Fatty Acid Synthesis as a Target for New Antimalarials

Jonathan Edward Urch

Submitted for examination for PhD
September 2004

Supervisors: Dr Colin Berry and Prof. John. Harwood

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

From the time I started my PhD project in the Autumn of 2000 to the final viva over four years later I have received a huge amount of help and support from too many people to mention.

Firstly, I would like to thank my supervisors, Colin Berry and John Harwood, who provided me with fantastic support, insightful discussions and an excellent environment to work in. They gave me exceptional opportunities to collaborate with national and international research groups and I was lucky enough to spend considerable time visiting laboratories in Europe and the USA. I am especially grateful that both Colin and John put so much time and effort in reading this thesis, ensuring writing went rapidly and without problems.

Secondly, thanks to all of those in the Berry and Harwood labs especially; Dave, Lowri, Tim, Gareth, John Kay, Marion, Irena, Andrea, Katrina, Janice, Duncan, Tracy and Tracey for their support, experience and insightful contributions. A special mention should go to Dave, Gareth and Tim for putting up with some of my terrible jokes and for some of the most entertaining skills I have ever seen on a football pitch. Thanks also to Graham White and his lab who inspired me to start a PhD.

Things would not have been so much fun without the following people; Mike, Geth, Huw, Iant, Aled, Jeff, Erryl, Franck, Andy D, Sally, Meg, Jo, Aga, Jim, and all of those who turned out to play football with the department and for the Accies. A big thankyou to Sam and Nikki for going down the pub to talk a load of old toot and play the odd game of cribbage.

I am indebted to my family who have made all of this possible. To Mum, Dad, Jodie, Joanna, Emily, Nan, Gran, Bampi, Tracy and Graham, John and Pat, I cannot thank you enough for all your love and support. Finally, none of this would have been possible without Jen, who made me laugh during the bad times, made sure I ate properly every day and who makes me so happy. This is dedicated to her.

Summary

Malaria can be regarded as one of the world's worst health problems and its incidence is rising inexorably. It already accounts for the deaths of approximately three children every minute. This situation is exacerbated by the increased frequency of parasite resistance to current antimalarial agents and necessitates the development of new drugs to combat this disease

P. falciparum possesses a plastid-like organelle, termed the apicoplast, which contains a small, highly reduced 35kb genome encoding tRNA, DNA polymerases and ribosomal proteins. Nuclear proteins are targeted to the apicoplast using clearly defined N-terminal signal and target peptide sequences. This led to the discovery that the apicoplast may be the site of at least two anabolic pathways; isoprenoid synthesis and Type II fatty acid synthesis (FAS). This system is also present in bacteria and plants and differs significantly from the Type I FAS system found in humans. This makes the pathway an attractive target for novel antimalarials. Furthermore, the antibiotic thiolactomycin, inhibits the growth of bacteria via inhibition of Type II FAS.

This project describes the characterisation of pfFabF and pfFabH, the β-ketoacyl-ACP synthase enzymes in the Type II FAS pathway of *P. falciparum*. One of these proteins, pfFabF, was shown to be the target of thiolactomycin inhibition. This enzyme was also capable of catalysing the final step of the unsaturated FAS pathway.

Thiolactomycin derivatives synthesised in this project were tested against cultures of the parasite and both the condensing enzymes of *P. falciparum*. Significant improvements in the inhibitory activity against parasite cultures and one of the condensing enzymes were achieved compared to thiolactomycin. Although the inhibitory effects of these compounds was only in the micromolar range, this study demonstrates that thiolactomycin derivatives may provide lead compounds for the inhibition of *P. falciparum* by targeting either of the β -ketoacyl-ACP synthase proteins.

Contents

Chapter 3 The cloning and expression of *P. falciparum* FabF

Chapter 4 Substrate specificity and characterisation of MBP-pfFabF

Chapter 5 The cloning and Expression of *Plasmodium falciparum* FabH

Chapter 6 Inhibition of *P. falciparum* fatty acid synthesis by thiolactomycin and its derivatives

Abbreviations used in this report

Å	Angstrom
ACCase	acetyl-CoA carboxylase
ACP	acyl carrier protein
APS	ammonium persulphate
ATP	adenosine triphosphate
β-mercaptoethanol	2-hydroxyethylmercaptan
bp	base pair
cDNA	complementary DNA
CDS	coding sequence
CER	cerulenin
CoA	coenzyme A
CPM	counts per minute
CTP	cytidine triphosphate
D	attenuance
DDT	dichlorodiphenyltrichloroethane
DEPC	di-ethylene polychloride
DMSO	dimethyl-sulphoxide
DNase	deoxyribonuclease
dNTPs	deoxyribonucleotide triphosphates
DTT	dithiothreitol
ecACP	*Escherichia coli* ACP
ecFabD	*Escherichia coli* FabD
EDTA	diaminoethanetetra-acetic acid disodium salt
Fab	fatty acid biosynthesis
FAME	fatty acid methyl ester
FAS	Fatty Acid Synthase
GPI	glycosyl-phosphatidylinositol
GTP	guanosine triphosphate
HPLC	high performance liquid chromatography
IPTG	isopropyl β-D-thiogalactopyranoside

KAS	β-ketoacyl-(ACP) synthase
kDa	kilodaltons
kPa	kilopascals
LB	Luria Bertani
MBP	maltose binding protein
MIC	minimum inhibitory concentration
MUFA	monounsaturated fatty acid
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NEAT proteins	Nuclear-encoded apicoplast-targeted proteins
Ni-AP	nickel-alkaline phosphatase conjugate
PDHC	pyruvate dehydrogenase complex
PUFAs	poly-unsaturated fatty acids
rbc	uninfected red blood cells
SAP	shrimp alkaline phosphatase
SDM	site directed mutagenesis
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSC	saline sodium citrate
siRNA	short interfering RNA
ssRNA	single stranded RNA
TBS	Tris buffered saline
TCA	trichloroacetic acid
TE	Tris buffer containing EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
THF	tetrahydrofuran
TNFα	tumour necrosis factor alpha
TLM	thiolactomycin
tRNA	transfer RNA
UFA	unsaturated fatty acid
UTP	uridine triphosphate
WHO	World Health Organisation

Chapter 1

1.1 The Malaria Problem

Malaria poses an enormous burden on public health throughout the world, causing between 1-3 million deaths every year (World Health Organisation (WHO), 2002). In perspective, this mortality rate is second only to tuberculosis for any communicable disease, including AIDS. Malaria remains the most widespread parasitic disease, affecting 40% of the world's population, with more than 2.4 billion people living in areas with moderate or high incidence of malaria (WHO, 2002).

The causative agents of malaria; *Plasmodium falciparum*, *P. vivax*, *P. ovale* and *P. malariae* are amongst the most serious of human pathogens. These single cell protozoa infect both human and mosquito hosts. *P. vivax* is the most abundant species in terms of human parasite infections and has the widest geographic distribution. However, *P. falciparum* is responsible for a much greater number of deaths and is, therefore, the most clinically important (WHO, 2002). Other species of *Plasmodium* infect mammals, birds and reptiles with serious consequences to the host population. Of particular interest is *Plasmodium berghei*, which infects both mice and rats and can be used as a model for human infection. Human malaria parasites are transmitted by *Anopheles* mosquitoes, of which there are more than 30 species.

Although present in more than 100 countries worldwide, over 90% of malaria cases are distributed in sub-Saharan Africa. The highest mortality rate in these regions occurs in children under 5 years old and may be responsible for up to 25% of total infant deaths (Greenwood *et al.*, 1987). It is estimated that every minute, three children die from malaria or malaria-related complications. The disease is widespread across the tropics in; Central and South America, South East Asia, the Indian

sub-continent and even some areas of the Western Mediterranean (Figure 1.1). The distribution is largely restricted by the habitats of the vector *Anopheles gambiae*, which feeds preferentially on humans.

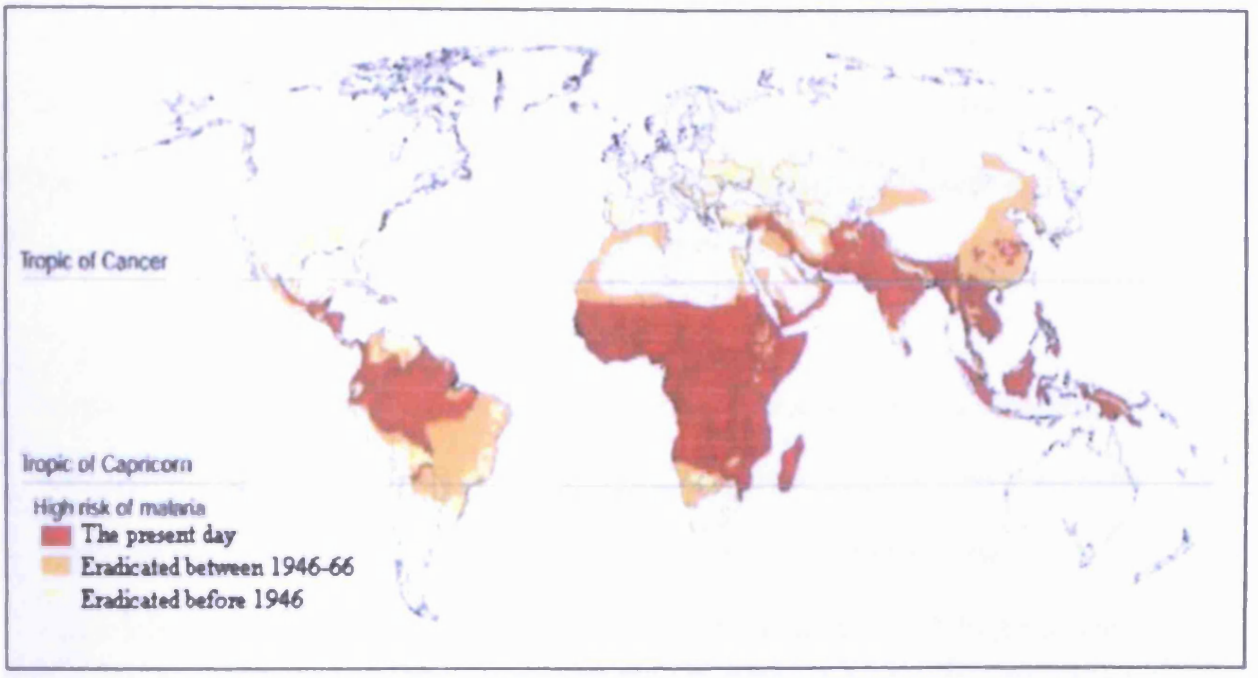


Figure 1.1 Changes in the global distribution of malaria

Malaria has been eradicated from the areas shown in yellow and orange in the last 60 years. These areas are mostly at risk from the resurgence of malaria and anopheles mosquito due to climate change. Reproduced from Sachs and Malaney (2002).

The social and economic implications in malaria endemic regions are staggering. The morbidity associated with malaria is responsible for a significant percentage of absenteeism from schools and work and has also been linked to impaired learning ability (Sachs and Malaney, 2002). *P. falciparum* also causes cerebral malaria, which, in 5-20% of surviving patients, can cause behavioural disorders and the impairment of initiating and planning tasks in later life. This affects the capabilities of the work force

and a study of income in malarious and non-malarious countries in 1995 showed that the average GDP in malarious countries in 1995 was US$1,526 compared to US$8,268 in non-malarious countries (Sachs and Malaney, 2002). The presence of malaria in tropical and sub-tropical regions correlates extensively with poverty in these areas. However, this correlation cannot solely be attributed to malaria, as poverty itself is partially responsible for some of the highest malaria transmission rates when funding for prevention methods such as bednets and insecticides is unavailable.

In recent years the incidence of malaria has seen a resurgence. Vector mosquitoes have become increasingly resistant to insecticides. Furthermore, the environmental impact of the most effective insecticides means that chemical interventions are frequently considered too toxic to use. In parallel, the emergence of drug resistant parasites has rendered many antimalarial compounds of moderate effectiveness (Olliaro, 2001). Resistance to chloroquine, the mainstay of malaria prophylaxis since the 1940s, was first reported in 1961, in highly endemic regions of South America (Columbia) and South East Asia (Thailand) (White, 1992). Subsequently, chloroquine resistance has spread to other endemic regions reaching West Africa in 1987 (Figure 1.2) (Reid *et al.*, 1998).

New drugs offer higher levels of protection as chemoprophylaxis but evidence suggests that they can have serious side effects at therapeutic doses e.g. neurological disturbances associated with mefloquine (Bem *et al.*, 1992; Hoebe *et al.*, 1997; Winstanley, 1996).

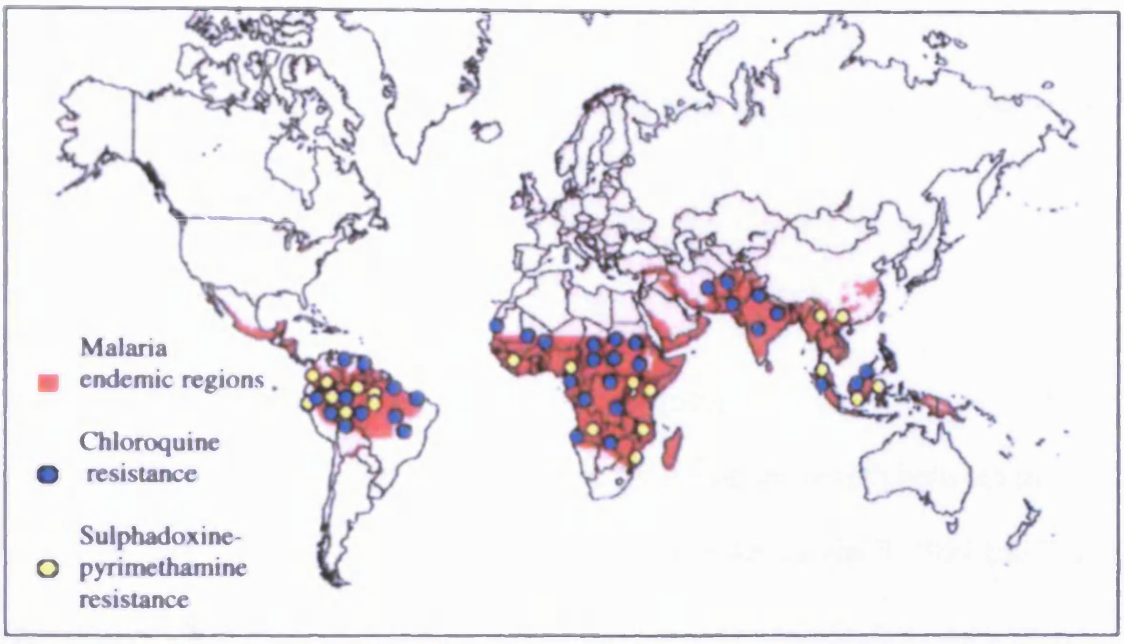


Figure 1.2 Drug resistance in malaria endemic regions

Resistance to the two most historically important antimalarial compounds is widespread. Resistance to chloroquine, a 4-amino quinoline, appeared approximately 50 years after it was first used. The anti-folate combination of sulphadoxine-pyrimethamine is the most widely used malaria drug (Newton and White, 1999), however, resistance occurs rapidly if this treatment is used excessively. Reproduced from Ridley, 2002.

Global climate changes are predicted to increase human population exposure to malaria. Simulations have shown that an increase of several degrees in the global climate would increase the mosquito population up to 100-fold in temperate regions (Martens *et al.*, 1995). Furthermore, conflicts and civil wars, industrialisation, increases in leisure travel to endemic areas and increases in the human population in endemic countries have resulted in a large, mobile reservoir of malaria. These factors

combined mean that there is a real risk that malaria could become an even greater problem in the tropics and spread further to areas of Australia, Southern USA and Southern Europe (Martens *et al.*, 1995), where it had been eradicated over the last century (Figure 1.1). Overall there is an urgent need for new, safe, inexpensive and effective antimalarial drugs.

1.2 Parasite lifecycle and the *Anopheles* mosquito

The malaria parasite has a complicated, multistage lifecycle split between the mosquito and human hosts (Figure 1.3). For overviews see Knell, 1991 and Phillips, 1983.

1.2.1 Pre-erythrocytic phase

Malaria parasites are introduced to the human host when an infected female mosquito takes a blood meal (to provide nutrients for maturation of her eggs). The bite from the mosquito releases an average of 2-3 infective cells termed sporozoites (Beier *et al.*, 1991), into the capillary bed of the skin, from where they migrate to invade the liver, being cleared from circulation in less than 30 minutes.

The sporozoites associate with the phagocytic Kupffer cell in the liver en route to the hepatocytes (Pradel and Frevert, 2001). In the hepatic parenchymal cells, the sporozoites mature into merozoites, which, in turn, mature into schizonts. The schizonts undergo several rounds of division, termed asexual schizogony, to form a large number of merozoites. In *P. falciparum* the hepatic phase lasts 6-12 days and ends as merozoites emerge from the liver. *P. ovale* and *P. vivax* have an extra

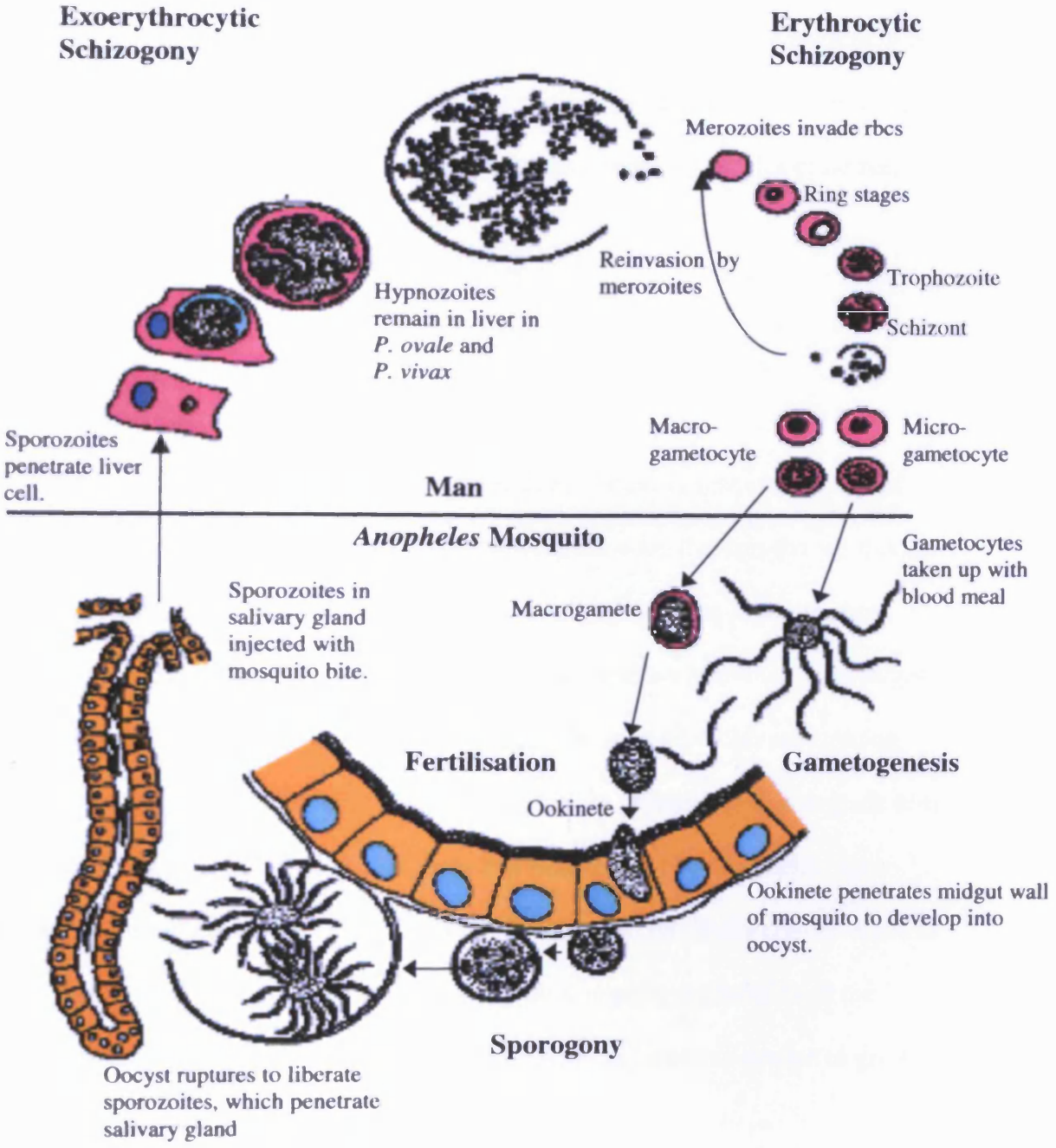


Figure 1.3 The lifecycle of human Plasmodium parasites

The diagram above represents the complete lifecycle of the four human malaria parasites. The latent hypnozoite stage in the liver is present only in two of the species as indicated above. Human cells are shown in pink, mosquito cells in orange and separate parasites in black. Adapted from Vickerman and Cox, (1967).

lifecycle stage in the liver. A proportion of the merozoites of these species become hypnozoites, laying dormant in liver cells and developing months or years later. It is these stages that are responsible for the characteristic relapses seen in both *P. vivax* and *P. ovale*. The hypnozoite stage is absent from the other two species of human malaria.

1.2.2 Erythrocytic cycles

The rupture of infected liver cells releases infective merozoites that are capable of invading erythrocytes in the bloodstream. Following invasion through the red blood cell (rbc) membrane, a single merozoite (Figure 1.4A), or in some *P. falciparum* infections, multiple merozoites (Figure 1.4B), develop into early and late ring stages; so-called because a distinctive open ring can clearly be seen after Giemsa staining. During this period, the ring parasites begin to synthesise proteinases that degrade host cell haemoglobin, in order to scavenge from this rich source of amino acids. Late rings mature into trophozoites (Figure 1.4C), which characteristically contain a single nucleus. During this trophozoite phase there is near-complete degradation of the erythrocyte's haemoglobin, providing the parasite with the nutrients needed to grow and divide into further merozoites.

Within 15-18h the trophozoite is fully developed and undergoes asexual schizogony to yield between 8-24 merozoites (Figure 1.4D). During this phase, the successful partitioning of organelles such as mitochondria, into each daughter cell is achieved (Waller *et al.*, 2000 and Striepen *et al.*, 2000). The rupture of the erythrocyte membrane releases these daughter merozoites (Figure 1.4E), causing wide distribution

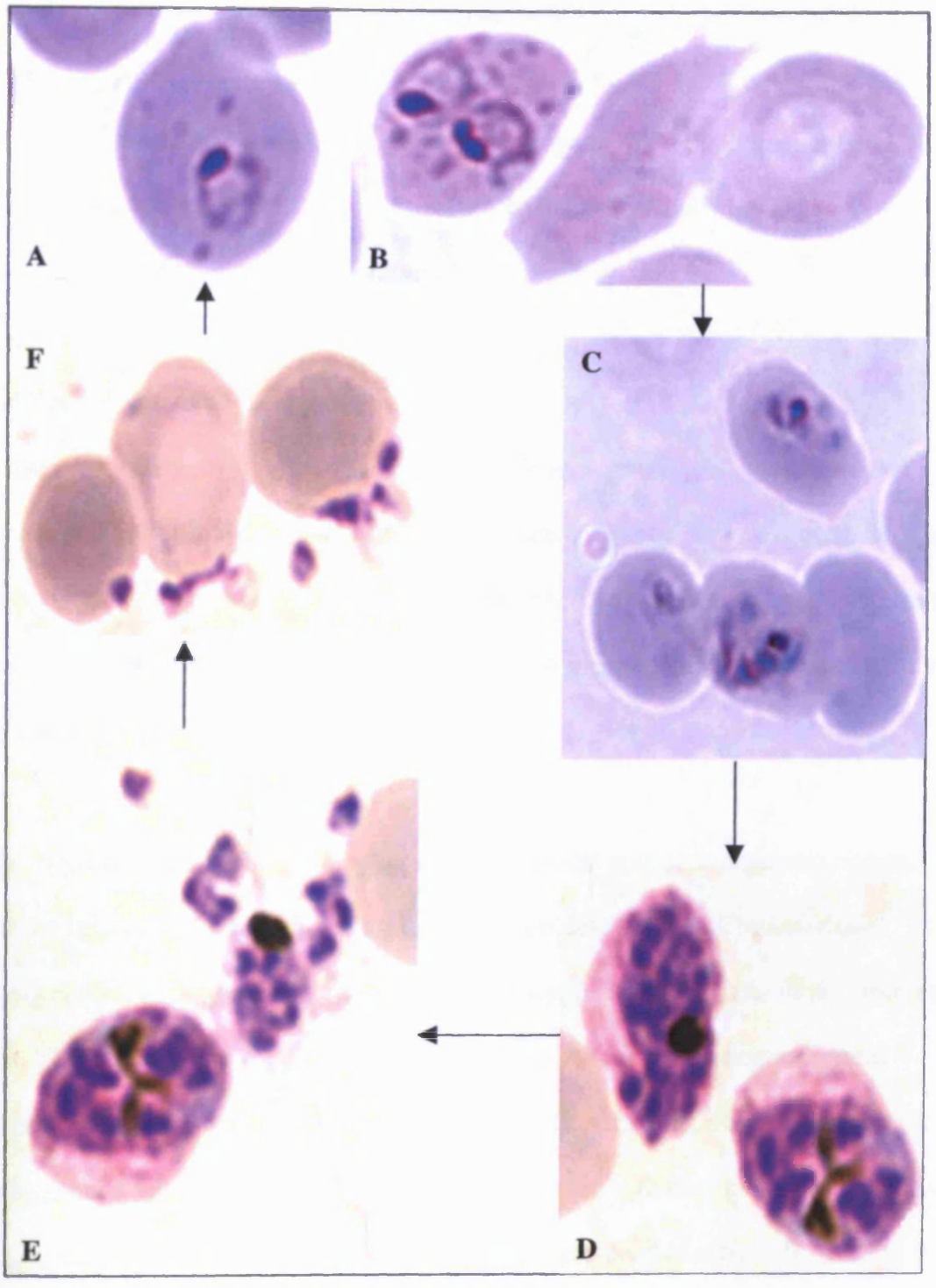


Figure 1.4 Erythrocytic stages of *P. falciparum* malaria
A = rbc infected with a single malaria parasite; B = a doubly infected ring adjacent to uninfected rbcs; C = trophozoites; D = late schizonts; E = rupture of irbc containing schizont; F = daughter merozoites invade uninfected rbc. Blood samples were taken from patients suffering from *P. falciparum* malaria at the Wellcome-Trust Centre, Kilifi, Kenya. Samples were smeared onto glass slides and stained with Giemsa. Photographs were captured using the Olympus BX41 microscope, with thanks to Prof. Alan Clarke (Cardiff University, UK).

of the merozoites and allowing them to invade other individual erythrocytes (Figure 1.4F). The asexual red blood cell cycle takes approximately 48h (tertian cycle) in *P. vivax* and *P. ovale* and 72h (quartan cycle) in *P. malariae*, whereas *P. falciparum* presents itself as an irregular tertian cycle (Knell, 1991).

1.2.3 Sexual stages of the parasite

Within infected rbc, some merozoites form male and female gametocytes, which develop no further in the human host but circulate until they die or a female mosquito ingests them as she feeds. The female gametocytes (or macrogametocytes) and male gametocytes (or microgametocytes) move through the mosquito digestive system until they reach the mid-gut.

Once in the mid-gut, the erythrocyte membrane is lost and gametogenesis occurs. Here the microgametocyte divides mitotically and develops until 16 individual microgametes, each possessing a nucleus and flagellum, detach themselves from each other. Thereafter, these male nuclei swim towards the fertile macrogametocyte, causing fertilisation several minutes after they became mobile. The resulting zygote matures into an ookinete within 12-18h and passes through the mid-gut cell wall 24h later. In the basal lamina of the cell wall, the ookinete divides repeatedly to form the oocyst; during the first 7 days, meiotic division occurs to produce an individual daughter sporozoite, this is followed by many rounds of mitosis producing several thousand sporozoites. Subsequently, the oocyst ruptures leading to the release of sporozoites into the haemocoel of the mosquito.

The sporozoites reach and penetrate the lumen of the salivary glands where they can remain indefinitely, until release during feeding infects a human host once more. The length of the parasite cycle in the vector mosquito is dependent on temperature and humidity, for instance, studies in *P. vivax* have shown that at 20°C the cycle takes 16 days but just 9 days at 24-25°C (Phillips, 1983).

1.3 Clinical manifestations of malaria

1.3.1 General Symptoms

No symptoms of malaria are observed in the pre-erythrocytic hepatic stage of infection, with all of the severe symptoms attributed to the asexual blood stages (Newton *et al.*, 2000). For this reason, the most useful antimalarials in severe infections are blood schizonticides.

Malaria is a complex disease, and the pathogenesis is still not completely understood. However, the most clinically recognisable symptom is the cyclic fever that may occur at regular intervals. Febrile episodes, which last for 2-3h, occur at 48h for *P. falciparum, P. vivax* and *P. ovale* but 72h for *P. malariae* (Knell, 1991). The initial symptoms of infection are similar to those of influenza or a hangover, including headaches, fever and general body pain. The similarity to these other conditions can cause the patient to delay seeking proper medical attention, leading to further complications. Cyclic fever is a result of the mass rupture of infected rbc, a process that becomes synchronised in the initial stages of infection and can lead to chronic anaemia in the patient.

The rupture of merozoites from infected erythrocytes, releases a malarial toxin (Mendis and Carter, 1995) that stimulates the release of cytokines, particularly tumour necrosis factor alpha (TNF-α) from T-cells and macrophages into the bloodstream (Karunaweera *et al.*, 1992). In turn, TNF-α induces the increase in serum levels of an endogenous pyrogen, interleukin-1, responsible for the characteristic malaria fever (Wahlgren *et al.*, 1995). In reality, fevers do not always follow a strong pattern because, frequently, the patient may have been bitten by more than one infected mosquito, complicating the infection and subsequent treatment.

1.3.2 Symptoms specific to *Plasmodium falciparum* malaria

P. falciparum causes more deaths per infection than any other human malaria species. The parasite displays several properties that cause more severe physiological effects in the host (WHO, 2002).

1) Hepatic Phase

Compared to other human *Plasmodium* species, *P. falciparum* produces a larger number of merozoites in the liver, resulting in a higher initial parasitaemia in the bloodstream (Despommier and Kerapelou, 1987).

2) Increased cyto-adherence

As *P. falciparum* reaches the late trophozoite stage, infected rbc become more adhesive to each other (agglutination), to uninfected erythrocytes (rosetting) and to the endothelial lining of capillaries (Udeinya *et al.*, 1981) (Figure 1.5). Increased cyto-adherence can cause congestion of the small blood vessels of the brain, lungs,

liver and kidneys that often causes serious organ damage and can lead to death. Agglutination is believed to provide protection to a proportion of the parasites because they are less likely to be removed from circulation by macrophages in the spleen or be attacked by circulating macrophages (Cranston *et al.*, 1984; Ramasamy, 1998). This helps to ensure the longevity of the parasite in the host.

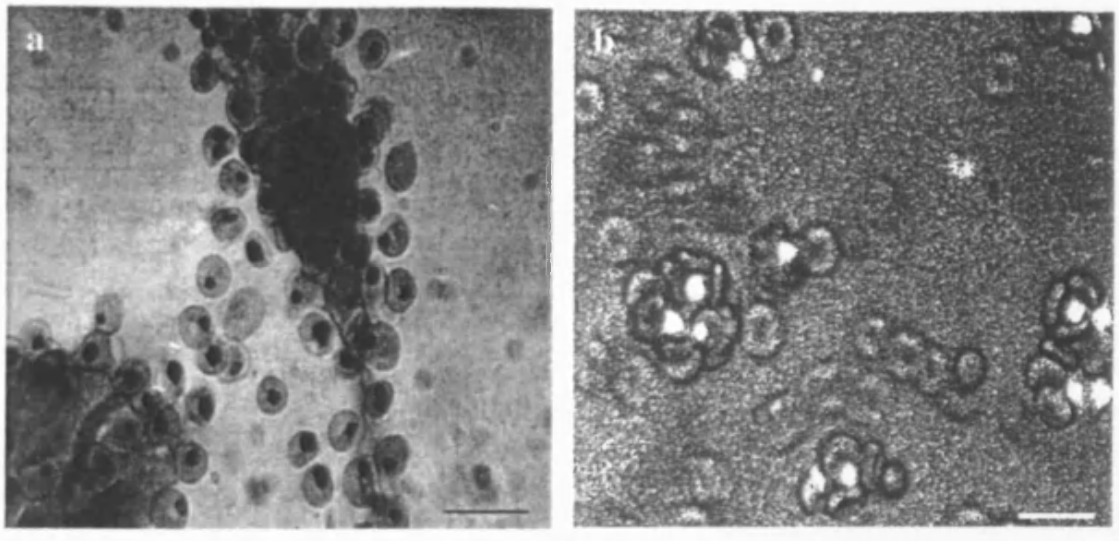


Figure 1.5 Agglutination and rosetting in *P. falciparum* infections

a) The cytoadherence of infected red blood cells to an endothelial membrane. The membrane structure can be observed at the bottom right of the image. b) The concomitant binding of uninfected red blood cells (rosetting) significantly contribute to the pathology of *P. falciparum* malaria. Scale bars = 20μm. Photos taken from Wahlgren *et al.*, 1994.

Cyto-adherence is caused by *P. falciparum* Erythrocytic Membrane Protein 1 (PfEMP-1). This is a family of 200-300kDa proteins encoded by a large number (50-150) of *var* genes (Baruch *et al.*, 1995; Smith *et al.*, 1995; Su *et al.*, 1995). Only

one of the *var* genes is expressed at any one time and expression is switched at regular periods to avoid destruction by the host immune response. The large number and variability of surface antigens make vaccine development difficult. These proteins, along with others, bind to endothelial surface receptors including the glycoprotein thrombospondin, intracellular adhesion molecule 1 (ICAM1) (Berendt *et al.*, 1989) and CD36 (Barnwell *et al.*, 1989; Doolan and Hoffman, 2001).

Infected cells may also bind to several uninfected red blood cells, a process termed rosetting (Wahlgren *et al.*, 1994). This process may provide merozoite progeny with a localised source for new infections, allowing for faster, more efficient proliferation of the parasite. Small (22-28kDa) proteins termed rosettins are produced by the parasite and displayed on the erythrocyte membrane, where they may interact with CD36 and A and B blood group components on uninfected blood cell membranes, resulting in adhesion (Wahlgren *et al.*, 1994 and 1995). It has also been suggested that *Pf*EMP1 could cause rosetting by binding to CR1 on uninfected cells (Doolan and Hoffman, 2001).

1.3.3 Cerebral Malaria

The two traits described in Section 1.3.2 lead to an increase in cerebral malaria, defined as a deep level of unconsciousness in the presence of a *P. falciparum* asexual parasitaemia (Newton *et al.*, 2000; Warrell and Molyneux 1990) and its associated complications include haemorrhages of surrounding blocked blood vessels and fitting. Cerebral malaria affects 575,000 African children every year and the mortality rate is 10-40%. Of those who survive, 5-20% exhibit long-lasting neurological disorders and

impaired intellectual functioning. Other chronic symptoms associated with increased cyto-adherence include liver failure, circulatory collapse, hypoglycaemia, anaemia, acute pulmonary oedema and kidney failure (Ramasamy, 1998).

1.3.4 *Plasmodium falciparum* malaria in pregnancy

Plasmodium falciparum malaria also causes complications to non-immune pregnant women (Nathwani *et al.*, 1992), mainly because the placenta is another common site of infected erythrocyte sequestration (Fried and Duffy, 1996). Occasionally, malaria parasites are able to pass through the placenta, infecting the foetus and causing severe complications. However, the foetus is generally well protected by maternal antibodies. Malaria infection in the mother or foetus can lead to maternal death and/or intrauterine death and if the foetus survives, low birth weight, which in turn may decrease the infant's chance of survival after birth (Warrell and Molyneux, 1990). In parallel, pregnant women have a higher incidence of malaria infection compared to other groups and this may be due to suppression of the immune system related to a decreased cellular immunity, especially within the third trimester (Sholapurkar *et al.*, 1990; Vleugels *et al.*, 1987). Ultimately, in order to decrease child and maternal mortality, new antimalarials must be safe and effective for administration to pregnant women.

1.4 *Anopheles* control

High rates of malaria transmission are attributed to large numbers of mosquitoes that efficiently transmit the malaria parasites. Therefore, controlling the *Anopheles* mosquito population should cause a decrease in the number of malaria cases.

Several strategies have been implemented with different levels of success throughout the world. The first programmes were aimed at the annihilation of the *Anopheles* mosquito population by destruction of their natural habitat and the use of insecticides. However, the World Health Organisation abandoned the mass spraying of houses and marshes with the most promising insecticide, DDT (dichlorodiphenyltrichloroethane), in 1969, although DDT is still a valuable tool in some malaria situations. Increasing evidence of harmful human (Schefczik and Buff, 1984) and environmental effects of DDT and an increase in mosquito resistance were reasons for this change in policy. Furthermore, resistance to DDT had arisen by the adaptation of the mosquito glutathione S-transferase system. Other insecticides have been tested but with little success, as resistance to second generation, organophosphorus insecticides has been linked to elevated levels of esterase in *Anopheles* (Hemingway, 1999).

Programmes to reduce the transmission rates of malaria have been more successful. The use of pyrethoid soaked nets and insect repellents have had significant impact on lowering human mortality and morbidity (WHO, 2002). Educational programmes, teaching people the best ways to avoid being bitten by mosquitoes, have also lowered rates of transmission. However, sustained effort and financial backing for these programmes have not been forthcoming, because they become expensive and difficult

to co-ordinate in developing countries that lack the infrastructure to support them. Therefore, many of these programmes have been dropped (WHO, 2002).

The production of transgenic mosquitoes that are unable to transmit the malaria parasite is an ambitious but achievable target. Mosquitoes that express antibodies to the malaria sporozoite have been generated in the laboratory and, when infected with *P. berghei*, such mosquitoes produce a much lower rate of transmission to mice (Ito *et al.*, 2002). However, the fitness of mosquito populations that have been reared in the laboratory and the retention of the antibody-encoding transgene in successive generations will have to be investigated fully.

1.5 The Apicomplexa

Apicomplexa are members of the eukaryotic Alveolata family, a group that also includes the ciliates and the dinoflagellates. The phylum Apicomplexa is composed of a diverse group of protozoan parasites that cause millions of human and animal deaths every year. The Apicomplexa includes the causative agents of malaria (*Plasmodium* spp.), toxoplasmosis (*Toxoplasma gondii*), coccidiosis of chickens (*Eimeria* spp.), babesiosis and East Coast cattle fever in cattle (*Babesia* spp. and *Theileria* spp., respectively) and cryptosporidiosis (*Cryptosporidium parvum*). A summary of the genera and the complications they cause are shown in Table 1.1. Unlike malaria, other apicomplexan parasites are not confined to the area surrounding the tropics, with the emergence of *Toxoplasma* and *Cryptosporidium* as opportunistic infections in HIV patients and the huge economic impact of drug-resistant *Eimeria* strains, also focusing attention on the identification of new drug targets.

Family	Genera and Species	Comments
Haemosporidae	*Plasmodium falciparum* *P. malariae* *P. ovale* *P. vivax*	Malaria causes 5 x 10^8 cases/year and 9 x 10^6 child deaths a year
Eimeriidae	*Cryptosporidium* spp.	Diarrhoea, causing problems in AIDS sufferers
	Cyclospora spp.	Traveller's diarrhoea, causing problems in AIDS sufferers
	Eimeria spp.	Poultry disease: UK spent £8.4M/year on control but still loses an estimated £20M/year (Williams, 1999)
	Sarcocystis spp.	Dog/Horse disease like toxoplasmosis
	Toxoplasma gondii	Cat disease, also a problem during human pregnancy or for AIDS sufferers. 50% of Americans exposed in their lifetime
Babesiidae	*Babesia* spp.	Cattle disease
Theileriidae	*Theileria* spp.	Cattle disease, occasionally infects humans
	Cytauxzoon spp.	Fatal cat disease with no cure

Table 1.1 Members of the Apicomplexa phylum

This Table describes some genera and species of the Apicomplexan phylum.

However, it is not a comprehensive list.

Despite the diverse range of primary and secondary hosts that these parasites infect, and the complicated lifecycles observed within this phylum, apicomplexan parasites share common subcellular and biochemical properties that are unique to the phylum. Of particular interest is the distinct organisation of organelles at the anterior region of the invading cell. This apical complex, from which the name of this group of organisms is derived, contains three secretory organelles that aid the invasion of the parasite into the host cell (Figure 1.6).

Figure 1.6 An electron micrograph of a *T. gondii* tachyzoite

A = apicoplast, C = conoid, D = dense granules, G =golgi body, Mi = micronemes, M = mitochondria, R = rhoptries. Adapted from Maréchal and Cesbron-Delauw, (2001).

1.6 The apicoplast

The diverse group of over 5000 organisms that compose the Apicomplexa share a very similar cellular morphology. As well as the usual eukaryotic organelles; a single nucleus, mitochondrion, rough endoplasmic reticulum and the golgi body, the parasites also harbour a number of unusual structures. These include the rhoptries, micronemes and the conoid (a cytoskeletal structure), which are organised at the apical region of each parasite cell prior to the invasion of the host cell (Figure 1.6). The recently discovered plastid-like organelle (apicoplast) in *P. falciparum* and other parasites is perhaps the most unexpected subcellular structure observed in these protozoa (Wilson *et al.*, 1996, Kohler *et al.*, 1997, Cai *et al.*, 2003).

1.6.1 Discovery of the apicoplast

The first suggestion of the organelle now known as the apicoplast, was the discovery of an extrachromosomal genetic element, first observed in electron micrographs of the avian malarial parasite *Plasmodium lophurae* (Kilejian *et al.*, 1975). This DNA was identified as a 35kb circular fragment that was also present in *Toxoplasma gondii* and *P. falciparum*. Originally the DNA was thought to be of mitochondrial origin. However, the mitochondrial genome was subsequently identified as a smaller 6kb fragment (Feagin, 1992). The identification of a separate mitochondrial genome suggested that, like in plants, the 35kb second extrachromosomal DNA fragment belonged to a plastid genome.

1.6.2 Identifying the location of a 35kb genetic fragment in the Apicomplexa

In early electron microscope studies of Apicomplexan parasites, the apicoplast did not go undetected. However, because the presence of a plastid was not expected the organelle was given several names including 'the golgi adjunct' and the '*Hohlzylinder*' or hollow cylinder (Sidall, 1992). The first evidence that this structure was also the location of the 35kb plastid genome was in *T. gondii* tachyzoites, when sequencing of a 35kb organellar genome allowed the design and *in situ* hybridisation of DNA probes to plastid DNA. These studies showed that, in each cell, there is a single, small (0.15-1.5μm) plastid that is closely associated with the golgi body. A recent study using apicoplast-targeted GFP and MitoTracker® dye, showed that the apicoplast and the mitochondrion were very closely associated in blood stages of *P. falciparum* (Waller *et al.*, 1999) (Figure 1.7).

Initial studies, utilising electron microscopy, showed that the apicoplast was surrounded by at least two membranes (McFadden *et al.*, 1996); with three, four and five membranes sometimes observed, results that were replicated in several other studies (McFadden *et al.*, 1997; Siddall, 1992). However, the debate surrounding the number of membranes was finally settled when four were consistently visualised surrounding the organelle, in electron micrographs (Kohler *et al.*, 1997) (Figure 1.8; Page 22). The four membranes surrounding the apicoplast indicate that the plastid may have been acquired by secondary symbiosis (Figure 1.9). This would have occurred when an Apicomplexan ancestor engulfed a eukaryotic photosynthetic organism. There is evidence that the photosynthetic eukaryote that was engulfed by the apicomplexan ancestor, was perhaps a red (rhodophyte) or green alga

(chlorophyte) and this issue is still being actively debated (Funes *et al.*, 2002 and 2003; Waller *et al.*, 2003b; for a review see Funes *et al.*, 2004).

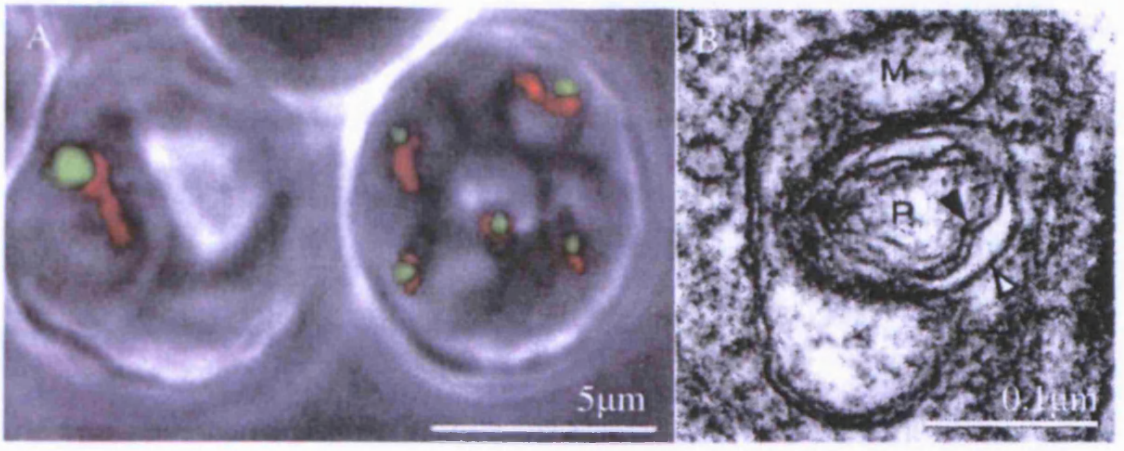


Figure 1.7 The closely associated apicoplast and mitochondria in *P. falciparum*

A) The close association of apicoplasts (green) and mitochondria (red) in two erythrocytic stages of *P. falciparum* (adapted from Waller *et al.*, 2000). GFP was targeted to the apicoplast by using an ACP $_{leader}$-GFP construct, which was then transfected and expressed in parasites. MitoTracker® Dye was used to define the mitochondria. The rbc on the left hand side is infected with a single ring stage parasite, whilst the rbc on the right is infected by more than one parasite.

B) Ultrastructure of the apicoplast and mitochondria of *Plasmodium falciparum* (Maréchal and Cesbron-Delauw, 2001). Labels are M = mitochondria, P = plastid (apicoplast). The white arrow shows the outermost membrane and the black arrow highlights the inner envelopes.

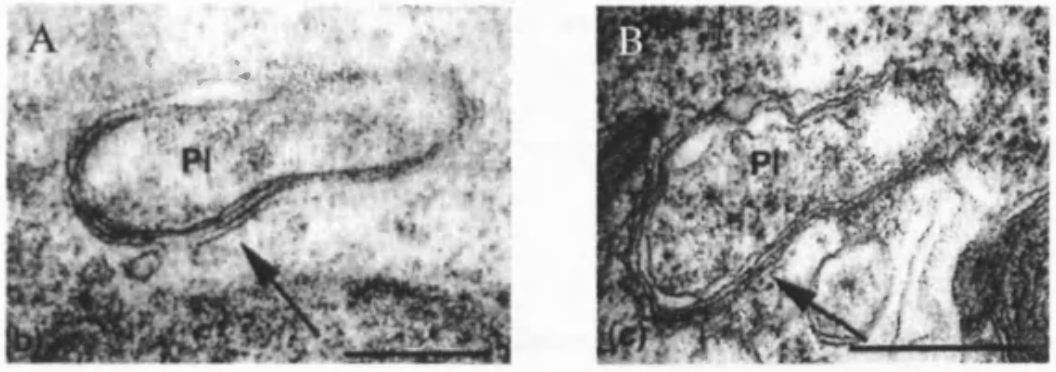


Figure 1.8 Electron microscope studies showing the apicoplast membranes

Ultrathin sections of *T. gondii* tachyzoites (A) and *P. falciparum* gametocyte (B) clearly show the four membranes (stained black with uranyl acetate) surrounding the apicoplast. These electron micrographs were taken from McFadden and Roos, (1999).

1.6.3 Genetic analysis of the apicoplast origin and function

The first genes sequenced from the circular DNA, encoded for subunits of RNA polymerase, *rpoB* and *rpoC* (McFadden *et al.,* 1991). These genes encode RNA polymerases similar to those observed in bacteria and chloroplast genomes and are distinct from nuclear and mitochondrial encoded RNA polymerases. The complete sequence of the 35kb DNA circle was elucidated in 1996 and described a 'stripped down' plastid DNA genome like that observed in non-photosynthetic plants (Wilson *et al.*, 1996). The complete lack of genes dedicated to photosynthesis suggests that the loss of the photosynthetic ability of the Apicomplexa occurred early in the evolution of the phylum. As well as genes encoding RNA polymerase, the circular DNA

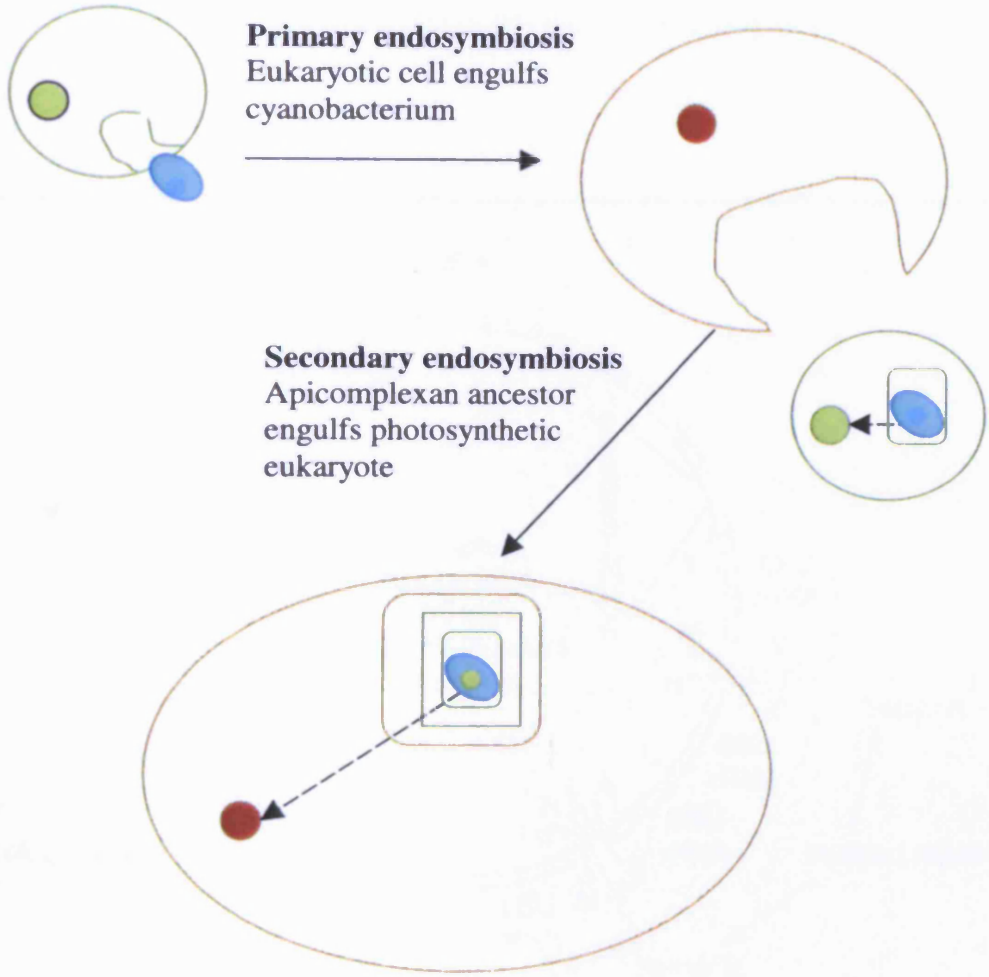


Figure 1.9 A possible origin of the apicoplast by secondary endosymbiosis

Dotted arrows indicate gene transfer from the nucleus of the engulfed organism to the nucleus of the new host.

contains a large number of 'housekeeping' genes encoding for a complete set of tRNAs, small- and large-ribosomal RNA subunits (arranged in plastid-like inverted repeats), ribosomal proteins and the elongating factor *tuf*A (Wilson *et al.*, 1996) (Figure 1.10). This suggests that the apicoplast retains the ability to synthesise proteins.

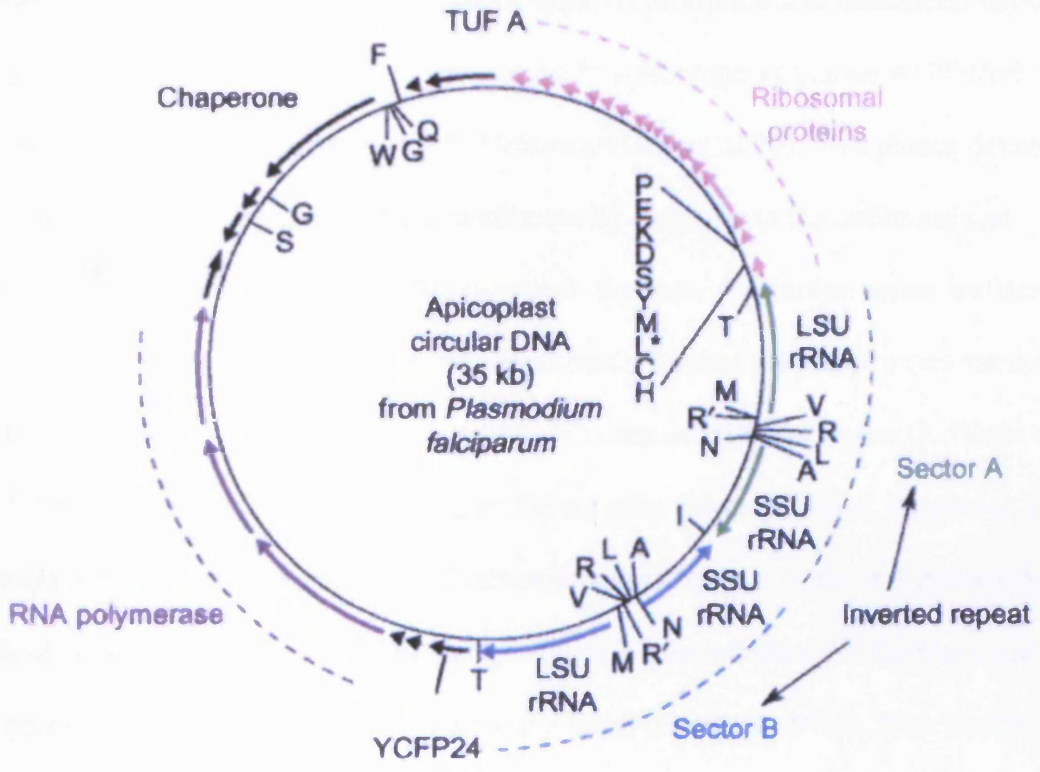


Figure 1.10 Organisation of the apicoplast genome

Taken from Maréchal and Cesbron-Delauw, (2001).

Only two proteins not involved in DNA replication, transcription and translation have been identified on the apicoplast genome (Ellis *et al.*, 2001; Wilson *et al.*, 1996); a protein related to *SufB*, which is involved in the production of Fe-S clusters in *E. coli* and *ClpC*, a chaperone/protease involved in the processing of nuclear-encoded

proteins targeted to plastids. The presence of these proteins suggested that, in Apicomplexa, nuclear-encoded proteins may be targeted to the apicoplast and that this organelle may be the location of at least one essential metabolic pathway.

1.7 Apicoplast function

Antibiotic drugs that disrupt genome replication, transcription and translation within the plastid trigger a 'delayed-death' phenotype in Apicomplexa parasites (Pfefferkorn and Borotz, 1994; Fichera *et al.*, 1995; Fichera and Roos, 1997). This phrase denotes that initially parasite cultures appear unaffected by exposure to the antibiotic and continue to grow at the expected rate. However, daughter merozoites arrest and die within new host cells, possibly due to an incorrectly formed parasitophorous vacuole. Furthermore, interruption of apicoplast protein synthesis by thiostrepton (Sullivan *et al.*, 2000), showed that, when mosquitoes fed on mice infected with *P. berghei* at 6% parasitaemia, thiostrepton reduced the development of oocysts in the mosquitoes to 10% of controls. This suggests that the apicoplast is also essential for development of the parasite during the sexual stage within the mosquito. Collectively, these studies highlight that, whatever the ultimate functions of the plastid, the apicoplast is essential for the maintenance of parasite infections.

Initial hypotheses of apicoplast function were based on the metabolic pathways observed in non-photosynthetic plastids present in some plants. These included the production of haem required for respiration, fatty acid biosynthesis, starch storage and the synthesis of aromatic amino acids. However, preliminary data obtained from the *Plasmodium falciparum* genome project, provided the first hard evidence of

apicoplast function. The research groups of Beck and McFadden utilised the genetic information to identify several nuclear-encoded proteins that contained a plastid target peptide motif, which was sufficient for transport to the apicoplast (Waller *et al.*, 2000; Jomaa *et al.*, 1999). These proteins belonged to two anabolic pathways, fatty acid biosynthesis and non-mevalonate isopentenyl diphosphate synthesis, found in bacteria, algae and plants. All of these nuclear-encoded apicoplast-targeted (NEAT) proteins contained a bipartite leader sequence at their N-terminus, consisting of a hydrophobic signal peptide (16-34 amino acids) terminating in a classical 'von Heijne' cleavage site followed by a typical plastidial transit peptide (30-44 amino acids) with a net positive charge (Waller *et al.*, 2000) (Figure 1.11). This distinct arrangement of amino acids is essential for the correct targeting and transport of proteins bound for the apicoplast (Waller *et al.*, 2000). Two programs have been developed that use known apicoplast targeted leader sequences to predict the likelihood of any individual *Plasmodium* protein being targeted to the apicoplast (Zuegge *et al.*, 2001; Foth *et al.*, 2003). In conjunction with the sequencing of the *P. falciparum* genome, this has allowed the identification of approximately 550 candidate apicoplast proteins (Ralph *et al.*, 2004).

Bioinformatics and the prevalence of pathways previously characterised in *E. coli* and plants have accelerated advances in the identification of whole and partial metabolic pathways within the apicoplast. In *P. falciparum*, proteins that constitute complete pathways for the synthesis of fatty acids, non-mevalonate isopentenyl diphosphate synthesis and a partial pathway for the synthesis of haem (Sato *et al.*, 2000; Sato and Wilson, 2002), are targeted to the apicoplast. The other enzymes in the haem

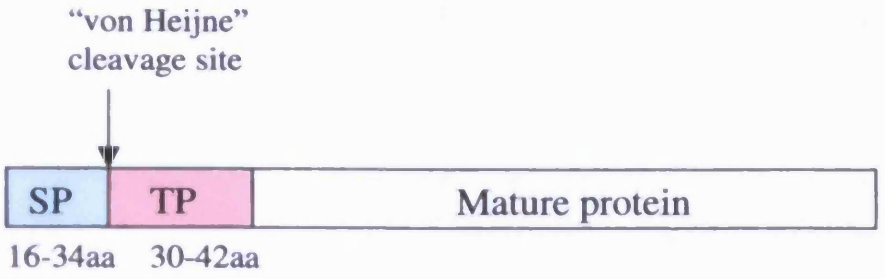


Figure 1.11 The bipartite signal sequence required for apicoplast targeting

An example of a nuclear-encoded apicoplast-targeted (NEAT) protein and its bipartite signal sequence that is both essential and sufficient for targeting to the apicoplast (Waller *et al.*, 1998, 2000). SP = signal peptide, necessary for entry of the protein into the secretory pathway (Waller *et al.*, 2000). Signal peptides are usually hydrophobic domains that end in a signal peptide cleavage site. TP = transit peptide, is required for subsequent targeting to the plastid and to avoid secretion into the parasitophorous vacuole (Waller *et al.*, 2000). Transit peptides of apicoplasts proteins have a net positive charge and are rich in asparagine and lysine and depleted in glutamate and aspartate (Foth *et al.*, 2003).

synthesis pathway are found in the cytosol and the mitochondria (Bonday *et al.*, 2000; Varadharajan *et al.*, 2002), which is closely-associated to the apicoplast during some stages of the parasite lifecycle (Figures 1.7 and 1.8). This may allow for the efficient exchange of substrates between the two organelles and the cytosol. All of these pathways include steps that differ significantly from those found in the human host and provide many potential targets for the development of specific inhibitors. The inhibition of these novel targets with existing antibacterial and herbicide compounds

has already provided one drug, fosmidomycin, an inhibitor of IspC, an enzyme in the non-mevalonate isopentenyl diphosphate pathway that has progressed to human trials (Lell *et al.*, 2003).

1.8 Fatty Acid Synthesis

Whilst most species derive some fatty acids from their diet, in many species *de novo* fatty acid synthesis (FAS) is essential. The end products of these pathways are integral components of cellular membranes with both structural and signalling functions and fatty acyl lipids are also important energy storage molecules. The reaction mechanisms of fatty acid biosynthesis are very similar in all organisms, taking place in either the cytosol of animals, yeast and bacteria or the plastids of plants and alga.

There are two types of FAS and these are distinguished by their structural organisation and catalytic mechanisms. Type I FAS of animals and yeast is catalysed by multifunctional proteins termed fatty acid synthase. In animals FAS consists of two identical, 272kDa polypeptides, with only the homodimer exhibiting activity. Each chain contains six catalytic and an acyl carrier protein (ACP) domain, arranged in the order; β-ketoacyl-ACP synthase, malonyl/acetyl transacylase, dehydratase, enoyl reductase, β-ketoacyl reductase, ACP and thioesterase. In contrast, the Type II FAS system present in plants, algae and bacteria contains many of the catalytic domains of Type I FAS as discrete enzymes that probably form a multi-enzyme complex. Some inhibitors of FAS are either Type I or Type II specific and can be exploited as antibiotics.

The first committed step in the Type II fatty acid synthesis pathway is the conversion of acetyl-CoA (a two carbon unit) to malonyl-CoA (a three carbon unit) catalysed by acetyl-CoA carboxylase (ACCase) (Figure 1.12). The malonyl group is then transferred between the phosphopantathiene prosthetic groups of coenzyme A and acyl carrier protein (ACP) by the malonyl-CoA:ACP transacylase (FabD - Fatty Acid Biosynthesis) to form malonyl-ACP. Acyl carrier protein is a small (~8kDa), acidic protein that remains bound to all of the intermediates of the FAS pathway and distinguishes such compounds from the intermediates of fatty acid oxidation, which remain bound to coenzyme A.

Binding sites of ACP are present in at least two of the fatty acid synthesis enzymes in *E. coli* (Zhang *et al.*, 2001& 2003). These sites are found on the surface of the protein and contain a number of highly conserved positively charged residues (Zhang *et al.*, 2001 & 2003). Disruption of the ACP recognition patch on these enzymes, by replacement of positive residues with negatively charged amino acids, caused a large decrease in the enzyme activity (Zhang *et al.*, 2003). This suggested that the docking of ACP in FAS enzymes is essential for the correct functioning of the pathway.

The initiating step of the FAS cycle is the condensation of malonyl-ACP with acetyl-CoA to form a four carbon β-ketoacyl-ACP group (acetoacetyl-ACP) catalysed by FabH – a β-ketoacyl-ACP synthase (Figure 1.12). The subsequent removal of the keto group to form a saturated hydrocarbon chain is achieved by a common mechanism found in nature. Firstly, the keto group is reduced by a β-ketoacyl-ACP reductase, then water is removed by a dehydratase and the resulting double bond of the enoyl group is reduced by an enoyl-ACP reductase (Figure 1.13). The product of

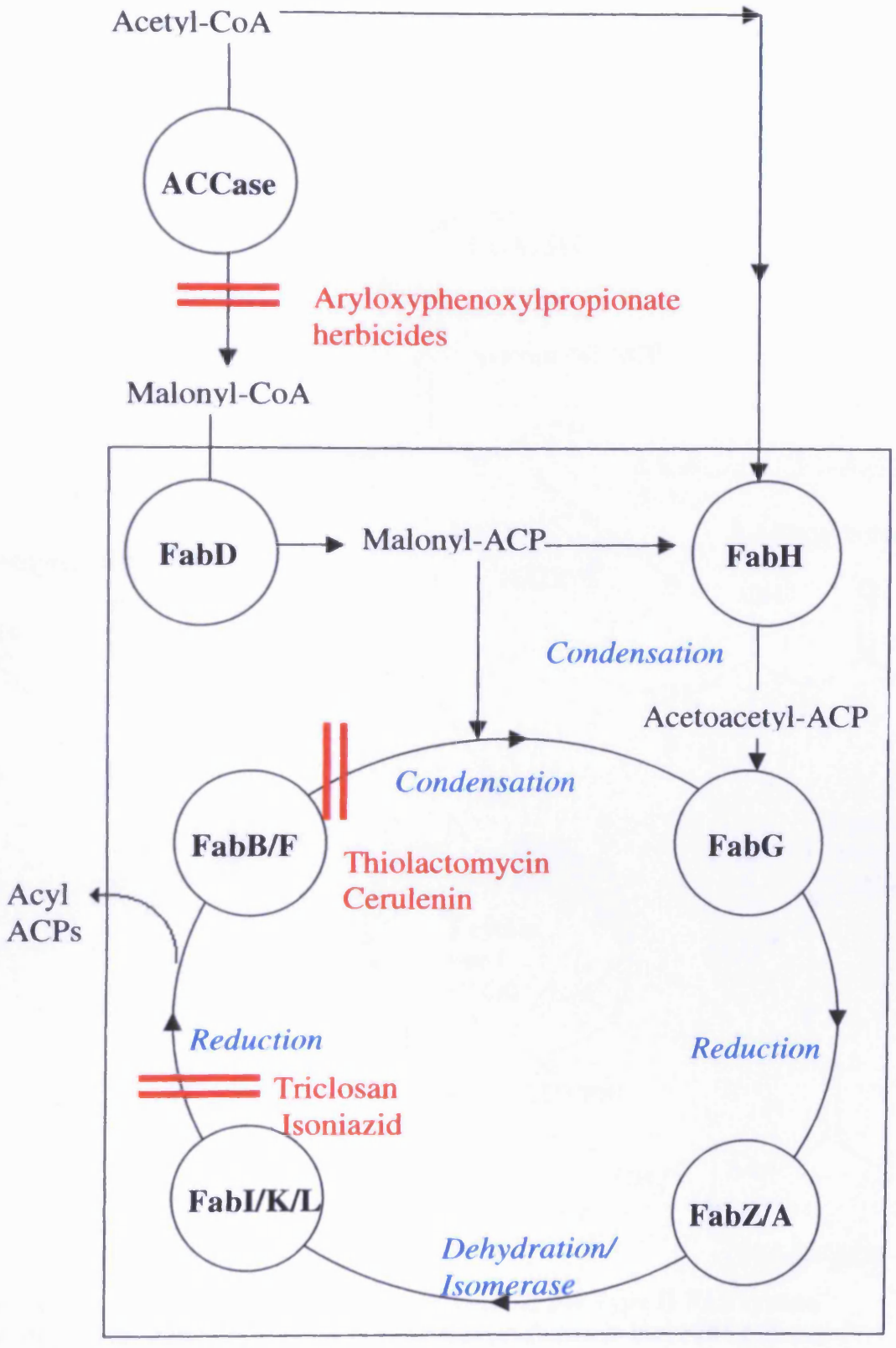


Figure 1.12 The Type II Fatty Acid Synthesis pathway
ACCase = Acetyl-CoA carboxylase. Inhibitors are shown in red and enzymes shown in bold. FabH catalyses the first condensation step between acetyl-CoA and malonyl-ACP. The FAS pathway carries out several cycles of condensation, reduction, dehydration and a second reduction to produce acyl-ACPs.

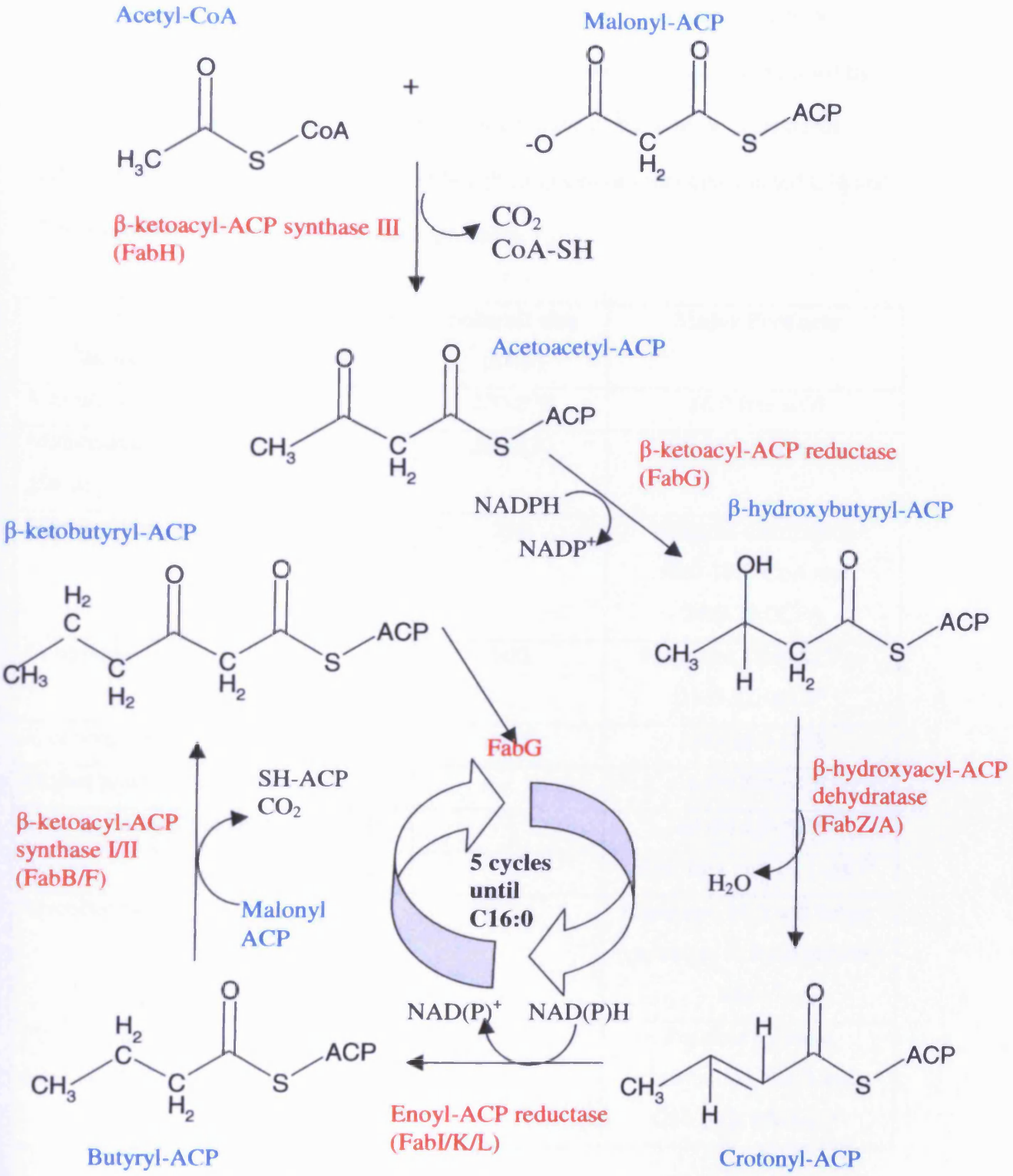


Figure 1.13 The individual enzymatic reactions of the Type II FAS system
Enzymes are shown in red, substrates and products shown in blue. The pathway proceeds through cycles of condensation, reduction, dehydration and a 2nd reduction. FabI/K/L are individual enoyl reductases that are all capable of catalysing the reduction step. They appear to be mutually exclusive as no organism has yet been found to contain more than one of the enoyl reductases.

the first cycle, butyryl-ACP, becomes the substrate for the subsequent rounds of condensation, reduction, dehydration and reduction, as the acyl chain is extended by two carbons per cycle (Figure 1.12). The final products of the fatty acid synthesis differ between organisms (Table 1.2), although saturated or monounsaturated C16 and C18 acyl chains are generally the major products.

Organism (tissue in brackets)	**Type**	**Subunit size (kDa)**	**Major Products**
Mammalian or avian (liver)	I	220-270	16:0 free acid
Mammalian (mammary gland)	I	220-270	4:0-16:0 free acids
Mycobacterium	I	290	Bimodal distribution 16:0-18:0-CoA and 24:0-26:0CoA
C. parvum	I	900	Elongates 16:0-ACP to 18:0-22:-ACP[a]
S. cerevisiae	I	180	16:0-18:0-CoA
Higher plant chloroplasts	II	-	16:0-18:0-ACP
Cuphea hookeriana seeds	II	-	10:0-14:0-ACP
E. coli	II	-	16:0, 16:1 and 18:1-ACP
Mycobacterium	II	-	Elongates 16:0 and longer precursors to form mycolic acids [b]
Plant and mammalian mitochondria	II	-	Possible bimodal distribution. C8:0 and C14:0-C16:0-ACP[c]

Table 1.2 The distribution of different types of FAS and their major products
Adapted, in part, from Lipid Biochemistry, 5th Edition, 2002, Blackwell Science.
[a] Zhu *et al.*, 2004. [b] Kremer *et al.*, 2000. [c] Yasuno *et al.*, 2004.

1.8.1 Determination of chain length

In Type I FAS, the growing acyl chain is attached to the flexible ACP domain, which allows it to interact with catalytic domains between and within each polypeptide subunit. In this system, the chain-length specificity of the thioesterase domain contributes significantly to the length of the final product. For example, in animals there is a twenty-fold decrease in thioesterase activity towards C14 compared to C16 (Naggert *et al.*, 1991) and this ensures that, despite the broad range of chain lengths accepted by most of the active domains, the main product of the pathway is palmitate. In contrast, the final products of Type II FAS appear to be more tightly controlled at the entry point of the cycle, by the β-ketoacyl-ACP synthase enzymes and can lead to the synthesis of branched chain fatty acids.

In bacteria and plants a second important regulator of fatty acid length is glycerol 3-phosphate acyltransferase (PlsB), an important enzyme in the synthesis of phospholipids. When *E. coli plsB* mutants are starved of glycerol, unusually long (20-22) carbon acyl chains accumulate, suggesting that under these conditions the condensing enzymes are capable of elongating further than C18. Therefore, the competition between PlsB and elongating condensing enzymes is important to the chain length of the acyl product.

The chain length of the products of Type II FAS in plants can be affected by thioesterase and acyl-CoA synthases, as well as glycerol 3-phosphate transacylase. As an example, a medium chain-length specific acyl-ACP thioesterase is responsible for the high proportion (>80%) of C10:0 (decanoate) and C12:0 (laurate) present in the triacylglycerol energy stores in *Cuphea lanceolata* seeds (Dörmann *et al.*, 1993).

Acyl-CoA transferases can activate free fatty acids, presumably following thioesterase activity, for essential reactions, such as the introduction of double bonds in eukaryotic systems by acyl-CoA desaturases.

1.8.2 Synthesis of unsaturated fatty acids

1.8.2.1 Bacterial

The ability to synthesise unsaturated fatty acids (UFA) in bacteria has been extensively studied in the model bacterium *E. coli* and the findings are discussed in detail in this section. Alternative bacterial UFA pathways are also described.

1.8.2.2 The FabB/FabA system

The gamma Proteobacteria (i.e. *E. coli* and its relatives) are capable of synthesising monounsaturated fatty acids (MUFAs) and possess two proteins that are not present in other bacteria families. These MUFA synthesising bacteria, contain a β-ketoacyl-ACP synthase (FabB) and a dehydratase (FabA) with isomerase activity that are essential for the incorporation of a double bond into the growing acyl chain (Figure 1.14).

The first step of UFA synthesis is performed by FabA, which like FabZ, is capable of catalysing the dehydration of β-hydroxyacyl-ACPs to produce the double bond containing enoyl-ACP. However, at the 10 carbon stage, FabA also catalyses the isomerisation of the *trans*-2 double bond to *cis*-3 (Heath and Rock, 1996a). In the saturated pathway, the natural route of the product following the dehydratase step, is the removal of the double bond by FabI, the enoyl-ACP reductase. However, in UFA synthesis, FabB shows a high specificity for the *cis*-3 decenoyl-ACP product and

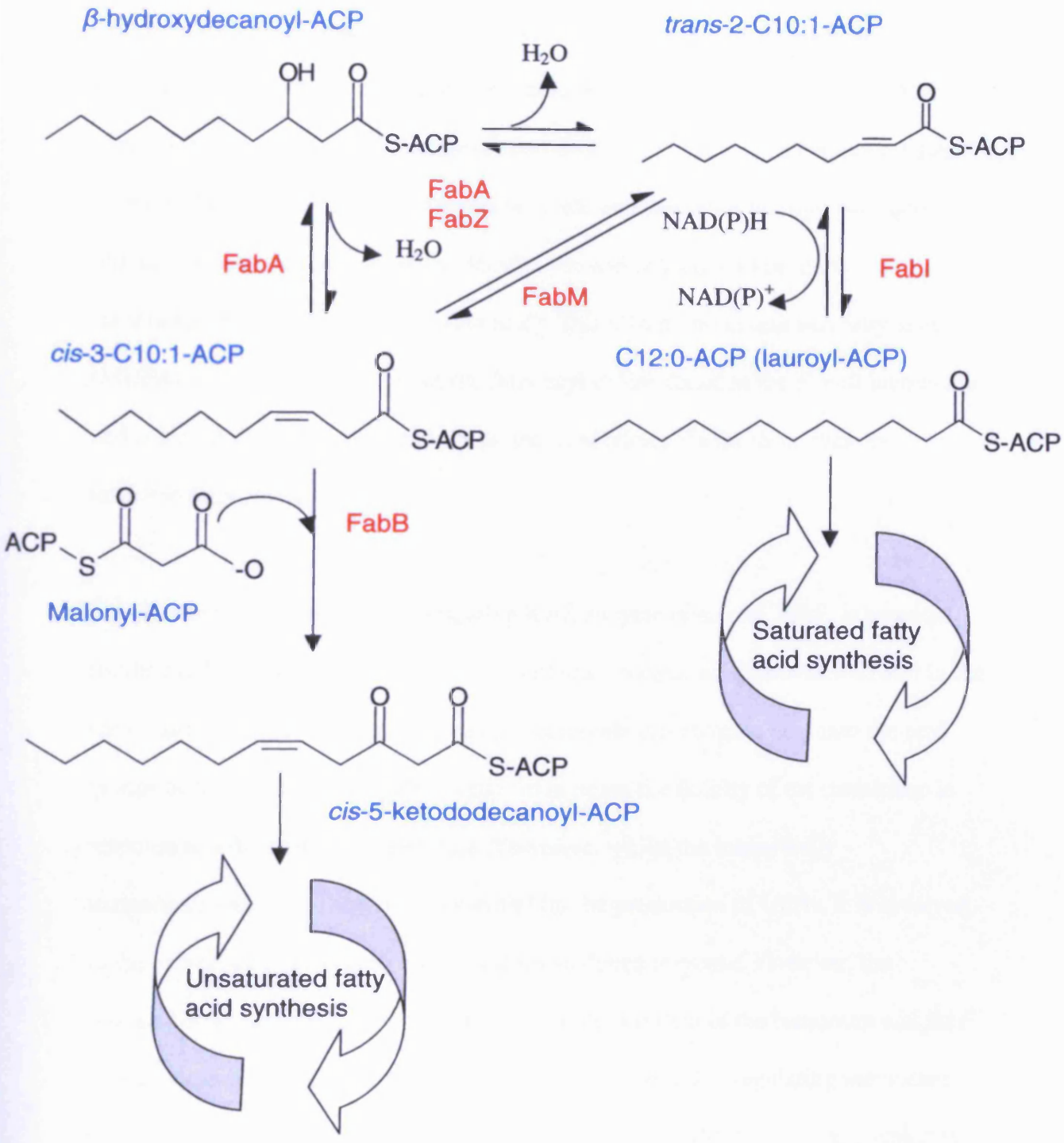


Figure 1.14 Unsaturated fatty acid synthesis in bacteria

Insertion of the double bond in unsaturated fatty synthesis is depicted on the left-hand scheme and the saturated pathway represented by the right-hand scheme. Enzymes are shown in red text; substrates/products are shown in blue text.

β-hydroxydecanoyl-ACP is shown at a branch point where it can continue along the saturated fatty acid synthesis pathway or if acted on by either isomerase (FabA/FabM) becomes a substrate for the unsaturated fatty acid synthesis pathway.

competes with FabI, which can only reduce *trans* double bonds. FabB is the only ketoacyl-ACP synthase that catalyses the condensation of the *cis*-3 decenoyl-ACP compound with malonyl-ACP to form *cis*-5 dodecanoyl-ACP. This compound then re-enters the "generic" fatty acid synthesis cycle and continues through two more rounds of condensation, reduction, dehydration and reduction to produce *cis*-9 hexadecenoyl-ACP (palmitoleic acid). This C16 monounsaturated fatty acid (MUFA) is a major component of the fatty acyl chains found in the *E. coli* membrane and is incorporated into bacterial lipids and membranes via the same route as saturated fatty acids.

The discovery that the second elongating KAS enzyme of *E. coli*, FabF, is essential for the synthesis of *cis*-11 octadecenoic acid (*cis*-vaccenic acid) provides a twist in the tail of bacterial UFA synthesis. Bacteria incorporate *cis*-vaccenic acid into the acyl groups of their membrane lipids in order to increase the fluidity of the membrane in response to a decrease in temperature. Therefore, whilst the intrinsically temperature-sensitive FabF is not essential for the production of UFAs, it is involved in the unsaturated pathway in a temperature mediated response. However, the production of vaccenic acid is not essential for the survival of the bacterium and *fabF* knockouts show no phenotype. In *E. coli*, a second method of regulating membrane fluidity acts by controlling the level of expression of the *fabA/fabB* operon, which is mediated by the transcriptional activator FadR and transcriptional repressor FabR (Hoang and Schweizer, 1997; Campbell *et al.*, 2001a; Zhang *et al.*, 2002).

A second dehydratase, called FabZ, cannot catalyse isomerisation of the C10 intermediate but is able to catalyse the dehydration of all other intermediates in

Type II FAS. Therefore, in *E. coli* and related bacteria, the combination of FabA and FabB activities is essential for the synthesis of UFAs and organisms containing a FabB homologue also contain a FabA enzyme.

1.8.2.3 FabM isomerase

The FabA/FabB system for the production of UFA is not widely distributed in bacteria (Campbell and Cronan, 2001b). Nevertheless, some gram-positive bacteria such as streptococci, do not contain FabA and FabB, yet synthesise UFA. In *Streptococcus pneumoniae*, the FabZ enzyme performs all of the dehydration steps of FAS. An enzyme unique to this family of bacteria, FabM, diverts fatty acids into the UFA pathway by isomerising *trans*-2 decenoyl-ACP to *cis*-3 decenoyl-ACP (Marrakchi *et al.*, 2002). Within this system FabF catalyses the elongation of the *cis*-3 decenoyl-ACP and all other unsaturated and saturated acyl-ACPs.

1.8.2.4 Desaturation of existing Fatty Acids

Aerobic desaturation of existing fatty acids on membrane phospholipids provides an alternative route to UFA synthesis in bacilli and cyanobacteria (Cronan and Campbell, 2001). These enzymes function outside of the Type II FAS pathway and in this way share similarities to UFA synthesis observed in eukaryotic organisms. The de Mendoza laboratory has shown that, *Bacillus subtilis*, contains a Δ5 desaturase (Altabe *et al.*, 2003). This enzyme can be induced following a sharp decrease in temperature and is thought to allow the bacterium to remodel its membrane lipids rapidly in a response to abrupt alterations in temperature (Cybulski *et al.*, 2002).

1.8.3 Eukaryotic UFA

So far, in this Section on UFA synthesis, focus has been directed towards bacterial UFA synthesis, which mainly occurs via an anaerobic pathway that is harboured only by a small group/s of bacteria. The most common mechanism of UFA synthesis is the aerobic pathway (see Section 1.8.2.4), which utilises molecular oxygen as an essential cofactor. This is by far the most widespread pathway and is present in bacilli (Altabe *et al.*, 2003), cyanobacteria, yeast, algae, higher plants, protozoa and animals. The aerobic and anaerobic systems appear to be mutually exclusive as no organism has been found to have both.

The desaturase enzymes that catalyse the insertion of double bonds are invariably membrane bound, and hence, characterisation of the process was initially slow. However, it is known that most of the double bonds are inserted at the Δ9 position, with some exceptions being Δ7 bonds in algae and the Δ5-desaturase in

Bacillus subtilis. Bloch and co-workers also found that molecular oxygen and NADH or NADPH were essential cofactors, which suggested a mechanism similar to many oxygenases. Most Δ9-desaturases readily accept saturated substrates from C14-C18 and are not entirely specific for stearate (C18:0) as first thought (Gibson *et al.*, 1993).

The animal Δ9-desaturase is found in the endoplasmic reticulum (ER), with the desaturase mainly enveloped within the membrane and only its active site exposed to the ER lumen. Two other membrane bound proteins, cytochrome b_5 and a NADH/cytochrome b_5 reductase, contribute to the reaction by generating the reducing potential required for the reaction. In contrast, the plant stearoyl desaturase is a soluble protein, present in the chloroplast stroma (the location of Type II FAS) and utilises an ACP bound substrate. These enzymes provide a mechanism for the production of monounsaturated fatty acids that are essential to all organisms. They also provide a stepping stone for the synthesis of polyunsaturated fatty acids (PUFAs), the production of which requires many subsequent enzyme catalysed reactions and varies considerably between different species (for a Review see Lipid Biochemistry, 2002, Ed. Gurr, Harwood and Frayn, Blackwell Science).

1.9 Apicomplexan Fatty Acid Synthesis

Until recently, it was suggested that *Plasmodium* spp. lacked the ability to carry out *de novo* fatty acid synthesis (FAS) (Vial and Ancelin, 1992). This was supported in studies with radiolabelled fatty acids and fatty acid precursors, which appeared to show that the parasite obtained its fatty acids by scavenging fatty acids from the host (Rock, 1971). However, plants and algae both perform *de novo* fatty acid synthesis in

their plastids, utilising dissociated enzymes grouped together in a multi-enzyme complex, termed Type II FAS.

Type II FAS is widespread in prokaryotes but in eukaryotes is limited to plastid containing organisms. Therefore, the discovery of Type II FAS-like proteins in the apicoplast and the ability of *P. falciparum* to incorporate acetate into fatty acids (Surolia and Surolia, 2001) suggested that Apicomplexa may contain a previously over-looked plastid-based FAS pathway. The presence of the apicoplast as the apparent site of Type II FAS highlights a novel target against the *Apicomplexa* phylum including *Plasmodium spp.*, *T. gondii* and *E. tenella* (Gleeson, 2000).

1.9.1 Initial discovery of FAS in Apicomplexans

Three genes that encoded proteins involved in Type II FAS were discovered 6 years ago in *P. falciparum* and *T. gondii* genomes (Waller *et al.*, 1998). These fatty acid biosynthesis proteins were homologous to proteins of the Type II FAS enzymes, acyl carrier protein (ACP), β-ketoacyl-ACP synthase III (FabH) and β-hydroxyacyl-ACP dehydratase (FabZ) and these enzymes are targeted to the apicoplast (Waller, 2000). The completion of the *P. falciparum* genome has facilitated the bioinformatic identification of all the proteins needed to reconstitute a complete Type II FAS pathway in apicomplexan parasites (Figure 1.15) (For a recent review see Ralph, *et al.*, 2004).

During the course of this project, the characterisation of most of the *P. falciparum* Type II FAS proteins was achieved by several research groups (Waters *et al,*. 2002;

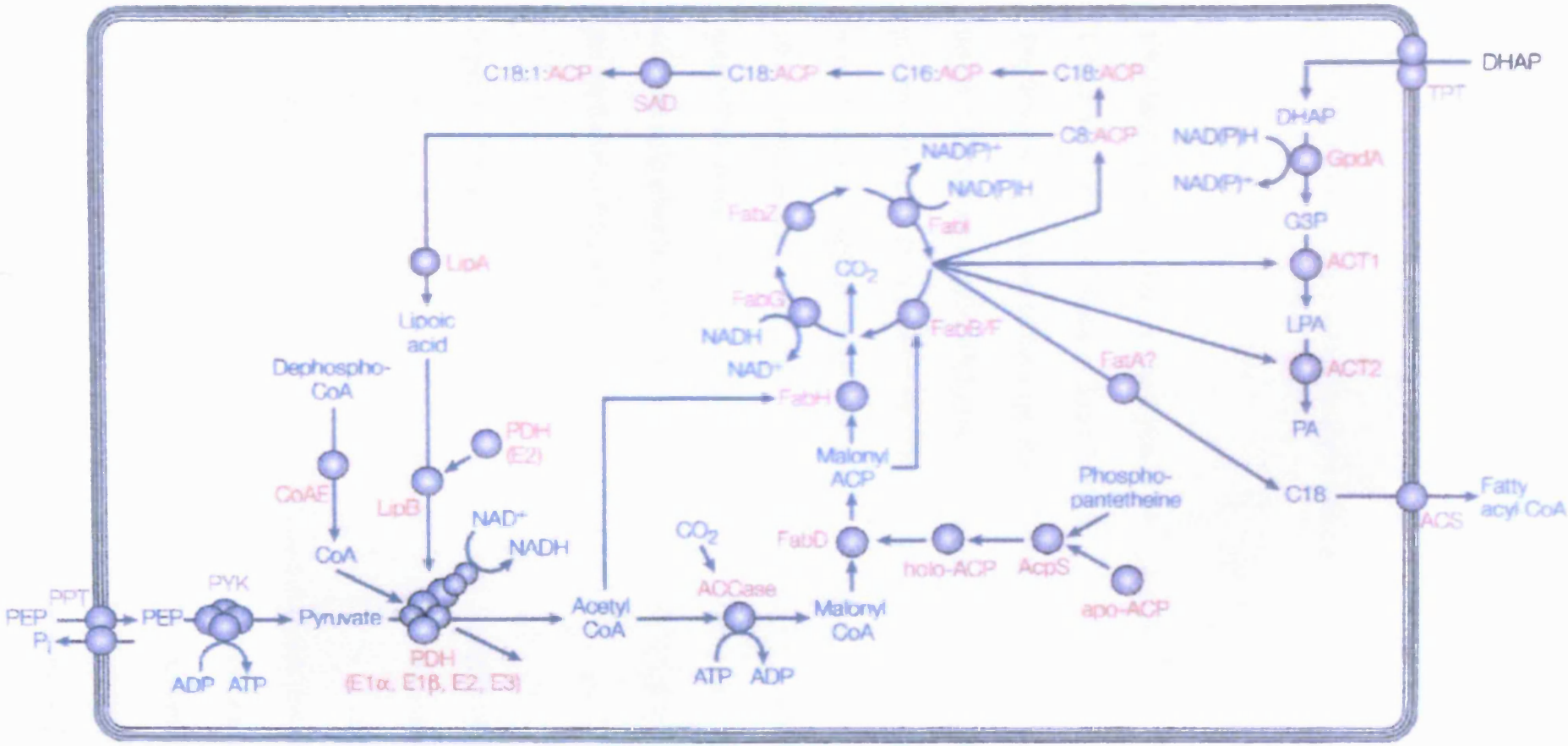


Figure 1.15 The apicoplast FAS pathway and contributing reactions

Enzyme names are shown in red, substrates and products are shown in blue. ACCase, acetyl-CoA carboxylase; acetyl CoA, acetyl coenzyme A; ACP, acyl carrier protein; ACS, acyl CoA synthetase; ACT1, glycerol-3-phosphate acyltransferase; ACT2, 1-acyl-glycerol-3-phosphate acyltransferase; ADP, adenosine diphosphate; ATP, adenosine triphosphate; FabB/F, β-ketoacyl-ACP synthase I/II; FabD, malonyl-CoA transacylase; FabH, β-ketoacyl-ACP synthase III; FabI enoyl-ACP reductase; FabZ, β-hydroxyacyl-ACP dehydratase; FatA, Acyl-ACP thioesterase; GpdA, glycerol-3-phospahte dehydrogenase; LipA, lipoic acid synthase; LipB, lipoate protein ligase; LPA, lysophosphatidic acid; NAD^+/NADH, nicotinamide adenosine; PA, phosphatidic acid; PDH, pyruvate dehydrogenase; PDH (E2), pyruvate dehydrogenase complex E2 subunit; PEP, phosphoenol pyruvate; Pi, inorganic phosphate; PP, pyrophosphate; PPT, phosphoenolpyruvate/phosphate translocator; PYK, pyruvate kinase; SAD, stearoyl-ACP desaturase; TPT, triose phosphate transporter. Adapted from Ralph *et al.*, 2004.

Prigge *et al.*, 2003; Pillai *et al.*, 2003; Surolia and Surolia 2001; Suguna *et al.*, 2003; Sharma *et al.*, 2003). To date, the only enzyme that has not been characterised is the FabB/F elongating condensing enzyme (β-ketoacyl-ACP synthase) and this provides a focus for the work described in Chapters 3 and 4.

1.9.2 Individual enzymes of the Type II FAS pathway

1.9.2.1 The initial reactions of Type II FAS

The first committed step of the Type II FAS pathway is the production of malonyl-CoA from acetyl-CoA by the large acetyl-CoA carboxylase (ACCase) protein (Figure 1.12). A nuclear-encoded apicoplast-targeted (NEAT) ACCase has been identified in both *P. falciparum* and *T. gondii* (Gardner *et al.*, 2002; Zuther *et al.*, 1999). However, the production of malonyl-CoA depends on the availability of acetyl-CoA as well as the presence of the catalytic enzyme. This begs the question: does the apicoplast have the ability to synthesise or import sufficient acetyl-CoA for the synthesis of fatty acids?

In plants, the main source of acetyl-CoA for plastidial FAS is pyruvate (Ke *et al.*, 2000; Bao, *et al.*, 2000), which is transported across the chloroplast membranes as phosphoenolpyruvate and converted to pyruvate within the plastid or formed by glycolysis *in situ*. Bioinformatic analyses performed within this project and by other groups, identified a putative phosphoenolpyruvate translocator and pyruvate kinase, which are predicted to be targeted to the apicoplast (J.Urch unpublished observation; Ralph *et al.*, 2004). These two proteins would provide a source of pyruvate to the plastid and, additionally, the kinase reaction would generate ATP. In plant plastids the

conversion of pyruvate to acetyl-CoA is catalysed by the multi-subunit pyruvate dehydrogenase enzyme. Genes encoding each of the subunit domains are present in the genomes of *P. falciparum* and *T. gondii* and the proteins they encode are located in the plastid (Ralph *et al.*, 2004). Therefore, the apicoplast appears to possess all of the substrates and enzymes required for the synthesis of acetyl-CoA and malonyl-CoA, the starting materials required for FAS.

ACP and the initiating condensing enzyme FabH were the first to be identified in apicomplexans because the *acp*P and *fab*H genes are located on Chromosome 2, the first *P. falciparum* chromosome to be sequenced in its entirety (Gardner *et al.*, 1998; Waller *et al.*, 1998). Both proteins were also identified due to their significant sequence homology to bacterial and plant proteins (for ACP 49% identity with *E. coli* ACP and 35% identity with *Spinacea oleracea* plastidial ACP; for FabH 38% identity with *E. coli* FabH and 43% identity with *Spinacea oleracea* FabH). ACP and FabH are both essential for the first condensation reaction of Type II FAS. However, prior to elongation, malonyl-CoA must first be transferred to ACP by the malonyl-CoA:ACP transacylase (FabD). The Prigge laboratory, recently identified, cloned and expressed *P. falciparum* FabD in a recombinant form (Waters *et al.*, 2002). In the aforementioned study, pfFabD was able to transfer efficiently the malonyl moiety from malonyl-CoA to either ecACP (*E. coli* ACP) or pfACP (*P. falciparum* ACP), to generate malonyl-ACP for the FabH reaction.

The pfFabH enzyme produced acetoacetyl-ACP from acetyl-CoA and malonyl-ACP and showed a slight preference for malonyl-pfACP compared to malonyl-ecACP (Prigge *et al.*, 2003). The low specificity towards ecACP or pfACP was not surprising

because the two proteins show 49% sequence identity. Furthermore, 6 of the 7 residues in helix α2 that are thought to be involved in the docking of ACP to enzymes involved in FAS (Zhang *et al.*, 2001), are conserved between the bacterial and malarial proteins.

1.9.2.2 The ketoacyl-ACP reductase

The next step of the FAS pathway is catalysed by a β-ketoacyl-ACP reductase, the only Type II FAS enzyme for which only a single homologue (FabG) has been detected. The β-ketoacyl-ACP reductase of *P. falciparum* was identified by its high sequence identity to both bacterial and plant FabG proteins (47% sequence identity to *E. coli* FabG and 48% to *Brassica napus* FabG) (Pillai *et al.*, 2003). This enzyme contains a long N-terminal extension of approximately 57 amino acids that is predicted to target the enzyme to the apicoplast. pfFabG also possesses the Ser, Tyr Lys residues involved in the catalytic triad of short chain alcohol dehydrogenases. Like all previously characterised FabG proteins, the *P. falciparum* is NADPH dependent and is unable to utilise NADH unlike the enoyl-ACP reductase (FabI – discussed later in this section). However, the structure and fold of the FabG and FabI reductases are similar, as observed previously in *Brassica napus*, both containing a typical NAD(P)-binding Rossman fold.

1.9.2.3 The dehydratase of Apicomplexan Type II FAS

The discovery of the enzyme responsible for the dehydration step was first reported in 2003 (Sharma *et al.*, 2003). This protein showed significant sequence homology to

bacterial and especially plant FabZ enzymes (41% identity to *E. coli* FabZ and 46% identity with *Arabadopsis thalania* FabZ) but less so with the *E. coli* FabA dehydratase/decenoyl isomerase (21% identity). On the basis of this phylogenetic analysis, the dehydratase was defined as pfFabZ. The dehydratase identified in *T. gondii* also shows significant sequence identity to FabZ, rather than FabA enzymes, suggesting that apicomplexan parasites may not contain the FabA/FabB system for the synthesis of unsaturated fatty acids. No FabM homologue has been identified within the complete *P. falciparum* genome. Nevertheless, the possibility that the *P. falciparum* FabZ can catalyse the isomerisation of the *trans*-2 enoyl product of the dehydration reaction cannot be eliminated because the kinetic characterisation was performed with four carbon, not 10 carbon, intermediates (Sharma *et al.*, 2003). In fact, although no conversion of the *trans* product to the *cis* product was observed, pfFabZ showed favourable kinetics towards the use of the enoyl product compared to the β-hydroxyacyl substrate, suggesting that the enoyl group could be a natural substrate for this enzyme.

1.9.2.4 Enoyl reductases of Apicomplexans

The final reaction of the Type II FAS pathway is catalysed by an enoyl-ACP reductase, of which several homologous proteins have been identified. FabI, enoyl-ACP reductase I, is the most widely distributed of these proteins and is present in plants and most bacteria. FabK, enoyl-ACP reductase II, is restricted to the Gram-positive bacteria and was first discovered in *Streptococcus pneumoniae* due to its lack of sensitivity to triclosan, an irreversible inhibitor of the FabI enzyme (Marrakchi *et al.*, 2001). A third homologue, FabL, was identified in *Bacillus subtilis*

(Heath and Rock, 2000) and this enzyme exhibits some sequence homology to FabI but not enough for it to be characterised as a FabI protein. FabL enzymes are not widely distributed in the bacteria.

A search of the *P. falciparum* genome with sequences of each of the three enoyl-ACP reductases identified a FabI homologue but no homologues of FabK or FabL (Surolia and Surolia, 2001). *P. falciparum* FabI is a 35kDa protein that contains a 77 amino acid N-terminal bipartite apicoplast targeting domain, reinforcing the idea that Type II FAS takes place in the apicoplast. Like the FabI proteins from plants and bacteria, pfFabI is sensitive to triclosan, a chlorinated *bis*-phenol regularly used as a bactericidal agent in toothpaste and detergents (Suguna *et al.*, 2001; Bhargava and Leonard, 1996). Triclosan was the first compound to demonstrate definitively that the inhibition of Type II FAS in Plasmodium was sufficient to clear blood stage malaria parasites from mice (Surolia and Surolia, 2001). The basis of triclosan inhibition of FabI is subsequently discussed in further detail in Section 1.11.2.b.

1.9.3 The function of Type II FAS in apicomplexan parasites

The fate of acyl-ACPs produced from the Type II FAS pathway is unknown. However, there is a high possibility that some octanoyl-ACP is converted to lipoic acid, an essential cofactor of the pyruvate dehydrogenase complex (PDHC) that aids in the transfer of an acetyl group to coenzyme A. The synthesis of lipoic acid from octanoyl-ACP occurs in both the mitochondria and plastids of plants (Yasuno and Wada, 2002). Firstly, octanoyl-ACP is converted to lipoyl-ACP by lipoic acid synthase (LipA) and then lipoic acid protein-ligase (LipB) catalyses the formation of a

covalent bond to a serine residue of PDHC. NEAT homologues of LipA and LipB proteins have been predicted in *P. falciparum* and *T. gondii* (Wrenger and Muller, 2004; Thomsen-Zieger *et al.*, 2003). In summary, a cyclic pathway incorporating TypeII FAS and lipoic acid synthesis may exist. Lipoic acid is an essential cofactor for the production of acetyl-CoA; one of the starting substrates required for the synthesis of acyl-ACPs and these acyl-ACPs may be utilised in the synthesis of lipoic acid in the apicoplast.

Fatty acids are also incorporated into phospholipids by esterification of glycerol hydroxyl groups. Phospholipids can be exported and assimilated into other cellular membranes or in this case may be important in maintaining the apicoplast membranes. Two genes encoding glycerol 3-phosphate acyl transferases have been identified in the *P. falciparum* genome. The products of these genes are predicted to be targeted to the apicoplast, suggesting that phosphatidic acid is synthesised in the plastid-like organelle. However, genes encoding the enzymes involved in the latter stages of phospholipid synthesis, for example those involved in phosphatidylcholine synthesis, are present in *P. falciparum* but appear to be targeted to the endoplasmic reticulum (Ancelin *et al.*, 2003). Several isoforms of a membrane protein, acyl-CoA synthetase, an enzyme that combines free fatty acids with coenzyme A and exports them to closely-associated organelles or the cytosol are predicted to be NEAT proteins. This would allow acyl chains produced by Type II FAS in the apicoplast to be utilised throughout the malaria parasite cell.

Finally, a NEAT protein that shows significant sequence identity to stearoyl-CoA desaturase, may be capable of utilising the products of Type II FAS to synthesise monounsaturated fatty acids such as palmitoleic acid (16:1) or oleic acid (18:1). This would mimic the eukaryotic method for the production of MUFA and, in the absence of UFA synthesis within the Type II pathway, could provide the only alternative route for the *de novo* synthesis of UFA in *Plasmodium.*

1. 10 Type II FAS Condensing Enzymes

β-Ketoacyl-ACP synthases (KAS) enzymes (Edwards *et al.*, 1997; Magnuson *et al.*, 1995; Jackowski and Rock, 1987) are responsible for the condensation reaction of FAS in plant chloroplasts, algae and bacteria (Lu *et al.*, 2004;Beeson *et al.*, 2001; Rock and Cronan, 1985). They perform the systematic addition of two carbon units to an existing carbon chain. New carbon-carbon bonds are created by carrying out Claisen condensations (Dewar and Dieter, 1988).

1.10.1 Condensing enzyme nomenclature

The nomenclature of Type II KAS enzymes can be somewhat confusing to the uninitiated. In *E. coli*, the model bacterial system, FabB is also designated KAS I, FabF as KAS II and FabH as KAS III. To avoid confusion, the Fab system has been adopted throughout the bacterial kingdom except in *Mycobacterium* species, which possess three KAS enzymes, mtFabH, KasA and KasB. As its name suggests mtFabH is a homologue of other bacterial and plant FabH enzymes and contains the Cys, His, Asn, catalytic triad observed in this family of enzymes. KasA and KasB exhibit greater homology to FabF than FabB enzymes. These enzymes accept long chain (C16-C28) acyl-ACP substrates and the products are used to synthesise mycolic acids,

essential components of the protective cell wall in mycobacteria (Schaeffer *et al.*, 2001; Kremer *et al.*, 2002).

In plants, the genes that encode condensing enzymes were historically designated KAS I, KAS II and KAS III and, therefore, are seldom referred to using the Fab nomenclature. Plant KAS III enzymes show significant homology to bacterial FabH enzymes and both carry out the same the initial condensation reaction of the Type II FAS cycle (Clough *et al.*, 1992; Jackowski *et al.*, 1989; Tsay *et al.*, 1992; Khandekar *et al.*, 2001). Plant KAS I and KAS II enzymes catalyse the chain elongation steps. KAS II appears to be essential for the production of stearate (C18:0) from C16-ACP, whereas KAS I will condense chains up to C14:0 with malonyl-ACP (Harwood, 1996). The Type II FAS pathway in plant plastids does not appear to be capable of synthesising unsaturated fatty acids and, accordingly, both KAS I and KAS II appear to show greater homology to bacterial FabF proteins than to FabB. Plants synthesise UFA downstream of the Type II FAS pathway, using the aerobic system for desaturation (Section 1.8.3).

1.10.2 Classification of KAS enzymes

Enzymes that catalyse the Claisen-condensation include KAS enzymes, alfalfa chalcone synthase and yeast thiolases. Comparison of the amino acid sequences of these enzymes identified a small number of conserved residues and led to the discovery of the active site cysteine (Moche *et al.*, 2001).

Recent analysis revealed that these enzymes can be divided into three structural and functional classes, based on the catalytic mechanism and residues of the active site,

substrate specificity and variations in the acyl carrier group (either CoA or ACP) (Moche *et al.*, 2001). Group 3 consists of the biosynthetic and degradative thiolases. FabB and FabF were placed in group 2. In *E. coli* these two enzymes share 38% identity with each other and limited identity with FabH (14.8% with FabB and 13.2% with FabF), which is placed in Group 1 with its distant relative alfalfa chalcone synthase (Moche *et al.*, 2001).

1.10.3 Substrate specificity of KAS enzymes

FabB, FabF and FabH-like proteins have overlapping substrate specificities. However, FabH catalyses the first condensation reaction between acetyl-CoA and malonyl-ACP, whereas FabB and FabF-like proteins perform the subsequent addition of two carbon units using acyl carrier protein activated substrates (Choi *et al.*, 2000; Harwood, 1996; Heath and Rock, 1996a and 1996b). In plastids, KAS I extends from C_4 to C_{16} in six rounds of elongation whereas KAS II performs extensions on longer fatty acid chains, carrying out an additional step to form C18:0 (Magnuson *et al.*, 1993). *In vitro* studies showed that FabB has the ability to perform the first condensation reaction, elongating from C_2 to C_4, suggesting that FabB is the only essential condensing enzyme in *E. coli* (Alberts *et al.*, 1972; Tsay *et al.*, 1992a). This hypothesis is supported by evidence that attempts to produce *fabB* knockouts in *E. coli* have been unsuccessful, whilst knockout *fabF* mutants can be produced (Rock and Cronan, 1996; Allen and Bartlett, 2000; Kutchma, *et al.*, 1999). Furthermore, Cronan's laboratory recently reported the first discovery of a *fabH* mutant lacking FabH activity (Lai and Cronan, 2003). The results showed that this mutant retained approximately 10% of the wildtype fatty acid

synthetic ability, suggesting that another enzyme is able to perform the C2-C4 condensation step *in vivo*.

1.10.4 Structural and active site motifs

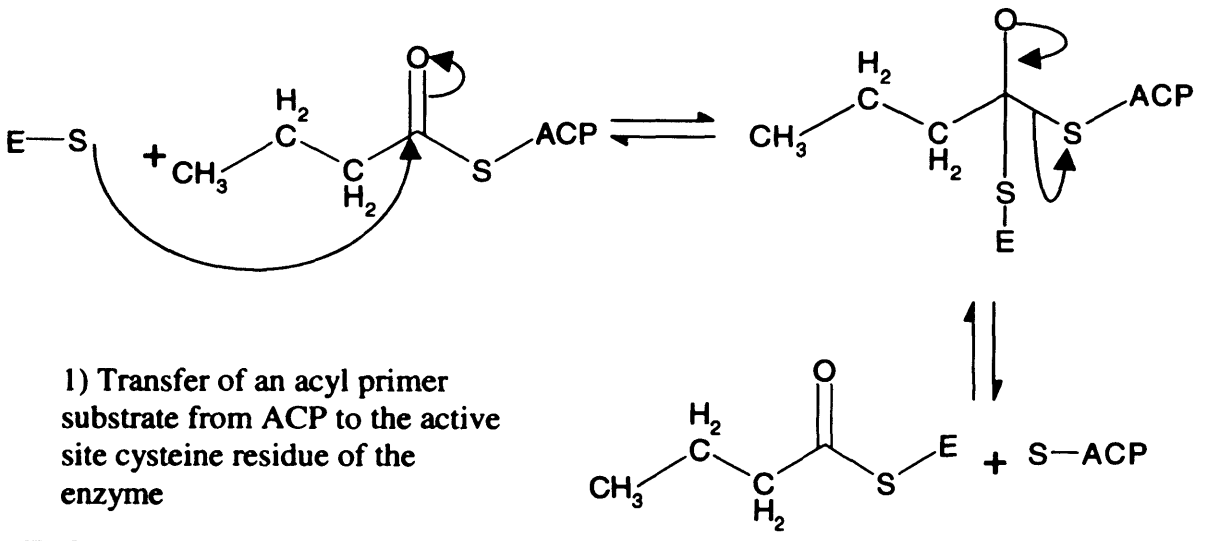
X-ray crystallography studies of all three β-ketoacyl-ACP synthases from *E. coli* and many other species suggest that the proteins share a similar fold and display some differences surrounding the active site (Olsen *et al.*, 1999; Huang *et al.*, 1998; Qiu *et al.*, 1999; Olsen *et al.*, 2001; Price *et al.*, 2001; Moche *et al.*, 1999; Moche *et al.*, 2001; Val *et al.*, 2000; Davies *et al.*, 2000; Qiu *et al.*, 2001; Scarsdale *et al.*, 2001). The quaternary structure of all KAS enzymes is a homodimer, the subunits associating by a 2-fold crystallographic axis. An extensive subunit interface region is formed and buried on dimerisation, resulting in a strong interaction between β-strands in the N-terminal regions that run antiparallel to each other, forming one continuous sheet (Price *et al.*, 2001; Moche *et al.*, 2001).

KAS enzymes exhibit overall similarities with the degrading enzyme, thiolase 1 from yeast and chalcone synthase (Moche *et al.*, 2001). Half of the subunit chain shows structural similarity, including the central β-sheets and the surrounding α-helixes, so that characteristic α-β-α-β-α layers are present. Outside of the conserved central region, the proteins show considerable differences in loop regions positioned at the bottom of the enzymes. These loop regions protect the active site from the external surroundings and may also result in the appropriate binding pockets for different hydrophobic substrates (Moche *et al.*, 2001).

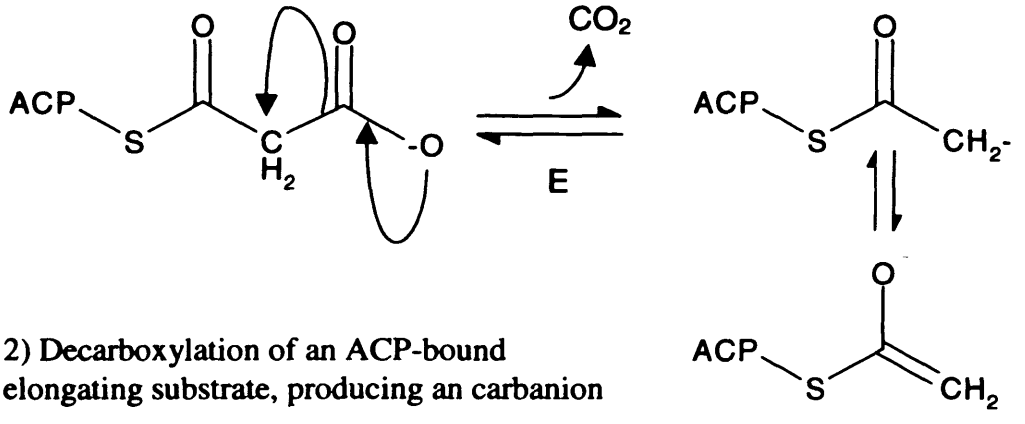
The active site cysteine residue resides at the carboxyl-terminus of an alpha helix dipole, resulting in an increase in the reducing capacity of the sulphydryl group. This interaction is essential for the activation of the sulphydryl group and the transfer of the acyl primer substrate from either coenzyme A or ACP. New carbon-carbon bonds are created by carrying out Claisen condensations (Dewar and Dieter, 1988). The reactions comprise three separate steps depicted in Figure 1.16. Briefly, the individual reactions are; the transfer of an acyl group from an acyl-ACP primer; the decarboxylation of malonyl-ACP the donor substrate and the condensation of the resulting carbanion with the enzyme bound acyl group.

1.10.5 Active site differences

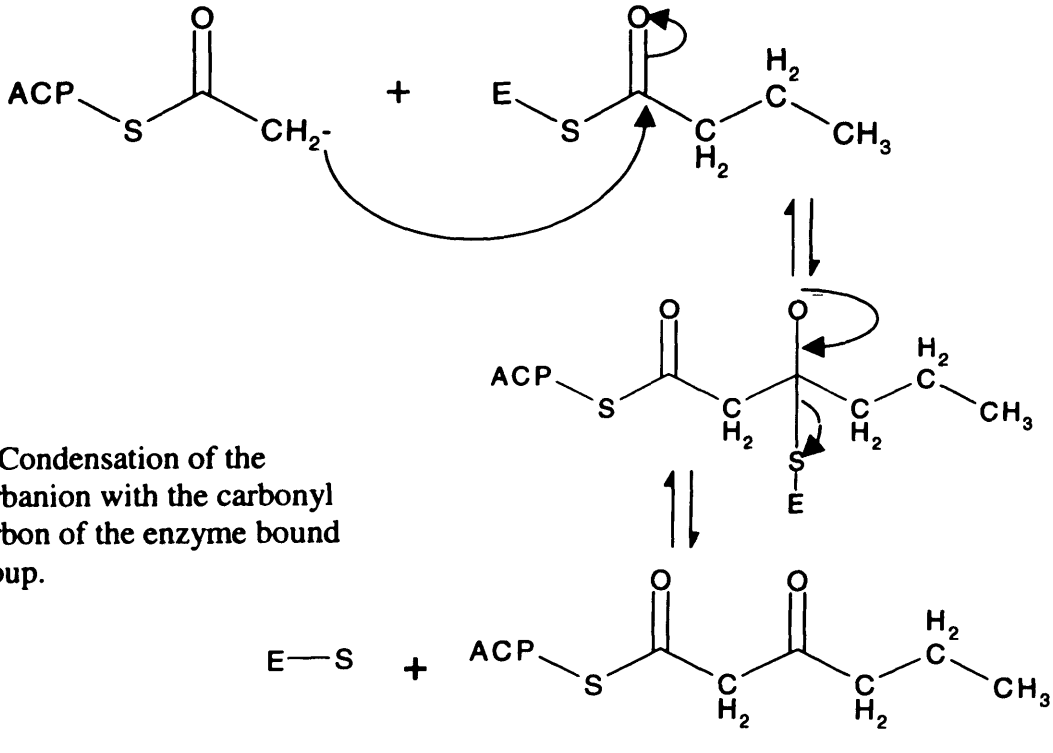
Alignments of FabB and FabF-like KAS enzymes highlighted several highly conserved residues and allowed the identification of the catalytic triad consisting of a cysteine and two histidines (Figure 1.17). There are, however, significant differences between the residues that constitute the active site of these enzymes and FabH enzymes. A catalytic triad of Cys112/His244/Asn274 (*E. coli* FabH numbering) comprise the active site of FabH enzymes, whereas FabB/FabF enzymes possess Cys163/His298/His333 (*E. coli* FabB numbering) motifs and an additional lysine that is essential for their correct positioning (Moche *et al.*, 2001; Olsen *et al.*, 2001; Price *et al.*, 2001; Siggaard-Anderson *et al.*, 1998).



1) Transfer of an acyl primer substrate from ACP to the active site cysteine residue of the enzyme



2) Decarboxylation of an ACP-bound elongating substrate, producing an carbanion



3) Condensation of the carbanion with the carbonyl carbon of the enzyme bound group.

Figure 1.16 Claisen condensation reactions as catalysed by KAS enzymes

Figure 1.17 Alignment of condensing enzymes shows regions of similarity.
Asterisks denote active site residues of either FabB /FabF or FabH enzymes. Conserved residues across all the enzyme types are shown against a black background. Residues of the same type (acidic D/E, basic R/K/H, hydroxyl containing S/T, small A/G, small hydrophobic L/I/V/M etc) are shown against a dark grey background. Regions that differ significantly between FabB /FabF are highlighted in cyan.

Hydrogen bonding to the His-His motif or the His-Asn motif stabilises the intermediate carbanion, prior to nucleophilic attack in condensation (Olsen *et al.*, 2001; Price *et al.*, 2001). Theoretically, histidine would be capable of performing the same role as the asparagine residue in FabH and reasons for the difference in active site residues between the two groups are yet to be determined.

1.10.6 Differences in the substrate-binding pocket

Structural differences are also observed outside the active site, most notably in the substrate binding pocket. In all KAS enzymes, the majority of the pocket is within one subunit, close to the dimer interface, with the second subunit making a small contribution to the binding cavity through several residues at conserved positions (Moche *et al.*, 2001). However, the substrate-binding pockets of FabB/FabF and FabH enzymes extend in different directions relative to the active site (Moche *et al.*, 1999). This probably results from the use of different acyl cofactors (CoA and ACP) and may explain the differences in sensitivity of KAS enzymes to the FAS inhibitor cerulenin (Olsen *et al.*, 1999).

Recently, Rock's laboratory has shown that surface interactions between Type II FAS enzymes and ACP contribute to the activity of this pathway (Zhang et al., 2001 and 2003). *E. coli* FabH and FabG both contain a positively charged/hydrophobic patch of arginine and lysine residues adjacent to the substrate binding tunnel. These interact with highly-conserved, negatively charged residues present in helix 2 of ACP. The physical surface of these two regions associate temporarily to form a coherent interface that contributes significantly to successful interaction.

1.11 Inhibition of Type II Fatty Acid Synthesis

The differences observed between the structure and organisation of Type I and Type II FAS, provide selective targets for inhibition of the Type II pathway present in pathogenic bacteria and parasites. These inhibitors were initially discovered due to their antibacterial properties and the large amount of research in this area is summarised below. Recent investigations of the anti-parasitic nature of these compounds is also included in this section.

1.11.1 Thiolactomycin

Thiolactomycin (TLM) (Figure 1.18) is a thiolactone that is a specific inhibitor of Type II β-ketoacyl-ACP synthases and was originally isolated from the soil bacteria *Nocardia* species 2-200 (Hayashi *et al.*, 1983; Miyakawa *et al.*, 1982; Noto *et al.*, 1982; Oishi *et al.*, 1982). The compound contains a single stereo centre at the C5 position and the bacteria synthesises only the (5R) enantiomer (Figure 1.18) (Sasaki *et al.*, 1982; Nishida *et al.*, 1986).

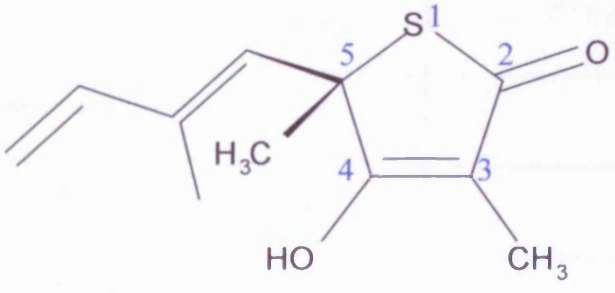


Figure 1.18 Structure of the natural antibiotic thiolactomycin

The numbers shown in blue are how each of the atoms of the thiolactone ring will be referred to within the rest of this report.

TLM is a selective inhibitor of Type II FAS. It has been shown to inhibit *E. coli* condensing enzymes FabB and FabF and to a lesser extent FabH and the equivalent enzymes from spinach and pea chloroplasts (Clough *et al.*, 1992; Jaworski *et al.*, 1989; Jones *et al.*, 1995; Jones *et al.*, 2000). In contrast, it has no effect on Type I FAS systems in *Saccharomyces cerevisiae* or rat liver (Hayashi *et al.*, 1983). This antibiotic is not toxic to mice at therapeutic doses and provides significant protection against urinary tract and intra-peritoneal bacterial infections (Miyakawa *et al.*, 1982). The potency of TLM in infected mice is achieved by rapid absorption and distribution to infective sites (Miyakawa *et al.*, 1982), this suggests favourable pharmacokinetics and suitability for drug administration. Thiolactomycin has been shown to inhibit the growth of several species *in situ* including the parasites *Plasmodium falciparum* and *Toxoplasma gondii* (Table 1.3).

Organism	IC_{50} of TLM (μM)
P. falciparum	50 [a]
T. gondii	200 [b]
Trypanosoma brucei	150 [c]
Mycobacterium tuberculosis	280 [d]
Pisum sativum (pea)	20 [e]
Spinacia oleasuceae	3.8 [f]

Table 1.3 Inhibition of growth observed with thiolactomycin

[a] Waller *et al.*, 1998 and 2003. [b] McFadden *et al.*, 1999. [c] Morita *et al.*, 2000.[d] Kremer *et al.*, 2000.[e] Jones *et al.*, 1995 and 2000. [f] Clough *et al.*, 1992.

Disruption of apicoplast function was first reported in *Toxoplasma gondii* using the prokaryotic gyrase inhibitor, ciprofloxacin and clindamycin, a prokaryotic transpeptidase inhibitor (Pfefferkorn *et al.*, 1992; Fichera *et al.*, 1995; Fichera and Roos, 1997). These compounds, directly or indirectly, inhibit the replication of the plastid genome and reduce its copy number compared to genomic DNA (Fichera and Roos, 1997). The inhibition of *T. gondii* parasites *in situ* exhibited distinct kinetics. There was no effect on tachyzoites growing within the first host cell. However, parasites within the second cycle failed to grow in newly infected cells, suggesting that disruption of apicoplast division and function decreases the viability of daughter merozoites.

An investigation of the action of thiolactomycin in blood stage cultures of *Plasmodium falciparum* was performed recently by McFadden's laboratory (Waller *et al.*, 2003). These studies showed that TLM did not display a delayed-death phenotype but inhibited growth of the parasite most effectively in early rings. Inhibition of growth during trophozoite stages was greatly reduced and TLM appeared to have little or no effect against schizonts. Similar results were observed with the enoyl reductase inhibitor, triclosan and this compound is also capable of clearing malaria infections in mice. These results suggest that *de novo* FAS is most important for the development of merozoites following the early erythrocytic stages.

1.11.1.1 Enzyme targets of thiolactomycin

In vitro, thiolactomycin inhibits all of the *E. coli* condensing enzymes (FabB, FabF and FabH), to various extents (Table 1.4).

Enzyme	IC_{50}
E. coli FabB [a]	6μM
E. coli FabF [a]	25μM
E. coli FabH [a]	110μM
Staphylococcus aureus FabF [b]	≈80μM

Table 1.4 The inhibition of bacterial condensing enzyme activity by thiolactomycin

[a] Price *et al.*, 2001. [b] Jackowski *et al.*, 2001.

As well as direct *in vitro* inhibition studies, *E. coli* strains that overexpress FabB exhibit some resistance to thiolactomycin (Tsay *et al.*, 1992b). In contrast, overexpression of FabH conferred no resistance to TLM (Tsay *et al.*, 1992b). A selection process used to identify TLM resistant *E. coli*, isolated only mutants with alterations in the *fabB* gene. No *fabF* or *fabH* mutants were observed. These combined data suggest that despite that *in vitro* the FabB enzyme exhibits lower sensitivity to TLM than FabF, FabB is the physiologically relevant target of TLM in this bacterium.

In *M. tuberculosis*, overexpression of KAS A and KAS B, respectively conferred resistance to TLM (Kremer *et al.*, 2000; Slayden *et al.*, 1996). The relative sensitivity of the KAS enzymes to TLM differs between organisms. However, FabH (KAS III) is

always reported as the least sensitive (Harwood, 1996; Jones *et al.*, 1995; Jones *et al.*, 2000; Price *et al.*, 2001). A general observation is that in higher plants sensitivity to TLM is greater in KAS II than KAS I but in bacteria FabB (KAS I) has greater sensitivity (Jones *et al.*, 1995; Jones *et al.*, 2000; Kremer *et al.*, 2000; Price *et al.*, 2001).

Compared to TLM, some thiolactomycin derivatives have shown greater inhibition of Type II FAS in higher plant chloroplasts (Jones *et al.*, 2000), *M. tuberculosis* (Kremer *et al.*, 2000) and *P. falciparum* cultures *in vitro* (Waller *et al.*, 2003). Generally, there is greater sensitivity towards derivatives with extended hydrocarbon chains (from 8-12 carbons in length) and the derivatives with the lowest IC_{50} against *P. falciparum* and *M. tuberculosis in vitro*, are unsaturated and contain some methyl side groups (Kremer *et al.*, 2000; Waller *et al.*, 2003). The link between increased side-chain length and increased efficacy is consistent with occupation of a large hydrophobic pocket, present in FabB and FabF proteins but absent in FabH. This pocket does not incorporate the growing acyl chain and its function is unknown. However, its absence in FabH may be the reason for the lack of this enzyme's sensitivity to TLM.

Compounds with more double bonds show an increased inhibitory activity compared to saturated chains of the same length, indicating that increased desaturation and/or rigidity of the side-chain might contribute to inhibition of the target enzymes (Harwood, 1996; Kremer *et al.*, 2000; Waller *et al.*, 2003). The results with thiolactomycin in *Mycobacterium* and pea, suggest that the hydrophobic crevice that may interact with the extending fatty acid chain of TLM derivatives is a common

feature to both FabB and FabF. These results provide information that contribute to the rational design of thiolactomycin based antibiotics.

1.11.1.2 Mechanism of thiolactomycin inhibition

Kinetic studies of *E. coli* FabB, elucidated the mechanism of thiolactomycin inhibition (Price *et al.*, 2001). Originally, it was thought that the sulphur group of TLM would form a covalent bond with the active site cysteine via a thioester exchange reaction. Incubation of TLM (20μM) with FabB 30min prior to the addition of the other assay components showed no decrease in enzyme activity compared to assays when TLM was added at the start point of the assay. To verify whether TLM inhibition is reversible or irreversible, a dilution experiment was performed. FabB was incubated with 100μM TLM, causing 95% inhibition. The TLM concentration was then diluted to 1μM and the activity of FabB was recovered to the level observed when continuously incubated with 1μM TLM. The decrease in inhibition indicated that TLM binding is reversible and results in no permanent interactions. This contrasted with the covalent inhibitor cerulenin, where a marked reduction in the activity of the pre-incubated FabB was observed (Price *et al.*, 2001). These data showed that TLM is a reversible inhibitor of the FabB enzyme, reinforcing results previously seen in cell extracts (Hayashi *et al.*, 1983).

The study in Rock's laboratory also calculated the affinity of TLM for each of the *E. coli* KAS enzymes (Price *et al.*, 2001). FabB exhibited the highest affinity (26μM), FabF showed intermediate affinity (60μM) and FabH had the lowest affinity (158μM) (Price *et al.*, 2001). This was consistent with IC_{50} values and results of other kinetic experiments, generated in the same study.

1.11.1.3 The structure of a FabB-TLM complex

The structure of TLM complexed to *E. coli* KAS I (FabB) was the first, and so far only, structure of a condensing enzyme with thiolactomycin bound to the active site (Figure 1.19).

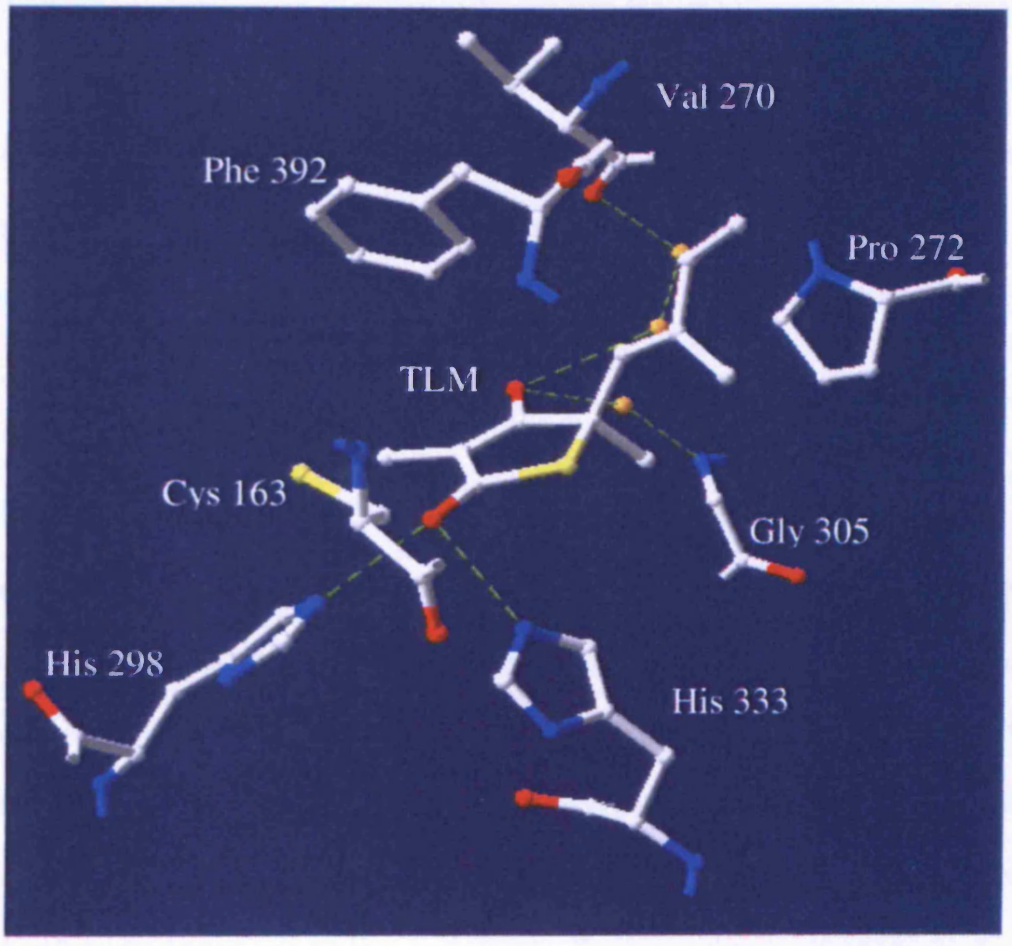


Figure 1.19 TLM bound in the active site of *E. coli* FabB

The inhibitor and key residues of *E. coli* FabB are shown. Hydrogen bonds are shown as green dashed lines. Water molecules are shown in orange. The 1FJ4.pdb file (Price *et al.*, 2001) was downloaded from the Brookhaven database and manipulated using SwissPDB viewer (Guex and Peitsch, 1997).

The structure shows that TLM makes a number of specific but non-covalent interactions with FabB. The isoprenoid group is sandwiched between two peptide bonds and an interaction between these three delocalised systems appears important to specificity. The C3 and C5 methyl side groups of TLM are sequestered within two hydrophobic pockets consisting of Phe229 & Phe392 and Pro272 & Phe290. Both the carbonyl and hydroxyl oxygens of TLM are involved with hydrogen bonding to the protein. The carbonyl oxygen forms hydrogen bonds to the two active site histidines (298 and 333, respectively), whilst the hydroxyl oxygen bonds via a lattice of water molecules to the carbonyl oxygen of valine 390 and the amide nitrogen of glycine 305 (Figure 1.19). Interestingly, the thiolactone sulphur does not make any specific interactions but lies adjacent to the sulphur of the active site Cys163 (Price *et al.*, 2001).

Differences in hydrogen-bonding probably contribute to the lower sensitivity of *E. coli* FabH to TLM, where the active site His333 is replaced with an asparagine (Davies *et al.*, 2000; Qiu *et al.*, 1999). The importance of the histidine in FabB was studied in a FabB[H333N] mutant that retained significant condensation activity but sensitivity to thiolactomycin was reduced by an order of magnitude. This establishes the importance of a His-His hydrogen bonding to the carbonyl oxygen of TLM for high affinity TLM binding (Price *et al.*, 2001). However, the specificity for two histidines is not completely understood because the asparagine could also donate a hydrogen bond to the carbonyl group of TLM. *E. coli* FabF (Huang *et al.*, 1998; Moche *et al.*, 2001), also shares the His-His active site motif with FabB. However, the reason for its lowered affinity for TLM is unknown.

The high affinity of TLM for FabB can be further explained by the minimal distortion caused by TLM in the active site. Local changes in FabB structure are observed but do not extend beyond the immediate vicinity of the active site, notably, the side chain of Cys163 moves 2.1Å to accommodate the TLM sulphur and His298 moves by 2Å to improve hydrogen bonding to the carbonyl group of TLM. Finally, the relocation of His-298 causes knock-on effects in the loop comprising residues 388-394, due to hydrogen bonding between the Nε-2 nitrogen of His298 and the main chain carbonyl of residue Phe390 (Price *et al.*, 2001).

1.11.1.4 Significance for the development of thiolactomycin as an antibiotic

The crystal structure of *E. coli* FabB-TLM shows that the isoprenoid moiety fits within a hydrophobic pocket of the enzyme, which has indiscernible function during fatty acid synthesis (Olsen *et al.*, 2001; Price *et al.*, 2001). Previously, it was suggested that the isoprenoid group might extend further along the active site (Jones *et al.*, 2000; Kremer *et al.*, 2000) but this is not the case. Furthermore, the structures of both FabB and FabF (Huang *et al.*, 1998) appear to contain a similar hydrophobic pocket that extends deeper into the enzyme, which may explain the higher sensitivity of these enzymes to longer chain TLM derivatives, as observed in *Mycobacterium* and pea (Jones *et al.*, 2000; Kohler *et al.*, 1997). An understanding of structural aspects of the KAS enzymes and their interactions with inhibitors, such as thiolactomycin derivatives, are clearly very important for developing thiolactomycin derivatives with increased antimicrobial activity.

1.11.1.5 Resistance to thiolactomycin

Selection for *E. coli* mutants that were resistant to TLM showed that a single base mutation in the *fabB* gene was responsible for increasing the resistance to TLM by an order of magnitude (Jackowski *et al.*, 2002). In these selection screens, no other *fabB* mutations or mutations of other genes were observed, suggesting that FabB is the physiologically relevant target of TLM in this organism. This mutation altered a highly conserved phenylalanine residue to a valine (F390V), without affecting the specific activity of the enzyme. *Bacillus subtilis* FabF contains a leucine at position 390 (Schujman *et al.*, 2001) and is naturally resistant to TLM. This suggested that both F390V and F390L substitutions may be responsible for TLM resistance. Structural models suggested that a steric clash between a side-chain methyl group of V390 and the branched ethene group of the TLM isoprenoid side chain. However, in the F390L model no steric clash was observed and the mode of TLM binding was similar to the wildtype. Furthermore, no *E. coli* FabB F390L mutants (generation of which would require only a single base mutation) were selected in TLM resistant screens and *Staphylococcus aureus* contains a F390L FabF that is highly sensitive to TLM. These data suggest that the FabF F390L alteration in *Bacillus subtilis* is not responsible (at least entirely) for the enzyme's resistance to TLM.

TLM resistance has also been linked to a mutation in the *emrR* gene, which causes upregulation of the *emrAB* efflux pump, a membrane transporter that confers resistance to structurally unrelated but moderately hydrophobic compounds (Furukawa *et al.*, 1993; Lomovskaya *et al.*, 1995). Abrogation of the *emrAB* efflux pump activity in *E. coli* knockouts lowered the MIC of TLM from 150μM to 25μM (Furukawa *et al.*, 1993). However, knockouts of all of the *tolC*-dependent efflux

pump family gave an MIC of 3μM (Jackowski *et al.*, 2002; Buchanan, 2001), suggesting that other unidentified efflux pumps may be involved in TLM resistance. Alterations in the TLM molecule that would give it less affinity for efflux pumps would be advantageous in treatment against gram-negative bacteria.

1.11.2 Other inhibitors of FAS as potential antibiotics

Thiolactomycin is not the only known inhibitor of FAS (see following section) and the potential of TLM as an antibiotic increases the relevance of other Type II FAS inhibitors. Antibiotics that block different steps of the FAS pathway could be combined, providing a 'multi-block' strategy and increasing their efficacy (White and Ollario, 1996). Used judiciously, combination therapy or the rotation of Type II FAS inhibitors through the global population could also lower the probability of resistance emerging to any individual drug.

1.11.2.1 Cerulenin

Cerulenin (CER) (Figure 1.20) is an irreversible inhibitor of FabB and FabF that forms a covalently bound complex with the active site cysteine of these enzymes (Price *et al.*, 2001). In a situation similar to TLM, FabH enzymes are generally insensitive to CER, even at high concentrations.

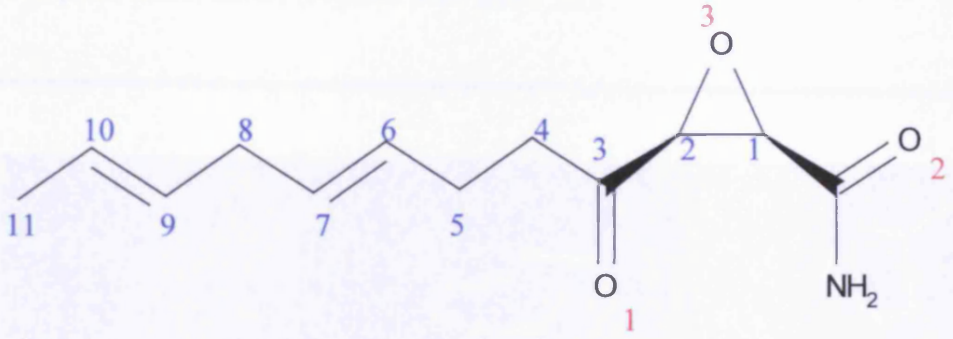


Figure 1.20 Structure of the natural antibiotic cerulenin

Carbon atoms are numbered in blue text and oxygen atoms in red text.

Cerulenin is not a selective antibiotic because it is also a potent inhibitor of the condensing reactions in Type I FAS (Funabashi *et al.*, 1989; Harwood, 1996). Furthermore, cerulenin has also been shown to affect feeding behaviour in mice, resulting in reduced food intake and body weight. Whilst cerulenin-based compounds are, therefore, excellent candidates in the fight against obesity and heart disease in the western World, these complications are undesirable for an antimalarial drug that may be used on a high proportion of malnourished patients (Loftus *et al.*, 2000). Structural studies, described below, with Type II condensing enzymes (Moche *et al.*, 1999; Price *et al.*, 2001) are central to improving selectivity of cerulenin.

Both FabB and FabF *E. coli* enzymes are inhibited by cerulenin, FabB being the most sensitive (Moche *et al.*, 1999; Price *et al.*, 2001). FabH is not effectively inhibited by cerulenin and was discovered because of this insensitivity (Harwood, 1996). The crystal structures of the *E. coli* FabB-CER and *E. coli* FabF-CER complexes show that cerulenin primarily occupies a different area to thiolactomycin (Moche *et al.*, 1999; Price *et al.*, 2001) (Figures 1.21 and 1.22).

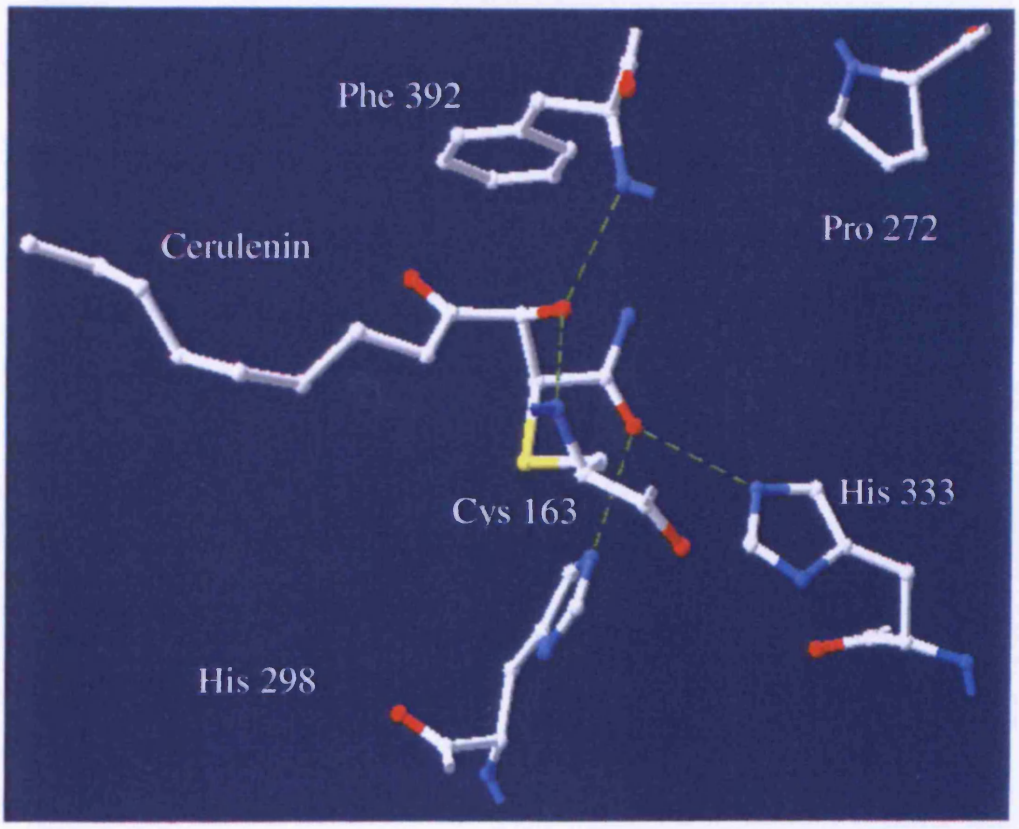


Figure 1.21 Cerulenin bound in the active site of *E. coli* FabB

Hydrogen bonds are shown as green dashed lines. Water molecules are shown in orange. The 1FJ8.pdb file (Price *et al.*, 2001) was manipulated using SwissPDB viewer (Guex and Peitsch, 1997).

Its extended acyl chain is located within a deep hydrophobic pocket at the dimer interface that is believed to harbour the hydrocarbon chain of the acyl intermediate during fatty acid synthesis (Olsen *et al.*, 2001). This hydrophobic pocket in FabH is shallower, suggesting a possible reason for its insensitivity to cerulenin (Qiu *et al.*, 1999).

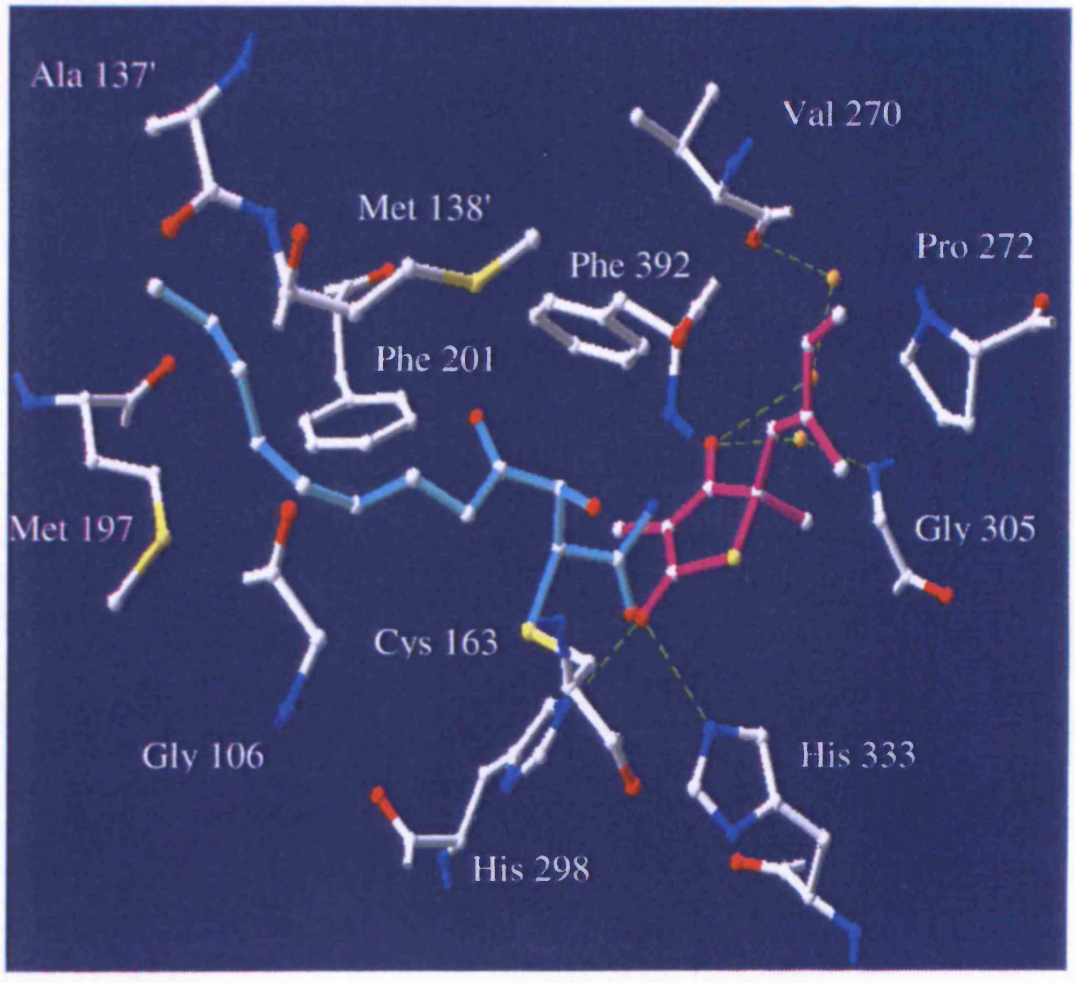


Figure 1.22 Thiolactomycin and cerulenin bound in the active site of *E. coli* FabB

Hydrogen bonds are shown as green dashed lines. Water molecules are shown in orange. Cerulenin is shown in cyan and TLM in magenta. The 1FJ4.pdb and 1FJ8.pdb files (Price *et al.*, 2001) were downloaded from the Brookhaven database and were then manipulated using SwissPDB viewer (Guex and Peitsch, 1997).

The covalent bond is formed between the central C2 of cerulenin and the active site Cys163 sulphur atom (*E. coli* numbering). The O2 atom of CER is positioned in the same position as the carbonyl atom of TLM and, as expected, forms hydrogen bonds with the Nε-2s from His298 and His333. Further hydrogen bonding occurs between the O3 atom of cerulenin and the main chain amide groups of Cys163 and Phe392 (Moche *et al.*, 1999; Price *et al.*, 2001).

Like TLM, cerulenin occupies the active site with minimal distortion of the FabB or FabF structures and the slight movements that do occur are similar to those observed in the TLM complex, described in Section 1.10.1.3 (Moche *et al.*, 1999; Price *et al.*, 2001). Interestingly, Ile108, a residue of the hydrophobic binding pocket, has to rotate to incorporate cerulenin in FabF but not FabB (Moche *et al.*, 1999 ; Price *et al.*, 2001). This suggests a possible reason for the lower sensitivity of FabF to cerulenin. Further demonstration of the importance of Ile108 was observed in a FabF [Ile108Phe] mutant that shows decreased activity with acyl-ACP substrates over 8 carbons in length. In the presence of 50μM cerulenin this mutant retains 90% of its activity (the wild-type enzyme showed no activity at this concentration) (Val *et al.*, 2000), suggesting that a single amino acid substitution may confer resistance to cerulenin.

1.11.2.2 Triclosan

Triclosan (Figure 1.23) is a broad-range antimicrobial used in many domestic products, such as toothpaste and detergents, due to its favourable safety profile (Bhargava and Leonard, 1996).

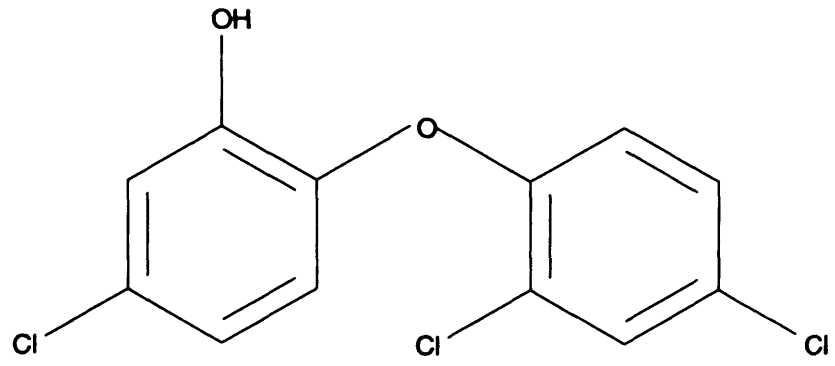


Figure 1.23 Structure of the antibiotic and antimalarial compound triclosan

Unfortunately, due to hydrophobic nature it is poorly absorbed when taken orally but can be effective when administered through an intravenous or intraperitoneal route. Recently the Type II FAS reducing enzyme, enoyl-ACP reductase (FabI in *E. coli*), was elucidated as the target of its antimicrobial activity (Heath *et al.*, 1999; Heath and Rock, 2000; McMurry *et al.*, 1998). This enzyme is responsible for the last step of the FAS cycle and is, therefore, an important site for chemical control (Heath and Rock, 1995). Triclosan inhibits FabI by promoting the binding of its oxidised cofactor, NAD^+, to the active site, forming a robust ternary complex of FabI-NAD^+-triclosan, effectively removing FabI from the FAS pathway (Heath *et al.*, 1999; Saguna *et al.*, 2001; Kapoor *et al.*, 2001).

Triclosan appears to be selective for Type II FAS enzymes and is able to inhibit significantly the growth of *T. gondii, in vitro* and *P. falciparum, in vitro* and *in vivo* (McLeod *et al.*, 2001; Surolia and Surolia, 2001). Furthermore, an enoyl-ACP reductase containing an apicoplast targeting sequence that was identified and purified from *P. falciparum*, is triclosan-sensitive (McLeod *et al.*, 2001; Surolia and Surolia,

2001). The IC_{50} against *P. falciparum* in culture is 0.7μM, much lower than that of thiolactomycin (50μM) and cerulenin (20μM) (Price *et al.*, 2001). Triclosan was also the first inhibitor of Type II FAS that was shown to clear *P. berghei* parasites from mice (Surolia and Surolia, 2001). The pfFabI was the first Apicomplexan fatty acid synthesis enzyme for which a 3D structure was solved (Perozzo *et al.*, 2002) and the co-ordinates have since been refined (Muench *et al.*, 2003). Complexes with triclosan and active derivatives have also been reported (Suguna *et al.*, 2001; Perozzo *et al.*, 2002; Kapoor *et al.*, 2004a, b, c) and these studies will contribute to the understanding or triclosan inhibition and may lead to the first anti-parasite drug that targets fatty acid synthesis.

During 30 years of use in many products, resistance to triclosan is uncommon but may be conferred by overexpression of FabI or a single mutation in the FabI gene (Heath and Rock, 2000; Kapoor *et al.*, 2004c). This suggests that, used judiciously, triclosan would be an effective addition to our arsenal of antimalarial Type II FAS inhibitors.

1.11.2.3 Aryloxyphenoxypropionate Herbicides

Aryloxyphenoxypropionate herbicides target acetyl CoA carboxylase (ACCase) localised in the chloroplasts of Graminaceae (Harwood, 1996). The ACCases of man, rat and yeast (and other plants) are insensitive to treatment by these compounds, allowing them to be used commercially as selective herbicides and implying a selectivity for particular ACCase structures (Harwood, 1996; Zuther *et al.*, 1999). They have been shown to inhibit the growth of the apicomplexan *T. gondii*, *in vitro* by targeting of its ACCase, via an unknown mechanism. The most promising of these

compounds showed 70% inhibition of growth of *T. gondii* cultures at 10μM concentrations (no IC_{50} data available) (Zuther *et al.*, 1999).

1.11.2.4 C_{18} Fatty Acids

Several fatty acids (oleic, elaidic, linoleic and linolenic) inhibit *P. falciparum* growth *in vitro* and also lower parasitaemia in murine models (Krugliak *et al.*, 1995). Nucleic acid and protein synthesis was inhibited after 30 minutes in both ring-stage and trophozoite infected cells and was irreversible (shown by pre-incubation with 200μg/ml oleic acid). However, despite rigorous tests, the mechanism of action is so far unknown (Krugliak *et al.*, 1995).

1.12 Aims of this Project

As outlined in this Chapter, β-ketoacyl ACP synthases from Apicomplexa are potentially very important targets for the development of new antibiotics (McFadden and Roos, 1999). In other organisms, most notably *E. coli*, these enzymes have been expressed and purified to allow characterisation and X-ray crystallography to investigate inhibitor binding (Edwards *et al.*, 1997; Jackowski and Rock, 1987). At the outset of this project only the gene for FabH (*pffabH*) had been identified (Gardner *et al.*, 1998; Waller *et al.*, 1998) in *P. falciparum*. The pfFabH protein had yet to be produced in a recombinant system and, therefore, had not been characterised. Furthermore, homologues of FabB and/or FabF, more promising targets for

thiolactomycin inhibition had not been identified. The aims of the project were to study *P. falciparum* FabB, FabF and FabH and their inhibition as follows -

- To clone, express in recombinant form and purify *P. falciparum* FabH.

- To identify the gene/s encoding FabB and/or FabF enzymes from *P. falciparum*, to successfully express and purify these protein/s in recombinant form.

- To perform studies on the substrate specificity and kinetic parameters of all the *P. falciparum* KAS enzymes.

- To perform inhibition studies of *P. falciparum* KAS enzymes with thiolactomycin and its derivatives and compare these to inhibition of the parasites *in vitro*.

- To carry out structural studies of thiolactomycin and/or the most promising TLM derivative complexed to the condensing enzymes, which would provide further information on interactions between the two molecules. This would facilitate further, rational, design of an effective antimalarial.

- To identify a new lead compound that targets *P. falciparum* fatty acid synthesis, which may be developed further as a potential drug for the treatment of malaria.

Chapter 2

2.1 Materials

2.1.1 Cloning and Expression Vectors

The pGEM-T vector was purchased from Promega Ltd. (Southampton, UK), the pET-22b vector was purchased from Novagen (AMS Biotechnology, Milton Keynes, UK), pGEX-4T-2 from Qiagen (West Sussex, UK) and pMALc2x from New England Biolabs (Beverly, MA, USA). A pMALc2x plasmid containing the full length sequence of *P. falciparum* ketoacyl-ACP synthase III was kindly donated by Dr. Sean Prigge (Johns Hopkins School of Public Health, Baltimore, USA). The pDS-His6-FabD plasmid was a kind gift from Dr. Roland Lange (Morphochem AG, Basel, Switzerland). The MG1655 wildtype and FabF knockout strains were kindly provided by the Blattner research group (Wisconsin, USA). The LCH68 strain that overexpresses acyl-ACP synthase was kindly provided by Dr. Charles Rock (St. Jude Children's Research Hospital, Memphis, USA).

2.1.2 *Escherichia coli* Host Strains

All genetic manipulations were carried out in the *E. coli* strain DH5α (Pharmacia Biotechnology Ltd., Milton Keynes, UK). *E. coli* expression studies were performed in the strains, BL21(DE3) pLysS, (purchased from AMS Biotechnology, Milton Keynes, UK) or BL21(DE3) Codon Plus-RIL (purchased from Stratagene, Amsterdam, The Netherlands). The genotype of each of the bacterial strains is detailed in Table 2.1.

Strain	Genotype	Reference
DH5α	*EndA1 hsdR17 glnV44 thi-1 recA1 gyrA relA1 Δ(lacIZYA-argF)U169 deoR (phi80dlacΔ(lacZ0M15)*	Hanahan, D. (1983)
BL21 (DE3) pLysS	*E. coliB dcm lon ompT hsdS(tB-mB-) gal lambda (DE3) lacIQ/placUV5-T7gene1*	Studier, F.W. and Moffat, B.A. (1986)
BL21 (DE3) Codon Plus - RIL	*E. coli* B F– *ompT hsdS*(r_B– m_B–) *dcm*+ Tet^r *gal λ* (DE3) *endA* Hte *[argU ileY leuW* Cam^r]	Studier, F.W. and Moffat, B.A. (1986)

Table 2.1. The genotypes of *E. coli* strains used for recombinant expression in this study

2.1.3 Media

Luria Bertani (LB) nutrient medium was prepared as described by the manufacturer's instructions (Sigma Chemicals Ltd., Poole, Dorset, UK). M9 minimal media was prepared as described in Sambrook *et al.*, 1989. Antibiotics, when present in 1.2% agar plates or media, were used at the following concentrations; ampicillin (100μg/ml), chloramphenicol (34μg/ml), kanamycin (25μg/ml). All media and constituents were sterilised either by autoclaving at 121°C (975kPa) for 20min or by filtration through 0.22μm nitrocellulose filters (Gelman Science, Ann Arbor, MI., USA).

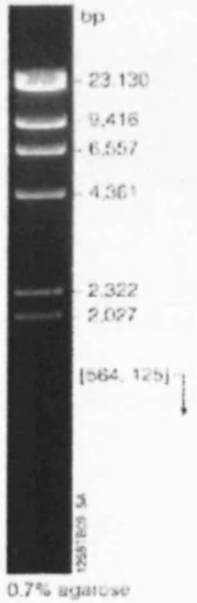
2.1.4 Molecular Weight Standards

Pre-stained high range molecular weight protein standards were purchased from BIORAD Laboratories, Hertfordshire, UK. The λ *Hin*dIII and ϕX174 *Hae*III DNA markers were purchased from Promega Ltd., Southampton, UK (Figure 2.1).

<u>High range molecular weight protein markers (kDa)</u>

Myosin (H chain)	200.0
Phosphorylase B	97.4
Bovine serum albumin	68.0
Ovalbumin	43.0
Carbonic Anhydrase	31.0
β-Lactoglobulin	18.4
Lysozyme	14.3

<u>Lambda DNA/*Hin*d III Markers</u>



<u>PhiX174 DNA/*Hae* III Markers</u>

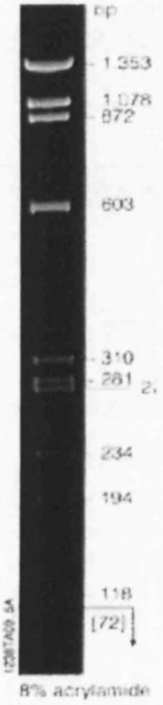


Figure 2.1 DNA markers visualised on agarose and acrylamide gels by UV light

Photos taken from the Promega Ltd. (Southampton, UK) webpage http://www.promega.com/techserv/apps/cloning/cloning4.htm on the 13/09/2004.

2.1.5 Other Materials

Routine laboratory chemicals were purchased from Sigma Chemicals Ltd., (Poole, Dorset, UK) and BDH Chemicals Ltd., (Poole, Dorset, UK). The suppliers of more specialised chemicals and equipment are listed below (Table 2.2). All radioactive compounds including [γ-^{32}P]-ATP, [1-^{14}C] acetyl-CoA, [2-^{14}C] malonyl-CoA were purchased from Amersham Pharmacia, Little Chalfont, UK.

Materials and Equipment	Companies and Addresses
Thiolactomycin (racemic mixture)	Sigma Chemicals Ltd., Poole, Dorset, UK
PCR Oligonucleotides	MWG Biotech, Ebersburg, Germany and Qiagen Ltd., West Sussex, UK
Restriction Endonucleases and DNA modifying enzymes	Promega Ltd., Southampton, UK and New England BioLabs, Beverly, MA, USA
Wizard Minipreps DNA Purification System	Promega Ltd., Southampton, UK
Acrylamide /Bisacrylamide (SDS-PAGE)	National Diagnostics, Manville, NJ, USA
HiTrap® SP 1ml ion exchange columns	Pharmacia LKB, Biotechnology, St Albans, Herts, UK

Table 2.2 Suppliers of other materials and equipment

The ring stage *P. falciparum* cDNA library was kindly provided by Prof. Steve Walsh of the Liverpool School of Tropical Medicine.

2.2 Methods

2.2.1 DNA Amplification by the Polymerase Chain Reaction (PCR)

PCRs were performed as first described in Mullis (1986) (Mullis *et al.*, 1986). Reaction mixtures were set-up with the final concentrations described below; containing 0.2mM dNTPs, PCR buffer (50mM KCl, 10mM Tris-HCl, pH9.0 at 25°C, 0.1% Tween20, 3mM $MgCl_2$) and 10pmoles of the appropriate primers were made up to a final volume of 20 or 50µl with nuclease-free water (Sigma Chemical Company, Poole, Dorset, UK). Either *Taq* polymerase (Promega Ltd., Southampton, UK), or *Pfu* polymerase (Stratagene, Amsterdam, The Netherlands), was added to a concentration of 2.5 units per 50µl reaction mixture. The tubes were spun briefly in a microfuge before the addition of a thin layer of mineral oil (Sigma Chemical Company, Dorset, UK), which served to minimise evaporation during thermocycling. The thermocycling steps were performed using a TRIO-Thermoblock thermal cycler (Biometra, Goettingen, Germany) or a Flexigene thermal cycler (Techne, Cambridge, UK). Denaturation, annealing, extension conditions and cycle numbers are described in the appropriate section of text. PCR products were analysed by submarine agarose gel electrophoresis (Section 2.2.6).

2.2.2 Ligation of DNA into Plasmid Vectors

DNA inserts and plasmid vectors were digested with the appropriate restriction endonucleases. Each restricted product was gel purified prior to ligation. In a typical ligation reaction, a vector to insert ratio of 1:4 (on the basis of molarity) was maintained in a final reaction volume of 10µl. DNA ligase and the appropriate buffer

(Promega Ltd., Southampton, UK) were added and the mixture was incubated for 16h at 16°C or for 3h at 4°C.

2.2.3 Creation of Blunt-Ended DNA Fragments with Klenow Enzyme

Reactions were made up to a minimal volume in nuclease-free water as follows;

DNA sample
1 Unit Klenow per μg DNA
33μM dNTPs
1 x Reaction buffer (Promega or New England Biolabs)

The samples were incubated at 25°C for 15min and the Klenow enzyme deactivated at 75°C for 20min.

2.2.4 Shrimp Alkaline Phosphatase (SAP)

Reactions were made up to the following conditions in a minimal volume of nuclease-free water;

DNA sample
1 unit of SAP (Roche, Basel, Switzerland) per picomole of DNA
1 x SAP buffer (from 10 x stock)

The reaction mixture was incubated at 37°C for 3h and the phosphatase inactivated at 75°C for 15min.

2.2.5 Isolation of Plasmid DNA

Isolation of plasmid DNA was performed in accordance with the instructions supplied with the Wizard ™ Miniprep Kit or Wizard ™ Maxiprep kits (Promega Ltd., Southampton, UK). This procedure is a modification of the Birnboim and Doly (Birnboim and Doly, 1979) alkaline lysis method and uses a resin purification step to yield high quality DNA.

2.2.6 DNA Modifying Enzymes and Submarine Agarose Gel Electrophoresis

DNA modifying enzymes were used according to the manufacturer's instructions. Restriction digests were analysed by submarine electrophoresis in 0.5% - 2% agarose gels containing ethidium bromide (1.3μM). Sizes of digested fragments were estimated from their migration relative to DNA markers of known lengths (Section 2.1.4.).

2.2.7 Purification of DNA Fragments from Agarose Gel

Following resolution in an agarose gel, the section of gel containing the DNA fragment of interest was excised and the agarose melted, by heating at 70°C. Extraction of the DNA was performed in accordance with the Qiagen DNA Purification System for DNA Fragments (Qiagen Ltd., West Sussex, UK).

2.2.8 Double Stranded DNA Sequencing

Double stranded DNA templates were sequenced by the dideoxy-chain termination method originally described by Sanger *et al.* (Sanger *et al.*, 1977). Automated sequencing of plasmid DNA was performed in the research support unit in the Cardiff School of Biosciences, Cardiff University. The ABI PRISM™ Dye Terminator Cycle Sequencing Ready BIG-DYE Reaction Kit (Perkin Elmer) was used to carry out the sequencing reactions. Briefly, the Terminator Ready Reaction solution (4µl) was mixed with 10ng of the appropriate primer and approximately 500ng of plasmid DNA in a final volume of 10µl. Thermal cycling (25 cycles of 96°C for 10s, 50°C for 5s, 60°C for 4min) was performed in an MWG Biotech Primus thermocycler (MWG Biotech, D-85560 Ebersberg, Germany), the extension products were then precipitated using sodium acetate and ethanol and the pellets resuspended in 50% (v/v) formamide. The whole of this sample was loaded onto a PE Biosystems ABI PRISM 3100 capillary DNA sequencer. Chromatograms produced by this machine were analysed using the BioEdit® computer program (Tom Hall, Department of Microbiology, North Carolina State University, USA).

2.2.9 Preparation of Competent *E. coli* Cells

Cells were grown for 16h in LB medium at 37°C with shaking and harvested by centrifugation (2,000 x g, 5min). The cells were resuspended in half of the original culture volume of 100mM $CaCl_2$/15% (v/v) glycerol and incubated on ice for 30min. The cells were harvested in a centrifuge as before and resuspended in $1/10^{th}$ of the original volume of the $CaCl_2$/glycerol mixture. The cell suspension was divided into

50µl or 100µl aliquots, snap frozen by dropping into liquid nitrogen and stored at -80°C.

2.2.10 Transformation of *E. coli* Cells by Plasmid DNA

Prior to transformation, an aliquot of competent cells (50µl) was thawed on ice. After thawing, 2µg of DNA was added to the competent cells. The cells were incubated on ice for 20min, subjected to heat shock at 42°C for 2min and placed on ice for a further 20min before the addition of 1ml of LB medium (containing no antibiotics). The tubes were then placed in shaking incubator at 37°C for 1h and subsequently harvested in a microcentrifuge at 8,000rpm for 1min. The resulting cell pellet was resuspended in 100µl of LB medium (containing no antibiotic) and spread onto LB agar plates containing the appropriate antibiotic for selection of colonies containing the plasmid. The plates were allowed to stand for 2min and incubated overnight at 37°C in an inverted position.

2.2.11 Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis

Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis (SDS-PAGE) was performed according to Laemmli (Laemmli, 1970). The separating gel was formed from 10-15% (w/v) acrylamide solution (acrylamide and N-N' methylene bisacrylamide in a ratio of 37:1 (w/w)). The gels also contained 375mM Tris-HCl buffer, pH8.8, 0.1% (w/v) SDS, 0.225% (w/v) ammonium persulphate, 13.2mM TEMED and water so that the final volume was 10ml. The stacking gel contained 5% (w/v) acrylamide, 0.14% (w/v) N-N' methylene bisacrylamide, 0.1% (w/v) SDS in

125mM Tris-HCl buffer, pH6.8, 0.45% (w/v) ammonium persulphate, 13.2mM TEMED and was made to a total volume of 5ml. The electrolyte running buffer contained 25mM Tris-HCl, pH8.3, 0.1% (w/v) SDS and 192mM glycine.

The 5x protein sample buffer consisted of 0.2% SDS, 8% glycerol, 0.003% (w/v) Bromophenol Blue and freshly added 2-hydroxyethylmercaptan (β-mercaptoethanol) (10% v/v) in 80mM Tris-HCl buffer, pH6.8. Gels were run at 200V and stained using Coomassie Blue solution (0.03% w/v in 50% v/v methanol and 10% v/v acetic acid) for 1–2h and destained in a wash solution containing methanol (25% v/v) and acetic acid (7% v/v). Pre-stained protein markers were subjected to simultaneous electrophoresis with the samples in order to determine molecular weights of samples.

2.2.12 Analytical Scale Expression in *E. coli* and Sample Preparation of Recombinant Protein

Cultures (10ml) of *E. coli* cells harbouring plasmids for expression were established by inoculation with a single colony picked from an overnight plate. The cultures were incubated at 37°C overnight, in LB medium containing antibiotics for the selection of the plasmid. An aliquot (0.5ml) from each culture was added to fresh medium (10ml). The cultures were incubated at 37°C or 42°C (200rpm) until a D_{600} of 0.4 was attained. Upon reaching the required D_{600}, protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG), to a final concentration of either 0.2mM, 0.4mM or 1mM. Samples (1ml) were removed at appropriate time intervals (up to 3h) and their D_{600} was measured. The samples were collected in a microcentrifuge at 14,000 x g for 3min and the cell pellet was resuspended in sterile

TE (10mM Tris-HCl pH8.0, 1mM EDTA) buffer to give a constant D_{600} of 10. Samples were subsequently stored at –20°C, prior to analysis by SDS-PAGE and Western blotting (Sections 2.2.11 and 2.2.15).

2.2.13 Analytical Scale Production of Soluble and Insoluble Fractions

Samples (1ml) of each *E. coli* culture to be analysed were collected in a microcentrifuge (10,000 x g, 3min) and the supernatant was discarded. Each cell pellet was resuspended in TE buffer (10mM Tris-HCl pH8.0, 1mM EDTA) to give an D_{600} of 10. A 100µl sample from each cell resuspension solution was repeatedly freeze-thawed (3X) and sonicated with a microtip on ice for 3 x 10s (Soniprobe, Output=7, Duty cycle=70%, Lucas Dawe Ultrasonics Ltd., London, UK). After sonication, insoluble material was collected in a microcentrifuge (14000 x g) for 3min. The pellets were resuspended in 100µl of 1x SDS sample buffer (containing 2% (v/v) β-mercaptoethanol). The supernatant was retained and 1/5th volume of 5x SDS sample buffer (containing 10% (v/v) β-mercaptoethanol) was added, this was the soluble fraction. Soluble and insoluble fractions were then boiled before examination by SDS-PAGE.

2.2.14 Preparation of Cell Pellets and Culture Media for SDS-PAGE Analysis

Cell pellets were thawed on ice and resuspended in 1x SDS sample buffer to achieve a D_{600} of 10. All samples were analysed using SDS-PAGE and Western blotting (Sections 2.2.11 and 2.2.15). Culture media (20µl) was mixed with 5x SDS protein sample buffer (4µl), containing 10% (v/v) β-mercaptoethanol.

2.2.15 Immunological Detection of Recombinant Proteins by Western Blotting

Following SDS polyacrylamide gel electrophoresis, the resolved proteins were transferred to a moistened nitrocellulose membrane measuring 60mm x 80mm (Amersham International plc, Bucks, UK). Prior to blotting, five rectangular pieces of 3MM filter paper (Whatman International Ltd., Maidstone, UK), measuring 60mm x 80mm were cut. Two pieces were soaked in Anode Solution 1 (300mM Tris-HCl, pH10.4, 20% methanol) and placed on to the anode of the semi-dry blotter (Sartorius Ltd., Epsom, Surrey, UK); a third was soaked in Anode Solution 2 (25mM Tris-HCl, pH10.4, 20% methanol) and placed on to the first two pieces of filter paper; the nitrocellulose membrane was placed upon these three pieces on the anode, followed by the SDS-PAGE gel; and the remaining two pieces of filter paper that had been soaked in Cathode Solution (25mM Tris-HCl, 20% methanol, 40mM Caproic Acid) were placed on top of the gel, before putting the cathode in place.

Resolved proteins were transferred to the nitrocellulose membrane by passing a current of 145mA through the semi-dry blotter for 45min. On completion of transfer, the membrane was washed for 10min in 10ml TBS (20mM Tris-HCl, pH8·0, 50mM NaCl), this was repeated three times. The membrane was blocked for 1h at room temperature using TBS containing 3% (w/v) bovine serum albumin. All subsequent steps were performed in TBS Tween Buffer (TBS with 0.05% (v/v) Tween20). The membrane was washed 3 x 5min before being shaken for 1h at room temperature in 10ml of TBS Tween containing the first antibody. The membrane was washed three times for 5min and agitated for 1h with TBS Tween containing the secondary antibody - an alkaline phosphatase conjugate (Sigma Chemical Company, Poole, UK) applied at a dilution of 1/5000. As before, the membrane was washed 3 x 5min in

TBS Tween before being stained with Nitro blue tetrazolium and 5-Bromo-4-chloro-3-indolyl phosphate (Bio-Rad Laboratories, Watford, Herts, UK) with continuous shaking until a colour reaction developed (usually within 5min). Immersing the developed membrane in distilled water stopped the staining reaction and the membrane was allowed to air-dry.

In expressed proteins that contained a His-tag a Nickel-alkaline phosphatase conjugate was added to the solution instead of the first antibody. Therefore, the addition of a secondary antibody was unnecessary and the blot was washed in TBS/Tween before being developed as described above.

2.2.16 Refolding of insoluble recombinant protein

Insoluble material containing the recombinant protein was collected by centrifugation at 16,000 x g for 40min. The pellet was resuspended in 200ml of buffer A (0.1M Tris-HCl, pH11.0, 50mM β-mercaptoethanol) and stirred at 4°C for 4h. Insoluble material was washed in this buffer and recollected by centrifugation. The washed pellet was resuspended in 10ml of buffer B (6M urea, 0.1M Tris-HCl, pH8.0, 1mM glycine, 1mM EDTA, 50mM β-mercaptoethanol) and stirred overnight in an attempt to resolubilise any recombinant material. Residual insoluble material was removed by centrifugation at 28,000 x g for 2h. The supernatant containing recombinant protein was placed in dialysis tubing and placed in 4l of buffer C (4M urea, 0.1M Tris-HCl, pH8.0, 1mM glycine, 1mM EDTA, 50mM β-mercaptoethanol) at 4°C for 4h with stirring. This was transferred to 4l of buffer D (2M urea, 0.1M Tris-HCl, pH8.0, 1mM glycine, 1mM EDTA, 50mM β-mercaptoethanol) at 4°C for 4h with stirring. Finally,

the remaining solution was placed in fresh dialysis tubing and transferred to 4l of 10mM Tris-HCl, pH8.5).

2.2.17 Screening by Hybridisation with Radioactive-labelled Probes

A colony blotting method was used to screen a plasmid library for colonies containing genes of interest.

A cDNA library (1.0µl) was transformed into DH5α cells (200µl), plated out on 8 LB-ampicillin plates and incubated overnight to give a good distribution of distinct colonies. Hybond-C nitrocellulose was carefully lowered onto each plate and was left for 3min. The orientation of the nitrocellulose to the plate was marked by piercing through the nitrocellulose into the agar, with a syringe needle filled with Indian ink. Some of each colony should stick to the nitrocellulose, which was carefully lifted from the plate and placed (colony side up) on a series of 3MM (Whatman International Ltd., Maidstone, UK) filter papers, soaked in the following solutions;

- Blotting solution: - 10% SDS for 3min.
- Denaturing solution: - 1.5M NaCl, 0.5M NaOH for 5min.
- Neutralising Solution: - 1.5M NaCl, 1.0M Tris-HCl, pH7.4 for 5min.
- 2x SSC for 5min.

The nitrocellulose blots were then dried for at least 30min on 3MM filter paper. Subsequently, blots were covered with dry filter paper, wrapped in tin foil and baked at 80°C for 2h. Blots were stored overnight at 4°C.

Dry nitrocellulose blots were moistened on filter paper covered with 2x SSC for 10min, before being transferred to pre-wash solution (5x SSC, 0.5% SDS, 1mM EDTA) for 10min. They were then wrapped in gauze and fitted into a hybridisation tube, before 25ml of pre-hybridisation solution (6x SSC, 5x Denhardt's solution (0.1% Ficoll, 0.1% polyvinylpyrilidone and 0.1% bovine serum albumin), 10% SDS, 20ng/ml of yeast tRNA) was added. The tube was left to rotate slowly at 55°C for 2h.

Gene specific oligonucleotides (see above) were labelled with ^{32}P in the following reaction;

2μl (50μg) of each single-stranded oligonucleotide

5μl of [γ-^{32}P]-ATP (Amersham)

2μl 10x Polynucleotide Kinase Buffer

1μl Polynucleotide Kinase

4μl nuclease-free H_2O

The reaction was mixed by pipetting and incubated at 37°C for 45min. Primers were then precipitated with ethanol and collected by centrifugation. The supernatant was discarded and the oligonucleotides were resuspended in 50μl of nuclease-free H_2O before addition to the hybridisation tubes containing the nitrocellulose blots and pre-hybridising solution. Hybridisation was left to proceed overnight at 55°C.

Nitrocellulose blots were carefully removed from the hybridisation tubes and washed with Washing solution (2x SSC, 0.1% SDS). Radiation levels on the blots were monitored using a Geiger counter. After three washes, the blots had radiation levels just above background, suggesting that most of the non-specifically bound probe had been washed away. Blots were air dried, wrapped in Saran wrap and placed in an AutoRad cassette (Genetic Research Instrumentation Ltd., Dunmow, UK), containing radiographic film (Kodak Ready pack, Kodak, New York, USA). The cassettes were stored overnight at -80°C to increase the sensitivity of the radiographic film. The film was removed from the cassette in the dark room and developed using a Gevamatic 60 automatic developer (AGFA-Gevaert, Brentford, UK).

Spots on the film were marked on the nitrocellulose blot by piercing with a syringe needle. The blot was then aligned with the original agar plate and colonies in the proximity were chosen for culturing, plasmid DNA purification and sequencing.

2.2.18 Single-stranded RNA synthesis

Template for single-stranded RNA (ssRNA) synthesis was prepared from pGEM-T plasmids containing pfFabF or LmKAS1/2 genes, within the multiple cloning site. These plasmids were transformed into *E. coli* DH5α cells (Section 2.2.10), plated onto fresh LB agar plates containing ampicillin and incubated at 37°C overnight. Single colonies were picked from these plates and used to inoculate 10ml of LB broth containing ampicillin. These cultures were grown overnight at 37°C in a shaking incubator (250rpm). Overnight cultures (4ml) were used to inoculate 400ml of LB broth containing ampicillin, which was grown at 37°C in a shaking incubator until a

D_{600} of 0.8. Microgram quantities of plasmid DNA were collected from these cultures using the DNA Purification Maxiprep Kit (Promega), according to the manufacturer's instructions.

Plasmid preparations were digested using the restriction enzyme *Pvu* II according to the manufacturer's instructions (New England Biolabs). This enzyme cuts approximately 100 bases upstream of both the T7 and SP6 promoter regions of the pGEM-T plasmid and has no cognate sites within the inserted genes. The digested plasmid was separated in a 1% agarose gel and visualised using ethidium bromide. DNA bands containing the pfFabF or LmKAS1/2 flanked by the T7 and SP6 promoter regions were excised from the gel and purified using the QiaQuick Gel Extraction Kit (Qiagen Ltd., West Sussex, UK). The concentration of these DNA fragments was analysed using a GenequantPro spectrophotometer (Pharmacia).

Single-strand RNA was synthesised from these fragments derived from the pGEM-T plasmid clones. In separate reactions, T7 and SP6 RNA polymerases (Ambion) were used to synthesise RNA from the respective promoter regions.

The reaction mixture is shown below;

Plasmid fragment (200ng/µl)	5µl
10mM NTPs (ATP, CTP, GTP, UTP)	8µl
RNA Polymerase Buffer (10x)	2µl
T7 or SP6 RNA Polymerase (Ambion)	1µl
DEPC treated H_2O	4µl
Total	20µl

The reaction was incubated at 37°C for 60min after which template DNA was digested by the addition of 1 unit of DNase I (Ambion) at 37°C for a further 15min. Production of ssRNA and template DNA degradation was analysed by running samples in a 1% agarose/ethidium bromide gel, visualised under UV light. The concentrations of RNA in these ssRNA products were analysed using a GenequantPro spectrophotometer (Novagen).

2.2.19 Double-stranded RNA production

Equal amounts of T7 and SP6 ssRNA were annealed by heating to 80°C and cooling slowly to 20°C over 4h. The concentration of ssRNA products was analysed (see above) and added in equimolar amounts to a 0.25ml eppendorf tube. In an attempt to optimise the annealing reaction, ssRNA products were purified using the RNAeasy Kit (Qiagen Ltd., West Sussex, UK) and resuspended in DEPC treated H_2O or 1x RNA polymerase buffer.

2.2.20 Design of short interfering RNA (siRNA) Oligomers

Short interfering RNA oligomers were designed using the protocol previously described by Elbashir. Briefly, only the coding sequence of each gene was used in the design process. The first 100 nucleotides downstream of the start codon were ignored and the remaining coding sequence was searched for regions containing 5'-AA(N_{19})UU-3', where N is any nucleotide. The G/C content of these potential siRNA sequences was calculated, those with a G/C content of 50% were chosen as prime candidates. Finally, these siRNA sequences were used in Blast searches against

both parasite and human genomes, to try to ensure that only a single gene is targeted. Lyophilised HPP grade siRNA oligos (40nmole) were ordered from Xeragon Inc. (www.xeragon.com) and stored at –20°C.

2.2.21 Preparation of Cell Extracts for Radioactive Assays

BL21 pLysS cells were grown in LB broth containing chloramphenicol. When the culture reached late exponential phase, cells were harvested by centrifugation at 9,000g, 4°C for 10min in the Beckman JLA 16.250 rotor. Cells were washed three times by resuspending the pellet in 10ml of lysis buffer (100mM Sodium phosphate, pH7.0, 5mM β-mercaptoethanol, 1mM EDTA) followed by centrifugation at 9,000g, 4°C for 5min in the Beckman JA-20 rotor. Following the third washing step, the cell pellet was resuspended in twice its wet weight of lysis buffer. The cell resuspension was sonicated with a MaxiTip probe on ice for 4 x 10s (Soniprobe, Output=7, Duty cycle=40%, Lucas Dawe Ultrasonics Ltd., London, UK). The cell lysate was cleared by centrifugation at 20,000g, 4°C, for 60min in the Beckman JA-20 rotor. The supernatant was then adjusted to 80% saturation with $(NH_4)_2SO_4$, left to stand for 5min and the precipitate collected by centrifugation at 20,000g at 4°C for 20min. The resultant supernatant was discarded and the pellet resuspended in 2ml of lysis buffer.

The protein concentration of this suspension was calculated using a Bradford Assay with bovine serum albumin standards ranging from 0.2-1.4mg/ml. Samples were diluted with H_2O until the absorbance lay between absorbance values corresponding to 0.2-0.8mg/ml standards.

2.2.22 Fatty Acid Synthase Activity Assays in *E. coli* Extracts

The assay was based on the method described previously Tsay *et al.*, 1992. The final concentrations of components of the assay were 1mM NADH, 1mM NADPH, 40μM acetyl-CoA, 25μM [2-^{14}C]malonyl-CoA (specific activity, 50.9μCi/μmol), 15μM *E. coli* acyl carrier protein (ACP), 1mM β-mercaptoethanol, 0.1mM Sodium Phosphate buffer pH7.0 and 25μg of *E. coli* cell extract, in a final assay of volume of 40μl. The acetyl-CoA and [2-^{14}C]malonyl-CoA components were mixed together and added to the tubes last to initiate the reaction. Reactions were performed in triplicate at 37°C for 30min and stopped by the addition of 8μl 5M KOH. In negative controls, KOH was added to tubes prior to the addition of [2-^{14}C]malonyl-CoA and acetyl-CoA, with subsequent incubation.

Reaction tubes were then heated to 65°C for 30min, cooled on ice and condensation collected in a microcentrifuge. Following acidification using 12μl of 20% (v/v) H_2SO_4 the resulting ^{14}C-labeled fatty acids were extracted with three 1ml volumes of petroleum ether. The organic extracts were pooled into glass vials, dried under N_2 (BOC gases, Surrey, UK) and resuspended in 1ml petroleum ether. Fractions of the resuspension were then added to scintillation vials, dried once more and 10ml of Opti-fluor scintillant (Packard, Groningen, Holland) was added. The incorporation of radiolabel into fatty acid chains was determined by using a Beckman 1209 Rackbeta Liquid Scintillation Counter (EG&G Wallac, Milton Keynes, UK). Samples were counted for 1min and corrected for quenching by the sample channels ratio method.

2.2.23 Conformationally Sensitive Gels

Prepared as shown below (Table 2.3)

Component	Separating Gel	
	0.5M urea, 13% acrylamide	2.5M urea, 15% acrylamide
40% acrylamide	4.9ml	5.6ml
Milli-Q H_2O	8.9ml	4.6ml
Trizma pre set pH9.0	1.4g	1.4g
8M urea	950µl	4.7ml
0.25M EDTA	60µl	60µl
TEMED	25µl	25µl
Ammonium Persulphate (100uM) (APS)	45µl	45µl

Table 2.3 Components of conformationally sensitive gels

All of the components of the separate gels, except the APS, were added to three conical flasks and the solutions degassed for 5min. The APS was then added and the gel poured and set.

Short chain acyl-ACPs were separated in 13% polyacrylamide gels containing 0.5M urea. Longer chain acyl-ACPs were separated in 15% polyacrylamide gels containing 2.5M urea. Temperature can have a significant effect on migration positions and care was taken to run these urea PAGE gels at 25°C.

2.2.24 Assay of FabB and FabF Condensing Enzymes and Inhibitor Studies

Elongating condensing enzymes from *E. coli* and *P. falciparum* were assayed for activity using a standard toluene extraction procedure (Garwin *et al.*, 1980; Price *et al.*, 2001.). To ensure complete reduction of ACP it was incubated with the DTT and Potassium phosphate buffer at 37°C for 20min. The remaining components, except the condensing enzyme, were then added to the reduced ACP mix and 16µl aliquoted into reaction tubes. Inhibitors were dissolved in 20% DMSO and 2µl added to the reaction tubes. Control reactions contained 2% DMSO. This solvent was chosen because, unlike ethanol, it had limited effect on the assay.

The reaction was started by the addition of 2µl of the condensing enzyme followed by incubation at 37°C for 20min. In assays where a lower concentration of the condensing enzyme was used, all dilutions were made in 100mM Sodium phosphate buffer, pH6.8. The final assay contained 100mM Potassium phosphate buffer, pH6.8, 100µM ACP, 300µM DTT, 1µg ecFabD, 50µM [2–^{14}C] malonyl-CoA, 50µM acyl-ACP in a 20µl volume. Addition of 400µl of reducing solution (see below) stopped the reaction. This was then vortexed vigorously for 10s and placed at 37°C for 40min. Reduced acyl chains were extracted into 400µl toluene with vigorous mixing. The upper phase was extracted and placed in scintillation vials containing 3ml of scintillation cocktail (Fisher). The decays per minute were measured using the Beckman 1209 Rackbeta Liquid Scintillation Counter (EG&G Wallac, Milton Keynes, UK).

Reducing Solution contained 100μM Potassium phosphate buffer, pH6.8, 400μM KCl, 30% tetrahydrofuran and 5mg/ml $NaBH_4$. Reducing solution was made in a fume cupboard, stirred and stored on ice prior to use.

2.2.25 Assay of FabH Enzyme Activity and Inhibitor Studies

The enzyme FabH condensing enzymes from *E. coli* and *P. falciparum* were assayed using a filter disc assay (Tsay *et al.*, 1992; Price *et al.*, 2001). The final assay contained 100mM Sodium phosphate buffer, pH6.6, 25μM ACP, 1mM β–mercaptoethanol, 1μg ecFabD, 65μM malonyl-CoA, 45μM [1-^{14}C] acetyl-CoA in a 40μl volume. To ensure complete reduction of ACP it was incubated with β–mercaptoethanol and Sodium phosphate buffer at 37°C for 30min. The remaining components were then added to the reduced ACP mix and 36μl pipetted into each reaction tube. Inhibitors were dissolved in 20% DMSO and 2μl added to the reaction tubes. Control reactions contained 2% DMSO.

The reaction was started by the addition of 2μl of FabH, followed by incubation at 37°C for 15min. To stop the reaction 35μl of the assay was spotted onto a Whatman 21mm filter disc. Filter discs were washed with gentle shaking in successive solutions of 10% trichloroacetic acid (TCA), 5% TCA and 1% TCA at 4°C for 20min, 20ml of wash solution per disc. To establish the total number of counts added to each tube, a control reaction was spotted onto a filter disc but was not washed in TCA. Filter discs were dried and placed in scintillation vials containing 3ml of scintillation cocktail (Fisher). The decays per minute were measured using the Beckman 1209 Rackbeta Liquid Scintillation Counter (EG&G Wallac, Milton Keynes, UK).

2.2.26 The overexpression and purification of Escherichia coli FabD

The pDS-His6-FabD plasmid was kindly provided by Dr. Roland Lange (Morphochem AG, Basel, Switzerland). The plasmid was transformed (Section 2.2.9) into competent *E. coli* M15 pREP4 cells prepared as described in Section 2.2.10 and plated on to LB agar plates containing ampicillin and kanamycin. A single colony was used to inoculate 10ml of LB media containing the aforementioned antibiotics and the culture was placed in a shaking incubator at 37°C for 16h. The overnight cultures (10ml) were used to inoculate 1litre of freshly prepared LB containing the appropriate antibiotics, in 2 x 1litre conical flasks. Cultures were placed in a shaking incubator at 30°C, 240rpm. When the cultures reached an $D_{600} = 0.6$, expression of *E. coli* FabD was induced by the addition of IPTG to a final concentration of 1mM. Cells were harvested at 24h post-induction, using a Beckman J series centrifuge (8,000 x g, 4°C, 15min). The resultant cell pellet was resuspended in 40ml of Lysis Buffer (50mM Sodium phosphate, pH7.5, 300mM NaCl, 10mM imidazole) per litre of culture. The cells were lysed by sonication, 60% amplitude for 3 x 30s bursts and the lysate centrifuged at 20,000 x g, 4°C for 15min. The supernatant was mixed with 5ml of Ni-NTA resin (Qiagen Ltd., West Sussex, UK), which had been equilibriated in Lysis Buffer and stirred slowly at 4°C for 1h. The resin solution was poured into a glass chromatography column and the resin was allowed to settle.

The column was washed with 20 column volumes of Lysis buffer and then with Wash Buffer (50mM Sodium phosphate, pH7.5, 300mM NaCl, 20mM imidazole) until the OD_{280} was less than 0.02. The His6-FabD protein was eluted across 20 x 1ml fractions using Elution Buffer (50mM Sodium phosphate, pH7.5, 300mM NaCl, 250mM imidazole). All fractions contained a single band of the expected size for *E. coli* FabD

and fractions 4-10 had an OD_{280} corresponding to that greater than 1mg/ml. These fractions were diluted in 50% of sterile glycerol and stored at –20°C until further use.

2.2.27 Preparation of Acyl-ACP Synthetase

An overnight culture of *E. coli* strain LCH68 was used to inoculate 1litre of fresh LB broth containing ampicillin that was grown to late log phase. All subsequent steps were performed at 4°C or on ice. Cells were collected by centrifugation at 8,000 x g for 20min and cell pellets resuspended in 2 x wet weight of 50mM Tris-HCl, pH8.0 containing 30µg/ml DNase I. Resuspended cells were ruptured in a French Press at 16,000psi and the homogenate collected in JA-20 centrifuge tubes. Cell debris and insoluble material was removed by spinning at 11,000rpm for 20min. The supernatant was adjusted to 10mM $MgCl_2$ and the membrane fraction collected at 80,000 x g for 90min.

The pellet was washed in a volume of 50mM Tris-HCl, pH8.0, 0.5M NaCl, 10mM $MgCl_2$ equal to the wet weight of the original cell pellet and the membrane fraction collected in an ultracentrifuge under the conditions described above. To help solubilise membrane proteins the pellet was resuspended in two thirds the volume used in the previous step of 50mM Tris-HCl, pH8.0, 2% Triton X-100, 10mM $MgCl_2$. The solution was spun in the ultracentrifuge for a final time and the supernatant retained for the purification of Acyl-ACP Synthetase.

A column packed with Blue Sepharose CL-6B was equilibrated with two column volumes of 50mM Tris-HCl, pH8.0, 2% Triton X-100. The supernatant was applied to

the column and the elutant collected. The column was then washed with 15ml 50mM Tris-HCl, pH8.0, 2% Triton X-100, 0.6M NaCl. Then the AAS protein was eluted using 50mM Tris-HCl, pH8.0, 2% Triton X-100, 0.5M KCSN and 1ml fractions collected for the analysis of activity. Each fraction was immediately dialysed against 2-4litres of 50mM Tris-HCl, pH8.0, 2% Triton X-100 to remove the potassium thiocyanate for a minimum of 4h. Each fraction was transferred to an eppendorf tube where ATP and $MgCl_2$ were added to final concentrations of 5mM.

The protein concentration of each fraction was determined using a Bradford Assay and fractions containing significant amounts of protein were assayed for AAS activity.

2.2.28 Extraction of Lipids and Production of Fatty Acid Methyl Esters (FAMEs) from *E. coli*

Cells were harvested by centrifugation at 8,000 x g for 10min at 4°C. The cell pellet was resuspended in 500µl of distilled water and transferred to glass test tubes. Chloroform/methanol (1:2), 2.5ml, was added to the cell resuspension, which was vortexed and left at room temperature for 15min. Then 1.25ml of chloroform and 1.25ml of Garbus reagent (2M KCl, 0.5M Sodium phosphate buffer, pH7.4) were added to the solution, which was vortexed. To ensure complete separation of the phases the test tubes were centrifuged for 5min at 1,500rpm (Baird Taitlock centrifuge). The upper aqueous phase was discarded and the lower phase was transferred to a clean test tube, where it was evaporated to dryness under N_2. The resulting precipitate was resuspended in 200µl of chloroform and 100µl of this was

removed to a glass capped vial and stored at -20°C for future studies. The remaining 100µl was placed in a methylating tube and evaporated to dryness under N_2. To produce fatty acid methyl esters 1ml of toluene and 2ml H_2SO_4 in dry methanol were added to the methylating tubes. The tightly capped methylating tubes were placed in a 70°C waterbath for 2h to ensure complete methylation of fatty acids. To stop the methylating reaction 4ml of 5% (w/v) NaCl was added to the methylating tubes. To extract the FAMEs 4ml of petroleum ether was added to the methylating tube, which was then vortexed and the upper phase removed to a clean test tube. The extraction was repeated and the upper phases pooled. To neutralise any remaining sulphuric acid, 4ml of 2% $KHCO_3$ was added to each tube and the upper phase removed to a clean test tube. HPLC grade sodium sulphate (0.5g) was added to absorb any water present in the sample and the solution was transferred to clean test tube. The solution was then evaporated to dryness under N_2 and resuspended in 25µl of HPLC grade hexane.

Gas chromatography of FAMEs were performed using an Agilent model 6890N series gas chromatograph equipped with splitless injector, autosampler and flame ionisation detector. A capillary column of fused silica coated with VF-23ms™ (0.25mm x 30m in length, 0.25µm film thickness; Varian) was used and run isothermally at 170°C. Helium was the carrier gas at an initial flow rate of 1.6ml/min. In all analyses, the detector was set at 300°C and the injector at 230°C.

2.2.29 Determination of minimum inhibitory concentrations (MIC)

The MICs of the TLM analogs against *E. coli* was determined by a broth microdilution method. ANS1 (*tolC*::Tn*10*, *metB1*, *relA1*, *spoT1*, λ^R, λ^-, *gyrA216*, F^-) is an *E. coli* strain that is defective in TolC-dependent type I secretion and efflux pumps (Buchanan, 2001). Strain ANS1 was grown to mid-log phase in 1% tryptone broth and then diluted 30,000 fold in the same medium. A 10µl aliquot of the diluted cell suspension (3,000 to 5,000 colony forming units) was used to inoculate each well of the 96 well plate (U-bottom with low evaporation lid) containing 100µl of tryptone broth with the indicated concentration of inhibitors. The plate was incubated at 37°C for 20 hours before being read by Fusion™ Universal Microplate Analyzer (Packard, Canada) at 600nm. The absorbance in wells treated with DMSO solvent control was considered as 100%.

Chapter 3

The Cloning and Expression of *P. falciparum* FabF

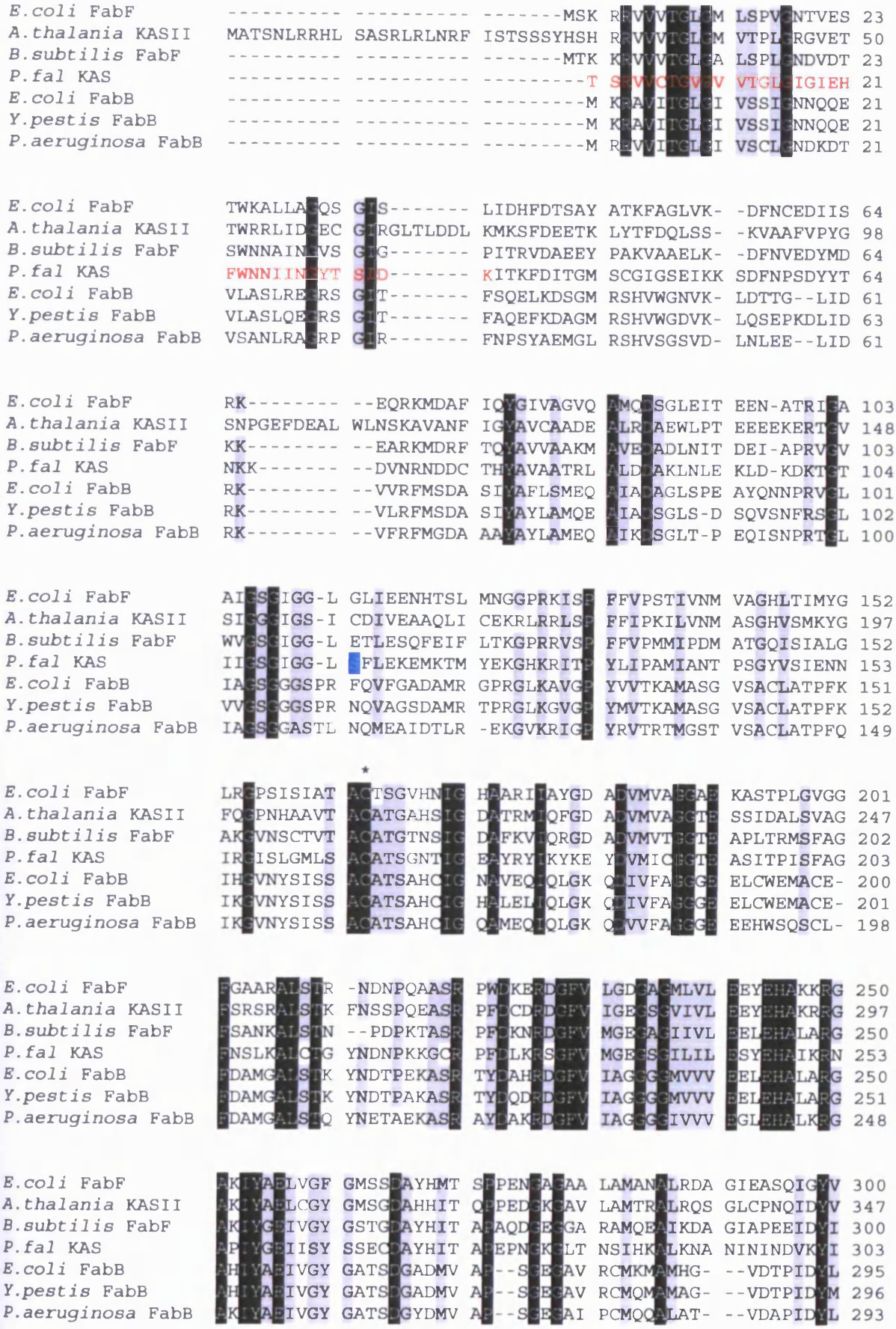
3.1 Introduction

In studies described in this chapter, a candidate gene that encodes a condensing enzyme from *Plasmodium falciparum* was identified by its sequence homology to plant and bacterial FabB and FabF protein sequences. Phylogenetic analysis of the translated protein suggested that it was more closely related to FabF enzymes than FabB enzymes and the protein was assigned as pfFabF. The pfFabF protein is encoded by a 1257 base pair gene, assigned as *pffabF*, on chromosome 6 of the *P. falciparum* genome. When translated, this gene encodes a protein of 418 amino acids that corresponds to the mature part of the enzyme. Our studies failed to locate the initiator ATG codon for the gene and any N-terminal signal sequence of this protein. Isolation of naturally occurring pfFabF protein was not considered a feasible method of yielding a sufficient quantity of protein for kinetic characterisation or structural analysis. Therefore, production of pfFabF was attempted using a recombinant system. In this study, the region corresponding to the mature pfFabF has successfully been produced in a recombinant form in *E. coli*. Initial protein expression studies showed pfFabF was contained solely within the insoluble fraction of lysed *E. coli* preparations. In an attempt to solubilise the protein, an N-terminal fusion to *E. coli* maltose binding protein was expressed. There was a limited but significant increase in the amount of soluble pfFabF protein and this was sufficient to yield enough active enzyme for initial characterisation of the enzyme and to allow assays with thiolactomycin and its derivatives.

3.2 Identification of a *P. falciparum* Elongating Condensing Enzyme

At the beginning of this project, genes encoding *P. falciparum* FabB and/or FabF proteins had yet to be identified. The tBlastN program (Altschul *et al.*, 1997) was used to search databases of *P. falciparum* contiguous DNA sequences with plant and bacterial FabB and FabF protein sequences. Within a 1662bp contiguous sequence located on chromosome 6 of the *P. falciparum* genome, initial studies identified a single candidate gene of 1152bp that included a stop codon at the 3' terminus. The deduced gene sequence encoded a protein of 383 amino acids that shared significant sequence identity with an assortment of plant and bacterial FabB and FabF enzymes (Figure 3.1 – shown as black text). However, the predicted protein sequence of this candidate gene was shorter at the N-terminus than mature proteins of the family and there was no obvious initiator ATG codon (Figure 3.1). In addition, for a protein believed to be targeted in the apicoplast, an N-terminal bipartite sequence would be expected for targeting it to the organelle.

Within the contiguous sequence, all three reading frames were searched upstream of the identified gene for an N-terminal exon encoding either an initiator methionine or a signal peptide. A 105bp extension encoding 35 amino acids that showed high sequence homology to bacterial FabB and FabF proteins was observed in a separate reading frame (Figure 3.1 – shown as red text). This indicated that a frameshift mutation may have been introduced during sequencing of the contiguous fragment at the Sanger Institute (responsible for sequencing *P. falciparum* Chromosome 6). This region also lacked an obvious initiator ATG codon and did not encode any signal and target sequences for transport of the mature protein to the apicoplast. Due to problems



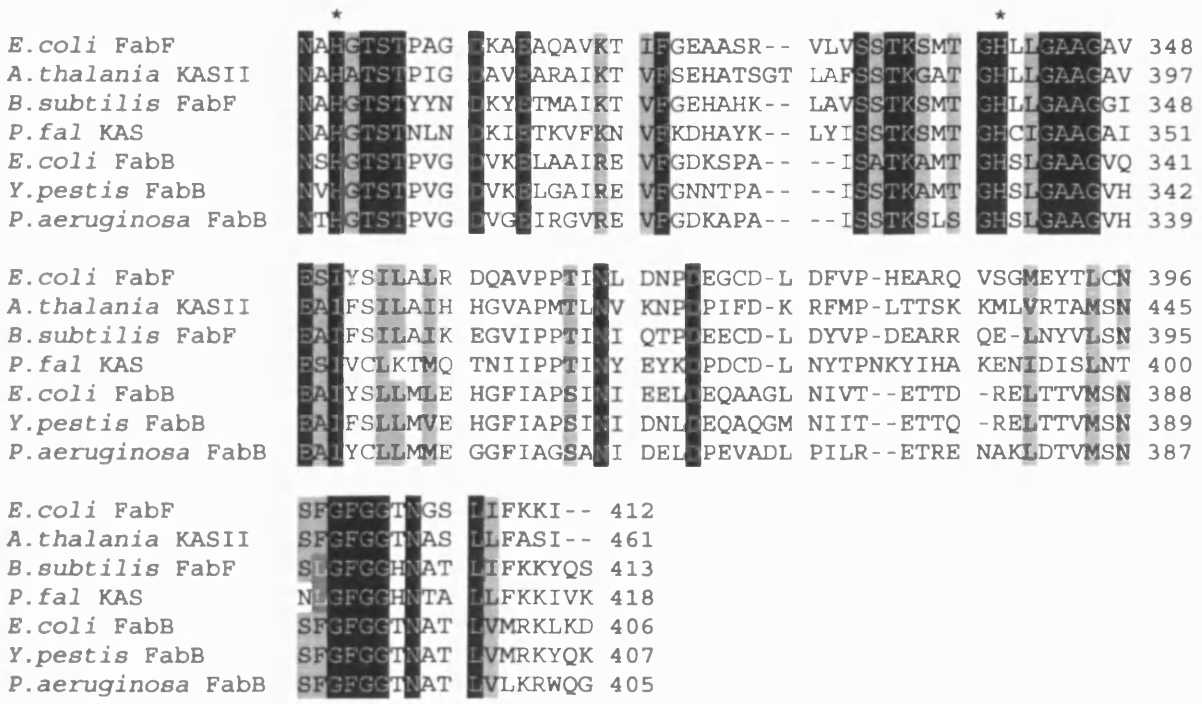


Figure 3.1 An alignment of bacterial, plant and *P. falciparum* condensing enzymes

The region of the mature pfFabF protein that was identified and cloned in Sections 3.4-3.5 is shown as black text. The N-terminal extension of 35 amino acids, cloned and assembled in Section 3.6-3.7 is shown as red text. The R114S mutation from the published sequence of pfFabF (PlasmoDB MAL6P1.165) as a result of a polymorphism is highlighted against a blue background. The L390 residue that may be involved in TLM resistance in *B. subtilis* is highlighted against a background of red. Residues of the catalytic triad are highlighted using *. The species names in full and the GenBank accession numbers of the sequences are: *Arabadopsis thalania* KAS II (NP_178533.2); *Bacillus subtilis* FabF (CAB12975); *Escherichia coli* FabB (P14926); *Escherichia coli* FabF (P39435); *Pseudomonas aeruginosa* FabB (AAC45620); *Yersinia pestis* FabB (CAC92996); *Plasmodium falciparum* KAS – this work and (NP_703921.1; PlasmoDB # MAL6P1.165). This alignment was produced using the ClustalW program (Thompson *et al.*, 1994), which is available at www.ebi.ac.uk and formatted using the Chroma program (Goodstadt and Ponting, 2001).

in amplifying the *pffabF* gene (Sections 3.3 and 3.4) it was decided not to clone this region (Section 3.6) until the rest of the gene was assembled (Section 3.5).

Homology searches with FabB and FabF proteins were performed to determine whether the candidate *P. falciparum* gene encoded a FabB or FabF enzyme. Due to the high sequence identity that usually exists between FabB and FabF proteins (37% between the two *E. coli* enzymes), it was difficult to come to an accurate conclusion as to which the *P. falciparum* protein resembled. Traditionally, FabB and FabF proteins have been characterised not by their sequence homology alone, but by their catalytic activity towards acyl-ACP substrates.

Like most plants and bacteria, *P. falciparum* contains a FabZ homologue but lacks a FabA homologue. The lack of a FabA dehydratase/isomerase within *P. falciparum* suggests that the condensing enzyme that was provisionally identified here, may belong to the FabF family (Section 1.8.2). *E. coli* FabF controls temperature-dependent regulation of fatty acid composition by elongating palmitoleic acid (C16:1) to *cis*-vaccenic acid (C18:1). FabF proteins are found in Type II FAS systems throughout the plant and bacterial kingdoms, whilst FabB proteins are limited to the Gram-negative bacteria.

3.2.1 Phylogenetic Analysis of the *P. falciparum* Condensing Enzyme

The amino acid sequence of the condensing enzyme from *P. falciparum* is closely related to both FabB and FabF proteins, thus hindering a clear classification of the protein. However, comparisons of condensing enzyme sequences from different organisms should elucidate the evolutionary relationships between these proteins.

Therefore, phylogenetic analysis of condensing enzyme amino acid sequences was performed using the CLUSTALW program (http://www.ebi.ac.uk /clustalw) and visualised using the Treeview program (Page, 1996). Sequences of; bacterial FabB proteins, bacterial FabF proteins, plant condensing enzymes, human and mouse putative mitochondrial Type II KAS and the KAS domains from human and mouse Type I FAS were used in the phylogenetic analysis.

The resulting tree (Figure 3.2) places the *P. falciparum* condensing enzyme within a bacterial FabF clade of proteins. The condensing enzymes the two apicomplexan parasites (*P. falciparum* and *Toxoplasma gondii*) appeared to be most closely related to the FabF protein from *Bacillus subtilis* within a clade that also contains the green algae *Synechocystis sp.*. The theoretical evidence discussed in this section suggests that the elongating condensing enzyme in *P. falciparum* is a FabF type enzyme. However, knowledge of enzyme substrate specificity and complementation studies of FabF knockouts is essential to classify definitively this enzyme. Until these data become available it was decided to designate the *P. falciparum* condensing enzyme pfFabF.

3.3 Cloning of a Truncated *pffabF* gene

Amplification of the gene encoding the N-terminal truncated form of *pffabF* was first attempted in a PCR using the cycling parameters and primers shown in Figure 3.3. A *P. falciparum* cDNA library (ring stage; provided by Prof. Steve Walsh, Liverpool School of Tropical Medicine) was used as the template. The forward primer, called *pffabF* F1 (Figure 3.3B), was designed to anneal to coding strand of the 5' region of

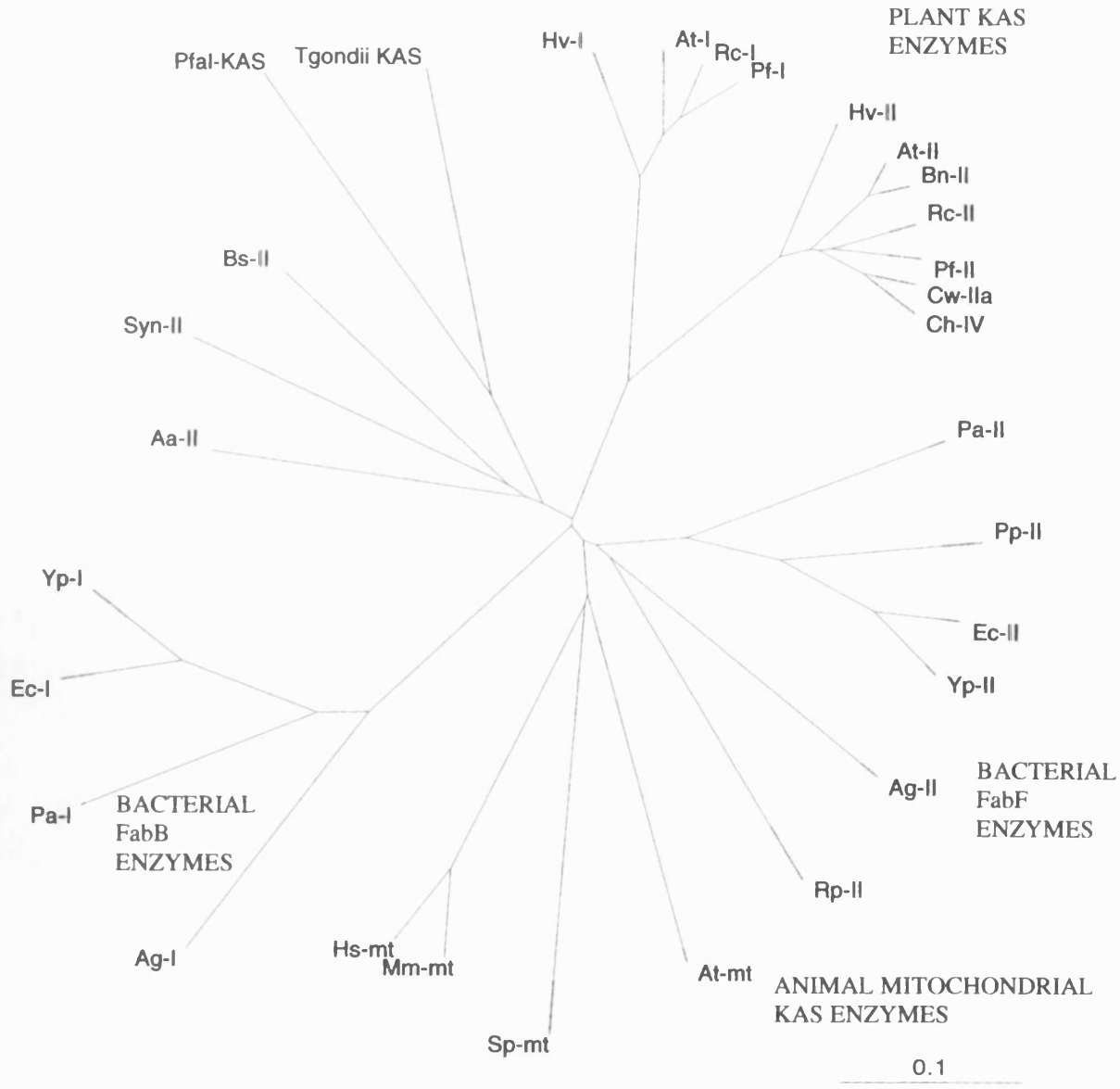
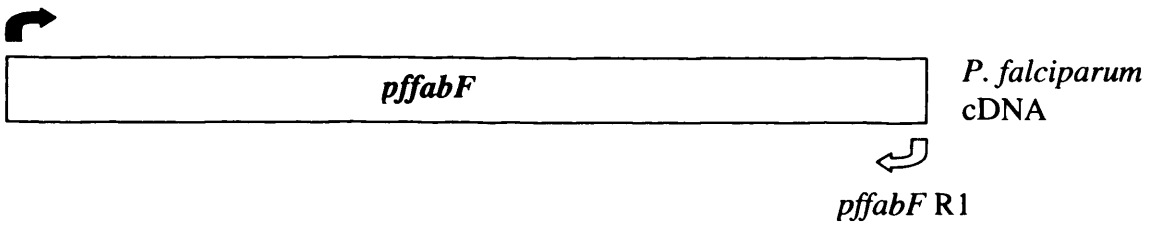


Figure 3.2 Phylogenetic tree of KAS enzymes based on their amino acid sequences.

The amino acid sequence of the mature *P. falciparum* KAS protein (416 amino acids) was aligned with the mature sequences of bacterial and plastid KAS enzymes, and mitochondria KAS proteins, using the CLUSTAL W (Thompson *et al.*, 2001) and Treeview programs (Page, *et al.*, 1997). Only the sequences of mature proteins were used in the analysis, N-terminal extensions that may encode transit and signal peptides were excluded. The full length species names for the abbreviated forms and GenBank accession numbers for the sequences are: *Agrobacterium tumefaciens*, Ag-I (AAL41174) and Ag-II (AAK86903); *Aquifex aeolicus*, Aa-II (AAC07574); *Arabadopsis thalania*, At-I (AAC49118), At-II (NP_178533.2), At-mt (BAB91181); *Bacillus subtilis*, Bs-II(CAB12975); *Brassica napus*, Bn-II(AAF61737); *Cuphea hookeriana*, Ch-IV (AAC68861); *Cuphea wrightti*, Cw-IIa (AAB37270); *Escherichia coli*, Ec-I (P14926) and Ec-II(P39435); *Homo sapiens*, Hs-mt (BAA91286); *Hordeum vulgare*, Hv-I (AAA58525) and Hv-IIa (CAA84022); *Mus musculus*, Ms-mt(BAB30490); *Pseudomonas aeruginosa*, Pa-I (AAC45620) and Pa-II (H83275); *Perilla frutescens*, Pf-I (AAC04691) and Pf-II (AAC04692); *Photbacterium profundum*, Pp-II (AAF04118); *Ricinus communis*, Rc-I (AAA33873) and Rc-II (AAA33872); *Rickettsia prowazekii*, Rp-II (CAA15192); *Schizosaccharomyces pombe*, Sp-mt (CAA21898); *Synechocystis sp.*, Syn-II (S75529); *Toxoplasma gondii*, Tgondii KAS (this work and TgTwinScan_0386, www.toxodb.org); *Yersinia pestis*, Yp-I (CAC92996) and Yp-II (NP_405182.1); *Plasmodium falciparum*, PfalKAS – this work and (NP_703921.1; PlasmoDB # MAL6P1.165).

A



B

***pffabF* F1**	5'-*GGATCC*AATTACAAAATTTGATATAACCGG-3'
***pffabF* R1**	5'-*GGATCC*ATCACTTTACAATTTTTTTGAAAA-3'

C

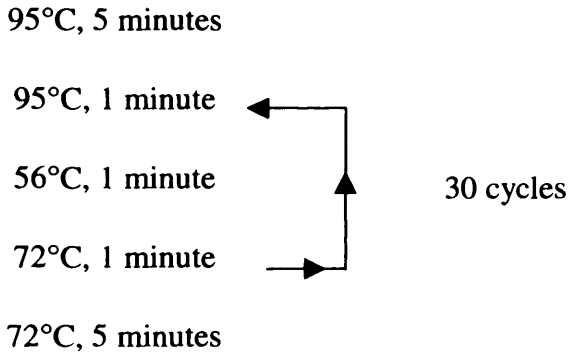


Figure 3.3 Primers and cycling parameters used in the initial attempt to amplify the *pffabf* gene

A The position of the oligonucleotide primers in relation to the truncated *pffabF* gene

B Sequences of the oligonucleotide primers. *Bam* HI sites are shown in italics.

Nucleotides complementary to the *pffabF* gene are underlined.

C Cycling parameters used in the initial attempt to amplify the *pffabF* gene

the truncated gene (Appendix 1). A reverse primer, called *pffabF* R1 (Figure 3.3B), was complementary to the non-coding strand of the 3' region and spanned the naturally occurring stop codon. The primer oligonucleotides were flanked by *Bam* HI restriction sites to facilitate subsequent subcloning of the PCR generated fragment.

Initial PCRs were performed using a wide range of conditions to amplify the entire sequence of the identified *pffabF* gene. Alterations in; annealing temperature, $MgCl_2$ concentrations, different types of DNA polymerase, the number of cycles and the length of individual cycle steps were unsuccessful in amplifying the full-length gene. This problem may be explained by the exceptionally high (~80%) A + T richness in the *P. falciparum* genome. Within the candidate *pffabF* gene, the overall A + T bias was found to be 70.1%, with some stretches of more than 100 nucleotides with an A + T bias up to 80%. This may cause the DNA polymerase to stall and amplification of the gene may be aborted, leading to insufficient product for the subsequent purification and cloning steps.

Amplification of the target sequence was not achieved under the conditions described above and it was apparent that a new cloning strategy was required. To improve the chances of amplifying sufficient product for cloning, the size of the amplicon was reduced. By reducing the size of the amplicon the chance that the DNA polymerase will extend to the end of the target sequence is increased. Reducing the amplicon size and yet still obtaining the desired protein coding sequence can be achieved by producing overlapping fragments of the gene that can be assembled to the full length product. This is normally achieved by designing oligonucleotide primers complementary to a region part way through the sequence to be amplified. These

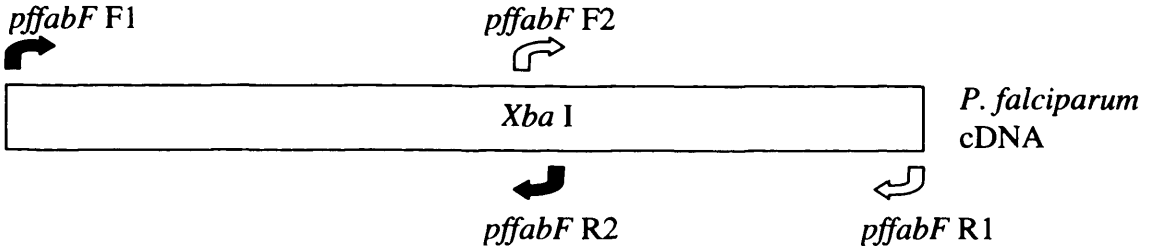
primers are then used in two separate reactions with primers complementary to the 5' and 3' regions of the target gene. This generates two fragments that can then be assembled to yield a full-length gene.

3.4 Adoption of a New Cloning Strategy

To increase the chance of amplifying the truncated *pffabF* gene, the target sequence was produced in two fragments. To facilitate the joining of the two fragments, the primers were designed to flank a naturally occurring *Xba* I site, approximately 620bp into the *pffabF* gene (Figure 3.4A). The first of these primers was termed *pffabF* R2 (Figure 3.4B), which complemented the non-coding strand in this region. It was used in conjunction with *pffabF* F1 (Figure 3.4B) to produce a 639bp fragment called *pffabF* front. The second of these primers, *pffabF* F2, complemented the coding strand around the *Xba* I site. It was used in conjunction with *pffabF* R1 to produce a 548bp fragment called *pffabF* back. The two primers flanking the *Xba* I site are not exact complements of each other (Appendix 1 showing the sequence of pfFabF with primers highlighted). They were designed so that the 3' end of the primer terminates with two C or G bases, the three hydrogen bonds formed between these bases and the template should improve the annealing efficiency. Each pair of primers was used in separate PCR reactions using identical cycling parameters. Expand High Fidelity *Taq* (Roche Molecular Biochemicals, Mannheim, Germany) was used as the DNA polymerase and *P. falciparum* cDNA library (as described previously) as the template.

Expand High Fidelity *Taq* does not add an overhang of adenosine residues required for the ligation into the pGEM-T vector. Consequently, 5 cycles with generic *Taq*

A



B

***pffabF* F2**	5'-CATCTTAAT*TCTAGA*ATCATACG -3'
***pffabF* R2**	5'-TATGAT*TCTAGA*ATTAAGATGCC -3'

C

95°C, 5 minutes

95°C, 1 minute

42°C, 1 minute

68°C, 2 minutes

40 cycles

68°C, 10 minutes

10°C, until removed from PCR block

Figure 3.4 Primers and cycling parameters used to amplify the *pffabF* gene in two segments

A The position of the oligonucleotide primers in relation to the truncated *pffabF* gene

B Sequences of the oligonucleotide primers. *Xba* I sites are shown in italics.

C Cycling parameters used to amplify the truncated ***pffabF*** gene in two fragments

polymerase (Sigma Chemicals Ltd., Poole, Dorset, UK) were performed to add adenosine overhangs. Following this second PCR, the amplified fragments were separated on an agarose gel, purified from the gel and ligated into the pGEM-T vector (Sections 2.2.2, 2.2.6 and 2.2.7). The ligation mix was transformed into competent *E. coli* DH5α cells as described in Section 2.2.10 and plated onto LB agar plates containing ampicillin. Colonies on the plate were screened using a PCR with either *pffabF* F1 and *pffabF* R2 or *pffabF* F2 and *pffabF* R1 under the conditions described in Figure 3.4C. Restriction enzyme digests with *Bam* HI, *Xba* I and a double digest with both of these restriction enzymes further confirmed the presence of the target insert (results not shown).

The orientation of the fragments with the pGEM-T vector was elucidated by sequencing using the vector specific M13F and M13R primers (Appendix 2) (results not shown). The sequencing results showed that the target sequences had been amplified and were flanked by *Bam* HI and *Xba* I restriction sites. Upon analysis of the sequence chromatograms produced from the *pffabF* front clone, it appeared that there was a single base mutation compared to the published *pffabF* sequence (http://plasmodb.org/plasmodb/servlet/sv?page=gene&source_id=MAL6P1.165&view=seqs). The adenine nucleotide at position 238 of truncated *pffabF* had been substituted with a cytosine, which resulted in a non-silent mutation of the 114th codon (of the entire gene) at the third position from AGA (arginine) to AGC (serine) (Figure 3.1 – highlighted against a blue background). However, this apparent mutation was found to be present in the products of four separate PCRs, which were cloned and sequenced independently. This suggested that the mutation has not been introduced by the polymerase chain reaction, but arose as a result of a polymorphism within the

particular *P. falciparum* cDNA library. No other mutations were observed in the clones containing the *pffabF* front region. The sequence chromatograms of the *pffabF* back construct showed no differences from the published *pffabF* gene and, therefore, both fragments were considered suitable for the construction of the gene encoding pfFabF.

3.5 Assembly of a Truncated form of the *pffabF* gene

Using the strategy illustrated in Figure 3.5, the two fragments were then reconstituted to form a truncated version of the *pffabF* gene within the pGEM-T vector. Both *pffabF* front and *pffabF* back constructs contain a *Xba* I restriction site at the region where the two fragments overlap. Furthermore, the pGEM-T vector contains several restriction enzyme sites on either side of its cloning site. Sequence chromatograms of the *pffabF* front and back clones elucidated the orientation of the inserts within the pGEM-T vector and this knowledge allowed the design of a simple strategy to assemble the *pffabF* gene.

The *pffabF* front construct was used to produce the donor fragment for ligation and the *pffabF* back plasmid acted as the recipient vector (Figure 3.5). Purified plasmid DNA from the *pffabF* front clone was cut with the restriction enzyme *Nco* I that recognises a single site within the cloning region of pGEM-T upstream of the inserted *pffabF* fragment. The linearised vector was treated with Klenow enzyme (Section 2.2.3) in the presence of dNTPs which filled in the overhangs produced by digestion with the restriction enzyme. To limit the probability of the blunt ends ligating to themselves in subsequent ligation reactions, the plasmid DNA was then treated with shrimp alkaline phosphatase (SAP) (Section 2.2.4). A fragment

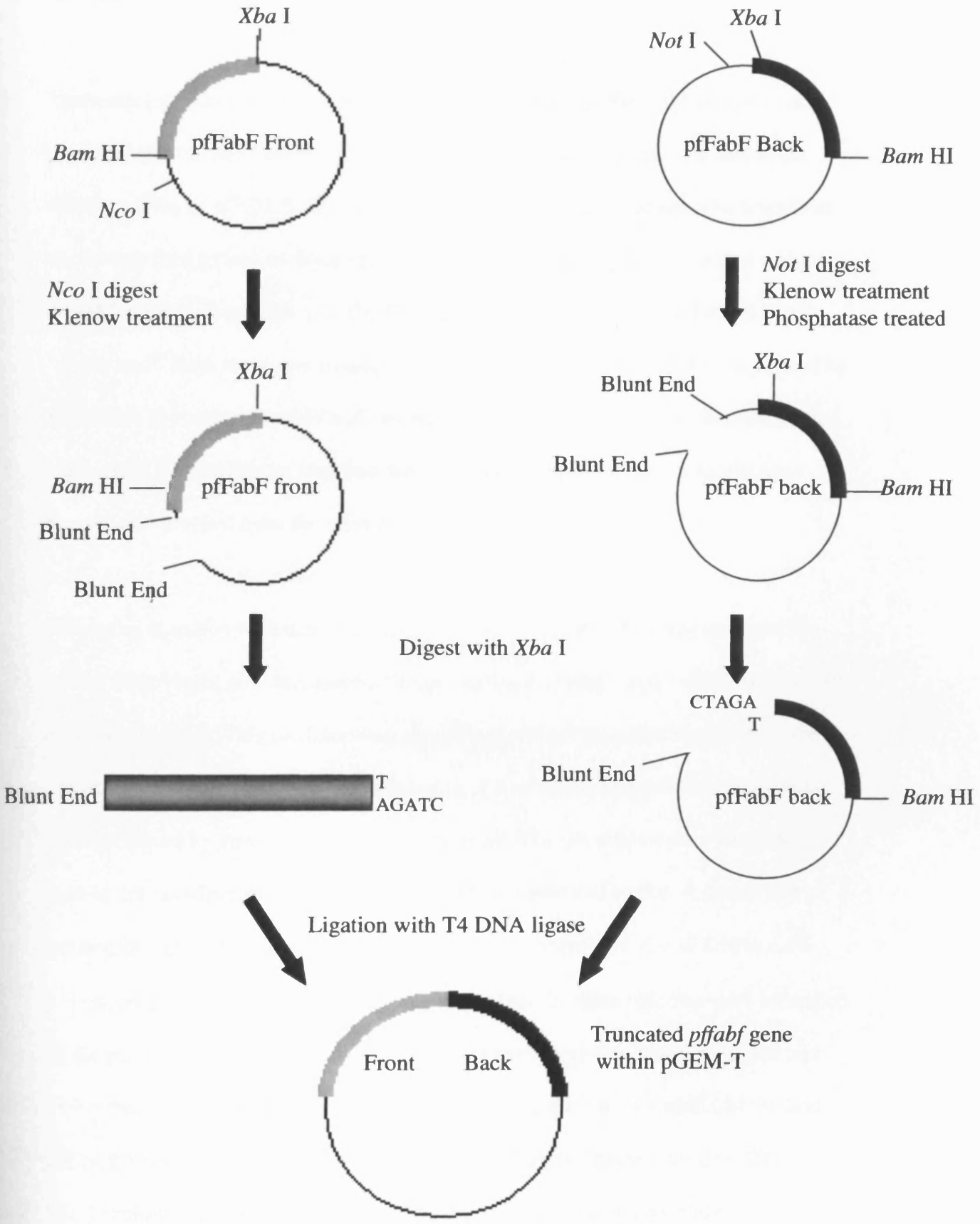


Figure 3.5 Method used for the initial cloning of the *pffabf* gene

containing the *pffabF* front sequence was then excised by cutting with the *Xba* I restriction enzyme (Figure 3.4).

The recipient vector was prepared under similar conditions. First, the plasmid was digested with the *Not* I restriction enzyme that recognised a unique site within the cloning region of pGEM-T upstream of the inserted *pffabF* fragment. The linearised vector was then treated as described above with Klenow enzyme to produce a blunt ended fragment. Digestion with the *Xba* I enzyme resulted in the production of a "sticky end". Both the donor fragment and the recipient vector were then separated by submarine gel electrophoresis and visualised under UV light. The two fragments that correspond to *pffabF* front fragment and the *pffabF* back linearised plasmid were excised and purified from the agarose gel as described in Section 2.2.7.

Following digestion with the *Xba* I enzyme, both the *pffabF* front fragment and the *pffabF* back vector now contained a "blunt end" and a "sticky end" with a compatible *Xba* I 5' overhang. This configuration forced insertion of the donor fragment in the correct orientation. The DNA concentrations of both donor and recipient fragments were estimated by visualisation on an agarose gel. The two fragments were added at a ratio of 1:1, to a ligation mix containing T4 DNA ligase and buffer. A proportion of the resulting ligation mix was then transformed into competent *E. coli* DH5α cells and plated onto LB agar plates containing ampicillin. Positive colonies were identified by the production of an amplicon of 1385bp (1166bp from the *pffabF* sequence and 219bp from the pGEM-T vector) in a PCR containing the vector specific M13F and M13R primers (Appendix 2) (results not shown). Double digests with *Bam* HI and *Xba* I excised two fragments 620bp and 530bp in length (results not shown),

corresponding to the front and back regions of the *pffabF* gene, suggesting that the *pffabF* gene was cloned into the pGEM-T vector. This was confirmed by sequencing. The resulting plasmid was named *pffabF* N-trunc because not all of the mature *pffabF* N-terminus had been cloned.

3.6 Identification of a 5' Extension of the *pffabF* gene

The truncated *pffabF* gene described and assembled above, encoded a protein that showed significant sequence identity with plant and bacterial FabF and FabB proteins but had a shorter N-terminus. As mentioned in Section 3.2, a region of the original contiguous fragment that encodes for an N-terminal extension of *pffabF* was identified in an alternative reading frame to the rest of the gene. This region encodes 35 amino acids that show high sequence identity to *E. coli* FabF (12/35 – 34% identity) and *B. subtilis* FabF (15/35 – 43%) and also contains 8 completely conserved amino acids (Figure 3.1 – shown as red text). The high sequence identity within the N-terminus of FabF proteins suggests that this region is part of the mature protein and may significantly contribute to the correct folding and catalytic activity of the enzyme.

Therefore, it was essential that the gene encoding the mature *P. falciparum* FabF protein was correctly identified. As described in Section 3.2, a region encoding an extension of 35 amino acids, with high sequence homology to both bacterial FabB and FabF proteins, was observed in a different reading frame of the contiguous sequence.

Confirmation of the frame-shift mutation in the original contiguous sequence was observed using the tBlastN program (Altschul *et al.*, 1997) to search databases containing *P. falciparum* contiguous DNA sequences with plant and bacterial FabB and FabF sequences. This search identified a single contiguous sequence located on chromosome 6 of the *P. falciparum* genome that had been recently added to the database. The newly identified 2637bp contiguous region contained the entire original contiguous sequence. However, the 5' region of the original contiguous fragment had been extended by amalgamating the overlapping regions of two contiguous fragments. This expanded region was searched for signal and target peptide sequences and an initiator methionine for the *pffabF* gene but no candidates for either were identified.

In comparison to this new sequence, a sequencing error within the original contiguous fragment corresponding to the deletion of an adenosine nucleotide, was revealed (Figure 3.6A). The nucleotide in question is positioned within a string of four adenosine nucleotides immediately upstream of the cloned *pffabF* front region. When the adenosine nucleotide was present, the corrected contiguous sequence encoded for a mature *pffabF* gene, with the exception of a missing methionine initiator codon, within the same reading frame.

The expanded region of the *pffabF* gene encoded a maximum extension of 35 amino acids before an upstream stop codon was reached (Figure 3.6B). High sequence identity was observed when these 35 amino acids were aligned with the N-terminal region of FabF proteins from other organisms, suggesting that the predicted extension is a correct one. The extended *pffabF* gene consisted of 1257bp encoding a protein of 418 amino acids that shows sequence identity along its entire length to other FabF

enzymes, which indicated that this new version of the gene encoded a mature pfFabF protein.

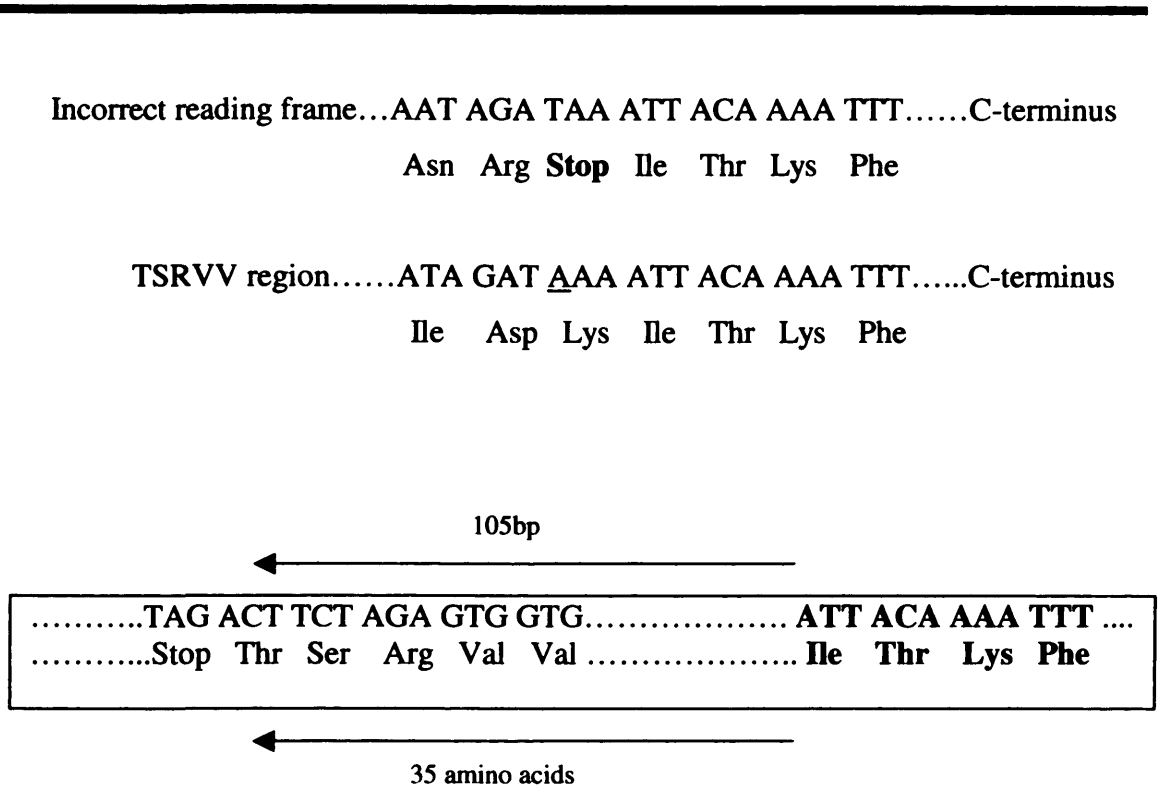


Figure 3.6 The corrected nucleotide sequence of *pffabF*

A i The original sequence of the contiguous fragment containing the *pffabF* gene. An erroneous stop codon (shown in bold text) is encoded, as a sequencing error has led to the absence of an adenosine residue. **ii** The corrected sequence of the contiguous fragment containing the *pffabF* gene. The missing adenosine residue is present (underlined) and instead of a stop codon, the CDS of pfFabF continues upstream. **B** Nucleotides and amino acids shown in bold text are contained within the *pffabF* front fragment previously cloned in Section 3.4. Nucleotides and amino acids shown in normal text are found in the upstream region of the *pffabF* gene. A stop codon is observed 105bp upstream of the *pffabF* front fragment.

3.7 Cloning of the 5' Extension of the *pffabF* Gene

It was decided to clone the 105 base pairs that encoded the 35 amino acid extension, to the 5' end of the *pffabF* N-trunc construct. The *pffabF* 5' extension was designated the TSRVV fragment after the first five amino acids within the coding sequence (CDS) (Figures 3.1 and 3.6; Appendix 1). The *pffabF* gene contained a single *Age* I site, 20 nucleotides downstream of its 5' terminus that was favourably positioned for the addition of a 5' extension (Appendix 1). Primers designed for the PCR amplification of the TSRVV fragment are shown in Figure 3.7. The first of these is a primer termed TSRVV F, which is complementary to the coding strand of the 5' region of the TSRVV fragment and is flanked by a *Bam* HI restriction site for cloning into an expression vector. This primer was used in conjunction with the primer termed TSRVV R, which is complementary to the non-coding strand of the region surrounding the *Age* I restriction site. A 145bp fragment, called TSRVV, was produced using the same cycling parameters as described in Figure 3.4C. As before, Roche High Fidelity *Taq* was used as the DNA polymerase and the *P. falciparum* cDNA library was used as template DNA. Expand High Fidelity *Taq* does not add an overhang of adenosine residues required for the ligation into the pGEM-T vector. Consequently, 5 cycles with generic *Taq* polymerase (Sigma) were performed to add an adenosine overhang.

TSRVV F	5'-GTCGACCGC*GGATCC*GACTTCTAGAGTGGTGTGC -3'
TSRVV R	5'-ACATGACATACCGGTTATATC -3'

Figure 3.7 Sequences of PCR primers used to amplify a 5' extension of the *pffabF* gene

A *Bam* HI site is shown in italics, *Age* I site is highlighted in bold text and underlined nucleotides are complementary to regions of the *pffabF* gene.

Following the second PCR, an amplified fragment of the correct size was visualised on an agarose gel under UV light, excised and purified from the gel. The fragment was ligated into the pGEM-T vector and transformed into *E. coli* DH5α competent cells. Colonies on the plate were screened using a PCR with the TSRVV F and TSRVV R primers under the conditions described above. Restriction enzyme digests with *Bam* HI, *Age* I and a double digest with both of these restriction enzymes further confirmed the presence of the target insert.

The orientation of the fragment within the pGEM-T vector was elucidated by sequencing using the vector specific M13 forward and reverse primers (Appendix 2). DNA sequencing also confirmed the absence of mutations compared to the contiguous sequence found on the *P. falciparum* genome database (www.plasmodb.org).

3.8 Assembly of a Gene Encoding the Complete Mature pfFabF Protein

The 145bp fragment was ready to be ligated on to the 5' region of the *pffabF* gene that was previously cloned into the pGEM-T plasmid (Section 3.5). However, sequencing of plasmid DNA from positive clones suggested that the TSRVV fragment was inserted in different orientations in different clones. Therefore, in order to simplify the assembly of the complete *pffabF* gene, a clone was chosen that had the TSRVV fragment inserted in the pGEM-T cloning site in the same orientation as the *pffabF* gene. Purified plasmid DNA from this clone was used in further procedures, allowing the assembly of the *pffabF* gene in two steps, as depicted in Figure 3.8.

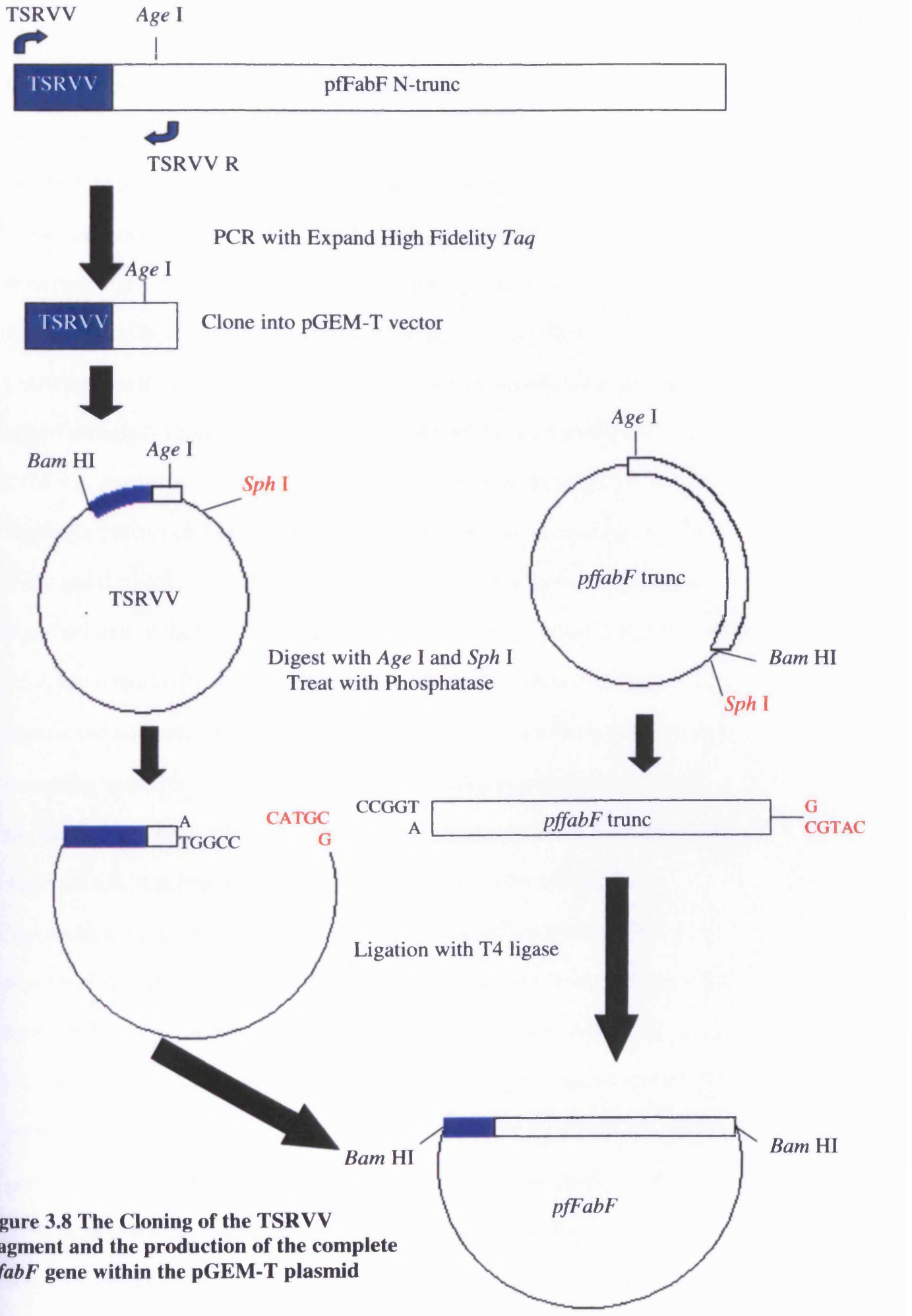


Figure 3.8 The Cloning of the TSRVV fragment and the production of the complete *pffabF* gene within the pGEM-T plasmid

In this strategy the pGEM-T TSRVV construct acted as the recipient plasmid for an excised fragment containing the remainder of the *pffabF* gene. First, both the pGEM-T TSRVV and pGEM-T *pffabF* plasmids were digested with *Age* I and *Sph* I, a restriction enzyme within the pGEM-T cloning site. The DNA fragments produced from each digest were separated in different lanes of an agarose gel (results not shown). The only visible fragment observed in the digest of pGEM-T TSRVV (~3200bp) and the smaller of two visible bands observed in the digested pGEM-T *pffabF* plasmid (~1200bp) were excised and purified. The vector pGEM-T TSRVV DNA was then treated with SAP (Section 2.2.4) to prevent the religation of any DNA fragments that had not been completely digested. The DNA concentrations of both donor and recipient fragments were estimated by visualisation on an agarose gel. Equal amounts of the two DNA fragments were added to a ligation mix containing T4 DNA ligase and buffer. A proportion of the resulting ligation mix was then transformed back into competent *E. coli* DH5α cells and plated onto LB agar plates containing ampicillin. Positive colonies were identified by the production of an amplicon of 1499bp (1280bp from the *pffabF* sequence and 219bp from the pGEM-T vector) in a PCR containing the vector specific M13F and M13R primers (Appendix 2) (results not shown). Sequencing of the purified plasmid DNA from positive clones confirmed that the TSRVV region and *pffabF* N-trunc fragment had ligated in the correct orientation, to form a gene encoding the mature pfFabF protein. Except for the polymorphism described in Section 3.4 and the region encoding a predicted bipartite signal sequence of the annotation on the PlasmoDB, the *pffabF* gene was identical to the sequence found on the *P. falciparum* genome database (PlasmoDB # MAL6P1.165). The *pffabF* gene within this construct was flanked by *Bam* HI sites and appeared in a suitable form for the production of pfFabF in a

recombinant system. The mature *pffabF* gene encodes a protein of 416 amino acids (followed by a natural stop codon) that shows significant identity throughout its entire amino acid sequence to plant and bacterial FabB and FabF proteins.

3.9 Recombinant Expression of pfFabF

3.9.1 Cloning of the *pffabF* gene into the pET15b expression vector

In previous studies, recombinant fatty acid synthesis proteins, including FabF from *E. coli*, had been successfully over-expressed in a pET vector system (Abbadi *et al.*, 2000; Choi *et al.*, 2000; Schaeffer *et al.*, 2001). This family of vectors utilises a promoter derived from the T7 bacteriophage that is not, therefore, recognised by the *E. coli* RNA polymerase. For the production of the recombinant protein, the expression vector must be transformed into an *E. coli* strain engineered to produce the T7 RNA polymerase, such as BL21 (DE3) pLysS. The pET vector provides high level, tightly regulated expression of the target gene, which is achieved by inducing the production of the T7 polymerase by addition of isopropyl β-D-thiogalactopyranoside (IPTG).

The pET15b vector encodes a tag of six histidine residues (His-tag) that is expressed as a short N-terminal fusion with the recombinant protein. The presence of the N-terminal His-tag is advantageous because it allows a convenient detection method of the recombinant protein using a nickel-alkaline phosphatase conjugate (Ni-AP). This was the method of choice as no anti-serum against pfFabF protein was available. Furthermore, the protein could be purified in one step by affinity chromatography on a nickel conjugated resin.

The gene encoding pfFabF was excised from the pGEM-T *pffabF* plasmid using the *Bam* HI restriction endonuclease. The fragment corresponding to the *pffabF* gene was purified from an agarose gel and ligated into the pET15b vector, which had been cut with *Bam* HI and treated with shrimp alkaline phosphatase. The ligation reaction was then transformed into *E. coli* DH5α cells and plated onto LB agar plates containing ampicillin. The ligation mix that produced the most positive colonies following transformation, contained digested pET15b and digested *pffabF* in a ratio of 2:1. Positive colonies were identified by the production of an amplicon of approximately 1500bp (1271bp from the *pffabF* gene and 266bp from pET15b – see Appendices 1 and 3. for sequences) in a PCR containing the vector specific primers T7 and SF1 (Appendix 2). Plasmid DNA was purified from several positive colonies because it is possible that the *pffabF* fragment could insert into the vector in two directions. The orientation of the *pffabF* insert within the pET15b vector was then elucidated by DNA sequencing using the T7 and SF1 primers, in separate reactions. A single clone that contained the *pffabF* gene in the correct orientation for expression was chosen and termed pET-*pffabF*.

3.9.2 Expression of the pfFabF protein in pLysS Cells

The BL21 (DE3) pLysS cells contained a plasmid that expresses lysozyme and can increase cell lysis under conditions that are less likely to cause destruction of cellular proteins. The pLysS plasmid also confers chloramphenicol resistance to the *E. coli*. The pET-*pffabF* plasmid was transformed into competent BL21 (DE3) pLysS cells in accordance with Section 2.2.10 and plated out onto LB agar containing ampicillin and

chloramphenicol. The resultant colonies were screened for their ability to overexpress a protein of the correct size after induction with IPTG.

A peptide mass program (Wilkins *et al.*, 1997) predicted the mass of the pfFabF protein including the N-terminal His-tag to be 48.8kDa. It was necessary to confirm that the pfFabF protein could be produced in a recombinant form in *E. coli*, so before large scale recombinant production of pfFabF could proceed, small scale inductions were attempted as described in Section 2.2.12. Samples were removed at 30min, 1h, 2h, 3h and 24h post-induction and analysed for the presence of recombinant protein by SDS-PAGE and western blotting, using the Ni-AP conjugate.

SDS-PAGE analysis of a culture grown at 37°C and induced with 1mM IPTG (Figure 3.9A), appeared to indicate the induction of several proteins of approximately the correct size. However, the Western blot showed a single reactive band that appeared 30min after the addition of IPTG and accumulated up to 3h post-induction (Figure 3.9B). The pfFabF protein appeared to be present solely in the insoluble fraction, as no immuno-reactive band was observed in the soluble fraction. This small-scale induction proved that the selected pET-*pffabF* clone was capable of producing recombinant pfFabF within an *E. coli* system.

However, it was considered preferable to produce pfFabF in a soluble form. Proteins that accumulate in the insoluble fraction may be produced too rapidly to be processed correctly in the cell and can aggregate together. In an attempt to increase the proportion of soluble pfFabF protein, a culture was grown at 42°C, which was designed to increase the expression of molecular chaperones as part of a heat shock

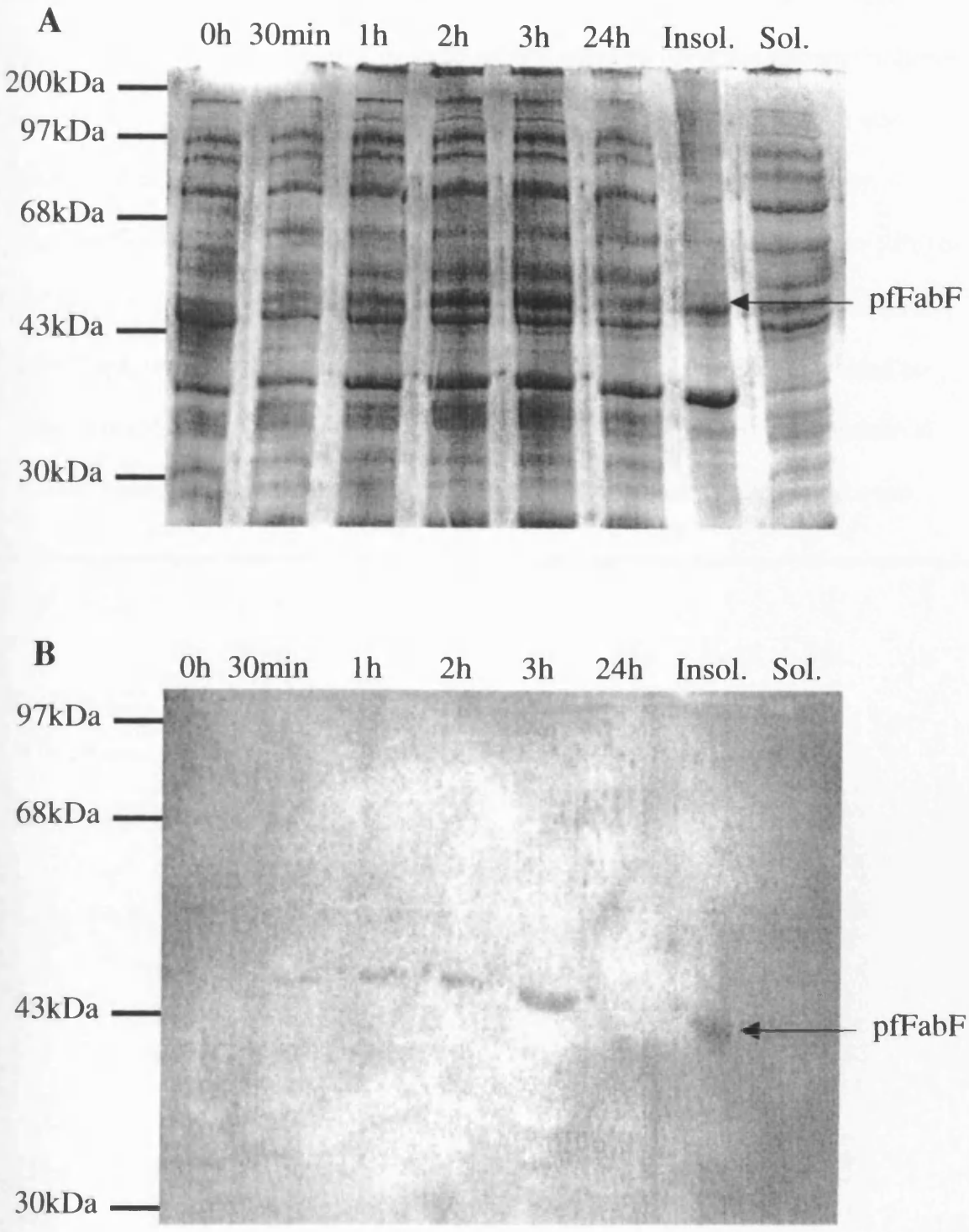


Figure 3.9 Time-course for the induction of pET-*pffabF* in BL21 (DE3) pLysS cells at 37°C

A= an SDS-PAGE gel, B=Western blot. Samples were removed from the culture at 0h (before the addition of IPTG). The other timepoints are at 30min, 1h, 2h, 3h and 24h post-induction. Samples were prepared as described in Sections 2.2.12 and 2.2.13.

response. Molecular chaperones such as GroEL, have been shown to aid correct protein folding and can increase the amount of soluble recombinant protein (Nollen *et al.*, 2001). Furthermore, to curb excessive production of pfFabF this culture was induced with 1/5th the amount of IPTG (0.2mM) compared to the initial study. A small-scale induction of this type suggested that a small proportion (less than 10%) of pfFabF was retained in the soluble fraction (Figure 3.10). However, under the same conditions, inductions on a larger scale (500ml cultures in 2 litre flasks) yielded no soluble pfFabF protein. In addition, compared to the 10ml cultures, the induction of pfFabF in larger cultures appeared to be diminished over time (results not shown).

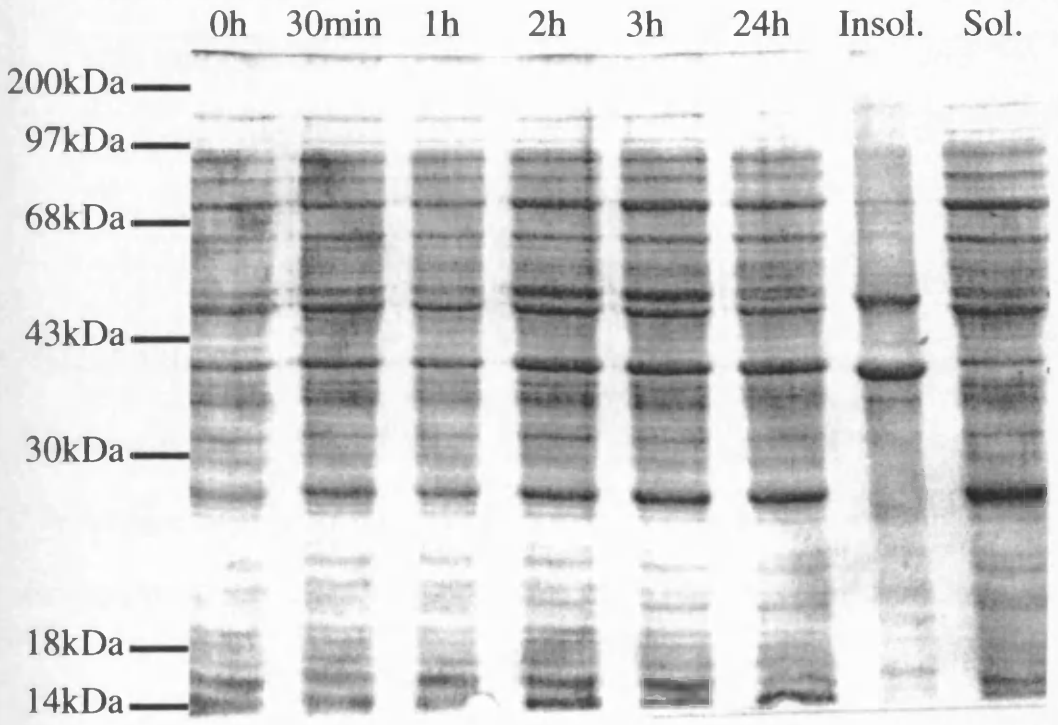


Figure 3.10 Time-course for the induction of pET-*pffabF* in BL21 (DE3) pLysS cells at 42°C

Samples were removed from the culture at 0h (before the addtion of IPTG). The other timepoints are at 30min, 1h, 2h, 3h and 24h post-induction. Samples were prepared as described in Sections 2.2.12 and 2.2.13.

Inductions were also performed at lower temperatures (20°C and 28°C) and with lower IPTG concentrations in an attempt to increase soluble pfFabF protein production (data not shown). However, for all inductions only low levels of *pffabF* expression were observed and the results were not reproducible.

3.10 Expression of pfFabF in Codon Plus Cells

Due to the high A + T nucleotide bias in the *P. falciparum* genome (up to 80% in some regions of the *pffabF* gene), the codon usage differs substantially from that of *E. coli* (Table 3.1).

Codon (amino acid)	*E. coli* [a]	*P. falciparum* [a]
AGG (Arg)/AGA (Arg)	0.91 [b]	76.3 [c]
AUA (Ile)	0.57 [b]	53.3 [c]
GGA (Gly)	1.97 [b]	43.5 [c]
CUA (Leu)	0.83 [b]	15.2 [c]

Table 3.1 Rare codons found in *E. coli* mRNA and their corresponding frequencies in *P. falciparum*

[a] Percentage usage of the codon relative to all other synonymous codons which code for the same amino acid. [b] Medigue *et al.*, 1991. [c] Nakamura *et al.*, 1998.

In particular, some commonly used *P. falciparum* codons for glycine, isoleucine, leucine and arginine are rarely observed in *E. coli* genes. Therefore, the population of tRNA molecules that recognise these rare *E. coli* codons may not be sufficient for the efficient production of some heterologous proteins. This tRNA deficiency can cause stalling of the ribosome on the mRNA strand, leading to diminished levels of

expression in *E. coli* and possibly termination of translation (Kane1995; Kurland and Gallant, 1996). The results obtained in BL21 (DE3) pLysS cells showed that under all conditions attempted only small amounts of pfFabF protein were induced, possibly due to the presence of rare codons. A search of the *pffabF* gene detected 47 codons rarely used in *E. coli*, with 8 of the 9 arginine codons (89%) being in this form. Furthermore, in two separate regions of the gene, a pair and a triplet of adjacent rare codons were observed, thus enhancing the probability of stalling during translation.

To overcome this problem plasmids that overexpress tRNA molecules recognising the rare codons of *E. coli* have been constructed by several groups (Baca and Hol, 2000; Novy *et al.*, 2001). When co-transformed into *E. coli* cells with the expression plasmid, these plasmids have been shown to enhance the expression of proteins from *P. falciparum* and other organisms. The *E. coli* strain BL21 (DE3) Codon + (RIL) contains a plasmid that overexpresses the following tRNAs; *arg*U, *ile*Y and *leu*W, recognising the arginine codons AGG and AGA, the isoleucine codon AUA and the leucine codon CUA. This strain was obtained from Novagen and competent cells made in accordance to Section 2.2.9. The pET-*pffabF* plasmid was transformed into these cells (Section 2.2.10) and colonies used to set-up small scale inductions. Unfortunately, no improvement in the proportion of soluble or the total amount of pfFabF was observed in these inductions compared to those performed with pLysS cells (data not shown).

3.11 Cloning of the *pffabF* Gene into the pMAL Vector System

The lack of solubility of the pfFabF protein poses questions over the ability of the *E. coli* system to produce active enzyme. An attempt to refold the protein from the insoluble fraction (Section 2.2.16; Hill *et al.*, 1996) and subsequent use of a Ni-NTA affinity column, failed to purify any substantial protein of the correct size. Western blots of the eluted fractions showed no reactive bands against the N-terminal His-tag suggesting that either the tag was not accessible to the Ni-AP conjugate or that the protein was absent. This initial refolding attempt suggested that it would be difficult to recover active pfFabF from the insoluble fraction.

A method of improving the solubility of proteins is to express them as fusions with other proteins that usually have a high solubility in the expression system. A screen of *E. coli* proteins showed that a maltose binding protein (MBP) had a very high solubility (Kapust and Waugh, 1999). This protein also has the advantage of binding with high affinity to maltose (Miller *et al.*, 1983), which could then be exploited to allow affinity purification of the fused protein. The pMAL vector family was designed so that heterologous proteins can be expressed as fusions with MBP at the N-terminus (New England Biolabs, Beverly, MA, USA). In an attempt to increase the solubility of the pfFabF protein it was decided to clone the *pffabF* gene into the pMALc2x plasmid (Appendix 4). The plasmid encodes a Factor Xa cleavage site between the MBP and the target protein (Figure 3.11) that can facilitate removal of the MBP after purification of the fused product.

The pGEM-T plasmid containing the *pffabF* gene was used to create a restriction endonuclease fragment for insertion into the pMALc2x vector. New primers were designed so that the *pffabF* product would be flanked by a proximal *Eco*R I restriction

endonuclease site and a distal *Sal* I site. This strategy had been used by Prigge's research group to insert the *P. falciparum* FabH (KAS III) gene into the pMALc2x

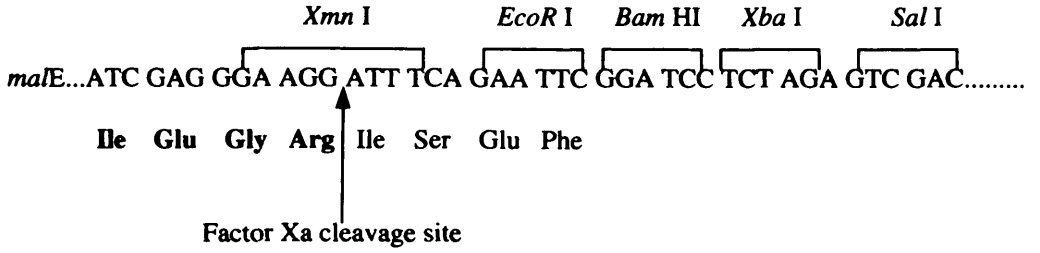


Figure 3.11 The cloning site of the pMALc2x vector

The position of the upstream *mal*E gene encoding *E. coli* maltose binding protein is shown. Amino acids of the Factor Xa recognition site are shown in bold text and the Factor Xa cleavage site highlighted. Unique restriction sites are indicated. The *pffabF* gene was cloned into the vector using a proximal *Eco*R I restriction site and a distal *Sal* I restriction site. (Waters *et al.*, 2002).

The first primer, *pffabF* MAL F, complemented exactly the same nucleotides of the *pffabF* gene as the TSRVV F primer (Figure 3.12). The 5' terminus of this primer contained an *EcoR* I site so that the restriction endonuclease digestion of the PCR product was possible. The second primer, *pffabF* MAL R was complementary to the 3' region of the *pffabF* gene and contained a *Sal* I restriction site. Six nucleotides were added to the 5' terminus of each primer to facilitate, if necessary, the efficient endonuclease cleavage of the PCR product.

A PCR was performed using the cycling parameters shown in Figure 3.4. The DNA polymerase used in this reaction was *Taq* (Sigma Chemicals Ltd., Poole, Dorset, UK) and the purified plasmid DNA of the pGEM-T vector containing the *pffabF* gene was

used as the template. A band of the correct size (1287bp) was purified from an agarose gel and ligated into an empty pGEM-T plasmid at 16°C overnight.

***pfFabF* MAL F**	5'-GGTGGT*GAATTC*ATGTCCACTTCTAGAGTGGTGTGC -3'
***pfFabF* MAL R**	5'-GGTGGTGTCGACATCACTTTACAATTTTTTTGAAAAG -3'

Figure 3.12 The position of the oligonucleotide primers in relation to the truncated *pfFabF* gene.

An *Eco*R I site is shown in italics, a *Sal* I site is highlighted in bold text and underlined nucleotides are complementary to regions of the *pfFabF* gene.

The ligation mix was used to transform competent DH5α cells and several of the resulting colonies were chosen and used to produce purified plasmid DNA. The presence of the *pffabF* gene was detected using a restriction enzyme digest of this plasmid DNA with *Eco*R I and *Sal* I (results not shown). Subsequent sequence chromatograms confirmed the presence of the *pffabF* gene and that no mutations had been introduced in the re-amplification of this gene.

A restriction endonuclease digested *pffabF* fragment was produced from the pGEM-T plasmid in an incubation containing *Eco*R I and *Sal* I and the resulting band purified from an agarose gel. The pMALc2x was prepared under the same conditions, followed by treatment with SAP and inactivation of the SAP enzyme. Both the *pffabF* fragment and the pMALc2x vector now exhibited complementary 5'-3' overhangs that will only allow the insertion of *pffabF* in the correct orientation. A ligation

reaction containing equimolar plasmid and insert DNA was performed under standard conditions.

The ligation mix was used to transform DH5α strain *E. coli* cells and plasmid DNA was purified from several colonies. DNA sequencing results of an individual clone suggested that the *pffabF* gene had inserted into the pMALc2x vector in the correct orientation and the open reading frame of the complete *pffabF* gene was present – this plasmid was termed pMAL-*pffabF*.

3.12 Expression of an MBP-pfFabF fusion protein

In preparation for the production of the protein in *E. coli* cells, the pMAL-*pffabF* plasmid was transformed into competent BL21 (DE3) Codon + RIL cells (Stratagene, Amsterdam, The Netherlands) and plated onto agar plates containing ampicillin and chloramphenicol. The resulting colonies were tested for their ability to produce the pfFabF protein in small scale inductions containing a final concentration of 0.4mM IPTG at 20°C for 16h. The expected size of the MBP-pfFabF fusion protein is 82.2kDa and the selected clone showed the largest induction of the pfFabF (fused to MBP) protein observed in any experiments to date (Figure 3.13). In addition, a significant proportion of the MBP-pfFabF fusion was retained in the soluble fraction (Figure 3.13) suggesting that these conditions were suitable for the production of pfFabF in a soluble form.

Since it had been shown that the pMAL-*pffabF* construct was capable of producing high levels of pfFabF in a recombinant form, a large scale induction was performed

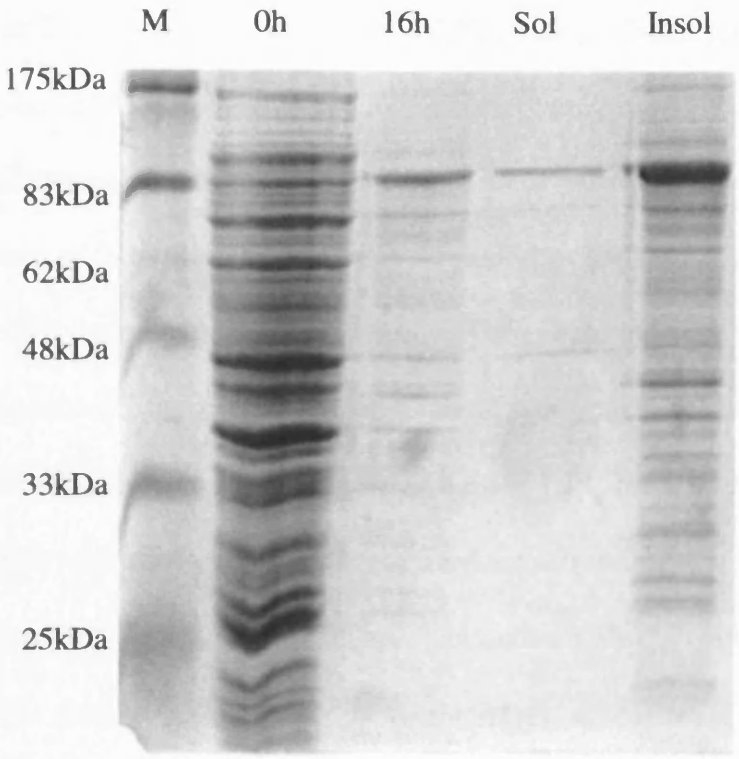


Figure 3.13 Time-course for the induction of pMAL-*pfFabF* in Codon Plus cells at 20°C

Samples were removed from the culture before the addition of 0.4mM IPTG at 0h and at the indicated timepoint post-induction. The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.

under the following conditions. Two 20ml volumes of LB, supplemented with ampicillin and chloramphenicol were inoculated with the BL21 (DE3) Codon + RIL containing the pMAL-*pffabF* plasmid and grown overnight at 37°C in a shaking incubator (250rpm). The following morning, 5ml of these cultures was used to inoculate each of four 2 litre baffle flasks, containing 500ml of pre-warmed LB medium with the aforementioned antibiotics. These cultures were placed at 37°C in a shaking incubator (250rpm) until the attenuance at 600nm was between 0.6-0.8. At this point, the flasks were placed in a water bath at 15°C for 10min and the induction was initiated by the addition of IPTG to a final concentration of 0.4mM. Flasks were then placed at 20°C with shaking (250rpm) and, after 16h, the cells were harvested by centrifugation (8000 x g, 20min, 4°C). At this stage, cell pellets could be stored at -20°C until required, with no adverse effect on the purification of the pfFabF protein. The harvested cells were then resuspended in 40ml of Column Buffer (20mM Tris-HCl, pH7.4, 200mM NaCl, 1mM EDTA). Lysis of the cells was achieved by three rounds of freeze/thawing and the intracellular contents disrupted by sonication (60% amplitude, 30s).

The soluble fraction of the lysed cells was collected by centrifugation (20,000 x g, 20min, 4°C) and filter-sterilised to remove any remaining intact cells. In an attempt to reduce the amount of proteolysis, Protease Inhibitor Cocktail (Sigma Chemicals Ltd., Poole, Dorset, UK) was added in accordance to the manufacturer's instructions and the sample stored on ice prior to further processing. SDS-PAGE and Western blot analysis of this fraction showed the presence of the MBP-pfFabF fusion protein.

3.13 Purification of the MBP-pfFabF fusion protein

The MBP-pfFabF fusion protein was purified from the myriad of other *E. coli* proteins using affinity chromatography with amylose resin. The MBP binds with high affinity to maltose subunits of the amylose chain (Miller *et al.*, 1983) and the protein can then be eluted by the addition of maltose to the Column Buffer.

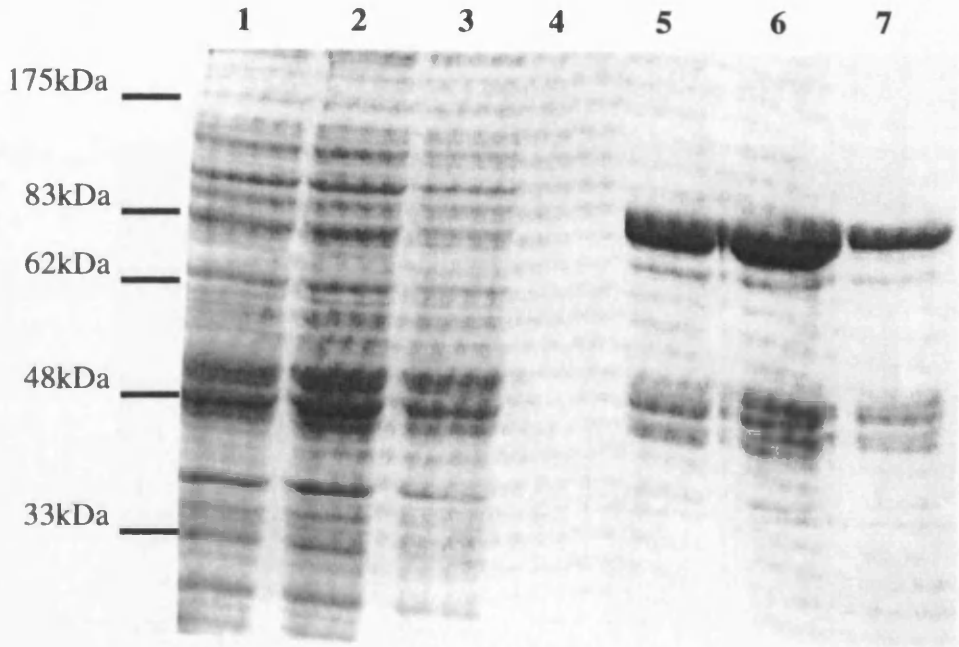
The amylose resin was originally stored in ethanol and was washed several times in equal volumes of Column Buffer at 4°C to remove any traces of the alcohol. A column containing 4ml of a 50% suspension of the resin was poured and equilibrated with at least ten column volumes of Column Buffer. The filtered cell lysate was then applied to the column and unbound proteins washed from the resin using Column Buffer until the attenuance of the flow through 280nm was ≤0.01. Proteins that had bound to the resin were eluted using 10 column volumes of Column Buffer containing 10mM maltose in 1ml fractions. The absorbance of each fraction at 280nm was recorded and used as an approximate indication of protein concentration (assuming a 1mg/ml solution had an absorbance of about 1.0). Typically, the highest protein concentration was observed in the second of these 1ml fractions and a significant amount was also observed in fractions 1 and 3-5. Typically, 1mg of purified, soluble MBP-pfFabF protein suitable for enzymatic studies was produced from each litre of culture. When 1ml fractions were collected the majority of the protein was eluted across Fractions 2-4, with the highest concentration (generally 1mg/ml) in Fraction 2. The resin was then regenerated and could be used several times for the purification of further MBP fusion protein.

SDS-PAGE analysis of the whole cell lysate, column flow through, washing steps and eluted fractions showed that substantial purification of the MBP-pfFabF fusion was achieved using affinity chromatography (Figure 3.14). One of the two major contaminating bands, corresponded to the expected size of free MBP (42.2kDa), whilst the second major contaminating band is slightly smaller than free MBP. Western blot analysis performed on these fractions using serum recognising MBP (New England BioLabs, Beverly, MA, USA) confirmed the presence of MBP in the two major contaminating bands (Figure 3.15). This suggested that the larger band corresponds to free MBP and the smaller protein may be a digestion fragment of MBP.

3.14 Attempted cleavage of the MBP-pfFabF fusion

The pMALc2x vector encodes a Factor Xa recognition site within the region between MBP and the N-terminally fused target protein (Figure 3.11). This protease site should facilitate separation of the pfFabF protein, leaving little or no vector encoded amino acids attached. Following successful cleavage, it is possible to separate pfFabF from free MBP by affinity chromatography on amylose resin, thus removing the contaminating band described above (Section 3.13).

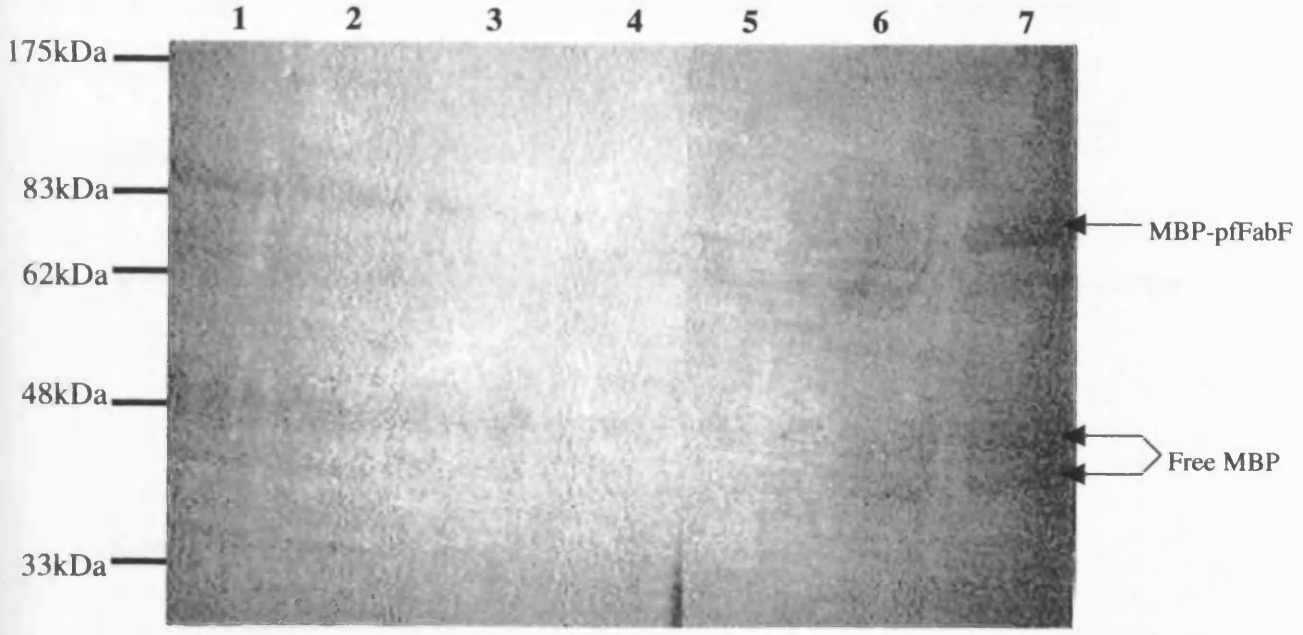
A small amount (20μg) of the eluted Fraction 2, which contained the highest concentration of the fusion protein, was used in cleavage experiments by incubation with 1μg of Factor Xa. A negative control containing no Factor Xa protease and a positive control containing a MBP-paramyosin fusion (New England BioLabs, Beverly, MA, USA) and Factor Xa were included in all experiments. No apparent



Lane	Sample
1	Soluble Fraction
2	Initial Flow Through (unbound protein)
3	Beginning of washing steps
4	End of washing steps
5	Eluted fraction 1
6	Eluted fraction 2
7	Eluted fraction 3

Figure 3.14 Purification steps for the MBP-FabF fusion protein

The soluble fraction of cells expressing the MBP-pfFabF fusion protein (Lane 1) was run through an amylose column and the effluent collected (Lane 2). The column was then washed with Column Buffer to remove any unbound protein (Lanes 3-4). 1ml fractions were eluted by the addition of 10mM maltose to the Column Buffer (Lanes 5-7). The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.

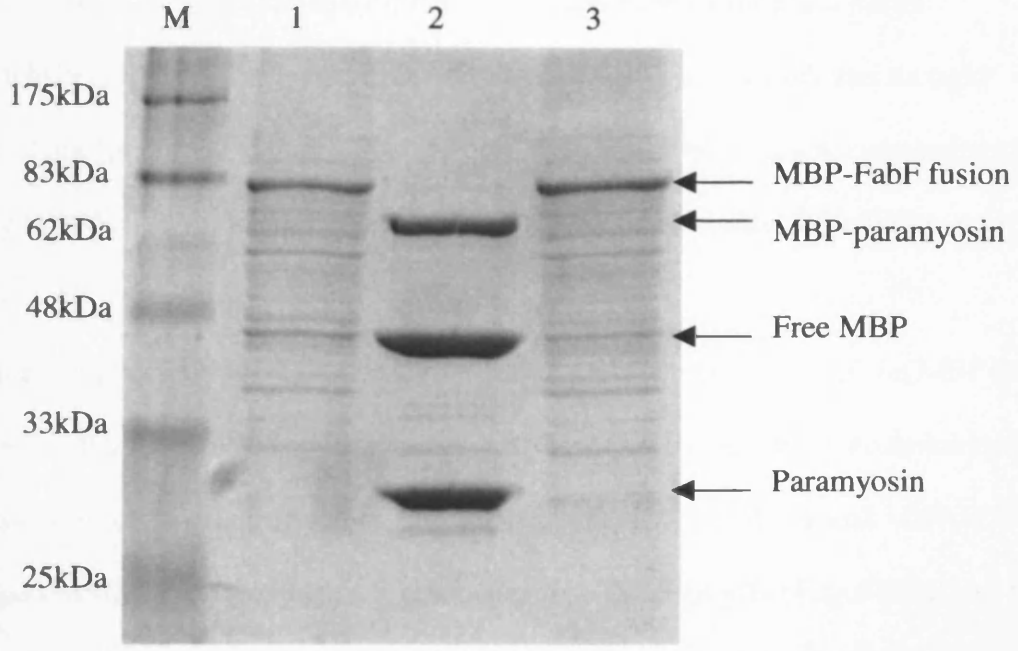


Lane	Sample
1	Soluble Fraction
2	Initial Flow Through (unbound protein)
3	Beginning of washing steps
4	End of washing steps
5	Eluted fraction 1
6	Eluted fraction 2
7	Eluted fraction 3

Figure 3.15 Purification steps of the MBP-FabF fusion protein

The soluble fraction of cells expressing the MBP-pfFabF fusion protein (Lane 1) was run through an amylose column and the effluent collected (Lane 2). The column was then washed with Column Buffer to remove any unbound protein (Lanes 3-4). 1ml fractions were eluted by the addition of 10mM maltose to the Column buffer (Lanes 5-7). The proteins contained in each sample were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. The western blot was washed and developed using anti-MBP serum (New England Biolabs, Beverly, MA) as described in Section 2.2.15.

cleavage of the MBP-pfFabF fusion was observed after an 8h incubation with Factor Xa (Figure 3.16, lane 3). However, under the same conditions, approximately half of the MBP-paramyosin fusion protein had been separated into its components, suggesting that the Factor Xa protease was active.



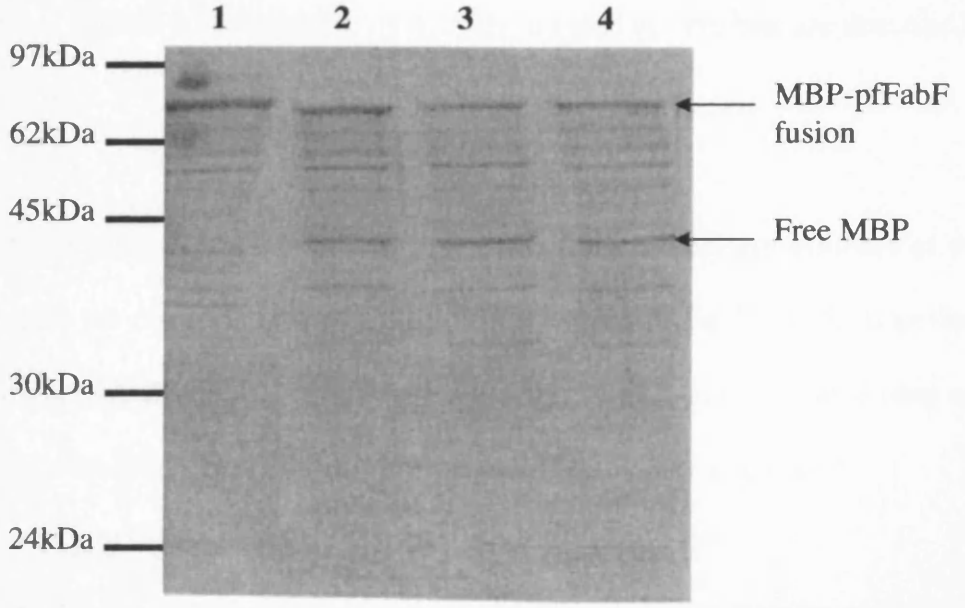
Lane	Sample
1	**MBP-pfFabF – No Factor Xa**
2	**MBP-paramyosin + Factor Xa**
3	**MBP-pfFabF + No Factor Xa**

Figure 3.16 Incubation of MBP-pfFabF fusion protein with 1µg of Factor Xa
MBP-pfFabF (20µg) and MBP-paramyosin (20µg) were incubated with 1µg of Factor Xa at 25°C for 8 hours. 1/20th of the proteins contained in each reaction were separated by SDS-PAGE and visualised by staining with Coomassie blue.

Factor Xa requires calcium ions to cleave target proteins efficiently and is supplied (New England BioLabs, Beverly, MA, USA) in a buffer containing 2mM $CaCl_2$. However, the MBP-pfFabF fusion protein was dissolved in elution buffer that contains a 1mM concentration of the metal chelating agent EDTA. This may lower the effective concentration of $CaCl_2$ and, thus, decrease the activity of Factor Xa. Therefore, cleavage of the MBP-pfFabF fusion was performed in a buffer supplemented with $CaCl_2$ to a final concentration of 3mM, which was thought sufficient for complete Factor Xa activity. The amount of Factor Xa was reduced from 1μg to 200ng, in an attempt to reduce any unspecific cleavage of the fusion protein.

Under these conditions, SDS-PAGE revealed significant cleavage of the MBP-pfFabF protein after 4h incubation with Factor Xa (Figure 3.17). However, accumulation of proteins of corresponding size to the cleavage products of pfFabF and MBP is not observed with time. Bands of a size smaller than those of pfFabF and MBP can be seen on the SDS-PAGE gel, suggesting that digestion of pfFabF by Factor Xa may occur at non-canonical sites within the protein. Factor Xa recognises specific sites and cleaves the natural substrate prothrombin after the protein sequence isoleucine-glutamate-glycine-arginine (IEGR) or, in some cases, isoleucine-aspartate-glycine-arginine (IDGR). A search of the pfFabF protein showed no sequences with obvious homology to these Factor Xa recognition sites. However, proteins can be cleaved at other basic residues depending on the context and several sequences containing lysine residues showed homology to Factor Xa recognition sites within pfFabF.

Thus, unfortunately, attempts to separate the pfFabF protein from MBP proved unsuccessful. However, even with a large N-terminal extension, the structure of the pfFabF protein may be unaffected and the formation of a homodimer with ketoacyl-ACP synthase activity could still be feasible. Studies of the enzyme activity of the MBP-pfFabF fusion protein are described in Chapter 4.



Lane	Sample
1	Zero timepoint
2	2 hour incubation
3	4 hour incubation
4	6 hour incubation

Figure 3.17 Incubation of MBP-pfFabF fusion protein with 200ng of Factor Xa and 3mM $CaCl_2$

MBP-pfFabF was incubated with 200ng of Factor Xa at 25°C with 3mM $CaCl_2$ added to each reaction, at the times indicated above. The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.

3.15 Discussion

The work presented in this chapter describes the identification of a gene encoding a condensing enzyme (pfFabF) from the malaria parasite, *Plasmodium falciparum.* In addition, it outlines the first reported cloning, expression and purification of soluble pfFabF in a recombinant form. In *P. falciparum* malaria and other organisms, this enzyme is believed to be the main target of the natural antibiotic thiolactomycin and, thus, it is advantageous to be able to study the purified, soluble protein. Studies of the activity of pfFabF and the inhibitory activity towards this enzyme are described in Chapter 4.

The *pffabF* gene, identified in this chapter, encodes a protein the sequence of which aligns with the complete amino acid sequences of mature FabF and FabB proteins from bacteria and plants (Figure 3.1). This region is sufficient for condensing enzyme activity in the FabB and FabF enzymes isolated from many organisms.

The *pffabF* gene identified in this project does not encode an initiator methionine residue or an N-terminal bipartite signal sequence that would target the protein to the apicoplast (the expected location of the Type II FAS pathway and pfFabF). Currently, the *Plasmodium* genome database (www.plasmodb.org) suggests two potential N-terminal sequences upstream of the region that encodes the mature protein, which starts with a TSRVV sequence (shown in bold).

These are;

MNIRCRKEKKKKKKKNINHVNSKRYAFIKRGIPGISKNYFKGFKLYNSREMK NLCE**TSRVV....mature protein**

or

MNIK**TSRVV....mature protein**

Both of these sequences start with an initiator methionine, usually required for successful translation of mRNA to produce proteins. The initiator methionine (encoded by the same ATG codon in both predictions) is 380 bases upstream of the mature protein sequence, on chromosome 6 of *P. falciparum*. However, neither of these N-terminal peptides would lead us to predict that the protein would be targeted to the apicoplast. The composition of these upstream regions does not resemble the characteristic bipartite signal sequence utilised for apicoplast targeting, that consists of a hydrophobic signal peptide followed by a positively charged target peptide. In both cases no obvious hydrophobic signal sequence is observed although in the larger peptide a positively charged region enriched with lysine and asparagine is present following a long stretch of lysine residues. Based on the presence of an N-terminal bipartite sequence, the PATS program (Zuegge *et al.*, 2001) is trained to differentiate between parasite proteins localised to the apicoplast and those that are localised elsewhere. Both of the N-terminal extensions published on PlasmoDB returned very low (0.023 and 0.025, respectively) readings, suggesting that neither would be sufficient for targeting of pfFabF to the apicoplast. Within this project, searches using a range of techniques were unsuccessful in identifying an N-terminal bipartite sequence of pfFabF.

The lack of success, by any group, to identify a 5' region of the *pffabF* gene is surprising. In this work, all of the searches used the ring stage cDNA library. Because the cDNA library is derived from mRNA sequences and contains at least one copy of the *pffabF* gene, identification of the 5' region should have been as simple as using the gene specific TSRVV R primer and an upstream vector primer in a PCR. The cDNA library used in this project was originally obtained from (Prof. Steve Walsh,

Liverpool School Tropical Medicine, UK) but had been used extensively and successfully in our laboratory for many years. Unfortunately, the supplying laboratory was unable to provide information on the plasmid vector used for cloning the library. Although, several common vector specific primers were tested in this work it was not possible to amplify fragments with these primers in combination with the gene specific TSRVV reverse primer to identify the N-terminus of the *pffabF* gene. I believe that there is a strong chance that pfFabF is targeted to the apicoplast, however, at the time of submission of this thesis, no apicoplast targeting sequence had been correctly identified.

The N-terminal signal sequence is not required for activity of other enzymes in the malaria Type II FAS pathway (Surolia and Surolia, 2001; Waters *et al.*, 2002; Prigge *et al.*, 2003; Pillai *et al.*, 2003; Sharma *et al.*, 2003). In fact, Prigge's laboratory showed that expression of the pfFabH (KAS III), with the bipartite signal attached, produced insoluble protein (Waters *et al.*, 2002). Deletion of the region of *pffabH* that encodes the signal sequence, allowed the expression of soluble protein (Waters *et al.*, 2002).

The finding that the condensing enzyme identified in this chapter was more closely related phylogenetically to FabF proteins than to FabB was not surprising. FabB enzymes have so far only been found in Gram-negative enterococciae bacteria and all of these organisms also contain a FabA-like dehydratase. The dehydratase in *P. falciparum* is more closely related to FabZ proteins (Sharma *et al.*, 2003) and, therefore, it is highly unlikely that *P. falciparum* contains a FabB homologue. However, the designation of this gene should help to avoid any confusion in the study

of this gene and its product. Following the completion of the *P. falciparum* genome project in 2002, a search for FabB, FabF and Type I FAS ketoacyl synthase domains was undertaken. No new candidate ketoacyl synthases were, or have since been, identified during the course of this work.

The *pffabF* gene encodes a protein that has significant identity (46% to *B. subtilis* FabF, 39% to *E. coli* FabF and 31% to *E. coli* FabB) to the β-ketoacyl-ACP synthase family (Figure 3.1). The catalytic triad of Cys163/His298/His333 (*E. coli* FabB numbering) is present as Cys165/His306/His343 and a highly conserved lysine residue (K328 - *E. coli* FabB numbering) that is essential for the correct positioning of His333 within the active site, is also conserved in pfFabF (K338).

The single base polymorphism in the *pffabF* gene identified during this work introduces a R114S mutation. An alignment of FabB and FabF proteins (Figure 3.1) showed that this residue was not conserved with the FabB/FabF family. Comparison of the structures of FabB and FabF proteins places the residues corresponding to R114 within the Nα1.1 alpha helix (nomenclature from Olsen *et al.*, 1999) (Figure 3.18), a long loop that is close to the dimer interface. Q113 (*E. coli* FabB numbering) is conserved in some FabB enzymes and appears to be the single residue of interest in this region (Olsen *et al.*, 1999). It is thought that the active site stretches from C163 to Q113, a distant of 16Å, which could reasonably hold the acyl substrates used by FabB enzymes. Q113 is not conserved in FabF-like enzymes (Figure 3.1) and the contribution of the R114 residue (F112 in *E. coli* FabB) has not been determined. However, the large number of different amino acids (E, D, K, R, H, L, S, T and G)

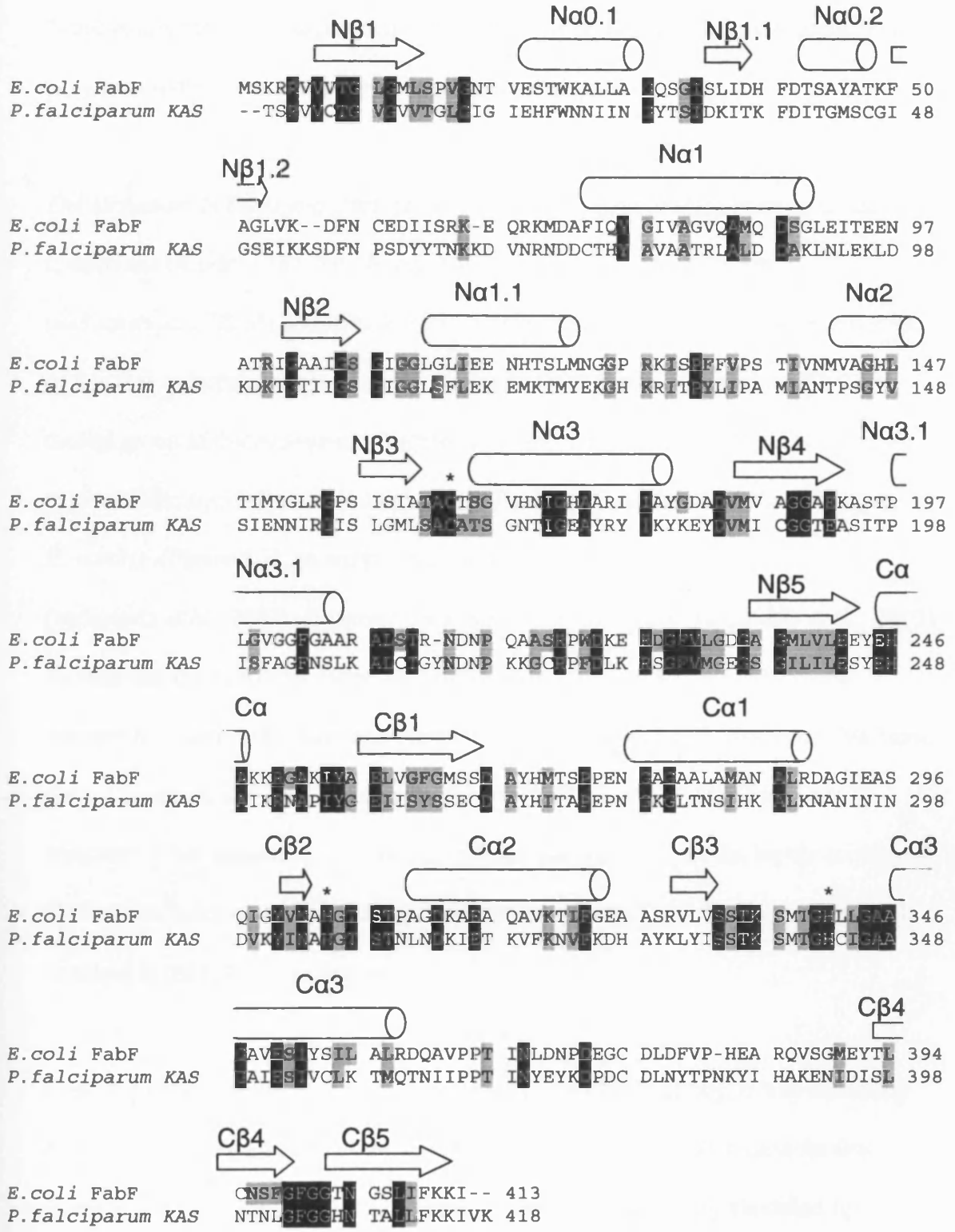


Figure 3.18 Secondary structure elements in *E. coli* FabF aligned with pfFabF
Active site residues are marked with asterisks. Highly conserved individual residues within the whole of the FabB/FabF family are highlighted in black. Groups of similar amino acids conserved within the whole of the FabB/FabF family are highlighted in grey (DE/HKF/IVLM/AGS/WFYH/ST). α helixes are shown by horizontal tubes and β sheets are shown by block arrows. The R114S alteration caused by a polymorphism is highlighted against a background of blue. Based on secondary structure elements first published in Olsen *et al.*, 1999.

found at this position, suggest that this residue has no central role to the condensing enzyme activity or substrate specificity of these enzymes.

The alignment of FabB and FabF proteins shows a region of identity close to the C-terminus (residues 387-395, *E. coli* FabB numbering). The structure of thiolactomycin (TLM) bound to *E. coli* FabB showed that the highly conserved F390 of this region forms a hydrophobic pocket with P272 that is occupied by the C10 methyl group of thiolactomycin (Figures 1.17 and 1.18). The pfFabF protein has a leucine in the equivalent position to *E. coli* FabB F390, as does the pfFabF enzyme of *B. subtilis* (Figure 3.1), an enzyme which is highly resistant to TLM (Jackowski *et al.*, 2002). However, Jackowski and colleagues (Jackowski *et al.*, 2002) showed that the F390V mutants were resistant to TLM due to steric hindrance between the valine side chain and the C10 methyl group of TLM. However, this steric clash was absent in F390L mutants and was not sufficient for inducing TLM resistance. This suggests that whilst the pfFabF enzyme replaces the highly conserved F390 with a leucine residue, it should still be sensitive to TLM. This hypothesis will be tested in the following chapter.

Even though the *pffabF* gene was represented in the cDNA library, it was extremely difficult to amplify the gene as a single fragment. The most likely reason for this problem was the A/T richness of the gene. Despite the gene being identified by several groups and the current focus on malaria fatty acid synthesis as a potential drug target no other research group has reported the cloning of the gene encoding this protein. It is thought that this may be due to the inherent problems in cloning this gene.

To overcome the lack of over-expression and the lack of soluble protein, two practical solutions were employed. The first was to transform the expression vector in to BL21 (DE3) Codon + (RIL) cells, which successfully overcame the lack of overexpression observed in other *E. coli* cells. Secondly, following discussions with Dr Sean Prigge (Johns Hopkins School of Public Health, Baltimore, USA) regarding production of another *P. falciparum* KAS enzyme, the *pffabF* gene was cloned into the pMALc2x vector, which, when expressed, added maltose binding protein to the N-terminus of mature region of pfFabF. This procedure helped to increase the amount of soluble protein, although a majority remained insoluble under various conditions. The maltose binding protein (MBP) system also allowed the efficient purification of the pfFabF from *E. coli* cellular proteins, with the only contaminating bands are free MBP or fragments produced by the breakdown of MBP.

The lack of success in cleaving the MBP from the N-terminus of pfFabF is disappointing. The FabD enzyme of malaria Type II FAS was also expressed as a MBP fusion. This enzyme could also not be efficiently cleaved from MBP but retained significant malonyl CoA:ACP transacylase activity *in vitro* (Prigge *et al.*, 2003). This indicates that it is possible that the MBP-pfFabF fusion may also retain enzyme activity. If there was time to repeat the cloning of pfFabF into an expression vector it would be advisable to express pfFabF as a fusion with MBP, but it would be advantageous to choose a highly specific protease site, such as enterokinase (pMALc2e plasmid, New England Biolabs, Beverly, MA, USA) or tobacco etch virus protease (Prigge *et al.*, 2003) between the two proteins in the hope of achieving complete cleavage of the fusion protein.

Chapter 4

Substrate Specificity and Characterisation of MBP-pfFabF

4.1 Initial studies of MBP-pfFabF Activity

A candidate gene encoding a putative FabF protein from *P. falciparum* was identified, cloned and expressed in an *E. coli* expression system as described in Chapter 3. The protein was produced with the 42kDa maltose binding protein (MBP) fused at the pfFabF N-terminus in an attempt to increase the solubility and as an aid to purify the protein. Attempts to remove the MBP by cleavage at an integral Factor Xa site located between the MBP and MBP-pfFabF domains were unsuccessful (Section 3.13) and the protein remained connected to MBP. For clarity and consistency, the pfFabF fusion protein with MBP is referred to as MBP-pfFabF throughout this Chapter.

FabB and/or FabF condensing enzymes of Type II FAS found in bacteria form homodimers that are essential for enzyme activity (Garwin *et al.*, 1980; Edwards *et al.*, 1997; Olsen *et al.*, 1999; Huang *et al.*, 1998; Moche *et al.*, 2001) and it may be assumed that the homologous enzyme from *P. falciparum* will also function as an active dimer. However, the presence of a large protein at the N-terminus of MBP-pfFabF may interfere with the formation of a homodimer or occlude the active site, leading to abrogation of activity. To test whether the enzyme was capable of performing the condensation reaction of fatty acid biosynthesis it was important to chose substrates that are utilised by a wide range Type II FAS β-ketoacyl-ACP synthases. Generally, in bacteria and plant plastids, palmitate (C16:0) is the primary saturated fatty acid produced by *de novo* synthesis Table 1.2. Palmitate is synthesised by the Claisen-condensation of malonyl-ACP with myristoyl-ACP (C14:0), a reaction catalysed in *E. coli* by both FabB and FabF enzymes. Therefore, in initial studies of

the *P. falciparum* KAS it seemed sensible to investigate enzyme activity in an assay independent of whether the enzyme showed FabB-or FabF-like substrate specificity i.e. an assay containing myristoyl-ACP as the acyl-ACP substrate and malonyl-ACP as the donor unit.

To determine if the purified MBP-pfFabF retained any KAS activity, an assay containing a large amount (1μg) of the MBP-pfFabF fusion was performed as described in Section 2.2.24. As a positive control for the substrates and conditions used in the assays, *E. coli* FabB (ecFabB) was kindly provided by the laboratory of Charles Rock (St. Jude Children's Research Hospital, Memphis). A significantly smaller amount of ecFabB (6.25ng), compared to MBP-pfFabF (1μg), was used in the assay because of its known activity under identical assay conditions. For the assay, ^{14}C labelled malonyl-ACP (Section 2.2.24) was generated by transacylation of [2-^{14}C]-malonyl-CoA to *E. coli* ACP (ecACP) catalysed by the *E. coli* FabD protein (ecFabD) (purified in Section 2.2.26). The radiolabelled malonyl-ACP is then able to donate a two carbon unit to an acyl-ACP substrate (myristoyl-ACP) to form a ketoacyl-ACP (Figures 1.12 & 1.13). In the assay, the ketoacyl group is then released from the phosphopantetheine arm of ACP by the reduction of the sulphydryl bond by sodium borohydride and tetrahydrofuran (Section 2.2.24). The resulting radiolabelled fatty acids were extracted in toluene, added to scintillation solution and the incorporation quantified by radioactive counting.

The results of the first assay of MBP-pfFabF activity are shown in Figure 4.1. The pfFabF enzyme catalysed the condensation of a large amount of [2-^{14}C]-malonyl-ACP with the myristoyl-ACP substrate, suggesting that the MBP-pfFabF fusion protein

exhibited a reasonable β-ketoacyl-ACP synthase activity. It must be noted that, although more radioactive products were observed in the MBP-pfFabF assay compared to that of ecFabB, there was over 100-fold more MBP-pfFabF protein than ecFabB in the assay. These data also suggested that myristoyl-ACP is a suitable acyl-ACP substrate of MBP-pfFabF that could be used for further assays.

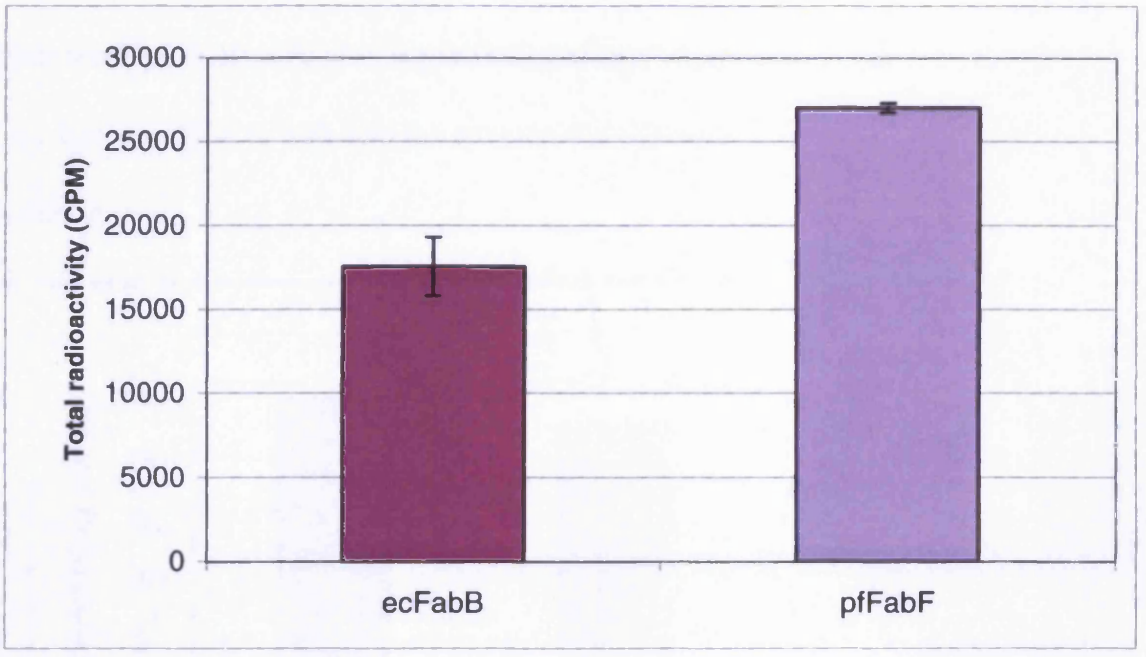


Figure 4.1 The ability of pfFabF to extend myristoyl-ACP

The incorporation of radiolabelled malonyl-ACP with ecFabB (■) and pfFabF(■) Each assay was performed in duplicate and errors are shown as the range of results from the mean. Assays contained 6.25ng of ecFabB and 1μg of pfFabF and the activities were measured using a radiochemical assay as described in Section 2.2.24.

Having established the activity of MBP-pfFabF it was important to store the enzyme in such a way as to retain as much activity for as long as possible. A comparison between the activity of; an aliquot of purified MBP-pfFabF that had been stored at 4°C for 30 days, MBP-pfFabF stored as solution in 50% glycerol at –20°C for 30 days and a sample of freshly prepared MBP-pfFabF was performed (Figure 4.2). When stored at -20°C in 50% glycerol for 30 days, greater than 90% activity was retained. However, storage at 4°C with no glycerol led to a loss of almost all of the condensing enzyme activity (2.6% remained) after 30 days. Using this knowledge, each time a fresh batch of MBP-pfFabF was purified, sterile glycerol was immediately added to a final concentration of 50% and the enzyme was stored at -20°C until required for further use.

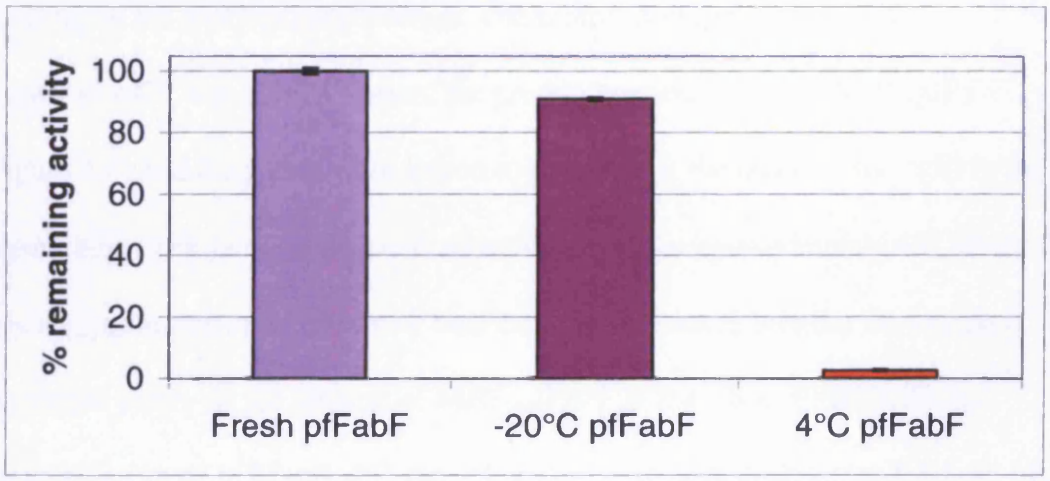


Figure 4.2 Effects of storage on MBP-pfFabF activity

The percentage remaining activity of freshly prepared MBP-pfFabF (■), MBP-pfFabF stored at -20°C for 30 days in 50% glycerol (■ or stored at 4°C for 30 days (■). Assays were performed in duplicate and errors are shown as the range of results from the mean. Each assay contained 1µg of the different pfFabF preparations and the activities were measured using a radiochemical assay as described in Section 2.2.24.

4.2 Refining the conditions of the MBP-pfFabF Assay

In Section 4.2 it was ascertained that the MBP-pfFabF fusion protein was capable of catalysing the condensing step in fatty acid biosynthesis. In order to perform studies of the acyl-ACP substrate specificity for MBP-pfFabF it was important to examine some basic assay conditions. The following experiment was to determine the linearity of the reaction with protein. This would also determine if the amounts of the substrates, malonyl-ACP or acyl-ACP, are limiting the amount of radiolabel incorporation. The activity of the MBP-pfFabF protein was tested at concentrations ranging from 1μg to 4.1ng whilst the concentrations of myristoyl-ACP (50μM), [2-^{14}C] malonyl-CoA (50μM), ACP (100μM) and ecFabD (1μg) remained constant. A 500ng/μl stock of MBP-pfFabF in 50% glycerol was mixed with 100mM Na-Phosphate buffer, pH6.8 containing 50% glycerol, to produce 1 in 3 serial dilutions of the MBP-pfFabF protein. The radioactivity incorporated into myristoyl-ACP was plotted against the protein concentrations of MBP-pfFabF (Figure 4.3) and the points were joined to elucidate if the reaction was still in the linear phase. The graph shows a straight line, rising at approximately 45° above the x axis suggesting that the amount of radiolabel incorporated into the fatty acids is dependent solely on the amount of MBP-pfFabF in the assay. Even in assays containing 1μg of MBP-pfFabF, the substrates were in abundance and did not impose any limit on the rate of the reaction. Therefore, there was confidence in using MBP-pfFabF within the range (4.1ng-1μg) in further assays. All subsequent assays contained 300ng of MBP-pfFabF protein as this amount was well within the linear phase of the reaction and was the lowest concentration (to conserve stocks of MBP-pfFabF) that incorporated enough radioactivity to produce a signal to background ratio adequate for accurate inhibition studies.

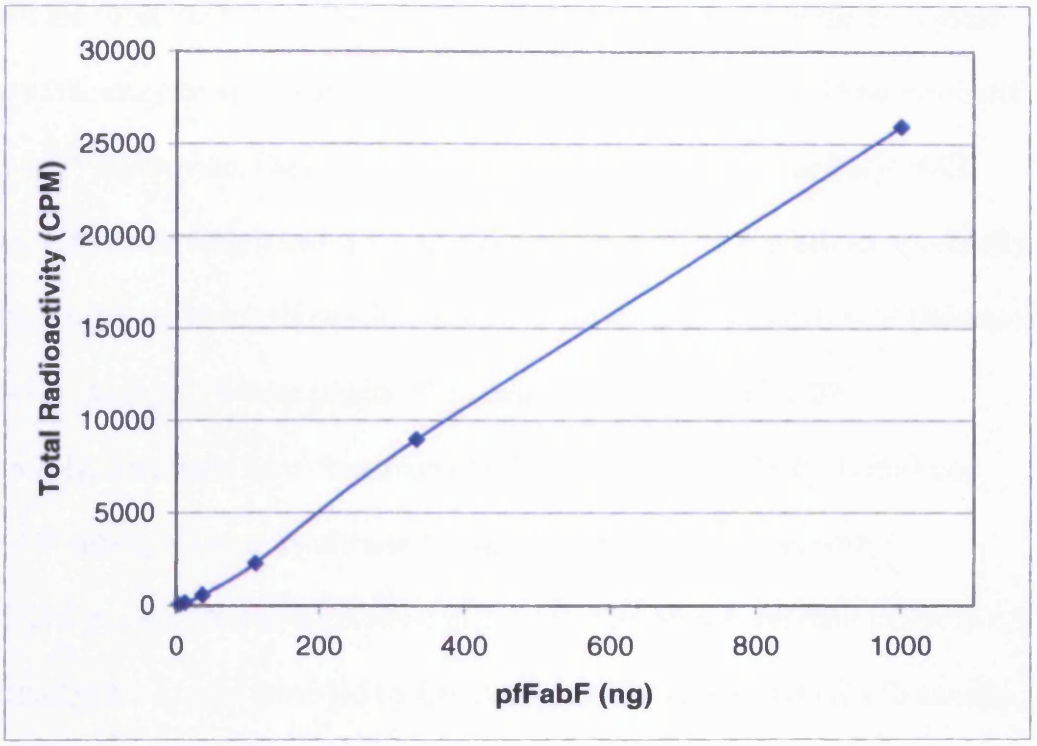


Figure 4.3 The amount of radiolabel observed in the acyl-ACP product with different amounts of pfFabF

Each assay was performed in duplicate and errors bars show the range of results from the mean. The activity of pfFabF was measured using the radiochemical assay as described in Section 2.2.24.

4.3 Acyl-ACP Substrate Specificity of MBP-pfFabF

Studies of the substrate specificity of MBP-pfFabF help to elucidate the biological function of the enzyme and improve the understanding of the fatty acid biosynthesis pathway in *P. falciparum*. FabF-like enzymes are able to use saturated acyl-ACP substrates C4-C14 in length and show an overlapping acyl-ACP substrate specificity with FabB enzymes, which show a lower activity towards C14:0 acyl-ACP (Edwards *et al.*, 1997). Acyl-ACPs were produced as described in Section 2.2.27. Unfortunately, within the time constraints of the project, only C14:0-, C16:0-and C18:0-ACP substrates were synthesised in time to be used in assays with the MBP-pfFabF protein (described in Section 2.2.24). The *Mycobacterium tuberculosis* protein (mtKasA – kindly provided by Charles Rock, St. Jude Children's Research Hospital, Memphis) was chosen as a positive control because it exhibits KAS activity across the range of acyl-ACPs tested (Kremer *et al,*. 2002).

The results (Figure 4.4) indicate that of the three acyl-ACP substrates tested, MBP-pfFabF had a much higher activity with myristoyl-ACP (C14) compared to the palmitoyl-ACP (C16) and stearoyl-ACP (C18) equivalents. Activity with palmitoyl-ACP and stearoyl-ACP was only 6% and 3% respectively, compared to that observed with myristoyl-ACP. This suggested that the longest saturated fatty acid produced in significant amounts in *P. falciparum in vivo* would be palmitate (16:0). However, these results also show that the enzyme is capable of elongating C16:0-and C18:0-ACP substrates, if they are available in sufficient quantities to compete with C14:0-ACP.

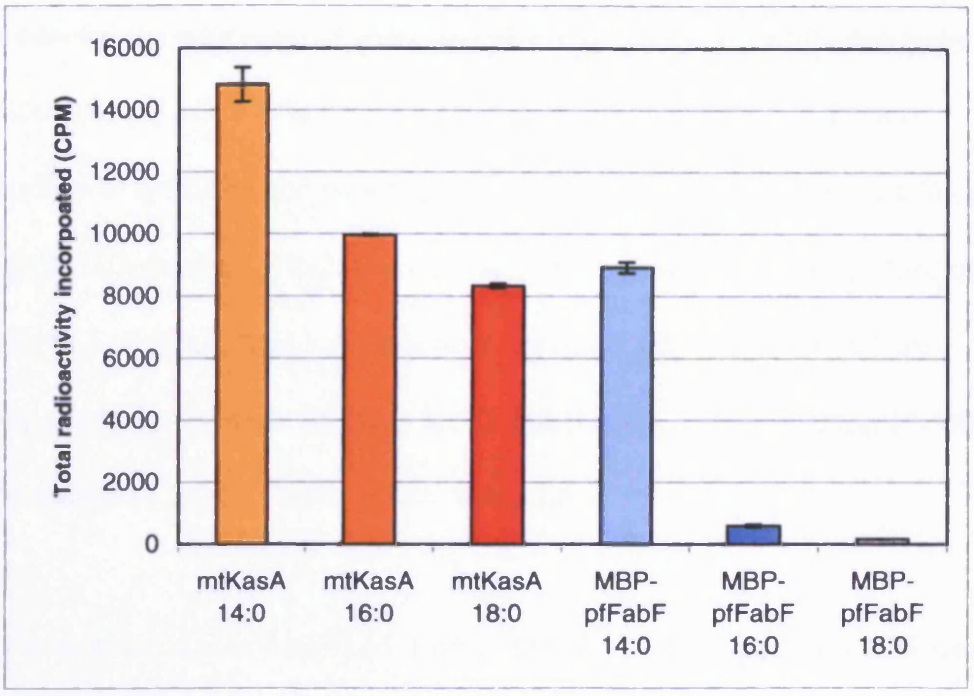


Figure 4.4 The acyl-ACP substrate specificity of MBP-pfFabF

The radioactivity of acyl-ACP products using mtKasA and myristoyl-ACP (14:0-ACP)(■), palmitoyl-ACP (16:0-ACP) (■) and stearoyl-ACP (18:0-ACP) (■) or MBP-pfFabF and myristoyl-ACP (14:0-ACP) (■), palmitoyl-ACP (16:0-ACP) (■) and 18:0-ACP (■) as the acyl-ACP substrates. The amounts of mtKasA and MBP-pfFabF used in each assay were 150ng and 300ng respectively. The activity of both mtKasA and pfFabF were measured using the radiochemical assay as described in Section 2.2.24.

4.4 Thiolactomycin Sensitivity of MBP-pfFabF

4.4.1 A general investigation of MBP-pfFabF sensitivity to Thiolactomycin

Thiolactomycin (TLM) is a thiolactone that is a specific inhibitor of Type II β-ketoacyl-ACP synthases and was originally isolated from the soil fungus *Nocardia* species 2-200 (Oishi *et al.*, 1982; Sasaki *et al.*, 1982; Noto *et al.*, 1982; Miyakawa *et al.*, 1982). In *E. coli*, all three KAS enzymes are inhibited by TLM with FabF exhibiting the highest sensitivity (IC_{50} 6μM), FabB is less sensitive (IC_{50} 25μM) and FabH minimally so (IC_{50} 110μM) (Price *et al.*, 2001).

Assays that contained 50μM of TLM, a concentration that inhibits over 50% of *E. coli* FabF and FabB activity were prepared to see if the MBP-pfFabF enzyme was highly or minimally sensitive to TLM. In these assays, naturally produced TLM (the (5R) enantiomer only) purified from cultures of the *Nocardia* fungus was kindly provided by Charles Rock (St. Jude Children's Research Hospital, Memphis) and ecFabB was used as a positive control for the inhibitory activity of TLM. The results (Figure 4.5) indicate that MBP-pfFabF is highly sensitive to TLM, as only 6.4% of activity remained when the assay was performed in the presence of 50μM TLM compared to the control containing no TLM. The MBP-pfFabF enzyme also appeared to be slightly more sensitive than ecFabB (9.7% activity remained) considering the amounts of enzymes present. The nature of the sensitivity of MBP-pfFabF to TLM is similar to bacterial FabF and FabB enzymes and unlike the low sensitivity observed with plant and bacterial FabH proteins (Price *et al.*, 2001, Jones *et al.*, 2000 and 2003, Prigge *et al.*, 2003).

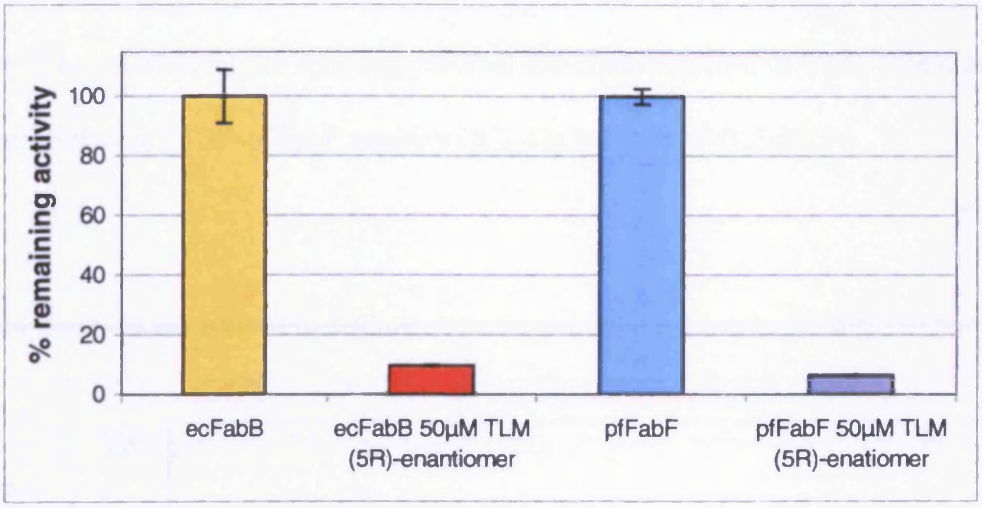


Figure 4.5 Inhibition of ecFabB and MBP-pfFabF with 50μM thiolactomycin

The incorporation of radiolabelled malonyl-ACP with no TLM ecFabB (■) and MBP-pfFabF(■) and in the presence of 50μM TLM, ecFabB (■) and MBP-pfFabF (■). Each assay was performed in duplicate and errors are shown the range of results from mean. Assays contained either 6.25ng of ecFabB or 1μg of MBP-pfFabF and the activity was measured using the radiochemical assay as described in Section 2.2.24.

4.4.2 Determination of the IC_{50} of TLM against MBP-pfFabF

In order to get a more accurate estimation of the sensitivity of MBP-pfFabF to TLM, assays containing variable concentrations of TLM were performed. Briefly, a 2mM stock of TLM was serially diluted in 40% DMSO to produce 1mM, 500μM, 200μM, 100μM, 50μM and 25μM stock solutions. When 2μl of each solution was added to the total assay volume of 40μl, the final TLM concentrations lay between 100μM-1.25μM. Assays containing no TLM (but with 2% DMSO) were used to

estimate the uninhibited activity. The final concentration of DMSO in all the assays was 2% and this solvent had a negligible effect on the enzyme activity (results not shown). The results (Figure 4.6) suggest that the concentration of TLM that caused 50% inhibition of MBP-pfFabF activity (IC_{50}) is between 3-3.5μM.

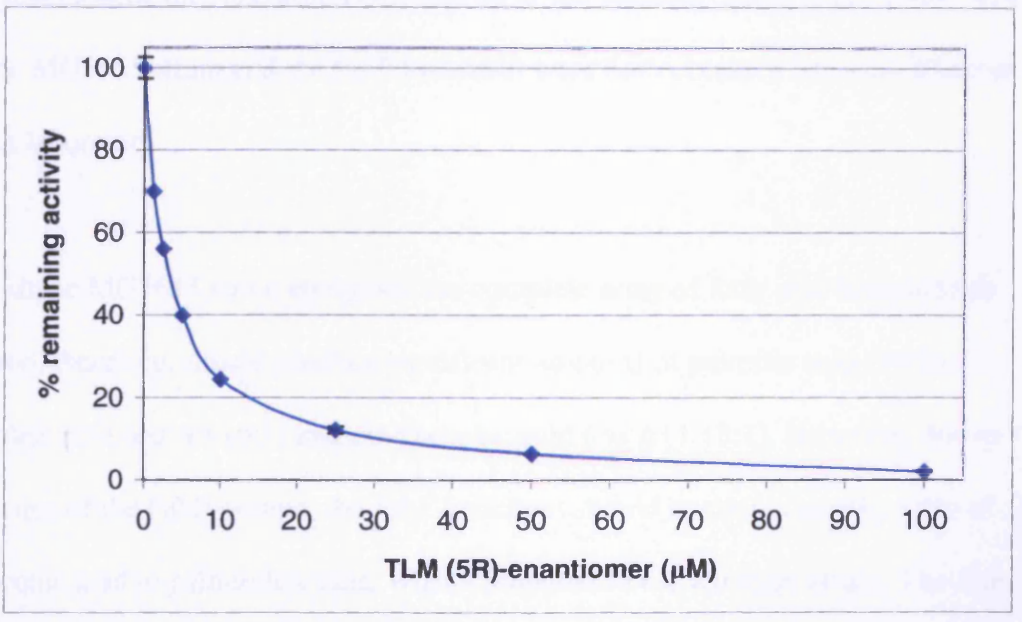


Figure 4.6 IC_{50} determination of TLM against MBP-pfFabF

Each assay was performed in duplicate and errors are shown as the range of results from the mean. Assays contained 300ng of MBP-pfFabF and the activity was measured using the radiochemical assay as described in Section 2.2.24.

These data indicate that the *P. falciparum* long chain β-ketoacyl-ACP synthase is inhibited by thiolactomycin and its sensitivity was comparable to that of *E. coli* FabF (IC_{50} = 6μM) (Price et al., 2001), which indicate that it may be a FabF-like enzyme.

4.5 Complementation of an *E. coli fabf* knockout with MBP-pfFabF

In order to determine if the *P. falciparum* long chain condensing enzyme was a FabF-like enzyme, complementation studies were performed using FabF knockouts of the *E. coli* strain MG1655. The Blattner laboratory (Wisconsin University) produces clones of this *E. coli* strain containing knockouts of all non-essential genes (Kang *et al.*, 2004). Loss of the FabF gene was successfully achieved by the insertion of a kanamycin cassette on the chromosomal DNA of the bacteria (Kang *et al.*, 2004). The wildtype MG1655 strain and the *fabF* knockout were both obtained from the Blattner research laboratory.

The wildtype MG1655 strain contained the complete array of fatty acid biosynthesis genes and, therefore, should produce significant amounts of palmitic acid (16:0), palmitoleic acid (*cis* Δ9 16:1) and *cis*-vaccenic acid (*cis* Δ11 18:1). However, due to the absence of the FabF protein, the *fabF* knockout should contain a smaller ratio of *cis*-vaccenic acid to palmitoleic acid, when compared to the wildtype strain. The aim of this study was to determine if the pMAL-*pffabF* expression plasmid could increase the amount of *cis* Δ11 18:1 (*cis*-vaccenic acid) incorporated into membrane lipids of the *fabF* knockout strain.

Competent MG1655 *fabF* knockout cells were produced as described in Section 2.2.9. These cells were transformed with the pMAL-*pffabF* plasmid as described in Section 2.2.10 and plated onto LB agar containing kanamycin and ampicillin. At the same time MG1655 wildtype and *fabF* knockout strains were streaked on to agar plates containing no antibiotics and kanamycin, respectively. The three agar plates containing the wildtype, *fabF* knockout and *fabF* knockout containing pMAL-*pffabF*,

were incubated at 37°C overnight. Single colonies from each of the three bacterial strains were used to inoculate 10ml each of M9 minimal medium (Section 2.1.3) containing the appropriate antibiotics. These cultures were grown overnight at 37°C in a shaking incubator and, the following morning, 500µl used to inoculate 10ml of freshly prepared M9 minimal medium. All of the cultures were grown at 37°C in a shaking incubator until the *fabF* knockout cells containing the pMAL-*pffabF* plasmid reached an D_{600} of 0.4 before expression of pMAL-*pffabF* was induced by the addition of IPTG to a final concentration of 0.4mM. As a control, 0.4mM IPTG was also added to the wildtype and *fabF* knockout strains. In an attempt to improve the expression of the MBP-pfFabF fusion protein, all of the cultures were placed in a shaking incubator at 20°C (Section 3.12). After the cultures had been incubated at 20°C for 24h, the D_{600} of each culture was determined. The cells were harvested and the membrane lipids extracted and analysed as described in Section 2.2.28. The major fatty acids extracted from the membrane lipids of each culture are shown in Table 4.1.

The percentages of *cis* Δ9 16:1 and *cis* Δ11 18:1 fatty acids, described in Tables 4.1 and 4.2, incorporated the proportions of *cis*-9, 10-methylene hexadecanoic acid and *cis*-11, 12-methylene octadecanoic acid (lactobacillic acid), respectively. These two cyclopropane fatty acids are synthesised directly from *cis* Δ9 16:1 and *cis* Δ11 18:1 via modification of the *cis* bond (For a review see Grogan and Cronan, 1997). They appear to be essential the survival of some bacteria species if a sudden drop in pH occurs and accumulate in bacteria cultures that reach stationary phase (Chang and Cronan, 1999).

Strain	Palmitate (16:0) (%)	Palmitoleate[a] (16:1) (%)	*cis*-vaccenate[b] (18:1) (%)
Wildtype	40.9	28.6	26.8
fabF knockout	42.2	48.9	4.8
fabF knockout pMAL-pfFabF	40.0	49.6	7.0*

Table 4.1 The fatty acid composition of MG1655 strains grown in M9 medium cell membranes

Percentages are shown as the proportion of the total extracted fatty acids from cultures of these cells grown for 24h at 20°C in M9 minimal medium. [a] The percentage for palmitoleate incorporates the proportion of *cis*-9, 10-methylene hexadecanoic acid, the cyclopropane derivative of this unsaturated fatty acid. [b] The percentage of *cis*-vaccenate incorporates the proportion of *cis*-11, 12-methylene octadecanoic acid, the cyclopropane derivative of this unsaturated fatty acid.
* Statistically significant when using in a two-tailed T-test with a t value = 0.0239.
Small amounts of myristate (14:0) and stearate (18:0) were also observed in the membrane lipids. An example of an annotated GC trace of wildtype *E. coli* is given in Appendix 5.

The proportion of palmitate within the fatty acids extracted from the membrane lipids of each strain was relatively constant. However, as expected the proportion of the unsaturated fatty acids and their cyclopropane derivatives differed significantly between the wildtype and both of the *fabF* knockout strains. A large reduction in the amount of *cis*-vaccenate and an increase in the proportion of palmitoleate was observed between the wildtype and *fabF* knockout strains. It appeared that some complementation of the *fabF* knockout had been achieved in the cells containing the pMAL-*pffabF* plasmid. The average amounts of *cis*-vaccenate observed in these

different samples (n=3) may appear to be similar. However, when the percentages were normalised, a two-tailed T-test suggested that there is a 97% confidence that the differences in the data sets did not arise by chance. Complete complementation of the fatty acid composition was not achieved, because a large difference between the amounts of 16:1 and 18:1 fatty acids in the wildtype and *fabF* knockout pMAL-*pffabF* strains was evident.

The rather modest increase in *cis*-vaccenic acid may be due to the low level of expression of MBP-pfFabF in the MG1655 *fabF* knockout strain, which is not optimised for protein expression studies. Furthermore, changes in the membrane fatty acid composition in *E. coli* are modulated by incorporation of *de novo* fatty acids into new membrane lipids, rather than alteration of fatty acids already attached to the membrane (Cronan, 2002). Between the time that pMAL-pfFabF expression was induced and the harvesting of the cultures, the *fabF* knockout pMAL-pfFabF cells had gone through approximately two cell divisions. This may limit the amount of *cis*-vaccenate introduced into cell membrane lipids.

The experiments described above were then repeated in LB medium instead of M9 minimal medium. In LB medium, bacteria often display accelerated growth thus allowing more cell divisions to occur in shorter space of time. In these experiments IPTG was added to a final concentration of 1mM when the cultures were at an D_{600} = 0.2. The cultures were placed at 20°C in a shaking incubator for 24h before the cells were harvested, the lipids extracted and fatty acid methyl esters (FAMEs)

analysed as described in Section 2.2.28. The proportions of fatty acids in the bacterial lipids are described in Table 4.2.

Strain	Palmitate (16:0) (%)	Palmitoleate[a] (16:1) (%)	*Cis*-vaccenate[b] (18:1) (%)
Wildtype	39.4	30.0	25.6
fabF knockout	39.2	52.3	3.0
fabF knockout pMAL-*pffabF*	37.3	50.2	8.3*

Table 4.2 The fatty acid composition of MG1655 strain grown in LB medium cell membranes

Percentages are shown as the proportion of the total extracted fatty acids from cultures of these cells grown for 24h at 20°C in LB medium. [a] The percentage for palmitoleate incorporates the proportion of *cis*-9, 10-methylene hexadecanoic acid, the cyclopropane derivative of this unsaturated fatty acid. [b] The percentage of *cis*-vaccenate incorporates the proportion of *cis*-11, 12-methylene octadecanoic acid, the cyclopropane derivative of this unsaturated fatty acid. * Statistically significant when using in a two-tailed T-test with a t value = 0.0014. Small amounts of myristate (14:0) and stearate (18:0) were also observed in the membrane lipids.

The fatty acid composition of the MG1655 cells shown in Table 4.2 showed that cultures of the *fabF* knockout pMAL-*pffabF* cells contained approximately two and a

half times the amount of *cis*-vaccenic acid compared to that seen in *fabF* knockout cells without the plasmid. These percentages were normalised by dividing by the smallest percentage of *cis*-vaccenate (2.72%) found in a *fabF* knockout culture. This gave the data shown below in Table 4.3.

	Wildtype	FabF knockout	FabF knockout pMAL-*pffabF*
Culture 1	10.14	1.25	3.08
Culture 2	8.97	1.00	3.42
Culture 3	9.10	1.07	2.62
Mean	**9.40**	**1.11**	**3.04**
SD	*0.52*	*0.10*	*0.33*

Table 4.3 The proportions of *cis*-vaccenate found in cultures of *E. coli* grown at 20°C for 20h

A two-tailed t-test was performed using Microsoft Excel to determine if the values in the *fabF* knockouts were statistically different from *fabF* knockouts containing the pMAL-*pfFabF* plasmid. * The average proportion of *cis*-vaccenate is statistically different from that observed for *fabF* knockouts. t value = 0.0014.

The data presented in Table 4.3 suggests that to the 99% confidence level the amount of *cis*-vaccenate in membrane lipids of *fabF* knockouts containing the pMAL-*pffabF* is significantly higher than the amounts present in *fabF* knockouts. The amount of

cis-vaccenate produced was approximately three times higher in the complemented cells compared to the knockouts

SDS-PAGE and Western blot analysis (Sections 2.2.13 and 2.2.15) was performed on cultures of the *fabF* knockout transfected with the pMAL-*pffabF* plasmid, to assess the expression of the MBP-pfFabF fusion protein (Figures 4.7 and 4.8). In the SDS-PAGE gel (Figure 4.7) no overexpression of any proteins could be seen. However, the Western blot of the same samples (Figure 4.8) using anti-MBP serum (New England Biolabs, Beverly, MA, USA) showed two major bands; the first at approximately 82kDa, as expected for the MBP-pfFabF protein and the second at ~ 42kDa corresponding to free MBP.

The Western blot (Figure 4.8, Page 171) showed the presence of other bands, which suggested that both the MBP-pfFabF fusion protein and the free MBP had been significantly digested in the cell. This is not surprising because, unlike cell lines optimised for protein expression, the MG1655 *fabF* knockout cells are not proteinase deficient. For instance, *ompT*, an outer-membrane proteinase cleaves between sequential basic amino acids and an analysis of the MBP-pfFabF protein showed that it contained four lysine-lysine motifs (Okuno *et al.*, 2002).

The Western blot also showed that only a rather small amount of the MBP-pfFabF fusion protein is observed in the soluble fraction. This may be due to the lack of a plasmid expressing rare codons (as discussed in Section 3.10) or because the cells have not been optimised for protein expression. The expression of the MBP-pfFabF protein in a recombinant form in *E. coli* has proved problematic (Sections 3.9-3.12)

and the yields of soluble, active protein were low, even when optimised. The low-level of MBP-pfFabF expressed in the *fabF* knockout cells, is likely to have limited the amount of palmitoleate elongated to *cis*-vaccenate and this could explain why a relatively small rise in *cis*-vaccenic acid is seen in these cells. However, the rise in *cis*-vaccenate was statistically significant and suggested that, like the FabF (KASII) proteins of bacteria and plants, the MBP-pfFabF protein is able to elongate palmitoleate to *cis*-vaccenate. Combined with the data obtained on the TLM sensitivity of this enzyme (Section 4.5), these complementation studies provide the first biochemical evidence that the condensing enzyme identified in *P. falciparum* is most likely to be a FabF enzyme.

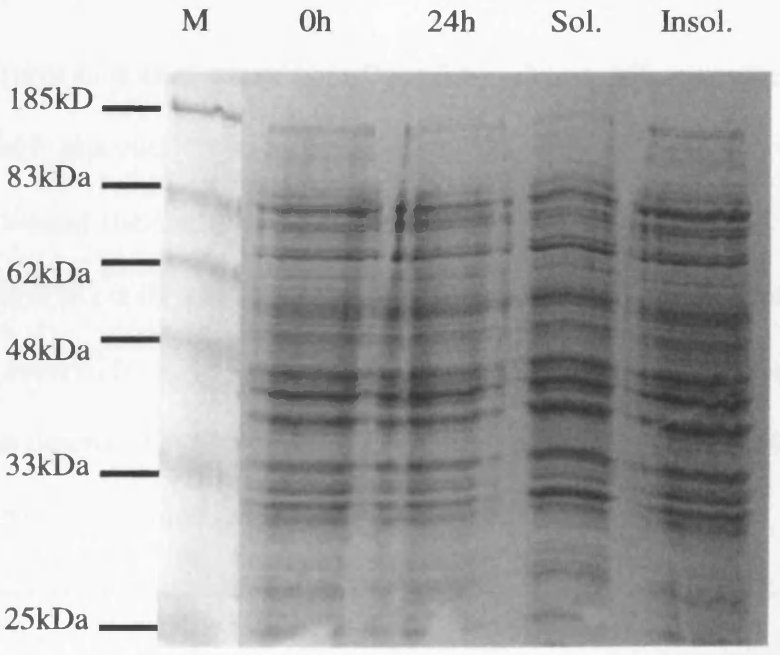


Figure 4.7 SDS-PAGE analysis of *E. coli fabF* knockout cells transfected with the pMAL-*pffabF* plasmid
Samples were removed from the culture before the addition of 1mM IPTG at 0h and at the indicated timepoint post-induction. Samples were prepared as described in Section 2.2.12 and 2.2.13. The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.

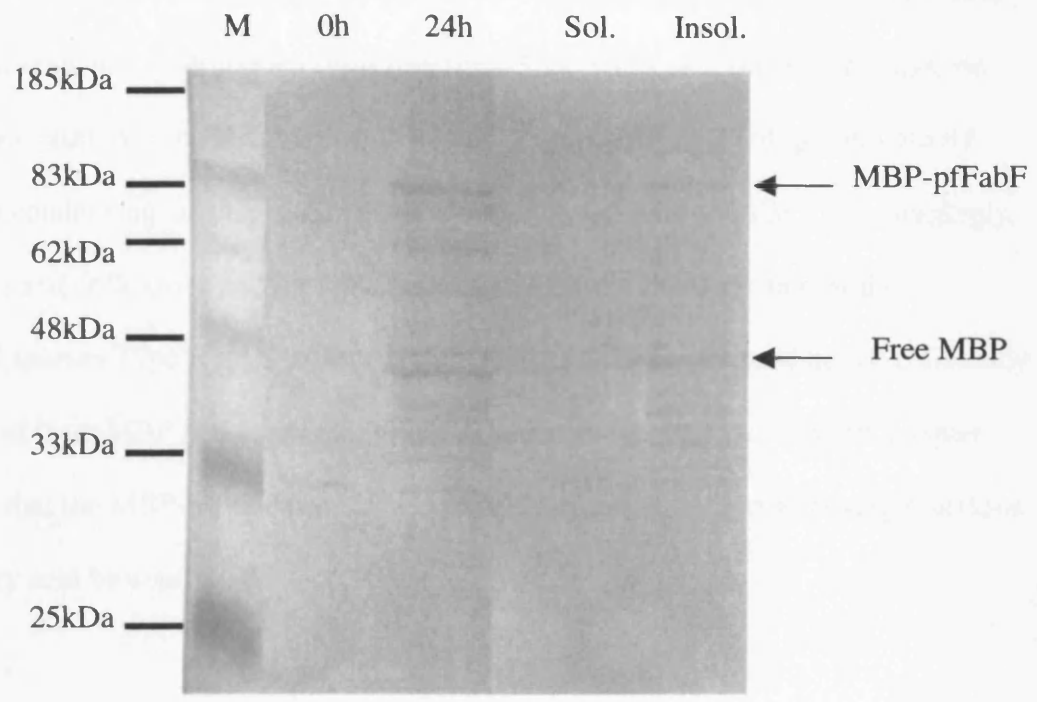


Figure 4.8 Western blot analysis of *E. coli fabF* knockout cells transfected with the pMAL-*pffabF* plasmid

M = molecular weight standards. Samples were removed from the culture before the addition of 1mM IPTG at 0h and at 24h post-induction. Soluble and insoluble fractions were prepared from the culture at the 24h post-induction timepoint. Samples were prepared as described in Sections 2.2.12-13 and the Western blot performed as described previously (Section 2.2.15).

4.6 Discussion

The *Plasmodium falciparum* β-ketoacyl-ACP synthase, pfFabF, was purified from *E. coli* cells in a recombinant form (Sections 3.11-3.13), as a fusion with maltose binding protein at the N-terminus. It was not known if the pfFabF protein would retain condensing enzyme activity whilst remaining bound to MBP. Encouragingly, Prigge and colleagues had previously reported that the FabD protein of the *P. falciparum* Type II FAS pathway retained activity when it could not be efficiently cleaved from MBP (Prigge *et al.*, 2003). *In vitro* studies presented in this chapter show that the MBP-pfFabF protein was capable of performing condensing reactions in fatty acid biosynthesis.

FabF proteins have several conserved residues, of which many have an important role in forming and maintaining the correct structure. The first of these in the primary sequence is a threonine residue, first identified in *Synechocystis* sp. FabF enzyme (Moche *et al.*, 2001), to which Thr7 of the mature pfFabF protein is analogous. This residue interacts with the side chain of another highly conserved residue, Glu245 (Glu247 in the *P. falciparum* enzyme) and combined, these residues are thought to be essential for the structural integrity of the protein (Moche *et al.*, 2001). Despite the closeness of Thr7 to the N-terminus of the mature pfFabF protein, there are 30 amino acids between Thr7 and the C-terminus of the maltose binding protein. It appears that this distance is sufficient so that MBP does not interfere with the correct folding of the pfFabF protein. Furthermore, since the ~42kDa MBP did not appear to affect the formation of an assumed pfFabF homodimer. Bacterial KAS enzymes form homodimers that are essential for condensing enzyme activity (Garwin *et al.*, 1980; Edwards *et al.*, 1997; Olsen *et al.*, 1999; Huang *et al.*, 1998; Moche *et al.*, 2001).

Whether the N-terminally fused MBP affected the specific activity of the enzyme cannot be determined. However, the calculated specific activity of the MBP-pfFabF protein was 25 +/- 0.48 pmoles/min/μg with C14:0-ACP (myristoyl-ACP) as the acyl-ACP substrate. The specific activities of *E. coli* FabB and FabF enzymes using the same C14:0-ACP substrate are reported to be 3pmoles/min/μg and 7pmoles/min/μg (Edwards *et al,*. 1997). Therefore, with regards to the C14:0-ACP substrate, the MBP-pfFabF protein exhibits a higher specific activity than both of the *E. coli* β-ketoacyl-ACP synthases. Assays of the *E. coli* FabF enzyme with C6:0, C8:0, C10:0, C12:0 acyl-ACP substrates show that it exhibits a similar specific activity towards all of the substrates ranging from 23-26pmoles/min/μg (Edwards *et al,*. 1997). So MBP-pfFabF appears to have a similar specific activity towards C14:0-ACP compared to that of ecFabF towards shorter, saturated acyl-ACP substrates. The specific activities of the *Mycobacterium tuberculosis* condensing enzymes, KasA and KasB, using C16:0-ACP as the acyl-ACP substrate were determined as 96pmol/min/μg and 29pmol/min/μg, respectively (Kremer *et al,*. 2002; Schaeffer *et al,*. 2001). Therefore, the MBP-pfFabF protein has a specific activity when using a C14:0-ACP substrate similar to the reported activities of bacterial Type II FAS condensing enzymes with acyl-ACP substrates of varying size.

These results may indicate that the fusion of MBP at the N-terminus of pfFabF does not affect the activity, or disrupt correct folding, of the condensing enzyme. Re-cloning of the *pffabF* gene into a vector that encodes an alternative protease site between MBP and pfFabF proteins is being considered as further work to this project.

This may allow separation of the two proteins and allow the determination of the specific activity of the purified pfFabF without its N-terminal fused protein.

A limited series of long-chain saturated acyl-ACPs was used to determine the substrate specificity of the MBP-pfFabF protein. The protein showed an enzyme activity fifteen-fold higher for myristoyl-ACP (C14:0) compared to that for palmitoyl-ACP (C16:0) and a fifty-fold increase in activity for myristoyl-ACP (C14:0) compared to that with the stearoyl-ACP (18:0) substrate. These data may indicate that, *in vivo*, the final product of the saturated fatty acid biosynthesis pathway in *P. falciparum* would be palmitate (C16:0).

Despite the evidence that palmitate may be the final saturated acyl product of Type II FAS, a bimodal distribution of C16 and shorter acyl products cannot be ruled out. Recent evidence suggests that the Type II β-ketoacyl-ACP synthase found in the mitochondria of *Arabadopsis thalania* predominantly produces a bimodal distribution of C8 and C14-16 acyl chains (Yasuno *et al.*, 2004). Very little is known about the Type II FAS system in mitochondria, although the pathway appears to be conserved in a wide range of organisms as homologues of the mitochondrial KAS enzyme of *Arabadopsis thalania* have been found identified in yeast, mammals and plants (Figure 3.2) (Yasuno *et al.*, 2004). There is no evidence that the malaria parasite contains a second Type II FAS system in its mitochondria. The synthesis of C8 acyl chains, a substrate for lipoic acid synthesis, by this enzyme suggests that it plays a major role in the synthesis of lipoic acid precursors in the mitochondria of *A. thalania*. Enzymes responsible for lipoic acid synthesis in *P. falciparum* are targeted to the apicoplast (Wrenger and Muller, 2004) and the pyruvate dehydrogenase complex,

which requires lipoic acid as an essential co-factor, is predicted to be targeted to the same organelle (Ralph *et al.*, 2004). Therefore, it seems plausible that the pfFabF enzyme may be involved in the synthesis of lipoic acid precursors in the apicoplast. To investigate this area further, a collaboration with Charles Rock (St. Jude Children's Hospital, Memphis) has been organised with the aim to synthesise and utilise C4-, C6-, C8-, C10- and C12-ACPs in assays containing MBP-pfFabF. A comparison of the relative specific activities of pfFabF with these substrates should provide information on the distribution of saturated acyl chains produced by the Type II FAS system of *P. falciparum*.

The ability of the MBP-pfFabF to participate in the synthesis of unsaturated fatty acids was also investigated. In Gram-negative enterococciae, the conversion of palmitoleic acid (C16:1) to *cis*-vaccenic acid (C18:1) at low temperatures is catalysed by FabF, in contrast the FabB enzyme lacks this activity. Complementation studies of an *E. coli* mutant lacking the *fabF* gene showed that at low temperatures the MBP-pfFabF protein was also capable of converting palmitoleic acid to *cis*-vaccenic acid, suggesting that the *P. falciparum* protein has been correctly assigned as a FabF-like enzyme. Whilst these experiments were only performed in triplicate a two-tailed T-test predicted that there can be over 99% confidence that this increase in *cis*-vaccenate did not arise by chance.

Like plant and bacterial elongating condensing enzymes MBP-pfFabF is sensitive to inhibition with the fungal metabolite thiolactomycin. A review of the scientific literature suggested that the IC_{50} of 3-3.5μM observed with MBP-pfFabF may be the lowest reported IC_{50} value observed for any enzyme inhibited by TLM. This provides

evidence that pfFabF is a physiologically relevant target of TLM in the malaria parasite and reinforces the concept that TLM may be a good lead compound to investigate the inhibition of Type II FAS in *P. falciparum.*

The sensitivity of MBP-pfFabF to TLM can be interpreted in the context of its amino acid sequence. The alignment of FabB and FabF proteins shows a region of identity close to the C-terminus (residues 387-395, *E. coli* FabB numbering) (Figure 3.1). The structure of thiolactomycin (TLM) bound to *E. coli* FabB showed that the highly conserved F390 of this region forms a hydrophobic pocket with P272 that is occupied by the C10 methyl group of thiolactomycin (Figures 1.18 and 1.19) (Price *et al.*, 2001). Jackowski and colleagues (Jackowski *et al.*, 2002) showed that an ecFabB F390V mutant was resistant to TLM due to steric hindrance between the valine side chain and the C10 methyl group of TLM. The pfFabF protein possesses a F390L substitution, as does *B. subtilis* FabF (Figure 3.1), an enzyme which is highly resistant to TLM. These observations may have indicated that pfFabF would be resistant to inhibition by TLM. However, this steric clash observed in the F390V mutant is not present in F390L proteins and the F390L replacement was not thought to be sufficient for inducing TLM resistance (Jackowski *et al.*, 2002). This explains why, despite the replacement of the highly conserved F390 with a leucine residue in the pfFabF enzyme, it remains sensitive to TLM.

Many of the assays described in this Chapter were only performed in duplicate. These assays use quite large quantities of radiolabelled [2-^{14}C]-malonyl CoA and lowering the number of assays reduced the cost and the risk of exposure to the radioactive isotope. Furthermore, these assays are very accurate and reproducible in independent

experiments, with variations from the means consistently less than 5% (Figures 4.1-4.6). Therefore, there is confidence that the data obtained from these assays were a true reflection of the enzyme activity.

Chapter 5

The Cloning and Expression of *Plasmodium falciparum* FabH

5.1 Overview

The *P. falciparum* FabH protein, also termed β-ketoacyl-ACP synthase III (KAS III), is encoded by a 1113 base pair gene located on chromosome 2 of the malaria parasite. Chromosome 2 was the first chromosome to be sequenced and annotated from a malaria parasite genome and led to the discovery of, FabH and ACP, the first proteins indicative of a Type II FAS pathway in *P. falciparum* (Waller *et al.*, 1998). When translated the *fabH* gene encodes a protein of 371 amino acids, of which approximately the first 49 residues correspond to a bipartite targeting sequence, which facilitates transport of FabH to the apicoplast. The FabH enzymes of other organisms catalyse the initial condensation reaction of the Type II FAS pathway, but are not substantially involved in chain elongation. Isolation of naturally-occurring FabH from the parasite apicoplast is not feasible and, like pfFabF (Chapter 3), production of pfFabH was attempted using a recombinant system. Cloning of the gene was problematic and the strategies undertaken are described in this Chapter. Eventually, an N-terminal truncation of pfFabH without the bipartite leader sequence was sub-cloned from a full-length version of the gene (kindly provided by Sean Prigge, Johns Hopkins School of Public Health, Baltimore, USA). The FabH protein was successfully produced in an *E. coli* expression system identical to that described for pfFabF (Chapter 3) and was subsequently purified as a fusion to maltose binding protein. pfFabH was shown to catalyse the initial condensation reaction of Type II FAS biosynthesis and its sensitivity to thiolactomycin was investigated.

5.2 Cloning of the *P. falciparum fabH* gene

At the start of this project, the sequence of the *P. falciparum fabH* gene had previously been identified following the annotation of Chromosome 2 (Gardner *et al.*, 1998, Waller *et al.*, 1998) from the parasite's genome. The *fabH* gene of *P. falciparum* consists of 1116bp intersected by 7 introns and encodes a protein of 371 amino acids that includes an N-terminal initiator methionine residue and a stop in the 3' region . The predicted protein, termed pfFabH, showed significant sequence homology with FabH proteins from bacteria and plants (Figure 5.1). Waller and colleagues observed that the pfFabH protein contained a bipartite leader sequence, consisting of a signal peptide and transit peptide sequence, both of which are essential for successful targeting to the apicoplast (Waller *et al.*, 1998 and 2000). The signal and transit peptide sequences are removed by a proteinase targeted to the apicoplast (van Dooren *et al.*, 2000; Waller *et al.*, 2000) and the processed protein is sufficient for enzyme activity (Waters *et al.*, 2002).

Isolation of naturally occurring pfFabH from the apicoplast organelle was not considered a feasible method of yielding a sufficient quantity of protein for kinetic characterisation or structural analysis. Therefore, production of pfFabH was attempted using a recombinant system in *Escherichia coli*. In order to express the pfFabH protein in a heterologous system, cloning of the *pffabH* gene was attempted.

Amplification of the *pffabH* gene was first attempted using the PCR parameters described in Figure 5.2. As in the amplification of *pffabF*, a ring stage *P. falciparum* cDNA library was used as the template. Two forward primers were designed to complement regions of the coding strand near the 5'region of the gene.

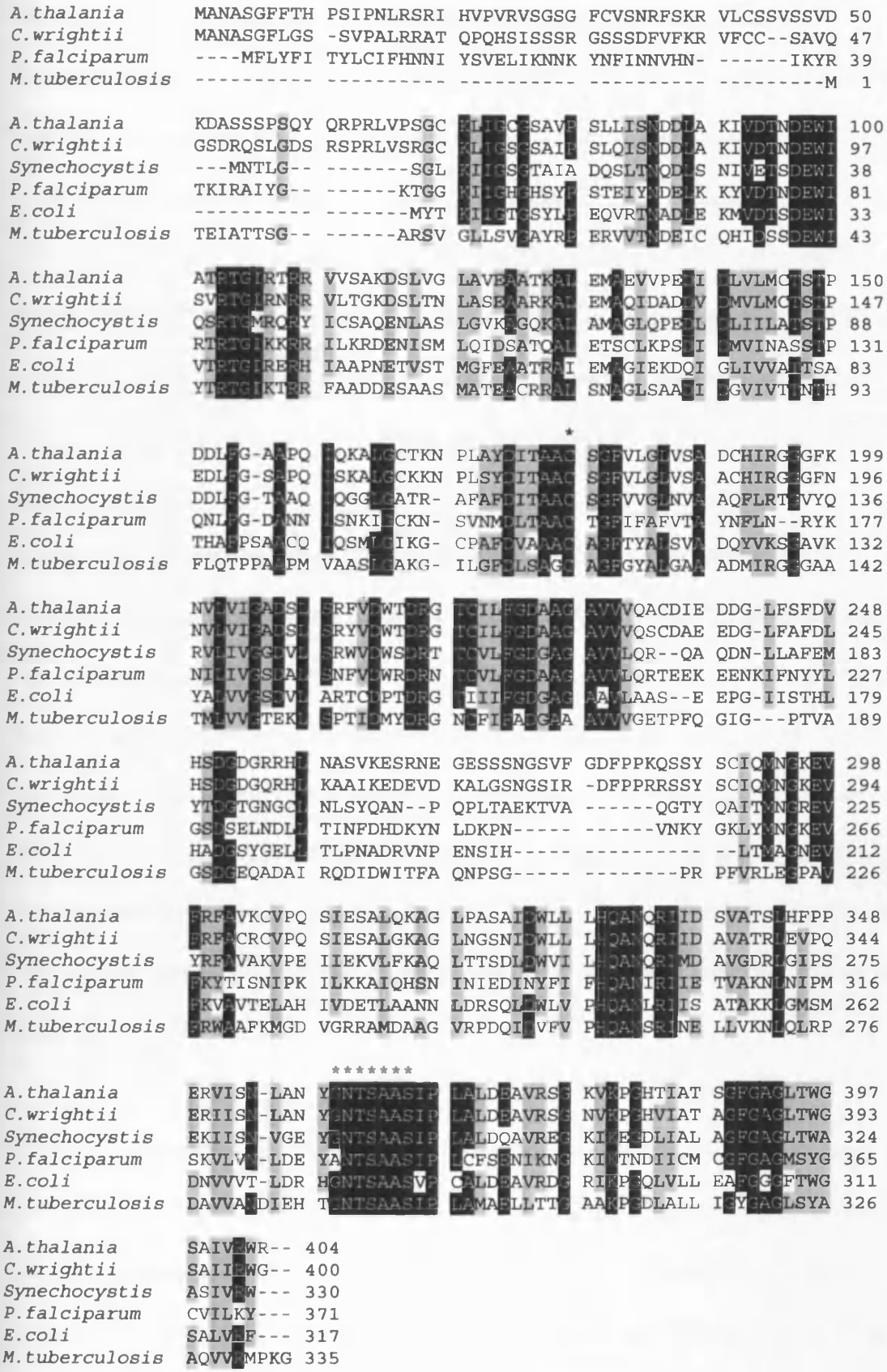


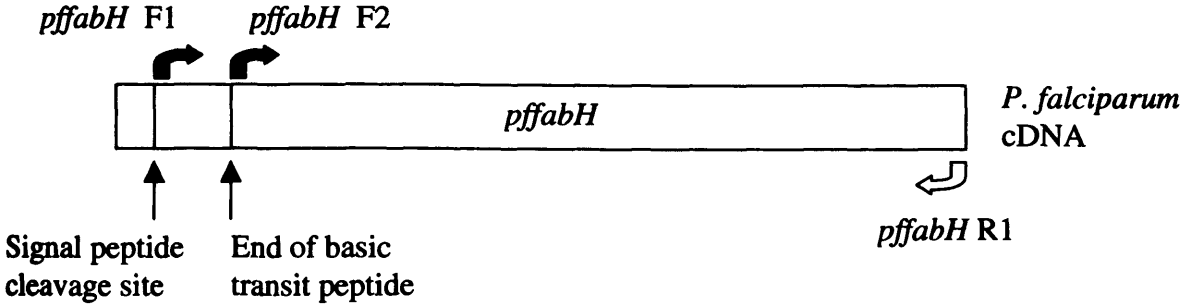
Figure 5.1 Alignment of *P. falciparum* FabH with plant, bacterial and homologues
This alignment was produced using the ClustalW program (Thompson *et al.*, 1994) and formatted using the Chroma program (Goodstadt and Ponting, 2001). The GenBank accession numbers were *Arabadopsis thalania* FabH (P49243); *Cuphea wrightii* (P49244); *Synechocystis sp.* (P73951); *Plasmodium falciparum* (AAC63960); *Escherichia coli* (P24249); *Mycobacterium tuberculosis* (O06399). The active site cysteine is marked with a black asterisk. Regulatory site residues are shown by red asterisks.

The first, *pffabH* F1 complements a region of the gene that is immediately downstream of the region encoding the signal peptide. The second primer, *pffabH* F2, was complementary to the region of the gene immediately downstream of the predicted signal peptide cleavage site. A reverse primer, called *pffabH* R1 (Figure 5.2), was complementary to the non-coding strand of the 3' region and incorporated the naturally occurring stop codon. A PCR containing the *pffabH* F2 and *pffabH* R1 primers should amplify a section of the gene that encodes amino acids 49-371, the mature region of the pfFabH protein (Figure 3.1).

The forward primers contained *Bam* HI restriction endonuclease sites at the 5' terminus and the reverse primer a *Xho* I site at the same position. These sites flank any full length fragment generated from the PCR and, in subsequent cloning steps, should facilitate insertion of the gene in the correct orientation within expression vectors. Initial PCRs were performed using a wide range of conditions to amplify these sections of the identified *pffabH* gene. Alterations in annealing temperature, $MgCl_2$ concentrations, different types of DNA polymerase, the number of cycles and the length of individual cycle steps were unsuccessful in amplifying these regions of the gene.

As described previously, in Section 3.3, the *P. falciparum* genome exhibits a large A+T bias (approximately 80%) which may lead to stalling of the DNA polymerase and PCR amplification of the gene may be aborted. Within the *pffabH* gene the overall A+T bias was found to be 74.2%, with some regions of the gene comprising up to 88% A+T nucleotides. The A+T bias of the *pffabH* gene, is higher than that of

A



B

***pffabH* F1**	5'-**GGATCC**GTAGAATTAATTAAAAATAATAAATATAAT -3'
***pffabH* F2**	5'-**GGATCC**GGAGGTAAAATAATAGGAC -3'
***pffabH* R1**	5'-*CTCGAG*TTAATATTTAAGTATAACGCATCC -3'

C

95°C, 5 minutes

95°C, 1 minute

42°C or 56°C, 1 minute

72°C, 1 minute

30 cycles

72°C, 5 minutes

Figure 5.2. Primers and cycling parameters used in the initial attempt to amplify the *pffabH* gene

A The position of the oligonucleotide primers in relation to the *pffabH* gene.

B Sequences of the oligonucleotide primers. *Bam* HI restriction endonuclease sites are shown in bold and a *Xho* I site in italics. Regions complementary to the sequence of the *pffabH* gene are underlined.

C Cycling parameters used in the initial attempt to amplify the *pffabH* gene.

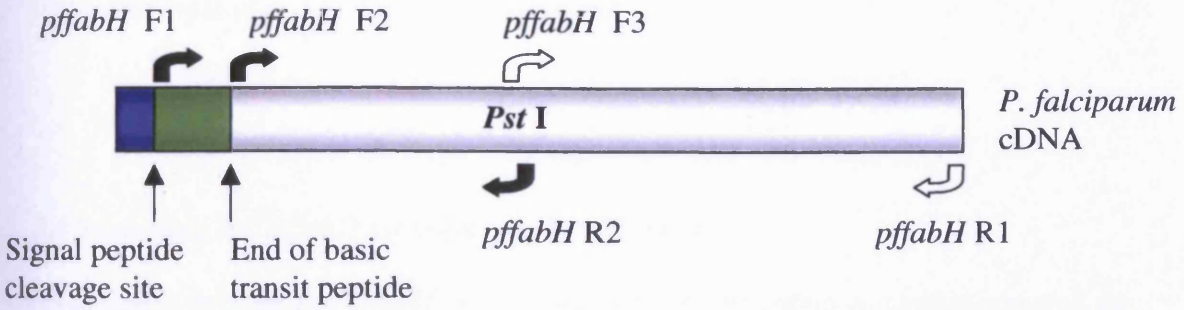
the *pffabF* gene, 70.1% (Section 3.3), which may suggest that the cloning of *pffabH* may be at least as, if not more, problematic than the cloning of *pffabF*.

5.3 Adoption of a new cloning strategy

As described in the cloning of the *pffabF* gene (Sections 3.3 –3.4), in order to improve the chances that the DNA polymerase will extend to the end of the target sequence, the size of the amplicon was reduced. In the *pffabH* gene this was achieved by designing two primer oligonucleotides that flank a single *Pst* I site, present at the centre of the gene (the cleavage site is after nucleotide 618; Appendix 6). The first of these primers, *pffabH* R2 (Figure 5.3), complemented the non-coding strand in this region and was used in conjunction with *pffabH* F2 (Figure 5.3). The second of these primers, *pffabH* F3, complemented the coding strand around the *Pst* I site and was used in conjunction with *pffabH* R1. The two primers surrounding the *Pst* I site are not exact complements of each other (Appendix 6). The *pffabH* F3 primer was designed so that the 3' end of the primer terminates with a guanine base, to improve the annealing efficiency, this meant the *pffabH* F3 primer one base longer than the *pffabH* R2 primer (Figure 5.3).

These primers were then used in two separate PCRs with the *pffabH* F1 and R1 as described above. It was hoped that this technique would generate two fragments that can then be assembled to yield a full length gene (Figure 5.3). Unfortunately, alterations in annealing temperature, $MgCl_2$ concentrations, different types of DNA polymerase, the number of cycles and the length of individual cycle steps were again unsuccessful in amplifying these regions of the gene. The conditions used to

A



B

***pffabH* F3**	5'- GGAGATGCTGCAGGTGCTG -3'
***pffabH* R2**	5'- AGCACCTGCAGCATCTCC -3'

C

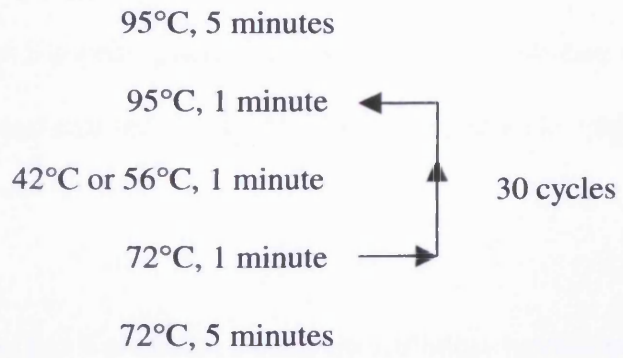


Figure 5.3. Primers and cycling parameters used in the second attempt to amplify the *pffabH* gene

A The position of the oligonucleotide primers in relation to the *pffabH* gene. An inherent *Pst* I site at nucleotide 618 is shown. Both the *pffabH* F3 and *pffabH* R2 complement the region around this site on the coding and non-coding strands, respectively.

B Sequences of the oligonucleotide primers.

C Cycling parameters used in the initial attempt to amplify the *pffabH* gene.

successfully amplify the *pffabF* gene (Section 3.4 and Figure 3.3) in three overlapping fragments were also used but with no success.

5.4 Screening for *pffabH* by colony hybridisation

As attempts to clone the *pffabH* gene using PCR amplification were unsuccessful, an alternative technique was utilised in an attempt to isolate this gene. Colony screening by hybridisation uses the ability of oligonucleotides to anneal and bind to homologous sequences to identify individual clones that contain the gene of interest. As described in Section 2.2.17, a *P. falciparum* cDNA library was used to transform competent *E. coli* DH5α cells (Section 2.1.2), which were plated onto eight LB agar plates containing ampicillin. This allowed a good distribution of approximately 4000 colonies across the plates and ensured that each colony retained a plasmid containing a cDNA fragment.

Colonies from individual plates were blotted onto nitrocellulose membranes and lysed gently. The oligonucleotide primers designed for the PCR amplification of *pffab*H (*pffab*H F1, F2, F3, R1 and R2) were then labelled with radioactive ^{32}P, using [γ-^{32}P]-ATP and polynucleotide kinase to phosphorylate the 3' termini. The radio-labelled oligonucleotide probes were added to the nitrocellulose blots under conditions, which allowed them to anneal to the DNA of lysed bacteria colonies (Section 2.2.17). Under the correct annealing conditions the probes should only bind to DNA sequences that are identical or very similar to themselves. The nitrocellulose blots were incubated with radiographic film to allow any colonies that had hybridised with the oligonucleotide probes to be identified.

When the hybridising temperature was at 55°C there were 25 spots across the 8 plates where the radiolabel was present. Due to the high number of colonies on each agar plate, in some cases several candidate colonies were identified for each radioactive spot. All of the identified colonies were used to inoculate 10ml of LB broth containing ampicillin. The plasmid DNA from each individual colony was then purified (Section 2.2.5) and the DNA sequenced (Section 2.2.9) using M13F and M13R primers. Unfortunately, the sequencing reaction with these primers was not particularly successful and analysis of the sequence chromatograms that were returned, showed that none of the sequences corresponded to the published sequence of *pffabH* mRNA (Genbank™ Accession # AF038929 (Waller *et al.*, 1998); PlasmoDB entry PFB0505c). The sequence of all the returned chromatograms was determined using a BlastN program (Altschul *et al.*, 1997) (www.ncbi.nlm.nih.gov/BLAST). This identified all of the returned sequences as fragments as vector plasmid DNA and suggested that the M13F/R primers were present on the plasmid but in an inappropriate position for the sequencing of cDNA inserts. Sequencing was performed using the gene specific primers *pffabH* F2 and *pffabH* R1. However, the sequence chromatograms showed that all of these sequencing reactions were unsuccessful.

The hybridisation step was performed at 60°C, in order to decrease the chance of non-specific binding. Fewer colonies were identified at this temperature and these colonies were cultured and their plasmid DNA isolated and sequenced, as described above. Once more, none of the sequencing chromatograms showed the presence of any DNA sequence corresponding to any region of the *pffabH* gene. Due to the lack

of success gathered from these sequencing results (and due to the fact that the sequence of the vector plasmid remained unknown) it was decided to use a different approach to clone the *pffabH* gene.

5.5 Engineering of *pffabH* to include *E. coli* preferred codons

To overcome the problems in cloning the *pffabH* gene that have been described previously, a new strategy was implemented. This involved the production of the *pffabH* gene using chemically synthesised oligonucleotides that overlap with each other along the entire length of the *pffabH* gene (Prapunwattana *et al.*, 1996; Pan *et al.*, 1999). A diagram that showing the strategy of this technique is shown in Figure 5.4. Oligonucleotides were designed to alternate between the coding and non-coding strand, which would allow them to anneal to each other using homologous overlapping regions. The missing regions of the gene could then be filled in by priming off the existing oligonucleotides in a PCR (Prapunwattana *et al.*, 1996; Pan *et al.*, 1999).

This technique offers the prospect of engineering novel features into the gene. In particular, the expression of pfFabF showed that the differences in codon usage between *P. falciparum* and *E. coli* can affect overexpression of recombinant proteins (Section 3.10). Thus, the *pffabH* gene was remodelled so that, for each amino acid the codon used most commonly (preferred codon) in *E. coli* was chosen. These changes did not alter the translated protein sequence but it was hoped that expression of the protein in the *E. coli* system might be enhanced.

A

B

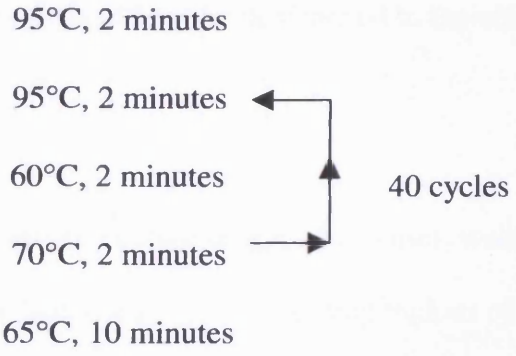


Figure 5.4 The assembly of the pffabH gene with *E. coli* preferred codons

A. Representation of the recursive PCR used to produced a synthetic version of the pfFabH gene. Oligonucleotides are shown in blue text, the black text represents the sequence of the pfFabH gene with codons replaced for preferred *E. coli* usage. In the PCR reaction only the overlapping oligonucleotides are present. The oligonucleotides in this diagram are a representation of those used in the actual synthesis of this gene (see Appendix 7). The red arrows represent priming off of the oligonucleotide by DNA polymerase.

B. The PCR conditions utilised to produce a full-length version of the pfFabH with *E. coli* preferred codons.

The original codons of the *pffabH* gene were replaced with *E. coli* preferred codons (Table 5.1) and the new sequence was searched for any restriction endonuclease sites generated by this change in sequence. The proposed codon changes would have generated a *Bam* HI site within the gene at position 243. This site was removed because *Bam* HI endonuclease was to be used in the cloning of the *pffabH* gene into expression vectors. This was achieved by altering an ATC codon to ATT, this converted the GGATCC sequence of the *Bam* HI site to GGATTC. Both the original ATC and replacement ATT codons encode isoleucine, so the amino acid sequence was unaffected. The percentage usage of the ATT codon compared to all isoleucine codons is relatively high (33.5%) and should not be detrimental to the effective translation of the gene (Table 5.1).

22 overlapping oligonucleotides, with an average length of 72 bases, were designed to span the *pffabH* gene in its entirety (Appendix 7). Overlapping regions of 19-21 nucleotides were implemented, to ensure that the annealing temperatures (T_m) between neighbouring oligonucleotides was within the range (58-62°C). 11 nucleotides complemented the coding strand and these were given the prefix "Coding" and numbered 1-11 from the 5'-3'. A separate group of 11 nucleotides complemented the non-coding strand and were given the prefix "Non-coding" and numbered 1-11 from the 5'-3' of this strand. Therefore, the two oligonucleotides that complemented the 5' terminus on both strands are termed "Coding 1" and "Non-coding 1". These oligonucleotides could be used to amplify the complete gene once it was fully assembled. In order, to anneal the oligonucleotides and produce a double-stranded DNA copy of the gene, a PCR reaction was performed. The PCR was

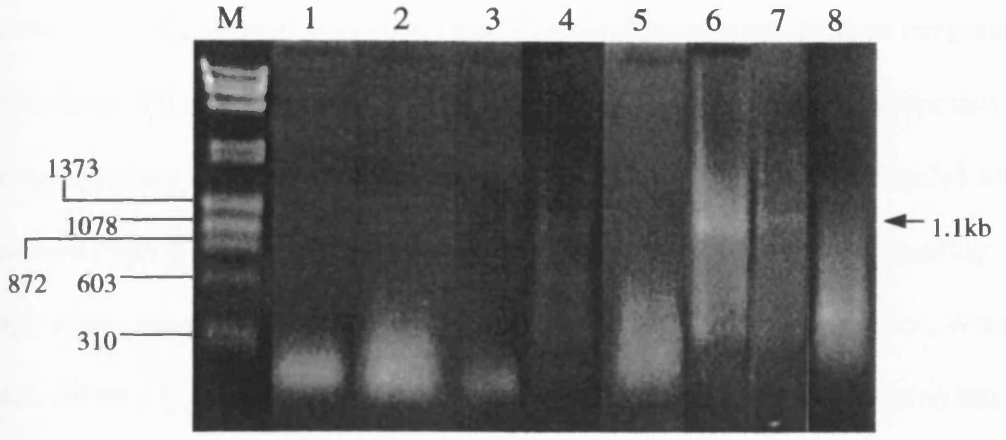
Amino Acid	Codon	% use [a]
Ala	gct	27.5
	gcc	16.1
	gca	24.0
	gcg	32.3
Arg	cgt	27.5
	cgc	16.1
	cga	24.0
	cgg	32.3
	aga	0.6
	agg	0.3
Asn	aat	17.2
	aac	82.8
Asp	gat	46.0
	gac	54.0
Cys	tgt	38.8
	tgc	61.2
Gln	caa	18.6
	cag	81.4
Glu	gaa	73.4
	gag	24.6
Gly	ggt	50.8
	ggc	42.8
	gga	2.0
	ggg	4.4
His	cat	29.8
	cac	70.2
Ile	att	33.5
	atc	65.9
	ata	0.6
Leu	tta	3.4
	ttg	5.5
	ctt	5.6
	ctc	8.0
	cta	0.8
	ctg	76.7
Lys	aaa	78.6
	aag	21.4
Met	atg	100.0
Phe	ttt	29.1
	ttc	70.9
Pro	cct	11.2
	ccc	1.6
	cca	15.3
	ccg	71.9
Ser	tct	32.4
	tcc	26.6
	tca	4.8
	tcg	7.4
	agt	4.5
	agc	24.3
Thr	act	29.1
	acc	53.6
	aca	4.7
	acg	12.7
Trp	tgg	100.0
Tyr	tat	35.2
	tac	64.8
Val	gtt	39.8
	gtc	13.5
	gta	20.0
	gtg	26.8

Table 5.1 The use of each codon for individual amino acids in highly expressed *E. coli* proteins

[a] Percentage usage of the codon relative to all other synonymous codons which code for the same amino acid. Table reproduced from Medigue *et al.*, 1991.

also capable of amplifying the full length gene so cloning of the *pffabH* gene in to the appropriate expression vector could be achieved. In this PCR, the external oligonucleotides (Coding 1 and Non-coding 1) would act as primers for the DNA polymerase and were, therefore, added in excess compared to the internal 20 oligonucleotides. The reaction mix contained 300nmoles of the external primers and 30nmoles of each internal primer. Some of the original conditions were altered (annealing temperature, $MgCl_2$ or $MgSO_4$ concentrations, different types of DNA polymerase, the number of cycles and the length of individual cycle steps) but were unsuccessful in amplifying the full-length gene. However, when visualised on an agarose gel, a faint band of the expected size (~1.1kb) of *pffabH* was observed in reactions that contained final concentrations of either 1μM BSA and 0.5% Tween 20 or 5% DMSO and 0.5% Tween 20 (Figure 5.5). Reactions containing 0.5% Tween 20 without BSA or DMSO did not produce a band of the expected size (data not shown).

As shown in Figure 5.5 the 1.1kb band is within a smeared region of the gel suggesting that DNA fragments of many sizes were present. This may be caused by the production of many PCR products of different sizes, in which case, the PCR conditions needed to be optimised towards the amplification of the target sequence. Subsequently, in an attempt to increase the stringency of the annealing step and to prevent mis-priming of the synthetic oligonucleotides, the annealing temperature of the PCR was raised to 64°C and then 68°C, and separately, the $MgSO_4$ concentrations lowered from its original 8mM concentration to 5mM and 2mM. Both of these alterations decreased the concentration of the 1.1kb fragment when visualised on an agarose gel. Raising and lowering the concentrations of Tween 20, DMSO and BSA also did not improve the production of an individual band of the expected size.



Lane	Sample
1	No additional reagents
2	1μM BSA
3	5% DMSO
4	0.5% Tween 20
5	BSA + DMSO
6	BSA + Tween 20
7	DMSO + Tween 20
8	BSA + DMSO + Tween 20

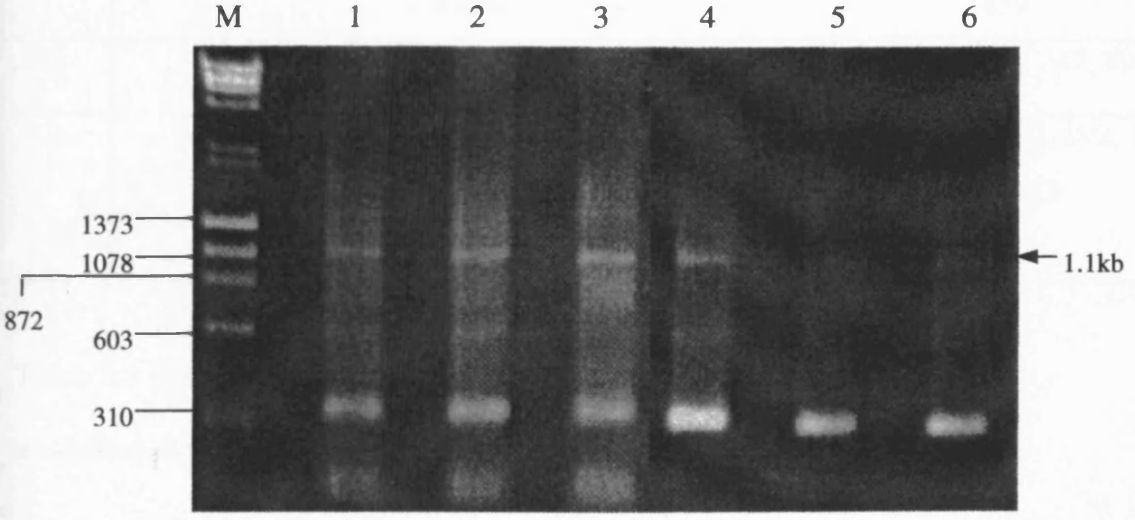
Figure 5.5 The effect of additions to the PCR mixture on the assembly and amplification of *pffabH* using synthetic oligonucleotides

The PCRs were performed under standard conditions. The annealing temperature was 60°C. 10μl of a 50μl reaction were run on a 1.5% agarose gel and visualised under UV light.

To be confident that a single band of the correct size could be excised from the agarose gel, a second PCR was performed that used the original PCR as template. The conditions of the PCR remained the same as described in Figure 5.4, except that in an attempt to reduce the number of spontaneous mutations, the number of cycles was decreased from 30 to 12. Reactions were performed at three different annealing temperatures using Vent and Expand High Fidelity *Taq* DNA polymerases. When visualised on a 1.5% agarose gel these reactions still showed some smearing but the 1.1kb band was more prominent (Figure 5.6) than in the earlier reactions (Figure 5.5). Two other DNA fragments of ~350bp and ~600bp were amplified but these were significantly different in size and caused no interference with the purification of the 1.1kb band.

The DNA polymerases used in both the first and second rounds of PCR did not add an overhang of adenosine residues required for the ligation into the pGEM-T vector. Consequently, 2 cycles with generic *Taq* polymerase (Sigma Chemicals Ltd., Poole, Dorset, UK) were performed to add an adenosine overhang. Following this third PCR, the amplified fragments were separated on an agarose gel, purified from the gel and ligated into the pGEM-T vector (Sections 2.2.2, 2.2.6 and 2.2.7). The ligation mix was transformed into competent *E. coli* DH5α cells as described in Section 2.2.10 and plated onto LB agar plates containing ampicillin. Colonies on the plate were screened for the presence of the 1.1kb insert using a PCR with the vector specific primers M13F and M13R (results not shown). The plasmid DNA from three colonies that contained an insert of the correct size was purified (Section 2.2.5) and sent for double stranded DNA sequencing using the M13F and M13R primers (Section 2.2.8).

The sequence chromatograms from each of the colonies revealed the presence of the full length *pffabH* gene and confirmed the presence of the *E. coli* preferred codons throughout the gene. However, the sequence of the gene from each colony contained a



Lane	Sample
1	Vent polymerase. Annealing temp. 42°C
2	Vent polymerase. Annealing temp. 56°C
3	Vent polymerase. Annealing temp. 60°C
4	Expand Taq. Annealing temp. 42°C
5	Expand Taq. Annealing temp. 56°C
6	Expand Taq. Annealing temp. 60°C

Figure 5.6 A second PCR amplification of *pffabH*

The PCRs were performed under standard conditions. The annealing temperature was 60°C. 10µl of a 50µl reaction were run on a 1.5% agarose gel and visualised under UV light.

significant number of mutations in various positions along the gene. A summary of the mutations is given in Table 5.2.

Clone	No. of mutations	Type of mutations	Position of mutations
1	9	7 single base deletions 1 deletion of 4 bases 1 single base insertion	27, 307, 336, 499, 765, 847, 892 979-982 439
3	7	7 single base deletions	300, 351, 372, 517, 747, 892, 1017
4	8	6 single base insertion 1 double base insertion 1 single base insertion	97, 124, 273, 492, 892, 1038 371-372 380

Table 5.2 Mutations observed in the sequence of *pffabH* gene with *E. coli* preferred codon usage

The relatively large number of mutations were distributed evenly through the sequence of each gene from the three separate colonies. This decreased the chances of piecing together a full length gene from sections of each clone. Interestingly, the single base deletion at position 892 was present in the DNA of all three clones. This suggests that, when the primers were annealing, one of the initial steps involved the Non-Coding 3 oligonucleotide n-1 product that lacked the cytosine base at this position. The large number of single base deletions suggest that the synthetic oligonucleotides contained many n-1 products at different positions and these remained when they became assimilated into the full length gene. If this was the case then the problem was inherent in the starting materials and was not a result of spontaneous mutations produced by the DNA polymerases. Therefore, reducing the

number of cycles and altering the PCR conditions would not be expected to have an effect. The single base mutations and insertions could be corrected by using site directed mutagenesis (SDM). However, the smallest number of corrections needed would be seven for Clone 3. This task would not be trivial as SDM depends on PCR amplification of the gene, which has proved extremely difficult with the *pffabH* gene. Therefore, it was decided that this technique would be very unlikely to yield a correct version the *pffabH* gene.

5.6 Cloning of a truncated version of *pffabH*

During the course of this project, other research groups were also attempting to clone, express and characterise the pfFabH protein of *P. falciparum*. The first paper to describe the cloning and expression of pfFabH was published by Sean Prigge's laboratory in 2002 (Waters *et al.*, 2002). This paper described the cloning of the full length *pffabH* gene using a cDNA library as template. A truncated form of the *pffabH* gene that encodes amino acids 50-371, i.e. without the bipartite signal peptide, was also cloned using an alternative forward primer. The products of these PCRs were digested with *Eco*R I and *Sal* I and ligated into the pMALc2x expression vector. The paper reported that the construct containing the full-length *pffabH* gene produced protein in an insoluble form and that the truncated form was used for all further studies (Waters *et al.*, 2002 and Prigge *et al.*, 2003).

The pMALc2x plasmid containing the full-length construct was obtained from Sean Prigge (Johns Hopkins School of Public Health, Baltimore, USA). In an attempt to amplify the truncated form of *pffabH*, described in the previous paragraph, this

plasmid was used as template in a PCR using two oligonucleotide primers; *pffabH*-trunc F and *pffabH*-trunc R (Figure 5.7). *Pfu* turbo DNA polymerase was used (Stratagene, Amsterdam, The Netherlands) and the cycling parameters are shown in Figure 5.7.

A

***pffabH* trunc F**	5'-GGTGGT**GAATTC**ATGTCCGGAGGTAAAATAATAGGAC -3'
***pffabH* trunc R**	5'-GGTGGT*GTCGAC*TTAATATTTAAGTATAACGCATCCATATG -3'

B

95°C, 5 minutes

95°C, 45 seconds

55°C, 45 seconds

65°C, 2.5 minute

30 cycles

65°C, 10 minutes

Figure 5.7 Primers and cycling parameters used in the initial amplification of the truncated *pffabH* gene

A Sequences of the oligonucleotide primers. An *Eco*R I restriction endonuclease site is shown in bold and a *Sal* I site in italics. Regions complementary to the sequence of the *pffabH* gene are underlined.

B Cycling parameters used to amplify the truncated version of the *pffabH* gene.

When the products from this PCR were separated on an agarose gel (Section 2.6), a single band of approximately 1000bp was observed. The expected size of the truncated form of *pffabH*, including exterior restriction enzyme sites is 999bp. *Pfu* turbo DNA polymerase does not add an overhang of adenosine residues required for the ligation into the pGEM-T vector. Consequently, 2 cycles with generic *Taq* polymerase (Sigma Chemicals Ltd., Poole, Dorset, UK) were performed to add an adenosine overhang. This band was excised from the gel, purified and ligated into the pGEM-T cloning vector (Sections 2.2 and 2.7). The ligation reaction was transformed into competent *E. coli* DH5α cells and plated onto LB agar plates containing ampicillin.

The resultant colonies were screened for the presence of the insert using a PCR containing the vector specific primers M13F and M13R (Appendix 2). Positive colonies were used to inoculate 10ml of LB medium containing ampicillin and the plasmid DNA was subsequently isolated from overnight cultures, as described previously (Section 2.2.5). To ensure that no mutations had occurred during the PCR, the plasmid DNA was sequenced using M13F and M13R primers. Analysis of the sequence chromatograms and subsequent comparison to the published sequence of the *pffabH* gene, showed that the truncated sequence was identical to the corresponding region of the gene. A proximal *Eco*R I site and a distal *Sal* I site (that were included in the oligonucleotide primers), flanked the truncated *pffabH* gene, which therefore, appeared to be in a suitable form for cloning into the pMALc2x vector and for the production of pfFabH protein in a recombinant system.

5.7 Cloning of the *pffabH* Gene into the pMAL Vector System

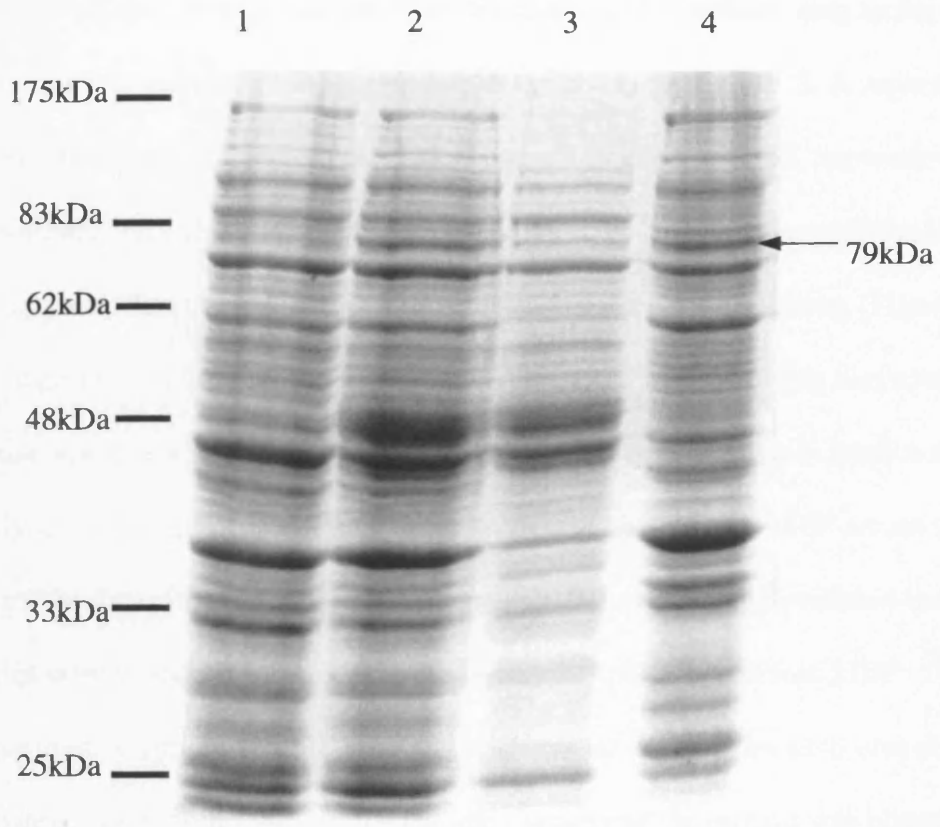
Following the successful production of soluble and active pfFabF protein using the pMAL system, it was decided that the truncated *pffabH* gene should be inserted into the pMALc2x plasmid. Furthermore, expression of truncated pfFabH in the pMAL system and the production of active protein was reported previously (Waters *et al.*, 2002). This suggests that the production of a catalytically active form of pfFabH should be a realistic possibility using this system.

The truncated *pffabH* gene was inserted into the pMALc2x vector using exactly the same method as described in Section 3.11. Briefly, both the pGEM-T plasmid containing the truncated *pffabH* gene and the pMALc2x plasmid were digested with the restriction endonucleases, *Eco*R I and *Sal* I. The digested fragments were separated on and purified from an agarose gel. The pMALc2x vector was then treated with SAP (Section 2.2.4) to prevent any religation of the plasmid during the ligation reaction. A ligation mixture containing equimolar amounts of plasmid and insert DNA was performed under standard conditions (Section 2.2.2.).

The ligation mix was used to transform DH5α *E. coli* and plasmid DNA was purified from several of the resulting colonies (Sections 2.2.10 and 2.2.5). DNA sequencing results of two separate clones suggested that the truncated *pffabH* gene was inserted into the pMALc2x vector in the correct orientation and the CDS of the truncated *pffabH* gene was present in its entirety – this plasmid was termed pMAL-*pffabH*.

5.8 Expression of an MBP-pfFabH fusion protein

By analogy to the production of the pfFabF protein (Section 3.10), the pMAL-*pffabH* plasmid was transformed into competent *E. coli* BL21 (DE3) Codon + RIL cells and plated out onto agar plates containing ampicillin and chloramphenicol. The resulting colonies were used to inoculate 2 litres of LB medium containing the aforementioned antibiotics. Expression of the MBP-pfFabH fusion protein was induced by the addition of IPTG under the conditions described in Section 3.12. These conditions produced a significant amount of MBP-pfFabH fusion protein after 16h (Figure 5.8), which was comparable to that observed with MBP-pfFabF (Figure 3.13). A peptide mass program (Wilkins *et al.*, 1997) predicted the mass of the MBP-pfFabH fusion protein to be 79.3kDa. A reasonable proportion of the fusion protein was retained in the soluble fraction, but the majority remained in an insoluble form (Figure 5.8). In preparation for the purification of the MBP-pfFabH protein, cells were harvested, lysed and the soluble fraction isolated as described in Section 3.12.



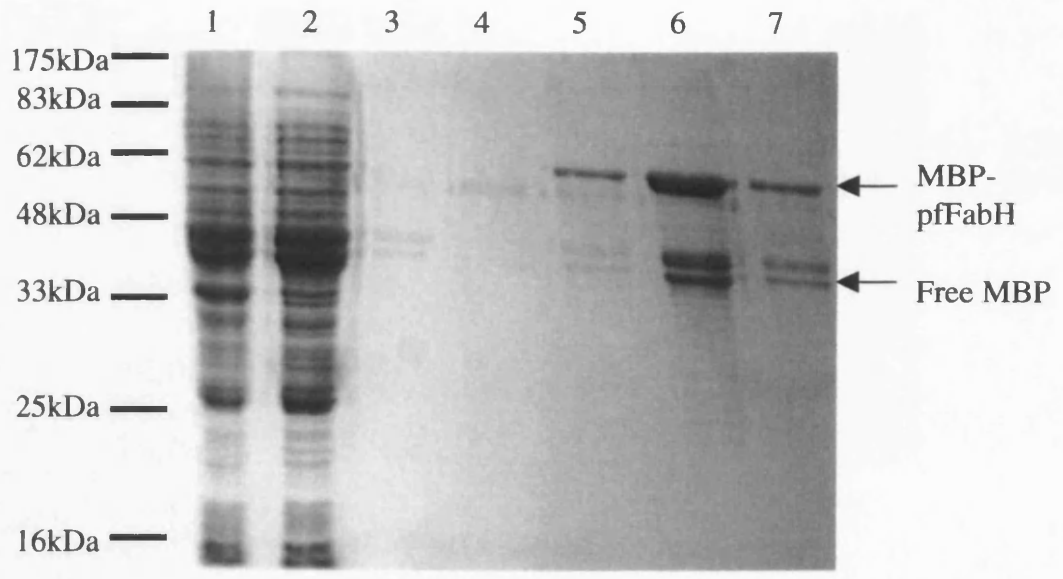
Lane	Sample
1	0h Prior to induction with IPTG
2	16h Post-induction
3	Soluble Fraction
4	Insoluble Fraction

Figure 5.8 Time-course for the induction of pMAL-pfFabH in Codon Plus cells at 20°C

Samples were removed from the culture before the addition of 0.4mM IPTG at 0h and at 16h post-induction. The proteins contained in each sample were separated by SDS-PAGE (12.5% acrylamide) and visualised by staining with Coomassie blue. Purified MBP-pfFabF (Chapter 3) and free MBP (supplied as MBP2* from New England Biolabs, Beverly, MA, USA) were included on the gel as comparisons of the expected size of MBP-pfFabH.

5.9 Purification of MBP-pfFabH in a soluble form

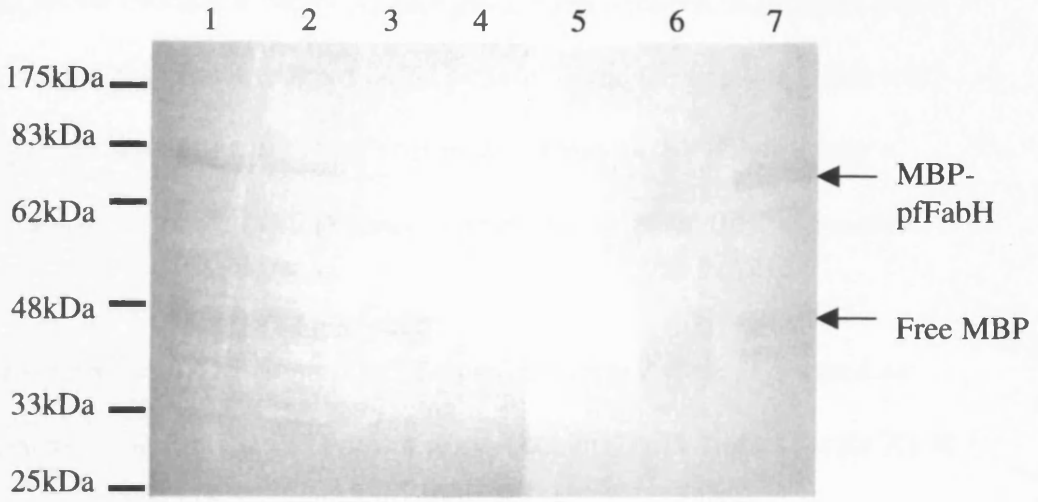
The MBP-pfFabH protein was purified from the soluble fraction using an amylose resin column under conditions previously described in Section 3.13. A separate column of amylose resin was used from the purification of pfFabF. Separation of the MBP-pfFabH protein and free MBP from native *E. coli* proteins proved successful, with only very few contaminating bands present in the eluted fractions (Figure 5.9). The major band in the purified fractions was MBP-pfFabH, showing that a relative increase in concentration had occurred. A Western blot of the same samples revealed that the major contaminating bands were recognised by the anti-MBP serum (New England Biolabs, Beverly, MA, USA) (Figure 5.10), which suggested that these proteins were produced by the degradation of MBP-pfFabH and free MBP. Approximately 1mg of soluble MBP-pfFabH was produced from each litre of culture and, when 1ml fractions were collected, the majority of the protein was eluted across Fractions 2-4.



Lane	Sample
1	Cell supernatant fraction
2	Column flow through
3	Begin of washing steps
4	End of washing steps
5	Eluted fraction 1
6	Eluted fraction 2
7	Eluted fraction 3

Figure 5.9 Affinity purification of MBP-pfFabH protein using an amylose column

The soluble fraction from cells that were harvested 16h post-induction with 0.4mM IPTG (Figure 5.7) (Lane 1) was run through an amylose column and the flow through collected (Lane 2). The column was then washed with Column Buffer to remove any unbound protein (Lanes 3-4). 1ml fractions were eluted by the addition of 10mM maltose to the Column Buffer (Lanes 5-7). The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.



Lane	Sample
1	Cell supernatant fraction
2	Column Flow Through
3	Begin of washing steps
4	End of washing steps
5	Eluted fraction 1
6	Eluted fraction 2
7	Eluted fraction 3

Figure 5.10 A Western blot of the affinity purification of MBP-pfFabH protein using an amylose column

The soluble fraction of cells expressing the MBP-pfFabH fusion protein (Lane 1) was run through an amylose column and the effluent collected (Lane 2). The column was then washed with Column Buffer to remove any unbound protein (Lanes 3-4). 1ml fractions were eluted by the addition of 10mM maltose to the Column Buffer (Lanes 5-7). The proteins contained in each sample were run on a 12.5% acrylamide SDS-PAGE gel, transferred to a nitrocellulose membrane and developed using the anti-MBP serum (Section 2.2.15).

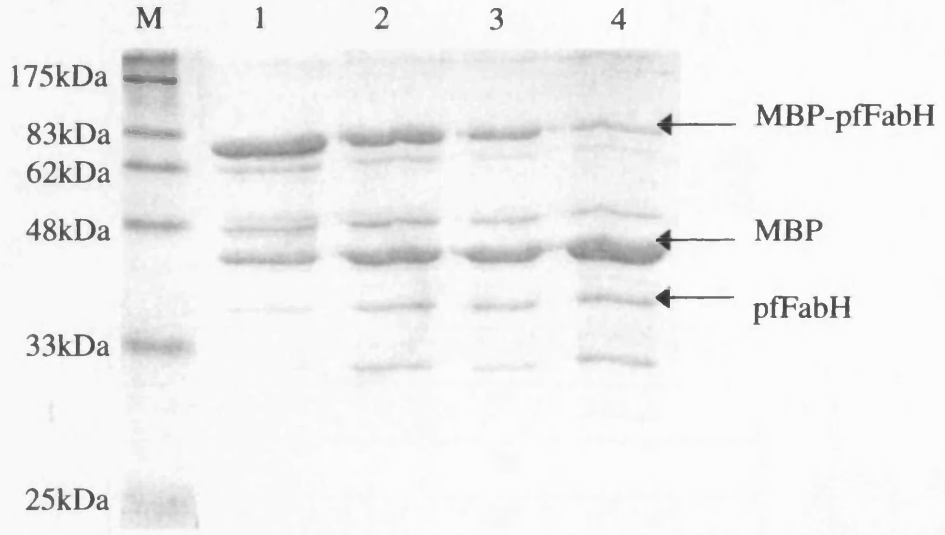
5.10 Attempted cleavage of the MBP-pfFabH fusion

The pMALc2x vector encodes a Factor Xa recognition site between the regions that encode MBP and the N-terminally fused target protein. In the current study, this was designed to facilitate separation of the pfFabH protein from the MBP, with only a small number of vector encoded amino acids remaining attached at the N-terminus.

Cleavage experiments with a small amount of eluted Fraction 2 were performed as described previously (Section 3.14). Samples were incubated with 1μg of Factor Xa at different timepoints and then separated on an SDS-PAGE gel (Figure 5.11). Following a 2h incubation with Factor Xa, the amount of MBP-pfFabH in the reaction had decreased and the amount of free MBP increased. A band of the expected size for free pfFabH was also observed to increase after a 2h incubation. A Western blot of the samples showed that this band was not detected by the anti-MBP serum and suggests that it was not a digestion product of free MBP (Figure 5.12). As the length of incubation was extended, the amount of MBP-pfFabH decreased in proportion to the increase in the quantity of free MBP (Figure 5.11, Lanes 3-4). However, at these times the band corresponding to pfFabH did not increase in size, suggesting that pfFabH had not accumulated. Proteins smaller than the predicted size of pfFabH, were observed on the gel and this suggested that digestion of free pfFabH protein had occurred. These smaller bands are not observed on the Western blot (Figure 5.12) and therefore were not related to MBP.

A search of the pfFabH protein sequence for Factor Xa cleavage sites (IEGR or IDGR) located one sequence with low similarity to the Factor Xa cleavage site – DELK – between residues 20-23. Cleavage at this site would produce a protein

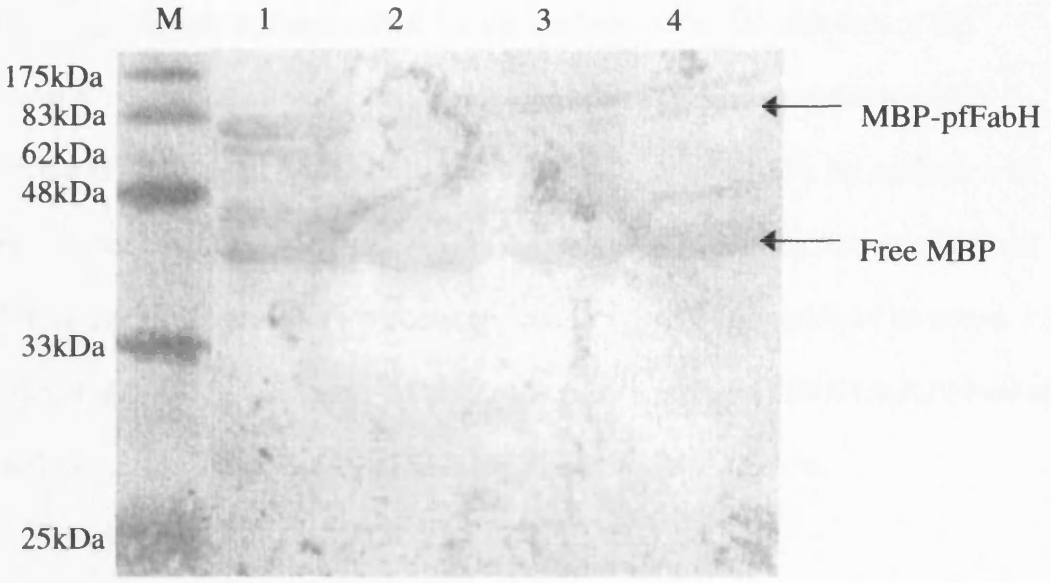
approximately 34kDa in size. The most prominent band that may correspond to a cleavage product of pfFabH appears smaller than the 33kDa marker, however, there is a faint band at ~ 34kDa. In conclusion, it was clear that the pfFabH protein was being degraded in longer incubations and, in order to maximise recovery of the truncated form, a 2h incubation was utilised.



Lane	Length of incubation with Factor Xa
1	0h
2	2h
3	8h
4	24h

Figure 5.11 Timecourse the digestion of the MBP-pfFabH fusion protein with 1µg of Factor Xa

Each reaction contained 10µg of the MBP-pfFabH protein, which was incubated with 1µg of Factor Xa for the times indicated above. The proteins contained in each sample were separated by SDS-PAGE and visualised by staining with Coomassie blue.



Lane	Length of incubation with Factor Xa
1	0h
2	2h
3	8h
4	24h

Figure 5.12 Timecourse of digestion of the MBP-pfFabH fusion protein with 1µg of Factor Xa

M = molecular weight markers. Each reaction contained 10µg of the MBP-pfFabH protein which was incubated with 1µg of Factor Xa for the times indicated above . The proteins contained in each sample were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. The western blot was washed and developed using the anti-MBP serum (New England Biolabs, Beverly, MA, USA) as described in Section 2.2.15.

The relative success of the cleavage of pfFabH from MBP should allow purification of the pfFabH protein, by using a second amylose resin column. The cleavage and purification steps may prove essential for the analysis of the 3D structure of the protein in X-ray crystallography studies. However, due to time constraints of this project it was decided not to pursue the purification of free pfFabH. By analogy with the enzyme activity of the MBP-pfFabF fusion protein (Chapter 4), the separation of MBP from the pfFabH protein may not be necessary for the formation of an active homodimer that displays ketoacyl-ACP synthase activity. Some initial explorations of the catalytic activity of the MBP-pfFabH fusion are described below.

5.11 Initial studies of MBP-pfFabH activity

Following the start of this PhD project but before preparation of recombinant MBP-pfFabH in this study, analyses of the substrate specificity and enzyme activity of pfFabH (KASIII) had already been reported in the literature (Waters *et al.*, 2002, Prigge *et al.*, 2003). These papers indicated that, like plant and bacterial FabH proteins, the pfFabH protein catalyses the initial step of fatty acid synthesis by condensing acetyl-CoA and malonyl-ACP to form β-ketobutyryl-ACP and CO_2 (Figure 1.13). Butyryl-CoA is also an acceptable substrate for pfFabH (Prigge *et al.*, 2003). However, in order to maximise enzyme activity, all inhibitor studies were performed using acetyl-CoA. These reactions also contained an excess of malonyl-CoA, ecFabD and ecACP in order to generate a sufficient concentration of malonyl-ACP for accurate assay (Section 2.2.25).

To determine if the purified MBP-pfFabH retained any KAS activity, an assay containing a large amount (1μg) of the MBP-pfFabH protein was performed as described in Section 2.2.25. As a positive control for the substrates and assay conditions, *E. coli* FabH (ecFabH) was kindly provided by the laboratory of Dr. Charles Rock (St. Jude Children's Research Hospital, Memphis). A significantly smaller amount of ecFabH (50ng), compared to MBP-pfFabH (1μg), was used in the assay because its activity under identical assay conditions had already been determined (Dr. Yong-Mei Zhang, St. Jude Children's Research Hospital, Memphis, personal communication). In the assay, ^{14}C labelled acetyl-CoA (Section 2.2.25) is condensed with malonyl-ACP to form radiolabelled β-ketobutyryl-ACP. Radioactivity that was incorporated into the ketoacyl-ACP is precipitated on to filter discs and any unreacted ^{14}C acetyl-CoA is separated by washing the discs in successive TCA solutions. The radioactivity is then measured by scintillation counting (Section 2.2.25). The results of this assay showed that the MBP-pfFabH fusion protein contained a significant amount of β-ketoacyl-ACP synthase activity (Figure 5.13).

Ideally, a comparison of the activity of the MBP-pfFabH fusion protein and free pfFabH - separated by Factor Xa cleavage and removal of the MBP by passing through an amylose resin – should be performed to ascertain if the N-terminal MBP affects β-ketoacyl-ACP synthase activity. It was decided that within the time constraints of the project and due to the relatively small amounts of purified soluble MBP-pfFabH produced, that this comparison would not be performed. However, due to the previous studies of pfFabH (Waters *et al.*, 2002 & Prigge *et al.*, 2003) a detailed characterisation of the enzyme was not necessary.

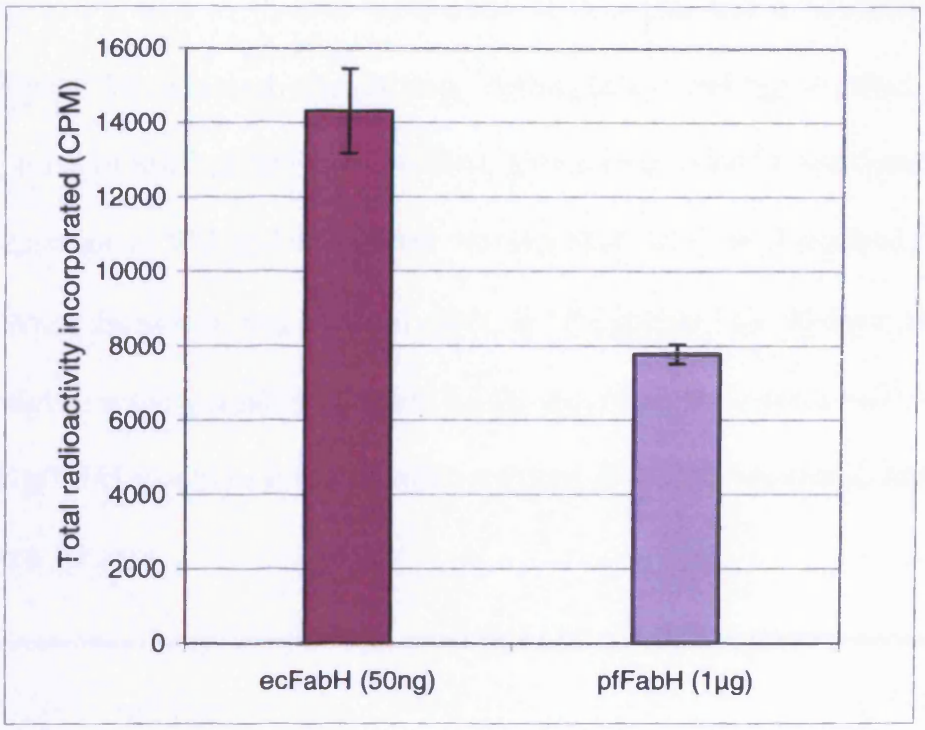


Figure 5.13 The activity of *P. falciparum* FabH protein

The incorporation of radiolabelled acetyl-CoA with ecFabH(■) and MBP-pfFabH(■). Each assay was performed in duplicate and errors are shown as errors from the mean. Assays contained 50ng of ecFabH or 1µg of MBP-pfFabH and the activities were measured using a radiochemical assay as described in Section 2.2.25.

Having established that the fusion protein had significant KAS activity, it was important to store the enzyme in such a way as to retain as much activity for as long as possible. An aliquot of purified MBP-pfFabH that had been stored at 4°C for 30 days was compared to MBP-pfFabH stored as solution in 50% glycerol at -20°C for

30 days and a sample of freshly prepared MBP-pfFabH (Figure 5.14). When stored at -20°C in 50% glycerol for 30 days, greater than 90% activity was retained. However, storage at 4°C with no glycerol led to a loss of more than half the condensing enzyme activity (30.7% remained) after 30 days. Putting this knowledge to effect, each time a fresh batch of MBP-pfFabH was purified, glycerol was added immediately to a final concentration of 50% and the enzyme was stored at -20°C until required for further use. When the protein was stored at -20°C in 50% glycerol for 90 days, only 55% of the original activity remained (Figure 5.14), suggesting that a fresh batch of MBP-pfFabH should be prepared when required, if the previous batch had been stored for over 30 days.

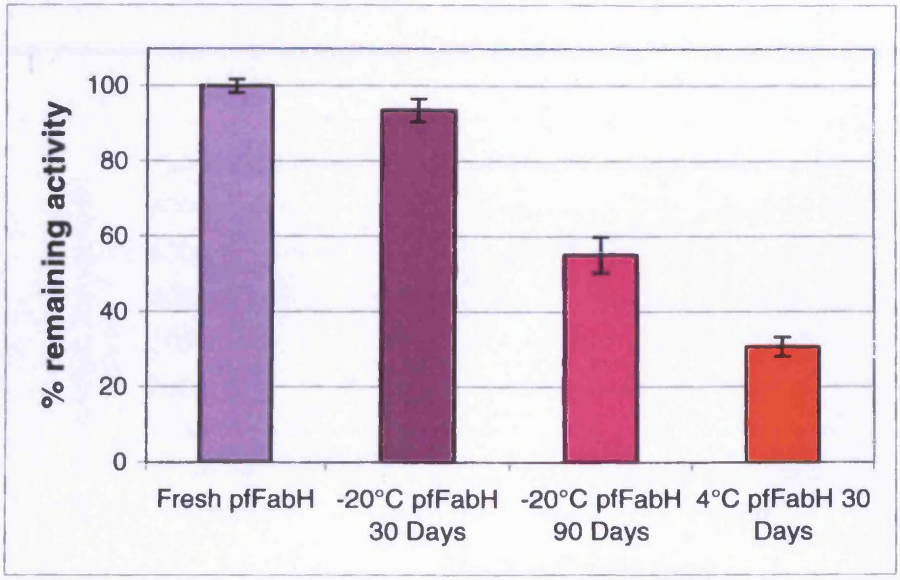


Figure 5.14 The effects of storage on MBP-pfFabH activity

The percentage remaining activity of freshly prepared MBP-pfFabH (■), MBP-pfFabH stored at -20°C for 30 days in 50% glycerol (■), stored for 90 days in 50% glycerol (■) or stored at 4°C for 30 days (■). Assays were performed in duplicate and errors are as the range of results from the mean. Each assay contained 1µg of the different MBP-pfFabH preparations and the activities were measured using a radiochemical assay as described in Section 2.2.25.

5.12 Refining the conditions of the MBP-pfFabH Assay

Linearity due to protein concentration was assessed at the normal substrate concentrations (Section 2.2.25). The activity of the MBP-pfFabH protein was tested at concentrations ranging from 1µg to 25ng whilst the concentrations of [1-^{14}C] acetyl-CoA, malonyl-CoA, ecACP and ecFabD remained constant.

To produce a gradient of protein concentrations (Figure 5.15), a 500ng/µl stock of MBP-pfFabH in 50% glycerol was diluted serially with appropriate amounts of 100mM Na-phosphate buffer, pH6.6 containing 50% glycerol. The incorporation of radiolabel in products bound to ACP was plotted against the protein concentrations of MBP-pfFabH (Figure 5.15).

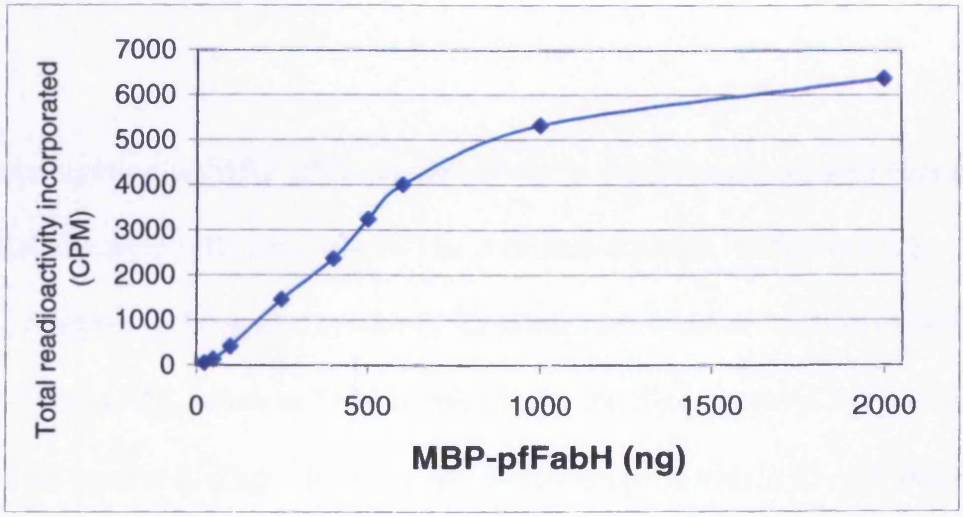


Figure 5.15 The amount of radiolabel observed in the acyl-ACP increases linearly with respect to the amount of MBP-pfFabH

Each assay was performed in duplicate and errors are shown as errors from the mean. The activity of MBP-pfFabH was measured using the radiochemical assay as described in Section 2.2.25.

The graph shows that the activity of the MBP-pfFabH enzyme increases linearity up to 600ng protein. Because the graph indicated that 1μg of MBP-pfFabH was outside the linear range, this could mean that the original calculation of β-ketoacyl-ACP synthase activity of MBP-pfFabH was underestimated.

The enzyme activity was recalculated using lower concentrations of MBP-pfFabH and these calculations produced a range of activities from 74-93pmoles min^{-1} μg^{-1}, suggesting that the original estimation (89.7pmoles min^{-1} μg^{-1}protein) was accurate. All subsequent assays were performed using 500ng of MBP-pfFabH protein because, at this concentration it was clear that the reaction was directly proportional to the amount of protein and not limited by the concentrations of substrates. This amount of MBP-pfFabH also incorporated enough radiolabel to produce a signal to background ratio adequate for accurate inhibition studies.

5.13 An investigation of MBP-pfFabH sensitivity to thiolactomycin and cerulenin

The bacterial metabolites thiolactomycin and cerulenin are both inhibitors of the condensing reactions in fatty acid synthesis. Cerulenin inhibits the reaction in both Type I and Type II FAS, whereas TLM is specific for the β-ketoacyl-ACP synthases of Type II FAS present in plants, bacteria and protozoan parasites. In *E. coli*, FabH is minimally sensitive to both TLM and cerulenin when compared to the inhibition of FabB and FabF enzymes (Price *et al.*, 2001). A very similar result was observed with TLM in pea plants (Jones *et al.*, 2000) (no data is available for cerulenin against the pea KAS enzymes), suggesting that the closely related pfFabH (Figure 5.1) might also be minimally sensitive to both these inhibitors. Prigge and colleagues have already

reported the inhibition of pfFabH by TLM (R enantiomer – the purified natural product) with an IC_{50} >330μM (Prigge *et al.*, 2003). However, his study used different conditions for the enzyme assay. Furthermore, a different source of TLM was also used in this assay and since it is known that different preparations of TLM can have quite different activities, it was decided that an evaluation of MBP-pfFabH sensitivity to TLM (and cerulenin) under the assay conditions of this study was important. These data would be central for comparison of the inhibitory activity of any TLM derivatives tested against the MBP-pfFabH enzyme.

Unfortunately, prior to testing the inhibitory effect of TLM on MBP-pfFabH it was not possible to gain access to the naturally produced TLM. An alternative source of TLM from (Sigma Chemicals, Poole, Dorset, UK) was obtained, as a synthetically produced, racemic mix of TLM. A comparison of the inhibitory effect of Sigma TLM and naturally produced TLM on the growth of the *E. coli* strain ANS1 was performed as described in Section 2.2.29. The ANS1 strain was used because it is deficient in all TolC-dependent type I secretion and multidrug efflux systems (Jackowski *et al.*, 2002). Upregulation of the *emrAB* efflux pump, a member of the above protein family caused TLM resistance in *E. coli* (Furukawa *et al.*, 1993; Lomovskaya *et al.*, 1995). Therefore, it was important to test the inhibitory activity of TLM against a strain, in which, the efflux pump activity had been abrogated. The *E. coli* ANS1 strain was kindly provided by Dr. Charles Rock (St. Jude Children's Hospital, Memphis, USA).

To achieve the same extent of inhibition as naturally produced thiolactomycin, the synthetic compound needed to be present at twice the concentration (Figure 5.16). This result suggested that the synthetic compound was truly a 50:50 racemic mixture

and the (5S)-enantiomer did not display an antagonistic effect. Therefore, TLM obtained from Sigma will be referred to as racemic TLM throughout the rest of this report.

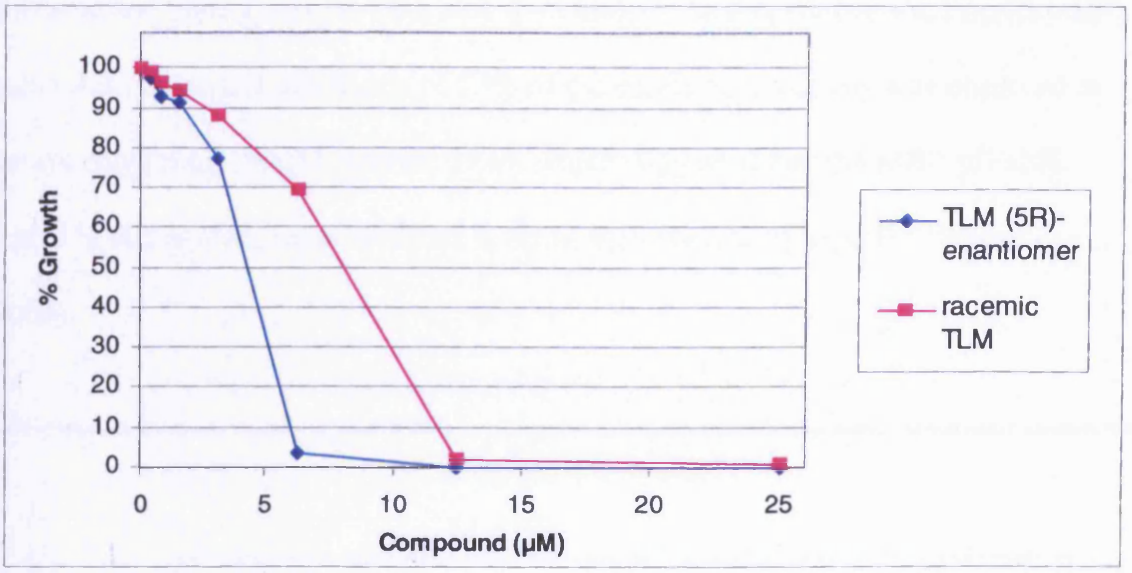


Figure 5.16 Inhibition of the growth of *E. coli* ANS1 strain

The racemic mixture and the (5R)-enantiomer of TLM were tested at 25μM, 12.5μM, 6.25μM, 3.13μM, 1.56μM, 0.78μM and 0.39μM concentrations formed by 1 in 2 serial dilutions from the highest concentration. The assay of bacterial growth was performed as described in Section 2.2.9. The *E. coli* ANS1 strain was kindly provided by Charles Rock (St. Jude Children's Research Hospital, Memphis, USA).

An assay that contained 100μM concentration of a synthetically produced, racemic TLM (Sigma Chemicals, Poole, Dorset, UK) was prepared. In a separate reaction 100μM of cerulenin was added to see if the MBP-pfFabH enzyme was highly or minimally sensitive to cerulenin (Sigma Chemicals, Poole, Dorset, UK) and would

indicate any differences between TLM and cerulenin sensitivity. The TLM obtained from Sigma was used because at the time of these assays there was no access to the natural product, (5R)-enantiomer only. The results indicated that the preparation of MBP-pfFabH was insensitive to cerulenin and minimally sensitive to TLM (Figure 5.17). The lack of inhibition by 100μM cerulenin is comparable to that of bacterial and plant FabH proteins and, therefore, no further studies were performed using this compound. Inhibition of 23% of the condensing activity was observed in assays containing 100μM racemic TLM, which suggested that the MBP-pfFabH tested in this study is more sensitive to TLM than previously reported (Prigge *et al.*, 2003).

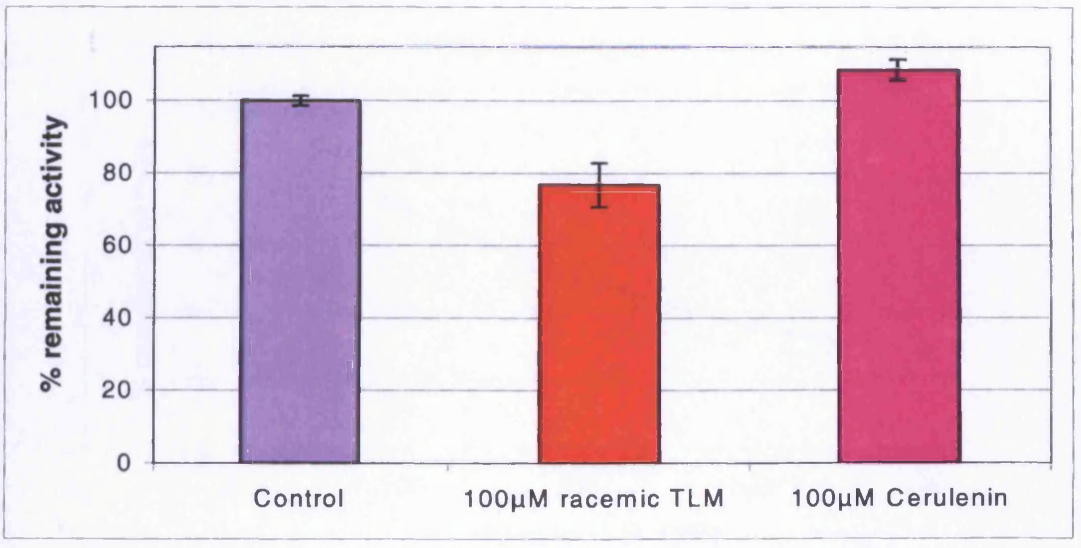


Figure 5.17 The inhibition of MBP-pfFabH with thiolactomycin and cerulenin
The incorporation of radiolabelled acetyl-CoA with no inhibitor (■), in the presence of 100μM Sigma TLM (■) or 100μM cerulenin (■). Each assay was performed in duplicate and errors are shown as errors from the mean. Assays contained 500ng of MBP-pfFabH and the activity was measured using the radiochemical assay as described in Section 2.2.25.

5.14 The determination of the IC_{50} of TLM against pfFabH

In order to get a more accurate estimation of the sensitivity of MBP-pfFabH to TLM, the IC_{50} of this compound against the enzyme was determined. Briefly, a 10mM stock of TLM was serially diluted in 40% DMSO to produce 8mM, 6mM, 4mM, 2mM, 1mM and 500µM stock solutions. The final DMSO concentrations were kept constant and up to 2% had no effect on the enzyme activity. When 2µl of each inhibitor solution was added to the total assay volume of 40µl, the final TLM concentrations lay between 25µM-400µM. Assays containing no TLM were used to estimate the uninhibited activity. The results (Figure 5.18) suggested that the concentration of racemic TLM that caused inhibition of 50% of pfFabF activity (IC_{50}) is approximately 200µM (the equivalent to 100µM TLM (R enantiomer only).

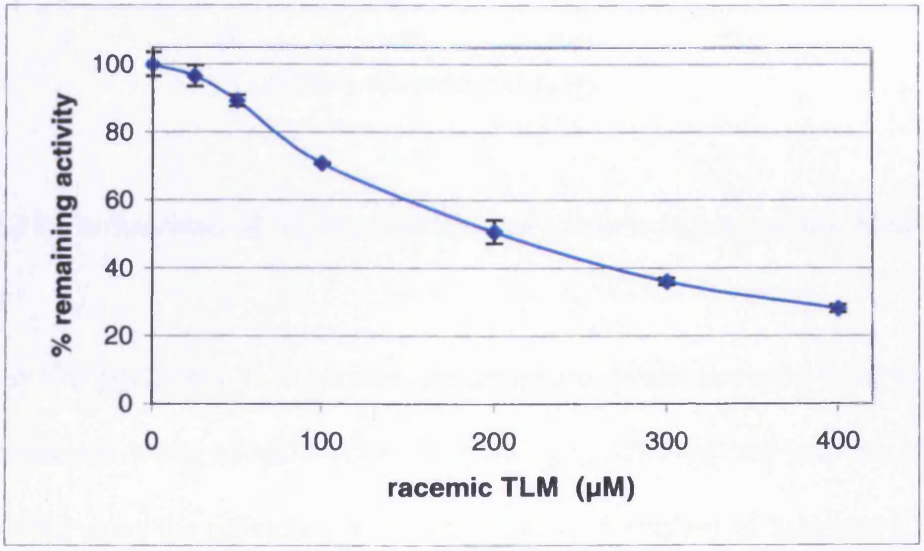


Figure 5.18 IC_{50} determination for TLM against MBP-pfFabH

Each assay was performed in duplicate and errors are shown as errors from the mean. Assays contained 500ng of MBP-pfFabH and the activity was measured using the radiochemical assay as described in Section 2.2.25.

These data indicated that the *P. falciparum* FabH (KASIII) enzyme was much less sensitive to thiolactomycin compared to the pfFabF enzyme (Figure 5.19), even after noting that the standard amounts of enzymes used in the two assays were slightly different.

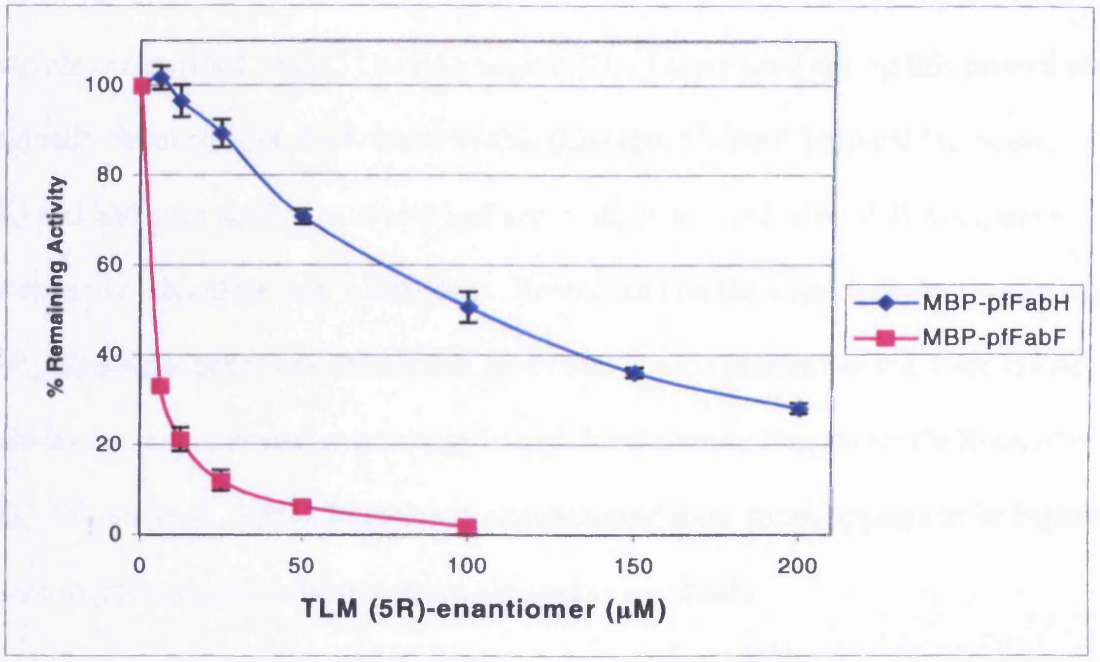


Figure 5.19 Comparison of TLM inhibition of the two *P. falciparum* KAS enzymes

Each assay was performed in duplicate and errors are shown as errors from the mean. Assays contained 500ng of MBP-pfFabH or 300ng of MBP-pfFabF and the activity was measured using the radiochemical techniques as described in Sections 2.2.24-2.2.25. The data points of the inhibition of MBP-pfFabH were converted from racemic TLM concentrations to naturally produced concentrations by dividing by two. These data are shown in separate graphs in Figure 4.6 and Figure 5.18.

5.15 Discussion

Amplification, and subsequent cloning, of the *pfFabH* gene was unsuccessful when a cDNA library was used as template. Amplification of the gene was only achieved when a copy of a plasmid containing the gene, kindly provided from Dr. Sean Prigge (Johns Hopkins, School of Public Health, Baltimore, USA) was used as the DNA template. A possible reason why cloning of *pfFabH* was unsuccessful using a cDNA template is described below. The ring stage cDNA library used during this project was originally obtained from Prof. Steve Walsh (Liverpool School Tropical Medicine, UK) and had been used extensively and successfully to clone several *P. falciparum* genes in our laboratory over many years. Recent data on the stage specific expression of *P. falciparum* genes has shown that genes that encode proteins of the Type II FAS pathway are not expressed at particularly high levels during ring stages (le Roch *et al.*, 2003; Waters *et al.*, 2002). In contrast, expression of these genes appears to be highest in late trophozoite and schizont stages (Waters *et al.*, 2002).

Assuming that the corresponding level of these proteins correlates with the level of mRNA this may explain why trophozoite and schizont stages are less sensitive to treatment with TLM, cerulenin and triclosan compared to ring stages (Waller *et al.*, 2003). If the level of transcription of these genes is low in ring stages it might be possible that the *pffabH* gene was not represented in the cDNA library. The high A+T bias of the *pffabH* gene (74.2%) may also have decreased the chances of obtaining a full length clone or any large fragments of the gene.

Following the successful cloning of *pffabH*, the enzyme was expressed and purified in the pMAL vector system, which had produced reasonable amounts of MBP-pfFabF

(Chapter 3, Section 3.11-13) and pfFabH (Waters *et al.*, 2002). This system produced approximately 1mg of soluble active MBP-pfFabH protein per litre of culture. Unfortunately, whilst cleavage of MBP from the N-terminus of pfFabH was achieved by treatment with Factor Xa, the pfFabH protein also appeared to be digested by this protease. Subsequently, the removal of MBP from pfFabH was not attempted with the consequence that the chances of producing a crystal structure of pfFabH would be reduced significantly. However, Prigge and colleagues, who separated MBP from pfFabH via an engineered TEV protease site, are endeavouring to obtain a crystal structure of pfFabH, (Prigge *et al.*, 2003: Sean Prigge, personal communication).

The MBP-pfFabH fusion protein was able to catalyse the initial condensing reaction of the Type II FAS cycle (Section 5.11). Compared to the radioactivity incorporated using a lower amount of ecFabH enzyme, the MBP-pfFabH enzyme may appear to have a poor β-ketoacyl-ACP synthase activity. However, the previously reported specific activity for pfFabH under similar conditions was 250pmoles min^{-1} μg^{-1} protein (Prigge *et al.*, 2003) and the calculated specific activity of the MBP-pfFabH protein in this study is 90pmoles min^{-1} μg^{-1}protein. These activities are of the same order and the discrepancy could be due to the slight differences in the assay conditions or to a lower activity of the MBP-pfFabH fusion protein purified in this project. There are several reports of the activity of the *E. coli* FabH enzyme, which suggest the specific activity of the bacterial enzyme is 5-7nmoles min^{-1} μg^{-1}protein (Khandekar *et al.*, 2000; Clough *et al.*, 1992). Compared to bacteria FabH enzymes, the *P. falciparum* FabH enzyme appears to have a significantly lower specific activity, although a comparison of K_m and K_{cat} values would give a more accurate insight into the catalytic abilities of these enzymes.

This Chapter describes the first study of the inhibitory effect of cerulenin on pfFabH. In a similar situation to plant and bacterial FabH enzymes, the *P. falciparum* enzyme appears to be insensitive to cerulenin treatment. The IC_{50} of 200 μM for TLM against MBP-pfFabH was also determined and showed that, in our experiments, the enzyme was more sensitive to this antibiotic than in the only previous report (>300 μM) (Prigge *et al.*, 2003). An IC_{50} of 200μM may indicate that there is a large window of opportunity in which to improve the binding of TLM to the enzyme, via alterations in to the structure of thiolactomycin. Recently, Prigge and colleagues have identified TLM-like compound that inhibit the pfFabH enzyme and also exhibit an increase inhibitory effect on parasites in culture (Prigge *et al.*, 2003). The inhibitory activities of TLM derivatives against the MBP-pfFabH protein are described in Chapter 6.

Chapter 6

Chapter 6 – Inhibition of *P. falciparum* Fatty Acid Synthesis by Thiolactomycin and its derivatives

6.1 Introduction

Prior to the start of this project, TLM had been shown to inhibit the growth of *P. falciparum* cultures, showing an LD_{50} = 50μM against the multi-drug resistant strain W2 mef (Waller *et al.*, 1998). Furthermore, *in vivo* TLM exhibited rapid absorption and wide distribution in tissues and protected against bacterial intraperitoneal and urinary tract infections in mice (Miyakawa *et al.*, 1982). Unlike cerulenin, which is a non-selective inhibitor of Type I and Type II FAS, thiolactomycin is a selective inhibitor of Type II FAS (Hayashi *et al.*, 1983 and 1984) and has negligible toxicity in mammals (Oishi *et al.*, 1982). These studies and the data presented in Chapters 4 and 5, suggested that TLM was a promising lead compound for the treatment of *P. falciparum* malaria. Against cultures of protozoan parasites, TLM exhibited LD_{50} values in the mid to high micromolar range (Table 6.1).

Organism	LD_{50} (μM)
P. falciparum	49[a]
T. brucei	~150[b]
T. gondii	100[c]

Table 6.1 The inhibition of protozoan parasites with TLM

[a] Waller *et al.*, 2003 (against the W2 mef strain of the parasite). [b] Morita *et al.*, 2000. [c] McFadden and Roos, 1999. All values are given as the dose that inhibited 50% growth of the parasite in *in vitro* studies as measured by incorporation of ^{3}H-hypoxanthine.

An increase in the inhibitory activity of TLM would be necessary before it could be considered for the treatment of parasitic infections. Analogues of TLM were first shown to have enhanced activity against Type II FAS in pea chloroplasts (Jones *et al.*, 1994 and 1995) and these analogues also showed an increase in activity against the FabB, FabF and FabH enzymes of pea plants (Jones *et al.*, 2000). These analogues contained substitutions of the isoprenoid moiety at position 5 of the thiolactone ring with longer hydrocarbon chains.

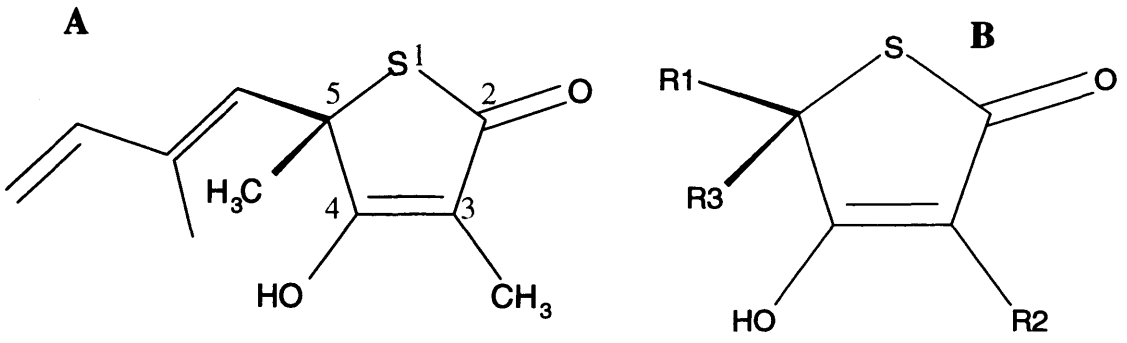


Figure 6.1 Structure of TLM and variable groups

A. The structure of the fungal metabolite TLM. Members of the thiolactone ring are numbered in blue text. B. The variable groups at the C5 and C3 positions as they will be referred to during the remainder of this report.

During the course of this project, two groups reported that similar substitutions at the 5 position showed increased activity against *Mycobacterium tuberculosis* (Kremer *et al.*, 2000; Douglas *et al.*, 2002; Senior *et al.*, 2003, 2004) and *P. falciparum* (Waller *et al.*, 2003) *in vitro*. The 3D-structure of the *E. coli* FabB-TLM complex showed the presence of a hydrophobic crevice, which is distinct from the hydrophobic groove that

usually accommodates the growing acyl chain (Figure 6.2) (Price *et al.*, 2001). In ecFabB this crevice is deeper than the isoprenoid group found on TLM and, therefore, may not be optimally filled by TLM. Assuming that the structures of the condensing enzymes are similar, this may explain the reason for the increased activity of TLM derivatives with longer, hydrophobic chains at the C5 position against malaria and trypanosome parasites.

Three compounds with similar structures to TLM have already been reported to have an increased activity against cultures of *P. falciparum* and significantly inhibit the pfFabH (KASIII) enzyme activity (Prigge *et al.*, 2003). Encouragingly, these results showed that derivatives of TLM can demonstrate a marked increase in the inhibitory activity against an individual condensing enzyme of *P. falciparum*, which in turn may increase the sensitivity of the parasite.

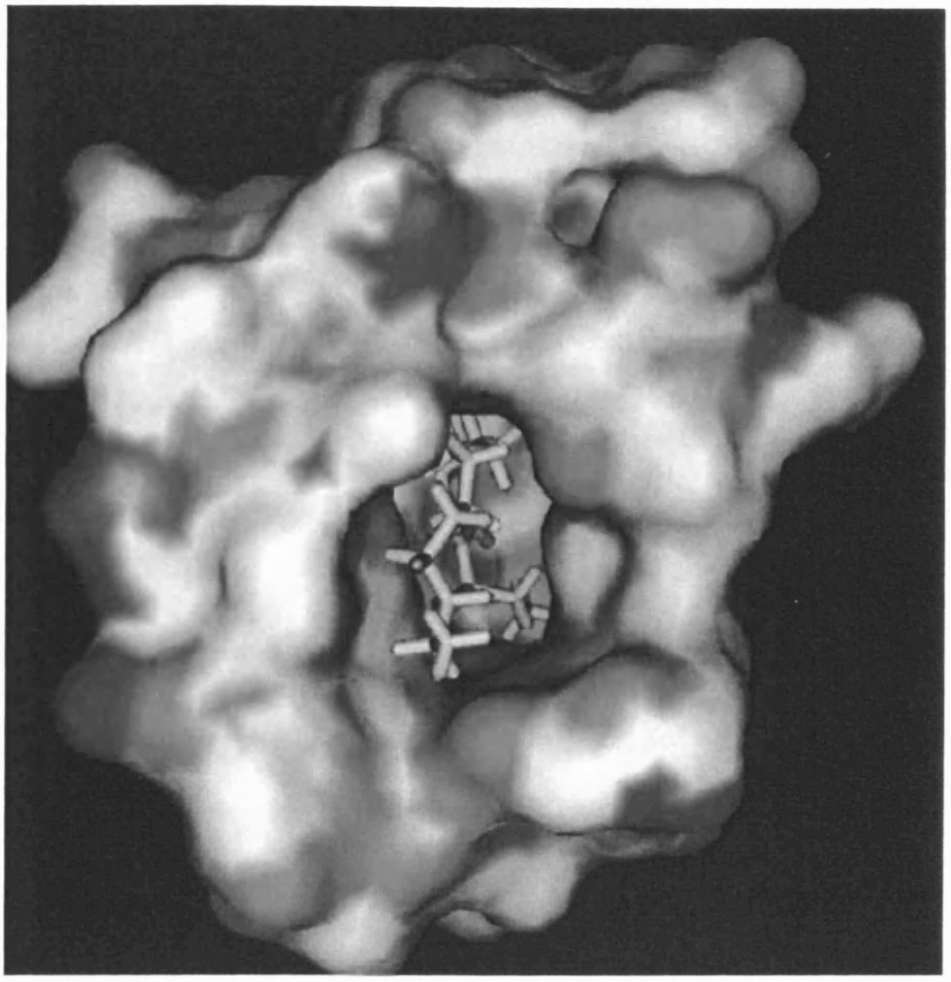


Figure 6.2 The hydrophobic crevice that incorporates the C5 chain of TLM and its derivatives

This diagram depicts compound 925, not TLM, docked in the active site of *E. coli* FabB (Jones *et al.*, 2004). The 1FJ4.pdb file (Price *et al.*, 2001) was downloaded from the Brookhaven database and was manipulated using FlexX Software and SwissPDB viewer (Guex and Peitsch, 1997). C5 decyl group is positioned in the same hydrophobic pocket that extends to the surface of the protein. This hydrophobic pocket is distinct from the one that incorporates the growing acyl chain (Price *et al.*, 2001).

6.2 Synthesis of TLM derivatives and determination of LD_{50}

At the outset of this work, collaborations with Dr. Simon Jones and Dr. Ian Gilbert (Welsh School of Pharmacy, Cardiff) and Prof. Reto Brun (Swiss Tropical Institute, Basel, Switzerland) were arranged with the aim to synthesise and assay analogues of TLM for anti-parasitic activity. All of the TLM analogues described subsequently were synthesised by Dr. Simon Jones, using methods or modifications of the methods described in Jones *et al.*, 2004. The synthesis of these compounds should produce equal ratios of the stereoisomers at position C5 and all of the TLM derivatives reported from this study are given as total concentrations of preparations, which would contain mixtures of the R and S enantiomers (Wang and Salvino, 1984). Prof. Reto Brun performed assays of the effect of TLM analogues on the growth of *in vitro* cultures of *P. falciparum,* as described previously (Jones *et al.*, 2004). These compounds were also assayed against cultures of *Trypanosoma brucei* and *Trypanosoma cruzi* (the causative agents of sleeping sickness and Chagas' disease, respectively). However, because this project focused on *P. falciparum*, the latter results are presented in Appendix 8.

6.3 TLM and derivatives of the C5 (R1) position

The first group of compounds that were synthesised were based on the series of TLM derivatives that exhibited increased activity against pea plastids (Jones *et al.*, 2000). At the outset of this work, these were the only TLM analogues that had been reported with any activity against a Type II FAS system, although later reports showed that these compounds also inhibited the growth of *P. falciparum* (Waller *et al.*, 2003) and *M. tuberculosis* (Kremer *et al.*, 2000). Thiolactomycin (compound 781) was

synthesised as described by Jones *et al.*, 2004). Compounds were also produced in which the isoprenoid moiety at position 5 of the thiolactone ring was substituted with hydrogen, hexyl, octyl, geranyl, decyl and hexadecyl groups. These inhibitory effect of these compounds were tested in *in situ* cultures of *P. falciparum* K1 strain (Jones *et al.*, 2004) and the results are shown in Table 6.2.

Several of these analogues showed greater activity than that of compound 781 (TLM) against the parasite; the best of this series, compound 777, displayed a 6-fold improvement. The findings from these studies suggested that replacement of the isoprenoid group with longer hydrocarbon chains led to enhanced activity against the parasite. Comparison of compounds 778 and 780 showed that a geranyl substitution was more effective against the parasite than a saturated, unbranched octyl group, suggesting that an increase in the number of double bonds and branched chains produced greater inhibition. This trend was subsequently observed in compounds tested against the *P. falciparum* W2mef strain (Waller *et al.*, 2003), although with different LD_{50} values compared to our studies using strain K1.

The increase in the inhibitory activity of substituted TLM compounds in cultures of *P. falciparum* does not prove that these derivatives target fatty acid synthesis or are better inhibitors of the target enzymes (possibly MBP-pfFabF or MBP-pfFabH). Thus, the increase in activity may be caused by more efficient diffusion of the more hydrophobic compounds through lipid bilayers (ref?). If the Type II FAS system in *P. falciparum* is located in the apicoplast, these compounds would have to pass through at least six membranes: - the erythrocyte membrane, the parasitophorous vacuolar membrane and the four membranes surrounding the apicoplast.

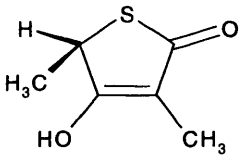
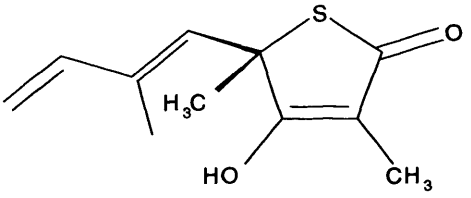
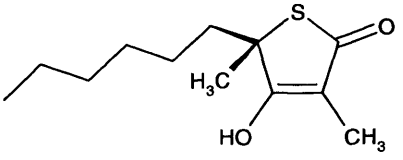
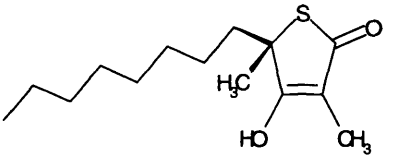
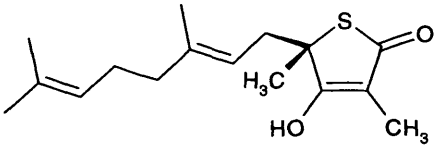
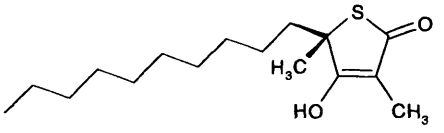
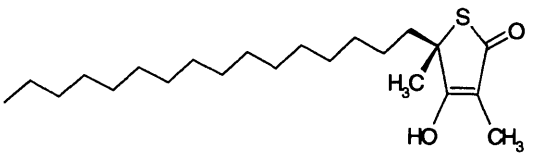
Compound No.	Description	Structure	LD_{50} (μM)
817	C5 Hydrogen		>347
781	C5 Isoprenoid (TLM)		143
779	C5 Hexyl		139
780	C5 Octyl		153
778	C5 Geranyl		40
776	C5 Decyl		36
777	C5 Hexadecyl		25

Table 6.2 Hydrocarbon substitutions at position 5 of TLM exhibit enhanced activity against *P. falciparum* cultures

The inhibition of the chloroquine and pyrimethamine resistant K1 strain of *P. falciparum* was measured using a ^{3}H-hypoxanthine assay (Jones *et al.*, 2004). TLM and its derivatives were dissolved in 2% DMSO. The LD_{50} values are calculated using the mean of two independent assays performed in duplicate.

Therefore, compounds with a significantly better ability to diffuse or be actively transported through these membranes would be expected to show a greater efficacy against *P. falciparum*.

6.4 Substitutions of the methyl group at the C3 (R2) position

6.4.1 Ethyl substitutions at the C3 (R2) position

The substitutions described at position 5 showed a modest improvement in anti-malarial activity. However, substitutions at other positions in the ring have not been fully investigated against any Type II FAS system. The detailed study of the crystal structure of *E. coli* FabB complexed with TLM (Price *et al.*, 2001), found that the two methyl substituents at the C3 (R2) and C5 (R3) positions were accommodated into two small hydrophobic pockets comprising Phe229 & Phe392 and Pro272 & Phe390 respectively (Figure 6.3). It has been proposed that the C3 methyl group occupies the position of the carboxylate carbon of the natural substrate malonyl-ACP (Price *et al.*, 2001). The two oxygen atoms of the malonyl-ACP carboxyl group protrude further into this environment (Price *et al.*, 2001), suggesting that the hydrophobic pocket surrounding the C3 methyl group was not optimally filled. Based on this information, a series of analogues (818-821 and 823), possessing an ethyl substituent at C5, were synthesised and tested against cultures of *P. falciparum* (Table 6.3).

With the exception of compound 818, all of the compounds demonstrated greater inhibition of *P. falciparum* cultures, compared to TLM. In parallel with the C3 methyl compounds (Table 6.2), C3 ethyl analogues that contained longer chains at C5 tended

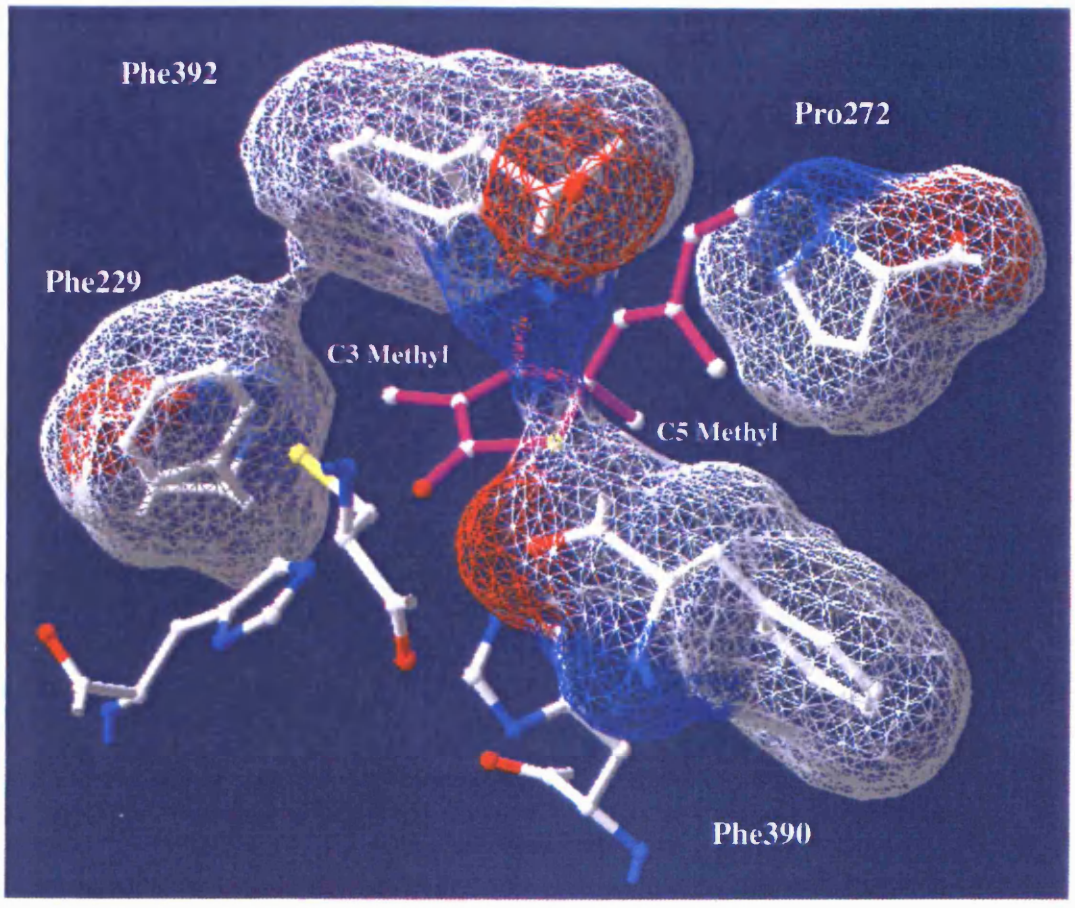


Figure 6.3 The methyl groups of TLM are positioned in small hydrophobic pockets

TLM is shown bound to the active site of *E. coli* FabB enzyme, with its two natural methyl groups nestled in two small, hydrophobic pockets formed by Phe390 % Pro272, and Phe392 & Phe229. The 1FJ4.pdb file (Price *et al.*, 2001) was downloaded from the Brookhaven database and was manipulated using SwissPDB viewer (Guex and Peitsch, 1997).

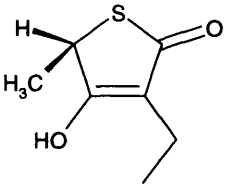
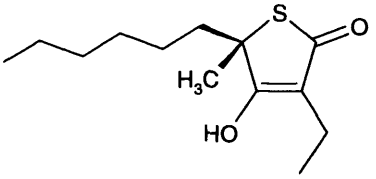
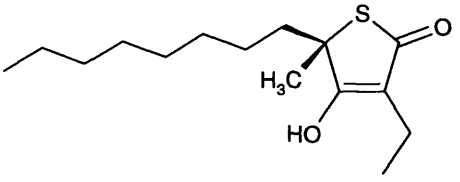
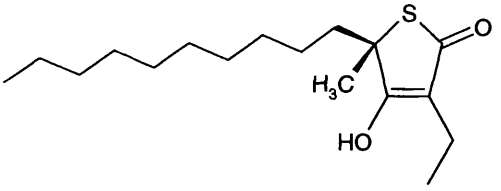
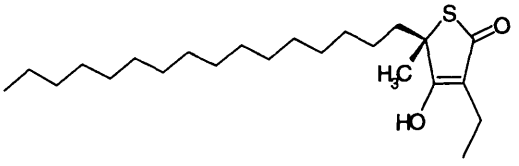
Compound No.	Description	Structure	LD_{50} (μM)
818	C5 Hydrogen C3 Ethyl		**233**
820	C5 Hexyl C3 Ethyl		**61**
823	C5 Octyl C3 Ethyl		**54**
821	C5 Decyl C3 Ethyl		**15**
819	C5 Hexadecyl C3 Ethyl		**19**

Table 6.3 Effect of ethyl substitutions at C3 of TLM on activity against *P. falciparum* cultures.

See the Legend of Table 6.2 for further details.

to show greater inhibition. The best analogues, 821 and 819 showed 10-fold and 8-fold increases in inhibition over TLM respectively. Analogues with C3 ethyl groups showed enhanced activity against the parasite when compared to those with C3

methyl groups (compare those in Table 6.3 to Table 6.2). This trend is highlighted when comparing the analogues with C5 octyl groups, 823 (54μM) shows a three-fold increase in activity compared to 780 (153μM). However, in analogues with longer C5 chains, the substitution of the C3 ethyl group caused a much smaller improvement in activity (compare 821 (15μM) to 776 (36μM) and 819 (19μM) to 777 (25μM). It is clear that the ethyl substituent is having a positive influence on the overall inhibitory activity against the parasite.

Compounds 817 and 818 (Table 6.2 and 6.3) possess a C5 (R1) hydrogen group exhibit LD_{50} values in the high micromolar range. Three other compounds (822, 924 and 926) that possess a hydrogen group at this position were synthesised an assayed against the malaria parasites (Table 6.4). In comparison to analogous compounds that contain C5 decyl groups (Table 6.4) these compounds show a significant decrease in the inhibition of growth in cultures of *P. falciparum*. All of the five compounds containing a C5 (R1) hydrogen group (Compounds 817, 818, 822, 924 and 926) showed lower inhibition of parasite cultures compared to TLM. Members of this series also displayed the lowest LD_{50} values observed in this project. These data provide further evidence that hydrophobic interactions, between the C5 group of TLM derivatives and the target protein in *P. falciparum*, make a significant contribution to the inhibitory activity of the compounds.

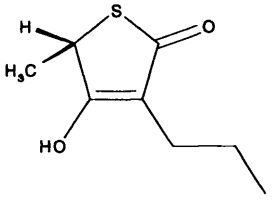
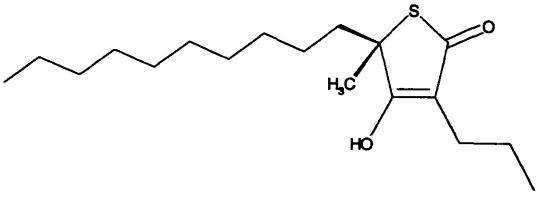
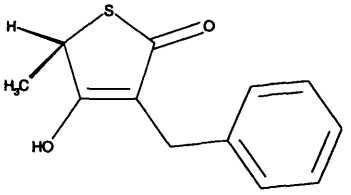
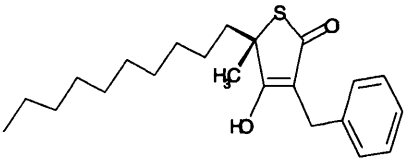
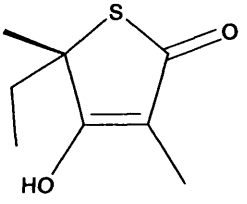
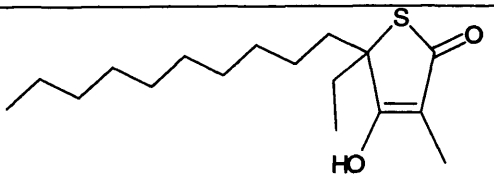
Compound No.	Description	Structure	LD_{50} (μM)
924	C5 (R1) Hydrogen C3 Propyl		**>290**
925	C5 (R1) Decyl C3 Propyl		**10**
926	C5 (R1) Hydrogen C3 Benzyl		**195**
927	C5 (R1)Decyl C3 Benzyl		**50**
822	C5 (R1) Hydrogen C5 (R2) Ethyl		**>316**
825	C5 (R1) Hydrogen C5 (R2) Ethyl		**65**

Table 6.4 Hydrogen substitutions at position 5 of TLM exhibit decreased activity against *P. falciparum* cultures

See the Legend of Table 6.2 for further details.

6.4.2 Longer hydrocarbon chains at the C3 (R2) position

The above data prompted the synthesis of a group of compounds with longer or larger hydrocarbon moieties at position C3. To achieve a direct comparison of these analogues the C5 chain length was kept constant. Since the most effective compound contained a C5 decyl chain (821:see above), while investigating the effect of altered C3 chain lengths, all of the analogues possessed a C5 decyl substituent. Substitutions at the C3 position were achieved with propyl, butyl and benzyl groups (Jones *et al.*, 2004) and these three new compounds were assayed against *P. falciparum* (Table 6.5). These data show that by extending the ethyl chain to a propyl chain an LD_{50} = 10μM was achieved for the first time in this project, giving a 14-fold increase in activity over TLM. Extending the chain further to butyl and benzyl groups resulted in a decrease in activity. This suggested that the propyl group of three carbons was the optimal chain length at the C3 position, beyond which activity was diminished.

To reinforce this assertion, compound 925 was docked into the active site of the *E. coli* FabB protein using FlexX software (Jones *et al.*, 2004) (Figure 6.4). The propyl group appeared to fit into the hydrophobic pocket with the C-terminus of the propyl chain in close proximity to the active site Cys163. This same active site cysteine is present in the MBP-pfFabF enzyme (Cys165- MBP-pfFabF numbering). However, one of the two highly-conserved residues (Pro272 and Phe390) that constitute the hydrophobic pocket where the C3 group is positioned, is replaced by leucine in pfFabF (Pro275 and Leu404 - pfFabF numbering). This suggests that there may be a larger hydrophobic pocket or that fewer hydrophobic interactions occur in this region. A model of pfFabF based on the *E. coli* FabB-TLM complex was constructed using Swiss PDB viewer and suggested that the active site architecture

was very similar and that a propyl chain at the C3 position of TLM could be accommodated into the active site of pfFabF (Results not shown).

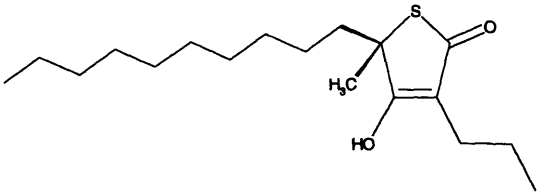
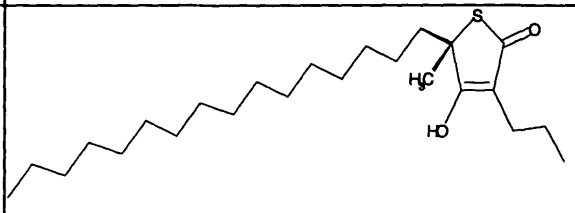
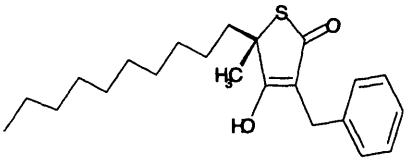
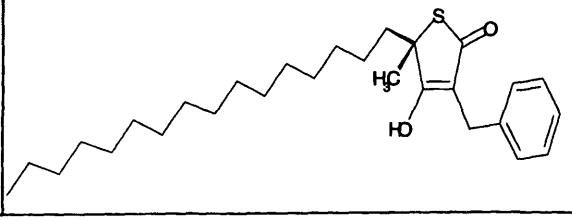
Compound No.	Description	Structure	LD_{50} (μM)
925	C5 Decyl C3 Propyl		**10**
1199	C5 Hexadecyl C3 Propyl		**6**
927	C5 Decyl C3 Benzyl		**50**
1200	C5 Hexadecyl C3 Benzyl		**7**

Table 6.5 A comparison of C5-decyl and C5-hexadecyl substitutions

See the Legend of Table 6.2 for further details.

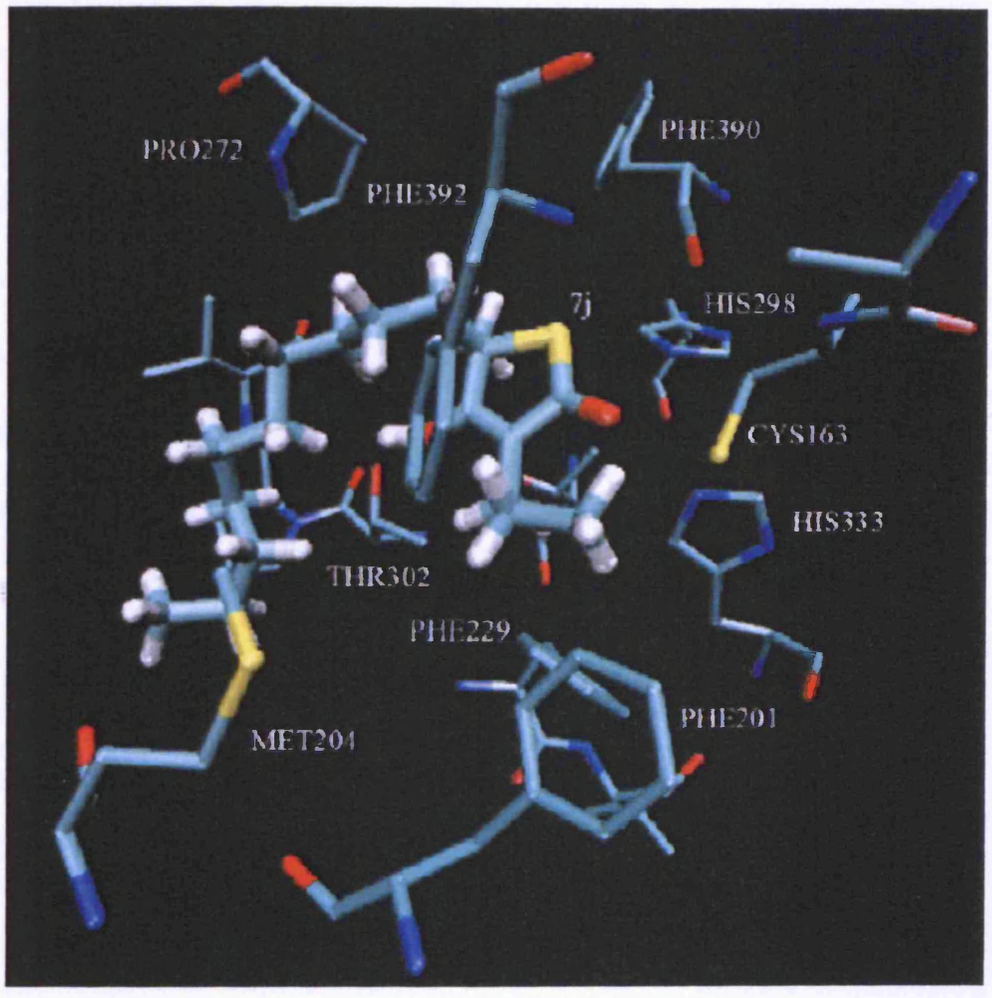


Figure 6.4 Compound 925 docked in the active site of *E. coli* FabB

The C3 propyl chain of this compound is positioned close to the active site Cys163 residue. Part of the C3 propyl group is obscured.

6.4.3 Hexadecyl substitutions

All of the compounds reported in the previous section possess C5 decyl groups. In the original batch of compounds (Section 6.3 & Table 6.2) a C5 hexadecyl group was shown to have enhanced activity compared to the analogue containing a C5 decyl group (compare 777 to 776). Therefore, because compound 925 contains a decyl chain and is the best inhibitor against *P. falciparum* cultures to date, it seemed logical to synthesise the derivative with a C5 hexadecyl chain combined with a C3 propyl. A derivative containing a C3 benzyl group was also synthesised and tested against *P. falciparum* cultures (Table 6.6). Both hexadecyl derivatives showed an increased in inhibition compared to the corresponding decyl derivatives. This provides further evidence that TLM derivatives with longer hydrocarbon side chains have an enhanced inhibitory effect on the growth of the malaria parasite.

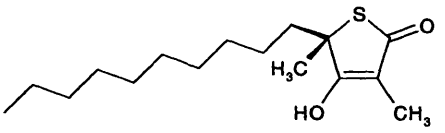
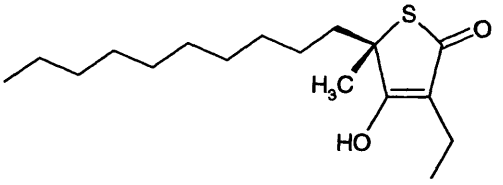
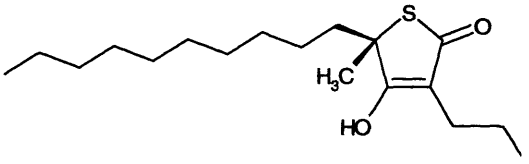
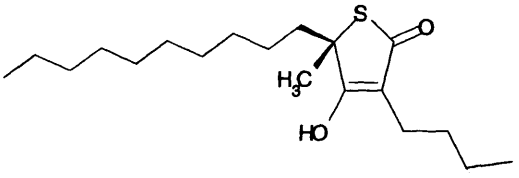
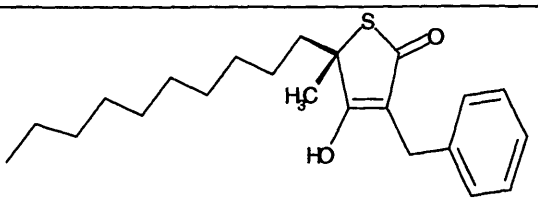
Compound No.	Description	Structure	LD_{50} (μM)
776	C5 Decyl C3 Methyl		**36**
821	C5 Decyl C3 Ethyl		**15**
925	C5 Decyl C3 Propyl		**10**
944	C5 Decyl C3 Butyl		**71**
927	C5 Decyl C3 Benzyl		**50**

Table 6.6 Substitutions at C3 of TLM exhibit variable activity against *P. falciparum* cultures

See the Legend of Table 6.2 for further details.

6.5 Substitution of the methyl group at the C5 (R3) position

The increased inhibitory activity shown by substitutions of the methyl group at position 3 suggested that an improvement might also be observed when substituting

the methyl group that is the second substituent (R3) at position 5. Two compounds with ethyl groups at C5 (R3) were synthesised (compounds 825 and 863) (Jones *et al.*, 2004). The removal of the C5 methyl group and its replacement with a hydrogen atom (compound 943) was also carried out. A comparison of the inhibition of *P. falciparum* by compounds containing hydrogen, methyl- or ethyl- groups at position R3 is shown in Table 6.7.

All of the analogues showed improved inhibition over TLM. However, direct comparisons of the analogues containing ethyl substituents and their methyl counterparts suggested that the presence of the ethyl group causes a decrease in activity (Table 6.7). This also appeared to be case in the analogue where the C5 methyl group was replaced by hydrogen (LD_{50} = 15μM for 821 and LD_{50} = 72μM for 943). Therefore, unlike the substitution of the methyl group at C3, modification of the C5 methyl group does not appear to offer any advantage for activity against *P. falciparum*. The two amino acids of *E. coli* FabB (Phe229 and Phe392) that constitute the hydrophobic pocket that the C5 methyl group rests in, are conserved in pfFabF (Phe232 and Phe404) and the areas surrounding these residues are also highly conserved.

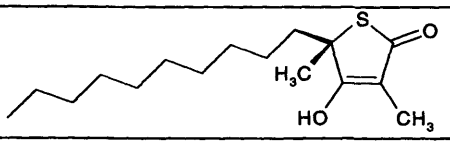
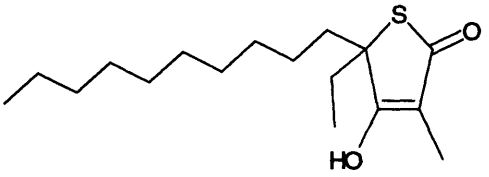
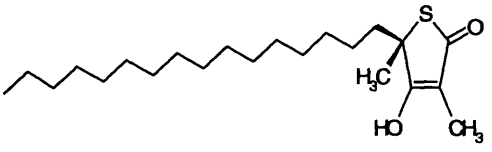
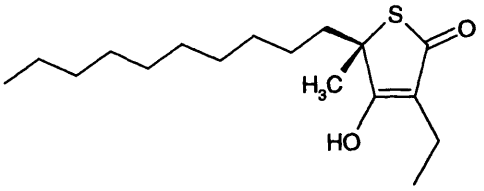
Compound No.	Description	Structure	LD_{50} (μM)
776	C5 R1 Decyl C5 R2 Methyl		**36**
825	C5 R1 Decyl C5 R2 Ethyl		**65**
777	C5 R1 Hexadecyl C5 R2 Methyl		**25**
863	C5 R1 Hexadecyl C5 R2 Ethyl		**35**
821	C5 R1 Decyl C3 Ethyl C5 R2 Methyl		**15**
943	C5 R1 Decyl C3 Ethyl C5 R2 H		**72**

Table 6.7 Substitution of the C5 (R3) methyl group of TLM causes a decrease in activity against *P. falciparum* cultures

See the Legend of Table 6.2 for further details.

6.6 Replacement of the sulphur heteroatom

6.6.1 Tetronic acids – substitution of the sulphur heteroatom with oxygen

The 3D structure of *E. coli* FabB complexed with TLM showed the thiolactone sulphur does not make any obvious specific interactions with the protein but sits

adjacent to the side chain of the active site Cys163 (Price *et al.*, 2001). However, the side chain of Cys163 moves 2.1 Angstroms to avoid clashing with the TLM sulphur. This suggested that replacing the thiolactone sulphur with a similar but smaller atom might cause less disruption of the protein. This could allow tighter binding of the TLM derivative and, thus, increased inhibition of the enzyme. Three analogues were synthesised where the thiolactone sulphur was replaced by an oxygen atom. This substitution would allow the shape and structure of the TLM analogues to remain essentially the same, whilst reducing the size of the atom at position 1 of the ring. These compounds (922, 923 and 945) were tested against *P. falciparum* cultures and were compared to the results observed with their sulphur containing counterparts (779, 821 and 945) (Table 6.8).

These compounds exhibited a reasonable activity against the parasite. Two of the three compounds showed a decrease in the activity against the parasite compared to the sulphur containing analogues; the third showed a small increase in activity. This suggested that the replacement of the sulphur thiolactone with an oxygen atom did not confer a significant advantage that would justify the replacement. Yields from the synthesis of oxygen-containing compounds were lower than those observed with the sulphur-containing compounds and it was decided not to continue to synthesise oxygen ring compounds.

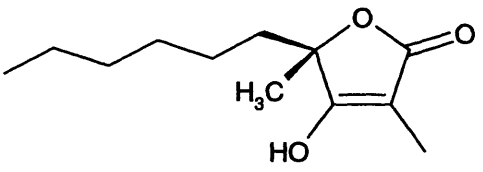
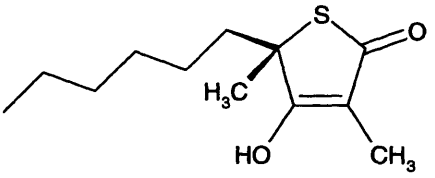
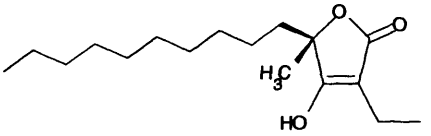
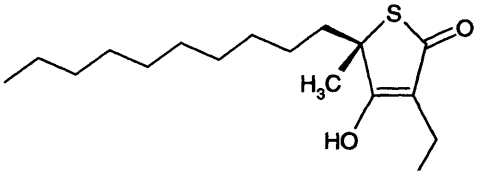
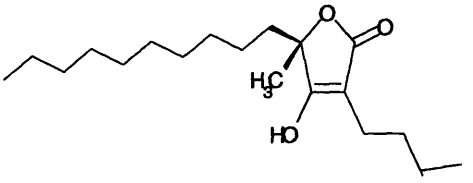
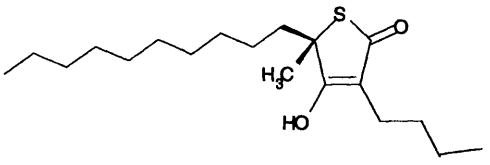
Compound No.	Description	Structure	LD_{50} (μM)
922	O1 C5 hexyl		**222**
779	S1 C5 Hexyl		**139**
923	O1 C5 Decyl C3 Ethyl		**94**
821	S1 C5 Decyl C3 Ethyl		**15**
945	O1 C5 Decyl C3 Butyl		**57**
944	S1 C5 Decyl C3 Butyl		**71**

Table 6.8 A comparison of sulphur and oxygen containing heterocycles

See the Legend of Table 6.2 for further details.

6.6.2 Replacement of the sulphur heteroatom with a nitrogen

An alternative to replacing the sulphur with an oxygen atom could be the introduction of a nitrogen atom. This might improve the activity against the parasite and would also provide the opportunity to extend from position 1 of the ring structure (Figure 6.5), which has so far been impossible.

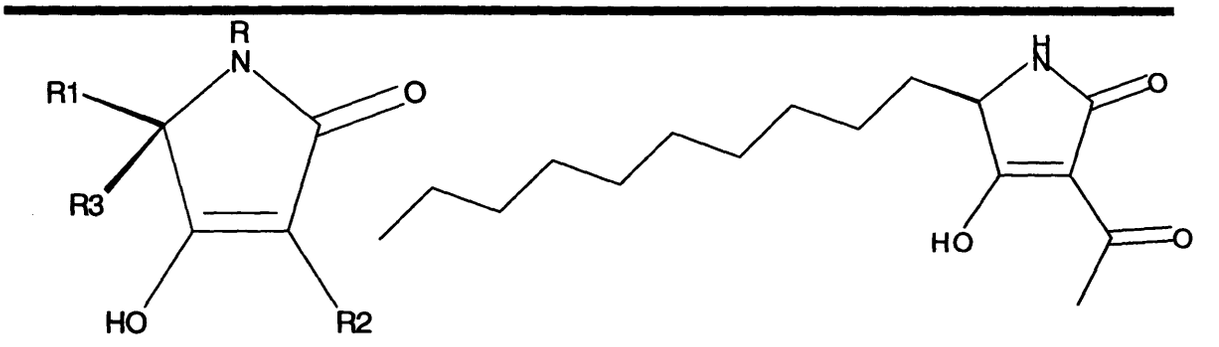


Figure 6.5 The general structure of nitrogen heterocycles and compound 1084

The synthesis of nitrogen heterocycles proved difficult since the adopted variation of the original scheme produced a very low yield (Simon Jones, personal communication). A new synthetic route was designed and a single compound, 1084, was produced in sufficient quantity and purity for biological testing. Unfortunately, as a result of the synthetic pathway the C3 methyl group was replaced by a ketone and the C5 methyl group was replaced by a hydrogen atom (Figure 6.5). Thus, with the nitrogen heterocycle compound, it is difficult to evaluate the effect of replacing the sulphur atom with nitrogen, as no direct comparison could be made with previous compounds.

Compound 1084 showed an LD_{50} of 63μM against cultures of *P. falciparum*, which is an improvement on TLM. The only compound tested previously that contained decyl and hydrogen moieties at position 5 is 943 (LD_{50} = 72μM), this compound had an ethyl group at the C3 position. Whilst there is a slight increase in inhibition observed with 1084 compared to 943, we cannot ascertain whether contributions from the heteroatom nitrogen or the ethyl to ketone substitution at C3 are solely or partially responsible for this increase.

6.7 Addressing problems in solubility

In preparation for biological assays each compound is dissolved in 20% DMSO before diluting to 2% DMSO in the final assay, where the solvent itself had no detectable effect. Some analogues with hexadecyl chains were turbid when dissolved at 2mM concentrations in 20% DMSO, which suggested that these compounds were not completely dissolved at these concentrations in this solvent. Other solvents (ethanol, tetrahydrofuran) were used to dissolve these compounds, however, these solvents had a significant impact on the assay itself (results not shown). These undesirable effects were to the detriment of the accuracy of the assays. Therefore, in order to improve the solubility of some of these compounds it was decided to attempt the synthesis of long chain C5 hydrocarbon chains containing ether or hydroxyl groups.

Three compounds, 1021, 1086 and 1195, all possessed C5 chains containing oxygen either as ether linkages or as a terminal hydroxyl group. These groups did not contain chain lengths equivalent to the hexadecyl chain observed in previous compounds, due

to the limited range of commercially available precursor molecules for use in synthesis. These compounds all dissolved at high concentrations in 2% DMSO and the solubility problem appeared to have been addressed. The three derivatives were all tested against *in vitro* cultures of *P. falciparum* (Table 6.9). These compounds all showed LD_{50} values in the 50-100μM range and did not appear to exhibit any increased inhibition compared to compounds with saturated hexadecyl C5 (R1) side groups.

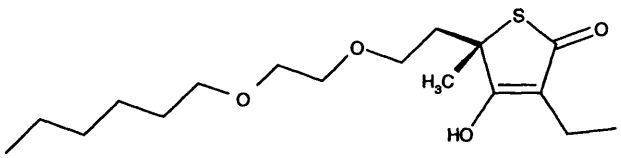
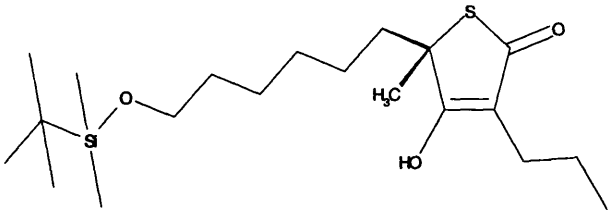
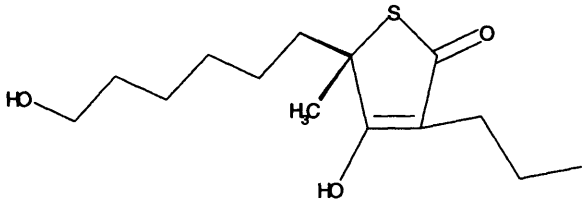
Compound No.	Structure	LD_{50} (μM)
1021		**61**
1086		**92**
1195		**>73**

Table 6.9 Attempts to overcome solubility in long chain C5 hydrocarbon chains

See the Legend of Table 6.2 for further details.

6.8 Extending from the 4-hydroxyl group

The structure of the *E. coli* FabB-TLM complex shows that the O2 oxygen at position 4 is involved with hydrogen bonding to the carbonyl oxygen of residue 270 and the amide nitrogen of residue 305 through a lattice of water molecules (Figure 6.6). These interactions may be central to the binding of TLM to the active site of the enzyme. The water molecules are positioned at the base of the active site tunnel that would incorporate the malonyl-ACP substrate and the phosphopantetheine arm of ACP. The synthesis of compounds with modifications of the C4 hydroxyl group would allow investigation and exploitation of possible interactions with the active site tunnel. Six compounds were synthesised utilising the C4 hydroxyl oxygen to generate an ether linkage to short and medium length hydrocarbon tethers. These were tested against the parasite as described before (Table 6.10).

All of the compounds showed an improvement in the level of inhibition of parasite cultures compared to TLM. Compound 1152 exhibited the best activity against *P. falciparum* cultures within this project, demonstrating the first LD_{50} ~1μM. The C3 propyl substitution continues to exert a positive influence on the efficacy of the compounds (compare 1023 to 1152). When compared to their hydroxyl containing counterparts (compare 776 LD_{50} = 36μM to 1023 LD_{50} = 6μM and 925 LD_{50} = 10μM to 1152 LD_{50} = 1μM), these analogues exhibited a six-fold and ten-fold increase in inhibitory activity. Compounds 1023 and 1152 had longer-chains that ended in an ethene group and this unsaturated bond may improve interactions with the hydrophobic tunnel. The replacement of the double bond with a benzyl group (compound 1192) slightly decreased the inhibitory activity. Suggesting that this large

group may cause some steric hindrance in the active site tunnel compared to the double bond of compound 1152.

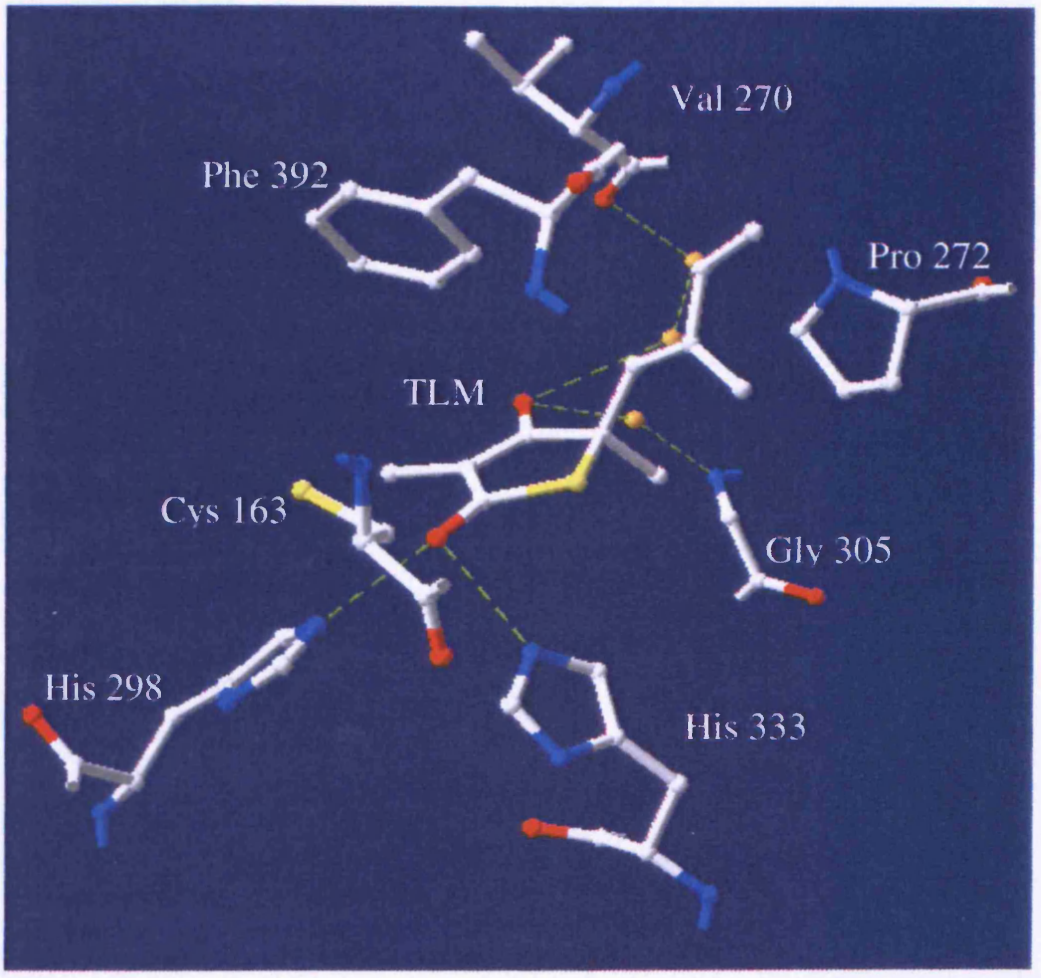


Figure 6.6 TLM bound in the active site of *E. coli* FabB

The inhibitor and key residues of *E. coli* FabB are shown. Hydrogen bonds are shown as green dashed lines. Water molecules are shown in orange. Hydrogen bonding between the C4 hydroxyl group of TLM and the backbone of two of the amino acid residues is shown. The 1FJ4.pdb file (Price *et al.*, 2001) was manipulated using SwissPDB viewer (Guex and Peitsch, 1997).

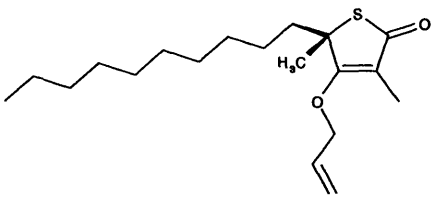
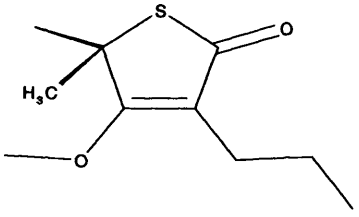
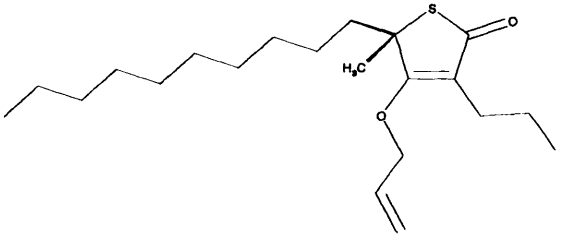
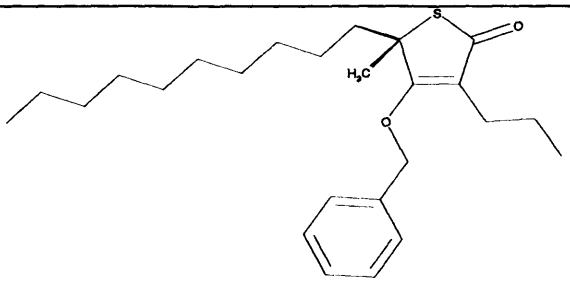
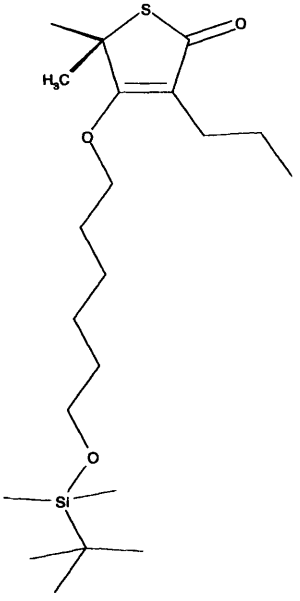
Compound No.	Structure	LD_{50} (μM)
1023		**6.3**
1140		**44**
1152		**1.3**
1192		**4**
1193		**2.7**

Table 6.10 (continued overleaf)

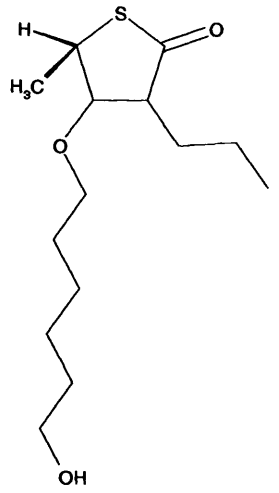
1210		**>18**

Table 6.10 Derivatives at the C4 position show a significant improvement in activity against the *P. falciparum* parasite *in vitro*

See the Legend of Table 6.2 for further details.

Compounds 1140, 1193 and 1210 all contained a C5 (R1) hydrogen group and cannot be directly compared to compounds 1023, 1152 and 1193, which all contain a C5 decyl side chain. Interestingly, compound 1193 that possessed an ether linked longer hydrocarbon chain and terminating in a silicon group showed the second best LD_{50} within this project suggesting that this group had favourable interactions between the target enzymes.

The presence of five different ether linked groups across the series of six compounds that showed improvement of activity compared to their hydroxyl counterparts, may indicate that the hydroxyl group contributes minimally to the binding of TLM and its derivatives to the physiological relevant target in *P. falciparum.*

6.9 Saturated ring structures (Dicarbonyl compounds)

Recently, Minnikin reported the synthesis of thiophenedione compounds (Figure 6.7) (Minnikin *et al.*, 2002). These compounds allow the addition of two groups at the C3 position thus permitting the investigation of further groups at this position. Such structures also differ from TLM in two major ways. Firstly, the C2, C3 and C4 of TLM is planar due to the double bond between C3 and C4, whilst the new thiolactone ring will form a non-planar structure that may significantly alter the position of the side groups. In turn, this may disrupt existing interactions between TLM and target proteins but also allows for the formation of new interactions. The dicarbonyl compounds should allow a decision to be made whether the planar structure of the ring is essential for the inhibitory activity of TLM and its derivatives. Secondly, the 4-hydroxyl is replaced with a carbonyl moiety, which will almost certainly have some effect on the biological activity of the compound. We have already reported that, by placing an alkyl chain on the 4-hydroxyl position through an ether linkage, there was an increase in the antimalarial efficacy of some compounds. It is likely that the carbonyl oxygen atom will be in a position very similar to the 4-hydroxyl oxygen, however, whilst the carbonyl oxygen retains its ability as a hydrogen bond acceptor, it has lost its ability as a donor to hydrogen bonds.

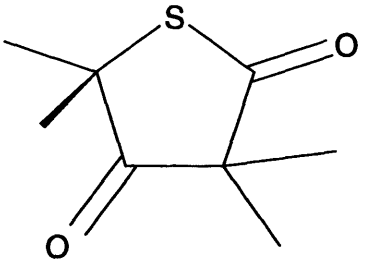


Figure 6.7 A thiophendione

Five saturated ring, dicarbonyl compounds were synthesised with varying substitutions at the C3 and C5 positions. These were tested against cultures of blood stage *P. falciparum* and the results are shown in Table 6.11.

Assays of the five compounds showed that saturated ring structures are capable of exhibiting a greater activity against the parasite compared to TLM. The replacement of a C5 hydrogen with a C5 decyl group increased the inhibitory activity once more (compare 1022 to 1020) and a propyl group. The addition of a propyl group at the C3 position (compare1041 to 1022) also caused a small increase in inhibition. Minnikin and colleagues reported the synthesis of a saturated compound with geranyl and methyl groups at position 3 and this compound displayed an increased activity against cultures of *M. tuberculosis* (Minnikin *et al.*, 2002). The latter compound also appeared to be inactive against the Type II FAS and mycolic acid biosynthesis systems in the bacteria, suggesting that the site of action of such a TLM derivative has yet to be elucidated. Compound 1194 differs from the Minnikin group compound by a substitution of the C3 methyl group with a C3 propyl group because we had already shown that this modification conferred has enhanced activity against *P. falciparum* (comparison of 1141 to 1022).

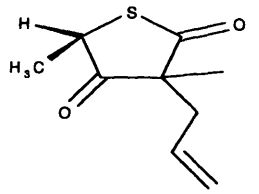
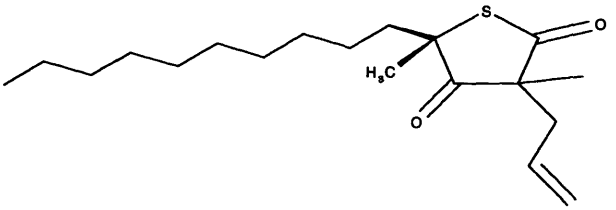
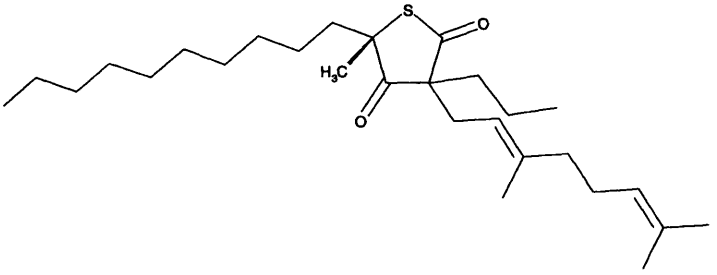
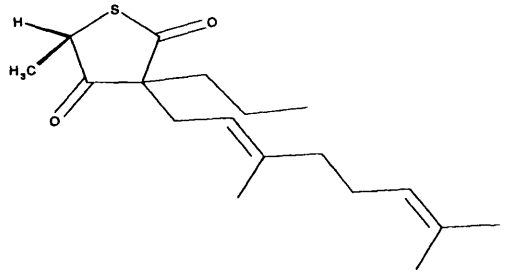
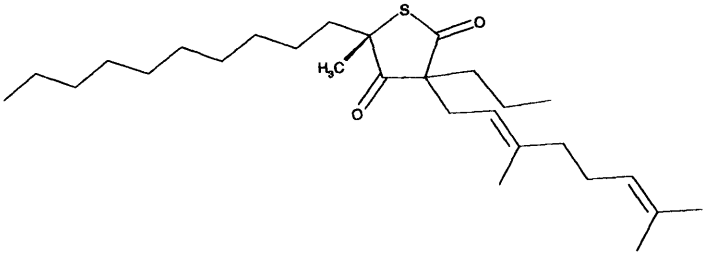
Compound No.	Structure	LD_{50} (μM)
1020		**62**
1022		**26**
1141		**19**
1194		**8.5**
1209		**>11**

Table 6.11 Open ring, di-carbonyl compounds show that the thiolactone ring is not essential for antimalarial activity

See the Legend of Table 6.2 for further details.

Compound 1194 showed the best inhibitory activity of all the saturated thiolactone ring derivatives. This compound contains only a C5 hydrogen group and may be

expected to have a lower activity than compound 1141. However, compound 1194 also possessed a geranyl group at the C3 position. Initial studies suggest that the conformation of this compound in the active site of pfFabF allows the positioning of the geranyl group in the hydrophobic pocket in which the C5 group of TLM is positioned (results not shown). This may explain why compound 1209 containing both a C5 decyl and C3 geranyl group exhibited decreased inhibition compared to 1194, as both hydrocarbon chains would not be able to fill the hydrophobic pocket simultaneously. These inhibitory activity exhibited by these five compounds suggest that the unsaturated ring structure and the 4-hydroxyl, present in TLM, are not essential for the inhibition of the target enzymes.

6.10 Activity of TLM derivatives against MBP-pfFabF, a *P. falciparum* condensing enzyme

Of the two *P. falciparum* condensing enzymes that have been identified (this work; Waller *et al.*, 1998; Prigge *et al.*, 2003) MBP-pfFabF appears to be much more sensitive to the antibiotic thiolactomycin than MBP-pfFabH. This may indicate that MBP-pfFabF is the physiologically relevant target of TLM. Derivatives of TLM have been shown to exhibit increased activity against Type II FAS in pea plants, although purified individual enzymes were not assayed in this paper (Jones *et al.*, 2000). The same derivatives exhibited an increase in inhibitory activity towards cultures of *P. falciparum* in erythrocytes (Waller *et al.*, 2003; work presented in this Chapter). So far in this Chapter, many TLM derivatives with alterations in almost all positions of the thiolactone ring have been described. In this section, all of TLM derivatives were tested against the MBP-pfFabF fusion protein to discover whether any displayed enhanced activity against the enzyme.

The activity of the MBP-pfFabF protein was tested separately in the presence of 100μM of each of the TLM derivatives. These assays should aid the identification of any derivatives that exhibit an inhibitory activity against the enzyme similar to TLM (IC_{50} = 3 μM (Figure 6.8)).

Unfortunately, none of the TLM derivatives showed similar or enhanced inhibitory activity against the MBP-pfFabF enzyme. To ascertain why there is an apparent lack of activity against this enzyme, several lines of inquiry were pursued. Firstly, the physical properties of the synthetic compounds i.e. purity, degradation during storage, mixtures of enantiomers skewed towards the inactive (5S)-enantiomer, should be investigated fully. Secondly, the structures of these compounds should be analysed to ascertain if there is any rational postulate that might explain their lack of activity against MBP-pfFabF.

The physical attributes of each TLM derivative were analysed immediately after synthesis (Jones *et al.*, 2004; Simon Jones personal communication). NMR and mass spectroscopy were used to test the purity, and confirm the identity, of each compound. In these studies, the synthesis of the thiolactone core was based on a method described in the first report of the synthesis of a racemic mixture of TLM (Wang and Salvino, 1984). Therefore, it was expected that both the (5S) and (5R) enantiomers would be equimolar. Subsequently, a test of the chirality of one of the compounds (925) showed that this derivative was a 50:50 racemic mixture. The chirality of other

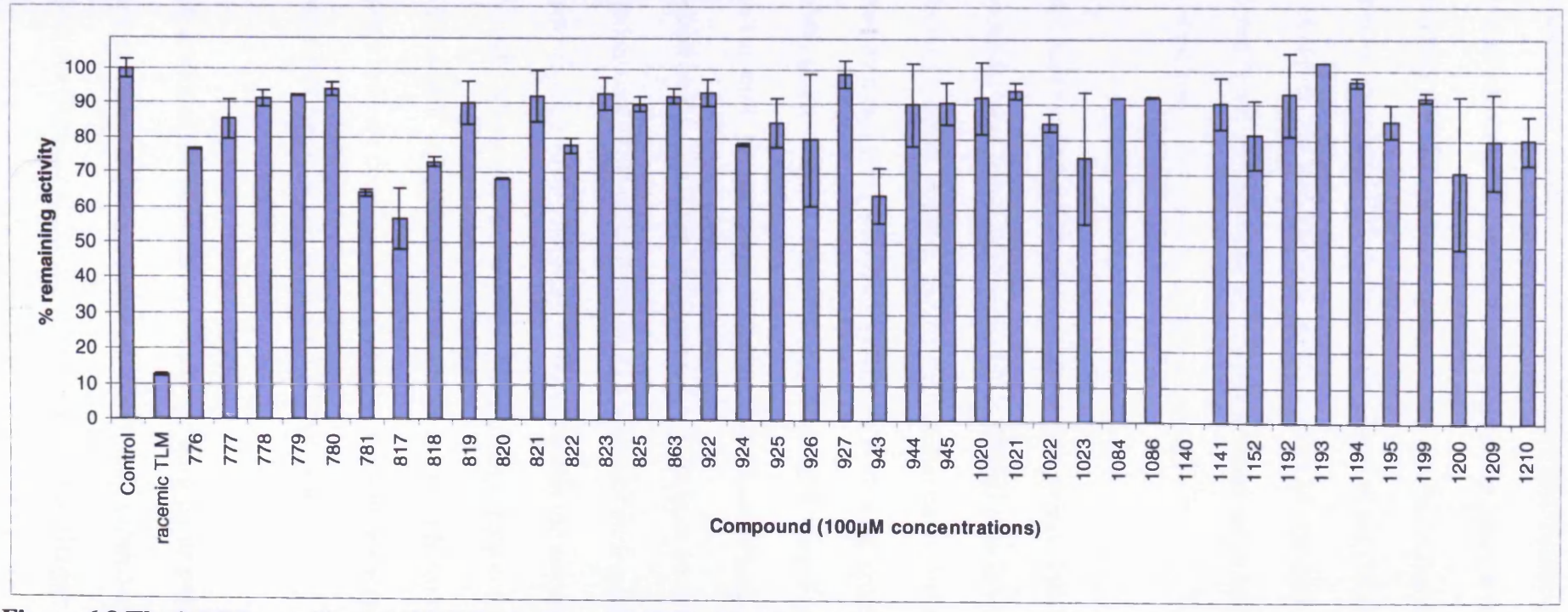


Figure 6.8 The inhibitory effect of TLM derivatives on MBP-pfFabF activity

Assays were performed as single point concentrations as described in Section 2.2.24. Compound 1140 was not tested against this enzyme because the stock was exhausted prior to the assays being performed. Assays were performed in duplicate and errors are shown as errors from the mean.

compounds was not determined as there were insufficient quantities of all the other TLM derivatives to be used in this test. However, there is no apparent chemical reason why a 50:50 racemic mixture would not be created and this is supported by the results for compound 925. Antagonistic effects of the (5S)-enantiomer cannot be ruled out. However, the 50:50 racemic mixture of TLM obtained from Sigma showed no evidence of antagonism from the (5S)-enantiomer as its inhibitory potency was half of that of the pure (5R)-enantiomer (Figure 5.16).

Compounds synthesised at a later stage of this project (1020-1210; 16 of the 40 compounds) were tested against the MBP-pfFabF protein within six months of their synthesis. However, some of the derivatives that were produced in early 2001 were assayed over two years after their synthesis. Following synthesis, compounds were generally stored in the absence of solvents at -80°C. Some compounds were stored in ethanol at -80°C for two years. The potential effects of long term storage in ethanol is described later in this Chapter (Section 6.12). At least 10 of the TLM derivatives were assayed against MBP-pfFabF within two weeks of their synthesis and these compounds showed no inhibitory activity towards the enzyme (compounds 1140-1210). The purity and physical properties of these compounds were tested at the time of the MBP-pfFabF assays and there was no evidence of degradation. Therefore, it appears unlikely that degradation of these compounds was an important factor in their apparent lack of activity against MBP-pfFabF.

The structures of the TLM derivatives tested during this project are diverse compared to those published in previous studies, which only showed alterations at the C3 and C5 positions (Jones *et al.*, 2000; Waller *et al.*, 2003; Kremer *et al.*, 2000). However,

despite the majority of these TLM derivatives showing increased inhibitory activity against *in vitro* cultures of *P. falciparum*, none of the compounds show increased activity against MBP-pfFabF. This may suggest that the alterations that have been made to the TLM structure decrease the affinity of these compounds for the active site of MBP-pfFabF. Of all of the side chains present in the structure of TLM, the isoprenoid group at C5 is only present in compound 781, a compound which we subsequently believed to have degraded significantly due to storage in ethanol. Recent evidence has shown that this group appears to be central for the inhibitory activity of TLM derivatives against bacterial FabB enzymes (Price *et al.*, 2001). A double bond of the isoprenoid group is positioned between two peptide bonds (391-392) and (271-272) the interaction between these three delocalised systems is important to TLM specificity in *E. coli*. This may suggest that the alterations made to the C5 side chains, in this project, increased the activity against *in situ* cultures of *P. falciparum* (Table 6.2) but caused a significantly decrease in the activity against MBP-pfFabF due to the lack of an isoprenoid group.

6.11 Activity of TLM derivatives against MBP-pfFabH

In an attempt to identify compounds that exhibited enhanced activity against the first condensing enzyme identified in *P. falciparum*, the activity of the MBP-pfFabH protein was assayed in separate reactions containing 100μM of all of the TLM derivatives described in this report. Racemic TLM showed an IC_{50} = 200μM (Figure 5.18), however inhibition of 30% of MBP-pfFabH activity was observed in the presence of 100μM of the compound. Therefore, any compound that showed similar or enhanced inhibitory activity towards MBP-pfFabH should be identified in

this assay. The results of assays containing 100μM of each of the TLM derivatives are shown in Figure 6.9.

The first batch of compounds (776-780) that were assayed against MBP-pfFabH contain a replacement of the C5 isoprenoid group of TLM with other hydrocarbon chains. One of these compounds, 777, showed a significant increase in inhibitory activity towards MBP-pfFabH. This compound possesses a hexadecyl ($C_{16}H_{33}$) side chain, the longest C5 hydrocarbon group synthesised during this project. Compounds containing a decyl or geranyl side chain at the C5 position (776 and 778, respectively) exhibited activity against MBP-pfFabH, but were not significant improvements compared to racemic TLM provided by Sigma. Shorter chain hydrocarbon groups at C5 had negligible activity.

A similar trend in the change of activity with C5 substitutions was seen in a second series of derivatives, which possess an ethyl substitution at the C3 position. Compounds with C5 hydrocarbon chains eight carbons, or fewer, in length, exhibited negligible inhibitory activity against MBP-pfFabH. However, compounds containing the ten carbon (decyl, compound 821) group at C5 inhibited approximately 40% of MBP-pfFabH activity, whilst the sixteen carbon (hexadecyl, compound 819) group inhibited over 60% of the enzyme activity (Figure 6.9). Compound 819 possesses a C3 ethyl group and the inhibition by this compound was slightly greater than that of compound 777, which possessed a C3 methyl group, however this difference was not statistically significant.

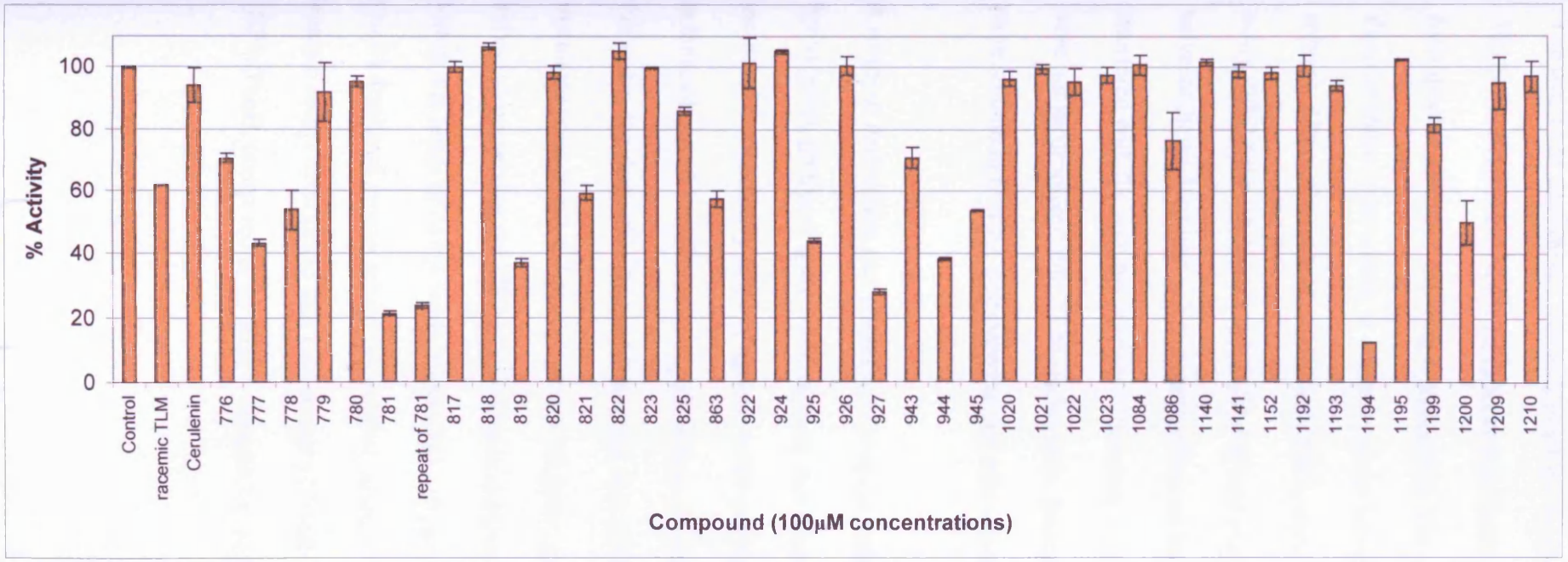


Figure 6.9 The inhibitory effect of TLM derivatives on MBP-pfFabH activity

Assays were performed as single point concentrations as described in Section 2.2.25. Assays were performed in duplicate and errors are shown as errors from the mean. The majority of errors were within 5% of the average inhibition.

The two TLM derivatives that exhibited the highest inhibitory effect against MBP-pfFabH possessed C5 hexadecyl side chains. Unfortunately, due to the very hydrophobic nature of the long hydrocarbon side chain, it was extremely difficult to dissolve these compounds, at concentrations utilised in the assays, using DMSO as a solvent. Alternative solvents (ethanol, THF) were evaluated for their suitability in the assay. Whilst the C5 hexadecyl TLM derivatives were soluble in both solvents, the solvents themselves had a detrimental effect on the assay, suggesting that they interfered with the enzyme activity. Therefore, in further studies, hexadecyl groups were not incorporated into TLM derivatives. Instead, decyl groups that were much more soluble in DMSO, compared to hexadecyl derivatives, were synthesised.

A series of derivatives that possessed C5 decyl hydrocarbon chains and different C3 hydrocarbon groups (Figure 6.10) showed that increasing the size of the C3 group increased the inhibitory activity against MBP-pfFabH. This trend was similar, in part, to those observed in cultures of *P. falciparum in situ* (Figure 6.10). However, within this series, the derivative that exhibited the best inhibitory activity against the whole organisms possessed a C3 propyl group, whilst in assays of MBP-pfFabH the derivative possessing the C3 benzyl group exhibited the highest inhibitory effect. Within this series of compounds, three (925-C3 propyl, 944-C3 butyl and 927-C3 benzyl) displayed greater inhibitory activity against MBP-pfFabH compared to TLM. Separate assays determined that compound 927 had an $IC_{50} \approx 50\mu M$ against MBP-pfFabH, compared to an $IC_{50} \approx 200\mu M$ for TLM.

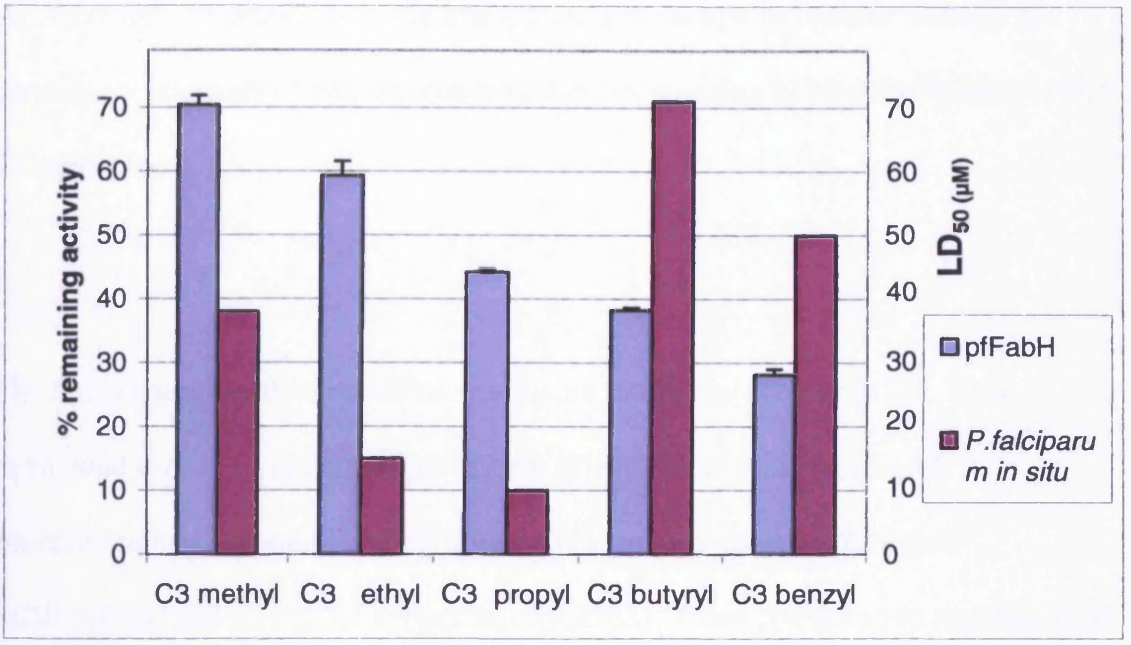


Figure 6.10 TLM derivatives with different C3 groups

Compounds with hydrocarbon chains of increasing size exhibit increased inhibitory activity against pfFabH. Increasing the side chain size from methyl to propyl also improved the LD_{50} value against the *in situ* cultures of *P. falciparum*. However, compounds possessing a C3 butyl and C3 benzyl group decreased the inhibitory activity of these compounds against the parasite.

Of the many alterations to the structure of thiolactomycin, only one derivative exhibited increased activity compared to that of compound 781 (synthetically prepared TLM). At a concentration of 100μM compound 1194 exhibited the highest inhibitory activity described so far in this Chapter, inhibiting 87.5% activity of MBP-pfFabH activity. The IC_{50} was approximately 18μM against MBP-pfFabH, an order of magnitude greater than that observed by with TLM. A 100μM concentration of this compound did not significantly inhibit cultures of *P. falciparum* and suggests

that there may be problems in the transport of this compound into or through the parasite, which would mean the compound is not reaching MBP-pfFabH localised in the apicoplast.

The final compound that exhibited significant inhibitory activity is 781. This compound should be structurally identical to thiolactomycin. At 100μM concentrations, compound 781 exhibited the second highest inhibition of MBP-pfFabH and an IC_{50} of approximately 25μM. When compared to racemic TLM (Sigma), compound 781 showed a large and significant increase in inhibitory activity towards MBP-pfFabH (Figure 6.9). One explanation of this could be a predominance of the active (5R)-enantiomer in compound 781. However, the chirality of compound 925 suggested that compounds produced using the same chemical route are a 50:50 racemic mixture. The chirality of compound 781 was not tested itself due to extremely low yields. Furthermore, if there was a predominance of the (5R)-enantiomer, the inhibitory activity against MBP-pfFabF (Figure 6.8) should be increased compared to that of racemic TLM, when in fact, it is much lower. Even if compound 781 was synthesised as on the (5R)-enantiomer is should not exhibit more than twice the activity of Sigma TLM. In combination, these data suggest that compound 781 is not thiolactomycin.

6.12 An investigation of compound 781 structure

Following the synthesis of compound 781 in early 2001, NMR and mass spectrometry were performed on the compound. These tests confirmed that compound 781 had an

identical structure to thiolactomycin. The compound was immediately tested against *in situ* cultures of the *P. falciparum* K1 strain and an LD_{50}=143μM was observed. When evidence was obtained that suggested compound 781 was not TLM, synthetic racemic TLM (Sigma Aldrich, Poole, UK) was also tested in cultures of the parasites. This racemic compound exhibited an LD_{50} >95μM, although a more accurate value was not obtained. This may indicate that when sent for testing against the K1 strain, compound 781 was a racemic mixture of TLM enantiomers and no significant degradation had occurred. The value of these LD_{50}s was lower than the 50μM previously reported against the W2mef strain (Waller *et al.*, 1998), however, strains may have significantly different susceptibilities to different inhibitors.

Assays of the inhibitory activity of compound 781 against MBP-pfFabF and MBP-pfFabH were performed in October 2003 and January 2004, respectively. During this long period of time between synthesis and testing, compound 781 had been stored at -80°C. However, in contrast to other compounds, compound 781 had not been stored dry but was dissolved in ethanol. An explanation behind this occurrence is explained herein. In preparation for use in biological and enzymatic assays, each compound was dissolved in ethanol because of the need to remove accurate μg quantities of the compound. An aliquot of the compound dissolved in ethanol was then removed and the ethanol evaporated under nitrogen. The compound could then be dissolved in the correct volume of DMSO to obtain the desired concentration. It appears that with respect to compound 781 the ethanol had not been evaporated from the original vial when an aliquot was removed for testing against *P. falciparum* cultures in early 2001. Had storage in ethanol for over two years altered the composition of compound 781?

6.12.1 The effect of storage in ethanol on compound 781

In an attempt to see if compound 781 had degraded whilst being stored in ethanol for a long period of time, the compound was analysed by mass spectrometry. The results (Appendix 9), when compared to the mass spectrometry data of the newly synthesised compound, showed a decrease in the peak at 211.0 that corresponds to the molecular weight of thiolactomycin. Several other peaks were observed but could not be easily identified because the remaining amount of 781 was insufficient for detailed NMR analysis.

The synthesis of a second batch of compound 781 would prove useful because it could be stored in ethanol and NMR and mass spectrometry could be performed on it at regular intervals alongside enzyme inhibitor studies. Frustratingly, it proved impossible to purify a second batch of compound 781 throughout the course of this project, despite several alternative synthetic routes being implemented. It appears that the isoprenoid group at the C5 position is remarkably difficult to synthesise compared to other hydrocarbon chain substitutions at C5 (work in this project and a personal communication from Charles Rock, St. Jude Children's Research Hospital, Memphis, USA).

As there were insufficient quantities of 781 for NMR analysis, racemic TLM (Sigma Aldrich, Poole, Dorset, UK) was stored in ethanol to determine if this had an effect on its inhibitory activity against MBP-pfFabH. Storage in ethanol for six weeks at 25°C showed an increase of 16% inhibitory activity compared to that not stored in ethanol (Figure 6.11), but this difference was not statistically significant. Analysis of this

compound in H^2 labelled ethanol suggested that some degradation had occurred but the main breakdown products have yet to be identified. The identification of the breakdown products is being actively pursued within this research project as it may yield a novel basis for the inhibition of pfFabH.

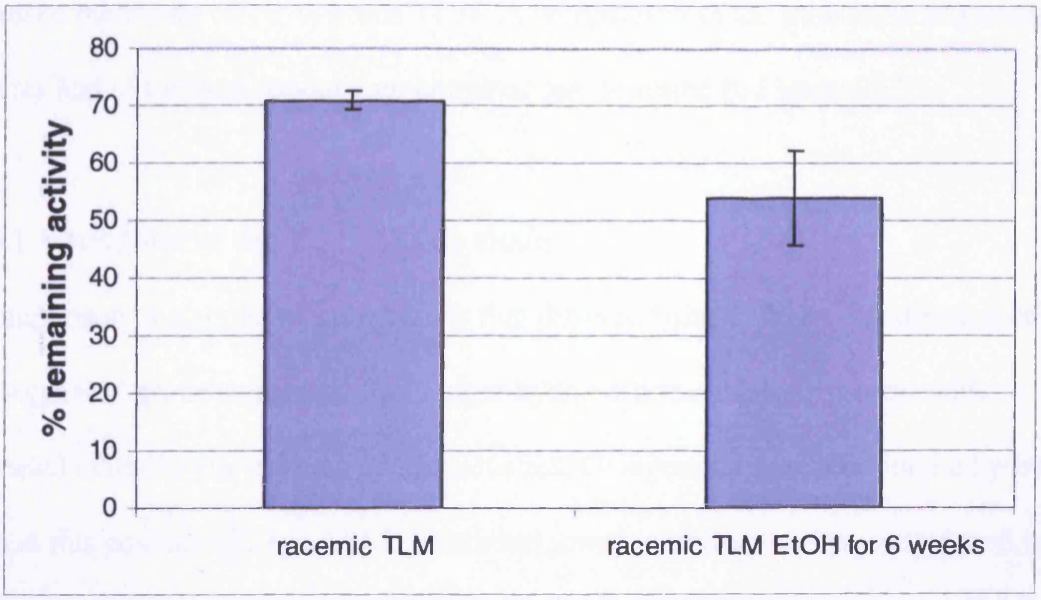


Figure 6.11 The effect of storage of TLM in ethanol on its inhibition of pfFabH

The assays were performed as described in Section 2.2.25, in triplicate and error bars show standard deviations from the mean.

6.13 Discussion

In this project, forty TLM derivatives were chemically synthesised and their inhibitory effect was evaluated in cultures of *P. falciparum* parasites and in assays of both the parasite's Type II FAS condensing enzymes, pfFabF and pfFabH. Encouragingly, against parasite cultures the majority of the derivatives showed enhanced inhibitory effect towards TLM. A comparison of the inhibition of parasite cultures and of the two condensing enzymes can be found in Figure 6.12.

6.13.1 Alterations of the C5 (R1) side chain

A comparison of a series of compounds that differed from TLM by substitution of the C5 isoprenoid group suggested that longer hydrocarbon chains correlated with increased inhibition in cultures of the parasites. Compounds that contained a hydrogen atom at this position (817 and 818) exhibited lower inhibitory activity compared to TLM whilst derivatives with saturated hydrocarbon chains longer than eight carbons (compound 780) show a much higher inhibitory effect (Table 6.2).

Increasing the size of the hydrocarbon group to ten carbon and sixteen carbon groups at the C5 (R1) position had an increased inhibitory effect on the parasites (Table 6.2-6.3). However, compounds containing hexadecyl groups were often not soluble in 2% DMSO, the solvent used in all biological and enzyme assays. Therefore, the results obtained with these compounds may not represent the maximal inhibitory effect that these compounds possess. Consider that if only 50% of these compounds were soluble, testing at 100μM concentrations means that only an effective concentration of 50μM would be available. To address this problem compounds were synthesised with C5 side chains designed to increase the solubility

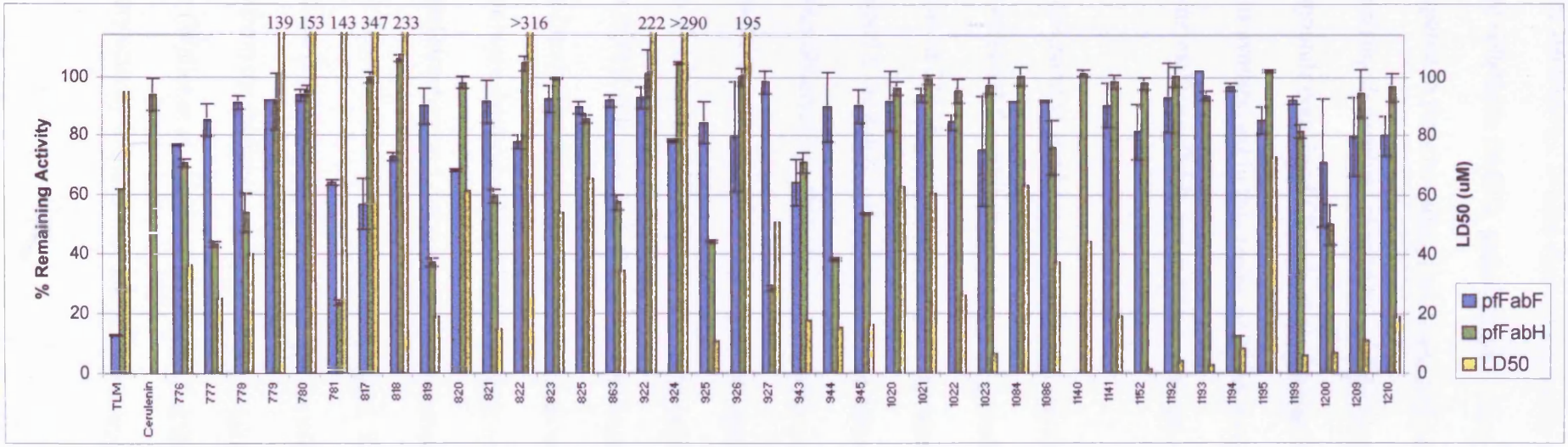


Figure 6.12 Combined chart showing the LD_{50} values and % inhibition of each TLM derivative synthesised in this project

pfFabF and pfFabH inhibition is shown as percentage remaining activity in assays containing 100μM of each compound is shown on the primary y axis. The inhibition of parasite cultures is shown as the LD_{50} values with μM concentrations of each derivative. LD_{50} values over 100μM are shown in μM above each bar.

of the derivatives. Whilst these compounds (1021, 1086 and 1195) appeared to be more soluble in DMSO, unfortunately, none of them exhibited increased activity compared to their hexadecyl analogues. However, direct comparison to hexadecyl possessing derivatives should be avoided because none of these more soluble compounds contained C5 side chains longer than 12 atoms. The length of the C5 side chains synthesised in this series of TLM derivatives was limited by the number of commercially available precursor molecules.

The unsaturated and branched chains at the C5 (R1) position of TLM did not appear to be essential for inhibition of parasite growth. However, the geranyl side group of compound 778 showed enhanced activity when compared to compound 780 that contained a saturated hydrocarbon of the same length. These data reinforce the previous observations that compounds with more unsaturated, branched hydrocarbon groups have an increased inhibitory effect against *P. falciparum in situ* (Waller *et al.*, 2003). The isoprenoid group of thiolactomycin consists of a four carbon backbone with a methyl side group on the second carbon from the thiolactone ring and two double bonds. Unfortunately, no compounds with a saturated butyl group at the C5 position were synthesised within this project and analysis of the contribution of the branched structure and unsaturation of the isoprenoid group cannot be obtained directly from the data detailed in this project. Waller and colleagues showed that TLM derivatives that possess a four carbon group with a methyl side group on the third carbon from the thiolactone ring and only a single double bond show a decrease in activity (Waller *et al.*, 2003). The removal of the double bond also decreased the inhibitory activity towards the malaria parasites.

The presence of the isoprenoid group at the C5 position of TLM may be essential for binding of the compound to the area surrounding the active site of pfFabF. None of the TLM derivatives synthesised during this project contained a C5 isoprenoid group, and all of these compounds were largely ineffective against the enzyme (Fig 6.9). Attempts in Cardiff to synthesise derivatives with C5 isoprenoid groups and other modifications were unsuccessful. This appears to be due to the difficulty of joining an isoprenoid group at the C5 position, due to the double bond at the secondary carbon of the precursor.

Interestingly, compound 1152, which exhibited the highest inhibition in cultures of the parasite (LD_{50} = 1.3μM) contained a modification of the C4 hydroxyl group to an ether linked unsaturated hydrocarbon. Structural studies of the *E. coli* FabB enzyme suggest that this group would be orientated towards the hydrophobic pocket that accommodates the growing acyl chain (Price *et al.*, 2001). The complex of the inhibitor cerulenin in the active site of the *E. coli* FabB enzyme suggested that the long hydrocarbon chain of this inhibitor is positioned in the same pocket (Price *et al.*, 2001). Therefore, hydrophobic interactions between the extended chain and the surrounding amino acids of the hydrophobic pocket might be responsible for the increase activity against of malaria parasites. However, when tested against the pfFabF enzyme at 100μM concentrations only ~20% of the total enzyme activity was inhibited.

Alterations at the C5 (R1) position also had some influence on the level of inhibition observed in the assays containing the pfFabH protein. Compounds containing C5 (R1) decyl or hexadecyl groups consistently showed increased inhibition of the enzyme

compared to those with shorter side chains. In some comparisons the hexadecyl derivatives appear to have a greater inhibitory effect (Figure 6.9 – compare 776 (decyl) to 777(hexadecyl); 821(decyl) to 819 (hexadecyl)) and in other cases decyl analogues had increased effect (compare 925 (decyl) to 1195 (hexadecyl); 927 (decyl) to 1200 (hexadecyl). This may suggest that these groups are having an effect outside of the active site area.

To date, the structure of TLM or any TLM derivative bound to a FabH enzyme has not been solved. Most likely because of the lower affinity of TLM for FabH proteins compared to FabB and FabF-like proteins. However, the hydrophobic pocket that accommodates the C5 (R1) isoprenoid group in *E. coli* FabB does not exist in the *E. coli* FabH protein. This suggests that these longer chain C5 (R1) derivatives would not fit into the active site of pfFabH, which shows high sequence identity to the *E. coli* FabH protein, without causing considerable changes in the protein structure.

There is evidence that bacterial and plant FabH proteins possess a highly conserved region of 7 amino acids near the C-terminus (Figure 3.1; *E. coli* numbering 273-279), which are involved in feedback inhibition of the enzyme by long chain acyl-ACPs and acyl-CoAs (Heath and Rock, 1996; Abbadi *et al.*, 2000). A N274D mutant of the *Cuphea wrightti* FabH enzyme abrogated the inhibitory effect of long chain acyl-ACPs whilst retaining all condensing enzyme activity. This conserved region is also present in the *P. falciparum* FabH protein and suggests that this enzyme may also be inhibited by the binding of long chain acyl groups. It is possible that the C10 and C16- acyl groups present in 100μM concentrations in these assays are interacting with this regulatory domain to inhibit enzyme activity. However, these acyl groups are not

bound to acyl carrier protein, which may contribute significantly to the interaction between the acyl-ACP group and the FabH protein. As further work to this project, the activity pfFabH protein should be tested in assays containing long chain acyl-ACPs to see if this feedback recognition exists in *P. falciparum*. Progress in the understanding of ACP docking to Type II FAS enzymes of *E. coli* (Zhang *et al.*, 2001 & 2003) may help to understand the interaction between the regulatory site and the acyl-ACP, with the aim to synthesise inhibitors of FabH proteins that bind to the regulatory site.

6.13.2 Alterations of the C3 (R2) side chain

Increasing the size of the C3 methyl group of TLM also increased the activity of the TLM derivatives towards the parasites. By extending the methyl group to a propyl group (compound 925) an LD_{50} of 10μM was achieved (Table 6.5). Increasing the propyl tether to a butyl or benzyl derivative resulted in a decrease in activity (LD_{50}s of 71μM and 50μM, respectively) (Table 6.5). Docking experiments of compound 925 in the active site *E. coli* FabB (the homologue of pfFabF the assumed target of TLM inhibition in *P. falciparum*) performed during this project suggested that the propyl group fits snugly into the active site of this enzyme, with the terminus of the propyl chain in close proximity to the active site Cys163. Unfortunately, possibly due to the lack of derivatives that possessed a C5 (R1) isoprenoid group, no increase in inhibitory activity was observed in substituted C3 groups towards the pfFabF protein.

Encouragingly, the increase in activity towards whole parasites may, in part, be explained by the increase inhibition of pfFabH by compounds containing extended C3 chains (Figure 3.11). The hydrophobic pocket in this enzyme that incorporates the C3

group is possibly larger than that of *E. coli* FabB because butyl and benzyl groups exhibited increased activity compared to the propyl derivative (Figure 3.11).

6.13.3 A comparison of TLM derivatives in biological and enzyme assays

The majority of TLM derivatives synthesised during this project were produced prior to the production of purified recombinant pfFabF and pfFabH. These compounds were, therefore, tested against cultures of the *P. falciparum* K1 strain prior to studies in assays containing the condensing enzymes. A general increase in the inhibitory activity of the TLM derivatives against whole parasites is apparent in Figure 6.12. The yellow bars show that derivatives synthesised at later stages in the project had considerably lower LD_{50} values compared to those synthesised at the start of the project. Interpretation of early *in situ* data combined with considerations of the TLM bound to condensing enzymes was used to design rationally compounds with moderate but increased inhibition of *P. falciparum* parasites. Encouragingly, this strategy appeared to be making advances towards finding a sub micromolar inhibitor of parasite cultures.

Previous work in this project suggested that pfFabF is inhibited by TLM with an IC_{50} of approximately 3μM. Therefore, it may have been expected that TLM derivatives that exhibited increased activity against the parasites would also show increased inhibition of the pfFabF enzyme. However, this was not reflected in the assays of the protein, possibly due to the lack of a C5 (R1) isoprenoid group (discussed in Section 6.13.1). Whilst TLM inhibits the purified pfFabF enzyme *in vitro*, it is possible that the best derivatives against whole parasites were targeting the pfFabH condensing enzyme. However, compounds 1152, 1192 and 1193 (LD_{50} values of 1.3μM, 2.7μM

and 4μM, respectively) had no inhibitory effect on the pfFabH enzyme at 100μM concentrations (Figure 6.9). This suggested that neither of the Type II FAS condensing enzymes were the targets of the best inhibitors of whole parasites and that inhibition of whole parasites with these TLM analogues must proceed via an as yet unidentified pathway.

Only compound 1194 appears to target effectively whole parasites (LD_{50} = 8.5μM) and one of the condensing enzymes, (IC_{50} = 18μM towards pfFabH). This compound contains no large C10 or C16 hydrophobic groups at the C5 position and, therefore, may be expected to show increased solubility and uptake into the parasite compared to some of the other derivatives. It may also exhibit increased absorption and bioavailability in the host, which are favourable drug characteristics. With Sigma TLM, the large discrepancy between the IC_{50} of 3μM towards pfFabF and the LD_{50} value of >95μM (Figure 6.12), suggests that TLM may not be able to reach, or accumulate in, the apicoplast of *P. falciparum*.

Chapter 7

Chapter 7 General Discussion

7.1 Studies of the *P. falciparum* FabF protein

Work presented in this project describes the first reported cloning and expression of pfFabF, a condensing enzyme of the fatty acid synthase of the malaria parasite *Plasmodium falciparum*. The protein shows significant identity with plant and bacterial FabB and FabF enzymes and an alignment of the primary sequences of these proteins identified the mature pfFabF protein (Figure 3.1). There is direct evidence that the *P. falciparum* FabH and ACP proteins are targeted to the apicoplast (Waller *et al.*, 1998 and 2000). Furthermore, genes of the remaining enzymes of the Type II FAS pathway all encode an N-terminal bipartite sequence, which is predicted to be sufficient for targeting of these proteins to the apicoplast (Surolia and Surolia, 2001; Waters *et al.*, 2002; Prigge *et al.*, 2003; Pillai *et al.*, 2003; Sharma *et al.*, 2003). However, attempts to identify a region of the *pffabF* gene that encoded a bipartite signal sequence were unsuccessful. Despite the inability to find an apicoplast targeting sequence associated with pfFabF, it is very likely that this protein is targeted to the apicoplast and interacts with the other enzymes of the Type II FAS pathway. The production of GFP-bound antibodies that recognise pfFabF is now possible following the purification of the protein described in this work (care should be exercised so that the antibodies do not recognise MBP). Staining of *P. falciparum* cells with these antibodies may provide the first conclusive evidence that pfFabF is located to the apicoplast.

Purified recombinant pfFabF exhibited significant β-ketoacyl-ACP synthase activity using myristoyl-ACP (C14:0) as the acyl-ACP substrate and malonyl-ACP as the donor substrate. A significant decrease in activity when using palmitoyl-ACP (C16:0) and stearoyl-ACP (C18:0) as acyl-ACP substrates (Figure 4.4) suggests that *in vivo* the predominant final

product of the Type II FAS pathway in *P. falciparum* may be palmitoyl-ACP, although small amounts of stearoyl-ACP and arachidoyl-ACP could be feasibly produced. Unfortunately, the ability of the enzyme to synthesise shorter acyl-ACPs was not investigated due to the lack of shorter chain acyl-ACP substrates.

The Type II FAS pathway of *Trypanosoma brucei* predominantly produces myristate (Morita *et al.*, 2000; Paul *et al.*, 2001), which is subsequently incorporated into glycosyl-phosphatidylinositol (GPI) anchors of the variant surface glycoprotein (VSG), a protein that significantly contributes to the parasite's ability to evade the host immune response. In *P. falciparum* many important surface proteins including the circumsporozoite protein which coats the invading sporozoite and merozoite surface proteins (MSP) 1 and 2 are also attached to the membrane by GPI anchors. An analysis of MSP-1 and MSP-2 anchors in *P. falciparum* suggests that myristate makes up a quarter of the *sn*-1 and *sn*-2 acyl groups and that palmitate is the predominant acyl group at these positions (Gerold *et al.*, 1996) (Figure 7.1). Modification of the inositol group by fatty acids also occurs in GPI anchors (Figure 7.1) and at least one report suggests that myristate is the predominant fatty acid attached to the inositol group in *P. falciparum* (Gerold *et al.*, 1996). Myristate comprises ~1% of the fatty acids present in serum (Paul *et al.*, 2001), and there is no detectable myristate in the membranes of non-infected or infected red blood cells (Omodeo-Salé *et al.*, 2003). Therefore, the parasite must either stimulate the host to produce the myristate it requires or synthesise myristate *de novo*. Therefore, the synthesis of myristate as a final product of the *P. falciparum* Type II FAS pathway should be investigated in more detail. This could be achieved by measuring the incorporation of radiolabelled precursors into fatty acids, especially in the fatty acids of GPI anchors and by testing the ability of pfFabF to elongate lauroyl-ACP to myristoyl-ACP *in vitro*.

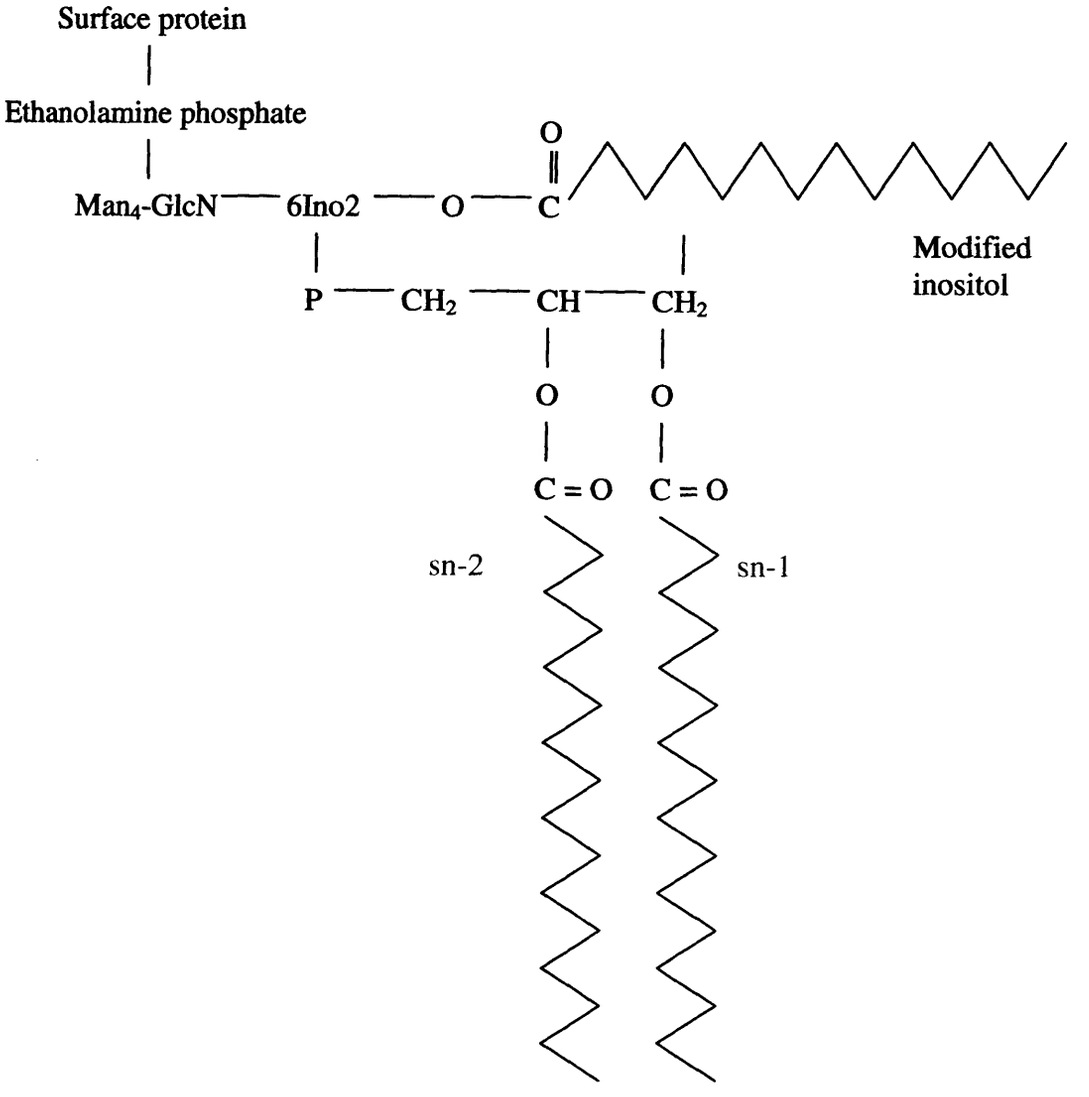


Figure 7.1 The basic structure of a *P. falciparum* GPI anchor

The structure is based on data described by analysis of the fatty acyl groups present on GPI anchors linked to merozoite surface proteins (MSP-1 and MSP-2) (Gerold *et al.*, 1996). Abbreviations; Ino (inositol); GlcN (N-acetyl glucosamine; Man (mannose); P (phosphate).

There have been two investigations into the incorporation of radiolabelled precursors into fatty acids of *P. falciparum*. The first report suggested that, in whole parasites, myristate was

the major product of the Type II FAS pathway with smaller amounts of decanoate (10:0) and laurate (12:0) also present (Surolia and Surolia, 2001). A separate investigation, in a cell free system, suggested that palmitate was the major product of this pathway with smaller amounts of myristate and laurate (Sharma *et al.*, 2003). This may indicate that myristate is the major fatty acid synthesised by the parasite *in vivo*. However, myristoyl-ACP synthesised in the cell-free system would not be incorporated into membrane lipids or GPI anchors, and may be elongated to palmitoyl-ACP in one final cycle of the pathway. The absence of longer, saturated chain fatty acids in the above reports, is in agreement with the huge decrease in the activity of pfFabF to elongate palmitoyl and stearoyl-ACPs *in vitro* (Figure 4.4).

Interestingly, in the cell-free system, whilst no decanoate was synthesised, a large band corresponding to the expected size of octanoate was present (Sharma *et al.*, 2003). This may indicate that octanoate is synthesised by the Type II FAS system of the malaria parasite but does not accumulate in whole parasites as it is converted into lipoic acid. Lipoic acid is an essential cofactor of the pyruvate dehydrogenase complex, predicted to be targeted to the apicoplast, and of several mitochondrial proteins essential for respiration.

7.1.2 Inhibition of pfFabF activity by TLM

The pfFabF enzyme was highly sensitive to the fungal metabolite thiolactomycin (IC_{50}=3μM; Figure 4.6), a compound that also inhibits growth in cultures of malaria parasites (this work; Waller *et al.*, 1998 and 2003). Data presented in this report indicate that TLM may inhibit growth of whole parasites via inhibition of the pfFabF enzyme. Proof that the pfFabF enzyme is inhibited by TLM, provides the further evidence of an *in vivo* target of TLM inhibition in malaria parasites. Identification of pfFabF as a target of TLM inhibition *in vivo* provides a

solid platform for further investigations into the potential of TLM or related compounds as candidates for antimalarial drugs. Information on the mechanism of action of TLM inhibition on the physiologically relevant target will be required before any potential drug could be licensed and reach the market.

7.2 The potential of TLM as a novel antimalarial drug

Thiolactomycin has several properties that make it a good lead compound for potential antimalarial drugs. This antibiotic is not toxic to mice at therapeutic doses and provides significant protection against bacterial infections (Miyakawa, *et al.*, 1982). The potency of TLM in infected mice is achieved by rapid absorption and distribution to infective sites (Miyakawa, *et al.*, 1982), this suggests favourable pharmacokinetics and suitability for drug administration. The discovery that triclosan inhibits pfFabI, a second enzyme of the Type II FAS pathway (Surolia and Surolia, 2001) provides the opportunity for a double block on the pathway which may decrease the emergence of resistance. The other proteins in this pathway are also attractive targets for antimalarial compounds. These enzymes are being characterised and a search for inhibitors of these enzymes is being actively pursued (Table 7.1) (Prigge *et al.*, 2003; Sharma *et al.*, 2003; Pillai *et al.*, 2003; Lakshmi Swarna Mukhi *et al.*, 2004).

Enzyme	Inhibitor	Inhibition details	Reference
FabF	Thiolactomycin	IC_{50} ~3μM	This thesis
FabH	1,2-Dithiole-3-one compounds	IC_{50} = 0.53-10.4μM	Prigge *et al.*, 2003
	Thiolactomycin	IC_{50} >330μM	
FabG	Possibly epigallocatechin gallate (EGCG)*	-	Zhang and Rock, 2004
FabZ	NAS-21	IC_{50} 100μM	Sharma *et al.*, 2003
	NAS-91	IC_{50} 7.4μM	
FabI	Triclosan	IC_{50} 700nM	Kapoor *et al.*, 2001
	2,2-dihydroxyphenyl ether	Ki 18nM	

Table 7.1 Known inhibitors of *P. falciparum* fatty acid synthesis enzymes

* The first reported inhibition of *E. coli* FabG by polyphenol extracts of green tea e.g. epigallocatechin gallate, was reported recently. The inhibitory effect of these compounds on the *Plasmodium falciparum* enzyme is yet to be investigated.

7.2.1 The inhibitory activity of TLM against malaria parasites

Inhibition of the growth of *P. falciparum* parasites by TLM (LD_{50} >95μM against K1 strain (this work) and LD_{50} = 49μM (Waller *et al.*, 2003) against the W2Mef strain) is not as high as the observed inhibition against pfFabF (IC_{50} = 3μM). This suggests that TLM may have significant difficulty in reaching, or accumulating in, the apicoplast, the subcellular location of the Type II FAS in *P. falciparum*. This may be due to poor uptake of the compound into infected erythrocytes or due to poor transport across the parasitophorous vacuole into the parasites where it must traverse the four membranes that surround the apicoplast.

Alternatively, the compound may be actively exported, in a similar mechanism thought to be responsible for the advent of resistance to chloroquine the mainstay malaria drug for 50 years (Sanchez *et al.*, 2003) and to TLM resistance in *E. coli* (Furukawa *et al.*, 1993), or be metabolised by the parasite.

Like thiolactomycin, triclosan is a hydrophobic compound that is poorly soluble in water, a property that has limited its clinical utility, although it is extensively used in other situations. Triclosan inhibits the enoyl reductase (FabI) enzyme of the Type II FAS system and is capable of clearing malaria parasites from mice (Surolia and Surolia, 2001). Like pfFabF, the FabI enzymes of *P. falciparum* and *T. gondii* are predicted to be targeted to the apicoplast. A recent investigation showed that when triclosan was linked to releasable octomers of arginine, the inhibitor was able to target the FabI enzyme of tachyzoites and encysted bradyzoites, causing a delayed death phenotype (Samuel *et al.*, 2003). It is extremely difficult for any compound to penetrate the encysted bradyzoite stages of *T. gondii* infection due to the large number of boundaries the compound needs to negotiate to reach the apicoplast, and there are no medicines that treat this form of the disease. Therefore, linking TLM to octoarginine groups may increase the ability of the compound to cross membranes in *P. falciparum* and may also increase its bioavailability in *T. gondii* parasites. This method may also improve the likelihood that TLM, or a derivative, may have better absorption and increased bioavailability when administered orally. These pharmacokinetic attributes would provide a significant advantage if a potential antimalarial was to be used in rural areas, where access to health care facilities is limited.

7.2.2 Synthesis of TLM or derivatives as potential antimalarial drugs

Several other questions about the suitability of TLM as a lead compound need to be addressed and investigated in further detail. These include whether the cost of mass production of TLM is commercially viable. Our studies suggest that it may be difficult to synthesise TLM and/or derivatives that also possess a C5 isoprenoid group from simple building blocks. Furthermore, synthesis of the pure (5R)-enantiomer is also a major consideration for the financial and clinical suitability of TLM. Recently, the first report of the synthesis of the (5R)-thiolactomycin enantiomer was reported using D-alanine as a chiral building block (McFadden, *et al.*, 2002). The route of synthesis used to produce this optimally enriched TLM was flexible and allowed for the substitution of groups at the C3 and C5 positions. This suggests that it may be possible to produce solely the (5R)-enantiomer of TLM derivatives. It must be noted that the removal of the double bond of the thiolactone ring, observed in the dicarbonyl TLM derivatives (Table 6.11) introduces a chiral centre at the C4 and C3 positions. This alteration would increase the number of stereoisomers of these derivatives and increase the problem of obtaining an optically pure compound.

An alternative route to the synthesis of TLM may be through the engineering of novel polyketide synthases that are able to produce large amounts of TLM or derivatives. In *Nocardia* sp., the fungi that produce natural TLM, there is evidence that a polyketide synthase (PKS) is involved in the synthesis of TLM (Brown *et al.*, 2003). An investigation of the areas of the *Nocardia* sp. genome that encode PKS modules should elucidate which PKS is responsible for the natural synthesis of TLM. Advances in the manipulation and engineering of novel PKS suggest that isolation of enzymes capable of synthesising large amounts of TLM from acetate precursors is an achievable goal (Khosla *et al.*, 1993; Watanabe *et al.*, 2003a, b). Finally, chemical alteration of naturally synthesised TLM may be

7.2.2 Synthesis of TLM or derivatives as potential antimalarial drugs

Several other questions about the suitability of TLM as a lead compound need to be addressed and investigated in further detail. These include whether the cost of mass production of TLM is commercially viable. Our studies suggest that it may be difficult to synthesise TLM and/or derivatives that also possess a C5 isoprenoid group from simple building blocks. Furthermore, synthesis of the pure (5R)-enantiomer is also a major consideration for the financial and clinical suitability of TLM. Recently, the first report of the synthesis of the (5R)-thiolactomycin enantiomer was reported using D-alanine as a chiral building block (McFadden, *et al.*, 2002). The route of synthesis used to produce this optimally enriched TLM was flexible and allowed for the substitution of groups at the C3 and C5 positions. This suggests that it may be possible to produce solely the (5R)-enantiomer of TLM derivatives. It must be noted that the removal of the double bond of the thiolactone ring, observed in the dicarbonyl TLM derivatives (Table 6.11) introduces a chiral centre at the C4 and C3 positions. This alteration would increase the number of stereoisomers of these derivatives and increase the problem of obtaining an optically pure compound.

An alternative route to the synthesis of TLM may be through the engineering of novel polyketide synthases that are able to produce large amounts of TLM or derivatives. In *Nocardia* sp., the fungi that produce natural TLM, there is evidence that a polyketide synthase (PKS) is involved in the synthesis of TLM (Brown *et al.*, 2003). An investigation of the areas of the *Nocardia* sp. genome that encode PKS modules should elucidate which PKS is responsible for the natural synthesis of TLM. Advances in the manipulation and engineering of novel PKS suggest that isolation of enzymes capable of synthesising large amounts of TLM from acetate precursors is an achievable goal (Khosla *et al.*, 1993; Watanabe *et al.*, 2003a, b). Finally, chemical alteration of naturally synthesised TLM may be

one method that could produce TLM derivatives containing C5 isoprenoid groups, overcoming the problems in chemically synthesising these derivatives. This technique is probably not a cost-effective method of producing TLM derivatives on a large scale, due to the low yield of TLM from cultures of *Nocardia* and the subsequent derivitisation. However, because of the presence of the natural C5 isoprenoid group, this route may provide many potent inhibitors of bacterial and Apicomplexan FabB and/or FabF proteins. The inhibitory effect of these compounds could then be determined before large scale production of the most promising was investigated through an alternative synthetic route.

7.2.3 Alternative targets of TLM inhibition in the human host

There is increasing evidence of a Type II FAS pathway in the mitochondria of mammals, plants and yeast (Wada *et al.*, 1997; Brody *et al.*, 1997; Schneider 1997a and 1997b; Torkko *et al.*, 2001;Yasuno and Wada, 2002; Miinalainen *et al.*, 2003; Zhang *et al.*, 2003; Yasuno *et al.*, 2004). In humans, the ACP and malonyl transferase components of this mitochondrial system have been identified. Pure recombinant proteins showed that the acyl transferase was capable of catalysing the efficient transfer of a malonyl group from coenzyme A to the mitochondrial ACP (Zhang *et al.*, 2003), suggesting that the component proteins are functional and interact with each other. To date, the only investigation of a mitochondrial condensing enzyme was performed on the *Arabadopsis thalania* protein (Yasuno *et al.*, 2004). This enzyme appeared to produce a bimodal distribution of acyl-ACP products with peaks at C8 and C14-C16 chain lengths, suggesting that this enzyme may be important in the synthesis of lipoic acid. The human genome does contain a gene that is predicted to encode a Type II FAS-like condensing enzyme (Figure 3.1) and this protein also contains a putative N-terminal mitochondrial target sequence (personal observation). It is possible that the

human mitochondrial enzymes may also be inhibited by TLM and, if this is the case, the level and effects of TLM inhibition in mammals will need to be fully investigated before TLM could be considered for clinical use.

There is growing evidence that Type I FAS is overexpressed in mammalian cancer cells that cause tumours in breast, prostate, stomach and colon tissue (Kuhajda *et al.*, 1994 and 2000; Li *et al.*, 2001; Thupari *et al.*, 2001; Swinnen *et al.*, 2002; Liu *et al.*, 2002; Kusakabe *et al.*, 2002; Visca *et al.*, 2003; Baron *et al.*, 2004). Studies have shown that overexpression of Type I FAS in prostate tumours is an early and common event in the development of cancerous cells and that tumours with higher levels of Type I FAS expression had a higher proliferative index (Swinnen *et al.*, 2002; Welsh *et al.*, 2001; Baron *et al.*, 2004). Cerulenin, an inhibitor of the β-ketoacyl-synthase activity of both Type I FAS and Type II FAS has been shown to inhibit the growth of colon carcinomas in mice *in vivo* (Huang *et al.*, 2000a and 2000b). Furthermore, triclosan, once thought to be a specific inhibitor of the enoyl-ACP reductase of Type II FAS, inhibits the enoyl reductase activity of Type I FAS with an IC_{50} of 10-20μM, compared to IC_{50} = 2μM for the *E. coli* FabI (Heath *et al.*, 1998) and IC_{50} = 66nM against *P. falciparum* FabI (Kapoor *et al.*, 2004a). Triclosan was also shown to have a cytotoxic effect against two breast cancer cell lines (LD_{50} = 17μM) and no effect against normal cells (Liu *et al.*, 2002).

7.2.4 Elucidating the target of TLM derivatives

It is clear that none of the TLM derivatives synthesised in this project had a high inhibitory effect on the pfFabF enzyme (Figure 6.8). However, some of these derivatives exhibited

significantly better inhibition of the pfFabH enzyme and this demonstrated the scope that exists for the design of novel TLM derivatives that target the fatty acid biosynthesis pathway. The production of compounds that inhibit the pfFabH enzyme should be actively pursued, in tandem with the inhibitors of the pfFabF enzyme, as pfFabH catalyses the initial condensation step of the Type II FAS pathway and is ideally suited to regulate the rate of fatty acid synthesis (Heath and Rock, 1996a and 1996b; Choi *et al.*, 2000a and 2000b).

Clearly, in whole parasites the target of some of the derivatives that did not inhibit either of the *P. falciparum* KAS enzymes, for example 1152 (Figure 6.12), remains to be elucidated. One way to identify the target of this compound may be to compare the expression profile of *P. falciparum* parasites in controls and cells grown in the presence of the inhibitor, using microarray technology. This approach could identify one or many targets for these inhibitors with unknown function. However, even if the identification of a single target and validating the mechanism of action was achieved, this process could take several years to complete. It was deemed too labour-intensive to attempt to identify the target of these TLM derivatives in the present study.

7.2.5 Activity of TLM and its derivatives against other parasites

The inhibitory effect of TLM derivatives against cultures of different protozoa parasites is shown in Appendix 8. Many of the TLM derivatives showed enhanced activity against these *in situ* cultures compared to racemic TLM. The best compounds against *Trypanosoma brucei*, *Trypanosoma cruzii* and *Leishmania donovani* showed 90-fold, 14-fold and 120-fold improvements in LD_{50} values compared to TLM. Against *L. donovani* compound 1199 exhibited the first and only submicromolar LD_{50} (500nM) observed in this project.

Identification of the *in vivo* target/s of this TLM derivative should be investigated further. Bioinformatic searches have identified putative genes that encode enzymes of the Type II FAS pathway in *Trypanosoma brucei*, *Trypanosoma cruzii* and *Leishmania donovani* including genes for the condensing enzymes FabH and FabB/F and these may be targets of inhibition by TLM and/or its derivatives. Also, there is biochemical evidence that suggests Trypanosomes can synthesise their own fatty acids (Morita *et al.*, 2000; Paul *et al.*, 2001 and 2004). The increased inhibitory activity of TLM derivatives against these parasitic protozoa suggests that there is significant scope to produce derivatives with enhanced activity against a range of organisms, not just *Plasmodium falciparum*. Studies of the inhibitory effect of these compounds on the FAS synthesis pathway or individual enzymes of these parasites may allow the identification of lead compounds in the treatment of many devastating parasitic diseases.

At the outset of this project six aims (outlined on Page 74) were used as a guide for the focus of this project. The first four of these aims relating to the identification, expression and characterisation of pfFabF and the testing of both the *P. falciparum* condensing enzymes against TLM derivatives were achieved partly or in full and are described in Chapters 3-6. An excellent kinetic characterisation of the pfFabH enzyme was achieved by a separate research group (Prigge *et al.*, 2003) prior to the production of soluble enzyme during this project. It was deemed unnecessary to repeat this work. A more detailed characterisation of the pfFabF kinetic parameters is still required and future work is planned. Unfortunately, structural studies of these enzymes were not possible due to the inability to cleave the N-terminally fused maltose binding protein. Whilst, the ultimate aim of finding a new lead antimalarial compound was not realised, these studies contribute significantly to our knowledge of the targets and mechanism of action of thiolactomycin against *Plasmodium falciparum*.

Bibliography

Abbadi, A., M. Brummel, B. S. Schutt, M. B. Slabaugh, R. Schuch and S. F. (2000). Reaction mechanism of recombinant 3-oxoacyl-(acyl-carrier-protein) synthase III from Cuphea wrightii embryo, a fatty acid synthase type II condensing enzyme. Biochem J 345: 153-60.

Abbadi, A., M. Brummel and F. Spener (2000). Knockout of the regulatory site of 3-ketoacyl-ACP synthase III enhances short- and medium-chain acyl-ACP synthesis. Plant J 24: 1-9.

Alberts, A. W., R. M. Bell and P. R. Vagelos (1972). Acyl carrier protein. XV. Studies of -ketoacyl-acyl carrier protein synthetase. J Biol Chem 247: 3190-8.

Allen, E. E. and D. H. Bartlett (2000). FabF is required for piezoregulation of cis-vaccenic acid levels and piezophillic growth of the deep-sea bacterium Photobacterium profundum Strain SS9. J Bacteriol 182: 1264-71.

Altabe, S. G., P. Aguilar, G. M. Caballero and D. de Mendoza (2003). The Bacillus subtilis acyl lipid desaturase is a delta5 desaturase. 185: 3228-31.

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

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

Augustijns, P., A. D'Hulst, J. Van Daele and R. Klinget (1996). Transport of artemisinin and sodium artesunate in Caco-2 intestinal epithelial cells. J Pharm Sci 85: 577-9.

Baca, A. M. and W. G. Hol (2000). Overcoming codon bias: a method for high-level overexpression of Plasmodium and other AT-rich parasite genes in *Escherichia coli*. Int J Parasitol 30: 113-8.

Bao, X., M. Focke, M. Pollard and J. Ohlrogge (2000). Understanding in vivo carbon precursor supply for fatty acid synthesis in leaf tissue. Plant J 22: 39-50.

Barnwell, J. W., A. S. Asch, R. L. Nachman, M. Yamaya, M. Aikawa and P. Ingravallo (1989). A human 88-kD membrane glycoprotein (CD36) functions in vitro as a receptor for a cytoadherence ligand on *Plasmodium falciparum*-infected erythrocytes. J Clin Invest 84: 765-72.

Baron, A., T. Migita, D. Tang and M. Loda (2004). Fatty acid synthase: a metabolic oncogene in prostate cancer? J Cell Biochem 91: 47-53.

Baruch, D. I., B. L. Pasloske, H. B. Singh, X. Bi, X. C. Ma, M. Feldman, T. F. Taraschi and R. J. Howard (1995). Cloning the *P. falciparum* Gene Encoding PfEMP1, a Malarial Variant Antigen and Adherence Receptor On the Surface of Parasitized Human Erythrocytes. Cell 82: 77.

Beeson, J. G., P. A. Winstanley, G. I. McFadden and G. V. Brown (2001). New agents to combat malaria. Nat Med 7: 149-50.

Beier, J. C., F. K. Onyango, J. K. Koros, M. Ramadhan, R. Ogwang, R. A. Wirtz, D. K. Koech and C. R. Roberts (1991). Quantitation of malaria sporozoites transmitted in vitro during salivation by wild Afrotropical Anopheles. Med Vet Entomol 5: 71-9.

Bem, J. L., L. Kerr and D. Stuerchler (1992). Mefloquine prophylaxis: an overview of spontaneous reports of severe psychiatric reactions and convulsions. J Trop Med Hyg 95: 167-79.

Berendt, A. R., D. L. Simmons, J. Tansey, C. I. Newbold and K. Marsh (1989). Intercellular adhesion molecule-1 is an endothelial cell adhesion receptor for *Plasmodium falciparum*. Nature 341: 57-9.

Bhargava, H. N. and P. A. Leonard (1996). Triclosan: applications and safety. Am J Infect Control 24: 209-18.

Birnboim, H. C. and J. Doly (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7: 1513-23.

Bonday, Z. Q., S. Dhanasekaran, P. N. Rangarajan and G. Padmanaban (2000). Import of host aminolevunate dehydratase into the malarial parasite. Nature Med 6: 893-903.

Brody, S., C. Oh, U. Hoja and E. Schweizer (1997). Mitochondrial acyl carrier protein is involved in lipoic acid synthesis in Saccharomyces cerevisiae. FEBS Lett 408: 217-20.

Brown, M. S., K. Akopiants, D. M. Resceck, H. A. McArthur, E. McCormick and K. A. Reynolds (2003). Biosynthetic origins of the natural product, thiolactomycin: a unique and selective inhibitor of type II dissociated fatty acid synthases. J Am Chem Soc 125: 10166-7.

Buchanan, S. K. (2001). Type I secretion and multidrug efflux: transport through the TolC channel-tunnel. Trends Biochem Sci 26: 3-6.

Cai, X., A. L. Fuller, L. R. McDougald and G. Zhu (2003). Apicoplast genome of the coccidian Eimeria tenella. Gene 321: 39-46.

Campbell, J. W. and J. E. Cronan, Jr. (2001). Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. Annu Rev Microbiol 55: 305-32.

Campbell, J. W. and J. E. Cronan, Jr. (2001). *Escherichia coli* FadR positively regulates transcription of the fabB fatty acid biosynthetic gene. J Bacteriol 183: 5982-90.

Chang, Y. Y. and J. E. Cronan (1999). Membrane cyclopropane fatty acid content is a major factor in acid resistance of *Escherichia coli*. Mol Microbiol 33: 249-59.

Chang, Y. Y., J. Eichel and J. E. Cronan, Jr. (2000). Metabolic instability of *Escherichia coli* cyclopropane fatty acid synthase is due to RpoH-dependent proteolysis. J Bacteriol 182: 4288-94.

Choi, K. H., R. J. Heath and C. O. Rock (2000). beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. J Bacteriol 182: 365-70.

Choi, K. H., L. Kremer, G. S. Besra and C. O. Rock (2000). Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from Mycobacterium tuberculosis. J Biol Chem 275: 28201-7.

Clough, R. C., A. L. Matthis, S. R. Barnum and J. G. Jaworski (1992). Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl- coenzyme A to initiate fatty acid synthesis. J Biol Chem 267: 20992-8.

Cranston, H. A., C. W. Boylan, G. L. Carroll, S. P. Sutera, J. R. Williamson, I. Y. Gluzman and D. J. Krogstad (1984). *Plasmodium falciparum* maturation abolishes physiologic red cell deformability. Science 223: 400-3.

Cybulski, L. E., D. Albanesi, M. C. Mansilla, S. Altabe, P. S. Aguilar and D. de Mendoza (2002). Mechanism of membrane fluidity optimization: isothermal control of the Bacillus subtilis acyl-lipid desaturase. Mol Microbiol 45: 1379-88.

Davies, C., R. J. Heath, S. W. White and C. O. Rock (2000). The 1.8 A crystal structure and active-site architecture of beta- ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. Structure Fold Des 8: 185-95.

Despommier, D. and J. Karapelou (1987). Parasite Life Cycles. Springer Verlag, New York.

Dewar, J. and K. M. Dieter (1988). Mechansim of the chain elongatio step in the biosynthesis of fatty acids. Biochemistry 27: 3302-8.

Doolan, D. L. and S. L. Hoffman (2001). DNA-based vaccines against malaria: status and promise of the Multi- Stage Malaria DNA Vaccine Operation. Int J Parasitol 31: 753-62.

Dormann, P., F. Spener and J. Ohlrogge (1993). Characterisation of two acyl-acyl carrier protein thioesterases from developing Cuphea seeds specific for medium-chain and oleoyl-acyl carrier protein. Planta 189: 425-432.

Douglas, J. D., S. J. Senior, C. Morehouse, B. Phetsukiri, I. B. Campbell, G. S. Besra and D. E. Minnikin (2002). Analogues of thiolactomycin: potential drugs with enhanced anti-mycobacterial activity. Microbiology 148: 3101-9.

Edwards, P., J. S. Nelsen, J. G. Metz and K. Dehesh (1997). Cloning of the fabF gene in an expression vector and in vitro characterization of recombinant fabF and fabB encoded enzymes from *Escherichia coli*. FEBS Lett 402: 62-6.

Ellis, K. E., B. Clough, J. W. Saldanha and R. J. Wilson (2001). Nifs and Sufs in malaria. Mol Microbiol 41: 973-81.

Feagin, J. E. (1992). The 6-kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. Mol Biochem Parasitol 52: 145-8.

Fichera, M. E., M. K. Bhopale and D. S. Roos (1995). In vitro assays elucidate peculiar kinetics of clindamycin action against Toxoplasma gondii. Antimicrob Agents Chemother 39: 1530-7.

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

Foth, B. J., S. A. Ralph, C. J. Tonkin, N. S. Struck, M. Fraunholz, D. S. Roos, A. F. Cowman and G. I. McFadden (2003). Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 299: 705-8.

Fried, M. and P. E. Duffy (1996). Adherence of *Plasmodium falciparum* to chondroitin sulfate A in the human placenta. Science 272: 1502-4.

Funabashi, H., A. Kawaguchi, H. Tomoda, S. Omura, S. Okuda and S. Iwasaki (1989). Binding site of cerulenin in fatty acid synthetase. J Biochem (Tokyo) 105: 751-5.

Funes, S., E. Davidson, A. Reyes-Prieto, S. Magallon, P. Herion, M. P. King and D. Gonzalez-Halphen (2002). A green algal apicoplast ancestor. Science 298: 2155.

Funes, S., E. Davidson, A. Reyes-Prieto, S. Magallon, P. Herion and M. P. King (2003). Response to a comment on "A green algal apicoplast ancestor". Science 301: 49b.

Funes, S., A. Reyes-Prieto, X. Perez-Martinez and D. Gonzalez-Halphen (2004). On the evolutionary origins of apicoplasts: revisiting the rhodophyte vs. chlorophyte controversy. Microbes Infect 6: 305-11.

Furukawa, H., J. T. Tsay, S. Jackowski, Y. Takamura and C. O. Rock (1993). Thiolactomycin resistance in *Escherichia coli* is associated with the multidrug resistance efflux pump encoded by emrAB. J Bacteriol 175: 3723-9.

Gardner, M. J., H. Tettelin, D. J. Carucci, L. M. Cummings, L. Aravind, E. V. Koonin, S. Shallom, T. Mason, K. Yu, C. Fujii, J. Pederson, K. Shen, J. Jing, C. Aston, Z. Lai, D. C. Schwartz, M. Pertea, S. Salzberg, L. Zhou, G. G. Sutton, R. Clayton, O. White, H. O. Smith, C. M. Fraser, S. L. Hoffman and et al. (1998). Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. Science 282: 1126-32.

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

Garwin, J. L., A. L. Klages and J. E. Cronan, Jr. (1980). Beta-ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. J Biol Chem 255: 3263-5.

Gerold, P., L. Schofield, M. J. Blackman, A. A. Holder and R. T. Schwarz (1996). Structural analysis of the glycosyl-phosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of PLasmoidum falciparum. Mol Biochem Parasitol 75: 131-143.

Gibson, K. J. (1993). Palmitoleate formation by soybean stearoyl-acyl carrier protein desaturase. Biochim Biophys Acta 1169: 231-5.

Gleeson, M. T. (2000). The plastid in Apicomplexa: what use it it ? Int J Parasitol 20: 1053-1070.

Goodstadt, L. and C. P. Ponting (2001). CHROMA: consensus-based colouring of multiple alignments for publication. Bioinformatics 17: 845-6.

Greenwood, B. and T. Mutabingwa (2002). Malaria in 2002. Nature 415: 670-2.

Grogan, D. W. and J. E. Cronan, Jr. (1997). Cyclopropane ring formation in membrane lipids of bacteria. Microbiol Mol Biol Rev 61: 429-41.

Guex, N. and M. C. Peitsch (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714-23.

Gurr, M. I., J. L. Harwood and K. N. Frayn (2002). Lipid Biochemistry: An Introduction. Oxford, UK, Blakwell Publishing Ltd.

Hanahan, D. (1983). Studies on the transformation of *Escherichia coli* with plasmids. J Mol Biol 166: 557-80.

Harwood, J. L. (1996). Recent advances in the biosynthesis of plant fatty acids. Biochim Biophys Acta 1301: 7-56.

Hayashi, T., O. Yamamoto, H. Sasaki, A. Kawaguchi and H. Okazaki (1983). Mechanism of action of the antibiotic thiolactomycin inhibition of fatty acid synthesis of *Escherichia coli*. Biochem Biophys Res Commun 115: 1108-13.

Hayashi, T., O. Yamamoto, H. Sasaki, H. Okazaki and A. Kawaguchi (1984). Inhibition of fatty acid synthesis by the antiobiotic thiolactomycin. J Antibiot (Tokyo) 37: 1456-61.

Heath, R. J. and C. O. Rock (1995). Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*. J Biol Chem 270: 26538-42.

Heath, R. J. and C. O. Rock (1996). Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in *Escherichia coli* fatty acid biosynthesis. J Biol Chem 271: 27795-801.

Heath, R. J. and C. O. Rock (1996). Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. J Biol Chem 271: 10996-11000.

Heath, R. J. and C. O. Rock (1996). Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. J Biol Chem 271: 1833-6.

Heath, R. J., Y. T. Yu, M. A. Shapiro, E. Olson and C. O. Rock (1998). Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis. J Biol Chem 273: 30316-20.

Heath, R. J., J. R. Rubin, D. R. Holland, E. Zhang, M. E. Snow and C. O. Rock (1999). Mechanism of triclosan inhibition of bacterial fatty acid synthesis. J Biol Chem 274: 11110-4.

Heath, R. J. and C. O. Rock (2000). A triclosan-resistant bacterial enzyme. Nature 406: 145-6.

Hemingway, J. (1999). Insecticide resistance in malaria vectors: a new approach to an old subject. Parassitologia 41: 315-8.

Hill, J., L. Tyas, L. H. Phylip, J. Kay, B. M. Dunn and C. Berry (1994). High level expression and characterisation of Plasmepsin II, an aspartic proteinase from *Plasmodium falciparum*. FEBS Lett 352: 155-8.

Hoang, T. T. and H. P. Schweizer (1997). Fatty acid biosynthesis in Pseudomonas aeruginosa: cloning and characterization of the fabAB operon encoding beta-hydroxyacyl-acyl carrier protein dehydratase (FabA) and beta-ketoacyl-acyl carrier protein synthase I (FabB). J Bacteriol 179: 5326-32.

Hoebe, C., J. de Munter and C. Thijs (1997). Adverse effects and compliance with mefloquine or proguanil antimalarial chemoprophylaxis. Eur J Clin Pharmacol 52: 269-75.

Huang, W., J. Jia, P. Edwards, K. Dehesh, G. Schneider and Y. Lindqvist (1998). Crystal structure of beta-ketoacyl-acyl carrier protein synthase II from E.coli reveals the molecular architecture of condensing enzymes. Embo J 17: 1183-91.

Huang, P., S. Zhu, S. Lu, Z. Dai and Y. Jin (2000). An experimental study on cerulenin induced apoptosis of human colonic cancer cells. Zhonghua Bing Li Xue Za Zhi 29: 115-8.

Huang, P., S. Zhu, S. Lu, L. Li, Z. Dai and Y. Jin (2000). Cerulenin inhibits growth of human colonic carcinoma in nude mice. Zhonghua Bing Li Xue Za Zhi 29: 435-8.

Ito, J., A. Ghosh, L. A. Moreira, E. A. Wimmer and M. Jacobs-Lorena (2002). Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature 417: 452-5.

Jackowski, S. and C. O. Rock (1987). Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. J Biol Chem 262: 7927-31.

Jackowski, S., C. Murphy, J. E. Cronan and C. O. Rock (1989). Acetoacetyl-acyl carrier protein synthase. J Biol Chem 264: 7624-9.

Jackowski, S., Y. M. Zhang, A. C. Price, S. W. White and C. O. Rock (2002). A missense mutation in the fabB (beta-ketoacyl-acyl carrier protein synthase I) gene confers tiolactomycin resistance to *Escherichia coli*. Antimicrob Agents Chemother 46: 1246-52.

Jaworski, J. G., R. C. Clough and S. R. Barnum (1989). A Cerulenin Insensitive Short Chain 3-Ketoacyl-Acyl Carrier Protein Synthase in Spincia. oleracea Leaves. Plant Physiology 90.

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

Jones, A. L., J. E. Dancer and J. L. Harwood (1994). The effect of thiolactomycin analogues on fatty acid synthesis in peas (Pisum sativum cv. Onward). Biochem Soc Trans 22: 258S.

Jones, A. L., J. E. Dancer and J. L. Harwood (1995). Effect of thiolactomycin on fatty acid synthesis in peas. Phytochemistry 39: 511-514.

Jones, A. L., D. Herbert, A. J. Rutter, J. E. Dancer and J. L. Harwood (2000). Novel inhibitors of the condensing enzymes of the type II fatty acid synthase of pea (Pisum sativum). Biochem J 347 Pt 1: 205-9.

Jones, A. L., A. M. Gane, D. Herbert, D. L. Willey, A. J. Rutter, P. Kille, J. E. Dancer and J. L. Harwood (2003). beta-Ketoacyl-acyl carrier protein synthase III from pea (Pisum sativum L.): properties, inhibition by a novel thiolactomycin analogue and isolation of a cDNA clone encoding the enzyme. Planta 216: 752-61.

Jones, S. M., J. E. Urch, R. Brun, J. L. Harwood, C. Berry and I. H. Gilbert (2004). Analogues of thiolactomycin as potential anti-malarial and anti-trypanosomal agents. Bioorg Med Chem 12: 683-92.

Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6: 494-500.

Kang, Y., T. Durfee, J. D. Glasner, Y. Qiu, D. Frisch, K. M. Winterberg and F. R. Blattner (2004). Systematic mutagenesis of the *Escherichia coli* genome. J Bacteriol 186: 4921-30.

Kapoor, M., M. J. Dar, A. Surolia and N. Surolia (2001). Kinetic determinants of the interaction of enoyl-ACP reductase from *Plasmodium falciparum* with its substrates and inhibitors. Biochem Biophys Res Commun 289: 832-7.

Kapoor, M., J. Gopalakrishnapai, N. Surolia and A. Surolia (2004). Mutational analysis of the triclosan-binding region of enoyl-ACP (acyl-carrier protein) reductase from *Plasmodium falciparum*. Biochem J 381: 735-41.

Kapoor, M., P. L. Mukhi, N. Surolia, K. Suguna and A. Surolia (2004). Kinetic and structural analysis of the increased affinity of enoyl-ACP (acyl-carrier protein) reductase for triclosan in the presence of NAD+. Biochem J 381: 725-33.

Kapoor, M., C. C. Reddy, M. V. Krishnasastry, N. Surolia and A. Surolia (2004). Slow-tight-binding inhibition of enoyl-acyl carrier protein reductase from *Plasmodium falciparum* by triclosan. Biochem J 381: 719-24.

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

Karunaweera, N. D., G. E. Grau, P. Gamage, R. Carter and K. N. Mendis (1992). Dynamics of fever and serum levels of tumor necrosis factor are closely associated during clinical paroxysms in Plasmodium vivax malaria. Proc Natl Acad Sci U S A 89: 3200-3.

Ke, J., T. N. Wen, B. J. Nikolau and E. S. Wurtele (2000). Coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. Plant Physiol. 122: 1057-71.

Khandekar, S. S., A. K. Konstantinidis, C. Silverman, C. A. Janson, D. E. McNulty, S. Nwagwu, G. S. Van Aller, M. L. Doyle, J. F. Kane, X. Qiu and J. Lonsdale (2000). Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. Biochem Biophys Res Commun 270: 100-7.

Khandekar, S. S., D. R. Gentry, G. S. Van Aller, P. Warren, H. Xiang, C. Silverman, M. L. Doyle, P. A. Chambers, A. K. Konstantinidis, M. Brandt, R. A. Daines and J. T. Lonsdale (2001). Identification, substrate specificity, and inhibition of the Streptococcus pneumoniae beta-ketoacyl-acyl carrier protein synthase III (FabH). J Biol Chem 276: 30024-30.

Khosla, C., R. McDaniel, S. Ebert-Khosla, R. Torres, D. H. Sherman, M. J. Bibb and D. A. Hopwood (1993). Genetic construction and funtional analysis of hybrid polyketide synthases ontaining heterologous acyl carrier proteins. J Bacteriol 175: 2197-2204.

Kilejian, A. (1975). Circular mitochondrial DNA from the avian malarial parasite Plasmodium lophurae. Biochim Biophys Acta 390: 276-284.

Knell, A. J. (1991). Malaria - A publication of the tropical programme of the Wellcome Trust. Oxford, UK, Oxford University Press.

Kohler, S., C. F. Delwiche, P. W. Denny, L. G. Tilney, P. Webster, R. J. Wilson, J. D. Palmer and D. S. Roos (1997). A plastid of probable green algal origin in Apicomplexan parasites. Science 275: 1485-9.

Kremer, L., J. D. Douglas, A. R. Baulard, C. Morehouse, M. R. Guy, D. Alland, L. G. Dover, J. H. Lakey, W. R. Jacobs, Jr., P. J. Brennan, D. E. Minnikin and G. S. Besra (2000). Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in Mycobacterium tuberculosis. J Biol Chem 275: 16857-64.

Kremer, L., K. M. Nampoothiri, S. Lesjean, L. G. Dover, S. Graham, J. Betts, P. J. Brennan, D. E. Minnikin, C. Locht and G. S. Besra (2001). Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of Mycobacterium tuberculosis fatty acid synthase II. J Biol Chem 276: 27967-74.

Kremer, L., L. G. Dover, S. Carrere, K. M. Nampoothiri, S. Lesjean, A. K. Brown, P. J. Brennan, D. E. Minnikin, C. Locht and G. S. Besra (2002). Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from Mycobacterium tuberculosis. Biochem J 364: 423-30.

Krugliak, M., E. Deharo, G. Shalmiev, M. Sauvain, C. Moretti and H. Ginsburg (1995). Antimalarial effects of C18 fatty acids on *Plasmodium falciparum* in culture and on Plasmodium vinckei petteri and Plasmodium yoelii nigeriensis in vivo. Exp Parasitol 81: 97-105.

Kuhajda, F. P., K. Jenner, F. D. Wood, R. A. Hennigar, L. B. Jacobs, J. D. Dick and G. R. Pasternack (1994). Fatty acid synthesis: a potential selective target for antineoplastic therapy. Proc Natl Acad Sci U S A 91: 6379-83.

Kuhajda, F. P., E. S. Pizer, J. N. Li, N. S. Mani, G. L. Frehywot and C. A. Townsend (2000). Synthesis and antitumor activity of an inhibitor of fatty acid synthase. Proc Natl Acad Sci U S A 97: 3450-4.

Kurland, C. and J. Gallant (1996). Errors of heterologous protein expression. Curr Opin Biotechnol 7: 489-93.

Kusakabe, T., A. Nashimoto, K. Honma and T. Suzuki (2002). Fatty acid synthase is highly expressed in carcinoma, adenoma and in regenerative epithelium and intestinal metaplasia of the stomach. Histopathology 40: 71-9.

Kutchma, A. J., T. T. Hoang and H. P. Schweizer (1999). Characterization of a Pseudomonas aeruginosa fatty acid biosynthetic gene cluster: purification of acyl carrier protein (ACP) and malonyl-coenzyme A:ACP transacylase (FabD). J Bacteriol 181: 5498-504.

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

Lai, C. Y. and J. E. Cronan (2003). Beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. J Biol Chem 278: 51494-503.

Lakshmi Swarna Mukhi, P., S. Kumar Sharma, M. Kapoor, N. Surolia, A. Surolia and K. Suguna (2004). Crystallization and preliminary crystallographic analysis of beta-hydroxyacyl ACP dehydratase (FabZ) from *Plasmodium falciparum*. Acta Crystallogr D Biol Crystallogr 60: 120-1.

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

Lell, B., R. Ruangweerayut, J. Wiesner, M. A. Missinou, A. Schindler, T. Baranek, M. Hintz, D. Hutchinson, H. Jomaa and P. G. Kremsner (2003). Fosmidomycin, a novel chemotherapeutic agent for malaria. Antimicrob Agents Chemother 47: 735-8.

Li, J. N., M. Gorospe, F. J. Chrest, T. S. Kumaravel, M. K. Evans, W. F. Han and E. S. Pizer (2001). Pharmacological inhibition of fatty acid synthase activity produces both cytostatic and cytotoxic effects modulated by p53. Cancer Res 61: 1493-9.

Liu, B., Y. Wang, K. L. Fillgrove and V. E. Anderson (2002). Triclosan inhibits enoyl-reductase of type I fatty acid synthase in vitro and is cytotoxic to MCF-7 and SKBr-3 breast cancer cells. Cancer Chemother Pharmacol 49: 187-93.

Loftus, T. M., D. E. Jaworsky, G. L. Frehywot, C. A. Townsend, G. V. Ronnett, M. D. Lane and F. P. Kuhajda (2000). Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. Science 288: 2379-81.

Lomovskaya, O., K. Lewis and A. Matin (1995). EmrR is a negative regulator of the *Escherichia coli* multidrug resistance pump EmrAB. J Bacteriol 177: 2328-34.

Lu, Y. J., Y. M. Zhang and C. O. Rock (2004). Product diversity and regulation of type II fatty acid synthases. Biochem Cell Biol 82: 145-55.

Magnuson, K., S. Jackowski, C. O. Rock and J. E. Cronan, Jr. (1993). Regulation of fatty acid biosynthesis in *Escherichia coli*. Microbiol Rev 57: 522-42.

Magnuson, K., M. R. Carey and J. E. Cronan, Jr. (1995). The putative fabJ gene of *Escherichia coli* fatty acid synthesis is the fabF gene. J Bacteriol 177: 3593-5.

Marechal, E. and M. F. Cesbron-Delauw (2001). The apicoplast: a new member of the plastid family. Trends Plant Sci 6: 200-5.

Marrakchi, H., Y. M. Zhang and C. O. Rock (2001). Mechanistic diversity and regulation of Type II fatty acid synthesis. Biochem Soc Trans 30: 1050-5.

Marrakchi, H., K. H. Choi and C. O. Rock (2002). A new mechanism for anaerobic unsaturated fatty acid formation in Streptococcus pneumoniae. J Biol Chem 277: 44809-16.

Martens, W. J., L. W. Niessen, J. Rotmans, T. H. Jetten and A. J. McMichael (1995). Potential impact of global climate change on malaria risk. Environ Health Perspect 103: 458-64.

McFadden, G. I., M. E. Reith, J. Munholland and N. Lang-Unnasch (1996). Plastid in human parasites. Nature 381: 482.

McFadden, G. I., R. F. Waller, M. E. Reith and N. Lang-Unnasch (1997). Plastids in apicomplexan parasites. Plant Syst Evol 11: 261-87.

McFadden, G. I. and D. S. Roos (1999). Apicomplexan plastids as drug targets. Trends Microbiol 7: 328-33.

McFadden, J. M., G. L. Frehywot and C. A. Townsend (2002). A flexible route to (5R)-thiolactomycin, a naturally occurring inhibitor of fatty acid synthesis. Org Lett 4: 3859-62.

McLeod, R., S. P. Muench, J. B. Rafferty, D. E. Kyle, E. J. Mui, M. J. Kirisits, D. G. Mack, C. W. Roberts, B. U. Samuel, R. E. Lyons, M. Dorris, W. K. Milhous and D. W. Rice (2001). Triclosan inhibits the growth of *Plasmodium falciparum* and Toxoplasma gondii by inhibition of Apicomplexan Fab I. Int J Parasitol 31: 109-13.

McMurry, L. M., M. Oethinger and S. B. Levy (1998). Triclosan targets lipid synthesis. Nature 394: 531-2.

Medigue, C., T. Rouxel, P. Vigier, A. Henaut and A. Danchin (1991). Evidence for horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol 222: 851-6.

Mendis, K. N. and R. Carter (1995). Clinical disease and pathogenesis in malatia. Parasitology Today 11: 1-16.

Miinalainen, I. J., Z. J. Chen, J. M. Torkko, P. L. Pirila, R. T. Sormunen, U. Bergmann, Y. M. Qin and J. K. Hiltunen (2003). Characterization of 2-enoyl thioester reductase from mammals. An ortholog of YBR026p/MRF1'p of the yeast mitochondrial fatty acid synthesis type II. J Biol Chem 278: 20154-61.

Miller, L. H., D. I. Baruch, K. Marsh and O. K. Doumbo (2002). The pathogenic basis of malaria. Nature 415: 673-9.

Miyakawa, S., K. Suzuki, T. Noto, Y. Harada and H. Okazaki (1982). Thiolactomycin, a new antibiotic. IV. Biological properties and chemotherapeutic activity in mice. J Antibiot (Tokyo) 35: 411-9.

Moche, M., G. Schneider, P. Edwards, K. Dehesh and Y. Lindqvist (1999). Structure of the complex between the antibiotic cerulenin and its target, beta-ketoacyl-acyl carrier protein synthase. J Biol Chem 274: 6031-4.

Moche, M., K. Dehesh, P. Edwards and Y. Lindqvist (2001). The crystal structure of beta-ketoacyl-acyl carrier protein synthase II from Synechocystis sp. at 1.54 A resolution and its relationship to other condensing enzymes. J Mol Biol 305: 491-503.

Morita, Y. S., K. S. Paul and P. T. Englund (2000). Specialized fatty acid synthesis in African trypanosomes: myristate for GPI anchors. Science 288: 140-3.

Muench, S. P., J. B. Rafferty, R. McLeod, D. W. Rice and S. T. Prigge (2003). Expression, purification and crystallization of the *Plasmodium falciparum* enoyl reductase. Acta Crystallogr D Biol Crystallogr 59: 1246-8.

Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 51: 263-73.

Naggert, J., A. Witkowski, B. Wessa and S. Smith (1991). Expression in *Escherichia coli*, purification and characterization of two mammalian thioesterases involved in fatty acid synthesis. Biochem J 273: 787-90.

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

Nathwani, D., P. F. Currie, J. G. Douglas, S. T. Green and N. C. Smith (1992). *Plasmodium falciparum* malaria in pregnancy: a review. Br J Obstet Gynaecol 99: 118-21.

Newton, P. and N. White (1999). Malaria: New developments in treatment and prevention. Annu. Rev Med 50: 179-192.

Newton, C. R., T. T. Hien and N. White (2000). Cerebral malaria. J Neurol Neurosurg Psychiatry 69: 433-41.

Nishida, I., A. Kawaguchi and M. Yamada (1986). Effect of thiolactomycin on the individual enzymes of the fatty acid synthase system in *Escherichia coli*. J Biochem (Tokyo) 99: 1447-54.

Nollen, E. A., F. A. Salomons, J. F. Brunsting, J. J. Want, O. C. Sibon and H. H. Kampinga (2001). Dynamic changes in the localization of thermally unfolded nuclear proteins associated with chaperone-dependent protection. Proc Natl Acad Sci U S A 98: 12038-43.

Noto, T., S. Miyakawa, H. Oishi, H. Endo and H. Okazaki (1982). Thiolactomycin, a new antibiotic. III. In vitro antibacterial activity. J Antibiot (Tokyo) 35: 401-10.

Novy, R., D. Drott, K. Yaeger and R. Mierendorf (2001). Overcoming the codon bias of *E. coli* for enhanced expression. Innovations (Novagen) 12: 1-3.

Oishi, H., T. Noto, H. Sasaki, K. Suzuki, T. Hayashi, H. Okazaki, K. Ando and M. Sawada (1982). Thiolactomycin, a new antibiotic. I. Taxonomy of the producing organism, fermentation and biological properties. J Antibiot (Tokyo) 35: 391-5.

Okuno, K., M. Yabuta, K. Kawanishi, K. Ohsuye, T. Ooi and S. Kinoshita (2002). Substrate specificity at the P1' site of *Escherichia coli* OmpT under denaturing conditions. Biosci Biotechnol Biochem 66: 127-34.

Olliaro, P. (2001). Mode of action and mechanisms of resistance for antimalarial drugs. Pharmacol Ther 89: 207-19.

Olsen, J. G., A. Kadziola, P. von Wettstein-Knowles, M. Siggaard-Andersen, Y. Lindquist and S. Larsen (1999). The X-ray crystal structure of beta-ketoacyl [acyl carrier protein] synthase I. FEBS Lett 460: 46-52.

Olsen, J. G., A. Kadziola, P. von Wettstein-Knowles, M. Siggaard-Andersen and S. Larsen (2001). Structures of beta-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate its Catalytic Machinery. Structure 9: 233-243.

Omodeo-Sale, F., A. Motti, N. Basilico, S. Parapini, P. Olliaro and D. Taramelli (2003). Accelerated senescence of human erythrocytes cultured with *Plasmodium falciparum*. Blood 102: 705-11.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357-8.

Pan, W., E. Ravot, R. Tolle, R. Frank, R. Mosbach, I. Turbachova and H. Bujard (1999). Vaccine candidate MSP-1 from *Plasmodium falciparum*: a redesigned 4917 bp polynucleotide enables synthesis and isolation of full-length protein from *Escherichia coli* and mammalian cells. Nucleic Acids Res 27: 1094-103.

Paul, K. S., D. Jiang, Y. S. Morita and P. T. Englund (2001). Fatty acid synthesis in African trypanosomes: a solution to the myristate mystery. Trends Parasitol 17: 381-7.

Perozzo, R., M. Kuo, A. S. Sidhu, J. T. Valiyaveettil, R. Bittman, W. R. Jacobs, Jr., D. A. Fidock and J. C. Sacchettini (2002). Structural elucidation of the specificity of the antibacterial agent triclosan for malarial enoyl acyl carrier protein reductase. J Biol Chem 277: 13106-14.

Pfefferkorn, E. R., R. F. Nothnagel and S. E. Borotz (1992). Parasiticidal effect of clindamycin on Toxoplasma gondii grown in cultured cells and selection of a drug-resistant mutant. Antimicrob Agents Chemother 36: 1091-6.

Pfefferkorn, E. R. and S. E. Borotz (1994). Comparison of mutants of Toxoplasma gondii selected for resistance to azithromycin, spiramycin, or clindamycin. Antimicrob Agents Chemother 38: 31-7.

Phillips, R. S. (1983). Malaria - Studies in Biology 53, Edward Arnold.

Pillai, S., C. Rajagopal, M. Kapoor, G. Kumar, A. Gupta and N. Surolia (2003). Functional characterization of beta-ketoacyl-ACP reductase (FabG) from *Plasmodium falciparum*. Biochem Biophys Res Commun 303: 387-92.

Pradel, G. and U. Frevert (2001). Malaria sporozoites actively enter and pass through rat Kupffer cells prior to hepatocyte invasion. Hepatology 33: 1154-65.

Prapunwattana, P., W. Sirawaraporn, Y. Yuthavong and D. V. Santi (1996). Chemical synthesis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. Mol Biochem Parasitol 83: 93-106.

Price, A. C., K. H. Choi, R. J. Heath, Z. Li, S. W. White and C. O. Rock (2001). Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. J Biol Chem 276: 6551-9.

Prigge, S. T., X. He, L. Gerena, N. C. Waters and K. A. Reynolds (2003). The initiating steps of a type II fatty acid synthase in *Plasmodium falciparum* are catalyzed by pfacp, pfmcat, and pfKASIII. Biochemistry 42: 1160-9.

Qiu, X., C. A. Janson, A. K. Konstantinidis, S. Nwagwu, C. Silverman, W. W. Smith, S. Khandekar, J. Lonsdale and S. S. Abdel-Meguid (1999). Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. J Biol Chem 274: 36465-71.

Qiu, X., C. A. Janson, W. W. Smith, M. Head, J. Lonsdale and A. K. Konstantinidis (2001). Refined structures of beta-ketoacyl-acyl carrier protein synthase III. J Mol Biol 307: 341-56.

Ralph, S. A., G. G. Van Dooren, R. F. Waller, M. J. Crawford, M. J. Fraunholz, B. J. Foth, C. J. Tonkin, D. S. Roos and G. I. McFadden (2004). Tropical infectious diseases: Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nat Rev Microbiol 2: 203-16.

Ramasamy, R. (1998). Molecular basis for evasion of host immunity and pathogenesis in malaria. Biochimica et Biophysica Acta 1406: 10-27.

Reid, A. J. C., C. J. M. Whitty, H. M. Ayles, R. M. Jennings, B. A. Bovill, J. M. Felton, R. H. Behrens, A. D. M. Bryceson and D. C. W. Mabey (1998). Malaria at Christmas: risks of prophylaxis versus risks of malaria. BMJ 317: 1506-1508.

Ridley, R. G. (2002). Medical need, scientific opportunity and the drive for antimalarial drugs. Nature 415: 686-93.

Rock, R. C. (1971). Incorporation of 14 C-labelled fatty acids into lipids of rhesus erythrocytes and Plasmodium knowlesi in vitro. Comp Biochem Physiol B 40: 893-906.

Rock, C. O. and J. J. E. Cronan (1985). Lipid metabolism in prokayotes. Biochemistry of lipids and membranes Benjamin-Cummings publishing co, Menlo Park, Calif.

Rock, C. O. and J. E. Cronan (1985). Lipid metabolism in prokaryotes. Biochemistry of lipids and membranes. D. E. Vance and J. E. Vance. Menol Park, CA, Benjamin/Cummins: 73-115.

Sachs, J. and P. Malaney (2002). The economic and social burden of malaria. Nature 415: 680-5.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular Cloning, A Laboratory Manual.

Samuel, B. U., B. Hearn, D. Mack, P. Wender, J. Rothbard, M. J. Kiristis, E. Mui, S. Wernimont, C. W. Roberts, S. P. Muench, D. W. Rice, S. T. Prigge, A. B. Law and R. McLeod (2003). Delivery of antimirobials into parasites. Proc Natl Acad Sci U S A 100: 14281-14286.

Sanchez, C. P., W. Stein and M. Lanzer (2003). Trans stimulation provides evidence for a drug efflux carrier as the mechanism of chloroquine resistance in *Plasmodium falciparum*. Biochemistry 42: 9383-94.

Sanger, F., S. Nicklen and A. R. Coulson (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74: 5463-7.

Sasaki, H., H. Oishi, T. Hayashi, I. Matasuura, K. Ando and M. Sawada (1982). Thiolactomycin, a new antibiotic. II. Structure elucidation. J Antibiot (Tokyo) 35: 396-400.

Sato, S., I. Tews and R. J. Wilson (2000). Impact of a plastid-bearing endocytobiont on apicomplexan genomes. Int J Parasitol 30: 427-39.

Sato, S. and R. J. Wilson (2002). The genome of the *Plasmodium falciparum* encodes an active aminolevunilic acid dehydratase. Curr Genet 40: 391-8.

Scarsdale, J. N., G. Kazanina, X. He, K. A. Reynolds and H. T. Wright (2001). Crystal Structure of the Mycobacterium tuberculosisbeta -Ketoacyl-Acyl Carrier Protein Synthase III. J Biol Chem 276: 20516-22.

Schaeffer, M. L., G. Agnihotri, C. Volker, H. Kallender, P. J. Brennan and J. T. Lonsdale (2001). Purification and biochemical characterization of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthases KasA and KasB. J Biol Chem 276: 47029-37.

Schefczik, K. and K. Buff (1984). The insecticide DDT decreases membrane potential and cell input resistance of cultured human liver cells. Biochim Biophys Acta 776: 337-9.

Schneider, R., B. Brors, F. Burger, S. Camrath and H. Weiss (1997). Two genes of the putative mitochondrial fatty acid synthase in the genome of Saccharomyces cerevisiae. Curr Genet 32: 384-8.

Schneider, R., B. Brors, M. Massow and H. Weiss (1997). Mitochondrial fatty acid synthesis: a relic of endosymbiontic origin and a specialized means for respiration. FEBS Lett 407: 249-52.

Schujman, G. E., K. H. Choi, S. Altabe, C. O. Rock and D. de Mendoza (2001). Response of Bacillus subtilis to cerulenin and acquisition of resistance. J Bacteriol 183: 3032-40.

Senior, S. J., Illarionov, P. A., Gurcha, S. S., Campbell, I. B., Schaeffer, M. L., Minnikin, D. E. and G. S. Besra (2003). Biphenyl-based analogues of thiolactomycin, active against *Mycobacterium tuberculosis* mtFabH fatty acid condensing enzyme. Bioorg Med Chem Lett.13: 3685-3688.

Senior, S. J., Illarionov, P. A., Gurcha, S. S., Campbell, I. B., Schaeffer, M. L., Minnikin, D. E. and G. S. Besra (2004). Acetylene-based analogues of thiolactomycin, active against *Mycobacterium tuberculosis* mtFabH fatty acid condensing enzyme. Bioorg Med Chem Lett.14; 373-376.

Sharma, S. K., M. Kapoor, T. N. Ramya, S. Kumar, G. Kumar, R. Modak, S. Sharma, N. Surolia and A. Surolia (2003). Identification, characterization, and inhibition of *Plasmodium falciparum* beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ). J Biol Chem 278: 45661-71.

Sholapurkar, S. L., R. C. Mahajan, A. N. Gupta and A. Wangoo (1990). Cellular immunity in pregnant and non-pregnant women with malarial infection. Asia Oceania J Obstet Gynaecol 16: 27-32.

Siddall, M. E. (1992). Hohlzylinders. Parasitology Today 8: 91.

Siggaard-Anderson, M., G. Bangera, J. G. Olsen and P. von Wettstein-Knowles (1998). Defining the functions of highly conserved reisdues in b-ketoacyl-ACP synthases. Advances in Plant Lipid Research Eds Sanchez, J., Cerda-Olmedo, E. and Martinez-Force, E. Iniversidad de sevilla. Secretariado de Publicaciones: 67 -70.

Slayden, R. A., R. E. Lee, J. W. Armour, A. M. Cooper, I. M. Orme, P. J. Brennan and G. S. Besra (1996). Antimycobacterial action of thiolactomycin: an inhibitor of fatty acid and mycolic acid synthesis. Antimicrob Agents Chemother 40: 2813-9.

Smith, J. D., C. E. Chitnis, A. G. Craig, D. J. Roberts, D. E. Hudson-Taylor, D. S. Peterson, R. Pinches, C. I. Newbold and L. H. Miller (1995). Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell 82: 101-10.

Striepen, B., M. J. Crawford, M. K. Shaw, L. G. Tilney, F. Seeber and D. S. Roos (2000). The plastid of toxoplasma gondii is divided by association with centromeres. J Cell Biol 151: 1423-34.

Studier, F. W. and B. A. Moffat (1986). Use of bacteriophage T7 polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189: 113-30.

Su, X. Z., V. M. Heatwole, S. P. Wertheimer, F. Guinet, J. A. Herrfeldt, D. S. Peterson, J. A. Ravetch and T. E. Wellems (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. Cell 82: 89-100.

Suguna, K., A. Surolia and N. Surolia (2001). Structural basis for triclosan and NAD binding to enoyl-ACP reductase of *Plasmodium falciparum*. Biochem Biophys Res Commun 283: 224-8.

Sullivan, M., J. Li, S. Kumar, M. J. Rogers and T. F. McCutchan (2000). Effects of interruption of apicoplast function on malaria infection, development, and transmission. Mol Biochem Parasitol 109: 17-23.

Surolia, N. and A. Surolia (2001). Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. Nat Med 7: 167-73.

Swinnen, J. V., T. Roskams, S. Joniau, H. Van Poppel, R. Oyen, L. Baert, W. Heyns and G. Verhoeven (2002). Overexpression of fatty acid synthase is an early and common event in the development of prostate cancer. Int J Cancer 98: 19-22.

Thompson, J. D., D. G. Higgins and T. J. Gibson (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-80.

Thomsen-Zieger, N., J. Schachtner and F. Seeber (2003). Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. FEBS Lett 547: 80-6.

Thupari, J. N., M. L. Pinn and F. P. Kuhajda (2001). Fatty acid synthase inhibition in human breast cancer cells leads to malonyl-CoA-induced inhibition of fatty acid oxidation and cytotoxicity. Biochem Biophys Res Commun 285: 217-23.

Torkko, J. M., K. T. Koivuranta, I. J. Miinalainen, A. I. Yagi, W. Schmitz, A. J. Kastaniotis, T. T. Airenne, A. Gurvitz and K. J. Hiltunen (2001). Candida tropicalis Etr1p and Saccharomyces cerevisiae Ybr026p (Mrf1'p), 2-enoyl thioester reductases essential for mitochondrial respiratory competence. Mol Cell Biol 21: 6243-53.

Tsay, J. T., W. Oh, T. J. Larson, S. Jackowski and C. O. Rock (1992). Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. J Biol Chem 267: 6807-14.

Tsay, J. T., C. O. Rock and S. Jackowski (1992). Overproduction of beta-ketoacyl-acyl carrier protein synthase I imparts thiolactomycin resistance to *Escherichia coli* K-12. J Bacteriol 174: 508-13.

Udeinya, I. J., J. A. Schmidt, M. Aikawa, L. H. Miller and I. Green (1981). Falciparum malaria-infected erythrocytes specifically bind to cultured human endothelial cells. Science 213: 555-7.

Val, D., G. Banu, K. Seshadri, Y. Lindqvist and K. Dehesh (2000). Re-engineering ketoacyl synthase specificity. Structure Fold Des 8: 565-6.

van Dooren, G. G., V. Su, M. C. D'Ombrain and G. I. McFadden (2002). Processing of an apicoplast leader sequence in *Plasmodium falciparum* and the identification of a putative leader cleavage enzyme. J Biol Chem 277: 23612-9.

Varadharajan, S., S. Dhanasekaran, Z. Q. Bonday, P. N. Rangarajan and G. Padmanaban (2002). Involvement of the aminolaevulinate synthase encoded by the parasite gene in de novo haem synthesis in *Plasmodium falciparum*. Biochem J 367: 321-7.

Vial, H. J. and M. L. Ancelin (1992). Malarial lipids. An overview. Subcell Biochem 18: 259-306.

Vickerman, K. and F. E. G. Cox (1967). The protozoa. London, John Murray.

Visca, P., V. Sebastiani, E. S. Pizer, C. Botti, P. De Carli, S. Filippi, S. Monaco and P. L. Alo (2003). Immunohistochemical expression and prognostic significance of FAS and GLUT1 in bladder carcinoma. Anticancer Res 23: 335-9.

Vleugels, M. P., W. M. Eling, R. Rolland and R. de Graaf (1987). Cortisol and loss of malaria immunity in human pregnancy. Br J Obstet Gynaecol 94: 758-64.

Wada, H., D. Shintani and J. Ohlrogge (1997). Why do mitochondria synthesize fatty acids? Evidence for involvement in lipoic acid production. Proc Natl Acad Sci U S A 94: 1591-6.

Wahlgren, M., V. Fernandez, C. Scholander and J. Carlson (1994). Rosetting. Parasitology Today 10: 73-94.

Wahlgren, M., J. S. Abrams, V. Fernandez, M. T. Bejarano, M. Azuma, M. Torii, M. Aikawa and R. J. Howard (1995). Adhesion of *Plasmodium falciparum*-infected erythrocytes to human cells

and secretion of cytokines (IL-1-beta, IL-1RA, IL-6, IL-8, IL-10, TGF beta, TNF alpha, G-CSF, GM-CSF. Scand J Immunol 42: 626-36.

Waller, R. F., P. J. Keeling, R. G. Donald, B. Striepen, E. Handman, N. Lang-Unnasch, A. F. Cowman, G. S. Besra, D. S. Roos and G. I. McFadden (1998). Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and *Plasmodium falciparum*. Proc Natl Acad Sci U S A 95: 12352-7.

Waller, R. F., M. F. Reed, A. F. Cowman and G. I. McFadden (2000). Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. Embo J 19: 1794-802.

Waller, R. F., S. A. Ralph, M. B. Reed, V. Su, J. D. Douglas, D. E. Minnikin, A. F. Cowman, G. S. Besra and G. I. McFadden (2003). A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. Antimicrob Agents Chemother 47: 297-301.

Waller, R. F., P. J. Keeling, G. G. van Dooren and G. I. McFadden (2003). Comment on "A green algal apicoplast ancestor". Science 301: 49; author reply 49.

Wang, C. L. J. and J. M. Salvino (1984). Total Synthesis of (+/-) thiolactomycin. Tetrahedron letters 25: 5243-6.

Warrell, D. A. and M. E. Molyneux (1990). Severe and complicated malaria. Trans R Soc Trop Med Hyg 84: 1-65.

Watanabe, K., M. A. Rude, C. T. Walsh and C. Khosla (2003). Engineered nbiosynthesis of an ansmyin polyketide preursor in *Escherichia coli*. Proc Natl Acad Sci U S A 100: 9774-8.

Watanabe, K., C. C. C. Wang, C. N. Boddy, D. E. Cane and C. Khosla (2003). Understanding substrate specificity of polyketide modules by generating hybrid multimodular domains. J Biol Chem 278: 42020-42026.

Waters, N. C., K. M. Kopydlowski, T. Guszczynski, L. Wei, P. Sellers, J. T. Ferlan, P. J. Lee, Z. Li, C. L. Woodard, S. Shallom, M. J. Gardner and S. T. Prigge (2002). Functional characterization of the acyl carrier protein (PfACP) and beta-ketoacyl ACP synthase III (PfKASIII) from *Plasmodium falciparum*. Mol Biochem Parasitol 123: 85-94.

Welsh, J. B., L. M. Sapinoso, A. I. Su, S. G. Kern, J. Wang-Rodriguez, C. A. Moskaluk, H. F. Frierson, Jr. and G. M. Hampton (2001). Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res 61: 5974-8.

White, N. J. (1992). Antimalarial drug resistance: the pace quickens. J Antimicrob Chemother 30: 571-85.

White, N. and P. Olliaro (1996). Strategies for the prevention of antimalrial drug resistance: Rationale for combination chemotherapy for malaria. Parasitology Today 12: 399-401.

WHO_Report (2002). Communicable Disease Report - Roll Back Malaria, http://www.who.int/infectious-disease-news/cds2002/chapter7.pdf. 2004.

Wilkins, M. R., I. Lindskog, E. Gasteiger, A. Bairoch, J. C. Sanchez, D. F. Hochstrasser and R. D. Appel (1997). Detailed peptide characterization using PEPTIDEMASS--a World-Wide-Web-accessible tool. Electrophoresis 18: 403-8.

Williams, R. B. (1999). A compartmentalised model for the estimation of the cost of coccidiosis to the world's chicken production industry. Int J Parasitol 29: 1209-29.

Wilson, R. J., P. W. Denny, P. R. Preiser, K. Rangachari, K. Roberts, A. Roy, A. Whyte, M. Strath, D. J. Moore, P. W. Moore and D. H. Williamson (1996). Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium falciparum. J Mol Biol 261: 155-72.

Winstanley, P. (1996). Mefloquine: the benefits outweight the risks. British J. of Pharmacology 42: 411-413.

Wrenger, C. and S. Muller (2004). The human malaria parasite *Plasmodium falciparum* has distinct organelle-specific lipoylation pathways. Mol Microbiol 53: 103-13.

Yasuno, R. and H. Wada (2002). The biosynthetic pathway for lipoic acid is present in plastids and mitochondria in Arabidopsis thaliana. FEBS Lett 517: 110-4.

Yasuno, R., P. von Wettstein-Knowles and H. Wada (2004). Identification and moleular characteisation of the B-ketoacyl -[acyl carrier portein] synthase component of the arabadopsis mitochondrial fatty acid synthase. J Biol Chem 279: 8242-51.

Zhang, Y. M., M. S. Rao, R. J. Heath, A. C. Price, A. J. Olson, C. O. Rock and S. W. White (2001). Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. J Biol Chem 276: 8231-8.

Zhang, Y. M., H. Marrakchi and C. O. Rock (2002). The FabR (YijC) transcription factor regulates unsaturated fatty acid biosynthesis in *Escherichia coli*. J Biol Chem 21: 21.

Zhang, Y. M., B. Wu, J. Zheng and C. O. Rock (2003). Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. J Biol Chem 278: 52935-43.

Zhang, L., A. K. Joshi and S. Smith (2003). Cloning, expression, characterization, and interaction of two components of a human mitochondrial fatty acid synthase. Malonyltransferase and acyl carrier protein. J Biol Chem 278: 40067-74.

Zhang, Y. M. and C. O. Rock (2004). Evaluation of epigallocatechin gallate and related plant polyphenols as inhibitors of the FabG and FabI reductases of bacterial type II fatty-acid synthase. J Biol Chem 279: 30994-1001.

Zhu, G., Y. Li, X. Cai, J. J. Millership, M. J. Marchewka and J. S. Keithly (2004). Expression and functional characterization of a giant Type I fatty acid synthase (CpFAS1) gene from Cryptosporidium parvum. Mol Biochem Parasitol 134: 127-35.

Zuegge, J., S. Ralph, M. Schmuker, G. I. McFadden and G. Schneider (2001). Deciphering apicoplast targeting signals--feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. Gene 280: 19-26.

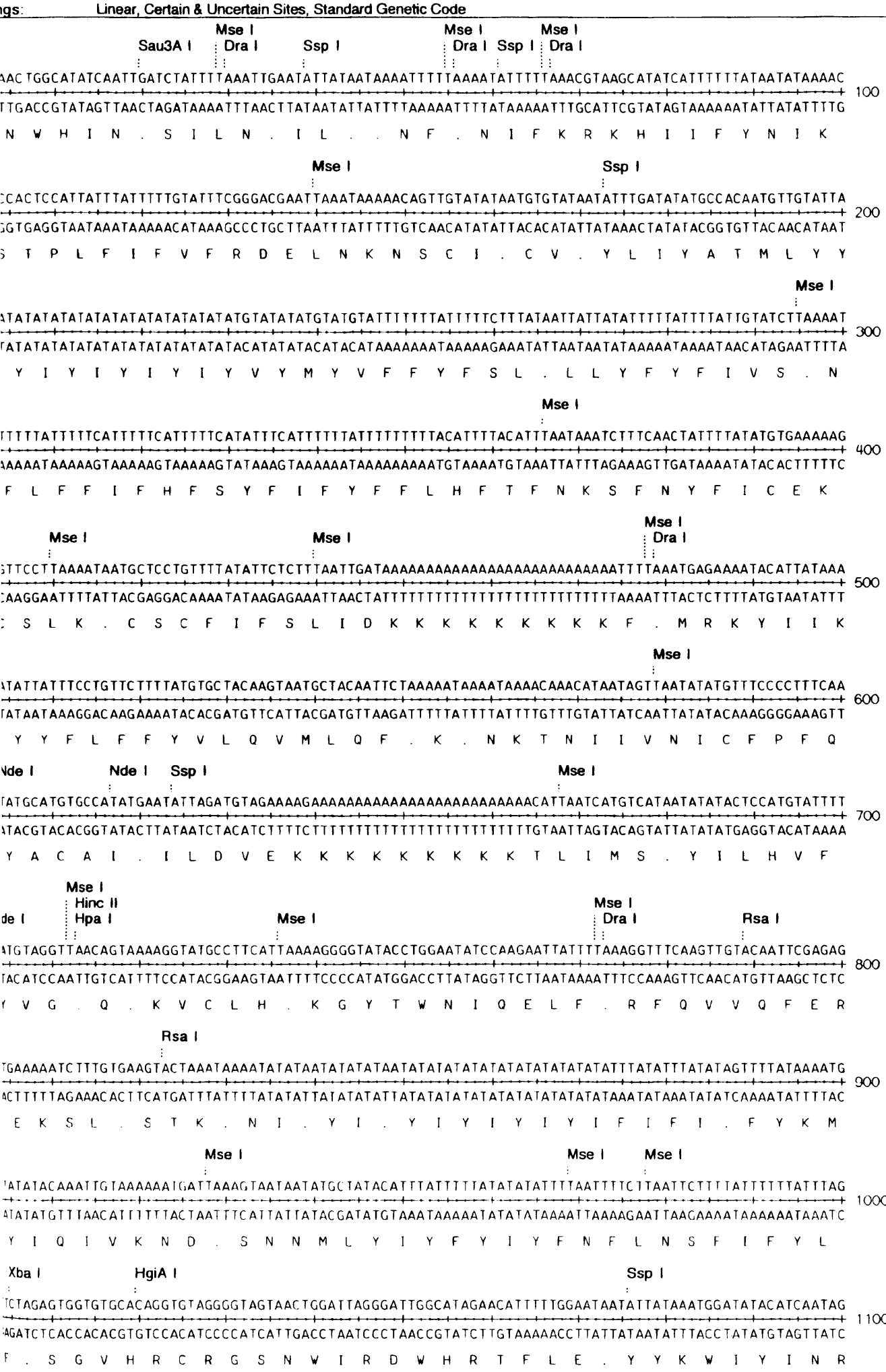
Zuther, E., J. J. Johnson, R. Haselkorn, R. McLeod and P. Gornicki (1999). Growth of Toxoplasma gondii is inhibited by aryloxyphenoxypropionate herbicides targeting acetyl-CoA carboxylase. Proc Natl Acad Sci U S A 96: 13387-92.

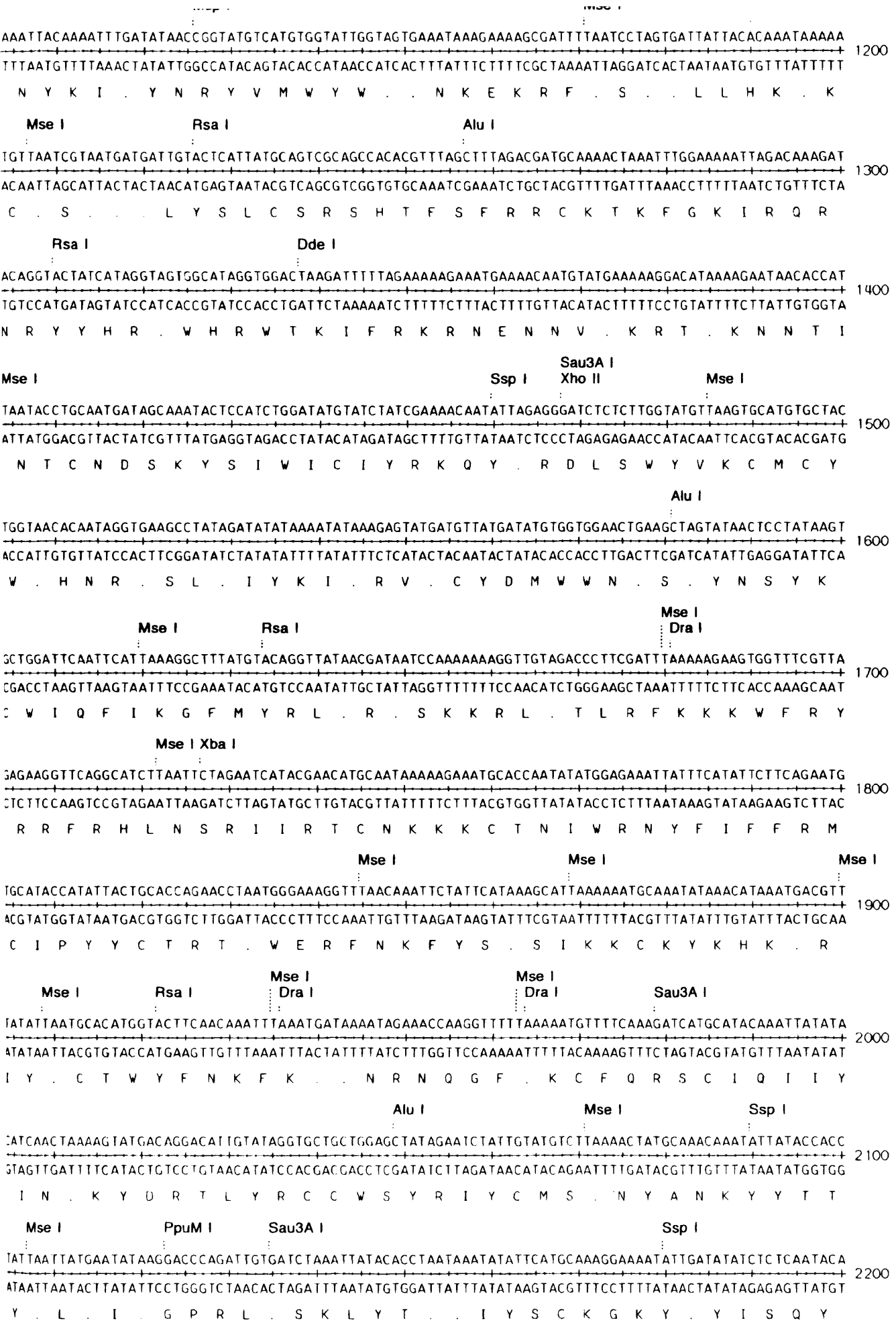
Appendices

List of Appendices

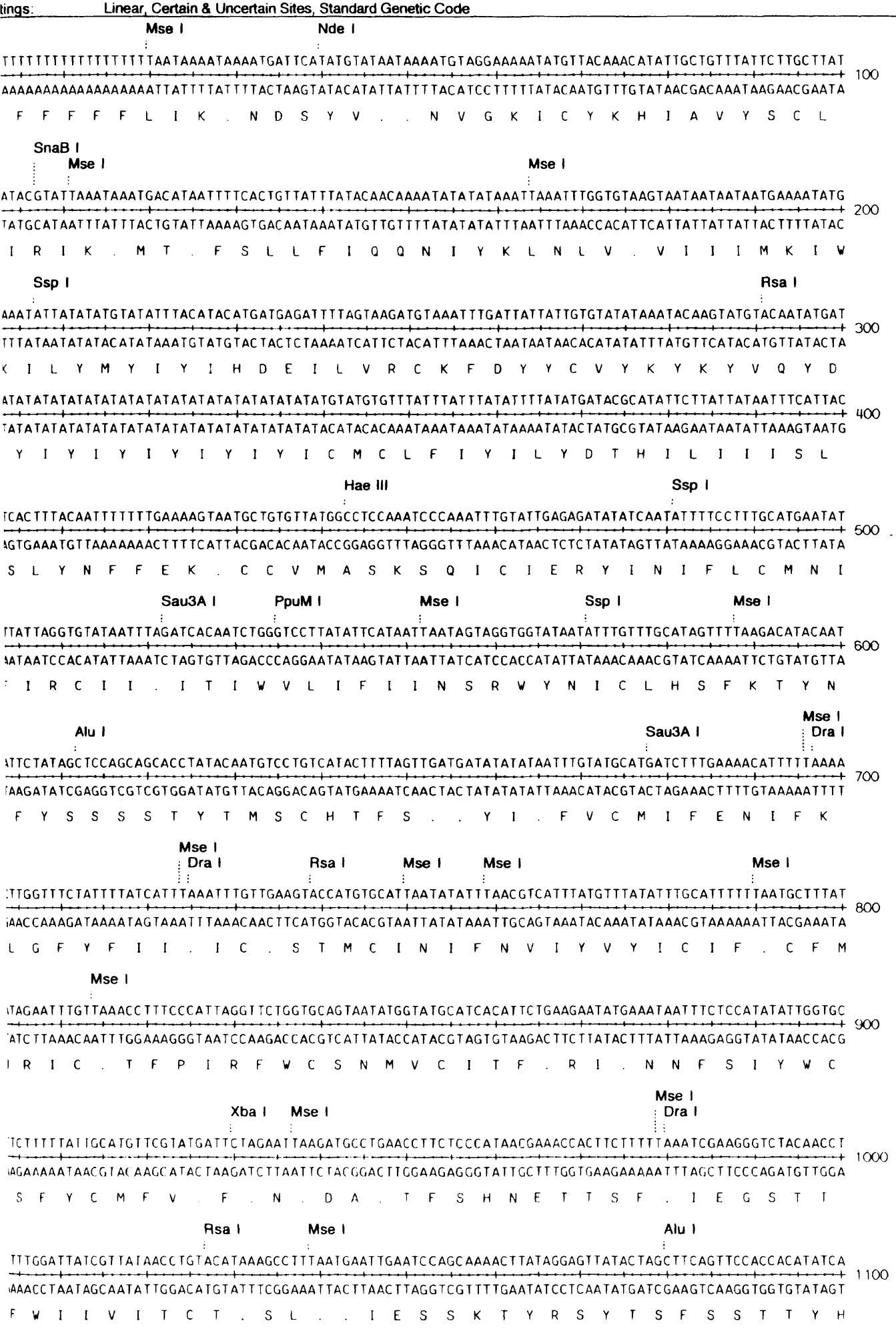
Publications that arose from work presented in this report

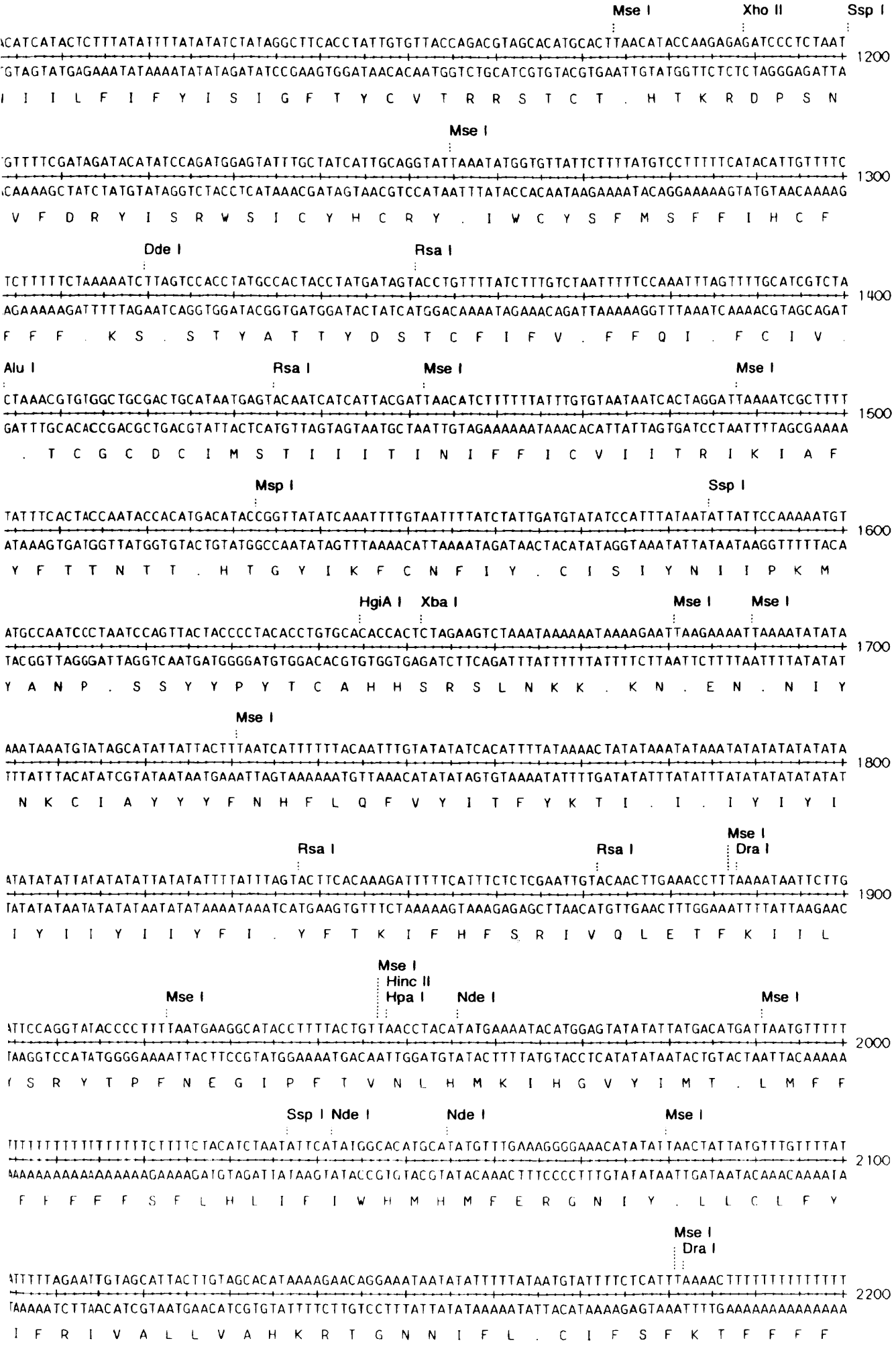
Appendix 1

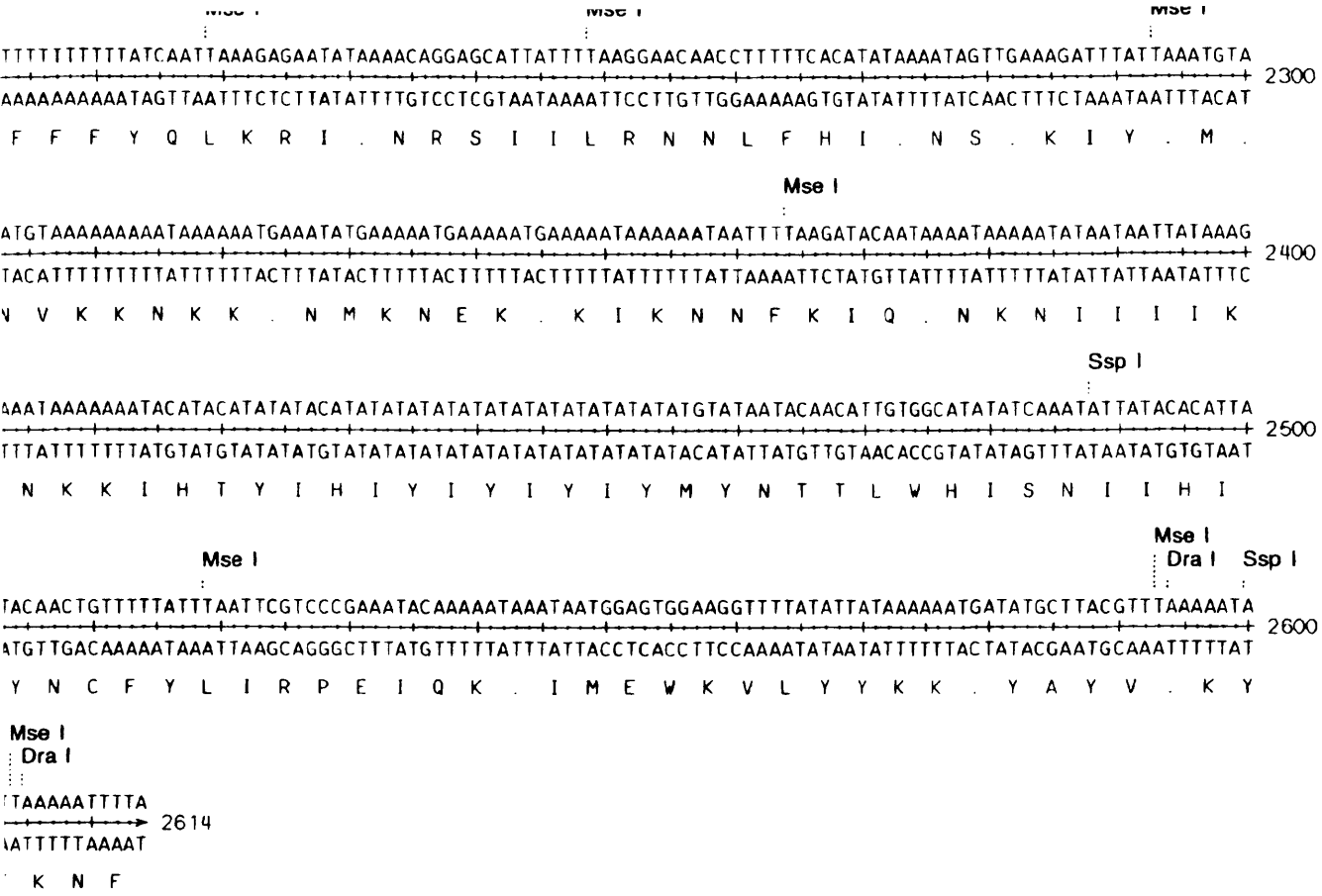

gs: Linear, Certain & Uncertain Sites, Standard Genetic Code
Sau3A I
Mse I
Dra I
Ssp I
Mse I
Dra I
Ssp I
Mse I
Dra I
Mse I
Ssp I
Mse I
Mse I
Mse I
Dra I
Nde I
Nde I
Ssp I
Mse I
Mse I
Hinc II
Hpa I
Mse I
Mse I
Dra I
Rsa I
Rsa I
Mse I
Mse I
Mse I
Xba I
HgiA I
Ssp I
100
200
300
400
500
600
700
800
900
1000
1100


Mse I
Rsa I
Alu I
Rsa I
Dde I
Mse I
Ssp I
Sau3A I
Xho II
Mse I
Alu I
Mse I
Rsa I
Mse I
Dra I
Mse I Xba I
Mse I
Mse I
Mse I
Mse I
Rsa I
Mse I
Dra I
Mse I
Dra I
Sau3A I
Alu I
Mse I
Ssp I
Mse I
PpuM I
Sau3A I
Ssp I
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200

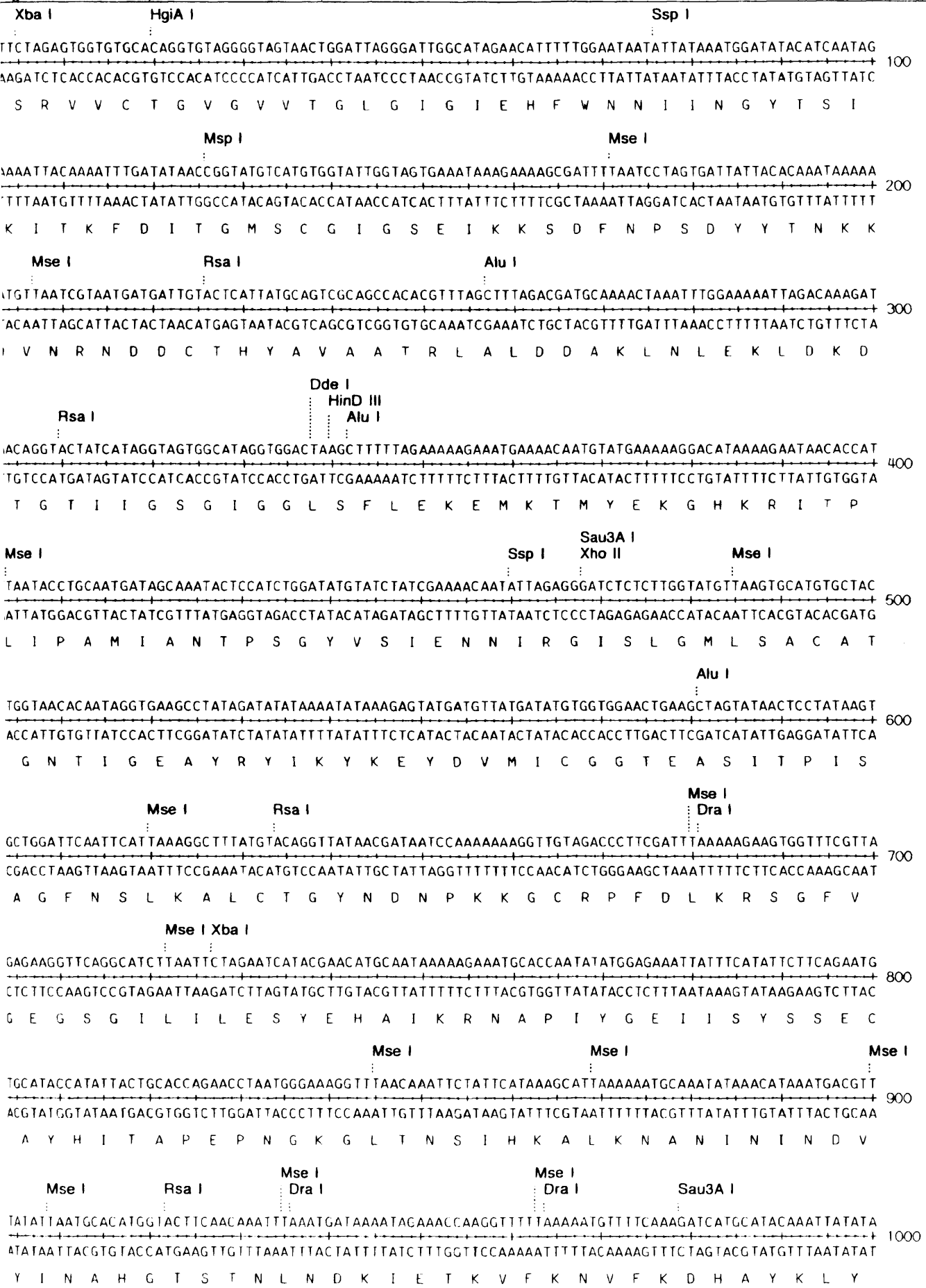
tings: Linear, Certain & Uncertain Sites, Standard Genetic Code

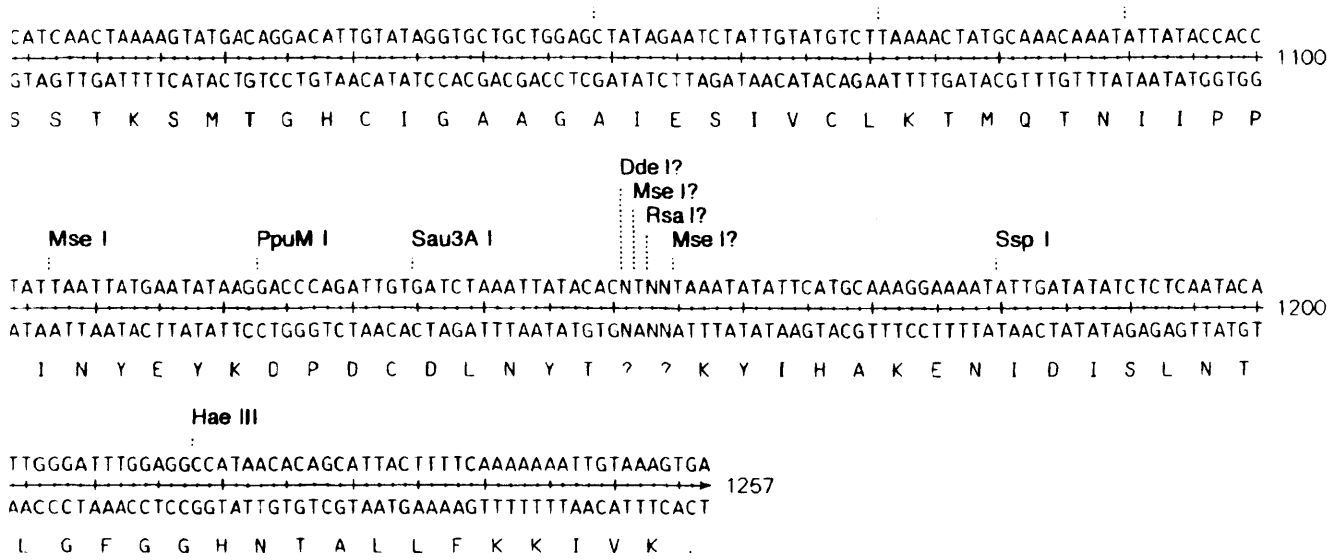



Mse I
Xho II
Ssp I
1200
Mse I
1300
Dde I
Rsa I
1400
Alu I
Rsa I
Mse I
Mse I
1500
Msp I
Ssp I
1600
HgiA I
Xba I
Mse I
Mse I
1700
Mse I
1800
Rsa I
Rsa I
Mse I
Dra I
1900
Mse I
Hinc II
Hpa I
Mse I
Nde I
Mse I
2000
Ssp I
Nde I
Nde I
Mse I
2100
Mse I
Dra I
2200


Mse I
Mse I
Mse I
TTTTTTTTTTATCAATTAAAGAGAATATAAAACAGGAGCATTATTTTAAGGAACAACCTTTTTCACATATAAAATAGTTGAAAGATTTATTAAATGTA
AAAAAAAAAATAGTTAATTTCTCTTATATTTTGTCCTCGTAATAAAATTCCTTGTTGGAAAAAGTGTATATTTTATCAACTTTCTAAATAATTTACAT
2300
F F F Y Q L K R I . N R S I I L R N N L F H I . N S . K I Y . M .
Mse I
ATGTAAAAAAAAATAAAAAATGAAATATGAAAAATGAAAAATGAAAAATAAAAAATAATTTTAAGATACAATAAAATAAAAATATAATAATTATAAAG
TACATTTTTTTTTATTTTTTACTTTATACTTTTTACTTTTTACTTTTTATTTTTTATTAAAATTCTATGTTATTTTATTTTTATATTATTAATATTTC
2400
V K K N K K . N M K N E K . K I K N N F K I Q . N K N I I I I K
Ssp I
AAATAAAAAAAATACATACATATATACATATATATATATATATATATATATATATATGTATAATACAACATTGTGGCATATATCAAATATTATACACATTA
TTTATTTTTTTATGTATGTATATATGTATATATATATATATATATATATATATATATACATATTATGTTGTAACACCGTATATAGTTTATAATATGTGTAAT
2500
N K K I H T Y I H I Y I Y I Y I Y M Y N T T L W H I S N I I H I
Mse I
Mse I
Dra I
Ssp I
TACAACTGTTTTTATTTAATTCGTCCCGAAATACAAAAATAAATAATGGAGTGGAAGGTTTTATATTATAAAAAATGATATGCTTACGTTTAAAAATA
ATGTTGACAAAAATAAATTAAGCAGGGCTTTATGTTTTTATTTATTACCTCACCTTCCAAAATATAATATTTTTTACTATACGAATGCAAATTTTTAT
2600
Y N C F Y L I R P E I Q K . I M E W K V L Y Y K K . Y A Y V . K Y
Mse I
Dra I
TAAAAATTTTA
ATTTTTAAAAT
2614
K N F

ings: Linear, Certain & Uncertain Sites, Standard Genetic Code




CATCAACTAAAAGTATGACAGGACATTGTATAGGTGCTGCTGGAGCTATAGAATCTATTGTATGTCTTAAAACTATGCAAACAAATATTATACCACC
1100
GTAGTTGATTTTCATACTGTCCTGTAACATATCCACGACGACCTCGATATCTTAGATAACATACAGAATTTTGATACGTTTGTTTATAATATGGTGG
S S T K S M T G H C I G A A G A I E S I V C L K T M Q T N I I P P
Dde I?
Mse I?
Rsa I?
Mse I
PpuM I
Sau3A I
Mse I?
Ssp I
TATTAATTATGAATATAAGGACCCAGATTGTGATCTAAATTATACACNTNNTAAATATATTCATGCAAAGGAAAATATTGATATATCTCTCAATACA
1200
ATAATTAATACTTATATTCCTGGGTCTAACACTAGATTTAATATGTGNANNATTTATATAAGTACGTTTCCTTTTATAACTATATAGAGAGTTATGT
I N Y E Y K D P D C D L N Y T ? ? K Y I H A K E N I D I S L N T
Hae III
TTGGGATTTGGAGGCCATAACACAGCATTACTTTTCAAAAAAATTGTAAAGTGA
1257
AACCCTAAACCTCCGGTATTGTGTCGTAATGAAAAGTTTTTTTAACATTTCACT
L G F G G H N T A L L F K K I V K .

Appendix 2

Appendix 2 Vector specific PCR primers used in this project

M13F: 5'-GTAAAACGACGGCCAGT-3'

M13R: 5'-CAGGAAACAGCTATGAC-3'

T7: 5'-TAATACGACTCACTATAGGG -3'

SF1: 5'-GCTAGTTATTGCTCAGCGG -3'

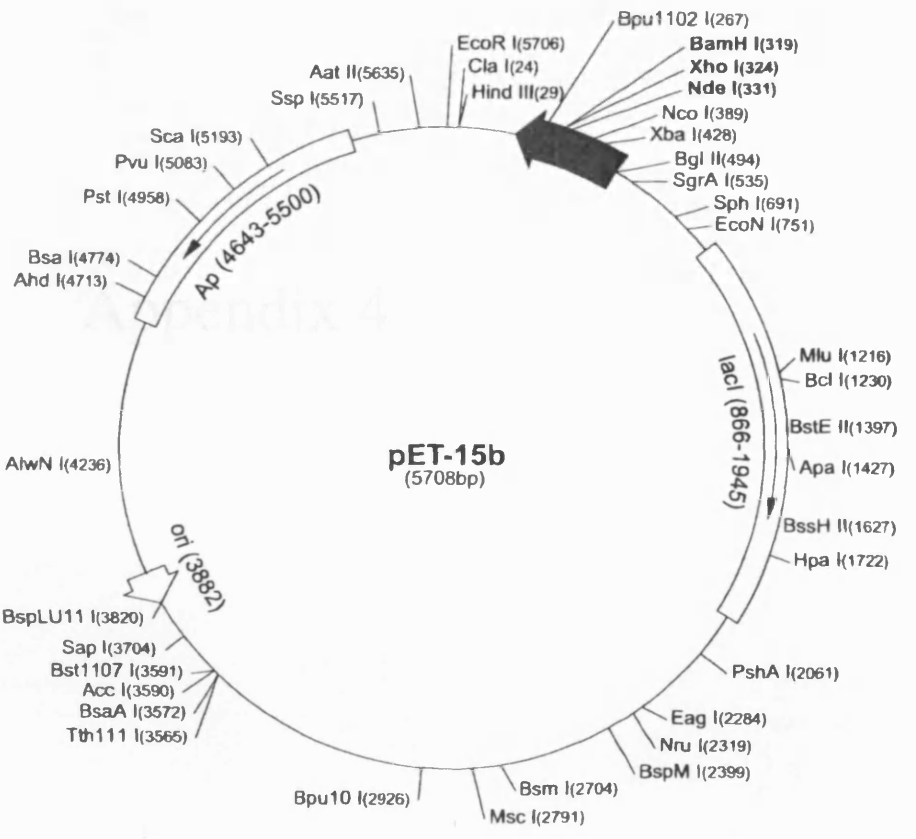
Appendix 3

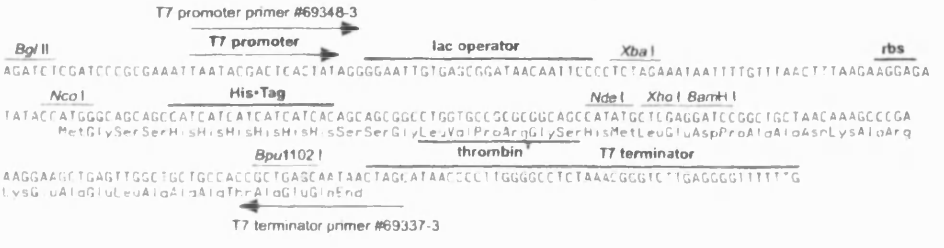
pET-15b Vector

TB045 5/99

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

pET-15b sequence landmarks	
T7 promoter	463-479
T7 transcription start	452
His•Tag coding sequence	362-380
Multiple cloning sites (*Nde* I - *Bam*H I)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
bla coding sequence	4643-5500





pET-15b cloning/expression region

Appendix 4

pMAL™-p2X

6,721 base pairs
Sequence file available at www.neb.com
See page 138 for ordering information.

Feature	Coordinates	Source
lacI^q	81-1163	*E. coli*
P_{tac}	1406-1433	–
expression ORF	1528-2991	–
malE	1528-2703	*E. coli*
MCS	2704-2809	–
*lacZ*α	2810-2991	–
bla (Ap^R)	3493-4353	*Tn3*
M13 origin	4395-4908	M13
origin	5019-5607	pMB1
rop	6228-6037	pMB1

ori = origin of replication
Ap = ampicillin

References

1. Guan, C. et al. (1987) *Gene* 67, 21–30.
2. Maina, C.V. et al. (1988) *Gene* 74, 365–373.
3. Riggs, P., in Ausubel, F.M. et al. (eds), *Current Prot. in Molecular Biol.* (1992) Greene Associates/Wiley Interscience, New York.
4. Zagursky, R.J. et al. (1984) *Gene* 27, 183–191.

pMAL-p2X is an *E. coli* plasmid cloning vector designed for recombinant protein expression and purification using the pMAL Protein Fusion and Purification System (NEB #E8000S) (1-3). It contains the pMB1 origin of replication from pBR322 and is maintained at a similar copy number to pBR322; in addition, pMAL-p2X also contains an M13 origin of replication (4).

The multiple cloning site (MCS) is positioned to allow translational fusion of the *E. coli* maltose binding protein (MBP, encoded by the *malE* gene) to the N-terminus of the cloned target protein. MBP's affinity for amylose allows easy purification of the fusion protein, and the MBP domain can be subsequently removed using Factor Xa protease (3). Cloning of the target gene at the MCS disrupts expression of *lacZ*α, allowing for insert screening by α-complementation.

Transcription of the gene fusion is controlled by the inducible "tac" promoter (P_{tac}). Basal expression from P_{tac} is minimized by the binding of the Lac repressor, encoded by the *lacI*^{q} gene, to the *lac* operator immediately downstream of P_{tac}. A portion of the *rrnB* operon containing two terminators, derived from the vector pKK233-2, prevents transcription originating from P_{tac} from interfering with plasmid functions.

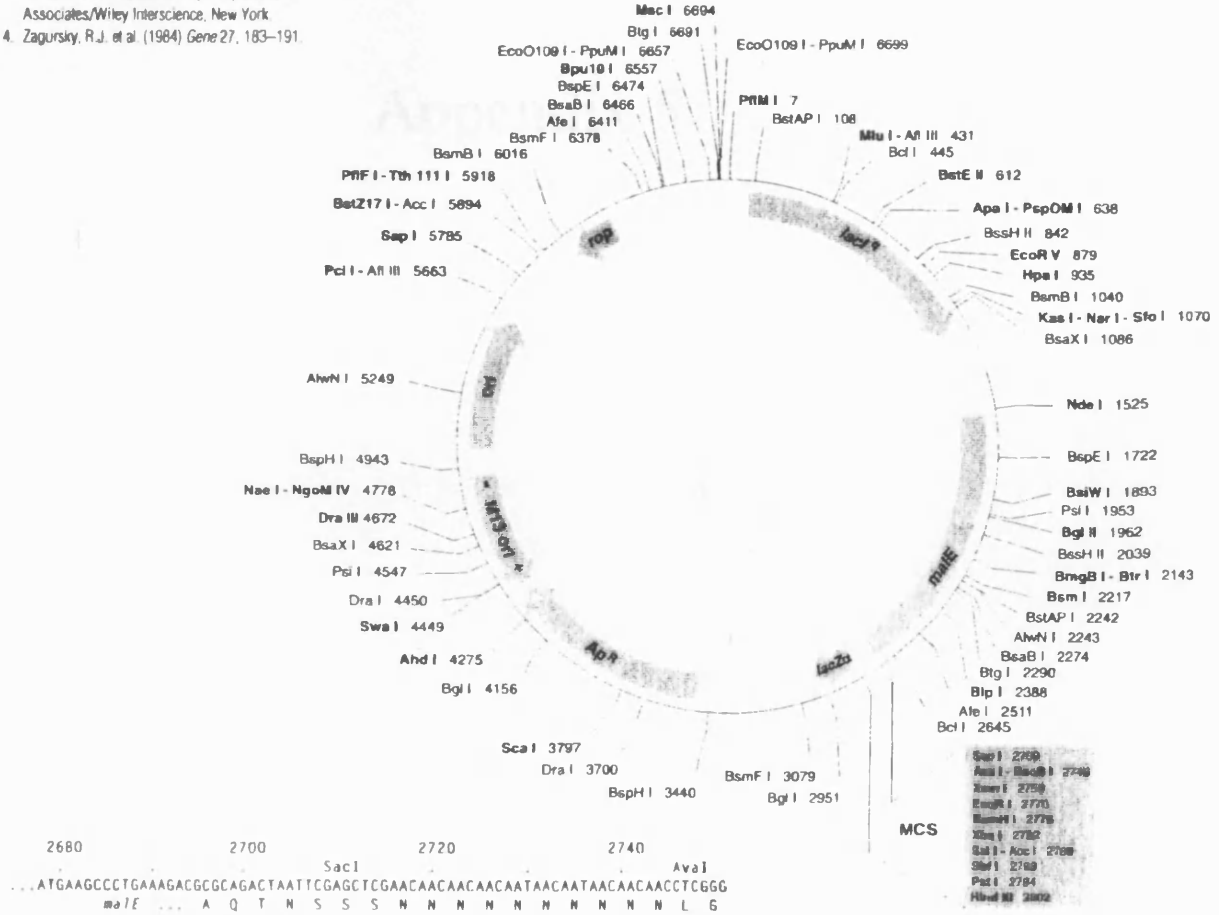
pMAL-p2E and pMAL-p2G are identical to pMAL-p2X except they replace the Factor Xa protease cleavage site with Enterokinase and Genenase I™ cleavage sites, respectively.

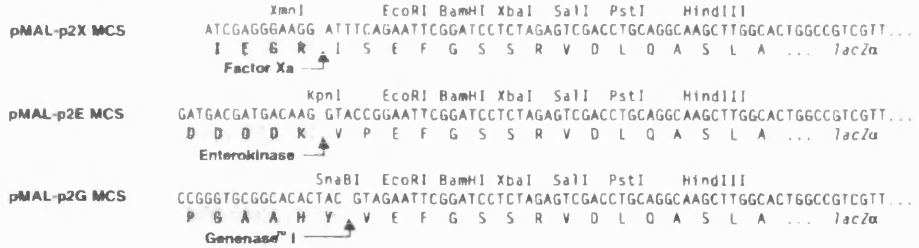
pMAL-c2-series vectors are identical to the pMAL-p2-series vectors above except for a deletion of the *malE* signal sequence (nt 1531-1605) (1).

Enzymes with unique restriction sites are shown in **bold** type and enzymes with two restriction sites are shown in regular type. The accompanying table shows restriction sites of those enzymes that cut a moderate number of times. Restriction site coordinates refer to the position of the 5′-most base on the top strand in each recognition sequence.

Open reading frame (ORF) coordinates are in the form "translational start – translational stop"; numbers refer to positions on the top (clockwise) strand, regardless of the direction of transcription and include the start and stop codons.

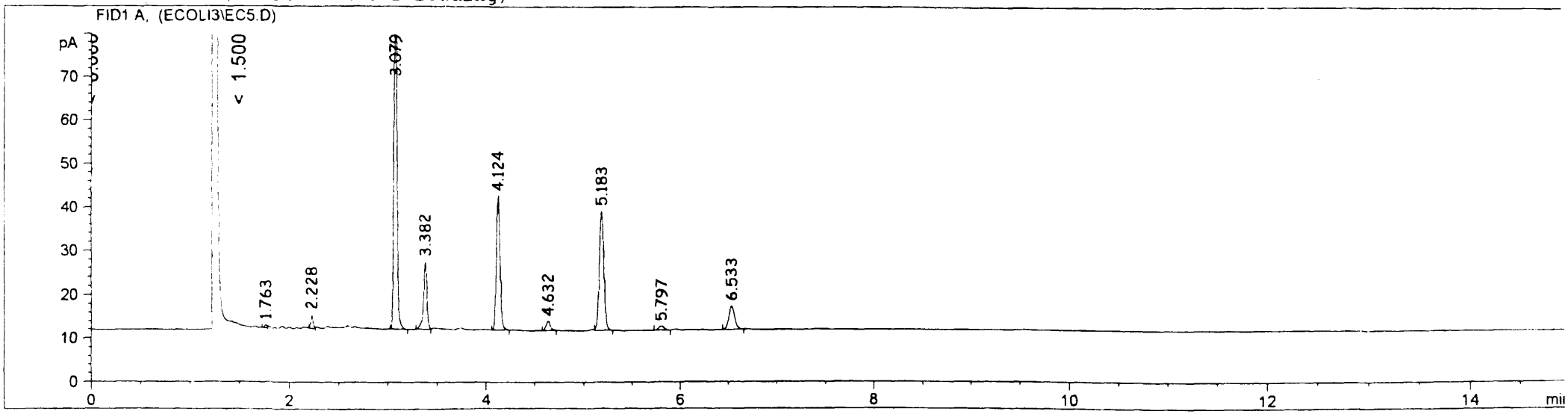
pMB1 origin of replication coordinates include the region from the -35 promoter sequence of the RNAII transcript to the RNA/DNA switch point. For the M13 origin, the arrow shows the direction of synthesis of the (+) strand, which gets packaged into phage particles. *bla* (Ap^R) gene coordinates include the signal sequence.





Appendix 5

```
Injection Date  : 30/03/2004 17:28:36 PM          Seq. Line :   5
Sample Name     :                                 Location : Vial 5
Acq. Operator   : IAG                                 Inj :   1
                                                Inj Volume : 1 µl
Acq. Method     : C:\HPCHEM\1\METHODS\MOR\WHEAT.M
Last changed    : 20/01/2004 11:37:40 PM by IAG
Analysis Method : C:\HPCHEM\1\METHODS\MOR\BLOLIP.M
Last changed    : 31/03/2004 10:50:18 PM by IAG
                  (modified after loading)
```



Area Percent Report

```
Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000

Signal 1: FID1 A,
```

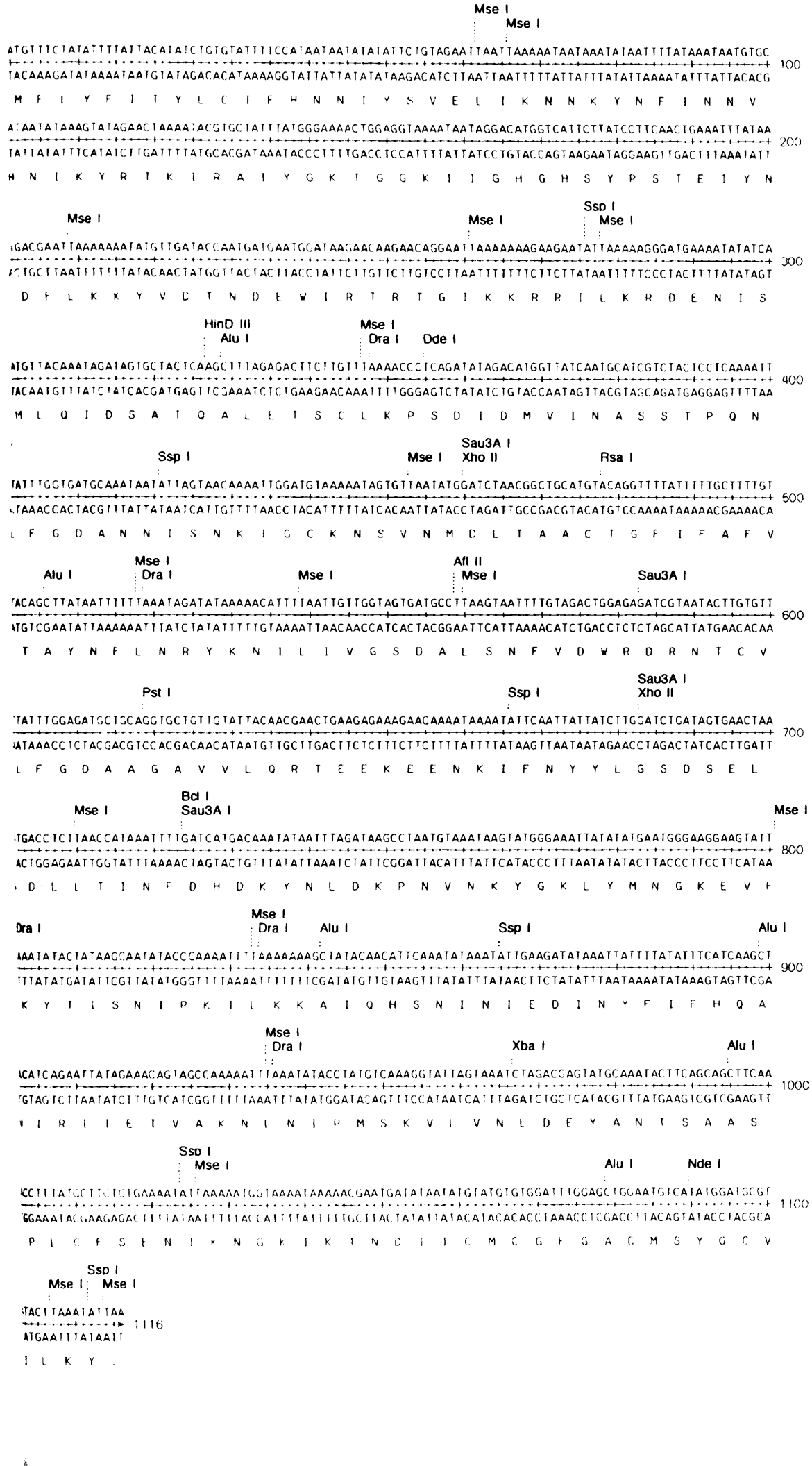
Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.763	VV	0.0273	1.26972	7.12452e-1	0.30392
2	2.228	VV	0.0259	5.01989	3.06323	1.20157

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
3	3.079	BB	0.0323	170.51694	81.81861	40.81539
4	3.382	PV	0.0366	37.08362	15.38176	8.87643
5	4.124	PB	0.0409	81.84461	30.87967	19.59054
6	4.632	BB	0.0465	6.59372	2.13839	1.57829
7	5.183	VB	0.0504	88.79726	27.29940	21.25475
8	5.797	BB	0.0566	3.64862	9.33687e-1	0.87334
9	6.533	BB	0.0671	23.00173	5.41026	5.50576
Totals :				417.77611	167.63746	

Results obtained with enhanced integrator!

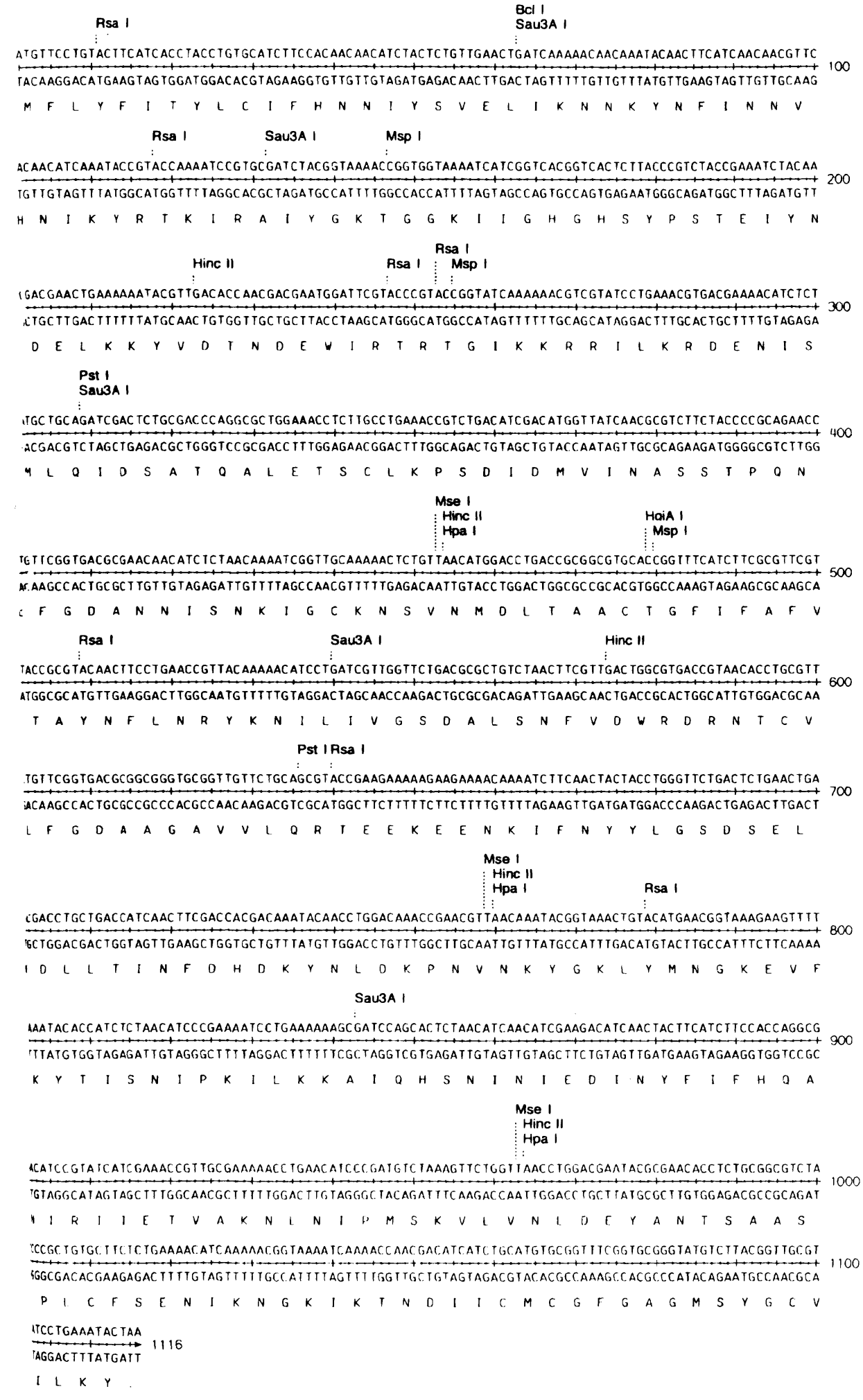
*** End of Report ***

Appendix 6


Mse I
Mse I
100
Ssp I
Mse I
Mse I
Mse I
300
HinD III
Alu I
Mse I
Dra I
Dde I
400
Sau3A I
Ssp I
Mse I
Xho II
Rsa I
500
Alu I
Mse I
Dra I
Mse I
Afl II
Mse I
Sau3A I
600
Sau3A I
Pst I
Ssp I
Xho II
700
Bcl I
Mse I
Sau3A I
Mse I
800
Mse I
Dra I
Dra I
Alu I
Ssp I
Alu I
900
Mse I
Dra I
Xba I
Alu I
1000
Ssp I
Mse I
Alu I
Nde I
1100
Ssp I
Mse I
Mse I
1116

Appendix 7

pfFabH Ecoli preferred codo Map (1 > 1116) Site and Sequence
Enzymes 47 of 436 enzymes (Filtered)
Settings: Linear, Certain Sites Only, Standard Genetic Code



Overlapping oligonucleotides for *E. coli* preferred codons of *pfFabH*

CODING 1

5'CTCGAGGATCCCATATGTTCCTGTACTTCATCACCTACCTGTGCATCTTCCACAACAACA
TC3'

CODING 2

5'CAACTTCATCAACAACGTTCACAACATCAAATACCGTACCAAAATCCGTGCGATCTACG
GTAAAACCGGT3'

CODING 3

5'CGTCTACCGAAATCTACAACGACGAACTGAAAAAATACGTTGACACCAACGACGAATGG
ATTCGTACCCGTACCGG3'

CODING 4

5'CATCTCTATGCTGCAGATCGACTCTGCGACCCAGGCGCTGGAAACCTCTTGCCTGAAACC
GTCTGACATCGACATGG3'

CODING 5

5'CGGTGACGCGAACAACATCTCTAACAAAATCGGTTGCAAAAACTCTGTTAACATGGACC
TGAC3'

CODING 6

5'CCGCGTACAACTTCCTGAACCGTTACAAAAACATCCTGATCGTTGGTTCTGACGCGCTGT
CTAACTTCGTTGACTGG3'

CODING 7

5'CGGTGACGCGGCGGGTGCGGTTGTTCTGCAGCGTACCGAAGAAAAAGAAGAAAACAAA
ATCTTC3'

CODING 8

5'ACTGAACGACCTGCTGACCATCAACTTCGACCACGACAAATACAACCTGGACAAACCGA
ACGTTAACAAATACGG3'

CODING 9

5'CAAATACACCATCTCTAACATCCCGAAAATCCTGAAAAAAGCGATCCAGCACTCTAACA
TCAACATCGAAGAC3'

CODING 10

5'CATCCGTATCATCGAAACCGTTGCGAAAAACCTGAACATCCCGATGTCTAAAGTTCTGGT
TAACCTGG3'

CODING 11

5'TATCCCGCTGTGCTTCTCTGAAAACATCAAAAACGGTAAAATCAAAACCAACGACATCA
TCTGCATGTGCGG3'

NON-CODING 11

5'TGAACGTTGTTGATGAAGTTGTATTTGTTGTTTTTGATCAGTTCAACAGAGTAGATGTTGT
TGTGGAAGATGC3'

NON-CODING 10

5'CGTTGTAGATTTCGGTAGACGGGTAAGAGTGACCGTGACCGATGATTTTACCACCGGTTT
TACCGTAGATCG3'

NON-CODING 9

5'CGATCTGCAGCATAGAGATGTTTTCGTCACGTTTCAGGATACGACGTTTTTTGATACCGG
TACGGGTACGAATCC3'

NON-CODING 8

5'GAGATGTTGTTCGCGTCACCGAACAGGTTCTGCGGGGTAGAAGACGCGTTGATAACCAT
GTCGATGTCAGACGG3'

NON-CODING 7

5'GGTTCAGGAAGTTGTACGCGGTAACGAACGCGAAGATGAAACCGGTGCACGCCGCGGTC
AGGTCCATGTTAACAGAG3'

NON-CODING 6

5'CCGCACCCGCCGCGTCACCGAACAGAACGCAGGTGTTACGGTCACGCCAGTCAACGAAG
TTAGACAGC3'

NON-CODING 5

5'GGTCAGCAGGTCGTTCAGTTCAGAGTCAGAACCCAGGTAGTAGTTGAAGATTTTGTTTTC
TTCTTTTTC3'

NON-CODING 4

5'GGATGTTAGAGATGGTGTATTTGAAAACTTCTTTACCGTTCATGTACAGTTTACCGTATTT
GTTAACGTTCGGT3'

NON-CODING 3

5'CGGTTTCGATGATACGGATGTTCGCCTGGTGGAAGATGAAGTAGTTGATGTCTTCGATGT
TGATGTTAGAG3'

NON-CODING 2

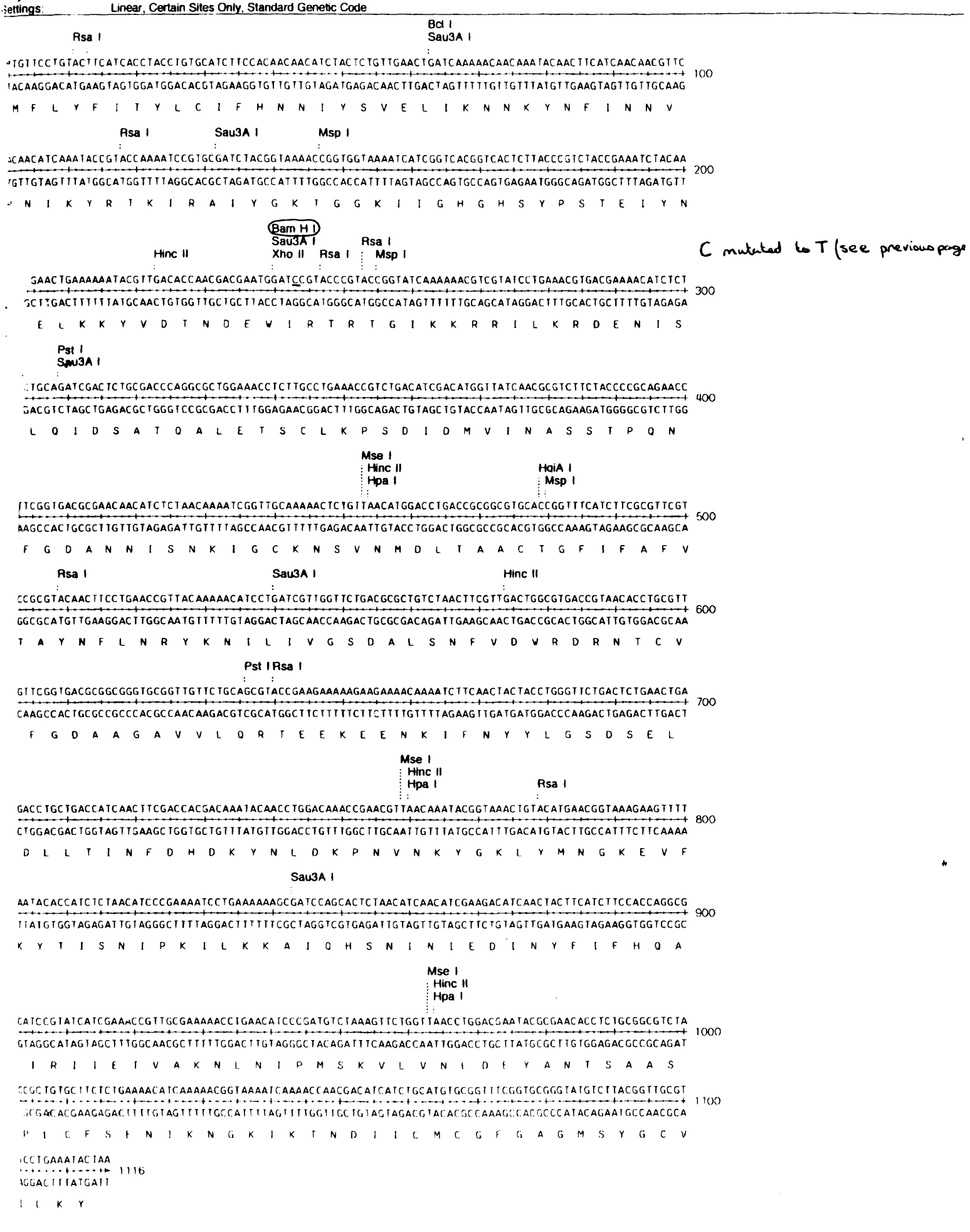
5'CAGAGAAGCACAGCGGGATAGACGCCGCAGAGGTGTTCGCGTATTCGTCCAGGTTAACC
AGAACTTTAG3'

NON-CODING 1 Start of complementary sequence to pdfab H

NdeI BamHI XhoI

5'CATATGGATCCCTCGAGTTAGTATTTCAGGATAACGCAACCGTAAGACATACCCGCACC
GAAACCGCACATGCAGATGATGTC3'

nzymes: 47 of 436 enzymes (Filtered)
ettings: Linear, Certain Sites Only, Standard Genetic Code

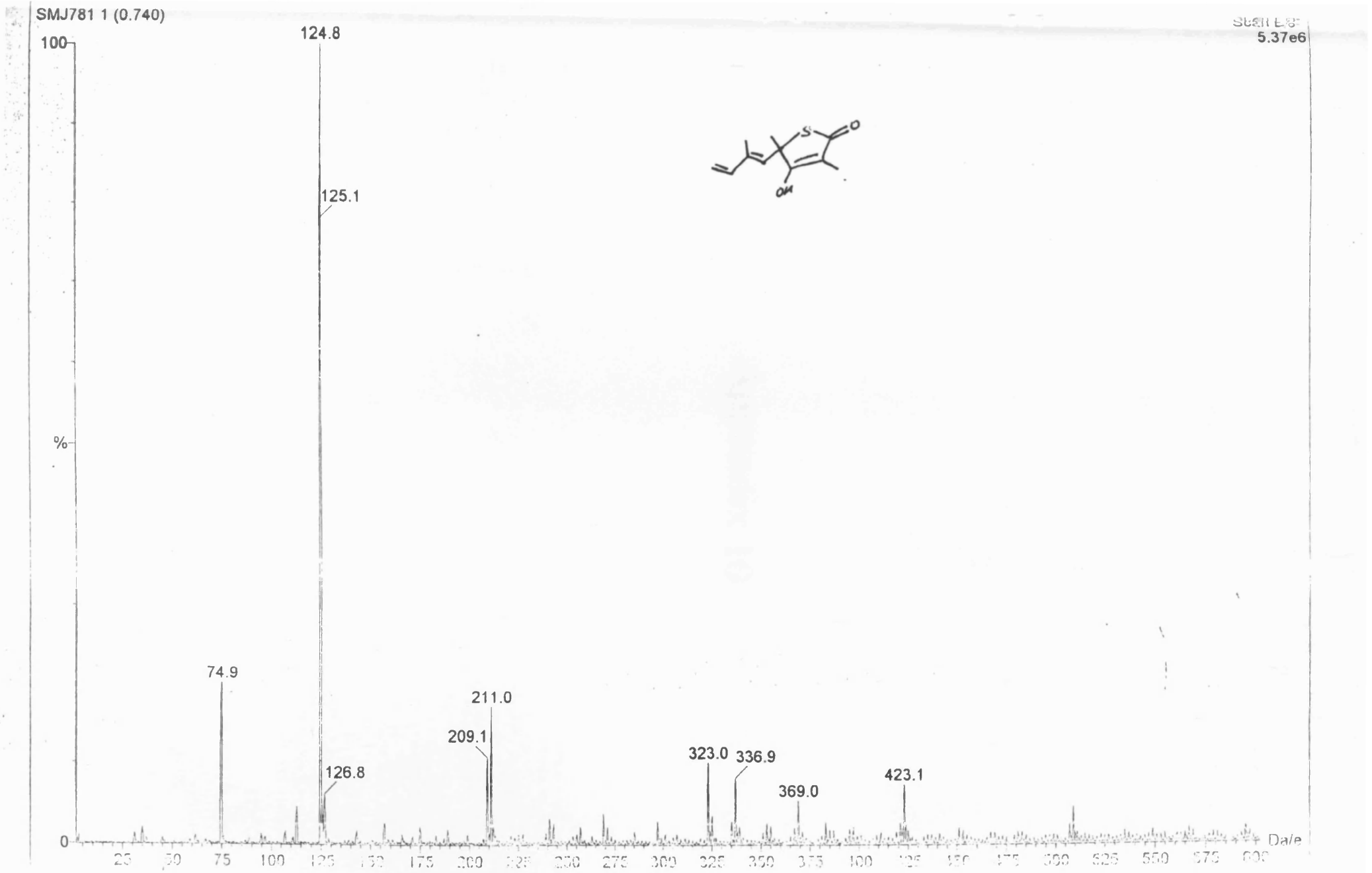


Appendix 8

Compound No.	*P. falciparum* LD_{50} (μM)	*T. brucei* LD_{50} (μM)	*T. cruzi* LD_{50} (μM)	*L. donovani* LD_{50} (μM)
Sigma TLM	>95	182	>142	56
776	36	171	68	>105
777	25	62	28	>83
778	40	190	68	>106.9
779	139	112	127	>131.3
780	153	194	>350	>116.9
781	143	256	>427	>142.6
817	347	>624	>624	Not tested
818	233	357	568	Not tested
819	19	47	43	Not tested
820	61	71	195	>123.7
821	15	97	54	>100.5
822	316	>568	>568	>189.6
823	54	157	165	>110.9
825	65.2	153	72	>100.5
863	34.4	130	49	>78.3
922	222.1	246	>423	>141.3
924	290.2	288	>522	>174.1
925	10.7	21	56	65.9
926	195.3	408	>408	>136.1
927	50.7	132	64	>83.2
943	17.57	**6**	164	97.0
944	15.3	8	77	77.4
945	16.1	13	70	>96.6
1020	62.4	79.2	>488.4	29.9
1021	60.5	119.8	195.4	9.3
1022	26.1	157	56	17.5
1023	6.3	167	31	10.5
1084	63	**6.7**	153	3.3
1086	37	**2.0**	150	**1.1**
1140	44	9	359	100
1141	19	9	47	82
1152	**1.3**	14	80	66
1192	**4**	>223	**13**	>74
1193	**2.7**	149	**14**	7
1194	8.5	55	21	12
1195	73	246	>110	88
1199	6	29	**13**	**0.5**
1200	7	32	**14**	**2**
1209	11	>200	>66	>66
1210	19	64	36	36

Appendix 8 Biological activity of TLM derivatives against parasitic protozoa
This table details all of the data obtained from assays against cultures of different parasites. The three best LD_{50} values against each organism are shown in bold. Assays were performed as described in Jones *et al.*, 2004.

Appendix 9


SMJ781 1 (0.740)
Scan ES-
5.37e6
100
%
0
124.8
125.1
74.9
126.8
211.0
209.1
323.0
336.9
369.0
423.1
S
O
OH
Da/e

Appendix 10

11 Generation of anti-complement prodrugs for specific delivery of active regulator to disease sites

C.L. Harris, C.E. Hughes, I.G. Goodfellow, B. Caterson, B.P. Morgan
Department of Medical Biochemistry, UWCM, Heath Park, Cardiff, CF14 4XN; Connective Tissue Biology Labs, Cardiff School of Biosciences, Cardiff University, CF10 3US; Division of Virology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, G11 5JR, UK.

We have developed a novel strategy for specific delivery of anti-complement prodrugs to sites of complement-mediated inflammatory disease. Fusion of the regulatory domains of complement inhibitors (CD55 and CD59) to the Fc domains of antibody extends plasma half life but can also substantially reduce biological activity when compared to native regulator or regulator released from the Fc following papain cleavage. By incorporating cleavage sites for MMPs and/or aggrecanases between CD55 and the Fc, we have engineered model prodrugs that would be cleaved at inflammatory sites, such as an inflamed joint, to release active inhibitor. Exposure of these prodrugs to relevant enzymes, either as pure enzyme or as supernatants from cytokine-stimulated chondrocytes, efficiently cleaved the reagents, releasing active CD55. Optimisation by modification of the enzyme cleavage site and antibody Fc are continuing to create a prodrug with essentially no systemic anti-complement activity that is efficiently cleaved at inflammatory sites to release the active inhibitor. Such agents represent a paradigm for the next generation of anti-complement therapeutics.

12 Study of wheat transformed with sense and anti-sense lipase genes

D.A.N. Edlin, P. Kille, C.M. Saunders, H.D. Jones* and J.L. Harwood
*School of Biosciences, Cardiff University, Cardiff, CF10 3US, UK and *Crop Performance and Improvement, IACR-Rothamsted, Harpenden, Herts, AL5 2JQ, UK*

Wheat is a member of the family Gramineae, which includes some of the world's most important crops. It is primarily used in the production of flour for bread, pasta and cakes. Milled wholemeal flours are susceptible to the action of bran lipases that lead to the production of rancid off-flavours, a loss in nutritional value and a reduced baking performance. Deterioration of wholemeal flours is initiated by bran lipases on triacylglycerols in the aleurone layer. Rapid-acting germ lipoxygenases then produce oxidative products from the non-esterified fatty acids released by the lipases when the flour is mixed with water. Wastage, due to lipase action, accounts for 5% of total wholemeal flour in the UK, which equates to an annual loss of approximately 6 million pounds sterling. We have identified a gene encoding lipase activity from wheat. This gene has been used to transform varieties of wheat, using the technique of particle bombardment, with sense and anti-sense copies of the wheat lipase gene under the control of the maize ubiquitin promoter.
We acknowledge the financial assistance of the BBSRC.

13 Fatty acid synthesis as a target for novel anti-malarials

J.E. Urch, C. Berry, J.L. Harwood
School of Biomedical Sciences, Cardiff University, P.O. Box 911, Museum Avenue, Cardiff, CF10 3US, UK

Malaria is a parasitic disease that kills over 1 million people every year. Increasing resistance to current anti-malarial agents is responsible for the resurgence of malaria and necessitates the development of new drugs to combat this disease.
Fatty acid synthesis is crucial for all living cells. Plants and bacteria utilise a dissociable (Type II) multi-enzyme fatty acid synthase (FAS) complex, which differs from the multi-functional FAS dimer found in mammals. The malaria parasite, Plasmodium falciparum, contains a Type II FAS system, which represents a new target for the development of antimalarial agents that would not compromise the host's lipid metabolism.
β-Ketoacyl-ACP synthases (KAS) catalyse the condensation reactions of FASs and are obvious potential targets. The antibiotic, thiolactomycin is an inhibitor of Type II KAS enzymes. A gene for one KAS isozyme from P. falciparum has been identified previously. We have identified a second P. falciparum KAS gene and have used it for expression studies in E. coli. The expressed KAS enzymes will then be used in further studies with thiolactomycin and derivatives in order to try and find better chemicals for malaria therapy.
Supported by; the Cardiff Partnership Fund, the Morgan E. Williams PhD Fellowship and WHO.

14 Signalling pathways in chondrocyte metabolism and the effects of *n*-3 fatty acid supplementation

S. Hurst, C.L. Curtis, S.G. Rees, B. Caterson, J.L. Harwood
School of Biosciences, Cardiff University, Cardiff CF10 3US, UK

Previous work has demonstrated that *n-3* polyunsaturated fatty acids (PUFAs) inhibit degradative and inflammatory aspects of arthritis. This study was aimed at elucidating signalling pathways involved in the observed inhibition of mRNA expression of degradative enzymes and inflammatory mediators by *n-3* PuFAs. Human RNA from arthritic cartilage (with or without fatty acid supplementation, and in the absence or presence of IL-1β) was extracted and compared using microarray technology. Genes from the MAPKinase pathway (egr-1, c-fos), NFκB pathway (iNOS, NFκB, IκB-α, c-myc), TGF-β pathway/anti-proliferation pathway (p19, p21, p57) and IL-2 were seen to increase in the presence of IL-1. Supplementing with *n-3* PUFA reduced the levels of mRNA for the above genes.
n-3 PUFA have previously been shown to affect the MAPK and NFκB pathways. Our results suggest that pathological signalling pathways associated with arthritic disease can be inhibited by such fatty acids. Studies are on going to elucidate further the relationship between these pathways and the effect of *n-3* PUFAs in alleviating arthritic symptoms.

Appendix 11

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 683-692

Analogues of thiolactomycin as potential anti-malarial and anti-trypanosomal agents

Simon M. Jones,[a] Jonathan E. Urch,[b] Reto Brun,[c] John L. Harwood,[b] Colin Berry[b] and Ian H. Gilbert[a,*]

[a]*Welsh School of Pharmacy, Cardiff University, Redwood Building, King Edward VII Avenue, Cardiff CF10 3XF, UK*
[b]*Cardiff School of Biosciences, Biomedical Building, Museum Avenue, Cardiff CF10 3US, UK*
[c]*Swiss Tropical Institute, Socinstrasse 57, CH-4002 Basel, Switzerland*

Received 10 July 2003; accepted 21 November 2003

Abstract—A series of analogues of the naturally occurring antibiotic thiolactomycin (TLM) have been synthesised and evaluated for their ability to inhibit the growth of the malaria parasite, *Plasmodium falciparum*. Thiolactomycin is an inhibitor of Type II fatty acid synthase which is found in plants and most prokaryotes, but not an inhibitor of Type I fatty acid synthase in mammals. A number of the analogues showed inhibition equal to or greater than TLM. The introduction of hydrophobic alkyl groups at the C3 and C5 positions of the thiolactone ring lead to increased inhibition, the best showing a fourteenfold increase in activity over TLM. In addition, some of the analogues showed activity when assayed against the parasitic protozoa, *Trypanosoma cruzi* and *Trypanosoma brucei*.

1. Introduction

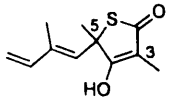
Malaria is by far the world's most important tropical disease. In many developing countries and in Africa especially, malaria exacts an enormous toll in lives, in medical costs, and in days of labour lost. At present, at least 300 million people are affected by malaria globally, and there are between 1 and 1.5 million malaria deaths annually.[1,2] Over the last several decades, the rapid development of drug-resistant strains has compounded the already serious health problems. Of the four known human malaria parasites, *Plasmodium falciparum* is the predominant cause of mortality, with 120 million new cases and 1 million deaths per year globally. It is this particular species, which has given rise to formidable drug-resistant strains, resulting in the urgent need for new chemotherapeutic agents. The search for new agents has recently been aided with the completion of the *P. falciparum* genome.[3] Detailed studies[4] of this genome have identified new potential drug and vaccine targets. One such target appears to be fatty acid biosynthesis of *P. falciparum*.

Fatty acid synthesis is a crucial function of living cells. The main steps in this process in animals are carried out by a single, multifunctional polypeptide fatty acid synthase (Type I FAS). In contrast, plants and bacteria utilise a dissociable multienzyme system (Type II FAS).[5] The structural differences between these systems are sufficient for the development of selective antibiotics targeted against the β-ketoacyl-acyl carrier protein synthase (KAS) and other individual enzymes of the Type II FAS.[6,7]

In plants and bacteria, there appear to be three KAS enzymes, denoted KAS I, KAS II and KAS III. The initial C2-C4 step is catalysed by KAS III; thereafter KAS I and KAS II are involved in chain elongation.[8] In most organisms with a Type II fatty acid synthase, KAS I and KAS II show very high similarity at the protein level. Recent discoveries acknowledge that *Plasmodium* synthesises fatty acids in the apicoplast which is a vestigial organelle thought to be derived from a chloroplast.[9-12] Not surprisingly, *Plasmodium* fatty acid synthase (FAS) is a Type II enzyme complex and, thus, differs markedly from human Type I FAS. Analysis of the recently published *P. falciparum* genome[3] reveals just two different KAS enzymes, with the KAS I and KAS II of typical Type II systems replaced by a single enzyme, that we and others, have denoted KAS1/2, and a separate KAS III.[8]

* Corresponding author. Fax: +44-29-2087-4149; e-mail: gilbertih@cf.ac.uk

0968-0896/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.11.023



1

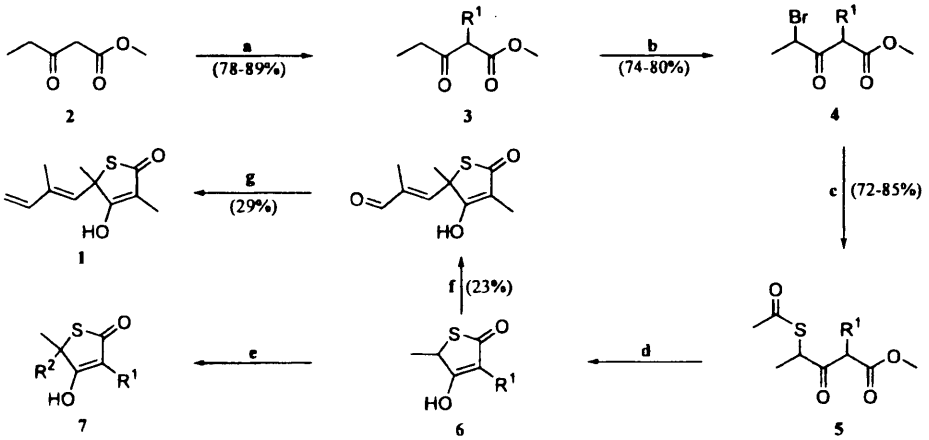
Figure 1. Structure of thiolactomycin.

A known inhibitor of the dissociable Type II FAS enzymes is the naturally occurring thiolactone antibiotic, thiolactomycin (TLM) **1**[13,14] (Fig. 1). Originally reported in 1982,[15] TLM has since become the most well-studied of the naturally occurring thiotetronic acids. The thiolactone has shown moderate in vitro activity against a broad spectrum of pathogens, including Gram-positive and Gram-negative[16] bacteria and *Mycobacterium tuberculosis*.[17] Furthermore, TLM analogues showed enhanced inhibition of pea (*Pisum sativum*) FAS,[18] and effective activity against *Staphylococcus aureus* and *Pasteurella multocida*.[19] Because of its negligible toxicity towards mammals,[15] TLM or its analogues are obvious candidates as new drugs for problem diseases such as malaria. In fact, recently Waller et al. have reported encouraging anti-malarial activity for TLM and various analogues.[20] Furthermore, because there are KAS isoforms in target organisms and each KAS isoform of Type II FAS complexes has been shown to be sensitive to TLM, then use of such inhibitors offers therapies with less chance of target-site resistance due to point mutations occurring readily. In this paper we present synthetic studies and findings in our attempts to improve the overall activity of TLM against *P. falciparum* cultured in red blood cells. Also studies against *Trypanosoma brucei rhodesiense* trypomastigotes and intracellular *Trypanosoma cruzi* amastigotes are reported.

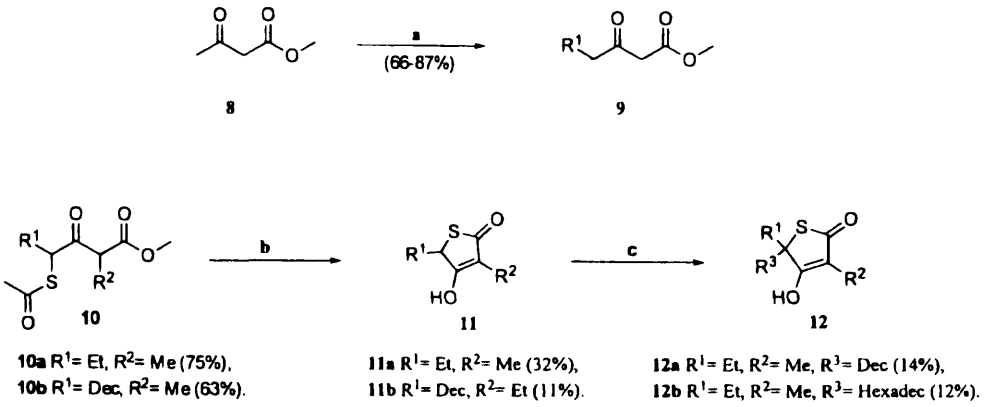
2. Chemistry

The first racemic total synthesis of TLM **1** was reported by Salvino et al. and involved the alkylation of a thiotetronic acid dianion with an isoprene cation equivalent.[21] A subsequent Wittig condensation with the resulting aldehyde afforded TLM **1**. The intermediate thiotetronic acid was prepared according to methodology developed by Benary.[22] Using a modification of this methodology we were able to synthesise TLM and a range of TLM analogues with varying chain length and saturation at the C5 and C3 positions of the ring (Scheme 1). Methyl propionylacetate **2** was treated with alkyl and benzyl halides in the presence of potassium carbonate as base to provide the alkylated products **3**. Treatment of the resulting β-ketoesters **3** with pyridinium tribromide in acetic acid afforded the bromides **4** as a mixture of diastereoisomers. Displacement of the bromides **4** proceeded smoothly using thiolacetic acid under basic conditions forming the thioacetates **5**. Cyclisation to the desired thiolactones **6** was accomplished upon treatment with aqueous potassium hydroxide solution. Isolated yields following flash chromatography were varied (30–93%). These derivatives were converted into the final compounds **7** after exposure to sodium hydride, butyllithium and the appropriate alkyl halides in dry THF at −78 °C. The alkylations tended to be generally low yielding; these findings are in agreement with previous publications in this area.[19] Attempts to protect the hydroxy functionality, in the hope that tying up of the free hydroxy group may improve the yields for the alkylation step, proved unsuccessful.

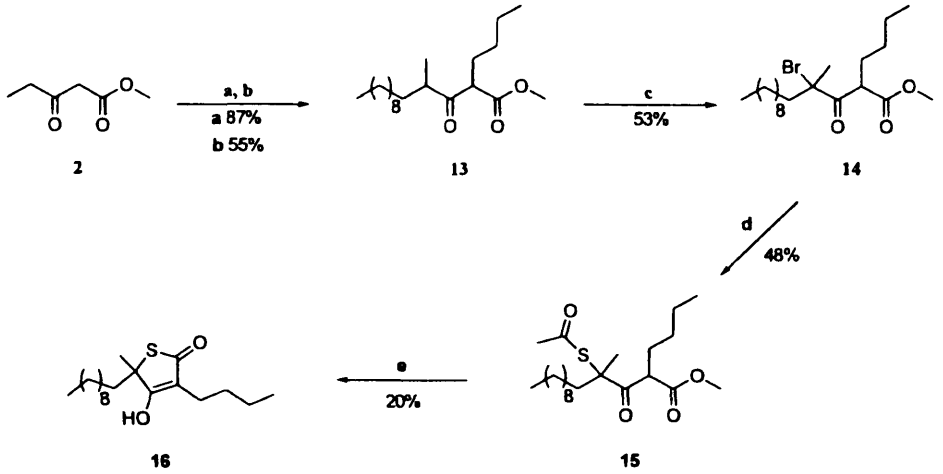
Derivatisation of the methyl substituent at the C5 position was accomplished as shown in Scheme 2. Following a literature procedure,[23] the β-ketoesters **9** were



Scheme 1. Reagents and conditions: (a) R^1X, K_2CO_3 (4 equiv), THF, reflux; (b) $PyBr_3$, AcOH, rt; (c) AcSH, Et_3N, EtOAc, rt; (d) KOH (2 equiv), H_2O/EtOH (1:1), 45 °C; (e) NaH (1.2 equiv), BuLi (1.1 equiv), R^2X, THF/DMPU (1:1), rt to −78 °C; (f) NaH, BuLi, $C_2H_5OCHC(CH_3)CHO$ (2 equiv), THF/Pyrimidone (1:1), rt to −78 °C; (g) BuLi (2 equiv), Ph_3P^+–CH_3Br,$^-$ THF, −78 °C to rt.



Scheme 2. Reagents and conditions: (a) NaH (1.2 equiv), BuLi (1.1 equiv), R^1I, 0 °C to rt; (b) KOH (2 equiv), H_2O/EtOH (1:1), 45 °C; (c) NaH (1.2 equiv), BuLi (1.1 equiv), R^3X, rt to −78 °C.



Scheme 3. Reagents and conditions: (a) NaH (1.2 equiv), BuLi (1.1 equiv), iododecane, 0 °C to rt; (b) iodobutane, K_2CO_3 (4 equiv), THF, reflux; (c) $PyBr_3$, AcOH, rt; (d) AcSH, Et_3N, EtOAc, rt; (e) KOH (2 equiv), H_2O/EtOH (1:1), 45 °C.

prepared in excellent yields by treatment of methyl acetoacetate **8** with sodium hydride, butyllithium and the relevant alkyl iodides. The corresponding β-ketoesters **10** were synthesised using the same methodology described in Scheme 1, and once formed were duly converted into the desired thiolactones **11** and **12**.

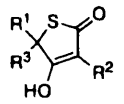
Using the same methodology described in Scheme 2, we were able to avoid the low yielding alkylation reaction at the C5 position with the construction of the desired carbon skeleton prior to cyclisation (Scheme 3). Subsequent alkylations of methyl propionylacetate **2** at the C4 and C2 positions respectively yielded the β-ketoester **13**. Bromination at the tertiary C4 position proceeded in reasonable yield (53%) forming the bromide **14**, which was duly converted to the thioacetate **15** in 48% yield. Treatment of the thioacetate **15** with potassium hydroxide delivered the derivitised thiolactone **16** (20%).

3. Results and discussion

All compounds were assayed in vitro for their activity against the malarial parasite, *P. falciparum* (Table 1). Furthermore, the compounds were also assayed against *T. brucei* and *T. cruzi* (Table 1). The screening procedures are briefly described in the Experimental (Section 5).

The basic structure of the thiolactomycin ring was retained with the analogues represented mainly by variations in the isoprenoid moiety and the methyl substituents at the C3 and C5 positions of the ring. Recently, thiolactomycin analogues have been tested in assays for fatty acid synthesis[18,20] and activities of the individual condensing enzymes.[17] The findings from these studies suggest that replacement of the isoprenoid side chain with longer tethers leads to an overall increase in inhibition of fatty acid synthesis. This notable trend is also evident in our own findings (Table 1). The analogues **7a** and **7b** showed inhibition comparable to that of thiolactomycin **1**. In comparison to **1**, the geranyl analogue **7c** showed almost a fourfold increase in activity, perhaps indicating that unsaturation in the hydrophobic chain enhances inhibition. Waller et al.[20] have reported a sixfold increase in efficacy for the geranyl analogue over TLM against the *P. falciparum* multidrug-resistant strain W2mef. The longer side-chain analogues **7d** (decyl) and **7e** (hexadecyl) also showed much improved activity over **1**, with **7e** registering a 6-fold (25 μM) increase in inhibition. The influence of the longer hydrophobic chain on activity is further

Table 1. Activities of compounds against *P. falciparum* cultured in red blood cells; *T. brucei* trypomastigotes; and *T. cruzi* amastigotes cultured in rat skeletal myoblasts



Compd	R^1	R^2	R^3	*P. falciparum* IC_{50} (μM)	*T. cruzi* IC_{50} (μM)	*T. brucei* IC_{50} (μM)
1	Me	Me	Isoprenoid	143	>427	256
6a	Me	Me	H	>347	>624	>624
6b	Me	Et	H	233	568	357
6c	Me	Pr	H	>290	>522	288
6d	Me	Bn	H	195	>408	408
7a	Me	Me	Hexyl	139	127	112
7b	Me	Me	Octyl	153	>350	194
7c	Me	Me	Geranyl	40	68	190
7d	Me	Me	Decyl	36	68	171
7e	Me	Me	Hexadecyl	25	28	62
7f	Me	Et	Hexyl	61	195	71
7g	Me	Et	Octyl	54	165	157
7h	Me	Et	Decyl	15	54	97
7i	Me	Et	Hexadecyl	19	43	47
7j	Me	Pr	Decyl	10	56	21
7k	Me	Bn	Decyl	50	64	132
11a	Et	Me	H	>316	>568	>568
11b	Decyl	Et	H	72	164	6
12a	Et	Me	Decyl	65	72	153
12b	Et	Me	Hexadecyl	35	49	130
16	Me	Bu	Decyl	71	77	8

Standards: For *P. falciparum*, chloroquine, IC_{50} = 0.126 μM; artemisinin, IC_{50} = 0.005 μM ; For *T. brucei*, melarsoprol, IC_{50} = 0.005 μM ; For *T. cruzi*, benznidazole, IC_{50} = 1.68 μM. The IC_{50} values are the means of four values of two independent assays done in duplicate. They were determined by linear interpolation between the two adjacent drug concentrations above and below the 50% incorporation line.

emphasised by considering the intermediate compounds **6a–d** and **11a**, all of which showed a marked decrease in inhibition (195 to >347) in comparison to **1**.

Recently, Price et al. have solved the crystal structure of TLM bound to *Escherichia coli* KAS I.[24,25] From this study a number of key interactions have been identified involving the thiolactone ring with the protein's active site. Specifically, the carbonyl oxygen at the C2 position binds with two histidines (His 298 and His 333). The isoprenoid moiety appears to occupy a hydrophobic crevice and is sandwiched between two pairs of amino acids (Gly 391 and Phe 392; Ala 271 and Pro 272). Interestingly, alignment of *P. falciparum* KAS1/2 with the *E. coli* KAS I indicates that these four amino acids are retained in the *P. falciparum* sequence (Gly 403 and Phe 404; Ala 274 and Pro 275).[25] Price et al.,[25] further observed that this crevice was not optimally filled by the isoprenoid side chain. Our findings lend support to this observation, and would suggest that alkyl chains of ten carbons and above better occupy this cleft, therefore increasing inhibition.

Price et al.[25] made a detailed study of the crystal structure of *E. coli* KAS I, and found that the two ring methyl substituents, at the C3 and C5 positions, are accommodated in hydrophobic pockets (the C3 methyl in a pocket bounded by Phe 229, Phe 392 and Thr 300; the C5 pocket by Pro 272 and Gly 305). The analogues **7f–i**, all possessing an ethyl substituent at the C3 position of the ring, all showed greater inhibition than TLM against *P. falciparum*. Once again, the longer chained analogues **7h** and **7i** gave the best inhibition, showing 10 and 8-fold increases in inhibition over TLM respectively. A further trend was also observed, analogues **7f–i** all displayed enhanced inhibition towards the parasite in comparison to analogues **7a–e** bearing a methyl substituent at C3 of the ring. This trend is highlighted when comparing the activities of analogues **7b** (153 μM) and **7g** (54 μM). Compound **7g** shows a 3-fold improvement in inhibition over **7b**. However with the more active compounds, there is a much smaller improvement in activity on addition of an ethyl group at the C3 position [compare **7h** (15 μM) with **7d** (36 μM); and **7i** (19 μM) with **7e** (25 μM)]. Clearly, the ethyl substituent is having a positive influence on the overall activity of the inhibitors. This led us to probe the C3 position further. Indeed, by extending the ethyl chain to a propyl group, and placing a decyl side chain at the C5 position (analogue **7j**) an IC_{50} of 10 μM was achieved, thus giving a 14-fold increase in activity over TLM. Increasing the propyl tether to a butyl derivative **16** resulted in a decrease in activity. A similar result was also obtained with the benzyl analogue **7k**, which showed an IC_{50} of 50 μM. Seemingly, it is plausible that we have reached an optimal chain length of three carbons at the C3 position, beyond which activity is diminished. To lend support to this assertion we were able to dock analogue **7j** into the *E. coli* KAS I using FlexX software (Figs 2 and 3). The propyl group appears to fit snugly into the hydrophobic pocket with the terminus of the propyl chain in close proximity to the Cys 163. This same amino acid is also present in the *P. falciparum* KAS1/2.[25]

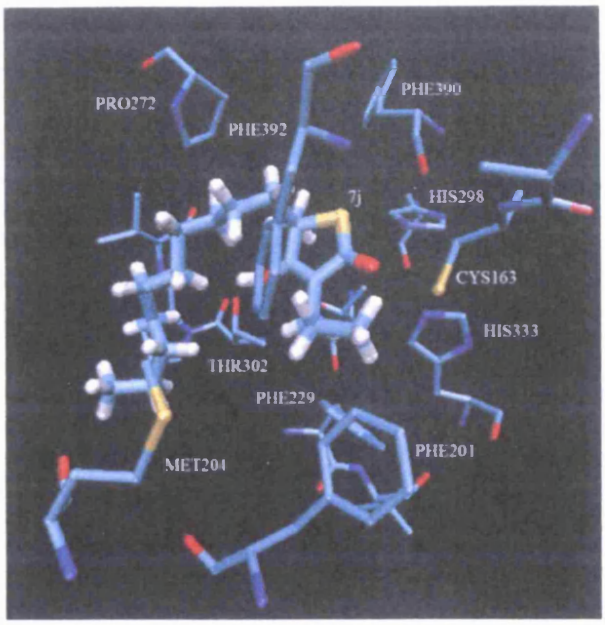


Figure 2. Compound 7j docked into the *E. coli* KAS I. Part of the 3-propyl group is obscured.

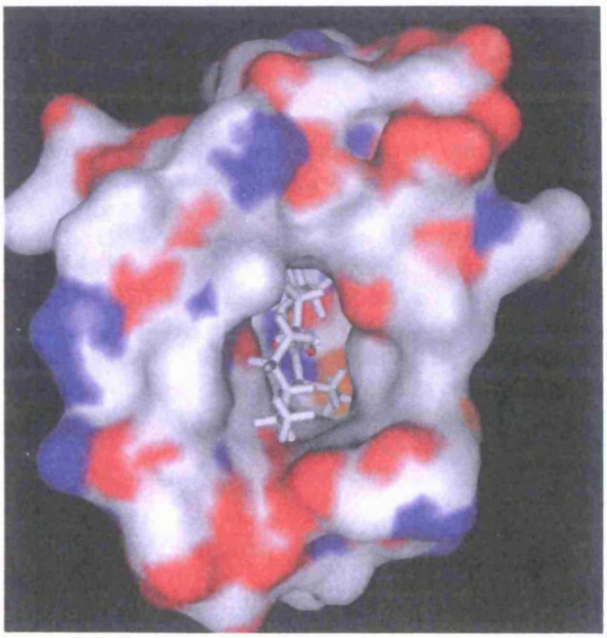


Figure 3. Inhibitor 7j in relation to the surface of *E. coli* KAS I.

In an attempt to improve the hydrophobic interactions of the methyl substituent at the C5 position, an ethyl substituent was introduced in place of the methyl group. Both analogues **12a** (65 μM) and **12b** (35 μM) showed improved inhibition over TLM (Table 1). However, a direct comparison of these analogues with there methyl counterparts [compare **12a** (65 μM) with **7d** (36 μM); and **12b** (35 μM) with **7e** (25 μM)] suggests that the presence of the ethyl substituent at C5 serves only to reduce the activity of the inhibitors. Likewise, removal of the C5 methyl substituent (analogue **11b**) also appears to reduce the inhibitory effects of the analogue. This is highlighted with a direct comparison of the activities for analogues **11b** (72 μM) and **7h** (15 μM). Clearly, modification of the methyl substituent at the C5 position does not appear as influential to the overall activity of the analogues as modification at the C3 position.

All our compounds were also assayed in vitro against the *T. brucei* parasite, which is the cause of human sleeping sickness and livestock disease in Africa. Recently, Englund[26] identified thiolactomycin as a promising lead for antitrypanosomal drug development. *T. brucei* requires large amounts of myristate for synthesis of glycosyl phosphatidylinositol anchored variable surface glycoprotein. Thiolactomycin was shown to effectively inhibit trypanosomal myristate synthesis in vitro with a reported IC_{50} of ~150 μM.[26] A number of our compounds showed inhibition greater than thiolactomycin **1** (IC_{50} 256 μM in our assays), with the best two [analogues **11b** (6 μM) and **16** (8 μM)] showing a 42-fold improvement in activity (Table 1). The most active analogue (**7j**) against the malaria parasite also showed improved inhibition against *T. brucei*, registering an IC_{50} of 21 μM.

In vitro assays against *T. cruzi*, the causative organism of Chagas disease, showed that thiolactomycin **1** was inactive with an IC_{50} of >427 μM (Table 1). Several of the analogues tended to show enhanced activity towards the *T. cruzi* parasite over thiolactomycin. Analogue **7e** showed the greatest inhibition with an IC_{50} of 28 μM.

Finally, none of the compounds showed any appreciable activity against the *Leishmania donovani* parasite. The lack of activity against *Leishmania* may be due to the parasite being an intracellular parasite dwelling within a parasitophorous vacuole within the host cell, significantly reducing accessibility of compounds.

4. Conclusion

We have reported the design and synthesis of a number of thiolactomycin analogues as agents against *P. falciparum*, a causative organism of malaria. The compounds were designed with the intention of targeting the Type II FAS utilised for fatty acid biosynthesis. Several analogues showed inhibitory effects greater than thiolactomycin **1**, with the best compound (**7j**) showing a 14-fold increase in activity (10 μM). Furthermore, all compounds were assayed against three other parasitic protozoa, *T. cruzi*, *T. brucei* and *L. donovani*. Varied activities were found against the *T. brucei* and *T. cruzi* parasites. However, the analogues showed no appreciable activity against *L. donovani*. Although the inhibitory effects of these analogues are only in the micromolar range, these show significantly better activity than thiolactomycin. This study has demonstrated the scope that exists for the design of new fatty acid biosynthesis inhibitors with enhanced activity against the parasites which cause malaria, African trypanosomiasis and Chagas disease. Significantly improved inhibitors could offer a valuable new way of controlling these important diseases.

5. Experimental

Melting points were determined using a Gallenkamp melting point apparatus and are uncorrected. Infrared spectra were recorded as thin films for liquid samples, or as Nujol mulls for solid samples, on a Perkin–Elmer 1600 series FTIR spectrophotometer using sodium chloride plates. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker Advance DPX300 spectrometer operating at 300 and 75 M Hz, respectively, with tetramethylsilane as internal standard, using deuterated chloroform purchased from Goss unless stated otherwise. Low-resolution mass spectra, that is, electrospray [ES], were recorded using a fisons VG Platform II spectrometer. High-resolution spectra were obtained on a VG ZAB spectrometer from the EPSRC Mass Spectrometry Service at Swansea University, UK. Microanalyses were obtained from the analytical and chemical consultancy services MEDAC LTD. All reactions were performed in predried apparatus under an atmosphere of nitrogen unless otherwise stated. Solvents and reagents were purchased from chemical companies and used without further purification. Dry solvents were generally purchased in sure sealed bottles stored over molecular sieves. Thin-layer chromatography (tlc) was performed on Merck silica gel $60F_{254}$ plates. Column chromatography was carried out using Fisons matrix silica 60 (35–70 micron).

5.1. Cyclisation, general procedure A

Potassium hydroxide (2.0 equiv) in water (10 mL) was added dropwise to a stirred solution of the thioacetate (1.0 equiv) in ethanol (20 mL) at ambient temperature. The resulting solution was warmed to 40 °C and stirred for 3 h. The ethanol was then evaporated and water (~10 mL) added. The aqueous layer was washed with diethyl ether (2×10 mL) and acidified to ph 1 with the addition of 2 M HCl (~10 mL). The aqueous layer was extracted with diethyl ether. The organic layers were washed with brine, dried over anhydrous magnesium sulfate and the solvent removed in vacuo. The crude residue was purified by flash column chromatography (5–20% ethyl acetate in hexanes). The coupling constants (J) are in Hz.

5.2. 4-Hydroxy-3,5-dimethyl-2(5H)-thiophenone (6a)

As described in procedure A, starting from **5a** (1.0 g, 4.58 mmol) and potassium hydroxide (0.5 g, 9.16 mmol), **6a** was obtained as a white solid (0.60 g, 90%); mp 131–132 °C, ν_{MAX}/cm^{-1} 3197, 2923, 2878, 1555, 1291 and 1186; δ_H (acetone-d_6) 1.66 (3H, d, $J=7.0$, 5-CH_3), 1.74 (3H, s, 3-CH_3), 4.32 (1H, q, $J=7.0$, 5-H) and 10.22 (1H, br s, OH); δ_C (acetone-d_6) 12.6 (3-CH_3), 24.5 (5-CH_3), 47.9 (4-CH), 115.5 (3-C), 183.6 (4-C) and 199.3 (CO); MS (ES^-) m/z 142.6 ($M-H,^-$ 100%); HRMS (ES^+) $(M+H)^+$ $C_6H_9O_2S$ requires 145.0323, found 145.0321. Anal. calcd for $C_6H_8O_2S.0.1H_2O$: C, 49.4; H, 5.7. Found: C, 49.3; H, 5.4%.

5.3. 3-Ethyl-4-hydroxy-5-methyl-2(5H)-thiophenone (6b)

As described in procedure A, starting from **5b** (3.52 g, 7.58 mmol) and potassium hydroxide (1.70 g, 15.17 mmol), **6b** was obtained as a white solid (1.07 g, 89%); mp 75–76 °C, ν_{MAX}/cm^{-1} 3188, 2923, 2869, 1563, 1291 and 1184; δ_H (acetone-d_6) 1.94 (3H, t, $J=7.2$, 2′-CH_3), 1.53 (3H, d, $J=7.0$, CH_3), 2.18 (2H, q, $J=7.2$, 1′-CH_2), 4.21 (1H, q, $J=7.0$, 5-CH) and 10.11 (1H, br s, OH); δ_C (acetone-d_6) 13.4 (2′-CH_3), 17.1 (1′-CH_2), 20.2 (5-CH_3), 43.4 (5-C), 117.1 (3-C), 179.2 (4-C) and 194.8 (CO); MS (ES^-) m/z 156.7 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_7H_9O_2S$ requires 157.0324, found 157.0323. Anal. calcd for $C_7H_{10}O_2S.0.1H_2O$: C, 52.5; H, 6.4. Found: C, 52.6; H, 6.4%.

5.4. 4-Hydroxy-5-methyl-3-propyl-2,5-dihydro-2-thiophene (6c)

As described in procedure A, starting from **5c** (3.50 g, 14.22 mmol) and potassium hydroxide (1.59 g, 28.44 mmol), **6c** was obtained as a colourless solid (1.04 g, 42%); mp 49–50 °C, δ_H (acetone-d_6) 0.91 (3H, t, $J=7.4$, 3′-CH_3), 1.41–1.54 (2H, m, 2′-CH_2), 1.65 (3H, d, $J=7.0$, CH_3), 2.22 (2H, t, $J=7.5$, 1′-CH_2) and 4.29 (1H, q, $J=7.0$, 5-CH); δ_C (acetone-d_6) 18.5 (3′-CH_3), 24.5 (CH_3), 26.6 (2′-CH_2), 29.8 (1′-CH_2), 47.6 (5-CH), 119.8 (3-C), 183.9 (4-C) and 199.1 (2-CO); MS (ES^-) m/z 170.9 ($M-H,^-$ 100%); HRMS (ES^+) $(M+NH_4)^+$ $C_8H_{12}O_2S$ requires 190.0896, found 190.0895. Anal. calcd for $C_8H_{12}O_2S$: C, 55.8; H, 7.0. Found: C, 55.4; H, 7.1%.

5.5. 3-Benzyl-4-hydroxy-5-methyl-2,5-dihydro-2-thiophene (6d)

As described in procedure A, starting from **5d** (6.00 g, 20.40 mmol) and potassium hydroxide (2.69 g, 40.80 mmol), **6d** was obtained as a white powder (1.32 g, 29%); mp 139–141 °C, δ_H (acetone-d_6) 1.59 (3H, d, $J=6.9$, CH_3), 3.50 (2H, s, CH_2Ph), 4.28 (1H, q, $J=6.9$, 5-CH), 7.08–7.21 (5H, m, ArH) and 10.33 (1H, br s, OH); δ_C (acetone-d_6) 20.2 (CH_3), 29.3 (CH_2Ph), 43.6 (5-CH), 115.0 (3-C), 127.1 (ArCH), 129.4 (ArCH), 129.5 (ArCH), 140.9 (ArC), 180.2 (4-C) and 194.6 (CO); MS (ES^-) m/z 219.5 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{12}H_{11}O_2S$ requires 219.0480, found 219.0484. Anal. calcd for $C_{12}H_{12}O_2S.0.1H_2O$: C, 64.9; H, 5.5. Found: C, 64.7; H, 5.5%.

5.6. 5-Ethyl-4-hydroxy-3-methyl-2,5-dihydro-2-thiophene (11a)

As described in procedure A, starting from **10a** (5.36 g, 23.10 mmol) and potassium hydroxide (2.59 g, 46.20 mmol), **11a** was obtained as a pale yellow solid (1.15 g, 32%); mp 95–96 °C, δ_H (acetone-d_6) 0.89 (3H, t, $J=7.4$, 2″-CH_3), 1.58 (3H, s, 1′-CH_3), 1.61–1.68 (1H, m, 1″-CH_AH_B), 2.07–2.17 (1H, m, 1″-CH_AH_B) and 4.14 (1H, dd, $J=6.2$ and 1.1, 5-CH); δ_C (acetone-d_6) 8.1 (7-CH_3), 11.4 (3-CH_3), 27.0 (6-CH_2), 50.8 (5-CH), 112.0 (3-C), 178.0 (4-C) and 195.2 (2-CO); MS (ES^-) m/z 156.8 ($M-H,^-$ 100%); HRMS (ES^+) $(M+NH_4)^+$ $C_7H_{14}NO_2S$ requires 176.0740; found 176.0740. Anal. calcd for $C_7H_{10}O_2S$: C, 53.1; H, 6.4. Found: C, 52.7; H, 6.4%.

5.7. 5-Decyl-3-ethyl-4-hydroxy-2(5H)-thiophenone (11b)

As described in procedure A, starting from **10b** (870 mg, 2.43 mmol) and potassium hydroxide (270 mg, 4.86 mmol), **11b** was obtained as a pale yellow solid (74 mg, 11%), mp 35–36 °C, δ_H 0.88 (3H, t, J = 6.8, 15-CH_3), 1.04 (3H, t, J = 7.4, 2′-CH_3), 1.26 (14H, m, 7×CH_2), 1.67–1.73 (2H, m, CH_2), 2.24–2.31 (4H, m, 6-CH_2 and 1′-CH_2) and 4.18 (1H, dd, J = 9.3 and 2.6, 5-CH); δ_C 13.1 (15-CH_3), 14.5 (2′-CH_3), 16.5, 23.1, 27.9, 29.6, 29.7, 29.8, 30.0, 32.3, 33.3 (all CH_2), 50.1 (5-CH), 118.0 (3-C), 179.2 (4-C) and 199.9 (2-CO); MS (ES^-) m/z 283.1 ($M-H,^-$ 100%); HRMS $(ES)^+$ $(M+H)^+$ $C_{16}H_{29}O_2S$ requires 285.1883, found 285.1892. Anal. calcd for $C_{16}H_{28}O_2S.0.2H_2O$: C, 66.7; H, 9.9. Found: C, 66.8; H, 10.1%.

5.8. 3-Butyl-5-decyl-4-hydroxy-5-methyl-2(5*H*)-thiophenone (16)

As described in procedure A, starting from **15** (2.84 g, 7.39 mmol) and potassium hydroxide (0.82 g, 14.79 mmol), **16** was obtained as a pale yellow oil (0.47 g, 20%), δ_H 0.89–0.93 (6H, m, 15-CH_3 and 4′-CH_3), 1.23–1.37 (20H, m, 10×CH_2), 1.72 (3H, s, CH_3), 1.93–1.96 (2H, m, 6-CH_2) and 2.29 (2H, t, J = 7.2, 1′-CH_2); δ_C 14.1, 14.3 (both CH_3), 22.8, 22.9, 23.1, 24.4, 25.4, 25.5, 26.4, 26.6, 29.7, 29.8, 29.9, 30.0, 32.0, 38.7 (all CH_2), 58.4 (5-C), 115.6 (3-C), 181.7 (4-C) and 198.6 (2-CO); MS (ES^-) m/z 325.4 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{19}H_{33}O_2S$ requires 325.2196, found 325.2199.

5.9. Alkylation, general procedure B

Thiophenone (1.0 equiv) was added portionwise to a stirred suspension of sodium hydride (1.2 equiv) in dry THF (4 mL) and DMPU (4 mL) at ambient temperature. The suspension was stirred for 0.5 h, then cooled to −78 °C and butyllithium (2.5 M solution in hexanes, 1.1 equiv) added dropwise. After stirring for 10 min, the mixture was allowed to warm to ambient temperature and stirred for 1 h. Once again the mixture was cooled to −78° and the halide (1.0 equiv) added dropwise. After stirring for 15 min, the mixture was allowed to warm to ambient temperature and stirred overnight. The reaction was quenched by the addition of 2 M HCl (10 mL) and the organic layer separated. The aqueous layer was then extracted with ethyl acetate (3×10 mL). The organic layers were washed with brine, dried over anhydrous magnesium sulfate and the solvent removed in vacuo. The crude residue was purified by flash column chromatography (0–20% ethyl acetate in hexanes).

5.10. Thiolactomycin 1

Butyllithium (1.6 mL of a 2.5 M solution in hexanes, 4.01 mmol) was added dropwise to a stirred suspension of the phosphonium salt (710 mg, 2.00 mmol) in dry THF (8 mL) at −78 °C. The resulting dark red solution was allowed to warm to ambient temperature and stirred for 1 h. The solution was then cooled to −78 °C and the aldehyde (430 mg, 2.00 mmol) added dropwise. A light brown suspension formed, which was then allowed to warm to ambient temperature and stirred overnight. 2 M HCl (10 mL) was added and the organic layer separated. The aqueous layer was extracted with ethyl acetate (3×10 mL). The combined organic solutions were washed with water (2×20 mL) and brine (2×20 mL), then dried and evaporated. Purification by column chromatography (20% ethyl acetate in hexanes) provided thiolactomycin **1** as a pale yellow oil (12 mg, 29% yield), ν_{MAX}/cm^{-1} 3600, 1700, 1630, 1450, 1380, 1325, 1280 and 1100; ^1H NMR δ 1.79 (3H, s, CH_3), 1.83 (3H, s, CH_3), 1.91 (3H, s, CH_3), 5.15 (1H, d, J = 11.1, 4′-H_a), 5.33 (1H, d, J = 17.2, 4′-H_b), 5.62 (1H, s, 1′-H) and 6.37 (1H, dd, J = 17.2 and 11.1, 3′-H); δ_C 7.6 (3-CH_3), 12.1 (CH_3), 29.9 (5-CH_3), 58.6 (5-C), 110.8 (3-C), 114.0 (CH_2), 129.8, 140.6, 141.1, 181.3 (4-C) and 198.6 (CO). Spectroscopic data was identical to those recorded in the literature.[21]

5.11. 5-Hexyl-4-hydroxy-3,5-dimethyl-2(5H)-thiophenone (7a)

As described in procedure B, starting from **6a** (500 mg, 3.46 mmol), sodium hydride (160 mg, 4.16 mmol), butyllithium (1.52 mL, 3.80 mmol) and bromohexane (0.48 mL, 3.46 mmol), **7a** was obtained as a colourless solid (80 mg, 10%); mp 62–64 °C, ν_{MAX}/cm^{-1} 3145, 2929, 1600, 1454, 1247 and 1010; δ_H 0.87 (3H, t, J = 6.5, 6′-CH_3), 1.26–1.34 (8H, m, 4×CH_2), 1.66 (3H, s, CH_3), 1.74 (3H, s, CH_3) and 1.85 (2H, t, J = 8.0, 1′-CH_2); δ_C 7.92 (6′-CH_3), 14.4 (3-CH_3), 22.9 (CH_2), 25.5 (5-CH_3), 26.4, 29.6, 32.0, 38.9 (all CH_2), 58.0 (5-C), 110.9 (3-C), 183.6 (4-C) and 198.4 (CO); MS (ES^-) m/z 227.2 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{12}H_{19}O_2S$ requires 227.1106, found 227.1106. Anal. calcd for $C_{12}H_{20}O_2S.0.2H_2O$: C, 62.1; H, 8.9. Found: C, 61.9; H, 8.9%.

5.12. 4-Hydroxy-3,5-dimethyl-5-octyl-2(5H)-thiophenone (7b)

As described in procedure B, starting from **6a** (500 mg, 3.46 mmol), sodium hydride (160 mg, 4.16 mmol), butyllithium (1.52 mL, 3.80 mmol) and 1-iodooctane (1.05 mL, 3.46 mmol), **7b** was obtained as a pale yellow solid (108 mg, 12%); mp 81–83 °C, ν_{MAX}/cm^{-1} 1611 and 1076; δ_H 0.97 (3H, t, J = 6.9, 8′-CH_3), 1.31 (10H, m, 5×CH_2), 1.75 (3H, s, CH_3), 1.83 (3H, s, CH_3) and 1.89–1.94 (2H, m, 1′-CH_2); δ_C 7.9 (8′-CH_3), 14.5 (3-CH_3), 23.0 (CH_2), 25.6 (5-CH_3), 26.3, 29.6, 29.8, 29.9, 32.2, 38.8 (all CH_2), 58.5 (5-C), 110.8 (3-C), 181.2 (4-C) and 198.3 (CO); MS (ES^-) m/z 255.2 ($M-H,^-$ 100%); HRMS (ES^+) $(M+H)^+$ $C_{14}H_{25}O_2S$ requires 257.1570, found 257.1568. Anal. calcd for $C_{14}H_{24}O_2S.0.1H_2O$: C, 65.1; H, 9.4. Found: C, 65.1; H, 9.6%. Spectroscopic data was identical to those recorded in the literature.[17]

5.13. 5-[(2*E*)-3,7-Dimethyl-2,6-octadienyl]-4-hydroxy-3,5-dimethyl-2(5H)-thiophenone (7c)

As described in procedure B, starting from **6a** (500 mg, 3.46 mmol), sodium hydride (160 mg, 4.16 mmol), butyllithium (1.52 mL, 3.80 mmol) and geranyl bromide (0.69 g, 3.46 mmol), **7c** was obtained as a thick yellow

oil (160 mg, 16%); ν_{MAX}/cm^{-1} 2972, 2927, 1600, 1450, 1383, 1280 and 969; δ_H 1.61 (3H, s, CH_3), 1.64 (3H, s, CH_3), 1.66 (3H, s, CH_3), 1.69 (3H, s, CH_3), 1.75 (3H, s, CH_3), 2.04–2.12 (4H, m, 2×CH_2), 2.52–2.69 (2H, m, CH_2), 5.08 (1H, m, CH) and 5.18 (1H, m, CH); δ_C 7.9 (CH_3), 16.9 (3-CH_3), 18.1 (CH_3), 25.2 (5-CH_3), 26.1 (CH_3), 26.8 (CH_2), 37.7, 40.2 (both CH_2), 58.2 (5-C), 110.7 (3-C), 119.2, 124.3 (both CH), 132.2, 140.5 (both C), 180.2 (4-C) and 203.2 (CO); MS (ES^-) *m/z* 279.1 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{16}H_{23}O_2S$ requires 279.1419, found 279.1416. Spectroscopic data was identical to those recorded in the literature.[17]

5.14. 5-Decyl-4-hydroxy-3,5-dimethyl-2(5H)-thiophenone (7d)

As described in procedure B, starting from **6a** (500 mg, 3.46 mmol), sodium hydride (160 mg, 4.16 mmol), butyllithium (1.52 mL, 3.80 mmol) and 1-iodododecane (0.73 mL, 3.46 mmol), **7d** was obtained as a white solid (110 mg, 10%); mp 74–75 °C, ν_{MAX}/cm^{-1} 3130, 2947, 1592, 1486, 1270, 1170 and 989; δ_H 0.82 (3H, t, $J=6.6$, 10′-CH_3), 1.20 (16H, m, 8×CH_2), 1.62 (3H, s, CH_3), 1.71 (3H, s, CH_3) and 1.80–1.85 (2H, m, 1′-CH_2); δ_C 7.9 (10′-CH_3), 14.5 (3-CH_3), 23.1 (CH_2), 25.6 (5-CH_3), 26.3, 29.7, 29.8, 29.9, 30.0, 32.3, 38.8 (all CH_2), 58.5 (5-C), 110.8 (3-C), 181.1 (4-C) and 198.3 (2-CO); MS (ES^-) *m/z* 283.0 ($M-H,^-$ 100%); HRMS (ES^+) $(M+NH_4)^+$ $C_{16}H_{32}NO_2S$ requires 302.2154, found 302.2149. Anal. calcd for $C_{16}H_{28}O_2S.0.2H_2O$: C, 66.7; H, 9.9. Found: C, 66.7; H, 9.9%.

5.15. 5-Hexadecyl-4-hydroxy-3,5-dimethyl-2(5H)-thiophenone (7e)

As described in procedure B, starting from **6a** (500 mg, 3.46 mmol), sodium hydride (160 mg, 4.16 mmol), butyllithium (1.52 mL, 3.80 mmol) and 1-bromohexadecane (1.05 mL, 3.46 mmol), **7e** was obtained as a white solid (87 mg, 7%), mp 56–57 °C, ν_{MAX}/cm^{-1} 2917, 2848, 1573, 1284, 1226 and 1145; δ_H 0.98 (3H, t, $J=6.9$, 16′-CH_3), 1.36 (28H, m, 14×CH_2), 1.77 (3H, s, CH_3), 1.85 (3H, s, CH_3) and 1.95 (2H, t, $J=7.9$, 1′-CH_2); δ_C 7.9 (16′-CH_3), 14.5 (3-CH_3), 23.1 (CH_2), 25.6 (5-CH_3), 26.3, 29.7, 29.8, 30.0, 30.1, 32.3, 38.9 (all CH_2), 58.2 (5-C), 110.9 (3-C), 180.1 (4-C) and 190.3 (2-CO); MS (ES^-) *m/z* 367.2 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{22}H_{39}O_2S$ requires 367.2671, found: 367.2666. Anal. calcd for $C_{22}H_{40}O_2S$: C, 71.7; H, 10.9. Found: C, 71.7; H, 11.2%.

5.16. 3-Ethyl-5-hexyl-4-hydroxy-5-methyl-2(5H)-thiophenone (7f)

As described in procedure B, starting from **6b** (400 mg, 2.53 mmol), sodium hydride (120 mg, 3.03 mmol), butyllithium (1.77 mL of a 1.57 M solution in hexanes, 2.78 mmol) and 1-iodohexaane (0.37 mL, 2.53 mmol), **7f** was obtained as a colourless oil (100 mg, 16%), δ_H 0.82 (3H, t, $J=6.8$, CH_3), 0.98 (3H, t, $J=7.5$, 2′-CH_3), 1.21 (8H, m, 4×CH_2), 1.62 (3H, s, CH_3), 1.80–1.83 (2H, m, CH_2) and 2.19 (2H, q, $J=7.5$, 1′-CH_2); δ_C 13.3 (CH_3), 14.4 (CH_3), 16.5 (CH_3), 23.0, 25.4, 26.3, 29.5, 31.9 and 38.8 (all CH_2), 58.1 (5-C), 116.9 (3-C), 180.4 (4-C) and 197.5 (2-CO); MS (ES^-) *m/z* 240.9 ($M-H,^-$ 100%). HRMS (ES^-) $(M-H)^-$ $C_{13}H_{21}O_2S$ requires 241.1262, found 241.1259. Anal. calcd for $C_{13}H_{22}O_2S.0.2H_2O$: C, 63.5; H, 9.2. Found: C, 63.6; H, 9.3%.

5.17. 3-Ethyl-4-hydroxy-5-methyl-5-octyl-2(5H)-thiophenone (7g)

As described in procedure B, starting from **6b** (300 mg, 1.89 mmol), sodium hydride (91 mg, 2.27 mmol), butyllithium (1.31 mL of a 1.57 M solution in hexanes, 2.07 mmol) and 1-iodooctane (0.31 mL, 1.89 mmol), **7g** was obtained as a pale yellow oil (40 mg, 18%); δ_H 0.92 (3H, t, $J=6.7$, 8″-CH_3), 1.09 (3H, t, $J=7.5$, 2″-CH_3), 1.30 (12H, m, 6×CH_2), 1.72 (3H, s, CH_3), 1.83–1.86 (2H, m, 1′-CH_2), 2.28 (2H, q, $J=7.5$, 1′-CH_2) and 7.68 (1H, br s, OH); δ_C 13.3, 14.5 and 16.4 (all CH_3), 23.0, 25.4, 26.4, 29.5, 29.6, 29.9, 32.2 and 38.8 (all CH_2), 57.8 (5-C), 116.9 (3-C), 179.9 (4-C) and 196.7 (5-C); MS (ES^-) *m/z* 269.1 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{15}H_{25}O_2S$ requires 269.1576, found 269.1578.

5.18. 5-Decyl-3-ethyl-4-hydroxy-5-methyl-2(5H)-thiophene (7h)

As described in procedure B, starting from **6b** (300 mg, 1.89 mmol), sodium hydride (91 mg, 2.27 mmol), butyllithium (1.31 mL of a 1.57 M solution in hexanes, 2.07 mmol) and 1-iododecane (0.40 mL, 1.89 mmol), **7h** was obtained as a colourless solid (125 mg, 22%), mp 36–38 °C, δ_H 0.93 (3H, t, $J=6.8$, 10″-CH_3), 1.10 (3H, t, $J=7.5$, 2′-CH_3), 1.30 (14H, m. 7×CH_2), 1.71 (3H, s, OCH_3), 1.89 (2H, m, CH_2), 2.00 (2H, m, CH_2), 2.25 (2H, q, $J=7.5$, 1′-CH_2) and 6.85 (1H, br s, OH); δ_C 13.3, 14.5 and 16.5 (all CH_3), 23.1, 25.5, 26.4, 29.6, 29.7, 29.8, 29.9, 30.0, 32.3 and 38.8 (all CH_2), 57.9 (5-C), 117.0 (3-C), 179.7 (4-C) and 198.2 (2-C); MS (ES^-) *m/z* 297.3 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{17}H_{29}O_2S$ requires 297.1889, found 297.1888. Anal. calcd for $C_{17}H_{30}O_2S$: C, 68.4; H, 10.1. Found: C, 68.1; H, 10.3%.

5.19. 3-Ethyl-5-hexadecyl-4-hydroxy-5-methyl-2(5H)-thiophenone (7i)

As described in procedure B, starting from **6b** (300 mg, 1.89 mmol), sodium hydride (91 mg, 2.27 mmol), butyllithium (1.31 mL of a 1.57 M solution in hexanes, 2.07 mmol) and 1-bromohexadecane (0.31 mL, 1.89 mmol), **7i** was obtained as a colourless solid (160 mg, 22%), mp 86–87 °C, ν_{MAX}/cm^{-1} 3183, 2916, 2848, 1580, 1471, 1330, 1251 and 1145; δ_H 1.00 (3H, t, $J=6.6$, 16″-CH_3), 1.16 (3H, t, $J=7.5$, 2′-CH_3), 1.37 (28H, m, 14×CH_2), 1.80 (3H, s, 5-CH_3), 1.97–2.01 (2H, m, 1″-CH_2), 2.36 (1H, q, $J=7.5$, 1′-CH_2) and 8.04 (1H, br s, OH); δ_C 13.3, 14.5, 16.5, 20.7, 23.1, 25.5, 26.4, 29.8, 29.8, 29.9, 29.9, 30.0, 30.1, 32.3, 38.9, 57.9 (5-C), 117.1 (3-C), 179.5 (4-C) and 196.9 (2-CO); MS (ES^-) *m/z* 381.2 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{23}H_{41}O_2S$ requires 381.2828, found 381.2835. Anal. calcd for $C_{23}H_{42}O_2S$: C, 72.2; H, 11.1. Found: C, 72.3; H, 11.2%.

5.20. 5-Decyl-4-hydroxy-5-methyl-3-propyl-2,5-dihydro-2-thiophene (7j)

As described in procedure B, starting from **6c** (250 mg, 1.45 mmol), sodium hydride (69 mg, 1.74 mmol), butyllithium (0.63 mL of a 2.5 M solution in hexanes, 1.59 mmol) and 1-iododecane (0.31 mL, 0.38 g, 1.45 mmol), **7j** was obtained as a yellow oil (57 mg, 12%); δ_H 0.93 (6H, m, 2×CH_3), 1.29 (14H, m, 7×CH_2), 1.44–1.60 (2H, m, CH_2), 1.70 (3H, s, CH_3), 1.92–1.94 (4H, m, 2×CH_2) and 2.25 (2H, t, J=7.3, 1′-CH_2); δ_C 14.1, 14.5 (both CH_3), 21.8 (CH_2), 23.0, 25.5, 26.0 (all CH_2), 26.6 (CH_3), 29.7, 29.8, 29.9, 30.0, 30.5, 34.5 and 38.8 (all CH_2), 68.4 (5-C), 115.3 (3-C), 180.8 (4-C) and 196.8 (2-CO); MS (ES^-) m/z 311.2 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{18}H_{32}O_2S$ requires 313.2196, found 313.2195. Anal. calcd for $C_{18}H_{32}O_2S$: C, 69.2; H, 10.3. Found: C, 69.3; H, 10.7%.

5.21. 3-Benzyl-5-decyl-4-hydroxy-5-methyl-2,5-dihydro-2-thiophene (7k)

As described in procedure B, starting from **6d** (250 mg, 1.13 mmol), sodium hydride (54 mg, 1.36 mmol), butyllithium (0.49 mL of a 2.5 M solution in hexanes, 1.24 mmol) and 1-iododecane (0.24 mL, 0.30 g, 1.13 mmol), **7k** was obtained as a yellow oil (16 mg, 4%); δ_H 0.85 (3H, t, J=6.8, 10′-CH_3), 1.20–1.21 (16H, m, 8×CH_2), 1.58 (3H, s, CH_3), 1.79 (2H, t, J=7.8, 1′-CH_2), 3.52 (2H, s, CH_2Ph) and 7.10–7.21 (5H, m, ArH); δ_C 14.5 (CH_3), 23.1 (CH_2), 25.5 (CH_2), 26.4 (CH_3), 29.3, 29.7, 29.8, 29.9, 30.0, 32.3 and 38.9 (all CH_2), 57.9 (5-C), 114.0 (3-C), 127.3, 128.9 and 129.4 (all ArCH), 137.9 (ArC), 180.5 (4-C) and 196.1 (2-CO); MS (ES^-) m/z 359.3 ($M-H,^-$ 100%); HRMS (ES^-) $(M-H)^-$ $C_{22}H_{31}O_2S$ requires 359.2045, found 359.2048.

5.22. 5-Decyl-5-ethyl-4-hydroxy-3-methyl-2,5-dihydro-2-thiophene (12a)

As described in procedure B, starting from **11a** (250 mg, 1.58 mmol), sodium hydride (75 mg, 1.51 mmol), butyllithium (0.69 mL of a 2.5 M solution in hexanes, 1.89 mmol) and 1-iododecane (0.33 mL, 0.38 g, 1.58 mmol), **12a** was obtained as a yellow oil (34 mg, 14%); δ_H 0.90–0.98 (6H, m, 2×CH_3), 1.29 (14H, m, 7×CH_2), 1.78 (3H, s, CH_3) and 1.89–2.00 (4H, m, 2×CH_2); δ_C 6.3 (CH_3), 7.9 (CH_3), 13.1 (CH_3), 21.6, 23.6, 28.2, 28.3, 28.4, 28.5, 28.6, 29.7, 30.8, 36.6 (all CH_2), 61.9 (5-C), 110.9 (3-C), 177.1 (4-C) and 191.2 (2-CO).

5.22.1. 5-Ethyl-hexadecyl-4-hydroxy-3-methyl-2,5-dihydro-2-thiophene (12b)

As described in procedure B, starting from **11a** (200 mg, 1.26 mmol), sodium hydride (60 mg, 1.51 mmol), butyllithium (0.70 mL of a 2.0 M solution in hexanes, 1.38 mmol) and 1-bromohexadecane (0.38 mL, 0.38 g, 1.26 mmol), **12b** was obtained as a yellow oil (57 mg, 12%); δ_H 0.98–1.06 (6H, m, 2×CH_3), 1.37 (26H, m, 13×CH_2), 1.85 (3H, s, CH_3), 1,98–2.04 (4H, m, 2×CH_2) and 6.80 (1H, br s, OH). Anal. calcd for $C_{23}H_{42}O_2S$: C, 72.2; H, 11.1. Found: C, 72.0; H, 11.4%.

5.24. Biological assays

5.24.1. Plasmodium falciparum. In vitro activity against erythrocytic stages of *P. falciparum* was determined using a ^{3}H-hypoxanthine incorporation assay,[27,28] using the chloroquine and pyrimethamine resistant K1 strain and the standard drugs chloroquine (Sigma C6628) and artemisinin (Sigma 36,159-3). Compounds were dissolved in DMSO at 10 mg/mL and added to parasite cultures incubated in RPMI 1640 medium without hypoxanthine, supplemented with HEPES (5.94 g/L), $NaHCO_3$ (2.1 g/L), neomycin (100 U/mL), Albumax^R (5 g/L) and washed human red cells A^+ at 2.5% haematocrit (0.3% parasitaemia). Serial doubling dilutions of each drug were prepared in 96-well microtiter plates and incubated in a humidified atmosphere at 37 °C; 4% CO_2, 3% O_2, 93% N_2.

After 48 h 50 μL of ^{3}H-hypoxanthine (=0.5 μCi) was added to each well of the plate. The plates were incubated for a further 24 h under the same conditions. The plates were then harvested with a Betaplate^{TM} cell harvester (Wallac, Zurich, Switzerland), and the red blood cells transferred onto a glass fiber filter then washed with distilled water. The dried filters were inserted into a plastic foil with 10 mL of scintillation fluid, and counted in a Betaplate^{TM} liquid scintillation counter (Wallac, Zurich, Switzerland). IC_{50} values were calculated from sigmoidal inhibition curves using Microsoft Excel.

5.24.2. *Trypanosoma cruzi*. Rat skeletal myoblasts (L-6 cells) were seeded in 96-well microtiter plates at 2000 cells/well/100 μL in RPMI 1640 medium with 10% FBS and 2 mM L-glutamine. After 24 h, 5000 trypomastigotes of *T. cruzi* [Tulahuen strain C2C4 containing the β-galactosidase (Lac Z) gene] were added in 100 μL per well with 2× of a serial drug dilution. The plates were incubated at 37 °C in 5% CO_2 for 4 days. Then the substrate CPRG/Nonidet was added to the wells. The colour reaction, which developed during the following 2–4 h, was read photometrically at 540 nm. From the sigmoidal inhibition curve IC_{50} values were calculated.

5.24.3. *Trypanosoma brucei rhodesiense*. Minimum Essential Medium (50 μL) supplemented according to Baltz et al.[29] with 2-mercaptoethanol and 15% heat-inactivated horse serum was added to each well of a 96-well microtiter plate. Serial drug dilutions were added to the wells. Then 50 μL of trypanosome suspension (*T.b. rhodesiense* STIB 900) was added to each well and the plate incubated at 37 °C under a 5% CO_2 atmosphere for 72 h. 10 μL Alamar Blue (Trinova, Giessen, Germany) was then added to each well and incubation continued for a further 2–4 h.[30] The plates are read in a microplate fluorescence scanner (Spectramax Gemini XS by Molecular Devices) using an excitation wavelength of 536 nm and an emission wavelength of 588 nm. From the sigmoidal inhibition curve IC_{50} values were calculated.

Acknowledgements

We would like to acknowledge the Cardiff Partnership Fund for financial support and the EPSRC National

Mass Spectrometry Service Centre Swansea for accurate mass spectrometry.

References and notes

1. WHO *Weekly Epidemiol. Rep.* **1997**, *72*, 269.
2. White, N. J. *Ann. Trop. Med. Parasitol.* **1998**, *92*, 449.
3. 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. A.; 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. *Nature* **2002**, *419*, 498.
4. 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. M.; Gardner, M. J.; Holder, A. A.; Sinden, R. E.; Yates, J. R.; Carucci, D. J. *Nature* **2002**, *419*, 520.
5. Harwood, J. L. *Biochim. Biophys. Acta-Lipids Lipid Metab.* **1996**, *7*, 1301.
6. Clough, B.; Rangachari, K.; Strath, M.; Preiser, P. R.; Wilson, R. *Protist* **1999**, *150*, 189.
7. Ralph, S. A.; D'Ombrain, M. C.; McFadden, G. I. *Drug Resist. Update* **2001**, *4*, 145.
8. Jackowski, S.; Rock, C. O. *J. Biol. Chem.* **1987**, *262*, 7927.
9. Kohler, S.; Delwiche, C. F.; Denny, P. W.; Tilney, L. G.; Webster, P.; Wilson, R. J. M.; Palmer, J. D.; Roos, D. S. *Science* **1997**, *275*, 1485.
10. McFadden, G. I.; Reith, M. E.; Munholland, J.; LangUnnasch, N. *Nature* **1996**, *381*, 482.
11. McFadden, G. I.; Waller, R. F. *Bioessays* **1997**, *19*, 1033.
12. Wilson, R. J. M.; Denny, P. W.; Preiser, P. R.; Rangachari, K.; Roberts, K.; Roy, A.; Whyte, A.; Strath, M.; Moore, D. J.; Moore, P. W.; Williamson, D. H. *J. Mol. Biol.* **1996**, *261*, 155.
13. Nishida, I.; Kawaguchi, A.; Yamada, M. *J. Biochem. (Tokyo)* **1986**, *99*, 1447.
14. Jackowski, S.; Murphy, C. M.; Cronan, J. E.; Rock, C. O. *J. Biol. Chem.* **1989**, *264*, 7624.
15. Oishi, H.; Noto, T.; Sasaki, H.; Suzuki, K.; Hayashi, T.; Okazaki, H.; Ando, K.; Sawada, M. *J. Antibiot.* **1982**, *35*, 391.
16. Noto, T.; Miyakawa, S.; Oishi, H.; Endo, H.; Okazaki, H. *J. Antibiot.* **1982**, *35*, 401.
17. Kremer, L.; Douglas, J. D.; Baulard, A. R.; Morehouse, C.; Guy, M. R.; Alland, D.; Dover, L. G.; Lakey, J. H.; Jacobs, W. R.; Brennan, P. J.; Minnikin, D. E.; Besra, G. S. *J. Biol. Chem.* **2000**, *275*, 16857.
18. Jones, A. L.; Herbert, D.; Rutter, A. J.; Dancer, J. E.; Harwood, J. L. *Biochem. J.* **2000**, *347*, 205.
19. Sakya, S. M.; Suarez-Contreras, M.; Dirlam, J. P.; O'Connell, T. N.; Hayashi, S. F.; Santoro, S. L.; Kamicker, B. J.; George, D. M.; Ziegler, C. B. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2751.
20. Waller, R. F.; Ralph, S. A.; Reed, M. B.; Su, V.; Douglas, J. D.; Minnikin, D. E.; Cowman, A. F.; Besra, G. S.; McFadden, G. I. *Antimicrob. Agents Chemother.* **2003**, *47*, 297.
21. Wang, C. L. J.; Salvino, J. M. *Tetrahedron Lett.* **1984**, *25*, 5243.
22. Benary, E. *Berichte* **1913**, *46*, 2103.
23. Josien, H.; Curran, D. P. *Tetrahedron* **1997**, *53*, 8881.
24. Price, A. C.; Zhang, Y. M.; Rock, C. O.; White, S. W. *Biochemistry* **2001**, *40*, 12772.
25. Price, A. C.; Choi, K. H.; Heath, R. J.; Li, Z. M.; White, S. W.; Rock, C. O. *J. Biol. Chem.* **2001**, *276*, 6551.
26. Morita, Y. S.; Paul, K. S.; Englund, P. T. *Science* **2000**, *288*, 140.
27. Desjardins, R. E.; Canfield, C. J.; Haynes, D.; Chulay, J. *Antimicrob. Agents Chemother.* **1979**, *16*, 710.
28. Matile, H.; Pink, J. R. L. Plasmodium Falciparum Malaria Parasite Cultures and Their Use in Immunology. In *Immunological Methods*; Academic: San Diego, 1990; p 221.
29. Baltz, T.; Baltz, D.; Giroud, C.; Crockett, J. *EMBO J.* **1985**, *4*, 1273.
30. Raz, B.; Iten, M.; Grether-Buhler, Y.; Kaminski, R.; Brun, R. *Acta Trop.* **1997**, *68*, 139.