statistically significant. Of the PLP-dependent enzyme transcripts, AspAT had the most stable expression levels across all three time points.

Of the pathways that involve the substrates of *Pf*Pdx1 and *Pf*Pdx2, some transcripts in the pentose phosphate pathway had increased expression, however the majority were not significantly affected (Table 4.5). Putative deoxyribose-phosphate aldolase (PF10_0210), which reversibly converts G3P to 2-deoxy-D-ribose 5-phosphate had increased abundance (Table 4.5). The transcript of a putative phosphoglucomutase (PF10_0122), which catalyse the interconversion of R5P to D-ribose 1-phosphate had increased expression. A putative glutamate dehydrogenase (PF14_0286) transcript - part of L-glutamine catabolism - had increased abundance with $\log_2$FC of 0.90 in the parasites at t_3 (result not shown). The fact that these transcripts were affected suggested that 4PEHz could have disrupted R5P, G3P as well as L-glutamine homeostasis, thereby resulting in transcriptional perturbations. 4PEHz is a weak inhibitor of spinach R5P isomerase with a K_i of 1.8 ± 0.2 mM [206, 218]. At 20 µM 4PEHz it was anticipated that *P. falciparum* R5P isomerases would be weakly affected. The transcript of a putative D-ribose 5-phosphate epimerase (PFE0730c) was not significantly affected during the course of 4PEHz treatment (Table 4.5), suggesting that 4PEHz did not affect R5P isomerases leading to changes in transcript expression.

Table 4.5: Differential expression of transcript in 4PEHz treated *P. falciparum* parasites. The differential expression levels of transcripts associated with PLP biosynthesis and salvage, as well as PLP-dependent enzyme transcripts during treatment of *P. falciparum* with 4PEHz are shown.

PlasmoDB ID	PlasmoDB description	log_2FC $UTt_1:Tt_1$	log_2FC $UTt_2:Tt_2$	log_2FC $UTt_3:Tt_3$
PLP metabolism				
PF11_0169	pyridoxine biosynthesis protein Pdx2 (Pdx2)	0.05	0.21	0.77*
PF14_0570	pyridoxal 5'-phosphate synthase, putative	0.52	0.29	0.41
MAL13P1.324	aldo-keto reductase, putative	0.44	-0.14	0.32
PF14_0088	aldo-keto reductase, putative	0.29	-0.11	0.26
PFF0775w	pyridoxine kinase	-0.30	0.21	0.23
PFF1025c	pyridoxine biosynthesis protein Pdx1 (Pdx1)	-0.27	0.19	-0.15
PF07_0059	4-nitrophenylphosphatase	0.20	0.15	-0.25
PLP-dependent enzymes				
PF14_0155	serine C-palmitoyltransferase (SPT), putative	0.22	0.17	1.48*
PF14_0557	conserved *Plasmodium* protein, unknown function	0.01	0.69	0.61
PFI0965w	pyridoxal 5'-phosphate dependent enzyme class III, putative	0.01	0.16	0.47
PFL1720w	serine hydroxymethyltransferase (SHMT), cytosolic	-0.07	-0.35	0.47
PFD0670c	lysine decarboxylase-like protein, putative	-0.15	-0.13	0.44
PF07_0068	cysteine desulfurase, putative	0.05	0.21	0.41
PFI1370c	phosphatidylserine decarboxylase (PSDC)	0.15	-0.21	-0.94*
PFL2210w	delta-aminolevulinic acid synthetase (ALAS)	-0.06	-0.08	-0.27
PF14_0534	serine hydroxymethyltransferase (SHMT), mitochondrial	1.72*	0.90	-0.37
PFB0200c	aspartate aminotransferase,aspartate transaminase (AspAT)	-0.15	0.15	0.19
PFF0435w	Ornithine aminotransferase (OAT)	-0.14	-0.14	0.12
Pentose phosphate pathway				
PF10_0122	phosphoglucomutase, putative	-0.08	0.08	0.94*
PF10_0210	deoxyribose-phosphate aldolase, putative	-0.02	-0.07	0.86*
PF14_0520	6-phosphogluconate dehydrogenase, decarboxylating, putative	0.24	-0.23	0.74
PF14_0466	Appr-1-p processing domain protein	0.09	0.06	0.42
PF13_0157	ribose-phosphate pyrophosphokinase, putative	0.20	-0.11	0.38
PF13_0143	phosphoribosylpyrophosphate synthetase	-0.06	0.11	0.33
PFL0960w	D-ribulose-5-phosphate 3-epimerase, putative	-0.18	-0.10	0.18
PFE0730c	D-ribose 5-phosphate epimerase, putative	0.05	0.05	0.02
PFF0530w	transketolase	-0.01	-0.15	-0.15
PF14_0511	glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase	-0.18	-0.04	-0.36

Footnotes: Significant differential expression with a log_2FC above or below 0.75 and -0.75, respectively, with $P \leq 0.05$ is indicated by *.

In conjunction with SPT and phospholipid metabolism, other transcripts such as putative patatin-like phospholipase (PFI1180w) had decreased expression with by 1.69-fold change (Table 4.6).

Similarly, within the phosphatidylserine pathway putative diacylglycerol kinase (PFI1485c) were also found to be less abundant during the t_3 time point (Table 4.6). Transcripts involved in fatty acid (FA) synthesis had increased abundance, favouring the malonyl- and palmitoyl-CoA route of FA synthesis. Transcripts of this pathway include a putative biotin carboxylase subunit of acetyl CoA carboxylase (PF14_0664) and acyl-CoA synthetase, PfACS9 (PFB0685c), with another putative long chain polyunsaturated fatty acid elongation enzyme (PFF0290w) (Table 4.6). Transcripts of the malonyl CoA-acyl carrier protein transacylase precursor (PF13_0066) also had increased expression forming part of the conversion of malonyl-CoA to malonyl-acyl carrier protein (ACP).

Some of the transcripts were also characterised by the GO term - acetyl-CoA biosynthetic process from pyruvate (GO:0006086) – having involvement in acetyl-CoA biosynthesis. The pyruvate dehydrogenase complex (PDC) converts pyruvate to acetyl-CoA, and three of the four components that constitute the PDC in *P. falciparum* had increased expression at the t_3 time point (Table 4.6). These include pyruvate dehydrogenase E1 alpha subunit (PF11_0256); lipoamide dehydrogenase (aLipDH, PF08_0066) and pyruvate dehydrogenase E1 beta subunit (PF14_0441). The other component of the PDC was a putative dihydrolipoamide acyltransferase (PFC0170c which had 0.40 log_2FC greater expression compared to the UTt_3 stage (results not shown). The log_2FC of the latter was not considered as significant, however, for a descriptive purpose was interpreted together within the PDC. Overall there was an increased enrichment of transcripts associated with the PDC.

Table 4.6: Transcripts associated with apicoplast fatty acid biosynthesis in *P.falciparum* affected by 4PEHz. Apicoplast-related transcripts with differential expression during t_3 are listed, with the associated GO terms.

PlasmoDB ID	PlasmoDB description	log_2FC $Tt_3:UTt_3$	adj.P.Val	PlasmoDB annotated processes	GO term
PF11_0256	pyruvate dehydrogenase E1 alpha subunit	1.65	4.28E-08	acetyl-CoA biosynthetic process from pyruvate	GO:0006084~acetyl-CoA metabolic process
PF08_0066	lipoamide dehydrogenase (aLipDH)	1.28	8.55E-05	acetyl-CoA biosynthetic process from pyruvate	GO:0006084~acetyl-CoA metabolic process
PF13_0066	malonyl CoA-acyl carrier protein transacylase precursor	1.19	1.76E-02	fatty acid biosynthetic process	GO:0004314~[acyl-carrier-protein] S-malonyltransferase activity
PFB0685c	acyl-CoA synthetase, PfACS9	1.07	3.47E-04	fatty acid metabolic process	
PF14_0441	pyruvate dehydrogenase E1 beta subunit	0.91	1.07E-04	acetyl-CoA biosynthetic process from pyruvate	GO:0009536~plastid
PFD1035w	steroid dehydrogenase, putative	0.82	2.52E-04	fatty acid biosynthetic process	GO:0006631~fatty acid metabolic process
PF14_0664	biotin carboxylase subunit of acetyl CoA carboxylase, putative	0.80	1.34E-03	fatty acid biosynthetic process	
PFI1180w	patatin-like phospholipase, putative	-0.76	7.06E-05	lipid metabolic process	GO:0016042~lipid catabolic process
PFC0260w	P-loop containing nucleoside triphospahte hydrolase, putative	-0.88	7.55E-04	metabolic process	
PFB0695c	acyl-CoA synthetase, PfACS8	-0.91	2.00E-04	fatty acid metabolic process	
PFI1485c	diacylglycerol kinase, putative	-1.39	8.62E-06	activation of protein kinase C activity by G-protein coupled receptor protein signaling pathway	GO:0007166~cell surface receptor linked signal transduction
PFL0035c	acyl-CoA synthetase, PfACS7	-1.89	2.32E-07	metabolic process	

Footnote: The adjusted P-value (adj. P-value) indicates the result of a moderated t-statistic test employed in LIMMA which normalises the standard errors from spot or probe results, and is interpreted as a normal t-statistic result.

Gene set enrichment analysis (GSEA) was performed to find functionally classified gene sets which were either positively or negatively correlated. The GSEA uses normalised M-values to rank gene sets, and find patterns of enrichment. Individual transcripts in these sets may not have statistically significant differential expression, however as cluster may reveal significant enrichment. Gene set

enriched by 4PEHz treatment were associated with cofactor and coenzyme metabolic processes (Table 4.7, Figure 4.9). Additional information regarding the identities of these transcripts is given in Appendix Table A4.2. Within this group there were several transcripts involved in folate metabolism and these folate-related transcripts were found to be enriched and positively correlated with 4PEHz treatment (Table A4.2). These included dihydrofolate synthase/folylpolyglutamate synthase (DHFS-FPGS), dihydropteroate synthetase (DHPS) and lipoamide dehydrogenase (aLipDH), with log_2FC of 0.49, 0.43 and 1.28 expression in t_3. The transcript of bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS, PFD0830w) had increased expression at t_3 (Table A4.2). DHFR-TS is not annotated as part of folate biosynthesis, however involved in recycling of folate, and was included in this cluster for interpretation (Table A4.2). Transcripts with the Kyoto encyclopaedia of genes and genomes (KEGG) term: Citric acid cycle (TCA cycle) gene set were enriched (Figure 4.9). These include the components of the PDC as well as an ATP-specific succinyl-CoA synthetase beta subunit (Figure 4.9). GSEA also revealed transcripts linked to the pentose phosphate pathway were enriched in the t_3 time point (Table 4.7). The core enriched components of this cluster included a putative deoxyribose-phosphate aldolase (PF10_0210), a putative decarboxylating 6-phosphogluconate dehydrogenase (PF14_0520) and TIM (PF14_0378). These transcripts had a log_2FC expression of 0.86, 0.74 and 0.5, respectively.

Table 4.7: Gene set enrichment analysis (GSEA) of Tt_3 compared to UTt_3.

Gene set name	Cluster size	Enrichment score	Normalised enrichment score	Nominal *P*-value	FDR q-value
GO.0006412. PROTEIN.BIOSYNTHESIS	97	0.50	1.93	0.000	0.012
GO.0051186. COFACTOR.METABOLIC.PROCESS	43	0.54	1.81	0.000	0.025
KEGG.PYRIMIDINE METABOLISM	28	0.58	1.78	0.000	0.028
GO.0006118. ELECTRON.TRANSPORT	34	0.55	1.77	0.002	0.031
GO.0008757. SADENOSYLMETHIONINEDEPENDENT. METHYLTRANSFERASE.ACTIVITY	16	0.61	1.62	0.018	0.083
GO.0016491. OXIDOREDUCTASE.ACTIVITY	19	0.56	1.59	0.015	0.088
KEGG.CITRATE CYCLE (TCA CYCLE)	13	0.64	1.59	0.015	0.091
GO.0006633. FATTY.ACID.BIOSYNTHESIS	4	0.91	1.59	0.002	0.092
GO.0006098. PENTOSEPHOSPHATE.SHUNT	6	0.75	1.54	0.017	0.126
GO.0051188. COFACTOR.BIOSYNTHETIC.PROCESS	24	0.52	1.52	0.032	0.143
GO.0051187. COFACTOR.CATABOLIC.PROCESS	16	0.56	1.48	0.047	0.176
KEGG.GLUTATHIONE METABOLISM	8	0.67	1.47	0.050	0.182
MPM.FOLATE BIOSYNTHESIS	5	0.78	1.47	0.031	0.185

Footnote: Gene sets enrichment scores were calculated with $P \leq 0.05$ and had a false discovery rate (FDR) $\leq$ 0.25 (25%)

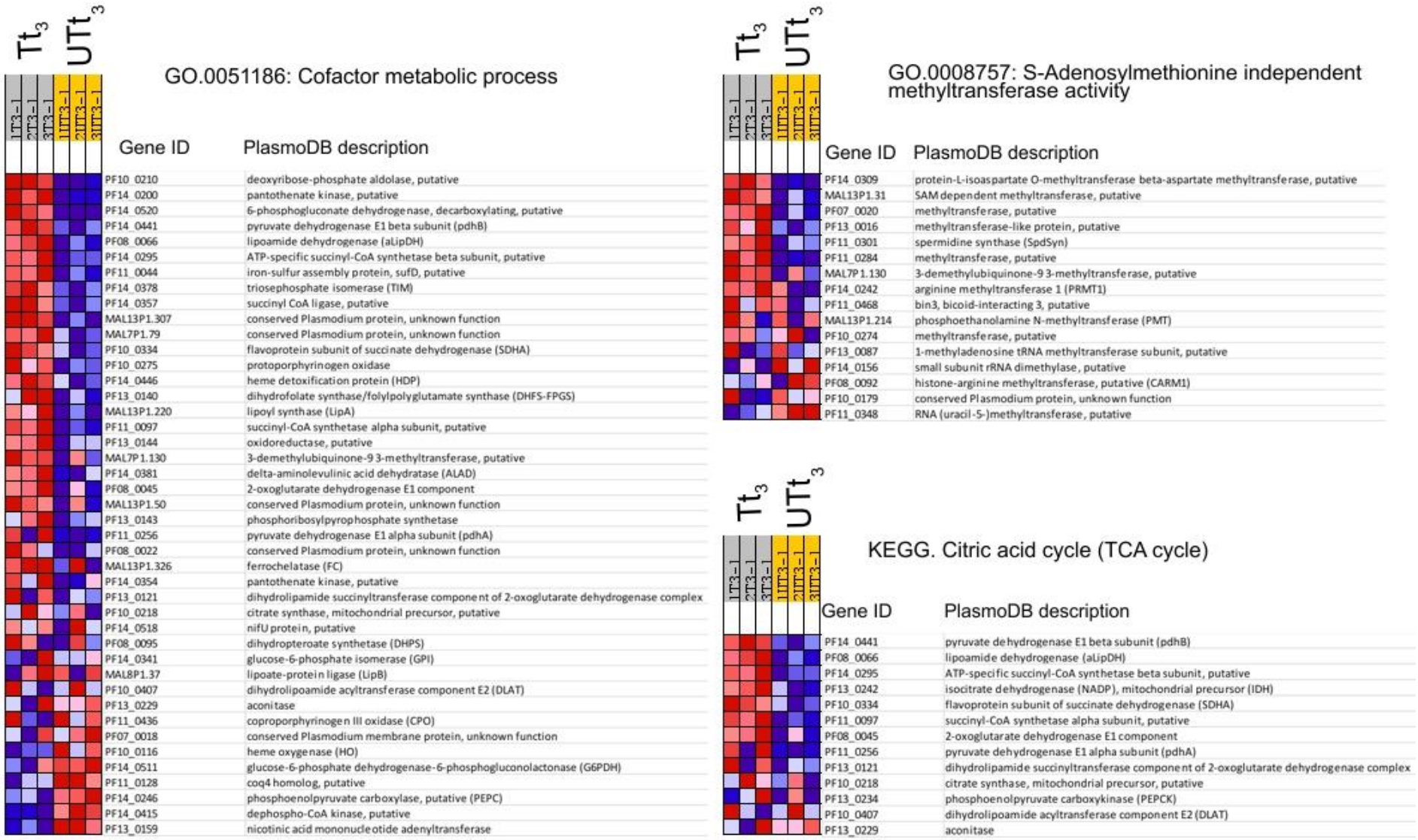


Figure 4.9: GSEA analysis of Tt_3 vs. UTt_3. Heat maps of gene sets found **to be enriched in GSEA. GSEA was calculated with $P \leq 0.05$ and had a false discovery rate (FDR) ≤ 0.25 (25%).** Red indicates increased transcript expression levels in the respective arrays, whereas blue indicates decreased transcript expression.

Transcripts within the purine metabolic pathway were also affected. Both MAL13P1.118 (3',5'-cyclic nucleotide phosphodiesterase, putative) and MAL8P1.150 (adenylyl cyclase beta, putative) had decreased expression at t_3 (Table 4A.1). The former transcript protein product is involved in the recycling of guanosine purine nucleosides, and converts 3', 5'-cyclic GMP (cGMP), a cell signalling metabolite, to GMP. In contrast, adenylyl cyclases convert ATP to 3', 5'-cyclic AMP (cAMP), a pivotal second messenger involved in signal transduction. AMP deaminase (MAL13P1.146), which convert AMP to IMP, had increased expression. Of the 14 genes which are responsive to oxidative stress (GO term: 0006979) only thioredoxin peroxidase 1 (Trx-Px1, PF14_0368) had increased abundance with a $\log_2$FC of 0.94 (Table 4A.1). A large proportion of transcripts which had decreased expression were kinases either involved in protein amino acid phosphorylation or in a phosphorous metabolic process (Table A4.1). A putative calmodulin transcript (PF10_0301) which had decreased expression was also correlated with phosphorylation. The predicted calmodulin transcript encodes for a small calcium-dependent protein critical for the activation of kinases.

4.3.6 Comparison to other transcriptomes

Transcriptomic fluctuations at the most prominent t_3 stage was compared to that of CQ treated *P. falciparum* (HB3), artesunate treatment of *P. falciparum* (FCR3) after 3h [262], febrile temperature changes (41°C, strain 3D7) [235], and a study which used WR99210 a lethal antifolate compound (Figure 4.10) [241]. The gene expression profiles were also compared to a large study in which 20 different compounds were tested on various late schizont stages of the parasites (3D7) [242]. These comparisons showed that 196 transcripts were uniquely expressed in 4PEHz-treated parasites (Figure 4.10). The majority of these unique transcripts, 54% (90/196), were hypotheticals with no definitive functions ascribed to the transcript products. Some transcripts (6.6%, 13/196) were associated with protein amino acid phosphorylation as well as cell cycle regulation and signal transduction pathways (Table A4.3 in the Appendix). Two unique transcripts are involved in cofactor biosynthetic processes; a putative hydroxyethylthiazole kinase (PFL1920c) and nicotinic acid mononucleotide adenyltransferase (PF13_0159). The former transcript is involved in intermediate steps leading to the production of thiamine pyrophosphate (TPP). The latter is involved in forming nicotinamide adenine dinucleotide (NAD^+).

Of the transcripts differentially expressed in t_3 during 4PEHz treatment, 14 corresponded to transcripts expressed during febrile temperature fluctuations. These were involved in protein phosphorylation, proteolysis and DNA replication. One hundred and twenty three transcripts

corresponded to CQ-treatment expression patterns, and these were involved in pathogenesis and proteolysis. Lipoamide dehydrogenase (aLipDH, PF08_0066) was also differently expressed in CQ-treated parasites, similar to 4PEHz treatment (Figure 4.10). Five hundred and twenty five transcripts also corresponded to a large 20-drug transcriptome study involving parasite-growth inhibitors such as atremisinin, quinine and trichostatin A (Figure 4.10) [242]. According to functional annotation clustering, these transcripts were mainly involved in protein amino acid phosphorylation, as well as cell division and chromosome partitioning (results not shown, annotation in DAVID [260, 261]).

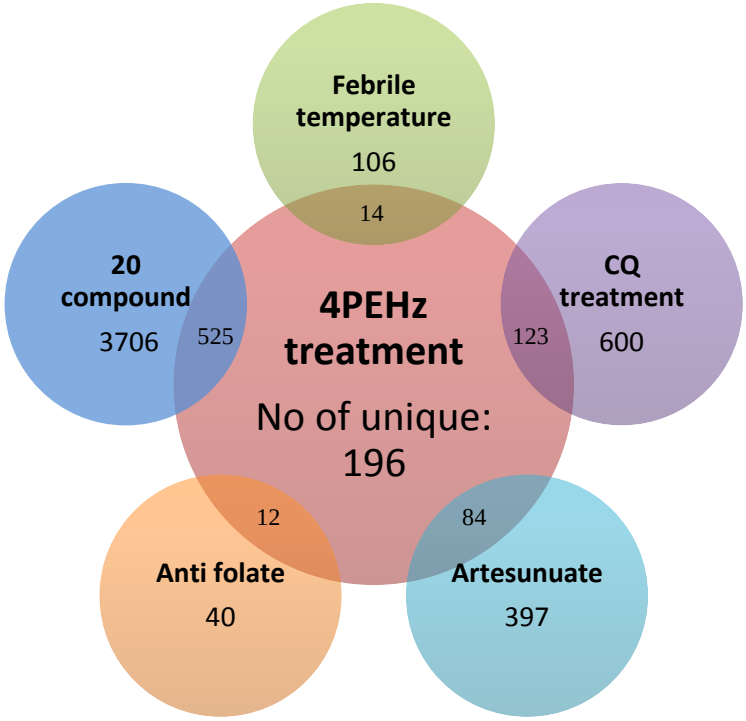


Figure 4.10: Venn diagram of differentially expressed transcripts during t_3 in 4PEHz-treated parasites. The differentially expressed were compared to transcriptome datasets of CQ, artesunate, febrile temperature treatment of *P. falciparum*.

The 4PEHz-treated transcriptome (525 transcripts) of the third time point was clustered, using hierarchical average linkage clustering, to corresponding transcripts from a 20-molecule growth inhibitor transcriptome dataset (Figure 4.11). Clustering revealed that 4PEHz had similar responses to staurosporine, PMSF and apicidin (Figure 4.12). A cluster of transcripts with decreased expression (Figure 4.11) were highly correlated with Pearson correlation factor of 0.93. This cluster of around 40 transcripts were ranked according to their GO-associated biological processes and were found to have involvement protein phosphorylation (GO:0006468), mismatch repair (GO:0006298) and response to DNA damage stimulus (GO:0006974). In the t_3 time point of 4PEHz-treated parasites several putative and known transcription factors had decreased expression (Appendix 4, Table 4A.1). The expression levels of these transcription factors in 4PEHz-treated

parasites were clustered with apicidin and staurosporine transcriptomes (Correlation 0.72, Figure 4.12). Transcripts with similar expression patterns in a sub-cluster (correlation 0.94) were associated with protein phosphorylation (*P* value = 0.008, GO term: 0043687), the cell cycle (*P* value = 0.018, GO term: 0007049) and cytokinesis (*P* value < 0.001, GO term: 0000910). A transcript encoding for a transcription factor, PF11_0442 (transcription factor with AP2 domains, AP2-O) had a decreased expression with $\log_2$FC of -2 in 4PEHz treated parasite, which corresponded with apicidin and staurosporine expression levels (Figure 4.12). Additional clustering was performed to determine the extent to which transcription factor expression was comparable from the different drug treatments (Figure 4.13). Expression of PFL0465c (zinc finger transcription factor (krox1) and PF11_0477 (CCAAT-box DNA binding protein subunit B) in 4PEHz-treated parasites was similarly decreased compared to apicidin transcript levels (Figure 4.13). In a study by Chaal *et al.* several Apicomplexan AP2 (ApiAP2) putative transcription factor proteins, which are thought to play a role in regulating stage specific gene expression, were shown to be affected by apicidin [263]. These results therefore also link up with transcriptome data from Hu *et al.* (Figure 4.13) suggesting disregulation of ApiAP2 transcription factors, moreover suggests that 4PEHz affects the same cluster of transcriptional regulators [242]. Through attenuation of PLP biosynthesis these transcription factors may be disregulated.

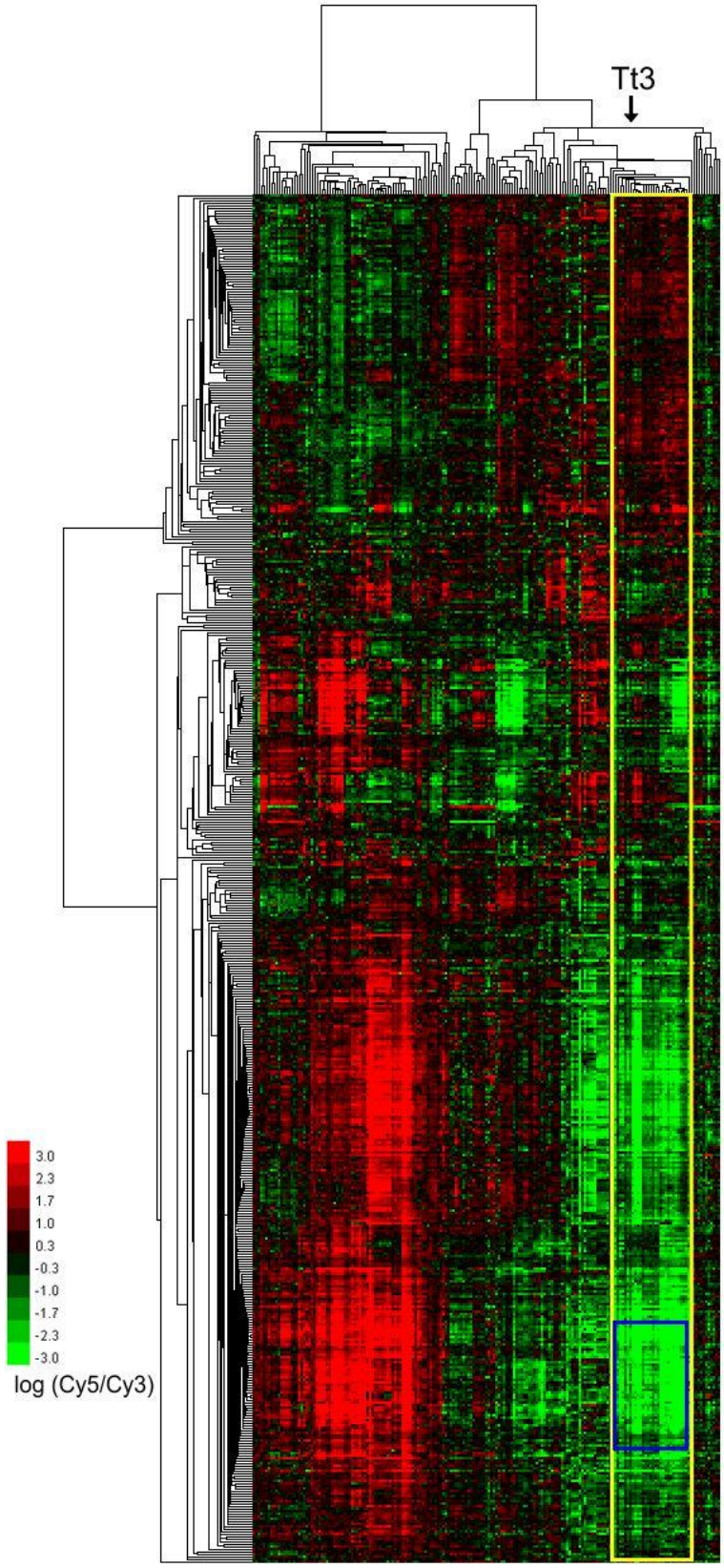


Figure 4.11: Transcriptional responses of *P. falciparum* treated with 4PEHz compared to other parasite growth inhibitors. The third time point t_3 (UT3 in figure) was clustered using hierarchical average linkage clustering with transcripts from a study involving 20 different growth parasite inhibitors. The boxed regions, which shows the correlation of 4PEHz responses, are expanded in the next figure. Clustering was performed in Cluster v 2.11 [255] and cluster trees were generated in TreeView v 1.1.6r2 [264, 265].

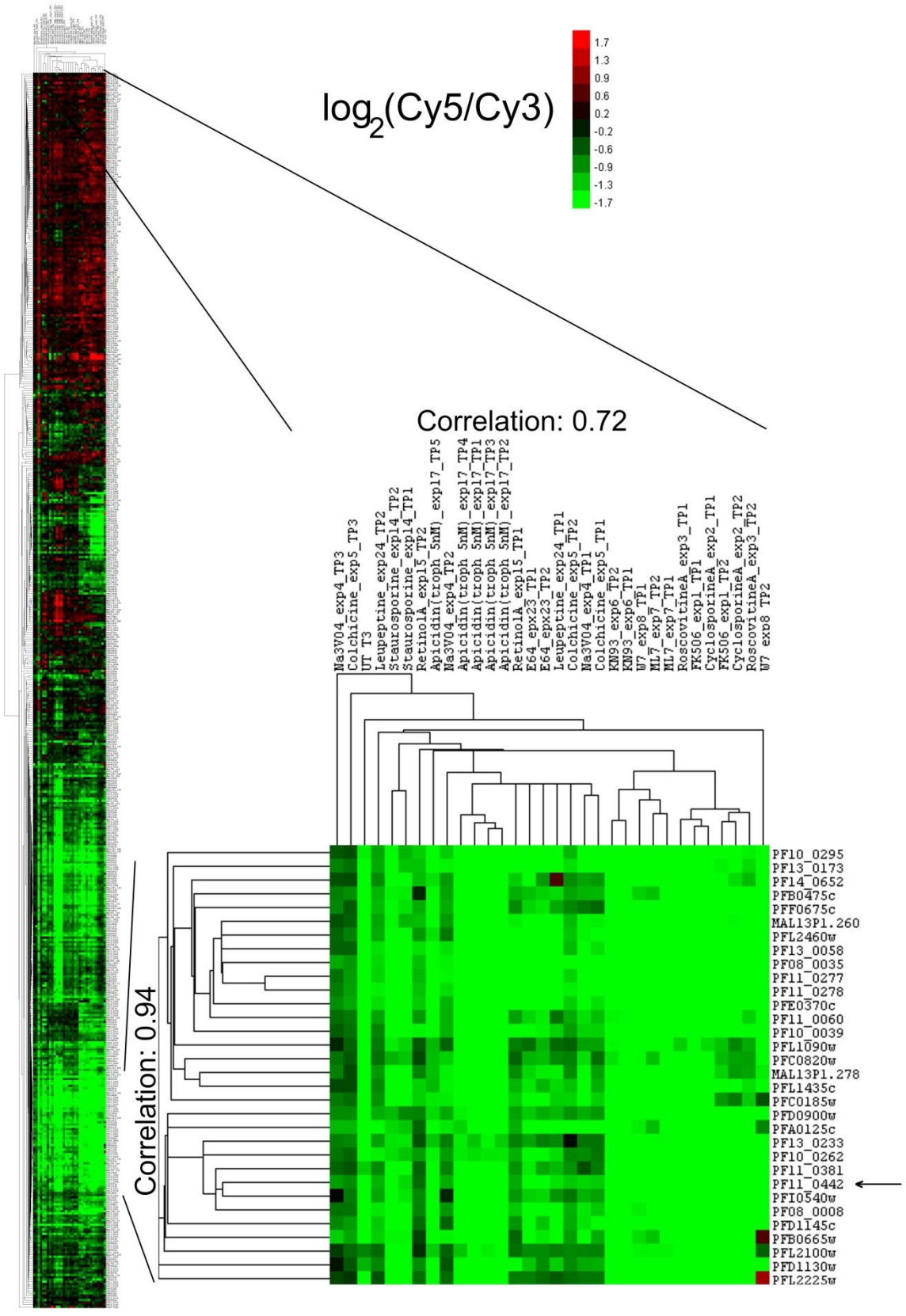


Figure 4.12: 4PEHz treament in t_3 was shown to correlate with apicidin and staurosporine treatment of parasites. Transcripts from 4PEHz-treated parasites in the third time point (UT3 T3) clustered with corresponding transcripts affected by staurosporine, apicidin and cyclosporine [242]. The expression pattern of a cluster (correlation: 0.94) revealed highly similar decreased expression patterns comparable between the treatments.

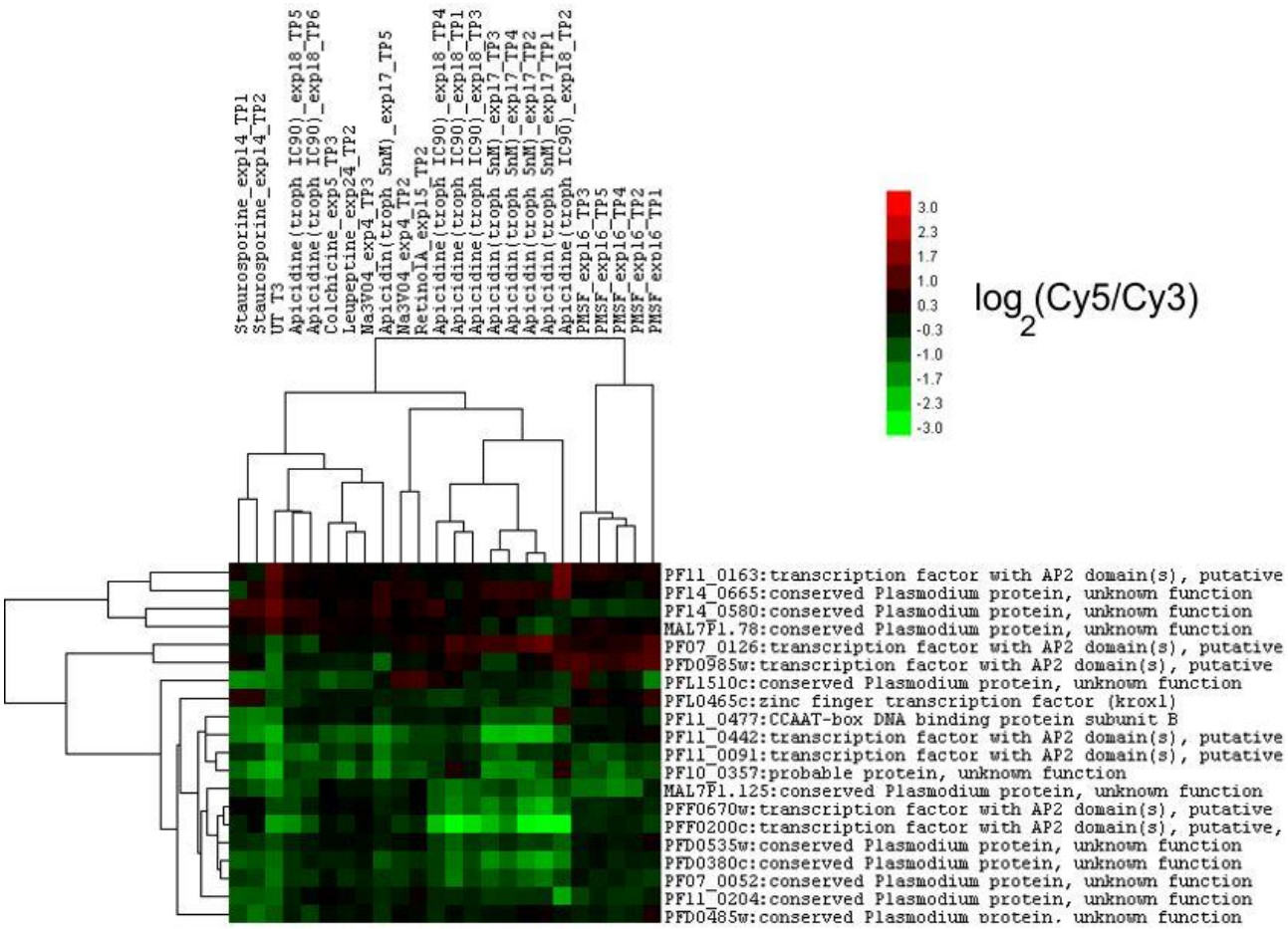


Figure 4.13: Transcription factors affected by 4PEHz-treatement. Several transcription factors with significant differential expression ($log_2FC > 0.75$ or < -0.75) in t_3 during 4PEHz-treatement were clustered with transcript expressions values from apicidin, PMSF and staurosporine-treated parasites [242]. Transcript expression was comparable to staurosporine as well as apicidin-related expression patterns, and the cluster had a correlation of 0.55.

4.3.7 Validation of microarrays using real-time PCR

Microarray data were validated using real-time PCR. The three reference target genes of decreased abundance; (PF14_0224), (PFB0695c), and (PFI1370c) were selected. Three transcripts which had increased expression in t_3 were selected; (PF11_0256), (PF14_0155) and (PF08_0066). Three housekeeping genes; LDH (PF13_0141), cyclophillin (cyc) (PFE0505w) and StSyn (PF07_0073) had geNorm scores values below 0.2 indicating that these reference genes had very stable expression, and had comparable expression in both treated and untreated parasites (Figure 4.14). From the expression of the reference genes expression in both treated and untreated parasites an expression normalisation factor is calculated, which is used to normalise expression of the targets relative to expression of the reference gene.

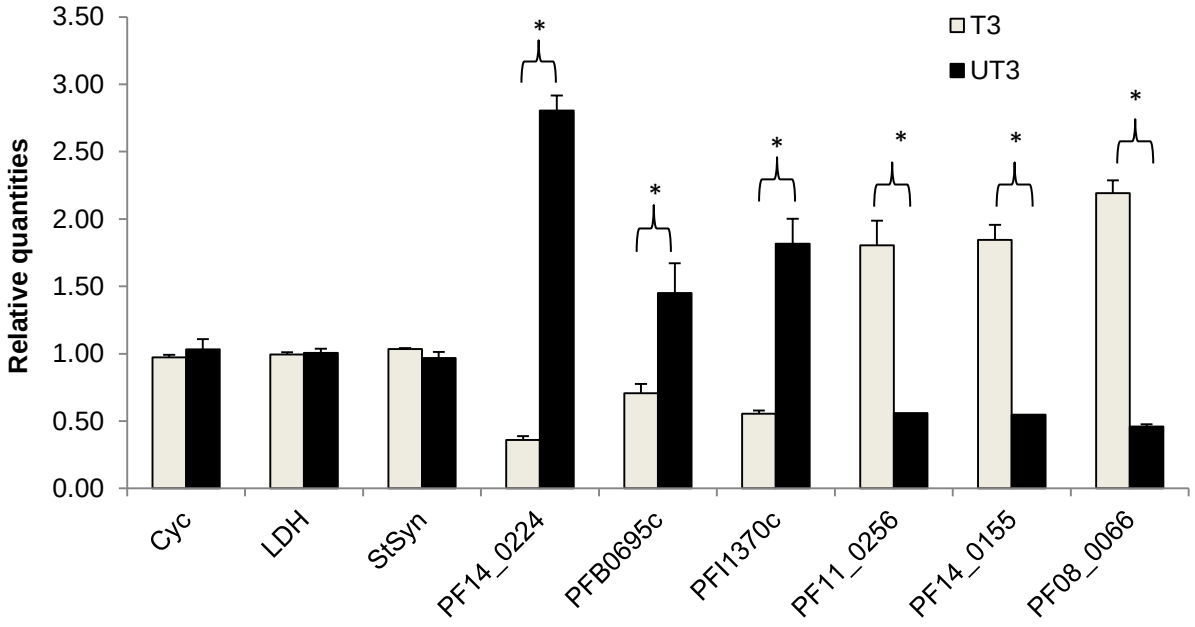


Figure 4.14: Real-time qRT-PCR validation of the t_3 time point transcript expressions from microarrays. Statistical significance is indicated with * which represents $P < 0.05$ in a two-tailed Student's t-test with unequal variance. Error bars indicate the SEM for two independent experiments performed in duplicate.

Six transcripts with differential expression in t_3 time points were selected for real-time PCR validation. Three target transcripts; PF14_0224, PFB0695c, and PFI1370c were had greater abundance in the UTt_3 treated parasite population compared to Tt_3 (Figure 4.14). The relative Tt_3:UTt_3 FC levels of these three transcripts were comparable to FC levels found in microarray comparisons (Table 4.8). Similarly, the three transcripts PF11_0256, PF14_0155 and PF08_0066 were shown to have greater levels in Tt_3 samples (Figure 4.14), and had Tt_3:UTt_3 FC values which correlated with microarray data (Table 4.8).

Table 4.8: Real-time PCR confirmation of microarray data. Six transcripts which had decreased expression and increased expression respectively were validated with real-time PCR. All six transcripts had FC levels comparable to microarray data from t_3.

Target transcript	PlasmoDB description	Real-time PCR Tt_3:UTt_3 FC	Tt_3:UTt_3 FC from microarray	
			FC	Log_2FC
PF14_0224	serine/threonine protein phosphatase	0.13	0.16	-2.65
PFB0695c	acyl-CoA synthetase, PfACS8	0.49	0.53	-0.91
PFI1370c	phosphatidylserine decarboxylase	0.31	0.52	-0.94
PF11_0256	pyruvate dehydrogenase E1 alpha subunit	3.23	3.13	1.65
PF14_0155	serine C-palmitoyltransferase, putative	3.38	2.79	1.48
PF08_0066	lipoamide dehydrogenase (aLipDH)	4.76	2.43	1.28

4.3.8 Proteome analysis

Proteins were extracted from 4PEHz treated and untreated *P. falciparum* parasites to study the effects of the compound on the parasite proteome. Protein extract samples were from the same treatment study used for microarray analysis. The proteins were separated on a 1D SDS-PAGE gel (Figure 4.15) after which the lanes were excised and submitted for MS-based proteomic quantitation and identification. A total of 33 unique proteins were detected in the Tt_3 sample which were not present in UTt_3 parasites. A summarised list is given in Table 4.7. A comprehensive list of proteins, with similarities scores and quantitative information regarding protein identity is given in Table A4.4 and Table A4.5 of Appendix 4.

GO annotation of the proteins present in 4PEHz-treated extracts suggested the involvement of proteins in single-stranded DNA specific endodeoxyribonuclease activity (GO:0000014), RNA-dependent DNA replication (GO term: 0006278), arginine biosynthetic (GO term: 0006526) as well as pyrimidine base biosynthesis (GO term: 0019856) (Appendix 4, Table A4.4). Some GO processes correlated with transcriptome data, and include activation of DNA repair processes with an uncharacterised conserved *Plasmodium* protein (PF13_0155) which is predicted to have involvement in double strand break repair (GO:0006302) and a putative ATP-dependent DNA helicase (PF14_0278). Carbamoyl phosphate synthetases (CPS), which are involved in L-glutamine catabolic processes, were also present in the treated protein fractions. This enzyme utilises ammonia to form carbamoyl phosphate from ATP and CO_2.

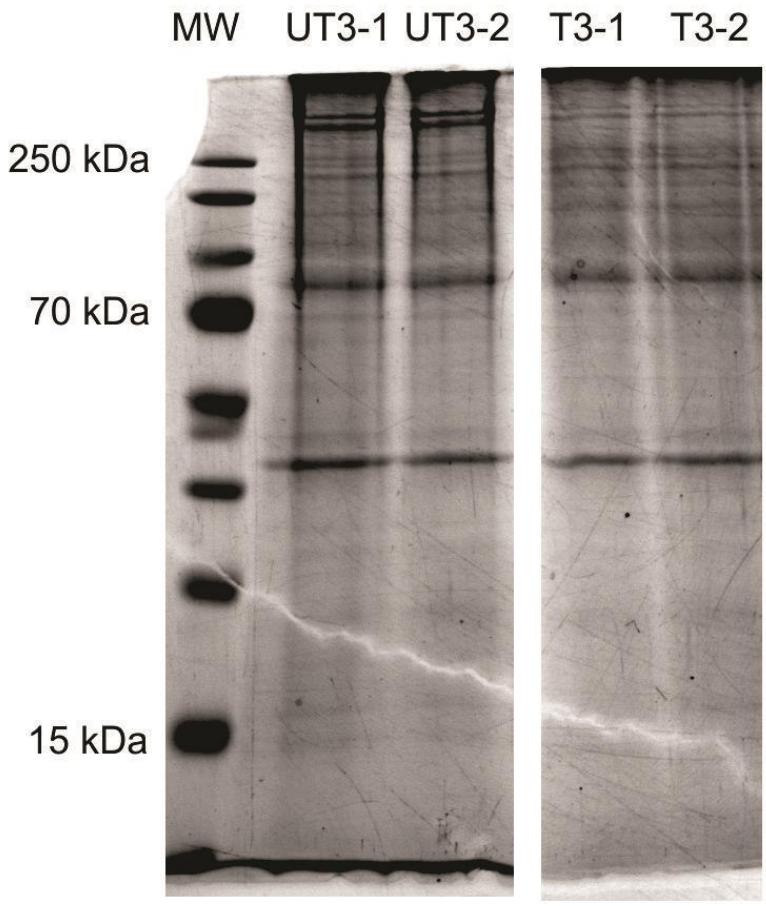


Figure 4.15: SDS-PAGE of protein extracted from untreated and treated 4PEHz parasites at time point t_3. Proteins were extracted from 4PEHz-treated and untreated parasites at time point t_3 of the microarray study. Proteins were quantified using a copper-based colourimetric assay, and samples containing 60 µg of protein were loaded and separated on a SDS-PAGE gel. Both lanes of the each sample were excised and submitted from MS-based protein identification.

Adenylyl cyclase beta (MAL8P1.150) protein was detected in the treated fractions, whereas this particular transcript had decreased expression in transcriptomic profiles of the t_3 stage (Table 4.9). Some of the proteins which correlated with transcriptomic expression patterns included a phosphoinositide-binding protein (MAL7P1.108) The catalytic subunit of DNA polymerase epsilon was also present in the treated protein fraction, corresponding to observed transcript expression data (Table 4.9). These results could suggest that expression of these proteins are regulated at a transcriptional level, rather than post transcriptionally.

Table 4.9: Summarized list of proteins identified in 4PEHz-treated parasites. Some proteins corresponded to differentially expressed transcripts from Tt_3 vs. UTt_3. For a comprehensive list of proteins see Appendix 4, Table A4.4 and Table A4.5.

PlasmoDB ID	Present in Tt_3 vs. UTt_3 transcitpomic data (log_2FC)	PlasmoDB Description
MAL13P1.278	-2.23	serine/threonine protein kinase, putative
MAL7P1.108	-0.94	phosphoinositide-binding protein, putative
MAL8P1.150	-1.50	adenylyl cyclase beta, putative
PF10_0184	-1.89	conserved *Plasmodium* protein, unknown function
PF10_0262	-3.35	conserved *Plasmodium* protein, unknown function
PF11_0417	-1.06	conserved *Plasmodium* protein, unknown function
PF13_0079	-2.56	conserved *Plasmodium* protein, unknown function
PF14_0123	1.12	conserved *Plasmodium* protein, unknown function
PFE0440w	-1.52	conserved *Plasmodium* protein, unknown function
PFF1470c	0.77	DNA polymerase epsilon, catalytic subunit a, putative
PFL0030c	-1.12	erythrocyte membrane protein 1, PfEMP1
PFL2505c	-2.07	conserved *Plasmodium* protein, unknown function
PF13_0080	-	telomerase reverse transcriptase, putative
PF13_0044	-	carbamoyl phosphate synthetase

4.3.9 Interaction networks

None of the proteins detected during 4PEHz treatment formed part of *Pf*Pdx1 or PLP association networks. Therefore, functional linkages were inferred from the transcriptome components of the interactome, and were evaluated based on their transcript expression patterns. Similar correlations from gene expression profiles were used to create a comprehensive interaction network in *P. falciparum* parasites [266]. Using STRING *(*Search Tool for the Retrieval of Interacting Genes/Proteins, v 9.0) [267, 268], most of the transcripts previously listed as PLP-dependent enzymes were flagged as part of the interaction network (Table 4.10). The discovery of Pdx1 and Pdx2 is relatively recent, and for this reason *Pf*Pdx2 has not been classified as an interacting partner of *Pf*Pdx1 in STRING (both *Pf*Pdx1 and *Pf*Pdx2 are not present in PlasmoMAP either). Interacting partners were identified based on previously reported literature or neighbouring transcripts located near to Pdx1, and are included in the interaction network. The predicted and known interacting proteins of *Pf*Pdx1 are shown in Figure 4.16 A, and listed in Table 4.10. There was an overall enrichment of genes associated with *Pf*Pdx1, with the majority of these having increased expression during t_3, however not statistically significant (or above log_2FC of 0.75). Phosphoglycerate kinase (PGK, PFI1105w) were found to be part of the interaction network of *Pf*Pdx1. This enzyme functions in glycolysis to reversibly convert 1,3-bisphosphoglycerate to 3-phosphoglycerate, and follows after the conversion of G3P to 1,3-bisphosphoglycerate by glyceraldehyde-3-phosphate dehydrogenase. This provides evidence that processes indirectly associated with the *Pf*Pdx1 substrate G3P were affected by 4PEHz.

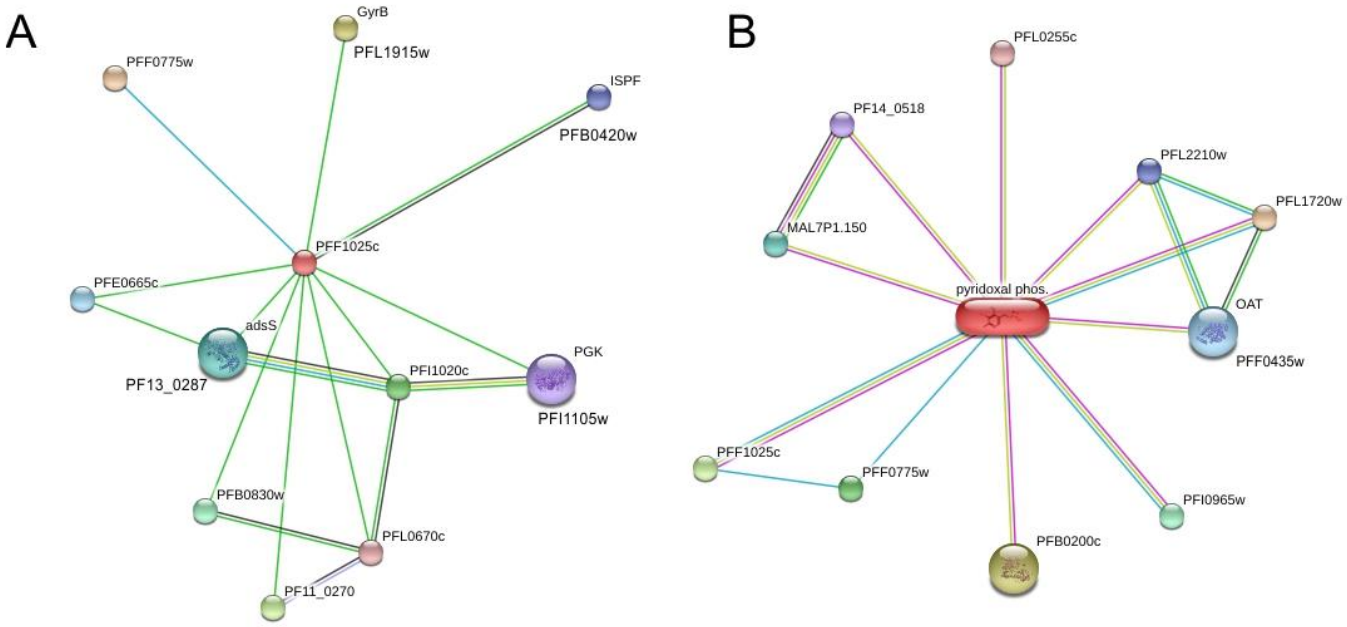


Figure 4.16: Predicted interaction networks of *Pf*Pdx1 and PLP. A) Interacting partners of *Pf*Pdx1 were identified in STRING. These interactions are based neighbouring locations of the transcripts in other organisms, as well as from derived from literature. **B)** The PLP chemical interaction network as predicted in STITCH. These nodes are linked based on neighbouring locations of genes in other genomes, such as SHMT (PFL1720w) and delta-aminolevulinic acid synthetase (PFL2210w). Putative homologues of OAT (PFF0435c) and delta-aminolevulinic acid synthetase (PFL2210w) in *Chlorobium luteolum* are predicted to interact with another [269].

Additional interaction partners which utilise PLP were identified using STITCH, which link chemical interaction groups with proteins and gene information (Table 4.10 and Figure 4.16 B) [269, 270]. PLP-dependent enzyme transcripts have been previously discussed, however STRING identified additional interacting partners based on association in curated databases as well as neighbouring locations of genes in other genomes. Of the predicted transcripts associated with PLP a putative pyridoxal 5'-phosphate dependent enzyme class III (PFI0965w) and SHMT (PFL1720w) were found to be enriched, having low level increased transcript expression. Delta-aminolevulinic acid synthetase (PFL2210w, ALAS) and SHMT (PFL1720w) are situated on the same loci in α-proteobacteria [269], however in *P. falciparum* the expression pattern of these two transcripts suggests that these are not co-regulated (Table 4.10).

Table 4.10: Predicted interacting partners of *Pf*Pdx1 (PFF1025c) and PLP. Transcript and proteins which form part of the *Pf*Pdx1 interaction network were identified using STRINGS. Interaction scores are derived from combined hits of the PFF1025c query based on neighbouring genes, text and database mining results as well as co-expression queries.

PlasmoDB ID	PlasmoDB description	log_2 FC Tt_3:UTt_3	Adj. P-value	Interaction score	Predicted GO process involvement
Interacting partners of *Pf*Pdx1 (PFF1025c) (STRINGS predictions)					
PFF0775w	pyridoxine kinase (PdxK)	0.234	0.117	0.899	salvage of pyridoxine, vitamin B6 biosynthetic
PFL1915w	DNA gyrase subunit B (GyrB)	0.588	0.013	0.717	DNA topological change
PF11_0270	threonyl-tRNA synthetase,Threonine--tRNA ligase (ThrRS)	0.316	0.067	0.596	tRNA aminoacylation,
PFI1020c	inosine-5'-monophosphate dehydrogenase	-0.183	0.264	0.595	oxidation-reduction process
PFB0830w	40S ribosomal protein S26e, putative	0.285	0.121	0.575	translation
PF13_0287	adenylosuccinate synthetase (adsS)	0.313	0.038	0.551	purine nucleotide biosynthetic process
PFE0665c	GTP binding protein, putative	0.574	0.002	0.532	barrier septum formation
PFB0420w	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF)	0.137	0.429	0.478	terpenoid biosynthetic process
PFI1105w	phosphoglycerate kinase (PGK)	0.719	0.003	0.463	glycolysis
PFL0670c	bifunctional aminoacyl-tRNA synthetase, putative	0.304	0.038	0.455	prolyl-tRNA aminoacylation
Interacting partners which utilise PLP (STITCH predictions)					
PFL1720w	serine hydroxymethyltransferase (SHMT)	0.469	0.004	0.994	L-serine and glycine metabolic process
PFB0200c	aspartate aminotransferase,aspartate transaminase (AspAT)	0.188	0.371	0.993	biosynthetic process
PFF1025c	pyridoxine biosynthetic enzyme (Pdx1)	-0.155	0.305	0.977	pyridoxal phosphate biosynthetic process
PFF0775w	pyridoxine kinase (PdxK)	0.234	0.117	0.957	null
PFI0965w	pyridoxal 5'-phosphate dependent enzyme class III, putative	0.471	0.038	0.952	null
MAL7P1.150	cysteine desulfurase, putative (NFS)	-0.348	0.050	0.934	metabolic process
PFF0435w	Ornithine aminotransferase (OAT)	0.120	0.415	0.92	null
PFL2210w	delta-aminolevulinic acid synthetase (ALAS)	-0.267	0.049	0.914	tetrapyrrole biosynthetic process
PF14_0518	nifU protein, putative	0.274	0.063	0.798	iron-sulphur cluster assembly
PFL0255c	UGA suppressor tRNA-associated antigenic protein, putative	0.427	0.080	0.787	null

4.4 Discussion

PLP is a vital co-factor in more than 100 different biochemical reactions, and perturbations in PLP biosynthesis are therefore expected to be severely detrimental to *P. falciparum* parasite growth. PLP plays a role in amino acid metabolism, and acts as a prosthetic group for enzymes predominantly involved amino acid metabolism. In the malaria parasite, PLP is required for the decarboxylation of L-ornithine to putrescine by ODC, where putrescine is subsequently used as scaffold diamine during the formation of polyamines spermidine and spermine [271]. Moreover, parasitic polyamine levels have been correlated to IDC of the parasite during asexual maturation, and depletion of polyamines results in parasite cytostasis [203, 272]. This further underscores the importance of PLP biosynthesis in the parasites. Apart from metabolic processes reliant on PLP, a recent antioxidant role has been ascribed to this molecule and is involved in elimination of ROS [141]. Considering the ubiquitous involvement of PLP in a multitude of different cellular processes, disruptions in PLP biosynthesis could easily result in escalated cellular stress responses. The involvement of PLP in combating oxidative stress would similarly suggest stress-related perturbations would be affected due to diminished PLP levels.

Consequences of PLP deficiency in humans include high homocysteine and cystathionine levels during methionine loading [273]. This was attributed to cystathionine β-synthases and cystathionase, both which are PLP-dependent enzymes which function to remove homocysteine. There is currently no evidence for the existence of similar proteins in *P. falciparum* parasites. In yeast (*S. cerevisiae*) depletion of PLP through inhibition of PdxK and knock-out of *de novo* PLP biosynthetic enzymes Pdx1 and Pdx2, resulted in DNA lesions and DNA damage responses [274]. This was thought to be related to reduced formation of N^5,N^{10}-methylene tetrahydrofolate (CH_2-THF) by SHMT, which in turn lead to reduced dTMP levels. These disruptions in nucleotide biosynthesis resulted in accumulation of uracil, and subsequently over-incorporation of this nucleobase into DNA, resulting in DNA lesions and genome instability [274].

Microarrays were utilised to determine the extent of transcriptome level responses caused by 4PEHz treatment. A common reference probe design was used in the microarrays. The reference pool consisted of cDNA from all the T and UT samples, and allows comparison of an individual transcript expression level from each condition in a separate array to a common reference. This indirect method is advantages as transcripts from different arrays, in this case time points as well as treatments, can be accurately compared. Normalisation of the common reference pool, thereby dye intensity, is necessary to account for between-array variations.

A large number of transcripts (787) were affected during the last t_3 time point, corresponding to around 10% of the known transcriptome. Previously classes of small molecule inhibitors that induced similar large changes (between ~5 – 10%) of the transcriptome included EGTA, trichostatin A, staurosporine, and apicidin [242]. These molecules were suggested to arrest the IDC of the parasites, with the exception of apicidin which was thought to deregulate the IDC owed to its HDAC inhibitory action [242]. Moreover, clustering of the transcript expression data from 4PEHz-treated parasites revealed gene expression correlations compared to apicidin and staurosporine treatment of schizonts [242]. The systematic changes in the transcriptome of 4PEHz-treated parasites from a few affected transcripts during the first two time points followed by large scale perturbations during the last time point, comparable to IDC arrest, suggested that 4PEHz had a profound effect on the parasites. These disruptions could have further led to transcriptional deregulation. This underscored the importance of PLP metabolism, more specifically *de novo* biosynthesis within the parasite.

As highlighted in during functional annotation and clustering of differential expressed genes in t_3, more than 10 transcription factors had decreased expression during 4PEHz treatment, and some of these transcript expression levels correlated to apicidin-treated parasites. PLP post-translationally affects the activity of transcription factors [275, 276], and also some transcription factors specific to PLP metabolism which modulate the expression of the Pdx1 and Pdx2 genes, as explained below [277]. PLP had been shown to inhibit the DNA binding capability of the HNF1 transcription factor in *E. coli* [276]. Moreover, PLP is post-translationally conjugated to mouse nuclear receptor interacting protein 140 (RIP140), and increases the co-repressive transcriptional activity, affecting gene expression [275]. The conjugated RIP140 was also shown to interact with HDAC [275]. Decreases in PLP levels of 4PEHz-treated parasites could either result in activation of certain transcription factors, or diminish of potential co-repressive functions (enhance activation), as expected for RIP140. The overall activation of transcription factors was reflected as compensatory responses which aimed to decreased transcription of transcription factors during 4PEHz treatment. This also provides a link between apicidin and 4PEHz treatment of parasites, these molecules could affect the same transcription factor, resulting in similar transcriptomics responses.

Wide-spread transcriptional perturbations during 4PEHz treatment of parasites could be linked to DNA damage repair processes. Topoisomerases regulate the topological states of DNA during transcription and repair [278]. PLP has been postulated to regulate topoisomerases and was shown to reversibly complex with DNA topoisomerases I in *Candida guilliermondii* by binding to the ε-amino group of an active site lysine [278]. Binding of PLP inhibits the DNA cleavage reaction catalysed by the enzyme. Depletion of PLP in 4PEHz-treated parasites could have led to release of

PLP from topoisomerases, activating DNA cleavage. The transcriptional responses were reflected as decreased expression of DNA topoisomerases (MAL13P1.328, DNA topoisomerase VI, b subunit, putative), and supports that DNA damage could have been caused during PLP depletion.

Transcripts of PLP-dependent enzymes were affected by 4PEHz treatment (Figure 4.17). This implies that transcriptomic perturbations could be functionally linked to PLP. This confirms that 4PEHz affected PLP metabolism and lends support to specific targeting of *Pf*Pdx1 by 4PEHz *in vitro*. Of the PLP-dependent enzyme transcripts a putative SHMT (PF14_0534) transcript had increased expression by 3-fold during the first 12 h of the treatment period. SHMT is involved in one carbon transfer reactions, and utilises serine to generate the CH_2-THF cofactor (Figure 4.17). There are two predicted SHMT isoforms in *P. falciparum;* PFL1720w and PF14_0534. The former transcript encodes for the cytosolic form of SHMT or *Pf*SHMTc [279]. The PF14_0534 transcript, which had increased expression during 4PEHz treatment, has 18% identity compared to *Pf*SHMTc, and is the mitochondrial form of SHMT [279]. Both *Pf*SHMTc and *Pf*SHMTm have stage-specific localisation within the parasite, and *Pf*SHMTc has been shown to predominantly have cytosolic localisation throughout the parasite IDC [279]. *Pf*SHMTm localises in the mitochondria of the parasite, and also in the apicoplast during the early trophozoite and schizont stages [279]. Intra-erythrocytic expression of *Pf*SHMTm transcript was previously shown to peak at 28 HPI and the lowest expression levels were at 7 HPI in the 3D7 strain [225]. The deviation of SHMTm transcript from known expression patterns highlighted the importance of this transcript and its dependence on PLP. During the t_1 stage only 25 transcripts were differentially expressed, compared to the t_3 stage with more than 700 affected transcripts. This perturbation in SHMT transcript expression is therefore particularly interesting.

GSEA showed that 4PEHz treatment of *P. falciparum* parasites results in positive enrichment of transcripts involved in THF metabolism. Three other transcripts formed part of this gene set which included dihydrofolate synthase/folylpolyglutamate synthase (PF13_0140, DHFS-FPGS), putative glycine cleavage T protein (GCVT) (PF13_0345), and lipoamide dehydrogenase (aLipDH) (PF08_0066). *P. falciparum* possesses two LipDH isoforms; mitochondrial (mLipDH) and apicoplast (aLipDH). The aLipDH has an apicoplast localisation signal, even though it is annotated to be involved in folate metabolism, aLipDH forms part of the PDC in the apicoplast [280, 281]. Mitochondrial SHMT provides glycine to the glycine cleavage complex (GCV) or glycine decarboxylase complex to generate additional THF [279, 282]. The increased enrichment of the putative glycine cleavage T protein (GCVT) transcript which forms part of the GCV also suggested compensatory THF metabolic responses. CH_2-THF donates a carbon in a reaction catalysed by thymidylate synthases (TS) which converts dUMP to dTMP. Recycling of the oxidised

dihydrofolate from TS is accomplished by dihydrofolate reductase (DHFR), which in *P. falciparum* is a component of a bifunctional complex consisting of DHFR and TS [283]. Surprisingly at t_3 the transcripts of the bifunctional DHFR-TS had increased expression to around 1.7-fold. These results suggests an increased requirement for folates and recycling thereof during 4PEHz treatment, and therefore presumably PLP depletion.

One consequence of PLP depletion is DNA lesions caused by disruptions in dTMP biosynthesis which is further associated with THF biosynthesis [274]. In 4PEHz-treated parasite these disruptions were linked to affected folate biosynthesis by SHMT and down-stream folate utilisation by TS. Disruptions in folate metabolism of parasites treated with 4PEHz could have been due to PLP depletion, resulting in decreased levels of CH_2-THF produced by SHMT, thereby affecting down-stream folate recycling processes. The sequential series in which transcript deviations were observed agrees with this notion. SHMT had increased expression during the t_1 time point, followed by compensatory increased in folate-related transcripts only at the t_3 time point. The onset of PLP depletion could have been rapid, triggering early compensatory responses, followed by a more dramatic rescue response which attempted to balance folate homeostasis.

*Pf*Pdx2 transcripts had increased expression during 4PEHz treatment (Figure 4.17). The *Pf*Pdx1 transcript was not shown to have significant transcript fluctuations throughout the monitored life cycle period. *Pf*Pdx1 and *Pf*Pdx2 are obligate heterodimers which assemble into the PLP synthase complex and the cellular abundance of each counterpart is expected to affect the activity of the other. Increased *Pf*Pdx2 transcript levels demonstrated that a metabolic perturbation can have feedback and results in transcript compensation directly affiliated with the producing enzyme. Presumably increased *Pf*Pdx2 transcript levels aims to increase *Pf*Pdx2 protein abundance, which would have an effect on the activity of the PLP synthase complex that produces PLP. *Pf*Pdx2 monomers are activated by *Pf*Pdx1, which should be inhibited by 4PEHz, potentially resulting in wasteful hydrolysis of L-glutamine and the generation of excess ammonia, and results supporting this speculation are discussed further below. Transcripts of pyridoxine/pyridoxal kinases were not affected by 4PEHz treatment, also corroborating with the notion that salvage of unphosphorylated PN and PL minimally contributes towards intracellular PLP levels (unpublished experiments, C. Wrenger).

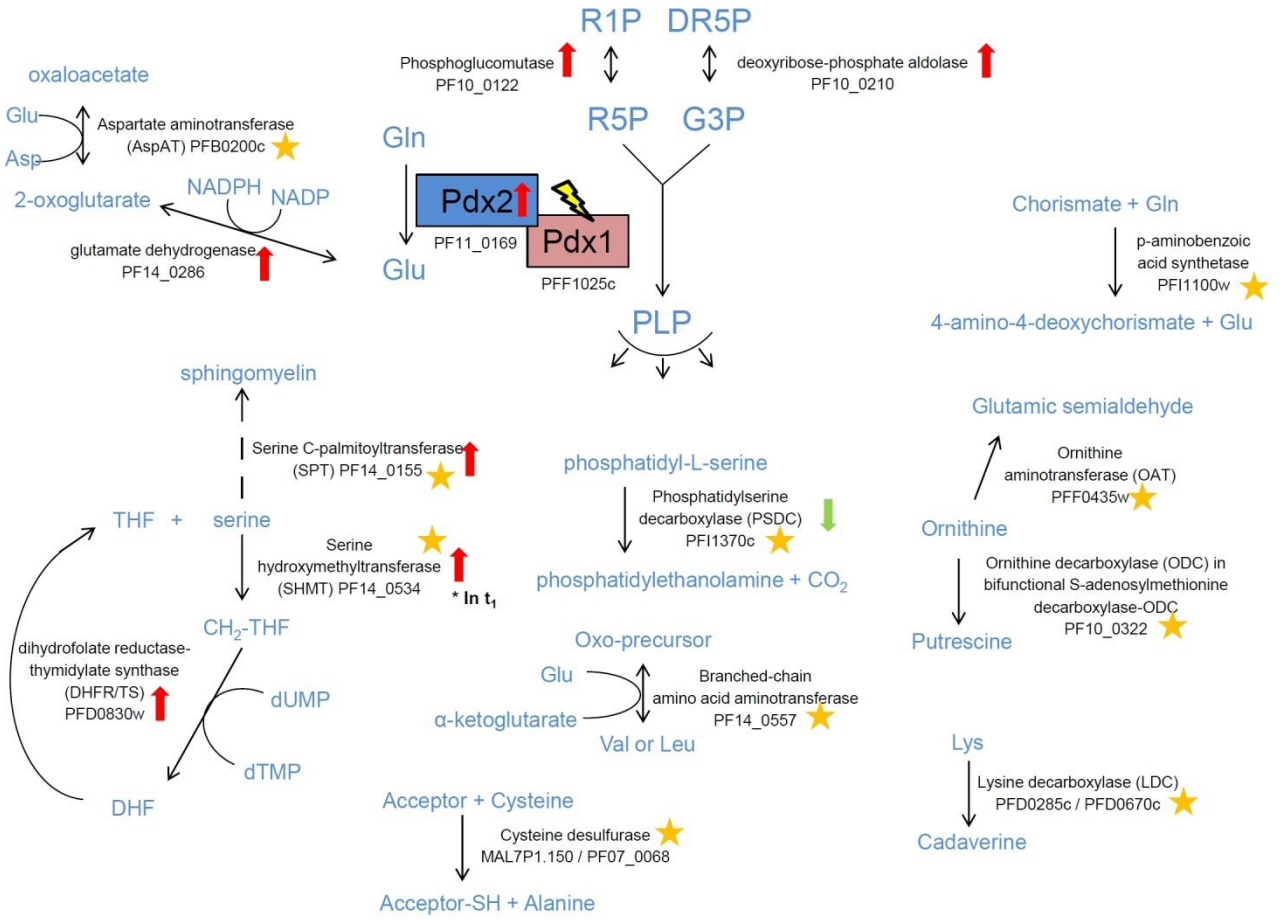


Figure 4.17: Effect of 4PEHz on PLP-dependent enzyme transcripts as well as Pdx1-related transcripts. The protein products of the transcripts indicated by a star are PLP-dependent enzymes. Arrows indicate the transcript expression during time point t_3 (except for SHMT which had increased expression in t_1).

In this study, 4PEHz was shown to affect *Pf*Pdx2 transcript levels, with no statistical significant changes in *Pf*Pdx1 transcript abundance. Several studies have shown that *P. falciparum* *Pf*Pdx1 and *Pf*Pdx2 are regulated transcriptionally, and often co-expressed [141, 242, 284]. Both *Pf*Pdx1 and *Pf*Pdx2 transcripts had increased expression in staurosporine and apicidin transcriptomes, and in the majority of time points sampled showed that these transcripts were co-regulated [242]. In response to cercosporin, parasites also exhibited increased expression of both *Pf*Pdx1 and *Pf*Pdx2 transcripts [141]. Isoleucine starvation also resulted in patterns of co-regulation of both *Pf*Pdx1 and *Pf*Pdx2 transcripts throughout the IDC and these transcripts had decreased expression compared to normal culture conditions [284]. The *Pf*Pdx1 and *Pf*Pdx2 transcripts were not significantly affected when parasites were exposed to increased temperature [235] or during hyperoxia [234], which suggests that these are not involved during general stress responses. Additional investigations into the co-regulation of *Pf*Pdx1 and *Pf*Pdx2 and the associated triggers involved in transcription expression could therefore be of interest. Interestingly, inhibition of both enzyme activities in bifunctional

*Pf*AdoMetDC/ODC resulted in increased *Pf*Pdx1 protein levels, however *Pf*Pdx2 was not detected [259]. This also supported the view that *Pf*Pdx1 could be regulated translationally.

PLP is capable of transcriptionally regulating expression of Pdx1 and Pdx2 transcripts, as well as other PLP-dependent enzymes, and this is thought to be mediated by specific transcription factors. A transcription factor termed PdxR was discovered in *Corynebacterium glutamicum* ATCC 13032 (NCBI Sequence ID: NP_600015.1) and is adjacent to Pdx1 and Pdx2 genes, *PdxS* and *PdxT*, respectively [277, 285]. Inactivation of PdxR resulted in decreased expression of these PLP synthase genes [277, 285]. DNA mobility shift assays confirmed that PdxR was capable of binding in the PdxR-Pdx1/Pdx2 intergenic region, however how PdxR senses cellular levels of PLP to affect biding could not be established [277]. In *Streptococcus pneumonia* the transcript expression of Pdx1 was decreased by exogenously supplied PLP and the presence of a PdxR-like transcription factor was shown to modulated these expression levels [286]. Another example of a transcription factor which is activated by PLP is the GabR transcription factor in *B. subtilis* [287]. GabR elegantly communicates PLP levels and affects the transcription of *GabT*, which encodes GABA-aminotransferases - a PLP-dependent enzyme [287, 288]. In the presence of PLP and GABA the GabR mediated expression of the *GabT* promoter [288]. These transcription factors, PdxR and GabR , share protein architecture and contain a short helix-turn-helix (HTH) domain, which is believed to bind DNA, with a C-terminal aminotransferase-like domain, which could mediate sensing of PLP [288]. The common HTH-fold linked to a PLP-recognising domain suggests that these transcription factors have similar functions to relay the cellular status of PLP concentrations to genes involved in PLP metabolism. BLAST analysis using PlasmoDB did not reveal any positive hits for PdxR-like or GabR transcription factors with similar sequences in *P. falciparum,* however does not exclude the possibility that a similar transcription factor might be present. *Pf*Pdx1 and *Pf*Pdx2 are situated on distant loci in *P. falciparum* (*Pf*Pdx1 on chromosome 6, *Pf*Pdx2 on chromosome 11). Further characterisation of upstream binding elements of *Pf*Pdx1 and *Pf*Pdx2 might reveal a common binding element, such as PdxR in *C. glutamicum,* which could regulate *Pf*Pdx1 and *Pf*Pdx2 expression. Additionally, uncharacterised sequences in *P. falciparum* containing putative aminotransferase or HTH domains could form part of a transcription factors which are able to recognise and relay cellular PLP levels.

The correlation between 4PEHz-treated and apicidin transcriptomes may be linked through a common transcription factor which is affected, such as an ApiAP2 family protein. ApiAP2 proteins are involved in stage-specific gene expression [263], however it is not known whether these are affected by PLP. Two independent microarray studies have also shown that apicidin growth inhibition of *P. falciparum* parasites results in increased transcription of either *Pf*Pdx1 or *Pf*Pdx2 or

both [242, 263]. The disregulation of transcription through inhibition of HDAC could affect a common transcription factor, which activates or derepresses PLP biosynthesis.

PLP biosynthesis is part of antioxidant responses in the parasite; more specifically is essential in response to ROS [141]. Interestingly, DHA resistance in *P. falciparum* parasites is associated with increased transcript expression of *Pf*Pdx2 compared to non-resistant cell lines [238]. Part of the mechanism of action of DHA is the release of singlet oxygen, which leads to formation of peroxides [289, 290]. Increased PLP levels could therefore play an essential role in resistance to DHA. This could also motivate the use of *Pf*Pdx1 or PLP synthase inhibitors in combination with DHA to prevent possible resistance development. Glutamate dehydrogenases involved in the synthesis of glutamate from 2-oxoglutarate and L-glutamine in a reaction requiring NADPH had increased expression due to 4PEHz treatment. This suggested that the transcripts associated with L-glutamine catabolism were more active, possibly to compensate for glutamine accumulation generated through lowered utilisation of the substrate by *Pf*Pdx2.

In this study transcripts of the pentose phosphate pathway (PPP) were also affected by 4PEHz. The deoxyribose-phosphate aldolase transcript which had increased expression encodes for an enzyme that reversibly converts G3P into 2-deoxy-D-ribose 5 phosphate (Figure 4.17). This could have been the cellular response towards increased G3P levels due to decreased utilisation of the substrate by *Pf*Pdx1. Another transcript of the PPP affected was phosphoglucomutase, which catalyses the interconversion of α-D-ribose 1-phosphate to R5P (Figure 4.17). This enzyme also catalyses the interconversion of α-D-glucose 1-phosphate to D-glucose 6-phosphate. Presumably disrupted R5P homeostasis could have led to increases in this transcript. During GSEA other transcripts with lower differential increased expression of the PPP were identified. A putative 6-phosphogluconate dehydrogenase (PF14_0520) and TIM are predicted to have involvement in the conversion of 6-phosphogluconate to Ru5P and isomerisation of G3P into DHAP. These perturbations in the PPP are directly linked to PLP metabolism through the substrate dependence of *Pf*Pdx1.

Other PLP-dependent enzymes affected by 4PEHz were putative serine C-palmitoyltransferases (SPT, PF14_0155) which had increased expression around 2.8-fold. The enzyme condenses serine with palmitoyl CoA to form 3-ketodihydrosphingosine, which is the precursor required to produce sphingomyelin. Sphingomyelin is believed to be one of the most prominently produced sphingolipids in *P. falciparum* [291]. SPT is a key enzyme that regulates not only sphingomyelin but also other sphingolipid levels within cells, and these are involved in signal cascades that can stimulate apoptosis or cell growth [291-293]. Moreover, previous evidence has suggested SPT are transcriptionally up-regulated in response to cellular stress including endotoxins [292]. SPT could

therefore be directly stimulated by PLP depletion, resulting in diminished levels of sphingomyelin precursors, which leads to activation or deactivation of signalling pathways. Of all the transcript with decreased expression during t_3 0.7% were involved in signalling pathways, however these were not sphingolipid-dependent, and mostly involved small GTPase mediated signal transduction factors.

Phosphatidylserine decarboxylase (PSDC, PFI1370c) had 1.9-fold decreased expression during t_3. PSDC utilise PLP to decarboxylate phosphatidyl-L-serine (PS) forming phosphatidylethanolamine (PE) and CO_2 in the process. Decreased levels of PSDC transcripts appeared to be linked to the down-regulation of a putative patatin-like phospholipase, the latter which uses PS for the production of L-2-lysophosphatidylethanolamine. Other transcripts involved in PS and PE metabolism included putative diacylglycerol kinases (PFI1485c) and putative patatin-like phospholipase (PFI1180w), and also had decreased abundance. Diacylglycerol kinases down-regulation suggested lowered levels of phosphatidate, a precursor of PS, which in turn would suggest lowered PS production. The down-regulation of a putative patatin-like phospholipase (PFI1180w) would suggest that triacylglycerol metabolism is favoured, with decreased 1,2-diacylglycerol production.

Expression patterns of mRNA transcripts are generally linked to delayed expression of the corresponding protein [294]. However, very few proteins were detected in 4PEHz treated parasites, and no definitive correlations could be made regarding linked expression patterns. Some proteins were involved in annotated processes such as DNA damage repair, which was similarly found to be active from transcriptome data. From transcriptome data glutamate dehydrogenases were shown to have increased expression which suggested L-glutamine catabolism was promoted. In the proteome the CPS protein was found in the treated parasite protein extractions and was not present in the untreated parasites. The transcriptomic expression of the corresponding transcript was not significant. CPS utilise L-glutamine together with CO_2 and ATP to generate carbamoyl phosphate [295]. Similar to *Pf*Pdx2 this enzyme is an amidotransferase, capable of providing ammonia through the hydrolysis of L-glutamine [296]. Two classes of CPS enzymes exist; CPS I and CPS II, the former is restricted to mitochondria and involved in arginine biosynthesis as well as the urea cycle. The latter CPS II has cytosolic expression and is involved in pyrimidine biosynthesis [295]. CPS I uses free ammonia instead as nitrogen source, whereas CPS II utilise L-glutamine. The CPS enzyme performs the first committed step in the urea cycle, and also provides the carboymyl phosphate precursor for synthesis of the pyrimidine UMP and arginine [297, 298]. The CPS protein identified during 4PEHz treatment represented a class II CPS which is associated with L-glutamine hydrolysis. This suggested that L-glutamine catabolic processes were active in the proteome, linking

with observations that other transcripts involved in L-glutamine catabolism were active in the transcriptome.

Drug efficacy and specificity cannot easily be assessed based on transcriptome data alone. Relatively few genes of low amplitude were expressed during CQ and the antifolate drug WR99210 (DHFR inhibitor) treatment of *P. falciparum* [240, 241]. These minor transcriptional changes were also not functionally linked to proposed drug targets [240]. In these cases the antagonists could affect very specific biologically-essential processes, thereby arresting transcriptomic response altogether, or simply there is no transcriptional response due to the drugs [240]. In stark contrast, other potent antimalarials such as artesunate resulted in 398 differentially expressed genes [262], and doxycycline induced responses were delayed however resulted in around 100 differentially expressed genes which were predominantly apicoplast-associated [207]. Similarly, artemisinin affected 270 transcripts with at least threefold change [242]. As in the case of co-inhibition of the bifunctional *S*-adenosylmethionine decarboxylase/ODC, 538 transcripts had differential expression, however transcripts directly linked to polyamine metabolism were affected, revealing a highly specific transcriptomic response [259]. This was also confirmed using proteomics and metabolomics [259]. Overall, dramatic transcriptional changes which appear unrelated may be of consequence due to initial target attenuation – possibly reflecting the cellular importance of the targeted enzyme *Pf*Pdx1. As demonstrated for 4PEHz, transcripts associated with PLP metabolism were affected, suggesting specificity of 4PEHz on *Pf*Pdx1. As a consequence additional indirectly-associated processes were affected and were suggested to be cellular compensatory responses due to disruptions in PLP biosynthesis. This supports the view that 4PEHz is specific for *Pf*Pdx1, and also highlights the dependence of the parasite on the PLP cofactor.

4.5 Conclusions

Inhibition of *Pf*Pdx1 using 4PEHz resulted in genome-wide perturbations within *P. falciparum* parasites. Several PLP-dependent enzyme transcripts had differential expression compared to untreated parasites, lending support to the proposed targeting of *Pf*Pdx1 by 4PEHz. Processes involved in folate biosynthesis and recycling were shown to be affected by 4PEHz. These results support previous observations that *Pf*Pdx1 is inhibited by 4PEHz resulting in PLP depletion, which in turn affected PLP-related processes.

Chapter 5

Concluding discussion

Malaria has plagued the world for many millennia. Efforts aimed at eradication have not been successful and at the present time more than a third of the human population still lives in malaria-stricken areas. Administration of antimalarial therapeutics and prophylaxis has been the most successful strategy, and is still the main-line defence we have against the parasites. Continual parasite drug-resistance has become a severe concern and drives the discovery and development of newer effective compounds. Previously successful antimalarials such as 4-aminoquinolines have set the precedent by exploiting endogenous parasite-specific processes such as haem detoxification. These compounds had good efficacy, were cheap and had limited host toxicity [299]. Amongst the successful antimalarials the antifolates have shown that endogenous parasite-only processes need not only be targeted in order to have a profound effect on treatment of malaria. However, in order to avoid targeting the host proteins additional pharmacokinetic optimisations are often required, and host-related metabolic perturbations may also require additional management complicating treatment. For these reasons the ideal candidates for target-based drug design remain parasite-specific processes. PLP biosynthesis is a parasite-specific process which is intricately regulated by two proteins – *Pf*Pdx1 and *Pf*Pdx2. The product of the PLP synthase is PLP - one of the most important cofactors in amino acid and polyamine metabolism as well as folate biosynthesis. The PLP synthase, or DXP-independent route of PLP biosynthesis was only recently discovered, and much regarding the regulation of this pathway within the parasite is not known. The absence of PLP biosynthetic pathways in humans gives a selective advantage for the exploitation of PLP biosynthesis in the malaria parasite.

PLP is involved in a myriad array of different metabolic processes and is a fascinating molecule. PLP can also facilitate non-enzymatic reactions [109]. Each chemical substituent on the PLP molecule is functionally essential; the phosphate group facilitates anchoring in PLP-dependent enzymes, whereas the pyridine ring facilitates delocalisation of negative charge at the C_α of reaction intermediates [300]. The principal component on PLP is the aldehyde group which forms the initial imine with embedded lysine residues in PLP-dependent proteins. Within the cellular milieu this group is relatively reactive and could be protected in order to avoid unnecessary conjugations. The Pdx1 and Pdx2 enzyme complex is therefore interesting within this regard, and questions arise as to how PLP is protected once synthesised. There is some evidence to suggest that PLP is tightly bound within Pdx1 [221]. In contrast, PLP is measured as a Tris-conjugate in Pdx1 reaction assays, and

this freely diffusible form suggests that PLP is not that tightly bound. How the PLP synthase complex sequesters and releases PLP remains to be discovered.

The objective of this study was to identify compounds that could interfere with *Pf*Pdx1 biosynthesis of PLP, and ultimately affect the proliferation of *P. falciparum* parasites. Computational approaches were performed to aid in the identification and selection of potential inhibitors of *Pf*Pdx1. A structure-based pharmacophore screening approach of *Pf*Pdx1 was followed. Initially several *Pf*Pdx1 homology models were created which shared homologous features compared to the *Pb*Pdx1 and *Tm*Pdx1 structures. The residues in the R5P-active site of *Pf*Pdx1 were shown to correspond to other structures, and were used to create structure-based *Pf*Pdx1 pharmacophores. Screening of these resulted in the identification of several ligands, as well as distinctive ligand scaffold groups. Docking was used to confirm predicted favourable binding in the R5P-active site of *Pf*Pdx1. The ligands that were identified had comparable or better predicted binding modes than R5P in *Pf*Pdx1. Some analogues of R5P, such as E4P and 4PEHz, were shown to have similar orientations in the active site of both *Pf*Pdx1 and *Pb*Pdx1.

The inhibitory capacity of *in silico*-identified ligands and other rationally selected compounds was determined on *Pf*Pdx1. Of the rationally selected compounds E4P showed some inhibitory activity. The E4P tetrose scaffold molecule had two specific chemical substituents which formed part of the molecule's pharmacophore; The C1 terminal aldehyde group, which was postulated to facilitate entry into the R5P-binding site (Figure 5.1). The exact mechanism is still speculative, and could involve initial imine formation with K83, thereby leading to disruption of R5P binding. The other pharmacophore on E4P was the phosphate moiety, which could have facilitated more favourable binding interactions within the R5P-binding site. Both these characteristics are shared by the substrates of *Pf*Pdx1 (Figure 5.1).

The limited efficacy of most of the *in silico*-identified compounds could have been related to the nature of the *Pf*Pdx1 protein. The initial requirement of Schiff base formation between R5P and K83 suggests that the aldehyde reactivity of R5P is an important requirement to enter the R5P active site [157]. The protein may therefore allow only molecules which conform to this prerequisite. Such features were also absent from the identified ligands. These are drug design considerations for future inhibitors which could be used to exploit this initial imine formation feature. There are numerous imine formation steps within the Pdx1 reaction mechanism, also recently proposed to occur during a C1 to C5 imine migration of R5P with K83 [160]. Compounds that contain chemical groups liable to imine formation such as ketones in place of hydroxyl substituents on the basic R5P scaffold may be promising inhibitors of *Pf*Pdx1. Sulphonate groups

were proposed to mimic phosphate groups of R5P, however whether these might do so *in vitro* is still questionable, as results from this study were inconclusive. A sulphonated compound with terminal aldehyde arrangement, plus the potential to interfere with imine formation, could have potential as a *Pf*Pdx1 inhibitor.

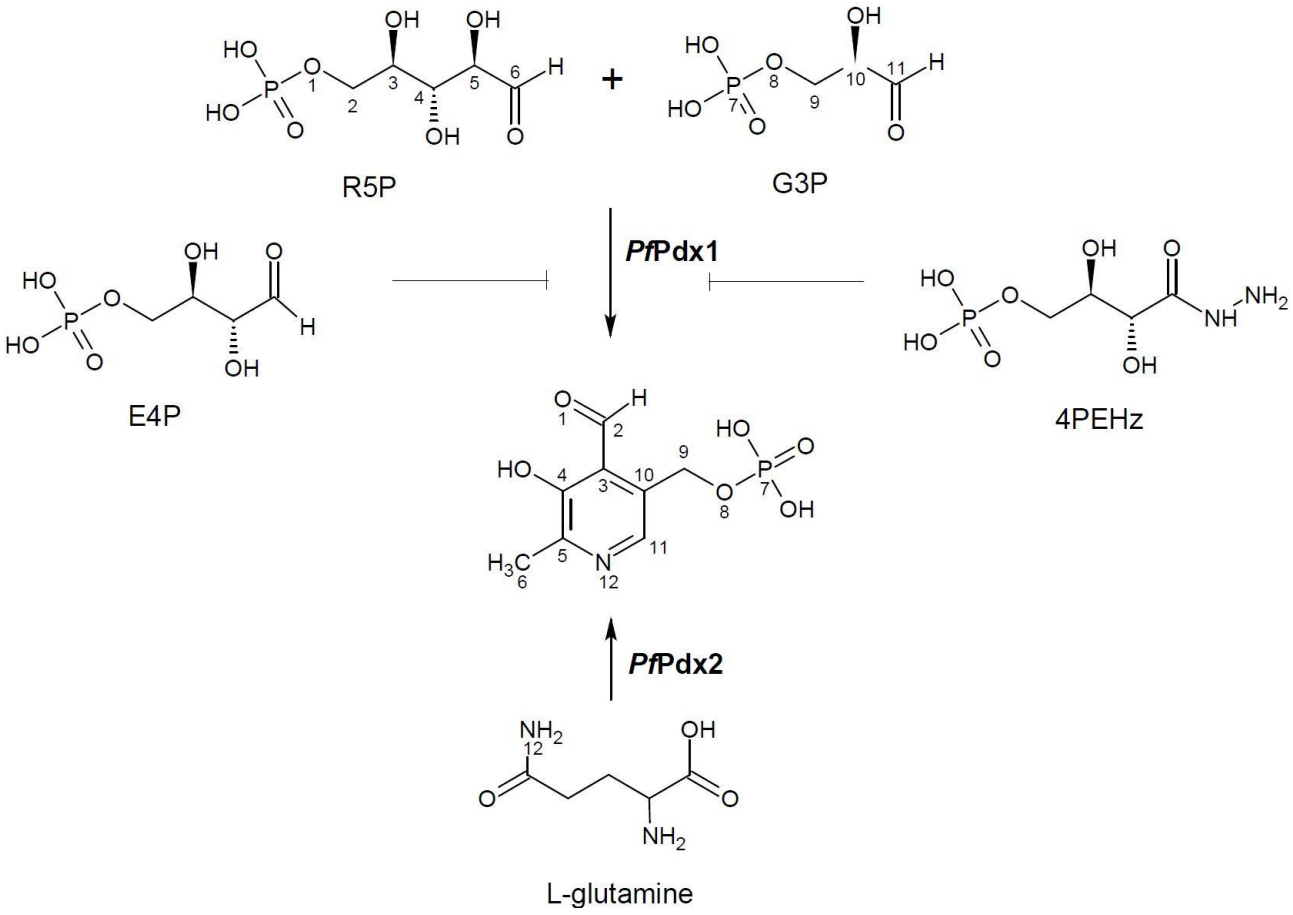


Figure 5.1: Summary of the formation of PLP by *Pf*Pdx1 and *Pf*Pdx2 with inhibitors. The PLP molecule is numbered according to the contribution of atoms from the R5P, G3P and L-glutamine substrates. E4P, R5P and G3P share chemical features such as terminal aldehyde groups and phosphate moieties. The 4PEHz inhibitor is based on the E4P scaffold with an additional hydrazide terminal group.

Analogues of E4P were selected and tested against *Pf*Pdx1. The 4PEHz compound showed appreciably inhibition of *Pf*Pdx1. Molecular docking supported the view that both E4P and 4PEHz could bind in the R5P-active site of *Pf*Pdx1. As part of the SAR the phosphate moiety of 4PEHz was established as part of the molecule's pharmacophore. This correlated with observations for E4P suggesting that this was an essential requirement for *Pf*Pdx1 inhibitors, and supported the prediction that 4PEHz occupies the same site as E4P. The increased inhibitory activity of 4PEHz when incubated with both *Pf*Pdx1 and *Pf*Pdx2 was linked to the cooperative binding of R5P in other PLP synthases and supported that 4PEHz could bind in the R5P-acitve site.

Pdx1 proteins share high degree of structural conservation, and a similar inhibitory MOA is therefore expected for E4P and 4PEHz against Pdx1 homologues in other organism. Both components of the PLP synthase are good drug targets. Inhibition of the *Pf*Pdx2 activity or inhibition of dimerization could be additional strategies to target the PLP synthase. The compounds identified here can also be utilised to manage other parasitic or bacterial infections. The causative agent of toxoplasmosis - *Toxoplasma gondii* – possess a functional *de novo* vitamin B_6 pathway which could similarly be exploited using E4P and 4PEHz [187]. Moreover *Mycobacterium tuberculosis* is solely reliant on the DXP-independent pathway for survival and knockout experiments have already shown that attenuation of the Pdx1 is a viable strategy to combat tuberculosis [301]. The importance of vitamin B_6 biosynthesis in plants could be exploited by using Pdx1 inhibitors as alternative herbicides, and considering the antioxidant properties of PLP which aids during abiotic stress, these in conjunction with other herbicides are expected to have a synergistic effect in management of unwanted crops [302].

The proliferation of *P. falciparum* parasites was analysed during treatment with the *Pf*Pdx1 inhibitor 4PEHz. The compound was detrimental to the proliferation of *P. falciparum* parasites *in vitro*. In order to establish whether inhibition was attributed to targeting of *Pf*Pdx1 and subsequently PLP biosynthesis, a cell line overexpressing *Pf*Pdx1 and *Pf*Pdx2 was tested with 4PEHz. These cells were significantly less susceptible to 4PEHz, and the parasite growth was not significantly perturbed compared to cells lacking additional copies of *Pf*Pdx1 and *Pf*Pdx2. These results suggested that 4PEHz had some specificity and was capable of targeting *Pf*Pdx1. To determine the global functional consequences of inhibiting *Pf*Pdx1 the transcriptome and proteome of the parasites treated with 4PEHz was analysed. Transcriptomic responses suggested that 4PEHz caused PLP-specific perturbations, however unrelated functional pathways were also affected. Attenuation of PLP biosynthesis could have caused major disruptions attributed to the ubiquitous involvement of PLP in cellular metabolism. Disruption of folate metabolism, which results in DNA lesion, has previously been linked to PLP depletion [274]. Corroborating with this, transcripts of enzymes involved in folate metabolism and recycling, such as SHMT, were found to be systematically affected during treatment of parasites with 4PEHz. This suggested that PLP biosynthesis was disrupted and also resulted in damage of DNA. In particular the *Pf*Pdx2 transcript expression was increased in 4PEHz-treated parasites. *Pf*Pdx2 has no other function apart from forming an obligate heterodimer with *Pf*Pdx1. Through guilt by association PLP biosynthesis was directly affected by 4PEHz suggesting that the compound had specificity for *Pf*Pdx1. Results from proteomic perturbations supported that L-glutamine catabolism was favoured, possibly due to decreased utilisation of this substrate by *Pf*Pdx2.

Whether PLP biosynthesis in *P. falciparum* is regulated transcriptionally, translationally, post-translationally is not known. Several investigations have shown that the transcripts of *Pf*Pdx1 and *Pf*Pdx2 are transcriptionally modulated, however were not always associated with general stress responses. This study supported the transcriptional regulation of *Pf*Pdx2, which was linked to inhibition of *Pf*Pdx1. Evidence from this study implicated that the PLP synthase activity could be regulated through inhibition by endogenous E4P. Both E4P, as well as D-xylulose 5-phosphate, is formed during the condensation of F6P and G3P by transketolase which are present in the parasites [216]. Transketolases play an essential role in the non-oxidative PPP branch by generating R5P from G3P and F6P via D-xylulose 5-phosphate [216]. D-xylulose 5-phosphate can be reversibly converted to Ru5P [215, 303], and in *P. falciparum* a putative Ru5P 3-epimerase (PFL0960w) has been implicated in this process [304]. The Ru5P is then isomerised to R5P by putative R5P epimerase (PFE0730c) [215]. Interestingly, the non-oxidative PPP pathway accounts for more than 80% of the carbon flux derived from [1-^{14}C]glucose which incorporated into nucleic acids [305, 306]. Moreover, the transcripts of enzymes involved in the non-oxidative PPP branch are co-ordinately expressed during the IDC of the parasites, highlighting the essential functional activity of this pathway [215]. Considering that E4P is the indirect by-product of the formation of R5P, appreciable levels of E4P are expected in the parasites, additionally also diverted into the shikimate pathway for the production of chorismate [216]. E4P metabolite levels in *P. falciparum* have not yet been established, nor have detectable levels been found in erythrocytes, and could be of interest for future investigations [307]. E4P could have a weak regulatory role on PLP biosynthesis, however other molecules or stereoisomers of E4P such as D-threose could also affect *Pf*Pdx1, and remains to be determined. Other stereoisomers of R5P which possess similar reactive groups, such as the product of the non-oxidative PPP arm - D-xylulose 5-phosphate - could also inhibit *Pf*Pdx1. Other means of regulating PLP biosynthesis includes the possibility of allosteric sites on the PLP synthase.

Based on results presented 4PEHz is a good therapeutic candidate, however additional chemical optimisation will be required to increase the potency of the compound before this compound is suited for clinical or *in vivo* investigations. Additional analogues of the lead E4P compound could also be created to further explore variations in the basic scaffold. The chemical reactivity of both the lead compound E4P and 4PEHz are undesirable properties, however the nature of the interaction with *Pf*Pdx1 was shown to require these functional groups. Both the substrates of *Pf*Pdx1 are also considered to be chemically reactive, and additional considerations should be made during lead optimisations.

Contributions of this study showed that the PLP biosynthetic pathway of the parasite is chemically tractable or druggable. Some questions remains as the influence of B_6 vitamer salvage which could diminish the effects of *Pf*Pdx1 inhibitors. Currently salvage is believed to minimally contribute towards PLP metabolism; the proliferation of parasites in PN-free culture medium is not affected [108]. Additionally, the efficacy of 4PEHz was tested in conditions which should promote B_6 vitamer salvage (5 µM PN in the culture media), and the erythrocyte is considered a rich source of other B_6 vitamers [125, 127]. Unphosphorylated B_6 vitamers can enter the parasite by diffusion, however the final phosphorylation reaction catalysed by PdxK requires ATP. The *de novo* biosynthesis of PLP utilises substrates from glycolysis and the PPP pathways, moreover is ATP-independent. This supports the notion that B_6 vitamer salvage is not the predominant route in order to obtain PLP. This also underscores the importance of the *de novo* PLP biosynthetic route in the parasites. Inhibition of this key pathway, in combination with other antimalarials, could be a feasible clinical combination to treat and eliminate *P. falciparum* infections. Overall results from this thesis support the hypothesis that inhibition of *Pf*Pdx1 in malaria parasites is detrimental and leads to perturbations in PLP-dependent processes. This study has contributed new knowledge regarding inhibitors that target *Pf*Pdx1 and the effects of inhibiting PLP biosynthesis in *P. falciparum* parasites.

References

1. Sachs, J. and Malaney, P. (2002), The economic and social burden of malaria. *Nature***. 415**, p. 680-685
2. Younis, Y., Douelle, F., Feng, T.-S., *et al.* (2012), 3,5-Diaryl-2-aminopyridines as a Novel Class of Orally Active Antimalarials Demonstrating Single Dose Cure in Mice and Clinical Candidate Potential. *Journal of Medicinal Chemistry***. 55**, p. 3479-3487
3. Alonso, P. L., Brown, G., Arevalo-Herrera, M., Binka, F., Chitnis, C., Collins, F., Doumbo, O. K., Greenwood, B., Hall, B. F., Levine, M. M., Mendis, K., Newman, R. D., Plowe, C. V., Rodríguez, M. H., Sinden, R., Slutsker, L. and Tanner, M. (2011), A Research Agenda to Underpin Malaria Eradication. *PLoS Med***. 8**, p. e1000406
4. Luft, F. C. (2004), Liver stage antigen and malaria. *Journal of molecular medicine***. 82**, p. 555-557
5. Gardiner, D., McCarthy, J. and Trenholme, K. (2005), Malaria in the post-genomics era: light at the end of the tunnel or just another train? *Postgraduate Medical Journal***. 81**, p. 505
6. Nerlich, A. G., Schraut, B., Dittrich, S., Jelinek, T. and Zink, A. R. (2008), *Plasmodium falciparum* in ancient Egypt. *Emerging Infectious Diseases***. 14**, p. 1317
7. Smith, D. C. and Sanford, L. B. (1985), Laveran's germ: the reception and use of a medical discovery. *The American journal of tropical medicine and hygiene***. 34**, p. 2-20
8. Hagan, P. and Chauhan, V. (1997), Ronald Ross and the problem of malaria. *Parasitology Today***. 13**, p. 290-295
9. Shortt, H. E., Fairley, N. H., Covell, G., Shute, P. G. and Garnham, P. C. C. (1949), Pre-erythrocytic Stage of *Plasmodium falciparum*. *British Medical Journal***. 2**, p. 1006-1008
10. Shortt, H. E. and Garnham, P. C. C. (1948), Persisting exo-erythrocytic cycle in *Plasmodium cynomolgi*. *British Medical Journal***. 1**, p. 1225-1228
11. Liu, W., Li, Y., Learn, G. H., Rudicell, R. S., Robertson, J. D., Keele, B. F., Ndjango, J. B. N., Sanz, C. M., Morgan, D. B. and Locatelli, S. (2010), Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. *Nature***. 467**, p. 420-425
12. Baron, J. M., Higgins, J. M. and Dzik, W. H. (2011), A Revised Timeline for the Origin of *Plasmodium falciparum* as a Human Pathogen. *Journal of Molecular Evolution*, p. 1-8
13. Mota, M. M. and Rodriguez, A. (2001), Migration through host cells by apicomplexan parasites. *Microbes and Infection***. 3**, p. 1123-1128
14. Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M. and Snow, R. W. (2004), The global distribution and population at risk of Malaria: Past, Present, and Future. *The Lancet, Infectious Diseases***. 4**, p. 327-336
15. Ittarat, W., Pickard, A. L., Rattanasinganchan, P., Wilairatana, P., Looareesuwan, S., Emery, K., Low, J., Udomsangpetch, R. and Meshnick, S. R. (2003), Recrudescence in artesunate-treated patients with falciparum malaria is dependent on parasite burden not on parasite factors. *Am J Trop Med Hyg***. 68**, p. 147-152
16. Bronner, U., Divis, P. C. S., Färnert, A. and Singh, B. (2009), Swedish traveller with *Plasmodium knowlesi* malaria after visiting Malaysian Borneo. *Malaria Journal***. 8**, p. 15
17. Collins, W. E. and Jeffery, G. M. (2007), *Plasmodium malariae*: parasite and disease. *Clinical Microbiology Reviews***. 20**, p. 579-592
18. Mueller, I., Zimmerman, P. A. and Reeder, J. C. (2007), Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. *Trends in Parasitology***. 23**, p. 278-283
19. Guinovart, C., Navia, M., Tanner, M. and Alonso, P. (2006), Malaria: burden of disease. *Current Molecular Medicine***. 6**, p. 137-140
20. Walther H, W. (2012), Global challenges of changing epidemiological patterns of malaria. *Acta Tropica***. 121**, p. 158-165

21. Baird, J. K. (2009), Resistance to therapies for infection by Plasmodium vivax. *Clin Microbiol Rev***. 22**, p. 508-534
22. White, N. J. and Imwong, M. (2012) Chapter Two - Relapse. In Advances in Parasitology (S.I. Hay, R. P. and Baird, J. K., eds.). pp. 113-150, Academic Press
23. White, N. (2011), Determinants of relapse periodicity in Plasmodium vivax malaria. *Malaria Journal***. 10**, p. 297
24. Tjitra, E., Anstey, N. M., Sugiarto, P., Warikar, N., Kenangalem, E., Karyana, M., Lampah, D. A. and Price, R. N. (2008), Multidrug-resistant *Plasmodium vivax* associated with severe and fatal malaria: a prospective study in Papua, Indonesia. *PLoS Medicine***. 5**, p. e128
25. Besansky, N. J., Hill, C. A. and Costantini, C. (2004), No accounting for taste: host preference in malaria vectors. *Trends in Parasitology***. 20**, p. 249-251
26. Koella, J. C. (1999), An evolutionary view of interactions between anopheline mosquitoes and malaria parasites *Microbes and Infection***. 1**, p. 303-308
27. Carey, A. F., Wang, G., Su, C. Y., Zwiebel, L. J. and Carlson, J. R. (2010), Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature***. 464**, p. 66-71
28. Verhulst, N. O., Qiu, Y. T., Beijleveld, H., Maliepaard, C., Knights, D., Schulz, S., Berg-Lyons, D., Lauber, C. L., Verduijn, W. and Haasnoot, G. W. (2011), Composition of Human Skin Microbiota Affects Attractiveness to Malaria Mosquitoes. *PLoS One***. 6**, p. e28991
29. Cowman, A. F. and Crabb, B. S. (2006), Invasion of red blood cells by malaria parasites. *Cell***. 124**, p. 755-766
30. Barnwell, J. W. (2001), Hepatic Kupffer Cells: The portal that permits infection of hepatocytes by malarial sporozoites? *Hepatology***. 33**, p. 1331-1333
31. Centre for Disease Control, http://www.dpd.cdc.gov/dpdx/hTML/ImageLibrary/Malaria_il.htm, 2009-02-25
32. Haldar, K. and Mohandas, N. (2007), Erythrocyte remodeling by malaria parasites. *Current opinion in hematology***. 14**, p. 203
33. Tilley, L., Dixon, M. W. A. and Kirk, K. (2011), The *Plasmodium falciparum*-infected red blood cell. *The International Journal of Biochemistry & Cell Biology***. 43**, p. 839-842
34. Baker, D. A. (2010), Malaria gametocytogenesis. *Molecular and biochemical parasitology***. 172**, p. 57-65
35. Baton, L. A. and Ranford-Cartwright, L. C. (2005), Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito. *Trends in Parasitology***. 21**, p. 573-580
36. Clark, I. A. and Cowden, W. B. (2003), The pathophysiology of falciparum malaria. *Pharmacology & Therapeutics***. 99**, p. 221-260
37. Mackintosh, C. L., Beeson, J. G. and Marsh, K. (2004), Clinical features and pathogenesis of severe malaria. *Trends in Parasitology***. 20**, p. 597-603
38. Bloland, P. B. (2001), Drug resistance in malaria (WHO/CDS/CSR/DRS/2001.4). *http://whqlibdoc.who.int/hq/2001/WHO_CDS_CSR_DRS_2001.4.pdf*, *World Health Organization*, p. pg. 27
39. World Malaria Report 2011 http://www.who.int/entity/malaria/world_malaria_report_2011/9789241564403_eng.pdf:9789241564403 World Health Organization, Geneva, Switzerland.
40. López, C., Saravia, C., Gomez, A., Hoebeke, J. and Patarroyo, M. A. (2010), Mechanisms of genetically-based resistance to malaria. *Gene***. 467**, p. 1-12
41. Murray, C. J. L., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., Fullman, N., Naghavi, M., Lozano, R. and Lopez, A. D. (2012), Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet***. 379**, p. 413-431
42. Gething, P., Patil, A., Smith, D., Guerra, C., Elyazar, I., Johnston, G., Tatem, A. and Hay, S. (2011), A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malaria Journal***. 10**, p. 378

43. Tsuji, M. and Zavala, F. (2001), Peptide-based subunit vaccines against pre-erythrocytic stages of malaria parasites. *Molecular Immunology***. 38**, p. 433-442
44. Greenwood, B. M., Fidock, D. A., Kyle, D. E., Kappe, S. H. I., Alonso, P. L., Collins, F. H. and Duffy, P. E. (2008), Malaria: progress, perils, and prospects for eradication. *The Journal of Clinical Investigation***. 118**, p. 1266-1276
45. Sharma, S. and Sulabha, S. (2008), Malaria vaccine: a current perspective. *Journal of Vector Borne Diseases***. 45**, p. 1-20
46. Francis, S. E., Sullivan, D. J. J. and Goldberg, D. E. (1997), Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annual Reviews in Microbiology***. 51**, p. 97-123
47. Cohen, J., Nussenzweig, V., Nussenzweig, R., Vekemans, J. and Leach, A. (2010), From the circumsporozoite protein to the RTS, S/AS candidate vaccine. *Human vaccines***. 6**, p. 90
48. Ménard, R., Heussler, V., Yuda, M. and Nussenzweig, V. (2008), *Plasmodium* pre-erythrocytic stages: what's new? *Trends in Parasitology***. 24**, p. 564-569
49. Vekemans, J., Leach, A. and Cohen, J. (2009), Development of the RTS, S/AS malaria candidate vaccine. *Vaccine***. 27**, p. G67-G71
50. Vogel, G. (2012), Disappointing Results Blunt Hopes for Malaria Vaccine. *Science***. 338**, p. 871-872
51. Agnandji, S. T., Lell, B., Fernandes, J. F., Abossolo, B. P., Methogo, B. G. N. O., Kabwende, A. L., Adegnika, A. A. and Mordmüller, B. (2012), A Phase 3 Trial of RTS,S/AS01 Malaria Vaccine in African Infants. *New England Journal of Medicine***. 367**, p. 2284-2295
52. Bongfen, S. E., Ntsama, P. M., Offner, S., Smith, T., Felger, I., Tanner, M., Alonso, P., Nebie, I., Romero, J. F., Silvie, O., Torgler, R. and Corradin, G. (2009), The N-terminal domain of *Plasmodium falciparum* circumsporozoite protein represents a target of protective immunity. *Vaccine***. 27**, p. 328-335
53. Carter, R. (2001), Transmission blocking malaria vaccines. *Vaccine***. 19**, p. 2309-2314
54. Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z. and Corbel, V. (2011), Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control? *Trends in Parasitology***. 27**, p. 91-98
55. World Health Organization, Malaria vector control and personal protection: report of a WHO study group. *WHO technical report series; No 936 - ISBN 92 4 120936 4 (ISBN 978 92 4 120936 6)*, p.
56. Okumu, F. O. and Moore, S. J. (2011), Combining indoor residual spraying and insecticide-treated nets for malaria control in Africa: a review of possible outcomes and an outline of suggestions for the future. *Malaria Journal***. 10**, p. 208
57. Dabire, R., Diabate, A., Baldet, T., Pare-Toe, L., Guiguemde, R., Ouedraogo, J.-B. and Skovmand, O. (2006), Personal protection of long lasting insecticide-treated nets in areas of Anopheles gambiae s.s. resistance to pyrethroids. *Malaria Journal***. 5**, p. 12
58. Turusov, V., Rakitsky, V. and Tomatis, L. (2002), Dichlorodiphenyltrichloroethane (DDT): ubiquity, persistence, and risks. *Environmental Health Perspectives***. 110**, p. 125
59. Walker, K. (2000), Cost-comparison of DDT and alternative insecticides for malaria control. *Medical and Veterinary Entomology***. 14**, p. 345-354
60. Sadasivaiah, S., Tozan, Y. and Breman, J. G. (2007), Dichlorodiphenyltrichloroethane (DDT) for indoor residual spraying in Africa: how can it be used for malaria control? *The American journal of tropical medicine and hygiene***. 77**, p. 249-263
61. Wong, M., Leung, A., Chan, J. and Choi, M. (2005), A review on the usage of POP pesticides in China, with emphasis on DDT loadings in human milk. *Chemosphere***. 60**, p. 740-752
62. Rogan, W. J. and Chen, A. (2005), Health risks and benefits of bis (4-chlorophenyl)-1, 1, 1-trichloroethane (DDT). *The Lancet***. 366**, p. 763-773

63. Hargreaves, K., Hunt, R. H., Brooke, B. D., Mthembu, J., Weeto, M. M., Awolola, T. S. and Coetzee, M. (2003), Anopheles arabiensis and An. quadriannulatus resistance to DDT in South Africa. *Medical and Veterinary Entomology***. 17**, p. 417-422
64. Takken, W. (2002), Do insecticide-treated bednets have an effect on malaria vectors? *Tropical Medicine and International Health***. 7**, p. 1022-1030
65. Maharaj, R., Mthembu, D. and Sharp, B. (2008), Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal. *South African Medical Journal***. 95**, p. 871
66. Hyde, J. E. (2005), Exploring the folate pathway in *Plasmodium falciparum*. *Acta Tropica***. 95**, p. 191-206
67. Petersen, I., Eastman, R. and Lanzer, M. (2011), Drug-resistant malaria: molecular mechanisms and implications for public health. *FEBS Letters*, p.
68. Hyde, J. E. and Müller, I. B. (2010), Antimalarial drugs: modes of action and mechanisms of parasite resistance. *Future Microbiology***. 5**, p. 1857-1873
69. Le Bras, J. and Durand, R. (2003), The mechanisms of resistance to antimalarial drugs in *Plasmodium falciparum*. *Fundamental & Clinical Pharmacology***. 17**, p. 147-153
70. Sanchez, C. P., Mayer, S., Nurhasanah, A., Stein, W. D. and Lanzer, M. (2011), Genetic linkage analyses redefine the roles of PfCRT and PfMDR1 in drug accumulation and susceptibility in *Plasmodium falciparum*. *Molecular Microbiology***. 82**, p. 865-878
71. Hill, D. R., Baird, J. K., Parise, M. E., Lewis, L. S., Ryan, E. T. and Magill, A. J. (2006), Primaquine: report from CDC expert meeting on malaria chemoprophylaxis I. *The American journal of tropical medicine and hygiene***. 75**, p. 402-415
72. Vale, N., Moreira, R. and Gomes, P. (2009), Primaquine revisited six decades after its discovery. *European Journal of Medicinal Chemistry***. 44**, p. 937-953
73. Shekalaghe, S., Drakeley, C., Gosling, R., Ndaro, A., van Meegeren, M., Enevold, A., Alifrangis, M., Mosha, F., Sauerwein, R. and Bousema, T. (2007), Primaquine Clears Submicroscopic *Plasmodium falciparum* Gametocytes that Persist after Treatment with Sulphadoxine-Pyrimethamine and Artesunate. *PLoS ONE***. 2**, p. e1023
74. Gosling, R. D., Okell, L., Mosha, J. and Chandramohan, D. (2011), The role of antimalarial treatment in the elimination of malaria. *Clinical Microbiology and Infection*, p.
75. White, N. J. Primaquine to prevent transmission of falciparum malaria. *The Lancet Infectious Diseases*, p.
76. White, N. (2008), The role of anti-malarial drugs in eliminating malaria. *Malaria Journal***. 7**, p. S8
77. Hyde, J. E. (2005), Drug-resistant malaria. *Trends in Parasitology***. 21**, p. 494-498
78. Olliaro, P. (2001), Mode of action and mechanisms of resistance for antimalarial drugs. *Pharmacology & Therapeutics***. 89**, p. 207-219
79. Sibley, C. H., Hyde, J. E., Sims, P. F. G., Plowe, C. V., Kublin, J. G., Mberu, E. K., Cowman, A. F., Winstanley, P. A., Watkins, W. M. and Nzila, A. M. (2001), Pyrimethamine–sulfadoxine resistance in *Plasmodium falciparum*: what next? *Trends in Parasitology***. 17**, p. 570-571
80. Wongsrichanalai, C., Pickard, A. L., Wernsdorfer, W. H. and Meshnick, S. R. (2002), Epidemiology of drug-resistant malaria. *Infectious Diseases***. 2**, p. 209-218
81. Yuthavong, Y., Tarnchompoo, B., Vilaivan, T., Chitnumsub, P., Kamchonwongpaisan, S., Charman, S. A., McLennan, D. N., White, K. L., Vivas, L., Bongard, E., Thongphanchang, C., Taweechai, S., Vanichtanankul, J., Rattanajak, R., Arwon, U., Fantauzzi, P., Yuvaniyama, J., Charman, W. N. and Matthews, D. (2012), Malarial dihydrofolate reductase as a paradigm for drug development against a resistance-compromised target. *Proceedings of the National Academy of Sciences***. 109**, p. 16823-16828
82. Wurtz, N., Pascual, A., Marin-Jauffre, A., Bouchiba, H., Benoit, N., Desbordes, M., Martelloni, M., de Santi, V., Richa, G., Taudon, N., Pradines, B. and Briolant, S. (2012), Early treatment failure during treatment of *Plasmodium falciparum* malaria with atovaquone-proguanil in the Republic of Ivory Coast. *Malaria Journal***. 11**, p. 146

83. Nosten, F., McGready, R., dAlessandro, U., Bonell, A., Verhoeff, F., Menendez, C., Mutabingwa, T. and Brabin, B. (2006), Antimalarial drugs in pregnancy: a review. *Current drug safety***. 1**, p. 1-15
84. Winstanley, P. and Ward, S. (2006) Malaria Chemotherapy. In Advances in Parasitology (David, H. M., ed.). pp. 47-76, Academic Press
85. Tinto, H., Guekoun, L., Zongo, I., Guiguemdé, R. T., D'Alessandro, U. and Ouédraogo, J. B. (2008), Chloroquine-resistance molecular markers (Pfcrt T76 and Pfmdr-1 Y86) and amodiaquine resistance in Burkina Faso. *Tropical Medicine & International Health***. 13**, p. 238-240
86. Vangapandu, S., Jain, M., Kaur, K., Patil, P., Patel, S. R. and Jain, R. (2007), Recent advances in antimalarial drug development. *Medicinal Research Reviews***. 27**, p. 65-107
87. Ding, X. C., Beck, H. P. and Raso, G. (2010), Plasmodium sensitivity to artemisinins: magic bullets hit elusive targets. *Trends in Parasitology*, p.
88. Haynes, R. K. and Krishna, S. (2004), Artemisinins: activities and actions. *Microbes and Infection***. 6**, p. 1339-1346
89. Briolant, S., Wurtz, N., Zettor, A., Rogier, C. and Pradines, B. (2010), Susceptibility of *Plasmodium falciparum* isolates to doxycycline is associated with pftetQ sequence polymorphisms and pftetQ and pfmdt copy numbers. *Journal of Infectious Diseases***. 201**, p. 153-159
90. Price, R. N., Nosten, F., Luxemburger, C., ter Kuile, F. O., Paiphun, L., Chongsuphajaisiddhi, T. and White, N. J. (1996), Effects of artemisinin derivatives on malaria transmissibility. *The Lancet***. 347**, p. 1654-1658
91. White, N. J. (1997), Assessment of the pharmacodynamic properties of antimalarial drugs *in vivo*. *Antimicrobial Agents and Chemotherapy***. 41**, p. 1413-1422
92. Bharati, A., Kar, M. and Sabat, S. C. (2012), Artemisinin Inhibits Chloroplast Electron Transport Activity: Mode of Action. *PLoS ONE***. 7**, p. e38942
93. Cui, L., Wang, Z., Jiang, H., Parker, D., Wang, H. and Su, X. Z. (2012), Lack of Association of the S769N Mutation in *Plasmodium falciparum* SERCA (PfATP6) with Resistance to Artemisinins. *Antimicrobial Agents and Chemotherapy***. 56**, p. 2546-2552
94. Dondorp, A. M., Yeung, S., White, L., Nguon, C., Day, N. P. J., Socheat, D. and von Seidlein, L. (2010), Artemisinin resistance: current status and scenarios for containment. *Nat Rev Micro***. 8**, p. 272-280
95. Bassat, Q., Mulenga, M., Tinto, H., Piola, P., Borrmann, S., Menéndez, C., Nambozi, M., Valéa, I., Nabasumba, C., Sasi, P., Bacchieri, A., Corsi, M., Ubben, D., Talisuna, A. and D'Alessandro, U. (2009), Dihydroartemisinin-Piperaquine and Artemether-Lumefantrine for Treating Uncomplicated Malaria in African Children: A Randomised, Non-Inferiority Trial. *PLoS ONE***. 4**, p. e7871
96. Hughes, J. P., Rees, S., Kalindjian, S. B. and Philpott, K. L. (2011), Principles of early drug discovery. *British Journal of Pharmacology***. 162**, p. 1239-1249
97. Frearson, J. A., Wyatt, P. G., Gilbert, I. H. and Fairlamb, A. H. (2007), Target assessment for antiparasitic drug discovery. *Trends in Parasitology***. 23**, p. 589-595
98. Sun, H. (2008), Pharmacophore-Based Virtual Screening. *Current Medicinal Chemistry***. 15**, p. 1018-1024
99. Guiguemde, W. A., Shelat, A. A., Garcia-Bustos, J. F., Diagana, T. T., Gamo, F. J. and Guy, R. K. (2012), Global Phenotypic Screening for Antimalarials. *Chemistry & Biology***. 19**, p. 116-129
100. Smith, C. (2003), Drug target validation: Hitting the target. *Nature***. 422**, p. 341-347
101. Djebali, S., Davis, C. A., Merkel, A., *et al.* (2012), Landscape of transcription in human cells. *Nature***. 489**, p. 101-108
102. Dharia, N. V., Chatterjee, A. and Winzeler, E. A. (2010), Genomics and systems biology in malaria drug discovery. *Current opinion in investigational drugs (London, England: 2000)***. 11**, p. 131

103. Garrido-Franco, M., Laber, B., Huber, R. and Clausen, T. (2001), Structural basis for the function of pyridoxine 5'-phosphate synthase. *Structure***. 9**, p. 245-253
104. Gengenbacher, M., Fitzpatrick, T. B., Raschle, T., Flicker, K., Sinning, I., Müller, S., Macheroux, P., Tews, I. and Kappes, B. (2006), Vitamin B6 biosynthesis by the malaria parasite *Plasmodium falciparum*. *Journal of Biological Chemistry***. 281**, p. 3633-3641
105. Müller, I. B., Wu, F., Bergmann , B., Knöckel, J., Gehring, H., Walter, R. D. and Wrenger, C. (2009), Poisoning pyridoxal 5-phosphate dependent enzymes: a new strategy to target the malaria parasite *Plasmodium falciparum*. *PLoS ONE***. 4**, p. 1-9
106. Fitzpatrick, T. B., Amrhein, N., Kappes, B., Macheroux, P., Tews, I. and Raschle, T. (2007), Two independent routes of *de novo* vitamin B_6 biosynthesis: not that different after all. *Biochemical Journal***. 407**, p. 1-13
107. Mooney, S., Leuendorf, J., Hendrickson, C. and Hellman, H. (2009), Vitamin B6: A long known compound of surprising complexity. *Molecules***. 14**, p. 329-351
108. Müller, I. B., Hyde, J. E. and Wrenger, C. (2010), Vitamin B metabolism in *Plasmodium falciparum* as a source of drug targets. *Trends in Parasitology***. 26**, p. 35-43
109. Toney, M. D. (2005), Reaction specificity in pyridoxal phosphate enzymes. *Archives of Biochemistry and Biophysics***. 433**, p. 279-287
110. Garrido-Franco, M., Laber, B., Huber, R. and Clausen, T. (2002), Enzyme-ligand complexes of pyridoxine 5'-phosphate synthase: Implications for substrate binding and catalysis. *Journal of Molecular Biology***. 321**, p. 601-612
111. Midttun, Ø., Hustad, S., Solheim, E., Schneede, J. and Ueland, P. M. (2005), Multianalyte quantification of vitamin B_6 and B_2 species in the nanomolar range in human plasma by liquid chromatography - tandem mass spectrometry. *Clinical Chemistry***. 51**, p. 1206-1216
112. Drewke, C. and Leistner, E. (2001), Biosynthesis of vitamin B_6 and structurally related derivatives. *Vitamins and Hormones***. 61**, p. 121-155
113. Hoffmann-Ostenhof, O., Cohn, W. E., Braunstein, A. E. and Karlson, P. (1970), IUPAC-IUB Commission on Biochemical Nomenclature - Nomenclature for vitamins B6 and related compounds. Tentative rules. *Biochemistry***. 9**, p. 4019-4021
114. IUPAC-IUB (1973), IUPAC-IUB Commission on Biochemical Nomenclature - Definitive nomeclature for vitamins B-6 and related compounds. *Pure and Applied Chemistry***. 33**, p. 445-452
115. Komatsu, S., Yanaka, N., Matsubara, K. and Kato, N. (2003), Antitumor effect of vitamin B_6 and its mechanims. *Biochemica et Biophysica Acta***. 1647**, p. 127-130
116. Ehrenshaft, M., Jenns, A. E., Chung, K. R. and Daub, M. E. (1998), *SOR1*, a gene required for photosensitizer and singlet oxygen resistance in *Cercospora* fungi, is highly conserved in divergent organisms. . *Molecular Cell***. 1**, p. 603-609
117. Ehrenshaft, M., Bilski, P., Li, M. Y., Chignell, C. F. and Daub, M. E. (1999), A highly conserved sequence is a novel gene involved in *de novo* vitamin B6 biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America***. 96**, p. 9374-9378
118. Ehrenshaft, M. and Daub, M. E. (2001), Isolation of PDX2, a second novel gene in the pyridoxine biosynthesis pathway of eukaryotes, archaebacteria, and a subset of eubacteria. *Journal of Bacteriology***. 183**, p. 3383-3390
119. Zhu, J., Burgner, J. W., Harms, E., Belitsky, B. R. and Smith, J. L. (2005), A new arrangment of $(\beta/\alpha)_8$ Barrels in the synthase subunit of PLP synthase. *Journal of Biological Chemistry***. 280**, p. 27914-27923
120. Osmani, A. H., May, G. S. and Osmani, S. A. (1999), The extremely conserved *pyroA* gene of *Aspergillus nidulans* is required for pyridoxine synthesis and is required indirectly for resistance to photosensitizers. *Journal of Biological Chemistry***. 274**, p. 23565-23569
121. Mukherjee, T., Hanes, J., Tews, I., Ealick, S. E. and Begley, T. P. (2011), Pyridoxal phosphate: Biosynthesis and catabolism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics***. 1814**, p. 1585-1596

122. di Salvo, M. L., Contestabile, R. and Safo, M. K. (2011), Vitamin B6 salvage enzymes: Mechanism, structure and regulation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics***. 1814**, p. 1597-1608
123. Fonda, M. L. (1992), Purification and characterization of vitamin B_6-phosphate phosphatase from human erythrocytes. *Journal of Biological Chemistry***. 267**, p. 15978-15983
124. McCormick, D. B. and Chen, H. (1999), Update on interconversions of vitamin B-6 with its coenzyme. *The Journal of Nutrition***. 129**, p. 325-327
125. Anderson, B. B., Fulford-Jones, C. E., Child, J. A., Beard, M. E. J. and Bateman, C. J. T. (1971), Conversion of vitamin B_6 compounds to active forms in the red blood cell. *Journal of Clinical Investigation***. 50**, p. 1901-1909
126. Anderson, B. B., Perry, G. M., Clements, J. E. and Greany, M. (1989), Rapid uptake and clearance of pyidoxine by red blood cells *in vivo*. *The American Journal of Clinical Nutrition***. 50**, p. 1059-1063
127. Mehansho, H. and Henderson, L. H. (1980), Transport and accumulation of pyridoxine and pyridoxal by erthrocytes. *Journal of Biological Chemistry***. 255**, p. 11901-11907
128. Gao, G. and Fonda, M. L. (1994), Kinetic analysis and chemical modification of vitamin B_6 phosphatase from human erythrocytes. *Journal of Biological Chemistry***. 269**, p. 7163-7168
129. Fonda, M. L. and Harker, M. S. (1982), Metabolism of pyridoxine and protein binding of the metabolites in human erythrocytes. *The American Journal of Clinical Nutrition***. 35**, p. 1391-1399
130. Benesch, R. E., Benesch, R., Renthal, R. D. and Maeda, N. (1972), Affinity labeling of the polyphosphate bining site of hemoglobin. *Biochemistry***. 11**, p. 3576-3582
131. Talwar, D., Quasim, T., McMillan, D. C., Kinsella, J., Williamson, C. and O'Reilly, D. S. J. (2003), Optimisation and validation of a sensitive high-performance liquid chromatography assay for routine measurement of pyridoxal 5-phosphate in human plasma and red cells using pre-column semicarbazide derivatisation. *Journal of Chromatography B***. 792**, p. 333-343
132. Wrenger, C., Eschbach, M., Müller, I. B., Warnecke, D. and Walter, R. D. (2005), Analysis of the vitamin B6 biosynthesis pathway in the human malaria parasite *Plasmodium falciparum*. *Journal of Biological Chemistry***. 280**, p. 5242-5248
133. Müller, I. B., Knöckel, J., Groves, M. R., Jordanova, R., Ealick, S. E., Walter, R. D. and Wrenger, C. (2008), The assembly of the plasmodial PLP synthase complex follows a defined course. *PLoS ONE***. 3**, p. 1-9
134. Müller, I. B., Knöckel, J., Eschbach, M., Bergmann , B., Walter, R. D. and Wrenger, C. (2010), Secretion of an acid phosphatase provides a possible mechanism to acquire host nutrients by *Plasmodium falciparum*. *Cellular Microbiology***. 12**, p. 677-691
135. Müller, S. and Kappes, B. (2007), Vitamin and cofactor biosynthesis pathways in *Plasmodium* and other apicomplean parasites. *Trends in Parasitology***. 23**, p. 112-121
136. Knöckel, J., Bergmann, B., Müller, I. B., Rathaur, S., Walter, R. D. and Wrenger, C. (2008), Filling the gap of intracellular dephosphorylation in the *Plasmodium falciparum* vitamin B1 biosynthesis. *Molecular & Biochemical Parasitology***. 157**, p. 241-243
137. Lasonder, E., Ishihama, Y., Andersen, J. S., Vermunt, A. M. W., Pain, A., Sauerwein, R. W., Eling, W. M. C., Hall, N., Waters, A. P., Stunnenberg, H. G. and Mann, M. (2002), Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature***. 419**, p. 537-542
138. Denslow, S. A., Walls, A. A. and Daub, M. E. (2005), Regulation of biosynthetic genes and antioxidant properties of vitamin B_6 vitamers during plants defense responses. *Physiological and Molecular Plant Pathology***. 66**, p. 244-255
139. Matxain, J. M., Padro, D., Ristilä, M., Strid, Å. and Eriksson, L. A. (2009), Evidence of high •OH radical quenching efficiency by vitamin B_6. *The Journal of Physical Chemistry B***. 113**, p. 9629-9632

140. Butzloff, S., Groves, M. R., Wrenger, C. and Müller, I. B. (2012), Cytometric quantification of singlet oxygen in the human malaria parasite *Plasmodium falciparum*. *Cytometry Part A***.** **81**, p. 698-703
141. Knöckel, J., Müller, I. B., Butzloff, S., Bergmann , B., Walter, R. D. and Wrenger, C. (2012), The antioxidative effect of de novo generated vitamin B6 in *Plasmodium falciparum* validated by protein interference. *Biochemical Journal***.** **443**, p. 397-405
142. Milord, F., Pépin, J., Ethier, L., Loko, L. and Mpia, B. (1992), Efficacy and toxicity of eflornithine for treatment of Trypanosoma brucei gambiense sleeping sickness. *The Lancet***.** **340**, p. 652-655
143. Wang, Q.-P., Jammoul, F., Duboc, A., Gong, J., Simonutti, M., Dubus, E., Craft, C. M., Ye, W., Sahel, J. A. and Picaud, S. (2008), Treatment of epilepsy: the GABA-transaminase inhibitor, vigabatrin, induces neuronal plasticity in the mouse retina. *European Journal of Neuroscience***.** **27**, p. 2177-2187
144. Amadasi, A., Bertoldi, M., Contestabile, R., Bettati, S., Cellini, B., Luigi di Salvo, M., Borri-Voltattorni, C., Bossa, F. and Mozzarelli, A. Pyridoxal 5-Phosphate Enzymes as Targets for Therapeutic Agents. *Current Medicinal Chemistry***.** **14**, p. 1291-1324
145. Kappes, B., Tews, I., Binter, A. and Macheroux, P. (2011), PLP-dependent enzymes as potential drug targets for protozoan diseases. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics***.** **1814**, p. 1567-1576
146. Plouffe, D., Brinker, A., McNamara, C., Henson, K., Kato, N., Kuhen, K., Nagle, A., Adrián, F., Matzen, J. T. and Anderson, P. (2008), In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proceedings of the National Academy of Sciences***.** **105**, p. 9059
147. Zartler, E. R. and Shapiro, M. J. (2005), Fragonomics: fragment-based drug discovery. *Current Opinion in Chemical Biology***.** **9**, p. 366-370
148. Leach, A. R. and Hann, M. M. (2011), Molecular complexity and fragment-based drug discovery: ten years on. *Current Opinion in Chemical Biology***.** **15**, p. 489-496
149. Swinney, D. C. and Anthony, J. (2011), How were new medicines discovered? *Nature Reviews Drug Discovery***.** **10**, p. 507-519
150. Lyne, P. D. (2002), Structure-based virtual screening: an overview. *Drug Discovery Today***.** **7**, p. 1047-1055
151. Nicola, G., Smith, C. A., Lucumi, E., Kuo, M. R., Karagyozov, L., Fidock, D. A., Sacchettini, J. C. and Abagyan, R. (2007), Discovery of novel inhibitors targeting enoyl–acyl carrier protein reductase in *Plasmodium falciparum* by structure-based virtual screening. *Biochemical and Biophysical Research Communications***.** **358**, p. 686-691
152. Haque, T. S., Skillman, A. G., Lee, C. E., Habashita, H., Gluzman, I. Y., Ewing, T. J. A., Goldberg, D. E., Kuntz, I. D. and Ellman, J. A. (1999), Potent, Low-Molecular-Weight Non-Peptide Inhibitors of Malarial Aspartyl Protease Plasmepsin II. *Journal of Medicinal Chemistry***.** **42**, p. 1428-1440
153. Kick, E. K., Roe, D. C., Skillman, A. G., Liu, G., Ewing, T. J., Sun, Y., Kuntz, I. D. and Ellman, J. A. (1997), Structure-based design and combinatorial chemistry yield low nanomolar inhibitors of cathepsin D. *Chem Biol***.** **4**, p. 297-307
154. Desai, P. V., Patny, A., Gut, J., Rosenthal, P. J., Tekwani, B., Srivastava, A. and Avery, M. (2006), Identification of Novel Parasitic Cysteine Protease Inhibitors by Use of Virtual Screening. 2. The Available Chemical Directory. *Journal of Medicinal Chemistry***.** **49**, p. 1576-1584
155. Yang, S.-Y. (2010), Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today***.** **15**, p. 444-450
156. Strohmeier, M., Raschle, T., Mazurkiewicz , J., Rippe, K., Sinning, I., Fitzpatrick, T. B. and Tews, I. (2006), Structure of a bacterial pyridoxal 5'-phosphate synthase complex. *Proceedings of the National Academy of Sciences of the United States of America***.** **103**, p. 19284-19289

157. Zein, F., Zhang, Y., Kang, Y. N., Burns, K., Begley, T. P. and Ealick, S. E. (2006), Structural insights into the mechanism of the PLP synthase holoenzyme from *Thermotoga maritima*. *Biochemistry***. 45**, p. 14609-14620
158. Knöckel, J., Jordanova, R., Müller, I. B., Wrenger, C. and Groves, M. R. (2009), Mobility of the conserved glycine 155 is required for formation of the active plasmodial dodecamer. *Biochimica et Biophysica Acta***. 1790**, p. 347-350
159. Flicker, K., Neuwirth, M., Strohmeier, M., Kappes, B., Tews, I. and Macheroux, P. (2007), Structural and thermodynamic insights into the assembly of the heteromeric pyridoxal phosphate synthase from *Plasmodium falciparum*. *Journal of Molecular Biology***. 374**, p. 732-748
160. Hanes, J. W., Keresztes, I. and Begley, T. P. (2008), 13C NMR snapshots of the complex reaction coordinate of pyridoxal phosphate synthase. *Nature Chemical Biology***. 4**, p. 425-430
161. Hanes, J. W., Burns, K. E., Hilmey, D. G., Chatterjee, A., Dorrestein, P. C. and Begley, T. P. (2008), Mechanistic studies on pyridoxal phosphate synthase: the reaction pathway leading to a chromophoric intermediate. *Journal of the American Chemical Society***. 130**, p. 3043-3052
162. Raschle, T., Arigoni, D., Brunisholz, R., Rechsteiner, H., Amrhein, N. and Fitzpatrick, T. B. (2007), Reaction mechanism of pyridoxal 5'-phosphate synthase. Detection of an enzyme-bound chromophoric intermediate. *Journal of Biological Chemistry***. 282**, p. 6098-6105
163. Zhang, X., Teng, Y., Liu, J., He, Y., Zhou, K., Chen, Y. and Zhou, C. (2010), Structural insights into the catalytic mechanism of the yeast pyridoxal 5-phosphate synthase Snz1. *Biochemical Journal***. 432**, p. 445-450
164. Guédez, G., Hipp, K., Windeisen, V., Derrer, B., Gengenbacher, M., Böttcher, B., Sinning, I., Kappes, B. and Tews, I. (2012), Assembly of the eukaryotic PLP-synthase complex from *Plasmodium* and activation of the Pdx1 enzyme. *Structure***. 20**, p. 172-184
165. Hillisch, A., Pineda, L. F. and Hilgenfeld, R. (2004), Utility of homology models in the drug discovery process. *Drug Discovery Today***. 9**, p. 659-669
166. Šali, A. and Blundell, T. L. (1993), Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology***. 234**, p. 779-815
167. Prlić, A., Bliven, S., Rose, P. W., Bluhm, W. F., Bizon, C., Godzik, A. and Bourne, P. E. (2010), Pre-calculated protein structure alignments at the RCSB PDB website. *Bioinformatics***. 26**, p. 2983-2985
168. Shindyalov, I. N. and Bourne, P. E. (1998), Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Engineering***. 11**, p. 739-747
169. Laskowski, R. A. (2009), PDBsum new things. *Nucleic Acids Research***. 37**, p. D355-D359
170. Laskowski, R. A. (2001), PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Research***. 29**, p. 221-222
171. Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993), PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography***. 26**, p. 283-291
172. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983), CHARMM: A program for macromolecular energy minimization and dynamics calculations. *Journal of Computational Chemistry***. 4**, p. 187-217
173. Cheng, T., Li, Q., Zhou, Z., Wang, Y. and Bryant, S. (2012), Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *The AAPS Journal***. 14**, p. 133-141
174. Irwin, J. J. and Shoichet, B. K. (2005), ZINC-a free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling***. 45**, p. 177-182
175. Kabsch, W. (1976), A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A***. 32**, p. 922-923

176. Venkatachalam, C., Jiang, X., Oldfield, T. and Waldman, M. (2003), LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling***. 21**, p. 289-307
177. Mayo, S. L., Olafson, B. D. and Goddard, W. A. (1990), DREIDING: A generic force field for molecular simulations. *Journal of Physical Chemistry***. 94**, p. 8897-8909
178. Krammer, A., Kirchhoff, P. D., Jiang, X., Venkatachalam, C. and Waldman, M. (2005), LigScore: a novel scoring function for predicting binding affinities. *Journal of Molecular Graphics and Modelling***. 23**, p. 395-407
179. Shen, M. and Sali, A. (2006), Statistical potential for assessment and prediction of protein structures. *Protein Science***. 15**, p. 2507-2524
180. Ho, B. and Brasseur, R. (2005), The Ramachandran plots of glycine and pre-proline. *BMC Structural Biology***. 5**, p. 14
181. MacArthur, M. W. and Thornton, J. M. (1991), Influence of proline residues on protein conformation. *Journal of Molecular Biology***. 218**, p. 397-412
182. Derrer, B., Windeisen, V., Guedez Rodriguez, G., Seidler, J., Gengenbacher, M., Lehmann, W. D., Rippe, K., Sinning, I., Tews, I. and Kappes, B. (2010), Defining the structural requirements of ribose 5-phosphate-binding and intersubunt cross-talk of the malarial pyridoxal 5-phopshate synthase. *FEBS Letters***. 584**, p. 4169-4174
183. Pal, D. and Chakrabarti, P. (2002), On residues in the disallowed region of the Ramachandran map. *Biopolymers***. 63**, p. 195-206
184. Lipinski, C. A., Lombardo, F., Dominy, B. W. and Feeney, P. J. (2001), Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews***. 46**, p. 3-26
185. Lipinski, C. A. (2004), Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies***. 1**, p. 337-341
186. Burger, P. B. (2009) Development of a dynamic receptor-based pharmacophore model of *Plasmodium falciparum* spermidine synthase for selective inhibitor identification. PhD thesis, University of Pretoria, Pretoria, http://upetd.up.ac.za/thesis/available/etd-05252009-220942/, viewed 2013-01-21
187. Knöckel, J., Müller, I. B., Bergmann, B., Walter, R. D. and Wrenger, C. (2007), The apicomplexan parasite *Toxoplasma gondii* generates pyridoxal phosphate *de novo*. *Molecular & Biochemical Parasitology***. 152**, p. 108-111
188. Galperin, M. Y. and Koonin, E. V. (1997), Seqeunce analysis of an exceptionally conserved operon suggests enzymes for a new link beteen histidine and purine biosynthesis. *Molecular Microbiology***. 24**, p. 443-445
189. Wolber, G., Seidel, T., Bendix, F. and Langer, T. (2008), Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discovery Today***. 13**, p. 23-29
190. Warren, G. L., Andrews, C. W., Capelli, A.-M., Clarke, B., LaLonde, J., Lambert, M. H., Lindvall, M., Nevins, N., Semus, S. F., Senger, S., Tedesco, G., Wall, I. D., Woolven, J. M., Peishoff, C. E. and Head, M. S. (2005), A Critical Assessment of Docking Programs and Scoring Functions. *Journal of Medicinal Chemistry***. 49**, p. 5912-5931
191. Davis, A. M., Keeling, D. J., Steele, J., Tomkinson, N. P. and Tinker, A. C. Components of Successful Lead Generation. *Current Topics in Medicinal Chemistry***. 5**, p. 421-439
192. Musicki, B. and Widlanski, T. S. (1990), Synthesis of carbohydrate sulfonates and sulfonate esters. *The Journal of Organic Chemistry***. 55**, p. 4231-4233
193. Böhm, H.-J., Flohr, A. and Stahl, M. (2004), Scaffold hopping. *Drug Discovery Today: Technologies***. 1**, p. 217-224
194. Burns, K. E., Xiang, Y., Kinsland, C. L., McLafferty, F. W. and Begley, T. P. (2005), Reconstitution and biochemical characterization of a new pyridoxal-5'-phosphate biosynthetic pathway. *Journal of the American Chemical Society***. 127**, p. 3682-3683

195. Tazuya, K., Adachi, Y., Masuda, K., Yamada and Kumaoka, H. (1995), Origin of the nitrogen atom of pyridoxine in *Saccharomyces cerevisiae*. *Biochemica et Biophysica Acta***. 1244**, p. 113-116
196. Hanes, J. W., Keresztes, I. and Begley, T. P. (2008), Trapping of a chromophoric intermediate in the Pdx1-catalyzed biosynthesis of pyridoxal 5'-phosphate. *Angewandte Chemie (International ed. in English)***. 47**, p. 2101-2105
197. Bradford, M. M. (1976), A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry***. 72**, p. 248-254
198. Laemmli, U. K. (1970), Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature***. 117**, p. 680-685
199. Sharma, S. and Dakshinamurti, K. (1992), Determination of vitamin B6 vitamers and pyridoxic acid in biological samples. *Journal of Chromatography: Biomedical Applications***. 578**, p. 45-51
200. Sambrook, J., Fritsch, E. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, pp. A1–A4. ed.)^eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
201. Trager, W. and Jensen, J. B. (1976), Human malaria parasites in continuous culture. *Science***. 193**, p. 673-675
202. Lambros, C. and Vanderberg, J. P. (1979), Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *The Journal of Parasitology***. 3**, p. 418-420
203. Das Gupta, R., Krause-Ihle, T., Bergmann, B., Müller, I. B., Khomutov, A. R., Müller, S., Walter, R. D. and Lüersen, K. (2005), 3-Aminooxy-1-aminopropane and derivatives have an antiproliferative effect on cultured *Plasmodium falciparum* by decreasing intracellular polyamine concentrations. *Antimicrobial Agents and Chemotherapy***. 49**, p. 2857-2864
204. Liu, J., Gluzman, I. Y., Drew, M. E. and Goldberg, D. E. (2005), The role of *Plasmodium falciparum* food vacuole plasmepsins. *Journal of Biological Chemistry***. 280**, p. 1432-1437
205. Brooks, H. B., Geeganage, S., Kahl, S. D., Montrose, C., Sittampalam, S., Smith, M. C. and Weidner, J. R. (2004) Basics of Enzymatic Assays for HTS. In Assay Guidance Manual (Sittampalam, G. S., Gal-Edd, N., Arkin, M., Auld, D., Austin, C., Bejcek, B., Glicksman, M., Inglese, J., Lemmon, V., Li, Z., McGee, J., McManus, O., Minor, L., Napper, A., Riss, T., Trask, O. J. and Weidner, J., eds.), Eli Lilly & Company and the National Center for Advancing Translational Sciences., Bethesda MD
206. Burgos, E. and Salmon, L. (2004), Synthesis and evaluation of new 4-phospho-D-erythronic acid derivatives as competitive inhibitors of spinach ribose-5-phosphate isomerase. *Tetrahedron Letters***. 45**, p. 753-756
207. Dahl, E. L., Shock, J. L., Shenai, B. R., Gut, J., DeRisi, J. L. and Rosenthal, P. J. (2006), Tetracyclines Specifically Target the Apicoplast of the Malaria Parasite *Plasmodium falciparum*. *Antimicrobial Agents and Chemotherapy***. 50**, p. 3124-3131
208. Blackmore, P. F., Williams, J. F. and MacLeod, J. K. (1976), Dimerization of erythrose 4-phosphate. *FEBS Letters***. 64**, p. 222-226
209. Parker, E. J., Bulloch, E. M. M., Jameson, G. B. and Abell, C. (2001), Substrate deactivation of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate 7-Phosphate synthase by erythrose 4-phosphate. *Biochemistry***. 40**, p. 14821-14828
210. Backhausen, J. E., Jöstingmeyer, P. and Scheibe, R. (1997), Competitive inhibition of spinach leaf phosphoglucose isomerase isoenzymes by erythrose 4-phosphate. *Plant Science***. 130**, p. 121-131
211. Lolis, E. and Petsko, G. A. (1990), Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-. ANG. resolution: implications for catalysis. *Biochemistry***. 29**, p. 6619-6625
212. Jeffery, C. J., Bahnson, B. J., Chien, W., Ringe, D. and Petsko, G. A. (2000), Crystal structure of rabbit phosphoglucose isomerase, a glycolytic enzyme that moonlights as

neuroleukin, autocrine motility factor, and differentiation mediator. *Biochemistry***. 39**, p. 955-964

213. Read, J., Pearce, J., Li, X., Muirhead, H., Chirgwin, J. and Davies, C. (2001), The crystal structure of human phosphoglucose isomerase at 1.6 Å resolution: implications for catalytic mechanism, cytokine activity and haemolytic anaemia1. *Journal of Molecular Biology***. 309**, p. 447-463
214. Roberts, C. W., Roberts, F., Lyons, R. E., Kirisits, M. J., Mui, E. J., Finnerty, J., Johnson, J. J., Ferguson, D. J. P., Coggins, J. R., Krell, T., Coombs, G. H., Milhous, W. K., Kyle, D. E., Tzipori, S., Barnwell, J., Dame, J. B., Jane, C. and McLeod, R. (2002), The Shikimate Pathway and Its Branches in Apicomplexan Parasites. *Journal of Infectious Diseases***. 185**, p. S25-S36
215. Bozdech, Z. and Ginsburg, H. (2005), Data mining of the transcriptome of *Plasmodium falciparum*: the pentose phosphate pathway and ancillary processes. *Malaria Journal***. 4**, p. 17
216. Joshi, S., Singh, A. R., Kumar, A., Misra, P. C., Siddiqi, M. I. and Saxena, J. K. (2008), Molecular cloning and characterization of *Plasmodium falciparum* transketolase. *Molecular & Biochemical Parasitology***. 160**, p. 32-41
217. Dieckmann, A. and Jung, A. (1986), Mechanisms of sulfadoxine resistance in *Plasmodium falciparum*. *Molecular & Biochemical Parasitology***. 19**, p. 143-147
218. Burgos, E. and Salmon, L. (2004), Synthesis and kinetic evaluation of 4-deoxy-4-phosphonomethyl-D-erythronate, the first hydrolytically stable and potent competitive inhibitor of ribose-5-phosphate isomerase. *Tetrahedron Letters***. 45**, p. 3465-3469
219. Raschle, T., Speziga, D., Kress, W., Moccand, C., Gehrig, P., Amrhein, N., Weber-Ban, E. and Fitzpatrick, T. B. (2009), Intersubunit cross-talk in pyridoxal 5'-phosphate synthase coordinated by the C terminus of the synthase subunit. *Journal of Biological Chemistry***. 284**, p. 7706-7718
220. Cassera, M. B., Hazleton, K. Z., Riegelhaupt, R. M., Merino, E. F., Luo, M., Akabas, M. H. and Schramm, V. L. (2008), Erythrocytic adenosine monophosphate as an alternative purine source in *Plasmodium facliparum*. *Journal of Biological Chemistry***. 283**, p. 32889-32899
221. Moccand, C., Kaufmann, M. and Fitzpatrick, T. B. (2011), It takes two to tango: defining an essential second active site in pyridoxal 5'-phosphate synthase. *PLoS ONE***. 6**, p. e16042
222. Birkholtz, L., van Brummelen, A. C., Clark, K., Niemand, J., Maréchal, E., Llinás, M. and Louw, A. I. (2008), Exploring functional genomics for drug target and therapeutics discovery in Plasmodia. *Acta Tropica***. 105**, p. 113-123
223. Bai, J. P. F. and Abernethy, D. R. (2013), Systems Pharmacology to Predict Drug Toxicity: Integration Across Levels of Biological Organization. *Annual Review of Pharmacology and Toxicology***. 53**, p. null
224. Baginsky, S., Hennig, L., Zimmermann, P. and Gruissem, W. (2010), Gene expression analysis, proteomics, and network discovery. *Plant Physiology***. 152**, p. 402-410
225. Bozdech, Z., Llinás, 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 Biology***. 1**, p. E5
226. Silvestrini, F., Bozdech, Z., Lanfrancotti, A., Giulio, E. D., Bultrini, E., Picci, L., deRisi, J. L., Pizzi, E. and Alano, P. (2005), Genome-wide identification of genes upregulated at the onset of gametocytogenesis in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology***. 143**, p. 100-110
227. Young, J. A., Fivelman, Q. L., Blair, P. L., de la Vega, P., Le Roch, K. G., Zhou, Y., Carucci, D. J., Baker, D. A. and Winzeler, E. A. (2005), The *Plasmodium falciparum* sexual development transcriptome: A microarray analysis using ontology-based pattern identification. *Molecular and Biochemical Parasitology***. 143**, p. 67-79
228. 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**, p. 1503-1508

229. Llinás, M., Bozdech, Z., Wong, E. D., Adai, A. T. and DeRisi, J. L. (2006), Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Research*. **34**, p. 1166-1173
230. Mackinnon, M. J., Li, J., Mok, S., Kortok, M. M., Marsh, K., Preiser, P. R. and Bozdech, Z. (2009), Comparative Transcriptional and Genomic Analysis of *Plasmodium falciparum* Field Isolates. *PLoS Pathog*. **5**, p. e1000644
231. Siau, A., Touré, F. S., Ouwe-Missi-Oukem-Boyer, O., Cicéron, L., Mahmoudi, N., Vaquero, C., Froissard, P., Bisvigou, U., Bisser, S., Coppée, J.-Y., Bischoff, E., David, P. H. and Mazier, D. (2007), Whole-Transcriptome Analysis of *Plasmodium falciparum* Field Isolates: Identification of New Pathogenicity Factors. *Journal of Infectious Diseases*. **196**, p. 1603-1612
232. Jiang, H., Patel, J. J., Yi, M., Mu, J., Ding, J., Stephens, R., Cooper, R. A., Ferdig, M. T. and Su, X.-z. (2008), Genome-wide compensatory changes accompany drug-selected mutations in the *Plasmodium falciparum* crt gene. *PLoS ONE*. **3**, p. e2484
233. Daily, J. P., Le Roch, K. G., Sarr, O., Ndiaye, D., Lukens, A., Zhou, Y., Ndir, O., Mboup, S., Sultan, A., Winzeler, E. A. and Wirth, D. F. (2005), *In Vivo* transcriptome of *Plasmodium falciparum* reveals overexpression of transcripts that encode surface proteins. *Journal of Infectious Diseases*. **191**, p. 1196-1203
234. Torrentino-Madamet, M., Almeras, L., Desplans, J., Priol, Y., Belghazi, M., Pophillat, M., Fourquet, P., Jammes, Y. and Parzy, D. (2011), Global response of *Plasmodium falciparum* to hyperoxia: a combined transcriptomic and proteomic approach. *Malaria Journal*. **10**, p. 4
235. Oakley, M. S. M., Kumar, S., Anantharaman, V., Zheng, H., Mahajan, B., Haynes, J. D., Moch, J. K., Fairhurst, R., McCutchan, T. F. and Aravind, L. (2007), Molecular factors and biochemical pathways induced by febrile temperature in intraerythrocytic *Plasmodium falciparum* parasites. *Infection and Immunity*. **75**, p. 2012
236. Kritsiriwuthinan, K., Chaotheing, S., Shaw, P., Wongsombat, C., Chavalitshewinkoon-Petmitr, P. and Kamchonwongpaisan, S. (2011), Global gene expression profiling of *Plasmodium falciparum* in response to the anti-malarial drug pyronaridine. *Malaria Journal*. **10**, p. 242
237. Brunner, R., Aissaoui, H., Boss, C., Bozdech, Z., Brun, R., Corminboeuf, O., Delahaye, S., Fischli, C., Heidmann, B., Kaiser, M., Kamber, J., Meyer, S., Papastogiannidis, P., Siegrist, R., Voss, T., Welford, R., Wittlin, S. and Binkert, C. (2012), Identification of a New Chemical Class of Antimalarials. *Journal of Infectious Diseases*. **206**, p. 735-743
238. Cui, L., Wang, Z., Miao, J., Miao, M., Chandra, R., Jiang, H., Su, X.-z. and Cui, L. (2012), Mechanisms of in vitro resistance to dihydroartemisinin in *Plasmodium falciparum*. *Molecular Microbiology*. **86**, p. 111-128
239. Dharia, N., Sidhu, A., Cassera, M., Westenberger, S., Bopp, S., Eastman, R., Plouffe, D., Batalov, S., Park, D., Volkman, S., Wirth, D., Zhou, Y., Fidock, D. and Winzeler, E. (2009), Use of high-density tiling microarrays to identify mutations globally and elucidate mechanisms of drug resistance in *Plasmodium falciparum*. *Genome Biology*. **10**, p. R21
240. Gunasekera, A. M., Myrick, A., Roch, K. L., Winzeler, E. and Wirth, D. F. (2007), *Plasmodium falciparum*: Genome wide perturbations in transcript profiles among mixed stage cultures after chloroquine treatment. *Experimental Parasitology*. **117**, p. 87-92
241. Ganesan, K., Ponmee, N., Jiang, L., Fowble, J. W., White, J., Kamchonwongpaisan, S., Yuthavong, Y., Wilairat, P. and Rathod, P. K. (2008), A genetically hard-wired metabolic transcriptome in *Plasmodium falciparum* fails to mount protective responses to lethal antifolates. *PLoS Pathogens*. **4**, p. e1000214
242. Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B. K., Haase, S., Engelberg, K., Cheemadan, S., Spielmann, T. and Preiser, P. R. (2009), Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. *Nature Biotechnology*. **28**, p. 91-98

243. Smit, S. (2010) Functional consequences of the inhibition of Malaria S-adenosylmethionine decarboxylase as a key regulator of polyamine and methionine metabolism. PhD thesis, University of Pretoria, Pretoria, http://upetd.up.ac.za/thesis/available/etd-06222011-081539/, viewed 2012-10-21
244. Van Brummelen, A. C. (2008) Functional genomics analysis of the effects of co-inhibition of the malarial S-adenosylmethionine decarboxylase/ornithine decarboxylase. PhD thesis, University of Pretoria, Pretoria, http://upetd.up.ac.za/thesis/available/etd-05302009-124548/, viewed 2012-10-21
245. Chomczynski, P. and Sacchi, N. (1987), Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry***. 162**, p. 156-159
246. Delibato, E., Gattuso, A., Minucci, A., Auricchio, B., De Medici, D., Toti, L., Castagnola, M., Capoluongo, E. and Gianfranceschi, M. V. (2009), PCR experion automated electrophoresis system to detect Listeria monocytogenes in foods. *Journal of Separation Science***. 32**, p. 3817-3821
247. Göhlmann, H. and Talloen, W. (2009) Gene expression studies using Affymetrix microarrays. Chapman & Hall/CRC
248. Smyth, G. (2005), Limma: linear models for microarray data. *Bioinformatics and computational biology solutions using R and bioconductor*, p. 397-420
249. Ritchie, M. E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A. and Smyth, G. K. (2007), A comparison of background correction methods for two-colour microarrays. *Bioinformatics***. 23**, p. 2700
250. Rainer, J., Sanchez-Cabo, F., Stocker, G., Sturn, A. and Trajanoski, Z. (2006), CARMAweb: comprehensive R-and bioconductor-based web service for microarray data analysis. *Nucleic Acids Research***. 34**, p. W498
251. Yang, M., Ruan, Q., Yang, J., Eckenrode, S., Wu, S., McIndoe, R. and She, J. (2001), A statistical method for flagging weak spots improves normalization and ratio estimates in microarrays. *Physiological Genomics***. 7**, p. 45
252. Smyth, G. K. and Speed, T. (2003), Normalization of cDNA microarray data. *Methods***. 31**, p. 265-273
253. Smyth, G. K., Michaud, J. and Scott, H. S. (2005), Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics***. 21**, p. 2067
254. Philip, L., Clotilde, C. R., Fourie, J., Abraham, L. and Dave, B. MADIBA: A web server toolkit for biological interpretation of *Plasmodium* and plant gene clusters. *BMC Genomics***. 9**, p.
255. Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998), Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences***. 95**, p. 14863
256. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. and Mesirov, J. P. (2005), Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America***. 102**, p. 15545-15550
257. Copois, V., Bibeau, F., Bascoul-Mollevi, C., Salvetat, N., Chalbos, P., Bareil, C., Candeil, L., Fraslon, C., Conseiller, E. and Granci, V. (2007), Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality. *Journal of Biotechnology***. 127**, p. 549-559
258. Denisov, V., Strong, W., Walder, M., Gingrich, J. and Wintz, H. (2008), Development and validation of RQI: An RNA quality indicator for the Experion automated electrophoresis system. *Bio-Rad Bulletin***. 5761**, p.
259. van Brummelen, A. C., Olszewski, K. L., Wilinski, D., Llinás, M., Louw, A. I. and Birkholtz, L. M. (2009), Co-inhibition of *Plasmodium falciparum* S-adenosylmethionine

decarboxylase/ornithine decarboxylase reveals perturbation-specific compensatory mechanisms by transcriptome, proteome, and metabolome analyses. *Journal of Biological Chemistry***. 284**, p. 4635

260. Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009), Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols***. 4**, p. 44-57
261. Sherman, B. T. and Lempicki, R. A. (2009), Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research***. 37**, p. 1-13
262. Natalang, O., Bischoff, E., Deplaine, G., Proux, C., Dillies, M. A., Sismeiro, O., Guigon, G., Bonnefoy, S., Patarapotikul, J. and Mercereau-Puijalon, O. (2008), Dynamic RNA profiling in *Plasmodium falciparum* synchronized blood stages exposed to lethal doses of artesunate. *BMC Genomics***. 9**, p. 388
263. Chaal, B. K., Gupta, A. P., Wastuwidyaningtyas, B. D., Luah, Y.-H. and Bozdech, Z. (2010), Histone Deacetylases Play a Major Role in the Transcriptional Regulation of the *Plasmodium falciparum* Life Cycle. *PLoS Pathog***. 6**, p. e1000737
264. Page, R. D. M. (1996), An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences***. 12**, p.
265. Saldanha, A. J. (2004), Java Treeview—extensible visualization of microarray data. *Bioinformatics***. 20**, p. 3246-3248
266. Date, S. V. and Stoeckert, C. J. (2006), Computational modeling of the *Plasmodium falciparum* interactome reveals protein function on a genome-wide scale. *Genome Research***. 16**, p. 542-549
267. Snel, B., Lehmann, G., Bork, P. and Huynen, M. A. (2000), STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res***. 28**, p. 3442-3444
268. Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L. J. and von Mering, C. (2011), The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res***. 39**, p. D561-568
269. Kuhn, M., Szklarczyk, D., Franceschini, A., von Mering, C., Jensen, L. J. and Bork, P. (2012), STITCH 3: zooming in on protein–chemical interactions. *Nucleic Acids Research***. 40**, p. D876-D880
270. Kuhn, M., von Mering, C., Campillos, M., Jensen, L. J. and Bork, P. (2008), STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res***. 36**, p. D684-688
271. Birkholtz, L., Joubert, F., Neitz, A. W. H. and Louw, A. I. (2003), Comparative properties of a three-dimensional model of *Plasmodium falciparum* ornithine decarboxylase. *Proteins: Structure, Function and Genetics***. 50**, p. 464-473
272. Assaraf, Y. G., Golenser, J., Spira, D. T., Messer, G. and Bachrach, U. (1987), Cytostatic effect of DL-alpha-difluoromethylornithine against *Plasmodium falciparum* and its reversal by diamines and spermidine. *Parasitology Research***. 73**, p. 313-318
273. Ubbink, J. B., van der Merwe, A., Delport, R., Allen, R. H., Stabler, S. P., Riezler, R. and Vermaak, W. (1996), The effect of a subnormal vitamin B-6 status on homocysteine metabolism. *Journal of Clinical Investigation***. 98**, p. 177
274. Kanellis, P., Gagliardi, M., Banath, J. P., Szilard, R. K., Nakada, S., Galicia, S., Sweeney, F. D., Cabelof, D. C., Olive, P. L. and Durocher, D. (2007), A screen for suppressors of gross chromosomal rearrangements identifies a conserved role for PLP in preventing DNA lesions. *PLoS Genetics***. 3**, p. e134
275. Huq, M. D. M., Tsai, N.-P., Lin, Y.-P., Higgins, L. and Wei, L.-N. (2007), Vitamin B6 conjugation to nuclear corepressor RIP140 and its role in gene regulation. *Nat Chem Biol***. 3**, p. 161-165

276. Oka, T., Sugitatsu, H., Nordin, H., Thakur, M. K., Aoyama, M., Sasagawa, T., Suzuki, I. and Tsuji, H. (2001), Pyridoxal 5′-phosphate inhibits DNA binding of HNF1. *Biochimica et Biophysica Acta (BBA) - General Subjects***. 1568**, p. 189-196
277. Jochmann, N., Götker, S. and Tauch, A. (2011), Positive transcriptional control of the pyridoxal phosphate biosynthesis genes pdxST by the MocR-type regulator PdxR of *Corynebacterium glutamicum* ATCC 13032. *Microbiology***. 157**, p. 77-88
278. Vermeersch, J. J., Christmann-Franck, S., Karabashyan, L. V., Fermandjian, S., Mirambeau, G. and Der Garabedian, P. A. (2004), Pyridoxal 5′-phosphate inactivates DNA topoisomerase IB by modifying the lysine general acid. *Nucleic Acids Research***. 32**, p. 5649-5657
279. Read, M., Muller, I., Mitchell, S., Sims, P. and Hyde, J. (2010), Dynamic subcellular localization of isoforms of the folate pathway enzyme serine hydroxymethyltransferase (SHMT) through the erythrocytic cycle of *Plasmodium falciparum*. *Malaria Journal***. 9**, p. 351
280. Salcedo, E., Sims, P. F. G. and Hyde, J. E. (2005), A glycine-cleavage complex as part of the folate one-carbon metabolism of *Plasmodium falciparum*. *Trends in Parasitology***. 21**, p. 406-411
281. McMillan, P. J., Stimmler, L. M., Foth, B. J., McFadden, G. I. and Müller, S. (2005), The human malaria parasite *Plasmodium falciparum* possesses two distinct dihydrolipoamide dehydrogenases. *Molecular Microbiology***. 55**, p. 27-38
282. Douce, R., Bourguignon, J., Neuburger, M. and Rébeillé, F. (2001), The glycine decarboxylase system: a fascinating complex. *Trends in Plant Science***. 6**, p. 167-176
283. Sirawaraporn, W., Sirawaraporn, R., Cowman, A. F., Yuthavong, Y. and Santi, D. V. (1990), Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*. *Biochemistry***. 29**, p. 10779-10785
284. Babbitt, S. E., Altenhofen, L., Cobbold, S. A., Istvan, E. S., Fennell, C., Doerig, C., Llinás, M. and Goldberg, D. E. (2012), *Plasmodium falciparum* responds to amino acid starvation by entering into a hibernatory state. *Proceedings of the National Academy of Sciences***. 109**, p. E3278–E3287
285. McHardy, A. C., Tauch, A., Rückert, C., Pühler, A. and Kalinowski, J. (2003), Genome-based analysis of biosynthetic aminotransferase genes of *Corynebacterium glutamicum*. *Journal of Biotechnology***. 104**, p. 229-240
286. El Qaidi, S., Yang, J., Zhang, J.-R., Metzger, D. W. and Bai, G. (2013), The Vitamin B6 biosynthesis pathway in Streptococcus pneumoniae is controlled by pyridoxal 5′-phosphate and the transcription factor PdxR, and impacts on ear infection. *Journal of Bacteriology*, p. Published ahead of print
287. Belitsky, B. R. and Sonenshein, A. L. (2002), GabR, a member of a novel protein family, regulates the utilization of γ-aminobutyrate in *Bacillus subtilis*. *Molecular Microbiology***. 45**, p. 569-583
288. Belitsky, B. R. (2004), *Bacillus subtilis* GabR, a Protein with DNA-binding and Aminotransferase Domains, is a PLP-dependent Transcriptional Regulator. *Journal of Molecular Biology***. 340**, p. 655-664
289. O'Neill, P. M., Barton, V. E. and Ward, S. A. (2010), The Molecular Mechanism of Action of Artemisinin—The Debate Continues. *Molecules***. 15**, p. 1705-1721
290. Jansen, F. H. (2010), The pharmaceutical death-ride of dihydroartemisinin. *Malaria Journal***. 9**, p. 212
291. Zhang, K., Bangs, J. D. and Beverley, S. M. (2010), Sphingolipids in parasitic protozoa. *Adv Exp Med Biol***. 688**, p. 238-248
292. Hanada, K. (2003), Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochim Biophys Acta***. 1632**, p. 16-30
293. Furuya, H., Shimizu, Y. and Kawamori, T. (2011), Sphingolipids in cancer. *Cancer Metastasis Rev***. 30**, p. 567-576

294. Le Roch, K. G., Johnson, J. R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S. F., Williamson, K. C., Holder, A. A., Carucci, D. J., Yates, J. R. and Winzeler, E. A. (2004), Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Research***. 14**, p. 2308-2318
295. Holden, H. M., Thoden, J. B. and Raushel, F. M. (1999), Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product. *Cellular and Molecular Life Sciences***. 56**, p. 507-522
296. Raushel, F. M., Thoden, J. B. and Holden, H. M. (1999), The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia. *Biochemistry***. 38**, p. 7891-7899
297. Flores, M. V. C. and Stewart, T. S. (1998), *Plasmodium falciparum*: A Microassay for the Malarial Carbamoyl Phosphate Synthetase. *Experimental Parasitology***. 88**, p. 243-245
298. Flores, M. V. C., O'Sullivan, W. J. and Stewart, T. S. (1994), Characterisation of the carbamoyl phosphate synthetase gene from *Plasmodium falciparum*. *Molecular and Biochemical Parasitology***. 68**, p. 315-318
299. Biagini, G. A., O'Neill, P. M., Nzila, A., Ward, S. A. and Bray, P. G. (2003), Antimalarial chemotherapy: young guns or back to the future? *Trends in Parasitology***. 19**, p. 479-487
300. Eliot, A. C. and Kirsch, J. F. (2004), Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annual Review of Biochemistry***. 73**, p. 383-415
301. Dick, T., Manjunatha, U., Kappes, B. and Gengenbacher, M. (2010), Vitamin B6 biosynthesis is essential for survival and virulence of *Mycobacterium tuberculosis*. *Molecular Microbiology***. 78**, p. 980-988
302. Titiz, O., Tambasco-Studart, M., Warzych, E., Apel, K., Amrhein, N., Laloi, C. and Fitzpatrick, T. B. (2006), PDX1 is essential for vitamin B6 biosynthesis, development and stress tolerance in Arabidopsis. *The Plant Journal***. 48**, p. 933-946
303. Akana, J., Fedorov, A. A., Fedorov, E., Novak, W. R. P., Babbitt, P. C., Almo, S. C. and Gerlt, J. A. (2006), D-Ribulose 5-Phosphate 3-Epimerase: Functional and Structural Relationships to Members of the Ribulose-Phosphate Binding $(\beta/\alpha)_8$-Barrel Superfamily. *Biochemistry***. 45**, p. 2493-2503
304. Caruthers, J., Bosch, J., Buckner, F., Van Voorhis, W., Myler, P., Worthey, E., Mehlin, C., Boni, E., DeTitta, G., Luft, J., Lauricella, A., Kalyuzhniy, O., Anderson, L., Zucker, F., Soltis, M. and Hol, W. G. J. (2006), Structure of a ribulose 5-phosphate 3-epimerase from *Plasmodium falciparum*. *Proteins: Structure, Function, and Bioinformatics***. 62**, p. 338-342
305. Atamna, H., Pascarmona, G. and Ginsburg, H. (1994), Hexose-monophosphate shunt activity in intact *Plasmodium falciparum*-infected erythrocytes and in free parasites. *Molecular and Biochemical Parasitology***. 67**, p. 79-89
306. Roth Jr, E. F., Ruprecht, R. M., Schulman, S., Vanderberg, J. and Olson, J. A. (1986), Ribose metabolism and nucleic acid synthesis in normal and glucose-6-phosphate dehydrogenase-deficient human erythrocytes infected with *Plasmodium falciparum*. *Journal of Clinical Investigation***. 77**, p. 1129-1135
307. Wamelink, M. M. C., Struys, E. A., Huck, J. H. J., Roos, B., van der Knaap, M. S., Jakobs, C. and Verhoeven, N. M. (2005), Quantification of sugar phosphate intermediates of the pentose phosphate pathway by LC–MS/MS: application to two new inherited defects of metabolism. *Journal of Chromatography B***. 823**, p. 18-25

Appendix 2

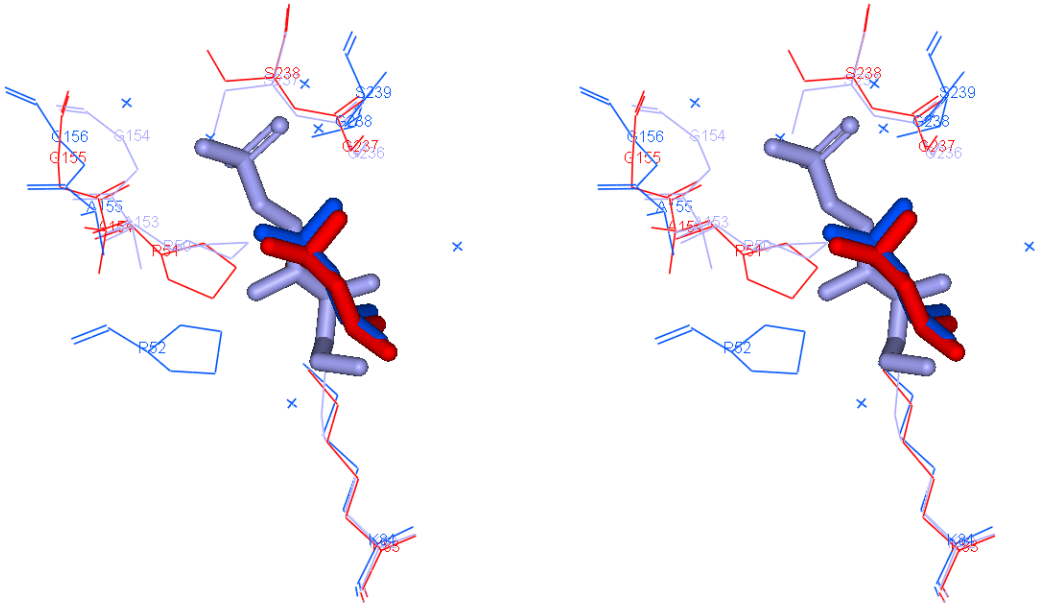


Figure A2.1: Stereoview of *Pf*Pdx1_MPD (red) superimposed with template structures *Tm*Pdx1 (2ISS, dark blue) and *Tt*Pdx1 (2ZBT, light blue). The proline residue (P52 in 2ZBT, P51 in 2ISS and P54 in *Pf*Pdx1) can be seen at the left of the ligand structure indicated as solid stick structure, and adopts a similar conformation to that of *Tm*Pdx1.

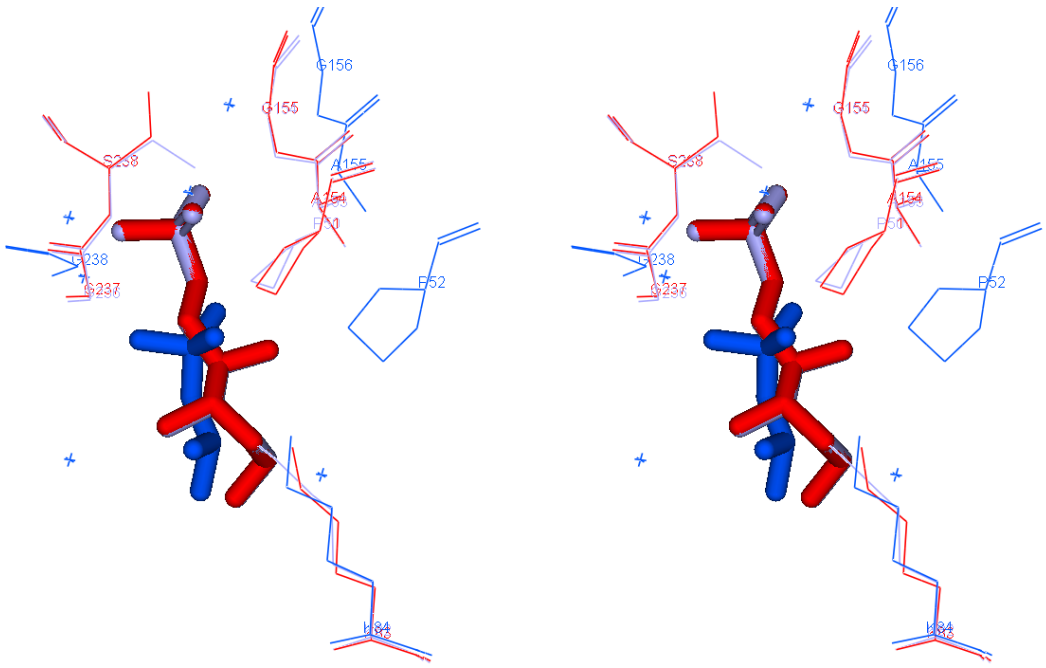


Figure A2.2: Stereoview of *Pf*Pdx1_Ru5P (red) superimposed with template structures *Tm*Pdx1 (2ISS, dark blue) and TtPdx1 (2ZBT, light blue). The proline residue (P52 in 2ZBT, P51 in 2ISS and *Pf*Pdx1) can be seen right of ligand structure indicated in hard solid line, and was comparable to that of *Tm*Pdx1. Similarly S237 was overlaid and more comparable to the *Tm*Pdx1 template.

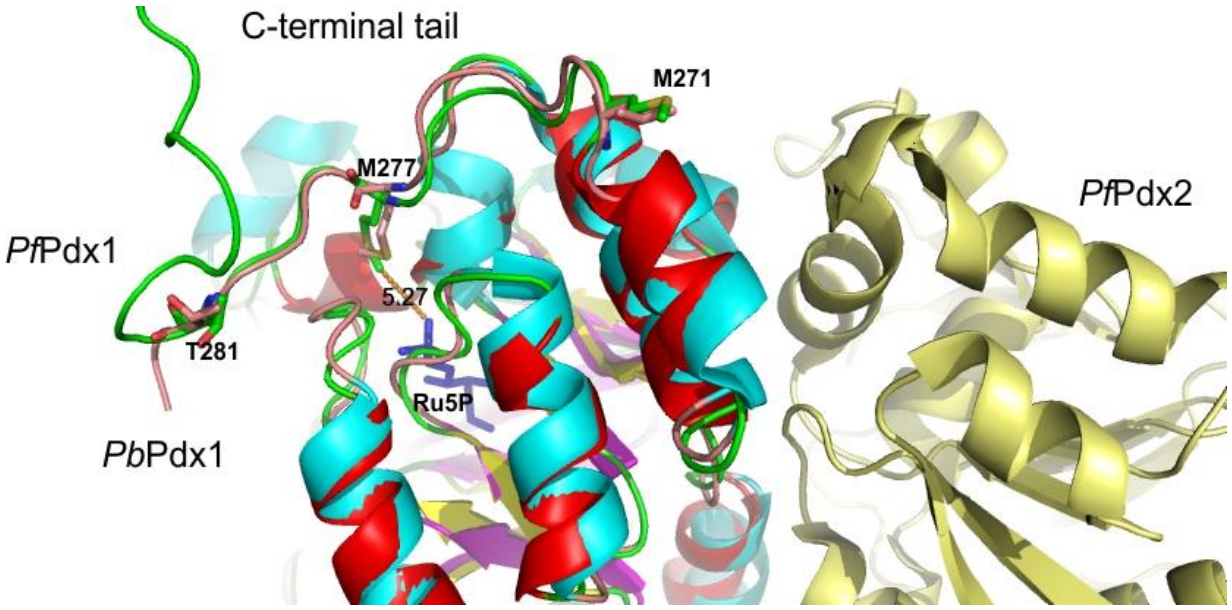


Figure A2.3: Superimposition of minimised *Pf*Pdx1 with *Pb*Pdx1 (PDB ID: 4ADS). The PbPdx1 component of the chimeric *Pb*Pdx1:*Pf*Pdx2 crystal structure was superimposed over the minimised *Pf*Pdx1_Ru5P model. Residues from M271 (*Pf*Pdx1 labelling) to T281 of *Pf*Pdx1 were aligned and spatially similar to those of *Pb*Pdx1. The M277 residue of *Pf*Pdx1 was located 5.27 Å from the phosphate group of the Ru5P substrate in the R5P-active site of *Pf*Pdx1.

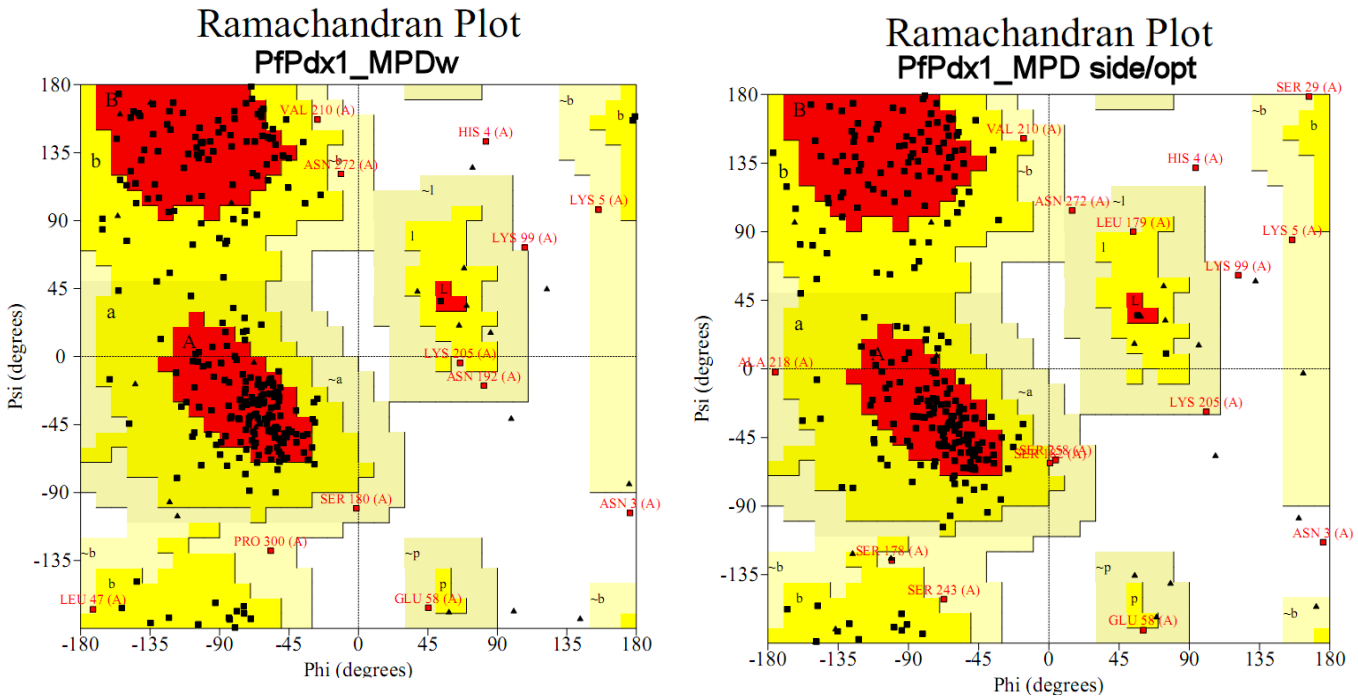


Figure A2.4: Ramachandran plots of *Pf*Pdx1 homology models after protein minimization.

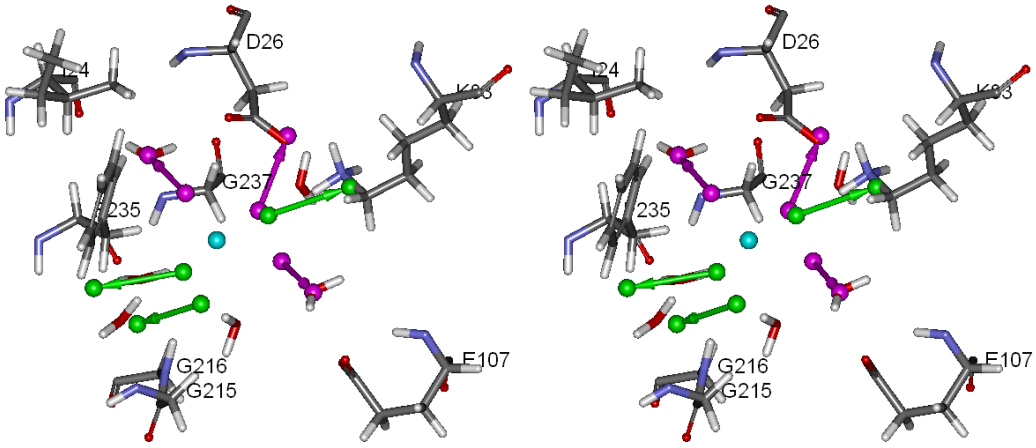


Figure A2.5: A stereo view of the *Pf*Pdx1_MPDw model pharmacophore, in which H-bond donors (green), H-bond acceptors (purple) and hydrophobic features (magenta) additionally involve water molecules, as listed in Table

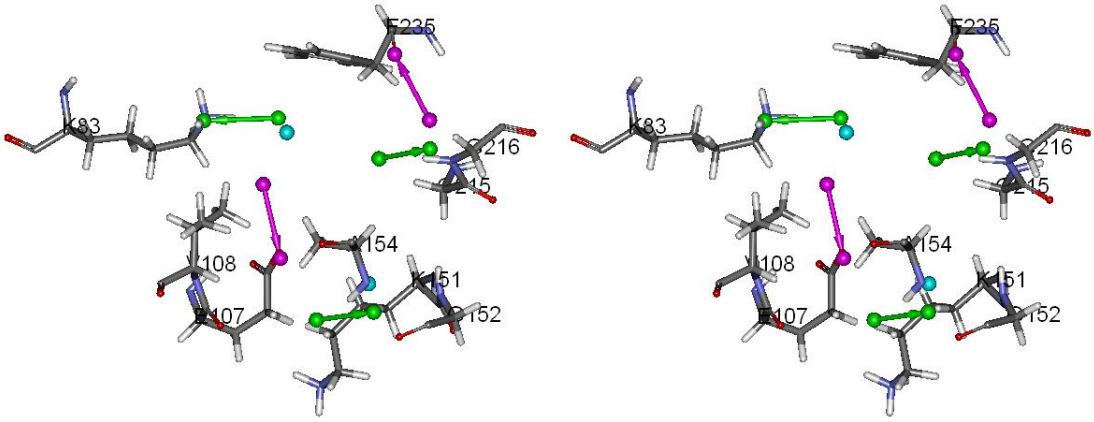


Figure A2.6: Representative stereo view of the *Pf*Pdx1_MPDsco pharmacophore.

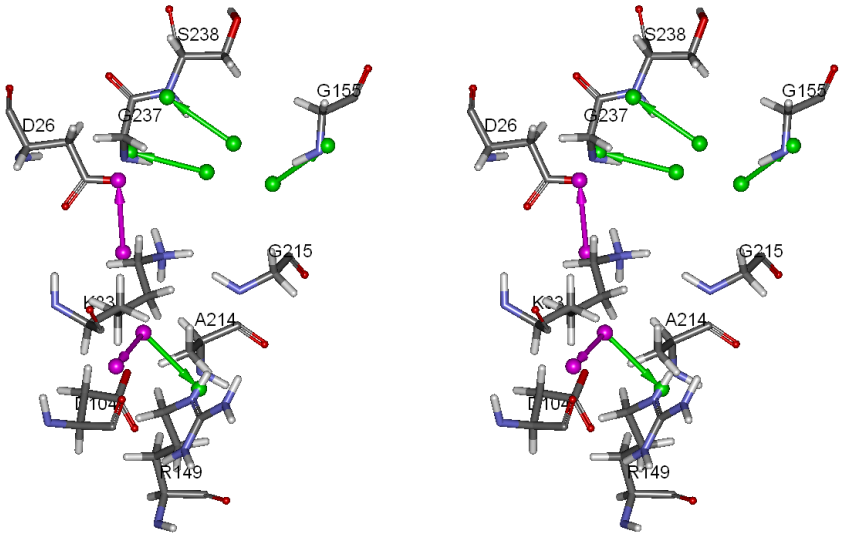


Figure A2.7: Stereo view of the pharmacophore from the *Pf*Pdx1_Ru5P model. Features were manually selected based on interactions between the R5P substrate (not shown in figure). This was termed a custom feature pharmacophore.

Appendix 3

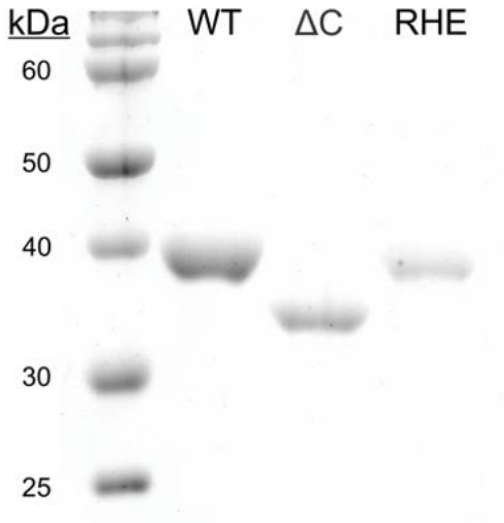


Figure A3.1: Coomassie stained SDS-polyacrylamide gel with 100 µg loaded each of *Pf*Pdx1 wildtype (WT), the C-terminal truncated form of *Pf*Pdx1 (ΔC) and the *Pf*Pdx1 RHE mutant. *Pf*Pdx1-ΔC was shown to be ~ 35 kDa compared to the wild-type (WT) *Pf*PDx1 at 37.5 kDa, confirming the 31 amino acid deletion.

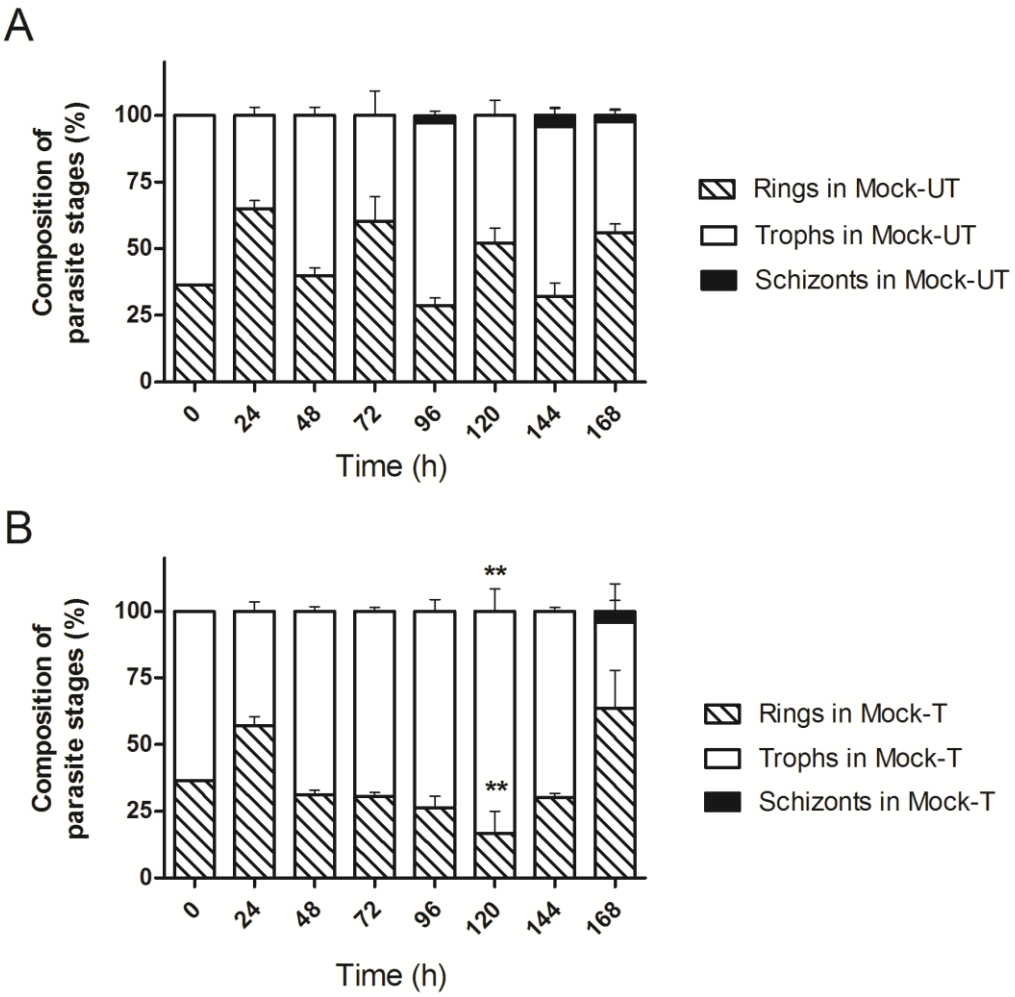


Figure A3.2: Morphological composition of parasite stages in mock control parasites during long-term exposure of 4PEHz. The parasite life-stage composition (percentage of rings, trophozoites or schizonts) of untreated **(A)** and treated **(B)** *P. falciparum* mock cells was determined microscopically during long term growth assays. Results represent average values from two independent experiments performed in triplicate, of which error bars indicate the SEM. Comparisons between treated and untreated mock parasites were made using an unpaired two-tailed Student's t-test in which ** represents statistical significance ($P < 0.05$). The treated mock parasites had significantly different ring and trophozoite life-stage compositions compared to the untreated controls parasites at 120 h.

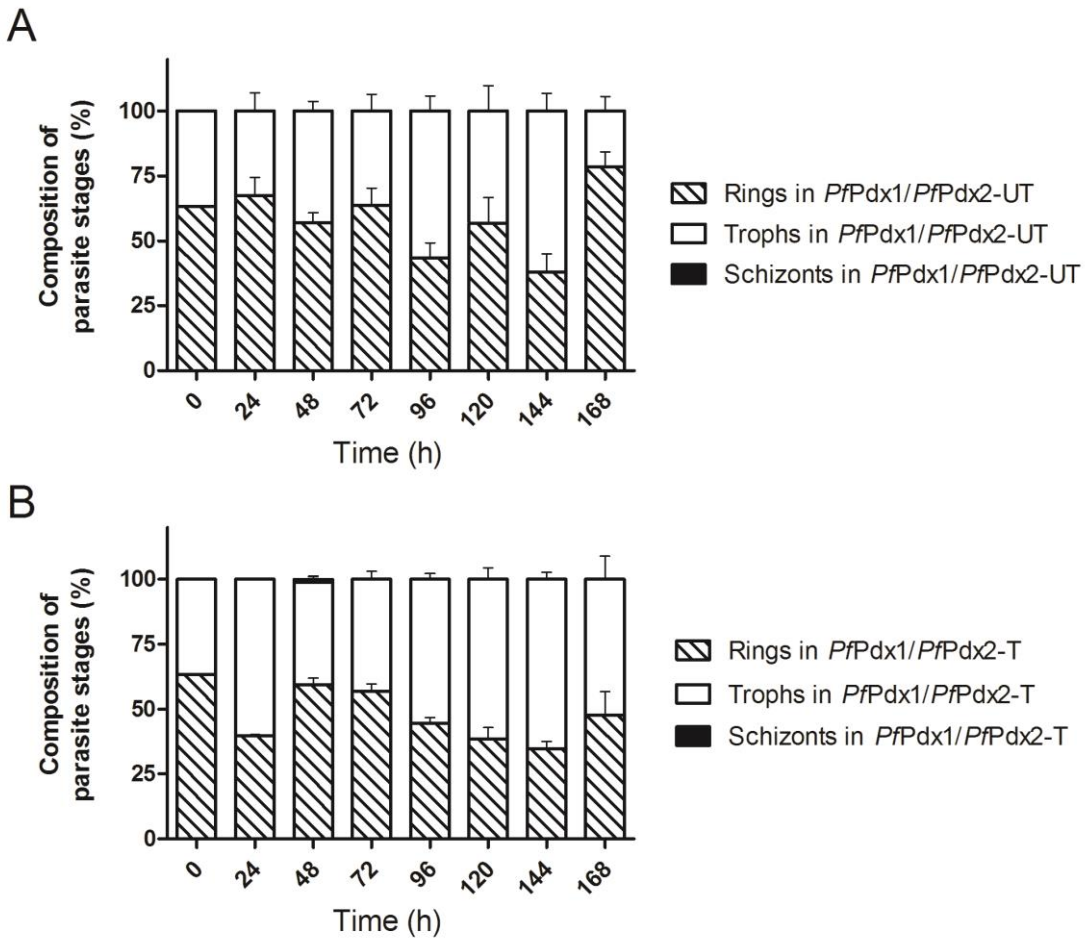


Figure A3.3: Morphological composition of parasite stages in *Pf*Pdx1/*Pf*Pdx2-complemented parasites during long-term exposure of 4PEHz. The parasite life-stage composition (percentage of rings, trophozoites or schizonts) of untreated **(A)** and treated **(B)** *P. falciparum Pf*Pdx1/*Pf*Pdx2-overexpressing parasites was determined microscopically during long term growth assays. Results represent average values from two independent experiments performed in triplicate, of which error bars indicate the SEM. Comparisons between treated and untreated mock parasites were made using an unpaired two-tailed Student's t-test in which ** represents statistical significance ($P < 0.05$). There were no significant differences in the life-stage compositions between untreated and treated *Pf*Pdx1/*Pf*Pdx2 parasites at any of the time points. This confirmed that complementation of PLP biosynthesis protected the parasites from the effects of 4PEHz, and the parasites grew normally compared to untreated cells.

Appendix 4

Table A4.1: Differentially expressed transcripts from t_3 and their associated metabolic processes

No	PlasmoDB ID	PlasmoDB description	log_2FC $Tt_3:UTt_3$	aj.P.Val	PlasmoDB annotated processes	Additional process involvement	GO term
Protein amino acid phosphorylation							
1	PF13_0211	calcium dependent protein kinase 5	-2.56	1.29E-07	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
2	PFI1285w	protein kinase, putative	-2.35	1.26E-06	protein amino acid phosphorylation		GO:0009536~plastid
3	MAL13P1.278	serine/threonine protein kinase, putative	-2.23	3.37E-07	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
4	PF11_0464	serine/threonine protein kinase, putative	-2.17	8.33E-08	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
5	PFB0665w	serine/threonine protein kinase, putative	-2.11	2.09E-05	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
6	PF11_0060	calcium/calmodulin-dependent protein kinase, putative	-2.09	7.20E-06	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
7	PFB0815w	calcium dependent protein kinase 1	-2.02	2.47E-07	protein amino acid phosphorylation		
8	PFL1885c	protein kinase 2	-1.91	7.20E-08	protein amino acid phosphorylation		
9	MAL13P1.114	conserved *Plasmodium* protein, unknown function	-1.80	4.53E-07	protein amino acid phosphorylation		GO:0009536~plastid
10	MAL7P1.18	serine/threonine protein kinase, putative	-1.66	6.20E-07	protein amino acid phosphorylation		
11	PFC0945w	protein kinase, putative	-1.58	5.34E-06	protein amino acid phosphorylation		
12	PFC0385c	serine/threonine protein kinase, putative	-1.55	8.77E-07	protein amino acid phosphorylation		
13	PFI0100c	serine/threonine protein kinase, FIKK family	-1.50	1.42E-06	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
14	PFL1110c	cAMP-dependent protein kinase regulatory subunit	-1.49	2.75E-07	regulation of protein amino acid phosphorylation		
15	PFD1165w	serine/threonine protein kinase, FIKK family	-1.28	8.62E-06	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
16	PF14_0346	cGMP-dependent protein kinase	-1.22	2.59E-05	protein amino acid phosphorylation	regulation of protein amino acid phosphorylation	
17	PFB0150c	protein kinase, putative	-1.16	1.51E-04	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
18	PFE0045c	serine/threonine protein kinase, FIKK family	-0.89	2.11E-03	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
19	PFL2280w	serine/threonine protein kinase, putative	-0.89	3.54E-04	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
20	PF14_0734	serine/threonine protein kinase, FIKK family	-0.89	1.42E-03	protein amino acid phosphorylation		
21	PFI0120c	serine/threonine protein kinase, FIKK family	-0.89	8.23E-05	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
22	PF07_0072	calcium dependent protein kinase 4	-0.89	8.53E-05	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
23	PF11_0281	protein phosphatase, putative	-0.86	9.81E-05	dephosphorylation		GO:0006793~phosphorus metabolic process
24	PFD1175w	serine/threonine protein kinase, FIKK family	-0.77	1.08E-04	protein amino acid phosphorylation		GO:0006468~protein amino acid

							phosphorylation
25	PFF0260w	serine/threonine protein kinase, putative	0.91	1.21E-03	protein amino acid phosphorylation		GO:0004674~protein serine/threonine kinase activity
Intracellular signaling pathway							
26	PFI1005w	ADP-ribosylation factor, putative	-2.07	8.59E-09	small GTPase mediated signal transduction		GO:0005525~GTP binding
27	MAL8P1.150	adenylyl cyclase beta, putative	-1.50	1.08E-07	cGMP biosynthetic process	intracellular signaling pathway	
28	PFI1485c	diacylglycerol kinase, putative	-1.39	8.62E-06	activation of protein kinase C activity by G-protein coupled receptor protein signaling pathway	intracellular signaling pathway	GO:0007166~cell surface receptor linked signal transduction
29	MAL13P1.118	3',5'-cyclic nucleotide phosphodiesterase, putative	-1.30	1.35E-05	signal transduction		GO:0004114~3'
30	MAL13P1.297	ADP-ribosylation factor, putative	-1.17	9.97E-05	protein amino acid ADP-ribosylation	small GTPase mediated signal transduction	GO:0006471~protein amino acid ADP-ribosylation
31	PF11_0107	conserved *Plasmodium* protein, unknown function	-1.00	1.91E-05	intracellular signaling pathway		
32	MAL7P1.108	phosphoinositide-binding protein, putative	-0.94	5.51E-05	cell communication		
33	PFL0475w	cGMP-specific phosphodiesterase	-0.85	1.22E-02	cGMP catabolic process	signal transduction	
34	PF11_0147	mitogen-activated protein kinase 2	-0.81	5.62E-05	MAPKKK cascade	protein amino acid phosphorylation	
35	PF10_0071	rhoGAP GTPase, putative	-0.76	5.36E-04	Rho protein signal transduction		GO:0007242~intracellular signaling cascade
36	MAL7P1.92	cysteine repeat modular protein 2	-0.76	9.18E-03	intracellular receptor mediated signaling pathway	intracellular transport	GO:0009536~plastid
37	MAL7P1.73	calcium/calmodulin-dependent protein kinase, putative	0.82	1.67E-02	protein amino acid phosphorylation	signal transduction	GO:0006468~protein amino acid phosphorylation
38	PFL2250c	RAC-beta serine/threonine protein kinase	0.96	6.69E-05	protein amino acid autophosphorylation		
39	PF10_0161	*Plasmodium* exported protein (PHISTc), unknown function	1.16	1.20E-04	small GTPase mediated signal transduction	translational initiation	
Regulation of cell cycle process							
40	PFD0900w	conserved *Plasmodium* protein, unknown function	-1.76	1.09E-04	cell cycle		
41	PFI1685w	cAMP-dependent protein kinase catalytic subunit	-1.67	4.76E-08	cell differentiation	protein amino acid phosphorylation	GO:0006468~protein amino acid phosphorylation
42	PFF0270c	cyclin dependent kinase binding protein, putative	0.79	1.61E-03	regulation of cell cycle	regulation of cell division	GO:0051726~regulation of cell cycle
43	PFI0810c	apicoplast Ufd1 precursor	0.79	1.08E-04	ER-associated protein catabolic process	regulation of cell cycle process	
44	PF14_0178	ubiquitin fusion degradation protein UFD1, putative	0.92	2.16E-04	ER-associated protein catabolic process	regulation of cell cycle process	
45	MAL13P1.279	protein kinase 5	0.92	4.35E-05	cell cycle	protein amino acid phosphorylation	GO:0006468~protein amino acid phosphorylation
46	PF14_0498	DER1-like protein	1.10	5.63E-05	ER-associated protein catabolic process	regulation of cell cycle process	
Apicoplast associated transcripts							
47	PF10_0138	conserved *Plasmodium* protein, unknown function	-3.69	2.92E-10	null		GO:0009536~plastid

48	MAL7P1.119	rhoptry-associated leucine zipper-like protein 1	-3.17	2.93E-09	null		GO:0009536~plastid
49	MAL7P1.141	conserved *Plasmodium* protein, unknown function	-3.16	9.73E-10	null		GO:0009536~plastid
50	MAL8P1.73	rhoptry neck protein 5, putative	-2.78	1.02E-09	null		GO:0009536~plastid
51	PF10_0119	conserved *Plasmodium* protein, unknown function	-2.53	7.16E-09	null		GO:0009536~plastid
52	PF14_0607	conserved *Plasmodium* membrane protein, unknown function	-2.23	4.83E-07	null		GO:0009536~plastid
53	PF10_0166	conserved *Plasmodium* protein, unknown function	-2.09	2.90E-07	null		GO:0009536~plastid
54	PFL2505c	conserved *Plasmodium* protein, unknown function	-2.07	4.83E-07	null		GO:0009536~plastid
55	PF13_0039	conserved *Plasmodium* protein, unknown function	-2.01	1.60E-07	null		GO:0009536~plastid
56	PF11_0304	conserved *Plasmodium* protein, unknown function	-1.93	5.83E-05	null		GO:0009536~plastid
57	PFE0575c	conserved *Plasmodium* protein, unknown function	-1.62	4.91E-07	null		GO:0009536~plastid
58	MAL13P1.49	conserved *Plasmodium* protein, unknown function	-1.57	1.48E-06	null		GO:0009536~plastid
59	PFL1835w	conserved *Plasmodium* protein, unknown function	-1.39	1.69E-06	null		GO:0009536~plastid
60	PF13_0125	conserved *Plasmodium* protein, unknown function	-1.32	8.67E-05	null		GO:0009536~plastid
61	PFL2410w	conserved *Plasmodium* membrane protein, unknown function	-1.24	1.79E-06	null		GO:0009536~plastid
62	PF08_0047	conserved *Plasmodium* protein, unknown function	-1.19	7.63E-07	null		GO:0009536~plastid
63	PFL0840c	conserved *Plasmodium* protein, unknown function	-1.14	9.24E-05	null		GO:0009536~plastid
64	PF11_0135	conserved *Plasmodium* protein, unknown function	-1.13	1.26E-03	null		GO:0009536~plastid
65	PFI1470c	leucine-rich repeat protein	-1.06	6.47E-06	null		GO:0009536~plastid
66	PFC0065c	alpha/beta hydrolase, putative	-1.01	5.04E-05	null		GO:0009536~plastid
67	MAL13P1.61	*Plasmodium* exported protein (hyp8), unknown function	-1.00	1.14E-03	null		GO:0009536~plastid
68	PF11_0231	conserved *Plasmodium* protein, unknown function	-0.99	1.63E-03	null		GO:0009536~plastid
69	PFB0110w	*Plasmodium* exported protein (hyp11), unknown function	-0.95	2.55E-02	null		GO:0009536~plastid
70	PFF0660w	conserved *Plasmodium* protein, unknown function	-0.92	8.97E-06	null		GO:0009536~plastid
71	PFL0875w	conserved *Plasmodium* protein, unknown function	-0.91	4.43E-04	null		GO:0009536~plastid
72	MAL13P1.70	conserved *Plasmodium* membrane protein, unknown function	-0.90	5.85E-05	null		GO:0009536~plastid
73	PF11_0470	conserved *Plasmodium* protein, unknown function	-0.88	5.10E-03	null		GO:0009536~plastid
74	MAL7P1.6	*Plasmodium* exported protein (hyp12), unknown	-0.79	3.77E-04	null		GO:0009536~plastid

		function					
75	PF14_0440	conserved *Plasmodium* membrane protein, unknown function	-0.75	5.40E-04	null		GO:0009536~plastid
76	PFE0215w	ATP-dependent helicase, putative	0.78	2.13E-03	null		GO:0009536~plastid
77	PFB0180w	5'-3' exonuclease, N-terminal resolvase-like domain, putative	0.80	8.80E-04	null		GO:0009536~plastid
78	PFB0395w	conserved *Plasmodium* protein, unknown function	0.83	1.50E-04	null		GO:0009536~plastid
79	PFL2330w	conserved *Plasmodium* protein, unknown function	0.89	1.08E-05	null		GO:0009536~plastid
80	MAL8P1.138	alpha/beta hydrolase, putative	0.89	1.13E-04	null		GO:0009536~plastid
81	PF11_0285	conserved *Plasmodium* protein, unknown function	0.89	2.09E-05	null		GO:0009536~plastid
82	PF14_0696	conserved *Plasmodium* protein, unknown function	0.92	7.11E-04	null		GO:0009536~plastid
83	PF11_0323	conserved *Plasmodium* protein, unknown function	0.94	1.81E-03	null		GO:0009536~plastid
84	MAL13P1.87	sec20 homolog, putative	0.99	2.36E-05	null		GO:0009536~plastid
85	PFL0700w	conserved *Plasmodium* protein, unknown function	1.05	1.34E-05	null		GO:0009536~plastid
86	PFF1090c	conserved *Plasmodium* membrane protein, unknown function	1.07	3.41E-05	null		GO:0009536~plastid
87	PFI0205w	conserved *Plasmodium* protein, unknown function	1.27	2.91E-04	null		GO:0009536~plastid
88	MAL13P1.342	conserved *Plasmodium* protein, unknown function	1.35	2.58E-06	null		GO:0009536~plastid
89	MAL13P1.254	conserved *Plasmodium* protein, unknown function	1.47	2.76E-05	null		GO:0009536~plastid
90	PF14_0155	serine C-palmitoyltransferase, putative	1.48	2.66E-05	biosynthetic process		GO:0009536~plastid
Pathogenesis and invasion							
91	MAL7P1.176	erythrocyte binding antigen-175	-4.93	2.99E-12	entry into host cell	pathogenesis	GO:0009405~pathogenesis
92	PF10_0346	merozoite surface protein 6	-3.33	6.50E-10	entry into host cell		
93	MAL13P1.176	reticulocyte binding protein 2 homologue b	-3.33	5.53E-09	entry into host cell		GO:0030260~entry into host cell
94	PFE0080c	rhoptry-associated protein 2	-3.03	2.99E-10	entry into host cell		
95	PF13_0198	reticulocyte binding protein 2 homologue a	-2.72	2.06E-06	entry into host cell		GO:0007155~cell adhesion
96	PFB0680w	rhoptry neck protein 6	-2.36	6.27E-08	entry into host		
97	PF10_0281	merozoite TRAP-like protein	-1.60	1.07E-06	entry into host cell		
98	PFF1420w	phosphatidylcholine-sterol acyltransferase precursor, putative	-1.60	4.83E-07	entry into host cell	lipid metabolic process	GO:0009536~plastid
99	PF11_0344	apical membrane antigen 1	-1.30	3.74E-06	entry into host cell	pathogenesis	
100	PF10_0345	merozoite surface protein 3	-0.92	2.48E-04	entry into host cell		
101	PF13_0197	merozoite surface protein 7 precursor	-0.87	2.77E-04	entry into host cell		
Transcription							
102	PF10_0357	probable protein, unknown function	-2.19	8.86E-09	transcription		
103	PF11_0442	transcription factor with AP2 domain(s), putative	-2.11	2.13E-08	regulation of transcription	DNA-dependent	
104	PFF0200c	transcription factor with AP2 domain(s), putative,SPE2-interacting protein	-1.89	7.63E-07	regulation of transcription	DNA-dependent	
105	PF11_0091	transcription factor with AP2 domain(s), putative	-1.74	4.76E-08	regulation of transcription	DNA-dependent	GO:0006355~regulation of transcription

106	MAL7P1.125	conserved *Plasmodium* protein, unknown function	-1.56	4.63E-07	regulation of transcription	DNA-dependent	
107	PF11_0477	CCAAT-box DNA binding protein subunit B	-1.36	6.13E-07	regulation of transcription	DNA-dependent	GO:0006355~regulation of transcription
108	PFD0985w	transcription factor with AP2 domain(s), putative	-1.35	9.88E-07	regulation of transcription	DNA-dependent	
109	PF07_0052	conserved *Plasmodium* protein, unknown function	-1.34	5.53E-05	regulation of transcription	DNA-dependent	
110	PFL1510c	conserved *Plasmodium* protein, unknown function	-1.31	1.50E-04	regulation of transcription	DNA-dependent	
111	PFD0380c	conserved *Plasmodium* protein, unknown function	-1.25	8.18E-06	regulation of transcription	DNA-dependent	
112	PF11_0204	conserved *Plasmodium* protein, unknown function	-1.15	2.23E-05	regulation of transcription	DNA-dependent	GO:0006355~regulation of transcription
113	PF07_0126	transcription factor with AP2 domain(s), putative	-1.07	1.72E-06	regulation of transcription	DNA-dependent	GO:0006355~regulation of transcription
114	PFD0485w	conserved *Plasmodium* protein, unknown function	-0.99	1.71E-05	regulation of transcription	DNA-dependent	GO:0043167~ion binding
115	PFD0535w	conserved *Plasmodium* protein, unknown function	-0.91	3.09E-05	regulation of transcription	DNA-dependent	
116	PFF0670w	transcription factor with AP2 domain(s), putative	-0.84	5.14E-05	regulation of transcription	DNA-dependent	
117	PFL0465c	zinc finger transcription factor (krox1)	-0.77	6.67E-03	regulation of transcription	DNA-dependent	GO:0006355~regulation of transcription
118	MAL7P1.78	conserved *Plasmodium* protein, unknown function	0.80	2.35E-04	regulation of transcription	DNA-dependent	GO:0006350~transcription
119	PF14_0665	conserved *Plasmodium* protein, unknown function	0.89	1.32E-04	regulation of transcription	DNA-dependent	
120	PF11_0163	transcription factor with AP2 domain(s), putative	1.05	8.98E-06	regulation of transcription	DNA-dependent	
121	PFI1630c	conserved *Plasmodium* protein, unknown function	1.18	1.14E-05	transcription		
122	PF14_0580	conserved *Plasmodium* protein, unknown function	1.20	1.21E-03	regulation of transcription	DNA-dependent	
Translation							
123	PFI0540w	conserved *Plasmodium* protein, unknown function	-3.34	1.37E-09	translation	translational initiation	
124	PF13_0173	conserved *Plasmodium* protein, unknown function	-2.64	4.85E-09	translational initiation		
125	PFD1100c	conserved *Plasmodium* protein, unknown function	-2.28	6.22E-07	translation	translational initiation	GO:0009536~plastid
126	PFI1675w	conserved *Plasmodium* protein, unknown function	-1.96	1.63E-07	translation		
127	PFL0570c	mitochondrial ribosomal protein S18 precursor, putative	-1.95	9.22E-07	translation		GO:0006412~translation
128	PFF0705c	conserved *Plasmodium* protein, unknown function	-1.66	3.17E-04	translation	translational initiation	
129	PFF1095w	leucyl tRNA synthase	-1.63	9.61E-07	leucyl-tRNA aminoacylation	translation	GO:0006399~tRNA metabolic process
130	MAL7P1.81	eukaryotic translation initiation factor 3 37.28 kDa subunit, putative	-1.12	1.38E-04	regulation of translational initiation		GO:0006417~regulation of translation
131	PFI1725w	*Plasmodium* exported protein, unknown function	-1.06	1.20E-05	translation	valyl-tRNA aminoacylation	
132	PFD0525w	conserved *Plasmodium* protein, unknown function	-0.94	2.00E-05	regulation of translation		
133	PFI0645w	elongation factor 1-beta	-0.89	8.60E-05	translational elongation		GO:0006412~translation
134	MAL7P1.17	conserved *Plasmodium* membrane protein,	-0.82	4.12E-05	translation	translational initiation	

		unknown function					
135	PFL2180w	mitochondrial ribosomal protein L3 precursor, putative	0.76	8.55E-05	translation		GO:0006412~translation
136	PFL2430c	eukaryotic translation initiation factor 2b, subunit 2, putative	0.80	5.59E-04	cellular metabolic process	translational initiation	GO:0006412~translation
137	MAL13P1.164	elongation factor Tu, putative	0.81	6.83E-05	translational elongation		GO:0006412~translation
138	PF13_0069	translation initiation factor IF-2, putative	0.82	1.10E-03	translation	translational initiation	GO:0006412~translation
139	PFA0480w	phenylalanyl-tRNA synthetase, putative	0.85	6.71E-04	phenylalanyl-tRNA aminoacylation	translation	
140	PFB0645c	mitochondrial ribosomal protein L13 precursor, putative	0.85	4.55E-05	translation		GO:0006412~translation
141	PF14_0606	mitochondrial ribosomal protein S6-2 precursor, putative	0.86	1.26E-02	translation		GO:0006412~translation
142	PFL1895w	mitochondrial ribosomal protein L23 precursor, putative	0.86	1.96E-04	translation		GO:0006412~translation
143	PF11_0182	conserved *Plasmodium* protein, unknown function	0.86	7.18E-05	translation	translational termination	
144	PF14_0289	mitochondrial ribosomal protein L17-2 precursor, putative	0.87	1.60E-05	translation		GO:0006412~translation
145	PFC0241w	conserved *Plasmodium* protein, unknown function	0.89	1.26E-04	translation		
146	PF13_0205	tryptophan--tRNA ligase, putative	0.89	8.20E-05	translation	tryptophanyl-tRNA aminoacylation	GO:0006399~tRNA metabolic process
147	PFI1645c	histidyl-tRNA synthetase, putative	0.93	3.74E-05	histidyl-tRNA aminoacylation	translation	GO:0006399~tRNA metabolic process
148	PFI1120c	conserved *Plasmodium* protein, unknown function	0.99	1.00E-04	translation	translational initiation	
149	PFL1910c	conserved *Plasmodium* protein, unknown function	2.05	1.96E-06	translation		
Antigenic variation, defense response and pathogenesis							
150	MAL13P1.60	erythrocyte binding antigen-140	-4.74	2.89E-11	cell-cell adhesion	pathogenesis	GO:0009405~pathogenesis
151	PFD0110w	normocyte binding protein 1,reticulocyte binding protein homologue 1	-3.72	2.53E-09	cell-cell adhesion		
152	PF10_0003	rifin	-3.70	6.50E-10	antigenic variation		GO:0006952~defense response
153	PFA0125c	erythrocyte binding antigen-181	-3.05	1.48E-08	cell-cell adhesion	entry into host cell	
154	PFD1155w	erythrocyte binding antigen-165	-3.00	2.56E-08	pathogenesis		GO:0009405~pathogenesis
155	MAL7P1.208	rhoptry-associated membrane antigen	-2.95	2.31E-08	attachment of GPI anchor to protein	cell-cell adhesion	
156	PFD0955w	apical merozoite protein	-2.79	6.50E-10	attachment of GPI anchor to protein		
157	PFC0120w	cytoadherence linked asexual protein 3.1	-2.70	2.56E-08	cytoadherence to microvasculature	mediated by parasite protein	
158	PFC0110w	cytoadherence linked asexual protein 3.2	-2.56	6.69E-08	cytoadherence to microvasculature	mediated by parasite protein	
159	PF08_0008	GPI-anchored micronemal antigen	-2.47	2.59E-06	attachment of GPI anchor to protein		
160	PFD0295c	apical sushi protein	-2.47	1.00E-07	attachment of GPI anchor to protein		
161	PFD1105w	asparagine-rich protein	-2.19	9.66E-07	cell-cell adhesion		
162	PF11_0373	conserved *Plasmodium* protein, unknown function	-2.17	4.02E-08	attachment of GPI anchor to protein		
163	PFF1545w	rifin	-2.09	4.93E-04	antigenic variation		
164	PF14_0325	conserved *Plasmodium* membrane protein, unknown function	-2.01	1.13E-04	attachment of GPI anchor to protein		
165	PFD1145c	reticulocyte binding protein homologue 5	-1.82	4.76E-08	cell-cell adhesion	entry into host	
166	PF14_0293	conserved *Plasmodium* protein, unknown function	-1.80	3.68E-07	attachment of GPI anchor to protein		

167	PFI1730w	cytoadherence linked asexual protein 9	-1.69	5.10E-08	cell-cell adhesion		
168	PFI1475w	merozoite surface protein 1	-1.65	3.72E-07	attachment of GPI anchor to protein	pathogenesis	GO:0009405~pathogenesis
169	PF10_0348	duffy binding-like merozoite surface protein	-1.46	4.09E-06	cell-cell adhesion	pathogenesis	GO:0009405~pathogenesis
170	PFF0615c	6-cysteine protein	-1.35	5.69E-07	attachment of GPI anchor to protein		
171	PFL1960w	erythrocyte membrane protein 1, PfEMP1	-1.32	1.90E-05	antigenic variation	cell-cell adhesion	
172	PF11_0008	erythrocyte membrane protein 1, PfEMP1	-1.23	2.58E-06	antigenic variation	cell-cell adhesion	GO:0006952~defense response
173	PFL0030c	erythrocyte membrane protein 1, PfEMP1	-1.12	1.42E-04	antigenic variation	cell-cell adhesion	
174	PF07_0132	rifin	-1.12	3.26E-03	antigenic variation		
175	PFB0300c	merozoite surface protein 2	-1.04	4.91E-05	attachment of GPI anchor to protein	cell adhesion	
176	PFB0570w	secreted protein altered thrombospondin repeat protein	-1.01	6.37E-06	cell-cell adhesion		GO:0009536~plastid
177	PFF0380w	conserved *Plasmodium* protein, unknown function	-1.01	3.51E-04	cell adhesion	phosphate transport	
178	PFF0995c	merozoite surface protein 10	-0.98	5.10E-04	attachment of GPI anchor to protein		
179	PFB0010w	erythrocyte membrane protein 1, PfEMP1	-0.92	7.65E-04	antigenic variation	cell-cell adhesion	GO:0006952~defense response
180	PF10_0404	rifin	-0.87	7.48E-05	antigenic variation		GO:0006952~defense response
181	PFL1970w	erythrocyte membrane protein 1, PfEMP1	-0.86	4.48E-03	pathogenesis		
182	PF14_0723	LCCL domain-containing protein	-0.82	3.58E-04	cell adhesion	signal transduction	
183	PF08_0142	erythrocyte membrane protein 1, PfEMP1	-0.81	1.25E-03	antigenic variation	cell-cell adhesion	GO:0009405~pathogenesis
184	PFA0005w	erythrocyte membrane protein 1, PfEMP1	-0.78	5.37E-03	antigenic variation	cell-cell adhesion	
185	PF11_0229	conserved *Plasmodium* protein, unknown function	0.87	1.54E-03	GPI anchor biosynthetic process	attachment of GPI anchor to protein	
186	PFF0975c	conserved *Plasmodium* protein, unknown function	0.88	1.10E-02	attachment of GPI anchor to protein		
187	MAL13P1.348	glycosylphosphatidylinositol anchor attachment 1 protein, putative	0.89	2.21E-04	GPI anchor biosynthetic process	attachment of GPI anchor to protein	
188	PFB0575c	conserved *Plasmodium* protein, unknown function	0.91	1.78E-04	cell adhesion	phosphate transport	
189	PFE1630w	rifin	0.95	2.23E-02	antigenic variation		
190	MAL7P1.50	erythrocyte membrane protein 1, PfEMP1	0.97	2.62E-04	antigenic variation	cell-cell adhesion	GO:0009405~pathogenesis
191	PFI0535w	phosphatidylinositol N-acetylglucosaminyltransferase, putative	0.99	7.77E-05	GPI anchor biosynthetic process		
Cytokinesis							
192	PF10_0039	membrane skeletal protein IMC1-related	-3.75	1.71E-10	cytoskeleton organization		
193	PFL2460w	coronin	-3.52	1.71E-10	cellular component movement	cytokinesis	GO:0000910~cytokinesis
194	PFL1435c	myosin D	-3.46	2.99E-10	actin filament organization	actin filament-based movement	GO:0007010~cytoskeleton organization
195	PF10_0368	dynamin-like protein	-1.88	8.61E-08	microtubule-based process	receptor-mediated endocytosis	GO:0006897~endocytosis
196	PF14_0626	dynein beta chain, putative	-1.35	3.33E-05	microtubule-based movement		GO:0007017~microtubule-based process
197	PFL2215w	actin I	-1.30	7.85E-07	cytoskeleton organization		
198	PFA0260c	cyclase-associated protein, putative	-1.21	5.19E-05	cytoskeleton organization		
199	PFL2165w	kinesin-like protein, putative	0.90	9.96E-05	microtubule-based movement		GO:0007017~microtubule-based process
200	PFE1420w	f-actin capping protein alpha subunit, putative	1.15	3.41E-05	actin cytoskeleton organization		GO:0007010~cytoskeleton organization
201	PFL0115w	dynein heavy chain, putative	1.51	1.42E-03	microtubule-based movement		GO:0007017~microtubule-based process
202	MAL7P1.162	dynein heavy chain, putative	1.81	1.29E-07	microtubule-based movement		GO:0007017~microtubule-based process
203	PFD1050w	alpha tubulin 2	2.71	2.68E-09	microtubule-based movement	protein polymerization	GO:0006461~protein complex assembly
DNA repair and replication							

204	PF14_0527	conserved *Plasmodium* protein, unknown function	-3.16	4.22E-09	DNA recombination	mismatch repair	
205	PF11_0383	conserved *Plasmodium* protein, unknown function	-1.68	6.58E-06	DNA-dependent DNA replication		
206	PFE0420c	guanidine nucleotide exchange factor	-1.49	6.09E-07	DNA packaging		GO:0006323~DNA packaging
207	MAL13P1.112	conserved *Plasmodium* protein, unknown function	-1.01	2.74E-04	DNA integration	DNA recombination	
208	PFI0510c	conserved *Plasmodium* protein, unknown function	-0.76	6.83E-05	DNA repair	DNA-dependent DNA replication	GO:0006259~DNA metabolic process
209	PFF1470c	DNA polymerase epsilon, catalytic subunit a, putative	0.77	2.87E-04	DNA-dependent DNA replication		GO:0006259~DNA metabolic process
210	PFE1345c	minchromosome maintenance (MCM) complex subunit, putative	0.79	1.59E-02	DNA replication	DNA-dependent DNA replication initiation	GO:0006259~DNA metabolic process
211	PF14_0366	small subunit DNA primase	0.87	4.41E-05	DNA replication	synthesis of RNA primer	GO:0006259~DNA metabolic process
212	PF11_0087	Rad51 homolog	0.89	6.81E-05	DNA recombination	DNA repair	
213	PFL2005w	replication factor C subunit 4	0.92	9.38E-05	DNA replication		GO:0006259~DNA metabolic process
214	PFI0530c	DNA primase large subunit, putative	0.93	6.71E-04	DNA replication	synthesis of RNA primer	
215	PFL0150w	origin recognition complex subunit 1	0.99	1.01E-04	DNA-dependent DNA replication initiation		
216	MAL13P1.191	conserved *Plasmodium* protein, unknown function	1.01	8.69E-04	DNA repair		
217	MAL13P1.346	DNA repair endonuclease, putative	1.09	3.22E-05	DNA repair		GO:0006259~DNA metabolic process
Fatty acid, lipid and phospholipid metabolism							
218	PF07_0005	lysophospholipase, putative	-1.57	1.71E-06	phospholipid metabolic process		GO:0006644~phospholipid metabolic process
219	PFI1015w	conserved *Plasmodium* protein, unknown function	-1.39	2.57E-06	lipid transport		
220	PFI1370c	phosphatidylserine decarboxylase	-0.94	1.09E-04	phospholipid biosynthetic process		GO:0006644~phospholipid metabolic process
221	PFB0695c	acyl-CoA synthetase, PfACS8	-0.91	2.00E-04	fatty acid metabolic process	long-chain fatty acid transport	
222	PFF0290w	long chain polyunsaturated fatty acid elongation enzyme, putative	-0.90	3.53E-04	fatty acid elongation	unsaturated fatty acid	GO:0006631~fatty acid metabolic process
223	PF13_0078	conserved *Plasmodium* membrane protein, unknown function	-0.86	1.81E-03	lipid catabolic process		
224	PFE0485w	phosphatidylinositol 4-kinase	-0.85	5.49E-05	phosphoinositide phosphorylation	phosphoinositide-mediated signaling	GO:0006644~phospholipid metabolic process
225	PF14_0020	choline kinase	-0.81	3.11E-05	lipid metabolic process	phosphatidylcholine biosynthetic process	
226	PF10_0015	acyl-CoA binding protein, isoform 1, ACBP1	-0.79	1.26E-02	fatty acid metabolic process		
227	PFL1055c	conserved *Plasmodium* membrane protein, unknown function	-0.78	9.58E-04	lipid catabolic process		
228	PFB0505c	beta-ketoacyl-ACP synthase III	-0.77	5.90E-04	fatty acid biosynthetic process		
229	PFI1180w	patatin-like phospholipase, putative	-0.76	7.06E-05	lipid metabolic process		GO:0016042~lipid catabolic process
230	PF14_0664	biotin carboxylase subunit of acetyl CoA carboxylase, putative	0.80	1.34E-03	fatty acid biosynthetic process		
231	PFD1035w	steroid dehydrogenase, putative	0.82	2.52E-04	fatty acid biosynthetic process		GO:0006631~fatty acid metabolic process
232	PF14_0441	pyruvate dehydrogenase E1 beta subunit	0.91	1.07E-04	acetyl-CoA biosynthetic process from pyruvate	fatty acid biosynthetic process	GO:0009536~plastid

233	PFB0685c	acyl-CoA synthetase, PfACS9	1.07	3.47E-04	fatty acid metabolic process	long-chain fatty acid transport	
234	PF14_0123	conserved *Plasmodium* protein, unknown function	1.12	1.46E-06	phosphatidylinositol metabolic process		GO:0006644~phospholipid metabolic process
235	PF13_0066	malonyl CoA-acyl carrier protein transacylase precursor	1.19	1.76E-02	fatty acid biosynthetic process		GO:0004314~[acyl-carrier-protein] S-malonyltransferase activity
236	PF11_0256	pyruvate dehydrogenase E1 alpha subunit	1.65	4.28E-08	acetyl-CoA biosynthetic process from pyruvate	fatty acid biosynthetic process	GO:0006084~acetyl-CoA metabolic process
Stress response							
237	PF13_0019	sodium/hydrogen exchanger, Na , H antiporter	-0.93	2.10E-05	response to drug	sodium ion transport	GO:0006811~ion transport
238	PFB0320c	iron-sulfur assembly protein, putative	-0.79	1.57E-04	iron-sulfur cluster assembly		GO:0016226~iron-sulfur cluster assembly
239	PF11_0169	pyridoxine biosynthesis protein Pdx2	0.77	5.98E-05	response to singlet oxygen	vitamin B6 biosynthetic process	
240	PF11_0044	iron-sulfur assembly protein, sufD, putative	0.92	6.29E-05	iron-sulfur cluster assembly		GO:0016226~iron-sulfur cluster assembly
241	PF14_0368	thioredoxin peroxidase 1	0.95	1.14E-04	DNA protection	cell redox homeostasis	
242	PF14_0186	conserved *Plasmodium* protein, unknown function	0.97	4.74E-03	cell redox homeostasis		
243	PF13_0021	small heat shock protein, putative	1.09	4.19E-05	response to heat	response to unfolded protein	GO:0006986~response to unfolded protein
244	MAL13P1.100	conserved *Plasmodium* protein, unknown function	1.11	1.01E-05	cell redox homeostasis		
245	PF08_0066	lipoamide dehydrogenase	1.28	8.55E-05	acetyl-CoA biosynthetic process from pyruvate	cell redox homeostasis	GO:0006084~acetyl-CoA metabolic process
Transport							
246	PF10_0262	conserved *Plasmodium* protein, unknown function	-3.35	6.38E-10	phosphate transport		
247	PF10_0059	conserved *Plasmodium* membrane protein, unknown function	-2.39	1.39E-08	transport		
248	PFL0245w	probable protein, unknown function	-1.59	4.29E-06	transport		GO:0009536~plastid
249	PFD1110w	glideosome associated protein with multiple membrane spans 2	-1.47	9.02E-07	transport		
250	PFL1700c	V-type K -independent H -translocating inorganic pyrophosphatase	-1.47	1.38E-07	proton transport		
251	PF13_0271	ABC transporter, (heavy metal transporter family), putative	-1.23	1.91E-06	transport		GO:0009536~plastid
252	PF14_0260	metabolite/drug transporter, putative	-1.19	1.14E-04	transport		
253	PFI0270w	conserved *Plasmodium* protein, unknown function	-1.15	1.78E-06	transport		
254	PFI1560c	conserved *Plasmodium* membrane protein, unknown function	-1.04	1.70E-04	transport		
255	PF10_0129	conserved *Plasmodium* protein, unknown function	-1.03	1.14E-05	phosphate transport		
256	PF11_0287	CRAL/TRIO domain-containing protein, putative	-0.95	4.31E-05	transport		
257	PF13_0238	kelch protein, putative	-0.81	4.59E-05	potassium ion transport		GO:0006811~ion transport
258	PF11_0340	conserved *Plasmodium* protein, unknown function	-0.77	2.10E-03	potassium ion transport		

259	PFD0725c	arsenical pump-driving ATPase, putative	0.80	1.14E-04	cellular metal ion homeostasis	ion transport	GO:0006873~cellular ion homeostasis
260	PF08_0097	conserved *Plasmodium* membrane protein, unknown function	0.82	3.93E-05	transport		
261	PF14_0133	SufC ATPase, putative	0.88	3.50E-05	transport		GO:0009536~plastid
262	PFF0170w	cation/H antiporter	0.90	3.83E-05	calcium ion transport		GO:0006811~ion transport
263	PF10_0216	zinc transporter, putative	1.03	1.70E-04	ion transport	transport	
264	PFI0400c	conserved *Plasmodium* membrane protein, unknown function	1.20	1.99E-06	ion transport		
Metabolic processes							
265	PFL0035c	acyl-CoA synthetase, PfACS7	-1.89	2.32E-07	metabolic process		
266	PF11_0190	haloacid dehalogenase-like hydrolase, putative	-1.13	8.38E-05	metabolic process		GO:0016791~phosphatase activity
267	PFC0260w	P-loop containing nucleoside triphospahte hydrolase, putative	-0.88	7.55E-04	metabolic process		
268	PF10_0147	FAD synthetase, putative	-0.79	7.18E-04	metabolic process		GO:0016779~nucleotidyltransferase activity
269	PF10_0036	N-acetyltransferase, putative	-0.77	8.59E-04	metabolic process		GO:0016407~acetyltransferase activity
270	PF11_0145	glyoxalase I	0.76	2.74E-04	methylglyoxal metabolic process		
271	MAL13P1.186	1-deoxy-D-xylulose 5-phosphate synthase	0.77	6.06E-05	metabolic process		GO:0016744~transferase activity
272	PFE1510c	triose phosphate transporter	0.80	2.67E-03	glucose import		GO:0009536~plastid
273	PF14_0060	glycerophodiester phosphodiesterase, putative	0.82	3.78E-04	glycerol metabolic process	intracellular signaling pathway	GO:0006071~glycerol metabolic process
274	PF13_0131	acetyltransferase, GNAT family, putative	0.86	2.38E-04	metabolic process		GO:0016407~acetyltransferase activity
275	PF10_0210	deoxyribose-phosphate aldolase, putative	0.86	1.01E-05	deoxyribonucleotide catabolic process	pentose-phosphate shunt	GO:0005996~monosaccharide metabolic process
276	PF11_0311	N-acetyl glucosamine phosphate mutase, putative	0.88	6.49E-05	carbohydrate metabolic process	glucosamine metabolic process	GO:0005996~monosaccharide metabolic process
277	PF14_0286	glutamate dehydrogenase, putative	0.90	2.92E-04	cellular amino acid metabolic process	oxidation reduction	GO:0055114~oxidation reduction
278	PF14_0357	succinyl CoA ligase, putative	0.91	3.26E-03	glycolysis	tricarboxylic acid cycle	GO:0005996~monosaccharide metabolic process
279	PFF1265w	oxidoreductase, short-chain dehydrogenase family, putative	0.91	1.12E-05	metabolic process		GO:0055114~oxidation reduction
280	PF14_0295	ATP-specific succinyl-CoA synthetase beta subunit, putative	0.92	6.79E-05	succinyl-CoA metabolic process	tricarboxylic acid cycle	GO:0006084~acetyl-CoA metabolic process
281	MAL13P1.319	indole-3-glycerol-phosphate synthase, putative	0.92	1.03E-03	metabolic process		GO:0016831~carboxy-lyase activity
282	PF10_0122	phosphoglucomutase, putative	0.94	3.07E-04	carbohydrate metabolic process		GO:0016866~intramolecular transferase activity
283	PFL0285w	targeted glyoxalase II	1.07	9.33E-06	methylglyoxal metabolic process		GO:0006081~cellular aldehyde metabolic process
Cofactor related processes							
284	PFD0830w	bifunctional dihydrofolate reductase-thymidylate synthase	0.76	4.08E-04	dTMP biosynthetic process	glycine biosynthetic process	
285	PFL1920c	hydroxyethylthiazole kinase, putative	0.76	2.19E-03	thiamin biosynthetic process		GO:0006766~vitamin metabolic process
Vesicle-mediated transport							
286	PF13_0090	ADP-ribosylation factor, putative	-1.92	1.09E-06	intracellular protein transport	small GTPase mediated signal transduction	GO:0006886~intracellular protein transport
287	PFI0515w	SNARE protein, putative	-1.08	3.55E-06	intracellular protein transport	regulation of vesicle fusion	GO:0016192~vesicle-mediated transport
288	PFL0885w	adaptor protein subunit, putative	-0.92	2.07E-05	intracellular protein transport	vesicle-mediated transport	GO:0006886~intracellular protein transport

289	PFB0480w	syntaxin, Qa-SNARE family	-0.90	6.25E-04	intracellular protein transport	regulation of vesicle fusion	
290	MAL13P1.135	SNARE protein, putative	-0.85	1.78E-04	intracellular protein transport	regulation of vesicle fusion	GO:0006886~intracellular protein transport
291	PFF0830w	alpha adaptin-like protein, putative	-0.84	6.76E-04	intracellular protein transport	protein complex assembly	GO:0006461~protein complex assembly
292	PF14_0034	translocation associated membrane protein, putative	1.07	1.84E-04	intracellular protein transmembrane transport		GO:0016021~integral to membrane
Post-translational protein modification							
293	PFE0340c	rhomboid protease ROM4	-1.34	2.28E-06	signal peptide processing		GO:0016021~integral to membrane
294	PF13_0301	ubiquitin conjugating enzyme, putative	-1.01	8.42E-05	post-translational protein modification	regulation of protein metabolic process	GO:0006508~proteolysis
295	PFE0835w	ubiquitin carboxyl-terminal hydrolase 2, putative	-0.96	1.67E-05	protein deubiquitination	ubiquitin-dependent protein catabolic process	GO:0006508~proteolysis
296	PFL2545c	*Plasmodium* exported protein, unknown function	-0.95	3.76E-04	post-translational protein modification	regulation of protein metabolic process	
297	PFL2100w	ubiquitin conjugating enzyme E2, putative	-0.88	2.84E-04	post-translational protein modification	regulation of protein metabolic process	forming carbon-nitrogen bonds
298	PFC0855w	ubiquitin conjugating enzyme, putative	0.77	1.25E-03	post-translational protein modification	regulation of protein metabolic process	
299	PFL1790w	ubiquitin-activating enzyme, putative	0.78	1.10E-04	ER-associated protein catabolic process	protein modification process	GO:0001882~nucleoside binding
300	PF08_0085	ubiquitin conjugating enzyme, putative	0.85	1.69E-05	post-translational protein modification	regulation of protein metabolic process	GO:0006508~proteolysis
301	PF10_0330	ubiquitin conjugating enzyme, putative	0.98	3.76E-05	post-translational protein modification	protein ubiquitination	GO:0006508~proteolysis
Proteolysis							
302	PF08_0108	plasmepsin X	-3.01	6.47E-10	proteolysis		GO:0006508~proteolysis
303	PFE0370c	subtilisin-like protease 1	-3.00	2.36E-08	proteolysis		GO:0006508~proteolysis
304	PFF1365c	HECT-domain (ubiquitin-transferase), putative	-2.69	9.72E-10	protein modification process		
305	PFD0230c	dipeptidyl peptidase 3	-2.17	6.31E-08	protein catabolic process	proteolysis	GO:0006508~proteolysis
306	PF10_0094	tubulin-tyrosine ligase, putative	-2.06	4.17E-08	protein modification process		GO:0016879~ligase activity
307	MAL8P1.126	serine protease, putative	-1.53	6.46E-06	proteolysis		GO:0006508~proteolysis
308	MAL13P1.38	conserved *Plasmodium* membrane protein, unknown function	-1.44	7.09E-07	proteolysis		
309	PF11_0381	subtilisin-like protease 2	-1.36	5.42E-06	protein maturation	proteolysis	
310	PFE0355c	subtilisin-like protease 3, putative	-1.08	1.54E-04	proteolysis		GO:0006508~proteolysis
311	PFE1235c	ubiquitin fusion degradation protein UFD1, putative	-1.04	2.55E-02	ubiquitin-dependent protein catabolic process		GO:0006508~proteolysis
312	PFB0585w	Leu/Phe-tRNA protein transferase, putative	-0.90	2.98E-05	protein catabolic process		GO:0009057~macromolecule catabolic process
313	MAL13P1.56	M1-family alanyl aminopeptidase	0.80	1.03E-02	proteolysis		GO:0006508~proteolysis
314	PFC0495w	plasmepsin VI	0.86	2.36E-02	proteolysis		
315	PFB0355c	serine repeat antigen 2	2.40	9.86E-07	immunoglobulin production	proteolysis	
Protein folding							
316	PF11_0509	ring-infected erythrocyte surface antigen	-3.03	2.93E-09	protein folding		GO:0006457~protein folding
317	PFD0095c	*Plasmodium* exported protein (PHISTb), unknown function	-2.38	7.17E-09	protein folding		
318	PF11_0513	DnaJ protein, putative	-1.52	3.86E-06	protein folding		GO:0006457~protein folding
319	PFD1170c	*Plasmodium* exported protein (PHISTb), unknown function	-1.17	3.58E-06	protein folding		

320	PF11_0175	heat shock protein 101	-0.96	3.08E-05	protein folding	protein hexamerization	GO:0006986~response to unfolded protein
321	PF08_0115	DnaJ protein, putative	-0.80	7.91E-04	protein folding		GO:0006457~protein folding
322	PFI0985c	chaperone, putative	0.78	9.21E-04	protein folding		
323	PFA0460c	tubulin-specific chaperone a, putative	1.11	8.42E-06	protein complex assembly	protein folding	
Protein transport							
324	PFD0720w	conserved ARM repeats protein, unknown function	-3.02	2.93E-09	protein import into nucleus		
325	PFA0155c	conserved *Plasmodium* protein, unknown function	-1.30	3.52E-06	exocytosis	protein transport	
326	PFL2220w	conserved *Plasmodium* protein, unknown function	-1.11	1.30E-05	endocytosis	intracellular protein transport	GO:0006886~intracellular protein transport
327	PF13_0024	conserved *Plasmodium* protein, unknown function	-0.87	2.24E-04	autophagy	protein transport	
328	MAL13P1.51	secretory complex protein 61 alpha,Rab GTPase 5b	-0.83	8.72E-05	protein transport	protein transport by the Sec complex	GO:0006355~regulation of transcription
329	PF14_0171	conserved *Plasmodium* protein, unknown function	-0.80	3.61E-02	autophagy	protein transport	
330	PFE1120w	conserved *Plasmodium* protein, unknown function	-0.79	1.49E-03	protein localization	protein transport	
331	PFF1125c	RNA-binding protein mei2 homologue, putative	0.75	3.98E-03	positive regulation of meiosis		GO:0010564~regulation of cell cycle process
332	PF14_0283	conserved *Plasmodium* protein, unknown function	0.81	6.23E-03	autophagy	cytolysis	GO:0006914~autophagy
333	PFF0155w	mitochondrial chaperone BCS1, putative	0.88	1.10E-02	protein complex assembly		GO:0006461~protein complex assembly
334	MAL8P1.14	mitochondrial inner membrane protein, putative	0.89	1.61E-03	protein insertion into membrane		GO:0008104~protein localization
Protein amino acid dephosphorylation							
335	PF14_0224	serine/threonine protein phosphatase	-2.65	7.57E-10	protein amino acid dephosphorylation		
336	PF08_0129	serine/threonine protein phosphatase, putative	-1.20	2.43E-04	protein amino acid dephosphorylation		GO:0006470~protein amino acid dephosphorylation
337	MAL8P1.108	protein phosphatase, putative	0.85	6.29E-05	protein amino acid dephosphorylation		GO:0006470~protein amino acid dephosphorylation
338	PF14_0280	phosphotyrosyl phosphatase activator, putative	1.01	5.96E-05	protein amino acid dephosphorylation		GO:0006470~protein amino acid dephosphorylation
Chromatin modification							
339	PF13_0152	transcriptional regulatory protein sir2a	-1.24	3.11E-04	chromatin silencing	protein amino acid deacetylation	GO:0006342~chromatin silencing
340	PF10_0211	conserved *Plasmodium* membrane protein, unknown function	-1.05	1.43E-06	ATP synthesis coupled electron transport	transport	
341	MAL13P1.328	DNA topoisomerase VI, b subunit, putative	-0.96	2.61E-04	null		GO:0001882~nucleoside binding
342	PFA0555c	UMP-CMP kinase, putative	-0.95	2.00E-05	nucleobase	nucleoside	
343	PFF0350w	MYND finger protein, putative	-0.93	6.94E-03	negative regulation of transcription from RNA polymerase II promoter		GO:0000122~negative regulation of transcription from RNA polymerase II promoter
344	PF14_0314	chromatin assembly factor 1 P55 subunit, putative	-0.89	6.05E-04	nucleosome assembly		GO:0006323~DNA packaging
345	PF13_0185	histone H3 variant, putative	-0.88	6.41E-05	chromatin assembly or disassembly	nucleosome assembly	GO:0006323~DNA packaging
346	PF11_0429	conserved *Plasmodium* protein, unknown	-0.81	1.65E-04	chromatin modification	regulation of transcription	GO:0043167~ion binding

		function					
347	PF14_0234	DNA-directed DNA polymerase, putative	-0.77	1.94E-04	null		GO:0001882~nucleoside binding
348	PFE1390w	RNA helicase 1	0.77	1.77E-04	null		GO:0001882~nucleoside binding
349	MAL13P1.146	AMP deaminase, putative	0.79	5.44E-04	AMP catabolic process	purine ribonucleoside monophosphate biosynthetic process	GO:0006163~purine nucleotide metabolic process
350	PFD0685c	chromosome associated protein, putative	0.84	1.57E-04	chromosome organization		GO:0000070~mitotic sister chromatid segregation
351	PF13_0349	nucleoside diphosphate kinase	0.89	1.55E-04	CTP biosynthetic process	GTP biosynthetic process	GO:0006163~purine nucleotide metabolic process
352	PFL1725w	ATP synthase beta chain, mitochondrial precursor, putative	0.93	1.10E-05	ATP synthesis coupled proton transport	hydrogen transport	GO:0006091~generation of precursor metabolites and energy
353	PF08_0042	ATP-dependent RNA helicase prh1, putative	1.04	1.06E-02	null		GO:0001882~nucleoside binding
354	PF14_0672	phosphodiesterase delta, putative	1.72	4.82E-06	cyclic nucleotide metabolic process	signal transduction	GO:0009123~nucleoside monophosphate metabolic process
RNA-related processes							
355	PFD0495c	conserved *Plasmodium* protein, unknown function	-2.44	4.87E-09	RNA splicing	mRNA processing	
356	PF11_0255	ribonucleoprotein, putative	-0.82	3.51E-04	RNA splicing	mRNA processing	
357	PFC0145c	conserved *Plasmodium* protein, unknown function	-0.81	2.10E-04	rRNA processing		
358	PF14_0028	pre-mRNA splicing factor, putative	0.80	2.06E-05	RNA processing		GO:0006396~RNA processing
359	PFL1525c	pre-mRNA splicing factor RNA helicase, putative	0.82	3.80E-05	RNA splicing	regulation of cell cycle	GO:0006396~RNA processing
360	PFE1245w	conserved *Plasmodium* protein, unknown function	0.92	2.68E-04	mRNA processing		GO:0043167~ion binding
361	PFD0700c	RNA binding protein, putative	1.48	8.85E-07	RNA metabolic process		
Other processes							
362	PFA0110w	ring-infected erythrocyte surface antigen	-3.77	1.15E-10	regulation of immune response		
363	PF14_0344	translocon component PTEX150	-1.26	2.85E-06	translocation of peptides or proteins into host cell cytoplasm		
364	PFL2285c	conserved *Plasmodium* protein, unknown function	-1.04	1.66E-04	tetrapyrrole biosynthetic process		GO:0018130~heterocycle biosynthetic process
365	PFD0930w	CGI-141 protein homolog, putative	-1.00	1.87E-06	vesicle-mediated transport		GO:0016192~vesicle-mediated transport
366	PF14_0054	conserved protein, unknown function	-1.00	2.34E-06	histone ubiquitination		GO:0043167~ion binding
367	PF13_0159	nicotinic acid mononucleotide adenyltransferase	-0.95	1.55E-04	NAD biosynthetic process		GO:0006732~coenzyme metabolic process
368	PFI0345w	GTPase activator, putative	-0.81	6.34E-04	regulation of Rab GTPase activity		GO:0032313~regulation of Rab GTPase activity
369	PF10_0158	conserved *Plasmodium* protein, unknown function	-0.78	1.06E-03	regulation of GTPase activity		
370	PF13_0221	conserved *Plasmodium* protein, unknown function	0.80	4.92E-03	multicellular organismal development		
371	PF10_0227	HORMA domain protein, putative	0.81	1.79E-03	mitosis		GO:0000087~M phase of mitotic cell cycle
372	PF14_0699	GTPase activator, putative	0.83	5.40E-04	regulation of Rab GTPase activity		GO:0032313~regulation of Rab GTPase activity
373	PFE0970w	cytochrome c oxidase assembly protein (heme A:	0.86	1.26E-05	heme O biosynthetic process		GO:0006778~porphyrin metabolic

		farnesyltransferase), putative					process
374	PF11_0295	farnesyl pyrophosphate synthase, putative	0.86	9.96E-05	isoprenoid biosynthetic process		GO:0006720~isoprenoid metabolic process
375	MAL13P1.183	conserved *Plasmodium* protein, unknown function	0.89	1.17E-02	respiratory electron transport chain		
376	PFB0220w	ubiE/COQ5 methyltransferase, putative	0.89	1.40E-04	quinone cofactor biosynthetic process		GO:0006732~coenzyme metabolic process
377	PF13_0353	NADH-cytochrome b5 reductase, putative	0.89	4.51E-05	null		GO:0055114~oxidation reduction
378	PFF1075w	conserved *Plasmodium* protein, unknown function	0.94	2.17E-04	vacuolar transport	vacuole fusion	
379	PF10_0374	Pf11-1 protein	1.07	5.96E-05	SRP-dependent cotranslational protein targeting to membrane	gamete generation	GO:0007276~gamete generation
380	MAL13P1.251	conserved *Plasmodium* protein, unknown function	1.26	1.40E-05	arginyl-tRNA aminoacylation		GO:0009536~plastid
Transcripts with hypothetical or unknown function							
381	PFL2520w	reticulocyte binding protein homologue 3, pseudogene	-4.63	8.53E-10	null		
382	PF08_0035	conserved *Plasmodium* protein, unknown function	-3.96	6.99E-10	null		
383	PF13_0058	conserved *Plasmodium* protein, unknown function	-3.93	6.47E-10	null		
384	PF07_0061	conserved *Plasmodium* protein, unknown function	-3.56	3.94E-09	null		
385	MAL8P1.70	conserved *Plasmodium* protein, unknown function	-3.49	2.93E-09	null		GO:0043167~ion binding
386	PF11_0268	kelch protein, putative	-3.44	6.50E-10	null		
387	PFE1285w	membrane skeletal protein IMC1-related	-3.22	2.73E-09	null		
388	PFA0440w	photosensitized INA-labeled protein 1, PhIL1, putative	-3.17	3.80E-10	null		
389	PF10_0170	conserved *Plasmodium* protein, unknown function	-3.16	3.04E-10	null		
390	PFC0185w	membrane skeletal protein IMC1-related	-3.09	2.57E-07	null		
391	PFC0830w	trophozoite stage antigen	-2.83	8.33E-08	null		
392	PFB0935w	cytoadherence linked asexual protein 2	-2.79	2.42E-08	null		GO:0020035~cytoadherence to microvasculature
393	PFI0675w	conserved *Plasmodium* protein, unknown function	-2.78	6.38E-10	null		
394	PFF0645c	integral membrane protein, putative	-2.78	3.21E-07	null		GO:0016021~integral to membrane
395	PFC1045c	conserved *Plasmodium* protein, unknown function	-2.74	3.22E-08	null		
396	MAL8P1.141	conserved *Plasmodium* protein, unknown function	-2.71	4.73E-09	null		
397	PFD0195c	conserved *Plasmodium* protein, unknown function	-2.70	3.31E-06	null		
398	PFB0120w	early transcribed membrane protein 2	-2.67	8.58E-09	null		
399	PF11_0039	early transcribed membrane protein 11.1	-2.60	3.58E-08	null		
400	PF13_0233	myosin A	-2.59	4.02E-08	null		GO:0005856~cytoskeleton
401	PF14_0572	conserved *Plasmodium* membrane protein, unknown function	-2.58	6.98E-08	null		

402	PF14_0343	conserved *Plasmodium* protein, unknown function	-2.58	1.65E-07	null		
403	PF13_0079	conserved *Plasmodium* protein, unknown function	-2.56	4.76E-08	null		
404	MAL13P1.268	conserved *Plasmodium* protein, unknown function	-2.54	1.99E-06	null		
405	PF14_0495	rhoptry neck protein 2	-2.51	2.05E-07	null		
406	PFI0410c	conserved *Plasmodium* protein, unknown function	-2.49	1.86E-08	null		
407	PFB0926c	*Plasmodium* exported protein (hyp2), unknown function	-2.45	5.30E-09	null		
408	PFB0670c	conserved *Plasmodium* protein, unknown function	-2.45	3.31E-06	null		
409	PFE0075c	rhoptry-associated protein 3	-2.42	8.59E-09	null		
410	PFI0690c	conserved *Plasmodium* protein, unknown function	-2.38	3.86E-05	null		
411	PF11_0277	conserved *Plasmodium* protein, unknown function	-2.35	9.09E-10	null		
412	PF10_0351	probable protein, unknown function	-2.35	1.17E-08	null		
413	PFF0870w	conserved *Plasmodium* membrane protein, unknown function	-2.34	2.71E-07	null		
414	PFI1800w	lysophospholipase, putative	-2.31	1.25E-06	null		
415	PF11_0213	conserved *Plasmodium* protein, unknown function	-2.28	2.56E-08	null		GO:0016021~integral to membrane
416	MAL7P1.229	cytoadherence linked asexual protein	-2.27	4.31E-09	null		
417	PF14_0337	conserved *Plasmodium* protein, unknown function	-2.25	5.01E-07	null		
418	PFL0300c	protein phosphatase, putative	-2.25	5.90E-08	null		
419	PF08_0058	conserved *Plasmodium* protein, unknown function	-2.24	5.53E-09	null		
420	PFF0710w	conserved *Plasmodium* protein, unknown function	-2.23	4.89E-08	null		
421	PF11_0168	rhoptry neck protein 4	-2.23	8.33E-08	null		
422	PFE0770w	conserved *Plasmodium* protein, unknown function	-2.20	1.63E-07	null		
423	PFD1150c	reticulocyte binding protein homologue 4	-2.18	1.89E-07	null		
424	PF10_0306	MORN repeat protein, putative	-2.15	4.22E-09	null		
425	PF14_0222	ankyrin, putative	-2.14	2.95E-06	null		
426	MAL7P1.66	mitochondrial ribosomal protein S5 precursor, putative	-2.12	8.30E-08	null		GO:0005840~ribosome
427	PF14_0102	rhoptry-associated protein 1	-2.10	1.20E-06	null		GO:0020007~apical complex
428	MAL13P1.260	alveolin, putative	-2.09	2.76E-07	null		
429	PFF0675c	myosin E, putative	-2.09	2.32E-07	null		GO:0008092~cytoskeletal protein binding
430	PFL2225w	myosin A tail domain interacting protein	-2.08	3.20E-09	null		
431	PFI0785c	sugar transporter, putative	-2.08	6.16E-07	null		GO:0008643~carbohydrate transport
432	PF10_0352	merozoite surface protein	-2.05	8.33E-08	null		
433	PF11_0528	conserved *Plasmodium* protein, unknown function	-2.03	1.04E-06	null		

434	PFC0355c	conserved *Plasmodium* protein, unknown function	-2.03	1.74E-07	null		
435	PF14_0018	*Plasmodium* exported protein (PHISTb), unknown function	-2.02	3.40E-09	null		
436	PFL1090w	glideosome-associated protein 45	-2.00	2.76E-06	null		
437	PF10_0135	conserved *Plasmodium* protein, unknown function	-1.99	9.40E-08	null		
438	PF10_0193	microtubule-associated protein 1 light chain 3, putative	-1.97	1.27E-08	null		
439	PFI1445w	high molecular weight rhoptry protein 2	-1.97	4.02E-08	null		
440	PFB0475c	conserved *Plasmodium* protein, unknown function	-1.96	3.58E-08	null		
441	PF11_0094	conserved *Plasmodium* protein, unknown function	-1.96	9.40E-08	null		
442	PF08_0130	rRNA processing WD-repeat protein, putative	-1.92	1.36E-06	null		
443	PFD0100c	surface-associated interspersed gene 4.1, (SURFIN4.1)	-1.90	9.64E-07	null		
444	MAL13P1.405	conserved protein, unknown function	-1.90	1.63E-07	null		
445	PF10_0146	conserved *Plasmodium* protein, unknown function	-1.89	9.40E-08	null		
446	PFE1260c	conserved *Plasmodium* protein, unknown function	-1.89	8.48E-08	null		
447	PF10_0184	conserved *Plasmodium* protein, unknown function	-1.89	2.56E-07	null		
448	PF10_0072	conserved *Plasmodium* protein, unknown function	-1.88	8.77E-07	null		GO:0043167~ion binding
449	PFL1785c	conserved *Plasmodium* protein, unknown function	-1.88	5.76E-09	null		
450	PFD0395c	conserved *Plasmodium* protein, unknown function	-1.87	9.90E-09	null		
451	PF07_0127	conserved *Plasmodium* protein, unknown function	-1.87	4.76E-08	null		
452	MAL13P1.94	conserved *Plasmodium* membrane protein, unknown function	-1.84	7.09E-08	null		
453	PF14_0652	conserved *Plasmodium* protein, unknown function	-1.82	3.95E-08	null		GO:0043167~ion binding
454	PF10_0220	phospholipid scramblase 1, putative	-1.80	1.84E-07	null		
455	PF14_0119	p1/s1 nuclease, putative	-1.80	8.11E-06	null		
456	PF11_0180	conserved *Plasmodium* protein, unknown function	-1.76	7.46E-07	null		
457	PFF1020c	conserved *Plasmodium* protein, unknown function	-1.76	5.57E-07	null		
458	PFL0060w	*Plasmodium* exported protein, unknown function	-1.74	8.68E-07	null		
459	PFE0175c	myosin B	-1.73	1.17E-06	null		GO:0005856~cytoskeleton
460	PF14_0732	*Plasmodium* exported protein (PHISTb), unknown function	-1.72	4.65E-07	null		
461	PF13_0255	conserved *Plasmodium* protein, unknown function	-1.70	3.13E-06	null		
462	PF14_0578	conserved *Plasmodium* protein, unknown	-1.68	1.42E-07	null		

		function					
463	PFI0265c	high molecular weight rhoptry protein 3	-1.67	7.22E-07	null		
464	PF10_0308	OTU-like cysteine protease, putative	-1.67	7.91E-08	null		
465	PFC0820w	conserved *Plasmodium* protein, unknown function	-1.67	8.59E-06	null		
466	PFE0365c	conserved *Plasmodium* protein, unknown function	-1.66	4.16E-05	null		
467	PFB0275w	metabolite/drug transporter, putative	-1.66	4.18E-07	null		
468	PF14_0586	conserved *Plasmodium* protein, unknown function	-1.64	2.26E-06	null		
469	PF11_0415	conserved *Plasmodium* protein, unknown function	-1.64	1.61E-07	null		
470	PF11_0267	kelch protein, putative	-1.63	7.74E-06	null		
471	PF13_0036	DnaJ protein, putative	-1.63	1.66E-06	null		
472	PFF0480w	conserved *Plasmodium* protein, unknown function	-1.62	1.54E-07	null		
473	PF10_0347	merozoite surface protein	-1.60	9.22E-07	null		
474	PFD0385w	conserved *Plasmodium* protein, unknown function	-1.60	6.44E-07	null		
475	PF10_0295	conserved *Plasmodium* protein, unknown function	-1.59	9.48E-05	null		
476	MAL8P1.144	AAA family ATPase, putative	-1.57	1.89E-07	null		GO:0006355~regulation of transcription
477	PFI1720w	gametocytogenesis-implicated protein	-1.57	2.03E-05	null		
478	PFE0440w	conserved *Plasmodium* protein, unknown function	-1.52	1.05E-05	null		
479	PFD0765w	RING zinc finger protein, putative	-1.52	6.05E-07	null		GO:0043167~ion binding
480	PFD0185c	conserved *Plasmodium* protein, unknown function	-1.52	5.19E-07	null		
481	PFF0990c	conserved *Plasmodium* protein, unknown function	-1.51	4.17E-07	null		
482	PFE1415w	cell cycle regulator with zn-finger domain, putative	-1.51	1.40E-05	null		GO:0043167~ion binding
483	PFC0370w	conserved *Plasmodium* protein, unknown function	-1.50	9.61E-07	null		
484	MAL13P1.102	conserved *Plasmodium* protein, unknown function	-1.50	4.84E-04	null		
485	PFI0360c	conserved *Plasmodium* protein, unknown function	-1.49	4.43E-07	null		
486	MAL7P1.29	conserved *Plasmodium* membrane protein, unknown function	-1.49	6.72E-06	null		
487	PF14_0660	protein phosphatase, putative	-1.48	1.91E-06	null		
488	MAL8P1.13	folate/biopterin transporter, putative	-1.47	1.27E-06	null		GO:0016021~integral to membrane
489	PF08_0119	conserved *Plasmodium* protein, unknown function	-1.47	1.66E-06	null		
490	PF10_0151	conserved *Plasmodium* protein, unknown function	-1.45	1.07E-05	null		
491	PF14_0044	conserved *Plasmodium* protein, unknown function	-1.44	3.22E-05	null		
492	PF13_0169	conserved *Plasmodium* protein, unknown	-1.42	1.68E-05	null		

		function					
493	PF14_0092	conserved *Plasmodium* membrane protein, unknown function	-1.41	2.05E-07	null		
494	PFF0485c	zinc finger protein, putative	-1.41	2.02E-07	null		GO:0016021~integral to membrane
495	PF10_0037	conserved *Plasmodium* protein, unknown function	-1.41	4.88E-06	null		
496	PF10_0223	conserved *Plasmodium* membrane protein, unknown function	-1.41	1.90E-06	null		
497	MAL13P1.331	conserved *Plasmodium* protein, unknown function	-1.41	3.74E-06	null		
498	MAL13P1.130	glideosome associated protein with multiple membrane spans 1	-1.40	1.57E-06	null		
499	PF14_0213	conserved *Plasmodium* protein, unknown function	-1.40	1.53E-05	null		
500	PFI0940c	PPPDE peptidase, putative	-1.40	4.01E-06	null		
501	PF14_0013	DnaJ protein, putative	-1.39	4.82E-06	null		GO:0016021~integral to membrane
502	MAL8P1.146	filament assembling protein, putative	-1.39	3.47E-04	null		GO:0005200~structural constituent of cytoskeleton
503	PF10_0292	conserved *Plasmodium* protein, unknown function	-1.38	9.18E-06	null		
504	PFA0265c	conserved *Plasmodium* protein, unknown function	-1.38	6.53E-06	null		
505	MAL13P1.245	conserved *Plasmodium* protein, unknown function	-1.37	1.15E-05	null		
506	PFB0106c	*Plasmodium* exported protein, unknown function	-1.37	1.99E-04	null		
507	PFL1025c	conserved *Plasmodium* protein, unknown function	-1.37	3.65E-07	null		
508	PFD0255w	ag-1 blood stage membrane protein homologue	-1.37	1.28E-06	null		
509	PFI0845w	conserved *Plasmodium* membrane protein, unknown function	-1.36	6.88E-06	null		
510	PF11_0193	conserved *Plasmodium* protein, unknown function	-1.36	1.33E-04	null		
511	PFI0705w	conserved *Plasmodium* protein, unknown function	-1.36	2.89E-06	null		
512	PF11_0300	conserved *Plasmodium* protein, unknown function	-1.36	6.53E-07	null		
513	PF08_0003	tryptophan/threonine-rich antigen	-1.36	1.17E-06	null		
514	PF14_0363	metacaspase-like protein	-1.35	2.41E-05	null		
515	PF11_0362	protein phosphatase, putative	-1.33	7.73E-07	null		
516	PFL1505c	conserved *Plasmodium* protein, unknown function	-1.32	6.62E-04	null		
517	PFL0315c	conserved *Plasmodium* protein, unknown function	-1.32	8.41E-07	null		
518	PF14_0135	conserved *Plasmodium* protein, unknown function	-1.31	4.71E-05	null		
519	PF11_0040	early transcribed membrane protein 11.2	-1.31	8.62E-06	null		
520	PF10_0248	conserved *Plasmodium* membrane protein, unknown function	-1.29	8.77E-06	null		
521	PF11_0167	conserved *Plasmodium* membrane protein,	-1.29	8.49E-07	null		GO:0043167~ion binding

		unknown function					
522	MAL7P1.170	*Plasmodium* exported protein, unknown function	-1.29	3.82E-07	null		
523	PF13_0162	conserved *Plasmodium* protein, unknown function	-1.29	2.61E-06	null		
524	PF08_0081	conserved *Plasmodium* protein, unknown function	-1.29	2.77E-06	null		
525	PF14_0110	rhomboid protease ROM8	-1.28	6.59E-07	null		
526	PF07_0019	conserved *Plasmodium* protein, unknown function	-1.27	7.73E-07	null		
527	PF10_0073	conserved *Plasmodium* protein, unknown function	-1.27	9.02E-06	null		
528	PF14_0374	CCAAT-binding transcription factor, putative	-1.27	5.78E-06	null		
529	PF10_0019	early transcribed membrane protein 10.1	-1.26	8.67E-05	null		
530	PFB0625w	conserved *Plasmodium* protein, unknown function	-1.23	6.86E-06	null		
531	PF14_0106	ankyrin-repeat protein, putative	-1.23	2.03E-04	null		
532	PF10_0254	conserved *Plasmodium* protein, unknown function	-1.23	1.80E-05	null		
533	PFI1555w	conserved *Plasmodium* protein, unknown function	-1.23	1.59E-06	null		
534	PF13_0161	conserved *Plasmodium* protein, unknown function	-1.23	8.65E-06	null		
535	PF11_0505	probable protein, unknown function	-1.23	7.09E-06	null		
536	PF08_0118	conserved *Plasmodium* protein, unknown function	-1.22	4.36E-06	null		GO:0043167~ion binding
537	PFE0650c	conserved *Plasmodium* protein, unknown function	-1.21	7.07E-06	null		
538	PFD1130w	conserved *Plasmodium* protein, unknown function	-1.21	5.85E-05	null		
539	MAL13P1.228	conserved *Plasmodium* protein, unknown function	-1.20	1.45E-06	null		
540	PFC0700c	conserved *Plasmodium* protein, unknown function	-1.19	7.99E-06	null		
541	PF11_0379	conserved *Plasmodium* protein, unknown function	-1.19	2.66E-03	null		
542	PF14_0045	conserved *Plasmodium* protein, unknown function	-1.19	2.74E-05	null		GO:0016021~integral to membrane
543	PFC0090w	*Plasmodium* exported protein, unknown function	-1.19	4.56E-04	null		
544	PFL1795c	conserved *Plasmodium* protein, unknown function	-1.18	7.24E-05	null		
545	PFC0345w	conserved *Plasmodium* protein, unknown function	-1.18	1.01E-05	null		
546	PFB0105c	*Plasmodium* exported protein (PHISTc), unknown function	-1.17	1.17E-04	null		
547	PFL1870c	sphingomyelin phosphodiesterase, putative	-1.17	5.92E-05	null		
548	MAL13P1.480	histidine-rich protein III	-1.17	2.19E-04	null		
549	PFC0235w	conserved *Plasmodium* protein, unknown function	-1.17	7.60E-06	null		
550	MAL13P1.215	conserved *Plasmodium* protein, unknown	-1.16	4.50E-05	null		

		function					
551	PFB0615c	conserved *Plasmodium* protein, unknown function	-1.16	2.41E-05	null		
552	PFI1040c	conserved *Plasmodium* membrane protein, unknown function	-1.16	4.44E-03	null		
553	PF10_0163	*Plasmodium* exported protein (PHISTc), unknown function	-1.16	4.91E-05	null		
554	PF10_0301	calmodulin, putative	-1.15	1.99E-04	null		GO:0043167~ion binding
555	PF07_0107	*Plasmodium* exported protein, unknown function	-1.15	8.38E-05	null		
556	PF14_0065	glideosome associated protein with multiple membrane spans 3, putative	-1.11	3.85E-06	null		
557	PF14_0016	early transcribed membrane protein 14.1	-1.11	3.35E-05	null		
558	PF10_0315	conserved *Plasmodium* protein, unknown function	-1.10	1.18E-04	null		
559	PFA0210c	conserved *Plasmodium* protein, unknown function	-1.10	9.25E-06	null		
560	PF14_0353	conserved *Plasmodium* protein, unknown function	-1.10	1.41E-05	null		
561	MAL13P1.413	membrane associated histidine-rich protein	-1.10	5.15E-04	null		
562	MAL13P1.380	conserved *Plasmodium* protein, unknown function	-1.09	1.40E-04	null		
563	MAL8P1.62	conserved *Plasmodium* protein, unknown function	-1.08	3.86E-06	null		
564	PFF1475c	conserved *Plasmodium* protein, unknown function	-1.08	7.54E-06	null		
565	MAL13P1.308	conserved *Plasmodium* protein, unknown function	-1.08	1.65E-06	null		
566	PF14_0388	conserved *Plasmodium* protein, unknown function	-1.08	1.67E-05	null		
567	PFE0785c	metabolite/drug transporter, putative	-1.08	5.61E-06	null		
568	PF14_0472	conserved *Plasmodium* protein, unknown function	-1.07	1.33E-04	null		
569	PF11_0417	conserved *Plasmodium* protein, unknown function	-1.06	6.25E-05	null		
570	PFI1380c	conserved *Plasmodium* protein, unknown function	-1.06	1.90E-03	null		
571	PFA0295c	conserved *Plasmodium* protein, unknown function	-1.05	1.22E-03	null		
572	PFE0200c	AN1-like zinc finger family protein	-1.05	8.48E-05	null		GO:0043167~ion binding
573	PFE0590w	conserved *Plasmodium* protein, unknown function	-1.05	1.33E-04	null		
574	MAL13P1.37	conserved *Plasmodium* protein, unknown function	-1.04	7.66E-05	null		GO:0043167~ion binding
575	PF13_0086	conserved *Plasmodium* protein, unknown function	-1.04	1.33E-04	null		
576	PFC0905c	oocyst capsule protein	-1.04	5.51E-05	null		
577	PF11_0172	folate/biopterin transporter, putative	-1.04	8.97E-05	null		
578	PFE0425w	conserved *Plasmodium* protein, unknown function	-1.03	2.13E-03	null		GO:0043167~ion binding

579	PFL2120w	conserved *Plasmodium* protein, unknown function	-1.03	2.77E-06	null		
580	MAL8P1.109	protein phosphatase, putative	-1.01	5.39E-05	null		
581	PFL0325w	Tat binding protein 1(TBP-1)-interacting protein, putative	-1.01	5.17E-03	null		
582	PFD0080c	*Plasmodium* exported protein (PHISTb), unknown function	-0.99	3.20E-05	null		
583	PFC0875w	ABC transporter, putative	-0.99	2.97E-05	null		
584	PF14_0149	conserved *Plasmodium* protein, unknown function	-0.99	1.22E-03	null		
585	PFI0885w	cytochrome b5-like heme/steroid binding protein, putative	-0.98	2.69E-05	null		
586	PFD0715c	conserved *Plasmodium* protein, unknown function	-0.98	3.86E-04	null		
587	PFD0940w	conserved *Plasmodium* protein, unknown function	-0.98	6.48E-05	null		
588	PFE1410c	conserved *Plasmodium* protein, unknown function	-0.97	1.93E-03	null		
589	PFL1565c	CG2-related protein, putative	-0.96	8.78E-06	null		
590	PF10_0156	conserved *Plasmodium* protein, unknown function	-0.96	1.92E-03	null		
591	PFL2110c	conserved protein, unknown function	-0.96	5.96E-03	null		
592	PF14_0143	atypical protein kinase, ABC-1 family, putative	-0.96	2.14E-05	null		
593	PF14_0530	ferlin, putative	-0.95	2.37E-04	null		
594	PFF0610c	PP-loop family protein, putative	-0.94	4.36E-05	null		
595	PF14_0173	cyclic nucleotide-binding protein, putative	-0.94	3.40E-05	null		
596	PFE1130w	conserved protein, unknown function	-0.93	4.12E-05	null		
597	PF14_0303	conserved *Plasmodium* protein, unknown function	-0.93	3.52E-04	null		
598	MAL7P1.7	*Plasmodium* exported protein (PHISTb), unknown function	-0.93	3.39E-04	null		
599	MAL13P1.264	conserved *Plasmodium* protein, unknown function	-0.93	1.90E-05	null		
600	PF14_0604	conserved *Plasmodium* protein, unknown function	-0.93	5.31E-04	null		
601	PFL1385c	merozoite surface protein 9	-0.93	1.24E-03	null		
602	PF10_0020	alpha/beta hydrolase, putative	-0.92	3.90E-04	null		
603	PFB0955w	stevor	-0.91	1.88E-04	null		GO:0005576~extracellular region
604	PF10_0283	conserved *Plasmodium* protein, unknown function	-0.91	4.60E-04	null		
605	PFB0190c	conserved *Plasmodium* protein, unknown function	-0.90	1.69E-04	null		
606	MAL13P1.188	conserved *Plasmodium* protein, unknown function	-0.90	1.47E-05	null		
607	MAL7P1.13	conserved *Plasmodium* protein, unknown function	-0.90	5.86E-04	null		
608	PF10_0342	probable protein, unknown function	-0.89	2.30E-03	null		
609	PFI0975c	conserved *Plasmodium* protein, unknown function	-0.89	1.45E-04	null		

610	PFF0295c	conserved *Plasmodium* protein, unknown function	-0.88	9.52E-05	null		
611	MAL7P1.222	rifin	-0.88	8.25E-03	null		
612	PF07_0028	conserved *Plasmodium* protein, unknown function	-0.88	2.00E-02	null		
613	PFL2205w	conserved *Plasmodium* protein, unknown function	-0.88	3.08E-04	null		
614	PF10_0244	formin 2, putative	-0.87	6.76E-05	null		GO:0043167~ion binding
615	PFD0300w	conserved *Plasmodium* protein, unknown function	-0.87	1.67E-03	null		
616	PF08_0082	conserved *Plasmodium* protein, unknown function	-0.87	6.97E-04	null		
617	PFL1945c	early transcribed membrane protein 12	-0.86	1.14E-03	null		
618	MAL8P1.12	conserved *Plasmodium* protein, unknown function	-0.86	3.45E-03	null		
619	PF10_0304	conserved *Plasmodium* protein, unknown function	-0.86	1.30E-03	null		
620	PFF0605c	conserved *Plasmodium* protein, unknown function	-0.86	2.03E-02	null		
621	PFL1650w	conserved *Plasmodium* protein, unknown function	-0.86	1.09E-04	null		
622	PF10_0023	*Plasmodium* exported protein (hyp16), unknown function	-0.85	3.86E-02	null		
623	PFD0910w	conserved *Plasmodium* protein, unknown function	-0.85	4.31E-03	null		
624	PF13_0055	conserved *Plasmodium* protein, unknown function	-0.85	2.14E-04	null		GO:0043167~ion binding
625	PFI0850w	conserved *Plasmodium* protein, unknown function	-0.84	4.89E-04	null		
626	PFB0145c	conserved *Plasmodium* protein, unknown function	-0.84	2.22E-05	null		
627	PFB0310c	merozoite surface protein 4	-0.84	4.16E-05	null		
628	PFF1460c	conserved *Plasmodium* protein, unknown function	-0.83	2.25E-04	null		
629	PF14_0703	conserved *Plasmodium* protein, unknown function	-0.83	1.25E-03	null		
630	PF10_0025	PF70 protein	-0.83	2.14E-04	null		
631	PFE1095w	conserved *Plasmodium* protein, unknown function	-0.83	4.31E-04	null		
632	PF10_0118	conserved *Plasmodium* protein, unknown function	-0.83	1.43E-02	null		
633	PF14_0196	tetratricopeptide repeat family protein, putative	-0.83	2.39E-04	null		
634	PF13_0303	regulator of chromosome condensation, putative	-0.82	3.30E-05	null		
635	MAL13P1.287	conserved *Plasmodium* protein, unknown function	-0.82	7.23E-04	null		
636	PF11_0390	conserved *Plasmodium* protein, unknown function	-0.82	5.30E-03	null		
637	PF10_0034	conserved *Plasmodium* protein, unknown function	-0.82	4.20E-04	null		

638	MAL13P1.137	conserved *Plasmodium* protein, unknown function	-0.81	1.47E-04	null		
639	PFI1735c	ring-exported protein 1	-0.81	1.57E-04	null		
640	MAL13P1.296	conserved *Plasmodium* protein, unknown function	-0.81	2.49E-04	null		
641	PF14_0151	RNA-binding protein Nova-1, putative	-0.81	8.68E-04	null		
642	PF11_0290	conserved *Plasmodium* protein, unknown function	-0.81	4.91E-05	null		
643	PFE0400w	SET domain protein, putative	-0.81	4.02E-04	null		
644	PFF1505w	TRAP-like protein, putative	-0.80	2.64E-04	null		
645	PF10_0185	conserved *Plasmodium* protein, unknown function	-0.80	4.02E-02	null		
646	PF14_0233	conserved *Plasmodium* protein, unknown function	-0.80	1.42E-04	null		
647	PFC0760c	conserved *Plasmodium* protein, unknown function	-0.80	1.88E-03	null		
648	PF10_0157	conserved *Plasmodium* protein, unknown function	-0.80	9.25E-03	null		
649	MAL13P1.175	MSP7-like protein, fragment	-0.79	7.11E-04	null		
650	PFI0355c	ATP-dependent heat shock protein, putative	-0.79	2.42E-03	null		GO:0009376~HslUV protease complex
651	MAL8P1.160	*Plasmodium* exported protein (hyp7), unknown function	-0.79	6.62E-03	null		
652	PF10_0024	*Plasmodium* exported protein (hyp2), unknown function	-0.78	1.29E-03	null		
653	PF14_0136	diphthamide synthesis protein, putative	-0.78	7.75E-05	null		
654	PF13_0285	inositol-polyphosphate 5-phosphatase, putative	-0.78	4.26E-04	null		GO:0004445~inositol-polyphosphate 5-phosphatase activity
655	PFI1755c	ring-exported protein 3	-0.78	8.67E-05	null		
656	MAL13P1.329	conserved *Plasmodium* membrane protein, unknown function	-0.77	1.71E-02	null		
657	PFD0580c	conserved *Plasmodium* protein, unknown function	-0.77	3.39E-02	null		
658	PFB0915w	liver stage antigen 3	-0.77	5.35E-03	null		
659	PFB0910w	*Plasmodium* exported protein, unknown function	-0.77	1.56E-04	null		
660	MAL8P1.209	var-like protein	-0.77	1.07E-02	null		
661	PFI0385c	P1 nuclease, putative	-0.77	6.62E-04	null		
662	PF08_0001	*Plasmodium* exported protein, unknown function	-0.76	5.99E-04	null		
663	PFC0130c	conserved *Plasmodium* protein, unknown function	-0.76	8.12E-04	null		
664	PF11_0179	conserved *Plasmodium* protein, unknown function	-0.76	8.87E-05	null		
665	PF14_0621	conserved *Plasmodium* protein, unknown function	-0.76	1.74E-04	null		
666	PFI1225w	conserved *Plasmodium* protein, unknown function	-0.76	6.58E-03	null		
667	PF11_0392	conserved *Plasmodium* protein, unknown function	-0.76	3.18E-02	null		
668	PF14_0140	conserved *Plasmodium* protein, unknown function	-0.76	1.51E-04	null		

669	PF10_0145	conserved *Plasmodium* protein, unknown function	-0.76	2.42E-03	null		GO:0043167~ion binding
670	MAL13P1.152	conserved *Plasmodium* protein, unknown function	-0.76	1.88E-02	null		
671	PF13_0064	conserved *Plasmodium* protein, unknown function	-0.75	2.72E-04	null		
672	PF14_0197	zinc finger protein, putative	-0.75	1.33E-04	null		
673	PFF0570c	conserved *Plasmodium* protein, unknown function	-0.75	2.61E-03	null		
674	PF11_0421	conserved *Plasmodium* protein, unknown function	-0.75	2.87E-04	null		
675	MAL13P1.127	conserved *Plasmodium* protein, unknown function	0.75	2.80E-04	null		
676	PFB0365w	conserved *Plasmodium* protein, unknown function	0.75	3.32E-05	null		
677	PFI0580c	falstatin	0.75	2.13E-04	null		
678	PF14_0032	conserved *Plasmodium* protein, unknown function	0.76	2.18E-03	null		
679	PFL2315c	conserved *Plasmodium* membrane protein, unknown function	0.76	4.06E-03	null		
680	MAL13P1.195	CPW-WPC family protein	0.76	3.78E-04	null		
681	PF10_0253	conserved *Plasmodium* protein, unknown function	0.77	6.24E-03	null		
682	PFI1430w	conserved *Plasmodium* protein, unknown function	0.77	1.03E-03	null		
683	PF07_0039	conserved *Plasmodium* protein, unknown function	0.77	2.78E-03	null		
684	PFF1070c	radical SAM protein, putative	0.77	2.08E-04	null		
685	MAL13P1.236	conserved *Plasmodium* protein, unknown function	0.78	2.31E-04	null		
686	PF08_0040	clp1-related protein, putative	0.78	9.61E-05	null		
687	PF11_0403	conserved *Plasmodium* protein, unknown function	0.78	2.42E-03	null		
688	PFD0415c	conserved *Plasmodium* protein, unknown function	0.79	4.32E-04	null		
689	PF14_0220	conserved *Plasmodium* protein, unknown function	0.79	1.14E-02	null		
690	PFE0760w	conserved *Plasmodium* protein, unknown function	0.79	3.66E-02	null		
691	PFC0915w	ATP-dependent RNA helicase, putative	0.79	3.40E-05	null		
692	PFI1610c	calcyclin binding protein, putative	0.79	5.14E-05	null		
693	PFC0570c	conserved *Plasmodium* protein, unknown function	0.80	5.35E-04	null		
694	PF14_0310	conserved *Plasmodium* protein, unknown function	0.80	1.15E-04	null		
695	PFL1140w	integral membrane protein, putative	0.80	5.28E-05	null		
696	PF11_0322	conserved *Plasmodium* protein, unknown function	0.80	9.39E-04	null		
697	PF08_0021	SPRY domain, putative	0.81	3.43E-03	null		

698	PFL1845c	calcyclin binding protein, putative	0.81	6.55E-04	null		
699	PF14_0103	conserved *Plasmodium* protein, unknown function	0.81	9.21E-05	null		
700	PF13_0114	conserved *Plasmodium* protein, unknown function	0.82	1.26E-02	null		
701	PF13_0194	probable protein, unknown function	0.82	2.87E-05	null		
702	PF10_0238	conserved *Plasmodium* protein, unknown function	0.82	5.75E-04	null		
703	MAL13P1.351	conserved *Plasmodium* protein, unknown function	0.82	1.35E-03	null		
704	PF13_0028	membrane integral peptidase, M50 family, putative	0.82	4.34E-04	null		
705	PF11_0357	zinc finger protein, putative	0.82	4.46E-03	null		GO:0043167~ion binding
706	PFF0640w	conserved *Plasmodium* protein, unknown function	0.82	2.80E-04	null		
707	PFA0235w	conserved *Plasmodium* protein, unknown function	0.83	1.27E-03	null		
708	PFI0605c	conserved *Plasmodium* protein, unknown function	0.83	8.28E-04	null		
709	PF11_0186	conserved *Plasmodium* protein, unknown function	0.83	1.98E-05	null		
710	PFL0610w	conserved *Plasmodium* protein, unknown function	0.84	2.72E-02	null		
711	PF14_0542	conserved protein, unknown function	0.84	2.83E-04	null		
712	MAL13P1.149	conserved *Plasmodium* protein, unknown function	0.85	7.75E-04	null		
713	PFI1330c	conserved *Plasmodium* protein, unknown function	0.85	2.16E-04	null		
714	MAL7P1.124	conserved *Plasmodium* protein, unknown function	0.86	3.98E-04	null		
715	PF13_0223	conserved protein, unknown function	0.86	4.55E-04	null		
716	PF13_0035	U3 small nucleolar RNA-associated protein 6, putative	0.86	4.35E-03	null		
717	PF07_0114	conserved *Plasmodium* protein, unknown function	0.86	2.33E-03	null		
718	PFL0800c	cell traversal protein for ookinetes and sporozoites	0.87	2.27E-02	null		
719	PF11_0308	conserved *Plasmodium* protein, unknown function	0.87	2.23E-04	null		
720	MAL8P1.56	conserved *Plasmodium* membrane protein, unknown function	0.87	7.73E-05	null		
721	MAL13P1.347	conserved *Plasmodium* protein, unknown function	0.88	3.64E-04	null		
722	PF10_0092	metallopeptidase, putative	0.88	2.05E-05	null		
723	PFL1065c	conserved *Plasmodium* protein, unknown function	0.88	3.70E-04	null		
724	MAL7P1.115	conserved *Plasmodium* membrane protein, unknown function	0.88	1.18E-02	null		
725	PF14_0176	conserved *Plasmodium* protein, unknown	0.91	2.62E-04	null		

		function					
726	PF07_0110	protein phosphatase, putative	0.91	2.69E-05	null		
727	PFI0440w	conserved *Plasmodium* membrane protein, unknown function	0.92	2.11E-03	null		
728	PFI1080w	dynein intermediate chain 2, ciliary	0.94	3.54E-04	null		
729	PFI1705w	phosphatidylinositol N-acetylglucosaminyltransferase subunit P, putative	0.94	4.56E-05	null		
730	MAL8P1.84	conserved *Plasmodium* protein, unknown function	0.96	5.92E-05	null		
731	PF07_0014	conserved *Plasmodium* protein, unknown function	0.97	1.88E-02	null		GO:0016021~integral to membrane
732	PFD0915w	conserved *Plasmodium* protein, unknown function	0.98	5.40E-04	null		
733	MAL8P1.111	JmjC domain containing protein	0.98	2.24E-05	null		
734	PFL0340w	conserved *Plasmodium* protein, unknown function	0.99	7.80E-05	null		
735	MAL8P1.97	hypothetical protein	1.00	6.32E-05	null		
736	PF07_0118	conserved *Plasmodium* membrane protein, unknown function	1.00	4.73E-05	null		
737	PF13_0202	conserved *Plasmodium* protein, unknown function	1.00	4.76E-05	null		
738	PFA0190c	actin-related protein	1.01	3.47E-04	null		
739	PFL0710w	conserved *Plasmodium* protein, unknown function	1.01	8.83E-06	null		
740	PF13_0266	conserved *Plasmodium* protein, unknown function	1.01	5.51E-04	null		
741	PF14_0358	41-2 protein antigen precursor,transport protein particle (TRAPP) component, Bet3	1.02	1.67E-04	null		GO:0016192~vesicle-mediated transport
742	PF14_0648	conserved *Plasmodium* protein, unknown function	1.02	3.65E-05	null		
743	PF13_0164	conserved *Plasmodium* protein, unknown function	1.03	3.99E-04	null		
744	PFI0950w	protein disulfide isomerase	1.03	2.79E-06	null		GO:0016860~intramolecular oxidoreductase activity
745	MAL13P1.182	conserved *Plasmodium* protein, unknown function	1.04	1.94E-05	null		
746	PF08_0099	acyl-CoA binding protein, putative	1.07	5.91E-05	null		GO:0005504~fatty acid binding
747	PFC0505c	conserved *Plasmodium* protein, unknown function	1.09	3.36E-04	null		
748	PFL1260w	hydrolase/phosphatase, putative	1.10	5.38E-05	null		
749	MAL13P1.106	probable protein, unknown function	1.13	2.67E-04	null		
750	PF10_0273	DHHC-type zinc finger protein, putative	1.16	1.17E-06	null		GO:0043167~ion binding
751	PFL1695c	conserved *Plasmodium* protein, unknown function	1.18	1.47E-04	null		
752	PF11_0486	merozoite adhesive erythrocytic binding protein	1.18	1.83E-04	null		GO:0020007~apical complex
753	PFI0615w	conserved *Plasmodium* protein, unknown function	1.20	9.89E-04	null		
754	PFL0995c	conserved *Plasmodium* protein, unknown function	1.20	2.03E-04	null		

755	PFE1540w	CPW-WPC family protein	1.27	2.09E-06	null		
756	PF14_0588	conserved *Plasmodium* protein, unknown function	1.28	7.01E-04	null		
757	PF13_0081	conserved *Plasmodium* protein, unknown function	1.29	3.48E-07	null		
758	PFL0070c	*Plasmodium* exported protein, unknown function	1.31	9.96E-04	null		
759	PF14_0631	conserved *Plasmodium* protein, unknown function	1.32	7.13E-06	null		
760	MAL13P1.197	probable protein, unknown function	1.33	7.27E-04	null		
761	PFE1450c	conserved *Plasmodium* protein, unknown function	1.34	4.79E-06	null		
762	PFF1065c	conserved *Plasmodium* protein, unknown function	1.34	1.40E-04	null		
763	PFF1520w	RESA-like protein	1.35	8.31E-03	null		
764	MAL8P1.6	early transcribed membrane protein 8	1.37	2.21E-06	null		
765	PFI1220w	conserved protein, unknown function	1.43	2.77E-06	null		
766	PF14_0305	leucine-rich repeat protein	1.46	3.85E-06	null		
767	PF14_0452	conserved *Plasmodium* protein, unknown function	1.49	8.00E-03	null		
768	PFL1750c	conserved *Plasmodium* protein, unknown function	1.50	2.79E-06	null		
769	PF11_0451	conserved *Plasmodium* protein, unknown function	1.58	1.95E-07	null		
770	PFC0680w	conserved *Plasmodium* protein, unknown function	1.59	6.00E-05	null		
771	PFI0980w	long chain fatty acid elongation enzyme, putative	1.60	1.23E-05	null		
772	PF14_0131	conserved *Plasmodium* protein, unknown function	1.60	2.79E-06	null		
773	MAL7P1.156	conserved *Plasmodium* protein, unknown function	1.61	2.56E-07	null		
774	PFL0795c	male development gene 1	1.65	4.18E-05	null		
775	MAL13P1.129	conserved *Plasmodium* protein, unknown function	1.66	1.59E-06	null		
776	PF10_0164	early transcribed membrane protein 10.3	1.67	4.17E-08	null		
777	PFI0315c	dynein light intermediate chain 2, cytosolic	1.69	2.83E-03	null		
778	PF14_0040	secreted ookinete adhesive protein	1.81	4.73E-05	null		
779	PF11_0092	mechanosensitive ion channel protein	1.84	3.09E-05	null		
780	PF14_0272	CPW-WPC family protein	1.84	2.31E-08	null		
781	PF11_0479	conserved *Plasmodium* protein, unknown function	1.89	7.35E-05	null		
782	PFL2385c	conserved *Plasmodium* protein, unknown function	2.09	1.44E-07	null		
783	PF14_0467	conserved *Plasmodium* protein, unknown function	2.16	3.50E-06	null		
784	PFE1205c	conserved *Plasmodium* membrane protein, unknown function	2.69	5.40E-09	null		
785	PF14_0708	probable protein, unknown function	3.00	2.31E-08	null		
786	PFI1520w	asparagine-rich antigen, putative	3.10	6.47E-10	null		
787	PFI1230c	conserved *Plasmodium* protein, unknown	3.43	4.15E-08	null		

		function					

Table A4.2: **Transcripts associated with cofactor metabolic processes and folate metabolism at Tt_3 parasites compared to UTt_3.** GSEA identified transcripts involved in cofactor metabolic process as well as folate metabolism were enriched in Tt_3 parasites compared to UTt_3. Listed below are the log_2FC of expressed transcripts belonging to these two different clusters.

PlasmoDB ID	PlasmoDB description	log_2FC Tt_3:UTt_3	aj.P.Val	PlasmoDB annotated processes	GO term
GO.0051186.COFACTOR.METABOLIC.PROCESS					
PF13_0159	nicotinic acid mononucleotide adenyltransferase	-0.95	1.55E-04	NAD biosynthetic process	GO:0006732~coenzyme metabolic process
PFB0320c	iron-sulfur assembly protein, putative	-0.79	1.57E-04	iron-sulfur cluster assembly	GO:0016226~iron-sulfur cluster assembly
PFL1920c	hydroxyethylthiazole kinase, putative	0.76	2.19E-03	thiamin biosynthetic process	GO:0006766~vitamin metabolic process
PFD0830w	bifunctional dihydrofolate reductase-thymidylate synthase	0.76	4.08E-04	dTMP biosynthetic process	
PF11_0169	SNO glutamine amidotransferase family protein	0.77	5.98E-05	response to singlet oxygen and vitamin B6 biosynthetic process	
PFE0970w	cytochrome c oxidase assembly protein (heme A: farnesyltransferase), putative	0.86	1.26E-05	heme O biosynthetic process	GO:0006778~porphyrin metabolic process
PF10_0210	deoxyribose-phosphate aldolase, putative	0.86	1.01E-05	deoxyribonucleotide catabolic process	GO:0005996~monosaccharide metabolic process
PFB0220w	ubiE/COQ5 methyltransferase, putative	0.89	1.40E-04	quinone cofactor biosynthetic process	GO:0006732~coenzyme metabolic process
PF11_0044	iron-sulfur assembly protein, sufD, putative	0.92	6.29E-05	iron-sulfur cluster assembly	GO:0016226~iron-sulfur cluster assembly

PF08_0066	lipoamide dehydrogenase	1.28	8.55E-05	acetyl-CoA biosynthetic process from pyruvate	GO:0006084~acetyl-CoA metabolic process
PF11_0256	pyruvate dehydrogenase E1 alpha subunit	1.65	4.28E-08	acetyl-CoA biosynthetic process from pyruvate	GO:0006084~acetyl-CoA metabolic process
PF14_0441	pyruvate dehydrogenase E1 beta subunit	0.91	1.07E-04	acetyl-CoA biosynthetic process from pyruvate	GO:0009536~plastid
PFB0220w	ubiE/COQ5 methyltransferase, putative	0.89	1.40E-04	quinone cofactor biosynthetic process	GO:0006732~coenzyme metabolic process
PF11_0044	iron-sulfur assembly protein, sufD, putative	0.92	6.29E-05	iron-sulfur cluster assembly	GO:0016226~iron-sulfur cluster assembly
MPM.FOLATE BIOSYNTHESIS					
PF08_0066	lipoamide dehydrogenase (aLipDH)	1.28	8.55E-05	cell redox homeostasis, oxidation-reduction process	GO:0051186~cofactor metabolic process
* PFD0830w	bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)	0.76	4.08E-04	dTMP biosynthetic process	glycine biosynthetic process
PF13_0345	glycine cleavage T protein, putative (GCVT)	0.32	2.95E-02	glycine catabolic process	
PF13_0140	dihydrofolate synthase/folylpolyglutamate synthase (DHFS-FPGS)	0.49	4.92E-03	biosynthetic process, folic acid-containing compound biosynthetic process	GO:0051186~cofactor metabolic process
PF08_0095	dihydropteroate synthetase (DHPS)	0.43	3.23E-02	folic acid-containing compound biosynthetic process	GO:0051188~cofactor biosynthetic process
MAL13P1.67	methionyl-tRNA formyltransferase, putative	0.30	1.54E-01	biosynthetic process,translational initiation	

Footnote: The adjusted P-value (adj. P-value) indicates a normalised standard error from spot or probe results, and is interpreted as a normal t-statistic result. * DHFR-TS (PFD0830w) was not part of the annotated cluster MPM.folate biosynthesis, however is part of folate recycling. This transcript was included for interpretation purposes.

Table A4.3: Uniquely expressed transcripts from 4PEHz-treated parasites at t_3. Comparison of differentially regulated transcripts from t_3 to microarray data of; CQ- ,artesunate- ,WR99210-treated *P. falciparum*, as well as heat-treated *P. falciparum*, including a study involving 20 different compounds which focuses on schizontal expression patterns. Comparisons Transcripts were uniquely expressed during 4PEHz treatment, and not found in other transcriptomes.

No.	PlasmoDB ID	PlasmoDB description	log_2FC Tt_3:UTt_3	aj.P.Val	PlasmoDB annotated processes	Additional process involvement	GO term
Protein amino acid dephosphorylation and signal transduction							
1	PFL0475w	cGMP-specific phosphodiesterase	-0.85	1.22E-02	cGMP catabolic process	signal transduction	
2	PF14_0723	LCCL domain-containing protein	-0.82	3.58E-04	cell adhesion	signal transduction	
3	PF14_0672	phosphodiesterase delta, putative	1.72	4.82E-06	cyclic nucleotide metabolic process	signal transduction	GO:0009123~nucleoside monophosphate metabolic process
4	MAL7P1.73	calcium/calmodulin-dependent protein kinase, putative	0.82	1.67E-02	protein amino acid phosphorylation	signal transduction	GO:0006468~protein amino acid phosphorylation
5	PF14_0498	DER1-like protein	1.10	5.63E-05	ER-associated protein catabolic process	regulation of cell cycle process	
6	PFF0270c	cyclin dependent kinase binding protein, putative	0.79	1.61E-03	regulation of cell cycle	regulation of cell division	GO:0051726~regulation of cell cycle
7	PFE0045c	serine/threonine protein kinase, FIKK family	-0.89	2.11E-03	protein amino acid phosphorylation		GO:0006468~protein amino acid phosphorylation
8	PFF0260w	serine/threonine protein kinase, putative	0.91	1.21E-03	protein amino acid phosphorylation		GO:0004674~protein serine/threonine kinase activity
9	MAL13P1.114	conserved *Plasmodium* protein, unknown function	-1.80	4.53E-07	protein amino acid phosphorylation		GO:0009536~plastid
10	PF08_0129	serine/threonine protein phosphatase, putative	-1.15	4.52E-05	protein amino acid dephosphorylation		GO:0006470~protein amino acid dephosphorylation
11	PF14_0224	serine/threonine protein phosphatase	-2.65	7.57E-10	protein amino acid dephosphorylation		
12	MAL13P1.297	ADP-ribosylation factor, putative	-1.17	9.97E-05	protein amino acid ADP-ribosylation	small GTPase mediated signal transduction	GO:0006471~protein amino acid ADP-ribosylation
13	PF11_0147	mitogen-activated protein kinase 2	-0.81	5.62E-05	MAPKKK cascade	protein amino acid phosphorylation	

Apicoplast associated transcripts							
14	MAL13P1.254	conserved *Plasmodium* protein, unknown function	1.47	2.76E-05	null		GO:0009536~plastid
15	MAL13P1.342	conserved *Plasmodium* protein, unknown function	1.35	2.58E-06	null		GO:0009536~plastid
16	MAL13P1.49	conserved *Plasmodium* protein, unknown function	-1.57	1.48E-06	null		GO:0009536~plastid
17	MAL13P1.87	sec20 homolog, putative	0.99	2.36E-05	null		GO:0009536~plastid
18	MAL7P1.6	*Plasmodium* exported protein (hyp12), unknown function	-0.79	3.77E-04	null		GO:0009536~plastid
19	MAL8P1.138	alpha/beta hydrolase, putative	1.39	2.04E-07	null		GO:0009536~plastid
20	MAL8P1.73	rhoptry neck protein 5, putative	-2.78	1.02E-09	null		GO:0009536~plastid
21	PF10_0119	conserved *Plasmodium* protein, unknown function	-2.53	7.16E-09	null		GO:0009536~plastid
22	PF11_0135	conserved *Plasmodium* protein, unknown function	-1.13	1.26E-03	null		GO:0009536~plastid
23	PF11_0285	conserved *Plasmodium* protein, unknown function	0.89	2.09E-05	null		GO:0009536~plastid
24	PF11_0304	conserved *Plasmodium* protein, unknown function	-1.93	5.83E-05	null		GO:0009536~plastid
25	PF11_0323	conserved *Plasmodium* protein, unknown function	0.94	1.81E-03	null		GO:0009536~plastid
26	PF13_0125	conserved *Plasmodium* protein, unknown function	-1.32	8.67E-05	null		GO:0009536~plastid
27	PFB0110w	*Plasmodium* exported protein (hyp11), unknown function	-0.95	2.55E-02	null		GO:0009536~plastid
28	PFE0575c	conserved *Plasmodium* protein, unknown function	-1.62	4.91E-07	null		GO:0009536~plastid
29	PFI0205w	conserved *Plasmodium* protein, unknown function	1.27	2.91E-04	null		GO:0009536~plastid
30	PFL0875w	conserved *Plasmodium* protein, unknown function	-0.91	4.43E-04	null		GO:0009536~plastid
Cytoskeleton organization and antigenic variation							
31	PFE1420w	f-actin capping protein alpha subunit, putative	1.15	3.41E-05	actin cytoskeleton organization		GO:0007010~cytoskeleton organization

32	PFF1545w	rifin	-2.09	4.93E-04	antigenic variation		
33	PFL0030c	erythrocyte membrane protein 1, PfEMP1	-1.12	1.42E-04	antigenic variation	cell-cell adhesion	
34	PF10_0003	rifin	-3.70	6.50E-10	antigenic variation		GO:0006952~defense response
35	MAL7P1.50	erythrocyte membrane protein 1, PfEMP1	0.97	2.62E-04	antigenic variation	cell-cell adhesion	GO:0009405~pathogenesis
36	PFB0010w	erythrocyte membrane protein 1, PfEMP1	-0.92	7.65E-04	antigenic variation	cell-cell adhesion	GO:0006952~defense response
37	PFF1420w	phosphatidylcholine-sterol acyltransferase precursor, putative	-1.60	4.83E-07	entry into host cell	lipid metabolic process	GO:0009536~plastid
38	PFD0955w	apical merozoite protein	-2.79	6.50E-10	attachment of GPI anchor to protein		
39	PFI0535w	phosphatidylinositol N-acetylglucosaminyltransferase, putative	0.99	7.77E-05	GPI anchor biosynthetic process		
40	PFD1105w	asparagine-rich protein	-2.19	9.66E-07	cell-cell adhesion		
41	PFI1730w	cytoadherence linked asexual protein 9	-1.69	5.10E-08	cell-cell adhesion		
42	PFE0175c	myosin B	-1.73	1.17E-06	null		GO:0005856~cytoskeleton
43	PFA0260c	cyclase-associated protein, putative	-1.21	5.19E-05	cytoskeleton organization		
44	PFB0680w	rhoptry neck protein 6	-2.36	6.27E-08	entry into host		
Metabolic process							
45	PFC0260w	P-loop containing nucleoside triphospahte hydrolase, putative	-0.88	7.55E-04	metabolic process		
46	PF11_0190	haloacid dehalogenase-like hydrolase, putative	-1.13	8.38E-05	metabolic process		GO:0016791~phosphatase activity
47	PFL0035c	acyl-CoA synthetase, PfACS7	-1.89	2.32E-07	metabolic process		
48	PF13_0066	malonyl CoA-acyl carrier protein transacylase precursor	1.19	1.76E-02	fatty acid biosynthetic process		GO:0004314~[acyl-carrier-protein] S-malonyltransferase activity
49	PFD1035w	steroid dehydrogenase, putative	0.82	2.52E-04	fatty acid biosynthetic process		GO:0006631~fatty acid metabolic process
50	PFB0505c	beta-ketoacyl-ACP synthase III	-0.77	5.90E-04	fatty acid biosynthetic process		

Transcription, translation and post translational modification							
51	PFD0700c	RNA binding protein, putative	1.48	8.85E-07	RNA metabolic process		
52	PF11_0383	conserved *Plasmodium* protein, unknown function	-1.68	6.58E-06	DNA-dependent DNA replication		
53	PF11_0429	conserved *Plasmodium* protein, unknown function	-0.81	1.65E-04	chromatin modification	regulation of transcription	GO:0043167~ion binding
54	PFI1630c	conserved *Plasmodium* protein, unknown function	1.18	1.14E-05	transcription		
55	PFC0241w	conserved *Plasmodium* protein, unknown function	0.89	1.26E-04	translation		
56	MAL13P1.251	conserved *Plasmodium* protein, unknown function	1.26	1.40E-05	arginyl-tRNA aminoacylation		GO:0009536~plastid
57	PFD1100c	conserved *Plasmodium* protein, unknown function	-2.28	6.22E-07	translation	translational initiation	GO:0009536~plastid
58	PFF0705c	conserved *Plasmodium* protein, unknown function	-1.66	3.17E-04	translation	translational initiation	
59	PFI1675w	conserved *Plasmodium* protein, unknown function	-1.96	1.63E-07	translation		
60	PFI1725w	*Plasmodium* exported protein, unknown function	-1.06	1.20E-05	translation	valyl-tRNA aminoacylation	
61	PFL1910c	conserved *Plasmodium* protein, unknown function	2.05	1.96E-06	translation		
62	PFL1895w	mitochondrial ribosomal protein L23 precursor, putative	0.86	1.96E-04	translation		GO:0006412~translation
63	PFL2545c	*Plasmodium* exported protein, unknown function	-0.95	3.76E-04	post-translational protein modification	regulation of protein metabolic process	
64	PFC0495w	plasmepsin VI	0.86	2.36E-02	proteolysis		
65	PF10_0094	tubulin-tyrosine ligase, putative	-2.06	4.17E-08	protein modification process		GO:0016879~ligase activity
Transport							
66	PF10_0059	conserved *Plasmodium* membrane protein, unknown function	-2.39	1.39E-08	transport		
67	PF11_0287	CRAL/TRIO domain-containing protein, putative	-0.95	4.31E-05	transport		
68	PF10_0216	zinc transporter, putative	1.03	1.70E-04	ion transport	transport	

69	PF11_0340	conserved *Plasmodium* protein, unknown function	-0.77	2.10E-03	potassium ion transport		
70	PFI0400c	conserved *Plasmodium* membrane protein, unknown function	1.20	1.99E-06	ion transport		
71	PFI1560c	conserved *Plasmodium* membrane protein, unknown function	-1.04	1.70E-04	transport		
72	PFF1075w	conserved *Plasmodium* protein, unknown function	0.94	2.17E-04	vacuolar transport	vacuole fusion	
Other processes							
73	MAL13P1.328	DNA topoisomerase VI, b subunit, putative	-0.96	2.61E-04	null		GO:0001882~nucleoside binding
74	PFI1005w	ADP-ribosylation factor, putative	-2.07	8.59E-09	small GTPase mediated signal transduction		GO:0005525~GTP binding
75	PFB0480w	syntaxin, Qa-SNARE family	-0.90	6.25E-04	intracellular protein transport	regulation of vesicle fusion	
76	PF14_0358	41-2 protein antigen precursor,transport protein particle (TRAPP) component, Bet3	1.02	1.67E-04	null		GO:0016192~vesicle-mediated transport
77	PFB0955w	stevor	-0.91	1.88E-04	null		GO:0005576~extracellular region
78	PFE1235c	ubiquitin fusion degradation protein UFD1, putative	-1.04	2.55E-02	ubiquitin-dependent protein catabolic process		GO:0006508~proteolysis
79	PF13_0159	nicotinic acid mononucleotide adenyltransferase	-0.95	1.55E-04	NAD biosynthetic process		GO:0006732~coenzyme metabolic process
80	PFL1920c	hydroxyethylthiazole kinase, putative	0.76	2.19E-03	thiamin biosynthetic process		GO:0006766~vitamin metabolic process
81	PF13_0021	small heat shock protein, putative	1.09	4.19E-05	response to heat	response to unfolded protein	GO:0006986~response to unfolded protein
82	MAL8P1.14	mitochondrial inner membrane protein, putative	0.89	1.61E-03	protein insertion into membrane		GO:0008104~protein localization
83	PF14_0034	translocation associated membrane protein, putative	1.07	1.84E-04	intracellular protein transmembrane transport		GO:0016021~integral to membrane
84	PF07_0014	conserved *Plasmodium* protein, unknown function	0.97	1.88E-02	null		GO:0016021~integral to membrane
85	PFF0485c	zinc finger protein, putative	-1.41	2.02E-07	null		GO:0016021~integral to membrane

86	PFL2285c	conserved *Plasmodium* protein, unknown function	-1.04	1.66E-04	tetrapyrrole biosynthetic process		GO:0018130~heterocycle biosynthetic process
87	PF10_0072	conserved *Plasmodium* protein, unknown function	-1.88	8.77E-07	null		GO:0043167~ion binding
88	PF10_0301	calmodulin, putative	-1.15	1.99E-04	null		GO:0043167~ion binding
89	PFD0765w	RING zinc finger protein, putative	-1.52	6.05E-07	null		GO:0043167~ion binding
Transcripts with hypothetical or unknown function							
90	MAL13P1.102	conserved *Plasmodium* protein, unknown function	-1.50	4.84E-04	null		
91	MAL13P1.106	probable protein, unknown function	1.13	2.67E-04	null		
92	MAL13P1.149	conserved *Plasmodium* protein, unknown function	0.85	7.75E-04	null		
93	MAL13P1.152	conserved *Plasmodium* protein, unknown function	-0.76	1.88E-02	null		
94	MAL13P1.175	MSP7-like protein, fragment	-0.79	7.11E-04	null		
95	MAL13P1.182	conserved *Plasmodium* protein, unknown function	1.04	1.94E-05	null		
96	MAL13P1.188	conserved *Plasmodium* protein, unknown function	-0.90	1.47E-05	null		
97	MAL13P1.215	conserved *Plasmodium* protein, unknown function	-1.16	4.50E-05	null		
98	MAL13P1.245	conserved *Plasmodium* protein, unknown function	-1.37	1.15E-05	null		
99	MAL13P1.287	conserved *Plasmodium* protein, unknown function	-0.82	7.23E-04	null		
100	MAL7P1.124	conserved *Plasmodium* protein, unknown function	0.86	3.98E-04	null		
101	MAL7P1.156	conserved *Plasmodium* protein, unknown function	1.61	2.56E-07	null		
102	MAL7P1.222	rifin	-0.88	8.25E-03	null		
103	MAL8P1.12	conserved *Plasmodium* protein, unknown function	-0.86	3.45E-03	null		
104	MAL8P1.160	*Plasmodium* exported protein (hyp7), unknown function	-0.79	6.62E-03	null		

105	MAL8P1.56	conserved *Plasmodium* membrane protein, unknown function	0.87	7.73E-05	null		
106	PF07_0028	conserved *Plasmodium* protein, unknown function	-0.88	2.00E-02	null		
107	PF07_0127	conserved *Plasmodium* protein, unknown function	-1.87	4.76E-08	null		
108	PF08_0082	conserved *Plasmodium* protein, unknown function	-0.87	6.97E-04	null		
109	PF10_0034	conserved *Plasmodium* protein, unknown function	-0.82	4.20E-04	null		
110	PF10_0092	metallopeptidase, putative	0.88	2.05E-05	null		
111	PF10_0135	conserved *Plasmodium* protein, unknown function	-1.99	9.40E-08	null		
112	PF10_0157	conserved *Plasmodium* protein, unknown function	-0.80	9.25E-03	null		
113	PF10_0184	conserved *Plasmodium* protein, unknown function	-1.89	2.56E-07	null		
114	PF10_0185	conserved *Plasmodium* protein, unknown function	-0.80	4.02E-02	null		
115	PF10_0238	conserved *Plasmodium* protein, unknown function	1.07	3.56E-06	null		
116	PF10_0248	conserved *Plasmodium* membrane protein, unknown function	-1.29	8.77E-06	null		
117	PF10_0254	conserved *Plasmodium* protein, unknown function	-1.23	1.80E-05	null		
118	PF10_0283	conserved *Plasmodium* protein, unknown function	-0.91	4.60E-04	null		
119	PF10_0342	probable protein, unknown function	-0.89	2.30E-03	null		
120	PF11_0180	conserved *Plasmodium* protein, unknown function	-1.76	7.46E-07	null		
121	PF11_0193	conserved *Plasmodium* protein, unknown function	-1.36	1.33E-04	null		
122	PF11_0300	conserved *Plasmodium* protein, unknown function	-1.36	6.53E-07	null		
123	PF11_0379	conserved *Plasmodium* protein, unknown function	-1.19	2.66E-03	null		
124	PF11_0390	conserved *Plasmodium* protein, unknown function	-0.82	5.30E-03	null		

125	PF11_0421	conserved *Plasmodium* protein, unknown function	-0.75	2.87E-04	null		
126	PF11_0451	conserved *Plasmodium* protein, unknown function	1.58	1.95E-07	null		
127	PF13_0114	conserved *Plasmodium* protein, unknown function	0.82	1.26E-02	null		
128	PF13_0194	probable protein, unknown function	0.82	2.87E-05	null		
129	PF13_0266	conserved *Plasmodium* protein, unknown function	1.01	5.51E-04	null		
130	PF14_0040	secreted ookinete adhesive protein	1.81	4.73E-05	null		
131	PF14_0044	conserved *Plasmodium* protein, unknown function	-1.44	3.22E-05	null		
132	PF14_0092	conserved *Plasmodium* membrane protein, unknown function	-1.41	2.05E-07	null		
133	PF14_0119	p1/s1 nuclease, putative	-1.80	8.11E-06	null		
134	PF14_0220	conserved *Plasmodium* protein, unknown function	0.79	1.14E-02	null		
135	PF14_0222	ankyrin, putative	-2.14	2.95E-06	null		
136	PF14_0303	conserved *Plasmodium* protein, unknown function	-0.93	3.52E-04	null		
137	PF14_0337	conserved *Plasmodium* protein, unknown function	-2.25	5.01E-07	null		
138	PF14_0353	conserved *Plasmodium* protein, unknown function	-1.10	1.41E-05	null		
139	PF14_0452	conserved *Plasmodium* protein, unknown function	1.49	8.00E-03	null		
140	PF14_0472	conserved *Plasmodium* protein, unknown function	-1.07	1.33E-04	null		
141	PF14_0530	ferlin, putative	-0.95	2.37E-04	null		
142	PF14_0572	conserved *Plasmodium* membrane protein, unknown function	-2.58	6.98E-08	null		
143	PF14_0578	conserved *Plasmodium* protein, unknown function	-1.68	1.42E-07	null		
144	PF14_0586	conserved *Plasmodium* protein, unknown function	-1.64	2.26E-06	null		

145	PF14_0604	conserved *Plasmodium* protein, unknown function	-0.93	5.31E-04	null		
146	PF14_0660	protein phosphatase, putative	-1.48	1.91E-06	null		
147	PFA0235w	conserved *Plasmodium* protein, unknown function	0.83	1.27E-03	null		
148	PFB0105c	*Plasmodium* exported protein (PHISTc), unknown function	-1.17	1.17E-04	null		
149	PFB0310c	merozoite surface protein 4	-0.84	4.16E-05	null		
150	PFB0625w	conserved *Plasmodium* protein, unknown function	-1.23	6.86E-06	null		
151	PFB0670c	conserved *Plasmodium* protein, unknown function	-2.45	3.31E-06	null		
152	PFC0370w	conserved *Plasmodium* protein, unknown function	-1.50	9.61E-07	null		
153	PFC0830w	trophozoite stage antigen	-2.83	8.33E-08	null		
154	PFC0905c	oocyst capsule protein	-1.04	5.51E-05	null		
155	PFC1045c	conserved *Plasmodium* protein, unknown function	-2.74	3.22E-08	null		
156	PFD0100c	surface-associated interspersed gene 4.1, (SURFIN4.1)	-1.90	9.64E-07	null		
157	PFD0195c	conserved *Plasmodium* protein, unknown function	-2.70	3.31E-06	null		
158	PFD0395c	conserved *Plasmodium* protein, unknown function	-1.87	9.90E-09	null		
159	PFD0580c	conserved *Plasmodium* protein, unknown function	-0.77	3.39E-02	null		
160	PFD0715c	conserved *Plasmodium* protein, unknown function	-0.98	3.86E-04	null		
161	PFD0910w	conserved *Plasmodium* protein, unknown function	-0.85	4.31E-03	null		
162	PFD0915w	conserved *Plasmodium* protein, unknown function	0.98	5.40E-04	null		
163	PFE0590w	conserved *Plasmodium* protein, unknown function	-1.05	1.33E-04	null		
164	PFE0760w	conserved *Plasmodium* protein, unknown function	0.79	3.66E-02	null		

165	PFE1095w	conserved *Plasmodium* protein, unknown function	-0.83	4.31E-04	null		
166	PFE1205c	conserved *Plasmodium* membrane protein, unknown function	2.69	5.40E-09	null		
167	PFE1410c	conserved *Plasmodium* protein, unknown function	-0.97	1.93E-03	null		
168	PFE1540w	CPW-WPC family protein	1.27	2.09E-06	null		
169	PFF0570c	conserved *Plasmodium* protein, unknown function	-0.75	2.61E-03	null		
170	PFF0605c	conserved *Plasmodium* protein, unknown function	-0.86	2.03E-02	null		
171	PFF0640w	conserved *Plasmodium* protein, unknown function	0.82	2.80E-04	null		
172	PFF0990c	conserved *Plasmodium* protein, unknown function	-1.51	4.17E-07	null		
173	PFF1065c	conserved *Plasmodium* protein, unknown function	1.34	1.40E-04	null		
174	PFF1475c	conserved *Plasmodium* protein, unknown function	-1.08	7.54E-06	null		
175	PFI0315c	dynein light intermediate chain 2, cytosolic	1.69	2.83E-03	null		
176	PFI0385c	P1 nuclease, putative	-0.77	6.62E-04	null		
177	PFI0440w	conserved *Plasmodium* membrane protein, unknown function	0.92	2.11E-03	null		
178	PFI0845w	conserved *Plasmodium* membrane protein, unknown function	-1.36	6.88E-06	null		
179	PFI0980w	long chain fatty acid elongation enzyme, putative	1.60	1.23E-05	null		
180	PFI1080w	dynein intermediate chain 2, ciliary	0.94	3.54E-04	null		
181	PFI1220w	conserved protein, unknown function	1.43	2.77E-06	null		
182	PFI1225w	conserved *Plasmodium* protein, unknown function	-0.76	6.58E-03	null		
183	PFI1230c	conserved *Plasmodium* protein, unknown function	3.43	4.15E-08	null		
184	PFI1330c	conserved *Plasmodium* protein, unknown function	0.85	2.16E-04	null		

185	PFI1380c	conserved *Plasmodium* protein, unknown function	-1.06	1.90E-03	null		
186	PFI1705w	phosphatidylinositol N-acetylglucosaminyltransferase subunit P, putative	0.94	4.56E-05	null		
187	PFI1735c	ring-exported protein 1	-0.81	1.57E-04	null		
188	PFI1800w	lysophospholipase, putative	-2.31	1.25E-06	null		
189	PFL0325w	Tat binding protein 1(TBP-1)-interacting protein, putative	-1.01	5.17E-03	null		
190	PFL0795c	male development gene 1	1.65	4.18E-05	null		
191	PFL0995c	conserved *Plasmodium* protein, unknown function	1.20	2.03E-04	null		
192	PFL1785c	conserved *Plasmodium* protein, unknown function	-1.88	5.76E-09	null		
193	PFL1870c	sphingomyelin phosphodiesterase, putative	-1.17	5.92E-05	null		
194	PFL1945c	early transcribed membrane protein 12	-0.86	1.14E-03	null		
195	PFL2110c	conserved protein, unknown function	-0.96	5.96E-03	null		
196	PFL2385c	conserved *Plasmodium* protein, unknown function	2.09	1.44E-07	null		

Table A4.4: Proteins present in protein extracts from 4PEHz-treated parasites.

Accession	Description	Sequest score	% Sequence coverage	Nr of peptides	MW (kDa)	Calc. pI	Predicted GO process
PFL0020w	erythrocyte membrane protein 1, PfEMP1	15.0	7.3	18	333.0	6.8	antigenic variation, cell-cell adhesion
MAL7P1.89	dynein heavy chain, putative	12.8	8.3	37	698.4	8.5	microtubule-based movement
PFB0870w	conserved *Plasmodium* protein, unknown function	12.3	13.6	23	283.4	9.2	metabolic process
MAL7P1.55	erythrocyte membrane protein 1, PfEMP1	12.3	9.9	18	256.4	5.8	pathogenesis
MAL13P1.133	conserved *Plasmodium* membrane protein, unknown function	10.5	11.7	43	658.6	9.3	null
MAL13P1.278	serine/threonine protein kinase, putative	10.2	11.6	35	475.4	8.9	protein phosphorylation
PF14_0722	cysteine repeat modular protein 4	9.5	8.6	36	700.5	8.2	null
PFD0020c	erythrocyte membrane protein 1, PfEMP1	8.8	8.5	21	398.0	5.9	pathogenesis
PFD0995c	erythrocyte membrane protein 1, PfEMP1	8.3	11.8	21	247.5	5.5	pathogenesis
PF14_0594	conserved *Plasmodium* protein, unknown function	7.3	9.1	20	395.4	6.6	null
PF13_0148	conserved *Plasmodium* protein, unknown function	6.7	10.6	33	575.1	8.9	null
PF13_0044	carbamoyl phosphate synthetase	6.7	7.4	10	273.3	6.6	nitrogen compound metabolic process
PFF0445w	conserved *Plasmodium* protein, unknown function	6.6	8.1	33	720.1	7.0	null
PF14_0668	conserved *Plasmodium* protein, unknown function	6.6	10.1	17	294.9	8.0	null
MAL8P1.127	conserved *Plasmodium* protein, unknown function	6.5	7.4	7	158.9	5.1	null
PFL0950c	aminophospholipid-transporting P-ATPase	6.4	8.4	8	179.0	7.1	ATP biosynthesis, phospholipid transport
PF13_0079	conserved *Plasmodium* protein, unknown function	6.2	13.8	17	240.4	8.7	null
PF13_0155	conserved *Plasmodium* protein, unknown function	5.7	6.6	15	319.9	7.5	double-strand break repair
PFC0930c	conserved *Plasmodium* protein, unknown function	5.4	13.3	15	221.4	9.5	null
PFF0285c	DNA repair-like protein, putative	4.6	12.3	19	267.8	8.6	null
PF11_0417	conserved *Plasmodium* protein, unknown function	4.6	7.3	9	219.7	9.6	null
PFE0440w	conserved *Plasmodium* protein, unknown function	4.1	8.4	14	370.0	5.4	null
PF13_0080	telomerase reverse transcriptase, putative	3.8	12.0	26	303.9	9.9	RNA-dependent DNA replication
PF07_0082	conserved *Plasmodium* membrane protein, unknown function	3.8	6.3	15	402.4	8.3	null
PF14_0318	conserved *Plasmodium* protein, unknown function	3.8	10.7	20	252.4	9.5	null
PF14_0278	ATP-dependent DNA helicase, putative	3.8	8.2	7	169.4	8.0	DNA recombination, DNA repair
PFL1130c	conserved *Plasmodium* protein, unknown function	3.7	5.0	14	544.0	8.8	null
PF13_0089	inositol polyphosphate kinase, putative	3.7	10.0	10	192.1	9.2	null
PF14_0160	conserved protein, unknown function	3.7	9.4	17	264.8	9.3	proteolysis
PFL2505c	conserved *Plasmodium* protein, unknown function	3.6	10.0	17	263.0	9.2	null
MAL8P1.150	adenylyl cyclase beta, putative	3.4	7.6	12	269.4	7.5	cyclic nucleotide biosynthetic process, intracellular signal transduction
MAL13P1.246	conserved *Plasmodium* membrane protein, unknown function	3.4	6.8	9	218.4	8.1	null
PF14_0165	conserved *Plasmodium* protein, unknown function	3.3	7.3	18	357.4	8.9	null

Table A4.5: Proteins present in untreated parasite extracts from the 4PEHz study.

Accession	Description	Sequest score	% Sequence coverage	Nr of peptides	MW (kDa)	Calc. pI	Predicted GO process
MAL8P1.208	rifin	54.3	12.4	3	41.1	8.8	NA
PF14_0101	conserved *Plasmodium* protein, unknown function	13.6	8.5	24	526.4	8.9	null
PF14_0627	40S ribosomal protein S3, putative	9.1	22.2	5	24.7	10.2	translation
PF07_0118	conserved *Plasmodium* membrane protein, unknown function	8.4	5.3	18	670.7	9.2	null
PFL0360c	conserved *Plasmodium* protein, unknown function	8.3	12.9	25	326.3	8.7	null
PFB0615c	conserved *Plasmodium* protein, unknown function	8.1	12.6	24	293.1	9.1	null
PFE0245c	conserved *Plasmodium* membrane protein, unknown function	8.1	6.5	17	355.0	5.4	null
PFL0405w	conserved *Plasmodium* protein, unknown function	8.0	7.0	18	533.3	9.0	null
MAL13P1.116	conserved *Plasmodium* membrane protein, unknown function	7.7	7.4	14	404.7	8.5	null
PF14_0419	conserved *Plasmodium* protein, unknown function	7.4	9.5	45	864.5	8.6	null
PFF1440w	SET domain protein, putative	7.2	10.7	51	795.5	8.6	null
PF08_0140	erythrocyte membrane protein 1, PfEMP1	7.1	8.9	18	340.4	5.8	pathogenesis
PF07_0037	Cg2 protein	6.2	8.0	15	325.3	7.8	null
MAL13P1.66	conserved *Plasmodium* protein, unknown function	6.2	12.9	24	345.5	9.3	null
PF14_0547	conserved *Plasmodium* protein, unknown function	6.1	6.6	15	370.8	8.9	null
PF14_0145	ubiquitin carboxyl-terminal hydrolase, putative	5.5	17.4	18	164.3	9.5	ubiquitin-dependent protein catabolic process
PF14_0337	conserved *Plasmodium* protein, unknown function	4.8	10.6	13	210.6	8.6	null
PFI0165c	DEAD/DEAH box helicase, putative	4.5	10.8	17	300.7	6.9	null
PFD1015c	erythrocyte membrane protein 1, PfEMP1	4.4	7.0	10	248.8	5.4	pathogenesis
PFD0965w	phosphatidylinositol 4-kinase, putative	4.3	8.5	28	611.1	9.5	NA
PFC0925w	conserved *Plasmodium* protein, unknown function	4.3	22.6	6	59.8	9.9	null
PF14_0073	conserved *Plasmodium* protein, unknown function	4.2	13.5	22	241.0	9.5	null
PF11_0481	tubulin-tyrosine ligase, putative	4.1	14.4	26	325.1	9.2	protein modification process
PF13_0298	conserved *Plasmodium* protein, unknown function	4.0	21.2	20	166.2	9.4	null
PFF0645c	integral membrane protein, putative	4.0	8.2	8	171.9	8.9	null
MAL7P1.15	conserved *Plasmodium* membrane protein, unknown function	3.9	10.0	30	513.5	8.4	null
PFL1280w	RAP protein, putative	3.8	9.6	10	179.7	9.2	null
PFB0560w	conserved *Plasmodium* protein, unknown function	3.7	7.7	22	477.6	8.3	null
PF10_0045	conserved *Plasmodium* membrane protein, unknown function	3.7	7.9	16	357.8	8.9	null
MAL13P1.26	conserved *Plasmodium* protein, unknown function	3.6	12.5	25	418.3	7.1	null
PFF0935c	conserved *Plasmodium* protein, unknown function	3.5	12.6	33	433.4	9.5	null