UNIVERSITY OF LIVERPOOL

A study of the mechanism of action and resistance of artemisinin antimalarials in *Plasmodium falciparum*

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

Matthew Phanchana

November 2016

Liverpool School of Tropical Medicine

Abstract

A study of the mechanism of action and resistance of artemisinin antimalarials in *Plasmodium falciparum*

Matthew Phanchana

Malaria remains a global health and economic issue affecting nearly half of the world's population. In the past decade, effective chemotherapy and vector control have been the major interventions used to control and reduce the burden of malaria. However, resistance to antimalarial drugs and insecticides is compromising the control and treatment strategies and the goal to eliminate malaria. In most malaria endemic countries, artemisinin combination therapies (ACTs) are the first line treatment for uncomplicated *Plasmodium falciparum* malaria, the most lethal cause of malaria. Despite the widespread use of artemisinin-based therapies, the mechanism of action of this class of drug remains elusive. Emergence of resistance to ACTs in South East Asia is a global concern for drug efficacy. In this thesis, a click chemistry coupled with mass spectrometry (MS) proteomics approach was used to identify the molecular targets of artemisinin in various stages of *P. falciparum* strain 3D7 and extensively applied to the candidate trioxolanes, a new class of fully-synthetic artemisinin-like drugs. Using artemisinin activity-based probes, a number of biological targets were identified, these targets derive from key biological pathways/process that include; haemoglobin metabolism, glycolysis, nucleic acid and protein biosynthesis, antioxidant defence and oxidative stress response. The identified fingerprint of biological targets was similar between semi synthetic artemisininin and fully synthetic next generation artemisinins. Identified biological targets were enriched with glutathionylated proteins, indicating that these proteins are vulnerable to endoperoxide antimalarial inhibition and loss of function. The shared protein targets or protein pattern of semisynthetic artemisinin and fully synthetic trioxolane suggest that they might share similar mechanism or action and, possibly, mechanism of resistance. This raises the concern of cross resistance between them. The ring stage parasites which showed the least sensitivity to artemisinins and associated with resistance to artemisinins have much less proteins identified, including the absence of proteins in haemoglobin metabolism and reduction in proteins of major pathways. These findings support the working hypothesis that artemisinin is most effective against later stages of the parasite in line with the activity of haemoglobin digestion, the main activator of artemisinins and other endoperoxides. The reduced sensitivity during the ring stage is possibly due to less activation of artemisinin. A whole genome sequence comparative approach was undertaken with parasites displaying phenotypic artemisinin resistance (slow clearance phenotype) derived from an experimental *in vivo* model of infection. Parasite genes that we correlated to the slow clearance phenotype included genes involved in the unfolded protein response pathway consistent with recent models of parasite resistance to artemisinins. The results presented in this thesis using chemical biology and omics technologies, have contributed to our understanding of the mechanism of action and resistance of endoperoxides and offer future research directions to study this important class of antimalarials.

Acknowledgements

I would like to acknowledge my supervisors, Prof Steve Ward and Prof Giancarlo Biagini, for their tireless guidance and all opportunities they have given. Working with malaria has never been easy but challenging. They have scientifically and personally supported me during the past years helping me to become a better me.

I also thank my advisory panel members, Dr Robert Harrison and Dr Simon Wagstaff for their constructive comments and reflections. I would like to extend my thanking to Dr Simon Wagstaff and the bioinformatics team; Dr John Archer, Dr Enrique Salsedo-Sora and Dr Gareth Weedall, for their contribution on genome analysis. I am indebted to Dr Hanafy Ismail for his helping hands and all the click techniques and cautions, Dr Gavin Laing and Dr Gemma Molyneux for their expertise on MS proteomics.

Life in the lab would not be this enjoyable without them! Eva, Arturas, Saif, Eilidh, Vera, and Mameow. The past and present members of the SAW/GB lab, Ashley, Paul, Ally, Jill, Grazia, Dave, Gavin, Ricky, Alison, Mary, and Angela, have made the lab became more than just a lab but a great environment and company. I am delighted to be your mate!

I owed a huge thank to Prof Mathirut Mungthin and Prof Peerapan Tan-ariya who opened the world of malaria research to me. All at the Parasitology group at PCM, you were and are always good friends and very supportive, especially HaMuay and Bee for laying the first stone of malaria work for me. Furthermore, I am thankful for Mahidol-Liverpool Chamlong Harinasuta PhD scholarship, and the selection panel for their beliefs in me.

Another life outside the school was just amazing because of these mates, Gluay, Oey, Yin, Fha, Smart, Mai, Big, Tor, T-family, Gam, Bear, Yayee, Pu, Nham, and the past and present members of Liverpool Thai Society. All the good and bad time I spent with you was precious. I would like to thank Lee for her generosity and all at Chaba Chaba for being good colleagues and sharing the great time with me.

Lastly, I would like to express my love and thank to my parents for their love, prayers and supports for all the thing that I do.

Contents|

Chapter 1

Introduction

1.1 Malaria

Malaria is among the highest mortality and morbidity rate with HIV/AIDS and tuberculosis for infectious diseases. Malaria is also the leading cause of deaths in post-neonatal along with pneumonia, diarrhoea, and injuries (WHO, 2016b). It is estimated 3 billion people are living in area at risk of malaria infection (WHO, 2015b). *Plasmodium* spp. are causative pathogens of malaria with *Plasmodium falciparum* responsible for most of death and severe malaria cases. Malaria prevalence is not limited to human only, but extended to other animals; avian, reptile, and other mammals. Distribution of malaria is in the tropic and sub-tropic regions and limited to distribution of its vectors *Anopheles* mosquitoes.

Malaria can be traced to the prehistoric period around 15-45 million year ago, when the first evidence of malaria was dating; the dating was uncertain due to the dating methods. *Plasmodium dominicana* was identified from mosquito fossil embedded in tertiary Dominican amber. It is also believed that the host would be galliformes (Poinar, 2005). However, human malaria is dated back to 2700 BC in the Chinese Canon of Medicine by Nei Ching. In the past, malaria was described as periodic fever or seasonal fever associated with splenomegaly, spleen enlargement (Cox, 2010). Although many historical documents recorded symptoms resembling later known malaria, the discovery of malaria parasite was in 1880 by French surgeon Charles Louis Alphonse Laveran. Not long after, Camillo Golgi had showed that there are more than one malarial species in 1886. Nearly twenty years later, British medic Sir Ronald Ross demonstrated that mosquitoes is the malaria vector in 1897. These discoveries led to the Nobel Prize in physiology or medicine awarded to Sir Ronald Ross, Camillo Golgi, and Charles Laveran in 1902, 1906, and 1907, respectively.

The distribution of malaria (Figure 1.1) is largely superimposed with the vector distribution as expected in typical vector borne diseases, and also linked with poverty (Gallup and Sachs, 2001). Other major factors affecting malaria distribution are including temperature and rainfall (Hay et al., 2004).

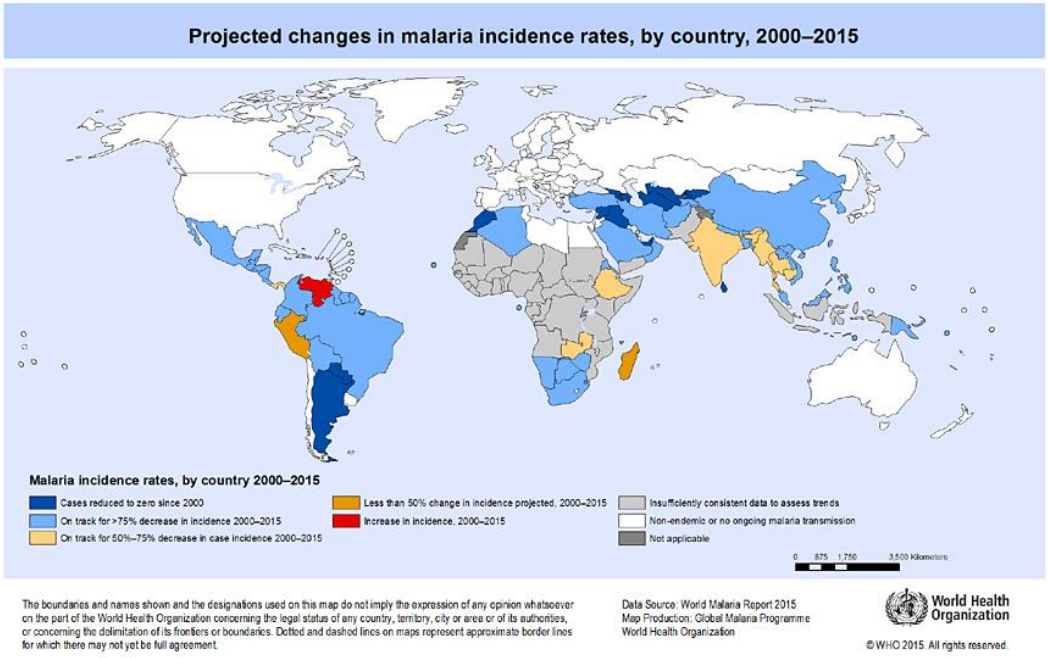


Figure 1.1 Malaria distribution change during 2000-2015 (taken from WHO)

1.1.1 Life cycle

Malaria parasites have a complex life cycle which requires two hosts to complete; an invertebrate host where sexual cycle takes place and a vertebrate host where asexual cycle occurs (Figure 1.2). Most human malaria parasites share similar life cycle with the exception of *Plasmodium vivax* and *Plasmodium ovale*. These species exhibit the dormant hepatic stage called **'hypnozoite'**, which is not detectable and is the cause of relapse malaria. Hypnozoite stage was first shown for relapse malaria in rhesus malaria *Plasmodium cynomolgi* in 1948, and that exoerythrocytic cycle exists (Shortt and Garnham, 1948).

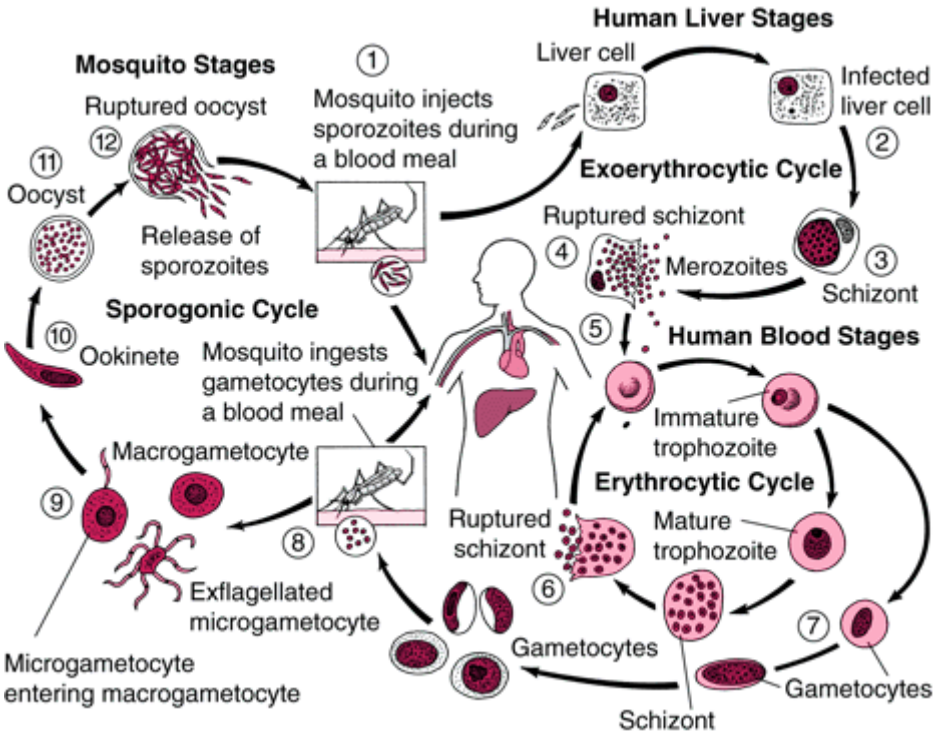


Figure 1.2 Plasmodium falciparum life cycle (taken from http://www.merckmanuals.com)

Female *Anopheles* mosquitoes, invertebrate vector, carrying infectious sporozoites stage transmit sporozoites to human during their blood meal. Once sporozoites are released into blood stream, they migrate to the liver, infect hepatocytes, and reproduce a high number of progeny or **'merozoites'** via the process called **'schizogony'**. In *Plasmodium vivax* and *Plasmodium ovale*, some sporozoites undergo dormancy and become hypnozoites. Mature merozoites in the vesicle called **'merosome'** release from the host hepatocyte by budding into blood vessels (sinusoids). This process coincides with inhibition of exposure to phosphatidylserine on the outer leaflet of host plasma membrane, preventing host immune response (Figure 1.3) (Sturm et al., 2006).

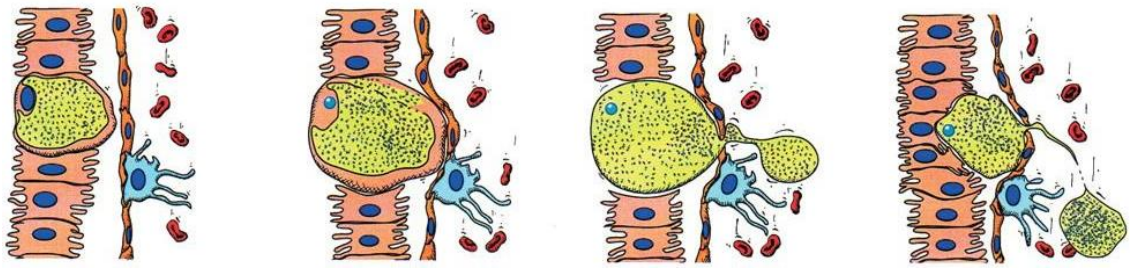


Figure 1.3 Hypothesis of the release of merosome to sinusoids (doi: 10.1126/science.1129720)

These merosomes travel to the lung, become smaller as they travel through blood vessels, and accumulate in the lung. It is believed that sequestration of merosomes in pulmonary capillaries ensures infectiveness and provides protection from host immune system. Merozoites are released from the

merosomes and invade host red blood cells (RBCs) marked the beginning of intraerythrocytic cycle (Baer et al., 2007).

Intraerythrocytic cycle is responsible for pathogenicity of malaria. Once merozoites invade host RBCs via a complex process (see Tardieux and Baum (2016) for review), they use host materials for their growth and development. The first stage of intraerythrocytic cycle is called **'trophozoite'**, but early trophozoite appearance is similar to ring shape, where cytoplasm is formed the ring part and nucleus is formed the setting part of the ring, when stained with Giemsa stain, so it is also called 'ring-shaped trophozoite' or **'ring stage'** (which will be used throughout to describe this early stage of trophozoite). Mature trophozoite is greater both in term of size, due to accumulation of biomass, and metabolisms. Then mature trophozoites proceed to schizogony process to produce a number of merozoites within the cells called **'schizont'**. Merozoites egress from host RBCs via a complex process.

This intraerythrocytic stage lasts between 42-49 h for *Plasmodium falciparum, Plasmodium vivax* and *Plasmodium ovale*, 72 h for *Plasmodium malariae*, and 24 h in *Plasmodium knowlesi*. Some trophozoites will commit gametogenesis upon stimulation of uncertain mechanism to become **'gametocyte'**, the transmission stage from human to *Anopheles* vectors, which will be picked up by the mosquitoes.

Microgametocyte (male) and **microgametocyte** (female) gametocytes are picked up by mosquitoes and move to the midgut, where male gametocytes undergo **'exfragellation'** process producing exflagellated microgametocytes which fuse with macrogametocytes and form the zygote called **'ookinete'**. Ookinetes migrate, in the special left-handed helical motion (Kan et al., 2014), to and invade mosquitoes midgut epithelial cells where they form **'oocyst'**. Within the oocyst, multiplication steps generate number of infectious sporozoites which later egress from the oocysts and migrate to mosquitoes' saliva gland. The whole life cycle of *Plasmodium* spp. completed when sporozoites are transmitted to human at the next blood meal.

1.2 Antimalarial drugs

Antimalarial drugs are one of key players in disease management and control. There are only a handful of licensed antimalarial drugs used in the clinic nowadays, although a remedy for malaria was introduced in the 17th century and established the modern years of antimalarial (Wells et al., 2015). In the past, plant extracts were the major sources of fever relief including cinchona extract and sweet wormwood extract, later known as cinchona alkaloids and artemisinin, respectively. In the past decades, only a few new antimalarials have been licensed (Figure 1.4).

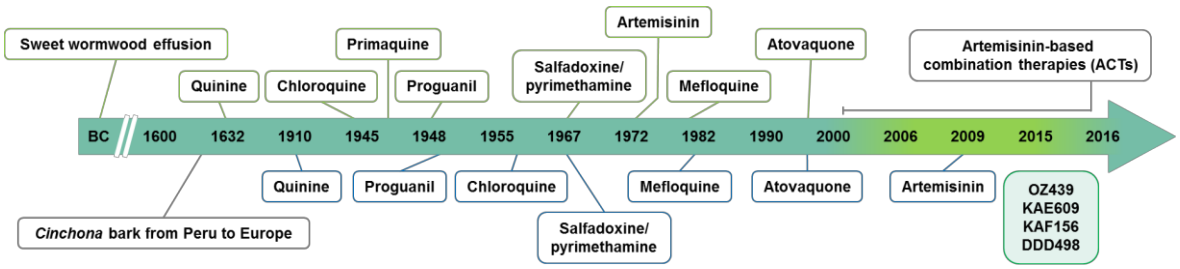
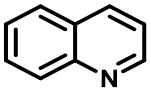


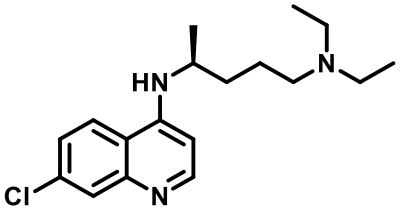
Figure 1.4 The timeline of antimalarial drug introduction and its resistance. Introductory and emerging resistance year of each antimalarial drug is shown in green and blue box, respectively.

1.2.1 4-aminoquinolines

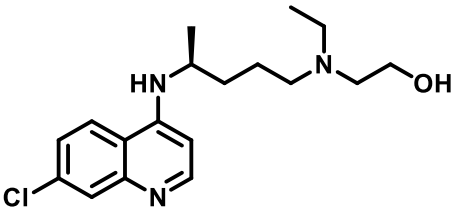
4-aminoquinoline is a group of antimalarial with quinoline as a core structure with amine group at position 4 (Figure 1.5) (Schlitzer, 2008). The 4-aminoquinoline chloroquine (Figure 1.5) is considered one of the most important antimalarial drugs owing its potent efficacy and low cost. Chloroquine had a pivotal role, together with dichlorodiphenyltrichloroethane (DTT), in the massive reduction of malaria burden in the 20th century (Committee on the Economics of Antimalarial Drugs, 2004). Mechanism of action of 4-aminiquinoline is believed, yet not fully understood, to be inhibition of crystallisation of haemozoin or malaria pigment.



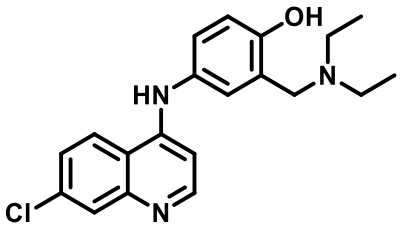
Quinoline



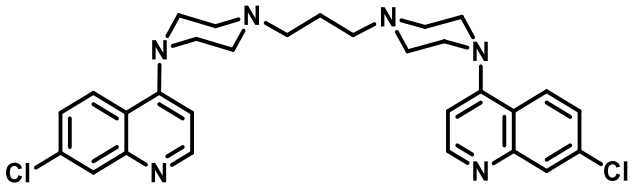
Chloroquine



Hydroxychloroquine



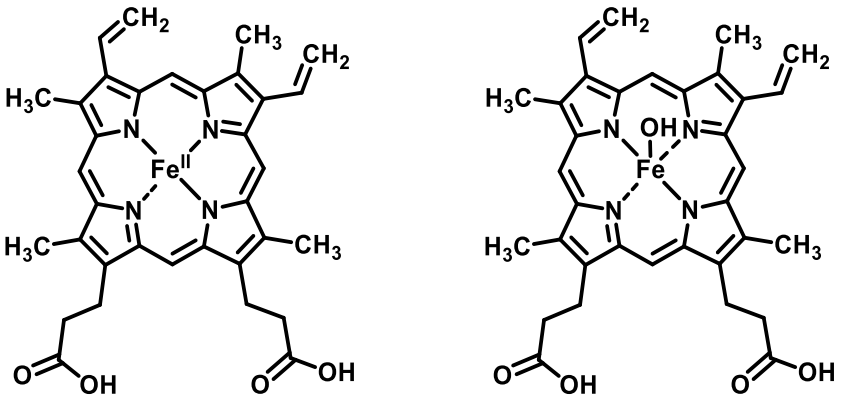
Amodiaquine



Piperaquine

Figure 1.5 4-aminoquinolines

During parasite intraerythrocytic cycle, parasites digest host haemoglobin for amino acids in acidic food vacuole (~pH 5) and generate free haem (Figure 1.6) as a by-product which is toxic to the parasites. Parasites overcome this toxic haem by crystallising haem to nontoxic insoluble haemozoin crystal, or malaria pigment (Figure 1.7) (Sullivan et al., 1996a).



Haem Haematin

Figure 1.6 Haem and haematin structure

This haemozoin crystallisation process is driven by histidine-rich proteins (HRPs) (Sullivan et al., 1996a) and/or haem detoxification protein (HDP) (Jani et al., 2008). Chloroquine, a singly or doubly-protonated form, accumulates in food vacuole of the parasites as a result of a weak base interaction (Yayon et al., 1985), but not the sole mechanism. The accumulation of protonated forms of chloroquine is essential and dependent on haem binding (Bray et al., 1999, Bray et al., 1998). Inhibition effect of haem crystallisation is due to binding of inhibitor on haem or haemozoin, rather than to the enzymes, preventing further crystallisation of haemozoin (Sullivan et al., 1996a), leading to the accumulation of toxic free haem and the parasite death (Slater and Cerami, 1992). In 2014, Erin L. Dodd and D. Scott Bohle successfully demonstrated that gallium protoporphyrin IX, haem analogue, forms complex with chloroquine confirmed by 1H NMR (Dodd and Bohle, 2014). This evidence supports haem-chloroquine complex hypothesis (Figure 1.8). Prior to this structural work, Sullivan and colleagues showed that ^{3}H-quinolines co-localised with haemozoin in parasite vacuole by autoradiography technique providing first evidence of haemozoin-quinoline complex (Sullivan et al., 1996b).

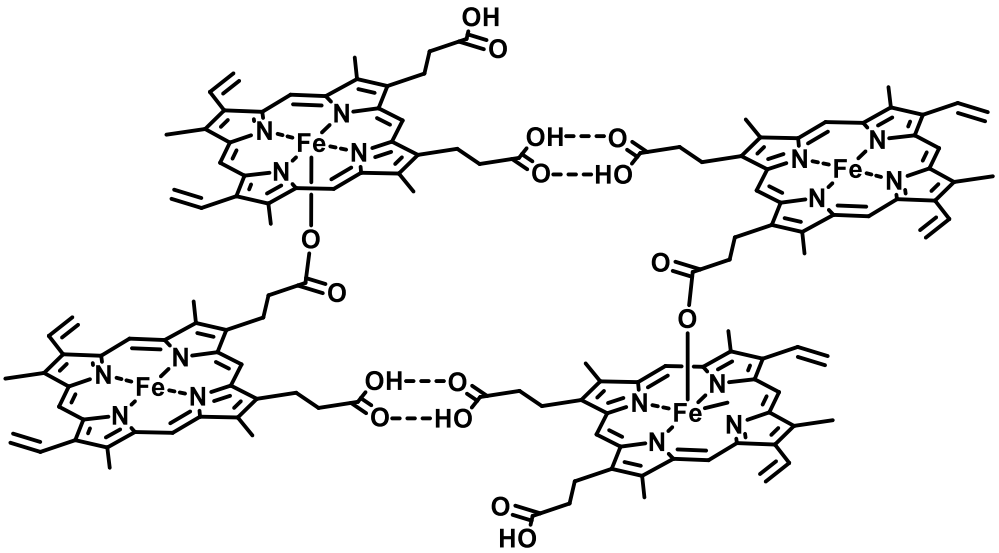


Figure 1.7 Haemozoin (β-haematin) structure

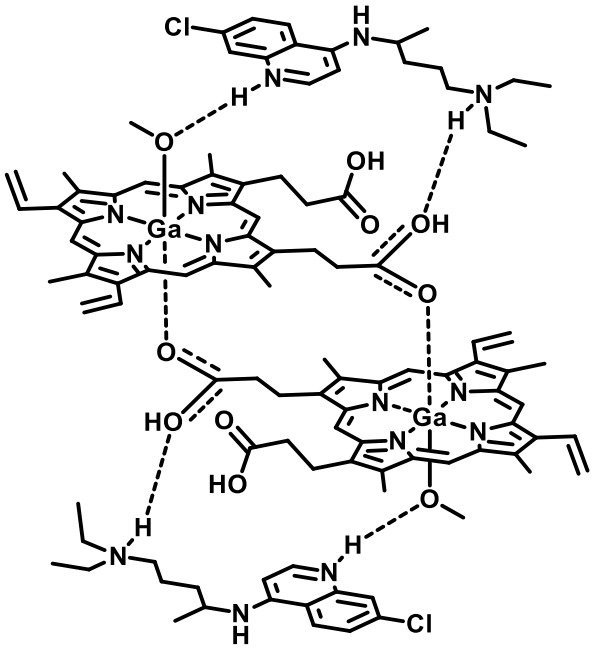


Figure 1.8 Haem-chloroquine complex

There is also evidence that chloroquine inhibits aspartic protease in vacuole at high concentration (~30 mM), however, this concentration is much higher than a typical pharmacological concentration of chloroquine in vacuole of 1-5 mM.

Amodiaquine (AQ) (Figure 1.5), aromatic substitution in quinoline side chain, is effective against low-level chloroquine-resistance parasites. It is used as monotherapy or with artesunate in artemisinin-based combination therapies (ACTs). There is a report of amodiaquine adverse effect, hepatotoxicity

and agranulocytosis, in a prolonged use for chemoprophylactic purpose or with HIV patients on zidovudine and/or cortrimoxazole, so it is not recommended to use amodiaquine with these patients (WHO, 2015a). In some western countries, amodiaquine is no longer available.

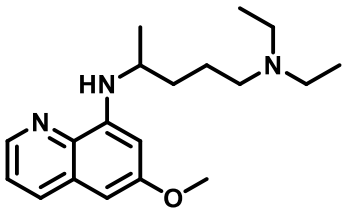
Piperaquine (PPQ) (Figure 1.5), bisquinoline, is reintroduced for malaria treatment in combination with DHA after its effectiveness declined due to the emergence of resistance in 1980s. It was first synthesised in 1960s by Chinese scientist. Clinical study has shown DHA-PPQ has higher efficiency and good safety profile for prevention of malaria in pregnancy specially when on monthly regimen (Kakuru et al., 2016). Another clinical study conducted in Cambodia also showed high efficiency and safety of DHA-PPQ in uncomplicated malaria with 96.9% cure rate at 28-day (Denis et al., 2002). It is not recommended to use DHA-PPQ in patients with congenital QT prolongation as high-dose of PPQ increases the risk of QT prolongation, by the same degree as chloroquine, especially when on high-fat meals. However, there is no evidence showing piperaquine-related cardiotoxicity (WHO, 2015a).

1.2.2 8-aminoquinolines

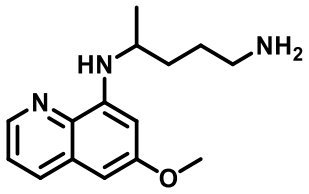
The 8-aminoquinolines (Figure 1.9) are antimalarial drugs with quinoline core structure with amine group at position 8 of quinoline ring. It was proposed that 8-aminoquinolines have similar mechanism of action to 4-aminoquinoline by inhibiting crystallisation of haematin (Vennerstrom et al., 1999).

Primaquine (Figure 1.9) is a sole licensed antimalarial effective against sexual stage of parasites and hypnozoites of *Plasmodium vivax* and *Plasmodium ovale*. For prevention of relapse in vivax and ovale malaria, 14-day course of 0.25-0.5 mg/kg bw daily is recommended in addition of ACT or chloroquine, where applicable, for blood stage treatment. A single-dose of 0.25 mg/kg is also used for transmission prevention in low-transmission area for *falciparum* malaria. Primaquine exhibits adverse effect of haemolysis in patients with G6PD deficiency, so it is recommended to check for G6PD status before prescription of 14-day course of primaquine but not necessary for single-dose prescription (WHO, 2015a, WHO, 2014a). Primaquine is a replacement of pamaquine which has high toxicity and withdrawn for clinical use.

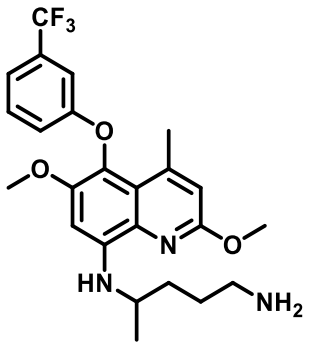
Tafenoquine (Figure 1.9) is currently in Phase III clinical trial evaluating for its safety for human use. Tafenoquine has proved its effectiveness against sexual stage parasite and hypnozoites (Vennerstrom et al., 1999, Li et al., 2014, Shanks et al., 2001). Benefit of tafenoquine over primaquine is its longer plasma half-life, ~2 weeks (Brueckner et al., 1998), thus a single dose of tafenoquine is sufficient for prevention of relapse malaria (Llanos-Cuentas et al., 2014). Like primaquine and other 8-aminoquinoline drugs, tafenoquine exhibits haemolysis effect in patients with G6PD deficiency.



Pamaquine



Primaquine



Tafenoquine

Figure 1.9 8-aminoquinolines

1.2.3 Arylaminoalcohols

Arylaminoalcohols is another class of important antimalarial drugs. It became the first class of antimalarial compounds isolated from plant extracts cinchona in 1820 by French chemists (Committee on the Economics of Antimalarial Drugs, 2004).

Quinine (Figure 1.10) is antimalarial drug derived from plant cinchona and has long been used for treatment of malaria. Quinine is currently not recommended for first line treatment of *Plasmodium falciparum* malaria but remains effective for treatment of other malaria and for malaria in first trimester of pregnancy, in combination with clindamycin (WHO, 2015a). However, quinine is used in combination with tetracycline for treatment of *Plasmodium falciparum* malaria when ACTs failed (Noedl et al., 2010) (Prof. Srivicha Krutsood, MD; personal communication). The recent study reveals quinine interferes with serotonin biosynthesis by either competitive binding with active site of tryptophan hydroxylase (TPH2), rate-limiting enzyme of serotonin biosynthesis, or binding to tryptophan transporter. This explains the adverse effect of quinine in patients with low plasma tryptophan (Islahudin et al., 2014).

Lumefantrine (Figure 1.10) is only used in combination with artemether (Coartem®) in ACTs as monotherapy of lumefantrine could develop resistance, therefore lumefantrine was never used in monotherapy (WHO, 2015a). The advantage of lumefantrine is its long elimination half-life of 6 days with peak plasma concentration at 4-6 h. This feature complements well the effect of artemether in clearing

parasites (Djimde and Lefevre, 2009). It has been reported that lumefantrine bioavailability can be increased by co-administration with fatty meal, milk is sufficient (Djimde and Lefevre, 2009).

Mefloquine (Figure 1.10) is used in combination with artesunate in ACTs for treatment of uncomplicated *falciparum* malaria, mostly in Asian countries. It is also used for chemoprophylaxis for malaria. However, its side effects including insomnia, depression, panic attack and hallucination limit its use for many circumstances (Mayxay et al., 2006).

Halofantrine (Figure 1.10) is not currently used for malaria treatment and/or prophylaxis in most countries due to its safety profile (prolonged QT-interval as in chloroquine and piperaquine) and lower efficiency than alternatives (Bindschedler et al., 2002).

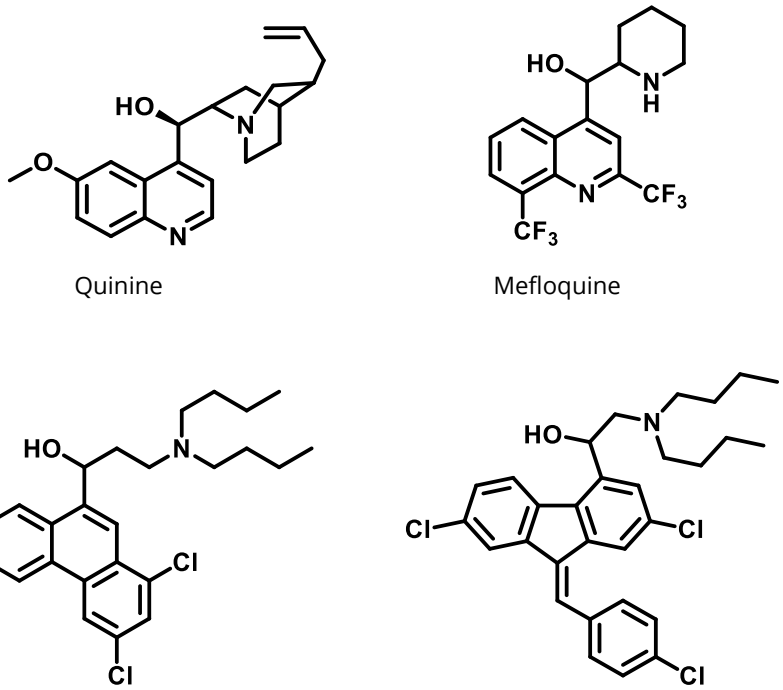


Figure 1.10 Arylaminoalcohols

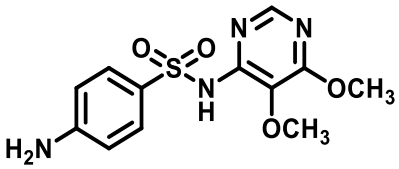
1.2.4 Antifolates

Antifolates (Figure 1.11) are used to inhibit folate metabolism mainly dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR). The effect of antifolate itself is not very potent hence it is used in combination to synergise the effect of other antimalarial compounds. Antifolates were not in a mainstream research focus as no new antifolate has been introduced to the clinical use or trial since its first introduction in 1940 (Nzila, 2006).

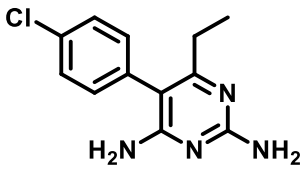
Cycloguanil (Figure 1.11) is active metabolite of proguanil inhibiting DHFR of the parasite by unclear mechanism. Proguanil is prescribed in combination with atovaquone (mitochondrial electron transport inhibitor) in a fixed-dose tablet Malarone®. Malarone® is the major FDA approved drug recommended for malaria chemoprophylaxis (Nixon et al., 2013).

Pyrimethamine (Figure 1.11) is DHFR inhibitor and the most widely used antifolates, usually in combination with sulfadrug sulfadoxine (SP). However, resistance to pyrimethamine and other antifolates reduce the efficacy of these drugs. The new regimen is to use SP in combination with artesunate in ACTs to slow the rate of resistance (White, 1999).

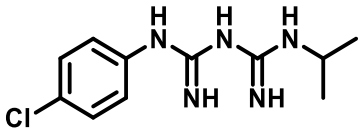
The use of some antifolate, dapsone and SP, is limited due to its haemolysis side effect on G6PD deficiency patients. However, SP is recommended for intermittent preventive treatment of malarial in pregnancy (IPTp) in highly endemic areas (WHO, 2015a).



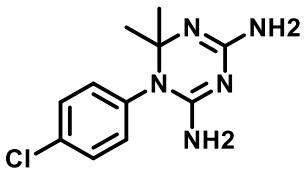
Sulfadoxine



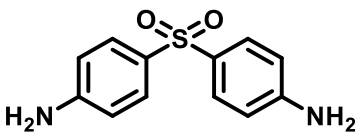
Pyrimethamine



Proguanil



Cycloguanil



Dapsone

Figure 1.11 Antifolates

1.2.5 Inhibitors of the respiratory chain

Atovaquone, the naphthoquinone, (Figure 1.12) is the only available inhibitor of electron transport in mitochondrion and is only prescribed in combination with proguanil as Malarone® because resistance to atovaquone is easily developed when used in monotherapy. Atovaquone and proguanil, or its active metabolite cycloguanil, have synergistic effect by disrupting mitochondrial membrane potential. Atovaquone/proguanil also active against early liver stage parasite, so it is used for prophylaxis. Due to its low solubility in water, it is recommended to supply with fatty meal. However, the cost of Malarone® is relatively high compared to other antimalarial drugs limiting its use in many endemic areas, it was estimated that Malarone® will become widespread after its patent expired in 2013 (Nixon et al., 2013) .

The mechanism of atovaquone was suggested by the mutations in cytochrome *b* gene. It was first found that atovaquone resistance is mediated by mutations in cytochrome *b* of cytochrome bc_1 complex (Srivastava et al., 1999). The binding site of atovaquone was confirmed by structural analysis showing that atovaquone binds to Q_o site in cytochrome *b* (Birth et al., 2014). These findings led to the development of new inhibitors for cytochrome bc_1 complex rationalised for Q_i site and could overcome atovaquone resistance (Capper et al., 2015).

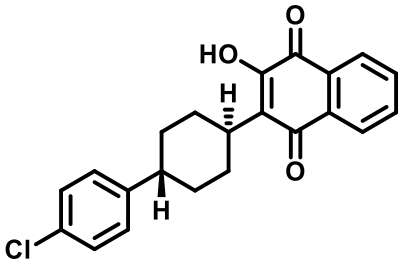
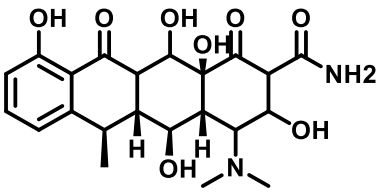


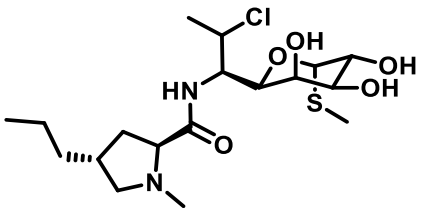
Figure 1.12 Atovaquone

1.2.6 Antibiotics

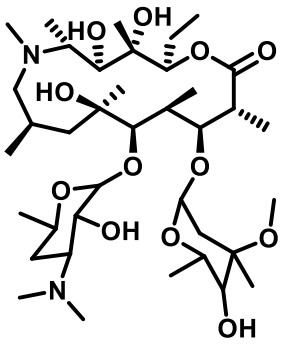
Antibiotics (Figure 1.13) for malaria treatment were used in combination with other antimalarial drugs e.g. ACTs. This is to address possible bacterial coinfection, usually in severe malaria. It is suggested that broad-spectrum antibiotic should be given to all malaria in children with suspected severe malaria (WHO, 2015a). In some situation, doxycycline can be given for chemoprophylaxis of malaria (Tan et al., 2011).



Doxycycline



Clindamycin



Azithromycin

Figure 1.13 Antibiotics used for treatment of malaria

1.2.7 Mechanisms of antimalarial resistance

1.2.7.1 Resistance to quinolines.

Among all antimalarial drugs, the 4-aminoquinolines have the best established mechanism of resistance, although not yet fully understood. A widely accepted mechanism of resistance to chloroquine is the reduced access to the haematin (Bray et al., 1998, Bray et al., 1999). The protonated forms of chloroquine are accumulated within the food vacuole of the parasites and led to haem-chloroquine complex formation preventing further crystallisation to a non-toxic haemozoin (Slater and Cerami, 1992, Slater, 1993, Sullivan et al., 1996b). The genetic determinant of chloroquine resistance is the chloroquine resistance transporter (*Pf*CRT) gene (Fidock et al., 2000). The resistance mechanism is mediated by efflux of chloroquine from the food vacuole by the CRT protein localised on the food vacuole membrane (Lehane and Kirk, 2008).

Another important genetic determinant for other quinolones compounds, including piperaquine, mefloquine, amodiaquine, lumefantrine, and halofantrine, is multidrug resistance protein 1 (*Pf*MDR1). It was shown that polymorphisms in *Pf*MDR1 are associated with resistance to multiple quinolone compounds but chloroquine (Reed et al., 2000). Furthermore, gene amplification of *Pf*MDR1 gene is believed to play a role in resistance to some quinoline compounds. It was demonstrated that gene copy number of *Pf*MDR1 is associated with *in vivo* resistance to mefloquine (Price et al., 2006). More recently, it has been proposed that polymorphisms in *Pf*CRT and *Pf*MDR1 have an interlink role in modulating parasite susceptibility to various antimalarial drugs including artemisinins and their clinical partner drugs, amodiaquine, lumefantrine, and mefloquine (Veiga et al., 2016).

1.2.7.2 Resistance to antifolates

The two enzymes targeted by antifolate antimalarials are dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR). As these antimalarials have a direct effect on the enzymes, resistance to antifolates is associated with point mutations in the genes coding enzymes. Resistance to DHFR antifolates is mediated by mutations in DHFR gene. Several point mutations including Ser180Asn, Asn51Ile, Cys59Arg, and Ile164Leu were reported to directly associate with DHFR antifolate resistance, and thus could be closely monitor for resistance (Gregson and Plowe, 2005, WHO, 2015a). DHPS resistance is linked with variations of 5 amino acids (436, 437, 518, 540, and 613) (Triglia et al., 1998). However, it was also demonstrated that folate and folate derivative levels affect the parasite sensitivity to antifolates (van Hensbroek et al., 1995, Wang et al., 1997).

1.2.7.3 Resistance to naphthoquinone

Resistance to naphthoquinone atovaquone was quickly acquired both *in vitro* and *in vivo* (Gassis and Rathod, 1996, Srivastava et al., 1999). It also observed from the clinical trials that relatively high

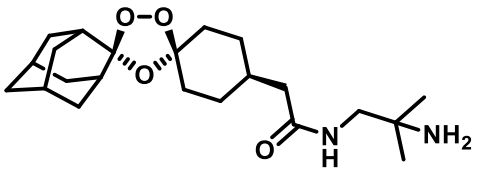
treatment failure rate is associated with atovaquone (Chiodini et al., 1995, Looareesuwan et al., 1996). Therefore, it is always used in combination with proguanil for a synergistic effect to reduce the failure rate. The target of atovaquone is the parasite Q_o site of the cytochrome bc_1 complex. The mutations in the Q_o site, especially near the highly conserved PEWY region, confer atovaquone resistant phenotype. In the clinic, the variations in cytochrome bc_1 at codon 268 are associated with Maralone® treatment failure (Fivelman et al., 2002). However, atovaquone resistant phenotype is with a fitness cost of reduced activity of the cytochrome bc_1 complex (Fisher et al., 2012).

1.3 Development of new antimalarial drugs

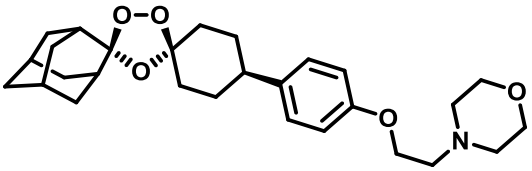
New antimalarial drugs are urgently needed as drug resistance to most current antimalarial drugs have been reported. Medicine for Malaria Venture (MMV) was founded to advocate the development of new antimalarial drugs through the partnerships. The goal of new antimalarial drugs is to avoid cross resistance, so novel therapeutic targets are also required. Furthermore MMV has published the target product profile (TPP) and target candidate profile (TCP) as a guideline for desirable new antimalarial drugs (Burrows et al., 2013). In the past years, antimalarial drugs in the pipeline are endoperoxides, aurora kinase inhibitors, *Pf*PI4K inhibitors, and some other targets (MMV global portfolio, accessed from www.mmv.org).

OZ277 or arterolane (Figure 1.14) is the first generation trioxolane fully synthesised antimalarial candidate (Vennerstrom et al., 2004) received regulatory review process according to MMV portfolio. It is effective against chloroquine-resistance parasite from field isolates (Kreidenweiss et al., 2006).

OZ439 or artefenomel (Figure 1.14) is a second generation synthetic trioxolane antimalarial candidate finishing phase II clinical trial with a good safety profile and a high efficiency on a single dose regimen. No serious adverse effect has been reported (Phyo et al., 2016). OZ439 is a promising candidate for single-dose regimen because it exhibits long elimination half-life of 46-62 h with maximum plasma concentration at 4 h. OZ439 also exhibits superior prophylaxis property to mefloquine. Administration of single dose 30 mg/kg provides total protection (Charman et al., 2011).



OZ277



OZ439

Figure 1.14 Fully synthetic endoperoxide trioxolanes

DDD107498 or DDD498 (Figure 1.15) is a newly developed antimalarial candidate led by University of Dundee Drug Discovery unit (DDU) and MMV. Preclinical study shows that DDD107498 has good oral bioavailability and long plasma elimination half-life (19 h). These properties are important for the use in resource-poor area and are required by target product profiles (TPP) and target candidate profiles (TCP) set by MMV (Burrows et al., 2013). DDD107498 showed similar potency to artemisinin *in vitro* and superior to artesunate *ex vivo* for both *Plasmodium falciparum* and *Plasmodium vivax* from clinical isolates. DDD107498 also inhibits gamete formation both male and female revealed by membrane feeding assay and direct mouse-to-mouse *Plasmodium berghei* model (Baragana et al., 2015). The study also shows that DDD107498 inhibits parasite translation elongation factor 2 (*Pf*eEF2) which is essential for protein synthesis (Baragana et al., 2015). These properties support DDD107498 as a promising antimalarial candidate.

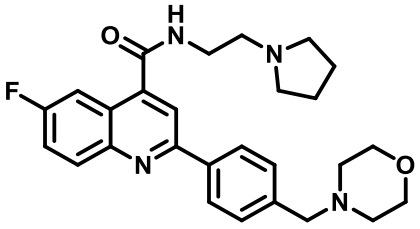


Figure 1.15 DDD107498

A novel class of antimalarial, imidazolopiperazines, was developed, and KAF156 (Figure 1.16) was selected from the lead optimised compounds. The development of KAF156 is led by the pharmaceutical company Novartis. KAF156 shows a promising characteristic for both therapeutic and prophylactic uses, and possibly transmission blocking (Kuhen et al., 2014). Preclinical study showed that *in vitro* induced-resistance to KAF156 is associated with SNPs in *Pf*CARL gene (Meister et al., 2011). Lately, other two

genes were reported with association to KAF156 resistant-phenotype, namely an UDP-galactose transporter (*Pf*UGT) and an acetyl-CoA transporter (*Pf*ACT) (Lim et al., 2016).

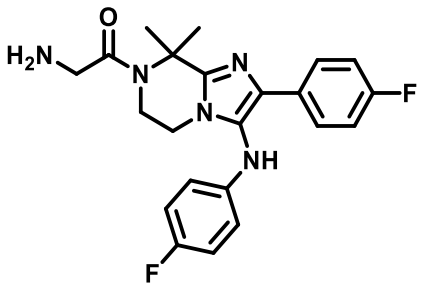


Figure 1.16 KAF156

Another antimalarial candidate developed by Novartis, Swiss THP and Wellcome, KAE609 (Figure 1.17), is currently assessed in clinical trials. The proposed mechanism of action of KAE609 is inhibition of protein synthesis as it was shown to reduce incorporation of radiolabelled Met/Cys within an h after exposure (Rottmann et al., 2010). Furthermore, the mutations in the P-type cation-transporter ATPase4 (*Pf*ATP4) encoding gene abolishes the antimalarial effect of KAE609 (Rottmann et al., 2010). The result from phase II clinical trial showed KAE609 has a high efficacy on both *Plasmodium falciparum* and *Plasmodium vivax* malaria, and is suitable for a once-daily regimen (White et al., 2014). In addition, a homology model study with yeast P-type ATPase (*Sc*PMA1) has shown that KAE609 directly inhibits the P-type ATPase activity, and computational molecular docking model also supports the hypothesis (Goldgof et al., 2016).

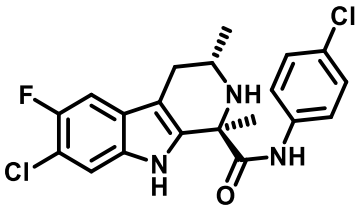


Figure 1.17 KAE609

1.4 Prevention and control

According to WHO world malaria report 2015, malaria prevention strategies involve vector control, chemoprevention, and case management. These strategies can be achieved by cost-effective interventions.

1.4.1 Vector control

Vector control plays a major role in a vector borne diseases control and prevention, cutting through the vector cycle can intervene and reduce disease burden in endemic areas very effectively. Main control measures for malaria vector control are using of insecticide-treated nets (ITNs) and indoor residue spraying (IRS).

ITNs are a protective barrier for human. ITNs have been proved to be superior to untreated nets for vector borne disease prevention. Main difference is seen when there is a hole in a bed net. Insecticide on the net kills, and repel, mosquitoes and other insects. This reduces the number of mosquitoes in house and also provides community-wide protection, when community adopted this intervention. Since 2000, ITN usage has been greatly increased in Africa (Figure 1.18). It is estimated that 55% of population in sub-Saharan Africa sleeping under the ITN in 2015, increasing from 2% in 2000. However, coverage goal is universal (100%).

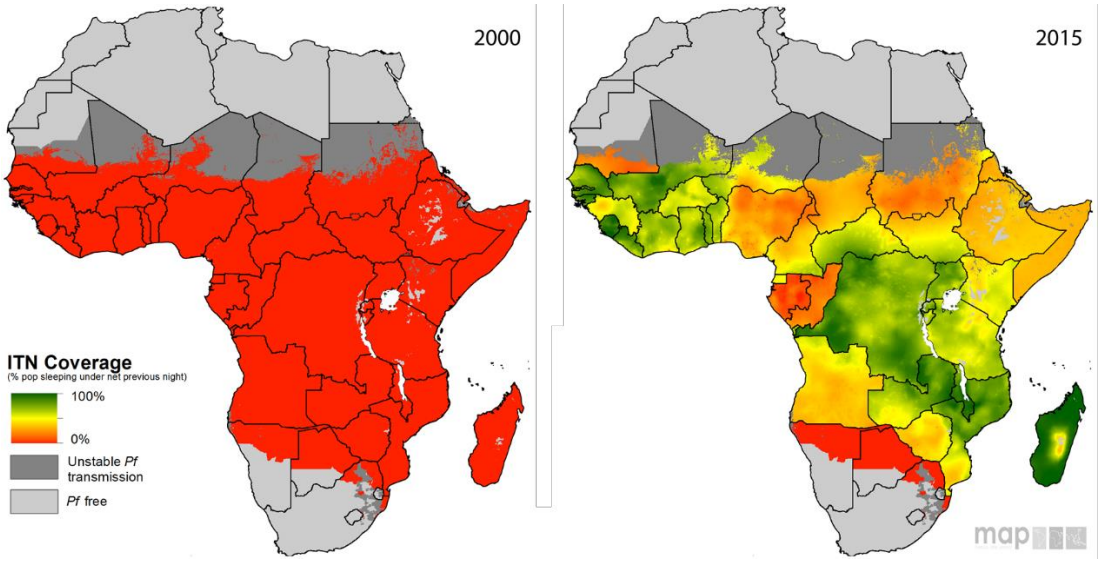


Figure 1.18 Insecticide-treated nets (ITNs) coverage in Africa in 2000 (left) and 2015 (right). Use under the Creative Commons Attribution 3.0 Unported License.

The first generation of ITNs provides 6-12 months protection, then reduces its protective power and requires re-treatment with insecticide. This process requires government involvement and is not normally carried out. Next generation ITNs, so called long-lasting insecticide-treated nets (LLINs), last

much longer than ITNs (up to 5 years), requiring no re-treatment and so less frequent distribution of the nets. LLINs rely mainly on pyrethroids class of compounds and incorporated into the net fibre given its long-lasting effect. Newer generation of LLINs uses piperonyl butoxide (PBO) and pyrethroid treatment. PBO itself is not an insecticide but synergists the effect of pyrethroid by sequester the pyrethroid effect until it is within the mosquitoes (Tungu et al., 2010).

Indoor residue spraying (IRS) is another factor in the control strategy. In the past dichlorodiphenyltrichloroethane (DDT) was extensively used before it was opted out from the control interventions due to health and environmental concerns (Chapin and Wasserstrom, 1981). However, the use of DDT and other control interventions had reduced malaria prevalence dramatically in the 1950s as part of WHO eradication program (Najera et al., 2011). Nowadays, the major classes of compounds for IRS are pyrethroids, organochlorines, organophosphates, and carbamates (WHO, 2012). IRS is still rolling in the endemic areas, however, insecticide resistance to pyrethroids affected the coverage of IRS due to less affordable costs of non-pyrethroids insecticides (WHO, 2015b).

1.4.2 Chemoprevention and malaria vaccine

Despite the fact that malaria vaccine is not currently available, chemoprevention for malaria is achieved through antimalarial drugs in some risk population i.e. in pregnancy. Intermittent preventive treatment in pregnancy (IPTp) and infant (IPTi) are to reduce the burden of malaria in pregnant women and infants. Sulfadoxin-pyrimethamine (SP) itself or in combination with amodiaquine is recommended in for those risk groups in highly endemic area (WHO, 2015a, WHO, 2015b).

The development of malaria vaccines is a great challenge due to antigenic variation of the parasites (Schwartz et al., 2012, Arama and Troye-Blomberg, 2014). The parasites have a sophisticated genetic regulation for surface antigen representation and the antigen switching facilitates host immune evasion (Kyes et al., 2007). The variation in antigen along with the parasite complex life cycle deter the vaccine development (Crompton et al., 2010). The most advanced malaria vaccine candidate is RTS,S from GlaxoSmithKline (GSK) and is receiving WHO policy and regulation reviewing process following the completion of phase III clinical trial in the risk populations in Africa (RTS, 2015). Other malaria vaccine progressing to clinical trials are including the sporozoite-based vaccines (*Pf*SPZ) and the modified vaccinia virus Ankara (MVA) with adenovirus vector encoding liver stage antigen (ME-TRAP) or (MVA ME-TRAP), according to the WHO rainbow table (WHO, 2016a) (Figure 1.19).

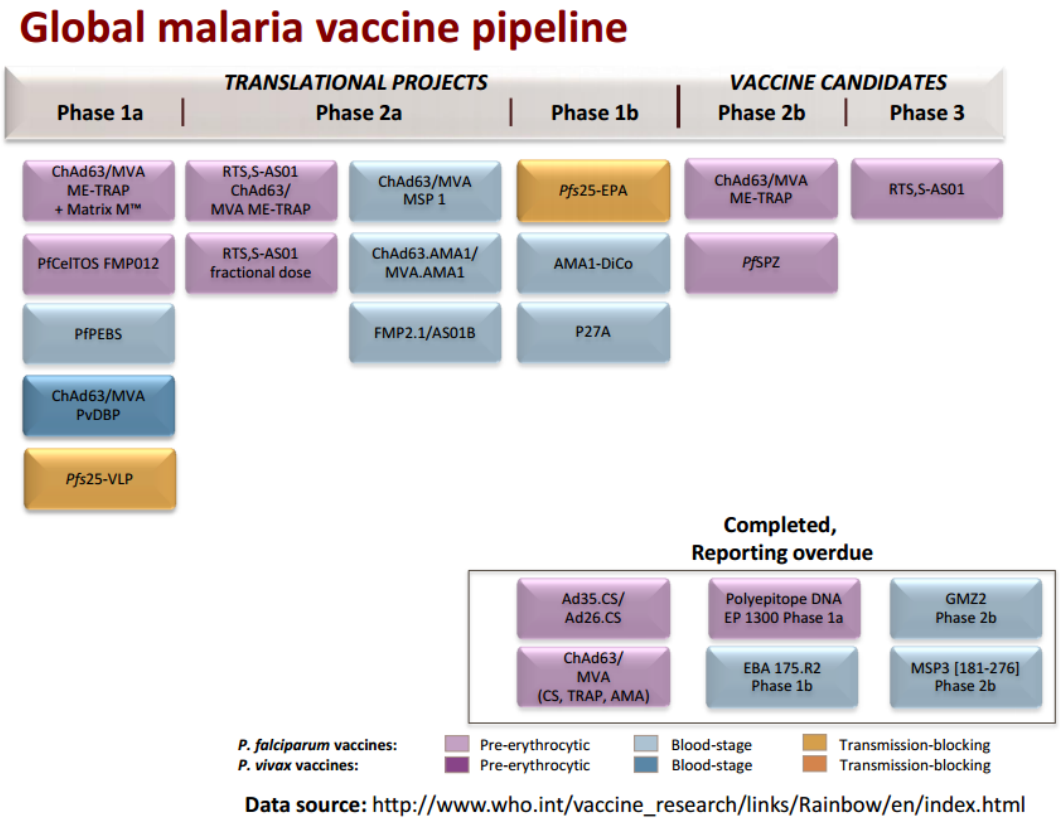


Figure 1.19 WHO rainbow table (accessed on 25th November 2016)

The RTS,S vaccine was developed based on the central repeat region of *Plasmodium falciparum* circumsporozoite protein and recombinantly (CSP) produced with hepatitis B surface antigen (HBsAg) in the yeast systems (Gordon et al., 1995). The RTS,S vaccine was optimised to co-injected with AS01 adjuvant which gives the best availability and immunogenicity of the vaccine (Kester et al., 2009). Unlike the RTS,S vaccine, *Pf*SPZ vaccine is based on life attenuated, metabolically active, non-dividing sporozoites (Richie et al., 2015), while MVA ME-TRAP was developed from liver stage antigen. Although there are dissimilarities among these vaccine candidates, it is seemingly that most vaccines in the developmental pipeline require booster shot to prolong the protection effect of the vaccine (Crompton et al., 2010), which would be a limitation in control measure.

1.4.3 Case management

Case management in highly endemic area is another important factor in malaria control. The gold standard of diagnostic for malaria is a microscopic technique, however, the technique requires experienced health workers which are limited. Therefore, suspected malaria cases is inflated by misdiagnosis. Rapid diagnostic tests (RDTs) are implemented to address this issue. Rapid and accurate diagnosis of malaria could reduce the disease burden and prolong the drugs efficacies (WHO, 2015b).

1.5 Malaria eradication program

Global attempts to eradicate malaria have been seen in the history. The biggest attempt was Global Malaria Eradication Programme (GMEP) launched by WHO in 1950s followed the successful attempts in regional eradication programmes. At the time, the use of DDT and antimalarial drug chloroquine was the major intervention of the programme. Although it was a promising effort, the program was abandoned in the late 1960s. After the program, increase in malaria prevalence was observed.

After the abandonment of GMEP in 1969, the attempts to reduce malaria, HIV/AIDS, tuberculosis (TB), and other neglected tropical diseases (NTDs) burdens were not totally opted. This is reflected by the dedicated goal set for these diseases in Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) in 2000 and 2016, respectively. In accordant with the goals, WHO has endorsed malaria eradication program worldwide in partnership with Bill and Melinda Gates foundation in 2007 (Tanner and de Savigny, 2008).

1.6 Artemisinins

Artemisinins are sesquiterpene lactones with an unusual endoperoxide bridge (Figure 1.20) considered the most important antimalarial drug class nowadays. It is recommended by World Health Organization (WHO) that artemisinin-based combination therapies (ACTs) should be used in malaria endemic areas as the first line treatment for uncomplicated falciparum malaria (WHO, 2015a).

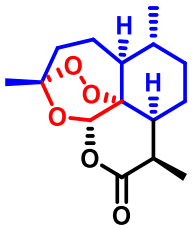


Figure 1.20 Artemisinin. Sesquiterpene (blue), lactone (black), 1,2,4-trioxane endoperoxide (red)

The antimalarial drug discovery project by Chinese scientists in 1967, under Project 523, led to the discovery of artemisinin from sweet wormwood extract, *Artemisia annua* L., by Professor Youyou Tu, who was awarded Nobel Prize in Physiology or Medicine in 2015, at the Institute of Chinese Meteria Medica, China Academy of Chinese Medical Sciences (CACAMS) in 1971 (Tu, 2011). Research team had screened traditional Chinese medicines recorded in literatures and found extract from sweet wormwood leaf exhibits antimalarial activity, but not consistently potent. They found that the traditional extraction method involved heating step might degrade the active ingredient, thus, cool extraction

method could improve potency of artemisinin. Extract from cool extraction method has 100% kill malaria in mouse and had been tested in humans.

Artemisinin and its derivatives (Figure 1.21) share an important core structure, 1,2,4-trioxane endoperoxide (shown in red in Figure 1.20), which is accounted for antimalarial activity. It is believed so because desoxyartemisin, lacking 1,2,4-trioxane, has no antimalarial activity *in vitro* and *in vivo*. The mechanisms by which artemisinins action are still uncertain. However, researchers have proposed mechanisms of activation of endoperoxide and mechanisms of action.

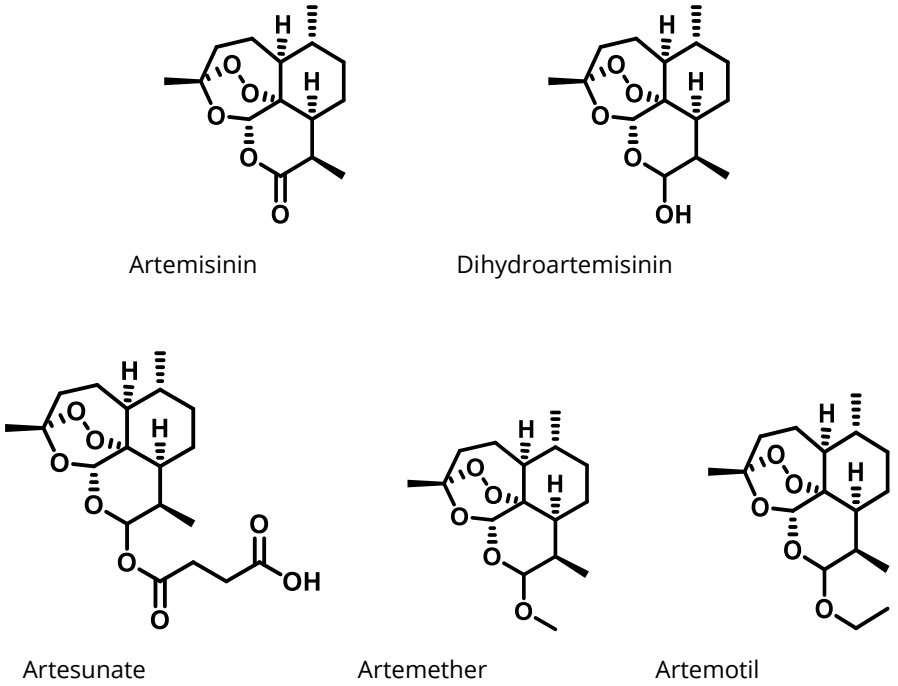


Figure 1.21 Artemisinins

Although artemisinins are very potent antimalarial drugs, their short elimination half-life is a major drawback and is not suitable for use in monotherapy. Therefore, artemisinins are recommended to be used in combination with other longer elimination half-life partner drugs to delay the resistance to artemisinins. The combinations are referred to as artemisinin-based combination therapies (ACTs) and is currently the first line treatment for uncomplicated falciparum malaria in most malaria endemic countries (WHO, 2015a).

Artemisinin resistance has been evidenced from South East Asia region (Phyo et al., 2012) despite an attempt has made to prevent or delay the resistance as a lessons learnt from chloroquine resistance. Interestingly, artemisinin resistance is dissimilar from other drugs resistance, the conventional *in vitro* drug inhibitory assay (IC_{50}) could not discriminate artemisinin susceptible and resistance in clinical

relevant phenotypes (Dondorp et al., 2009). It has been shown that *in vivo* artemisinin resistance is associated with the delayed parasite clearance (Dondorp et al., 2009, Phyo et al., 2012), and the delayed or prolonged developmental cycle during the early stage of parasites is a plausible mechanism of resistance to artemisinins (Teuscher et al., 2010).

1.6.1 Bioactivation of artemisinins

Artemisinins are activated to carbon centred-radicals within the parasite, especially in digestive vacuole where its activators are prominent. It is controversial how artemisinins are activated in the parasite. Three mechanisms of artemisinin activation have been proposed; reductive scission, open peroxide, and cofactor model.

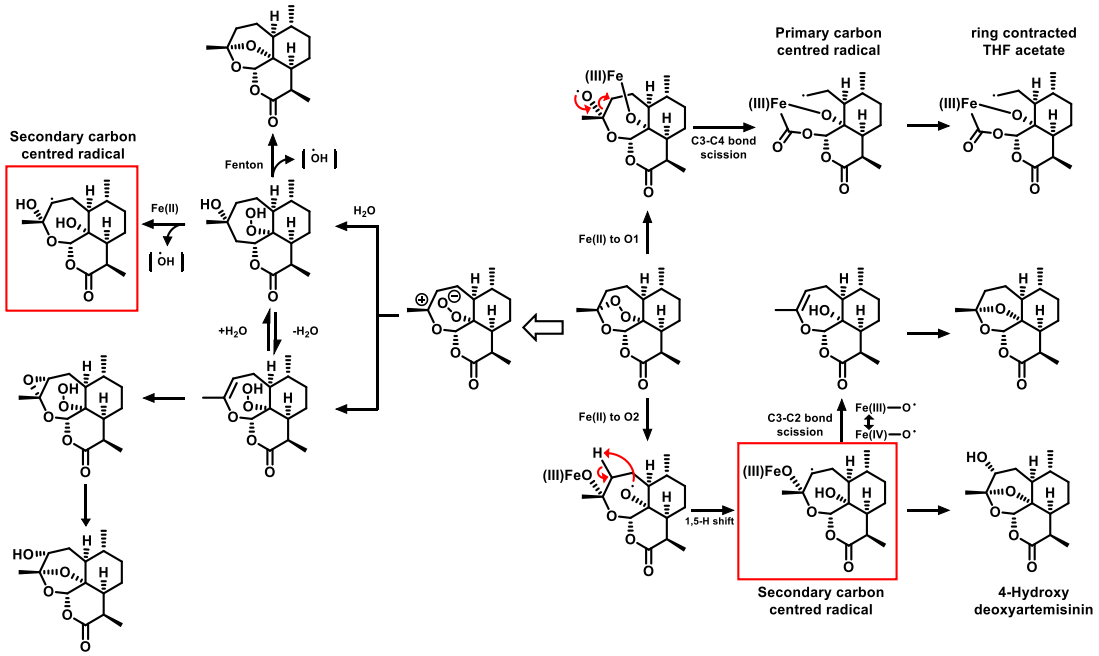


Figure 1.22 Proposed mechanism of artemisinin activation (adapted from O'Neill et al. (2010))

Reductive scission model proposes (Figure 1.22, right) that artemisinin is activated by binding of ferrous iron to either O1 or O2 in 1,2,4-trioxane endoperoxide, and rearrangement to primary or secondary carbon centred radical via β-scission or 1,5-H shift. In support of this mechanism, there is evidence of the existence of these radical intermediates (Butler et al., 1998, O'Neill et al., 2000, Wu et al., 1998). These carbon centred radicals are believed to alkylate parasite proteins or haems.

The open peroxide model proposes that iron acts as a Lewis acid to open the trioxane endoperoxide bridge via ionic activation to generate subsequent reactive oxygen species. These reactive oxygen species are capable of protein oxidation.

The reductive scission and open peroxide models require iron for activation which is believed to derive from haem-iron and free iron. A recent study showed that haem is main activator (>80%) of artemisinin in early ring stage (Xie et al., 2016). This hypothesis is supported by detectable active falcipains, haemoglobin digestive enzymes, during early ring stage (Xie et al., 2016). Free iron also plays a role in activation of artemisinin as iron chelator, deferoxamine (DFO) showed antagonistic effect to artemisinin (Stocks et al., 2007, Xie et al., 2016) but did not abolish the inhibitory effect (Eckstein-Ludwig et al., 2003). These studies lend support to a role for free iron during artemisinin activation.

The cofactor model proposes that the trioxane bridge is activated by redox-active flavoenzymes generating reactive oxygen species (Haynes et al., 2012). Flavoenzymes involved in the model are not limited to mitochondrial flavoenzymes but also cytosolic flavoenzymes. The cofactor model has reduced its importance as haem-iron and free iron have been evidenced to play important role in activation of artemisinin (Xie et al., 2016), which are not required by the cofactor model.

1.6.2 Postulated mechanisms of action and resistance of artemisinin

With an attempt to address the mechanism of action and resistance of artemisinin and derivatives, many studies have been conducted to propose the mechanism or targets, including translationally controlled tumour protein (TCTP), sarco/endoplasmic reticulum Ca^{2+} ATPase (SERCA or *Pf*ATP6), and phosphatidylinositol-3-kinase (PI3K), whereas kelch propeller domain (K13) gene was proposed as a prominent artemisinin resistance marker.

1.6.2.1 Translationally controlled tumour protein (TCTP)

The first evidence of TCTP as the artemisinin target was suggested by Bhisutthibhan et al. (1998). Tritiated-dihydroartemisin ([^{3}H]-DHA) was used to identified DHA protein target *in situ*. The 25-kDA protein was detected by autoradiographic detection of tritium and identified by matching sequence (Bhisutthibhan et al., 1998). Further study revealed molecular interaction of artemisinin with TCTP, the binding sites of artemisinin in TCTP were identified using MS, surface plasmon resonance spectroscopy, and bioinformatics approaches (Eichhorn et al., 2013). Also the interaction between artemisinin, TCTP and haemin in anti-cancer was also established (Bhisutthibhan et al., 1998, Eichhorn et al., 2013). These finding suggested artemisinin binding might interfere with cytokine-like activity of TCTP (Eichhorn et al., 2013). However, the mechanism of artemisinin on TCTP is yet unknown.

1.6.2.2 Sarco/endoplasmic reticulum $Ca2^+$ ATPase (SERCA)

A study conducted in *Xenopus laevis* oocyst proposed that artemisinin is a direct inhibitor of SERCA (*Pf*ATP6) *in vitro*. Eckstein-Lugwig and colleagues showed that *Xenopus* membrane expressing *Pf*ATP6 protein alongside with other transporters are normally functional. Artemisinin, but not quinine or

chloroquine, inhibits the function of *Pf*ATP6, but not other transporters, with high specificity. This effect is comparable to thapsigargin, a known specific inhibitor of *Pf*ATP6. They also showed free iron chelator deferoxamine (DFO) antagonises the effect of artemisinin on *Pf*ATP6 but not thapsigargin as thapsigargin does not require iron for activation (Eckstein-Ludwig et al., 2003).

However, the hypothesis of *Pf*ATP6 as the target of artemisinin was challenged by the work by David-Bosne et al. (2016). They showed that *Pf*ATP6 expressed in *Xenopus* oocyst is not affected by artemisinin and also suggested that *Xenopus* expression system is not suitable for the study and should be used with high caution. In addition, the relation between artemisinin resistance and mutations in *Pf*ATP6 could not be established in clinical studies (Cui et al., 2012, Phompradit et al., 2011, Brasil et al., 2012, Tanabe et al., 2011).

1.6.2.3 Kelch propeller domain (K13)

K13 gene had been identified as a major marker for artemisinin resistance. Association between K13 mutation and resistant trait of *Plasmodium falciparum* was identified from genome-wide association study (GWAS) on more than 1,000 clinical cases from Cambodia and Africa (Ariey et al., 2014).

The major mutations in K13 propeller domain are C580Y, R539T and Y493H. These mutation patterns are strongly associated with slow clearance half-life (>5 h) (Ashley et al., 2014). The association between K13 and resistant trait was also confirmed by *in vitro* ring survival assay (RSA_{0-3}) study conducted with transgenic parasites and clinical isolates (Straimer et al., 2015). Although mutations in K13 gene show correlation to artemisinin resistance, other putative genes (*Pf*CRT, *Pf*CTP, *Pf*MDR1, *Pf*MRP1, and ABC transporter) or putative target of artemisinin (*Pf*ATPase6) were not acquired mutation during 5-year selection of artemisinin resistant parasite, F32-ART5 (Menard et al., 2016).

After the K13-propeller domain was identified as a prominent marker for artemisinin resistance, many other studies used K13 mutations to monitor artemisinin resistance worldwide (Conrad et al., 2014, Isozumi et al., 2015). A recent study monitoring polymorphisms in the K13 gene reveals a distribution of K13 polymorphisms. South East Asia are distributed into 2 regions according to haplotypes of mutation. Cambodia, Vietnam, and Laos are included in one region. Thailand, Myanmar, and China form another region (Figure 1.23). Africa shows rare nonsynonymous mutations, different from Asia. A common A578S mutation in Africa region shows no correlation with delayed parasite clearance as assayed by *in vitro* ring survival assay and clinical outcome (Menard et al., 2016).

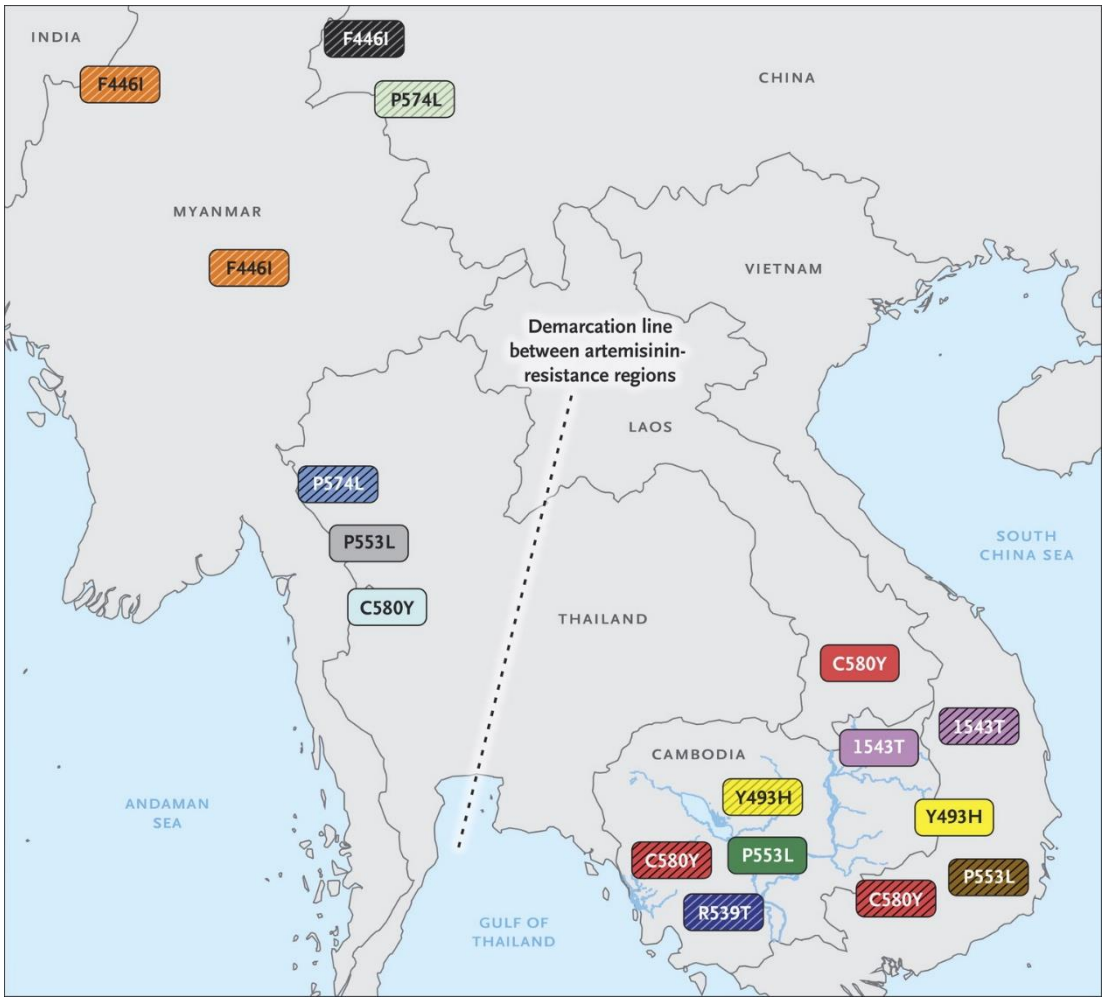


Figure 1.23 Distribution of K13-propeller domain polymorphisms in Asia
(Reproduced with permission from Menard et al. (2016), Copyright Massachusetts Medical Society)

1.6.2.4 Phosphatidylinositol-3-kinase (PI3K)

In addition to other proposed targets of artemisinin, phosphatidylinositol-3-kinase (PI3K) was also suggested as the direct target of artemisinin. An increase in PI3K level is associated with mutation in K13 gene. It was proposed that PI3K in artemisinin susceptible parasite is ubiquitinated, upon binding with K13 protein, and undergone degradation by proteasome system reducing the phosphatidylinositol-3-phosphate (PI3P) level. On the contrary, K13 mutation blocks the interaction between PI3K and K13 proteins preventing PI3K degradation and increasing PI3P level (Mbengue et al., 2015). PI3K plays roles in signalling pathway and trafficking in the parasites (Vaid et al., 2010). Its role in signalling might involve in resistance mediation (Mbengue et al., 2015).

1.6.2.5 Delayed parasite life cycle (dormant ring)

It was proposed that resistance to artemisinin is linked with delayed growth in the early stage of the parasite life cycle (dormant ring). This dormant ring was protected from the effects of artemisinin and resumed normal growth in the absent of drug (Teuscher et al., 2010, Witkowski et al., 2010, Cheng et al.,

2012). Later on, this delayed parasite life cycle was observed in the clinic as a reflection of delayed parasite clearance half-life (Phyo et al., 2012, Ashley et al., 2014). The hypothesis led to the development of *in vitro* assays to determine sensitivity to artemisinin. Basically, the *in vitro* assays mimicked the clinical episode of artemisinin exposure and monitored the development of parasites. (Witkowski et al., 2013, Hott et al., 2015). It is believed that the short plasma half-life of artemisinins and dormant ring parasites led to the recrudescent after the treatment (Codd et al., 2011). Therefore, prolonged artemisinin regimen could reduce the treatment failures (de Vries and Dien, 1996, Codd et al., 2011, Ashley et al., 2014). In addition, the delayed parasite clearance half-life was associated with mutations in the K13 gene (Ariey et al., 2014). However, what causes the parasite to enter dormancy is yet unknown.

Following the recent studies, WHO has defined the new working definition of artemisinin resistance (Box 1). This definition is based on therapeutic studies of ACTs, clinical trials, and K13 sequencing.

Box 1 Working definition of artemisinin resistance (WHO, 2014b)

Suspected partial artemisinin resistance is defined as:

- ≥ 5% of patients carrying K13 resistance-associated mutations; or
- ≥ 10% of patients with persistent parasitaemia by microscopy on day 3 after treatment with ACT or artesunate monotherapy; or
- ≥ 10% of patients with a parasite clearance half-life of ≥ 5 h after treatment with ACT or artesunate monotherapy.

Confirmed partial artemisinin resistance is defined as:

≥ 5% of patients carrying K13 resistance-associated mutations, all of whom have been found, after treatment with ACT or artesunate monotherapy, to have either persistent parasitaemia by microscopy on day 3, or a parasite clearance half-life of ≥ 5 h.

1.7 Approaches toward understanding mode of action

Understanding mode of action or target of the antimalarial drugs is important for management of use and developing the novel antimalarial drugs. In order to study mode of action of the drugs, many approaches have currently been used including *in silico* assignment, stage specificity determination, gene annotation, and *in vitro* evolution of drug resistant parasite (McNamara and Winzeler, 2011). Unfortunately, many of the techniques applicable with other organisms were incapable in *Plasmodium falciparum* e.g. RNAi and microRNA (McNamara and Winzeler, 2011). A technique such as *in vitro* or possibly *in vivo* evolution of drug resistance is very time consuming and rather expensive, however, it was used extensively in the past to successfully identify the target of chloroquine (Fidock et al., 2000) and, more recently, the targets of the new antimalaril candidates (Lim et al., 2016). The technique has it

own limitation for compounds with multiple targets, which is hypothesised as artemisnin mode of action, and that resistance is gained by indirect mechanism.

In this study a chemical biology approach using the modified artemisinin activity-based probes and trioxolane activity-based probe (click chemistry) and mass spectrometry proteomics were undertaken to identify the protein partners of these compounds. Click chemistry is a highly selective chemical reaction that favours only a direction of reaction. Click reaction is achievable via copper-catalysed and copper-free reaction; both reactions were adopted in this study (see detailed review in Chapter 3). In addition to chemical biology and proteomic approaches, whole genome sequencing by next-generation sequencing was employed to identify SNPs or genes associated with resistant phenotypes from *in vivo* artesunate-induced resistant parasite.

1.8 Aims

The aims of this work are to:

1. Provide proof of concept (PoF) for click chemistry-based approach to study molecular targets and mechanism of actions.
2. Identify the artemisinin molecular targets in various stages of *Plasmodium falciparum* strains 3D7 parasite.
3. Identify the targets of fully synthetic endoperoxide trioxolane in *Plasmodium falciparum* strain 3D7 parasite.
4. Identify SNPs or genes associated with artesunate resistant phenotype.

Chapter 2

General Materials and Methods

This chapter outlined general materials and methods used throughout the study in details. Specific experimental protocols can be found in the particular chapter.

2.1 Chemicals

All chemicals used in the study were supplied from Sigma-Aldrich unless otherwise stated.

2.2 Preparation of HEPES

HEPES or (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) was used as a buffer in complete media. One molar HEPES solution pH 7.4 was prepared by mixing 238.3 g HEPES in 800 mL distilled water on magnetic stirrer until completely dissolved. The pH was adjusted by adding 5 N NaOH to achieve pH 7.4. Then volume was adjusted to 1 L by distilled water. Final solution was filtered sterile using 0.2 µm PES filter unit (Thermo Scientific, 566-0020) and stored at 4°C.

2.3 Preparation of hypoxanthine

Hypoxanthine was supplemented in Albumax-I-based media as both RPMI-1640 and Albumax-I lack hypoxanthine, a precursor metabolite for purine biosynthesis. Four millimolar hypoxanthine was prepared by mixing 272.2 mg hypoxanthine in 400 mL distilled water. The pH was adjusted using 5 N NaOH to pH 12 and adjusted total volume using distilled water to 500 mL. The final solution was filtered sterile using 0.2 µm PES filter unit and stored at 4°C.

2.4 Preparation of human serum

Human sera were obtained from Royal Liverpool University Hospital and allowed to coagulate at room temperature (RT) overnight. Serum from each donor was distributed to 50 mL aliquots and centrifuged to sediment any remaining cells at 3500g 5 min at RT. Human serum which appeared red due to haemolysis was discarded. Then human serum suitable for parasite culture from each donor was pooled together; each pool contained more than 10 donors, and distributed to 50 mL aliquots. Prepared human sera were heat inactivated at 56°C for 1 h and stored at -20°C until used. Each batch of human serum was tested for sterility by examining the turbidity and microscopically checked after 24 h incubation at 37°C. The batch passing sterility test was used in complete medium without filtration.

2.5 Preparation of sorbitol

The sorbitol solution was used to synchronise ring stage parasites. Five per cent w/v sorbitol solution was prepared in distilled water by mixing 25 g of sorbitol in 500 mL distilled water and mixed well. The final solution was filtered sterile using 0.2 µm PES filter unit and stored at 4°C. Sorbitol solution was pre-warmed prior to use.

2.6 Preparation of saponin

The saponin solution was used for RBC membrane lysis. The saponin solution was prepared in 0.15% w/v by mixing 0.75 g saponin in 500 mL D-PBS and mixed well. The final solution was filtered sterile using 0.2 µm PES filter unit and stored at 4°C. All process was carefully carried out as saponin is irritation to respiratory system.

2.7 Preparation of cryopreservation solution

The cryopreservation solution was prepared by mixing 140 mL of glycerol with 360 mL of 4.2% sorbitol in 0.9% NaCl solution. The cryopreservation solution was prepared by weighting 15.12 g of sorbitol and 3.24 g of NaCl and mixed with 360 mL of distilled water until completely dissolve. This give final concentration of 28% glycerol, 3% sorbitol, and 0.65% NaCl. The final solution was filtered sterile using 0.2 PES filter unit and stored at 4°C.

2.8 Preparation of TE buffer

TE buffer or Tris-EDTA buffer was prepared by mixing 1 M Tris and 0.5 M EDTA pH 8 together and adjusted the pH to 7.4 by HCl. 0.5 M EDTA buffer was prepared by adding 18.6 g of EDTA to 80 mL distilled water adjusted to pH 8 to dissolve fully, and adjusted to 100 mL by distilled water. 1M Tris was prepared by adding 60.57 g of Trizma® base to 400 mL distilled water, adjusted to pH 7.4, and adjusted volume to 500 mL by distilled water. Then 10 mL of 1 M Tris and 2 mL 0.5 M EDTA pH 8 were added to 900 mL distilled water and adjusted the pH to 7.4, if needed, and adjusted volume to 1 L by distilled water. The fiinal solution was filtered sterile using 0.2 µm PES filter unit and stored at 4°C.

2.9 Preparation of lysis buffer

The lysis buffer was prepared at 10X concentration for long term storage and diluted to 1X solution with sterile distilled water when required. To prepare 500 mL of 10X lysis buffer, 100 mL of 1 M Tris pH 7.4 and 500 µL of 0.5 M EDTA pH 8 (both reagents were prepared with the same method as for TE buffer, section 2.8. Then 4 mL of Triton-X 100 and 400 mg of saponin were added to the mixture and adjusted volume to 500 mL with distilled water. The 10X lysis buffer was filtered sterile with 0.2 µm PES filter unit and stored at 4°C until use.

2.10 Preparation of complete media

There are two types of complete media (CM) used in the study: human serum-based and Albumax-I-based complete media. Albumax-I-based complete media were prepared from base RPMI-1640 (Sigma, R8758) supplemented with 12.5 mL 1M HEPES, 5 mL 4 mM hypoxanthine pH 12, 25 mL 5% Albumax-I (Life Technologies, 11020-039), and 200 µL 10 mg/mL gentamycin solution. As all components were sterile-filtered, no sterilisation process was needed. Human serum-based media were prepared from based RPMI-1640 supplemented with 12.5 mL 1 M HEPES, 50 ml human serum, and 200 µL 10 mg/mL gentamycin, and filtered sterile with 0.2 µm PES filter unit, if required. Complete media were stored at RT and used within 2 weeks after preparation.

2.11 Preparation of washed red blood cells

Whole blood O^+ supplied in citrate-phosphate-dextrose bag was obtained from national blood services and aseptically distributed to 25 mL aliquots. Prior to use, the whole blood was washed 3 times with RPMI-1640 to remove all buffy coat. Washed blood was stored at 4°C and used within 1 week.

2.12 Retrieval of parasite from cryopreservation

The parasites used in the study were preserved and stored in vapour phase of liquid nitrogen at -190°C for long-term storage. Parasites were retrieved from cryo storage using single step 3.5% $NaCl_2$ method. Cryogenic vials containing cryogenic preserved parasites were thawed in 37°C water bath. The contents were then aseptically transferred to 15 mL sterile centrifuge tubes. An equal volume of 3.5% $NaCl_2$ was gently added drop-by-drop whilst gently mixed and allowed to equilibrate at 37°C for 5 min. Parasite pellet was obtained by centrifuged at 500g 5 min at RT. Then aliquot of 1 mL aseptically processed human serum was added and equilibrated at 37°C for 30 min, then washed with 10 mL RPMI-1640. Culture was initialised by adding 100 µL of fresh washed RBCs, 10 mL complete medium, and gassed as per normal culture.

2.13 Parasite culture

In vitro parasite culture was performed followed the method developed by Trager and Jensen (1976) with some modifications. The culture was maintained in T50 flask (Nunc, Thermo Scientific) at 2% haematocrit in 50 mL complete medium. The culture was maintained in physiological conditions at 37°C in physiological atmosphere by flushing with 4% oxygen, 3% carbon dioxide, 93% nitrogen gas mixture via sterile serological pipet. The day-to-day parasite culture was maintained at 1-2% parasitaemia, daily examined for parasitaemia and stages by thin smear stained with 10% Giemsa solution (BDH, improved R66 solution) at 1000X magnification power. Once parasitaemia was higher that 2%, excessive culture

was taken out and supplied with equal volume of washed RBCs to reduce parasitaemia, unless higher parasitaemia was desired.

2.14 Synchronisation of ring stage parasites by sorbitol

The 5% w/v sorbitol solution was used to eliminate trophozoite stage parasites in the culture. Parasite pellet was obtained by centrifugation at 800g 5 min at RT. Then 5 volumes of pre-warmed 5% sorbitol solution at 37°C was added to parasite pellet and mixed well. Parasite pellet was allowed to incubate with sorbitol solution for 10 min at 37°C. After incubation, parasite pellet was obtained by centrifugation and sorbitol removed by aspiration. Parasite pellet was washed to completely remove sorbitol solution by 35 mL RPMI-1640. Then synchronised parasite was maintained in normal culture condition. Parasite remains synchronous for up to 5 erythrocytic cycles if carefully maintained. However, synchronous period is also dependent on parasite strain.

2.15 Establishment of cryopreservation stock

Cryopreservation stock of parasite culture was always establised within the first 2 cycles of parasite culture after culture initiation to ensure parasite integrity. Cryopreservation stock was prepared by adding 2 volumes of cryopreservation solution to the washed parasite pellet, gently drop-by-drop whilst gently mixed, and transferred to a sterile cryogenic vial. Cryogenic vial was directly transferred to Statebourne cryogenic tank.

2.16 Parasite strains and isolates

The parasite used in this study was mainly *Plasmodium falciparum* strain 3D7. Proteomic works of ring and trophzoite stage were performed with 3D7 parasite.

Parasites used in whole genome sequencing were *Plasmodium falciparum* parasites infected in humanised SCID mouse model. Parasite TE616 parent isolate was used to infect mouse and treated with clinical relevant 3-day corse artesunate monotherapy (10-50 mg/kg). Then blood was collected and injected to the new mouse to generate the passage 2 parasite TE1201. These process was repeated for 22 passages to induce resistance to artesunate. Parasite clearance half-life was determined for every passage to monitor resistant phenotype to artesunate. The experiment to generate this parasite line was conducted at GSK open lab at Tres Cantos, Spain. Then cryopreserved parasites were shipped to LSTM (phenotype data not shown). Parasites were retrieved, adapted to culture, and extracted for gDNA as described in section 2.12, 2.13, and 2.19.

Table 2.1 Plasmodium falciparum parasite isolates from GSK

Isolate	Passage	Artesunate phenotype	Note
TE616	Parent	Sensitive	
TE1201	2	Sensitive	
TE1211	3	Sensitive	
TE1214	4	Sensitive	Failed to retrieve
TE1281	5	Sensitive	Failed to retrieve
TE1304	6	Sensitive	
TE1328	8	Sensitive	
TE1368	10	Sensitive	
TE1373	11	Sensitive	
TE1389	11	Sensitive	
TE1411	13	Sensitive	
TE1419	14	Resistant	
TE1423	15	Resistant	
TE1432	16	Resistant	
TE1435	17	Resistant	
TE1436	18	Resistant	
TE1439	19	Resistant	
TE1475	22	Resistant	

2.17 Conventional IC_{50} assay

The half maximal inhibitory concentration (IC_{50}) of antimalarial compounds was determined using conventional fluorometric-based IC_{50} assay. The stock compounds were generally prepared in DMSO at 10 mM concentration. Assay plates were created using Hamilton® Microlab Star robotic platform. The highest compound concentration was 1 µM and 1:4 serially diluted to achieve 8 concentration levels. Blank CM was used in positive growth control (assay negative control) and 1 µM artemisinin was used in negative growth control (assay positive control). The assay plate has layout as in Figure 2.1.

The parasite inoculum was prepared to 1% haematocrit and 2% parasitaemia. An aliquot of 50 µL of parasite inoculum was added to each pre-loaded well and incubated in modular chamber and gassed with culture gas for 2 min. After 48 h, assay was terminated by freezing the assay plate at -20°C overnight or until further processed.

The parasite growth was measured in reflection of nucleic acid contents detected by nucleic acid fluorescent dye SYBR Green I in 1X lysis buffer. An aliquot of 100 µL 1X SYBR Green I in 1X lysis buffer was added to each well and incubated in dark condition for 1 h before measured by Thermo Scientific

VarioSkan™ microplate reader. SYBR Green I was excited at 497 nm and fluorescent emission detected at 520 nm. The IC_{50} value was determined by regression model incorporated in 'IC50' package in R programming or GraphPad Prism® software.

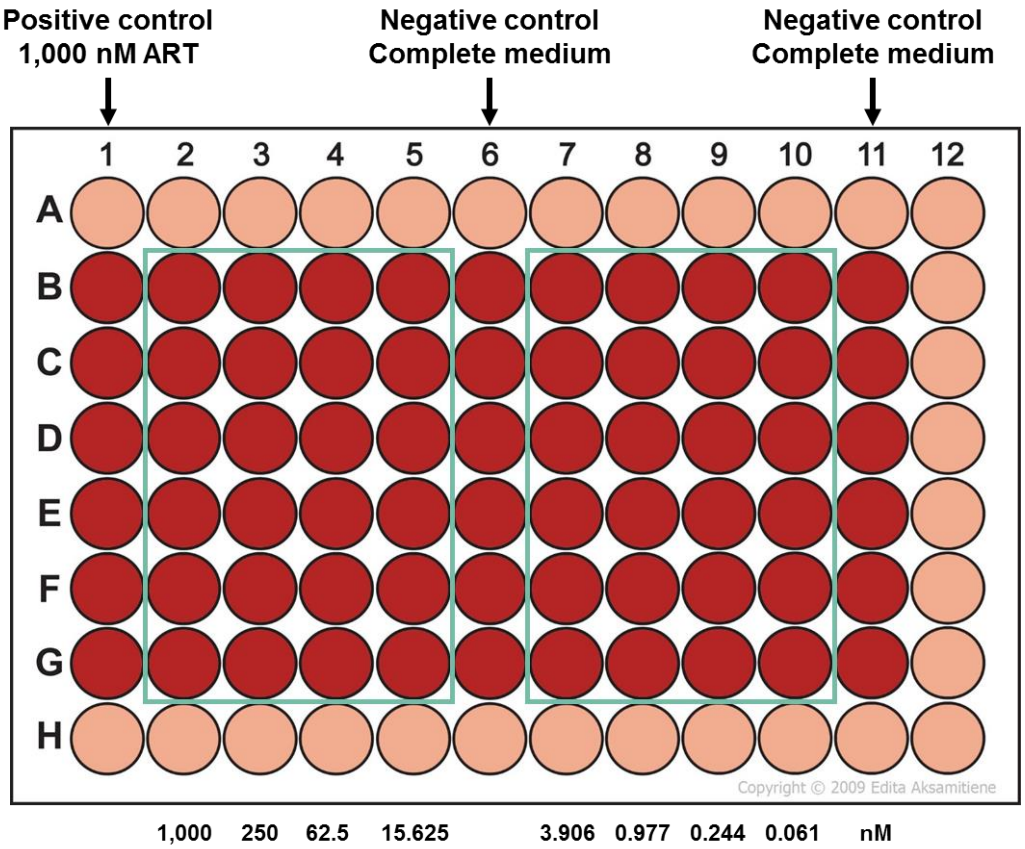


Figure 2.1 IC_{50} assay plate layout. Each assy was performed in triplicate wells. The outermost wells are avoided and filled with complete medium to minimise the edge effect.

2.18 RBC lysis of infected RBC by saponin

The 0.15% w/v saponin solution was used for RBC membrane lysis. Parasite culture was first pelleted by centrifugation at 800g 5 min at RT. Then 5 volumes of 0.15% saponin solution was added to parasite culture pellet and left on for 5 min on ice to prevent degradation of proteins or DNA by exogenous proteases or DNases. After incubation, free parasite pellet was obtained by centrifugation at 3500g 5 min at 4°C. Free parasite pellet was then washed with 10 mL to 35 mL D-PBS, 3 times to remove lysed cells. Free parasite pellet was finally resuspended in either TE buffer or D-PBS supplemented with 1X Roche cOmplete™ EDTA-free protease inhibitor according to subsequent applications.

2.19 Extraction of genomic DNA

All genomic DNA (gDNA) in this study was extracted from the trophozotite stage parasite cultures by phenol-chloroform method. The saponin solution was used to purify parasites from RBCs and

resuspended in TE buffer. This parasite pellet can be stored at -80°C until further process. Free parasite pellet was allowed to defrost at 4°C until completely thawed and complete lysis was assisted by adding 20 µL of protenase K solution (Qiagen) and heated to 65°C for 5 h, then treated with 20 µg/mL of RNase A for 10 min at RT to eliminate RNA contamination.

The extract was added to 1 mL phenol:chloroform:isomoyl alcohol solution (25:24:1) in 1.5 mL microcentrifuge tube, vortexed, and centrifuged at 17,000g 2 min at RT. Then aqueous phase (upper layer) was carefully transferred to a new clean microcentrifuge tube containing 1 mL phenol:chloroform:isomoyl alcohol solution. This step was repeated 3 times to minimise protein contamination in gDNA samples. The final aqueous phase was collected, and 750 µL cold absolute methanol was added to precipitate gDNA overnight at -20°C. Genomic DNA was collected from methanol precipitation step by centrifugation at 17,000g 20 min at 4°C. Methanol was completely removed by aspiration and pellet was air-dried at 50°C until completely dry as methanol residue can interfere some enzymes in the subsequent steps. Genomic DNA pellet was recovered by adding adequate (~50 µL) volume of nuclease-free water pre-warmed at 42°C. This gDNA solution can be stored at 4°C or -20°C or -80°C depending on storage time. Genomic DNA was checked for purity and concentration by spectrophotometry (NanoDrop™, Thermo Scientific) and agarose gel electrophorysis.

2.20 Spectrophotometry analysis of genomic DNA

Quality and quantity of genomic DNA was assessed by determination of protein and nucleic acid contents in the sample by NanoDrop™ spectrophotometer (Thermo Scientific). An aliquot of 2 µL gDNA sample was loaded onto NanoDrop™ head and read for absorbance at 260 nm for nucleotides measurement and 280 nm for proteins, phenols, and other contaminants measurement. The 260/280 ratio was used to determined purity of gDNA sample, ratio of more than 1.8 is acceptable for most application including PCR and sequencing purposes. NanoDrop™ quantification of gDNA is not accepted by Centre for Genomic Research (CGR), University of Liverpool (UoL). CGR required quantification of gDNA by fluorometric-based which was generously performed by CGR staff using Qubit® fluorometer (Thermo Scientific).

2.21 Agarose gel electrophoresis

Agarose gel electrophoresis was used to detect PCR products or check for genomic DNA quality. To prepare 1% agarose gel in tris-borate-EDTA (TBE) buffer, 1 g of agarose was added to 100 mL TBE buffer and heated in medium heat microwave for 2 min and mixed gently until completely dissolved. Then 10 µL SYBR Safe nucleic stain was added to gel solution and gently mixed. From this point onwards, exposure to light was avoided. The agarose gel solution was allowed to cool down before casted in the

gel tray and comb inserted. Agarose gel solution was allowed to polymerise at RT for 30 min. The comb was removed from the casted gel befor sample loading.

Thermo Scientific GeneRuler™ 1 kb plus DNA ladder was used for size comparison, an aliquot of 3-5 µL ladder was used depending on the well volume. DNA sample was mixed with 6X sample loading buffer (Thermo Scientific) to achieve 1X dye solution before loading to the agarose gel. Gel electrophoresis was performed at 80 V constant for 1.15 h for 7 cm gel. The gel was imaged by UV transilluminator gel documentation system (Syngene) via GeneSys software.

2.22 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyse protein samples. Thermo Scientific Novex™ NuPAGE™ Bis-Tris 4-12% pre-cast gel was used in the present study. Running buffer used was NuPAGE™ MES SDS at 1X prepared from 20X commercial stock solution by distilled water.

Protein samples were incubated with corresponding volume of 4X NuPAGE™ LDS sample buffer and 10X NuPAGE™ reducing agent and heated at 90°C for 10 min. Then up to 15 µL of samples was loaded to the gel. Electrophoresis was performed at 120 V constant for 2.15 h in the presence of 1X NuPAGE™ antioxidant and imaged with Bio-Rad Gel Doc™ EZ gel documentation system or GE ImageQuant™ LAS 4000 imaging system. Blue LED excitation and the Y515 BP Cy2 filter setting was used to detect signal from Alexa Fluor® 488.

2.23 *In vitro* parasite treatment with chemical probes

In each experiment, ten 50-mL flasks were used for each probe and also for DMSO control. Parasite of more than 10% parasitaemia was used. *Plasmodium falciparum* parasite culture is described in section 2.13. Parasites from all flasks were pooled together and re-distributed to aliquots of 2% haematocrit in 50 mL complete medium to obtain an equal parasitaemia in each experimental condition.

Parasite cultures were treated with either 1 µM active probe or inactive probes or an equal volume of DMSO for 6 h at normol culture conditions. After 6 h, all flasks from respective condition were pooled together and parasite pellet was sedimented by centrifugation at 1500g 5 min at RT. Then RBC lysis was performed by adding 35 mL of 0.15% saponin solution and incubated for 5 min at 4°C to allow maximum lysis of RBCs. The free parasites were collected by centrifugation at 3,500g 5 min at 4°C followed by 3 washing steps with 50 ml of Ca^{2+} and Mg^{2+}-free phosphate buffer saline solution. Then the free parasites were suspended in 1 mL of D-PBS containing Roche cOmplete™ protease inhibitor and stored at -80°C until process.

2.24 Protein samples preparation and quantification

Protein samples from clickable-probes treatment were process as follows. Free parasite in proteinase inhibitor D-PBS solution was briefly sonicated to assist cell lysis using probe sonicator (30% power). Cell debris was separated by centrifugation at 1,600g 20 min at 4°C and supernatant was collected.

The Bradford protein quantification assay was performed to quantify protein concentration in each sample. Bovine serum albumin (BSA) standard ranging from 2-0.125 mg/mL was prepared in triplicate in 96-well microwell plate. Each well contained 5 µL protein sample (or BSA standard solution), 45 µL D-PBS, and 250 µL Bradford reagent (Bio-Rad). Then the protein concentration was adjusted to appropriated concentration (1.5-2 mg/mL) using D-PBS. Concentration adjusted protein samples were ready for click chemistry reaction.

2.25 Concentration of protein samples

In order to concentrate protein in the samples, Christ rotational vacuum concentrator was used. Protein samples in 2 mL screw-cap tubes were placed in vacuum concentrator with lid off. A vacuum state was generated by external vacuum pump equipped with cooler condenser operating at -50°C. Duration for concentration was depended on initial sample volume and desired final sample volume. To completely remove solvent in the samples, in typical 250 µL solution, samples were placed in the system for 6 h.

2.26 Click reaction

In this study copper-catalysed and copper-free click reaction were used. For copper-catalysed reaction, biotin-azide (PEG4 carboxamide-6-Azidohexanyl Biotin) was supplied from Life Technologies. Protein samples were prepared in 500-µL aliquot and processed as follows. Aa aliquot of 5.65 µL of 5 mM biotin-azide in DMSO was added to each protein sample and vortexed, followed by 11.3 µL of 50 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), no vortex applied, 34 µL of 1.7 mM Tris[(1- benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA), vortexed, and 11.3 µL of 50 mM $CuSO_4$, vortexed. Then reaction was allowed to incubate at RT for 1 h under dark condition and gently mixed every 15 min. For copper-free click reaction, Click-IT® biotin dibenzocyclooctyne (DIBO) alkyne (Life Technologies) was used instead of biotin-azide. The aliquot of 5.65 µL of 5 mM Click-IT® biotin DIBO alkyne was added to protein samples and vortexed. Copper-free click reaction was allowed to occur at RT for 1 h in the absent of light and gently mixed every 15 min.

After 1 h, respective samples were pooled together and centrifuged at 6,500g 4 min at 4°C to pellet protein adducts. Then 750 µL of ice-cold absolute HPLC-grade methanol was added to protein pellet and briefly sonicated for 3-4 s, repeated methanol wash step twice. Then protein pellet was

resuspended in 650 µL of 2.5% SDS in D-PBS and briefly sonicated 3-4 s 3 times to dissolve protein pellet. Samples were then heated in 90°C heat block for 5 min to denature proteins in the samples followed by pulse sonication 2, twice. Cell debris and insoluble proteins were pelleted by centrifugation at 6,500g 4 min at RT. The supernatant was collected and adjusted volume to 3.5 mL by D-PBS. Samples were stored at -20°C until further process.

To fluorescently label protein in the samples by click reaction, protein samples from azide probes treatments were used. An aliquot of 50 µL protein samples was mixed with 5 µL of 200 µM Click-IT® Alexa Fluor® 488 DIBO alkyne (Invitrogen, USA) in D-PBS and incubated for 1 h at RT with gentle mixing every 15 min interval. Then reaction was terminated by adding 500 µL absolute cold methanol to precipitate protein. The samples were centrifuged to obtain protein pellet and washed 3 times with cold methanol. Protein pellet was then resuspended in 1X NuPAGE™ LDS sample buffer (Invitrogen, USA) supplemented with 1X NuPAGE™ sample reducing agent (Invitrogen, USA) and heated at 90°C in heat block for 10 min. Samples were ready for gel electrophoresis analysis (section 2.22).

2.27 Streptavidin enrichment, protein alkylation and reduction

Protein samples from click reaction step were thawed and adjusted the volume to 8.3 mL by D-PBS. Then 200 µL of pre-washed streptavidin agarose beads (Pierce, Thermo Scientific) in D-PBS, containing 100-µL of 50% slurry streptavidin, was added to the solution and mixed on end-over rotator for 1.5 h. After incubation with beads, samples were centrifuged at 1,400g 2 min at RT, and supernatant was discarded. The remaining beads were transferred to microbio-spin column (Bio-Rad) and washed with 1 mL of 1% SDS in D-PBS, 6 M urea in D-PBS, and D-PBS, each for 3 times. Washed beads were transferred to 2-mL screw-cap tube using minimal D-PBS and centrifuged at 1,400g 2 min at RT. Beads were then resuspended in 500 µL of 6 M urea in D-PBS. Then 25 µL of 200 mM DL-dithiothreitol (DTT) was added to the samples and heated in 65°C heat block for 15 min. An aliquot of 25 µL of 500 mM 3-indoleacetic acid (IAA) was added and mixed on medium speed end-over rotator for 30 min at RT under dark condition. After incubation, samples were centrifuged at 1,400g 2 min at RT and discarded the supernatant. The remaining beads were washed with 1 mL D-PBS.

To perform protein digestion, washed beads were resuspended in 200 µL of 2 M urea. Then 2 µL of 100 mM $CaCl_2$ and 5 µL of 0.5 mg/mL sequencing-grade trypsin (Promega) were added to the samples and allowed to incubate in orbital shaker (250 rpm) overnight at 37°C. Elution of tryptic peptides was performed using microbio-spin column (Bio-Rad). Then 17 µL of 90% formic acid was added to tryptic peptide. At this step samples were ready for peptide sequencing by HPLC-MS/MS and can be stored at -80°C.

2.28 HPLC-MS/MS analysis

Peptide sequencing was performed on ultra-high-performance liquid chromatography coupled tandem mass spectrometry system (UHPLC-MS/MS). UHPLC used in the study was Thermo Scientific UltiMate™ 3000LC chromatography system. Mass spectrometer used in the study was Themo Scietific LTQ Orbitrap™ Velos using Xcalibur™ software v2.1 (Thermo Scientific). Peptides sample were injected onto analytical column (Dionex Acclaim® PepMap™ RSLC C18, 2 µm, 100 Å, 75 µm i.d. x 15 cm, nanoViper™.), which was maintained at 35°C and at a nanoflow rate of 0.3 µL/min. Peptides were separated over linear chromatographic gradients composed of buffer A (2.5% ACN: 0.1% formic acid) and buffer B (90% ACN: 0.1% formic acid). Two gradients, 60 (3-50% buffer B in 40 min) and 120 minutes (3-60% buffer B in 90 min), were employed for analysis. Full scan MS spectra were acquired over the *m/z* range of 350-2000 in positive polarity mode by the LTQ Orbitrap™ Velos at a resolution of 30,000. A data dependent Top20 collision induced dissociation (CID) data acquisition method was used. The ion-trap operated with CID MSMS on the 20 most intense ions (above the minimum MS signal threshold of 500 counts).

2.29 Protein Identification

Protein identification was performed by MASCOT and SEQUEST search engine via Thermo Scientific Proteome Discoverer™ v1.4. The database used in the study was SwissProt database for *Plasmodium falciparum*. Raw spectrum files from mass spectrometer were imported to the software and processed with following MASCOT parameters: precursor mass tolerance of 10 ppm, fragment ion tolerance 0.8 Da with one tryptic missed cleavage permitted. Carbamidomethyl (C) was set as a static modification with oxidation of methionine (M) and deamidation (N,Q) set as dynamic modifications. A decoy database was searched and relaxed peptide confidence filters applied to the dataset (ion scores $p < 0.05$ / FDR 5%).

The protein list was then filtered by high confidence peptide and 2 peptide-per-protein filters to increase confidence of protein identification. For each paired experiment, proteins identified by DMSO and desoxyartemisinin probe were excluded from the artemisinin activity-based probe proteome.

2.30 Computational protein-ligand docking

Computational protein-ligand docking was performed using Autodock Vina software either on standalone or built-in in Chimera software suite. Three-dimensional protein structures were obtained from PDB database. Any bound inhibitor was removed from the structure and polar hydrogens were added by MGL tools or Dock Prep software. Area of interest was defined manually with either in AutoDock tools or Chimera. AutoDock Vina calculates the results according to the user inputs. Docking results were visualised by Chimera.

Chapter 3

The use of click-chemistry to define the proteomic targets of the antimalarial artemisinin: validation of a pairwise strategy

3.1 Introduction

Proteomics is a branch of the post-genomics era –omics technology. The suffix –*ome* comes from latin root *oma* meaning mass. Proteomics is thus a study of large amounts of proteins. The term 'proteome' was first coined in 1994 (Patterson & Aebersold 2003), and there are 2 major proteomic approaches; untargeted proteomics and targeted proteomics. Untargeted proteomics aims to study all proteins in a given set of conditions, whereas targeted proteomics focuses on a given set of pre-determined proteins in a particular condition. In this study, an untargeted proteomics approach using click chemistry was employed to study the potential molecular target(s) of artemisinin action against *Plasmodium* for the first time.

During the early years of proteomic research, two-dimensional gel electrophoresis (2DE), developed in 1975, was the most commonly used separation technique. This technique relies on the principle that each protein will migrate on a polyacrylamide gel matrix at different rates based on their isoelectric points (which is charge dependent) in the first dimension and their relative molecular weights in the second dimension (O'Farrell 1975). In the late 1960s a mass spectrometry method was developed for protein and, more specifically, peptide sequencing. However, peptide sequencing only really came into its own as an analytical technique when the tandem mass spectrometry technique was invented (Biemann 1986; Hunt et al. 1986). Since then, mass spectrometry-based proteomics has become mainstream and together with protein sequence databases, has significantly advanced research in this field.

Click chemistry is a relatively new approach in synthetic chemistry developed by Professor Barry Sharpless's group at The Scripps Research Institute in the USA. He was awarded the Nobel Prize in chemistry in 2001 for the development of chirally catalysed oxidation reactions (also known as Sharpless epoxidation and Sharpless asymmetric dihydroxylation). Briefly, it is a type of reaction that can synthesise complex molecules from small units in a stereospecific manner with high yields. The reaction itself should also display a large thermodynamic driving force to favour the formation of a single reaction product. Thus the click reaction provides a good platform for complex molecule synthesis.

Subsequently this approach has been modified to allow chemical tagging of target biomolecules coupled with targeted pull down purification and identification, the “click chemistry” strategy.

3.1.1 Mass spectrometry-based proteomics

The birth of the mass spectrometer (MS) was closely related to the study of the atom by Sir Joseph John Thomson, 1906 Nobel laureate in physics, which led to development of the mass spectrograph by Francis William Aston. Mass spectrometry (MS) was used extensively by physicists and chemists in the early years and, later on, caught biologists’ attention for biomolecular studies in the late 1950s (Figure 3.1) (Yates III 2011).

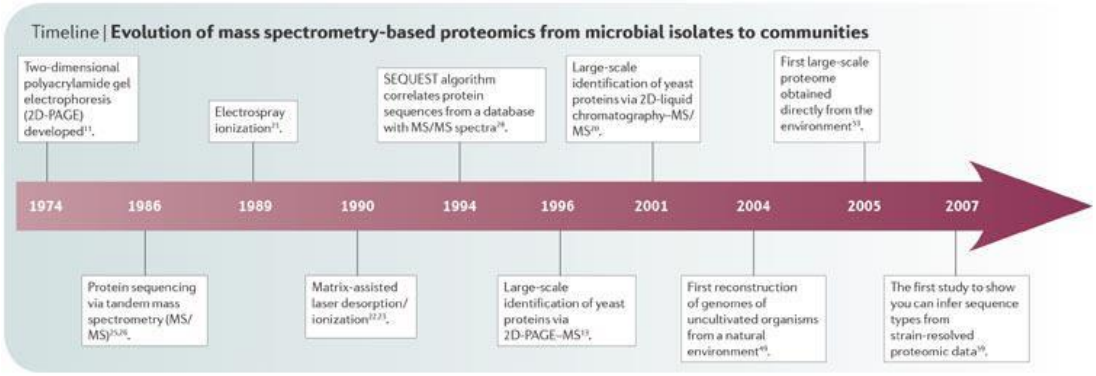


Figure 3.1 Evolution of mass spectrometry-based proteomics from microbial isolates to communities. (Taken from VerBerkmoes et al. (2009). License permission no. 3973081315319).

In general, a mass spectrometer detects the mass of charged molecules, giving the technique its name. There are three main elements in the MS technique - ionisation, mass analysis and ion detection. Samples for MS analysis are required to be ionised in the initial step of MS analysis. Those ionised molecules then travel through the path of the mass analyser where they are separated in some way before finally reaching the detector. MS results are reported in the form of mass-to-charge ratios (*m/z*) (Todd, 1991).

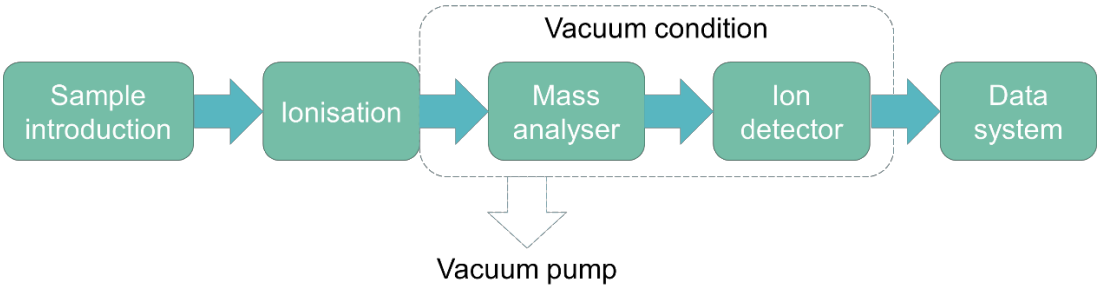


Figure 3.2 Mass spectrometer schematic

Ionisation is the first critical step in order to make a sample detectable by MS. Ionisation can be performed by several methods. In biological applications, soft ionisation methods are commonly preferred i.e. electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI). ESI is one of the most popular ionisation methods, due to the fact that it can be easily equipped on-line with a separation platform, usually liquid chromatography, and it is able to generate single and multiple charged masses, which broaden masses analysed. John B. Fenn and Toichi Tanaka, inventors of ESI, were awarded the Nobel Prize in Chemistry in 2002. On the other hand, MALDI sample preparation is more complicated. It requires co-crystallisation with an organic matrix. This matrix, when excited by a laser at an appropriate wavelength, will then ionise the sample within the matrix into a gas phase. Masses charged by MALDI are usually singlets (Finehout and Lee, 2004). More recently, MALDI has been applied in a MALDI imaging technique (Ly et al., 2016, Aichler and Walch, 2015, Weaver and Hummon, 2013). This technique is advantageous in pathophysiological research, as a desired area within a tissue can be analysed whilst being imaged.

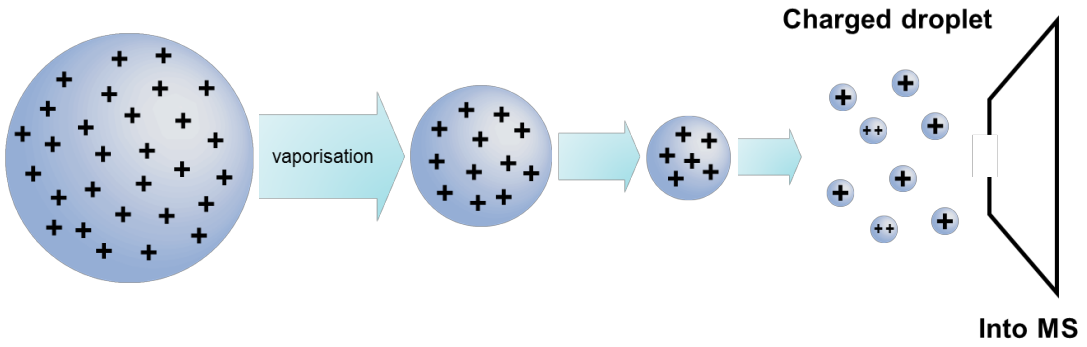


Figure 3.3 Principle of electrospray ionisation (ESI)

The mass analyser is another important step in MS analysis. This component also determines the resolution of the MS. Resolution in mass spectrometry is directly related to number of masses which can be analysed by MS. Resolution of MS is usually determined by full width of the peak at half of its maximum height or FWHM (IUPAC, 1997). Popular mass analysing techniques are quardrupole/iontrap, time-of-flight (TOF), orbitrap, and Fourier transform-ion cyclotron resonance (FTIC), in the order of increasing mass resolution.

In modern MS techniques, masses are usually analysed more than once, termed tandem mass spectrometry (MS/MS). Charged masses are first analysed and allowed to collide with inert gas such as argon or helium in a collision cell to generate fragmented ions or daughter ions. This process is called collision-induced dissociation (CID). Each molecule has specific patterns of fragmentation and this pattern can be used to specifically identify the molecule according to its fragmentation pattern.

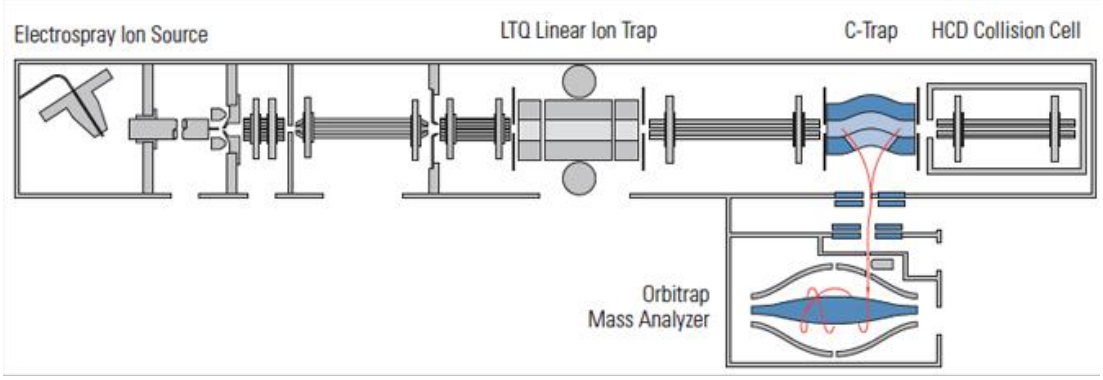


Figure 3.4 Schematic diagram of Thermo Scientific LTQ Orbitrap™ Velos (taken from www.planetorbitrap.com)

The mass detector is the final destination of masses being analysed by MS. Charged masses that pass through the ion path of the mass analyser hit the detector, usually an electron multiplier or microchannel plate, generating an electron cascade magnifying signal in the form of a detectable current. This detection can be calibrated to read out the specific mass/charge.

In order to increase the number of detectable analytes, MS or MS/MS is usually coupled with liquid chromatography (LC) to separate analytes in the sample according to their chemical properties before analysis by MS. LC deployed in biological applications are mainly high-performance or ultra-high performance (HPLC or UHPLC) depending on the pressure level used in the separation. Pressure in LC system is mainly affected by column size and flow rate.

In the present study, ultra-high performance liquid chromatography coupled tandem mass spectrometry was employed (UHPLC-MS/MS) for peptide sequencing. This technique provides high sensitivity of mass analysis and an easy to operate platform.

Peptide sequencing, not protein sequencing, by MS is a fast growing field in light of computer-assisted technology and bioinformatics. Before being analysed, protein samples need to be digested to create peptide fragments by use of certain proteases, usually trypsin. These peptide fragments are then sequenced by MS. In the past, peptide sequencing had been achieved manually by comparing peaks in a mass spectrum, known as *de novo* peptide sequencing (for detail see review in Steen and Mann (2004)). Nowadays, *de novo* peptide sequencing and peptide sequencing matches are performed computationally by software or via database searching. Peptide database searching is a leap forward in proteomic research. A peptide database is created from a genome sequence database by predicting all possible peptide sequences in an organism. Nonetheless, a limitation of the database searching approach as that not every organism has its genome sequenced. Common database peptide searching algorithms are PeptideSearch, Sonar MS/MS, MASCOT, and SEQUEST (Steen and Mann, 2004).

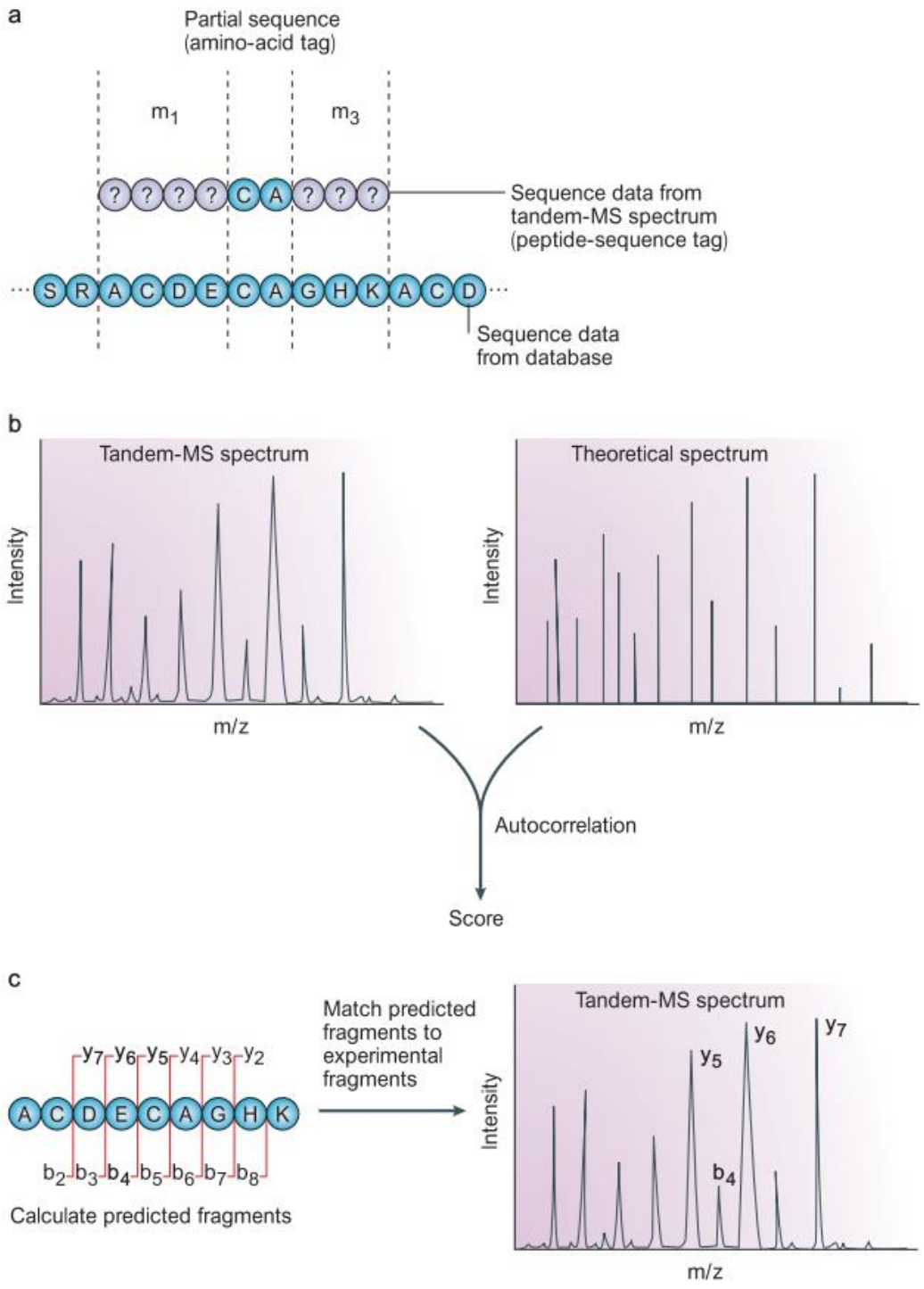


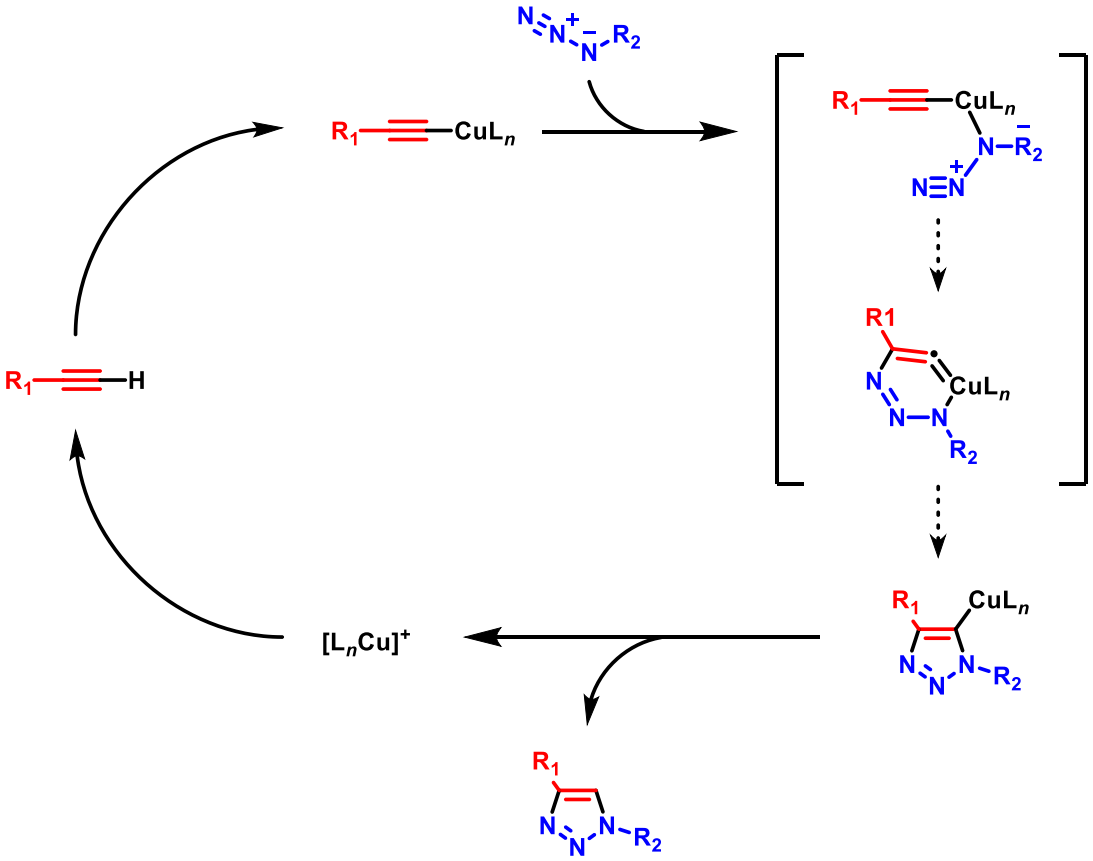
Figure 3.5 Peptide sequencing by MS using database matching. (Taken from Steen and Mann (2004). License permission no. 3973090357128)

3.1.2 Click chemistry

Click chemistry was originally developed as a synthetic approach to generate complex molecules from small simple molecules but has recently been applied to biological and pharmaceutical research as it provides high specificity and selectivity (Kolb et al., 2001, Moses and Moorhouse, 2007). In biological applications, click reactions are also referred to as bio-orthogonal reactions. There are a number of limitations to which reactions can be bio-orthogonal. They have to be insensitive to water, nucleophilic attack, redox, and biological enzymes and able to occur under physiological conditions (pH and temperature) (Jewett and Bertozzi, 2010). These criteria leave only a few compatible reactions and functional groups. Among these are azides and alkynes, thanks to their inert nature and unique chemistry.

To allow bio-orthogonal reactions between alkynes and azides to occur in a biological system, copper-catalysed and copper-free reactions are mainly used (Ramil and Lin, 2013). Both reactions result in production of 1,2,3-triazole via [3+2] cycloaddition (also called Huisgen's 1,3-dipolar cycloaddition). The difference between these reactions is the types of alkyne used which governs the need of a copper catalyst. In a copper-catalysed reaction, terminal alkyne reacts with an azide in the presence of copper and a stabilising ligand (to increase the rate constant of the reaction). When cyclooctyne is used in the reaction, there is no need to add a catalyst as its 8-membered ring strain is the driving force for the cycloaddition reaction (Jewett and Bertozzi, 2010, Subramanian et al., 2014).

Sharpless and colleagues (2002) first proposed the mechanism of Cu-catalysed 1,3 dipolar cycloaddition as a stepwise process involving copper-alkyne intermediates (Rostovtsev et al., 2002), a hypothesis which was subsequently proved in 2013 by Fokin and Worrell (Scheme 3.1). The presence of Cu as a catalyst in the reaction increases the reaction's rate constant enormously (Worrell et al., 2013, Rostovtsev et al., 2002).



Scheme 3.1 Proposed mechanism of 1,3-dipolar cycloaddition by copper-catalysed click reaction (Adapted from Worrell et al. (2013), Rostovtsev et al. (2002))

For copper-free click reactions, cyclooctyne was used instead of a terminal alkyne. The alkyne moiety in cyclooctynes has an internal bond angle of 160°, compared to 180° for a typical alkyne, generating a ring strain and elevating its ground state energy (Jewett and Bertozzi, 2010). This ring strain can be increased by adding fluorine atoms onto the ring (Sletten and Bertozzi, 2009), or by incorporating benzene rings to cyclooctyne to create dibenzocyclooctyne (Ning et al., 2008) (Figure 3.6). There are continuing attempts to generate new reagents for copper-free click reactions, with the aim to increase the rate of the reaction (Sletten and Bertozzi, 2009, Ramil and Lin, 2013). However, cyclooctyne-based molecules share the same reaction mechanism (Scheme 3.2).

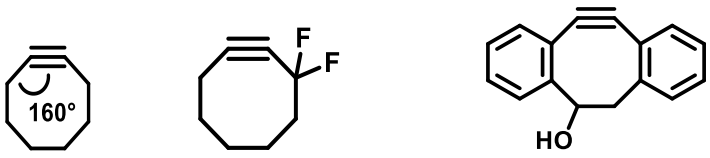
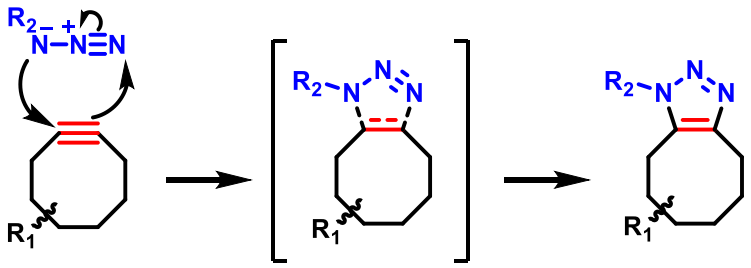


Figure 3.6 From left, cycloocytne, cyclooctyne with 2 fluorine atoms (DIFO), and dibenzocyclooctyne (DIBO)



Scheme 3.2 Proposed mechanism of 1,3-dipolar cycloaddition of cyclooctyne and azide

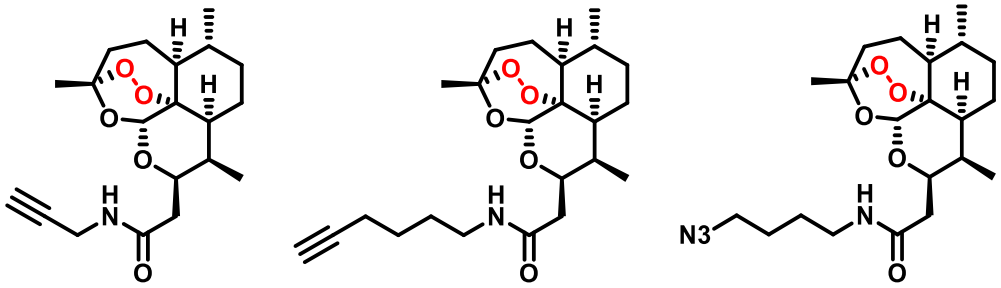
3.2 Experimental

This section will briefly outline the materials and methods used in this chapter. Detailed materials and method can be found in Chapter 2.

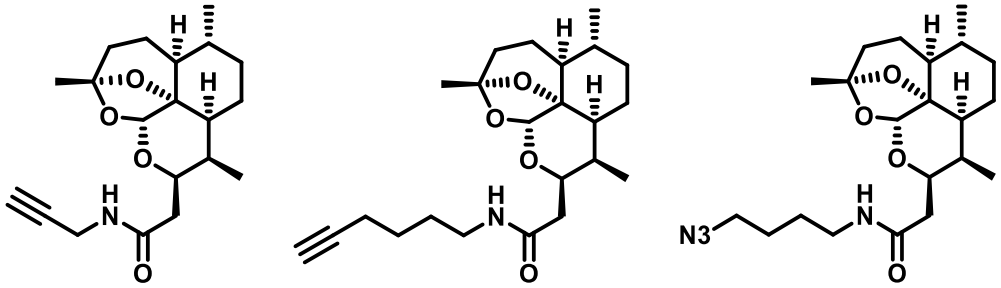
3.2.1 Chemical probes

Probes used in this study were synthesised by Sitthivut Charoensutthivarakul and Michael H.L. Wong under a supervision of Paul O'Neill at Department of Chemistry, University of Liverpool. The rationale of probe design was to mimic the structure of artemisinin with a clickable functional group, either a terminal alkyne or azide.

In brief, a series of chemical probes were synthesised in 6-10 steps, in moderate yields. The chemicals used in this synthesis are commercially available, including dihydroartemisinin, and they were used without any prior purification. The intermediates and products from each step were purified by chromatography or crystallisation and fully characterised, as appropriate, by ^{1}H NMR, ^{13}C NMR, high resolution MS and elemental analysis providing the purity of >95%. The chemical probes were then transferred to LSTM and kept under atmospheric conditions in the absent of light prior to its uses. The synthetic schemes for those probes are published elsewhere (Ismail et al. 2016a; Ismail et al. 2016b). For each active probe containing the peroxide bridge and inactive desoxy control probe was also produced.



ART- 15 SCR-A-17 SCR-A-10



ART-21 SCR-A-21 SCR-A-22

Figure 3.7 Structures of chemical probes

3.2.2 Parasite treatment

Briefly, trophozoite stage *Plasmodium falciparum* 3D7 parasites were maintained as described in section 2.13 and synchronised for the ring stage by sorbitol (section 2.14). Then parasites were maintained until they reached the desired parasitaemia and life-cycle stage. Parasites were treated with 1 µM probe or DMSO (as negative control) for 6 h under normal culture condition (section 2.23). Then parasites were freed from RBC by saponin lysis (section 2.18), washed with D-PBS, resuspended in D-PBS with 1X cOmplete™ EDTA-free protease inhibitor (Roche), and stored at -80°C until further processing.

3.2.3 Click reaction and peptide preparation

Treated parasites samples were retrieved at RT and cell lysis was assisted with probe sonication. The sample protein concentration was determined using the Bradford assay (section 2.24). The corresponding samples were adjusted for equal concentration prior to initiating click reactions.

The click reaction was performed following the protocol described in section 2.26. Briefly, for copper-catalysed click reactions, biotin-azide, TCEP, TBTA, and $CuSO_4$ were added sequentially. For copper-free click reactions, biotin-DIBO alkyne was added to the samples. The samples were then incubated at RT for an hour, away from light and with gentle mixing every 15 min. After the incubation period, pellets were obtained by centrifugation and washed three times with 750 µL cold absolute methanol. Pellets

were then resuspended in 650 µL 2.5% SDS in D-PBS and heated to 90ºC. Undissolved pellets were removed after centrifugation and supernatants were adjusted with D-PBS to 3.5 mL and stored at -20ºC.

Protein samples then exposed to a streptavidin enrichment step to remove unbound proteins followed by alkylation and reduction, respectively. Briefly, samples collected from the click reaction were adjusted to 8.3 mL D-PBS and incubated with 200 µL streptavidin agarose beads for 1.5 h. After incubation the beads were collected, and washed respectively with 1% SDS, 6 M urea, and D-PBS, each 3 times to completely remove unbound proteins. Then on-bead alkylation with DTT and reduction with IAA were performed before overnight incubation with trypsin for digestion. Tryptic digestion was terminated with 90% formic acid prior to MS analysis. For the detailed protocol see Chapter 2.

3.2.4 Click reaction for SDS-PAGE analysis

SDS-PAGE was used to visualise protein labelling profiles from the protein extracts. Protein extracts from probe treatments were incubated with 20 µM Click-IT® Alexa Fluor® 488 DIBO alkyne for 1 h and washed trice with 1 mL cold absolute methanol to remove unbound dye (section 2.26). Protein samples were then analysed by NuPAGE™ Bis-Tris 4-12% gel following the method described in section 2.22. Fluorescence was detected using SYBR green setting in GE ImageQuant™ 4000 system (section 2.26). Next the gel was stained with Instant Blue coomasie-based staining for 1 h and then de-stained overnight, coomasie images was captured using a Bio-Rad Gel Doc™ EZ gel documentary system (section 2.22).

3.3 Results

3.3.1 Antimalarial activity of chemical probes

The antimalarial activity of the probes was tested using the conventional SYBR Green I-based IC_{50} assay (section 2.17). All synthetic artemisinin activity-based probes were active exhibited a similar antimalarial activity level compared to artemisinin and dihydroartemisinin, whereas desoxyartemisinin (non-peroxidic) probes provide no antimalarial activity as expected at the concentrations tested. Activity of the probes was checked regularly to ensure that the chemical constituents of the probes remained the same throughout the following experiments. IC_{50} values for artemisinin and dihydroartemisinin were 17.3 nM (95% CI = 13.46-22.24 nM) and 2.92 nM (95% CI = 2.33-3.65 nM), respectively. IC_{50} values for artemisinin-containing probes, namely ART-15, SCR-A-10, and SCR-A-17 were 18.04 nM (95% CI = 12.70-25.57 nM), 5.26 nM (95% CI = 3.92-8.41 nM), and 5.74 (95% CI = 4.22-655 nM), respectively (Figure 3.8).

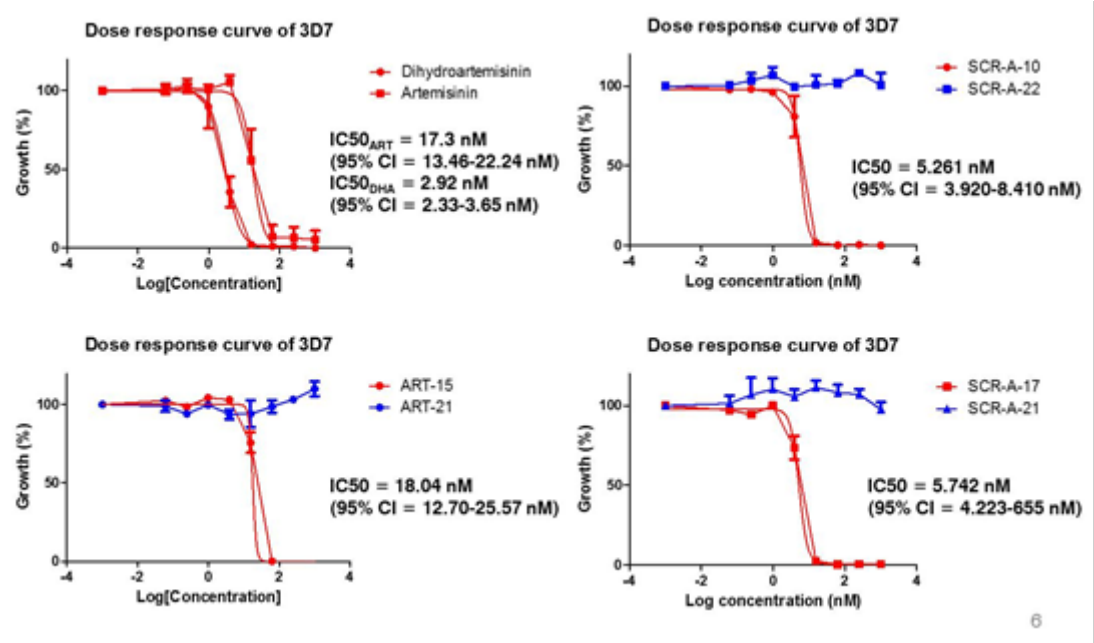


Figure 3.8 Dose response curves of the chemical probes

3.3.2 Gel electrophoresis

Using the protocol for click reaction and fluorescent gel analysis and SDS-PAGE techniques described in section 2.22 and 2.26. Protein extracts from trophozoite stage parasites, post-treatment with azide probes, were subjected to a copper-free click reaction with Click-IT® Alexa Fluor® 488 DIBO alkyne and separated through a 4-12% Bis-Tris gel. The gel was then imaged under blue LED excitation and the Y515 BP Cy2 filter using a GE ImageQuant™ LAS 4000 biomolecular imager.

Figure 3.9 shows clicked proteomes from two individual trophozoite protein experiments labelled with fluorescent conjugated DIBO alkyne or azide. It demonstrated that the artemisinin-based probe SCR-A-10 and ART-15 labelled more proteins while their inactive counterpart SCR-A-22 and ART-21 were not able to label proteins after a 1 h click reaction. Coomasie staining gels confirmed that this selective labelling effect was not due to unequal loading of proteins but a direct function of the activity of the probes themselves.

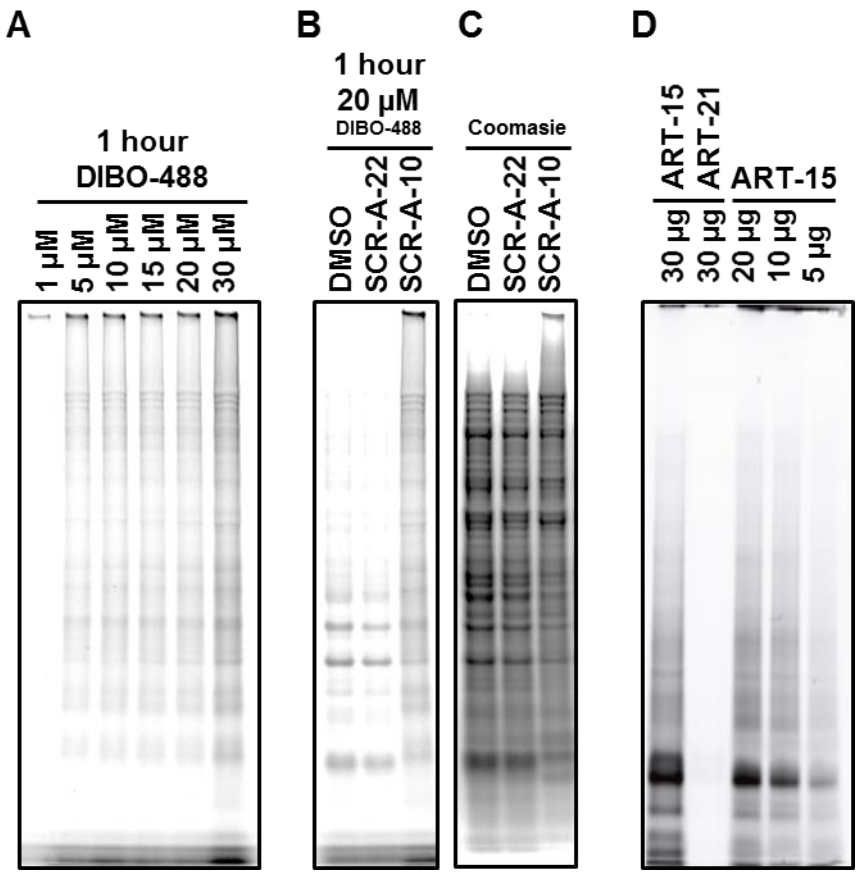


Figure 3.9 Copper-free click chemistry visualised by SDS-PAGE. (A) DIBO alkyne-488 concentration titration range from 1 µM to 30 µM. (B) DIBO alkyne-488 labelling profiles of DMSO, desoxyartemisinin probe SCR-A-22, and artemisinin activity-based probe SCR-A-10 treatments. (C) Coomasie staining of the same gel in B. (D) Trifunctional biotin-azide-rhodamine labelling profile of artemisinin activity-based probe (ART-17) and desoxyartemisinin probe (ART-21) and protein concentration variation, gel image Dr Hanafy Ishmail.

3.3.3 Protein concentration

The protein concentration of each sample was determined using a standard Bradford protein quantification assay. The BSA Protein standard was prepared prior to each use. Sample concentrations from corresponding treatments were adjusted to concentrations of 2 or 1.5 or 1 or 0.5 mg/ml, where applicable, prior to initiation of the click reaction.

3.3.4 Protein IDs and gene IDs mapping

Results returned from the protein identification step by MASCOT and SEQUEST were in UniProt ID format, which was not fully compatible with some databases, so mapping UniProt ID to systematic gene ID was performed in the PlasmoDB database. Any missing gene IDs were manually integrated into the list. Lists of identified and tagged proteins can be found in Chapters 4 and Chapter 5. Gene IDs or UniProt protein IDs were used where applicable, however gene IDs were used as the first choice.

3.3.5 Artemisinin-activity based probes vs control probes and DMSO

Artemisinin probes and their desoxy counterparts have been introduced as a pairwise strategy in the study. Drug inhibition assays have already shown that there was a clear difference between them in terms of antimalarial activity (Figure 3.8). HPLC-MS/MS analysis confirmed the stringency of the strategy.

Protein extracts from parasites treated with DMSO, artemisinin-activity based probes, and desoxyartemisinin probes were processed and sequenced by HPLC-MS/MS as previously described in Chapter 2. A clear difference is seen in the numbers of proteins identified from artemisinin-activity based probes, desoxyartemisinin probes and the DMSO control group (Figure 3.10). Protein hits identified from the DMSO group and desoxyartemisinin group were considered as background proteins and were therefore excluded from artemisinin-activity based probes protein hits when further analysed.

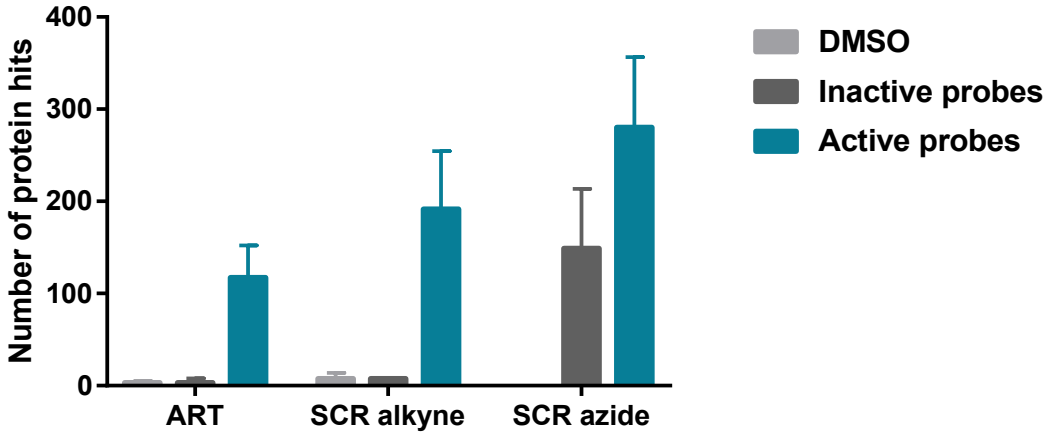


Figure 3.10 Bar graph showing number of protein *hits identified by different probes*

3.3.6 Difference between short and long linker click probes

Two different lengths of carbon linker were used in the study. ART probes possess a 1-carbon linker between the artemisinin moiety and a tag while SCR probes are equipped with a 4-carbon linker. These 2 probe lengths led to a differences in the number of alkylated protein hits, however they largely overlapped. SCR probes with a 4-carbon linker labelled more proteins than the 1-carbon (ART-15) probe albeit insignificantly (p-value >0.05).

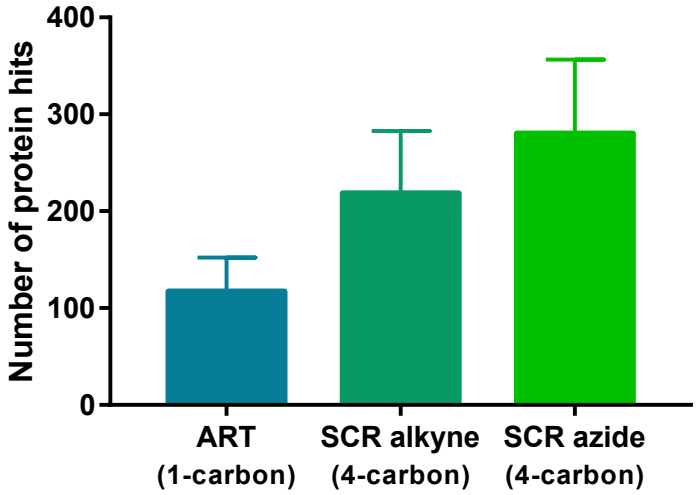


Figure 3.11 Number of protein hits identified by 1-carbon linker (ART) probe and 4-carbon linker (SCR) probes. These results were total number of protein identified by artemisinin activity-based probes before subtracting background protein from DMSO and desoxyartemisinin probes.

However, when considering unique protein hits identified by each artemisinin activity-based probe, this difference in the number of protein hits identified was smaller (Figure 3.12).

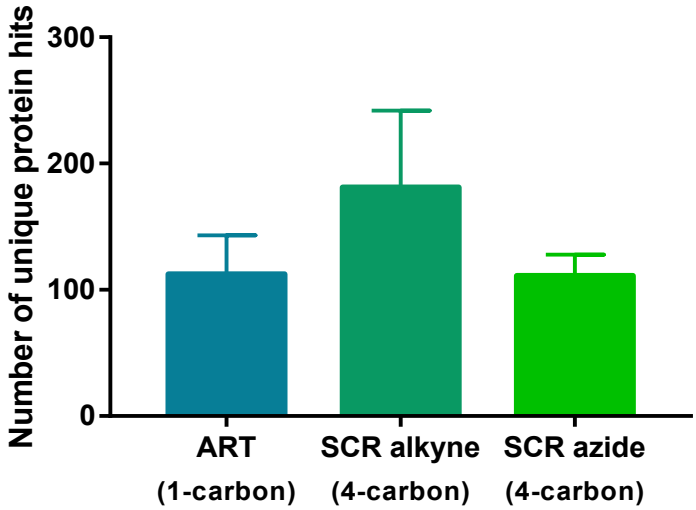


Figure 3.12 Number of unique protein hits identified by artemisinin activity-based probe. Protein hits identified by DMSO and desoxyartemisinin probe were excluded.

3.3.7 Copper-catalysed vs copper-free click reaction

Two versions of the alkyne probes: 1-carbon linker and 4-carbon linker, were used in the copper-catalysed click reaction. For copper-free reactions, only the 4-carbon linker azide probe was used.

The numbers of unique protein hits identified from 1-carbon linker and 4-carbon linker alkyne probes were 86 and 125 proteins respectively. Both probes labelled 45.5% (66 proteins) in common (Figure 3.15). The 4-carbon linker alkyne and azide probes labelled 125 and 121 proteins, respectively, and shared 20% common proteins identified (Figure 3.13).

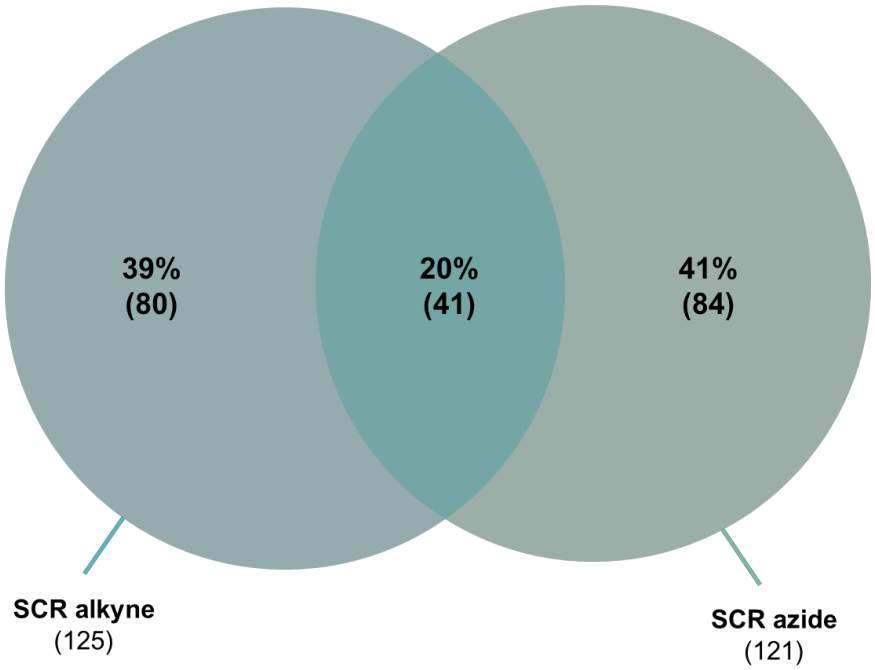


Figure 3.13 Proportional Venn diagram showing unique protein hits by SCR alkyne and SCR azide artemisinin activity-based probes

Both artemisinin probes with an alkyne or azide tag labelled a similar number of proteins. However, both reactions resulted in different levels of background protein labelling. Both 1-carbon linker and 4-carbon linker desoxyartemisinin alkyne probes labelled only a few proteins that were considered as a background, whereas the 4-carbon linker desoxyartemisinin azide probe labelled many more background proteins, up to 100 (Figure 3.10, Table A1.3 and Table A1.4).

3.4 Discussion

In this section 2 pairwise strategies are discussed. In the first analysis, artemisinin activity-based probes (active) and desoxyartemisinin probes (inactive) are compared and discussed thereafter the copper-catalysed and copper-free reaction are considered.

3.4.1 Antimalarial activity of chemical probes

It is widely accepted that the unusual peroxide structure in the artemisinins, or other synthetic endoperoxides contributes to the potent antimalarial activity of these compounds, since their non-peroxidic equivalents exhibit no antimalarial effects (Kaiser et al. 2007; O'Neill, Barton & Ward 2010). It was also demonstrated here by conventional IC_{50} assays that the artemisinin activity-based probes have retained antimalarial activity similar to parent artemisinin and dihydroartemisinin, while their desoxyartemisinin counterparts were inactive against asexual stages of *Plasmodium falciparum* 3D7 (Figure 3.8). Importantly the addition of the click liker did not compromise antimalarial activity. This is in accordance with previous studies which confirms that desoxyartemisinin is unable to generate carbon centred radicals which leads to the complete loss of antimalarial activity (O'Neill et al., 2010). Desoxyartemisinin probes could provide non-activity controls for background protein labelling increasing labelling confidence of artemisinin activity-based probes. The use of these two types of control, DMSO and desoxy inactive probes, is a unique strength of the strategy deployed here in this study.

Results from the HPLC-MS/MS approach also supported the importance of the 1,2,4 trioxane structure hypothesis and confirmed the value of the pairwise analysis strategy. DMSO and desoxyartemisinin probes labelled a small number of proteins from parasite extracts, while artemisinin probes labelled an order of magnitude more proteins (Figure 3.10). Some proteins have been reported previously to directly bind to streptavidin, which was used in the pull-down process in the study. For example, histone proteins were reported to directly bind to streptavidin independently of biotin. It is also evident that histone proteins are post-translationally modified by biotinylation (Bailey et al., 2008). For this reason, histone proteins were excluded from the protein hits identified by artemisinin-activity based probes. Other sequences known for their high affinity to streptavidin including HPM, HPQ, HPYP, and HPFP motifs were also omitted from the protein hits identified by artemisinin-activity based probes (Lam and Lebl, 1992).

3.4.2 Probe linker length and protein number

It is hypothesised that the 2 different lengths of carbon linker between a pharmacophore and a tag can result in different numbers of protein hits being identified. It is possibly due to the fact that a longer linker probe may provide more accessibility to fit inside the target proteins while being able to present

clickable functional group, alkyne or azide, for click reaction. Results from HPLC-MS/MS analysis revealed that the probes with longer linkers, SCR, labelled more protein than the shorter linker counterpart, ART, albeit not significantly (p-value >0.05).

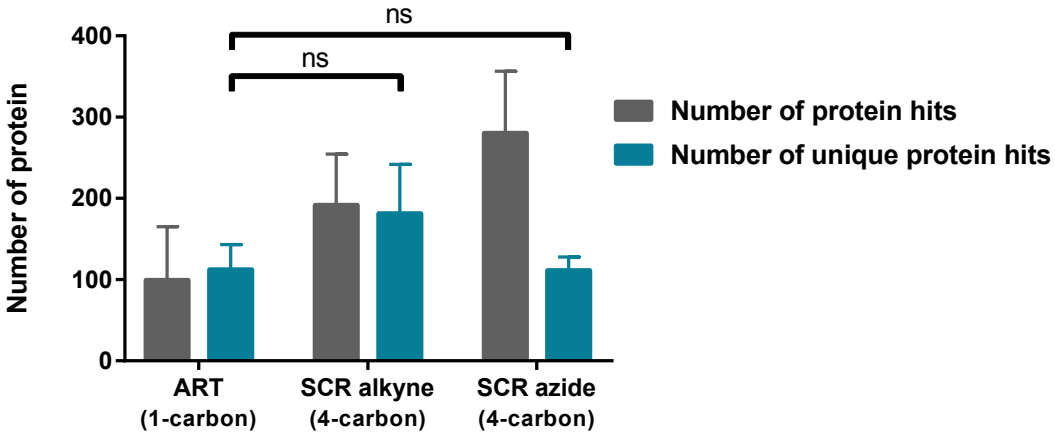


Figure 3.14 Number of proteins identified by artemisinin activity-based probes. Blue bar showing all proteins identified by artemisinin activity-based probes and red bar showing unique proteins identified artemisinin activity-based probe (control subtracted).

In contrast to the working hypothesis, the result shows that there is no significant difference in number of proteins identified from 1-carbon linker and 4-carbon linker probes. In addition, proteins identified by 1-carbon and 4-carbon alkyne probes overlapped by 45.5% (Figure 3.15). This result suggested that the use of a longer linker in the chemical probe can improve the efficiency of protein labelling while maintain the selectivity of artemisinin activity-based probes.

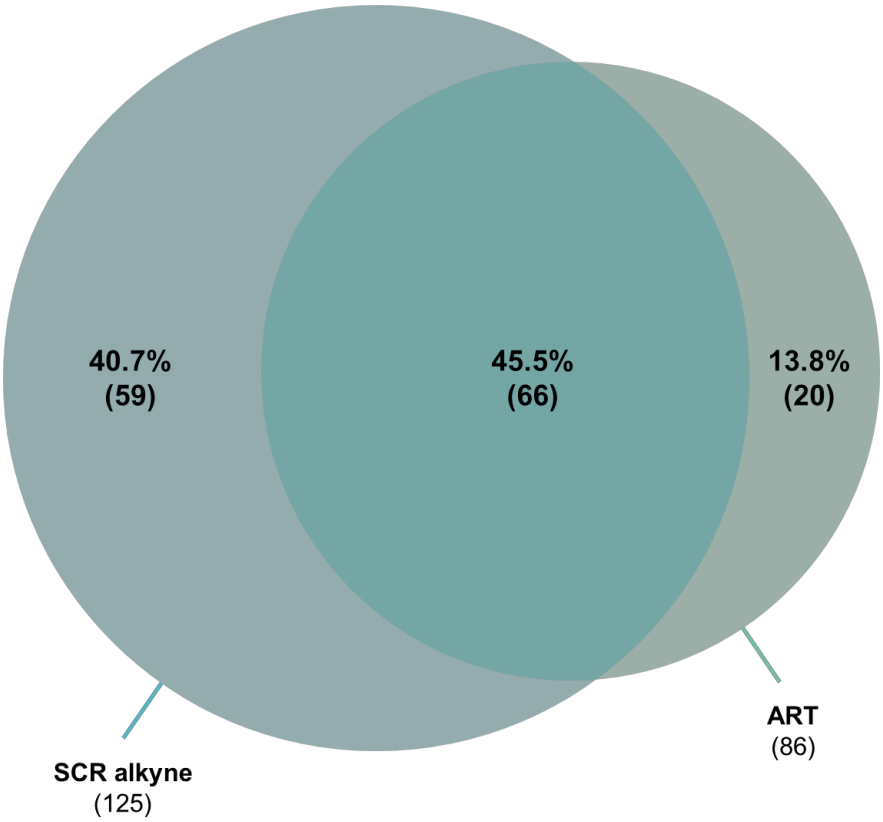


Figure 3.15 Proportional Venn diagram showing unique protein hits by ART and SCR alkyne artemisinin activity-based probe

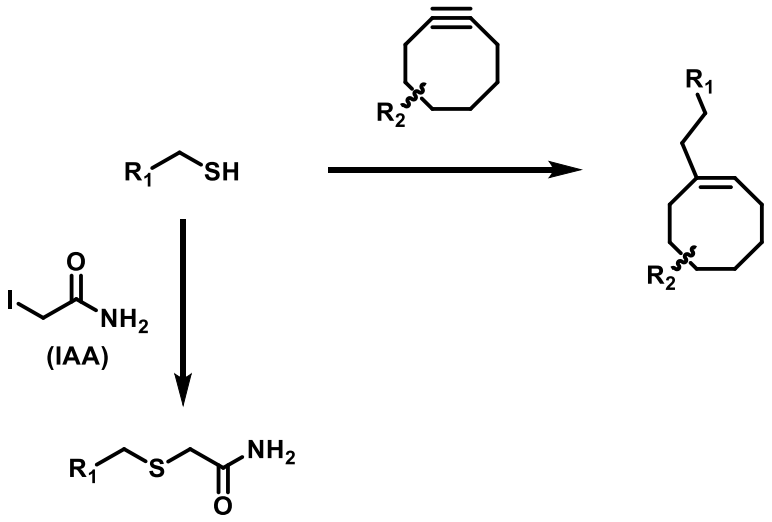
3.4.3 Copper-catalysed vs copper-free click reactions

Since click reactions can be performed either with or without copper through the Huisgen's 1,3-dipolar cycloaddition, both chemistries were included in the study as a chemistry pairwise comparison strategy. The presence of copper in the reaction can be toxic to living cells (Ramil and Lin, 2013) and it was unclear if copper could act like iron in cleaving any remaining active peroxide containing drug that might remain post processing and therefore generating adducts that were an artefact of the processing. In this section, the aim was to study the necessity of copper as a catalyst in the biorthogonal click chemistry. Although click reactions in this study were performed *in vitro*, where copper causes much less interference and toxicity, both chemistries were adopted to compare the labelling profiles.

The results from 4-carbon alkyne and azide artemisinin activity-based probes showed that both probes identify a similar number of protein (125 and 121 proteins) but only 20% of proteins were identified by both sets of conditions. Also the higher background protein labelling from desoxyartemisinin probe treatment might discourage the use of copper-free click reactions. However, the copper-free click reaction is preferred in most applications due to its higher sensitivity and lower toxicity in living cells (Jewett and Bertozzi, 2010). The difference of both chemistry approaches lies in the reactivity of the chemicals used in each reaction.

For copper-catalysed click reactions, $CuSO_4$ was used to catalysed the 1,3 Huisgen's cycloaddition with TBTA as a ligand for copper and TCEP as a reducing agent for proteins. Considering the chemistry basis of the reaction, the terminal alkyne is a relatively stable form of carbon-carbon triple bond with its bond angle of 180º while the azide is known as a good leaving group that is sensitive to a nucleophilic attack (Jewett and Bertozzi, 2010). These two functional groups are not reactive without a catalyst due to the large energy gap between ground and transition states. To overcome this energy gap and drive the reaction forward, Huisgen introduced [3+2] cycloaddition or 1,3-dipolar cycloaddition, by increasing the reaction temperature to higher than 100ºC, however this is not applicable in biological systems (Huisgen, 1963). Sharpless and colleagues have proposed that this 1,3-dipolar cycloaddition can be achieved by using copper as a catalyst instead of heat (Himo et al., 2004). This copper-catalysed 1,3-dipolar cycloaddition allows the reaction to occur at physiological temperatures and is feasible for biological applications.

As opposed to copper-catalysed click reactions, copper-free click reactions are less toxic to living cells (Ramil and Lin, 2013). However, when considering chemicals used in the reaction, some nonspecific reactions can be detected as the reagent used in the reaction is quite reactive, thus allowing the reaction to occur without a catalyst (Ramil and Lin, 2013). In a copper-free click reaction, 1,3-dipolar cycloaddition is driven by the ring strain of the cyclooctyne-based reagent (Jewett and Bertozzi, 2010). This strain is generated by the alkyne bond angle in the cyclooctyne moiety which is 160º instead of 180º (Figure 3.6). This stress in the molecule reduces the energy gap between ground state and transition state, increasing the rate constant of the reaction (Sletten and Bertozzi, 2009, Jewett and Bertozzi, 2010, Ramil and Lin, 2013). Nonetheless, using DIBO-biotin and other cyclooctynes in the reaction could lead to nonspecific binding (Yao et al., 2012, Subramanian et al., 2014) as DIBO can react and conjugate to not only azide but also to free cysteine residues via a thiol-yne reaction (van Geel et al., 2012). This nonspecific labelling can be reduced by adding reducing agents, e.g. β-mercaptoethanol or iodoacetamide (IAA), to alkylate the terminal thiol group of cysteine (van Geel et al., 2012, Chen and Wu, 2016, Tian et al., 2016) (Scheme 3.3).



Scheme 3.3 Protecting effect of reducing agent to cysteine residue from cyclooctyne. (adapted from van Geel et al. (2012))

To control for any copper activation of the artemisinin, the copper-free click reaction was introduced to address this. Results from the 4-carbon tagged alkyne and azide artemisinin activity-based probes showed 20% overlap in the labelled proteome (41 proteins). However, it was unlikely that copper activation of artemisinin was an issue as the copper-catalysed click reaction was performed in a reducing environment in the presence of TCEP (tris(2-carboxyethyl)phosphine) (Meshnick et al., 1989). The reducing environment was shown to antagonise the effect of artemisinin (Krungkrai and Yuthavong, 1987, Meshnick et al., 1989). In addition, several washing steps after probe treatment should remove free non-activated artemisinin probes, therefore there should be no artemisinin available for activation by copper during the click reaction.

The background from non-specific labelling by DIBO alkyne could be minimised by reducing the concentration of DIBO. When the concentration of DIBO in the reaction was reduced from 50 μM to 20 μM, no proteins were identified from desoxyartemisinin probe treatments. (This experimental part had been performed in association with Dr Hanafy Ismail) (Ismail et al., 2016a).

3.4.4 Protein localisation coverage

There is a concern that this approach will not be able to label possible target proteins localised on membranes. Protein hits identified by artemisinin-activity based probes were subjected to search for membrane localised proteins (GO term 0016020; membrane). Sixty-nine proteins uniquely identified by artemisinin activity-based probes were annotated with GO membrane (Table 4.7). This shows that this approach can cover a wide range of protein with multiple localisations. This is made possible due to the fact that the sonication extraction method adopted is powerful enough to disrupt membranes and release membrane localised proteins. However, it is worth noting that membrane localised proteins

might be either disaggregated or part of smaller soluble membrane fractions (Rosenberg and McIntosh, 1968).

3.4.5 Protein labelling is not due to abundance bias

In order to address the criticism that protein labelling is not specific to the targets for artemisinin but a measure if protein abundance in the parasite preparation, a search against protein abundance levels was evaluated using data on abundance published by Pease et al. (2013). The analysis showed that the artemisinin tagged proteins identified by this approach were distributed equally across all expression levels and did not form an identifiable clusters corresponding to highly expressed proteins (Figure 3.16).

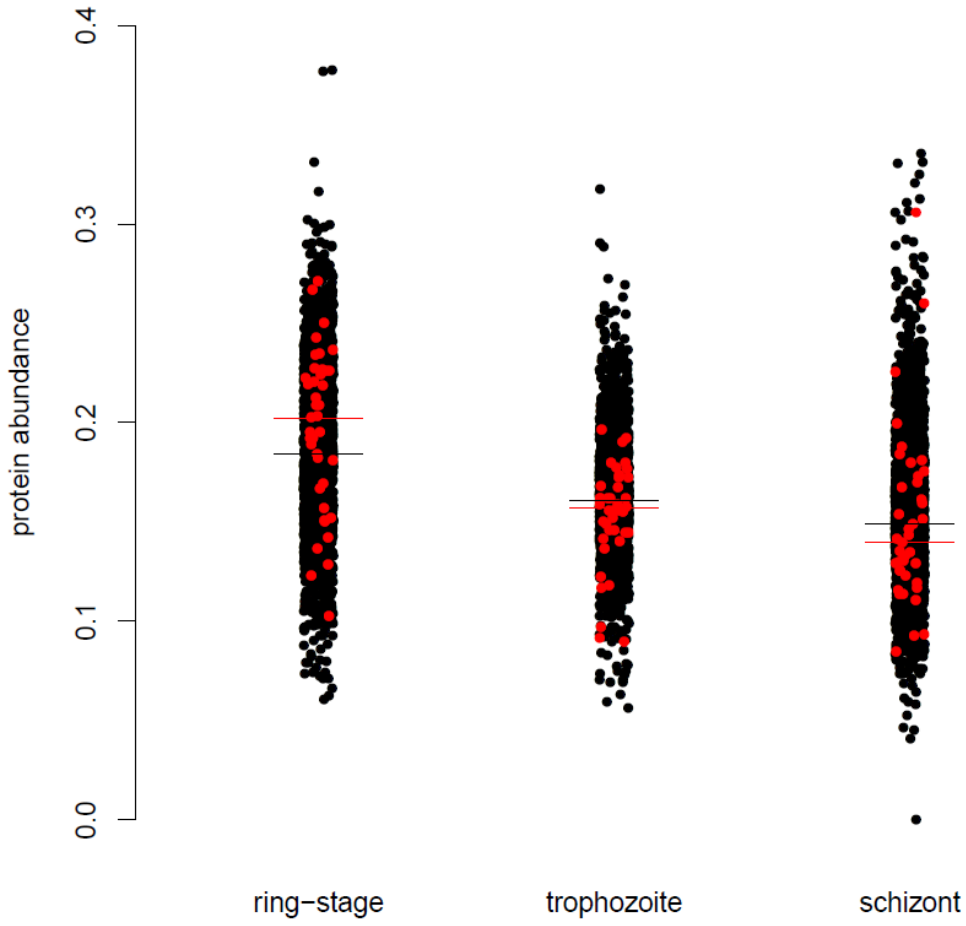


Figure 3.16 Proteins expression level. Red dots represent top 40 proteins from the hits. Black dots represent all 2605 proteins in (Pease et al., 2013). This figure was generated by Dr Gareth Weedall (LSTM bioinformatics unit).

3.5 Conclusions

The advances in mass spectrometry-based proteomics has been extensively used in biological research thanks to its ease of use and high sensitivity. Recently developed click chemistry is a growing application in chemical biological research. Combining these two powerful tools together has provided a new approach for proteomic identification of potential drug targets with high specificity as previously discussed throughout this chapter. Both chemical and activity pairwise analysis strategies introduced in the study provided high confidence in the proteins identified as potential artemisinin molecular targets. Copper-catalysed and copper-free click reactions were both capable of tagging and pulling down potential targets albeit with different efficiencies.

Chapter 4

Proteins identified by artemisinin activity-based probes of trophozoite stage parasites

4.1 Introduction

The mechanism of action and resistance of artemisinin and its derivatives has remained poorly understood for long time, despite these drugs being on the centre stage of antimalarial drug research and clinical use. Reduced sensitivity to artemisinins and ACTs treatment failures are threatening global attempts to control and eradicate malaria. Many studies have proposed molecular targets of artemisinin and its derivatives (see Chapter 1). One hypothesis is that these drugs have a global effect on parasite protein alkylation, however experiments showing this have not been established. In this study, click chemistry and MS proteomics approaches were applied to identify molecular targets of artemisinin.

Click chemistry-based approach has been developed first as chemical synthetic approach and later applied to biological research. Click chemistry approach had successfully used to identify the protein partners of phosphatidylinositol 3,4,5-triphosphate or PIP_3 (Rowland et al., 2011), the targets of serine protease inhibitors (Adibekian et al., 2012), and the molecular targets of artemisinin in *Plasmodium falciparum* parasites (Wang et al., 2015). The work by Rowland et al. (2011) used synthetic modified PIP_3 with alkyne tag, copper-catalysed click reaction, and LC-MS/MS to identify the partner proteins. Using aforementioned approach, 265 proteins were identified as PIP_3-binding proteins. A work by Adibekian et al. (2012) used the modified serine protease inhibitors to confirm that hydrolases lysophospholipase 1 (LYPLA1) and lysophospholipase 2 (LYPLA2) are selectively inhibited by the proteases in mouse model. More recently, click chemistry approach was applied to identify molecular targets of artemisinin (Wang et al., 2015). A copper-catalysed click reaction with a modified artemisinin alkyne probe successfully identified 124 artemisinin targets in mixed stage parasites (Wang et al., 2015). In addition, copper-free reaction was applied for *in vivo* imaging (Baskin et al., 2007). This growing application of click chemistry in biological research will advance the drug target research with better click reagents in the future.

In this study, it has been shown in Chapter 3 that click chemistry proteomic approach using artemisinin activity-based probes is a robust platform to study molecular targets of artemisinin. In this chapter, proteins identified from trophozoite stage parasites will be discussed in context of artemisinin mechanism of action and possible resistance mechanism. It was shown that the trophozoite stage of *Plasmodium falciparum* is much more sensitive to artemisinin than the ring stages (Klonis et al., 2013). A working hypothesis is that artemisinin is first activated by haem or free iron in the parasites as iron

chelators and hamoglobinase inhibitor have antagonistic effect to artemisinin (Stocks et al., 2007, Xie et al., 2016), then the global damage of the parasite proteins is caused by alkylation by activated artemisinin. The benefits of performing an experiment with the trophozoite stage parasites are relatively high biomass, easy to purify by saponin lysis, and high metabolic activity of the parasites. This higher sensitivity to artemisinin of the trophozoite stage parasites could provide some insight into how artemisinin affects the parasite. In this chapter, probe-labelled proteome or '**endoperoxome**' is identified and discussed.

4.2 Experimental

4.2.1 Parasite treatment and protein processing

In order to identify artemisinin targets, chemically modified artemisinin and desoxyartemisinin probes (Figure 3.7) were used to treat the parasites at synchronous trophozoite stage (~18-24 hpi) for 6 h at 1 μM concentration. Parasites were lysed and released from RBCs by saponin lysis (section 2.18) and proteins were extracted by sonication (section 2.24). Protein extracts underwent click copper-catalysed or copper-free click reactions according to the probe structures (section 2.26). Protein-drug complexes were purified by streptavidin affinity and cleaned to remove unbound proteins (section 2.27). Proteins were digested on beads, and peptides were sequenced by MS (section 2.28).

4.2.2 Protein identification

Peptide sequencing spectra were used to identify protein ID from respective samples. RAW spectra were processed in Thermo Scientific Proteome™ 1.4 and proteins were identified by SEQUEST and MASCOT (section 2.29). Proteins identified by artemisinin activity-based probes were subtracted with background proteins identified by corresponding desoxyartemisinin probe and DMSO. Histone proteins and streptavidin binding motifs were also subtracted from the list. Confidence of proteins was assigned to each protein according to the number of probes identified.

4.3 Results

Artemisinin activity-based probes were used to treat trophozoite stage *Plasmodium falciparum* 3D7 parasites for 6 h at 1 μM. Then proteins were extracted, processed with click reaction, sequenced and identified. Using this procedure 237 proteins were identified (Table 4.1). The identified proteins were assigned confidence terms, defined by how frequently proteins were identified by artemisinin activity-based probes. Proteins identified from 8-10 experiments, and 6-7 experiments were considered very high and high confidence, respectively. Proteins assigned with moderate confidence were identified from 5-4 experiments, whereas proteins identified from 3 experiments were considered low confidence. Proteins identified from 1 or 2 experiments were regarded as noise (Table A1.6).

Table 4.1 Protein hits identified by artemisinin activity-based probes (proteins in bold were identified from more than one probe)

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1459400	Q8IKN7	conserved Plasmodium protein, unknown function	Very high	-
PF3D7_1438100	Q8IL86	secretory complex protein 62	Very high	-
PF3D7_1432100	Q8ILE3	voltage-dependent anion-selective channel protein, putative	Very high	-
PF3D7_1410700	Q8ILY8	conserved Plasmodium protein, unknown function	Very high	-
PF3D7_1410400	Q8ILZ1	rhoptry-associated protein 1	Very high	+
PF3D7_1408100	Q8IM15	plasmepsin III	Very high	+
PF3D7_1408000	Q8I6V3	plasmepsin II	Very high	+
PF3D7_1360900	C0H5J5	polyadenylate-binding protein, putative	Very high	+
PF3D7_1311900	Q76NM6	V-type proton ATPase catalytic subunit A	Very high	+
PF3D7_1311800	Q8IEK1	M1-family alanyl aminopeptidase	Very high	+
PF3D7_1306200	Q8IEQ3	conserved Plasmodium protein, unknown function	Very high	-
PF3D7_1301700	Q8IEJ0	Plasmodium exported protein (hyp8), unknown function	Very high	-
PF3D7_1237700	Q8I546	conserved protein, unknown function	Very high	-
PF3D7_1211800	Q7KQK2	polyubiquitin	Very high	+
PF3D7_1121600	Q8IIF0	exported protein 1	Very high	+
PF3D7_1118200	Q8III6	heat shock protein 90, putative	Very high	-
PF3D7_1108700	Q8IIR6	heat shock protein DnaJ homologue Pfj2	Very high	+
PF3D7_1105800	Q8IIU5	conserved Apicomplexan protein, unknown function	Very high	-
PF3D7_1104400	Q8IIV8	thioredoxin, putative	Very high	-
PF3D7_1016300	Q8I6U8	glycophorin binding protein	Very high	+
PF3D7_1010700	Q8IJT8	dolichyl-phosphate-mannose protein mannosyltransferase, putative	Very high	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1008900	Q8IJV6	adenylate kinase	Very high	+
PF3D7_0936800	Q8I2F2	Plasmodium exported protein (PHISTc), unknown function	Very high	-
PF3D7_0903200	C0H516	ras-related protein RAB7	Very high	+
PF3D7_0823800	Q8IB72	DnaJ protein, putative	Very high	-
PF3D7_0807300	Q7K6B0	ras-related protein Rab-18	Very high	+
PF3D7_0720400	Q8IBP8	ferrodoxin reductase-like protein	Very high	+
PF3D7_0702400	Q8IC43	small exported membrane protein 1	Very high	-
PF3D7_0629200	C6KTC7	DnaJ protein, putative	Very high	-
PF3D7_0628300	C6KTB9	choline/ethanolaminephosphotransferase, putative	Very high	-
PF3D7_0532100	Q8I3F3	early transcribed membrane protein 5	Very high	-
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1	Very high	+
PF3D7_0501200	Q8I488	parasite-infected erythrocyte surface protein	Very high	-
PF3D7_0309600	O00806	60S acidic ribosomal protein P2	Very high	+
PF3D7_0207600	Q9TY95	serine repeat antigen 5	Very high	+
PF3D7_1471100	Q8IKC8	exported protein 2	High	+
PF3D7_1468700	Q8IKF0	eukaryotic initiation factor 4A	High	+
PF3D7_1456800	Q8IKR1	V-type H(+)-translocating pyrophosphatase, putative	High	-
PF3D7_1454400	Q8IKT5	aminopeptidase P	High	+
PF3D7_1451800	Q8IKV8	sortilin	High	-
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase	High	-
PF3D7_1421200	Q8ILN8	40S ribosomal protein S25	High	-
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	High	+

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1347500	Q8IDM3	DNA/RNA-binding protein Alba 4	High	-
PF3D7_1345700	Q8I6T2	isocitrate dehydrogenase [NADP], mitochondrial	High	+
PF3D7_1344800	Q8IDP8	aspartate carbamoyltransferase	High	-
PF3D7_1330400	Q8IE22	ER lumen protein retaining receptor 1, putative, unspecified product	High	-
PF3D7_1328300	Q8IE43	conserved Plasmodium protein, unknown function	High	-
PF3D7_1325100	Q8IE67	phosphoribosylpyrophosphate synthetase	High	+
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase	High	+
PF3D7_1246200	Q8I4X0	actin I	High	+
PF3D7_1231100	Q8I5A9	ras-related protein Rab-2	High	-
PF3D7_1224300	Q8I5H4	polyadenylate-binding protein, putative	High	+
PF3D7_1222300	Q8I0V4	endoplasmin, putative	High	+
PF3D7_1143200	Q8IHT4	DnaJ protein, putative	High	-
PF3D7_1134000	Q8II24	heat shock protein 70	High	+
PF3D7_1130200	Q8II61	60S ribosomal protein P0	High	+
PF3D7_1129100	Q8II72	parasitophorous vacuolar protein 1	High	-
PF3D7_1129000	Q8II73	spermidine synthase	High	-
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4	High	+
PF3D7_1116700	Q8IIJ9	dipeptidyl aminopeptidase 1	High	+
PF3D7_1105000	Q8IIV2	histone H4	High	+
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2	High	+
PF3D7_1020900	Q7KQL3	ADP-ribosylation factor	High	+
PF3D7_1015600	Q8IJN9	heat shock protein 60	High	+

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase	High	+
PF3D7_0936000	C0H592	ring-exported protein 2	High	-
PF3D7_0930300	Q8I0U8	merozoite surface protein 1	High	+
PF3D7_0929400	C0H571	high molecular weight rhoptry protein 2	High	-
PF3D7_0927900	Q8I2N0	phosphatidylserine decarboxylase	High	-
PF3D7_0919100	Q8I2W2	DnaJ protein, putative	High	-
PF3D7_0914700	Q8I305	major facilitator superfamily-related transporter, putative	High	-
PF3D7_0912400	Q8I328	alkaline phosphatase, putative	High	-
PF3D7_0823200	Q8IB66	RNA-binding protein, putative	High	+
PF3D7_0818200	C0H4V6	14-3-3 protein	High	+
PF3D7_0814200	Q8IAX8	DNA/RNA-binding protein Alba 1	High	-
PF3D7_0731600	Q8I6Z1	acyl-CoA synthetase	High	+
PF3D7_0721100	Q8IBP0	conserved Plasmodium protein, unknown function	High	-
PF3D7_0709700	Q8IBZ2	lysophospholipase, putative	High	-
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter	High	-
PF3D7_0706500	C0H4L6	conserved Plasmodium protein, unknown function	High	-
PF3D7_0608800	Q6LFH8	ornithine aminotransferase	High	+
PF3D7_0601900	C6KSL9	conserved Plasmodium protein, unknown function	High	-
PF3D7_0532300	Q8I3F1	Plasmodium exported protein (PHISTb), unknown function	High	-
PF3D7_0501300	Q8I487	skeleton-binding protein 1	High	+
PF3D7_0501000	Q8I490	Plasmodium exported protein, unknown function	High	-
PF3D7_0500800	Q8I492	mature parasite-infected erythrocyte surface antigen	High	+

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_0424600	Q8IFM0	Plasmodium exported protein (PHISTb), unknown function	High	+
PF3D7_0422400	Q8IFP2	40S ribosomal protein S19	High	+
PF3D7_0402000	Q8I206	Plasmodium exported protein (PHISTa), unknown function	High	-
PF3D7_0316800	O77395	40S ribosomal protein S15A, putative	High	-
PF3D7_0106300	Q76NN8	calcium-transporting ATPase	High	-
PF3D7_1466400	Q8IKH2	transcription factor with AP2 domain(s)	Moderate	-
PF3D7_1465900	Q8IKH8	40S ribosomal protein S3	Moderate	+
PF3D7_1447700	Q8IKZ8	conserved Plasmodium protein, unknown function	Moderate	-
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	Moderate	+
PF3D7_1434800	Q8ILB6	mitochondrial acidic protein MAM33, putative	Moderate	-
PF3D7_1429600	Q8ILG8	conserved Plasmodium protein, unknown function	Moderate	-
PF3D7_1424100	Q8ILL3	60S ribosomal protein L5, putative	Moderate	+
PF3D7_1407800	Q8IM16	plasmepsin IV	Moderate	+
PF3D7_1405600	Q8IM38	ribonucleoside-diphosphate reductase small chain, putative	Moderate	+
PF3D7_1359400	Q8IDB7	CUGBP Elav-like family member 1	Moderate	+
PF3D7_1357000	Q8I0P6	elongation factor 1-alpha	Moderate	+
PF3D7_1354500	Q8IDF6	adenylosuccinate synthetase	Moderate	+
PF3D7_1353900	Q8IDG2	proteasome subunit alpha type-7, putative	Moderate	+
PF3D7_1353100	Q8IDG9	Plasmodium exported protein, unknown function	Moderate	-
PF3D7_1346100	Q8IDN6	protein transport protein SEC61 subunit alpha	Moderate	-
PF3D7_1343000	Q8IDQ9	phosphoethanolamine N-methyltransferase	Moderate	+
PF3D7_1304500	Q8IES0	small heat shock protein, putative	Moderate	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1242800	Q8I501	rab specific GDP dissociation inhibitor	Moderate	+
PF3D7_1242700	Q8I502	40S ribosomal protein S17, putative	Moderate	-
PF3D7_1241700	Q8I512	replication factor C subunit 4, putative	Moderate	-
PF3D7_1232100	Q8I0V3	60 kDa chaperonin	Moderate	+
PF3D7_1229400	Q8I5C5	macrophage migration inhibitory factor	Moderate	+
PF3D7_1228600	Q8I5D2	merozoite surface protein 9	Moderate	-
PF3D7_1223100	Q7KQK0	cAMP-dependent protein kinase regulatory subunit	Moderate	+
PF3D7_1212000	Q8I5T2	glutathione peroxidase-like thioredoxin peroxidase	Moderate	+
PF3D7_1211900	Q8I5T3	non-SERCA-type Ca2+ -transporting P-ATPase	Moderate	+
PF3D7_1211400	Q7KQK3	heat shock protein DNAJ homologue Pfj4	Moderate	-
PF3D7_1203700	Q8I608	nucleosome assembly protein	Moderate	+
PF3D7_1201000	Q8I635	Plasmodium exported protein (PHISTb), unknown function	Moderate	-
PF3D7_1149400	Q8IHM9	Plasmodium exported protein, unknown function	Moderate	-
PF3D7_1134100	Q8II23	protein disulfide isomerase	Moderate	+
PF3D7_1124700	Q8IIB6	GrpE protein homolog, mitochondrial, putative	Moderate	+
PF3D7_1124600	Q8IIB7	ethanolamine kinase	Moderate	+
PF3D7_1119000	Q8IIH7	acyl-CoA-binding protein, putative	Moderate	-
PF3D7_1117300	Q8IIJ4	conserved Plasmodium protein, unknown function	Moderate	-
PF3D7_1116800	Q8IIJ8	heat shock protein 101	Moderate	-
PF3D7_1115600	Q8IIK8	peptidyl-prolyl cis-trans isomerase	Moderate	+
PF3D7_1113300	Q8IIM9	UDP-galactose transporter, putative	Moderate	-
PF3D7_1037300	Q8IJ34	ADP/ATP transporter on adenylate translocase	Moderate	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1019900	Q8IJK2	autophagy-related protein 8	Moderate	-
PF3D7_1015900	Q8IJN7	enolase	Moderate	+
PF3D7_1011400	Q8IJT1	proteasome subunit beta type-5	Moderate	+
PF3D7_1008700	Q7KQL5	tubulin beta chain	Moderate	-
PF3D7_0934500	Q8I2H3	V-type proton ATPase subunit E, putative	Moderate	-
PF3D7_0922500	P27362	phosphoglycerate kinase	Moderate	+
PF3D7_0919400	Q8I2V9	protein disulfide isomerase	Moderate	+
PF3D7_0917900	Q8I2X4	heat shock protein 70	Moderate	+
PF3D7_0907400	Q8I377	ATP-dependent protease ATPase subunit ClpY	Moderate	-
PF3D7_0831400	C0H4Z7	Plasmodium exported protein, unknown function	Moderate	-
PF3D7_0827900	C0H4Y6	protein disulfide isomerase	Moderate	+
PF3D7_0826700	Q8IBA0	receptor for activated c kinase	Moderate	+
PF3D7_0824400	Q8IB78	nucleoside transporter 2	Moderate	-
PF3D7_0821000	Q8IB44	conserved Plasmodium protein, unknown function	Moderate	-
PF3D7_0813900	Q8IAX5	40S ribosomal protein S16, putative	Moderate	-
PF3D7_0807900	Q8IAR7	tyrosine--tRNA ligase	Moderate	-
PF3D7_0802000	Q8IAM0	glutamate dehydrogenase, putative	Moderate	-
PF3D7_0727400	Q8IBI3	proteasome subunit alpha type-5, putative	Moderate	+
PF3D7_0708400	Q8IC05	heat shock protein 90	Moderate	+
PF3D7_0624600	C6KT82	SNF2 helicase, putative	Moderate	-
PF3D7_0624000	C6KT76	hexokinase	Moderate	+
PF3D7_0621200	C6KT50	pyridoxine biosynthesis protein PDX1	Moderate	+

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_0610400	C6KSV0	histone H3	Moderate	-
PF3D7_0608300	C6KST1	conserved Plasmodium protein, unknown function	Moderate	-
PF3D7_0532400	Q8I3F0	lysine-rich membrane-associated PHISTb protein	Moderate	-
PF3D7_0517000	Q8I3T8	60S ribosomal protein L12, putative	Moderate	+
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2	Moderate	-
PF3D7_0512600	Q7K6A8	ras-related protein Rab-1B	Moderate	+
PF3D7_0509000	Q8I0X0	SNAP protein (soluble N-ethylmaleimide-sensitive factor attachment protein), putative	Moderate	-
PF3D7_0501600	Q8I484	rhoptry-associated protein 2	Moderate	+
PF3D7_0416800	Q8I1S0	small GTP-binding protein sar1	Moderate	-
PF3D7_0406100	Q6ZMA8	V-type proton ATPase subunit B	Moderate	+
PF3D7_0316700	O77388	protein YOP1, putative	Moderate	-
PF3D7_0307100	O97249	40S ribosomal protein S12, putative	Moderate	+
PF3D7_0303600	Q8I224	plasmoredoxin	Moderate	+
PF3D7_0204700	Q7KWJ5	hexose transporter	Moderate	-
PF3D7_0108000	Q8I261	proteasome subunit beta type-3, putative	Moderate	+
PF3D7_0105200	Q8I289	RAP protein, putative	Moderate	-
PF3D7_1474600	Q8IK92	conserved Plasmodium membrane protein, unknown function	Low	-
PF3D7_1468100	Q8IKF6	conserved Plasmodium protein, unknown function	Low	-
PF3D7_1464700	Q8IKJ0	ATP synthase (C/AC39) subunit, putative	Low	-
PF3D7_1453700	Q8IKU1	co-chaperone p23	Low	-
PF3D7_1451100	Q8IKW5	elongation factor 2	Low	+
PF3D7_1424400	Q8ILL2	60S ribosomal protein L7-3, putative	Low	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1420700	Q8ILP3	surface protein P113	Low	-
PF3D7_1419700	Q8ILQ3	conserved Plasmodium protein, unknown function	Low	-
PF3D7_1417500	Q8ILS0	H/ACA ribonucleoprotein complex subunit 4, putative	Low	-
PF3D7_1412300	Q8ILX1	nuclear transport factor 2, putative	Low	+
PF3D7_1408600	Q8IM10	40S ribosomal protein S8e, putative	Low	+
PF3D7_1370300	C0H5L9	membrane associated histidine-rich protein	Low	-
PF3D7_1368100	Q8ID28	26S proteasome regulatory subunit RPN11, putative	Low	-
PF3D7_1361900	P61074	proliferating cell nuclear antigen 1	Low	+
PF3D7_1361800	C0H5J9	conserved Plasmodium protein, unknown function	Low	-
PF3D7_1355100	Q8IDF0	DNA replication licensing factor MCM6	Low	-
PF3D7_1342000	Q8IDR9	40S ribosomal protein S6	Low	-
PF3D7_1338200	Q8IDV1	60S ribosomal protein L6-2, putative	Low	-
PF3D7_1331700	Q8IE10	glutamine--tRNA ligase, putative	Low	+
PF3D7_1318800	Q8IEC8	translocation protein SEC63	Low	-
PF3D7_1252100	Q8I4R5	rhoptry neck protein 3	Low	-
PF3D7_1246800	Q8I4W4	signal recognition particle receptor, beta subunit	Low	-
PF3D7_1238100	Q8I542	calcyclin binding protein, putative	Low	+
PF3D7_1226900	Q8I5E9	conserved Plasmodium protein, unknown function	Low	-
PF3D7_1149000	Q8IHN4	antigen 332, DBL-like protein	Low	+
PF3D7_1136500	C6S3F7	casein kinase 1	Low	+
PF3D7_1132200	Q8II43	T-complex protein 1 subunit alpha	Low	-
PF3D7_1126200	Q8IIA2	40S ribosomal protein S18, putative	Low	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_1126000	Q8IIA4	threonine--tRNA ligase	Low	+
PF3D7_1118500	Q8III3	nucleolar protein 56, putative	Low	-
PF3D7_1105100	Q8IIV1	histone H2B	Low	+
PF3D7_1104000	Q8IIW2	phenylalanine--tRNA ligase beta subunit	Low	-
PF3D7_1021900	Q8IJI4	PHAX domain-containing protein, putative	Low	-
PF3D7_1017900	Q8IJM0	26S proteasome regulatory subunit p55, putative	Low	-
PF3D7_1010300	Q8IJU2	succinate dehydrogenase subunit 4, putative	Low	-
PF3D7_1008800	Q8IJV7	nucleolar protein 5, putative	Low	-
PF3D7_1006700	Q8IJX4	conserved Plasmodium protein, unknown function	Low	-
PF3D7_0935800	Q8I2G2	cytoadherence linked asexual protein 9	Low	+
PF3D7_0933600	Q8I2I2	mitochondrial-processing peptidase subunit beta, putative	Low	-
PF3D7_0931800	Q8I0U7	proteasome subunit beta type-6, putative	Low	+
PF3D7_0925900	Q8I2Q0	conserved Plasmodium protein, unknown function	Low	-
PF3D7_0918300	Q8I2X0	eukaryotic translation initiation factor 3 subunit F, putative	Low	-
PF3D7_0905400	Q8I395	high molecular weight rhoptry protein 3	Low	+
PF3D7_0903700	Q6ZLZ9	alpha tubulin 1	Low	-
PF3D7_0818900	Q8IB24	heat shock protein 70	Low	+
PF3D7_0807500	Q8IAR3	proteasome subunit alpha type-6, putative	Low	+
PF3D7_0801800	Q8IAL6	mannose-6-phosphate isomerase, putative	Low	-
PF3D7_0714000	Q8IBV7	histone H2B variant	Low	+
PF3D7_0702500	Q8IC42	Plasmodium exported protein, unknown function	Low	-
PF3D7_0631900	C6KTF1	stevor	Low	-

Gene ID	UniProt ID	Product Description	Confidence	GSH
PF3D7_0627800	C6KTB4	acetyl-CoA synthetase, putative	Low	-
PF3D7_0626800	C6KTA4	pyruvate kinase	Low	+
PF3D7_0623500	C6KT71	superoxide dismutase [Fe]	Low	-
PF3D7_0619400	C6KT34	cell division cycle protein 48 homologue, putative	Low	+
PF3D7_0617800	C6KT18	histone H2A	Low	-
PF3D7_0614300	C6KSY4	major facilitator superfamily-related transporter, putative	Low	-
PF3D7_0524000	Q8I3M5	karyopherin beta	Low	-
PF3D7_0523100	Q8I3N3	mitochondrial-processing peptidase subunit alpha, putative	Low	+
PF3D7_0520900	P50250	adenosylhomocysteinase	Low	+
PF3D7_0507100	Q8I431	60S ribosomal protein L4	Low	-
PF3D7_0503400	Q8I467	actin-depolymerizing factor 1	Low	+
PF3D7_0415900	C0H4A6	60S ribosomal protein L15, putative	Low	-
PF3D7_0316600	O77389	formate-nitrite transporter	Low	-
PF3D7_0310600	Q9NFE6	eukaryotic translation initiation factor 3 subunit K, putative, unspecified product	Low	-
PF3D7_0310400	O77361	parasite-infected erythrocyte surface protein	Low	-
PF3D7_0108300	Q8I259	conserved Plasmodium protein, unknown function	Low	-
PF3D7_0102900	Q8I2B1	aspartate--tRNA ligase	Low	-
PF3D7_0102200	Q8I0U6	ring-infected erythrocyte surface antigen	Low	+

4.3.1 Pathway enrichment

GO biological process showed significant enrichment (p-value <0.05) in some metabolic pathways with protein refolding and haemoglobin metabolic process at the top. Other important parasite pathways including cell redox homeostasis, glycolysis, purine metabolism, protein folding, transport, signalling, translation, and proteolysis were also enriched (Figure 4.1).

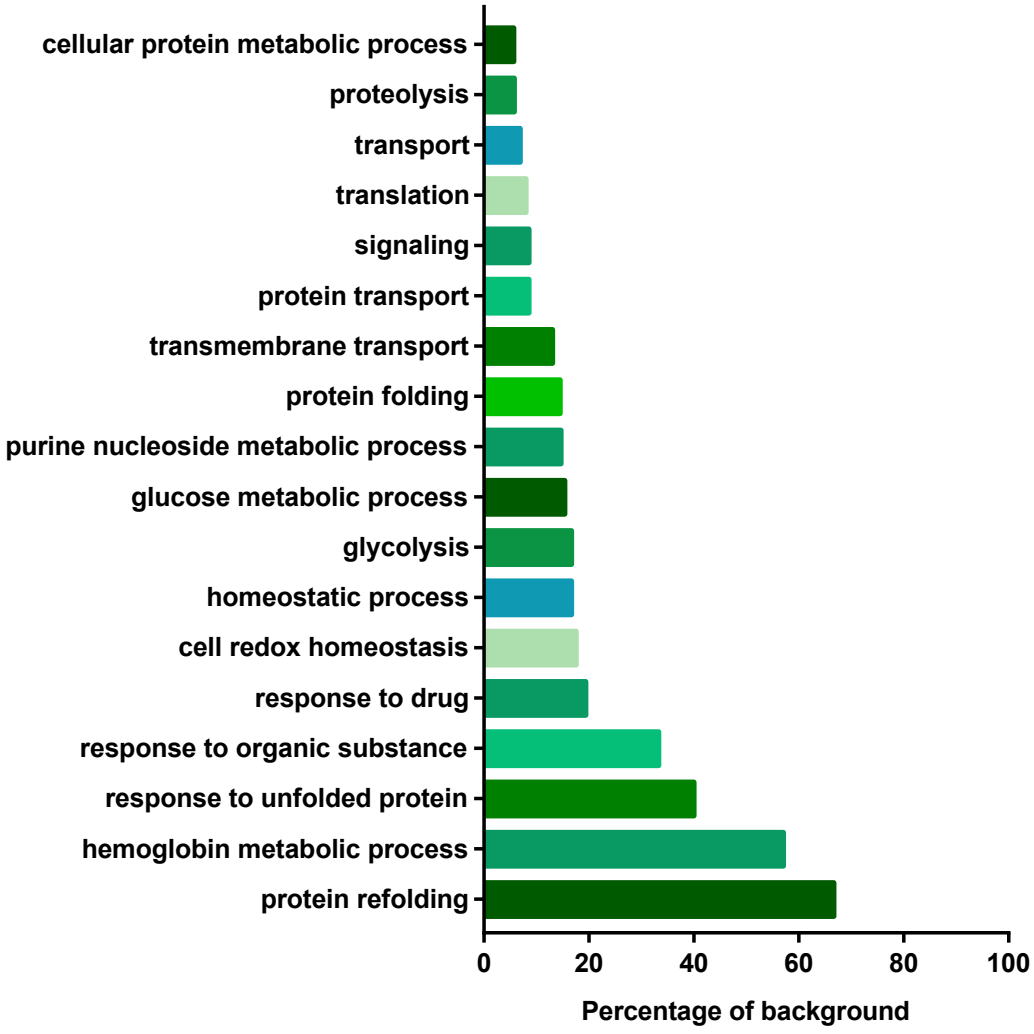


Figure 4.1 GO biological process enrichment of protein hits.

4.3.2 Protein previously proposed as possible antimalarial targets

From 237 total proteins identified, 18 proteins have previously been reported as possible molecular targets of antimalarial drugs. These include proteins identified by computational modelling and biochemical evidences (Table 4.2).

Table 4.2 Protein hits previously proposed as possible antimalarial targets

Gene ID	UniProt ID	Product Description	References
PF3D7_0106300	Q76NN8	calcium-transporting ATPase (ATP6)	(Eckstein-Ludwig et al., 2003, Muller and Hyde, 2010)
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1 (MDR1)	(Ding et al., 2011)
PF3D7_0608800	Q6LFH8	ornithine aminotransferase (OAT)	(Berger, 2000, Ludin et al., 2012)
PF3D7_0626800	C6KTA4	pyruvate kinase (PyrK)	(Crowther et al., 2010)
PF3D7_0629200	C6KTC7	DnaJ protein, putative	(Crowther et al., 2010)
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter (CRT)	(Ding et al., 2011, Muller and Hyde, 2010)
PF3D7_0823800	Q8IB72	DnaJ protein, putative	(Crowther et al., 2010)
PF3D7_1129000	Q8II73	spermidine synthase (SpdSyn)	(Plata et al., 2010, Huthmacher et al., 2010)
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase (LDH)	(Plata et al., 2010)
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	(Huthmacher et al., 2010)
PF3D7_1407800	Q8IM16	plasmepsin IV (PM4)	(Rosenthal, 1998)
PF3D7_1408000	Q8I6V3	plasmepsin II	(Rosenthal, 1998)
PF3D7_1408100	Q8IM15	plasmepsin III (HAP)	(Rosenthal, 1998)
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase (HGPRT)	(Yeh et al., 2004, Crowther et al., 2010, Plata et al., 2010, Huthmacher et al., 2010, Downie et al., 2008)
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase (FBPA)	(Huthmacher et al., 2010, Plata et al., 2010, Crowther et al., 2010, Yeh et al., 2004)
PF3D7_0709700	Q8IBZ2	lysophospholipase, putative	(Zidovetzki et al., 1994)
PF3D7_0802000	Q8IAM0	glutamate dehydrogenase, putative (GDH3)	(Aparicio et al., 2010, Werner et al., 2005)
PF3D7_1344800	Q8IDP8	aspartate carbamoyltransferase (ATCase)	(Maria Belen et al., 2011)

4.3.3 Protein concentration of artemisinin and desoxyartemisinin treated parasites

Concentration of protein extracts from parasite treatment was determined by Bradford protein assay before click reaction to normalise the starting concentration. Figure 4.2 showed protein concentrations of protein extract in paired treatment of artemisinin activity-based (active) probes and desoxyartemisinin (inactive) probes. A reduction trend was observed from most of the paired experiments, Wilcoxon test showed significant difference (p-value <0.005).

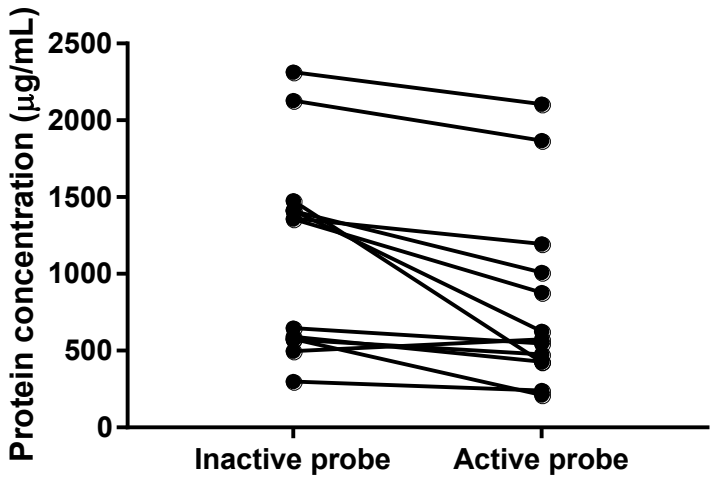


Figure 4.2 Protein concentrations from each artemisinin activity-based probes vs desoxyartemisinin probes treatments. Each point showed protein concentration and the line showed paired experiments.

4.4 Discussion

In this section, protein hits identified by artemisinin activity-based probes from trophozoite stage parasites will be discussed according to their metabolic pathways. The discussion will mainly focus on major metabolic pathways including haemoglobin digestion, glycolysis, nucleic acid and protein biosynthesis, ubiquitin-proteasome system, transporter protein, antioxidant system, and parasite-host interaction.

4.4.1 Artemisinin disrupts haemoglobin metabolism

Artemisinin activity-based probes identified 6 enzymes in haemoglobin metabolism of *Plasmodium falciparum*, namely, plasmepsin II, plasmepsin III (HAP), plasmepsin IV, aminopeptidase P (APP), dipeptidyl aminopeptidase 1 (DPAP1), and M-1 family alanyl aminopeptidase (M1AAP) (Figure 4.3).

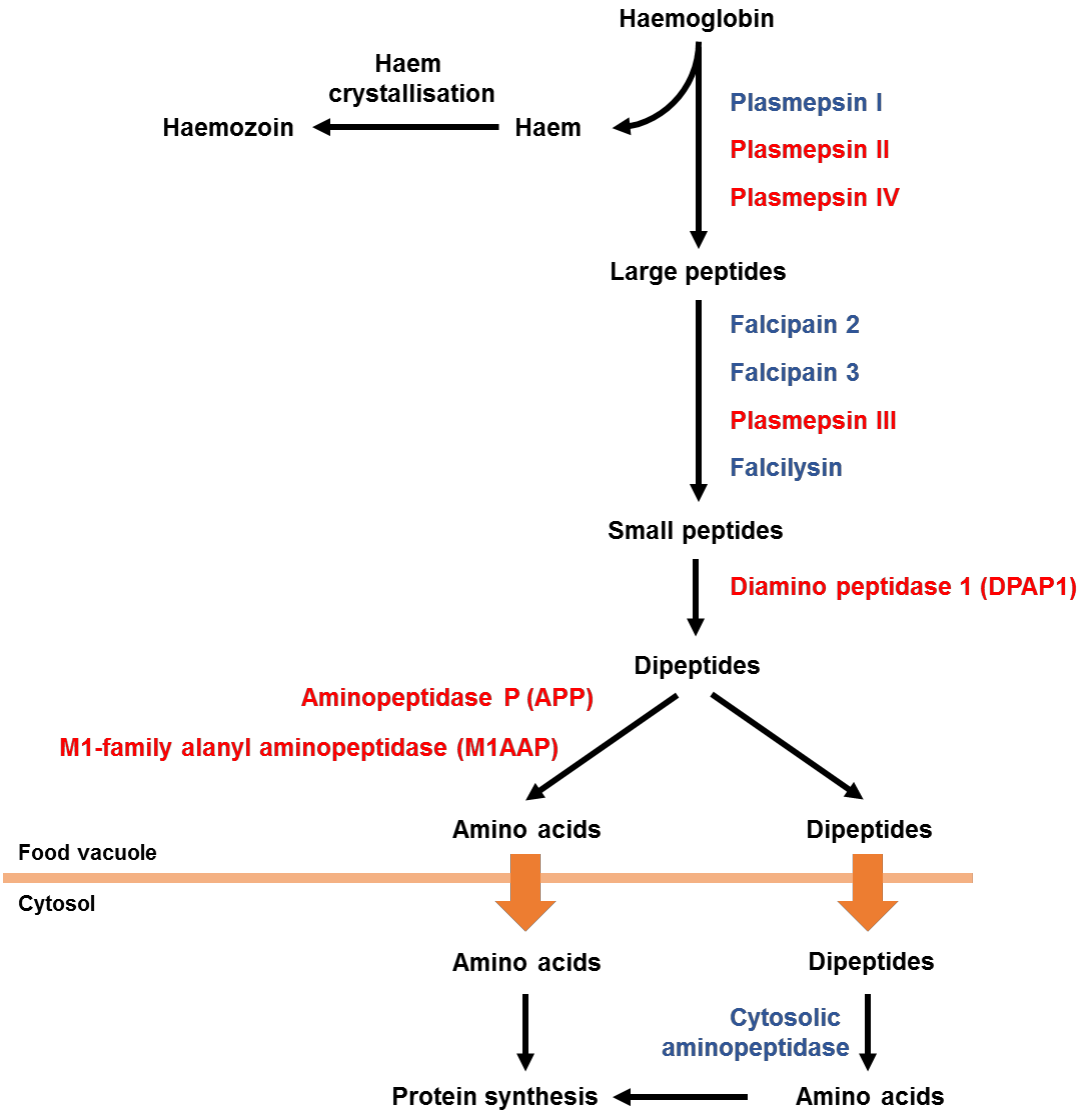


Figure 4.3 Haemoglobin catabolic pathway in Plasmodium falciparum (adapted from MPMP database). Enzymes are indicated in blue. Red indicates enzymes identified from experiment.

Plasmodium falciparum relies on haemoglobin uptake and digestion as a main source of amino acids for its growth and development. Stage-specific proteomic analysis revealed haemoglobin metabolism is a major player in trophozoite stage parasites (Florens et al., 2002). However, human adult haemoglobin lacks one essential amino acid, isoleucine, which parasites needs to obtain directly from the plasma or culture medium (Babbitt et al., 2012, Liu et al., 2006). There is evidence that parasites cultured in isoleucine-free medium enter a hibernatory state, but are able to resume normal growth after re-supplement of isoleucine in the culture medium. Another study also suggested that inhibition of the enzymes falcipain 2, plasmepsin I and plasmepsin IV in haemoglobin metabolism is lethal to the parasite (Liu et al., 2006). This reflects the important role of haemoglobin metabolism on parasite survival. The parasite also exhibits redundant pathways for haemoglobin digestion to ensure amino acid supply. Falcipain 2 and plasmepsins pathways were reported to be redundant. Parasites with either falcipain 2

or plasmepsins knockout survive in isoleucine-supplement culture medium, while falcipain 2 knockout parasite treated with potent plasmepsins inhibitor pepstatin A was killed (Babbitt et al., 2012). Furthermore, the haemoglobin digestion is essential for activation of artemisinin. This was demonstrated by inhibitions of the haemoglobinase enzymes or knockout of falcipain 2 led to significant reduction of artemisinin activity in the trophozoite stage parasites (Klonis et al., 2011).

Plasmepsins are aspartic proteases found in the parasite food vacuole and are involved in haemoglobin catabolism. Two major types of plasmepsins in *Plasmodium falciparum*, plasmepsin I and plasmepsin II, are believed to be crucial for early processing of haemoglobin in order for subsequent proteolysis to occur. Both enzymes can cleave the bond between Phe33-Leu34 in each monomer of dimeric haemoglobin, but not in native tetrameric haemoglobin. Both enzymes are synthesised as proenzymes, thus needing activation through removal of the pro-domain in the food vacuole via an as yet unknown mechanism (Coombs et al., 2001).

In general, aspartic proteases' catalytic mechanism involves an activated water molecule in its active site and does not form the covalent intermediate during cleavage, unlike serine or cysteine proteases that require specific amino acids as attacking nucleophiles. Therefore, aspartic protease inhibitors usually contain hydroxyl moieties to bind in place of catalytic water molecules (Coombs et al., 2001). A commonly known aspartic protease inhibitor, pepstatin, also contain hydroxyl moieties (Figure 4.4).

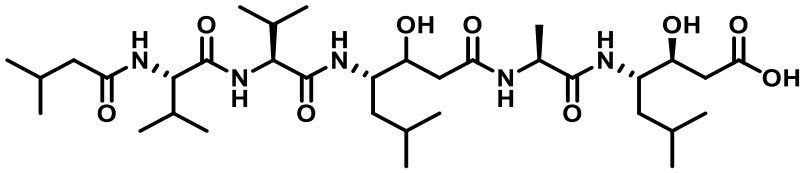


Figure 4.4 Common aspartic proteases inhibitor pepstatin

According to O'Neill and colleagues (2010), activated artemisinin also contains hydroxyl moiety (Figure 4.5) which might interfere with the catalytic site of plasmepsins.

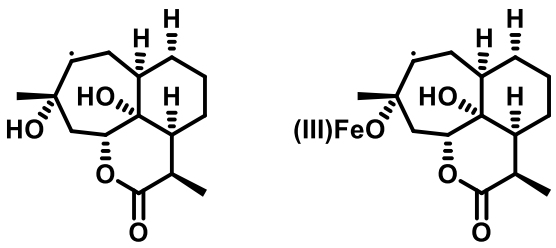


Figure 4.5 Secondary carbon centred radicals of activated artemisinin

To further investigate the binding potential of artemisinin and carbon-centred radical artemisinin to plasmepsin II, computational docking modelling was performed using AutoDock Vina plug-in in Chimera software suite. The docking model of artemisinin and its carbon-centred radical revealed that both compounds fit in the same pocket as the known inhibitor pepstatin A (Figure 4.6).

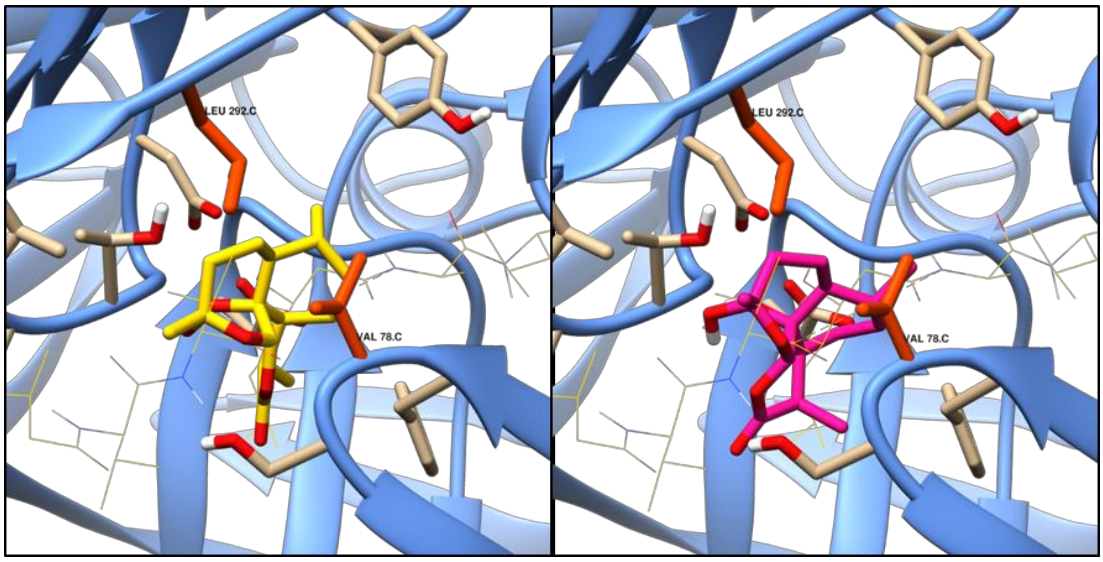


Figure 4.6 Computational docking model of artemisinin (yellow, left panel) and carbon centred radical artemisinin (pink, right panel) with plasmepsin II (blue ribbon), Cys292 and Val78 residues forming close pocket are shown in orange. Pepstatin A is shown in yellow wire. The best fit pose was shown, binding energy of -8.4 and -10.3 kcal/mol for artemisinin and carbon centred radical artemisinin, respectively.

The structural and activity studies of plasmepsin II and its synthetic inhibitors show that relatively smaller compounds have a greater effect on inhibition of parasite growth, although they exhibit similar potency in enzyme inhibition. Another feature of smaller compounds is they favour parasite plasmepsin II to its human analogue, cathepsin D (Silva et al., 1996). Artemisinin is much smaller than pepstatin A. Together with the docking model, these findings suggest artemisinin's potential inhibitory effect on plasmepsin II.

A metabolomic study by Cobbold *et al*. (2016) revealed that levels of haemoglobin-derived peptides prolyl glutamate (PE), prolyl-aspartate (PD), prolyl-glutamyl-glutamate (PEE), and aspartylleucyl-histidine (DLH) suddenly reduce following DHA treatment, suggesting haemoglobin catabolic process had been perturbed by DHA (Figure 4.7). The fact that haemoglobin catabolism is increased in accordance with parasite development, together with effects of artemisinins on other biological processes, might explain why artemisinins are more potent in later stage than in ring stage parasite.

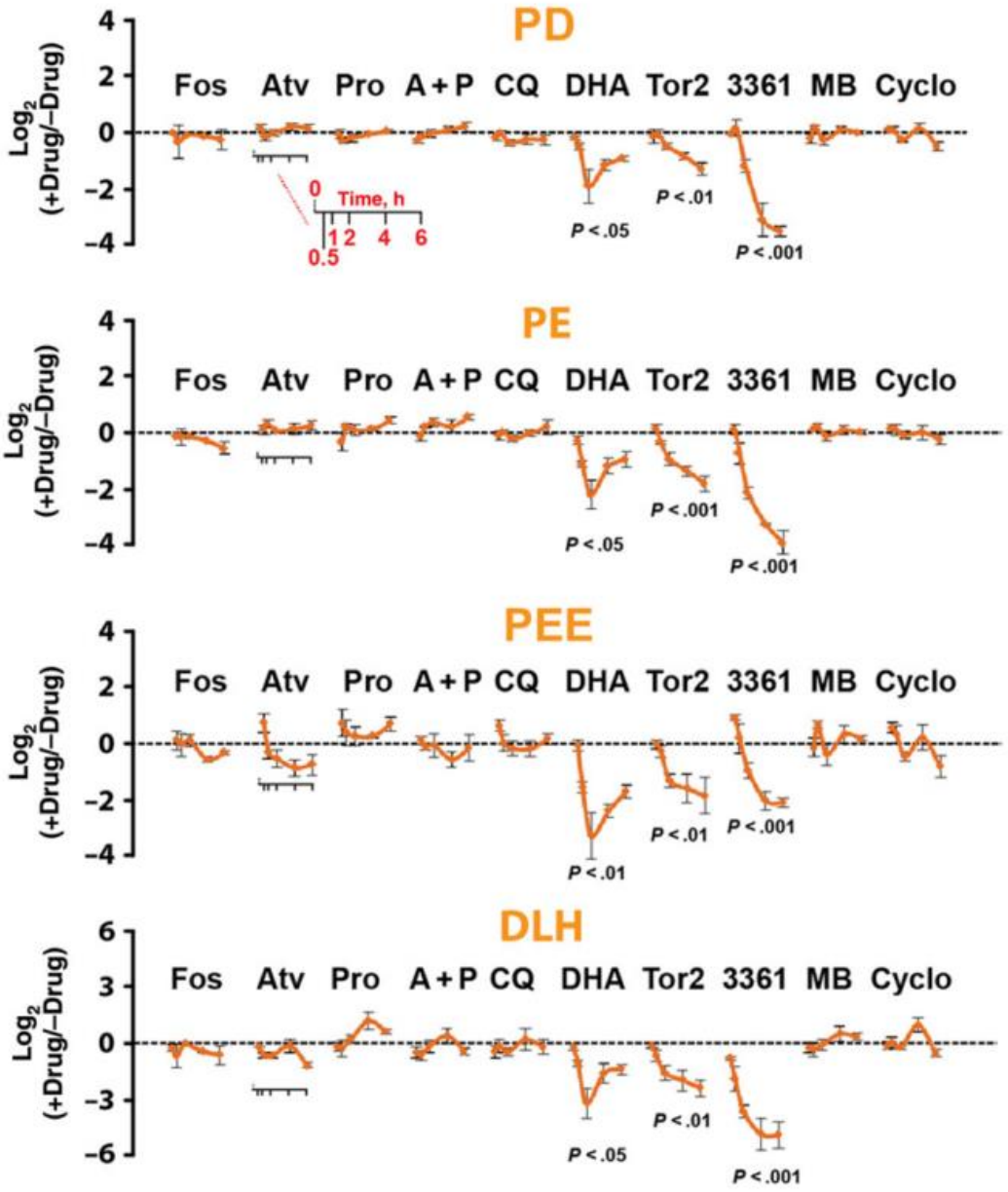


Figure 4.7 Perturbations to haemoglobin-derived peptides after antimalarial drugs treatment (Permission licence number 3852701483027; DOI: 10.1093/infdis/jiv372).

Dipeptidyl-aminopeptidase 1 (DPAP1) is a chloride-activated enzyme catalysing amide bond hydrolysis at acidic pH. There is evidence that DPAP1 fulfils the function of endopeptidase and aminopeptidase in the food vacuole, increasing amino acids production (Wang et al., 2011). Another work by Klemba et al. (2004) suggested DPAP1 is essential for asexual stage parasite proliferation as it is impossible to be genetically deleted (by single crossover). Targeting this enzyme could be detrimental to parasite survival and proliferation.

Two other enzymes identified by artemisinin-activity based probes, **aminopeptidase P (APP)** and **M1-family alanyl aminopeptidase (M1AAP)**, are involved in hydrolysis of dipeptides to amino acids. It was believed that haemoglobin degradation in the food vacuole is ceased at dipeptide level, with dipeptides then transported to the cytosol for further digestion to produce amino acids (Gavigan et al., 2001, Kolakovich et al., 1997). However, it was reported that two out of four parasite aminopeptidases are localised and enzymatically active in food vacuole (Dalal and Klemba, 2007). Both vacuolar aminopeptidases were detected by artemisinin-activity based probes, suggesting these enzymes are plausible artemisinin targets.

It has been observed that haemozoin content was markedly, even if not significantly, reduced after artemisinin treatment as compared to desoxyartemisinin treatment (data not shown). It is still controversial that artemisinins inhibit haem crystallisation process. To the best of our knowledge it is unlikely that artemisinins interfere with haemozoin production in parasites at concentrations which don't affect proliferation, as reflected by ^{3}H-hypoxanthine incorporation (Asawamahasakda et al., 1994) (Figure 4.8). It is possible that reduction of haemozoin production is due to inhibition of the haemoglobin catabolic process, which in turn reduces haem and haemozoin production. Unlike artemisinins, cysteine protease inhibitor E64 exhibits inhibitory effect on haemozoin production but not on ^{3}H-hypoxanthine incorporation (Figure 4.8) (Asawamahasakda et al., 1994).

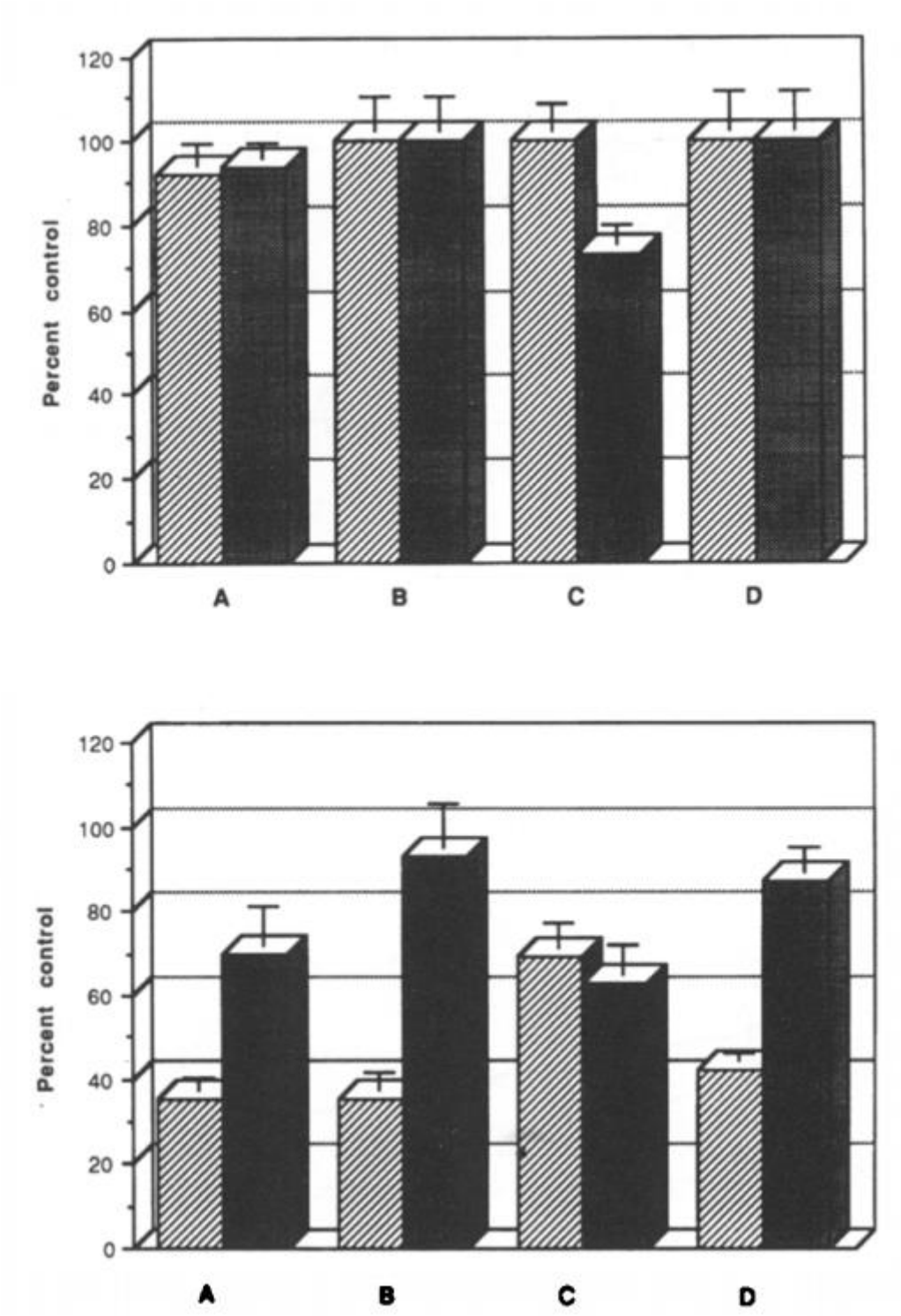


Figure 4.8 Inhibition of hypoxanthine uptake (hatched bars) and haemozoin production (solid bars) in infected RBCs carrying rings (top) or trophozoites (bottom) by 0.5 μM CQ (A), 0.5 μM ART (B), 50 μM E64 (C), and 50 μM pepstatin A (D). (Permission license number 3961370378723).

Although the parasite haemoglobin catabolic process involves redundancy to ensure stable and interruptible amino acid supply for survival and proliferation, artemisinin potentially inhibits 6 crucially important enzymes in the pathway (Figure 4.3). Sensitivity to artemisinin is commensurate with activity

of the haemoglobin catabolic process in asexual stage parasite. As there is no clear evidence that artemisinins inhibit or target haemozoin production, it is presumed that reduction in haemozoin content is due to reduction in haemoglobin degradation and reduced metabolic processes due to parasite death.

4.4.2 Glycolysis pathway

Glycolysis is an essential pathway for cells to maintain their normal functions and viability and is a major source of energy for *Plasmodium* parasites. Proteins involved in the glycolysis pathway identified from the experiment were hexose transporter (HT), hexokinase (HKX), enolase, phosphoglycerate kinase, pyruvate kinase (PryK), and L-lactate dehydrogenase (LDH) (Figure 4.9).

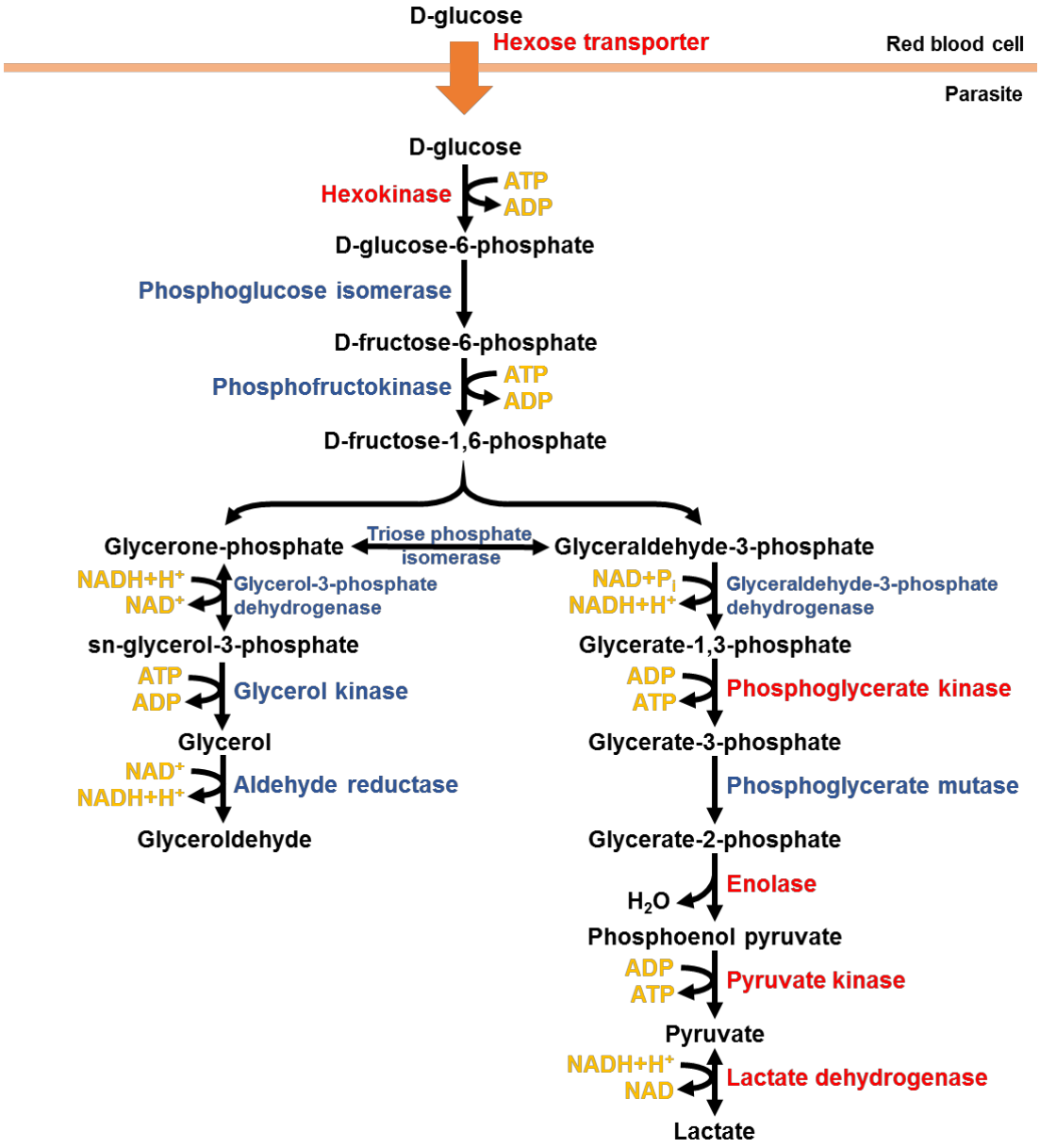


Figure 4.9 Plasmodium falciparum glycolysis pathway (adapted from MPMP). Proteins in red were labelled by artemisinin activity-based probes.

Work by Cobbold and colleagues (2016) shows that DHA had no effect on glycolysis metabolites, glucose-6-phosphate, PEP or pyruvate (Figure 4.10). The study used targeted metabolomics approach to study changes in metabolite levels upon treatment with antimalarial drugs. DHA, atovaquone, and chloroquine had no effect on glycolysis metabolite levels, while antimalarial candidate compound 3361, a glucose transport inhibitor, significantly affected metabolite levels, including glucose, glucose-6-phosphate, phosphoenol pyruvate, and pyruvate, reflecting its inhibitory effect on enzymes in the glycolysis pathway (Cobbold et al., 2016). This is plausible because a crucial process in glycolysis in *Plasmodium falciparum* parasite is glucose transport. Glucose transport has to be inhibited to result in 50% glycolytic flux (van Niekerk et al., 2016). Although hexose transporter (HT) was labelled by ART-15

probe but not other artemisinin-activity based probes. It is not always necessary that artemisinin binding to any protein or transporter would inhibit the function of the protein or not up to the level that can interfere glycolysis pathway. Also glycolysis is a crucial metabolic process in the parasite and could be highly self-protective as no major antimalarial drug, apart from 3361, has an effect on metabolites level (Cobbold et al., 2016) (Figure 4.10). However, the concentration of DHA used by Cobbold et al. (2016) was 40 nM compared to 1,000 nM used by this study. At 40 nM, the inhibitory effect of artemisinin on glucose transport might be minimal and could not be detected by the setting. Further experiment is required to investigate this in a more relevant concentration of artemisinin.

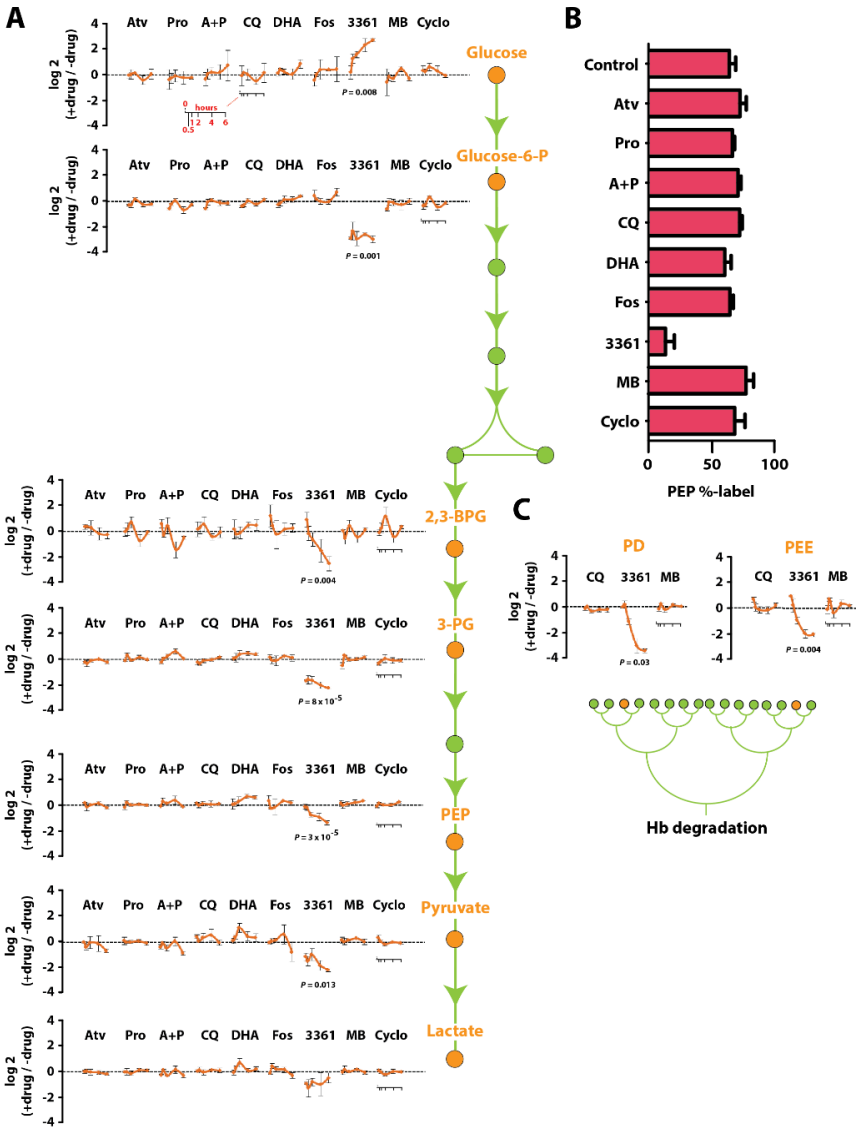


Figure 4.10 Metabolites in glycolysis pathway affected by antimalarial drugs (Permission licence number 3852701483027; DOI: 10.1093/infdis/jiv372)

4.4.3 Nucleic acid and protein biosynthesis pathway

Proteins involved in the translation process identified by artemisinin activity-based probes are listed in Table 4.3. Spermidine synthase, ornithine aminotransferase, hypoxanthine-guanine phosphoribosyltransferase (HGPRT), adenylate kinase (AK1), elongation factor 2 (eEF2), and initiation factor 4A (eIF4a) were also identified from the experiments.

Table 4.3 Protein identified by artemisinin activity-based probes involved in translation process (GO:translation)

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_0309600	O00806	60S acidic ribosomal protein P2	Very high
PF3D7_1421200	Q8ILN8	40S ribosomal protein S25	High
PF3D7_1130200	Q8II61	60S ribosomal protein P0	High
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2	High
PF3D7_0706500	C0H4L6	conserved Plasmodium protein, unknown function	High
PF3D7_0422400	Q8IFP2	40S ribosomal protein S19	High
PF3D7_0316800	O77395	40S ribosomal protein S15A, putative	High
PF3D7_1465900	Q8IKH8	40S ribosomal protein S3	Moderate
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	Moderate
PF3D7_1424100	Q8ILL3	60S ribosomal protein L5, putative	Moderate
PF3D7_1357000	Q8I0P6	elongation factor 1-alpha	Moderate
PF3D7_1242700	Q8I502	40S ribosomal protein S17, putative	Moderate
PF3D7_0813900	Q8IAX5	40S ribosomal protein S16, putative	Moderate
PF3D7_0807900	Q8IAR7	tyrosine--tRNA ligase	Moderate
PF3D7_0517000	Q8I3T8	60S ribosomal protein L12, putative	Moderate
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2	Moderate
PF3D7_0307100	O97249	40S ribosomal protein S12, putative	Moderate
PF3D7_1451100	Q8IKW5	elongation factor 2	Low
PF3D7_1424400	Q8ILL2	60S ribosomal protein L7-3, putative	Low
PF3D7_1408600	Q8IM10	40S ribosomal protein S8e, putative	Low
PF3D7_1342000	Q8IDR9	40S ribosomal protein S6	Low
PF3D7_1338200	Q8IDV1	60S ribosomal protein L6-2, putative	Low
PF3D7_1331700	Q8IE10	glutamine--tRNA ligase, putative	Low
PF3D7_1126200	Q8IIA2	40S ribosomal protein S18, putative	Low
PF3D7_1126000	Q8IIA4	threonine--tRNA ligase	Low
PF3D7_1104000	Q8IIW2	phenylalanine--tRNA ligase beta subunit	Low
PF3D7_0918300	Q8I2X0	eukaryotic translation initiation factor 3 subunit F, putative	Low

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_0507100	Q8I431	60S ribosomal protein L4	Low
PF3D7_0415900	C0H4A6	60S ribosomal protein L15, putative	Low
PF3D7_0102900	Q8I2B1	aspartate--tRNA ligase	Low

Protein biosynthesis pathways are another possible target of artemisinins, as artemisinin-activity based probes labelled 21 ribosomal protein subunits at mostly high to low confidence, with only one ribosomal subunit with very high confidence Table 4.3). It was also observed that protein concentrations in paired treatments were different. Protein concentrations in artemisinin activity-based probes treatments were lower that desoxyartemisinin or DMSO treatments (p-value <.005) (Figure 4.2). Morphologic evidence by electron micrographs showed that parasites treated with artemisinin and dihydroartemisinin lost their ribosomes 4 h post-treatment (Maeno et al., 1993). Another study revealed that artemisinin rapidly reduced incorporation of radioactive labelled isoleucine to newly synthesised protein at concentrations as low as 50 nM, without reduction of ^{3}H-hypoxanthine incorporation (Gu et al., 1983). However, a study performed in cell-free system suggested artemisinin and other antimalarial drugs have no effect on the protein elongation machinery (Ferreras et al., 2002), but does not exclude the possibility of inhibition of translation initiation processes. These evidence support the hypothesis that artemisinins interfere with parasite protein biosynthesis and strengthen the results obtained in the study reported here.

Nonetheless, it is also possible that artemisinin binds to ribosome subunits non-specifically and could not exhibit any inhibitory effect as it lacks specificity to ribosome subunits, none of ribosomal proteins were identified by all probes, and number of ribosome in the parasite could overcome the effect of promiscuous binding of artemisinin. However, it is too early to draw any conclusion here on protein translation inhibition by artemisinin as more experiments are needed to confirm this, but this proteomic approach has provided some insight.

Ornithine aminotransferase (*Pf*OAT) is a protein of particular interest. It catalyses a very narrow process in the amino acid metabolic pathway and is also involved in redox mechanisms. The enzyme catalyses the conversion of L-glutamate 5-semialdehyde to ornithine and is regulated by thioredoxin (Trx) via a unique Cys163 in the structural loop (residue 147-170) (Jortzik et al., 2010). Two cysteine residues, Cys154 and Cys163, are unique to *Plasmodium falciparum*, although there is conservation in the sequences. However, this structural loop is not defined by electron density map, so it does not appear in crystal structures (3LG0 and 3NTJ) suggesting it has a flexible structure (Jortzik et al., 2010).

To further explore the finding that artemisinin probes bind to *Pf*OAT, computational docking has been performed. The model predicted artemisinin fits in the pocket near the structural loop containing the

unique cysteine residues with binding energy of -7.9 kcal/mol. The model suggested possible competitive binding with the OAT regulator thioredoxin (Trx) and could interfere with its activity (section A2.2.3). Taken together, this phenomenon could affect protein function. However, further experiments are needed to investigate if artemisinin inhibits protein function in the physiological environment, presumably by binding cysteine 163 in the structural loop. Site-directed mutagenesis of the active site residues could give some insight into this protein function and artemisinins' effect.

Spermidine synthase is an enzyme involved in methionine and polyamine metabolism, catalysing propylamine transfer from S-adenosylmethioninamine to putrescine, and was identified with high confidence by artemisinin activity-based probes. The enzyme also produces spermidine which is important for activation of translation initiation factor 5A (eIF5A). Inhibition of this enzyme affects parasite development by arresting parasite at early trophozoite stage (Becker et al., 2010). Most recently, Sprenger et al (2016) had demonstrated that inhibition of spermidine synthase is a hierarchical process where the substrate binding site (dcAdoMet) must be occupied first, after which the product binding site (putrescine) can be occupied. Several known inhibitors of spermidine synthase have been biochemically characterised. Although artemisinin was not included in the work by Sprenger et al. (2016), chemical perspective of compound structure suggested artemisinin is more similar to spermidine synthase inhibitors AdoDATA and NAC used in their study. Both compounds are also the most potent spermidine synthase inhibitors tested (Sprenger et al., 2016).

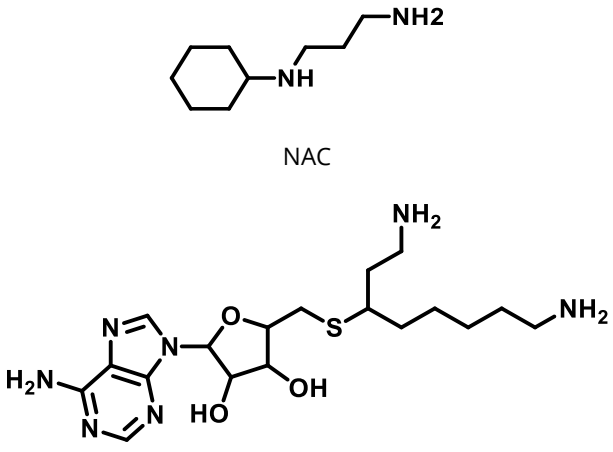


Figure 4.11 Plasmodium falciparum spermidine synthase inhibitors, NAC (top) and AdoDATO (bottom).

Another important function of spermidine synthase lies in its product spermidine, activator of translation initiation factor 5A (eIF5A), which was not identified by artemisinin activity-based probes. The unique production of amino acid hypusine of Lys residue in eIF5A requires aminobutyl moiety from spermidine to activate the eIF5A (Molitor et al., 2004). If artemisinin does inhibit spermidine synthase,

this possibly synergises the effect of artemisinin on eIF4A (identified with high confidence) in translation initiation process of protein synthesis in the parasite.

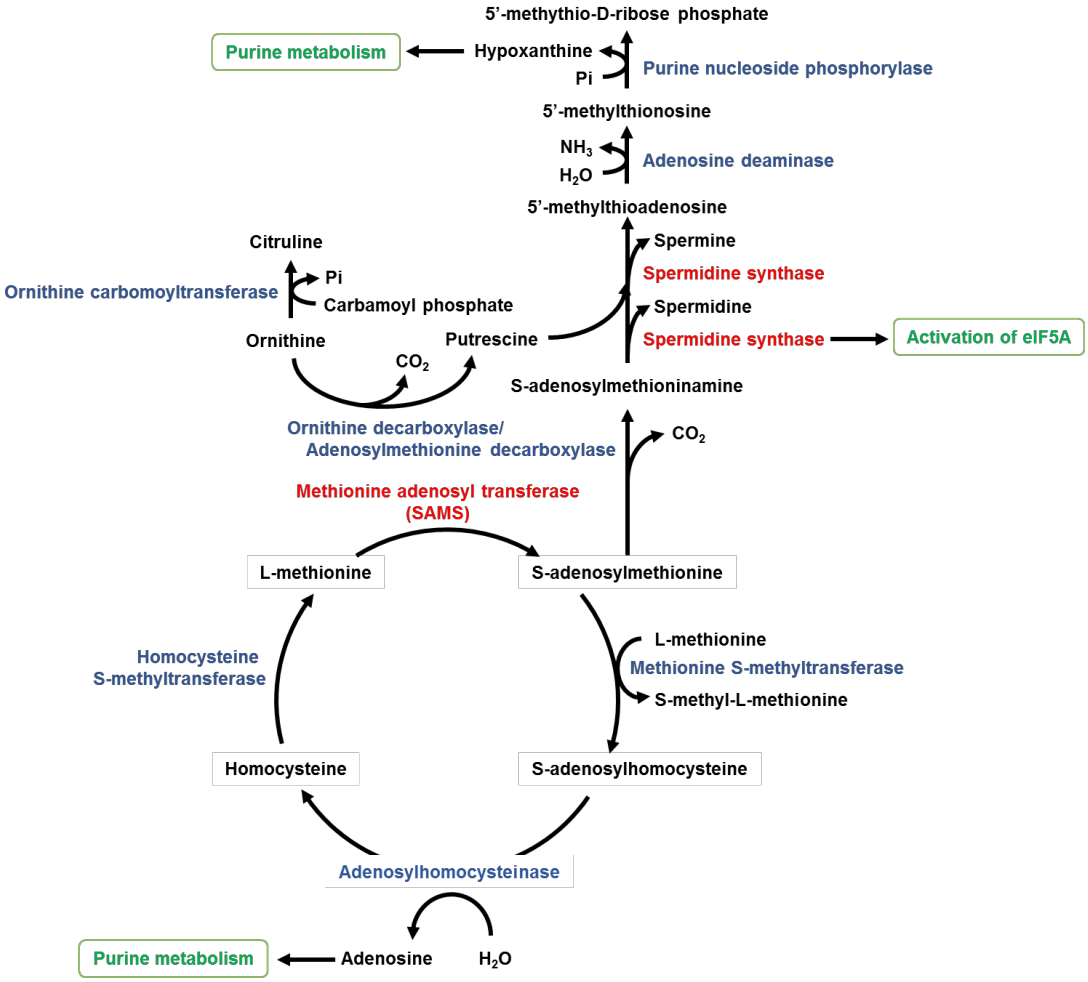


Figure 4.12 Methionine and polyamine metabolism (adapted from MPMP). Detected enzyme is indicated in red.

The evidence from metabolic study suggested that DHA affects pyrimidine biosynthesis as reflected by a significant reduction in the pathway metabolites including aspartate, carbamoyl aspartate, orotate, and UMP. Atovaquone also has an effect on pyrimidine biosynthesis. However, it is likely that artemisinins and atovaquone might have different mechanisms of action as they have different metabolic profile changes. Atovaquone treatment resulted in accumulation of dihydroorotate and carbamoyl-L-aspartate suggesting inhibition of dihydroorotate dehydrogenase enzyme (Biagini et al., 2012), while DHA resulted in depletion of the metabolites (Cobbold et al., 2016). However, artemisinin activity-based probes were not able to identify any enzyme directly involved in pyrimidine biosynthesis pathway, suggesting artemisinin might have indirect effects on pyrimidine pathway. On the other hand, enzymes in the purine biosynthesis pathway were specifically labelled by artemisinin activity-based probes.

The purine biosynthesis pathway was enriched by GO enrichment analysis (Figure 4.1). **Adenylate kinase (AK1)** and **hypoxanthine-guanine phosphoribosyltransferase (HGPRT)** are the 2 proteins in the purine metabolism pathway identified by the probes with very high and moderate confidence, respectively.

Adenylate kinase (AK) catalyses reversible conversion of 2 ADP and ATP + AMP. This process is part of energy metabolism in the parasite. Although metabolomics work by Cobbold and co-workers (2016) did not show any significant alteration in purine metabolites levels, adenylate kinase also plays an important role, together with GTP:AMP phosphotransferase, in cellular energy metabolism. It is possible that artemisinin would have an effect on the parasite energy metabolism in the wider aspect through adenylate kinase.

Hypoxanthine-guanine phosphoribosyltransferase (HGPRT or HGXPRT) is an enzyme in purine salvage pathway of the parasite (lacking *de novo* purine synthesis). Therefore, the parasites purine nucleobases rely completely on salvage pathway (Downie et al., 2008). It was shown that hypoxanthine is an essential factor for the parasite development (Asahi et al., 1996) and inhibition of this enzyme by purine analogues was lethal to the parasites (Keough et al., 2006). This finding suggested artemisinin might have an effect on purine biosynthesis pathway.

To evaluate the binding affinity of artemisinin to HGPRT, computational docking modelling was performed by AutoDock Vina software. The best fit pose has the binding energy of -10.3 kcal/mol and fit in the same binding site as a known inhibitor immucillin (S-SerMe-ImmH) (Hazleton et al., 2012) (Figure 4.13 and A2.2.6). However, more experiments are required for further investigation as this enzyme is a promising drug target.

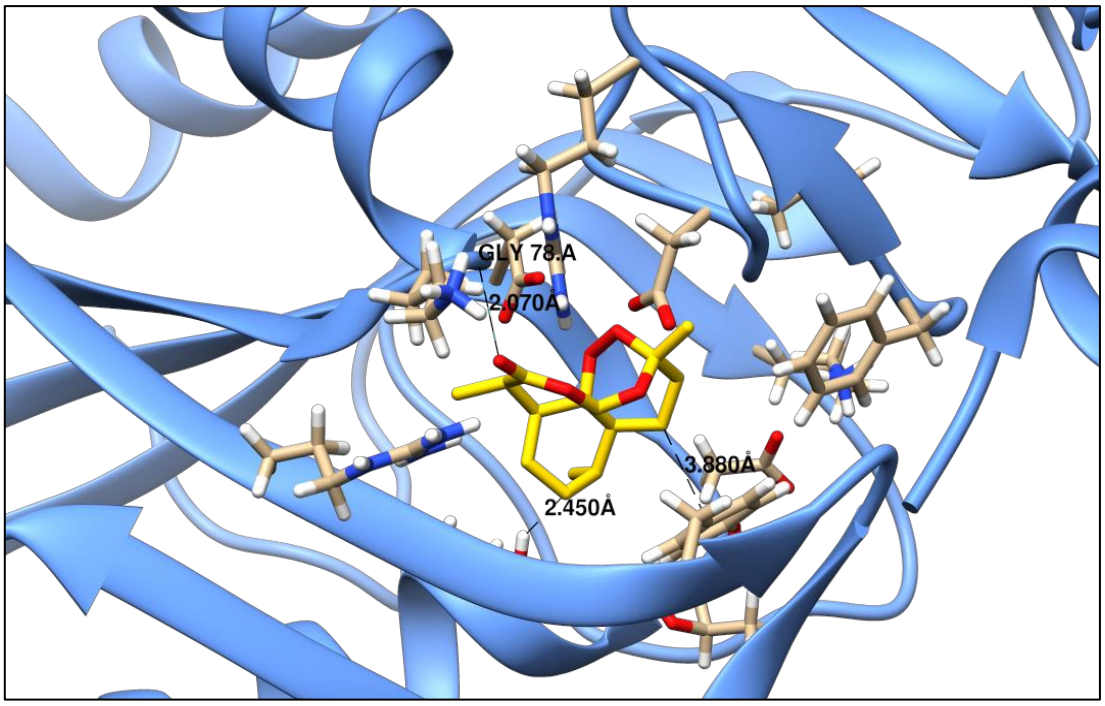


Figure 4.13 Computational molecular docking of PfHGXPRT (blue) and artemisinin (yellow). HGXPRT structure with inhibitor S-SerMe-ImmH (3OZG) was obtained from PDB database. Cocrystallised inhibitor was removed and artemisinin was docked.

4.4.4 Chaperone protein and unfolded protein response

It has long been proposed that artemisinin rapid killing of the parasites are a result of catastrophic effect on parasite proteins. Activated artemisinin or artemisinin carbon-centred radicals alkylate many parasite proteins averting them from normal functions, thus many artemisinin mechanism of action studies have focused on this hypothesis. The results obtained here showing that chaperonins and unfolded protein response proteins are enriched in probe-labelled proteome suggest that parasites possibly overcome this catastrophic effect by boosting its protein recycling process via an unfolded protein response. Transcriptome analysis of parasite treated with artesunate also revealed chaperone and chaperone-related genes are upregulated within a 90-180 min period after treatment (Natalang et al., 2008). These findings are supported by a population transcriptomic study that revealed how the artemisinin resistance phenotype is associated with elevated expression levels of proteins involved in the unfolded protein response in the parasite, including *Plasmodium* Oxidative Stress Complex (PROSC) and TCP-1 Ring Complex (TRiC) (Mok et al., 2015). PROSC proteins are listed in Table 4.4. Artemisinin activity-based probes identified 5 out of 8 PROSC proteins, suggesting artemisinin action is strongly correlated with disruption of protein involved in response to unfolded proteins. However, none of TRiC proteins were identified by the artemisinin activity-based probes. Artemisinin effect on this parasite process might synergise with the effect on other parasite pathways.

Table 4.4 Plasmodium Oxidative Stress Complex (PROSC) proteins (taken from (Mok et al., 2015)). Proteins in bold were identified from the experiment.

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1010700	Q8IJT8	dolichyl-phosphate-mannose protein mannosyltransferase, putative	very high
PF3D7_1222300	Q8I0V4	endoplasmin, putative (GRP94)	high
PF3D7_1115600	Q8IIK8	peptidyl-prolyl cis-trans isomerase (CYP19B)	high
PF3D7_0827900	C0H4Y6	protein disulfide isomerase (PDI8)	moderate
PF3D7_1134100	Q8II23	protein disulfide isomerase (PDI-11)	moderate
PF3D7_1344200	C0H5H0	heat shock protein 110, putative (HSP110)	not detected
PF3D7_1437900	Q8IL88	HSP40, subfamily A, putative (ERdj3)	not detected
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	not detected

The two proteins in PROSC identified by artemisinin activity-based probes were **protein disulfide isomerase 8 (PDI8) and 11 (PDI11)**, named after their chromosomal location, are 2 of 4 disulfide isomerases annotated in *Plasmodium falciparum* genome, localised in the parasite ER and are part of the ER-associated degradation (ERAD) pathway. The structures of PDI8 is highly conserved among *Plasmodium* species. It has 2 thioredoxin domains while PDI11 has 1 thioredoxin domain as predicted by SMART protein database or 2 thioredoxin domains as predicted by the Pfam and HMMs databases (Mahajan et al., 2006). These proteins have oxidoreductase/isomerase and chaperone activity, assisting protein folding via disulfide-dependent conformation (Mahajan et al., 2006, Mouray et al., 2007). *Plasmodium* PDIs have been reported to be targeted, but not solely targeted, by the compound DS61(Mouray et al., 2007). These proteins contain thioredoxin domain(s) suggesting their involvement in response to oxidative stress. Although PDI11 contains HPM motif in the sequence which suggests direct binding to streptavidin, PDI11 was never identified by any control experiment.

Endoplasmin (GRP94) is another protein involved in ERAD pathway dealing with misfolded protein stress in ER. The ERAD pathway in *Plasmodium* parasites and other apicomplexan protozoa are much reduced compared to other eukaryotes and are relatively sensitive to inhibition of the pathway enzymes. Small molecule Inhibition of signal peptide peptidase (SPP), an enzyme in ERAD, has a lethal effect to *Plasmodium falciparum* (Harbut et al., 2012). It has been shown that some compounds directly bind to and inhibit SPP in the parasite. In addition, SPP mutant also renders parasite less sensitive to those compounds by an order of magnitude (Harbut et al., 2012).

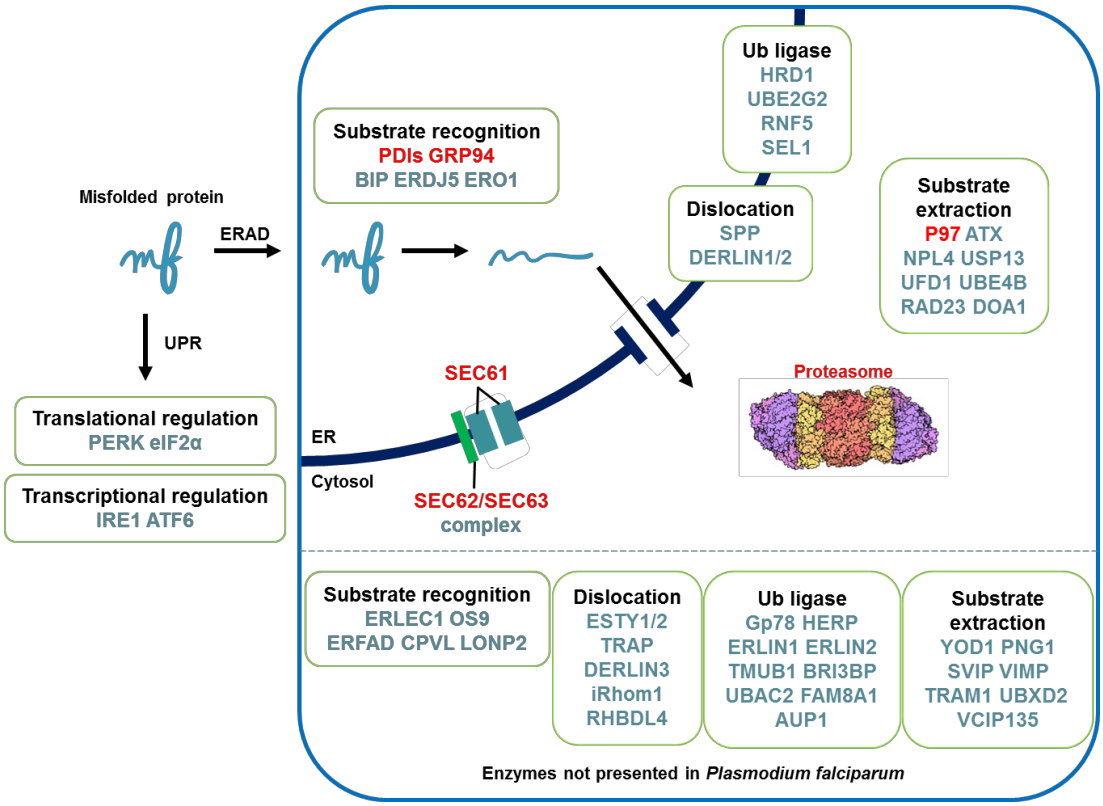


Figure 4.14 Endoplasmic reticulum-assisted degradation (ERAD) pathway in Plasmodium parasites (adapted from Harbut et al. (2012)). Proteins indicated in red were identified by the probes.

System analysis of the chaperone network in *Plasmodium falciparum* by Pavithra et al. (2007) revealed that *Plasmodium* parasites have 97 chaperone proteins (Pavithra et al., 2007). From this total number, artemisinin activity-based probes identified 16 chaperone proteins (16.5%) with various confidence (Table 4.5). This finding supports the hypothesis that artemisinin alkylates parasite proteins and that parasite boosted its response to unfolded proteins (Natalang et al., 2008). Not only that parasite increases response to the unfolded proteins, but the chaperone network itself might be affected by artemisinin.

Table 4.5 Identified proteins associated with chaperone network involved in chaperone network (Pavithra et al., 2007)

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1118200	Q8III6	heat shock protein 90, putative	Very high
PF3D7_1108700	Q8IIR6	heat shock protein DnaJ homologue Pfj2	Very high
PF3D7_0823800	Q8IB72	DnaJ protein, putative	Very high
PF3D7_0629200	C6KTC7	DnaJ protein, putative	Very high
PF3D7_1015600	Q8IJN9	heat shock protein 60	High
PF3D7_0919100	Q8I2W2	DnaJ protein, putative	High
PF3D7_1116800	Q8IIJ8	heat shock protein 101	Moderate
PF3D7_1115600	Q8IIK8	peptidyl-prolyl cis-trans isomerase	Moderate
PF3D7_0917900	Q8I2X4	heat shock protein 70	Moderate
PF3D7_0907400	Q8I377	ATP-dependent protease ATPase subunit ClpY	Moderate
PF3D7_0827900	C0H4Y6	protein disulfide isomerase	Moderate
PF3D7_0708400	Q8IC05	heat shock protein 90	Moderate
PF3D7_0532400	Q8I3F0	lysine-rich membrane-associated PHISTb protein	Moderate
PF3D7_1453700	Q8IKU1	co-chaperone p23	Low
PF3D7_0818900	Q8IB24	heat shock protein 70	Low
PF3D7_0102200	Q8I0U6	ring-infected erythrocyte surface antigen	Low

As discussed above, artemisinin activity-based probes bind, and possibly interfere, with many proteins involved in parasite response to misfolded proteins and unfolded protein, including ERAD and chaperone network, suggesting artemisinin gains it rapid killing effect by interrupting this important process. It has been shown that the unfolded protein response process is crucial to parasite survival, with inhibition of this process leading to apoptosis-like cell death (Rathore et al., 2015).

4.4.5 Ubiquitin-proteasome system

The ubiquitin-proteasome system is responsible for protein degradation in the parasite and thus recycling unwanted proteins. The system comprises of an initial step where ubiquitin is transferred to designated proteins, then ubiquitinated proteins undergo degradation by the proteasome complex. This process is important for parasite survival as it degrades non-functional proteins including damaged proteins. Ubiquitin-proteasome is interlaced with unfolded protein response where unfolded proteins are usually refolded first by chaperone-assisted system or heavily damaged proteins undergone degradation by ubiquitin-proteasome system. Artemisinin activity-based probe SCR-A-17 and SCR-A-10 labelled **polyubiquitin (*Pf*pUB)**, one of the proteins involved in the ubiquitin-proteasome system. Some

proteasome subunits were labelled by artemisinin activity-based probes, usually with moderate and low confidence (Table 4.6).

Table 4.6 Proteins in ubiquitin-proteasome system identified by artemisinin activity-based probes

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1353900	Q8IDG2	proteasome subunit alpha type-7, putative	Moderate
PF3D7_1011400	Q8IJT1	proteasome subunit beta type-5	Moderate
PF3D7_0907400	Q8I377	ATP-dependent protease ATPase subunit ClpY	Moderate
PF3D7_0727400	Q8IBI3	proteasome subunit alpha type-5, putative	Moderate
PF3D7_0108000	Q8I261	proteasome subunit beta type-3, putative	Moderate
PF3D7_0931800	Q8I0U7	proteasome subunit beta type-6, putative	Low
PF3D7_0807500	Q8IAR3	proteasome subunit alpha type-6, putative	Low
PF3D7_1368100	Q8ID28	26S proteasome regulatory subunit RPN11, putative	Low
PF3D7_1017900	Q8IJM0	26S proteasome regulatory subunit p55, putative	Low

Upon exposure to artemisinin, parasite proteins are subjected to stress and ubiquitination leading to activation of the proteasome system. It has been demonstrated that artemisinin-susceptible parasites had higher level of ubiquitinated proteins than in resistant parasite (Dogovski et al., 2015). This finding is suggestive of some possibilities. Firstly, artemisinin generates vast protein damage as evidenced by this study. Secondly, accumulation of ubiquitinated proteins are due to proteasomes were also affected by artemisinin. Lastly, resistant parasites may be able to protect against protein damage or are better able to cope with the stress generated by artemisinin. Proteasome inhibition has a synergistic effect with artemisinin in both artemisinin susceptible and resistant parasites, highlighting the proteasome as a promising therapeutic target (Dogovski et al., 2015).

4.4.6 Transporter proteins

Transporter proteins are important for the parasite to maintain adequate supplies of nutrients, metabolites, and ions by controlling transport across the membranes. Normal function of transporter proteins is required for cellular development and survival. Artemisinin activity-based probes identified many transporter proteins localised on many major membranes including parasite membrane, ER, and food vacuole (Table 4.7). Transporter proteins are also important for retaining normal physiological membrane potential.

Artemisinin, derivatives, and synthetic tetraoxane RKA182 have been demonstrated to cause rapid depolarisation of plasma and mitochondrial membranes. The latter is possibly due to disruption of transporter proteins or other components rather than mitochondrial electron transport chain (ETC) components as direct enzyme assays on ETC components failed to establish any inhibition effect (Antoine et al., 2014). In accordance with previous finding, artemisinin activity-based probes labelled some transporter proteins involved in maintaining membrane potentials, V-type ATPase subunits A (very high), V-type H(+)-translocating pyrophosphatase (high), and V-type ATPase subunits E and V-type ATPase subunits B (moderate). It is possible that artemisinin binding to these transporter proteins alters protein conformation resulting in influx or efflux of ions across membranes and depolarises the membrane potential.

A recent study has claimed that dormant parasite can maintain the mitochondrial membrane potential when exposed to ~700 nM (200 ng/mL) of DHA and are able to recover from dormancy, inferring a resistance phenotype, while parasite whose mitochondrial membrane potential has been abolished failed to recover (Peatey et al., 2015). This is suggestive of membrane potential as a possible therapeutic target and modulator of resistance phenotype. In the present study, **voltage-dependent anion-selective channel protein** was the only protein identified with very high confidence localised on mitochondrial outer membrane with suggestive ion transport function which could contribute to mitochondrial membrane potential generation. Other proteins localised on mitochondrial membrane identified by the artemisinin activity-based probes included the ADP/ATP transporter subunit of adenylate translocase (ADT) (moderate) and GrpE protein homolog (MGE1) (moderate) (Table 4.1 and Table 4.7).

Table 4.7 Protein hits associated with GO term transport

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1438100	Q8IL86	secretory complex protein 62	Very high
PF3D7_1311900	Q76NM6	V-type proton ATPase catalytic subunit A	Very high
PF3D7_0903200	C0H516	ras-related protein RAB7	Very high
PF3D7_0807300	Q7K6B0	ras-related protein Rab-18	Very high
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1	Very high
PF3D7_1432100	Q8ILE3	voltage-dependent anion-selective channel protein, putative	Very high
PF3D7_0106300	Q76NN8	calcium-transporting ATPase	High
PF3D7_0501300	Q8I487	skeleton-binding protein 1	High
PF3D7_1231100	Q8I5A9	ras-related protein Rab-2	High

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4	High
PF3D7_1020900	Q7KQL3	ADP-ribosylation factor	High
PF3D7_1015600	Q8IJN9	heat shock protein 60	High
PF3D7_1451800	Q8IKV8	sortilin	High
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter	High
PF3D7_0914700	Q8I305	major facilitator superfamily-related transporter, putative	High
PF3D7_0601900	C6KSL9	conserved Plasmodium protein, unknown function	High
PF3D7_1456800	Q8IKR1	V-type H(+)-translocating pyrophosphatase, putative	High
PF3D7_1330400	Q8IE22	ER lumen protein retaining receptor 1, putative, unspecified product	High
PF3D7_1346100	Q8IDN6	protein transport protein SEC61 subunit alpha	Moderate
PF3D7_1242800	Q8I501	rab specific GDP dissociation inhibitor	Moderate
PF3D7_1211900	Q8I5T3	non-SERCA-type Ca2+ -transporting P-ATPase	Moderate
PF3D7_1116800	Q8IIJ8	heat shock protein 101	Moderate
PF3D7_0824400	Q8IB78	nucleoside transporter 2	Moderate
PF3D7_1113300	Q8IIM9	UDP-galactose transporter, putative	Moderate
PF3D7_0934500	Q8I2H3	V-type proton ATPase subunit E, putative	Moderate
PF3D7_0512600	Q7K6A8	ras-related protein Rab-1B	Moderate
PF3D7_0416800	Q8I1S0	small GTP-binding protein sar1	Moderate
PF3D7_0406100	Q6ZMA8	V-type proton ATPase subunit B	Moderate
PF3D7_0204700	Q7KWJ5	hexose transporter	Moderate
PF3D7_1037300	Q8IJ34	ADP/ATP transporter on adenylate translocase	Moderate
PF3D7_0509000	Q8I0X0	SNAP protein (soluble N-ethylmaleimide-sensitive factor attachment protein), putative	Moderate
PF3D7_1124700	Q8IIB6	GrpE protein homolog, mitochondrial, putative	Moderate
PF3D7_1464700	Q8IKJ0	ATP synthase (C/AC39) subunit, putative	Low
PF3D7_0524000	Q8I3M5	karyopherin beta	Low
PF3D7_1412300	Q8ILX1	nuclear transport factor 2, putative	Low
PF3D7_1010300	Q8IJU2	succinate dehydrogenase subunit 4, putative	Low
PF3D7_0316600	O77389	formate-nitrite transporter	Low

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_1318800	Q8IEC8	translocation protein SEC63	Low
PF3D7_1246800	Q8I4W4	signal recognition particle receptor, beta subunit	Low
PF3D7_1006700	Q8IJX4	conserved Plasmodium protein, unknown function	Low
PF3D7_0933600	Q8I2I2	mitochondrial-processing peptidase subunit beta, putative	Low
PF3D7_0614300	C6KSY4	major facilitator superfamily-related transporter, putative	Low
PF3D7_0523100	Q8I3N3	mitochondrial-processing peptidase subunit alpha, putative	Low

One protein of particular interest is **sarco/endoplasmic reticulum Ca^{2+} ATPase (SERCA-type ATPase or *Pf*ATP6)**. This protein has previously been proposed as molecular target of artemisinin (Eckstein-Ludwig et al., 2003). The hypothesis is backed by several pieces of computational works (Naik et al., 2011, Shandilya et al., 2013). They reported artemisinin, artemisinin derivatives, and Fe-artemisinin bind to the protein similarly to specific inhibitor thapsigargin and interfere with protein function. A separate experiment on *Xenopus* oocyst with WT and mutant protein sequences showed that mutations in *Pf*ATP6 L263D, L263E, and L263K render the protein insensitive to artemisinin (Uhlemann et al., 2005, Uhlemann et al., 2012). In contrast, another study, showed *Pf*ATP6, rabbit SERCA1a, and E255L SERCA1a (reversed substitution of *Pf*ATP6) expressed on *Xenopus* oocyst showed no sensitivity to artemisinin, even up to 50 µM (David-Bosne et al., 2016). Transgenic parasite expressing WT *Pf*ATP6 or L263E mutant *Pf*ATP6 showed near significant difference in artemisinin and derivatives IC_{50} values (Valderramos et al., 2010). A study by Arnou et al. (2011) has confirmed that *Pf*ATP6 is SERCA but with unique pharmacological profile. They also showed that artemisinin fails to exhibit inhibitory effect on purified *Pf*ATP6 and rabbit SERCA1a as assayed by *in vitro* enzyme activity (Arnou et al., 2011). In relation with bench results, studied from field isolate parasites failed to establish any correlation between mutations in *Pf*ATP6 and reduced sensitivity to artemisinin (Cui et al., 2012, Phompradit et al., 2014, Na-Bangchang et al., 2013, Phompradit et al., 2011) while two other studies also showed parasites with mutations in *Pf*ATP6 remained sensitive to artemisinin (Tanabe et al., 2011, Brasil et al., 2012).

Although many studies argue against *Pf*ATP6 as an artemisinin target, in addition to the computational works, all artemisinin activity-based probes used labelled *Pf*ATP6 both in this study and in a separate

study (Wang et al., 2015). However, how artemisinin affects *Pf*ATP6 has to be further studied and this topic remains a point of debate.

It is worth noting that **endoplasmic reticulum-resident calcium binding protein (ERC)** was also identified by the probes (considered as noise). This protein is closely related to mammalian calumenin, a CREC (**C**ab45, **R**eticulo-calbin, **E**RC-55, and **C**alumenin) protein family. Although CREC has low affinity to calcium, compared to other calcium binding proteins, ERC may play a role in calcium storage and regulation in the parasite ER. Parasite ERC has been proposed as a direct target of some synthetic endoperoxides, N-89 and N-251 (Morita et al., 2012). Chemical proteomic approach showed synthetic endoperoxides have high affinity to ERC in asexual stage of *Plasmodium falciparum* parasite (Morita et al., 2012).

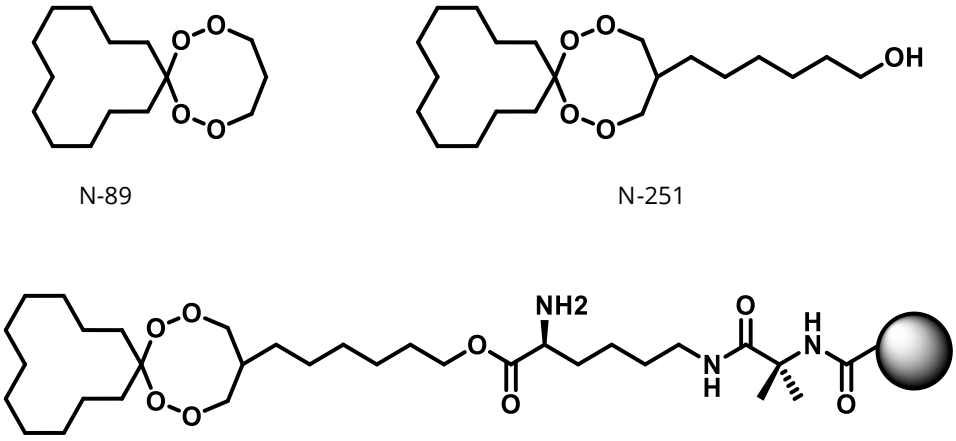


Figure 4.15 Synthetic endoperoxide N-89 (top left), N-251 (top right), and N-346-based column (bottom) (adapted from (Sato et al., 2011) and (Morita et al., 2012)).

Artemisinin might have effects on calcium homeostasis in the parasite as evidenced by ERC and *Pf*ATP6 are proteins involved in calcium homeostasis of the cell were labelled by artemisinin activity-based probes. Presumably artemisinin abolishes cellular calcium homeostasis which is important and fundamental in parasite and other eukaryotes. Cellular calcium served as secondary messenger in signalling pathways and also in gene transcription regulation. In normal physiological condition, cells maintain free cellular calcium in optimum level and stored excess calcium in ER or other organelles. In *Plasmodium* parasites, calcium is also important for parasite egress from RBC (Glushakova et al., 2013) and invasion into RBC (Alleva and Kirk, 2001). However, excessive free calcium in cytoplasm can lead to adverse effect or induce cell death (Ermak and Davies, 2002).

Another transporter protein identified from the experiment was **chloroquine resistance transporter (*Pf*CRT).** It is a transporter protein located on the food vacuole membrane which has been widely studied, as mutations in this gene are associated with chloroquine resistance, hence its name. Mutation in *Pf*CRT confers resistance to chloroquine supposedly by reduced accumulation of chloroquine in food vacuole due to leakage.

Artemisinin and chloroquine are believed to share a similar cellular localisation in parasite food vacuole. Genome wide association study (GWAS) study showed that polymorphisms in *Pf*CRT gene are strongly associated with artemisinin resistance (Miotto et al., 2015). This is suggestive of *Pf*CRT being involved either directly or indirectly with artemisinin's mechanism of resistance. In *Pf*CRT mutant parasite conferring chloroquine resistance, chloroquine is actively transported out of the food vacuole, preventing chloroquine from sequestering haemozoin. This process is believed to be coupled with ion exchange and proton leakage (Lehane and Kirk, 2008, Juge et al., 2015). If artemisinin shares a similar mechanism of resistance conferred by CRT, this process may prevent artemisinin activation by a haem-mediated process (Xie et al., 2016) or its effects on haemoglobin digestion (section 4.4.1).

CRT also transport ions and proton across food vacuole membranes and could contribute the generation of membrane potential across the food vacuole membrane (Juge et al., 2015). Artemisinin binding to CRT may cause membranes to leak. It has been shown that proton leakage is associated with chloroquine resistance (Lehane and Kirk, 2008). Not only drugs and ions that are transported by CRT, glutathione has been reported to be transported by CRT and linked with chloroquine resistance (Patzewitz et al., 2013).

Multidrug resistance protein 1 (MDR1) is another transporter protein identified. The gene encoded for MDR1 protein is associated with susceptibility to multiple antimalarial drugs. Disruption of one copy of *Pf*MDR1 gene renders parasite more sensitive to mefloquine, lumefantrine, halofantrine, quinine, and artemisinin, but not chloroquine (Sidhu et al., 2006).

Although MDR1 and CRT and their association with resistance to many antimalarial drugs were established (Veiga et al., 2016), a clear evidence for artemisinin of action and resistance to these proteins remain elusive and yet to further explored.

4.4.7 Antioxidant defence system

The malaria parasite is subjected to oxidative stress throughout its life cycle. To withstand a highly oxidative environment, the parasite possesses a large and complex antioxidant system (Bozdech and Ginsburg, 2004) (Figure 4.16). The system serves the parasite as a redox switch, redox homeostasis, and antioxidant defence system.

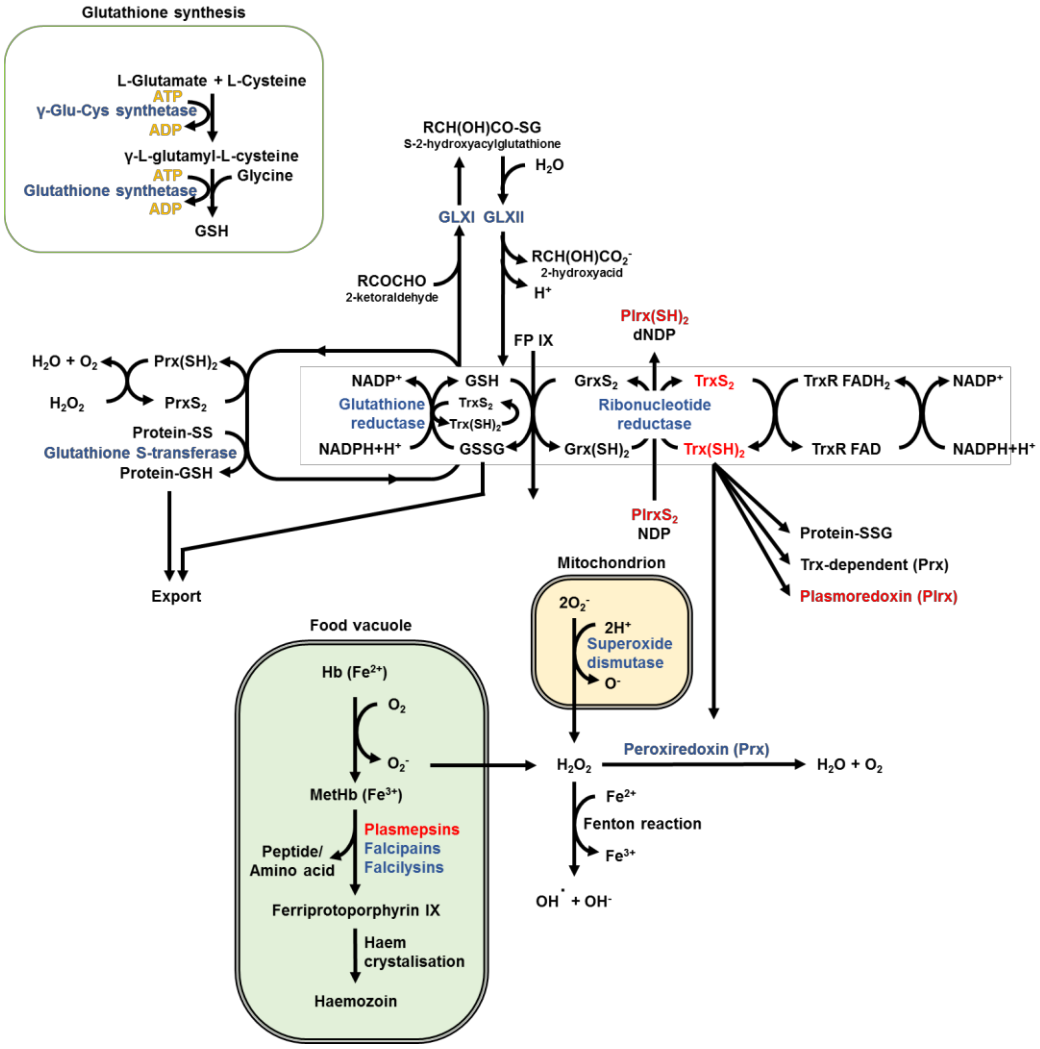


Figure 4.16 Antioxidant system in Plasmodium falciparum. Proteins identified from the experiments indicated in red.

Artemisinin activity-based probes identified a number of proteins in antioxidant systems in *Plasmodium falciparum* (Table 4.8). It has been hypothesised that artemisinin kills the parasite because ROS generated by artemisinin overwhelm the redox balance of these systems. If the defence systems themselves are affected by artemisinin, this could be catastrophic and rapidly kill the parasite, which is a key feature of artemisinin. Furthermore, nearly 50% of proteins identified by artemisinin activity-based probes (very high-moderate confidence) contain glutathione binding site (Kehr et al., 2011). This

implicated artemisinin disturbs antioxidant system in the parasites and that cysteine residue is possibly the main target of alkylation by carbon centred artemisinin radicals.

Table 4.8 Protein hits involved in cell redox homeostasis and antioxidant system (GO term)

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_0823800	Q8IB72	DnaJ protein, putative	very high
PF3D7_1104400	Q8IIV8	conserved protein, unknown function	very high
PF3D7_1108700	Q8IIR6	heat shock protein DnaJ homologue Pfj2	very high
PF3D7_1212000	Q8I5T2	glutathione peroxidase-like thioredoxin peroxidase (TPx(Gl))	high
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	high
PF3D7_0303600	Q8I224	plasmoredoxin (Plrx)	high
PF3D7_0827900	C0H4Y6	protein disulfide isomerase (PDI8)	moderate

4.4.8 Parasite-host interaction and protein export

A number of parasite proteins localised on the RBC membrane and involved in parasite-host interaction were labelled from the experiment with various confidence (Table 4.9 and Figure 4.17).

Table 4.9 Protein hits associated with GO Maurer's cleft, PHIST protein family, and host-parasite interaction

Gene ID	UniProt ID	Product Description	Confidence
PF3D7_0501200	Q8I488	parasite-infected erythrocyte surface protein (PIESP2)	very high
PF3D7_0702400	Q8IC43	small exported membrane protein 1 (SEMP1)	very high
PF3D7_0936000	C0H592	ring-exported protein 2 (REX2)	very high
PF3D7_0936800	Q8I2F2	Plasmodium exported protein (PHISTc), unknown function	very high
PF3D7_0416800	Q8I1S0	small GTP-binding protein sar1 (SAR1)	high
PF3D7_0501300	Q8I487	skeleton-binding protein 1 (SBP1)	high
PF3D7_0601900	C6KSL9	conserved Plasmodium protein, unknown function	high
PF3D7_0424600	Q8IFM0	Plasmodium exported protein (PHISTb), unknown function	high
PF3D7_1201000	Q8I635	Plasmodium exported protein (PHISTb), unknown function	high
PF3D7_0532300	Q8I3F1	Plasmodium exported protein (PHISTb), unknown function	high
PF3D7_0402000	Q8I206	Plasmodium exported protein (PHISTa), unknown function	high
PF3D7_0501000	Q8I490	Plasmodium exported protein, unknown function	moderate
PF3D7_1149000	Q8IHN4	antigen 332, DBL-like protein (Pf332)	moderate

Although there is no established evidence showing that mechanism of artemisinin involves host-parasite interaction, these finding suggested artemisinin could have effects on this important process.

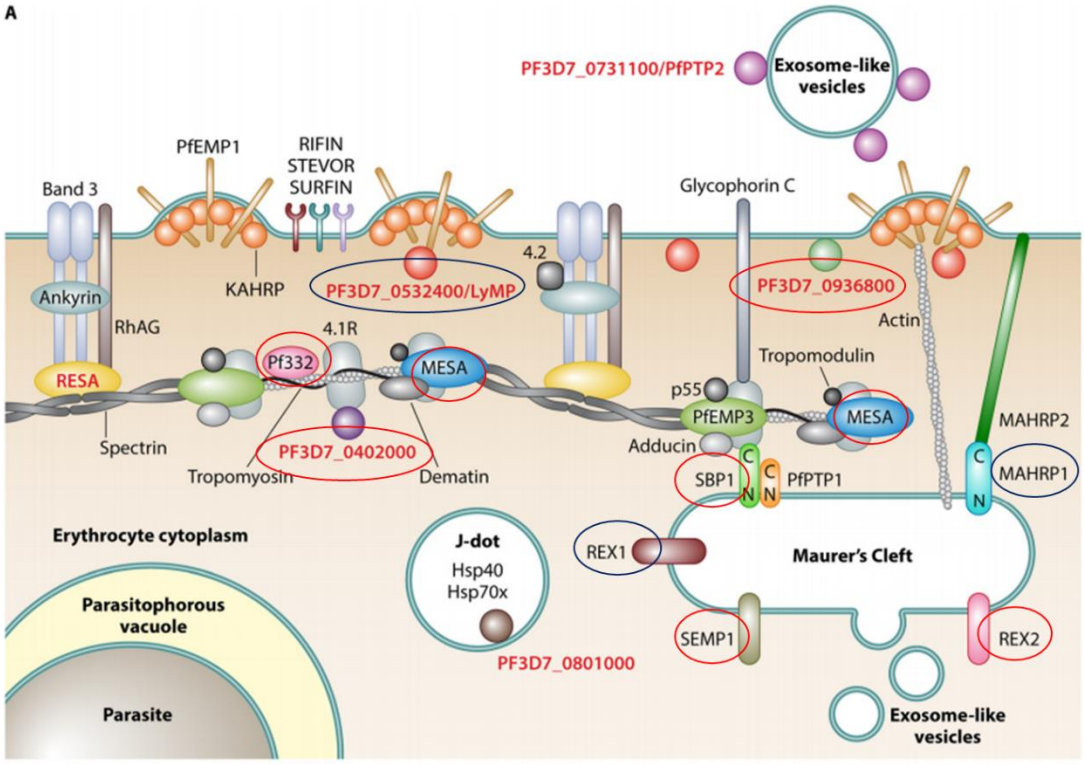


Figure 4.17 Parasite exported proteins (Warncke et al., 2016). Proteins in circles were identified from the experiments.

4.5 Conclusions

Artemisinins are most effective in mature trophozoite stage in *Plasmodium falciparum* as its activator ferrous haem is abundant (Klonis et al., 2013). Proteome of trophozoite parasites treated with artemisinin activity-based probes revealed many proteins, spanning most major parasite metabolic pathways, were irreversibly bound to artemisinin. Possible pathways implicated include haemoglobin metabolism, protein biosynthesis, nucleic acid, transport of metabolites and ions, and membrane potential as previously discussed in this chapter. These findings support the idea that artemisinin's mechanism of action is not through single inhibitory pathway but due to a global effect of parasite protein alkylation causing unfunctional proteins or possible induction of cell death pathway. However, further study is required to elucidate effect(s) of artemisinin on selected pathways or enzymes.

Limitation in the study are including parasite material used for the treatment, low coverage of lysine and arginine-rich proteins, contamination from host protein, and high background from copper-free click reaction. In this study a tremendous amount of parasite is required to obtain a substantial amount of protein for click reaction and subsequent purification. Also high background from copper-free click reaction due to highly reactive reagent DIBO, but could be prevented by lowering the concentration and reducing the proteins before click reaction. Although saponin lysis was introduced to purify the free parasites and to reduce the host haemoglobin contamination, the detectable level of host haemoglobin is inevitable and could cause ion suppression during peptide sequencing step by MS. As peptide preparation involved trypsin digestion, proteins with high lysine and arginine could undergo extensive digestion and not detectable by MS.

Chapter 5

The "endoperoxome" of the ring stage parasite

5.1 Introduction

Previous studies have suggested that different stages of the *Plasmodium falciparum* intra-erythrocytic life cycle respond differentially to artemisinin exposure (Figure 5.1) with significant implications for the K13 dependent resistance mechanism. The data suggest that the very early ring stage (~0-3 hpi) parasite is hypersensitive to artemisinin whereas mid-late ring stage (~3-18 hpi) parasite exhibits high artemisinin tolerance and the trophozoite stage (~18-40 hpi) parasite see a restoration in parasite susceptibility to artemisinin action (Klonis et al., 2013). In addition, resistance to artemisinin is associated with delayed or prolonged ring stage (3-18 hpi) which has the least sensitivity to artemisinin (Cheng et al., 2012) and is believed to be mediated by K13 mutations (Ariey et al., 2014, Tilley et al., 2016).

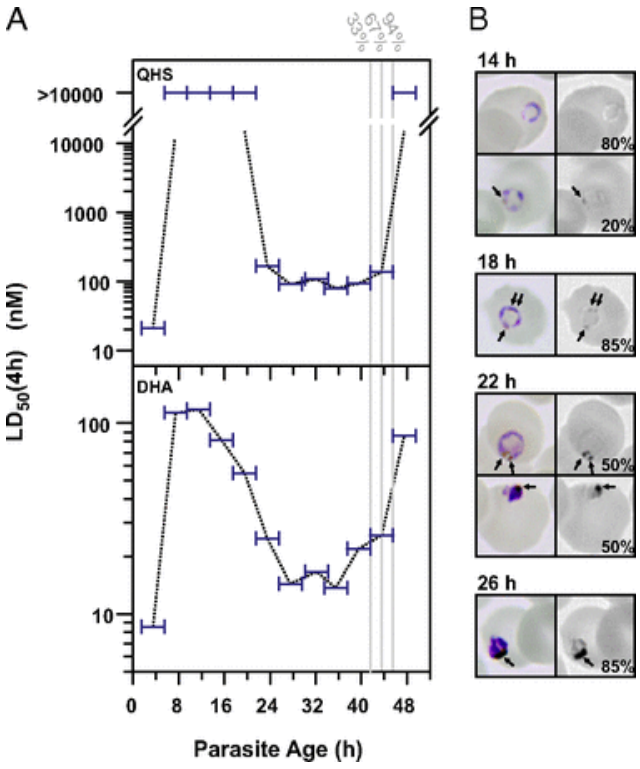


Figure 5.1 Parasite sensitivity fluctuations to short pulses of artemisinin (taken from Klonis et al. (2013). (A) A tightly synchronized 3D7 culture (>90% of parasites within a 1-hour time window) was subjected to 4-hour drug pulses every 4 h. Horizontal bars correspond to the range of ages (of ≥50% of the parasite population) during each assay. The grey vertical lines indicate where a significant fraction (indicated above the lines) progressed to the next cycle. (B) Haemozoin formation during the ring to trophozoite transition. (Left) Colour images of Giemsa-stained parasites from A. (Right) Blue channel of the Giemsa images showing dark puncta that correspond to haemozoin (arrows). The percentages of represented morphologies are indicated (n ≥ 19).

The proposed mechanism of artemisinin resistance mediated by K13 is via the interaction between K13 protein and phosphatydylinositol-3-kinase (PI3K), a proposed target of artemisinin. It has been suggested that the wildtype K13 protein interacts with PI3K and leads to polyubiquitination and degradation of PI3K, while the K13 mutant protein hinders the interaction between the two proteins preventing PI3K degradation and increasing level of phosphatidylinositol-3-phosphate (PI3P). However, the consequential mechanism of this observation is as yet unknown (Mbengue et al., 2015).

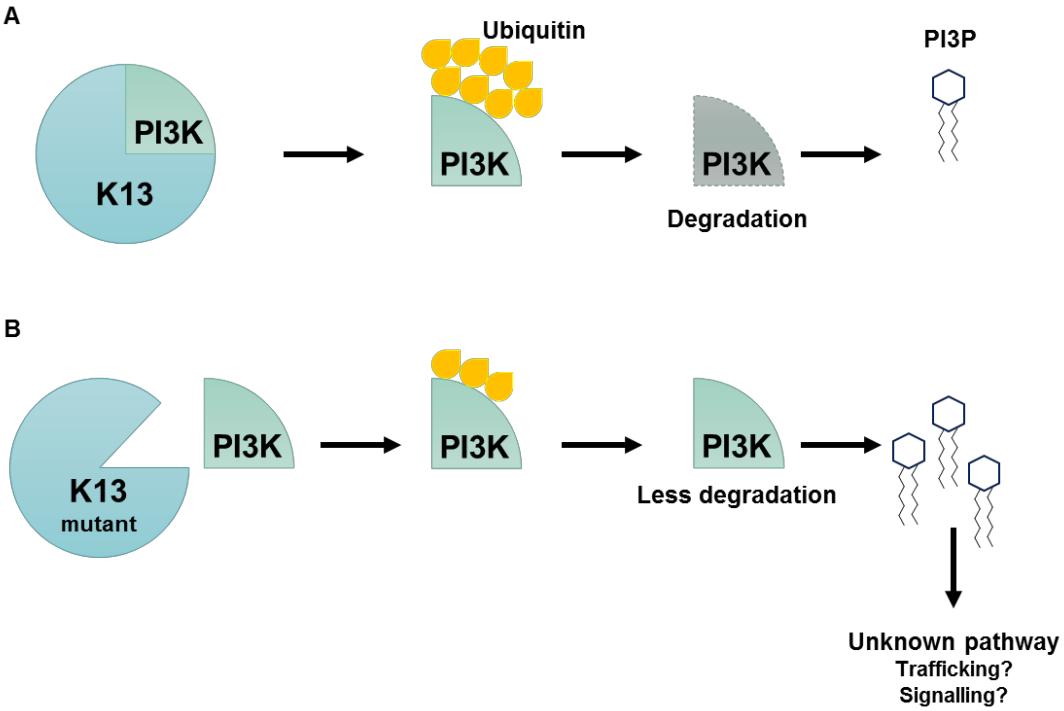


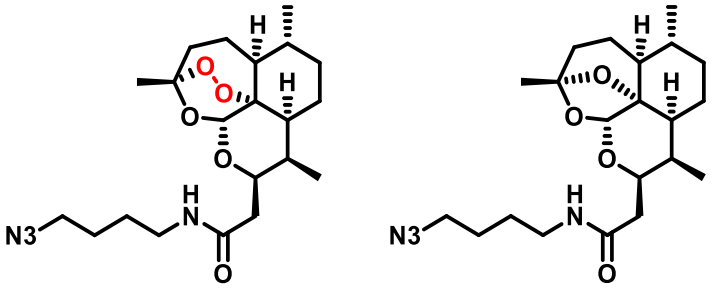
Figure 5.2 Proposed mechanism of artemisinin resistance via K13 mutation and PI3K interaction (adapted from Mbengue et al. (2015)). (A) K13 and PI3K interaction leads to PI3K degradation by ubiquitin-mediated proteolysis resulting in low level of PI3P and artemisinin sensitive phenotype. (B) K13 mutation results in reduced interaction with PI3K, therefore less ubiquitination and degradation of PI3K. However, how increased level of PI3P is associated with resistance phenotype is still unknown.

In Chapter 4 click probes were used to identify proteins that could be tagged by artemisinin. Through inference there is a strong possibility that one or more of these drug protein interaction is important to antimalarial activity so an attempt has been made to relate the targeted proteins to plausible mechanism of action. In the light of the information on differential susceptibility in erythrocytic parasites, analysis of the artemisinin protein labelling profile of ring stage parasites could provide some perspective on what underlies the differential sensitivity to artemisinin during the asexual developmental cycle and could assist in prioritising protein targets of greatest relevance or linkage with the exquisite sensitivity to these drugs.

5.2 Experimental

5.2.1 Chemical probes

In order to study the molecular protein targets of artemisinin in ring stage parasites, azide tagged artemisinin activity-based probes (active SCR-A-10 and control SCR-A - 22) were used to treat the synchronous staged parasites. The probe synthetic schemes were published elsewhere (Ismail et al., 2016a, Ismail et al., 2016b) and briefly described in section 3.2.1.



SCR-A-10 SCR-A-22

Figure 5.3 Probes structures used in ring parasites treatments

5.2.2 Parasite treatment

Plasmodium falciparum strain 3D7 parasite was cultured as described in section 2.13, and synchronised for ring stage (section 2.14). The synchronous parasites were allowed to recover from sorbitol treatment for at least 1 cycle. Probes treatments were performed when ring stage (~6-12 hpi) parasites reached ~10% parasitaemia with 1 µM probes for 6 h at normal parasite culture conditions (section 2.23). In order to obtain substantial amount of protein from ring stage parasite, 15-20 flasks of the parasite were used per condition. After parasite treatment, protein extraction, copper-free click reaction (section 2.26) and protein processing for MS analysis (section 2.27 and 2.28) were performed.

5.3 Results

In this section, ring stage proteins identified by the artemisinin activity-based probe are presented and compered with the results from trophozoite stage parasites (described in detail in Chapters 3 and 4).

5.3.1 Number of protein hits identified from ring and trophozoite stage parasites using the azide click probe

Because experiments with ring stage parasite were only performed with the azide probes, protein hits from ring and trophozoite stages identified by azide artemisinin activity-based probe were first compared. MS results showed that there were 70 proteins identified from ring stage, while probe labelled 176 proteins from trophozoite stage parasites despite equivalent parasite numbers being used in each experiment. Among these, only 10 of the 70 proteins from the ring stage were found in trophozoite probe-labelled proteome.

When considering all 255 protein hits identified from trophozoite stage by all probes, 38 of the 70 proteins ring stage proteins overlapped, resulting in 32 proteins uniquely identified from (drug insensitive) ring stage parasite using this azide probe (Table 5.1).

Confidence of the proteins identified from ring stage parasites was differently defined from the trophozoite stage proteome because only one type of probe was used. Protein confidence in ring stages was defined by the frequency proteins were identified from repeated experiments. Any protein that was identified from 4 experiments (100%), 3 experiments, and 2 was considered high, moderate, and low confidence, respectively. Proteins only identified by one experiment were considered background and they mostly overlapped with the background proteome from desoxyartemisinin treatment (Table A1.5). This result again highlighted the advantage of using the pairwise strategy (active probe versus inactive control probe) as described in Chapter 3.

Table 5.1 Protein hits identified from ring stage parasite by azide artemisinin activity-based probe

Gene ID	UniProt ID	Product Description	Confidence	Note
PF3D7_0317600	O77381	40S ribosomal protein S11, putative (RPS11)	high	Unique in ring
PF3D7_0517000	Q8I3T8	60S ribosomal protein L12, putative	high	Unique in ring
PF3D7_0618300	C6KT23	60S ribosomal protein L27a, putative	high	Unique in ring
PF3D7_0619400	C6KT34	cell division cycle protein 48 homologue, putative	high	Unique in ring
PF3D7_0903700	Q6ZLZ9	alpha tubulin 1	high	Unique in ring
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	high	Unique in ring
PF3D7_0922500	P27362	phosphoglycerate kinase (PGK)	high	Unique in ring
PF3D7_1008700	Q7KQL5	tubulin beta chain	high	Unique in ring
PF3D7_1019400	Q8IJK8	60S ribosomal protein L30e, putative	high	Unique in ring
PF3D7_1310700	Q8IEK9	RNA-binding protein, putative	high	Unique in ring
PF3D7_1318800	Q8IEC8	translocation protein SEC63 (SEC63)	high	Unique in ring
PF3D7_1351400	Q8IDI5	60S ribosomal protein L17, putative	high	Unique in ring
PF3D7_1426000	Q8ILK3	60S ribosomal protein L21 (RPL21)	high	Unique in ring
PF3D7_0525800	Q8I3K7	inner membrane complex protein 1g, putative (IMC1g)	Moderate	Unique in ring
PF3D7_0624000	C6KT76	hexokinase (HK)	Moderate	Unique in ring
PF3D7_1003500	Q8IK02	40S ribosomal protein S20e, putative	Moderate	Unique in ring
PF3D7_1011800	Q8IJS7	PRE-binding protein (PREBP)	Moderate	Unique in ring
PF3D7_1015900	Q8IJN7	enolase (ENO)	Moderate	Unique in ring
PF3D7_1105400	Q8IIU8	40S ribosomal protein S4, putative	Moderate	Unique in ring
PF3D7_1126200	Q8IIA2	40S ribosomal protein S18, putative	Moderate	Unique in ring
PF3D7_1306400	Q8IEQ1	26S protease regulatory subunit 10B, putative (RPT4)	Moderate	Unique in ring

Gene ID	UniProt ID	Product Description	Confidence	Note
PF3D7_1317800	C0H5C2	40S ribosomal protein S19 (RPS19)	Moderate	Unique in ring
PF3D7_1338300	Q8IDV0	elongation factor 1-gamma, putative	Moderate	Unique in ring
PF3D7_1358800	Q8IDB0	40S ribosomal protein S15 (RPS15)	Moderate	Unique in ring
PF3D7_1414300	Q8ILV2	60S ribosomal protein L10, putative	Moderate	Unique in ring
PF3D7_1438900	Q8IL80	thioredoxin peroxidase 1 (Trx-Px1)	Moderate	Unique in ring
PF3D7_0719600	Q8IBQ6	60S ribosomal protein L11a, putative	low	Unique in ring
PF3D7_1142500	Q8IHU0	60S ribosomal protein L28 (RPL28)	low	Unique in ring
PF3D7_1309100	Q8IEM3	60S ribosomal protein L24, putative	low	Unique in ring
PF3D7_0922200	Q7K6A4	S-adenosylmethionine synthetase (SAMS)	high	Identified from trophozoite
PF3D7_0106300	Q76NN8	calcium-transporting ATPase (ATP6)	high	Identified from trophozoite
PF3D7_0415900	C0H4A6	60S ribosomal protein L15, putative	high	Identified from trophozoite
PF3D7_0422400	Q8IFP2	40S ribosomal protein S19 (RPS19)	high	Identified from trophozoite
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1 (MDR1)	high	Identified from trophozoite
PF3D7_0721600	Q8IBN5	40S ribosomal protein S5, putative	high	Identified from trophozoite
PF3D7_0813900	Q8IAX5	40S ribosomal protein S16, putative	high	Identified from trophozoite
PF3D7_0814200	Q8IAX8	DNA/RNA-binding protein Alba 1 (ALBA1)	high	Identified from trophozoite
PF3D7_0827900	C0H4Y6	protein disulfide isomerase (PDI8)	high	Identified from trophozoite
PF3D7_0929400	C0H571	high molecular weight rhoptry protein 2 (RhopH2)	high	Identified from trophozoite
PF3D7_1011400	Q8IJT1	proteasome subunit beta type-5	high	Identified from trophozoite
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase (HGPRT)	high	Identified from trophozoite
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2 (RPS2)	high	Identified from trophozoite
PF3D7_1117300	Q8IIJ4	conserved Plasmodium protein, unknown function	high	Identified from trophozoite

Gene ID	UniProt ID	Product Description	Confidence	Note
PF3D7_1129000	Q8II73	spermidine synthase (SpdSyn)	high	Identified from trophozoite
PF3D7_1224300	Q8I5H4	polyadenylate-binding protein, putative (PABP)	high	Identified from trophozoite
PF3D7_1237700	Q8I546	conserved protein, unknown function	high	Identified from trophozoite
PF3D7_1325100	Q8IE67	phosphoribosylpyrophosphate synthetase	high	Identified from trophozoite
PF3D7_1341200	Q8IDS6	60S ribosomal protein L18, putative	high	Identified from trophozoite
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	high	Identified from trophozoite
PF3D7_1408600	Q8IM10	40S ribosomal protein S8e, putative	high	Identified from trophozoite
PF3D7_1421200	Q8ILN8	40S ribosomal protein S25 (RPS25)	high	Identified from trophozoite
PF3D7_1424100	Q8ILL3	60S ribosomal protein L5, putative	high	Identified from trophozoite
PF3D7_1424400	Q8ILL2	60S ribosomal protein L7-3, putative	high	Identified from trophozoite
PF3D7_1468700	Q8IKF0	eukaryotic initiation factor 4A (eIF4A)	high	Identified from trophozoite
PF3D7_0628300	C6KTB9	choline/ethanolaminephosphotransferase, putative (CEPT)	high	Identified from trophozoite
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter (CRT)	high	Identified from trophozoite
PF3D7_0823800	Q8IB72	DnaJ protein, putative	high	Identified from trophozoite
PF3D7_1105800	Q8IIU5	conserved Plasmodium protein, unknown function	high	Identified from trophozoite
PF3D7_1211800	Q7KQK2,Q9U5M1	polyubiquitin (PfpUB)	high	Identified from trophozoite
PF3D7_1459400	Q8IKN7	conserved Plasmodium protein, unknown function	high	Identified from trophozoite
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase (FBPA)	Moderate	Identified from trophozoite
PF3D7_1471100	Q8IKC8	exported protein 2 (EXP2)	Moderate	Identified from trophozoite
PF3D7_1019900	Q8IJK2	autophagy-related protein 8 (ATG8)	moderate	Identified from trophozoite
PF3D7_1143200	Q8IHT4	DnaJ protein, putative	moderate	Identified from trophozoite
PF3D7_0309600	O00806	60S acidic ribosomal protein P2 (PfP2)	low	Identified from trophozoite

Gene ID	UniProt ID	Product Description	Confidence	Note
PF3D7_1346100	Q8IDN6	protein transport protein SEC61 subunit alpha (SEC61)	low	Identified from trophozoite
PF3D7_0207600	Q9TY95	serine repeat antigen 5 (SERA5)	low	Identified from trophozoite
PF3D7_1451800	Q8IKV8	sortilin	low	Identified from trophozoite

5.3.2 Ring and trophozoite stages divergent protein profiles

Not only were the total number of protein hits identified by the azide artemisinin activity-based probe difference between rings and trophozoites the protein profiles displayed minimum overlap. Both proteomes shared only 4.2% overlap or 10 proteins (Figure 5.4). The overlap between ring and trophozoite probe-labelled proteomes became larger (18.2% or 38 proteins) when compared with trophozoite protein hits identified by all artemisinin activity-based probes (Figure 5.5). This resulted in 32 proteins uniquely identified from ring stage parasite.

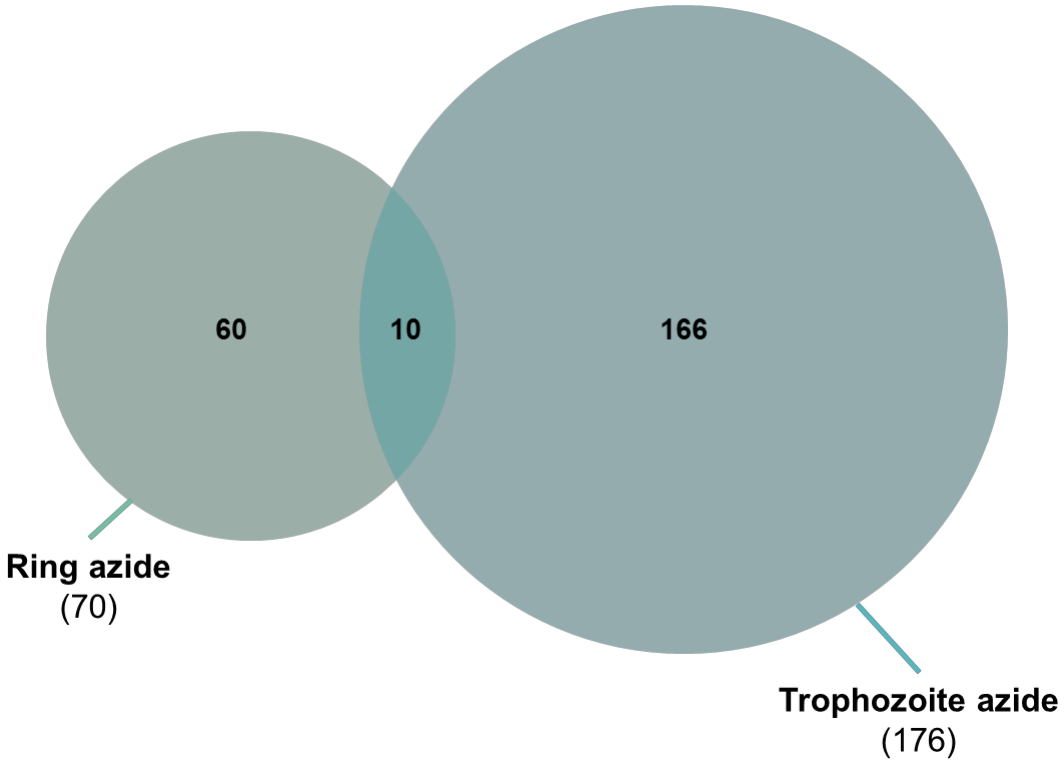


Figure 5.4 Proportional Venn diagram showing proteins identified by azide probe from ring and trophozoite stages

Figure 5.5 Proportional Venn diagram showing proteins identified by azide probe from ring stage and by all probes from trophozoite stages

5.3.3 GO biological enrichment of ring stage protein hits

To evaluate which biological processes were mostly affected, the 70 ring stage protein hits were enriched for pathway by GO enrichment. The majority of the proteins identified in ring stages were ribosomal proteins (27%). Gene expression, hexose catabolism, purine nucleoside metabolism, and cell redox homeostasis were also enriched from the ring stage probe-labelled proteome (Figure 5.6).

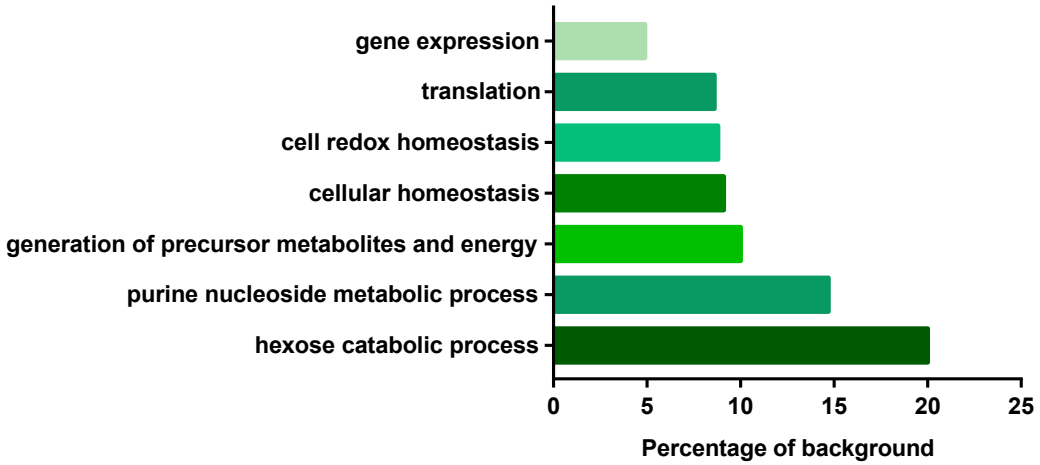


Figure 5.6 GO enrichment by biological process of artemisinin activity-based probe protein hits

5.4 Discussion

In this section, proteins hits are discussed according to their metabolic pathway with a focus on some uniquely identified proteins from ring stage proteome and when compared with trophozoite stage parasites.

5.4.1 Less proteins were identified in ring stages by the azide probe

By comparing the protein hits from ring and trophozoite stages as identified by the azide artemisinin activity-based probe a 60% reduction in total number of proteins identified was noted (from 176 to 70) (Figure 5.4). This finding is in line with data suggesting that ring stage parasite (~3-18 hpi) are much less sensitive to artemisinins than trophozoite stage or very early ring stage (0-3hpi) parasite (Klonis et al., 2013). One argument is that this parasite stage has relatively low biochemical activity and as such fewer pathways were active, or flux through pathways is reduced (Shaw et al., 2015). Interestingly the fact that some protein adducts were identified does suggest some degree of peroxide activation.

During the early stage of parasite development, parasite metabolism is relatively low compared to the later stages including the process of haemoglobin metabolism, which therefore reduces haem production and activation of artemisinin. Lower protein expression and total protein numbers in ring stages (Le Roch et al., 2004) could also contribute to the low numbers of tagged proteins seen here. However, very early ring stage parasites (~0-3 hpi) are reported to be hypersensitive to artemisinin (Klonis et al., 2013) and there is a suggestion that haemoglobin digestion is active at these very early stage of the life cycle (Xie et al., 2016). The data presented here would suggest that if the claim of Xie et al. (2016) is correct the process of haemoglobin degradation must be switched off after 3 hpi or it would be difficult to reconcile with the massive reduction in protein labelling seen in the 3-18 hpi ring stage parasites recorded here. Alternatively, there could be some protective mechanism operational in these older rings but there is little literature data to support this.

5.4.2 Haemoglobin metabolic pathway is not affected in ring stage

Unlike the trophozoite stage probe-labelled proteome, none of proteins in haemoglobin metabolism pathway were identified in ring stage endoperoxome. As previously discussed in Chapter 4, haemoglobin metabolism was heavily labelled by artemisinin as would be expected as there is substantial literature evidence implicating artemisinin haem activation as a central feature of its antimalarial activity. The absence of tagged haemoglobin metabolic enzymes in ring stage probe-labelled proteome supports the argument for reduced sensitivity to artemisinin in ring stage (3-18 hpi) stage parasite. This data can be considered as further validation of the click pull down strategy for investigating mechanisms of drug action even against the backdrop of promiscuous labelling and complex data interpretation. It is unclear from this study what the source of probe activation is that has

led to the generation of the 70 or so protein drug adducts identified but it is unlikely to be freshly generated haem from haemoglobin degradation.

5.4.3 Response to the unfolded protein and ERAD pathway

There were 4 proteins adducted that are involved in response to the unfolded protein response (UPR) in ER and ERAD pathways identified from ring stage proteome, namely protein disulfide isomerase 8 (PDI8), protein transport protein SEC61 subunit alpha (SEC61), translocation protein SEC63, (SEC63), and cell division cycle protein 48 homologue, putative (CDC48 or P97). The latter 2 were not identified in the trophozoite stage parasite study.

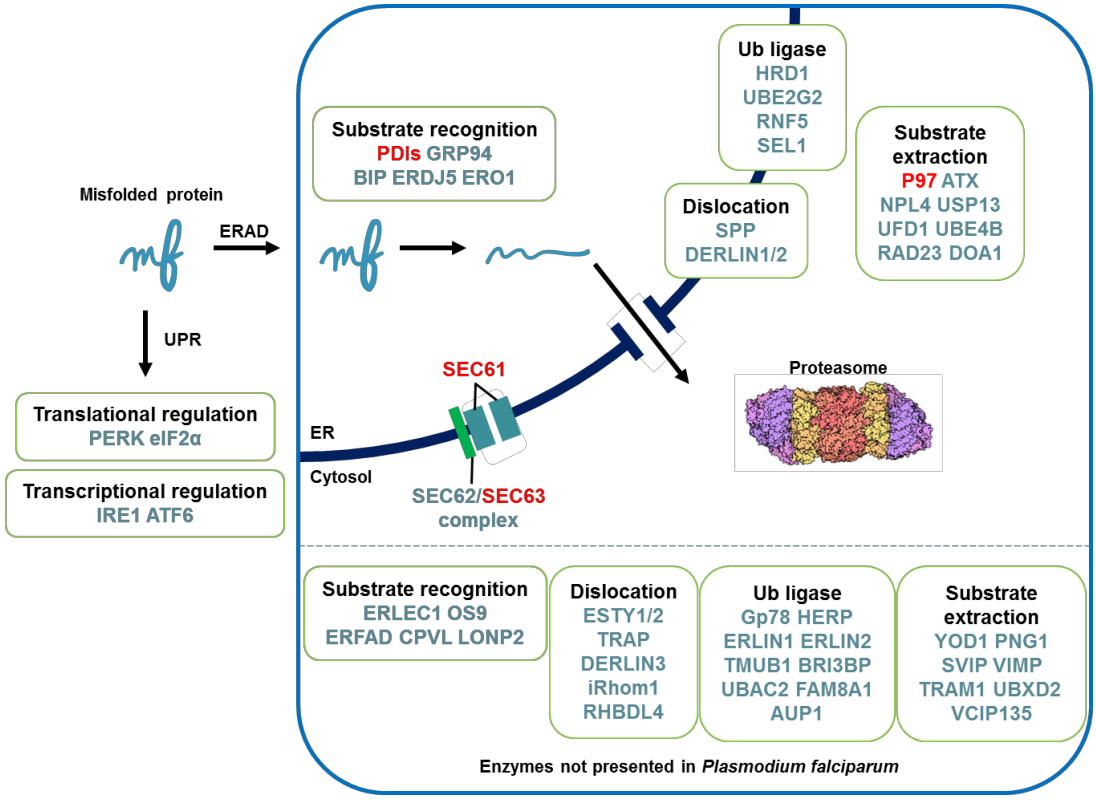


Figure 5.7 Endoplasmic reticulum associated degradation (ERAD) pathway. Proteins indecated in red were identified from the experiment.

As previously discussed in Chapter 4, the ERAD pathway in ER and chaperone system are critically important for the parasite. Inhibition of the pathways enzymes severely affects parasite growth. There was a reduction in the proportion of chaperone proteins identified in ring stages (4.08% or 4 proteins) compared to trophozoites (20.4% or 20 proteins) as shown in Table 4.5. This reduction in chaperones labelling by the artemisinin activity-based probe might underlie some reduced susceptibility of the ring stages to artemisinin. In addition, resistance to artemisinin is associated with upregulation of proteins involved in the unfolded protein response suggesting UPR as the (direct or indirect) target for

artemisinin (Mok et al., 2015). Also proteasome inhibition have a synergistic effect with artemisinin during ring stages and can overcome the artemisinin resistance (Dogovski et al., 2015). This data is intriguing and suggests the need for a more careful characterisation of the stage specific functioning of this important biochemical pathway.

Cell division cycle protein 48 homologue, putative (CDC48 or P97) was a protein uniquely identified from the ring stage study. Although there is no data on this protein in *Plasmodium falciparum*, its homologues have been studied in mammalian cells and zebrafish. A study from cancer cells revealed inhibition of P97 by the specific inhibitor N^2,N^4-dibenzylquinazoline-2,4-diamine (DBeQ) blocks degradation of unfolded proteins by the ERAD pathway, and induces the caspase pathway inhibiting cell proliferation (Chou et al., 2011). In addition, overexpression of the CDC48 gene by low temperature in zebrafish cells increases cell proliferation but the expression level was unaffected in a Tyr805Ala CDC48 mutant (Imamura et al., 2003). Both studies implicated importance of P97 in controlling cell proliferation and suggest that it is a promising drug target for cancer and also for possibly malaria? More recently, the P97 human homologue VCP structures were resolved in complex with an inhibitor by cryo-electronmicroscopy technique (Banerjee et al., 2016). The study in human P97 revealed the inhibitor interferes the movement of P97 preventing it from catalysing the reaction.

5.4.4 Transport proteins

The work by Antoine et al. (2014) showed that artemisinins affect the membrane potential of the trophozoite stage parasites and the trophozoite probe-labelled proteome supports these findings. Due to the fact that none of V-type proton pumps or other transport proteins involved in membrane potential generation was labelled in the ring stage, despite some activation, suggests that ring stage parasites can possibly maintain membrane potential in the presence of artemisinin stress or that activation occurs at sites away from the mitochondrion. Both ring and trophozoite stage parasites have similar membrane potentials of -90±3 mV and -93±10 mV, respectively (Mikkelsen et al., 1986). This difference could also point to the role of the mitochondria in trophozoites which are susceptible to drug action compared to rings that are not.

The other 2 transport proteins ***Pf*ATP6** and ***Pf*CRT** were also pulled-out from the ring stage proteome. These 2 proteins might contribute to mechanism of action of artemisinin as previously discussed in Chapter 4. However, more solid evidence on these proteins are required to draw any clear conclusion.

5.4.5 Protein and nucleic acids biosynthesis

Similar to the trophozoite stage probe-labelled proteome, ribosomal subunits were promiscuously labelled by artemisinin activity-based probe (15 proteins, Table 5.1). Along with ribosomal subunits,

elongation factor 1-gamma was labelled in the ring stage, whereas elongation factor 2 (eEF2) was identified in the trophozoite stage. The possibility of these proteins having a role in the mechanism of artemisinin action is discussed in Chapter 4.

5.4.6 Postulated mechanism of reduced sensitivity to artemisinin in ring stages

The results from ring and trophozoite stages endoperoxomes together with existing knowledge on parasite biology suggested a plausible mechanism that might underlie different sensitivity of various stages of the parasites to peroxides as follows. During ring stages, haemoglobin degradation is minimal, so less artemisinin is activated (Klonis et al., 2011, Shaw et al., 2015). Less activation of artemisinin is reflected in less proteins identified by artemisinin activity-based probes. Furthermore, the major pathways targeted by artemisinin in the trophozoite stages and implicated in drug action are less affected or not operational at all in the ring stages, including haemoglobin metabolism and elements of the unfolded protein response. The absence of any proteins involved in haemoglobin metabolism in the ring stage parasite endoperoxome suggests the pathway is essentially switched of in these rings or significantly down regulated compared to the trophozoite stage. It is important to note that if haemoglobin degradation and heam generation is central to drug activation and parasite death these would be the first proteins an activated drug would see within the parasite, hence the expected adduct formation (Figure 5.8).

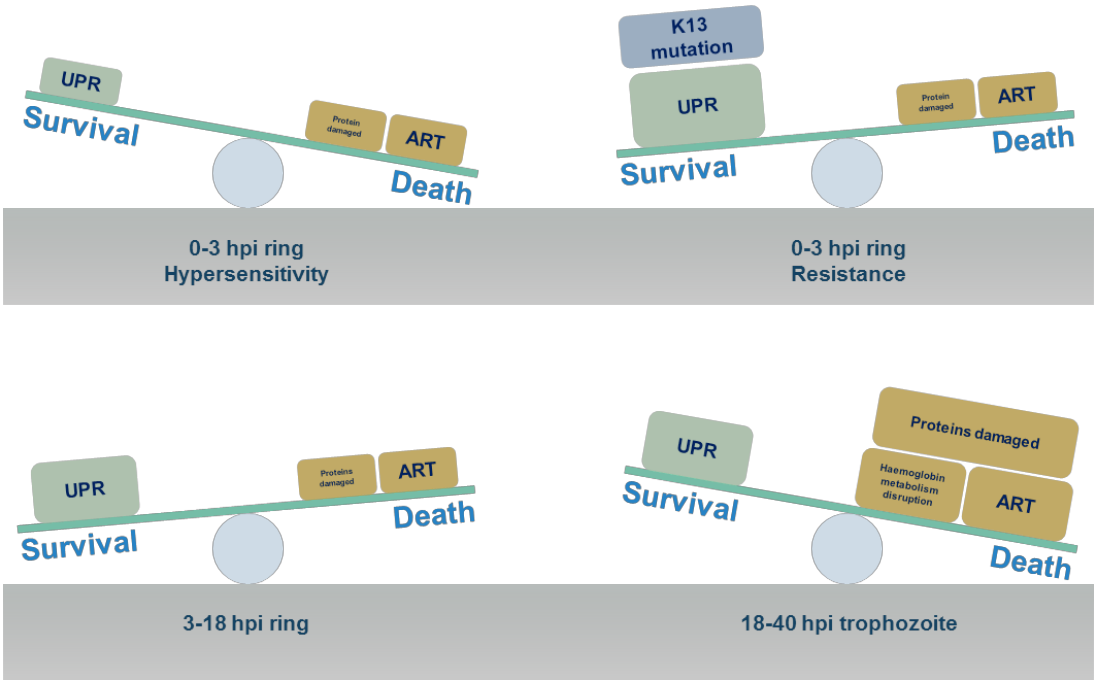


Figure 5.8 Postulated mechanism of different sensitivity to artemisinin in different stages of parasite. (A) During very early ring stage, the hypersensitivity to artemisinin is possibly due to low unfolded protein response (UPR) system and could not overcome the effects of protein damages caused by artemisinin (even though low activation of artemisinin). (B) K13 mutation and up-regulation of the UPR in the resistance parasites could overcome the effects of artemisinin during very early ring stage. (C) At ring stage, the parasites become insensitive to artemisinin is possibly due to low activation of artemisinin, less protein alkylation and slightly more UPR system. (D) During trophozoite stage, artemisinin activation is increased, more damage to the proteins including haemoglobin digestion so the UPR system and other stress response could not overcome the effects of artemisinin.

5.5 Conclusions

There are serious limitations conducting experiment on ring stage parasite particularly related to low protein abundance per unit parasite, challenging and inefficient parasite recovery after saponin lysis, and synchronisation to confidently generate very early ring stages (~0-3 hpi) on mass. Parasite biomass in the ring stage is obviously much lower than in the trophozoite stage and therefore these experiments required much more parasite material to generate equivalent amount of proteins for click reactions. In order to purify ring stage parasite from the culture and to reduce host protein contamination, saponin lysis had to be used and this led to a significant loss of parasites. Saponin lysis has much poorer lysis efficiency in ring stages compered to trophozoite stage parasite, also the absence of pigment made recovery from centrifugation after saponin lysis much less efficient. Significant efforts (over almost a year) were made to generate very young synchronous rings 0-3 hpi without success.

Despite all these technical difficulties it has been possible to confidently define the endoperoxome of the mid–late ring stage parasite (3-8hpi). A limited number of proteins were identified and the overlap with the trophozoite was minimal. This is useful on several levels, the differential susceptibility if rings and trophozoites is linked with very different pull down profiles moreover the proteins where there is overlap may represent promiscuous irrelevant targets for further investigation, i.e. an artefact of the pull down strategy. Alternatively, the fact that we still see some activation in these relatively insensitive ring stages and that key components of pathways implicated in drug action are tagged may suggest that a more details stage specific global “omic” analysis of the parasite could help unlock the basis of the differential drug sensitivity patterns that have been reported.

Chapter 6

The "endoperoxome" of a fully synthetic endoperoxide activity-based probe in trophozoite stage malaria parasites

6.1 Introduction

Resistance to the current antimalarial drugs and emerging resistance to the first line treatment artemisinin based combination therapies (ACTs) compromises all treatment and control interventions. Therefore, new antimalarials are urgently needed. Following the success of artemisinin in the reduction of malarial burden, a group of new antimalarial fully synthetic peroxide drugs was developed based on the same core structure of endoperoxide artemisinins, namely the 1,2,4-trioxolanes (Figure 1.14). The first generation of fully synthetic trioxolane, OZ277, exhibits good antimalarial activity *in vitro* and *in vivo*, but poor bioavailability (Vennerstrom et al., 2004, Kreidenweiss et al., 2006) and instability in parasitised blood. The next generation trioxolane, OZ439, was developed to improve the bioavailability (~20 h elimination half-life) and blood stability (Charman et al., 2011) and thus drug is in late phase clinical trials. It is hoped that such drugs will reduce the recrudescent rate and treatment failure associated with short half-life peroxides such as artemisinins and OZ277.

Like artemisinin, the mechanism of action of these trioxolanes remains elusive, however, it was shown that like the semisynthetic artemisinins, trioxolanes are activated and via iron species leading to radical formation (Fugi et al., 2010). The work by Fugi et al. (2010) demonstrated that the radical spin trap 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) and its analogues antagonised the effect of artemisinin and OZ277 *in vitro*. This finding implies that antimalarial activity of artemisinin and OZ277 is radical dependent. The activators of fully synthetic trioxolanes are believed to be free iron and haem-iron (Stocks et al., 2007). Fully synthetic trioxolanes seems to be more effective than semisynthetic artemisinins because activated artemisinins undergoes self-quenching process more so than that of fully synthetic trioxolane (Fugi et al., 2010).

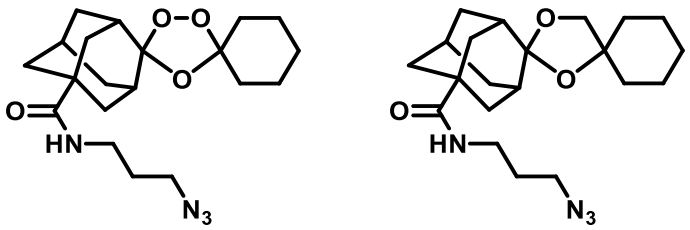
Although activation of artemisinins and fully synthetic trioxolanes shared many similarities in terms of mechanism of activation it is unclear if the specific end targets of drug action are the same. This has important implications for cross resistance patterns and long term utility of second and third generation peroxide antimalarials. Following the successful identification of artemisinin targeted protein partners (the endoperoxome) from trophozoite and ring stage falciparum parasites using artemisinin activity-based probes as demonstrated in Chapters 3, 4, and 5, the same conceptual approach was applied to study the molecular targets (endoperoxome) of these fully synthetic trioxolane compounds. The central

question was do the synthetic trioxalones and the semi-synthetic artemisinins share common intra-parasitic protein targets or not?

6.2 Experimental

6.2.1 Chemical probes

In this study, trioxolane probes were synthesised mimicking fully synthetic trioxolane core structure (Figure 1.14). Only azide linker probes were developed for use in this study. Again we adopted the pairwise approach with an active trioxalone (P4) compared directly with an inactive equivalent (CP4).



P4 CP4

Figure 6.1 Chemical probes used. P4 active probe contains 1,2,4 trioxolane structure, while CP4 inactive control probe is non-epoxidic counterpart.

6.2.2 *In vitro* activity of chemical probes

Prior to the parasite treatment with the probes, antimalarial activity of the probes was evaluated using standard SYBR Green I-based assay (section 2.17). Assay plates containing test compounds were created using a Hamilton robotic platform in 1:4 dilutions, the final concentration ranged from 1,000 nM to 0.061 nM. Parasite inoculum (1% haematocrit and 2% parasitaemia) was added to each well and incubated under normal physiological condition. The assay was terminated by freezing the plates at -20°C and nucleic acid content was measured by SYBR Green I nucleic acid stain. IC_{50} was calculated by GraphPad Prism® 7.0 software.

6.2.3 Parasite treatment

Probe treatment was performed as described in section 2.23. Each treatment was performed with 10 flasks of synchronous trophozoite stage parasites at 1 µM for 6 h under normal culture conditions. After incubation, parasites were released from RBCs by saponin lysis (section 2.18), washed with D-PBS thrice, resuspended in 1X protease Roche cOmplete™ protease inhibitor in D-PBS, and stored at -80°C for further processing.

6.2.4 Protein extraction and sequencing

Parasite protein extraction was performed by sonication (section 2.24). Parasite samples were allowed to defrost at RT and then sonicated briefly for 5 s, thrice. Parasite debris was obtained by centrifugation and soluble proteins were collected for and prepared for the click reaction.

Copper-free click reaction was performed as detailed in section 2.26. DIBO-alkyne in DMSO was added directly to each sample to a final concentration of 20 µM and incubated for 1 h in the absence of light and with gentle mixing every 15 min to ensure homogeneity throughout the reaction. Conjugated proteins were sedimented by centrifugation at 6,500g, 4 min, at 4°C. The protein pellet was then washed 3 times with cold absolute methanol to remove excess non-labelled proteins and reagents. Streptavidin affinity enrichment was performed to purify probe-labelled proteins (section 2.27), followed by on bead protein reduction and alkylation with DTT and IAA. Processed protein samples were digested with sequencing grade modified trypsin overnight and reactions were terminated with 90% formic acid. Tryptic peptide samples were sequenced by HPLC-MS/MS method (section 2.28). Protein identification was performed via Thermo Scientific Proteome Discoverer™ 1.4 software.

6.3 Results

6.3.1 Activity of chemical probes

Like the semisynthetic artemisinin probes, the modified fully synthetic probes were first tested for antimalarial activity *in vitro* by standard SYBR Green I-based assay. IC_{50} values for the P4 probe was 7.675 nM (SE = 1.58, n = 2) while CP4 has no activity *in vitro* against *Plasmodium falciparum* 3D7 strain (Figure 6.2), in the concentration range tested.

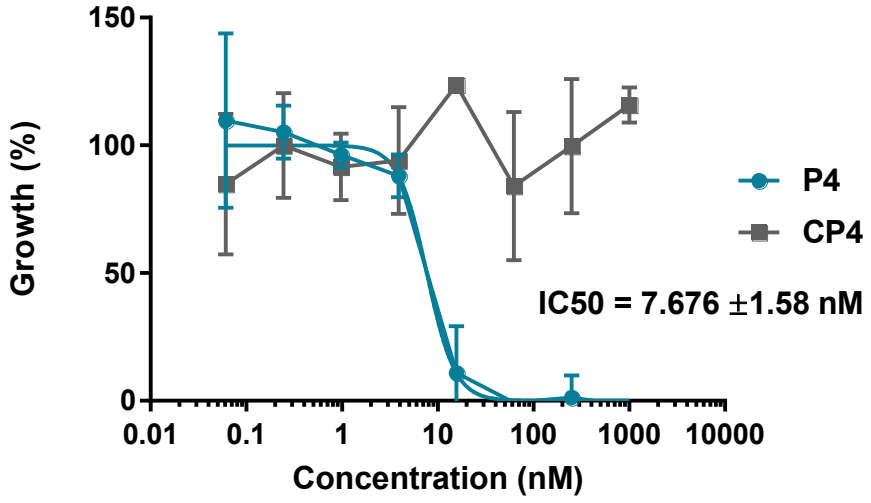


Figure 6.2 Dose response curve of trioxolane activity-based probe (P4).

6.3.2 Proteins identified by trioxolane activity-based probe

The optimised protocol reduced the non-specific binding to non-detectable levels (no proteins were identified from DMSO and inactive probe incubations). By using aforementioned method, 54 proteins were identified with this trioxolane activity-based probes as drug protein adducts. The proteins identified by fully synthetic trioxolane probe were largely overlapped with endoperoxome of artemisinin in trophozoite stage parasites (Figure 6.3). The probe-labelled proteome was then manually interrogated for S-glutathionylation potential (Kehr et al., 2011) (Figure 6.4).

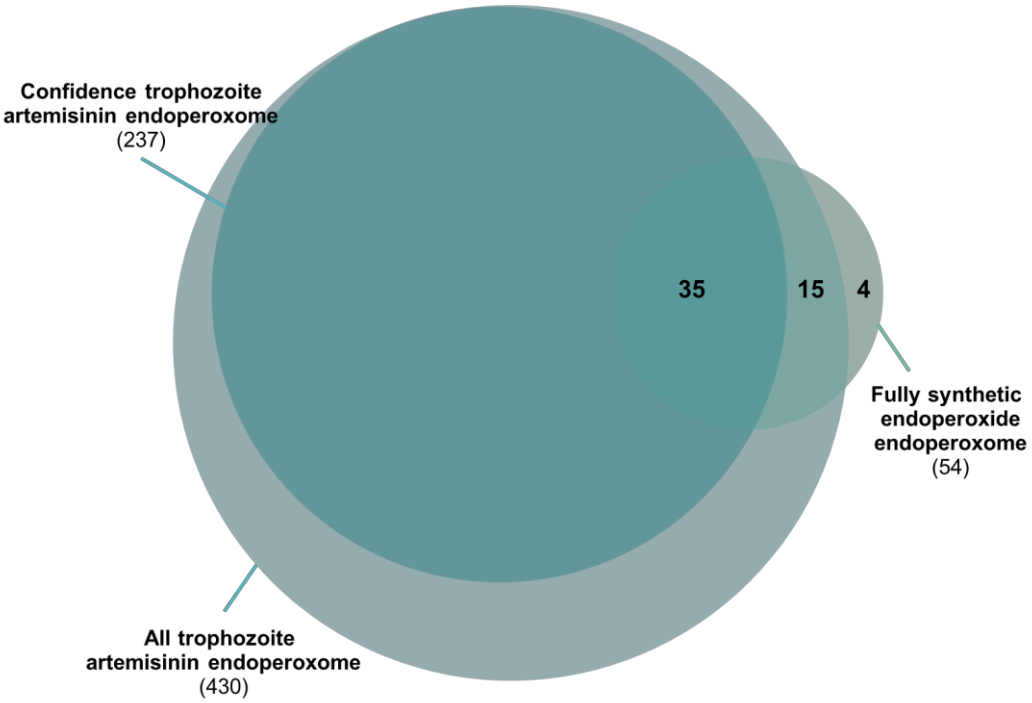


Figure 6.3 Proportional Venn diagram showing number of proteins identified by semi-synthetic and fully-synthetic exdoperoxide probes from trophozoite stage parastites.

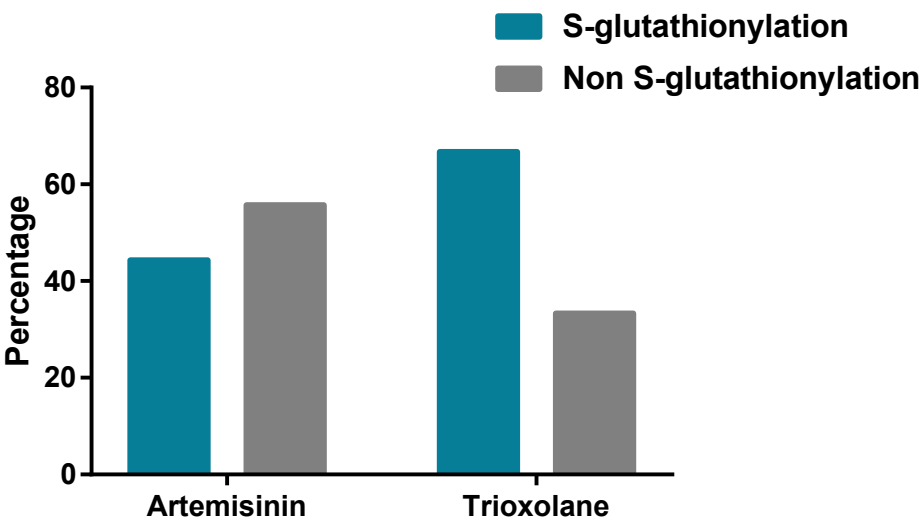


Figure 6.4 Protein S-glutathionylation percentage of artemisinin and trioxolane probes-labelled proteomes

Table 6.1 Proteins identified by trioxolane activity-based probe (2 repeats)

Gene ID	UniProt ID	Product Description	Artemisinin Acitivity-Based Probes	GSH Binding
PF3D7_0203000	O96133	repetitive organellar protein, putative	Trophozoite and ring	-
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase	Trophozoite and ring	+
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase	Trophozoite and ring	-
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2	Trophozoite and ring	+
PF3D7_0102200	Q8I0U6	ring-infected erythrocyte surface antigen	Trophozoite	+
PF3D7_0309600	O00806	60S acidic ribosomal protein P2	Trophozoite	+
PF3D7_0406100	Q6ZMA8	V-type proton ATPase subunit B	Trophozoite	+
PF3D7_0503400	Q8I467	actin-depolymerizing factor 1	Trophozoite	+
PF3D7_0520900	P50250	adenosylhomocysteinase	Trophozoite	+
PF3D7_0617800	C6KT18	histone H2A	Trophozoite	-
PF3D7_0624000	C6KT76	hexokinase	Trophozoite	+
PF3D7_0708400	Q8IC05	heat shock protein 90	Trophozoite	+
PF3D7_0903700	Q6ZLZ9	alpha tubulin 1	Trophozoite	-
PF3D7_0917900	Q8I2X4	heat shock protein 70	Trophozoite	+
PF3D7_1008700	Q7KQL5	tubulin beta chain	Trophozoite	-
PF3D7_1015900	Q8IJN7	enolase	Trophozoite	+
PF3D7_1121600	Q8IIF0	exported protein 1	Trophozoite	+
PF3D7_1136500	C6S3F7	casein kinase 1	Trophozoite	-
PF3D7_1228600	Q8I5D2	merozoite surface protein 9	Trophozoite	-
PF3D7_1246200	Q8I4X0	actin I	Trophozoite	+
PF3D7_1311800	Q8IEK1	M1-family alanyl aminopeptidase	Trophozoite	+

Gene ID	UniProt ID	Product Description	Artemisinin Acitivity-Based Probes	GSH Binding
PF3D7_1311900	Q76NM6	V-type proton ATPase catalytic subunit A	Trophozoite	+
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase	Trophozoite	+
PF3D7_1357000	Q8I0P6	elongation factor 1-alpha	Trophozoite	+
PF3D7_1361900	P61074	proliferating cell nuclear antigen 1	Trophozoite	+
PF3D7_1405600	Q8IM38	ribonucleoside-diphosphate reductase small chain, putative	Trophozoite	+
PF3D7_1408000	Q8I6V3	plasmepsin II	Trophozoite	+
PF3D7_0608800	Q6LFH8	ornithine aminotransferase	Trophozoite	+
PF3D7_0818900	Q8IB24	heat shock protein 70	Trophozoite	+
PF3D7_0922500	P27362	phosphoglycerate kinase	Trophozoite	+
PF3D7_0930300	Q8I0U8	merozoite surface protein 1	Trophozoite	+
PF3D7_1020900	Q7KQL3	ADP-ribosylation factor	Trophozoite	+
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4	Trophozoite	+
PF3D7_1130200	Q8II61	60S ribosomal protein P0	Trophozoite	+
PF3D7_1354500	Q8IDF6	adenylosuccinate synthetase	Trophozoite	+
PF3D7_0308200	O77323	T-complex protein 1 subunit eta	Trophozoite noise	+
PF3D7_0318300	O77374	conserved Plasmodium protein, unknown function	Trophozoite noise	-
PF3D7_0322900	O97313	40S ribosomal protein S3A, putative	Trophozoite noise	-
PF3D7_0417200	Q8I1R6	bifunctional dihydrofolate reductase-thymidylate synthase	Trophozoite noise	+
PF3D7_0711000	P46468	AAA family ATPase, CDC48 subfamily	Trophozoite noise	-
PF3D7_0919000	Q8I2W3	nucleosome assembly protein	Trophozoite noise	+
PF3D7_1035200	Q03400	S-antigen	Trophozoite noise	-
PF3D7_1213800	Q8I5R7	proline--tRNA ligase	Trophozoite noise	+

Gene ID	UniProt ID	Product Description	Artemisinin Acitivity-Based Probes	GSH Binding
PF3D7_1342600	Q8IDR3	myosin A	Trophozoite noise	-
PF3D7_1407900	Q7KQM4	plasmepsin I	Trophozoite noise	+
PF3D7_1427900	Q8ILI6	conserved protein, unknown function	Trophozoite noise	-
PF3D7_1434300	Q8ILC1	Hsp70/Hsp90 organizing protein	Trophozoite noise	-
PF3D7_1436000	Q8ILA4	glucose-6-phosphate isomerase	Trophozoite noise	+
PF3D7_1439900	Q7KQM0	triosephosphate isomerase	Trophozoite noise	+
PF3D7_1453800	Q8IKU0	glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase	Trophozoite noise	-
PF3D7_0605800	C6KSQ6	DNA repair protein RAD50, putative	none	-
PF3D7_0731500	Q8IBE8	erythrocyte binding antigen-175	none	-
PF3D7_1419300	Q8ILQ7	glutathione S-transferase	none	+
PF3D7_API0030			none	-

6.4 Discussion

6.4.1 Activity pairwise strategy

IC_{50}s of P4 and CP4 probes showed a similar pattern as seen with artemisinin activity-based probes and their non-peroxidic partners (section 3.3.1). P4 containing the 1,2,4-trioxolane structure and azide functional group exhibited antimalarial activity at a similar potency to artemisinin and dihydroartemisinin (Figure 6.2). This structure was also classified as an endoperoxide as with artemisinins. Earlier studies also reported that iron chelators (Stocks et al., 2007, Meshnick et al., 1993) and radical spin trap TEMPO (Fugi et al., 2010) reagents antagonised the antimalarial effects of the trioxolanes and that activation of these compounds is iron and/or haem dependent. These experiments confirm that the artemisinins and other trioxolanes have similar activation processes and, possibly, similar/identical mechanisms of action.

6.4.2 Trioxolane probe shared similar profile with artemisinin probes

Comparison of the proteins adducted by the trioxolane probe (Figure 6.3) with equivalent data from the semi-synthetic probes (Table 4.1) showed that 66% of the proteins identified by trioxolane activity-based probe were also identified by artemisinin activity-based probes from trophozoite stage or up to 94% when included the proteins identified by artemisinin activity-based probe from a few experiments (regarded as noise, Table A1.6). This finding strongly suggested that the fully synthetic trioxolane-based compounds might share similar mechanism of action with artemisinins. The parasites' key metabolic pathways including glycolysis, haemoglobin digestion, nucleic acid and protein biosynthesis were affected by both the artemisinin and trioxolane (Table 6.1). Similar to artemisinin, proteins identified from the trioxolane activity-based probe were 68% S-glutathionylation proteins (Figure 6.4), suggesting that alkylation at cysteine residues might play an important role in mechanism of action of endoperoxide compounds. Furthermore the endoperoxome of the trioxolane and the artemisinins were quantitatively very similar (Figure 6.5) again suggesting significant overlap in the activation of these chemically distinct peroxides.

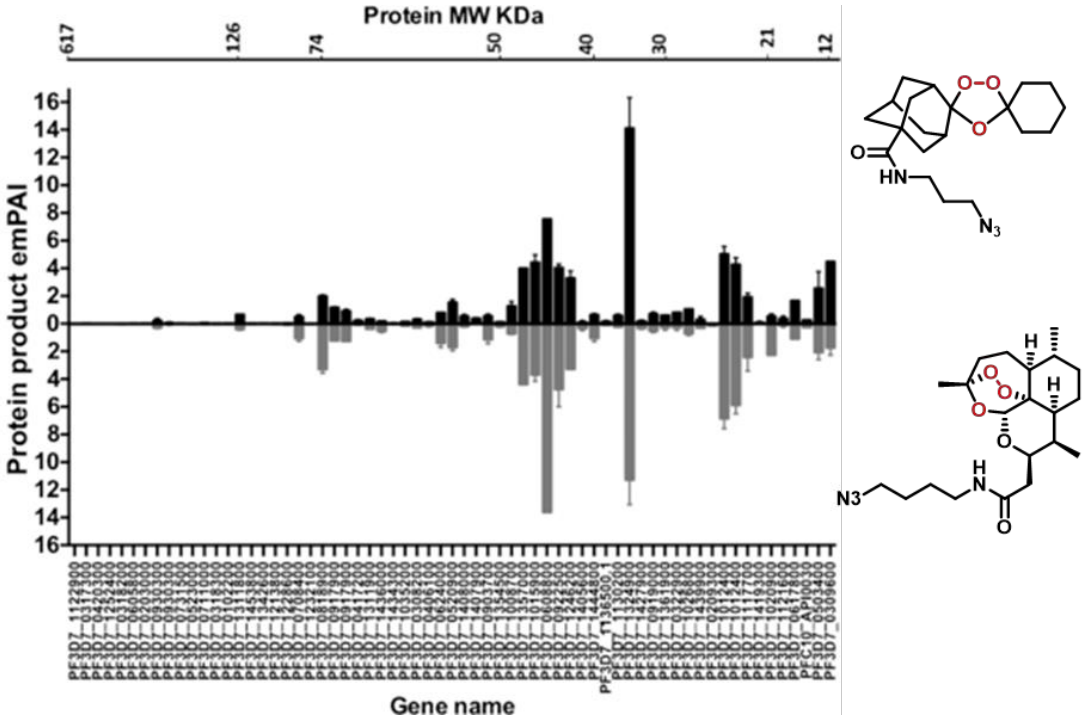
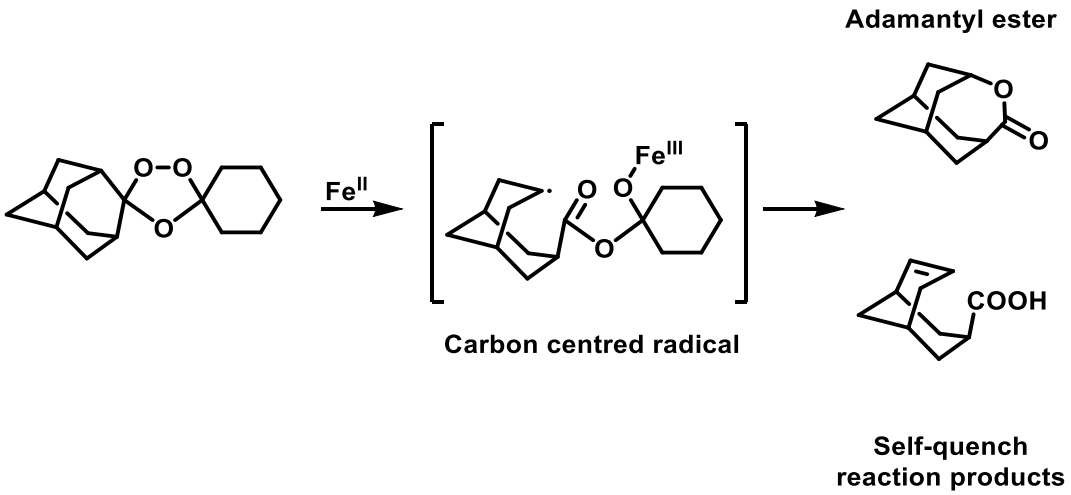


Figure 6.5 Comparison of protein adducts identified by fully synthetic (top) and semi-synthetic (bottom) endoperoxides. Proteins are sorted according to their molecular weight. This figure was generated by Dr Hanafy Ismail.

The shared mechanism of action raises the concern of possible cross resistance among this group of compounds. Promisingly, there is no report of resistance or reduced sensitivity to OZ439 from the clinical trials to date but it is very early in development (Phyo et al., 2016). It is believed that the long elimination half-life of OZ439 is effective for current artemisinin resistant parasites with mutations in K13 gene (mediated through delayed development in early stage of the life cycle). Artemisinins with a short half-life, it is argued, do not provide adequate antimalarial cover once the K13 mutant has made very early stage rings insensitive to peroxide action (Phyo et al., 2016).

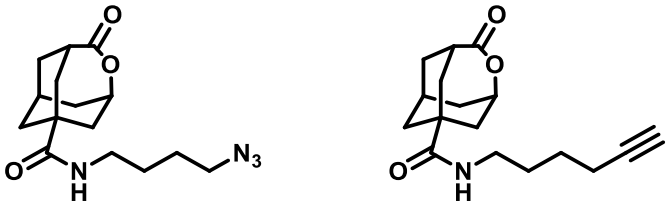
6.4.3 Adamantyl ester has no antimalarial activity and possibly not involved in protein alkylation

There is a concern that alkylation profiles identified by trioxolane probe was not due to the radical of trioxolane radicals but adamantyl ester, the by-product of fully-synthetic trioxolane activation (Scheme 6.1).



Scheme 6.1 Activation of trioxolane by iron (II), adapted from Fugi et al. (2010).

Therefore adamantyl ester probes (Figure 6.6) were synthesised and evaluated for *in vitro* antimalarial activity. Preliminary results showed that both compounds failed to inhibit the parasite growth at up to 1,000 nM, and unlikely to alkylate parasite proteins. However, further validation of the results by MS approach is required.



SCR-AD-07 SCR-AD-08

Figure 6.6 Adamantyl ester click probes.

6.5 Conclusions

The data presented in this chapter confirm that the endoperoxome of the fully syhthetic peroxide trioxalones share significant overlap with that of the semi-synthetic artemisinins. This overlap in adducted proteins suggests commonality in terms of mechanisms and sites of activation and possibly the same mechanism of parasite killing. It is notable that the probe tagging pattern were very similar both in terms both qualitatively (the specific proteins tagged) and quantitatively (the amount of protein adduct). The observation that again nearly 70% of these proteins can be glutathinoylated points to a common reactivity profile. This data does raise some concerns over potential cross-resistance mechanisms and the scope for development of second and third generation peroxides. Are all peroxide classes just a vehicle for delivering the endoperoxide warhead? However, this needs to be considered in the context of the potential for modified and improved human pharmacokinetics that may make the K13 mutant irrelevant in a clinical context.

Chapter 7

Comparative genomics of artemisinin resistant *Plasmodium falciparum*

7.1 Introduction

The technology advancement in sequencing has accelerated genome research in the past years and coined the term next-generation sequencing (NGS). The benefits over the traditional Sanger sequencing are affordability and speed. However, most NGS platforms resulted in relatively shorter reads and higher error rates.

Prior to NGS technology, whole genomes were sequenced by Sanger sequencing, shotgun sequencing, and yeast artificial chromosome (YAC) systems (Figure 7.1), but these techniques were limited to relatively small genomes, they were technically difficult, and they were expensive (Goodwin et al., 2016). The first genome ever sequenced was bacteriophage ΦX174 using the plus and minus method (Sanger et al., 1977). The first free-living organism to have its complete genome sequenced was the bacteria *Haemophilus influenzae* Rd using shotgun sequencing (Fleischmann et al., 1995), then in 1996 the first eukaryote genome, *Saccharomyces cerevisiae,* was completed (Goffeau et al., 1996). Not long after, the first multicellular organisms, *Caenorhabditis elegans,* genome was completed in 1998 by shotgun sequencing (Consortium, 1998). The *Plasmodium falciparum* genome was completed in 2002 with worldwide collaboration. Unlike many other genomes, the malaria genome is AT rich with multiple repeats. These unusual features of the *Plasmodium* genome poses technological and sequence processing challenges (Gardner et al., 2002).

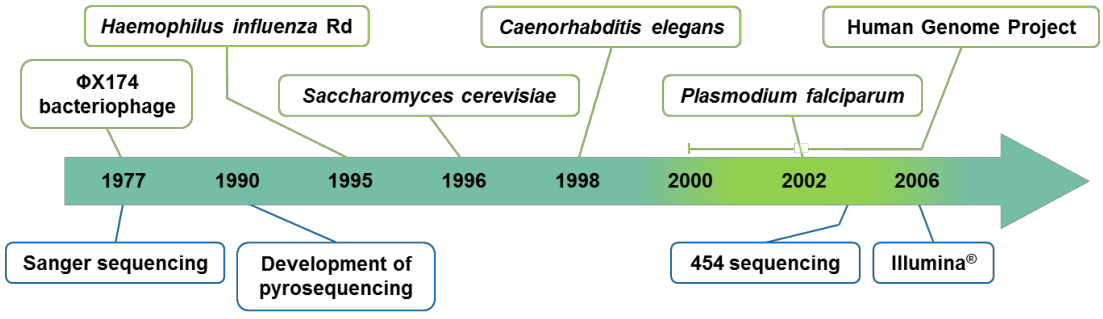


Figure 7.1 Timeline of key events in genome sequencing

Next-generation sequencing arrived in the mid-2000s when the 454 sequencing method hit the market (Figure 7.1). The 454 sequencing is based on sequencing-by-synthesis (SBS), the newly incorporated nucleotide releases a pyrophosphate group (PPi) and a subsequent enzymatic reaction emits detectable

light, also referred to pyrosequencing (Rothberg and Leamon, 2008). Most NGS technologies rely on this SBS principle but might differ in the specific of the technique (Shendure and Ji, 2008).

7.1.1 Next-generation sequencing (NGS)

Next-generation sequencing (NGS) is a high-throughput sequencing technology. It is considerably faster and cheaper than the conventional Sanger sequencing technique, which relies on chain termination of DNA strands. NGS technologies include Illumina®, Roche 454, Ion Torrent™, and SOLiD® sequencing technologies. Each technology has its advantages and disadvantages, but mostly rely on sequencing-by-synthesis principle (SBS). Most of NGS results in a very large data output but usually of short read length (~35-700 bp) (Goodwin et al., 2016).

7.1.2 Illumina® sequencing

Illumina® sequencing uses the sequencing-by-synthesis (SBS) technique. The technology involves 3 keys steps: library preparation, cluster generation, and sequencing-by-synthesis. Briefly, DNA samples (library) is prepared by excising DNA to a specific length and adding an adaptor to each end. The adaptor region is important for subsequent steps. Cluster generation generates clonal clusters of DNA via a bridge amplification process. Then each cluster is sequenced simultaneously.

Library preparation is to prepare the DNA template for subsequent sequencing. The DNA template is simultaneously ligated and tagged with adapter using transposon, this process is called tagmentation. Then reduced cycle amplification adds additional motifs including sequencing primers, indexes, and flow cell complementary oligos (Figure 7.2).

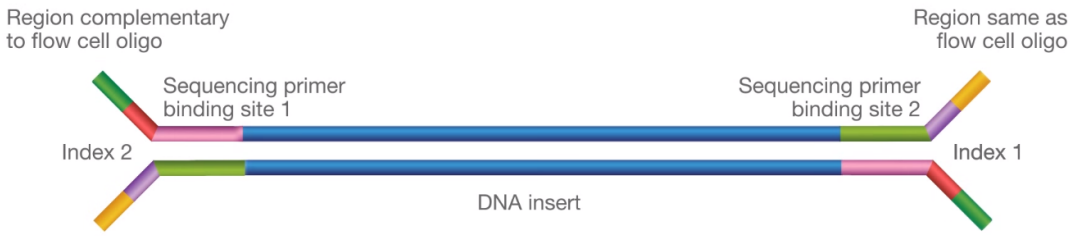


Figure 7.2 Modified DNA molecule (taken from Illumina® website)

Cluster amplification is via bridge amplification (Figure 7.3). Briefly, the DNA molecule is hybridised to the flow cell via an adaptor flow cell complementary oligo at the end of each modified DNA molecule. Then polymerase synthesises complementary stand ①. Both strands are denatured and the template is washed out ②. The remaining strand is folded over and hybridised to the flow cell oligo, forming a bridge structure and the polymerase synthesises complementary a strand ③, then the double strand is

denatured leaving each strand attached to the flow cell ④. The process is repeated to generate clonal clusters ⑤. After bridge amplification, the reverse strand is washed out from the flow cell, and the forward strand is sequenced ⑥.

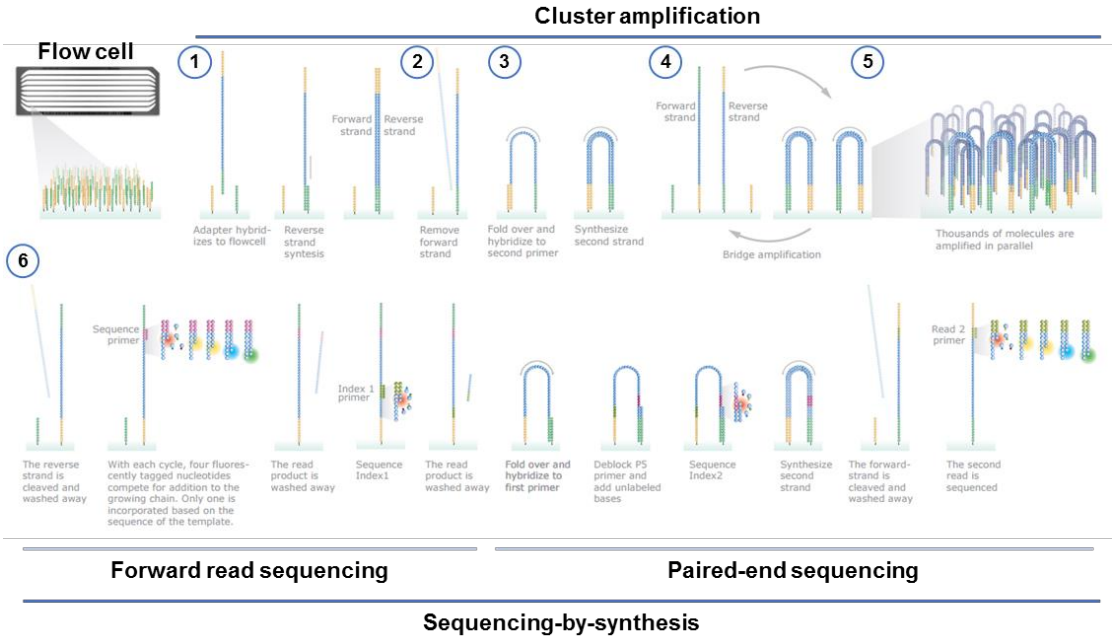


Figure 7.3 Illumina® sequencing technology (adapted from Illumina® website)

The forward strand is sequenced by sequencing-by-synthesis technique. Fluorescent-tagged nucleotide is incorporated into the newly synthesised strand and fluorescence is activated and recorded in each cycle. Then fluorescent tag is cleaved allowing new complement nucleotide to incorporate into the molecule. This process is repeated for a defined number of cycles resulting in equal read lengths for every cluster. Then product strand is denatured and washed off. The remaining strand is folded over and hybridised with another flow cell oligo. Polymerase synthesises complementary reverse strand onto the flow cell oligo. Both strands are denatured, and the forward strand is cleaved and washed out leaving the reverse strand tethered on the flow cell. The reverse strand is sequenced. This process is called paired-end sequencing (Figure 7.3). This paired-end sequence can be used to verify sequencing quality.

In this study NGS was used to sequence the whole genomes of artesunate-induced resistant parasites developed *in vivo* SCID mouse model. The NGS is a high-throughput technology providing a greater opportunity to study a subtle change in the genomes.

7.2 Experimental

7.2.1 Parasite isolates used in the study

All parasite isolates used in this study were generated in collaboration with the GSK Open Lab, Tres Cantos, Spain, and generously given to LSTM. Parasites isolates were generated by subsequently repeated infection to SCID mice under artesunate dug pressure *in vivo*.

Briefly, *Plasmodium falciparum* strain 3D7 parasites were injected to immunocompromised SCID mouse, initiating passage 1, and treated with artesunate monotherapy, 10 mg/kg for 3 days. Then mouse blood was collected and subsequently injected to another SCID mouse. The parasitaemia was monitored regularly in every passage to generate parasite clearance half-life. Parasite clearance half-life was plotted for each passage to establish the resistance break-point (data not shown). Parasite clearance times remained constant under drug pressure until passage 9 when the clearance time more than doubled. This slower clearance rate was maintained in subsequent passaged despite an increase artesunate dose of 50mg/kg (the highest dose tested) and was stable in the absence of drug selection over 3 years (GSK Tres cantos personnel communication). This phenotype was very similar to the slow clearance phenotype seen in clinical studies in S.E. Asia.

7.2.2 Parasite culture

Eighteen sequentially selected parasite isolates were received at LSTM and cryopreserved until retrieved for culture. Parasites were initiated from cryopreservation stock and maintained as described in Chapter 2 section 2.13. Parasite pellets were collected once culture exhibited >10% parasitaemia at 2% haematocrit and preserved in 1 mL TE buffer at -80°C until extraction.

7.2.3 Genomic DNA extraction and purification

Parasite genomic DNA was extracted from parasite pellets preserved in TE buffer by adding 20 µL of proteinase K solution and heated to 65ºC for 5 h. An aliquot of 1 mL phenol:chloroform:isomoeyl alcohol (25:24:1) was added to extract DNA samples and centrifuged at 17,000g for 1 min at RT, then the upper layer (aqueous phase) was collected into a new tube. The phenol/chloroform purification step was repeated for 3 times, then 750 µL of molecular grade absolute ethanol was added to each sample to precipitate genomic DNA which was then stored overnight at -20ºC. Genomic DNA was retrieved by centrifugation to precipitate the genomic DNA sample in absolute ethanol at 17000g for 20 min at 4ºC. The supernatant was removed and the pellet was dried at 50ºC for 30-45 min until completely dry as ethanol can interfere subsequent PCR. Genomic DNA was resuspended in 50 µL pre-warmed molecular grade water.

7.2.4 Genomic DNA quality control

The required quality of genomic DNA required for Illumina® platform sequencing is very high, only genomic DNA with a 260/280 ratio >1.8 and a 260/230 >2.0 was submitted for sequencing. NanoDrop™ was used to determine the 260/280 ratio (section 2.20). Agarose gel electrophoresis was used to check for genomic DNA quality, RNA contamination and genomic DNA integrity (section 2.21). A Qubit™ flourometric quantitation kit was used to determine DNA quantity (performed by the Centre for Genomic Research, UoL).

7.2.5 Genomic DNA sequencing

Purified genomic DNAs were submitted for sequencing using the Illumina® HiSeq™ 2500 platform at the Centre for Genomic Research (CGR), University of Liverpool (UoL). Illumina® TruSeq™ DNA sample preparation kit was used to prepare the library from PCR-free gDNA samples. Sequencing was performed in 2 different batches, LIMS2208 and LIMS4913.

7.2.6 Sequence analysis

This experimental part was performed by LSTM's Bioinformatics Unit; Dr Simon Wagstaff, Dr Enrique Salsedo-Sora, Dr John Archer, and Dr Gareth Weedall all contributed to this analysis. The analysis pipeline is shown in Figure 7.4.

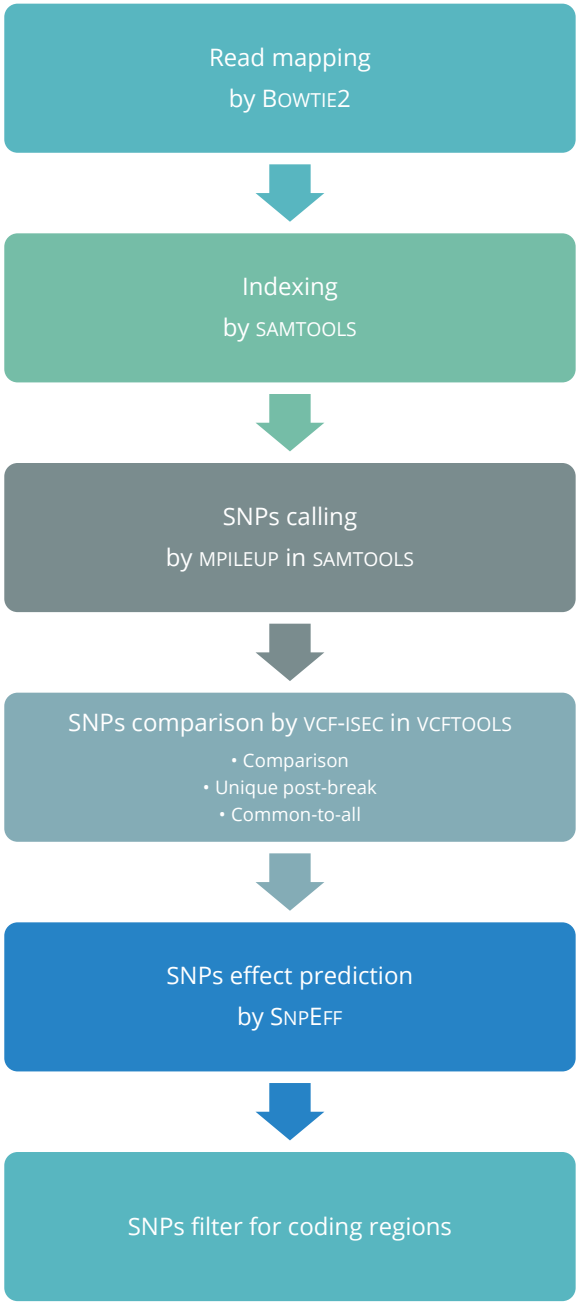


Figure 7.4 Analysis pipeline

Briefly, sequences retrieved from CGR were in FASTQ format including paired-end reads (forward and reverse reads) and unpaired reads. The sequences were already trimmed for Illumina® adapter sequences using CUTADAPT version 1.2.1 (Martin, 2011) and further trimmed with a quality score of >20 by SICKLE version 1.200 (Joshi and Fass, 2011). Only paired-end reads were used for mapping by BOWTIE2 software with the default setting to the reference genome sequence *Plasmodium falciparum* 3D7 version 28.0 (PlasmoDB). The resulting SAM files were converted to BAM files and indexed according to 3D7 reference genome. The indexed BAM files were subsequently called for variation by MPILEUP function in SAMTOOLS and returned VCF files. The VCF files were used for subsequent analysis.

To compare SNPs between susceptible and resistant parasites, a bespoke bioinformatics pipeline including custom Perl and R scripts was developed for SNP calling, SNP annotation, and SNP visualisation as follows. SNP calling – forward and reverse FASTQ files from a total 16 samples were mapped to the reference genome sequence *Plasmodium falciparum* 3D7 version 28.0 (PlasmoDB) using BOWTIE2 with default parameters. Each alignment output (BAM) file was indexed then analysed by SAMTOOLS MPILEUP to produce a single variant call format (VCF) output describing the positions, frequency and quality of each variant detected relative to the reference genome. A break point (defined as loss of susceptibility to artesunate; data not shown) was identified between samples 9 and 10 defining 9 isolates (Samples 1-9) and 7 isolates (10-16) as pre and post-break point sample groups. Pre and post break point VCF files were then intersected using the following commands to output merged pre and post break files retaining SNPs present in at least 7 of 9 samples (pre) and every samples (post).

```
vcf-isec -f -n +7 Sample1 Sample2 Sample3 Sample 4 Sample5 Sample6 Sample7 Sample8
Sample9 | bgzip -c > pre_break.vcf.gz


vcf-isec -f -n +7 Sample10 Sample11 Sample12 Sample13 Sample14 Sample15 Sample16 |
bgzip -c > post_break.vcf.gz
```

In the final part of the SNP calling pipeline, single files or groups were further intersected to output the following comparisons summarised in table 7.1:

```
Pattern #1 "Common_to_all" – this command reports snps present in sample groups
both before (samples 1-9) and after (samples 10-16) the breakpoint.
vcf-isec -f -n =2 post_break.vcf.gz pre_break.vcf.gz

Pattern #2 "Unique post break point" - this command reports SNPs present in all
samples after the break point (samples 10-16) but not present in any of the samples
before the break point (samples 1-9).
vcf-merge Sample1 Sample2 Sample3 Sample 4 Sample5 Sample6 Sample7 Sample8 Sample9
| bgzip -c > merge.vcf.gz
vcf-isec -f -c post_break.vcf.gz merge.vcf.gz

#Pattern #3 – "Comparison" – this command reports SNPs that are present in the post
break point file that are not present in the pre-break point file i.e. output
positions that are common to all samples after the break point (samples 10-16) but
that do not occur in all (7 of 9 samples) before the break point.
vcf-isec -f -c post_break.vcf.gz pre_break.vcf.gz;
```

SNP Annotation – Functional effects and the impact of SNPs on coding sequences were predicted by SNPEFF software (Cingolani et al., 2012) using information contained in the corresponding GFF files for the reference genome available on PlasmoDB and output as separate VCF files with functional annotation appended. Original and SNPEFF annotated VCF files were used as input for a bespoke R script written to output the SNPs comparison results for visualisation (section 7.3.7).

Table 7.1 SNP calling setting for VCF-ISEC function

	SNP presented in	
	Before the break point (9 isolates)	**After the break point (7 isolates)**
Unique post-break point	0	Any
Common-to-all	At least 2	At least 2
Comparison	7 out of 9	All

7.2.7 GO enrichment and network analysis

GO enrichment analysis is an algorithm to identify which GO terms are overrepresented or underrepresented in the set of genes. GO enrichment and network analysis were performed by Cluego plugin (Bindea et al., 2009) in Cytoscape 3.0 software (Shannon et al., 2003) using the following settings; GO biological process terms/KEGG pathways, enrichment analysis, Benjamini-Hochberg correction, and GO term fusion. Cluego plugin accepts only UniProt ID, therefore *Plasmodium* systematic gene IDs were converted to corresponding UniProt IDs in the PlasmoDB database and UniProt database. Psuedogenes were not included in the analysis as no protein products were provided.

7.3 Results

All genomes were analysed in 3 ways, namely unique post-break point, common-to-all, and comparison, in addition to global analysis of SNPs. Unique post-break point analysis only focuses on any SNPs that presented after the post-break point (resistance phenotype) isolates but not before the break point (susceptible phenotype), while common-to-all analyses any SNPs that differ from the 3D7 reference genome but are common in all studied isolates. Comparison analysis compares all SNPs in all studied isolates (section 7.2.6). Results from comparison analysis were of particular interest as there are associated with developing resistant phenotype.

7.3.1 Genomic DNA quality

The quality of genomic DNA samples was determined by NanoDrop™ spectrometer and quantified by Qubit™ spectrometry. The measured results were included in Table 7.2. Contamination of proteins and other contaminants were determined from 260/280 ratio. All genomic DNA samples met the minimum quality requirements of the CGR for Illumina® platform with the exception of TE1411, but CGR accepted the sample upon quality assessed by gel electrophoresis results (Figure 7.5). Agarose gel electrophoresis was also used to assess genomic DNA quality for contamination of RNA (Figure 7.5 and Figure 7.6)

Table 7.2 Genomic DNA quality and quantity

		NanoDrop™ Quality			
Batch	**Isolate**	**A260**	**A280**	**260/280**	**260/230**
LIMS2008	TE616	16.06	7.453	2.15	2.53
	TE1201	16.253	7.336	2.22	2.65
	TE1211	18.261	8.597	2.12	2.48
	TE1389	2.585	1.253	2.06	2.15
	TE1411	4.646	2.293	2.03	1.77
	TE1419	14.303	6.624	2.16	2.7
	TE1423	2.694	1.354	1.99	2.03
	TE1439	0.965	0.442	2.18	1.96
	TE1475	4.612	2.369	1.95	1.94
LIMS4913	TE1304	0.594	0.315	1.88	2.93
	TE1328	0.922	0.495	1.86	2.57
	TE1368	1.159	0.563	2.06	6.23
	TE1373	1.129	0.616	1.83	2.45
	TE1432	0.426	0.23	1.85	3.48
	TE1435	1.806	0.96	1.88	2.5
	TE1436	4.372	2.069	2.11	2.73

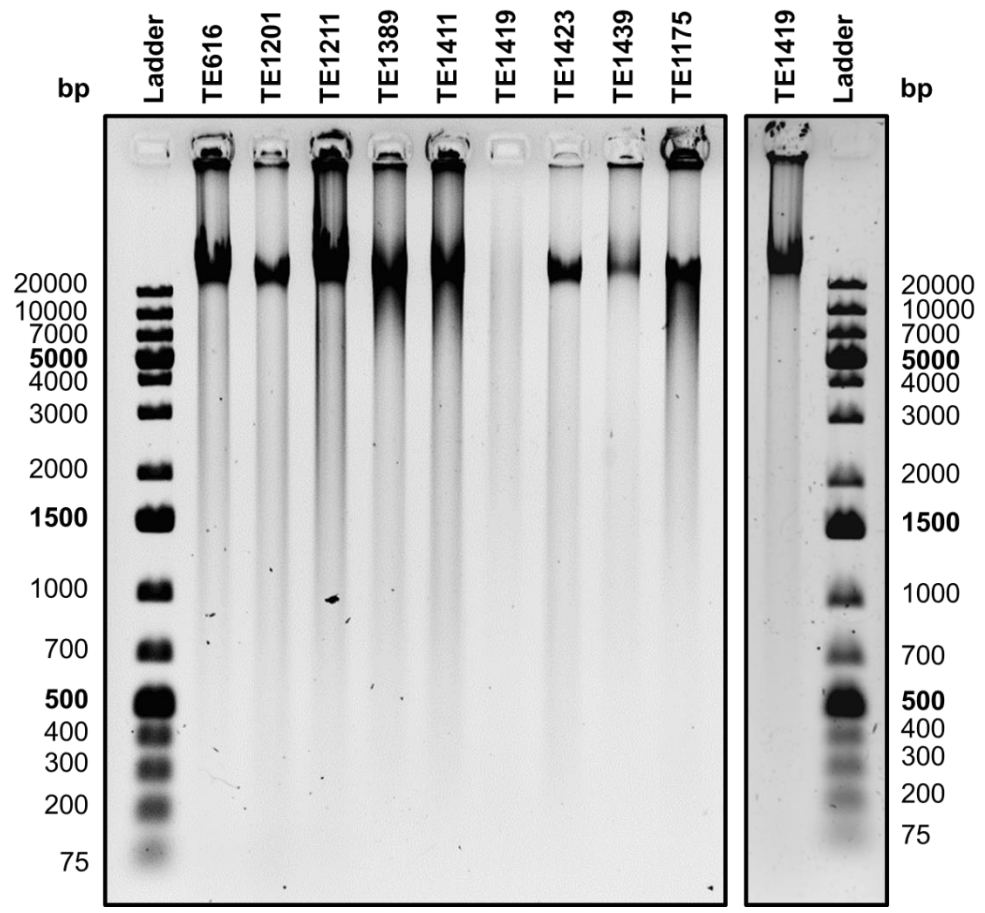


Figure 7.5 Agarose gel electrophoresis result of batch LIMS2208 samples, 1% agarose in 1X TBE buffer, 10 µL sample volume, 80V constant for 1 h 15 min.

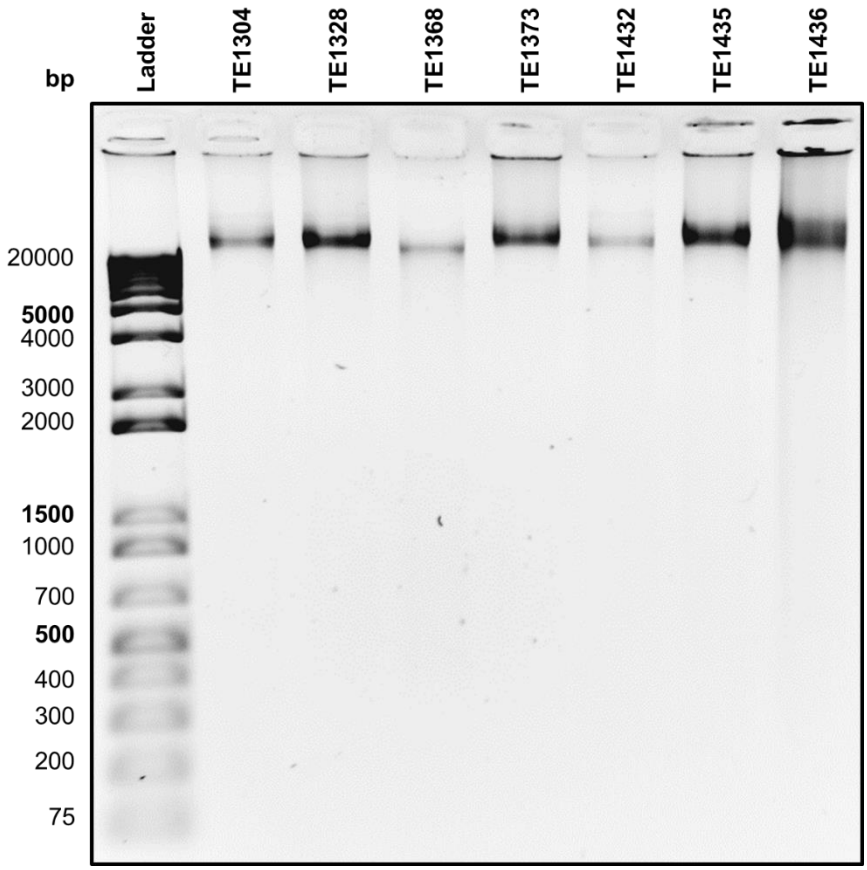


Figure 7.6 Agarose gel electrophoresis result of batch LIMS2208 samples, 1% agarose in 1X TBE buffer, 10 µL sample volume, 80V constant for 1 h.

7.3.2 Sequencing results

According to the library preparation, adapter sequences were introduced to genomic DNA samples required removal via trimming process. The reads were further trimmed to remove low quality section of reads (section 7.2.6). The trimming results are shown in Table A3.1. The number of reads from each sample is shown in Figure 7.7, with an average of ~45 million reads per sample. The average read quality score for each sample is shown in Table A3.1.

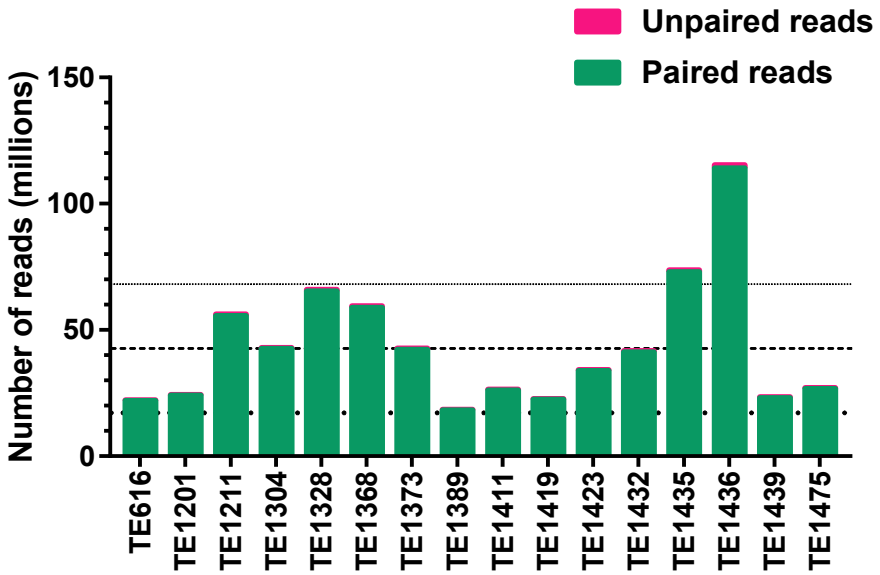


Figure 7.7 Number of reads for each sample. Paired reads are shown in green, unpaired reads are shown in grey. Thick dashed line represents mean of paired reads number and SD represented by thin dashed lines.

7.3.3 Genome mapping and SNP analysis

Sequencing reads were mapped to the reference genome *Plasmodium falciparum* strain 3D7 version 28.0 (PlasmoDB) with mapping coverage for each sample ranging from 85-98% (Table 7.3, section A3.2 and A3.3)

Table 7.3 Read mapping coverage

Sample	Mapping Coverage (%)
TE616	96.09
TE1201	98.31
TE1211	92.71
TE1304	95.34
TE1328	98.43
TE1368	96.10
TE1373	96.11
TE1389	93.14
TE1411	85.60
TE1419	98.39
TE1423	98.12
TE1432	95.98
TE1435	98.45
TE1436	90.47
TE1439	97.66
TE1475	98.05

7.3.4 SNPs number and distribution

The total SNP number was 2,851 positions, among these 806 SNPs were within gene regions (Table 7.4).

Table 7.4 Number of SNPs from analyses

Analysis	No. of SNPs	No. of SNPs in gene regions
Unique post-break	6	1
Common-to-all	288	62
Comparison	2,557	743
Total	2,851	806

The SNP number in each chromosome is commensurate with the chromosome size (Figure 7.8). It is unlikely that there is a selective pressure on a particular chromosome as SNP numbers in each

chromosome are in accordance with chromosome size and equally distributed across all chromosomes (Figure 7.9).

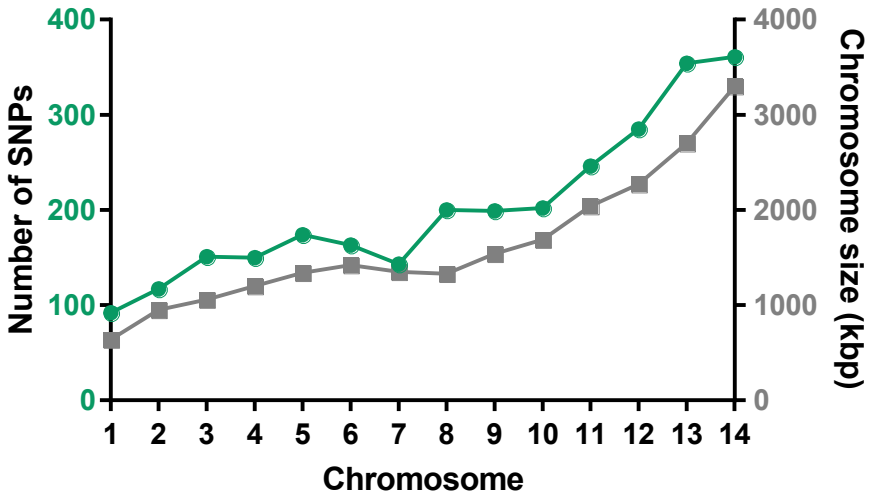


Figure 7.8 Comparison between chromosome size and SNPs in each chromosome. Total number of 2557 SNPs were plotted (SNPs from comparison analysis).

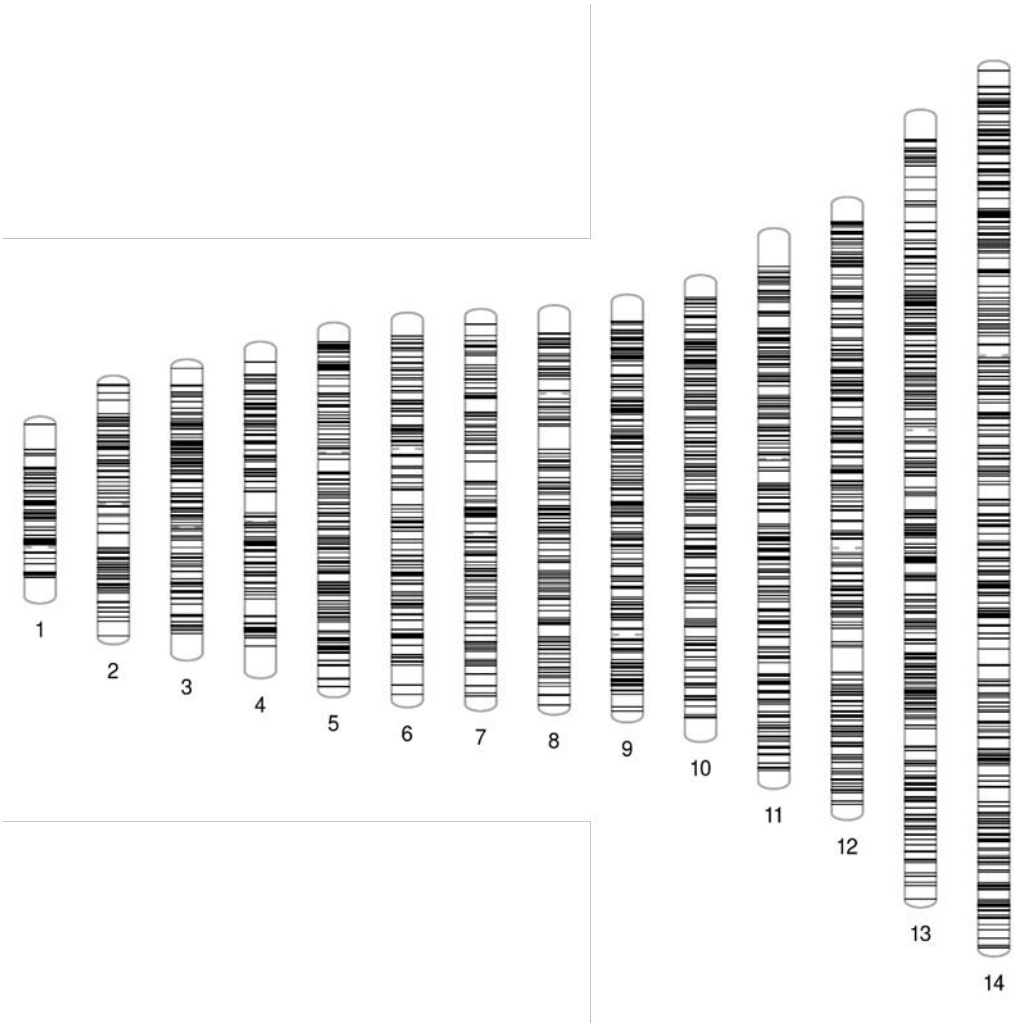


Figure 7.9 SNPs distribution on 14 chromosomes of Plasmodium falciparum (generated by online tool PhenoGram at http://visualization.ritchielab.psu.edu). This figure showed SNPs were distributed randomly on chromosomes and that sequencing was well covered the whole genome of the parasite.

7.3.5 Global analysis of SNPs

Global analysis of SNPs aimed to determine the selective pressure on particular pathways by GO enrichment analysis. The results of a 733 SNP search showed that cell communication, signal transduction, and unfolded protein response systems are significantly enriched by artesunate drug pressure (Table 7.5, Figure 7.10, and Figure 7.11).

Table 7.5 GO enrichment of 733 proteins from 743 SNPs by Cluego plugin

ID	Name	Bgd count	Pct of bgd	P-value	Benjamini
GO:0043170	macromolecule metabolic process	157	13.57	0.00000014	0.0000098
GO:0019538	protein metabolic process	106	15.75	0.00000013	0.000018
GO:0007154	cell communication	23	30.26	0.00000058	0.000027
GO:0044267	cellular protein metabolic process	90	15.52	0.0000033	0.000093
GO:0044260	cellular macromolecule metabolic process	140	13.31	0.0000029	0.0001
GO:0010467	gene expression	90	13.65	0.00036	0.0073
GO:0009187	cyclic nucleotide metabolic process	4	80.00	0.00045	0.0078
GO:0006508	proteolysis	29	19.21	0.00033	0.0078
GO:0006511	ubiquitin-dependent protein catabolic process	12	26.67	0.0011	0.014
GO:0006810	transport	43	15.93	0.001	0.014
GO:0006464	cellular protein modification process	37	16.67	0.00097	0.015
GO:0006796	phosphate-containing compound metabolic process	64	14.04	0.002	0.021
GO:0035556	intracellular signal transduction	13	24.07	0.0019	0.022
GO:0018108	peptidyl-tyrosine phosphorylation	6	37.50	0.0031	0.031
GO:0016192	vesicle-mediated transport	14	21.88	0.0034	0.032
GO:0071702	organic substance transport	22	17.74	0.0048	0.042
GO:0006468	protein phosphorylation	20	18.02	0.0059	0.046
GO:0043412	macromolecule modification	40	14.81	0.0058	0.047

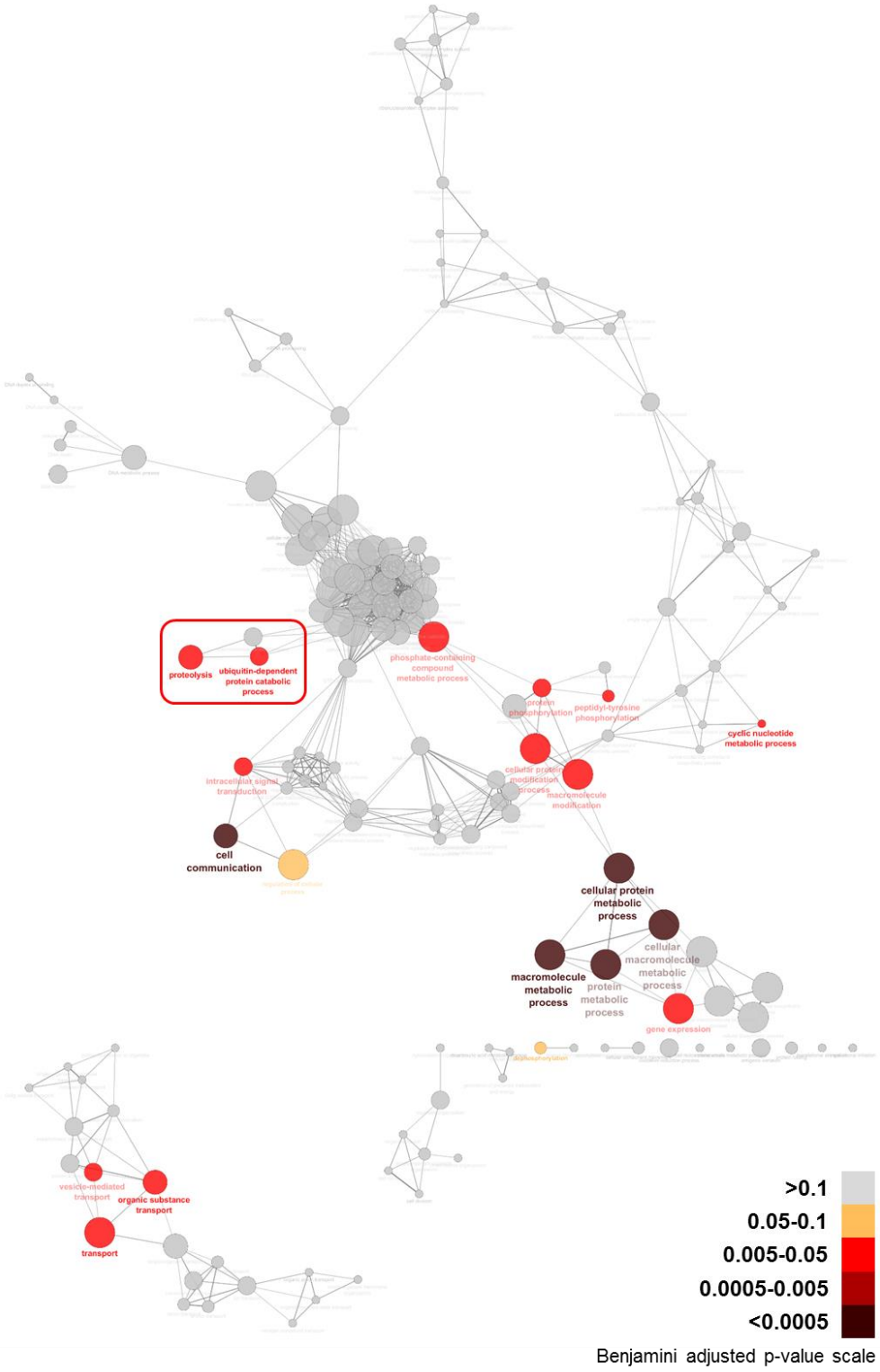


Figure 7.10 Network analysis of 733 proteins encoded from corresponding SNP containing genes. Nodes shown in colours are significantly enriched (Benjamin p-value <0.05), non-significant nodes are greyed out. Network analysis performed by Cluego plugin in Cytoscape 3.0. Nodes in red rectangle expanded in Figure 7.11.

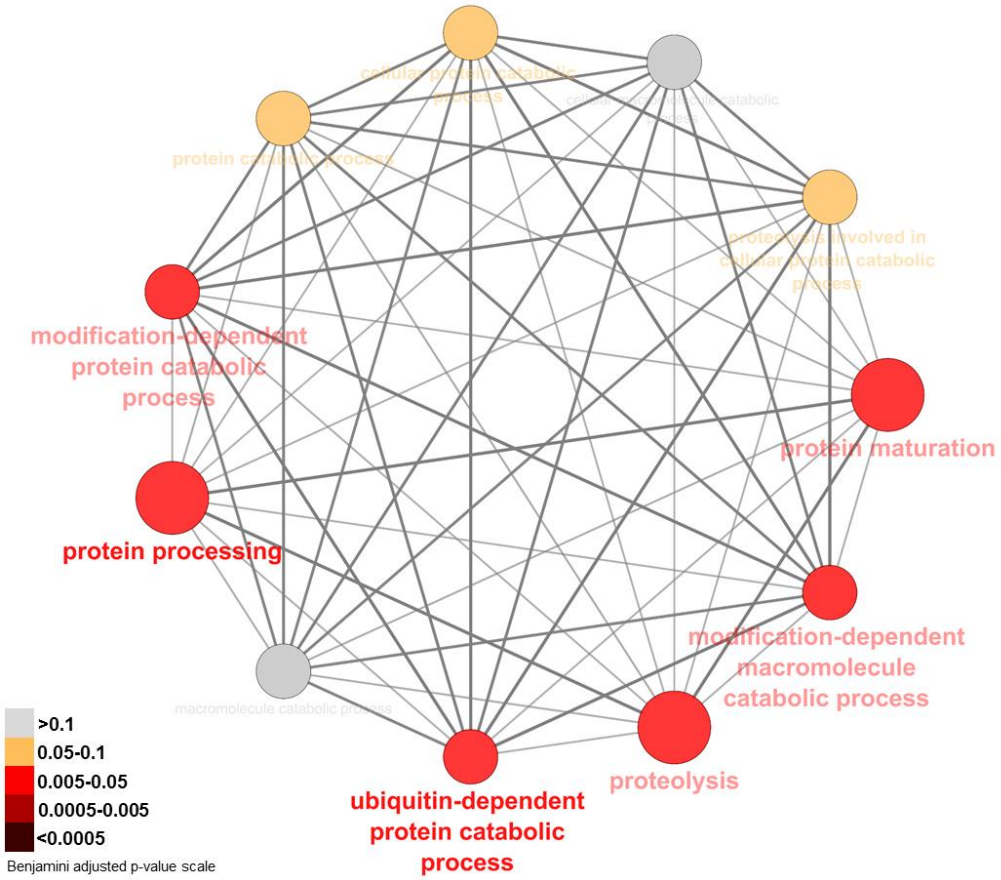


Figure 7.11 Expanded network of proteolysis and ubiquitin-mediated process shown in Figure 7.10.

7.3.6 Genes similarly identified from proteomics experiments

SNP associated genes from comparison analysis were cross referenced with the proteins identified using artemisinin activity-based probes from ring and trophozoite stages (Chapters 3-6). From 743 genes, 30 protein encoded genes were identified from the proteomic experiments (Table 7.6).

Table 7.6 SNPs associated genes similarly identified from proteomic experiments

Gene ID	UniProt ID	Product Description	Stage
PF3D7_0619400	C6KT34	cell division cycle protein 48 homologue, putative	Trophozoite and ring
PF3D7_0628300	C6KTB9	choline/ethanolaminephosphotransferase, putative	Trophozoite and ring
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter	Trophozoite and ring
PF3D7_0929400	C0H571	high molecular weight rhoptry protein 2	Trophozoite and ring
PF3D7_1008700	Q7KQL5	tubulin beta chain	Trophozoite and ring

Gene ID	UniProt ID	Product Description	Stage
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase	Trophozoite and ring
PF3D7_1129000	Q8II73	spermidine synthase	Trophozoite and ring
PF3D7_0500800	Q8I492	mature parasite-infected erythrocyte surface antigen	Trophozoite
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2	Trophozoite
PF3D7_0720400	Q8IBP8	ferrodoxin reductase-like protein	Trophozoite
PF3D7_0727400	Q8IBI3	proteasome subunit alpha type-5, putative	Trophozoite
PF3D7_0826700	Q8IBA0	receptor for activated c kinase	Trophozoite
PF3D7_0903200	C0H516	ras-related protein RAB7	Trophozoite
PF3D7_0905400	Q8I395	high molecular weight rhoptry protein 3	Trophozoite
PF3D7_0919400	Q8I2V9	protein disulfide isomerase	Trophozoite
PF3D7_0935800	Q8I2G2	cytoadherence linked asexual protein 9	Trophozoite
PF3D7_1010700	Q8IJT8	dolichyl-phosphate-mannose protein mannosyltransferase, putative	Trophozoite
PF3D7_1116800	Q8IIJ8	heat shock protein 101	Trophozoite
PF3D7_1118500	Q8III3	nucleolar protein 56, putative	Trophozoite
PF3D7_1211400	Q7KQK3	heat shock protein DNAJ homologue Pfj4	Trophozoite
PF3D7_1242700	Q8I502	40S ribosomal protein S17, putative	Trophozoite
PF3D7_1252100	Q8I4R5	rhoptry neck protein 3	Trophozoite
PF3D7_1353900	Q8IDG2	proteasome subunit alpha type-7, putative	Trophozoite
PF3D7_1361800	C0H5J9	conserved Plasmodium protein, unknown function	Trophozoite
PF3D7_1438100	Q8IL86	secretory complex protein 62	Trophozoite
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	Trophozoite
PF3D7_1474600	Q8IK92	conserved Plasmodium membrane protein, unknown function	Trophozoite
PF3D7_0317600	O77381	40S ribosomal protein S11, putative	Ring
PF3D7_1142500	Q8IHU0	60S ribosomal protein L28	Ring
PF3D7_1351400	Q8IDI5	60S ribosomal protein L17, putative	Ring

7.3.7 Visualisation of the SNP readout

A specific R script was used to generate SNP readout from the final SNP effect prediction VCF files. This graphical readout is for graphical purpose with associated information from VCF file (Figure 7.12).

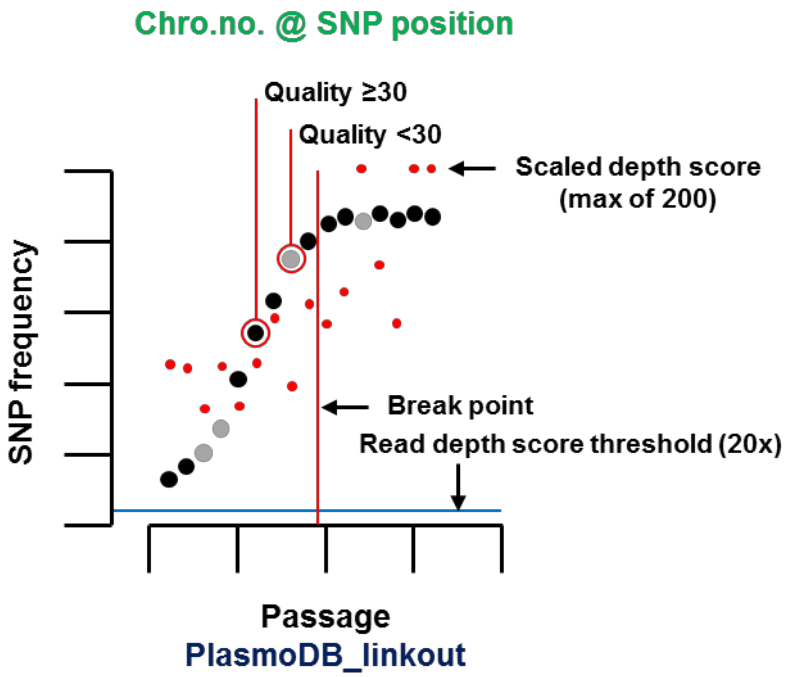


Figure 7.12 SNP comparison readout generated by R script. Vertical red line indicates resistance break point. Y-axis represents SNP frequency expressed as the sum of all non-reference SNPs/all SNPs. Horizontal blue line indicates scaled depth score threshold of (20x), while red dots represent scaled depth score to maximum of 200. SNPs frequencies represented in black when quality score ≥30, and greyed out if quality is <30. SNP position and chromosome number printed in green if SNP position is within gene coding region, black if within non-coding region of gene, and greyed out when positioned outside gene region. Link to corresponding gene in PlasmoDB website provided when SNP identified in gene region.

7.3.8 Unique post-break point analysis

The unique post-break point analysis aims to search for any SNPs that emerge after the defined resistance break point and is the most stringent analysis applied. Analysis allows zero SNPs before the break point, only SNPs detected after the resistance break point are reported from the analysis.

Search for SNPs that are unique in post-break point samples returned 6 SNPs, however, none of these SNPs were in coding regions of any gene. One SNP was within a gene but not in a coding region, 5 SNPs were in intergenic regions (Figure 7.13).

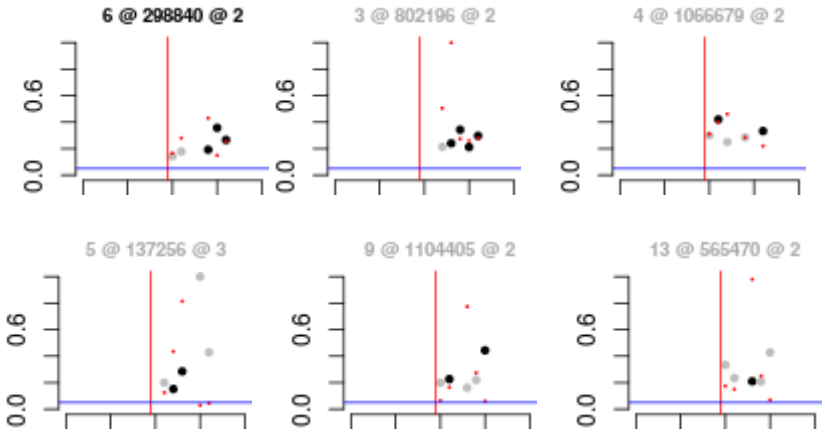


Figure 7.13 Unique post-break point SNPs

Results from vcf file with predicted SNP effects from SnpEff were presented for each SNP position in Box 7.1 to Box 7.6.

Box 7.1 VCF result for SNP 298840 on chromosome 6

```
Pf3D7_06_v3   298840 .      TTATATATATATATATATATATATATAT
       TTATATATATATATATATATATATATAT      18.5   .
       INDEL;IS=2,0.060606;DP=33;VDB=1.701979e-
01;AF1=0.5;AC1=1;DP4=9,9,2,1;MQ=39;FQ=21.5;PV4=1,1,1,1;SF=0,1;ANN=TTATATATATATATATA
TATATATATAT|downstream_gene_variant|MODIFIER|PF3D7_0607000|PF3D7_0607000|transcript
|rna_PF3D7_0607000-
1|protein_coding||c.*2071_*2072delAT|||||2071|,TTATATATATATATATATATATATATAT|downstr
eam_gene_variant|MODIFIER|PF3D7_0607200|PF3D7_0607200|transcript|rna_PF3D7_0607200-
1|protein_coding||c.*3005_*3006delAT|||||3006|,TTATATATATATATATATATATATATAT|intron_
variant|MODIFIER|PF3D7_0607100|PF3D7_0607100|transcript|rna_PF3D7_0607100-
1|protein_coding|4/15|c.492-13_492-12delAT||||||      GT:PL:DP:GQ
       0/1:56,0,255:21:59
```

Predicted affected genes:

- MYND finger protein, putative (PF3D7_0607100), intron variant
- Translation initiation factor IF-2, putative (PF3D7_0607000), downstream gene variant
- RING zinc finger protein, putative (PF3D7_0607200), downstream gene variant

Box 7.2 VCF result for SNP 802196 on chromosome 3


```
Pf3D7_03_v3   802196 .      AATATATATATATATATATATATATATATATA
      AATATATATATATATATATATATATATATATA 19.5   .
      INDEL;IS=14,0.138614;DP=101;VDB=1.922956e-
01;AF1=0.5;AC1=1;DP4=46,13,12,4;MQ=39;FQ=22.5;PV4=0.75,1,0.36,0.035;SF=0,1;ANN=AATA
TATATATATATATATATATATATATATA|downstream_gene_variant|MODIFIER|PF3D7_0319000|PF3D7_0
319000|transcript|rna_PF3D7_0319000-
1|protein_coding||c.*1057_*1058delTA|||||1057|,AATATATATATATATATATATATATATATATA|dow
nstream_gene_variant|MODIFIER|PF3D7_0319100|PF3D7_0319100|transcript|rna_PF3D7_0319
100-
1|protein_coding||c.*31_*32delTA|||||32|,AATATATATATATATATATATATATATATATA|downstrea
m_gene_variant|MODIFIER|PF3D7_0319200|PF3D7_0319200|transcript|rna_PF3D7_0319200-
1|protein_coding||c.*1363_*1364delTA|||||1364|,AATATATATATATATATATATATATATATATA|int
ergenic_region|MODIFIER|PF3D7_0319000-PF3D7_0319100|PF3D7_0319000-
PF3D7_0319100|intergenic_region|PF3D7_0319000-
PF3D7_0319100|||n.802228_802229delTA||||||     GT:PL:DP:GQ   0/1:57,0,255:75:60
```


Predicted affected genes:

- P-type ATPase, putative (PF3D7_0319000), downstream gene variant
- E3 ubiquitin-protein ligase RBX1, putative (PF3D7_0319100), downstream gene variant
- Endonuclease/exonuclease/phosphatase family protein, putative (PF3D7_0319200), downstream gene variant

Box 7.3 VCF result for SNP 1066679 on chromosome 4


```
Pf3D7_04_v3   1066679      .      AATATATATATATATATATATATATATATAT
      AATATATATATATATATATATATATATATAT     28.5   .
      INDEL;IS=8,0.129032;DP=62;VDB=3.266121e-
01;AF1=0.5;AC1=1;DP4=8,20,7,5;MQ=38;FQ=31.5;PV4=0.091,1,1,0.076;SF=0,1;ANN=AATATATA
TATATATATATATATATATATATATAT|upstream_gene_variant|MODIFIER|PF3D7_0423600|PF3D7_0423
600|transcript|rna_PF3D7_0423600-1|protein_coding||c.-84_-
83insAT|||||84|,AATATATATATATATATATATATATATATAT|downstream_gene_variant|MODIFIE
R|PF3D7_0423700|PF3D7_0423700|transcript|rna_PF3D7_0423700-
1|protein_coding||c.*4723_*4724insAT|||||4723|,AATATATATATATATATATATATATATATAT|
intergenic_region|MODIFIER|PF3D7_0423600-PF3D7_0423700|PF3D7_0423600-
PF3D7_0423700|intergenic_region|PF3D7_0423600-
PF3D7_0423700|||n.1066711_1066712insAT||||||   GT:PL:DP:GQ   0/1:66,0,255:40:69
```


Predicted affected genes:

- Conserved Plasmodium protein, unknown function (PF3D7_0423600), upstream gene variant
- Early transcribed membrane protein 4 (PF3D7_0423700), downstream gene region.

Box 7.4 VCF result for SNP 137256 on chromosome 5

```
Pf3D7_05_v3   137256 .      TTATATATATATATATATATATATATATATATATATAT
       TTATATATATATATATATATATATATATATATATATAT      10.8   .
       INDEL;IS=1,0.040000;DP=25;VDB=4.391815e-
02;AF1=0.5;AC1=1;DP4=3,5,2,0;MQ=40;FQ=13.7;PV4=0.44,1,1,0.49;SF=0,1;ANN=TTATATATATA
TATATATATATATATATATATATATATAT|upstream_gene_variant|MODIFIER|PF3D7_0503300|PF3D7_05
03300|transcript|rna_PF3D7_0503300-1|protein_coding||c.-1364_-
1363insAT|||||1364|,TTATATATATATATATATATATATATATATATATATAT|upstream_gene_variant|
MODIFIER|PF3D7_0503400|PF3D7_0503400|transcript|rna_PF3D7_0503400-
1|protein_coding||c.-3417_-
3416insAT|||||3416|,TTATATATATATATATATATATATATATATATATATAT|downstream_gene_varian
t|MODIFIER|PF3D7_0503200|PF3D7_0503200|transcript|rna_PF3D7_0503200-
1|protein_coding||c.*3901_*3902insAT|||||3902|,TTATATATATATATATATATATATATATATATAT
ATAT|intergenic_region|MODIFIER|PF3D7_0503300-PF3D7_0503400|PF3D7_0503300-
PF3D7_0503400|intergenic_region|PF3D7_0503300-
PF3D7_0503400|||n.137293_137294insAT||||||     GT:PL:DP:GQ   0/1:48,0,149:10:50
```

Predicted affected genes:

- Serine/arginine-rich splicing factor 12 (PF3D7_0503300), upstream gene variant
- Actin-depolymerizing factor 1 (PF3D7_0503400), upstream gene variant
- Conserved Plasmodium protein, unknown function (PF3D7_0503200), downstream gene variant

Box 7.5 VCF result for SNP 1104405 on chromosome 9

```
Pf3D7_09_v3   1104405      .      TTATATATATATATATATATATAT
       TTATATATATATATATATATAT   10.8   .
       INDEL;IS=2,0.153846;DP=13;VDB=9.639229e-
02;AF1=0.5;AC1=1;DP4=3,5,0,2;MQ=42;FQ=13.7;PV4=1,5.6e-
07,1,1;SF=0,1;ANN=TTATATATATATATATATATAT|upstream_gene_variant|MODIFIER|PF3D7_092
7100|PF3D7_0927100|transcript|rna_PF3D7_0927100-1|protein_coding||c.-3599_-
3598delAT|||||3598|,TTATATATATATATATATATAT|upstream_gene_variant|MODIFIER|PF3D7_0
927200|PF3D7_0927200|transcript|rna_PF3D7_0927200-1|protein_coding||c.-2714_-
2713delAT|||||2714|,TTATATATATATATATATATAT|downstream_gene_variant|MODIFIER|PF3D7
_0927000|PF3D7_0927000|transcript|rna_PF3D7_0927000-
1|protein_coding||c.*4679_*4680delAT|||||4679|,TTATATATATATATATATATAT|intergenic_
region|MODIFIER|PF3D7_0927100-PF3D7_0927200|PF3D7_0927100-
PF3D7_0927200|intergenic_region|PF3D7_0927100-
PF3D7_0927200|||n.1104429_1104430delAT||||||   GT:PL:DP:GQ   0/1:48,0,230:10:50
```

Predicted affected genes:

- conserved Plasmodium protein, unknown function (PF3D7_0927100), upstream gene variant
- zinc finger protein, putative (PF3D7_0927200), upstream gene variant

Box 7.6 VCF result for SNP 565470 on chromosome 13

```
Pf3D7_13_v3   565470 .      TTATATATATATATATATATATATATATAT
       TTATATATATATATATATATATATATATATAT      18.5   .
       INDEL;IS=4,0.114286;DP=35;VDB=2.156019e-
02;AF1=0.5;AC1=1;DP4=1,9,3,2;MQ=42;FQ=21.5;PV4=0.077,1,1,1;SF=0,1;ANN=TTATATATATATA
TATATATATATATATATATAT|upstream_gene_variant|MODIFIER|PF3D7_1313200|PF3D7_1313200|tr
anscript|rna_PF3D7_1313200-1|protein_coding||c.-735_-
734insAT|||||734|,TTATATATATATATATATATATATATATATAT|upstream_gene_variant|MODIFIER
|PF3D7_1313300|PF3D7_1313300|transcript|rna_PF3D7_1313300-1|protein_coding||c.-
3163_-
3162insAT|||||3162|,TTATATATATATATATATATATATATATATAT|downstream_gene_variant|MODI
FIER|PF3D7_1313100|PF3D7_1313100|transcript|rna_PF3D7_1313100-
1|protein_coding||c.*250_*251insAT|||||251|,TTATATATATATATATATATATATATATATAT|down
stream_gene_variant|MODIFIER|PF3D7_1313400|PF3D7_1313400|transcript|rna_PF3D7_13134
00-
1|protein_coding||c.*4338_*4339insAT|||||4338|,TTATATATATATATATATATATATATATATAT|i
ntergenic_region|MODIFIER|PF3D7_1313100-PF3D7_1313200|PF3D7_1313100-
PF3D7_1313200|intergenic_region|PF3D7_1313100-
PF3D7_1313200|||n.565501_565502insAT||||||     GT:PL:DP:GQ   0/1:56,0,177:15:59
```

Predicted affected genes:

- methionyl-tRNA formyltransferase, putative (PF3D7_1313200), upstream gene variant
- peptidyl-prolyl cis-trans isomerase, putative (PF3D7_1313300), upstream gene variant
- conserved Plasmodium protein, unknown function (PF3D7_1313100), downstream gene variant
- DEAD box helicase, putative (PF3D7_1313400), downstream gene variant

7.3.9 Common-to-all

The algorithm used in this analysis was designed to call any SNPs in studied isolates that are different from the reference 3D7 strain but common in all studied isolates (Table 7.1). This information is in general a normal genetic variation in an organism and might not contribute to artemisinin resistance phenotypes at all. Because it showed all SNPs that are fixed throughout the parasite line, regardless of phenotypes. So genes associated with these SNPs might not be important or play direct role in resistance to artemisinin.

In the studied population 288 SNPs were reported as common-to-all SNPs as compared to the reference 3D7 genome, 22 of which were SNPs within annotated coding sequence of genes (Figure 7.14).

Fifthteen genes were identified from 22 SNPs as some genes contain more than 1 SNP (Table 7.7). Among these, 6 genes encode for conserved proteins with unknown function, 5 gene products are classified as parasite exported proteins (Table 7.7 in **bold**), including *Pf*EMP1 and rifin. These gene products are parasite surface antigens involved in pathogenicity of parasites.

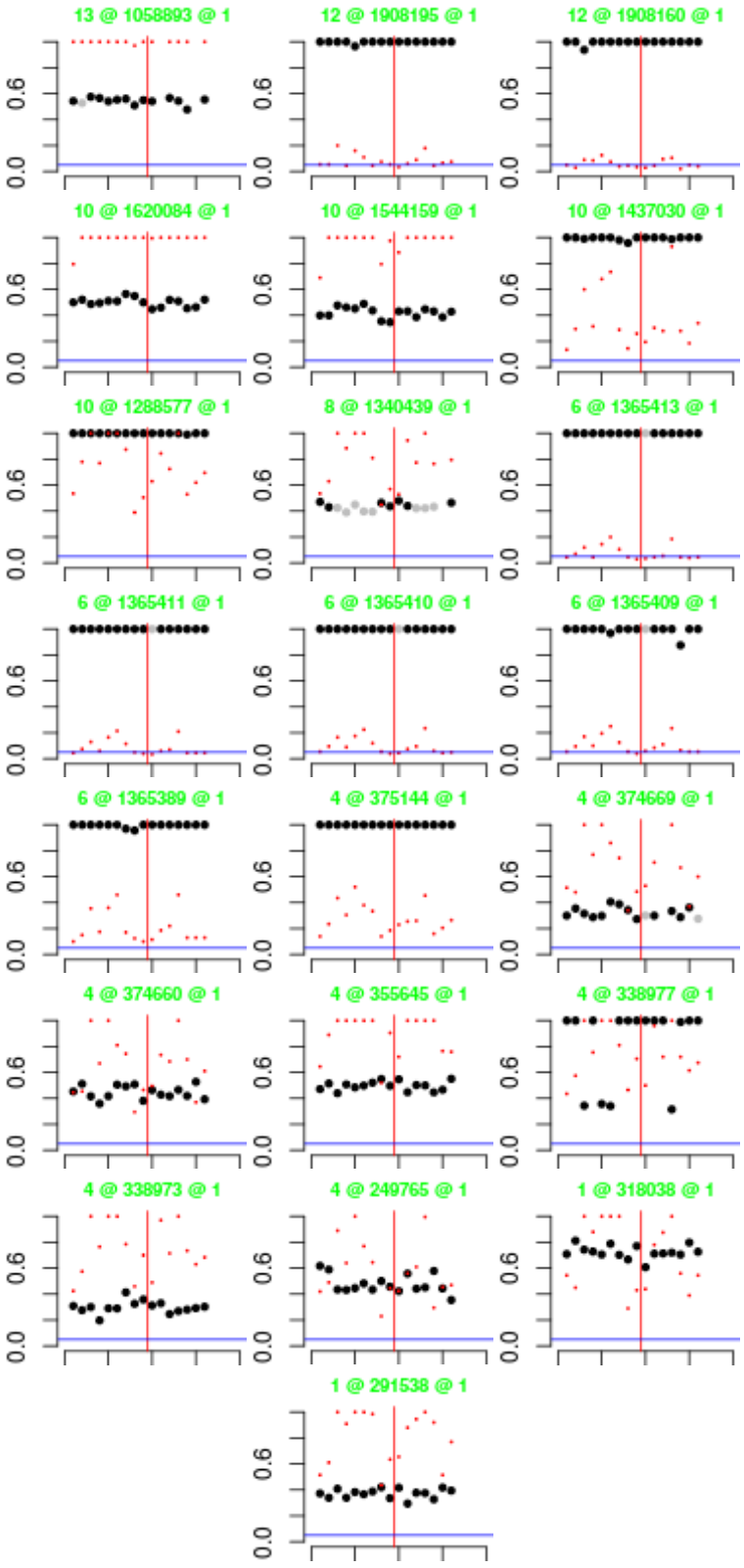


Figure 7.14 Twenty-two SNPs located within coding region from common-to-all analysis

*Table 7.7 Genes identified from common-to-all SNPs in studied isolates. Genes in **bold** are classified as exported protein (GO term: parasite exported protein).*

			SNP Effect Prediction					
Gene ID	**UniProt ID**	**Product Description**	**Total**	**High**	**Moderate**	**Low**	**Modifier**	**Genomic Location**
PF3D7_0632500	**C6KTF7**	**erythrocyte membrane protein 1, PfEMP1 (VAR)**	**6**	**1**	**4**	**1**	**-**	**Pf3D7_06_v3: 1,353,946 - 1,366,430 (-)**
PF3D7_0407600	Q9U0L0	conserved Plasmodium protein, unknown function	3	-	3	-	-	Pf3D7_04_v3: 373,039 - 376,677 (+)
PF3D7_1245800	Q8I4X4	epsin, putative	3	-	2	1	-	Pf3D7_12_v3: 1,908,149 - 1,909,764 (+)
PF3D7_0406500	C0H491	conserved Plasmodium protein, unknown function	2	-	2	-	-	Pf3D7_04_v3: 334,282 - 344,054 (-)
PF3D7_1036400	A0A143ZZD7	liver stage antigen 1 (LSA1)	2	-	1	1	-	Pf3D7_10_v3: 1,436,316 - 1,439,804 (+)
PF3D7_1038400	**Q8I6U6**	**gametocyte-specific protein (Pf11-1)**	**2**	**-**	**1**	**1**	**-**	**Pf3D7_10_v3: 1,519,021 - 1,547,825 (+)**
PF3D7_1325400	Q8IE65	conserved Plasmodium protein, unknown function	2	-	1	1	-	Pf3D7_13_v3: 1,052,751 - 1,063,310 (+)
PF3D7_1040600	**Q8IJ03**	**rifin (RIF)**	**2**	**-**	**1**	**1**	**-**	**Pf3D7_10_v3: 1,619,086 - 1,620,244 (+)**
PF3D7_0106900	Q8I273	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, putative (IspD)	1	-	1	-	-	Pf3D7_01_v3: 290,386 - 292,590 (+)
PF3D7_0404600	**Q8I1Y3**	**conserved Plasmodium membrane protein, unknown function**	**1**	**1**	**-**	**-**	**-**	**Pf3D7_04_v3: 247,731 - 260,147 (+)**

			SNP Effect Prediction					
Gene ID	**UniProt ID**	**Product Description**	**Total**	**High**	**Moderate**	**Low**	**Modifier**	**Genomic Location**
PF3D7_0406900	C0H493	conserved Plasmodium protein, unknown function	1	-	1	-	-	Pf3D7_04_v3: 353,655 - 359,878 (-)
PF3D7_1032000	Q8IJ86	ribosome maturation factor RimM, putative (RimM)	1	-	1	-	-	Pf3D7_10_v3: 1,287,715 - 1,290,045 (-)
PF3D7_1146200	Q8IHQ6	conserved Plasmodium protein, unknown function	1	-	-	1	-	Pf3D7_11_v3: 1,826,203 - 1,826,792 (-)
PF3D7_0107600	Q8I265	serine/threonine protein kinase, putative	1	-	1	-	1	Pf3D7_01_v3: 314,618 - 319,405 (+)
PF3D7_0831300	**C0H4Z6**	**Plasmodium exported protein, unknown function (GEXP13)**	**1**	**-**	**1**	**-**	**1**	**Pf3D7_08_v3: 1,339,659 - 1,342,306 (+)**
		Total	**55**	**2**	**18**	**7**	**28**	

Box 7.7 VCF result for SNP 291538 on chromosome 1 (IspD gene)

```
Pf3D7_01_v3   291538 .      A      G      204.22 .
      AC1=1;AC=9;AF1=0.5;AN=18;DP4=617,381,320,283;DP=1758;FQ=157;MQ=27;PV4=0.046,1
,0.095,1;RPB=-7.531736e+00;SF=0,1;VDB=1.988432e-
03;ANN=G|missense_variant|MODERATE|PF3D7_0106900|PF3D7_0106900|transcript|rna_PF3D7
_0106900-
1|protein_coding|1/1|c.1153A>G|p.Asn385Asp|1153/2205|1153/2205|385/734||,G|upstream
_gene_variant|MODIFIER|PF3D7_0107000|PF3D7_0107000|transcript|rna_PF3D7_0107000-
1|protein_coding||c.-
2104A>G|||||2104|,G|downstream_gene_variant|MODIFIER|PF3D7_0106800|PF3D7_0106800|tr
anscript|rna_PF3D7_0106800-
1|protein_coding||c.*2620A>G|||||2620|,G|downstream_gene_variant|MODIFIER|PF3D7_010
7100|PF3D7_0107100|transcript|rna_PF3D7_0107100-
1|protein_coding||c.*3762T>C|||||3762|  GT:DP:PL:GQ   0/1:94:184,0,255:99
      0/1:109:246,0,255:99      0/1:311:255,0,255:99      0/1:165:235,0,255:99
      0/1:329:255,0,255:99      0/1:221:255,0,255:99      0/1:178:255,0,255:99
      0/1:81:198,0,255:99 0/1:113:225,0,255:99
```

7.3.10 Comparison

This analysis aimed to search for SNPs that change as the resistance phenotype developed under sequential *in vivo* passage. Figure 7.15 shows interesting SNP trends as the resistance phenotype changes. Searching using this criterion (Table 7.1) returned 18 SNPs that developed as the resistance phenotype evolved and became fixed in the resistance isolates, and have predicted SNP effects. Eight SNPs were within the coding regions and have high, moderate, or low effects on the gene product as predicted by SNP effect prediction, SNPEFF software. The other 10 SNPs were within the genes but not in the coding region and have modifier effects as predicted by SNPEFF (Cingolani et al., 2012) (Table 7.8).

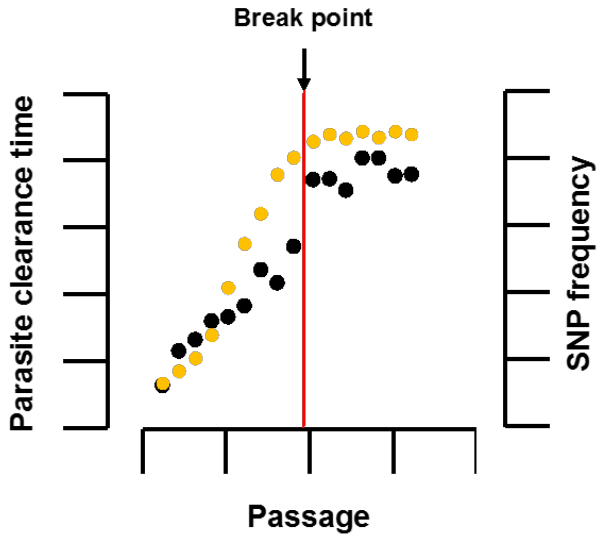


Figure 7.15 Interested pattern of SNP change over phenotype. The yellow dots represent parasite clearance half-life of each passage, whereas black dots indicate SNP frequency in corresponding passage. The resistance break point is shown in vertical red line.

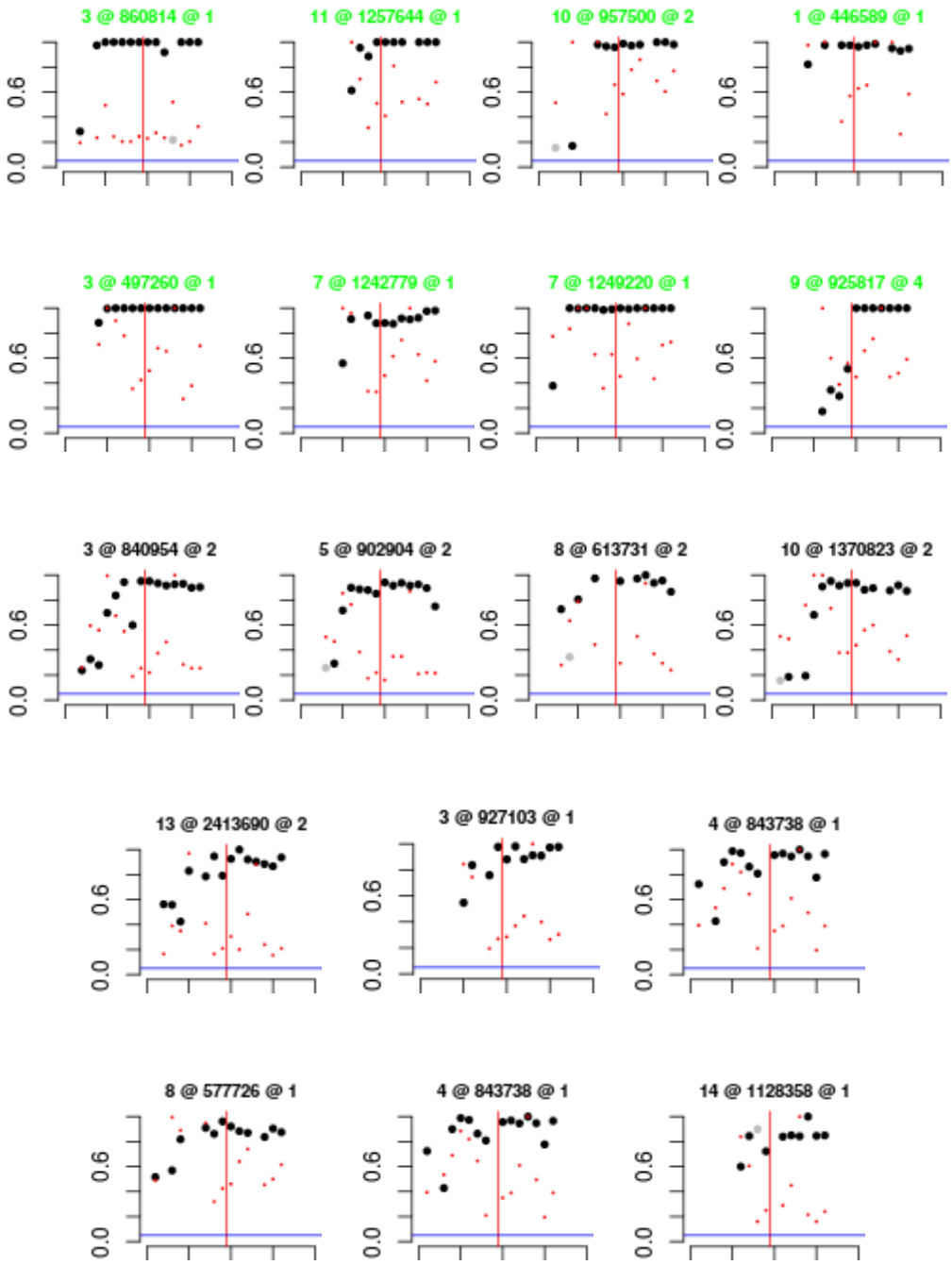


Figure 7.16 SNP readouts from comparison analysis.

*Table 7.8 Gene with SNPs developed with resistance phenotype in studied isolates. The gene in **bold** was identified from the proteomics experiments.*

			SNP Effect Prediction					
Gene ID	**UniProt ID**	**Product Description**	**Total**	**High**	**Moderate**	**Low**	**Modifier**	**Genomic Location**
PF3D7_0111800	B9ZSJ0	eukaryotic translation initiation factor 4E, putative	1	-	1	-	-	Pf3D7_01_v3: 446,026 - 448,849 (+)
PF3D7_0311500	Q9NFE5	conserved Plasmodium protein, unknown function	1	-	1	-	-	Pf3D7_03_v3: 497,118 - 497,363 (-)
PF3D7_0320100	O97323	protein transport protein SEC22 (SEC22)	1	-	-	-	1	Pf3D7_03_v3: 840,362 - 841,758 (+)
PF3D7_0320500	O97284	nicotinamidase, putative (Nico)	1	1	-	-	-	Pf3D7_03_v3: 860,813 - 862,102 (+)
PF3D7_0321900	C0H483	cyclic amine resistance locus protein (CARL)	3	-	-	-	3	Pf3D7_03_v3: 922,974 - 927,942 (+)
PF3D7_0418800	C0H4B4	conserved Plasmodium protein, unknown function	1	-	-	-	1	Pf3D7_04_v3: 843,562 - 845,169 (+)
PF3D7_0522300	C0H4F7	S-adenosylmethionine-dependent methyltransferase, putative	1	-	-	-	1	Pf3D7_05_v3: 902,613 - 904,844 (-)
PF3D7_0729100	C0H4Q1	conserved Plasmodium protein, unknown function	1	-	1	-	-	Pf3D7_07_v3: 1,238,446 - 1,246,188 (+)
PF3D7_0729300	Q8IBG6	60S ribosomal export protein NMD3, putative (NMD3)	1	-	1	-	-	Pf3D7_07_v3: 1,248,837 - 1,251,503 (-)
PF3D7_0811400	C0H4U0	conserved protein, unknown function	2	-	-	-	2	Pf3D7_08_v3: 575,986 - 578,370 (+)

			SNP Effect Prediction					
Gene ID	**UniProt ID**	**Product Description**	**Total**	**High**	**Moderate**	**Low**	**Modifier**	**Genomic Location**
PF3D7_0812220	A0A143ZZW7	GTPase, putative	1	-	-	-	1	Pf3D7_08_v3: 613,325 - 614,862 (-)
PF3D7_0919400	**Q8I2V9**	**protein disulfide isomerase (PDI9)**	**1**	**-**	**-**	**-**	**1**	**Pf3D7_09_v3: 795,539 - 797,470 (+)**
PF3D7_0922800	C0H552	conserved Plasmodium protein, unknown function	1	-	-	1	-	Pf3D7_09_v3: 922,292 - 935,372 (-)
PF3D7_1022800	Q8IJH7	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (GcpE)	1	-	1	-	-	Pf3D7_10_v3: 955,883 - 958,357 (+)
PF3D7_1034500	Q8IJ65	conserved Plasmodium protein, unknown function	1	-	-	-	1	Pf3D7_10_v3: 1,370,731 - 1,376,131 (+)
PF3D7_1132400	Q8II41	conserved Plasmodium membrane protein, unknown function	1	1	-	-	-	Pf3D7_11_v3: 1,255,286 - 1,259,671 (-)
PF3D7_1360500	Q8IDA0	guanylyl cyclase beta (GCbeta)	1	-	-	-	1	Pf3D7_13_v3: 2,413,324 - 2,425,261 (+)
PF3D7_1428600	Q8ILH8	peptide chain release factor 1, putative	1	-	-	-	1	Pf3D7_14_v3: 1,127,042 - 1,129,031 (+)
		Total	**21**	**2**	**5**	**1**	**13**	

Box 7.8 VCF result for SNP 957500 on chromosome 10

```
Pf3D7_10_v3   957500 .      GATAATAATAATAATAATAATAATAATAATA
       GATAATAATAATAATAATAATAATAATAATA,GATAATAATAATAATAATAATAATA    143.95 .
       AC1=1;AC=7,1;AF1=0.5;AN=10;DP4=96,85,168,168;DP=756;FQ=5.09;INDEL;IS=7,0.0679
61;MQ=42;PV4=0.34,0.092,0.00036,0.15;SF=0,1;VDB=2.553269e-
01;ANN=GATAATAATAATAATAATAATAATAATAATAATA|disruptive_inframe_insertion|MODERATE|PF3
D7_1022800|PF3D7_1022800|transcript|rna_PF3D7_1022800-
1|protein_coding|1/1|c.1646_1648dupATA|p.Asn549dup|1649/2475|1649/2475|550/824||,GA
TAATAATAATAATAATAATAATAATA|disruptive_inframe_deletion|MODERATE|PF3D7_1022800|PF3D7
_1022800|transcript|rna_PF3D7_1022800-
1|protein_coding|1/1|c.1646_1648delATA|p.Asn549del|1646/2475|1646/2475|549/824||,GA
TAATAATAATAATAATAATAATAATA|upstream_gene_variant|MODIFIER|PF3D7_1022700|PF3D7_10227
00|transcript|rna_PF3D7_1022700-1|protein_coding||c.-4951_-
4949delTAT|||||4949|,GATAATAATAATAATAATAATAATAATAATA|upstream_gene_variant|MODIF
IER|PF3D7_1022700|PF3D7_1022700|transcript|rna_PF3D7_1022700-1|protein_coding||c.-
4952_-
4951insTAT|||||4952|,GATAATAATAATAATAATAATAATA|downstream_gene_variant|MODIFIER|
PF3D7_1022900|PF3D7_1022900|transcript|rna_PF3D7_1022900-
1|protein_coding||c.*1409_*1411delTAT|||||1411|,GATAATAATAATAATAATAATAATAATAATA|
downstream_gene_variant|MODIFIER|PF3D7_1022900|PF3D7_1022900|transcript|rna_PF3D7_1
022900-
1|protein_coding||c.*1408_*1409insTAT|||||1408|,GATAATAATAATAATAATAATAATA|downst
ream_gene_variant|MODIFIER|PF3D7_1023000|PF3D7_1023000|transcript|rna_PF3D7_1023000
-
1|protein_coding||c.*3536_*3538delTAT|||||3538|,GATAATAATAATAATAATAATAATAATAATA|
downstream_gene_variant|MODIFIER|PF3D7_1023000|PF3D7_1023000|transcript|rna_PF3D7_1
023000-1|protein_coding||c.*3535_*3536insTAT|||||3535|      GT:DP:PL:GQ   .
       0/2:71:38,.,.,0,.,255:37   .      0/1:135:112,0,255,.,.,.:99 .      .
       1/1:155:255,255,0,.,.,.:99 1/1:59:255,91,0,.,.,.:99
       1/1:97:255,139,0,.,.,.:99
```

Box 7.9 VCF result for SNP 1257644 on chromosome 11

```
Pf3D7_11_v3   1257644       .      CTTTTTTTTTTT  CTTTTTTTTT    122.15 .
       AC1=2;AC=8;AF1=1;AN=8;DP4=41,34,153,191;DP=508;FQ=-
39.5;INDEL;IS=83,0.410891;MQ=41;PV4=0.15,1,0.3,0.21;SF=0,1;VDB=7.086270e-
02;ANN=CTTTTTTTTT|frameshift_variant|HIGH|PF3D7_1132400|PF3D7_1132400|transcript|rn
a_PF3D7_1132400-
1|protein_coding|1/1|c.2017_2018delAA|p.Lys676fs|2018/4386|2017/4386|673/1461||,CTT
TTTTTTT|upstream_gene_variant|MODIFIER|PF3D7_1132500|PF3D7_1132500|transcript|rna_P
F3D7_1132500-1|protein_coding||c.-4105_-
4104delTT|||||4105|,CTTTTTTTTT|downstream_gene_variant|MODIFIER|PF3D7_1132300|PF3D7
_1132300|transcript|rna_PF3D7_1132300-
1|protein_coding||c.*3760_*3761delTT|||||3760|;LOF=(PF3D7_1132400|PF3D7_1132400|1|1
.00)   GT:GQ:DP:PL   .      .      .      .      .      1/1:9:166:124,5,0
       1/1:99:110:188,247,0       1/1:99:53:147,57,0  1/1:99:90:191,255,0
```

Box 7.10 VCF result for SNP 895153 on chromosome 7

```
Pf3D7_07_v3   895153 .      GATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA
      GATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 154.28 .
      AC1=2;AC=10;AF1=1;AN=10;DP4=0,0,20,11;DP=832;FQ=-
46.5;INDEL;IS=2,0.012987;MQ=42;SF=0,1;VDB=5.749530e-
02;ANN=GATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA|disruptive_inframe_del
etion|MODERATE|PF3D7_0720700|PF3D7_0720700|transcript|rna_PF3D7_0720700-
1|protein_coding|3/4|c.2997_2999delTAA|p.Asn1000del|2997/6501|2997/6501|999/2166||,
GATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA|upstream_gene_variant|MODIFIE
R|PF3D7_0720800|PF3D7_0720800|transcript|rna_PF3D7_0720800-1|protein_coding||c.-
4545_-4543delTAA|||||4545| GT:PL:DP:GQ   1/1:146,12,0:4:21   1/1:119,9,0:3:16
      1/1:217,24,0:8:45   1/1:245,24,0:8:45   .      1/1:247,24,0:8:45   .      .
      .
```

Box 7.11 VCF result for SNP 446589 on chromosome 1

```
Pf3D7_01_v3   446589 .      CAATAATAATAATAATAATAATAATAATAATA
      CAATAATAATAATAATAATAATAATAATAATAATA     214.75 .
      AC1=1;AC=7;AF1=0.5;AN=8;DP4=10,15,163,166;DP=606;FQ=126;INDEL;IS=74,0.379487;
MQ=41;PV4=0.8,1,1,0.0033;SF=0,1;VDB=1.023815e-
01;ANN=CAATAATAATAATAATAATAATAATAATAATAATA|disruptive_inframe_insertion|MODERATE|PF
3D7_0111800|PF3D7_0111800|transcript|rna_PF3D7_0111800-
1|protein_coding|1/8|c.593_595dupATA|p.Asn198dup|596/1863|596/1863|199/620||,CAATAA
TAATAATAATAATAATAATAATAATAATA|upstream_gene_variant|MODIFIER|PF3D7_0111600|PF3D7_01
11600|transcript|rna_PF3D7_0111600-1|protein_coding||c.-3332_-
3331insTAT|||||3332|,CAATAATAATAATAATAATAATAATAATAATAATA|upstream_gene_variant|MODI
FIER|PF3D7_0111700|PF3D7_0111700|transcript|rna_PF3D7_0111700-1|protein_coding||c.-
2683_-
2682insTAT|||||2683|,CAATAATAATAATAATAATAATAATAATAATAATA|downstream_gene_variant|MO
DIFIER|PF3D7_0111900|PF3D7_0111900|transcript|rna_PF3D7_0111900-
1|protein_coding||c.*3422_*3423insTAT|||||3422|,CAATAATAATAATAATAATAATAATAATAATAATA
|downstream_gene_variant|MODIFIER|PF3D7_0112000|PF3D7_0112000|transcript|rna_PF3D7_
0112000-1|protein_coding||c.*4210_*4211insTAT|||||4210|     GT:GQ:PL:DP   .      .
      .      0/1:99:255,0,161:107       .      1/1:99:255,255,0:128       .
      1/1:99:255,84,0:42  1/1:99:255,149,0:77
```

Box 7.12 VCF result for SNP 497260 on chromosome 3

```
Pf3D7_03_v3   497260 .      C      G      218.67 .
      AC1=2;AC=12;AF1=1;AN=12;DP4=7,10,509,349;DP=919;FQ=-
65;MQ=33;PV4=0.43,0.21,0.14,0.053;RPB=-1.030801e-01;SF=0,1;VDB=5.187392e-
01;ANN=G|missense_variant|MODERATE|PF3D7_0311500|PF3D7_0311500|transcript|rna_PF3D7
_0311500-
1|protein_coding|1/1|c.104G>C|p.Gly35Ala|104/246|104/246|35/81||,G|upstream_gene_va
riant|MODIFIER|PF3D7_0311600|PF3D7_0311600|transcript|rna_PF3D7_0311600-
1|protein_coding||c.-
2342C>G|||||2342|,G|downstream_gene_variant|MODIFIER|PF3D7_0311400|PF3D7_0311400|tr
anscript|rna_PF3D7_0311400-1|protein_coding||c.*750C>G|||||750|    GT:DP:PL:GQ   .
      .      .      1/1:138:235,38,0:73 1/1:271:255,255,0:99
      1/1:169:255,255,0:99       1/1:149:255,255,0:99       1/1:68:255,205,0:99
      1/1:80:255,241,0:99
```

Box 7.13 VCF result for SNP 925818 on chromosome 9

```
Pf3D7_09_v3   925817 .      T      A      174.75 .
      AC1=1;AC=4;AF1=0.5;AN=8;DP4=207,147,79,75;DP=526;FQ=40;MQ=42;PV4=0.35,1,0.09,
1;RPB=-5.044023e-01;SF=0,1;VDB=3.139301e-
01;ANN=A|synonymous_variant|LOW|PF3D7_0922800|PF3D7_0922800|transcript|rna_PF3D7_09
22800-
1|protein_coding|3/3|c.9243A>T|p.Val3081Val|9243/12768|9243/12768|3081/4255||,A|ups
tream_gene_variant|MODIFIER|PF3D7_0922700|PF3D7_0922700|transcript|rna_PF3D7_092270
0-1|protein_coding||c.-4515A>T|||||4515|       GT:PL:DP:GQ   .      .      .      .
      .      0/1:67,0,255:212:70 0/1:255,0,255:113:99       0/1:242,0,255:74:99
      0/1:255,0,255:109:99
```

Box 7.14 VCF result for 860814 on chromosome 3

```
Pf3D7_03_v3   860814 .      T      C      202.0  .
      AC1=1;AC=13;AF1=0.5;AN=14;DP4=9,17,146,112;DP=365;FQ=92;MQ=39;PV4=0.71,4.4e-
05,1,1;RPB=-7.850690e-01;SF=0,1;VDB=1.122838e-
01;ANN=C|start_lost|HIGH|PF3D7_0320500|PF3D7_0320500|transcript|rna_PF3D7_0320500-
1|protein_coding|1/1|c.2T>C|p.Met1?|2/1290|2/1290|1/429||,C|upstream_gene_variant|M
ODIFIER|PF3D7_0320400|PF3D7_0320400|transcript|rna_PF3D7_0320400-
1|protein_coding||c.-
3310A>G|||||3310|,C|upstream_gene_variant|MODIFIER|PF3D7_0320700|PF3D7_0320700|tran
script|rna_PF3D7_0320700-1|protein_coding||c.-
4200T>C|||||4200|,C|downstream_gene_variant|MODIFIER|PF3D7_0320600|PF3D7_0320600|tr
anscript|rna_PF3D7_0320600-
1|protein_coding||c.*1927A>G|||||1927|;LOF=(PF3D7_0320500|PF3D7_0320500|1|1.00)
      GT:GQ:PL:DP   .      0/1:99:119,0,255:35 .      1/1:99:255,91,0:40
      1/1:99:255,217,0:72 1/1:99:255,114,0:38 1/1:99:248,87,0:29
      1/1:99:255,96,0:33  1/1:99:255,111,0:37
```

7.4 Discussion

7.4.1 Sequence quality

The overall quality of the sequences was acceptable for further analysis. The average read quality score was all over 30, read length was between 80-100 bp (section A3.1). Read mapping used in the study systematically reduced the errors and complications from *de novo* genome assembly. Even though PCR-free library preparation is ideal for genomes that have AT-rich to reduce PCR associated errors, PCR-free DNA library preparation kit was not available at the time of experiment. Read and mapping coverage for each sample was considered to be in the good range between 66-182% (Table A3.2) and 85-98% (Table 7.3 and section A3.3), respectively.

7.4.2 Global analysis of SNPs

SNPs didn't show any selection toward any particular chromosome but a more random distribution (Figure 7.8 and Figure 7.9). Proteolysis and ubiquitin-dependent protein catabolic process were enriched from the 733 proteins (Table 7.5, Figure 7.10 and Figure 7.11). These systems are involved in the later stages of the unfolded protein response (UPR). It was reported that upregulation in the UPR system is associated with resistance to artemisinin (Mok et al., 2015). It has been proposed that resistance to

artemisinin is due to the parasites' ability to overcome the effect of stress generated by artemisinin via the UPR system (Mok et al., 2015, Paloque et al., 2016, Dogovski et al., 2015), together with delays in the erythrocytic developmental cycle (Hott et al., 2015, Teuscher et al., 2010). The result shown here suggests that there is a possible selective pressure on the UPR and ubiquitin-proteasome system under artesunate pressure *in vivo* and supports the current working hypothesis. The comparative proteomic results from ring and trophozoite stage parasites labelled with artemisinin activity-based probes in earlier chapters also supported the UPR as a potential resistance mechanism (section 5.4.1 and 5.4.3). Other enriched biological processes were gene expression, transport, and cell communication (Table 7.5 and Figure 7.10).

Interestingly, there seemed to be a selection pressure toward asparagine rich proteins. Nearly 50% of SNP associated genes were asparagine rich proteins (Table 7.9), although the organism has remarkably ~30% asparagine repeats in its proteome (Muralidharan and Goldberg, 2013). This unusual protein feature is evolutionarily conserved in the *Plasmodium falciparum* and *Plasmodium reichenowi* genomes but apparently not in other *Plasmodium* species, although some are as AT-rich (Muralidharan and Goldberg, 2013). The *Plasmodium falciparum* genome is known for its AT-rich hallmark and that increases the chance of tri-nucleotide repeats, especially AAT (coding for asparagine) (Aravind et al., 2003, Muralidharan and Goldberg, 2013). Even though the benefits of these repeats are uncertain the evolutionary conservation of the feature implies it is of biological importance and confers some benefit. One hypothesis is that asparagine repeats can lead to new protein domains or functions under heat shock proteins and stress responses (Muralidharan and Goldberg, 2013). Generally, proteins with repeated sequences are prone to aggregation and loss of function, but that was not the case for *Plasmodium falciparum* asparagine-rich proteins, at least with *Pf*CDK2 (PF3D7_0923500); it was shown that *Pf*CDK2 remains functional in the presence of *Plasmodium falciparum* heat shock protein 110 (*Pf*HSP110), but not in the *Pf*HSP110 knockdown parasite (Muralidharan et al., 2012). Our limited knowledge on this particular area is intriguing offers a possible link with drug resistance mechanisms.

Table 7.9 Number of genes with various asparagine runs (from 743 genes in comparison analysis)

Number of asparagine runs	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Number of genes	375	244	188	164	142	121	96	80

7.4.3 Unique post-break point

The unique post-break point analysis, the most stringent analysis, returned only one SNP within the MYND finger protein, putative gene region (PF3D7_0607100). This SNP is not presented in the 3D7 reference genome or any of the susceptible isolates and emerged straight after the break point (passage 9), although not in the coding region. Other unique post-break point SNPs were intergenic regions,

therefore is concluded that they should have no direct effect on product proteins that could confer resistance.

These SNPs seemed to be tandem repeat (TR) and indel variations which are the most common variations in the *Plasmodium falciparum* genome (Miles et al., 2016). These variations are possibly involved in the epigenetic regulation; however, there is no clear established mechanism of drug resistance associated with epigenetic regulation in the *Plasmodium falciparum*.

7.4.4 Common-to-all

In this section SNPs that appeared commonly in all study isolates but were different from the 3D7 reference strain from common-to-all analysis are discussed (Table 7.7). These SNPs are unlikely to contribute to resistance to artemisinin phenotype observed in the mouse model but possibly benefit the parasites' survival in the humanised SCID model.

7.4.4.1 SNPs in erythrocyte membrane protein 1 (*Pf*EMP1) gene

*Pf*EMP1 is a protein encoded by the multigene *var* family, so far there are 62 *Pf*EMP1 encoded genes reported in PlasmoDB, including *var*, and *var*-like. *Pf*EMP1 proteins play a role in cytoadhesion of infected RBCs to human endothelial cells. This phenomenon is linked with parasite pathogenicity and major clinical manifestations. Only single *Pf*EMP1 gene is expressed by each infected RBC at any given time. Genes in the *Pf*EMP1 family exhibit high polymorphism which provides parasites the tools to fight and/or evade host immunity. Together with clonal variation switching, *Pf*EMP1 expression patterns are numerous.

Genes encoding for *Pf*EMP1 are structured in a 2-exon structure. Exon 1 encodes for the external part of *Pf*EMP1 while exon 2 encodes for the transmembrane region and internal parts. Exon 2 sequences are extensively conserved among the parasites as this part is not beneficial in terms of parasite pathogenicity but rather provides the membrane anchoring structure for the protein. Unlike exon 2, exon 1 sequences are highly diverse, and its products contribute to cytoadhesion profile of parasites.

Then general structure of the *Pf*EMP1 protein is shown in Figure 7.17. Intercellular components of *Pf*EMP1 contains an acidic terminal sequence (ATS) which is believed to act as an anchoring system for *Pf*EMP1. However, Mayer et al. (2012) suggests that ATS associated with PHIST-type domain (PF3D7_0936800) might act in the parasites intra-erythrocyte protein network. Although *Pf*EMP1 is part of parasite's knob structure, there is no evidence showing interactions with knob-associated histidine-rich protein (KAHPR), an important protein for knob formation (Watermeyer et al., 2016, Hviid and

Jensen, 2015). The ATS is connected with a transmembrane region, thrombomodulin (TM), which links it to the extracellular segment.

The extracellular segment of *Pf*EMP1 is highly diverse. The N-terminus segment is located at the end of the extracellular segment followed by a duffy binding-like domain (DBL), specifically DBL1α, and a cysteine-rich interdomain region (CIDR) followed by various DBL domains (Hviid and Jensen, 2015).

Five SNPs were identified from the *Pf*EMP1 gene, PF3D7_0632500, at positions 1365389, 1365409, 1365410, 1365411, and 1365413. All these SNPs' positions are located in the exon 1 region (1356045 to 1366430). This finding was not surprising as the gene has a high number of SNPs in this area (PlasmoDB recorded 4238 SNPs). This increases diversity in the final protein product and enhances the chance to evade host recognition. Although *Pf*EMP1 has high sequence diversity, there is evidence that it exhibits structural conservation to maintain its interaction with partner ligands, e.g. CIDRα1:EPCR (endothelial protein C receptor) (Lau et al., 2015).

Figure 7.17 PfEMP1 protein general structure

7.4.4.2 SNPs in rifin gene (PF3D7_1040600)

Rifin is a parasite protein that modifies the host RBC and might play a role in parasite-host interactions e.g. binding to host cells (Kyes et al., 1999, Maier et al., 2009). The protein rifin is encoded by a *var* multigene family *rif*, like *Pf*EMP1, that comprises the largest copy in the *Plasmodium* genome (Joannin et al., 2008). PlasmoDB searches returned 157 *rif* genes, excluding pseudogenes. Rifin is not as widely studied as *Pf*EMP1 but it is likely that rifin has a similar role in parasite pathogenicity based on its presentation on host RBC surface, and the number of genes in the family. Rifin protein are part of knob structures of infected RBC but there is no evidence that the rifin protein interacts directly with *Pf*EMP1.

7.4.4.3 SNP in IspD gene (PF3D7_0106900)

IspD is an enzyme in non-mevalonate pathway in the *Plasmodium* parasite catalysing the production of 4-(Cytidine 5'-diphospho)-2-C-methyl-D-erythritol. The enzyme was reported as a therapeutic target (Wu et al., 2015, Imlay et al., 2015). MMV-08138 was shown to directly inhibit IspD by competitive binding with cytidine triphosphate (CTP) (Imlay et al., 2015), furthermore mutations in IspD gene rendered parasite resistance to the inhibitor (Wu et al., 2015). There is no evidence showing that mutations in the IspD gene are associated with artemisinin resistance. According to the predicted phenotype by SNPEFF

software, SNP at position 291,538 was a missense variant, 1153 A→G (Asn385Asp). The result presented here is also not suggestive of a resistance determinant as this SNP presented in all studied isolates regardless of the phenotypes.

The relevance of the changes highlighted by this type of analysis are probably more linked to 3D7 adaptations to stable infection in the SCID mouse model and should be investigated with that in mind.

7.4.5 Comparison

In this section SNPs with predicted SNP effects will be discussed in more detail with the aim of trying to identify plausible links with a mechanism of resistance to artesunate. These mutations were apparently fixed in the resistant parasite isolates post break.

7.4.5.1 SNP in IspG gene (PF3D7_1022800)

SNP in gene encoding for 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, IspG or GcpE (PF3D7_1022800), showed development in accordance with resistance phenotype and were fixed after the post-break point (Figure 7.18).

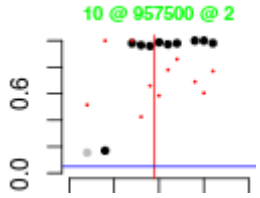


Figure 7.18 SNP readout of position 957500 on chromosome 10 (PF3D7_1022800)

IspG is the penultimate enzyme in non-mevalonate isoprenoid synthetic pathway in the apicoplast, non-photosynthetic plastid (Hecht et al., 2001). IspG catalyses 2C-methyl-D-erythritol 2,4-cyclodiphosphate conversion into 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate which in turn is converted to isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) by 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, IspH, is the last enzyme in the pathway (Figure 7.19). Structural studies in bacterial IspG and IspH proteins revealed both IspG and IspH have 2 domains, a TIM barrel domain (A) and 4Fe-4S cluster domain (B). The TIM barrel domain (A) acts as a diphosphate binding site while the 4Fe-4S cluster domain (B) is involved in reduction (Lee et al., 2010, Guerra et al., 2014).

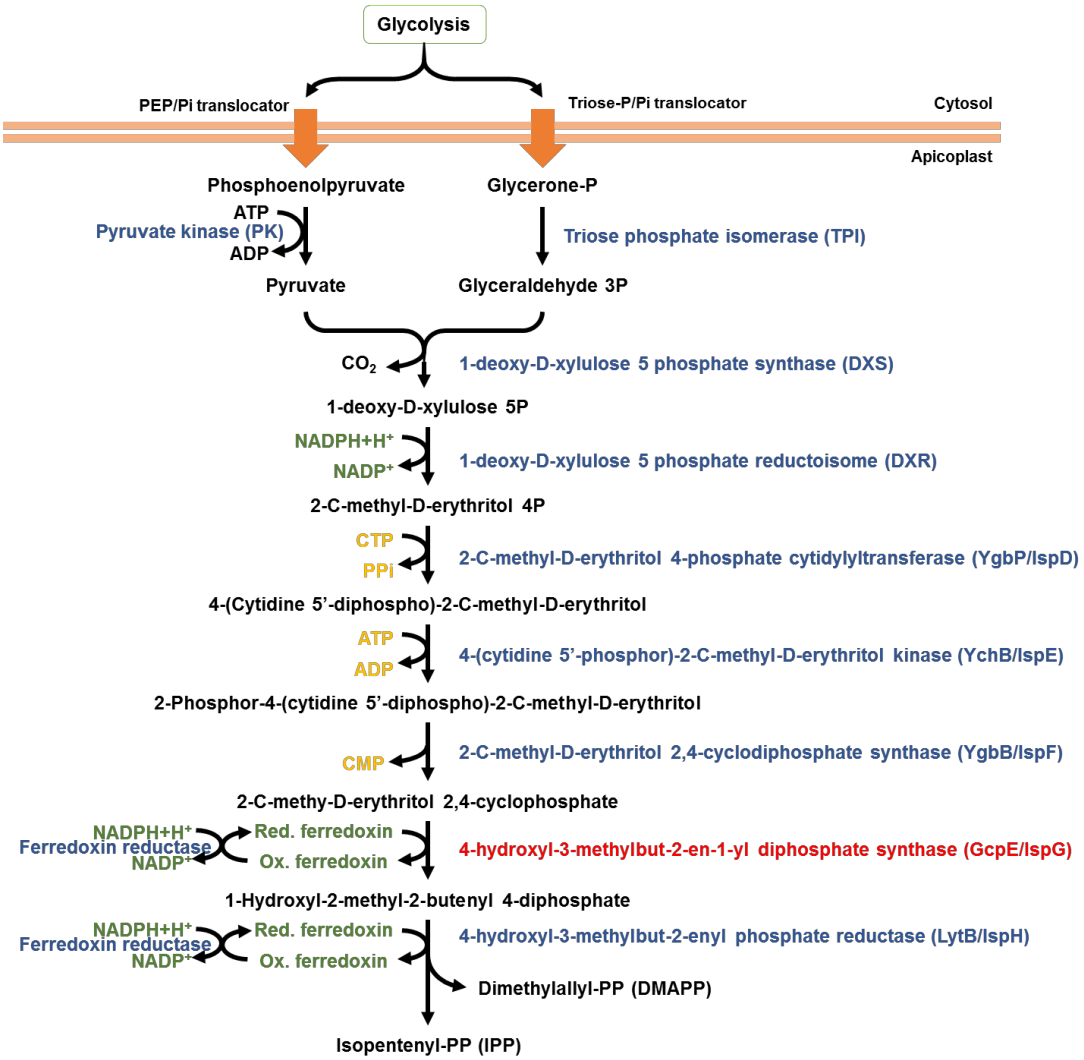


Figure 7.19 Non-mevalonate (MEP) isoprenoid biosynthesis pathway in Plasmodium falciparum (adapted from MPMP database). Protein indicated in red harbours SNP associated with artesunate resistance.

Sequence analysis of the *Plasmodium falciparum* IspG sequence and other IspG sequences, including plant *Arabidopsis thaliana*, bacteria *Escherichia coli*, *Aquifex aeolicus*, and *Bacillus anthracis*, revealed that plants and parasite have an additional sequence (A*) of around 400-500 amino acids between TIM barrel domain (A) and 4Fe-4S cluster domain (B) (Figure 7.20). However, the additional sequence (A*) is less conserved suggesting more of structural role rather than catalytic role (section A3.6) (Liu et al., 2012).

Figure 7.20 Sequence alignment of P. falciparum, A. thaliana, A aeolicus, B. anthracis, and E. coli

In order to form an active enzyme, bacterial IspGs forms $(AB)_2$ homodimer structure from 2 TIM domains and 2 4Fe-4S cluster domains. The TIM barrel domain of one monomer forms an active site with the

4Fe-4S cluster of another monomer, while plant and plasmodial IspG have the addition domain (A*) and function as 3-domain monomer (Figure 7.21) (Liu et al., 2012).

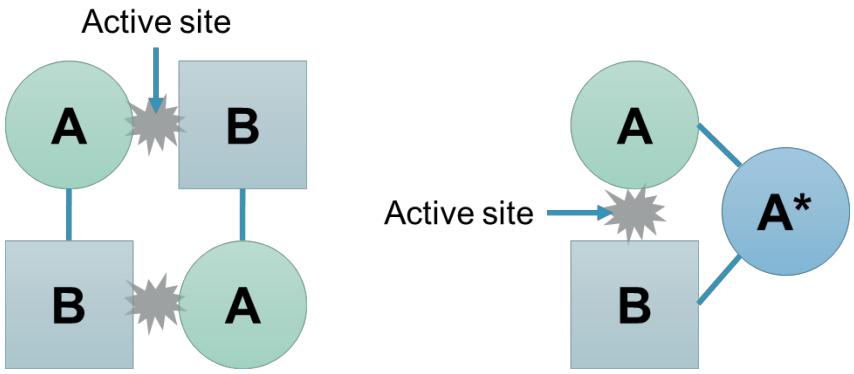


Figure 7.21 Diagram showing $(AB)_2$ homodimer structure of bacterial IspG (left) and proposed plasmodial IspG monomer (right).

IspG utilises reduced ferredoxin in order to catalyse conversion of 2C-methyl-D-erythritol 2,4-cyclodiphosphate into 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate, as does IspH. Ferredoxin is coupled with ferredoxin-$NADP^+$ reductase (FNR), or FNR/fd, and serves as electron transport and/or redox system for the apicoplast. A recent study shows that SNPs in ferredoxin, apicoplast ribosomal protein S10 (arps10), multidrug resistance protein 2 (mdr2), and chloroquine resistance transporter (crt) are strongly associated with artemisinin resistance (Miotto et al., 2015). As these two proteins, IspG and ferredoxin, are closely linked together. Mutations in these genes might confer the delayed parasite clearance phenotype but a mechanism for this is unclear.

IspG interacts with Kae1 (kinase-associated endopeptidase 1) and might be involved in transcriptional regulation (*Chlamydia trachomatis*). Work by Mallari et al. (2014) has demonstrated that Kae1 is expressed in the cytoplasm and apicoplast of *Plasmodium falciparum* and interacts with IspG, EF-Ts, and RNA helicase in apicoplast. These protein complexes further interact with other proteins in the apicoplast and might (control or) play a role in protein translation in the apicoplast. Immuoprecipitation experiment also shows that the Kae1-EF-Ts protein complex interacts with arps10, another SNP showing strong association with artemisinin resistance (Miotto et al., 2015).

SNPs at position 957500 in the IspG gene were classified as disruptive in frame insertion/deletions according to SNP effect prediction by SNPEFF software (Box 7.8). These mutations resulted in an insertion or deletion of the amino acid asparagine, N, at position 549 which was located within a repeated asparagine sequence in domain A* (Figure 7.22) and was not a known lethal mutation. The *Plasmodium falciparum* genome is known for asparagine repeats, but the evolutional and functional purposes of this

asparagine repeats are still elusive (Muralidharan and Goldberg, 2013). However, the structural linker can play a role in maintaining protein symmetry which is important for protein function. For example, the study of P97 proteins revealed a linker between the P97 subunits is essential for catalytic activity of the enzyme. The truncated linkers resulted in reduced activity of ATPase activity (Tang and Xia, 2016). Although, it is possible that mutation in this position may contribute to catalytic activity of the enzyme, further specific experiment on this enzyme is needed to evaluate the protein function and resistance phenotype.

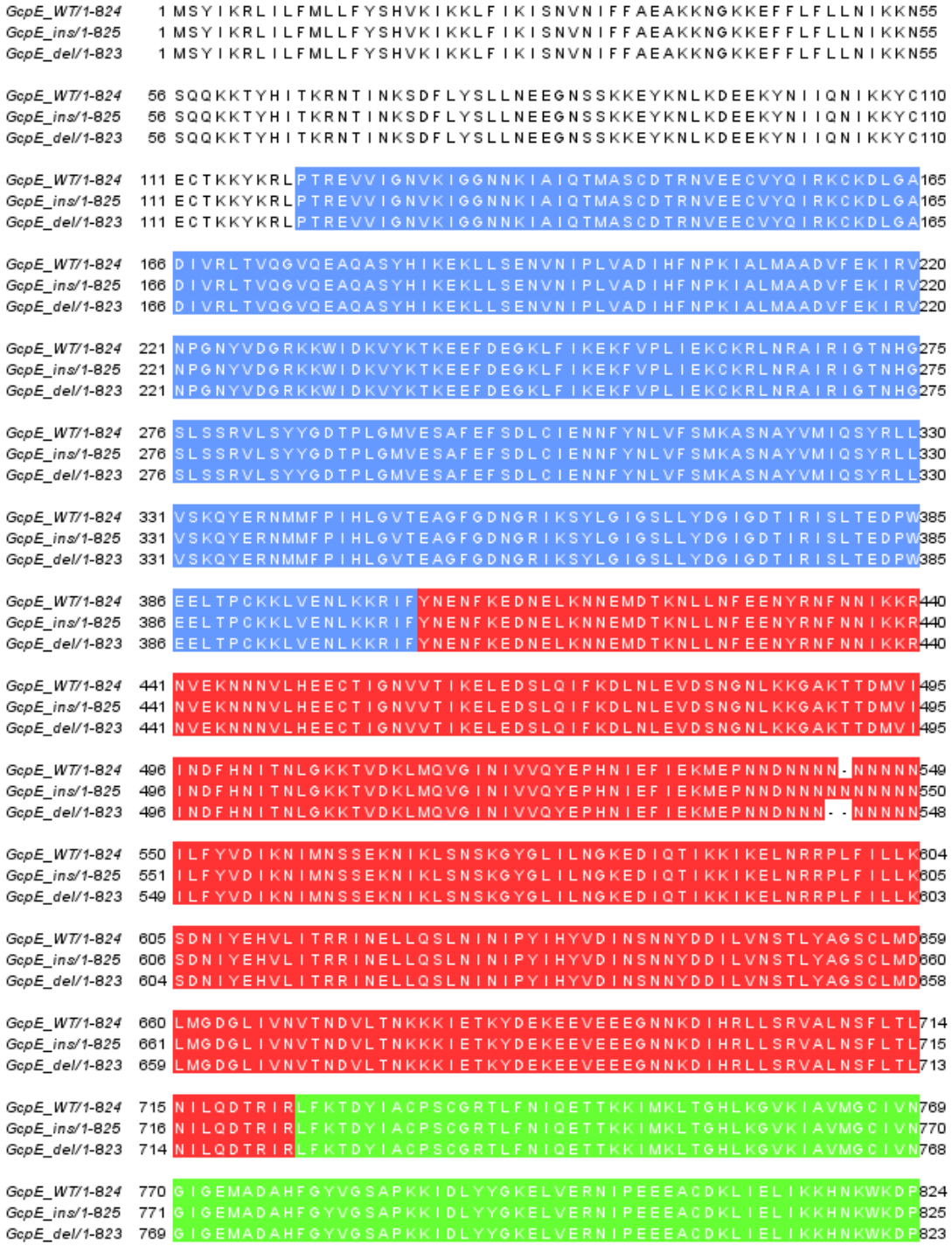


Figure 7.22 GcpE protein sequences. Wildtype (WT) on top, insertion (ins) sequence in the middle, and deletion (del) sequence on the bottom. Domain A is shown in blue, domain A in red, and domain B in green. Affected amino acid is located within low complexity region of the protein containing asparagine repeat.*

7.4.5.2 SNP in PF3D7_1132400 gene

Another SNP that developed with increased parasite clearance half-life and fixed with the resistance phenotype is the apicoplast localised conserved Plasmodium membrane protein, (PF3D7_1132400, UniProt Q8II41). This protein is curated as unknown function but InterPro domain predicts a chloroquine resistance transporter (CRT)-like domain in the gene product at amino acid 844-980. Mutation at this position resulted in a frame shift variant and generated a stop codon at position 681, therefore the CRT-like domain was lost. This variation was unlikely to be a sequencing error as it has very good depth coverage and read alignment (Box 7.8, Box 7.9, and Figure 7.24). The CRT-like domain in this protein is possibly involved in drug transportation across the apicoplast membrane, the absence of this domain might reduce the stress generated by artemisinin within apicoplast which is a critical organelle in the parasite. However, biological evidence of this mutant protein will be needed to provide more detail on the function of this protein.

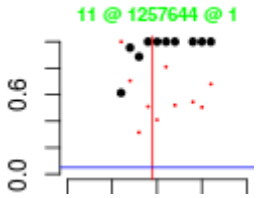


Figure 7.23 SNP readout of position 1257644 on chromosome 11 (PF3D7_1132400)

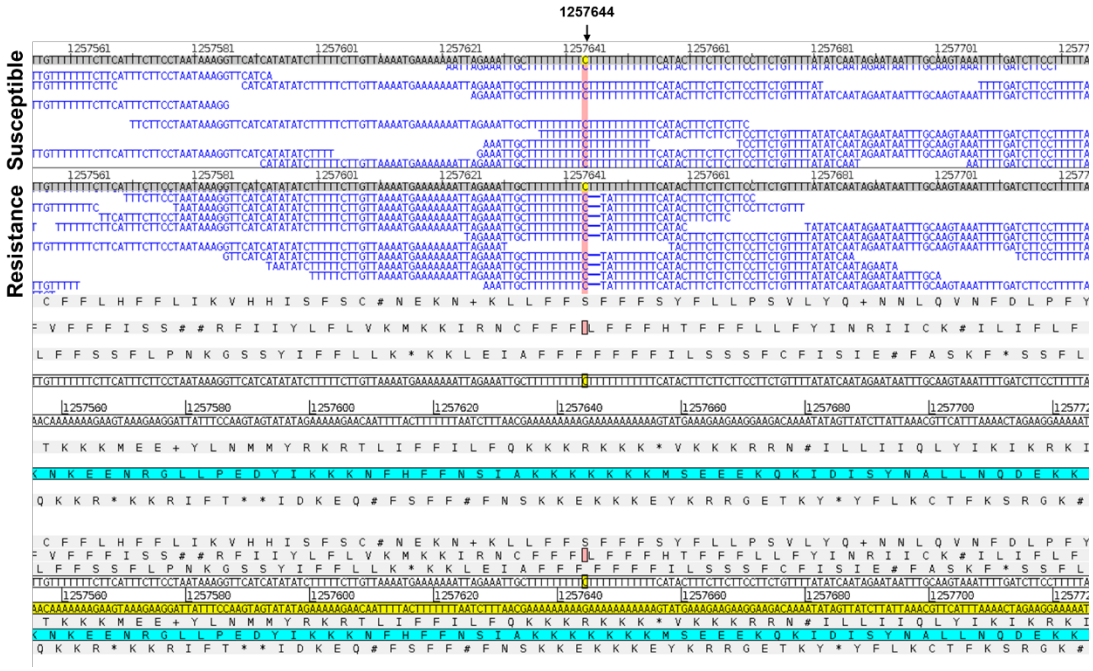


Figure 7.24 Reads alignment (BAM) of studied isolates at position 1257644 on chromosome 11.

7.4.5.3 SNP in eIF4e putative gene (PF3D7_0111800)

The SNP position 446,589 on chromosome 1 (Figure 7.25) belongs to the PF3D7_0111800 gene encoded for eukaryotic translation initiation factor 4E, putative (eIF4e, UniProt B9ZSJ0). SNPEFF predicted for a disruptive in frame insertion of asparagine at position 198 (Box 7.11) within the low complexity region (LCR) as identified by the SEG program (Wootton, 1994). WT protein contains a 10 asparagine repeat run in this region, while the mutation inserted 1 additional asparagine into this region. Similar to SNP position 957,500 on chromosome 10 encoded for the IspG protein, variations in low complexity regions of the protein possibly contribute to the response to antimalarial. It was evidence that the variations in the pattern of asparagine repeats in low complexity region of *Pf*MDR1, a prominent antimalarial resistance determinant, is associated with several antimalarial responses (Okombo et al., 2013). The parasite eIF4e plays a role in translation initiation steps with other eIF4s and are regulated via PI3K/AKT pathway (Sonenberg and Hinnebusch, 2009), which was proposed as a direct target of artemisinins (Mbengue et al., 2015). A more detailed evidence and enquiry is required to get a greater understanding of the relevance if this finding.

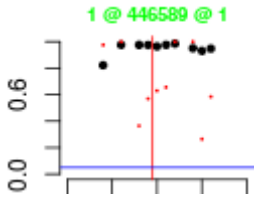


Figure 7.25 SNP readout of position 446589 on chromosome 1 (PF3D7_0111800)

7.4.5.4 SNP in PF3D7_0311500 gene

The SNP at position 497,260 on chromosome 3 resulted in nonsynonymous mutation, Gly35Ala, in PF3D7_0311500 encoded for the conserved *Plasmodium* protein of unknown function (Q9NFE5). The gene is conserved among *Plasmodium* species (~70-98% identity), and nearly 50% identical to *Toxoplasma* proteins. However, there is little known about this gene or its protein product. The transcript encodes for an 81 amino acid long protein with 2 transmembrane domains. Further experiments such as genetic modification of this gene and its association with artemisinin resistance is required.

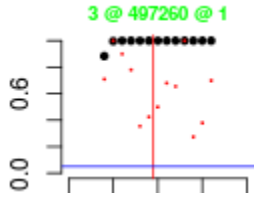


Figure 7.26 SNP readout of position 497260 on chromosome 3 (PF3D7_0311500)

7.4.5.5 SNP in PF3D7_0922800 gene

The SNP at position 925,817 on chromosome 9 is localised in PF3D7_0922800 encoded for another conserved *Plasmodium* protein, unknown function (C0H552). This SNP resulted in a synonymous mutation, Val3081Val, thus it has no biological implication to the protein sequence.

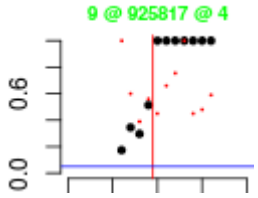


Figure 7.27 SNP readout of position 925817 on chromosome 9 (PF3D7_0922800)

7.4.5.6 SNP in the nicotinamidase gene (PF3D7_0320500)

The SNP at position 860,814 on chromosome 3 is in a nicotinamidase encoded gene (PF3D7_0320500, UniProt O97284). It has predicted with a high SNP effect causing loss of a start codon (Box 7.14). This gene orthologue in *Plasmodium berghei* ANKA strain (PBANKA_121800) is classified as likely to be dispensable (identified from PlasmoGEM database) (Gomes et al., 2015, Schwach et al., 2015), therefore the absence of this gene product is unlikely to affect parasite survival under normal conditions. This SNP appeared quite early in the parasite line (after the passage 6), midway before the defined artesunate resistance-break point (Figure 7.28). Nicotinamidase is a parasite enzyme involved in reversible catalytic process of nicotinic acid and nicotinamide, and it is therefore involved in redox metabolism in the parasite (Figure 7.29). The enzyme is absent in the host RBC.

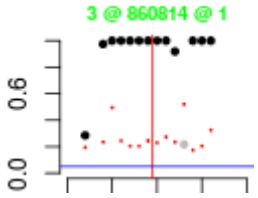


Figure 7.28 SNP readout of position 860814 on chromosome 3 (PF3D7_0320500)

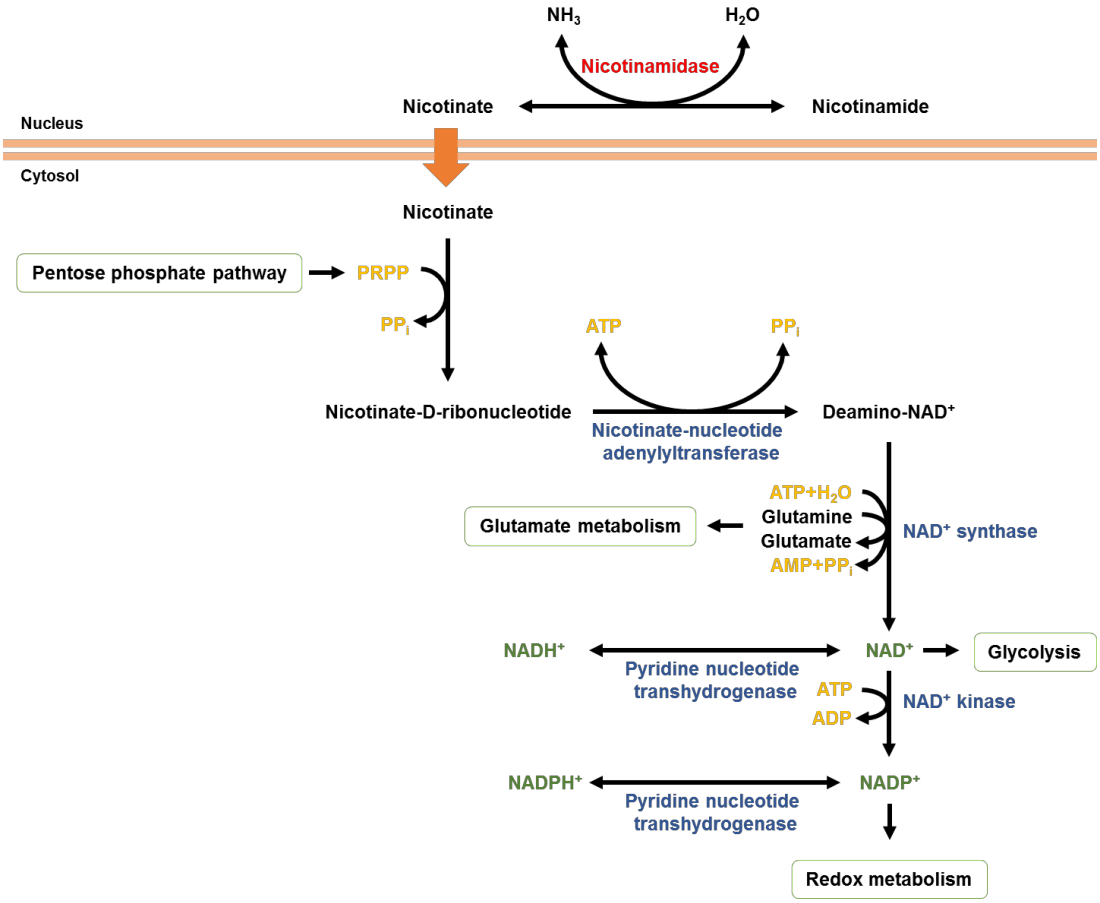


Figure 7.29 Nicotinate and nicotinamide metabolism in Plasmodium falciparum (adapted from MPMP database).

7.5 Conclusions

The mouse model system has long been used in the disease biology. In malaria research, mouse malaria, *Plasmodium berghei*, has sometimes been used as a representative of *Plasmodium falciparum*. In this study the humanised SCID mouse model and *Plasmodium falciparum* system were used to generate artesunate resistant parasites after clinical relevant artesunate monotherapy by repeated exposure to artesunate and serial passage over 23 generations. Although the system closely represented clinical exposures, the system is obviously not completely same as to human system. However, in this system artesunate rapidly reduced parasite biomass after administration of 10 mg/kg with a relatively constant parasite clearance rate. However, between parasite passage 8 and 9 this clearance rate reduced by more than half and this phenotype was maintained under exposures up to 50 mg/kg (the maximum dose investigated) and was stable in the absence of drug pressure. From an *in vivo* phenotypic perspective this looks very similar to artesunate resistance as it is seen in the clinical environment n S.E Asia.

Interestingly, the published artemisinin resistance marker kelch propeller (K13) protein (PF3D7_1343700) was not identified from the current studied population, as a determinant of the

phenotype. What this suggests is that there are more than one way in which the *in vivo* parasite response to artesunate exposure can be modified and highlights the need to look beyond K13 for resistance determinants in the clinical environment. Several interesting patters did emerge from this series of related parasite passage lines. Some of these were presumably linked to parasite adaptation to life within the SCID mouse whereas other presumably do help parasites tolerate artemisinins in some way.

The approach is not without its limitations. It is expensive to conduct and takes many months/years to complete and it is essentially one sequential experiment. Coupled with natural and random genome variations over time this makes interrogation of the data and final interpretation very difficult. However, having said that the experiment resulted in a very clear phenotype and our inability to link this conclusively with one or more key genomic changes is both intriguing and important message to the "omics" community.

Perhaps in future an alternative strategy will be required. One limitation in this study is the sample size so the study could not really be considered as a genome-wide association study (GWAS), therefore only comparative genomic approaches could be applied to the analysis. Mapping the relatively large genome is computational demanding, so all studied isolates were mapped against 3D7 reference genome. However, mapping passage isolates with its parents isolate would be the ideal situation to see progressive mutation in the parasite line. This information was not available.

Despite these limitations the study has highlighted again the importance of the unfolded protein response (UPR) pathway and redox metabolism acquisition of resistance and this could be the focus of further more detailed study.

Chapter 8

General Discussion

Plasmodium falciparum the deadliest form of malaria pathogens remains a global economic and health burdens. The pursuit of eradication of malaria relies on vector control, management, and chemotherapy. Antimalarial drugs efficacies were challenged by resistant parasites; chloroquine was a good example, it was once the most effective and most affordable antimalarial drug, and raised the hope for successful eradication. However, the widespread of resistance to chloroquine had rendered it useless in the field. This could repeat with the present most important antimalarial drug artemisinins, if mechanism of action and resistance remains unknown.

In the past decade, artemisinins became first line treatment for uncomplicated *falciparum* malaria as other antimalarial drugs are ineffective. Artemisinins are currently used in combination with other antimalarial as artemisinin combination therapies (ACTs). ACTs have been described to have favourable features; fast-acting artemisinins combined with a long acting partner drug. Artemisinins monotherapies failure rate had increased due to short half-life of artemisinins in plasma, therefore any parasites that survive the surge of artemisinin remain infectious. This could be easily addressed by long half-life antimalarial drug e.g. mefloquine or pyrimethamine/sulfadoxin.

Although ACTs are globally implemented along vector control programs, malaria prevalence is not greatly reduced in most endemic areas. This is due to resistance of parasite and vector. Unlike other antimalarial drugs, artemisinins resistance is not easily determined by conventional IC_{50} value and mostly dependent on clinical features. In the past years, reduced sensitivity to ACTs was evidenced in the fields as reflected by delayed parasite clearance half-life. ACTs treatment failures were also reported. Recently, artemisinin *in vitro* resistance phenotype is assessed by ring survival assay (RSA) or delayed clearance assay (DCA). More recently, mutations in K13 gene was reported as a major artemisinin resistance marker as it is strongly correlated with delayed parasite clearance half-life in clinical samples. However, it seems that mutations in K13 gene is geographical divergence (Menard et al., 2016) and not utterly accurate prediction (Wilairat et al., 2016). K13 marker is useful for artemisinin resistance monitoring in Southeast Asia where reduced sensitivity to artemisinin is a topical concern.

In this thesis, chemically modified artemisinin activity-based probes and desoxyartemisinin probes were used to study molecular targets of artemisinin in different stages of *Plasmodium falciparum* strain 3D7. By using these probes as activity pairwise strategy, parasite protein targets of artemisinin were

confidently identified by mass spectrometry-based proteomics approach. Chemical pairwise strategy of copper-catalysed and copper-free click reaction was also introduced to compare protein labelling profiles. The stringency of these pairwise strategies discussed in Chapter 3 of this thesis. Probe-labelled proteome (or 'endoperoxome') from trophozoite stage parasite was discussed in the context of plausible artemisinin targets and mechanism of action in Chapter 4. Then copper-free click reaction was selected to study artemisinin molecular targets in ring stage parasite. Although the background labelling from copper-free click reaction was higher than copper-catalysed click reaction, slightly higher sensitivity, less toxicity of the reaction and the absence of copper ion were the pluses, and activity pairwise strategy was also in place to increase confidence of protein labelling. Proteins identified from ring stage parasites were discussed in Chapter 5 in the context of plausible artemisinin mechanism, differential response to artemisinin, and possible mechanism of resistance. In Chapter 6, an artemisinin alternative 1,2,4-trioxolane sharing similar core structure endoperoxide was used to study for its molecular targets. Lastly in Chapter 7, artemisinin susceptible and resistant parasites whole genomes were compared to identify similarities and differences that would underline the phenotypes.

Artemisinin activity-based probes and desoxyartemisinin probes were tested for antimalarial activity. As expected, artemisinin activity-based probes exhibit potent antimalarial activity similar to artemisinin and dihydroartemisinin, while all desoxyartemisinin probes have no antimalarial activity. Probe-labelled proteomes were compared and clearly showed that artemisinin activity-based probes labelled more proteins that desoxyartemisinin. An order of magnitude of difference was achieved by alkyne probes via copper-catalysed click reaction. The 2 to 3-fold difference was obtained by azide probes via copper-free click reaction. Similar observation was also depicted by protein gel electrophoresis. This finding established the firm foundation for modified probes and click reaction for pharmacological study.

The trophozoite stage parasite is the most sensitive to artemisinin (Klonis et al., 2013), therefore target identification of artemisinin in this stage can provide some information on how artemisinin gains its potent activity in the parasite. Artemisinin activity-based probes labelled proteins involved in many major parasite biological pathways including haemoglobin digestion, glycolysis, redox metabolism, nucleic acid and protein synthesis metabolism, membrane transport, and unfolded protein response. This finding supports the hypothesis that artemisinin killing effect is a cluster bomb. This catastrophic effect is initialised by activation of artemisinin to activated artemisinin, secondary carbon centred radicals, by predominantly haem-mediated process and free iron within the parasite vacuole (Stocks et al., 2007, O'Neill et al., 2010, Tilley et al., 2016, Xie et al., 2016).

The endoperoxome from ring stage parasites showed reduced number of protein partners compared to tropozoite stage. In addition, the absence of haemoglobin digestion pathway in ring stage implies

that the pathway is one of the major contributors of artemisinin action during later stage of parasites. Along with the reduction in protein partners in other pathway e.g. unfolded protein response (UPR) which was suggested as one of resistant mediator, redox metabolism, and nucleic acid and protein biosynthesis pathways. This finding highlights a possible mechanism of reduced sensitivity to artemisinin during early developmental stage. However, the reduced sensitivity is in accordance with the level of haemoglobin degradation activity during this stage. Therefore, parasites whose life cycle are prolonged during this stage can escape from the artemisinins pressure and is believed to be a mechanism of resistance to artemisinins.

To further investigate the molecular targets of the fully synthetic endoperoxide antimalarial candidates, a set of trioxolane endoperoxide activity-based probes were synthesised and used to identify the protein partners. Interestingly, the endoperoxome of the fully synthetic endoperoxide shares similar profile with artemisinin. This raises a concern over a cross resistance between endoperoxide group of compounds as they may have similar or possibly identical mechanism of action and resistance. However, no resistance have been reported from the clinical trial yet thanks to its much longer elimination half-life over artemisinins.

Finally, whole genome sequences of artesunate susceptible and resistant parasites were compared to identify genes associated with the phenotypes. Although variation in kelch 13 (K13) gene, a prominent artemisinins resistant marker, was not identified from the studied population, a selection towards genes involved in unfolded protein response and cell communication was observed from the analysis of SNPs. Strikingly, the indels variation were more common than the SNP, and asparagine repeat-rich proteins were enriched from the genes associated with SNPs. This area is very interesting for further study for possible mechanism of drug resistance.

In summary, the chemical biology by click chemistry coupled with mass spectrometry proteomic approach was proved very useful for target validation of antimalarial drug artemisinin. These experiments have provided the strong pieces of evidence showing artemisinin molecular targets and agreed with proposed mechanism of artemisinin. Even though it is not enough to conclude the whole mechanism of action but a good starting point for further validation of target proteins. In addition, the whole genome sequencing comparative approach has shown that some genetic and protein features might be overlooked for mechanism of resistance to antimalarial drugs. Together these findings help guiding and shaping a future research.

Future perspective

The results presented in this work showed that modified artemisinin and click chemistry-based proteomics approach can be used to study molecular targets of antimalarial drug. Further study with lower concentration, relevant to physiological condition, and shorter time period would provide some information on what happens during the very early stage of drug action as artemisinins gain a killing effect within the very first hours of administration (Sanz et al., 2012). In addition, using this approach to study molecular targets in the parasites with well-defined resistance phenotype would help to understand the mechanism of resistance to artemisinins. To further extend our knowledge on artemisinin action throughout the intraerytrocytic life cycle of the parasites, targets identification in gametocyte stages would fulfil the understanding of different effects of artemisinin between early and late gametocyte (in-house data not published).

As this click chemistry coupled with MS proteomic approach provided a number of protein targets of artemisinin in form of drug-protein adducts, further validations of these artemisinin partner proteins are required for better understanding the killing effect of artemisinins and possible resistance mechanism to artemisinins. The genetic modification approach such as CRISPR technique is an emerging platform for *Plasmodium falciparum* and proven very useful and successful in this organism over other techniques (Lee and Fidock, 2014). By using CRISPR to knockout or insert particular genes and study its consequences on artemisinins sensitivity could be very useful. Another interesting genetic approach is genetic modification in *Plasmodium berghei* from PlasmoGEM project. The project has generated a large number of genetically modified *Plasmodium berghei* strains and could be used to study the effect of overexpression or deletion of interested genes in the mouse malaria model. The metabolomics, either global or targeted, would provide some more information on loss of function of an enzyme and a pathway as metabolites are linked with the protein functions (Tiwari et al., 2016, Creek et al., 2016). Metabolomics studies have been facilitated the studies of drug target validation (Cobbold et al., 2016, Becker et al., 2010) and becoming more mainstream in drug target validation as more tools and databases become available.

Appendix 1

Proteins identified from the proteomics experiments

Table A1.1 Proteins identified from DMSO treatments

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	5
PF3D7_0818900	Q8IB24	heat shock protein 70 (HSP70)	4
PF3D7_1105100	Q8IIV1	histone H2B (H2B)	3
PF3D7_0617800	C6KT18	Histone H2A	3
PF3D7_1343000	Q8IDQ9	Phosphoethanolamine N-methyltransferase	2
PF3D7_0827900	C0H4Y6	Protein disulfide-isomerase	2
PF3D7_0608800	Q6LFH8	ornithine aminotransferase (OAT)	1
PF3D7_0922200	Q7K6A4	S-adenosylmethionine synthetase (SAMS)	1
PF3D7_0922500	P27362	phosphoglycerate kinase (PGK)	1
PF3D7_0930300	Q8I0U8	merozoite surface protein 1 (MSP1)	1
PF3D7_1015900	Q8IJN7	enolase (ENO)	1
PF3D7_1027300	Q8IJD0	peroxiredoxin (nPrx)	1
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase (LDH)	1
PF3D7_1462800	Q8IKK7	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	1
PF3D7_1357000	Q8I0P6	Elongation factor 1-alpha	1

Table A1.2 Protein identified by 1-carbon alkyne desoxyartemisinin probe (ART-21) on trophozoite stage parasite treatment

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0917900	Q8I2X4	Heat shock protein 70	2
PF3D7_0617800	C6KT18	histone H2A (H2A)	1
PF3D7_0818900	Q8IB24	heat shock protein 70 (HSP70)	1
PF3D7_0827900	C0H4Y6	protein disulfide isomerase (PDI8)	1
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	1
PF3D7_0930300	Q8I0U8	merozoite surface protein 1 (MSP1)	1
PF3D7_1027300	Q8IJD0	peroxiredoxin (nPrx)	1
PF3D7_1104400	Q8IIV8	conserved protein, unknown function	1
PF3D7_1105100	Q8IIV1	histone H2B (H2B)	1
PF3D7_1129000	Q8II73	spermidine synthase (SpdSyn)	1
PF3D7_1222300	Q8I0V4	endoplasmin, putative (GRP94)	1
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	1

Table A1.3 Proteins identified from 4-carbon alkyne desoxyartemisinin probe (SCR-A-21) on trophozoite stage parasite treatment

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0818900	Q8IB24	heat shock protein 70 (HSP70)	2
PF3D7_1015900	Q8IJN7	enolase (ENO)	2
PF3D7_1462800	Q8IKK7	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	2
PF3D7_1357000	Q8I0P6	Elongation factor 1-alpha	2
PF3D7_0608800	Q6LFH8	ornithine aminotransferase (OAT)	1
PF3D7_0617800	C6KT18	histone H2A (H2A)	1
PF3D7_0708400	Q8IC05	heat shock protein 90 (HSP90)	1
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	1
PF3D7_0922200	Q7K6A4	S-adenosylmethionine synthetase (SAMS)	1
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase (LDH)	1
PF3D7_1343000	Q8IDQ9	phosphoethanolamine N-methyltransferase (PMT)	1

Table A1.4 Proteins identified by 4-carbon azide desoxyartemisinin probe (SCR-A-22) on trophozoite stage parasite treatment

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0307200	O97250	60S ribosomal protein L7, putative	4
PF3D7_0317600	O77381	40S ribosomal protein S11, putative (RPS11)	4
PF3D7_0322900	O97313	40S ribosomal protein S3A, putative	4
PF3D7_0415900	C0H4A6	60S ribosomal protein L15, putative	4
PF3D7_0422400	Q8IFP2	40S ribosomal protein S19 (RPS19)	4
PF3D7_0507100	Q8I431	60S ribosomal protein L4 (RPL4)	4
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2 (RPL2)	4
PF3D7_0520900	P50250	S-adenosyl-L-homocysteine hydrolase (SAHH)	4
PF3D7_0608800	Q6LFH8	ornithine aminotransferase (OAT)	4
PF3D7_0614500	C6KSY6	60S ribosomal protein L19 (RPL19)	4
PF3D7_0618300	C6KT23	60S ribosomal protein L27a, putative	4
PF3D7_0626800	C6KTA4	pyruvate kinase (PyrK)	4
PF3D7_0708400	Q8IC05	heat shock protein 90 (HSP90)	4
PF3D7_0710600	Q8IBY4	60S ribosomal protein L34 (RPL34)	4
PF3D7_0721600	Q8IBN5	40S ribosomal protein S5, putative	4
PF3D7_0813900	Q8IAX5	40S ribosomal protein S16, putative	4
PF3D7_0818900	Q8IB24	heat shock protein 70 (HSP70)	4
PF3D7_0903700	Q6ZLZ9	alpha tubulin 1	4
PF3D7_0922200	Q7K6A4	S-adenosylmethionine synthetase (SAMS)	4
PF3D7_0922500	P27362	phosphoglycerate kinase (PGK)	4
PF3D7_1008700	Q7KQL5	tubulin beta chain	4

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1019400	Q8IJK8	60S ribosomal protein L30e, putative	4
PF3D7_1027800	Q8IJC6	60S ribosomal protein L3 (RPL3)	4
PF3D7_1105100	Q8IIV1	histone H2B (H2B)	4
PF3D7_1105400	Q8IIU8	40S ribosomal protein S4, putative	4
PF3D7_1126200	Q8IIA2	40S ribosomal protein S18, putative	4
PF3D7_1212000	Q8I5T2	glutathione peroxidase-like thioredoxin peroxidase (TPx(Gl))	4
PF3D7_1235600	Q8I566	serine hydroxymethyltransferase (SHMT)	4
PF3D7_1246200	Q8I4X0	actin I (ACT1)	4
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase (LDH)	4
PF3D7_1341200	Q8IDS6	60S ribosomal protein L18, putative	4
PF3D7_1351400	Q8IDI5	60S ribosomal protein L17, putative	4
PF3D7_1368200	Q8I6Z4	ABC transporter E family member 1, putative (ABCE1)	4
PF3D7_1421200	Q8ILN8	40S ribosomal protein S25 (RPS25)	4
PF3D7_1424100	Q8ILL3	60S ribosomal protein L5, putative	4
PF3D7_1424400	Q8ILL2	60S ribosomal protein L7-3, putative	4
PF3D7_1426000	Q8ILK3	60S ribosomal protein L21 (RPL21)	4
PF3D7_1431700	Q8ILE8	60S ribosomal protein L14, putative	4
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	4
PF3D7_1451100	Q8IKW5	elongation factor 2 (eEF2)	4
PF3D7_1460700	Q8IKM5	60S ribosomal protein L27 (RPL27)	4
PF3D7_1462800	Q8IKK7	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	4
PF3D7_1465900	Q8IKH8	40S ribosomal protein S3	4

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1357000	Q8I0P6	Elongation factor 1-alpha	4
PF3D7_0205900	O96153	26S proteasome regulatory subunit RPN1, putative (RPN1)	3
PF3D7_0209800	Q9TY94	ATP-dependent RNA helicase UAP56 (UAP56)	3
PF3D7_0214000	O96220	T-complex protein 1 subunit theta (CCT8)	3
PF3D7_0306800	O97247	T-complex protein 1 subunit beta (CCT2)	3
PF3D7_0308200	O77323	T-complex protein 1 subunit eta (CCT7)	3
PF3D7_0309600	O00806	60S acidic ribosomal protein P2 (PfP2)	3
PF3D7_0316800	O77395	40S ribosomal protein S15A, putative	3
PF3D7_0320300	O97282	T-complex protein 1 subunit epsilon (CCT5)	3
PF3D7_0322000	Q76NN7	peptidyl-prolyl cis-trans isomerase (CYP19A)	3
PF3D7_0401800	Q8I207	Plasmodium exported protein (PHISTb), unknown function (PfD80)	3
PF3D7_0406100	Q6ZMA8	V-type proton ATPase subunit B	3
PF3D7_0424600	Q8IFM0	Plasmodium exported protein (PHISTb), unknown function	3
PF3D7_0503400	Q8I467	actin-depolymerizing factor 1 (ADF1)	3
PF3D7_0511800	Q8I3Y8	inositol-3-phosphate synthase (INO1)	3
PF3D7_0513300	Q8I3X4	purine nucleoside phosphorylase (PNP)	3
PF3D7_0517000	Q8I3T8	60S ribosomal protein L12, putative	3
PF3D7_0517700	Q8I3T1	eukaryotic translation initiation factor 3 subunit B, putative (EIF3B)	3
PF3D7_0524000	Q8I3M5	karyopherin beta (KASbeta)	3
PF3D7_0525100	Q8I3L4	acyl-CoA synthetase (ACS10)	3
PF3D7_0532300	Q8I3F1	Plasmodium exported protein (PHISTb), unknown function	3
PF3D7_0606700	C6KSR5	coatomer alpha subunit, putative	3

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0608700	C6KST5	T-complex protein 1 subunit zeta (CCT6)	3
PF3D7_0617800	C6KT18	histone H2A (H2A)	3
PF3D7_0619400	C6KT34	cell division cycle protein 48 homologue, putative	3
PF3D7_0624000	C6KT76	hexokinase (HK)	3
PF3D7_0627500	C6KTB1	protein DJ-1 (DJ1)	3
PF3D7_0703500	Q8IC35	erythrocyte membrane-associated antigen	3
PF3D7_0708800	Q8IC01	heat shock protein 110 (HSP110c)	3
PF3D7_0716800	Q8IBT2	eukaryotic translation initiation factor 3 subunit I, putative (EIF3I)	3
PF3D7_0719600	Q8IBQ6	60S ribosomal protein L11a, putative	3
PF3D7_0722400	Q8IBM9	GTP-binding protein, putative	3
PF3D7_0812400	Q8IAW0	karyopherin alpha (KARalpha)	3
PF3D7_0813300	C0H4U4	conserved Plasmodium protein, unknown function	3
PF3D7_0814000	Q8IAX6	60S ribosomal protein L13-2, putative	3
PF3D7_0814200	Q8IAX8	DNA/RNA-binding protein Alba 1 (ALBA1)	3
PF3D7_0818200	C0H4V6	14-3-3 protein (14-3-3I)	3
PF3D7_0826700	Q8IBA0	receptor for activated c kinase (RACK)	3
PF3D7_0827900	C0H4Y6	protein disulfide isomerase (PDI8)	3
PF3D7_0831400	C0H4Z7	Plasmodium exported protein, unknown function	3
PF3D7_0905400	Q8I395	high molecular weight rhoptry protein 3 (RhopH3)	3
PF3D7_0915400	Q8I2Z8	6-phosphofructokinase (PFK9)	3
PF3D7_0917900	Q8I2X4	heat shock protein 70 (HSP70-2)	3
PF3D7_0929400	C0H571	high molecular weight rhoptry protein 2 (RhopH2)	3

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0930300	Q8I0U8	merozoite surface protein 1 (MSP1)	3
PF3D7_0934500	Q8I2H3	V-type proton ATPase subunit E, putative	3
PF3D7_0934800	Q7K6A0	cAMP-dependent protein kinase catalytic subunit (PKAc)	3
PF3D7_1003500	Q8IK02	40S ribosomal protein S20e, putative	3
PF3D7_1004000	Q8IJZ7	60S ribosomal protein L13, putative	3
PF3D7_1006800	Q8IJX3	RNA-binding protein, putative	3
PF3D7_1007900	Q8IJW4	eukaryotic translation initiation factor 3 subunit D, putative (EIF3D)	3
PF3D7_1008400	Q8IJW0	26S protease regulatory subunit 4, putative (RPT2)	3
PF3D7_1011800	Q8IJS7	PRE-binding protein (PREBP)	3
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase (HGPRT)	3
PF3D7_1015900	Q8IJN7	enolase (ENO)	3
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2 (RPS2)	3
PF3D7_1027300	Q8IJD0	peroxiredoxin (nPrx)	3
PF3D7_1029600	Q8IJA9	adenosine deaminase (ADA)	3
PF3D7_1033400	Q8IJ74	haloacid dehalogenase-like hydrolase (HAD1)	3
PF3D7_1034900	Q8IJ60	methionine--tRNA ligase (MRScyt)	3
PF3D7_1037300	Q8IJ34	ADP/ATP transporter on adenylate translocase (ADT)	3
PF3D7_1105000	Q8IIV2	histone H4 (H4)	3
PF3D7_1108400	Q8IIR9	casein kinase 2, alpha subunit (CK2alpha)	3
PF3D7_1116800	Q8IIJ8	heat shock protein 101 (HSP101)	3
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4 (RAN)	3
PF3D7_1120100	Q8IIG6	phosphoglycerate mutase, putative (PGM1)	3

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1129000	Q8II73	spermidine synthase (SpdSyn)	3
PF3D7_1130200	Q8II61	60S ribosomal protein P0 (PfP0)	3
PF3D7_1134000	Q8II24	heat shock protein 70 (HSP70-3)	3
PF3D7_1145400	Q8IHR4	dynamin-like protein (DYN1)	3
PF3D7_1206200	Q8I5Y3	eukaryotic translation initiation factor 3 subunit C, putative (EIF3C)	3
PF3D7_1211900	Q8I5T3	non-SERCA-type Ca2+ -transporting P-ATPase (ATP4)	3
PF3D7_1212700	Q8I5S6	eukaryotic translation initiation factor 3 subunit A, putative (EIF3A)	3
PF3D7_1219100	Q8I5L6	clathrin heavy chain, putative	3
PF3D7_1222300	Q8I0V4	endoplasmin, putative (GRP94)	3
PF3D7_1224300	Q8I5H4	polyadenylate-binding protein, putative (PABP)	3
PF3D7_1229500	Q8I5C4	T-complex protein 1 subunit gamma (CCT3)	3
PF3D7_1248900	Q8I4U5	26S protease regulatory subunit 8, putative (RPT6)	3
PF3D7_1306400	Q8IEQ1	26S protease regulatory subunit 10B, putative (RPT4)	3
PF3D7_1308200	Q8IEN3	carbamoyl phosphate synthetase (cpsSII)	3
PF3D7_1309100	Q8IEM3	60S ribosomal protein L24, putative	3
PF3D7_1311800	Q8IEK1	M1-family alanyl aminopeptidase (M1AAP)	3
PF3D7_1311900	Q76NM6	V-type proton ATPase catalytic subunit A (vapA)	3
PF3D7_1317800	C0H5C2	40S ribosomal protein S19 (RPS19)	3
PF3D7_1318800	Q8IEC8	translocation protein SEC63 (SEC63)	3
PF3D7_1325100	Q8IE67	phosphoribosylpyrophosphate synthetase	3
PF3D7_1331800	Q8IE09	60S ribosomal protein L23, putative	3
PF3D7_1332900	Q8IDZ9	isoleucine--tRNA ligase, putative	3

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1338200	Q8IDV1	60S ribosomal protein L6-2, putative	3
PF3D7_1338300	Q8IDV0	elongation factor 1-gamma, putative	3
PF3D7_1342000	Q8IDR9	40S ribosomal protein S6	3
PF3D7_1343000	Q8IDQ9	phosphoethanolamine N-methyltransferase (PMT)	3
PF3D7_1346100	Q8IDN6	protein transport protein SEC61 subunit alpha (SEC61)	3
PF3D7_1347500	Q8IDM3	DNA/RNA-binding protein Alba 4 (ALBA4)	3
PF3D7_1349200	Q8IDK7	glutamate--tRNA ligase, putative	3
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	3
PF3D7_1357800	C0H5I7	T-complex protein 1 subunit delta (CCT4)	3
PF3D7_1357900	Q8IDC6	pyrroline-5-carboxylate reductase	3
PF3D7_1358800	Q8IDB0	40S ribosomal protein S15 (RPS15)	3
PF3D7_1407100	Q8IM23	rRNA 2'-O-methyltransferase fibrillarin, putative (NOP1)	3
PF3D7_1408600	Q8IM10	40S ribosomal protein S8e, putative	3
PF3D7_1410600	Q8ILY9	eukaryotic translation initiation factor 2 gamma subunit, putative	3
PF3D7_1428300	Q8ILI2	proliferation-associated protein 2g4, putative	3
PF3D7_1437900	Q8IL88	HSP40, subfamily A, putative (ERdj3)	3
PF3D7_1438900	Q8IL80	thioredoxin peroxidase 1 (Trx-Px1)	3
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase (FBPA)	3
PF3D7_1468700	Q8IKF0	eukaryotic initiation factor 4A (eIF4A)	3
PF3D7_1471100	Q8IKC8	exported protein 2 (EXP2)	3
PF3D7_1136500.1	Q8IHZ9	Casein kinase I	3
PF3D7_0106300	Q76NN8	Calcium-transporting ATPase	3

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_0214100	O96221	protein transport protein SEC31 (SEC31)	2
PF3D7_0217800	O96258	40S ribosomal protein S26 (RPS26)	2
PF3D7_0302500	O77310	cytoadherence linked asexual protein 3.1 (CLAG3.1)	2
PF3D7_0302900	O77312	exportin-1, putative	2
PF3D7_0402000	Q8I206	Plasmodium exported protein (PHISTa), unknown function	2
PF3D7_0413600	Q8I1V1	26S protease regulatory subunit 6B, putative (RPT3)	2
PF3D7_0500800	Q8I492	mature parasite-infected erythrocyte surface antigen (MESA)	2
PF3D7_0501000	Q8I490	Plasmodium exported protein, unknown function	2
PF3D7_0513600	Q8I0W8	deoxyribodipyrimidine photo-lyase, putative	2
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1 (MDR1)	2
PF3D7_0528200	Q8I3I5	eukaryotic translation initiation factor 3 subunit E, putative (EIF3E)	2
PF3D7_0601200	C6KSL5	Pfmc-2TM Maurer's cleft two transmembrane protein (MC-2TM)	2
PF3D7_0705700	C0H4K8	40S ribosomal protein S29, putative	2
PF3D7_0706400	C0H4L5	60S ribosomal protein L37 (RPL37)	2
PF3D7_0822600	Q8IB60	protein transport protein SEC23 (SEC23)	2
PF3D7_0903900	Q8I3B0	60S ribosomal protein L32 (RPL32)	2
PF3D7_0913200	Q8I320	elongation factor 1-beta (EF-1beta)	2
PF3D7_0919000	Q8I2W3	nucleosome assembly protein (NAPS)	2
PF3D7_1004400	Q8IJZ3	RNA-binding protein, putative	2
PF3D7_1015800	Q8IJN8	ribonucleotide reductase small subunit, putative	2
PF3D7_1103100	Q8IIX0	60S acidic ribosomal protein P1, putative (RPP1)	2
PF3D7_1129200	Q8II71	26S proteasome regulatory subunit RPN7, putative (RPN7)	2

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1130400	Q8II60	26S protease regulatory subunit 6A, putative (RPT5)	2
PF3D7_1132200	Q8II43	T-complex protein 1 subunit alpha (TCP1)	2
PF3D7_1142600	Q8IHT9	60S ribosomal protein L35ae, putative	2
PF3D7_1143300	Q8IHT3	DNA-directed RNA polymerase I, putative	2
PF3D7_1302800	Q8IET7	40S ribosomal protein S7, putative	2
PF3D7_1311500	Q8IEK3	26S protease regulatory subunit 7, putative (RPT1)	2
PF3D7_1323100	Q8IE85	60S ribosomal protein L6, putative	2
PF3D7_1414300	Q8ILV2	60S ribosomal protein L10, putative	2
PF3D7_1437200	Q8IL94	ribonucleoside-diphosphate reductase, large subunit, putative	2
PF3D7_1441200	Q8IL58	60S ribosomal protein L1, putative	2
PF3D7_1454700	Q8IKT2	6-phosphogluconate dehydrogenase, decarboxylating, putative	2
PF3D7_0302200	O77309	cytoadherence linked asexual protein 3.2 (CLAG3.2)	1
PF3D7_0312800	O77364	60S ribosomal protein L26, putative	1
PF3D7_0512600	Q7K6A8,Q9BHN1	ras-related protein Rab-1B (RAB1b)	1
PF3D7_0610400	C6KSV0	histone H3 (H3)	1
PF3D7_0621200	C6KT50	pyridoxine biosynthesis protein PDX1 (PDX1)	1
PF3D7_0807300	Q7K6B0	ras-related protein Rab-18 (RAB18)	1
PF3D7_0912400	Q8I328	alkaline phosphatase, putative	1
PF3D7_0914700	Q8I305	transporter, putative	1
PF3D7_1010600	Q8IJT9	eukaryotic translation initiation factor 2 beta subunit, putative	1
PF3D7_1011400	Q8IJT1	proteasome subunit beta type-5	1
PF3D7_1124600	Q8IIB7	ethanolamine kinase (EK)	1

Gene ID	UniProt ID	Product Description	Frequency
PF3D7_1142500	Q8IHU0	60S ribosomal protein L28 (RPL28)	1
PF3D7_1229400	Q8I5C5	macrophage migration inhibitory factor (MIF)	1
PF3D7_1237700	Q8I546	conserved protein, unknown function	1
PF3D7_1252100	Q8I4R5	rhoptry neck protein 3 (RON3)	1
PF3D7_1323400	Q8IE82	60S ribosomal protein L23 (RPL23)	1
PF3D7_1328100	Q8I6T3	proteasome subunit beta type-7, putative	1
PF3D7_1341300	C0H5G3	60S ribosomal protein L18-2, putative	1
PF3D7_1342400	Q8IDR5	casein kinase II beta chain (CK2beta2)	1
PF3D7_1353100	Q8IDG9	Plasmodium exported protein, unknown function	1
PF3D7_1401800	Q8IM71	choline kinase (CK)	1
PF3D7_1410400	Q8ILZ1	rhoptry-associated protein 1 (RAP1)	1
PF3D7_1442300	Q8IL48	tRNA binding protein, putative	1
PF3D7_1445900	Q8IL13	ATP-dependent RNA helicase DDX5, putative (DDX5)	1
PF3D7_1456800	Q8IKR1	V-type H(+)-translocating pyrophosphatase, putative (VP1)	1
PF3D7_0102900	Q8I2B1	Aspartate--tRNA ligase	1

Table A1.5 Proteins identified by 4-carbon azide desoxyartemisinin probe (SCR-A-22) on ring stage parasite treatment

Gene ID	UniProt ID	Product description	Frequency
PF3D7_0322900	O97313	40S ribosomal protein S3A, putative	4
PF3D7_0507100	Q8I431	60S ribosomal protein L4 (RPL4)	4
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2 (RPL2)	4
PF3D7_0608800	Q6LFH8	ornithine aminotransferase (OAT)	4
PF3D7_0617800	C6KT18	histone H2A (H2A)	4
PF3D7_0626800	C6KTA4	pyruvate kinase (PyrK)	4
PF3D7_0818900	Q8IB24	heat shock protein 70 (HSP70)	4
PF3D7_1004000	Q8IJZ7	60S ribosomal protein L13, putative	4
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	4
PF3D7_1451100	Q8IKW5	elongation factor 2 (eEF2)	4
PF3D7_1462800	Q8IKK7	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	4
PF3D7_1465900	Q8IKH8	40S ribosomal protein S3	4
PF3D7_1357000	Q8I0P6	Elongation factor 1-alpha	4
PF3D7_0818200	C0H4V6	14-3-3 protein (14-3-3I)	3
PF3D7_1027800	Q8IJC6	60S ribosomal protein L3 (RPL3)	3
PF3D7_1105100	Q8IIV1	histone H2B (H2B)	3
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4 (RAN)	3
PF3D7_1343000	Q8IDQ9	phosphoethanolamine N-methyltransferase (PMT)	3
PF3D7_0322000	Q76NN7	peptidyl-prolyl cis-trans isomerase (CYP19A)	2
PF3D7_0708400	Q8IC05	heat shock protein 90 (HSP90)	2
PF3D7_1246200	Q8I4X0	actin I (ACT1)	2

Gene ID	UniProt ID	Product description	Frequency
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase (LDH)	2
PF3D7_0826700	Q8IBA0	Receptor for activated c kinase	1

Table A1.6 All protein identified from artemisinin activity-based probes treatment ranking by identified frequency (n=12)

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0207600	Q9TY95	serine repeat antigen 5	10	Very high
PF3D7_0629200	C6KTC7	DnaJ protein, putative	10	Very high
PF3D7_1306200	Q8IEQ3	conserved Plasmodium protein, unknown function	10	Very high
PF3D7_0523000	Q7K6A5	multidrug resistance protein 1	10	Very high
PF3D7_1121600	Q8IIF0	exported protein 1	10	Very high
PF3D7_0702400	Q8IC43	small exported membrane protein 1	9	Very high
PF3D7_0720400	Q8IBP8	ferrodoxin reductase-like protein	9	Very high
PF3D7_1008900	Q8IJV6	adenylate kinase	9	Very high
PF3D7_1118200	Q8III6	heat shock protein 90, putative	9	Very high
PF3D7_1438100	Q8IL86	secretory complex protein 62	9	Very high
PF3D7_1459400	Q8IKN7	conserved Plasmodium protein, unknown function	9	Very high
PF3D7_0501200	Q8I488	parasite-infected erythrocyte surface protein	8	Very high
PF3D7_0532100	Q8I3F3	early transcribed membrane protein 5	8	Very high
PF3D7_0628300	C6KTB9	choline/ethanolaminephosphotransferase, putative	8	Very high
PF3D7_0807300	Q7K6B0	ras-related protein Rab-18	8	Very high
PF3D7_0903200	C0H516	ras-related protein RAB7	8	Very high
PF3D7_0936800	Q8I2F2	Plasmodium exported protein (PHISTc), unknown function	8	Very high
PF3D7_1016300	Q8I6U8	glycophorin binding protein	8	Very high
PF3D7_1104400	Q8IIV8	thioredoxin, putative	8	Very high
PF3D7_1108700	Q8IIR6	heat shock protein DnaJ homologue Pfj2	8	Very high
PF3D7_1211800	Q7KQK2	polyubiquitin	8	Very high

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1301700	Q8IEJ0	Plasmodium exported protein (hyp8), unknown function	8	Very high
PF3D7_1408100	Q8IM15	plasmepsin III	8	Very high
PF3D7_1410400	Q8ILZ1	rhoptry-associated protein 1	8	Very high
PF3D7_1432100	Q8ILE3	voltage-dependent anion-selective channel protein, putative	8	Very high
PF3D7_0309600	O00806	60S acidic ribosomal protein P2	8	Very high
PF3D7_0709000	Q8IBZ9	chloroquine resistance transporter	7	High
PF3D7_0823200	Q8IB66	RNA-binding protein, putative	7	High
PF3D7_0912400	Q8I328	alkaline phosphatase, putative	7	High
PF3D7_0914700	Q8I305	major facilitator superfamily-related transporter, putative	7	High
PF3D7_0927900	Q8I2N0	phosphatidylserine decarboxylase	7	High
PF3D7_1345700	Q8I6T2	isocitrate dehydrogenase [NADP], mitochondrial	7	High
PF3D7_1347500	Q8IDM3	DNA/RNA-binding protein Alba 4	7	High
PF3D7_0936000	C0H592	ring-exported protein 2	7	High
PF3D7_0608800	Q6LFH8	ornithine aminotransferase	7	High
PF3D7_1231100	Q8I5A9	ras-related protein Rab-2	7	High
PF3D7_1105800	Q8IIU5	conserved Apicomplexan protein, unknown function	10	Very high
PF3D7_1408000	Q8I6V3	plasmepsin II	10	Very high
PF3D7_0823800	Q8IB72	DnaJ protein, putative	9	Very high
PF3D7_1010700	Q8IJT8	dolichyl-phosphate-mannose protein mannosyltransferase, putative	9	Very high
PF3D7_1237700	Q8I546	conserved protein, unknown function	9	Very high
PF3D7_1410700	Q8ILY8	conserved Plasmodium protein, unknown function	9	Very high
PF3D7_1360900	C0H5J5	polyadenylate-binding protein, putative	8	Very high

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1311800	Q8IEK1	M1-family alanyl aminopeptidase	8	Very high
PF3D7_1311900	Q76NM6	V-type proton ATPase catalytic subunit A	8	Very high
PF3D7_0316800	O77395	40S ribosomal protein S15A, putative	7	High
PF3D7_0500800	Q8I492	mature parasite-infected erythrocyte surface antigen	7	High
PF3D7_0731600	Q8I6Z1	acyl-CoA synthetase	7	High
PF3D7_1105000	Q8IIV2	histone H4	7	High
PF3D7_1143200	Q8IHT4	DnaJ protein, putative	7	High
PF3D7_1325100	Q8IE67	phosphoribosylpyrophosphate synthetase	7	High
PF3D7_1456800	Q8IKR1	V-type H(+)-translocating pyrophosphatase, putative	7	High
PF3D7_1468700	Q8IKF0	eukaryotic initiation factor 4A	7	High
PF3D7_1471100	Q8IKC8	exported protein 2	7	High
PF3D7_1020900	Q7KQL3	ADP-ribosylation factor	7	High
PF3D7_1246200	Q8I4X0	actin I	7	High
PF3D7_1324900	Q76NM3	L-lactate dehydrogenase	7	High
PF3D7_1330400	Q8IE22	ER lumen protein retaining receptor 1, putative, unspecified product	7	High
PF3D7_1330400	Q8IE22	ER lumen protein retaining receptor 1, putative, unspecified product	7	High
PF3D7_0402000	Q8I206	Plasmodium exported protein (PHISTa), unknown function	6	High
PF3D7_0422400	Q8IFP2	40S ribosomal protein S19	6	High
PF3D7_0424600	Q8IFM0	Plasmodium exported protein (PHISTb), unknown function	6	High
PF3D7_0501000	Q8I490	Plasmodium exported protein, unknown function	6	High
PF3D7_0501300	Q8I487	skeleton-binding protein 1	6	High
PF3D7_0532300	Q8I3F1	Plasmodium exported protein (PHISTb), unknown function	6	High

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0601900	C6KSL9	conserved Plasmodium protein, unknown function	6	High
PF3D7_0706500	C0H4L6	conserved Plasmodium protein, unknown function	6	High
PF3D7_0709700	Q8IBZ2	lysophospholipase, putative	6	High
PF3D7_0721100	Q8IBP0	conserved Plasmodium protein, unknown function	6	High
PF3D7_0814200	Q8IAX8	DNA/RNA-binding protein Alba 1	6	High
PF3D7_0818200	C0H4V6	14-3-3 protein	6	High
PF3D7_0919100	Q8I2W2	DnaJ protein, putative	6	High
PF3D7_0929400	C0H571	high molecular weight rhoptry protein 2	6	High
PF3D7_1015600	Q8IJN9	heat shock protein 60	6	High
PF3D7_1116700	Q8IIJ9	dipeptidyl aminopeptidase 1	6	High
PF3D7_1129000	Q8II73	spermidine synthase	6	High
PF3D7_1129100	Q8II72	parasitophorous vacuolar protein 1	6	High
PF3D7_1134000	Q8II24	heat shock protein 70	6	High
PF3D7_1222300	Q8I0V4	endoplasmin, putative	6	High
PF3D7_1224300	Q8I5H4	polyadenylate-binding protein, putative	6	High
PF3D7_1328300	Q8IE43	conserved Plasmodium protein, unknown function	6	High
PF3D7_1344800	Q8IDP8	aspartate carbamoyltransferase	6	High
PF3D7_1352500	Q8IDH5	thioredoxin-related protein, putative	6	High
PF3D7_1451800	Q8IKV8	sortilin	6	High
PF3D7_1454400	Q8IKT5	aminopeptidase P	6	High
PF3D7_0106300	Q76NN8	Calcium-transporting ATPase	6	High
PF3D7_1421200	Q8ILN8	40S ribosomal protein S25	6	High

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0930300	Q8I0U8	merozoite surface protein 1	6	High
PF3D7_1012400	Q8IJS1	hypoxanthine-guanine phosphoribosyltransferase	6	High
PF3D7_1026800	Q8IJD4	40S ribosomal protein S2	6	High
PF3D7_1117700	Q7KQK6	GTP-binding nuclear protein RAN/TC4	6	High
PF3D7_1130200	Q8II61	60S ribosomal protein P0	6	High
PF3D7_1444800	Q7KQL9	fructose-bisphosphate aldolase	6	High
PF3D7_0303600	Q8I224	plasmoredoxin	5	Moderate
PF3D7_0416800	Q8I1S0	small GTP-binding protein sar1	5	Moderate
PF3D7_0501600	Q8I484	rhoptry-associated protein 2	5	Moderate
PF3D7_0610400	C6KSV0	histone H3	5	Moderate
PF3D7_0802000	Q8IAM0	glutamate dehydrogenase, putative	5	Moderate
PF3D7_0813900	Q8IAX5	40S ribosomal protein S16, putative	5	Moderate
PF3D7_0824400	Q8IB78	nucleoside transporter 2	5	Moderate
PF3D7_0907400	Q8I377	ATP-dependent protease ATPase subunit ClpY	5	Moderate
PF3D7_0919400	Q8I2V9	protein disulfide isomerase	5	Moderate
PF3D7_0934500	Q8I2H3	V-type proton ATPase subunit E, putative	5	Moderate
PF3D7_1011400	Q8IJT1	proteasome subunit beta type-5	5	Moderate
PF3D7_1019900	Q8IJK2	autophagy-related protein 8	5	Moderate
PF3D7_1037300	Q8IJ34	ADP/ATP transporter on adenylate translocase	5	Moderate
PF3D7_1115600	Q8IIK8	peptidyl-prolyl cis-trans isomerase	5	Moderate
PF3D7_1201000	Q8I635	Plasmodium exported protein (PHISTb), unknown function	5	Moderate
PF3D7_1203700	Q8I608	nucleosome assembly protein	5	Moderate

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1211400	Q7KQK3	heat shock protein DNAJ homologue Pfj4	5	Moderate
PF3D7_1211900	Q8I5T3	non-SERCA-type Ca2+ -transporting P-ATPase	5	Moderate
PF3D7_1212000	Q8I5T2	glutathione peroxidase-like thioredoxin peroxidase	5	Moderate
PF3D7_1223100	Q7KQK0	cAMP-dependent protein kinase regulatory subunit	5	Moderate
PF3D7_1343000	Q8IDQ9	phosphoethanolamine N-methyltransferase	5	Moderate
PF3D7_1346100	Q8IDN6	protein transport protein SEC61 subunit alpha	5	Moderate
PF3D7_1353900	Q8IDG2	proteasome subunit alpha type-7, putative	5	Moderate
PF3D7_1359400	Q8IDB7	CUGBP Elav-like family member 1	5	Moderate
PF3D7_1407800	Q8IM16	plasmepsin IV	5	Moderate
PF3D7_1434800	Q8ILB6	mitochondrial acidic protein MAM33, putative	5	Moderate
PF3D7_1465900	Q8IKH8	40S ribosomal protein S3	5	Moderate
PF3D7_1466400	Q8IKH2	transcription factor with AP2 domain(s)	5	Moderate
PF3D7_0105200	Q8I289	RAP protein, putative	5	Moderate
PF3D7_0406100	Q6ZMA8	V-type proton ATPase subunit B	5	Moderate
PF3D7_0708400	Q8IC05	heat shock protein 90	5	Moderate
PF3D7_1405600	Q8IM38	ribonucleoside-diphosphate reductase small chain, putative	5	Moderate
PF3D7_1228600	Q8I5D2	merozoite surface protein 9	5	Moderate
PF3D7_1232100	Q8I0V3	60 kDa chaperonin	5	Moderate
PF3D7_0204700	Q7KWJ5	hexose transporter	4	Moderate
PF3D7_0316700	O77388	protein YOP1, putative	4	Moderate
PF3D7_0509000	Q8I0X0	SNAP protein (soluble N-ethylmaleimide-sensitive factor attachment protein), putative	4	Moderate
PF3D7_0512600	Q7K6A8	ras-related protein Rab-1B	4	Moderate

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0516900	Q8I3T9	60S ribosomal protein L2	4	Moderate
PF3D7_0517000	Q8I3T8	60S ribosomal protein L12, putative	4	Moderate
PF3D7_0532400	Q8I3F0	lysine-rich membrane-associated PHISTb protein	4	Moderate
PF3D7_0608300	C6KST1	conserved Plasmodium protein, unknown function	4	Moderate
PF3D7_0621200	C6KT50	pyridoxine biosynthesis protein PDX1	4	Moderate
PF3D7_0624600	C6KT82	SNF2 helicase, putative	4	Moderate
PF3D7_0727400	Q8IBI3	proteasome subunit alpha type-5, putative	4	Moderate
PF3D7_0807900	Q8IAR7	tyrosine--tRNA ligase	4	Moderate
PF3D7_0821000	Q8IB44	conserved Plasmodium protein, unknown function	4	Moderate
PF3D7_0827900	C0H4Y6	protein disulfide isomerase	4	Moderate
PF3D7_0831400	C0H4Z7	Plasmodium exported protein, unknown function	4	Moderate
PF3D7_1113300	Q8IIM9	UDP-galactose transporter, putative	4	Moderate
PF3D7_1116800	Q8IIJ8	heat shock protein 101	4	Moderate
PF3D7_1117300	Q8IIJ4	conserved Plasmodium protein, unknown function	4	Moderate
PF3D7_1119000	Q8IIH7	acyl-CoA-binding protein, putative	4	Moderate
PF3D7_1124600	Q8IIB7	ethanolamine kinase	4	Moderate
PF3D7_1124700	Q8IIB6	GrpE protein homolog, mitochondrial, putative	4	Moderate
PF3D7_1134100	Q8II23	protein disulfide isomerase	4	Moderate
PF3D7_1149400	Q8IHM9	Plasmodium exported protein, unknown function	4	Moderate
PF3D7_1229400	Q8I5C5	macrophage migration inhibitory factor	4	Moderate
PF3D7_1241700	Q8I512	replication factor C subunit 4, putative	4	Moderate
PF3D7_1242700	Q8I502	40S ribosomal protein S17, putative	4	Moderate

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1242800	Q8I501	rab specific GDP dissociation inhibitor	4	Moderate
PF3D7_1304500	Q8IES0	small heat shock protein, putative	4	Moderate
PF3D7_1353100	Q8IDG9	Plasmodium exported protein, unknown function	4	Moderate
PF3D7_1424100	Q8ILL3	60S ribosomal protein L5, putative	4	Moderate
PF3D7_1429600	Q8ILG8	conserved Plasmodium protein, unknown function	4	Moderate
PF3D7_1447000	Q8IL02	40S ribosomal protein S5	4	Moderate
PF3D7_1447700	Q8IKZ8	conserved Plasmodium protein, unknown function	4	Moderate
PF3D7_0108000	Q8I261	Proteasome subunit beta type	4	Moderate
PF3D7_0307100	O97249	40S ribosomal protein S12, putative	4	Moderate
PF3D7_0624000	C6KT76	hexokinase	4	Moderate
PF3D7_0826700	Q8IBA0	receptor for activated c kinase	4	Moderate
PF3D7_0917900	Q8I2X4	heat shock protein 70	4	Moderate
PF3D7_0922500	P27362	phosphoglycerate kinase	4	Moderate
PF3D7_1008700	Q7KQL5	tubulin beta chain	4	Moderate
PF3D7_1015900	Q8IJN7	enolase	4	Moderate
PF3D7_1354500	Q8IDF6	adenylosuccinate synthetase	4	Moderate
PF3D7_1357000	Q8I0P6	Elongation factor 1-alpha	4	Moderate
PF3D7_0310400	O77361	parasite-infected erythrocyte surface protein	3	Low
PF3D7_0316600	O77389	formate-nitrite transporter	3	Low
PF3D7_0415900	C0H4A6	60S ribosomal protein L15, putative	3	Low
PF3D7_0507100	Q8I431	60S ribosomal protein L4	3	Low
PF3D7_0523100	Q8I3N3	mitochondrial-processing peptidase subunit alpha, putative	3	Low

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0524000	Q8I3M5	karyopherin beta	3	Low
PF3D7_0614300	C6KSY4	major facilitator superfamily-related transporter, putative	3	Low
PF3D7_0619400	C6KT34	cell division cycle protein 48 homologue, putative	3	Low
PF3D7_0623500	C6KT71	superoxide dismutase [Fe]	3	Low
PF3D7_0626800	C6KTA4	pyruvate kinase	3	Low
PF3D7_0627800	C6KTB4	acetyl-CoA synthetase, putative	3	Low
PF3D7_0631900	C6KTF1	stevor	3	Low
PF3D7_0702500	Q8IC42	Plasmodium exported protein, unknown function	3	Low
PF3D7_0714000	Q8IBV7	histone H2B variant	3	Low
PF3D7_0801800	Q8IAL6	mannose-6-phosphate isomerase, putative	3	Low
PF3D7_0807500	Q8IAR3	proteasome subunit alpha type-6, putative	3	Low
PF3D7_0905400	Q8I395	high molecular weight rhoptry protein 3	3	Low
PF3D7_0918300	Q8I2X0	eukaryotic translation initiation factor 3 subunit F, putative	3	Low
PF3D7_0925900	Q8I2Q0	conserved Plasmodium protein, unknown function	3	Low
PF3D7_0931800	Q8I0U7	proteasome subunit beta type-6, putative	3	Low
PF3D7_0933600	Q8I2I2	mitochondrial-processing peptidase subunit beta, putative	3	Low
PF3D7_0935800	Q8I2G2	cytoadherence linked asexual protein 9	3	Low
PF3D7_1006700	Q8IJX4	conserved Plasmodium protein, unknown function	3	Low
PF3D7_1008800	Q8IJV7	nucleolar protein 5, putative	3	Low
PF3D7_1010300	Q8IJU2	succinate dehydrogenase subunit 4, putative	3	Low
PF3D7_1017900	Q8IJM0	26S proteasome regulatory subunit p55, putative	3	Low
PF3D7_1021900	Q8IJI4	PHAX domain-containing protein, putative	3	Low

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1104000	Q8IIW2	phenylalanine--tRNA ligase beta subunit	3	Low
PF3D7_1105100	Q8IIV1	histone H2B	3	Low
PF3D7_1118500	Q8III3	nucleolar protein 56, putative	3	Low
PF3D7_1126000	Q8IIA4	threonine--tRNA ligase	3	Low
PF3D7_1126200	Q8IIA2	40S ribosomal protein S18, putative	3	Low
PF3D7_1132200	Q8II43	T-complex protein 1 subunit alpha	3	Low
PF3D7_1149000	Q8IHN4	antigen 332, DBL-like protein	3	Low
PF3D7_1226900	Q8I5E9	conserved Plasmodium protein, unknown function	3	Low
PF3D7_1238100	Q8I542	calcyclin binding protein, putative	3	Low
PF3D7_1246800	Q8I4W4	signal recognition particle receptor, beta subunit	3	Low
PF3D7_1252100	Q8I4R5	rhoptry neck protein 3	3	Low
PF3D7_1318800	Q8IEC8	translocation protein SEC63	3	Low
PF3D7_1331700	Q8IE10	glutamine--tRNA ligase, putative	3	Low
PF3D7_1338200	Q8IDV1	60S ribosomal protein L6-2, putative	3	Low
PF3D7_1342000	Q8IDR9	40S ribosomal protein S6	3	Low
PF3D7_1355100	Q8IDF0	DNA replication licensing factor MCM6	3	Low
PF3D7_1361800	C0H5J9	conserved Plasmodium protein, unknown function	3	Low
PF3D7_1368100	Q8ID28	26S proteasome regulatory subunit RPN11, putative	3	Low
PF3D7_1370300	C0H5L9	membrane associated histidine-rich protein	3	Low
PF3D7_1408600	Q8IM10	40S ribosomal protein S8e, putative	3	Low
PF3D7_1412300	Q8ILX1	nuclear transport factor 2, putative	3	Low
PF3D7_1417500	Q8ILS0	H/ACA ribonucleoprotein complex subunit 4, putative	3	Low

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1419700	Q8ILQ3	conserved Plasmodium protein, unknown function	3	Low
PF3D7_1420700	Q8ILP3	surface protein P113	3	Low
PF3D7_1424400	Q8ILL2	60S ribosomal protein L7-3, putative	3	Low
PF3D7_1451100	Q8IKW5	elongation factor 2	3	Low
PF3D7_1453700	Q8IKU1	co-chaperone p23	3	Low
PF3D7_1464700	Q8IKJ0	ATP synthase (C/AC39) subunit, putative	3	Low
PF3D7_1468100	Q8IKF6	conserved Plasmodium protein, unknown function	3	Low
PF3D7_1474600	Q8IK92	conserved Plasmodium membrane protein, unknown function	3	Low
PF3D7_0108300	Q8I259	Uncharacterized protein	3	Low
PF3D7_0102900	Q8I2B1	Aspartate--tRNA ligase	3	Low
PF3D7_0310600	Q9NFE6	eukaryotic translation initiation factor 3 subunit K, putative, unspecified product	3	Low
PF3D7_0310600	Q9NFE6	eukaryotic translation initiation factor 3 subunit K, putative, unspecified product	3	Low
PF3D7_0503400	Q8I467	actin-depolymerizing factor 1	3	Low
PF3D7_0520900	P50250	adenosylhomocysteinase	3	Low
PF3D7_0617800	C6KT18	histone H2A	3	Low
PF3D7_0818900	Q8IB24	heat shock protein 70	3	Low
PF3D7_1361900	P61074	proliferating cell nuclear antigen 1	3	Low
PF3D7_0102200	Q8I0U6	Ring-infected erythrocyte surface antigen	3	Low
PF3D7_1136500	Q8I274	Casein kinase I	3	Low
PF3D7_0903700	Q6ZLZ9	alpha tubulin 1	3	Low
PF3D7_0207500	Q9TY96	serine repeat antigen 6	2	Noise
PF3D7_0209800	Q9TY94	ATP-dependent RNA helicase UAP56	2	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0212900	O96210	leucyl/phenylalanyl-tRNA--protein transferase, putative	2	Noise
PF3D7_0214100	O96221	protein transport protein SEC31	2	Noise
PF3D7_0217500	P62344	calcium-dependent protein kinase 1	2	Noise
PF3D7_0220000	O96275	liver stage antigen 3	2	Noise
PF3D7_0302900	O77312	exportin-1, putative	2	Noise
PF3D7_0305600	O97240	AP endonuclease (DNA-[apurinic or apyrimidinic site] lyase), putative	2	Noise
PF3D7_0308600	O77325	pre-mRNA-processing factor 19, putative	2	Noise
PF3D7_0309500	O77330	asparagine synthetase [glutamine-hydrolyzing], putative	2	Noise
PF3D7_0312800	O77364	60S ribosomal protein L26, putative	2	Noise
PF3D7_0317600	O77381	40S ribosomal protein S11, putative	2	Noise
PF3D7_0320900	O97320	histone H2A variant, putative	2	Noise
PF3D7_0322000	Q76NN7	peptidyl-prolyl cis-trans isomerase	2	Noise
PF3D7_0401800	Q8I207	Plasmodium exported protein (PHISTb), unknown function	2	Noise
PF3D7_0405400	Q8I1X5	pre-mRNA-processing-splicing factor 8, putative	2	Noise
PF3D7_0405600	Q8I1X3	conserved Plasmodium membrane protein, unknown function	2	Noise
PF3D7_0418200	Q8I1Q5	eukaryotic translation initiation factor 3 subunit M, putative	2	Noise
PF3D7_0422700	Q8IFN9	eukaryotic initiation factor 4A-III, putative	2	Noise
PF3D7_0511800	Q8I3Y8	inositol-3-phosphate synthase	2	Noise
PF3D7_0523400	Q8I3N0	DnaJ protein, putative	2	Noise
PF3D7_0608500	C6KST3	proteasome subunit alpha type-2, putative	2	Noise
PF3D7_0608700	C6KST5	T-complex protein 1 subunit zeta	2	Noise
PF3D7_0614500	C6KSY6	60S ribosomal protein L19	2	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0617000	C6KT11	mitochondrial import receptor subunit TOM40, putative	2	Noise
PF3D7_0617900	C6KT19	histone H3 variant	2	Noise
PF3D7_0627500	C6KTB1	protein DJ-1	2	Noise
PF3D7_0627700	C6KTB3	transportin	2	Noise
PF3D7_0701600	Q8IC51	Pfmc-2TM Maurer's cleft two transmembrane protein	2	Noise
PF3D7_0708800	Q8IC01	heat shock protein 110	2	Noise
PF3D7_0719600	Q8IBQ6	60S ribosomal protein L11a, putative	2	Noise
PF3D7_0721600	Q8IBN5	40S ribosomal protein S5, putative	2	Noise
PF3D7_0813300	C0H4U4	conserved Plasmodium protein, unknown function	2	Noise
PF3D7_0819600	Q8IB31	conserved Plasmodium protein, unknown function	2	Noise
PF3D7_0821400	Q8IB48	conserved Plasmodium protein, unknown function	2	Noise
PF3D7_0821700	Q8IB51	60S ribosomal protein L22, putative	2	Noise
PF3D7_0913200	Q8I320	elongation factor 1-beta	2	Noise
PF3D7_0918000	Q8I2X3	glideosome-associated protein 50	2	Noise
PF3D7_0922200	Q7K6A4	S-adenosylmethionine synthetase	2	Noise
PF3D7_0932800	Q8I2I8	importin alpha re-exporter, putative	2	Noise
PF3D7_0935900	Q8I2G1	ring-exported protein 1	2	Noise
PF3D7_1007700	Q8IJW6	transcription factor with AP2 domain(s)	2	Noise
PF3D7_1010600	Q8IJT9	eukaryotic translation initiation factor 2 subunit beta	2	Noise
PF3D7_1011800	Q8IJS7	PRE-binding protein	2	Noise
PF3D7_1029600	Q8IJA9	adenosine deaminase	2	Noise
PF3D7_1106000	Q8IIU3	RuvB-like helicase 2	2	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1108600	Q8IIR7	endoplasmic reticulum-resident calcium binding protein	2	Noise
PF3D7_1132800	Q8II36	aquaglyceroporin	2	Noise
PF3D7_1137300	Q8IHZ2	CLPTM1 domain-containing protein, putative	2	Noise
PF3D7_1238800	Q8I535	acyl-CoA synthetase	2	Noise
PF3D7_1302000	Q8IEI6	EMP1-trafficking protein	2	Noise
PF3D7_1306600	Q8IEP9	V-type proton ATPase subunit H, putative	2	Noise
PF3D7_1311500	Q8IEK3	26S protease regulatory subunit 7, putative	2	Noise
PF3D7_1323400	Q8IE82	60S ribosomal protein L23	2	Noise
PF3D7_1328100	Q8I6T3	proteasome subunit beta type-7, putative	2	Noise
PF3D7_1341200	Q8IDS6	60S ribosomal protein L18, putative	2	Noise
PF3D7_1341300	C0H5G3	60S ribosomal protein L18-2, putative	2	Noise
PF3D7_1342400	Q8IDR5	casein kinase II beta chain	2	Noise
PF3D7_1349200	Q8IDK7	glutamate--tRNA ligase, putative	2	Noise
PF3D7_1357800	C0H5I7	T-complex protein 1 subunit delta	2	Noise
PF3D7_1358800	Q8IDB0	40S ribosomal protein S15	2	Noise
PF3D7_1407900	Q7KQM4	plasmepsin I	2	Noise
PF3D7_1409800	Q8ILZ7	CUGBP Elav-like family member 2, putative	2	Noise
PF3D7_1414300	Q8ILV2	60S ribosomal protein L10, putative	2	Noise
PF3D7_1437200	Q8IL94	ribonucleoside-diphosphate reductase large subunit, putative	2	Noise
PF3D7_1437900	Q8IL88	HSP40, subfamily A, putative	2	Noise
PF3D7_1439000	Q8IL79	copper transporter	2	Noise
PF3D7_1441200	Q8IL58	60S ribosomal protein L1, putative	2	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1444300	Q8IL28	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative	2	Noise
PF3D7_1445900	Q8IL13	ATP-dependent RNA helicase DDX5, putative	2	Noise
PF3D7_1460600	Q8IKM6	inner membrane complex sub-compartment protein 3	2	Noise
PF3D7_1466300	Q8IKH3	26S proteasome regulatory subunit RPN2, putative	2	Noise
PF3D7_1470900	Q8IKC9	proteasome subunit beta type-2, putative	2	Noise
PF3D7_0106100	Q8I280	V-type proton ATPase subunit C, putative	2	Noise
PF3D7_0103200	Q8I2A8	Nucleoside transporter 4	2	Noise
PF3D7_0919000	Q8I2W3	nucleosome assembly protein	2	Noise
PF3D7_0209300	P62368	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	2	Noise
PF3D7_0308200	O77323	T-complex protein 1 subunit eta	2	Noise
PF3D7_0318200	O77375	DNA-directed RNA polymerase II subunit RPB1	2	Noise
PF3D7_0318300	O77374	conserved Plasmodium protein, unknown function	2	Noise
PF3D7_0322900	O97313	40S ribosomal protein S3A, putative	2	Noise
PF3D7_0417200	Q8I1R6	bifunctional dihydrofolate reductase-thymidylate synthase	2	Noise
PF3D7_0711000	P46468	AAA family ATPase, CDC48 subfamily	2	Noise
PF3D7_0729900	Q8IBG1	dynein heavy chain, putative	2	Noise
PF3D7_1035200	Q03400	S-antigen	2	Noise
PF3D7_1213800	Q8I5R7	proline--tRNA ligase	2	Noise
PF3D7_1252400	Q8I4R2	reticulocyte binding protein homologue 3, pseudogene	2	Noise
PF3D7_1427900	Q8ILI6	conserved protein, unknown function	2	Noise
PF3D7_1434300	Q8ILC1	Hsp70/Hsp90 organizing protein	2	Noise
PF3D7_1436000	Q8ILA4	glucose-6-phosphate isomerase	2	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1439900	Q7KQM0	triosephosphate isomerase	2	Noise
PF3D7_1453800	Q8IKU0	glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase	2	Noise
PF3D7_API02700	N/A	apicoplast ribosomal protein S12	2	Noise
PF3D7_0206800	P50498	merozoite surface protein 2	1	Noise
PF3D7_0207800	O96165	serine repeat antigen 3	1	Noise
PF3D7_0213700	O96217	conserved protein, unknown function	1	Noise
PF3D7_0214000	O96220	T-complex protein 1 subunit theta	1	Noise
PF3D7_0218000	O96260	replication factor C subunit 2, putative	1	Noise
PF3D7_0302500	O77310	cytoadherence linked asexual protein 3.1	1	Noise
PF3D7_0306800	O97247	T-complex protein 1 subunit beta	1	Noise
PF3D7_0306900	O97248	40S ribosomal protein S23, putative	1	Noise
PF3D7_0317000	O77396	proteasome subunit alpha type-3, putative	1	Noise
PF3D7_0320300	O97282	T-complex protein 1 subunit epsilon	1	Noise
PF3D7_0324100	O97307	Pfmc-2TM Maurer's cleft two transmembrane protein	1	Noise
PF3D7_0413600	Q8I1V1	26S protease regulatory subunit 6B, putative	1	Noise
PF3D7_0505800	Q8I444	small ubiquitin-related modifier	1	Noise
PF3D7_0513300	Q8I3X4	purine nucleoside phosphorylase	1	Noise
PF3D7_0513600	Q8I0W8	deoxyribodipyrimidine photo-lyase, putative	1	Noise
PF3D7_0517700	Q8I3T1	eukaryotic translation initiation factor 3 subunit B, putative	1	Noise
PF3D7_0518300	C0H4E8	proteasome subunit beta type-1, putative	1	Noise
PF3D7_0525100	Q8I3L4	acyl-CoA synthetase	1	Noise
PF3D7_0527000	Q8I3J5	DNA replication licensing factor MCM3, putative	1	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_0527500	Q8I3J0	Hsc70-interacting protein	1	Noise
PF3D7_0528200	Q8I3I5	eukaryotic translation initiation factor 3 subunit E, putative	1	Noise
PF3D7_0601200	C6KSL5	Pfmc-2TM Maurer's cleft two transmembrane protein	1	Noise
PF3D7_0605100	C6KSP9	RNA-binding protein, putative	1	Noise
PF3D7_0617200	C6KT13	conserved Plasmodium protein, unknown function	1	Noise
PF3D7_0705700	C0H4K8	40S ribosomal protein S29, putative	1	Noise
PF3D7_0728000	Q8IBH7	eukaryotic translation initiation factor 2 subunit alpha, putative	1	Noise
PF3D7_0812400	Q8IAW0	karyopherin alpha	1	Noise
PF3D7_0814000	Q8IAX6	60S ribosomal protein L13-2, putative	1	Noise
PF3D7_0822600	Q8IB60	protein transport protein SEC23	1	Noise
PF3D7_0823900	Q8IB73	dicarboxylate/tricarboxylate carrier	1	Noise
PF3D7_0903900	Q8I3B0	60S ribosomal protein L32	1	Noise
PF3D7_0904900	Q8I3A0	copper-transporting ATPase	1	Noise
PF3D7_0915400	Q8I2Z8	6-phosphofructokinase	1	Noise
PF3D7_0916500	C0H540	ubiquitin fusion degradation protein 1	1	Noise
PF3D7_0923900	Q8I2R8	RNA-binding protein, putative	1	Noise
PF3D7_0929200	C0H570	RNA-binding protein, putative	1	Noise
PF3D7_1001200	Q8IK24	acyl-CoA binding protein, isoform 2, ACBP2	1	Noise
PF3D7_1003500	Q8IK02	40S ribosomal protein S20e, putative	1	Noise
PF3D7_1004000	Q8IJZ7	60S ribosomal protein L13, putative	1	Noise
PF3D7_1012900	Q8IJR6	autophagy-related protein 18, putative	1	Noise
PF3D7_1019400	Q8IJK8	60S ribosomal protein L30e, putative	1	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1027300	Q8IJD0	peroxiredoxin	1	Noise
PF3D7_1033200	Q8IJ76	early transcribed membrane protein 10.2	1	Noise
PF3D7_1033400	Q8IJ74	haloacid dehalogenase-like hydrolase	1	Noise
PF3D7_1035700	Q8IJ52	duffy binding-like merozoite surface protein	1	Noise
PF3D7_1036900	Q8IJ39	conserved Plasmodium protein, unknown function	1	Noise
PF3D7_1103100	Q8IIX0	60S acidic ribosomal protein P1, putative	1	Noise
PF3D7_1104200	Q8IIW0	chromatin remodeling protein	1	Noise
PF3D7_1108400	Q8IIR9	casein kinase 2, alpha subunit	1	Noise
PF3D7_1109900	Q8I713	60S ribosomal protein L36	1	Noise
PF3D7_1115300	Q8I6U5	cysteine proteinase falcipain 2b	1	Noise
PF3D7_1115700	Q8I6U4	cysteine proteinase falcipain 2a	1	Noise
PF3D7_1116500	Q8IIK1	folate transporter 2	1	Noise
PF3D7_1120100	Q8IIG6	phosphoglycerate mutase, putative	1	Noise
PF3D7_1127000	Q8II93	protein phosphatase, putative	1	Noise
PF3D7_1128100	Q8II82	prefoldin subunit 5, putative	1	Noise
PF3D7_1129200	Q8II71	26S proteasome regulatory subunit RPN7, putative	1	Noise
PF3D7_1130400	Q8II60	26S protease regulatory subunit 6A, putative	1	Noise
PF3D7_1135900	Q8II05	3-oxo-5-alpha-steroid 4-dehydrogenase, putative	1	Noise
PF3D7_1142500	Q8IHU0	60S ribosomal protein L28	1	Noise
PF3D7_1144000	Q8IHS5	40S ribosomal protein S21	1	Noise
PF3D7_1227100	Q8I5E7	DNA helicase 60	1	Noise
PF3D7_1229500	Q8I5C4	T-complex protein 1 subunit gamma	1	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1239600	Q8I527	hydroxyethylthiazole kinase	1	Noise
PF3D7_1302800	Q8IET7	40S ribosomal protein S7, putative	1	Noise
PF3D7_1309100	Q8IEM3	60S ribosomal protein L24, putative	1	Noise
PF3D7_1323100	Q8IE85	60S ribosomal protein L6, putative	1	Noise
PF3D7_1331600	C0H5F1	protein tyrosine phosphatase-like protein, putative	1	Noise
PF3D7_1332900	Q8IDZ9	isoleucine--tRNA ligase, putative	1	Noise
PF3D7_1338100	Q8IDV2	26S proteasome regulatory subunit RPN3, putative	1	Noise
PF3D7_1338300	Q8IDV0	elongation factor 1-gamma, putative	1	Noise
PF3D7_1341900	Q8IDS0	V-type proton ATPase subunit D, putative	1	Noise
PF3D7_1342600	Q8IDR3	myosin A	1	Noise
PF3D7_1344200	C0H5H0	heat shock protein 110, putative	1	Noise
PF3D7_1346300	Q8IDN4	DNA/RNA-binding protein Alba 2	1	Noise
PF3D7_1347200	Q8IDM6	nucleoside transporter 1	1	Noise
PF3D7_1358700	Q8IDB8	YOP1-like protein, putative	1	Noise
PF3D7_1364100	Q8ID66	6-cysteine protein	1	Noise
PF3D7_1365500	Q8I6T0	aminomethyltransferase, mitochondrial, putative	1	Noise
PF3D7_1401800	Q8IM71	choline kinase	1	Noise
PF3D7_1417800	Q8ILR7	DNA replication licensing factor MCM2	1	Noise
PF3D7_1419200	Q8ILQ8	thioredoxin-like protein, putative	1	Noise
PF3D7_1426000	Q8ILK3	60S ribosomal protein L21	1	Noise
PF3D7_1429800	Q8ILG6	coatamer beta subunit, putative	1	Noise
PF3D7_1437000	Q8IL96	N-acetyltransferase, putative	1	Noise

Gene ID	UniProt ID	Product Description	Frequency	Confidence
PF3D7_1438900	Q8IL80	thioredoxin peroxidase 1	1	Noise
PF3D7_1446200	Q8IL11	M17 leucyl aminopeptidase	1	Noise
PF3D7_1447900	Q8IKZ6	multidrug resistance protein 2	1	Noise
PF3D7_1454700	Q8IKT2	6-phosphogluconate dehydrogenase, decarboxylating, putative	1	Noise
PF3D7_1456700	Q8IKR2	conserved Plasmodium protein, unknown function	1	Noise
PF3D7_1457000	Q8IKQ9	signal peptide peptidase	1	Noise
PF3D7_1462300	Q8IKL1	conserved Plasmodium protein, unknown function	1	Noise
PF3D7_1462800	Q8IKK7	glyceraldehyde-3-phosphate dehydrogenase	1	Noise
PF3D7_1463200	Q8IKK4	replication factor C subunit 3, putative	1	Noise
PF3D7_1474800	Q8IK90	proteasome subunit alpha type-1, putative	1	Noise
PF3D7_1477300	Q8IK62	Plasmodium exported protein (PHIST), unknown function	1	Noise
PF3D7_0109800	Q8I246	Phenylalanine--tRNA ligase alpha subunit	1	Noise
PF3D7_0111500	Q8I231	UMP-CMP kinase, putative	1	Noise

Appendix 2

Molecular docking of protein-ligand binding

A2.1 Method

A2.1.1 Ligand and protein structures

Artemisinin PDB file was obtained from www.wiki.jmole.org, while carbon centred artemisinin radical (section 1.6.1 and Figure 4.5) was generated in ChemDraw Profession 15.1 and optimised for structural energy using MM2 protocol in Chem3D 15.1 software. All protein 3D structures were retrieved from PDB database. Structures with bound inhibitors are preferable.

A2.1.2 Molecular docking

Computational docking model was performed by AutoDock Vina plugin in the UCSF Chimera software suite. The known inhibitor of each protein was removed manually. Structure was prepared for docking by Dock Prep tool (Dunbrack, 2002); remove solvents, add hydrogens and charges, default settings were selected unless solvents or ions known for binding requirement.

AutoDock Vina settings were default. Search space was defined in vicinity of the known inhibitor binding sites and not larger than 2700 $Å^3$.

A2.2 Results and discussion

A2.2.1 3D structure of carbon centred radical

The input 3D structure was optimised for the most stable conformation by MM2 function. The total 258 iterations result was as follows;

Figure A2.1 MM2 structural optimisation result for carbon centred radical artemisinin

```
Minimization terminated normally because the gradient norm is less than the
minimum gradient norm
  Stretch: 2.6257
  Bend: 20.5500
  Stretch-Bend: 1.4170
  Torsion: 11.9120
  Non-1,4 VDW: -5.4153
  1,4 VDW: 21.2078
  Dipole/Dipole: 3.3767
  Total Energy: 55.6740 kcal/mol
  Calculation ended
```

Figure A2.2 MM2 structural optimisation result for artemisinin

```
Iteration 325: Minimization terminated normally because the gradient norm is less
than the minimum gradient norm
  Stretch: 3.1754
  Bend: 19.5958
  Stretch-Bend: 1.6894
  Torsion: 8.6174
  Non-1,4 VDW: -3.4642
  1,4 VDW: 23.5598
  Dipole/Dipole: 13.3861
  Total Energy: 66.5597 kcal/mol
  Calculation ended
```

A2.2.2 Docking model of plasmepsin II (PDB 1W6I)

Parameters:

- Removed ligand: pepstatin A
- Ligand docking: artemisinin
- XYZ coordinate: 15,100,60
- XYZ size: 20,30,20

Table A2.1 AutoDock Vina result for plasmepsin II with artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-8.4	0	0	1	1	1
-8.0	1.681	4.197	1	1	1
-7.3	4.199	6.878	2	2	1
-7.0	10.969	13.339	2	2	1
-6.9	4.601	6.224	2	2	2
-6.8	2.284	5.714	0	0	0
-6.7	4.387	6.679	0	0	0
-6.6	5.312	8.193	1	1	1

Parameters:

- Removed ligand: pepstatin A
- Ligand docking: carbon centred radical artemisinin
- XYZ coordinate: 15,100,60
- XYZ size: 20,30,20

Table A2.2 AutoDock Vina result for plasmepsin II with carbon centred radical artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-7.1	0.0	0.0	0	0	0
-7.0	2.851	5.278	0	0	0
-6.8	3.156	6.108	2	1	2
-6.7	3.698	7.023	8	4	6
-6.7	4.385	7.33	2	1	2
-6.7	3.517	5.297	10	5	8
-6.5	3.67	6.317	2	1	2
-6.3	1.771	4.651	6	3	6
-6.3	2.196	4.206	4	2	4

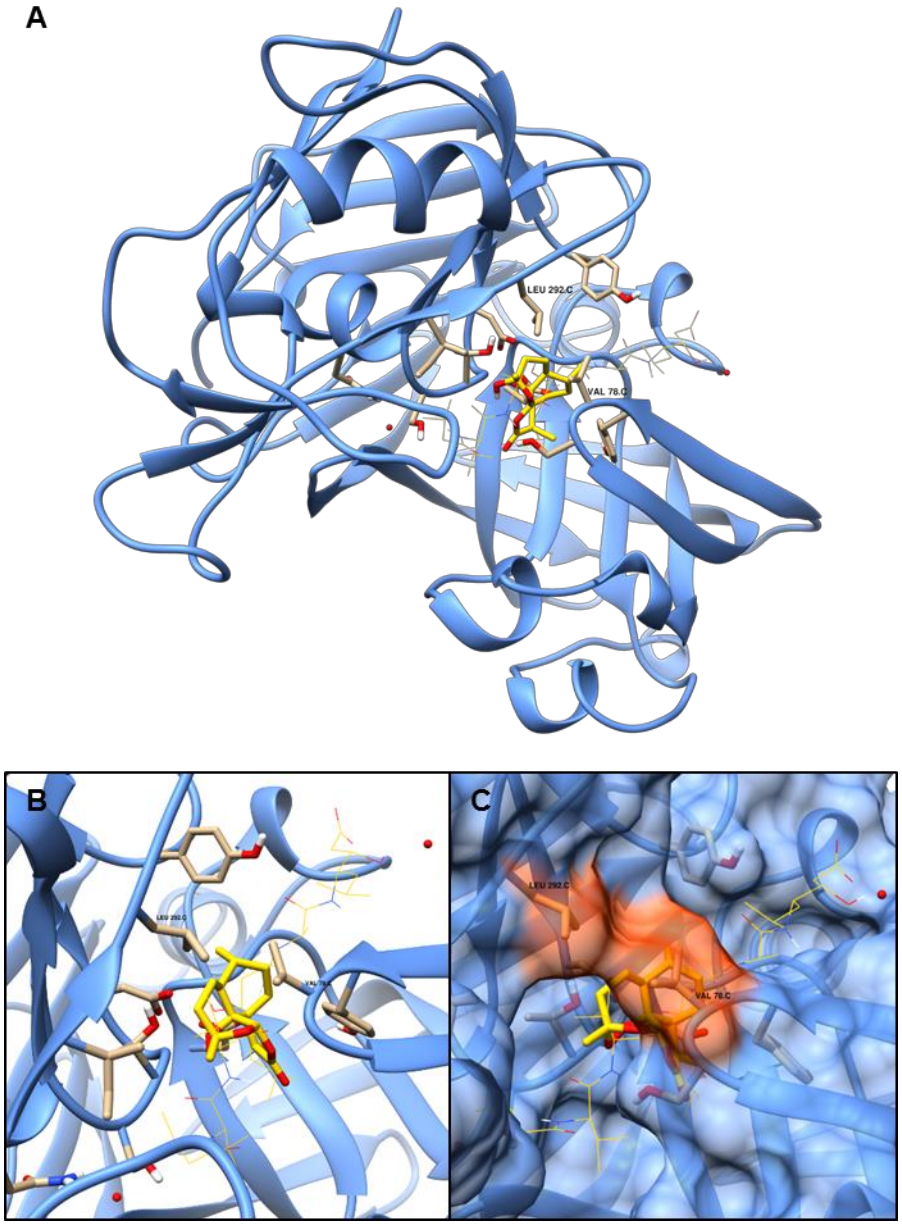


Figure A2.3 Molecular docking model of Plasmodium falciparum plasmepsin II with carbon centred radical artemisinin (carbon shown in yellow), bound inhibitor pepstatin A shown in wire yellow. (B) Artemisinin binding pocket shared with pepstatin A. (C) Close pocket formed by Val78 and Leu292 are shown in orange.

A2.2.3 Ornithine aminotransferase (*Pf*OAT, PDB 3LG0)

Parameters:

- Removed ligand: no bound ligand
- Ligand docking: artemisinin
- XYZ coordinate: 20,65,-20
- XYZ size: 30,30,30

Table A2.3 AutoDock Vina result for PfOAT with artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-7.9	0	0	3	2	2
-7.7	23.067	25.157	3	2	2
-7.7	1.905	3.292	3	2	2
-7.6	22.507	24.564	2	2	1
-7.5	1.372	4.058	3	3	3
-7.4	1.639	4.267	1	1	1
-7.3	23.724	25.517	1	1	1
-7.2	2.35	3.45	1	1	1
-7.1	22.532	24.477	1	1	1

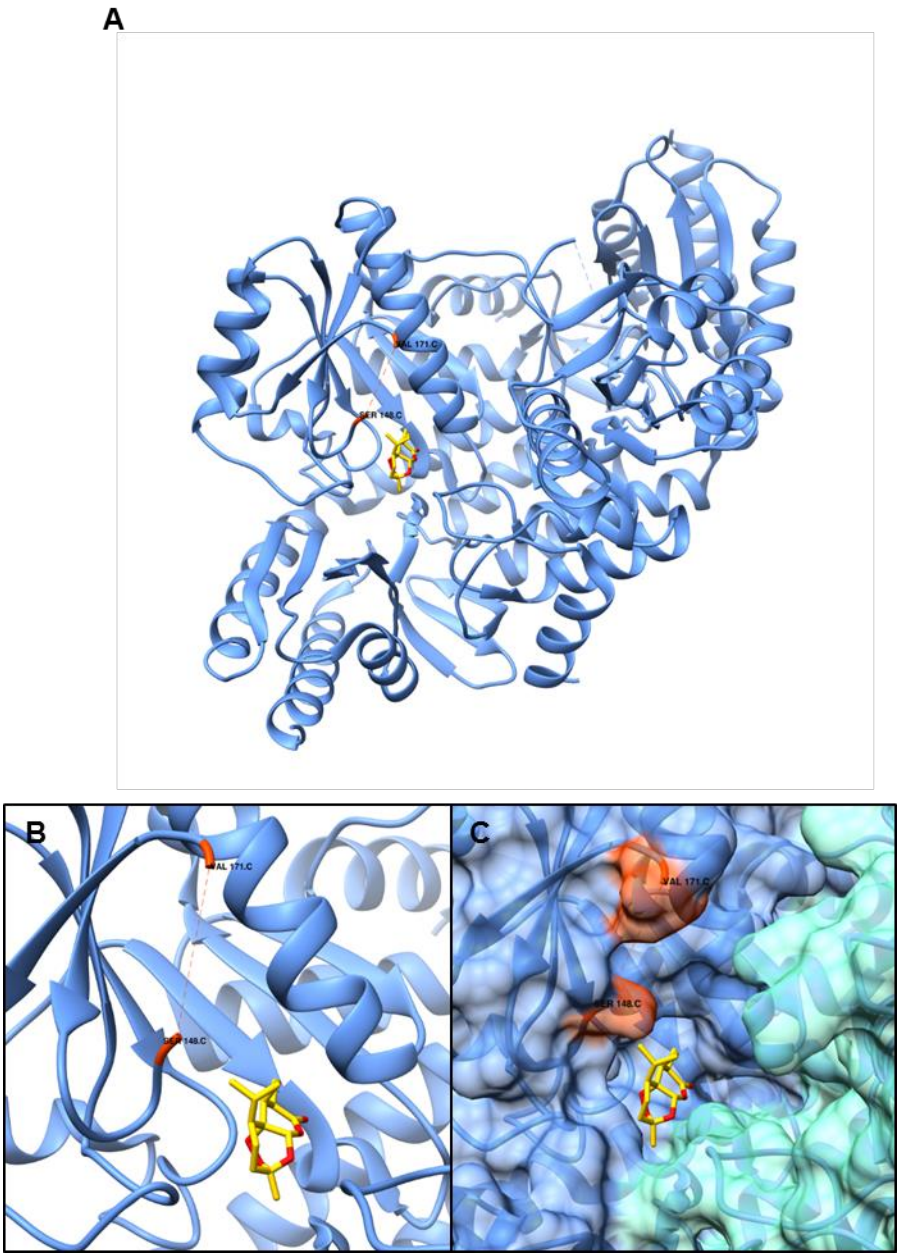


Figure A2.4 (A) Molecular docking of Plasmodium falciparum ornithine aminotransferase, OAT (PDB 3LG0) with artemisinin best fit pose (carbon shown in yellow). Residue 149-170 are shown in dashed orange line. (B) Close up image of artemisinin binding site near Ser291. (C) Binding pocket of OAT possibly under flexible loop between residue 149-170.

A2.2.4 Cell division cycle protein 48 (*Pf*CDC48)

Homology model

- Template: human P97 (PDB 5FTK),
- Homology model server: SWISS-MODEL

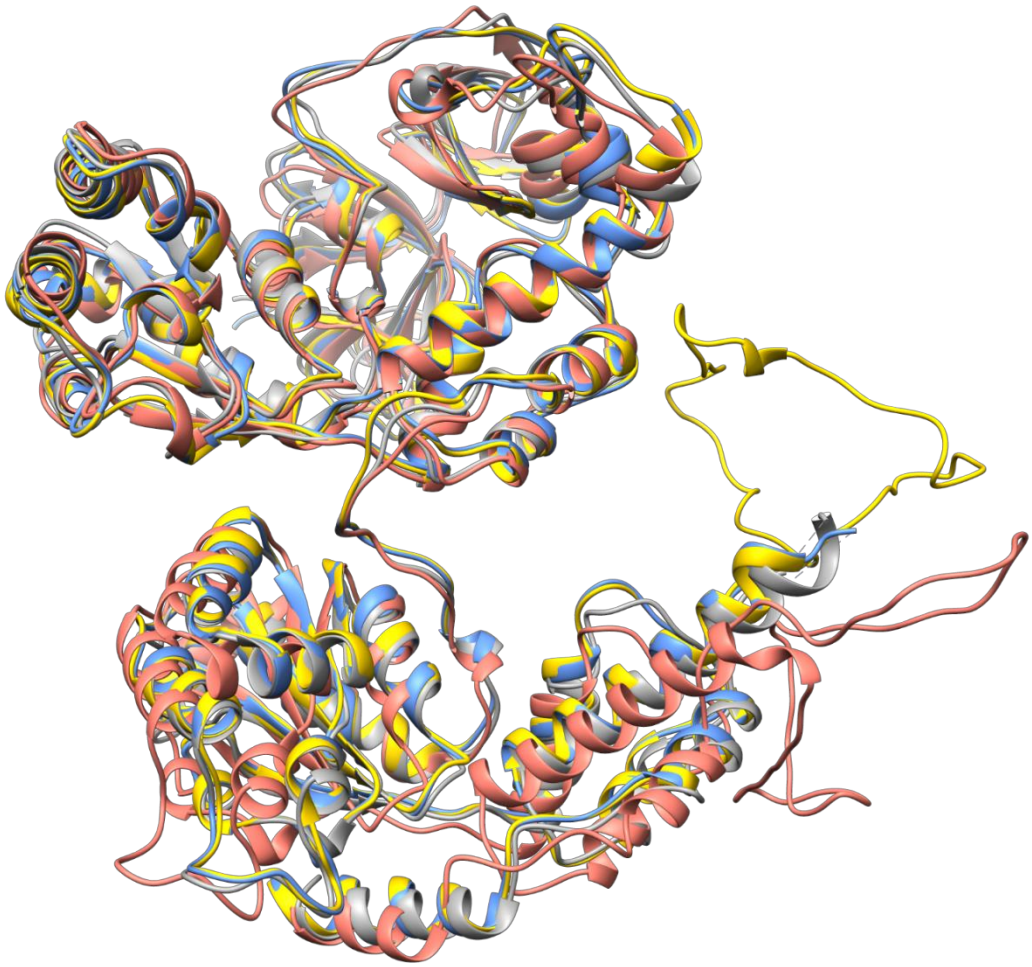


Figure A2.5 Comparison of homology models. Only chain A of each crystal structure are shown. PfCDC48 homology model from ModBase is shown in pink. PfCDC48 modelled by SWISS-MODEL is shown in yellow. Human P97 (PDB 5FTK) is shown in blue (used for SWISS-MODEL). Mouse P97 (PDB 3CF3) is shown in grey. Both crystal structure, 5FTK and 3CF3, are lacking a region between residue 707-726 due to flexibility of the protein sequence.

Parameters:

- Removed ligand: no bound ligand
- Ligand docking: artemisinin
- XYZ coordinate: 65, 130, 280
- XYZ size: 30,30,30

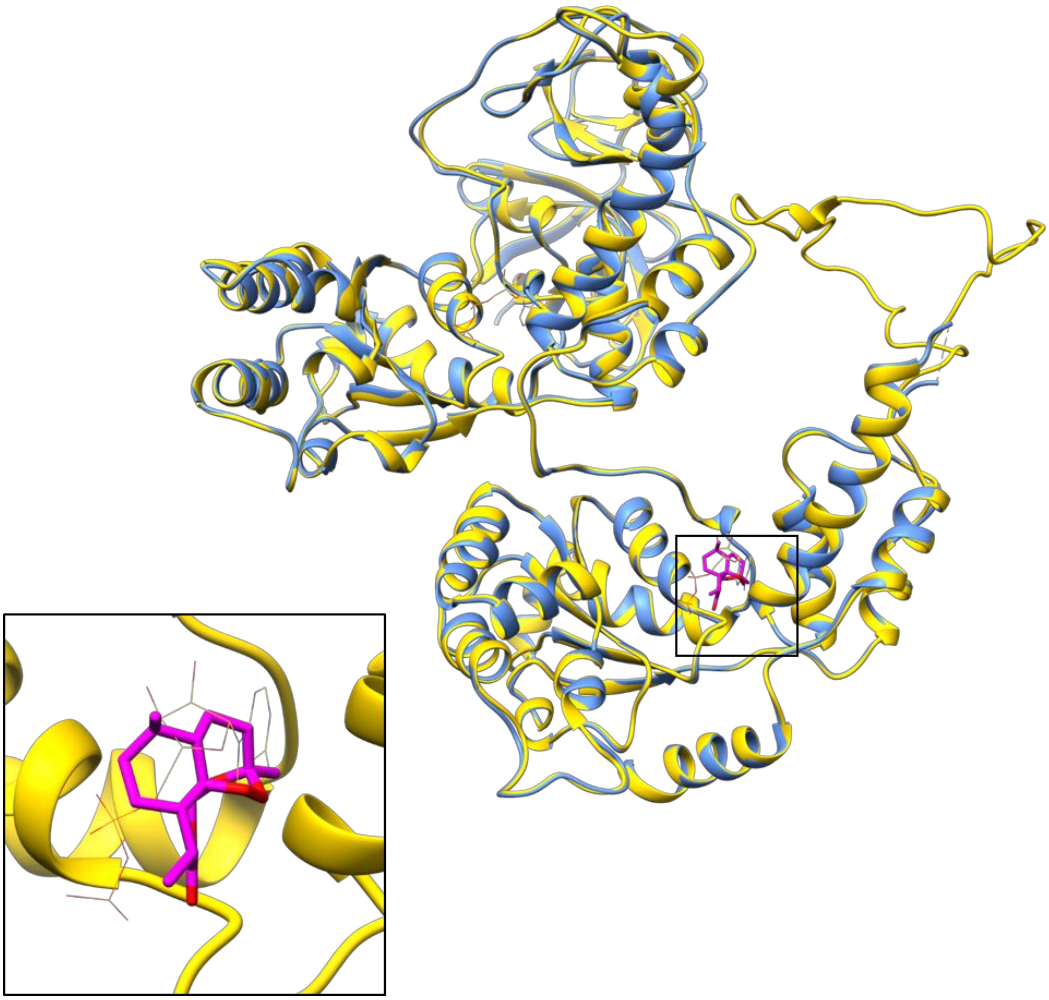


Figure A2.6 Molecular docking of artemisinin (carbon shown in magenta) with homology model of PfCDC48 shown in yellow (SWISS-MODEL). The human P97 template used to generate homology model is shown in blue (PDB 5FTK). The bound inhibitor Is shown in yellow wire (inset).

Table A2.4 AutoDock Vina result for PfCDC48 with artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-7.2	0	0	2	1	2
-7.1	1.627	4.121	0	0	0
-6.9	1.926	5.264	1	1	1
-6.5	1.524	2.228	1	1	1
-6.4	1.703	3.982	0	0	0
-6.0	5.736	8.277	5	4	5
-6.0	1.922	3.913	0	0	0
-5.9	21.264	23.172	1	1	1
-5.9	1.611	3.9	0	0	0

A2.2.5 Spermidine synthase (PDB 2I7C)

Parameters:

- Removed ligand: AdoDATO
- Ligand docking: artemisinin
- XYZ coordinate: 17,115,30 (chain A)
- XYZ size: 20,20,20

Table A2.5 AutoDock Vina result for HGXPRT with artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-4.3	0.0	0.0	0	0	0
-4.2	1.36	4.021	0	0	0
-4.1	1.452	3.899	0	0	0
-4.0	5.118	6.704	1	1	1
-4.0	2.266	5.209	0	0	0
-3.9	1.99	4.693	1	1	1
-3.3	1.941	4.859	0	0	0
-3.1	4.54	7.144	2	2	1
-2.2	15.792	17.267	3	2	2

A2.2.6 Hypoxanthine guanine xanthine phosphoribosyltransferase (HGXPRT, PDB 3OZG)

Parameters:

- Removed ligand: S-SerMe-ImmH phosphonate
- Remaining ligand: Mg^{2+}
- Ligand docking: artemisinin
- XYZ coordinate: 60,35,75 (chain A)
- XYZ size: 20,15,15

Table A2.6 AutoDock Vina result for HGXPRT with artemisinin

Score	RMSD lb	RMSD ub	Hbond all	Hbond ligand atoms	Hbond receptor atoms
-10.3	0.0	0.0	2	1	2
-9.8	1.728	4.163	2	1	2
-9.5	2.085	4.019	2	1	2
-9.3	2.11	4.93	3	2	2
-9.1	1.517	3.964	2	1	2
-8.7	2.03	4.244	2	1	2
-8.7	1.651	3.856	2	1	2
-8.4	1.941	4.395	0	0	0
-8.2	1.776	2.705	0	0	0

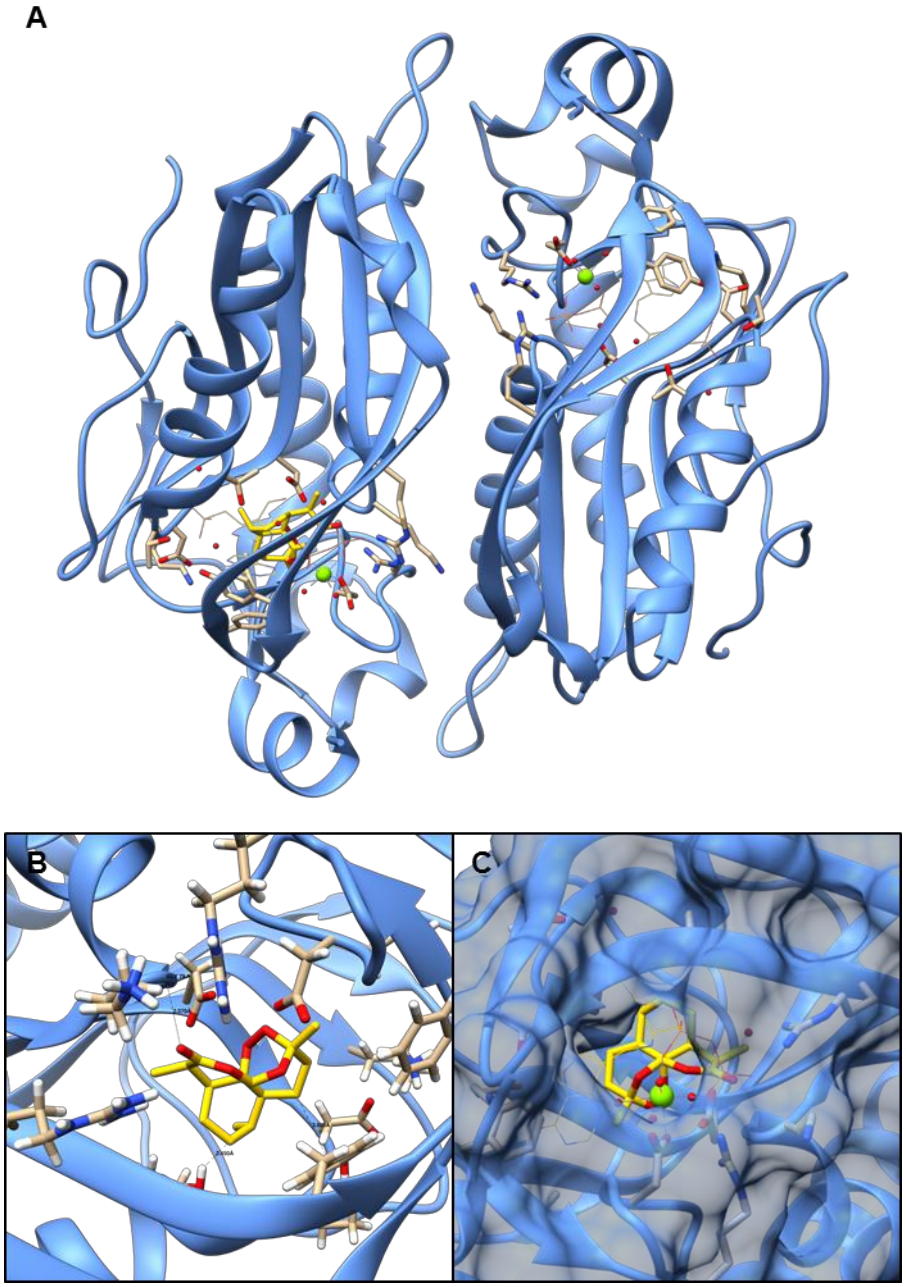


Figure A 2.7 Molecular docking model of PfHGXPRT with artemisinin

Appendix 3

Comparative genomics

A3.1 Sequencing read summary

Table A3.1 Trimmed reads summary

Batch	Sample	Read type	Number of Reads	Maximum Read Length (bp.)	Mean Read Length (bp.)	Read Length SD (bp.)	Minimum Base Quality Score	Maximum Base Quality Score	Mean Base Quality Score	Base Quality Score Standard Deviation	A content (%)	C content (%)	G content (%)	T content (%)	N content (%)
LIMS2208	TE616	R0	264155	100	85.9	24.7	2	41	34.6	6.7	39.3	11.4	10.7	38.2	0.2
		R1	10977540	100	95.4	14.9	2	41	37	5.6	40.2	10	10.2	39.2	0.4
		R2	10977540	100	93.9	17.4	2	41	37.3	5	40.7	10	10.3	39	0
	TE1201	R0	270284	100	85	25.4	2	41	34.4	6.8	38.3	12	10.7	38.8	0.2
		R1	12050734	100	95.7	14.4	2	41	37	5.7	39.2	10.7	10.6	39.2	0.4
		R2	12050734	100	94.5	16.5	2	41	37.3	5	39.7	10.6	10.7	39	0
	TE1211	R0	596719	100	86.8	24.1	2	41	34.4	6.8	36.8	13.8	12.3	37	0.2
		R1	27788412	100	95.9	14.2	2	41	36.9	5.7	38.4	11.5	11.4	38.3	0.4
		R2	27788412	100	94.7	16.2	2	41	37.2	5.1	38.9	11.5	11.5	38.2	0
	TE1389	R0	222182	100	84.8	25.7	2	41	34.4	6.9	37.5	12.9	11.7	37.6	0.2
		R1	9091159	100	95.5	14.8	2	41	36.9	5.7	38.7	11.1	11	38.7	0.4
		R2	9091159	100	94.3	16.8	2	41	37.3	5	39.2	11.1	11.1	38.6	0

Batch	Sample	Read type	Number of Reads	Maximum Read Length (bp.)	Mean Read Length (bp.)	Read Length SD (bp.)	Minimum Base Quality Score	Maximum Base Quality Score	Mean Base Quality Score	Base Quality Score Standard Deviation	A content (%)	C content (%)	G content (%)	T content (%)	N content (%)
	TE1411	R0	382010	100	86.4	24.6	2	41	34.7	6.8	36.7	13.9	12.8	36.3	0.2
		R1	13022321	100	95.5	14.8	2	41	36.8	5.8	38.2	11.9	11.8	37.7	0.4
		R2	13022321	100	94.5	16.6	2	41	37.2	5.1	38.5	11.9	11.9	37.7	0
	TE1419	R0	278956	100	86	24.8	2	41	34.6	6.7	39.4	11.3	10.6	38.5	0.2
		R1	11194907	100	95.4	14.9	2	41	37	5.6	39.9	10.2	10.3	39.2	0.4
		R2	11194907	100	94	17.4	2	41	37.3	5	40.5	10.2	10.4	39	0
	TE1432	R0	401561	100	86	24.8	2	41	34.6	6.8	38.7	11.8	10.8	38.4	0.2
		R1	16878116	100	95.5	14.8	2	41	37	5.7	39.4	10.5	10.5	39.1	0.4
		R2	16878116	100	94.2	17	2	41	37.3	5	39.9	10.5	10.7	38.9	0
	TE1439	R0	292244	100	85.8	24.9	2	41	34.5	6.8	38.4	12.1	11	38.3	0.2
		R1	11541471	100	95.6	14.7	2	41	36.9	5.7	38.9	11	10.9	38.8	0.4
		R2	11541471	100	94.4	16.7	2	41	37.2	5	39.5	11	11.1	38.4	0
	TE1475	R0	356017	100	86.4	24.5	2	41	34.7	6.7	38.7	11.8	10.9	38.4	0.2
		R1	13325079	100	95.7	14.5	2	41	36.9	5.7	39.1	10.9	10.8	38.8	0.4
		R2	13325079	100	94.5	16.6	2	41	37.3	5	39.5	10.9	11	38.6	0
LIMS4913	TE1304	R0	164219	100	89.1	22.3	2	41	35.9	6.1	37.2	12.9	11.6	38.3	0
		R1	8605982	100	95.8	14.3	2	41	37.6	4.7	39.1	10.9	10.8	39.3	0
		R2	8605982	100	95.3	15.1	2	41	37.7	4.6	39.7	10.8	10.8	38.7	0

Batch	Sample	Read type	Number of Reads	Maximum Read Length (bp.)	Mean Read Length (bp.)	Read Length SD (bp.)	Minimum Base Quality Score	Maximum Base Quality Score	Mean Base Quality Score	Base Quality Score Standard Deviation	A content (%)	C content (%)	G content (%)	T content (%)	N content (%)
		R0	278406	100	84.3	24.6	2	41	34.6	6.2	39	10	8.5	42.5	0
		R1	12656341	100	95.7	14.2	2	41	37.3	4.7	40.8	9.5	9.3	40.4	0
		R2	12656341	100	93.6	18	2	41	37	5	41.5	9.5	9.4	39.6	0
	TE1328	R0	246638	100	90.9	20.7	2	41	36	6	36.6	13.5	11.7	38.2	0
		R1	12854121	100	96.8	12.6	2	41	37.7	4.6	38.5	11.2	11.1	39.2	0
		R2	12854121	100	95.9	14.2	2	41	37.8	4.6	39.1	11.1	11.1	38.6	0
		R0	421340	100	84.9	24	2	41	34.6	6.1	38.8	10.2	8.6	42.4	0
		R1	19783597	100	96.1	13.5	2	41	37.3	4.7	40.4	9.7	9.4	40.5	0
		R2	19783597	100	94	17.5	2	41	37	5	41.1	9.6	9.6	39.7	0
	TE1368	R0	271644	100	89	22.6	2	41	35.8	6.1	36.2	14.2	12.3	37.3	0
		R1	12303744	100	96.1	13.9	2	41	37.6	4.8	38.1	11.7	11.7	38.5	0
		R2	12303744	100	95.4	15	2	41	37.6	4.7	38.7	11.7	11.6	38	0
		R0	429962	100	84.8	24.2	2	41	34.5	6.2	38.3	10.9	9.1	41.7	0
		R1	17085191	100	95.9	13.8	2	41	37.2	4.7	40.1	10.2	9.9	39.9	0
		R2	17085191	100	93.6	18	2	41	36.9	5.1	40.8	10.1	10	39.1	0
	TE1373	R0	149981	100	89.3	22.1	2	41	35.9	6	37.6	12.5	11.2	38.6	0
		R1	8269167	100	95.8	14.3	2	41	37.7	4.7	39.4	10.5	10.5	39.5	0
		R2	8269167	100	95.2	15.3	2	41	37.8	4.6	40.1	10.5	10.6	38.8	0

Batch	Sample	Read type	Number of Reads	Maximum Read Length (bp.)	Mean Read Length (bp.)	Read Length SD (bp.)	Minimum Base Quality Score	Maximum Base Quality Score	Mean Base Quality Score	Base Quality Score Standard Deviation	A content (%)	C content (%)	G content (%)	T content (%)	N content (%)
		R0	273857	100	84	24.8	2	41	34.7	6.1	39.4	9.5	8.2	42.9	0
		R1	12836726	100	95.6	14.4	2	41	37.3	4.7	41.2	9.2	9	40.6	0
		R2	12836726	100	93.5	18.2	2	41	37.1	5	42	9.1	9.2	39.7	0
	TE1432	R0	157838	100	89.7	21.9	2	41	35.9	6.1	37.2	13.1	11.7	38	0
		R1	7978437	100	96	14	2	41	37.6	4.7	39	10.9	11	39.1	0
		R2	7978437	100	95.4	14.9	2	41	37.7	4.7	39.7	10.9	11.1	38.3	0
		R0	280284	100	84.7	24.3	2	41	34.6	6.1	39.2	10	8.6	42.2	0
		R1	12624168	100	95.7	14.2	2	41	37.3	4.7	40.8	9.5	9.3	40.3	0
		R2	12624168	100	93.6	18	2	41	37	5	41.7	9.4	9.5	39.4	0
	TE1435	R0	293041	100	91	20.6	2	41	36	6	36.5	13.5	11.7	38.2	0
		R1	14524603	100	96.9	12.5	2	41	37.7	4.7	38.3	11.3	11.2	39.2	0
		R2	14524603	100	96	14	2	41	37.7	4.7	38.9	11.3	11.3	38.5	0
		R0	479907	100	85	24.1	2	41	34.5	6.2	38.8	10.5	8.7	42.1	0
		R1	21927102	100	96.1	13.4	2	41	37.3	4.7	40.2	9.8	9.6	40.4	0
		R2	21927102	100	94.1	17.4	2	41	37	5	40.9	9.8	9.7	39.6	0
	TE1436	R0	230412	100	88.4	23	2	41	35.9	6	37.5	12.4	11.8	38.2	0
		R1	10986472	100	95	15.4	2	41	37.5	4.8	39.3	10.7	10.8	39.2	0
		R2	10986472	100	94.8	15.8	2	41	37.7	4.7	40	10.7	10.8	38.5	0

Batch	Sample	Read type	Number of Reads	Maximum Read Length (bp.)	Mean Read Length (bp.)	Read Length SD (bp.)	Minimum Base Quality Score	Maximum Base Quality Score	Mean Base Quality Score	Base Quality Score Standard Deviation	A content (%)	C content (%)	G content (%)	T content (%)	N content (%)
		R0	372670	100	84.3	24.8	2	41	34.8	6.1	39.1	9.9	8.7	42.3	0
		R1	15488183	100	95.5	14.5	2	41	37.2	4.7	40.8	9.6	9.6	40	0
		R2	15488183	100	93.4	18.2	2	41	37	5	41.6	9.6	9.7	39.1	0

Note R0 is unpaired read, R1 and R2 are forward and reverse paired-reads, respectively.

A3.2 Read coverage

Read coverage was calculated as;

$$\frac{(R1 + R2) \times \bar{X}_{R1R2}}{Genome\ size}$$

Where genome size is 22,900,000 bp.

Table A3.2 Read coverage for each sample

Sample	Read Coverage (%)
TE616	90.74447
TE1201	100.0895
TE1211	231.287
TE1389	75.34943
TE1411	108.0455
TE1419	92.59019
TE1432	139.8157
TE1439	95.75893
TE1475	110.6738
TE1304	71.81673
	104.6221
TE1328	108.1655
	164.2298
TE1368	102.8894
	141.3818
TE1373	68.96991
	106.0011
TE1432	66.6844
	104.3561
TE1435	122.3492
	182.1194
TE1436	91.05818
	127.7606

A3.3 Read mapping coverage

The read mapping coverage was calculated by BOWTIE2 and reported in BOWTIE stat file.

TE616

10977540 reads; of these:
  10977540 (100.00%) were paired; of these:
    1015848 (9.25%) aligned concordantly 0 times
    8667646 (78.96%) aligned concordantly exactly 1 time
    1294046 (11.79%) aligned concordantly >1 times

    1015848 pairs aligned concordantly 0 times; of these:
      389301 (38.32%) aligned discordantly 1 time

    626547 pairs aligned 0 times concordantly or discordantly; of these:
      1253094 mates make up the pairs; of these:
        858139 (68.48%) aligned 0 times
        231410 (18.47%) aligned exactly 1 time
        163545 (13.05%) aligned >1 times
96.09% overall alignment rate

TE1201

12050734 reads; of these:
  12050734 (100.00%) were paired; of these:
    1089637 (9.04%) aligned concordantly 0 times
    9538048 (79.15%) aligned concordantly exactly 1 time
    1423049 (11.81%) aligned concordantly >1 times

    1089637 pairs aligned concordantly 0 times; of these:
      593953 (54.51%) aligned discordantly 1 time

    495684 pairs aligned 0 times concordantly or discordantly; of these:
      991368 mates make up the pairs; of these:
        408100 (41.17%) aligned 0 times
        357402 (36.05%) aligned exactly 1 time
        225866 (22.78%) aligned >1 times
98.31% overall alignment rate

TE1211

27788412 reads; of these:

27788412 (100.00%) were paired; of these:

2770231 (9.97%) aligned concordantly 0 times

22001621 (79.18%) aligned concordantly exactly 1 time

3016560 (10.86%) aligned concordantly >1 times

2770231 pairs aligned concordantly 0 times; of these:

496302 (17.92%) aligned discordantly 1 time

2273929 pairs aligned 0 times concordantly or discordantly; of these:

4547858 mates make up the pairs; of these:

4053423 (89.13%) aligned 0 times

271589 (5.97%) aligned exactly 1 time

222846 (4.90%) aligned >1 times

92.71% overall alignment rate

TE1304

21262323 reads; of these:

21262323 (100.00%) were paired; of these:

5573327 (26.21%) aligned concordantly 0 times

13573208 (63.84%) aligned concordantly exactly 1 time

2115788 (9.95%) aligned concordantly >1 times

5573327 pairs aligned concordantly 0 times; of these:

3094753 (55.53%) aligned discordantly 1 time

2478574 pairs aligned 0 times concordantly or discordantly; of these:

4957148 mates make up the pairs; of these:

1983178 (40.01%) aligned 0 times

1934555 (39.03%) aligned exactly 1 time

1039415 (20.97%) aligned >1 times

95.34% overall alignment rate

TE1328

32637718 reads; of these:
  32637718 (100.00%) were paired; of these:
    1581027 (4.84%) aligned concordantly 0 times
    27087844 (83.00%) aligned concordantly exactly 1 time
    3968847 (12.16%) aligned concordantly >1 times

    1581027 pairs aligned concordantly 0 times; of these:
      693492 (43.86%) aligned discordantly 1 time

    887535 pairs aligned 0 times concordantly or discordantly; of these:
      1775070 mates make up the pairs; of these:
        1027722 (57.90%) aligned 0 times
        441018 (24.85%) aligned exactly 1 time
        306330 (17.26%) aligned >1 times
98.43% overall alignment rate

TE1368

29388935 reads; of these:
  29388935 (100.00%) were paired; of these:
    4996138 (17.00%) aligned concordantly 0 times
    20720980 (70.51%) aligned concordantly exactly 1 time
    3671817 (12.49%) aligned concordantly >1 times

    4996138 pairs aligned concordantly 0 times; of these:
      2475748 (49.55%) aligned discordantly 1 time

    2520390 pairs aligned 0 times concordantly or discordantly; of these:
      5040780 mates make up the pairs; of these:
        2295015 (45.53%) aligned 0 times
        1630959 (32.36%) aligned exactly 1 time
        1114806 (22.12%) aligned >1 times
96.10% overall alignment rate

TE1373

21105893 reads; of these:
  21105893 (100.00%) were paired; of these:

4620803 (21.89%) aligned concordantly 0 times

14333349 (67.91%) aligned concordantly exactly 1 time

2151741 (10.19%) aligned concordantly >1 times

4620803 pairs aligned concordantly 0 times; of these:

2611912 (56.53%) aligned discordantly 1 time

2008891 pairs aligned 0 times concordantly or discordantly; of these:

4017782 mates make up the pairs; of these:

1640599 (40.83%) aligned 0 times

1567983 (39.03%) aligned exactly 1 time

809200 (20.14%) aligned >1 times

96.11% overall alignment rate

TE1389

9091159 reads; of these:

9091159 (100.00%) were paired; of these:

1500572 (16.51%) aligned concordantly 0 times

6601422 (72.61%) aligned concordantly exactly 1 time

989165 (10.88%) aligned concordantly >1 times

1500572 pairs aligned concordantly 0 times; of these:

582253 (38.80%) aligned discordantly 1 time

918319 pairs aligned 0 times concordantly or discordantly; of these:

1836638 mates make up the pairs; of these:

1246709 (67.88%) aligned 0 times

372883 (20.30%) aligned exactly 1 time

217046 (11.82%) aligned >1 times

93.14% overall alignment rate

TE1411

13022321 reads; of these:

13022321 (100.00%) were paired; of these:

2763738 (21.22%) aligned concordantly 0 times

8958299 (68.79%) aligned concordantly exactly 1 time

1300284 (9.99%) aligned concordantly >1 times

2763738 pairs aligned concordantly 0 times; of these:

549605 (19.89%) aligned discordantly 1 time

2214133 pairs aligned 0 times concordantly or discordantly; of these:

4428266 mates make up the pairs; of these:

3750837 (84.70%) aligned 0 times

425074 (9.60%) aligned exactly 1 time

252355 (5.70%) aligned >1 times

85.60% overall alignment rate

TE1419

11194907 reads; of these:

11194907 (100.00%) were paired; of these:

1077527 (9.63%) aligned concordantly 0 times

8814137 (78.73%) aligned concordantly exactly 1 time

1303243 (11.64%) aligned concordantly >1 times

1077527 pairs aligned concordantly 0 times; of these:

611219 (56.72%) aligned discordantly 1 time

466308 pairs aligned 0 times concordantly or discordantly; of these:

932616 mates make up the pairs; of these:

360204 (38.62%) aligned 0 times

338219 (36.27%) aligned exactly 1 time

234193 (25.11%) aligned >1 times

98.39% overall alignment rate

TE1423

16878116 reads; of these:

16878116 (100.00%) were paired; of these:

1801694 (10.67%) aligned concordantly 0 times

13056598 (77.36%) aligned concordantly exactly 1 time

2019824 (11.97%) aligned concordantly >1 times

1801694 pairs aligned concordantly 0 times; of these:

992077 (55.06%) aligned discordantly 1 time

809617 pairs aligned 0 times concordantly or discordantly; of these:

1619234 mates make up the pairs; of these:

636092 (39.28%) aligned 0 times

584354 (36.09%) aligned exactly 1 time

398788 (24.63%) aligned >1 times

98.12% overall alignment rate

TE1432

20602605 reads; of these:

20602605 (100.00%) were paired; of these:

4517564 (21.93%) aligned concordantly 0 times

13903087 (67.48%) aligned concordantly exactly 1 time

2181954 (10.59%) aligned concordantly >1 times

4517564 pairs aligned concordantly 0 times; of these:

2447466 (54.18%) aligned discordantly 1 time

2070098 pairs aligned 0 times concordantly or discordantly; of these:

4140196 mates make up the pairs; of these:

1655202 (39.98%) aligned 0 times

1581143 (38.19%) aligned exactly 1 time

903851 (21.83%) aligned >1 times

95.98% overall alignment rate

TE1435

36451705 reads; of these:

36451705 (100.00%) were paired; of these:

1511582 (4.15%) aligned concordantly 0 times

30435822 (83.50%) aligned concordantly exactly 1 time

4504301 (12.36%) aligned concordantly >1 times

1511582 pairs aligned concordantly 0 times; of these:

610942 (40.42%) aligned discordantly 1 time

900640 pairs aligned 0 times concordantly or discordantly; of these:

1801280 mates make up the pairs; of these:

1131747 (62.83%) aligned 0 times

373890 (20.76%) aligned exactly 1 time

295643 (16.41%) aligned >1 times

98.45% overall alignment rate

TE1436

26474655 reads; of these:

26474655 (100.00%) were paired; of these:

15646532 (59.10%) aligned concordantly 0 times

9048895 (34.18%) aligned concordantly exactly 1 time

1779228 (6.72%) aligned concordantly >1 times

15646532 pairs aligned concordantly 0 times; of these:

9178412 (58.66%) aligned discordantly 1 time

6468120 pairs aligned 0 times concordantly or discordantly; of these:

12936240 mates make up the pairs; of these:

5046379 (39.01%) aligned 0 times

5451541 (42.14%) aligned exactly 1 time

2438320 (18.85%) aligned >1 times

90.47% overall alignment rate

TE1439

11541471 reads; of these:

11541471 (100.00%) were paired; of these:

1603009 (13.89%) aligned concordantly 0 times

8593516 (74.46%) aligned concordantly exactly 1 time

1344946 (11.65%) aligned concordantly >1 times

1603009 pairs aligned concordantly 0 times; of these:

894900 (55.83%) aligned discordantly 1 time

708109 pairs aligned 0 times concordantly or discordantly; of these:

1416218 mates make up the pairs; of these:

539362 (38.08%) aligned 0 times

533875 (37.70%) aligned exactly 1 time

342981 (24.22%) aligned >1 times

97.66% overall alignment rate

TE1475

13325079 reads; of these:

13325079 (100.00%) were paired; of these:

1400361 (10.51%) aligned concordantly 0 times

10387800 (77.96%) aligned concordantly exactly 1 time

1536918 (11.53%) aligned concordantly >1 times

1400361 pairs aligned concordantly 0 times; of these:

762493 (54.45%) aligned discordantly 1 time

637868 pairs aligned 0 times concordantly or discordantly; of these:

1275736 mates make up the pairs; of these:

518827 (40.67%) aligned 0 times

459426 (36.01%) aligned exactly 1 time

297483 (23.32%) aligned >1 times

98.05% overall alignment rate

A3.4 SnpEff software

The classification of SNP effect prediction by SnpEff software was implemented from the VCF annotation standard 'ANN' field version 4.3a (28th October 2015) as follows;

Table A3.3 SNP effect prediction

Sequence	Ontology Term
HIGH	chromosome_number_variation
HIGH	exon_loss_variant
HIGH	frameshift_variant
HIGH	rare_amino_acid_variant
HIGH	splice_acceptor_variant
HIGH	splice_donor_variant
HIGH	start_lost
HIGH	stop_gained
HIGH	stop_lost
HIGH	transcript_ablation
MODERATE	3_prime_UTR_truncation+exon_loss
MODERATE	5_prime_UTR_truncation+exon_loss_variant
MODERATE	coding_sequence_variant
MODERATE	disruptive_inframe_deletion
MODERATE	disruptive_inframe_insertion
MODERATE	inframe_deletion
MODERATE	inframe_insertion
MODERATE	missense_variant
MODERATE	regulatory_region_ablation
MODERATE	splice_region_variant
MODERATE	TFBS_ablation
LOW	5_prime_UTR_premature start_codon_gain_variant
LOW	initiator_codon_variant
LOW	splice_region_variant
LOW	start_retained
LOW	stop_retained_variant
LOW	synonymous_variant
MODIFIER	3_prime_UTR_variant

MODIFIER	5_prime_UTR_variant
MODIFIER	coding_sequence_variant
MODIFIER	conserved_intergenic_variant
MODIFIER	conserved_intron_variant
MODIFIER	downstream_gene_variant
MODIFIER	exon_variant
MODIFIER	feature_elongation
MODIFIER	feature_truncation
MODIFIER	gene_variant
MODIFIER	intergenic_region
MODIFIER	intragenic_variant
MODIFIER	intron_variant
MODIFIER	mature_miRNA_variant
MODIFIER	miRNA
MODIFIER	NMD_transcript_variant
MODIFIER	non_coding_transcript_exon_variant
MODIFIER	non_coding_transcript_variant
MODIFIER	regulatory_region_amplification
MODIFIER	regulatory_region_variant
MODIFIER	TF_binding_site_variant
MODIFIER	TFBS_amplification
MODIFIER	transcript_amplification
MODIFIER	transcript_variant
MODIFIER	upstream_gene_variant

A3.5 Genomic DNA library preparation

Genomic DNA samples were summited to CGR, UoL, for sequencing. All the library preparation step was performed by the CGR members, Dr Pia Koldkjaer and Dr Anita Lucaci. DNA libraries were prepared by commercial kit, Illumina® TruSeq™ DNA sample preparation kit. Adapters, indexes, and primer sequences were ligated to fragmented DNA sample. Prepared libraries were measured by Agilent 2100 Bioanalyzer for fragment lengths. Then the range of desired insertion sizes was determined for each sample. Next, size selection of prepared DNA libraries by gel-based assay, Pippen™ Prep (Sage Science) was performed to obtain desired length fragments, and confirmed by Agilent 2100 Bioanalyzer.

During library preparation step, there were 6 and 5 PCR cycles involved for LIMS2208 and LIMS4913, respectively. Defined fragment length libraries for LIMS2208 and LIMS4319 were 420-650 bp and 400-700 pb, respectively.

Table A3.4 Illumina® quality scores

Quality Score	Error Probability
Q40	0.0001 (1 in 10,000)
Q30	0.001 (1 in 1,000)
Q20	0.01 (1 in 100)
Q10	0.1 (1 in 10)

A3.6 Comparative sequence analysis of IspG

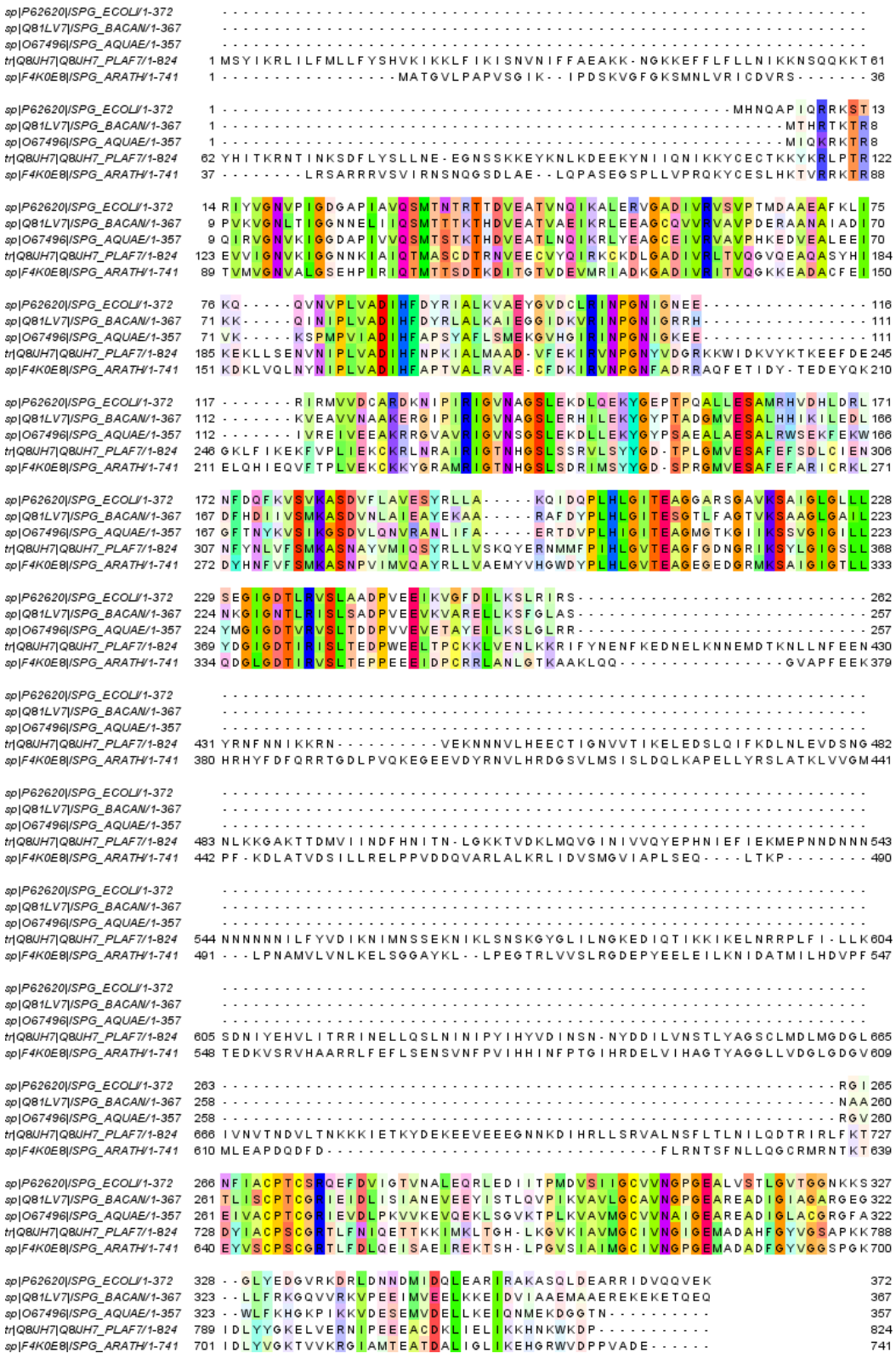


Figure A3.1 Protein sequence alignment comparison of IspG of Escherichia coli (ECOLI), Oryza sativa (BACAN), Aquifex aeolicus (AQUE), Plasmodium falciparum (PLAF7), and Arabidopsis thaliana (ARATH), from top to bottom. Sequence alignment was performed on Clustal Omega web server and coloured by Taylor conservation pattern. All compared protein sequences share similar domain A and B, while parasite (PLAF7) and plant (ARATH) have inserted domain in between.


sp|Q8KG23|/SPG_CHLTE/1-746
sp|Q72TR2|/SPG_LEPIC/1-663
tr|Q8UH7|Q8UH7_PLAF7/1-824
tr|A7AUN2|A7AUN2_BABBO/1-772
tr|B8C7D5|B8C7D5_THAPS/1-686
tr|A9SJI1|A9SJI1_PHYPA/1-691
sp|Q6KBJ4|/SPG_ORYSJ/1-744
tr|A9CBE3|A9CBE3_GINBI/1-741
sp|F4K0E8|/SPG_ARATH/1-741

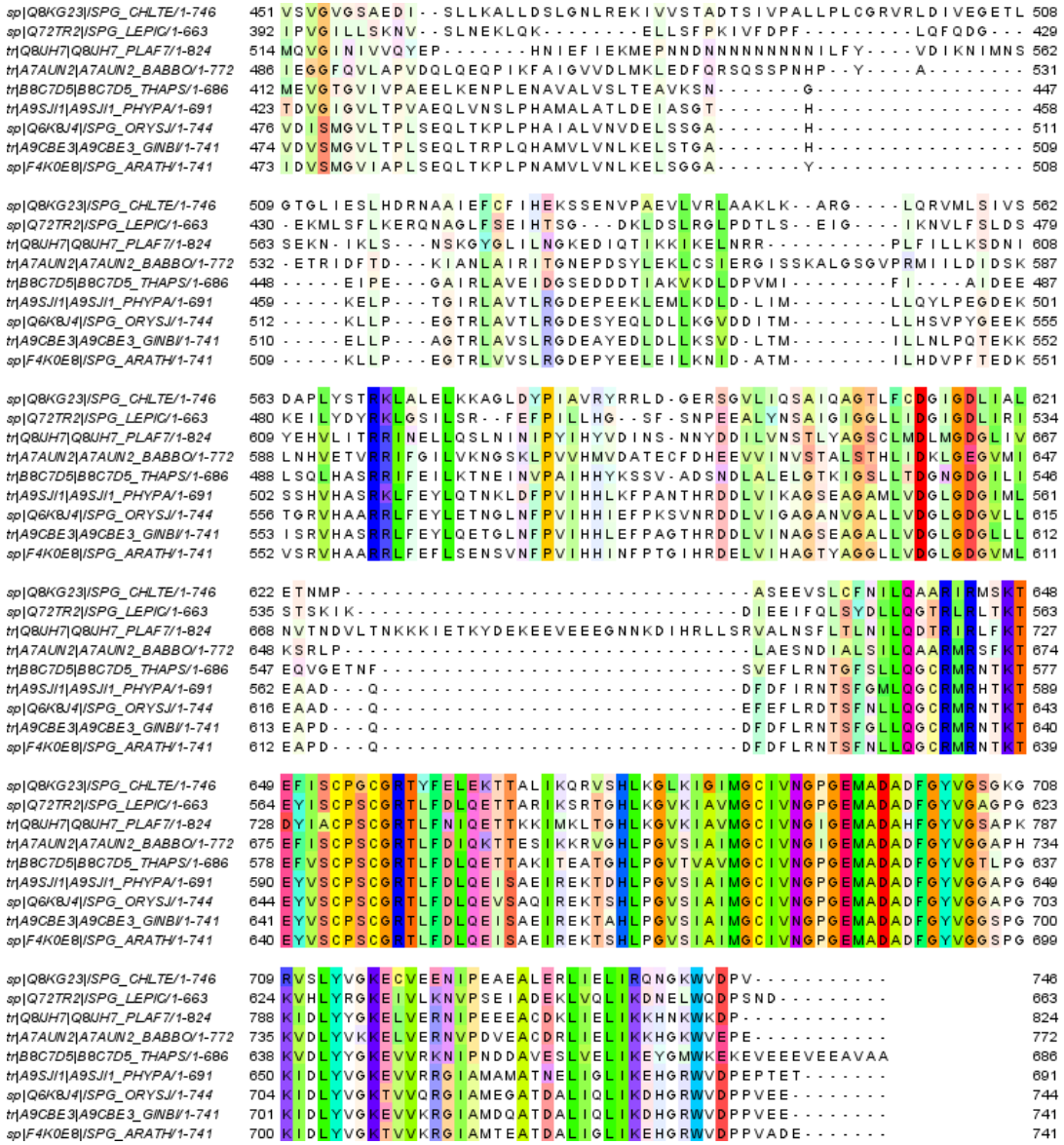


Figure A3.2 Protein sequence comparison of parasites and plant IspG. From top to bottom, Chlorobium tepidum (CHLTE), Osornophryne guacamayo (LEPIG), Plasmodium falciparum (PLAF7), Babesia bovis (BABBO), Thalassiosira pseudonana (THAPS), Physcomitrella patens subspecies patens (PHYPA), Oryza sativa subspecies japonica (ORYSJ), Ginkgo biloba (GINBI), and Arabidopsis thaliana (ARATH). Sequence alignment was performed on Clustal Omega web server and coloured by Taylor conservation pattern. Conservation pattern shows that domain A and B are highly conserved among all species while domain A is less conserved among all species.*

References|

ADIBEKIAN, A., MARTIN, B. R., CHANG, J. W., HSU, K. L., TSUBOI, K., BACHOVCHIN, D. A., SPEERS, A. E., BROWN, S. J., SPICER, T., FERNANDEZ-VEGA, V., FERGUSON, J., HODDER, P. S., ROSEN, H. & CRAVATT, B. F. 2012. Confirming Target Engagement for Reversible Inhibitors in Vivo by Kinetically Tuned Activity-Based Probes. *Journal of the American Chemical Society,* 134**,** 10345-10348.

AICHLER, M. & WALCH, A. 2015. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Lab Invest,* 95**,** 422-31.

ALLEVA, L. M. & KIRK, K. 2001. Calcium regulation in the intraerythrocytic malaria parasite Plasmodium falciparum. *Mol Biochem Parasitol,* 117**,** 121-8.

ANTOINE, T., FISHER, N., AMEWU, R., O'NEILL, P. M., WARD, S. A. & BIAGINI, G. A. 2014. Rapid kill of malaria parasites by artemisinin and semi-synthetic endoperoxides involves ROS-dependent depolarization of the membrane potential. *J Antimicrob Chemother,* 69**,** 1005-16.

APARICIO, I. M., MARIN-MENENDEZ, A., BELL, A. & ENGEL, P. C. 2010. Susceptibility of Plasmodium falciparum to glutamate dehydrogenase inhibitors--a possible new antimalarial target. *Mol Biochem Parasitol,* 172**,** 152-5.

ARAMA, C. & TROYE-BLOMBERG, M. 2014. The path of malaria vaccine development: challenges and perspectives. *J Intern Med,* 275**,** 456-66.

ARAVIND, L., IYER, L. M., WELLEMS, T. E. & MILLER, L. H. 2003. Plasmodium Biology: Genomic Gleanings. *Cell,* 115**,** 771-785.

ARIEY, F., WITKOWSKI, B., AMARATUNGA, C., BEGHAIN, J., LANGLOIS, A. C., KHIM, N., KIM, S., DURU, V., BOUCHIER, C., MA, L., LIM, P., LEANG, R., DUONG, S., SRENG, S., SUON, S., CHUOR, C. M., BOUT, D. M., MENARD, S., ROGERS, W. O., GENTON, B., FANDEUR, T., MIOTTO, O., RINGWALD, P., LE BRAS, J., BERRY, A., BARALE, J. C., FAIRHURST, R. M., BENOIT-VICAL, F., MERCEREAU-PUIJALON, O. & MENARD, D. 2014. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature,* 505**,** 50-5.

ARNOU, B., MONTIGNY, C., MORTH, J. P., NISSEN, P., JAXEL, C., MOLLER, J. V. & MAIRE, M. 2011. The Plasmodium falciparum Ca(2+)-ATPase PfATP6: insensitive to artemisinin, but a potential drug target. *Biochem Soc Trans,* 39**,** 823-31.

ASAHI, H., KANAZAWA, T., KAJIHARA, Y., TAKAHASHI, K. & TAKAHASHI, T. 1996. Hypoxanthine: a low molecular weight factor essential for growth of erythrocytic Plasmodium falciparum in a serum-free medium. *Parasitology,* 113 (Pt 1)**,** 19-23.

ASAWAMAHASAKDA, W., ITTARAT, I., CHANG, C. C., MCELROY, P. & MESHNICK, S. R. 1994. Effects of antimalarials and protease inhibitors on plasmodial hemozoin production. *Mol Biochem Parasitol,* 67**,** 183-91.

ASHLEY, E. A., DHORDA, M., FAIRHURST, R. M., AMARATUNGA, C., LIM, P., SUON, S., SRENG, S., ANDERSON, J. M., MAO, S., SAM, B., SOPHA, C., CHUOR, C. M., NGUON, C., SOVANNAROTH, S., PUKRITTAYAKAMEE, S., JITTAMALA, P., CHOTIVANICH, K., CHUTASMIT, K., SUCHATSOONTHORN, C., RUNCHAROEN, R., HIEN, T. T., THUY-NHIEN, N. T., THANH, N. V., PHU, N. H., HTUT, Y., HAN, K. T., AYE, K. H., MOKUOLU, O. A., OLAOSEBIKAN, R. R., FOLARANMI, O. O., MAYXAY, M., KHANTHAVONG, M., HONGVANTHONG, B., NEWTON, P. N., ONYAMBOKO, M. A., FANELLO, C. I., TSHEFU, A. K., MISHRA, N., VALECHA, N., PHYO, A. P., NOSTEN, F., YI, P., TRIPURA, R., BORRMANN, S., BASHRAHEIL, M., PESHU, J., FAIZ, M. A., GHOSE, A., HOSSAIN, M. A., SAMAD, R., RAHMAN, M. R., HASAN, M. M., ISLAM, A., MIOTTO, O., AMATO, R., MACINNIS, B., STALKER, J., KWIATKOWSKI, D. P., BOZDECH, Z., JEEYAPANT, A., CHEAH, P. Y., SAKULTHAEW, T., CHALK, J., INTHARABUT, B., SILAMUT, K., LEE, S. J., VIHOKHERN, B., KUNASOL, C., IMWONG, M., TARNING, J., TAYLOR, W. J., YEUNG, S., WOODROW, C. J., FLEGG, J. A., DAS, D., SMITH, J., VENKATESAN, M., PLOWE, C. V., STEPNIEWSKA, K., GUERIN, P. J., DONDORP, A. M., DAY, N. P., WHITE, N. J. & TRACKING RESISTANCE TO ARTEMISININ, C. 2014. Spread of artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med,* 371**,** 411-23.

BABBITT, S. E., ALTENHOFEN, L., COBBOLD, S. A., ISTVAN, E. S., FENNELL, C., DOERIG, C., LLINAS, M. & GOLDBERG, D. E. 2012. Plasmodium falciparum responds to amino acid starvation by entering into a hibernatory state. *Proc Natl Acad Sci U S A,* 109**,** E3278-87.

BAER, K., KLOTZ, C., KAPPE, S. H., SCHNIEDER, T. & FREVERT, U. 2007. Release of hepatic Plasmodium yoelii merozoites into the pulmonary microvasculature. *PLoS Pathog,* 3**,** e171.

BAILEY, L. M., IVANOV, R. A., WALLACE, J. C. & POLYAK, S. W. 2008. Artifactual detection of biotin on histones by streptavidin. *Anal Biochem,* 373**,** 71-7.

BANERJEE, S., BARTESAGHI, A., MERK, A., RAO, P., BULFER, S. L., YAN, Y., GREEN, N., MROCZKOWSKI, B., NEITZ, R. J., WIPF, P., FALCONIERI, V., DESHAIES, R. J., MILNE, J. L., HURYN, D., ARKIN, M. & SUBRAMANIAM, S. 2016. 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science,* 351**,** 871-5.

BARAGANA, B., HALLYBURTON, I., LEE, M. C., NORCROSS, N. R., GRIMALDI, R., OTTO, T. D., PROTO, W. R., BLAGBOROUGH, A. M., MEISTER, S., WIRJANATA, G., RUECKER, A., UPTON, L. M., ABRAHAM, T. S., ALMEIDA, M. J., PRADHAN, A., PORZELLE, A., MARTINEZ, M. S., BOLSCHER, J. M., WOODLAND, A., NORVAL, S., ZUCCOTTO, F., THOMAS, J., SIMEONS, F., STOJANOVSKI, L., OSUNA-CABELLO, M., BROCK, P. M., CHURCHER, T. S., SALA, K. A., ZAKUTANSKY, S. E., JIMENEZ-DIAZ, M. B., SANZ, L. M., RILEY, J., BASAK, R., CAMPBELL, M., AVERY, V. M., SAUERWEIN, R. W., DECHERING, K. J., NOVIYANTI, R., CAMPO, B., FREARSON, J. A., ANGULO-BARTUREN, I., FERRER-BAZAGA, S., GAMO, F. J., WYATT, P. G., LEROY, D., SIEGL, P., DELVES, M. J., KYLE, D. E., WITTLIN, S., MARFURT, J., PRICE, R. N., SINDEN, R. E., WINZELER, E. A., CHARMAN, S. A., BEBREVSKA, L., GRAY, D. W., CAMPBELL, S., FAIRLAMB, A. H., WILLIS, P. A., RAYNER, J. C., FIDOCK, D. A., READ, K. D. & GILBERT, I. H. 2015. A novel multiple-stage antimalarial agent that inhibits protein synthesis. *Nature,* 522**,** 315-20.

BASKIN, J. M., PRESCHER, J. A., LAUGHLIN, S. T., AGARD, N. J., CHANG, P. V., MILLER, I. A., LO, A., CODELLI, J. A. & BERTOZZI, C. R. 2007. Copper-free click chemistry for dynamic in vivo imaging. *Proc Natl Acad Sci U S A,* 104**,** 16793-7.

BECKER, J. V., MTWISHA, L., CRAMPTON, B. G., STOYCHEV, S., VAN BRUMMELEN, A. C., REEKSTING, S., LOUW, A. I., BIRKHOLTZ, L. M. & MANCAMA, D. T. 2010. Plasmodium falciparum spermidine synthase inhibition results in unique perturbation-specific effects observed on transcript, protein and metabolite levels. *BMC Genomics,* 11**,** 235.

BERGER, B. J. 2000. Antimalarial Activities of Aminooxy Compounds. *Antimicrobial Agents and Chemotherapy,* 44**,** 2540-2542.

BHISUTTHIBHAN, J., PAN, X. Q., HOSSLER, P. A., WALKER, D. J., YOWELL, C. A., CARLTON, J., DAME, J. B. & MESHNICK, S. R. 1998. The Plasmodium falciparum translationally controlled tumor protein homolog and its reaction with the antimalarial drug artemisinin. *J Biol Chem,* 273**,** 16192-8.

BIAGINI, G. A., FISHER, N., SHONE, A. E., MUBARAKI, M. A., SRIVASTAVA, A., HILL, A., ANTOINE, T., WARMAN, A. J., DAVIES, J., PIDATHALA, C., AMEWU, R. K., LEUNG, S. C., SHARMA, R., GIBBONS, P., HONG, D. W., PACOREL, B., LAWRENSON, A. S., CHAROENSUTTHIVARAKUL, S., TAYLOR, L., BERGER, O., MBEKEANI, A., STOCKS, P. A., NIXON, G. L., CHADWICK, J., HEMINGWAY, J., DELVES, M. J., SINDEN, R. E., ZEEMAN, A. M., KOCKEN, C. H., BERRY, N. G., O'NEILL, P. M. & WARD, S. A. 2012. Generation of quinolone antimalarials targeting the Plasmodium falciparum mitochondrial respiratory chain for the treatment and prophylaxis of malaria. *Proc Natl Acad Sci U S A,* 109**,** 8298-303.

BINDEA, G., MLECNIK, B., HACKL, H., CHAROENTONG, P., TOSOLINI, M., KIRILOVSKY, A., FRIDMAN, W. H., PAGES, F., TRAJANOSKI, Z. & GALON, J. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics,* 25**,** 1091-3.

BINDSCHEDLER, M., LEFEVRE, G., DEGEN, P. & SIOUFI, A. 2002. Comparison of the cardiac effects of the antimalarials co-artemether and halofantrine in healthy participants. *Am J Trop Med Hyg,* 66**,** 293-8.

BIRTH, D., KAO, W. C. & HUNTE, C. 2014. Structural analysis of atovaquone-inhibited cytochrome bc1 complex reveals the molecular basis of antimalarial drug action. *Nat Commun,* 5**,** 4029.

BOZDECH, Z. & GINSBURG, H. 2004. Antioxidant defense in Plasmodium falciparum--data mining of the transcriptome. *Malar J,* 3**,** 23.

BRASIL, L. W., AREAS, A. L., MELO, G. C., OLIVEIRA, C. M., ALECRIM, M. G., LACERDA, M. V., O'BRIEN, C., OELEMANN, W. M. & ZALIS, M. G. 2012. Pfatp6 molecular profile of Plasmodium falciparum isolates in the western Brazilian Amazon. *Malar J,* 11**,** 111.

BRAY, P. G., JANNEH, O., RAYNES, K. J., MUNGTHIN, M., GINSBURG, H. & WARD, S. A. 1999. Cellular uptake of chloroquine is dependent on binding to ferriprotoporphyrin IX and is independent of NHE activity in Plasmodium falciparum. *J Cell Biol,* 145**,** 363-76.

BRAY, P. G., MUNGTHIN, M., RIDLEY, R. G. & WARD, S. A. 1998. Access to hematin: the basis of chloroquine resistance. *Mol Pharmacol,* 54**,** 170-9.

BRUECKNER, R. P., LASSETER, K. C., LIN, E. T. & SCHUSTER, B. G. 1998. First-time-in-humans safety and pharmacokinetics of WR 238605, a new antimalarial. *Am J Trop Med Hyg,* 58**,** 645-9.

BURROWS, J. N., VAN HUIJSDUIJNEN, R. H., MOHRLE, J. J., OEUVRAY, C. & WELLS, T. N. 2013. Designing the next generation of medicines for malaria control and eradication. *Malar J,* 12**,** 187.

BUTLER, A. R., GILBERT, B. C., HULME, P., IRVINE, L. R., RENTON, L. & WHITWOOD, A. C. 1998. EPR evidence for the involvement of free radicals in the iron-catalysed decomposition of qinghaosu (artemisinin) and some derivatives; antimalarial action of some polycyclic endoperoxides. *Free Radic Res,* 28**,** 471-6.

CAPPER, M. J., O'NEILL, P. M., FISHER, N., STRANGE, R. W., MOSS, D., WARD, S. A., BERRY, N. G., LAWRENSON, A. S., HASNAIN, S. S., BIAGINI, G. A. & ANTONYUK, S. V. 2015. Antimalarial 4(1H)-pyridones bind to the Qi site of cytochrome bc1. *Proc Natl Acad Sci U S A,* 112**,** 755-60.

CHAPIN, G. & WASSERSTROM, R. 1981. Agricultural production and malaria resurgence in Central America and India. *Nature,* 293**,** 181-185.

CHARMAN, S. A., ARBE-BARNES, S., BATHURST, I. C., BRUN, R., CAMPBELL, M., CHARMAN, W. N., CHIU, F. C., CHOLLET, J., CRAFT, J. C., CREEK, D. J., DONG, Y., MATILE, H., MAURER, M., MORIZZI, J., NGUYEN, T., PAPASTOGIANNIDIS, P., SCHEURER, C., SHACKLEFORD, D. M., SRIRAGHAVAN, K., STINGELIN, L., TANG, Y., URWYLER, H., WANG, X., WHITE, K. L., WITTLIN, S., ZHOU, L. & VENNERSTROM, J. L. 2011. Synthetic ozonide drug candidate OZ439 offers new hope for a single-dose cure of uncomplicated malaria. *Proc Natl Acad Sci U S A,* 108**,** 4400-5.

CHEN, X. & WU, Y. W. 2016. Selective chemical labeling of proteins. *Org Biomol Chem,* 14**,** 5417-39.

CHENG, Q., KYLE, D. E. & GATTON, M. L. 2012. Artemisinin resistance in Plasmodium falciparum: A process linked to dormancy? *Int J Parasitol Drugs Drug Resist,* 2**,** 249-255.

CHIODINI, P. L., CONLON, C. P., HUTCHINSON, D. B. A., FARQUHAR, J. A., HALL, A. P., PETO, T. E. A., BIRLEY, H. & WARRELL, D. A. 1995. Evaluation of atovaquone in the treatment of patients with uncomplicated Plasmodium falciparum malaria. *Journal of Antimicrobial Chemotherapy,* 36**,** 1073-1078.

CHOU, T. F., BROWN, S. J., MINOND, D., NORDIN, B. E., LI, K., JONES, A. C., CHASE, P., PORUBSKY, P. R., STOLTZ, B. M., SCHOENEN, F. J., PATRICELLI, M. P., HODDER, P., ROSEN, H. & DESHAIES, R. J. 2011. Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proc Natl Acad Sci U S A,* 108**,** 4834-9.

CINGOLANI, P., PLATTS, A., WANG LE, L., COON, M., NGUYEN, T., WANG, L., LAND, S. J., LU, X. & RUDEN, D. M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin),* 6**,** 80-92.

COBBOLD, S. A., CHUA, H. H., NIJAGAL, B., CREEK, D. J., RALPH, S. A. & MCCONVILLE, M. J. 2016. Metabolic Dysregulation Induced in Plasmodium falciparum by Dihydroartemisinin and Other Front-Line Antimalarial Drugs. *J Infect Dis,* 213**,** 276-86.

CODD, A., TEUSCHER, F., KYLE, D. E., CHENG, Q. & GATTON, M. L. 2011. Artemisinin-induced parasite dormancy: a plausible mechanism for treatment failure. *Malar J,* 10**,** 56.

COMMITTEE ON THE ECONOMICS OF ANTIMALARIAL DRUGS, B. O. G. H. 2004. Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance. *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance* [Online]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25009879

https://www.nap.edu/catalog/11017/saving-lives-buying-time-economics-of-malaria-drugs-in-an 2016].

CONRAD, M. D., BIGIRA, V., KAPISI, J., MUHINDO, M., KAMYA, M. R., HAVLIR, D. V., DORSEY, G. & ROSENTHAL, P. J. 2014. Polymorphisms in K13 and falcipain-2 associated with artemisinin resistance are not prevalent in Plasmodium falciparum isolated from Ugandan children. *PLoS One,* 9**,** e105690.

CONSORTIUM, C. E. S. 1998. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science,* 282**,** 2012-8.

COOMBS, G. H., GOLDBERG, D. E., KLEMBA, M., BERRY, C., KAY, J. & MOTTRAM, J. C. 2001. Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. *Trends in Parasitology,* 17**,** 532-537.

COX, F. E. 2010. History of the discovery of the malaria parasites and their vectors. *Parasit Vectors,* 3**,** 5.

CREEK, D. J., CHUA, H. H., COBBOLD, S. A., NIJAGAL, B., MACRAE, J. I., DICKERMAN, B. K., GILSON, P. R., RALPH, S. A. & MCCONVILLE, M. J. 2016. Metabolomics-based screening of the Malaria Box reveals both novel and established mechanisms of action. *Antimicrob Agents Chemother*.

CROMPTON, P. D., PIERCE, S. K. & MILLER, L. H. 2010. Advances and challenges in malaria vaccine development. *J Clin Invest,* 120**,** 4168-78.

CROWTHER, G. J., SHANMUGAM, D., CARMONA, S. J., DOYLE, M. A., HERTZ-FOWLER, C., BERRIMAN, M., NWAKA, S., RALPH, S. A., ROOS, D. S., VAN VOORHIS, W. C. & AGUERO, F. 2010. Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. *PLoS Negl Trop Dis,* 4**,** e804.

CUI, L., WANG, Z., JIANG, H., PARKER, D., WANG, H., SU, X. Z. & CUI, L. 2012. Lack of association of the S769N mutation in Plasmodium falciparum SERCA (PfATP6) with resistance to artemisinins. *Antimicrob Agents Chemother,* 56**,** 2546-52.

DALAL, S. & KLEMBA, M. 2007. Roles for two aminopeptidases in vacuolar hemoglobin catabolism in Plasmodium falciparum. *J Biol Chem,* 282**,** 35978-87.

DAVID-BOSNE, S., CLAUSEN, M. V., POULSEN, H., MOLLER, J. V., NISSEN, P. & LE MAIRE, M. 2016. Reappraising the effects of artemisinin on the ATPase activity of PfATP6 and SERCA1a E255L expressed in Xenopus laevis oocytes. *Nat Struct Mol Biol,* 23**,** 1-2.

DE VRIES, P. J. & DIEN, T. K. 1996. Clinical pharmacology and therapeutic potential of artemisinin and its derivatives in the treatment of malaria. *Drugs,* 52**,** 818-36.

DENIS, M. B., DAVIS, T. M., HEWITT, S., INCARDONA, S., NIMOL, K., FANDEUR, T., PORAVUTH, Y., LIM, C. & SOCHEAT, D. 2002. Efficacy and safety of dihydroartemisinin-piperaquine (Artekin) in Cambodian children and adults with uncomplicated falciparum malaria. *Clin Infect Dis,* 35**,** 1469-76.

DING, X. C., BECK, H. P. & RASO, G. 2011. Plasmodium sensitivity to artemisinins: magic bullets hit elusive targets. *Trends Parasitol,* 27**,** 73-81.

DJIMDE, A. & LEFEVRE, G. 2009. Understanding the pharmacokinetics of Coartem. *Malar J,* 8 Suppl 1**,** S4.

DODD, E. L. & BOHLE, D. S. 2014. Orienting the heterocyclic periphery: a structural model for chloroquine's antimalarial activity. *Chem Commun (Camb),* 50**,** 13765-8.

DOGOVSKI, C., XIE, S. C., BURGIO, G., BRIDGFORD, J., MOK, S., MCCAW, J. M., CHOTIVANICH, K., KENNY, S., GNADIG, N., STRAIMER, J., BOZDECH, Z., FIDOCK, D. A., SIMPSON, J. A., DONDORP, A. M., FOOTE, S., KLONIS, N. & TILLEY, L. 2015. Targeting the cell stress response of Plasmodium falciparum to overcome artemisinin resistance. *PLoS Biol,* 13**,** e1002132.

DONDORP, A. M., NOSTEN, F., YI, P., DAS, D., PHYO, A. P., TARNING, J., LWIN, K. M., ARIEY, F., HANPITHAKPONG, W., LEE, S. J., RINGWALD, P., SILAMUT, K., IMWONG, M., CHOTIVANICH, K., LIM, P., HERDMAN, T., AN, S. S., YEUNG, S., SINGHASIVANON, P., DAY, N. P., LINDEGARDH, N., SOCHEAT, D. & WHITE, N. J. 2009. Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med,* 361**,** 455-67.

DOWNIE, M. J., KIRK, K. & MAMOUN, C. B. 2008. Purine salvage pathways in the intraerythrocytic malaria parasite Plasmodium falciparum. *Eukaryot Cell,* 7**,** 1231-7.

DUNBRACK, R. L., JR. 2002. Rotamer libraries in the 21st century. *Curr Opin Struct Biol,* 12**,** 431-40.

ECKSTEIN-LUDWIG, U., WEBB, R. J., VAN GOETHEM, I. D., EAST, J. M., LEE, A. G., KIMURA, M., O'NEILL, P. M., BRAY, P. G., WARD, S. A. & KRISHNA, S. 2003. Artemisinins target the SERCA of Plasmodium falciparum. *Nature,* 424**,** 957-61.

EICHHORN, T., WINTER, D., BUCHELE, B., DIRDJAJA, N., FRANK, M., LEHMANN, W. D., MERTENS, R., KRAUTH-SIEGEL, R. L., SIMMET, T., GRANZIN, J. & EFFERTH, T. 2013. Molecular interaction of artemisinin with translationally controlled tumor protein (TCTP) of Plasmodium falciparum. *Biochem Pharmacol,* 85**,** 38-45.

ERMAK, G. & DAVIES, K. J. 2002. Calcium and oxidative stress: from cell signaling to cell death. *Mol Immunol,* 38**,** 713-21.

FERRERAS, A., TRIANA, L., SANCHEZ, E. & HERRERA, F. 2002. Effect of antimalarial drugs on plasmodia cell-free protein synthesis. *Mem Inst Oswaldo Cruz,* 97**,** 377-80.

FIDOCK, D. A., NOMURA, T., TALLEY, A. K., COOPER, R. A., DZEKUNOV, S. M., FERDIG, M. T., URSOS, L. M., SIDHU, A. B., NAUDE, B., DEITSCH, K. W., SU, X. Z., WOOTTON, J. C., ROEPE, P. D. & WELLEMS, T. E. 2000. Mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell,* 6**,** 861-71.

FINEHOUT, E. J. & LEE, K. H. 2004. An introduction to mass spectrometry applications in biological research. *Biochem Mol Biol Educ,* 32**,** 93-100.

FISHER, N., ABD MAJID, R., ANTOINE, T., AL-HELAL, M., WARMAN, A. J., JOHNSON, D. J., LAWRENSON, A. S., RANSON, H., O'NEILL, P. M., WARD, S. A. & BIAGINI, G. A. 2012. Cytochrome b mutation Y268S conferring atovaquone resistance phenotype in malaria parasite results in reduced parasite bc1 catalytic turnover and protein expression. *J Biol Chem,* 287**,** 9731-41.

FIVELMAN, Q. L., BUTCHER, G. A., ADAGU, I. S., WARHURST, D. C. & PASVOL, G. 2002. Malarone treatment failure and in vitro confirmation of resistance of Plasmodium falciparum isolate from Lagos, Nigeria. *Malar J,* 1**,** 1.

FLEISCHMANN, R. D., ADAMS, M. D., WHITE, O., CLAYTON, R. A., KIRKNESS, E. F., KERLAVAGE, A. R., BULT, C. J., TOMB, J. F., DOUGHERTY, B. A., MERRICK, J. M. & ET AL. 1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science,* 269**,** 496-512.

FLORENS, L., WASHBURN, M. P., RAINE, J. D., ANTHONY, R. M., GRAINGER, M., HAYNES, J. D., MOCH, J. K., MUSTER, N., SACCI, J. B., TABB, D. L., WITNEY, A. A., WOLTERS, D., WU, Y., GARDNER, M. J., HOLDER, A. A., SINDEN, R. E., YATES, J. R. & CARUCCI, D. J. 2002. A proteomic view of the Plasmodium falciparum life cycle. *Nature,* 419**,** 520-6.

FUGI, M. A., WITTLIN, S., DONG, Y. & VENNERSTROM, J. L. 2010. Probing the antimalarial mechanism of artemisinin and OZ277 (arterolane) with nonperoxidic isosteres and nitroxyl radicals. *Antimicrob Agents Chemother,* 54**,** 1042-6.

GALLUP, J. L. & SACHS, J. D. 2001. The economic burden of malaria. *Am J Trop Med Hyg,* 64**,** 85-96.

GARDNER, M. J., HALL, N., FUNG, E., WHITE, O., BERRIMAN, M., HYMAN, R. W., CARLTON, J. M., PAIN, A., NELSON, K. E., BOWMAN, S., PAULSEN, I. T., JAMES, K., EISEN, J. A., RUTHERFORD, K., SALZBERG, S. L., CRAIG, A., KYES, S., CHAN, M. S., NENE, V., SHALLOM, S. J., SUH, B., PETERSON, J., ANGIUOLI, S., PERTEA, M., ALLEN, J., SELENGUT, J., HAFT, D., MATHER, M. W., VAIDYA, A. B., MARTIN, D. M., FAIRLAMB, A. H., FRAUNHOLZ, M. J., ROOS, D. S., RALPH, S. A., MCFADDEN, G. I., CUMMINGS, L. M., SUBRAMANIAN, G. M., MUNGALL, C., VENTER, J. C., CARUCCI, D. J., HOFFMAN, S. L., NEWBOLD, C., DAVIS, R. W., FRASER, C. M. & BARRELL, B. 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature,* 419**,** 498-511.

GASSIS, S. & RATHOD, P. K. 1996. Frequency of drug resistance in Plasmodium falciparum: a nonsynergistic combination of 5-fluoroorotate and atovaquone suppresses in vitro resistance. *Antimicrob Agents Chemother,* 40**,** 914-9.

GAVIGAN, C. S., DALTON, J. P. & BELL, A. 2001. The role of aminopeptidases in haemoglobin degradation in Plasmodium falciparum-infected erythrocytes. *Molecular and Biochemical Parasitology,* 117**,** 37-48.

GLUSHAKOVA, S., LIZUNOV, V., BLANK, P. S., MELIKOV, K., HUMPHREY, G. & ZIMMERBERG, J. 2013. Cytoplasmic free Ca2+ is essential for multiple steps in malaria parasite egress from infected erythrocytes. *Malar J,* 12**,** 41.

GOFFEAU, A., BARRELL, B. G., BUSSEY, H., DAVIS, R. W., DUJON, B., FELDMANN, H., GALIBERT, F., HOHEISEL, J. D., JACQ, C., JOHNSTON, M., LOUIS, E. J., MEWES, H. W., MURAKAMI, Y., PHILIPPSEN, P., TETTELIN, H. & OLIVER, S. G. 1996. Life with 6000 genes. *Science,* 274**,** 546, 563-7.

GOLDGOF, G. M., DURRANT, J. D., OTTILIE, S., VIGIL, E., ALLEN, K. E., GUNAWAN, F., KOSTYLEV, M., HENDERSON, K. A., YANG, J., SCHENKEN, J., LAMONTE, G. M., MANARY, M. J., MURAO, A., NACHON, M., STANHOPE, R., PRESCOTT, M., MCNAMARA, C. W., SLAYMAN, C. W., AMARO, R. E., SUZUKI, Y. & WINZELER, E. A. 2016. Comparative chemical genomics reveal that the spiroindolone antimalarial KAE609 (Cipargamin) is a P-type ATPase inhibitor. *Sci Rep,* 6**,** 27806.

GOMES, A. R., BUSHELL, E., SCHWACH, F., GIRLING, G., ANAR, B., QUAIL, M. A., HERD, C., PFANDER, C., MODRZYNSKA, K., RAYNER, J. C. & BILLKER, O. 2015. A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. *Cell Host Microbe,* 17**,** 404-13.

GOODWIN, S., MCPHERSON, J. D. & MCCOMBIE, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet,* 17**,** 333-51.

GORDON, D. M., MCGOVERN, T. W., KRZYCH, U., COHEN, J. C., SCHNEIDER, I., LACHANCE, R., HEPPNER, D. G., YUAN, G., HOLLINGDALE, M., SLAOUI, M. & ET AL. 1995. Safety, immunogenicity, and efficacy of a recombinantly produced Plasmodium falciparum circumsporozoite protein-hepatitis B surface antigen subunit vaccine. *J Infect Dis,* 171**,** 1576-85.

GREGSON, A. & PLOWE, C. V. 2005. Mechanisms of resistance of malaria parasites to antifolates. *Pharmacol Rev,* 57**,** 117-45.

GU, H. M., WARHURST, D. C. & PETERS, W. 1983. Rapid action of Qinghaosu and related drugs on incorporation of [3H]isoleucine by Plasmodium falciparum in vitro. *Biochemical Pharmacology,* 32**,** 2463-2466.

GUERRA, F., WANG, K., LI, J., WANG, W., LIU, Y. L., AMIN, S. & OLDFIELD, E. 2014. Inhibition of the 4Fe-4S Proteins IspG and IspH: an EPR, ENDOR and HYSCORE Investigation. *Chem Sci,* 5**,** 1642-1649.

HARBUT, M. B., PATEL, B. A., YEUNG, B. K., MCNAMARA, C. W., BRIGHT, A. T., BALLARD, J., SUPEK, F., GOLDE, T. E., WINZELER, E. A., DIAGANA, T. T. & GREENBAUM, D. C. 2012. Targeting the ERAD pathway via inhibition of signal peptide peptidase for antiparasitic therapeutic design. *Proc Natl Acad Sci U S A,* 109**,** 21486-91.

HAY, S. I., GUERRA, C. A., TATEM, A. J., NOOR, A. M. & SNOW, R. W. 2004. The global distribution and population at risk of malaria: past, present, and future. *The Lancet Infectious Diseases,* 4**,** 327-336.

HAYNES, R. K., CHEU, K. W., CHAN, H. W., WONG, H. N., LI, K. Y., TANG, M. M., CHEN, M. J., GUO, Z. F., GUO, Z. H., SINNIAH, K., WITTE, A. B., COGHI, P. & MONTI, D. 2012. Interactions between artemisinins and other antimalarial drugs in relation to the cofactor model--a unifying proposal for drug action. *ChemMedChem,* 7**,** 2204-26.

HAZLETON, K. Z., HO, M. C., CASSERA, M. B., CLINCH, K., CRUMP, D. R., ROSARIO, I., JR., MERINO, E. F., ALMO, S. C., TYLER, P. C. & SCHRAMM, V. L. 2012. Acyclic immucillin phosphonates: second-generation inhibitors of Plasmodium falciparum hypoxanthine-guanine-xanthine phosphoribosyltransferase. *Chem Biol,* 19**,** 721-30.

HECHT, S., EISENREICH, W., ADAM, P., AMSLINGER, S., KIS, K., BACHER, A., ARIGONI, D. & ROHDICH, F. 2001. Studies on the nonmevalonate pathway to terpenes: the role of the GcpE (IspG) protein. *Proc Natl Acad Sci U S A,* 98**,** 14837-42.

HIMO, F., LOVELL, T., HILGRAF, R., ROSTOVTSEV, V. V., NOODLEMAN, L., SHARPLESS, K. B. & FOKIN, V. V. 2004. Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *Journal of the American Chemical Society,* 127**,** 210-216.

HOTT, A., CASANDRA, D., SPARKS, K. N., MORTON, L. C., CASTANARES, G. G., RUTTER, A. & KYLE, D. E. 2015. Artemisinin-resistant Plasmodium falciparum Exhibit Altered Patterns of Development in Infected Erythrocytes. *Antimicrob Agents Chemother*.

HUISGEN, R. 1963. 1,3-Dipolar Cycloadditions. Past and Future. *Angewandte Chemie International Edition in English,* 2**,** 565-598.

HUTHMACHER, C., HOPPE, A., BULIK, S. & HOLZHUTTER, H. G. 2010. Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis. *BMC Syst Biol,* 4**,** 120.

HVIID, L. & JENSEN, A. T. 2015. PfEMP1 - A Parasite Protein Family of Key Importance in Plasmodium falciparum Malaria Immunity and Pathogenesis. *Adv Parasitol,* 88**,** 51-84.

IMAMURA, S., OJIMA, N. & YAMASHITA, M. 2003. Cold-inducible expression of the cell division cycle gene CDC48 and its promotion of cell proliferation during cold acclimation in zebrafish cells. *FEBS Lett,* 549**,** 14-20.

IMLAY, L. S., ARMSTRONG, C. M., MASTERS, M. C., LI, T., PRICE, K. E., EDWARDS, R. L., MANN, K. M., LI, L. X., STALLINGS, C. L., BERRY, N. G., O'NEILL, P. M. & ODOM, A. R. 2015. Plasmodium IspD (2-C-Methyl-D-erythritol 4-Phosphate Cytidyltransferase), an Essential and Druggable Antimalarial Target. *ACS Infect Dis,* 1**,** 157-167.

ISLAHUDIN, F., TINDALL, S. M., MELLOR, I. R., SWIFT, K., CHRISTENSEN, H. E., FONE, K. C., PLEASS, R. J., TING, K. N. & AVERY, S. V. 2014. The antimalarial drug quinine interferes with serotonin biosynthesis and action. *Sci Rep,* 4**,** 3618.

ISMAIL, H. M., BARTON, V., PHANCHANA, M., CHAROENSUTTHIVARAKUL, S., WONG, M. H., HEMINGWAY, J., BIAGINI, G. A., O'NEILL, P. M. & WARD, S. A. 2016a. Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites Plasmodium falciparum 3D7. *Proc Natl Acad Sci U S A,* 113**,** 2080-5.

ISMAIL, H. M., BARTON, V. E., PANCHANA, M., CHAROENSUTTHIVARAKUL, S., BIAGINI, G. A., WARD, S. A. & O'NEILL, P. M. 2016b. A Click Chemistry-Based Proteomic Approach Reveals that 1,2,4-Trioxolane and Artemisinin Antimalarials Share a Common Protein Alkylation Profile. *Angew Chem Int Ed Engl,* 55**,** 6401-5.

ISOZUMI, R., UEMURA, H., KIMATA, I., ICHINOSE, Y., LOGEDI, J., OMAR, A. H. & KANEKO, A. 2015. Novel Mutations in K13 Propeller Gene of Artemisinin-Resistant Plasmodium falciparum. *Emerg Infect Dis,* 21**,** 490-2.

IUPAC 1997. *Compendium of Chemical Terminology (the Gold Book)*, IUPAC.

JANI, D., NAGARKATTI, R., BEATTY, W., ANGEL, R., SLEBODNICK, C., ANDERSEN, J., KUMAR, S. & RATHORE, D. 2008. HDP-a novel heme detoxification protein from the malaria parasite. *PLoS Pathog,* 4**,** e1000053.

JEWETT, J. C. & BERTOZZI, C. R. 2010. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society Reviews,* 39**,** 1272-1279.

JOANNIN, N., ABHIMAN, S., SONNHAMMER, E. L. & WAHLGREN, M. 2008. Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics,* 9**,** 19.

JORTZIK, E., FRITZ-WOLF, K., STURM, N., HIPP, M., RAHLFS, S. & BECKER, K. 2010. Redox regulation of Plasmodium falciparum ornithine delta-aminotransferase. *J Mol Biol,* 402**,** 445-59.

JOSHI, N. A. & FASS, J. N. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.200).

JUGE, N., MORIYAMA, S., MIYAJI, T., KAWAKAMI, M., IWAI, H., FUKUI, T., NELSON, N., OMOTE, H. & MORIYAMA, Y. 2015. Plasmodium falciparum chloroquine resistance transporter is a H+-coupled polyspecific nutrient and drug exporter. *Proc Natl Acad Sci U S A,* 112**,** 3356-61.

KAKURU, A., JAGANNATHAN, P., MUHINDO, M. K., NATUREEBA, P., AWORI, P., NAKALEMBE, M., OPIRA, B., OLWOCH, P., ATEGEKA, J., NAYEBARE, P., CLARK, T. D., FEENEY, M. E., CHARLEBOIS, E. D., RIZZUTO, G., MUEHLENBACHS, A., HAVLIR, D. V., KAMYA, M. R. & DORSEY, G. 2016. Dihydroartemisinin-Piperaquine for the Prevention of Malaria in Pregnancy. *N Engl J Med,* 374**,** 928-39.

KAN, A., TAN, Y. H., ANGRISANO, F., HANSSEN, E., ROGERS, K. L., WHITEHEAD, L., MOLLARD, V. P., COZIJNSEN, A., DELVES, M. J., CRAWFORD, S., SINDEN, R. E., MCFADDEN, G. I., LECKIE, C., BAILEY, J. & BAUM, J. 2014. Quantitative analysis of Plasmodium ookinete motion in three dimensions suggests a critical role for cell shape in the biomechanics of malaria parasite gliding motility. *Cell Microbiol,* 16**,** 734-50.

KEHR, S., JORTZIK, E., DELAHUNTY, C., YATES, J. R., 3RD, RAHLFS, S. & BECKER, K. 2011. Protein S-glutathionylation in malaria parasites. *Antioxid Redox Signal,* 15**,** 2855-65.

KEOUGH, D. T., SKINNER-ADAMS, T., JONES, M. K., NG, A. L., BRERETON, I. M., GUDDAT, L. W. & DE JERSEY, J. 2006. Lead compounds for antimalarial chemotherapy: purine base analogs discriminate between human and P. falciparum 6-oxopurine phosphoribosyltransferases. *J Med Chem,* 49**,** 7479-86.

KESTER, K. E., CUMMINGS, J. F., OFORI-ANYINAM, O., OCKENHOUSE, C. F., KRZYCH, U., MORIS, P., SCHWENK, R., NIELSEN, R. A., DEBEBE, Z., PINELIS, E., JUOMPAN, L., WILLIAMS, J., DOWLER, M., STEWART, V. A., WIRTZ, R. A., DUBOIS, M. C., LIEVENS, M., COHEN, J., BALLOU, W. R., HEPPNER, D. G., JR. & RTS, S. V. E. G. 2009. Randomized, double-blind, phase 2a trial of falciparum malaria vaccines RTS,S/AS01B and RTS,S/AS02A in malaria-naive adults: safety, efficacy, and immunologic associates of protection. *J Infect Dis,* 200**,** 337-46.

KLEMBA, M., GLUZMAN, I. & GOLDBERG, D. E. 2004. A Plasmodium falciparum dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J Biol Chem,* 279**,** 43000-7.

KLONIS, N., CRESPO-ORTIZ, M. P., BOTTOVA, I., ABU-BAKAR, N., KENNY, S., ROSENTHAL, P. J. & TILLEY, L. 2011. Artemisinin activity against Plasmodium falciparum requires hemoglobin uptake and digestion. *Proc Natl Acad Sci U S A,* 108**,** 11405-10.

KLONIS, N., XIE, S. C., MCCAW, J. M., CRESPO-ORTIZ, M. P., ZALOUMIS, S. G., SIMPSON, J. A. & TILLEY, L. 2013. Altered temporal response of malaria parasites determines differential sensitivity to artemisinin. *Proc Natl Acad Sci U S A,* 110**,** 5157-62.

KOLAKOVICH, K. A., GLUZMAN, I. Y., DUFFIN, K. L. & GOLDBERG, D. E. 1997. Generation of hemoglobin peptides in the acidic digestive vacuole of Plasmodium falciparum implicates peptide transport in amino acid production. *Molecular and Biochemical Parasitology,* 87**,** 123-135.

KOLB, H. C., FINN, M. G. & SHARPLESS, K. B. 2001. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie-International Edition,* 40**,** 2004-+.

KREIDENWEISS, A., MORDMULLER, B., KRISHNA, S. & KREMSNER, P. G. 2006. Antimalarial activity of a synthetic endoperoxide (RBx-11160/OZ277) against Plasmodium falciparum isolates from Gabon. *Antimicrob Agents Chemother,* 50**,** 1535-7.

KRUNGKRAI, S. R. & YUTHAVONG, Y. 1987. The antimalarial action on Plasmodium falciparum of qinghaosu and artesunate in combination with agents which modulate oxidant stress. *Transactions of The Royal Society of Tropical Medicine and Hygiene,* 81**,** 710-714.

KUHEN, K. L., CHATTERJEE, A. K., ROTTMANN, M., GAGARING, K., BORBOA, R., BUENVIAJE, J., CHEN, Z., FRANCEK, C., WU, T., NAGLE, A., BARNES, S. W., PLOUFFE, D., LEE, M. C., FIDOCK, D. A., GRAUMANS, W., VAN DE VEGTE-BOLMER, M., VAN GEMERT, G. J., WIRJANATA, G., SEBAYANG, B., MARFURT, J., RUSSELL, B., SUWANARUSK, R., PRICE, R. N., NOSTEN, F., TUNGTAENG, A., GETTAYACAMIN, M., SATTABONGKOT, J., TAYLOR, J., WALKER, J. R., TULLY, D., PATRA, K. P., FLANNERY, E. L., VINETZ, J. M., RENIA, L., SAUERWEIN, R. W., WINZELER, E. A., GLYNNE, R. J. & DIAGANA, T. T. 2014. KAF156 is an antimalarial clinical candidate with potential for use in prophylaxis, treatment, and prevention of disease transmission. *Antimicrob Agents Chemother,* 58**,** 5060-7.

KYES, S. A., KRAEMER, S. M. & SMITH, J. D. 2007. Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family. *Eukaryot Cell,* 6**,** 1511-20.

KYES, S. A., ROWE, J. A., KRIEK, N. & NEWBOLD, C. I. 1999. Rifins: A second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. *Proceedings of the National Academy of Sciences,* 96**,** 9333-9338.

LAM, K. S. & LEBL, M. 1992. Streptavidin and avidin recognize peptide ligands with different motifs. *ImmunoMethods,* 1**,** 11-15.

LAU, C. K., TURNER, L., JESPERSEN, J. S., LOWE, E. D., PETERSEN, B., WANG, C. W., PETERSEN, J. E., LUSINGU, J., THEANDER, T. G., LAVSTSEN, T. & HIGGINS, M. K. 2015. Structural conservation despite huge sequence diversity allows EPCR binding by the PfEMP1 family implicated in severe childhood malaria. *Cell Host Microbe,* 17**,** 118-29.

LE ROCH, K. G., JOHNSON, J. R., FLORENS, L., ZHOU, Y., SANTROSYAN, A., GRAINGER, M., YAN, S. F., WILLIAMSON, K. C., HOLDER, A. A., CARUCCI, D. J., YATES, J. R., 3RD & WINZELER, E. A. 2004. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome Res,* 14**,** 2308-18.

LEE, M., GRAWERT, T., QUITTERER, F., ROHDICH, F., EPPINGER, J., EISENREICH, W., BACHER, A. & GROLL, M. 2010. Biosynthesis of isoprenoids: crystal structure of the [4Fe-4S] cluster protein IspG. *J Mol Biol,* 404**,** 600-10.

LEE, M. C. & FIDOCK, D. A. 2014. CRISPR-mediated genome editing of Plasmodium falciparum malaria parasites. *Genome Med,* 6**,** 63.

LEHANE, A. M. & KIRK, K. 2008. Chloroquine resistance-conferring mutations in pfcrt give rise to a chloroquine-associated H+ leak from the malaria parasite's digestive vacuole. *Antimicrob Agents Chemother,* 52**,** 4374-80.

LI, Q., O'NEIL, M., XIE, L., CARIDHA, D., ZENG, Q., ZHANG, J., PYBUS, B., HICKMAN, M. & MELENDEZ, V. 2014. Assessment of the prophylactic activity and pharmacokinetic profile of oral tafenoquine compared to primaquine for inhibition of liver stage malaria infections. *Malar J,* 13**,** 141.

LIM, M. Y., LAMONTE, G., LEE, M. C., REIMER, C., TAN, B. H., COREY, V., TJAHJADI, B. F., CHUA, A., NACHON, M., WINTJENS, R., GEDECK, P., MALLERET, B., RENIA, L., BONAMY, G. M., HO, P. C., YEUNG, B. K., CHOW, E. D., LIM, L., FIDOCK, D. A., DIAGANA, T. T., WINZELER, E. A. & BIFANI, P. 2016. UDP-galactose and acetyl-CoA transporters as Plasmodium multidrug resistance genes. *Nat Microbiol***,** 16166.

LIU, J., ISTVAN, E. S., GLUZMAN, I. Y., GROSS, J. & GOLDBERG, D. E. 2006. Plasmodium falciparum ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proc Natl Acad Sci U S A,* 103**,** 8840-5.

LIU, Y. L., GUERRA, F., WANG, K., WANG, W., LI, J., HUANG, C., ZHU, W., HOULIHAN, K., LI, Z., ZHANG, Y., NAIR, S. K. & OLDFIELD, E. 2012. Structure, function and inhibition of the two- and three-domain 4Fe-4S IspG proteins. *Proc Natl Acad Sci U S A,* 109**,** 8558-63.

LLANOS-CUENTAS, A., LACERDA, M. V., RUEANGWEERAYUT, R., KRUDSOOD, S., GUPTA, S. K., KOCHAR, S. K., ARTHUR, P., CHUENCHOM, N., MOHRLE, J. J., DUPARC, S., UGWUEGBULAM, C., KLEIM, J. P., CARTER, N., GREEN, J. A. & KELLAM, L. 2014. Tafenoquine plus chloroquine for the treatment and relapse prevention of Plasmodium vivax malaria (DETECTIVE): a multicentre, double-blind, randomised, phase 2b dose-selection study. *Lancet,* 383**,** 1049-58.

LOOAREESUWAN, S., VIRAVAN, C., WEBSTER, H. K., KYLE, D. E., HUTCHINSON, D. B. & CANFIELD, C. J. 1996. Clinical studies of atovaquone, alone or in combination with other antimalarial drugs, for treatment of acute uncomplicated malaria in Thailand. *Am J Trop Med Hyg,* 54**,** 62-6.

LUDIN, P., WOODCROFT, B., RALPH, S. A. & MASER, P. 2012. In silico prediction of antimalarial drug target candidates. *Int J Parasitol Drugs Drug Resist,* 2**,** 191-9.

LY, A., BUCK, A., BALLUFF, B., SUN, N., GORZOLKA, K., FEUCHTINGER, A., JANSSEN, K. P., KUPPEN, P. J., VAN DE VELDE, C. J., WEIRICH, G., ERLMEIER, F., LANGER, R., AUBELE, M., ZITZELSBERGER, H., MCDONNELL, L., AICHLER, M. & WALCH, A. 2016. High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. *Nat Protoc,* 11**,** 1428-43.

MAENO, Y., TOYOSHIMA, T., FUJIOKA, H., ITO, Y., MESHNICK, S. R., BENAKIS, A., MILHOUS, W. K. & AIKAWA, M. 1993. Morphologic effects of artemisinin in Plasmodium falciparum. *Am J Trop Med Hyg,* 49**,** 485-91.

MAHAJAN, B., NOIVA, R., YADAVA, A., ZHENG, H., MAJAM, V., MOHAN, K. V., MOCH, J. K., HAYNES, J. D., NAKHASI, H. & KUMAR, S. 2006. Protein disulfide isomerase assisted protein folding in malaria parasites. *Int J Parasitol,* 36**,** 1037-48.

MAIER, A. G., COOKE, B. M., COWMAN, A. F. & TILLEY, L. 2009. Malaria parasite proteins that remodel the host erythrocyte. *Nat Rev Microbiol,* 7**,** 341-54.

MALLARI, J. P., OKSMAN, A., VAUPEL, B. & GOLDBERG, D. E. 2014. Kinase-associated endopeptidase 1 (Kae1) participates in an atypical ribosome-associated complex in the apicoplast of Plasmodium falciparum. *J Biol Chem,* 289**,** 30025-39.

MARIA BELEN, C., YONG, Z., KEITH, Z. H. & VERN, L. S. 2011. Purine and Pyrimidine Pathways as Targets in Plasmodium falciparum. *Current Topics in Medicinal Chemistry,* 11**,** 2103-2115.

MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,* 17.

MAYER, C., SLATER, L., ERAT, M. C., KONRAT, R. & VAKONAKIS, I. 2012. Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J Biol Chem,* 287**,** 7182-9.

MAYXAY, M., THONGPRASEUTH, V., KHANTHAVONG, M., LINDEGARDH, N., BARENDS, M., KEOLA, S., PONGVONGSA, T., PHOMPIDA, S., PHETSOUVANH, R., STEPNIEWSKA, K., WHITE, N. J. & NEWTON, P. N. 2006. An open, randomized comparison of artesunate plus mefloquine vs. dihydroartemisinin-piperaquine for the treatment of uncomplicated Plasmodium falciparum malaria in the Lao People's Democratic Republic (Laos). *Trop Med Int Health,* 11**,** 1157-65.

MBENGUE, A., BHATTACHARJEE, S., PANDHARKAR, T., LIU, H., ESTIU, G., STAHELIN, R. V., RIZK, S. S., NJIMOH, D. L., RYAN, Y., CHOTIVANICH, K., NGUON, C., GHORBAL, M., LOPEZ-RUBIO, J. J., PFRENDER, M., EMRICH, S., MOHANDAS, N., DONDORP, A. M., WIEST, O. & HALDAR, K. 2015. A molecular mechanism of artemisinin resistance in Plasmodium falciparum malaria. *Nature,* 520**,** 683-7.

MCNAMARA, C. & WINZELER, E. A. 2011. Target identification and validation of novel antimalarials. *Future Microbiol,* 6**,** 693-704.

MEISTER, S., PLOUFFE, D. M., KUHEN, K. L., BONAMY, G. M., WU, T., BARNES, S. W., BOPP, S. E., BORBOA, R., BRIGHT, A. T., CHE, J., COHEN, S., DHARIA, N. V., GAGARING, K., GETTAYACAMIN, M., GORDON, P., GROESSL, T., KATO, N., LEE, M. C., MCNAMARA, C. W., FIDOCK, D. A., NAGLE, A., NAM, T. G., RICHMOND, W., ROLAND, J., ROTTMANN, M., ZHOU, B., FROISSARD, P., GLYNNE, R. J., MAZIER, D., SATTABONGKOT, J., SCHULTZ, P. G., TUNTLAND, T., WALKER, J. R., ZHOU, Y., CHATTERJEE, A., DIAGANA, T. T. & WINZELER, E. A. 2011. Imaging of Plasmodium liver stages to drive next-generation antimalarial drug discovery. *Science,* 334**,** 1372-7.

MENARD, D., KHIM, N., BEGHAIN, J., ADEGNIKA, A. A., SHAFIUL-ALAM, M., AMODU, O., RAHIM-AWAB, G., BARNADAS, C., BERRY, A., BOUM, Y., BUSTOS, M. D., CAO, J., CHEN, J. H., COLLET, L., CUI, L., THAKUR, G. D., DIEYE, A., DJALLE, D., DORKENOO, M. A., EBOUMBOU-MOUKOKO, C. E., ESPINO, F. E., FANDEUR, T., FERREIRA-DA-CRUZ, M. F., FOLA, A. A., FUEHRER, H. P., HASSAN, A. M., HERRERA, S., HONGVANTHONG, B., HOUZE, S., IBRAHIM, M. L., JAHIRUL-KARIM, M., JIANG, L., KANO, S., ALI-KHAN, W., KHANTHAVONG, M., KREMSNER, P. G., LACERDA, M., LEANG, R., LEELAWONG, M., LI, M., LIN, K., MAZARATI, J. B., MENARD, S., MORLAIS, I., MUHINDO-MAVOKO, H., MUSSET, L., NA-BANGCHANG, K., NAMBOZI, M., NIARE, K., NOEDL, H., OUEDRAOGO, J. B., PILLAI, D. R., PRADINES, B., QUANG-PHUC, B., RAMHARTER, M., RANDRIANARIVELOJOSIA, M., SATTABONGKOT, J., SHEIKH-OMAR, A., SILUE, K. D., SIRIMA, S. B., SUTHERLAND, C., SYAFRUDDIN, D., TAHAR, R., TANG, L. H., TOURE, O. A., TSHIBANGU-WA-TSHIBANGU, P., VIGAN-WOMAS, I., WARSAME, M., WINI, L., ZAKERI, S., KIM, S., EAM, R., BERNE, L., KHEAN, C., CHY, S., KEN, M., LOCH, K., CANIER, L., DURU, V., LEGRAND, E., BARALE, J. C., STOKES, B., STRAIMER, J., WITKOWSKI, B., FIDOCK, D. A., ROGIER, C., RINGWALD, P., ARIEY, F., MERCEREAU-PUIJALON, O. & CONSORTIUM, K. 2016. A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms. *N Engl J Med,* 374**,** 2453-64.

MESHNICK, S. R., TSANG, T. W., LIN, F. B., PAN, H. Z., CHANG, C. N., KUYPERS, F., CHIU, D. & LUBIN, B. 1989. Activated oxygen mediates the antimalarial activity of qinghaosu. *Prog Clin Biol Res,* 313**,** 95-104.

MESHNICK, S. R., YANG, Y. Z., LIMA, V., KUYPERS, F., KAMCHONWONGPAISAN, S. & YUTHAVONG, Y. 1993. Iron-dependent free radical generation from the antimalarial agent artemisinin (qinghaosu). *Antimicrob Agents Chemother,* 37**,** 1108-14.

MIKKELSEN, R. B., WALLACH, D. F., VAN DOREN, E. & NILLNI, E. A. 1986. Membrane potential of erythrocytic stages of Plasmodium chabaudi free of the host cell membrane. *Mol Biochem Parasitol,* 21**,** 83-92.

MILES, A., IQBAL, Z., VAUTERIN, P., PEARSON, R., CAMPINO, S., THERON, M., GOULD, K., MEAD, D., DRURY, E., O'BRIEN, J., RUANO RUBIO, V., MACINNIS, B., MWANGI, J., SAMARAKOON, U., RANFORD-CARTWRIGHT, L., FERDIG, M., HAYTON, K., SU, X. Z., WELLEMS, T., RAYNER, J., MCVEAN, G. & KWIATKOWSKI, D. 2016. Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. *Genome Res,* 26**,** 1288-99.

MIOTTO, O., AMATO, R., ASHLEY, E. A., MACINNIS, B., ALMAGRO-GARCIA, J., AMARATUNGA, C., LIM, P., MEAD, D., OYOLA, S. O., DHORDA, M., IMWONG, M., WOODROW, C., MANSKE, M., STALKER, J., DRURY, E., CAMPINO, S., AMENGA-ETEGO, L., THANH, T. N., TRAN, H. T., RINGWALD, P., BETHELL, D., NOSTEN, F., PHYO, A. P., PUKRITTAYAKAMEE, S., CHOTIVANICH, K., CHUOR, C. M., NGUON, C., SUON, S., SRENG, S., NEWTON, P. N., MAYXAY, M., KHANTHAVONG, M., HONGVANTHONG, B., HTUT, Y., HAN, K. T., KYAW, M. P., FAIZ, M. A., FANELLO, C. I., ONYAMBOKO, M., MOKUOLU, O. A., JACOB, C. G., TAKALA-HARRISON, S., PLOWE, C. V., DAY, N. P., DONDORP, A. M., SPENCER, C. C., MCVEAN, G., FAIRHURST, R. M., WHITE, N. J. & KWIATKOWSKI, D. P. 2015. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nat Genet,* 47**,** 226-34.

MOK, S., ASHLEY, E. A., FERREIRA, P. E., ZHU, L., LIN, Z., YEO, T., CHOTIVANICH, K., IMWONG, M., PUKRITTAYAKAMEE, S., DHORDA, M., NGUON, C., LIM, P., AMARATUNGA, C., SUON, S., HIEN, T. T., HTUT, Y., FAIZ, M. A., ONYAMBOKO, M. A., MAYXAY, M., NEWTON, P. N., TRIPURA, R., WOODROW, C. J., MIOTTO, O., KWIATKOWSKI, D. P., NOSTEN, F., DAY, N. P., PREISER, P. R., WHITE, N. J., DONDORP, A. M., FAIRHURST, R. M. & BOZDECH, Z. 2015. Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science,* 347**,** 431-5.

MOLITOR, I. M., KNOBEL, S., DANG, C., SPIELMANN, T., ALLERA, A. & KONIG, G. M. 2004. Translation initiation factor eIF-5A from Plasmodium falciparum. *Mol Biochem Parasitol,* 137**,** 65-74.

MORITA, M., SANAI, H., HIRAMOTO, A., SATO, A., HIRAOKA, O., SAKURA, T., KANEKO, O., MASUYAMA, A., NOJIMA, M., WATAYA, Y. & KIM, H. S. 2012. Plasmodium falciparum endoplasmic reticulum-resident calcium binding protein is a possible target of synthetic antimalarial endoperoxides, N-89 and N-251. *J Proteome Res,* 11**,** 5704-11.

MOSES, J. E. & MOORHOUSE, A. D. 2007. The growing applications of click chemistry. *Chem Soc Rev,* 36**,** 1249-62.

MOURAY, E., MOUTIEZ, M., GIRAULT, S., SERGHERAERT, C., FLORENT, I. & GRELLIER, P. 2007. Biochemical properties and cellular localization of Plasmodium falciparum protein disulfide isomerase. *Biochimie,* 89**,** 337-46.

MULLER, I. B. & HYDE, J. E. 2010. Antimalarial drugs: modes of action and mechanisms of parasite resistance. *Future Microbiol,* 5**,** 1857-73.

MURALIDHARAN, V. & GOLDBERG, D. E. 2013. Asparagine repeats in Plasmodium falciparum proteins: good for nothing? *PLoS Pathog,* 9**,** e1003488.

MURALIDHARAN, V., OKSMAN, A., PAL, P., LINDQUIST, S. & GOLDBERG, D. E. 2012. Plasmodium falciparum heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nat Commun,* 3**,** 1310.

NA-BANGCHANG, K., MUHAMAD, P., RUAENGWEERAYUT, R., CHAIJAROENKUL, W. & KARBWANG, J. 2013. Identification of resistance of Plasmodium falciparum to artesunate-mefloquine combination in an area along the Thai-Myanmar border: integration of clinico-parasitological response, systemic drug exposure, and in vitro parasite sensitivity. *Malar J,* 12**,** 263.

NAIK, P. K., SRIVASTAVA, M., BAJAJ, P., JAIN, S., DUBEY, A., RANJAN, P., KUMAR, R. & SINGH, H. 2011. The binding modes and binding affinities of artemisinin derivatives with Plasmodium falciparum Ca2+-ATPase (PfATP6). *J Mol Model,* 17**,** 333-57.

NAJERA, J. A., GONZALEZ-SILVA, M. & ALONSO, P. L. 2011. Some lessons for the future from the Global Malaria Eradication Programme (1955-1969). *PLoS Med,* 8**,** e1000412.

NATALANG, O., BISCHOFF, E., DEPLAINE, G., PROUX, C., DILLIES, M. A., SISMEIRO, O., GUIGON, G., BONNEFOY, S., PATARAPOTIKUL, J., MERCEREAU-PUIJALON, O., COPPEE, J. Y. & DAVID, P. H. 2008. Dynamic RNA profiling in Plasmodium falciparum synchronized blood stages exposed to lethal doses of artesunate. *BMC Genomics,* 9**,** 388.

NING, X., GUO, J., WOLFERT, M. A. & BOONS, G. J. 2008. Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast huisgen cycloadditions. *Angew Chem Int Ed Engl,* 47**,** 2253-5.

NIXON, G. L., MOSS, D. M., SHONE, A. E., LALLOO, D. G., FISHER, N., O'NEILL, P. M., WARD, S. A. & BIAGINI, G. A. 2013. Antimalarial pharmacology and therapeutics of atovaquone. *J Antimicrob Chemother,* 68**,** 977-85.

NOEDL, H., SE, Y., SRIWICHAI, S., SCHAECHER, K., TEJA-ISAVADHARM, P., SMITH, B., RUTVISUTTINUNT, W., BETHELL, D., SURASRI, S., FUKUDA, M. M., SOCHEAT, D. & CHAN THAP, L. 2010. Artemisinin resistance in Cambodia: a clinical trial designed to address an emerging problem in Southeast Asia. *Clin Infect Dis,* 51**,** e82-9.

NZILA, A. 2006. The past, present and future of antifolates in the treatment of Plasmodium falciparum infection. *J Antimicrob Chemother,* 57**,** 1043-54.

O'NEILL, P. M., BARTON, V. E. & WARD, S. A. 2010. The molecular mechanism of action of artemisinin--the debate continues. *Molecules,* 15**,** 1705-21.

O'NEILL, P. M., BISHOP, L. P. D., SEARLE, N. L., MAGGS, J. L., WARD, S. A., PARK, B. K. & MABBS, F. 2000. Biomimetic Fe(II)-Mediated Degradation of Arteflene (Ro-42-1611). The First EPR Spin-Trapping Evidence for the Previously Postulated Secondary Carbon-Centered Cyclohexyl Radical. *The Journal of Organic Chemistry,* 65**,** 1578-1582.

OKOMBO, J., ABDI, A. I., KIARA, S. M., MWAI, L., POLE, L., SUTHERLAND, C. J., NZILA, A. & OCHOLA-OYIER, L. I. 2013. Repeat polymorphisms in the low-complexity regions of Plasmodium falciparum ABC

transporters and associations with in vitro antimalarial responses. *Antimicrob Agents Chemother,* 57**,** 6196-204.

PALOQUE, L., RAMADANI, A. P., MERCEREAU-PUIJALON, O., AUGEREAU, J. M. & BENOIT-VICAL, F. 2016. Plasmodium falciparum: multifaceted resistance to artemisinins. *Malar J,* 15**,** 149.

PATZEWITZ, E. M., SALCEDO-SORA, J. E., WONG, E. H., SETHIA, S., STOCKS, P. A., MAUGHAN, S. C., MURRAY, J. A., KRISHNA, S., BRAY, P. G., WARD, S. A. & MULLER, S. 2013. Glutathione transport: a new role for PfCRT in chloroquine resistance. *Antioxid Redox Signal,* 19**,** 683-95.

PAVITHRA, S. R., KUMAR, R. & TATU, U. 2007. Systems analysis of chaperone networks in the malarial parasite Plasmodium falciparum. *PLoS Comput Biol,* 3**,** 1701-15.

PEASE, B. N., HUTTLIN, E. L., JEDRYCHOWSKI, M. P., TALEVICH, E., HARMON, J., DILLMAN, T., KANNAN, N., DOERIG, C., CHAKRABARTI, R., GYGI, S. P. & CHAKRABARTI, D. 2013. Global analysis of protein expression and phosphorylation of three stages of Plasmodium falciparum intraerythrocytic development. *J Proteome Res,* 12**,** 4028-45.

PEATEY, C. L., CHAVCHICH, M., CHEN, N., GRESTY, K. J., GRAY, K. A., GATTON, M. L., WATERS, N. C. & CHENG, Q. 2015. Mitochondrial Membrane Potential in a Small Subset of Artemisinin-Induced Dormant Plasmodium falciparum Parasites In Vitro. *J Infect Dis,* 212**,** 426-34.

PHOMPRADIT, P., MUHAMAD, P., WISEDPANICHKIJ, R., CHAIJAROENKUL, W. & NA-BANGCHANG, K. 2014. Four years' monitoring of in vitro sensitivity and candidate molecular markers of resistance of Plasmodium falciparum to artesunate-mefloquine combination in the Thai-Myanmar border. *Malar J,* 13**,** 23.

PHOMPRADIT, P., WISEDPANICHKIJ, R., MUHAMAD, P., CHAIJAROENKUL, W. & NA-BANGCHANG, K. 2011. Molecular analysis of pfatp6 and pfmdr1 polymorphisms and their association with in vitro sensitivity in Plasmodium falciparum isolates from the Thai-Myanmar border. *Acta Trop,* 120**,** 130-5.

PHYO, A. P., JITTAMALA, P., NOSTEN, F. H., PUKRITTAYAKAMEE, S., IMWONG, M., WHITE, N. J., DUPARC, S., MACINTYRE, F., BAKER, M. & MOHRLE, J. J. 2016. Antimalarial activity of artefenomel (OZ439), a novel synthetic antimalarial endoperoxide, in patients with Plasmodium falciparum and Plasmodium vivax malaria: an open-label phase 2 trial. *Lancet Infect Dis,* 16**,** 61-9.

PHYO, A. P., NKHOMA, S., STEPNIEWSKA, K., ASHLEY, E. A., NAIR, S., MCGREADY, R., LER MOO, C., AL-SAAI, S., DONDORP, A. M., LWIN, K. M., SINGHASIVANON, P., DAY, N. P. J., WHITE, N. J., ANDERSON, T. J. C. & NOSTEN, F. 2012. Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *The Lancet,* 379**,** 1960-1966.

PLATA, G., HSIAO, T. L., OLSZEWSKI, K. L., LLINAS, M. & VITKUP, D. 2010. Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network. *Mol Syst Biol,* 6**,** 408.

POINAR, G., JR. 2005. Plasmodium dominicana n. sp. (Plasmodiidae: Haemospororida) from Tertiary Dominican amber. *Syst Parasitol,* 61**,** 47-52.

PRICE, R. N., UHLEMANN, A. C., VAN VUGT, M., BROCKMAN, A., HUTAGALUNG, R., NAIR, S., NASH, D., SINGHASIVANON, P., ANDERSON, T. J., KRISHNA, S., WHITE, N. J. & NOSTEN, F. 2006. Molecular and pharmacological determinants of the therapeutic response to artemether-lumefantrine in multidrug-resistant Plasmodium falciparum malaria. *Clin Infect Dis,* 42**,** 1570-7.

RAMIL, C. P. & LIN, Q. 2013. Bioorthogonal chemistry: strategies and recent developments. *Chem Commun (Camb),* 49**,** 11007-22.

RATHORE, S., DATTA, G., KAUR, I., MALHOTRA, P. & MOHMMED, A. 2015. Disruption of cellular homeostasis induces organelle stress and triggers apoptosis like cell-death pathways in malaria parasite. *Cell Death Dis,* 6**,** e1803.

REED, M. B., SALIBA, K. J., CARUANA, S. R., KIRK, K. & COWMAN, A. F. 2000. Pgh1 modulates sensitivity and resistance to multiple antimalarials in Plasmodium falciparum. *Nature,* 403**,** 906-9.

RICHIE, T. L., BILLINGSLEY, P. F., SIM, B. K., JAMES, E. R., CHAKRAVARTY, S., EPSTEIN, J. E., LYKE, K. E., MORDMULLER, B., ALONSO, P., DUFFY, P. E., DOUMBO, O. K., SAUERWEIN, R. W., TANNER, M., ABDULLA, S., KREMSNER, P. G., SEDER, R. A. & HOFFMAN, S. L. 2015. Progress with Plasmodium falciparum sporozoite (PfSPZ)-based malaria vaccines. *Vaccine,* 33**,** 7452-61.

ROSENBERG, S. A. & MCINTOSH, J. R. 1968. Erythrocyte membranes: Effects of sonication. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 163**,** 285-289.

ROSENTHAL, P. J. 1998. Proteases of malaria parasites: new targets for chemotherapy. *Emerg Infect Dis,* 4**,** 49-57.

ROSTOVTSEV, V. V., GREEN, L. G., FOKIN, V. V. & SHARPLESS, K. B. 2002. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angewandte Chemie International Edition,* 41**,** 2596-2599.

ROTHBERG, J. M. & LEAMON, J. H. 2008. The development and impact of 454 sequencing. *Nat Biotechnol,* 26**,** 1117-24.

ROTTMANN, M., MCNAMARA, C., YEUNG, B. K., LEE, M. C., ZOU, B., RUSSELL, B., SEITZ, P., PLOUFFE, D. M., DHARIA, N. V., TAN, J., COHEN, S. B., SPENCER, K. R., GONZALEZ-PAEZ, G. E., LAKSHMINARAYANA, S. B., GOH, A., SUWANARUSK, R., JEGLA, T., SCHMITT, E. K., BECK, H. P., BRUN, R., NOSTEN, F., RENIA, L., DARTOIS, V., KELLER, T. H., FIDOCK, D. A., WINZELER, E. A. & DIAGANA, T. T. 2010. Spiroindolones, a potent compound class for the treatment of malaria. *Science,* 329**,** 1175-80.

ROWLAND, M. M., BOSTIC, H. E., GONG, D., SPEERS, A. E., LUCAS, N., CHO, W., CRAVATT, B. F. & BEST, M. D. 2011. Phosphatidylinositol 3,4,5-trisphosphate activity probes for the labeling and proteomic characterization of protein binding partners. *Biochemistry,* 50**,** 11143-61.

RTS, S. C. T. P. 2015. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *The Lancet,* 386**,** 31-45.

SANGER, F., AIR, G. M., BARRELL, B. G., BROWN, N. L., COULSON, A. R., FIDDES, J. C., HUTCHISON, C. A., SLOCOMBE, P. M. & SMITH, M. 1977. Nucleotide sequence of bacteriophage [phi]X174 DNA. *Nature,* 265**,** 687-695.

SANZ, L. M., CRESPO, B., DE-COZAR, C., DING, X. C., LLERGO, J. L., BURROWS, J. N., GARCIA-BUSTOS, J. F. & GAMO, F. J. 2012. P. falciparum in vitro killing rates allow to discriminate between different antimalarial mode-of-action. *PLoS One,* 7**,** e30949.

SATO, A., HIRAMOTO, A., MORITA, M., MATSUMOTO, M., KOMICH, Y., NAKASE, Y., TANIGAWA, N., HIRAOKA, O., HIRAMOTO, K., HAYATSU, H., HIGAKI, K., KAWAI, S., MASUYAMA, A., NOJIMA, M., WATAYA, Y. & KIM, H. S. 2011. Antimalarial activity of endoperoxide compound 6-(1,2,6,7-tetraoxaspiro[7.11]nonadec-4-yl)hexan-1-ol. *Parasitol Int,* 60**,** 270-3.

SCHLITZER, M. 2008. Antimalarial drugs - what is in use and what is in the pipeline. *Arch Pharm (Weinheim),* 341**,** 149-63.

SCHWACH, F., BUSHELL, E., GOMES, A. R., ANAR, B., GIRLING, G., HERD, C., RAYNER, J. C. & BILLKER, O. 2015. PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. *Nucleic Acids Res,* 43**,** D1176-82.

SCHWARTZ, L., BROWN, G. V., GENTON, B. & MOORTHY, V. S. 2012. A review of malaria vaccine clinical projects based on the WHO rainbow table. *Malar J,* 11**,** 11.

SHANDILYA, A., CHACKO, S., JAYARAM, B. & GHOSH, I. 2013. A plausible mechanism for the antimalarial activity of artemisinin: A computational approach. *Sci Rep,* 3**,** 2513.

SHANKS, G. D., OLOO, A. J., ALEMAN, G. M., OHRT, C., KLOTZ, F. W., BRAITMAN, D., HORTON, J. & BRUECKNER, R. 2001. A new primaquine analogue, tafenoquine (WR 238605), for prophylaxis against Plasmodium falciparum malaria. *Clin Infect Dis,* 33**,** 1968-74.

SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res,* 13**,** 2498-504.

SHAW, P. J., CHAOTHEING, S., KAEWPROMMAL, P., PIRIYAPONGSA, J., WONGSOMBAT, C., SUWANNAKITTI, N., KOONYOSYING, P., UTHAIPIBULL, C., YUTHAVONG, Y. & KAMCHONWONGPAISAN, S. 2015. Plasmodium parasites mount an arrest response to dihydroartemisinin, as revealed by whole transcriptome shotgun sequencing (RNA-seq) and microarray study. *BMC Genomics,* 16**,** 830.

SHENDURE, J. & JI, H. 2008. Next-generation DNA sequencing. *Nat Biotechnol,* 26**,** 1135-45.

SHORTT, H. E. & GARNHAM, P. C. 1948. Pre-erythrocytic stage in mammalian malaria parasites. *Nature,* 161**,** 126.

SIDHU, A. B., UHLEMANN, A. C., VALDERRAMOS, S. G., VALDERRAMOS, J. C., KRISHNA, S. & FIDOCK, D. A. 2006. Decreasing pfmdr1 copy number in plasmodium falciparum malaria heightens susceptibility to mefloquine, lumefantrine, halofantrine, quinine, and artemisinin. *J Infect Dis,* 194**,** 528-35.

SILVA, A. M., LEE, A. Y., GULNIK, S. V., MAIER, P., COLLINS, J., BHAT, T. N., COLLINS, P. J., CACHAU, R. E., LUKER, K. E., GLUZMAN, I. Y., FRANCIS, S. E., OKSMAN, A., GOLDBERG, D. E. & ERICKSON, J. W. 1996. Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 10034-10039.

SLATER, A. F. 1993. Chloroquine: mechanism of drug action and resistance in Plasmodium falciparum. *Pharmacol Ther,* 57**,** 203-35.

SLATER, A. F. & CERAMI, A. 1992. Inhibition by chloroquine of a novel haem polymerase enzyme activity in malaria trophozoites. *Nature,* 355**,** 167-9.

SLETTEN, E. M. & BERTOZZI, C. R. 2009. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angewandte Chemie International Edition,* 48**,** 6974-6998.
SONENBERG, N. & HINNEBUSCH, A. G. 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell,* 136**,** 731-45.
SPRENGER, J., CAREY, J., SVENSSON, B., WENGEL, V. & PERSSON, L. 2016. Binding and Inhibition of Spermidine Synthase from Plasmodium falciparum and Implications for In Vitro Inhibitor Testing. *PLoS One,* 11**,** e0163442.
SRIVASTAVA, I. K., MORRISEY, J. M., DARROUZET, E., DALDAL, F. & VAIDYA, A. B. 1999. Resistance mutations reveal the atovaquone-binding domain of cytochrome b in malaria parasites. *Mol Microbiol,* 33**,** 704-11.
STEEN, H. & MANN, M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol,* 5**,** 699-711.
STOCKS, P. A., BRAY, P. G., BARTON, V. E., AL-HELAL, M., JONES, M., ARAUJO, N. C., GIBBONS, P., WARD, S. A., HUGHES, R. H., BIAGINI, G. A., DAVIES, J., AMEWU, R., MERCER, A. E., ELLIS, G. & O'NEILL, P. M. 2007. Evidence for a common non-heme chelatable-iron-dependent activation mechanism for semisynthetic and synthetic endoperoxide antimalarial drugs. *Angew Chem Int Ed Engl,* 46**,** 6278-83.
STRAIMER, J., GNADIG, N. F., WITKOWSKI, B., AMARATUNGA, C., DURU, V., RAMADANI, A. P., DACHEUX, M., KHIM, N., ZHANG, L., LAM, S., GREGORY, P. D., URNOV, F. D., MERCEREAU-PUIJALON, O., BENOIT-VICAL, F., FAIRHURST, R. M., MENARD, D. & FIDOCK, D. A. 2015. Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates. *Science,* 347**,** 428-31.
STURM, A., AMINO, R., VAN DE SAND, C., REGEN, T., RETZLAFF, S., RENNENBERG, A., KRUEGER, A., POLLOK, J. M., MENARD, R. & HEUSSLER, V. T. 2006. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science,* 313**,** 1287-90.
SUBRAMANIAN, N., SREEMANTHULA, J. B., BALAJI, B., KANWAR, J. R., BISWAS, J. & KRISHNAKUMAR, S. 2014. A strain-promoted alkyne-azide cycloaddition (SPAAC) reaction of a novel EpCAM aptamer-fluorescent conjugate for imaging of cancer cells. *Chem Commun (Camb),* 50**,** 11810-3.
SULLIVAN, D. J., GLUZMAN, I. Y. & GOLDBERG, D. E. 1996a. Plasmodium Hemozoin Formation Mediated by Histidine-Rich Proteins. *Science,* 271**,** 219-222.
SULLIVAN, D. J., JR., GLUZMAN, I. Y., RUSSELL, D. G. & GOLDBERG, D. E. 1996b. On the molecular mechanism of chloroquine's antimalarial action. *Proc Natl Acad Sci U S A,* 93**,** 11865-70.
TAN, K. R., MAGILL, A. J., PARISE, M. E., ARGUIN, P. M., CENTERS FOR DISEASE, C. & PREVENTION 2011. Doxycycline for malaria chemoprophylaxis and treatment: report from the CDC expert meeting on malaria chemoprophylaxis. *Am J Trop Med Hyg,* 84**,** 517-31.
TANABE, K., ZAKERI, S., PALACPAC, N. M., AFSHARPAD, M., RANDRIANARIVELOJOSIA, M., KANEKO, A., MARMA, A. S., HORII, T. & MITA, T. 2011. Spontaneous mutations in the Plasmodium falciparum sarcoplasmic/endoplasmic reticulum Ca2+-ATPase (PfATP6) gene among geographically widespread parasite populations unexposed to artemisinin-based combination therapies. *Antimicrob Agents Chemother,* 55**,** 94-100.
TANG, W. K. & XIA, D. 2016. Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. *Sci Rep,* 6**,** 20037.
TANNER, M. & DE SAVIGNY, D. 2008. Malaria eradication back on the table. *Bull World Health Organ,* 86**,** 82.
TARDIEUX, I. & BAUM, J. 2016. Reassessing the mechanics of parasite motility and host-cell invasion. *J Cell Biol,* 214**,** 507-15.
TEUSCHER, F., GATTON, M. L., CHEN, N., PETERS, J., KYLE, D. E. & CHENG, Q. 2010. Artemisinin-induced dormancy in plasmodium falciparum: duration, recovery rates, and implications in treatment failure. *J Infect Dis,* 202**,** 1362-8.
TIAN, H., SAKMAR, T. P. & HUBER, T. 2016. A simple method for enhancing the bioorthogonality of cyclooctyne reagent. *Chem Commun (Camb),* 52**,** 5451-4.
TILLEY, L., STRAIMER, J., GNADIG, N. F., RALPH, S. A. & FIDOCK, D. A. 2016. Artemisinin Action and Resistance in Plasmodium falciparum. *Trends Parasitol*.
TIWARI, N. K., REYNOLDS, P. J. & CALDERON, A. I. 2016. Preliminary LC-MS Based Screening for Inhibitors of Plasmodium falciparum Thioredoxin Reductase (PfTrxR) among a Set of Antimalarials from the Malaria Box. *Molecules,* 21**,** 424.
TODD, J. F. J. 1991. Recommendations for nomenclature and symbolism for mass spectroscopy. *Pure and applied chemistry,* 63**,** 1541-1566.

TRIGLIA, T., WANG, P., SIMS, P. F., HYDE, J. E. & COWMAN, A. F. 1998. Allelic exchange at the endogenous genomic locus in Plasmodium falciparum proves the role of dihydropteroate synthase in sulfadoxine-resistant malaria. *Embo j,* 17**,** 3807-15.

TU, Y. 2011. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med,* 17**,** 1217-20.

TUNGU, P., MAGESA, S., MAXWELL, C., MALIMA, R., MASUE, D., SUDI, W., MYAMBA, J., PIGEON, O. & ROWLAND, M. 2010. Evaluation of PermaNet 3.0 a deltamethrin-PBO combination net against Anopheles gambiae and pyrethroid resistant Culex quinquefasciatus mosquitoes: an experimental hut trial in Tanzania. *Malar J,* 9**,** 21.

UHLEMANN, A.-C., CAMERON, A., ECKSTEIN-LUDWIG, U., FISCHBARG, J., ISEROVICH, P., ZUNIGA, F. A., EAST, M., LEE, A., BRADY, L., HAYNES, R. K. & KRISHNA, S. 2012. A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nat Struct Mol Biol,* 19**,** 264-264.

UHLEMANN, A. C., CAMERON, A., ECKSTEIN-LUDWIG, U., FISCHBARG, J., ISEROVICH, P., ZUNIGA, F. A., EAST, M., LEE, A., BRADY, L., HAYNES, R. K. & KRISHNA, S. 2005. A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nat Struct Mol Biol,* 12**,** 628-9.

VAID, A., RANJAN, R., SMYTHE, W. A., HOPPE, H. C. & SHARMA, P. 2010. PfPI3K, a phosphatidylinositol-3 kinase from Plasmodium falciparum, is exported to the host erythrocyte and is involved in hemoglobin trafficking. *Blood,* 115**,** 2500-7.

VALDERRAMOS, S. G., SCANFELD, D., UHLEMANN, A. C., FIDOCK, D. A. & KRISHNA, S. 2010. Investigations into the role of the Plasmodium falciparum SERCA (PfATP6) L263E mutation in artemisinin action and resistance. *Antimicrob Agents Chemother,* 54**,** 3842-52.

VAN GEEL, R., PRUIJN, G. J., VAN DELFT, F. L. & BOELENS, W. C. 2012. Preventing thiol-yne addition improves the specificity of strain-promoted azide-alkyne cycloaddition. *Bioconjug Chem,* 23**,** 392-8.

VAN HENSBROEK, M. B., MORRIS-JONES, S., MEISNER, S., JAFFAR, S., BAYO, L., DACKOUR, R., PHILLIPS, C. & GREENWOOD, B. M. 1995. Iron, but not folic acid, combined with effective antimalarial therapy promotes haematological recovery in African children after acute falciparum malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 89**,** 672-676.

VAN NIEKERK, D. D., PENKLER, G. P., DU TOIT, F. & SNOEP, J. L. 2016. Targeting glycolysis in the malaria parasite Plasmodium falciparum. *FEBS J,* 283**,** 634-46.

VEIGA, M. I., DHINGRA, S. K., HENRICH, P. P., STRAIMER, J., GNADIG, N., UHLEMANN, A. C., MARTIN, R. E., LEHANE, A. M. & FIDOCK, D. A. 2016. Globally prevalent PfMDR1 mutations modulate Plasmodium falciparum susceptibility to artemisinin-based combination therapies. *Nat Commun,* 7**,** 11553.

VENNERSTROM, J. L., ARBE-BARNES, S., BRUN, R., CHARMAN, S. A., CHIU, F. C., CHOLLET, J., DONG, Y., DORN, A., HUNZIKER, D., MATILE, H., MCINTOSH, K., PADMANILAYAM, M., SANTO TOMAS, J., SCHEURER, C., SCORNEAUX, B., TANG, Y., URWYLER, H., WITTLIN, S. & CHARMAN, W. N. 2004. Identification of an antimalarial synthetic trioxolane drug development candidate. *Nature,* 430**,** 900-4.

VENNERSTROM, J. L., NUZUM, E. O., MILLER, R. E., DORN, A., GERENA, L., DANDE, P. A., ELLIS, W. Y., RIDLEY, R. G. & MILHOUS, W. K. 1999. 8-Aminoquinolines active against blood stage Plasmodium falciparum in vitro inhibit hematin polymerization. *Antimicrob Agents Chemother,* 43**,** 598-602.

VERBERKMOES, N. C., DENEF, V. J., HETTICH, R. L. & BANFIELD, J. F. 2009. Systems Biology: Functional analysis of natural microbial consortia using community proteomics. *Nat Rev Micro,* 7**,** 196-205.

WANG, F., KRAI, P., DEU, E., BIBB, B., LAURITZEN, C., PEDERSEN, J., BOGYO, M. & KLEMBA, M. 2011. Biochemical characterization of Plasmodium falciparum dipeptidyl aminopeptidase 1. *Mol Biochem Parasitol,* 175**,** 10-20.

WANG, J., ZHANG, C. J., CHIA, W. N., LOH, C. C., LI, Z., LEE, Y. M., HE, Y., YUAN, L. X., LIM, T. K., LIU, M., LIEW, C. X., LEE, Y. Q., ZHANG, J., LU, N., LIM, C. T., HUA, Z. C., LIU, B., SHEN, H. M., TAN, K. S. & LIN, Q. 2015. Haem-activated promiscuous targeting of artemisinin in Plasmodium falciparum. *Nat Commun,* 6**,** 10111.

WANG, P., SIMS, P. F. & HYDE, J. E. 1997. A modified in vitro sulfadoxine susceptibility assay for Plasmodium falciparum suitable for investigating Fansidar resistance. *Parasitology,* 115 (Pt 3)**,** 223-30.

WARNCKE, J. D., VAKONAKIS, I. & BECK, H. P. 2016. Plasmodium Helical Interspersed Subtelomeric (PHIST) Proteins, at the Center of Host Cell Remodeling. *Microbiol Mol Biol Rev,* 80**,** 905-27.

WATERMEYER, J. M., HALE, V. L., HACKETT, F., CLARE, D. K., CUTTS, E. E., VAKONAKIS, I., FLECK, R. A., BLACKMAN, M. J. & SAIBIL, H. R. 2016. A spiral scaffold underlies cytoadherent knobs in Plasmodium falciparum-infected erythrocytes. *Blood,* 127**,** 343-51.

WEAVER, E. M. & HUMMON, A. B. 2013. Imaging mass spectrometry: from tissue sections to cell cultures. *Adv Drug Deliv Rev,* 65**,** 1039-55.

WELLS, T. N., HOOFT VAN HUIJSDUIJNEN, R. & VAN VOORHIS, W. C. 2015. Malaria medicines: a glass half full? *Nat Rev Drug Discov,* 14**,** 424-42.

WERNER, C., STUBBS, M. T., KRAUTH-SIEGEL, R. L. & KLEBE, G. 2005. The crystal structure of Plasmodium falciparum glutamate dehydrogenase, a putative target for novel antimalarial drugs. *J Mol Biol,* 349**,** 597-607.

WHITE, N. J. 1999. Delaying antimalarial drug resistance with combination chemotherapy. *Parassitologia,* 41**,** 301-8.

WHITE, N. J., PUKRITTAYAKAMEE, S., PHYO, A. P., RUEANGWEERAYUT, R., NOSTEN, F., JITTAMALA, P., JEEYAPANT, A., JAIN, J. P., LEFEVRE, G., LI, R., MAGNUSSON, B., DIAGANA, T. T. & LEONG, F. J. 2014. Spiroindolone KAE609 for falciparum and vivax malaria. *N Engl J Med,* 371**,** 403-10.

WHO 2012. Global plan for insecticide resistance management in malaria vectors. Geneva.

WHO 2014a. Safety of 8-aminoquinoline antimalarial medicines. Geneva: WHO.

WHO 2014b. Status report on artemisinin resistance.

WHO 2015a. Guidelines for the treatment of malaria. Third edition. Third edition ed. Geneva: World Health Organization.

WHO 2015b. World malaria report 2015. Geneva.

WHO 2016a. The rainbow tables. *In:* WHO (ed.).

WHO 2016b. World health statistics 2016: monitoring health for the SDGs, sustainable development goals. Geneva.

WILAIRAT, P., KUMPORNSIN, K. & CHOOKAJORN, T. 2016. Plasmodium falciparum malaria: Convergent evolutionary trajectories towards delayed clearance following artemisinin treatment. *Med Hypotheses,* 90**,** 19-22.

WITKOWSKI, B., AMARATUNGA, C., KHIM, N., SRENG, S., CHIM, P., KIM, S., LIM, P., MAO, S., SOPHA, C., SAM, B., ANDERSON, J. M., DUONG, S., CHUOR, C. M., TAYLOR, W. R. J., SUON, S., MERCEREAU-PUIJALON, O., FAIRHURST, R. M. & MENARD, D. 2013. Novel phenotypic assays for the detection of artemisinin-resistant Plasmodium falciparum malaria in Cambodia: in-vitro and ex-vivo drug-response studies. *The Lancet Infectious Diseases,* 13**,** 1043-1049.

WITKOWSKI, B., LELIEVRE, J., BARRAGAN, M. J., LAURENT, V., SU, X. Z., BERRY, A. & BENOIT-VICAL, F. 2010. Increased tolerance to artemisinin in Plasmodium falciparum is mediated by a quiescence mechanism. *Antimicrob Agents Chemother,* 54**,** 1872-7.

WOOTTON, J. C. 1994. Non-globular domains in protein sequences: Automated segmentation using complexity measures. *Computers & Chemistry,* 18**,** 269-285.

WORRELL, B. T., MALIK, J. A. & FOKIN, V. V. 2013. Direct evidence of a dinuclear copper intermediate in Cu(I)-catalyzed azide-alkyne cycloadditions. *Science,* 340**,** 457-60.

WU, W.-M., WU, Y., WU, Y.-L., YAO, Z.-J., ZHOU, C.-M., LI, Y. & SHAN, F. 1998. Unified Mechanistic Framework for the Fe(II)-Induced Cleavage of Qinghaosu and Derivatives/Analogues. The First Spin-Trapping Evidence for the Previously Postulated Secondary C-4 Radical. *Journal of the American Chemical Society,* 120**,** 3316-3325.

WU, W., HERRERA, Z., EBERT, D., BASKA, K., CHO, S. H., DERISI, J. L. & YEH, E. 2015. A chemical rescue screen identifies a Plasmodium falciparum apicoplast inhibitor targeting MEP isoprenoid precursor biosynthesis. *Antimicrob Agents Chemother,* 59**,** 356-64.

XIE, S. C., DOGOVSKI, C., HANSSEN, E., CHIU, F., YANG, T., CRESPO, M. P., STAFFORD, C., BATINOVIC, S., TEGUH, S., CHARMAN, S., KLONIS, N. & TILLEY, L. 2016. Haemoglobin degradation underpins the sensitivity of early ring stage Plasmodium falciparum to artemisinins. *J Cell Sci,* 129**,** 406-16.

YAO, J. Z., UTTAMAPINANT, C., POLOUKHTINE, A., BASKIN, J. M., CODELLI, J. A., SLETTEN, E. M., BERTOZZI, C. R., POPIK, V. V. & TING, A. Y. 2012. Fluorophore targeting to cellular proteins via enzyme-mediated azide ligation and strain-promoted cycloaddition. *J Am Chem Soc,* 134**,** 3720-8.

YAYON, A., CABANTCHIK, Z. I. & GINSBURG, H. 1985. Susceptibility of human malaria parasites to chloroquine is pH dependent. *Proc Natl Acad Sci U S A,* 82**,** 2784-8.

YEH, I., HANEKAMP, T., TSOKA, S., KARP, P. D. & ALTMAN, R. B. 2004. Computational analysis of Plasmodium falciparum metabolism: organizing genomic information to facilitate drug discovery. *Genome Res,* 14**,** 917-24.

ZIDOVETZKI, R., SHERMAN, I. W., PRUDHOMME, J. & CRAWFORD, J. 1994. Inhibition of Plasmodium falciparum lysophospholipase by anti-malarial drugs and sulphydryl reagents. *Parasitology,* 108 (Pt 3)**,** 249-55.

Publications|

ISMAIL, H. M., BARTON, V., **PHANCHANA, M**., CHAROENSUTTHIVARAKUL, S., WONG, M. H., HEMINGWAY, J., BIAGINI, G. A., O'NEILL, P. M. & WARD, S. A. 2016a. Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites Plasmodium falciparum 3D7. *Proc Natl Acad Sci U S A,* 113**,** 2080-5.

ISMAIL, H. M., BARTON, V. E., **PHANCHANA, M**., CHAROENSUTTHIVARAKUL, S., BIAGINI, G. A., WARD, S. A. & O'NEILL, P. M. 2016b. A Click Chemistry-Based Proteomic Approach Reveals that 1,2,4-Trioxolane and Artemisinin Antimalarials Share a Common Protein Alkylation Profile. *Angew Chem Int Ed Engl,* 55**,** 6401-5.