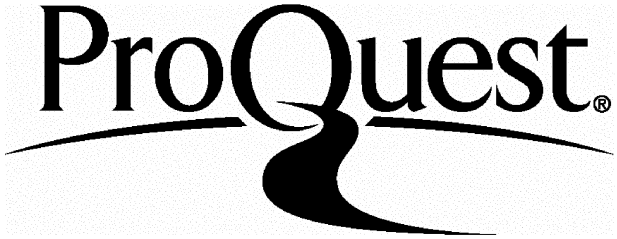
Genetic manipulation and gene expression in the human malaria parasite *Plasmodium falciparum*.

Stuart James Lucas

A thesis submitted in partial fulfillment of
the requirements of University College London
for the degree of Doctor of Philosophy

Division of Parasitology,
National Institute for Medical Research
Mill Hill, London, UK.

March 2003



Acknowledgements

Many thanks to everyone in the Holder lab, past and present, for their support and companionship throughout the course of this project. Particular thanks to Helen Taylor for helping me to find my feet when I first started in the lab and for many useful discussions, to Muni Grainger for excellent training in culture techniques and for supplying parasites on numerous occasions, and to Irene Ling for valuable advice and running sequencing gels. Thanks also to Mike Blackman, Bill Jarra, Hannah Polson, Barbara Clough, Justin Pachebat, Chairat Uthaipibull, Sola Ogun, Terry Scott-Finnigan, Mallika Kaviratne, Brian Yim Lim and everyone else for advice and help. Thanks to Peter Preiser for valuable comments on my mid-term report, and to Ariel Achtman for patient friendship, humour, and discussions about everything under the sun. Special thanks to my supervisor Tony Holder, for letting me loose in his lab, for reading this thesis, and for encouragement and advice throughout the project. Thanks to NIMR photo-graphics for their work, and to D. J. Atkinson and the Division of Neurobiology & Development, NIMR, for letting me borrow their GenePulser so often. Thanks to Rod King, Ann Ager and especially Kristine Cronhelm for ensuring that all the administrative aspects of being a student ran smoothly.

Most of all I want to thank my gorgeous wife, Pippa, for continual love, affirmation and support, for being the major wage-earner, and for lots of fun! Thanks to my parents and in-laws for their encouragement and concern, and to my friends for letting me enthuse about science occasionally. A massive thank you to God for my abilities and for motivation and perspective.

This research was funded by a Medical Research Council studentship.

Index of Contents

Index of Figures

Abbreviations

A, C, G, T	Adenine, Cytosine, Guanine and Thiamine
AMA-1	Apical Membrane Antigen 1
BSA	Bovine Serum Albumin
CAM	Calmodulin
CAT	Chloroamphenicol O-acetyltransferase
cpm	counts per minute
DHFR-TS	Dihydrofolate reductase/thymidylate synthase
EST	Expressed Sequence Tag
FACS	Fluorescence Activated Cell Sorting
GA3PDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green Fluorescent Protein
hDHFR	human Dihydrofolate reductase
HDL	High Density Lipoprotein
hr	hour
ICAM-1	Inter-cellular Adhesion Molecule 1
KAHRP	Knob-associated Histidine Rich Protein
kb(p)	Kilobase(pair)s
kDa	KiloDaltons
K_m	Michaelis constant
LSC	Liquid Scintillation Counting
MRE	Metal Response Element

MSP-1	Merozoite Surface Protein 1
MSP-7	Merozoite Surface Protein 7
orf	open reading frame
PcDT	*Plasmodium chabaudi* DHFR-TS
PCR	Polymerase Chain Reaction
PfEMP1	*Plasmodium falciparum* Erythrocyte Membrane Protein 1
PfEMP3	*Plasmodium falciparum* Erythrocyte Membrane Protein 3
rpm	revolutions per minute
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SDS-PAGE	Sodium Dodecyl Sulphate/Polyacrylamide Gel Electrophoresis
SV40	Simian Virus 40
TARE	Telomere Associated Repetitive Element
TBP	TATA Binding Protein
TdT	Terminal deoxyribonucleotide transferase
TgDHFR-TS	*Toxoplasma gondii* DHFR-TS
TLC	Thin layer chromatography
TRAP	Thrombospondin Related Anonymous Protein
TVM/TVN	Tubo-vesicular membrane network
UTR	Untranslated region

Abstract

Molecular genetic analysis of the most important malaria pathogen, *Plasmodium falciparum,* presents numerous challenges. The parasite is relatively refractory towards transfection, and the nature of its transcriptional regulation mechanisms remain to be elucidated.

Both existing techniques and new approaches for the transfection of parasites within red blood cells were assessed, including electroporation, lipofection, pre-loading erythrocytes and applying morpholino antisense oligonucleotides. Of these only electroporation proved reliable, and although the efficiency of transfection was improved by optimization of this method it remained around 0.1%. In an attempt to identify transfected cells from the untransfected background, a plasmid was designed to express a c-myc epitope-tagged protein of the Rifin family on the surface of erythrocytes infected with transfected parasites. However, in transient transfection studies, expression of the Rifin was too low to detect at either the mRNA or the protein level.

This Rifin expression problem raised the issue of transcriptional control in *P. falciparum.* Before promoter activity could be accurately measured, a quantitative assay for the chloramphenicol acetyltransferase (CAT) reporter gene product was established. A range of *P. falciparum* 5' untranslated regions from genes of biological interest were then cloned and used to drive CAT expression in a transient transfection system. Clear differences in promoter strength and stage specificity were observed, and regions containing important control sequences for the *enolase, kahrp* and *msp1* genes were identified. The expression patterns of several promoters differed from their predicted stage specificity; however, patterns much closer to their biological activity were observed after plasmid replication by the parasite, indicating a role for epigenetic factors in regulating gene expression.

Stable transfection was established using selection for the human dihydrofolate reductase (hDHFR) drug resistance gene, and this strategy was used to try and express the Rifin gene either with c-myc tag or fused to Green Fluorescent Protein.

1. Introduction

1.1 A Global Health Problem

1.1.1. The malaria disease burden

Malaria today stands not only among the world's most serious infectious diseases, but as one of the most pressing health problems of any kind. By conservative estimates, over a million people die of malaria in Africa alone in a typical year (Snow et al, 1999), and the disease is also endemic in many other areas throughout the tropics. In total, about 40% of the world's population live in malarious areas, with highest risk groups including children under five years of age, pregnant women, and adults with no previous exposure to the infection. Increasing international travel has also led to a steady rise in cases of the disease reported in currently non-malarious countries.

Furthermore, the mortality rate is only the tip of the iceberg – about 200 million non-fatal clinical cases per year represent a vast loss of earnings and drain on health resources, primarily in developing countries ill-equipped to cope with it. Moreover, there is epidemiological evidence that in malaria endemic areas the death rate by other causes increases, presumably due to the strain constant malaria infection places on the host (Molineaux, 1997).

1.1.2. Intervention strategies

Treatment of malaria dates back to the mid-seventeenth century in Central and South America, where ground *Cinchona* bark was employed as a natural remedy to the disease. Following its identification as the most active ingredient of the bark by Pelletier and Caventou in 1820 (described in Turner & Woodward, 1953), quinine and its synthetic derivatives (chloroquine, amodiaquone, primaquine and mefloquine) have been employed as anti-malarials ever since. Unfortunately, their effectiveness has been greatly diminished by the rapid development of drug-resistant malaria over the last 30 years. The impact of resistance to chloroquine, the frontline drug in many developing countries, is seen in

numerous hospital- and population-based studies from the 1980s where malaria mortality increased by more than twofold in as little as two years (Trape et al, 2002). Alternative drugs currently available are more expensive, and many also show signs of developing resistance in the field. In a repeat of history, the Chinese plant *Artimisia annua* L has recently been found to have potent anti-malarial activity, and a new group of anti-malarial compounds are currently being developed from it. However, only time will tell whether they can be produced cheaply enough to control malaria in the developing world, and whether drug resistance will again prove to be a confounding factor.

The essential role of the *Anopheles* mosquito in the spread of malaria was discovered by Sir Ronald Ross in 1897 (Ross, 1897), and controlling the vector has been another approach employed against the disease. Extensive spraying with the insecticide DDT in the 1950s eradicated malaria from much of Europe, but the same approach has not been successful in the tropics, where the areas to be covered are larger and the mosquitoes breed much more freely. Simple physical measures to keep the insects away, such as using bed nets and eliminating pools of standing water from the vicinity of habitation, have had a limited effect, but are unlikely ever to remove malaria altogether from endemic areas.

As yet there is no viable vaccine against malaria. The huge burden of the disease outlined above and the inadequacy of current interventions make biological research into this infection a priority. Molecular biology in particular has the potential to reveal the mechanisms by which disease is mediated, elucidate the interactions between the parasite and the host, and identify potential targets for new drugs or vaccines.

1.2 Biology of the Malaria Parasite

1.2.1 Life-cycle and morphology

Malaria is caused by protozoan parasites of the genus *Plasmodium* and the phylum *Apicomplexa,* a name referring to the distinctive organelles located at the apical end of their invasive stages. All *Plasmodia* are obligate parasites which alternate between a vertebrate host and an insect vector, the latter being essential for transmitting the parasite from one

host to another. A wide range of animals including mammals, birds and lizards can be infected by specific *Plasmodium* species, each of which seems to be restricted to one vertebrate host. Molecular phylogenetic analysis finds that the mammalian *Plasmodium* lineages cluster according to the taxa of their hosts; this suggests that *Plasmodium* invaded an early mammal at a single point in evolutionary history, and has evolved over millions of years alongside mammals, to give the separate species seen today (Perkins & Schall, 2002).

The four parasites known to infect humans are *P. falciparum, P. malariae, P. ovale* and *P. vivax.* Of these *P. falciparum* is responsible for the vast majority of malaria deaths, and is therefore the focus of much malaria research, including this project. Female mosquitoes of the *Anopheles gambiae* complex are responsible for transmitting the infection, which is a major factor limiting the distribution of the disease: the mosquito-borne parasites can only survive where the climate is hot (not dropping below 18°C), while the mosquitoes themselves require a wet environment.

The complex life-cycle of *P. falciparum* is summarized in fig 1.1. During a blood meal from an infected human, a mosquito ingests male and female **gametocytes,** which then combine sexually within the mosquito gut and develop into a motile stage called the **ookinete**. This form is able to penetrate the gut epithelial cells to the basal lamina of the midgut epithelium, where it transforms into a stationary **oocyst**. After a few days the oocyst bursts, releasing thousands of **sporozoites** which migrate to the mosquito's salivary glands, ready to be injected into a new host when the mosquito next feeds.

Once in the human bloodstream, the sporozoites find their way to the liver and invade hepatocytes within a few minutes. Some malaria species such as *P. vivax* can form a dormant liver stage which may be reactivated months or years later, but in *P. falciparum* all the invaded cells mature over a few days into **hepatic schizonts,** which then lyse to release thousands of **hepatic merozoites.**

This signals the start of the asexual erythrocytic reproductive cycle. Again within a few minutes, merozoites invade mature red blood cells (erythrocytes), within which they grow and develop for about 48 hours. After this time each erythrocyte lyses to release about 16 new merozoites and start the cycle again. The invasion begins with merozoite attachment to

the erythrocyte membrane, followed by reorientation of the merozoite to direct its apical end towards the target cell. As the merozoite disgorges the contents of its distinctive apical organelles (such as the rhoptries and micronemes) the erythrocyte membrane invaginates, eventually completely surrounding the parasite to form a compartment called the parasitophorous vacuole. Within this compartment the parasite passes through several distinct morphological stages (fig 1.2); young **trophozoites** have a cup- or signet ring-shaped appearance, and are often simply referred to as **rings.** By 20-24 hours after invasion the parasite begins to grow in size, digest haemoglobin from the host cell, and develop a pigment-containing vacuole; this is the **mature trophozoite.** Later on as the nucleus divides it is referred to as a **schizont,** which matures until all the new individual merozoites can be clearly distinguished just prior to rupture of the parasitophorous vacuolar and host cell membranes.

The parasite can theoretically pass through the erythrocytic cycle an indefinite number of times, increasing parasite number exponentially. In practice, a proportion of parasites undergo a developmental switch, and after schizont maturation and rupture all the merozoites released from these parasites grow into sexually differentiated male and female gametocytes after re-invasion (Bruce et al, 1990). As the parasitaemia rises to include more than about 20% of the available erythrocytes, the rate of gametocytogenesis increases dramatically. The gametocytes are then available to infect a mosquito and start the whole life-cycle again.

With such a complex life-cycle – two hosts, both sexual and asexual reproduction, and the numerous different invasive and intracellular forms of the parasite – there are many stages which could be examined as a target for intervention. One of the most important, on which this project is focused, is the erythrocytic reproductive cycle. This is the part of the life-cycle responsible for all the pathology of the disease in humans; it is also when the parasite is most exposed to the host immune system, and therefore to potential interventions. On a more practical note, it is more accessible to research than other stages as it can be cultured *in vitro,* as discussed in section 1.2.3.

1.2.2 Pathology of *P. falciparum* malaria

The intraerythrocytic development of the parasite affects the host in many ways, and has in turn been affected by the host immune response. The fevers at 48-hour intervals which are characteristic of *P. falciparum* infection are induced following simultaneous schizont rupture and re-invasion by each generation of merozoites; conversely, it has been proposed that febrile temperatures are at least partially responsible for synchronizing the parasites *in vivo* (Kwiatkowski, 1989). This synchronization may benefit the parasite as it minimizes the period during which free merozoites – the only extracellular stage of the erythrocytic cycle – are exposed to the immune system. If the disease is unchecked by the immune system or treatment, more and more of the host's erythrocytes will be lysed during each cycle leading to severe malarial anaemia and/or metabolic acidosis.

Parasitized erythrocytes undergo changes in morphology, becoming bumpy and more rigid than healthy red blood cells, and as such might be expected to be cleared from the circulation in the spleen and destroyed. However, *P. falciparum* has evolved the ability to 'cytoadhere' to endothelial cells, causing parasitized erythrocytes to sequester in the microvasculature of many organs rather than circulating. A related property of some *P. falciparum* parasites is 'rosetting,' where healthy red blood cells clump around a parasitized erythrocyte. These two phenomena are involved in many of the possible complications that make *P. falciparum* malaria so dangerous; sequestration of parasitized erythrocytes in the microvasculature can lead to single or multi-organ failure. One of the most dangerous complications of *P. falciparum* infection arises when a similar process in the brain leads to cerebral malaria, which is fatal in 20% of cases.

1.2.3 *P. falciparum* as an experimental system

The erythrocytic stages of *P. falciparum* are currently the only malarial parasites which can be maintained for extended periods *in vitro.* Trager and Jensen (1976) first established conditions for culturing these forms using mammalian cell culture medium, freshly donated human erythrocytes, and a CO_2 enriched atmosphere designed to mimic that in the

microvasculature. As long as the medium is replaced frequently and the parasites are diluted periodically with unparasitized erythrocytes, the erythrocytic cycle can be repeated indefinitely, greatly facilitating research. In contrast, the sexual cycle requires gametocytes (which can be generated *in vitro*) to be transferred to live mosquitoes, with the associated safety risks, and the other mammalian malarias studied as models of the human disease all require a rodent or primate host.

Cultured parasites can be readily visualized by making a thin blood smear and staining with Giemsa's stain (fig 1.3). For simplicity, in this project the parasite forms shown in the 3-21hr pictures are described as 'rings,' those in the 24-33hr pictures as 'trophozoites,' and those in the 36-48hr pictures as schizonts.

However, *in vitro* culture is not a perfect representation of the situation during an infection. For example, *P. falciparum* cultures are not naturally synchronous, although methods have been developed to impose synchronicity by exploiting the fragility of late trophozoites and schizonts relative to less mature parasites (Freeman & Holder, 1983; Haynes & Moch, 2002). More importantly, the immune response is completely absent *in vitro;* repeated cycling of parasites in the absence of immune pressure can rapidly accumulate mutations in genes that are not essential for parasite growth, but are involved in immune evasion *in vivo.* One well documented example of this is the truncation of chromosome 9, which occurs with surprising frequency in culture (Shirley et al, 1990). These mutants and many parasite lines widely used in culture have significant loss of function phenotypes, such as the inability to form 'knobs' (structures at the parasitized erythrocyte surface which are essential for cytoadherence) or to develop into gametocytes.

These caveats must be borne in mind, but parasites grown *in vitro* nevertheless provide an excellent resource for molecular biological analysis of *P. falciparum.* The parasite clone used throughout this project is 3D7, which is knob positive, and able to form gametocytes; it is also the strain sequenced by the malaria genome project.

1.3 Molecular Biology of *P. falciparum* Intraerythrocytic Stages

1.3.1 Metabolism & development within the host cell

Over the last 25 years, substantial advances have been made in understanding the molecular mechanisms which underlie malarial parasitism of the erythrocyte. In the invasion process, Merozoite Surface Protein 1 (PfMSP-1) has been shown to have a critical role; it is found all over the surface of free merozoites, but is shed during erythrocyte invasion by proteolytic processing. Blocking this processing step with specific antibodies prevents invasion *in vitro* and is associated with protective immunity *in vivo* (Blackman et al, 1994; Egan et al, 1996).

Once inside the erythrocyte, *P. falciparum* is heavily dependent on glucose uptake from the blood, and exports large quantities of lactate, indicating that glycolysis is its major source of metabolic energy (Sherman, 1979). Although all the genes encoding enzymes involved in the citric acid cycle are present in the genome, their protein products are only detected in gametocytes and mosquito stages (Florens et al, 2002). Additionally, the parasite ingests and digests haemoglobin from host erythrocytes as it grows, until by the schizont stage all the haemoglobin is digested and the parasite fills the host cell. This process releases potentially toxic haem species which the parasite packages into a semi-crystalline aggregate called haemazoin – the pigment seen in the vacuoles of trophozoites and schizonts, which is released into the medium on schizont rupture. Many of the enzymes which mediate these processes have been identified and are being evaluated as potential drug targets (reviewed in Subbayya et al, 1997).

As well as carrying out its own complex developmental programme, the parasite also extensively re-models the host erythrocyte. Soon after invasion an assembly of membrane tubules and vesicles known as the tubo-vesicular membrane network (TVM or TVN) is formed within the erythrocyte cytoplasm (Elmendorf & Haldar, 1994). The TVN extends from the parasitophorous vacuolar membrane throughout the host cell, and is associated with the export of malarial proteins to the host erythrocyte (Behari & Haldar, 1994). *P. falciparum* proteins may have many roles in the host cell, but one which is certainly important is the induction of new permeation pathways (NPPs) in the erythrocyte plasma

membrane, by which the parasite can take up a wide variety of molecules from outside the cell (reviewed by Kirk, 2001). Another important and well researched role is the pathologically important modification of the erythrocyte surface.

1.3.2 Modification of the erythrocyte surface

As described in section 1.2.2, parasitized erythrocytes are able to cytoadhere, a process mediated by 'knob' structures in the erythrocyte plasma membrane. Several *P. falciparum* proteins have key roles in these structures. Knob-associated histidine rich protein (KAHRP or HRP-1) is found in large quantities beneath knobs on the cytoplasmic side of the plasma membrane, and binds to host cell cytoskeletal elements. Loss of the KAHRP gene abolishes knob formation, and cytoadherence under flow conditions (Crabb et al, 1997a). Also present at the base of knobs in smaller quantities is Erythrocyte Membrane Protein 3 (PfEMP-3). Truncation of this protein does not prevent knob formation, but does block transport to the knob of Erythrocyte Membrane Protein 1 (PfEMP-1; Waterkeyn et al, 2000). PfEMP-1 is a large, multi-domain protein that has a cytoplasmic tail which interacts with KAHRP (Oh et al, 2000), a transmembrane region, and then several extracellular domains which mediate binding to host cell surface proteins such as CD36, ICAM-1 (Berendt et al, 1989) and chondroitin sulphate A (Buffet et al, 1999).

Its location external to the erythrocyte should make PfEMP-1 one of the most vulnerable parasite proteins to the immune system. However, *P. falciparum* evades immune recognition by a mechanism of antigenic variation (Brown & Brown, 1965). PfEMP-1 is actually a set of 60 different proteins produced by the *var* multigene family, only one of which is expressed on the surface of each parasite (Chen et al, 1998). Each life-cycle some parasites switch to express a different *var* gene from their parent, so whenever an immune response is mounted against one particular PfEMP-1 many of the parasites escape. A large amount of current research is directed at both the adhesive and immunological characteristics of PfEMP-1.

1.4 Transfection of *Plasmodium* Species

1.4.1 Development of malaria transfection systems

Transfection has proved to be an invaluable tool in studying the molecular biology of many different organisms. By introducing tailor-made DNA sequences and taking advantage of naturally occurring processes such as DNA recombination, it allows for a variety of valuable experiments: knocking genes out, overexpressing or conditionally expressing them, introducing mutated genes, and many more. Although transfection is a routine technique in organisms such as *Escherichia coli* and *Saccharomyces cerevisiae,* it has only recently been applied to *Plasmodium* species, and is still a far from straightforward procedure.

The first malaria species to be transfected was the avian parasite *P. gallinaceum* (Goonewardene et al, 1993). A plasmid containing the gene for firefly luciferase inserted in frame within the sexual stage specific *pgs28* gene was introduced into fertilized gametes and zygotes by electroporation, and luciferase activity could be detected from the parasites 24hr later. Transfection of intracellular parasite forms proved much more difficult, but was eventually obtained in *P. falciparum* by Wu et al (1995), again using electroporation and a plasmid carrying the enzyme chloramphenicol acetyltransferase (CAT) flanked by 5' and 3' non-coding sequences from two parasite genes, *Pfhrp2* and *Pfhsp86*. Reporter gene activity was detected transiently from 24-72hr post-transfection, after which it dropped rapidly, presumably due to failure of the parasite to maintain the plasmid.

The next major breakthrough came in the rodent malaria *P. berghei,* where Van Dijk et al (1995) were able to maintain a plasmid after electroporation by drug selection - the plasmid, which contained the mutant allele of the *P. berghei* Dihydrofolate reductase/thymidylate synthase (DHFR-TS) gene, conferred resistance to pyrimethamine. In the first instance the plasmids were maintained episomally, but it was then found that linearized plasmids could integrate into the genome by homologous recombination (Van Dijk et al, 1996). This achievement was then repeated in *P. falciparum* (Wu et al, 1996; Crabb & Cowman, 1996); in this case the *Toxoplasma gondii* DHFR-TS was used, only circular plasmids could be

maintained, and the efficiency of integration was poor. Nevertheless, the ability to insert sequences into the genome at specific sites has proved very valuable (see section 1.4.4).

1.4.2 Technical and biological considerations and improvements

Despite the years that have past since the breakthroughs described above, transfecting malaria parasites remains problematic. Electroporation is the only widely reproduced transfection technique, and even in this case the proportion of parasites which actually take up DNA is very low and variable; optimistic estimates put it at 1 in a 1000 (Crabb & Cowman, 1997) and many are much lower. One problem is that 50-70% of the parasites are rendered non-viable by the electroporation, although this can be avoided by electroporating plasmid DNA into uninfected erythrocytes, and then adding parasites which take up the DNA during or after invasion of the pre-loaded cells (Deitsch et al, 2001). Furthermore, van Dijk et al (1997) showed in *P. berghei* that while episomal plasmid DNA is replicated by the parasite, it does not segregate uniformly between merozoites during schizogony, and is therefore rapidly diluted out of the culture in successive erythrocytic cycles. Low and transient levels of plasmid mean that, while the presence of a reporter gene can be detected, any other phenotypic effects are masked by the huge background of untransfected parasites.

In a drug selected stable transfectant culture a much higher proportion of parasites carry the transgene, so phenotypes are easier to detect. However, segregation of episomes is still inefficient and the drug selection only kills parasites in the late trophozoite and schizont stages, so there will always be some untransfected background. In most cases this can be accounted for, but in some it could be a confounding factor. The backgorund can be eliminated by selecting for plasmid integration into the genome, but this is not trivial, as discussed below. Furthermore, the efficiency of stable transfection is even lower than that of transient; approximately one in a million of the transfected parasites go on to generate the stable culture (B. Crabb, personal communication). The reasons for this inefficiency are unclear, although the health of the parasites being transfected is thought to be of paramount importance. Also, recombination, either between the plasmid and the genome or between multiple copies of the plasmid to form concatamers (O'Donnell et al, 2001b), appears to occur in all stable transfectants. The reliability of this process is unknown, but it has been successfully exploited by Kadekoppala et al (2001), who showed that multiple plasmids

containing homologous sequences could be co-transfected into *P. falciparum* and would then recombine and co-segregate. This allows multiple expression cassettes to be inserted into *P. falciparum* on separate plasmids, avoiding the difficulties inherent in cloning large AT-rich pieces of DNA.

Such low transfection efficiency means that it takes several weeks to obtain stable transfectants. Protracted culture times make stable transfection not only laborious, but also unsuitable for analyzing transgenes that confer deleterious phenotypes; in such cases the parasites never grow up again after applying drug pressure. Moreover, extended periods of pyrimethamine pressure can lead to selection of pyrimethamine resistant mutants, rather than parasites containing plasmids. Fidock & Wellems (1997) solved this last problem by using the human dihydrofolate reductase (hDHFR) gene, which unlike its apicomplexan homologues confers resistance to the drug WR99210; no parasite resistance to this drug has ever been observed. Since then, blasticidin S deaminase (Mamoun et al, 1999), neomycin phosphotransferase II (Mamoun et al, 1999) and puromycin-N-acetyltransferase (de Koning-Ward et al, 2001) have also been used as selectable markers for *P. falciparum* transfection, allowing multiple genetic manipulations of the same parasite clone.

Culture times are even longer if integration into the genome is required, as the transfectants are either cloned by limiting dilution or cycled with and without drug to eliminate parasites with episomes, which can take months (or years in some cases). Recently, however, two groups have reported mechanisms for rapidly selecting for homologous integration. The first is a positive/positive mechanism, by Wang et al (2002), using a plasmid with a second drug resistance gene fused to the C-terminus of the gene of interest. The plasmid copy of the gene has no promoter, so only if the plasmid recombines into the genome is the drug resistance gene expressed. The second is a positive/negative selection system developed by Duraisingh et al (2002), which uses a plasmid with both a hDHFR cassette to select for the presence of the transgene, and a second thymidine kinase cassette. Parasites which express thymidine kinase are sensitive to ganciclovir; the enzyme converts the parent drug, which is not toxic to the parasite, into destructive metabolites. Therefore any transfected cell which contains the entire plasmid, either as an episome or a single crossover integrant, will be killed by ganciclovir. However, the hDHFR cassette is flanked by two parasite genomic sequences so there is a possibility of rare double crossover recombination events, where the

hDHFR cassette is inserted between the genomic copies of its two flanking sequences and the rest of the plasmid is lost due to its poor segregation. These parasites therefore survive ganciclovir treatment. The great advantage of this system is that it allows not just insertion into the genome, but replacement of a section of genomic sequence by the hDHFR cassette. For example, if the genomic copies of the flanking sequences are found immediately before and after an orf of interest, that orf will be completely deleted, rather than just interrupted as is the case in knockouts produced by single crossover recombination.

Other refinements have also been made to *Plasmodium* transfection; green fluorescent protein (GFP) has been shown to function in both *P. falciparum* (Van Wye & Haldar, 1997) and *P. berghei* (de Koning-Ward et al, 1998), providing a valuable *in vivo* reporter system. Also, transfection has now been applied to another rodent malaria parasite, *P. yoelii* (Mota et al, 2001), and to two *Plasmodium* species which are studied as primate models of malaria: *P. knowlesi* (van der Wel et al, 1997) and *P. cynomolgi* (Kocken et al, 1999). The recent development of *in vitro* culture for erythrocytic stages of *P. knowlesi* (Kocken et al, 2002) for the first time allows analysis of transfected malaria parasites both *in vivo* and *in vitro.*

1.4.3 Use of transfection to analyze gene expression

The techniques of malaria transfection have facilitated new research into the mechanisms controlling gene expression. In one of the earliest transfection studies, Crabb & Cowman (1996) produced plasmids containing the CAT orf preceded by varying lengths of the 5' untranslated regions (UTRs) of three genes, *P. falciparum* calmodulin and the DHFR-TS genes from *P. falciparum* and the rodent parasite *P. chabaudi.* By measuring the CAT activity expressed in *P. falciparum* parasites transiently transfected with each plasmid, regions of each promoter which were required for efficient gene expression were identified. A similar approach was used by Horrocks and Lanzer (1999a) to characterize sequences essential to the GBP130 promoter. A more comprehensive study by Dechering et al (1999) on the promoters of two sexual stage genes, *pfs16* and *pfs25*, both identified sequences important to the activity of each promoter and defined their stage specificity – the *pfs16* promoter is induced in early gametocytes, while *pfs25* is not active until after transfer to the mosquito vector. However, placing a promoter on an episome can also result in aberrant

nucleosome organization and a concomitant loss of stage specificity, as observed for that of GBP130 (Horrocks and Lanzer, 1999b).

Apart from these insights into transcription in *P. falciparum,* considerable progress has also been made in understanding protein trafficking. Expression of reporter genes fused to putative targeting sequences has allowed the destination of specific proteins to be identified, either by subcellular fractionation (Burghaus & Lingelbach, 2001) or if GFP was the chosen reporter by fluorescence microscopy (e.g. Waller et al, 2000). In this latter study it was demonstrated that a putative plastid targeting sequence (PTS) did indeed direct proteins to the apicoplast of malaria parasites. Further transfections using modified versions of the PTS have allowed the general characteristics of such sequences to be clearly defined, enabling hundreds of genes coding for plastid targeted proteins to be identified (Foth et al, 2003). The possibility of observing GFP *in vivo* combined with fluorescence recovery after photobleaching (FRAP; Wickham et al, 2001) has been used to define more clearly compartments involved in protein transport; KAHRP travels to the knobs via the Maurer's clefts, a compartment just below the parasitized erythrocyte membrane (Wickham et al, 2001), while the parasitophorous vacuole has been shown to consist of multiple partially connected subcompartments (Adisa et al, 2003).

1.4.4 Use of transfection to explore gene function

Identifying the function of newly discovered genes is one of the biggest challenges in malaria research today, and transfection is one of the few approaches that can facilitate this. The archetypal experiment of this sort is the genetic knockout, in which a target gene is disrupted by homologous recombination with a plasmid containing some identical sequence, eliminating the expression of that gene. The resulting phenotypes can be very informative, demonstrated first by the KAHRP knockout in *P. falciparum* (cf. section 1.3.2; Crabb et al, 1997a), and the knockouts of two sporozoite proteins in *P. berghei.* Sporozoites lacking TRAP ('thrombospondin related anonymous protein' – Sultan et al, 1997) lost their gliding motility and became non-invasive, while CSP knockout parasites (circumsporozoite protein – Menard et al, 1997) showed inhibited sporozoite formation in the oocyst. Since these findings, knockouts have been made of many more *P. falciparum* genes, including those

encoding PfEMP3 (cf. section 1.3.2; Waterkeyn et al, 2000) and the serine repeat antigen (SERA) family (Miller et al, 2002). However, many knockouts have had no discernible phenotype, suggesting that the target gene is functionally redundant, while knockouts of other genes have not be obtained despite repeated attempts, indicating that such genes are essential to parasite growth. In neither case is the functional role of the target gene made any clearer.

In the light of these limitations of gene knockouts, which induce a loss of function, it is also important to use transfection for functional complementation and gain of function strategies, which are currently less well developed in *P. falciparum.* Complementation can be by allelic replacement; O'Donnell et al (2001a) successfully replaced the C-terminal 19 kDa domain of PfMSP-1 with its homologue from *P. chabaudi*, showing that the two were functionally interchangeable and that specific antibodies against this domain form a major component of a protective host immune response. Alternatively, where a knockout has been obtained, another copy of the targeted gene can be introduced episomally to complement for the lost gene, as has been achieved for TRAP in *P. berghei* (Sultan et al, 2001). The usefulness of gain of function experiments has recently been shown by Sidhu et al (2002), who expressed naturally occurring *pfcrt* gene mutants in chloroquine sensitive parasites, proving that these mutants confer clinically relevant drug resistance.

In summary, the immense potential of transfection to unravel many key questions in *P. falciparum* biology is beginning to be realized. However, there are still many technical issues to be resolved before it can fulfill this potential.

1.5 *P. falciparum* Molecular Genetics and Genomics

1.5.1 The eukaryotic paradigm for gene expression

Throughout biology, essential regulation in the timing and level of gene expression occurs at the level of transcription. The standard model of eukaryotic transcription emphasizes the roles of DNA sequences located immediately upstream of each open reading frame in this regulation, and seems to apply to all eukaryotes examined (reviewed in Lewin, 1997).

Generally, multiple sequence elements are required for efficient transcription, some of which form the *promoter,* which is typically located a short distance upstream of the transcription initiation site. The function of the promoter is highly dependent on its location, as it can often be rendered inactive by inserting a small number of bases between it and the start of transcription. In contrast, *enhancer* elements may be several kilobases up or downstream of the transcription initiation site in either orientation and yet still stimulate transcription. In most cases the promoter is essential and sufficient for a limited basal level of transcription, with enhancers being required to increase transcription to a useful level.

Both promoters and enhancers contain short sequence motifs that form binding sites for specific protein transcription factors. No sequence/factor combination is essential to all genes, but one of the most ubiquitous is the TATA box/TATA-binding protein (TBP) pair. The TATA box is an 8bp consensus composed entirely of As and Ts, flanked by short GC-rich tracts, and is centred about 25bp upstream of the transcriptional start site. It has been implicated in specifying the location of this site, as some of the minority of promoters that lack a TATA box initiate transcription at multiple positions. Other motifs provide specificity in transcription; for example, the metal response element (MRE) found in human metalliothionein promoters and enhancers binds a specific transcription factor which is somehow activated when free heavy metal ions are present, causing the genes which contain MREs to be upregulated in these circumstances.

In addition to the sequences 5' of an open reading frame that control its expression, there are also critical sequences in the 3' untranslated region (UTR). These signal the end of transcription and the polyadenylation of the resulting RNA by a specific terminal deoxyribonucleotide transferase (TdT). RNAs which do not have the 'poly-A tail' that this provides are prone to rapid degradation by cellular enzymes; the importance of this system in malaria parasites has been demonstrated in *P. gallinaceum* by Golightly et al (2000). Gene expression levels are therefore dependent on both the strength of the promoter sequences 5' of the open reading frame, and the stability of the mRNA conferred chiefly by the 3' sequences.

1.5.2 *P. falciparum* gene expression

The detailed structure of eukaryotic promoters varies from species to species, although most share the general features described above. Characterizing promoters in *P. falciparum* has, however, proved difficult, despite the fact that many genes have been found to have a tightly regulated transcription pattern. This is at least partly due to the extreme AT-richness of non-coding sequences, frequently in excess of 90%, making most of the genome look like a possible TATA box. A gene homologous to the TBP of other eukaryotes has been identified, but it is unusually divergent (only 42% identity to the archetypal yeast protein) and has not yet been successfully cloned and expressed to test its function (McAndrew et al, 1993).

Also, although the promoter regions of a number of genes have been characterized (reviewed by Horrocks et al, 1998) no conserved sequence motifs important to their transcription have been unequivocally identified (see, e.g., discussion in Crabb and Cowman, 1996). It may be that there are conserved structural elements in *P. falciparum* promoters that are not necessarily conserved at the sequence level; alternatively, the parasite's transcription mechanism may be highly divergent from that of other eukaryotes. Certainly viral promoters such as that of SV40 which are ubiquitously active in higher eukaryotes are inactive in *P. falciparum,* while conversely *P. falciparum* promoters are inactive in COS-7 cells (Crabb and Cowman, 1996). This interesting situation suggests that the transcription process might be a good target for drug design, as the drugs would be unlikely to affect transcription in the host. Furthermore, in order to express transgenes in *P. falciparum,* 5' and 3' control sequences from the parasite must be placed adjacent to the open reading frames. A small number of such sequences have been identified as suitable for use in transfection (reviewed in Waterkeyn et al, 1999) including ones from two rodent malaria species, *P. berghei* and *P. chabaudi,* demonstrating that there is at least partial conservation of the transcription mechanism within the genus.

1.5.3 The genome project

The *P. falciparum* genome project was initiated in 1996 in recognition of the boost that knowing the entire parasite genomic DNA sequence would give to the search for effective treatments. A parasite clone was selected for sequencing – 3D7, which is widely used in laboratory research while still resembling field isolates phenotypically. The genome was published in an almost complete form in 2002 (Gardner et al, 2002), with the data being made generally available in a convenient form via an online database, PlasmoDB (Bahl et al, 2002). Even before the publication, such data as were complete could be accessed through this and other bioinformatics websites, including the entirety of chromosomes 2 and 3 by 1999 (Gardner et al, 1999; Bowman et al, 1999). The genomic data have greatly facilitated the rapid identification of potentially important genes, the characterization of multi-gene families, and the description of general characteristics of the *P. falciparum* genome. Additionally, a number of studies have been carried out to look at gene expression across the whole genome; for example looking at stage-specific mRNA levels using microarrays (Mamoun et al, 2001) or the presence of proteins by mass spectrometry (Florens et al, 2002).

1.5.4 Findings related to gene structure and organization

P. falciparum has 14 chromosomes, varying in length from 0.643 to 3.29 million bases. Chromosomal structure is substantially similar to that of other eukaryotes, with telomeres at each end and a centromeric region in the middle which contain no predicted open reading frames. This similarity persists at the level of individual genes, with each protein coding region being flanked by its own control sequences (as opposed to prokaryotes, where several proteins are often expressed from a single 'operon') and introns being present in many genes. However, a major distinctive feature of the genome is its AT-richness, with in total 80.6% of the sequence made up of these two bases. This is extraordinarily high, especially considering that coding regions require a greater proportion of Gs and Cs in order to code for all amino acids; many non-coding sequences are therefore over 90% AT. Interestingly, across the genome, long poly-dA and poly-dT tracts seem to occur more frequently than would be expected in the regions immediately 5' of open reading frames (Horrocks et al, 1998), suggesting that they might have some role in transcriptional regulation.

In sequencing projects of other organisms, probable functions of genes can often be inferred from the similarity of their translated protein sequences to known genes from other species, but in *P. falciparum* over 60% of the 5,300 predicted open reading frames have no homologues in any gene database. The AT-bias of the genome may have contributed to this problem by influencing the evolution of *P. falciparum* protein sequences to diverge from other organisms. Deciphering the functions of all these genes will be a huge challenge.

1.5.5 Sub-telomeric regions

Another interesting feature of the *P. falciparum* genome is the conserved structure of the sub-telomeric regions on all 14 chromosomes. Telomeres form the very ends of each chromosome, but beneath these are found several different telomere-associated repetitive elements (TAREs), each TARE being made up a of short consensus sequence repeated up to hundreds of times, with some alterations in the repeats. The length of each TARE varies from chromosome to chromosome but all are always present, leading to the suggestion that they may have a role in chromatin structure or in segregation during DNA replication. While screening a genomic library for sequences which enhance the efficiency of transfection, O'Donnell et al (2002) found that a section of TARE 6 – also called *rep20* – causes plasmids containing it to associate with chromosome ends during schizogony, thereby increasing the efficiency of segregation. This sequence may therefore be an anchor point for proteins responsible for separating the sister chromatids during mitosis.

Immediately centromeric of the TAREs are found members of four multigene families – *var*, *rif*, *stevor* and *varC*. As noted in section 1.3.2, *var* encodes parasitized erythrocyte surface proteins which mediate cytoadherence and undergo antigenic variation. Less is known about the other families, but intriguingly *rif* gene products (called Rifins) are also thought to be transported to the erythrocyte surface; they are labelled by surface radioiodination of parasitized erythrocytes and are lost when intact parasitized erythrocytes are treated with trypsin (Kyes et al, 1999). Additionally, genetically identical parasite clones express different subsets of Rifins, suggesting that they may be undergoing antigenic variation. Rifins are much smaller than PfEMP-1 – 35-45 kDa rather than 250-350 kDa – but much more numerous; there are over 200 putative Rifin genes in total in the genome, with several at every chromosome end. Their presence in such numbers implies that they must fulfill an

important role, and their presence on the erythrocyte surface suggests that they may be potential vaccine candidates. The much smaller *stevor* family (18-20 genes) is closely related to *rif,* and its proteins have recently been localized to the Maurer's clefts (Kaviratne et al, 2002). Like *rif,* the function of *stevor* is as yet unknown. The *varC* genes show homology to the C-terminal domain of *var* and are thought to encode nuclear proteins (Bischoff et al, 2000).

The presence of three multigene families involved in producing proteins which are exported to the erythrocyte plasma membrane in the sub-telomeric regions has led to speculation that some of the TAREs have some functional importance for gene regulation. One possibility is that they in some way promote antigenic variation, either by providing binding sites for the proteins which activate or suppress particular genes, or by promoting recombination to generate new protein sequences.

1.6 General Aims

Among the many important areas awaiting more exploration in *P. falciparum*, molecular genetics is one of the foremost. It underlies the entire biology of the organism and yet is relatively poorly understood compared with other eukaryotes, despite the information provided by the genome project. Transfection is one of the most powerful tools for the molecular geneticist, and being able to employ it effectively will be crucial in making best use of the genome data, particularly in elucidating the mechanisms of gene regulation and the functions of the many genes whose role is unknown. Therefore this project has the following aims:

A) To optimize current transfection techniques such that the efficiency of transfection is improved.
B) To investigate new approaches to transfection which bypass the efficiency problem.
C) To develop an assay appropriate to measuring gene expression levels.
D) To use transfection to investigate transcriptional control mechanisms and gene functions.

Fig 1.1: **Simplified view of the life cycle of malaria parasites.** In *P. falciparum,* the mammalian host is human and the mosquito species involved is *A. gambiae.* See text for a full description of the cycle. Figure prepared by Frank Johnson (NIMR photographics).

Plasmodium Life Cycle

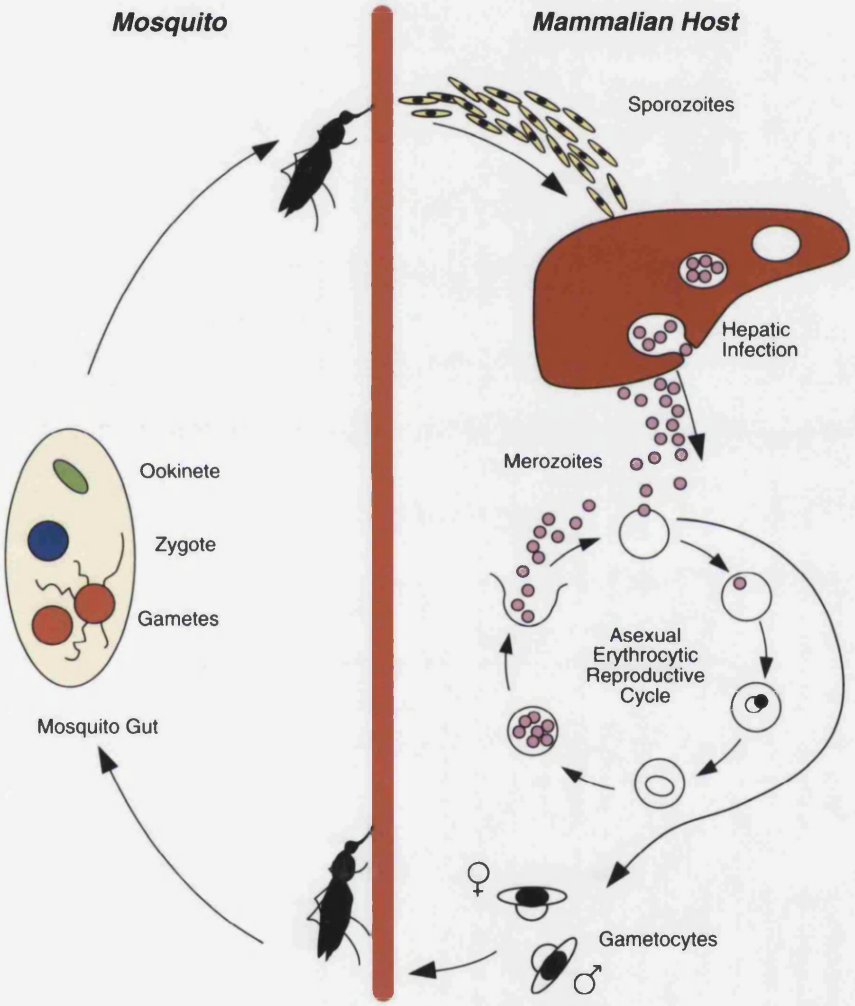


Fig 1.2: **Morphological features of the erythrocytic forms of *P. falciparum.*** Diagrams by L. H. Bannister (Bannister et al, 2001). For the trophozoite stages (**B.** & **C.**) the erythrocyte and parasitophorous vacuole that surround the parasite have been omitted for clarity. The cytostomes labelled in these diagrams are the sites at which the parasite ingests material from the host erythrocyte.
Black silhouettes show the approximate relative sizes of the different forms.

A. Merozoite.
B. Early 'ring' stage of the trophozoite.
C. Mature trophozoite.
D. Schizont.

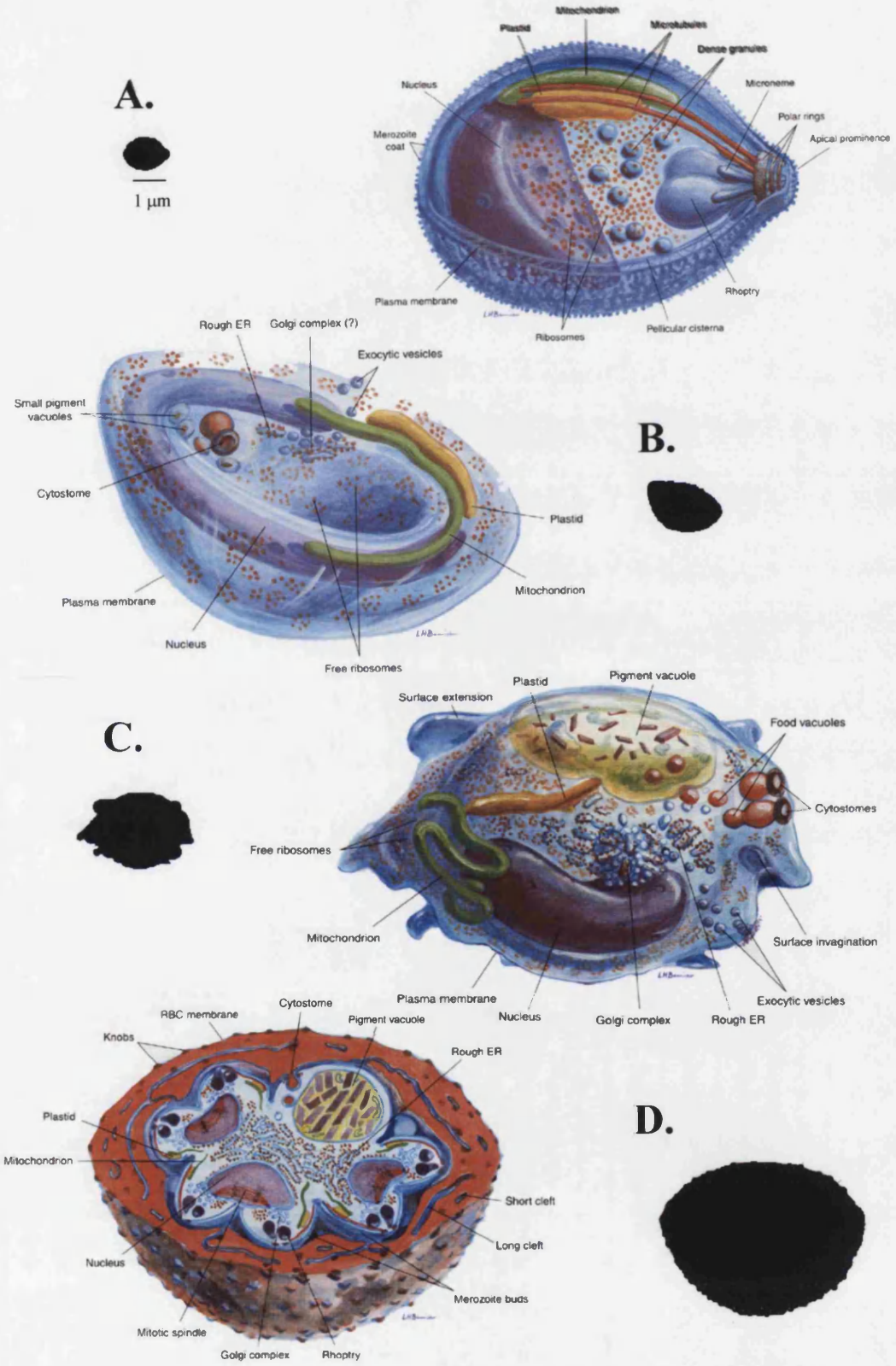
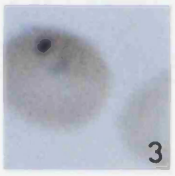
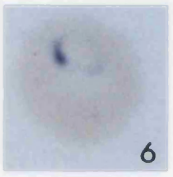
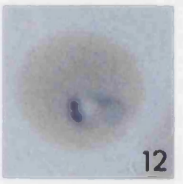

A.
1 μm
Mitochondrion
Plastid
Microtubules
Dense granules
Nucleus
Microneme
Polar rings
Merozoite coat
Apical prominence
Plasma membrane
Rhoptry
Ribosomes
Pellicular cisterna
B.
Rough ER
Golgi complex (?)
Exocytic vesicles
Small pigment vacuoles
Cytostome
Plastid
Mitochondrion
Plasma membrane
Nucleus
Free ribosomes
C.
Plastid
Pigment vacuole
Surface extension
Food vacuoles
Cytostomes
Free ribosomes
Mitochondrion
Surface invagination
Plasma membrane
Nucleus
Golgi complex
Rough ER
Exocytic vesicles
D.
Cytostome
RBC membrane
Pigment vacuole
Knobs
Rough ER
Plastid
Mitochondrion
Short cleft
Long cleft
Nucleus
Merozoite buds
Mitotic spindle
Golgi complex
Rhoptry

Fig 1.3: **Giemsa stained *P. falciparum* intra-erythrocytic forms in culture.** Figure from Freeman & Holder (1983). This series of sixteen photographs follows the development of asexual stages within the erythrocyte. The numbers at the bottom right-hand corner of each picture indicate the number of hours after invasion at which each photograph was taken.

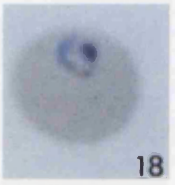

3


6


9


12

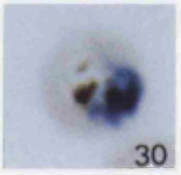

15


18


21


24

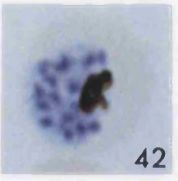

27

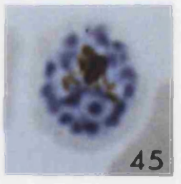

30

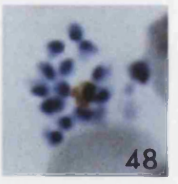

33


36


39


42


45


48

2. *Materials and Methods*

2.1 General Materials and Hardware

2.1.1. Chemicals

Throughout the project, general chemicals and ingredients for buffers were obtained from Sigma unless otherwise stated.

2.1.2. Solutions

Commonly used solutions included:

TE	10mM Tris-HCl, 1mM EDTA in H_2O, pH 7.6
1 x TBE	90mM Tris-Borate, 2mM EDTA in H_2O, pH 8.0
1 x SSC	150mM NaCl, 330mM Na citrate in H_2O, pH 7.0
PBS	137mM NaCl, 3mM KCl, 8mM Na_2HPO_4, 1.5mM KH_2PO_4 in H_2O

Solutions specific to particular protocols are described in the relevant sections.

2.1.3. Centrifuges

When handling parasites, a Sigma 4-10 bench-top centrifuge with rotor Nr.11140 was used (cf. sections 2.2.2, 2.2.3, 2.3.4, 2.4.1, 2.6.1, 2.6.2, & 2.8.1). For smaller volumes, a Sigma 112 microcentrifuge with rotor Nr. 12026 was used (cf. sections 2.2.4 & 2.4.1).
In all other sections, the microcentrifuge used was a Heraeus Biofuge Pico with rotor #3325 B, except when refrigeration was required; in these cases the model used was a Sigma 1K15 with rotor Nr.12024-H.
E. coli cultures were spun down in a Beckman J6-B refrigerated centrifuge with rotor TY.JS 5 and disposable centrifuge tubes obtained from Corning.
Ultracentrifugation of proteins (section 2.8.1) was carried out using a Beckman TL-100 ultracentrifuge with rotor TLS-55 and polycarbonate tubes (Beckman #349622)

2.2. *Plasmodium falciparum* Culture

2.2.1. Parasites and Media

Clone 3D7 parasites were originally obtained from D. Walliker.
Media:

RPMI w/o - RPMI 1640 enriched with 24mM $NaHCO_3$, 25mM HEPES, 25μg/ml gentamycin, 16μg/ml hypoxanthine, and 1.6 mg/ml Glucose, supplied by GibcoBRL.

RPMI Alb - As above supplemented with 5% w/v Albumax I (GibcoBRL)

RPMI HS - As above supplemented with 10% v/v pooled AB+ human serum, obtained from the National Blood Transfusion Service.

To all media, L-glutamine was added to a final concentration of 0.1% w/v just before use, after which unused medium was stored at 4°C for no more than 1 month.

Human erythrocytes were obtained from the National Blood Transfusion Service, and stored at 4°C for no more than 1 month. The blood group varied but group B blood was avoided, and usually blood was not washed before use.

2.2.2. Continuous culture

The culture method used was a modification of that of Trager and Jensen (1976). *P. falciparum* asexual stages were maintained in human erythrocytes at up to 15% parasitaemia and 0.5-1% haematocrit in RPMI Alb. The maximum culture volume was 250ml per 600ml tissue culture flask. Cultures were gassed with a mixture of 7% CO_2, 5% O_2 and 88% N_2 and incubated at 37°C. When it was necessary to replace the medium (every 2 days for a culture at 5% parasitaemia), the parasitized erythrocytes were spun down at 2000rpm/815g for 3 minutes, and the old medium aspirated off.

Cultures were analysed by taking thin blood smears, staining with Giemsa's stain diluted 1:10 in water for 10 minutes, and viewing by light microscopy.

2.2.3. Synchronization

Solutions:

Sorbitol - 5% w/v sorbitol in PBS. Stored at 4°C.

70% Percoll - Percoll (Amersham Biosciences) was mixed 9:1 with 10 x PBS, and 7 volumes of this solution added to 3 volumes RPMI w/o. Stored at 4°C.

Crude synchronization was maintained by centrifuging the cells at 2000rpm/815g for 3 minutes, aspirating off the culture medium, and resuspending the pellet in sorbitol to lyse the late stages. After 10 minutes incubation at 37°C, the cells were centrifuged again and the sorbitol replaced by fresh culture medium.

Parasites were tightly synchronized by the Percoll method (Freeman and Holder, 1983). Parasites were centrifuged as above, washed twice with RPMI w/o to remove all traces of Albumax, resuspended at 20% haematocrit in RPMI w/o, and centrifuged through 70% Percoll for 11 minutes at 2200rpm/1000g to isolate the schizonts. These were then added to fresh red blood cells and cultured for a period of 1-3 hours at 37°C in a shaking incubator at 250rpm (to reduce the occurrence of multiple invasions of the same erythrocyte) after which any remaining schizonts were lysed by sorbitol treatment.

2.2.4. Cryopreservation

Cryoprotectant: 38% w/v glycerol, 2.9% w/v sorbitol, 110mM NaCl in H_2O, sterilized by passing through a 0.22μm pore filter. Stored at 4°C.

P. falciparum ring stages were stored stably in liquid nitrogen by the method of Phillips and Wilson (1978). Briefly, cultures were grown to a high ring stage parasitaemia, centrifuged at 2000rpm/815g for 3 minutes and washed twice with PBS before resuspending at 50% haematocrit in RPMI Alb. 250μl aliquots of the parasites were transferred to screw topped polypropylene vials and mixed 1:1 with cryoprotectant, which was added dropwise with gentle mixing, after which the vials were plunged into $N_{2(l)}$.

To resuscitate, parasitized erythrocytes were thawed rapidly at 37°C, and microcentrifuged at 1500rpm for 2 minutes. The supernatant was removed and the cells resuspended, centrifuged and resuspended again in sequentially reducing concentrations of sorbitol in PBS: 20%, 10%, 7.5%, 5% and 2.5% w/v. Finally the cells were resuspended in culture media.

2.3. Transfection of *Plasmodium falciparum*

2.3.1. Preparation of plasmid DNA

Luria-Bertani (LB) broth: 1% w/v Bacto-Tryptone, 0.5% w/v Bacto-Yeast Extract, 170mM NaCl in H_2O, sterilized by autoclaving

LB-Agar: As above with 1.5% w/v Agar

LB-Amp: As LB broth with 100µg/ml Ampicillin

LB-Amp/Agar: As LB broth with 100µg/ml Ampicillin and 1.5% w/v Agar

Circular, supercoiled plasmid DNA was prepared for transfection from *E. coli* glycerol stocks using the QiaFilter plasmid maxiprep kit (Qiagen). *E. coli* were streaked onto fresh LB-Amp/Agar plates, and incubated overnight at 37°C. Individual colonies were used to inoculate 5ml LB-Amp minicultures; after incubation for 7-8 hours at 37°C with shaking at 250rpm, 250µl of these were added to 100ml LB-Amp cultures which were incubated overnight under the same conditions.

Following the kit protocol, the bacteria were centrifuged to a pellet, resuspended and broken down by alkaline lysis. The lysate was rapidly cleared of protein and genomic DNA precipitate by passing it through a QiaFilter cartridge, and the cleared lysate added to a Qiagen Tip 100 column, which retains plasmid. The column was washed and then the plasmid eluted, followed by isopropanol precipitation of the plasmid to complete the purification. The isopropanol pellet of each plasmid was washed with 70% ethanol, air dried for 5 minutes, and then redissolved in 100µl 10mM Tris-HCl buffer, pH 8.0.

Two microlitres of each plasmid solution was diluted 1 in 250 in water and tested for purity and concentration by measuring absorbance at 260 and 280nm on a UNICAM UV1 spectrophotometer. The ratio of these two values, 260/280, was typically between 1.75 and 1.8, the latter being the value for pure double stranded DNA. The concentration

was calculated on the basis that a 50ug/ml solution of double stranded DNA has an absorbance of 1 at 260nm, and was typically between 2 and 5mg/ml in the undiluted plasmid solution. Where applicable, the plasmid solution was diluted with 10mM Tris-HCl Buffer, pH 8.0, to a final concentration of 3.33mg/ml.

2.3.2. Reagent based methods

SuperFect reagent was obtained from Qiagen and employed as recommended by the manufacturer. Five micrograms of plasmid DNA was diluted to a total volume of 150µl in RPMI w/o, to which 20µl SuperFect reagent was added. After vortexing for 10s, the samples were incubated for 5-10 minutes at room temperature and then mixed with 1ml RPMI Alb. This was immediately added drop by drop to 200µl packed cell volume of parasitized erythrocytes, which had been resuspended in 4ml RPMI Alb. After gassing and incubating for 3 hours at 37°C, RPMI Alb was added to bring the total culture volume to 15ml, and the cultures incubated for a further 40 hours before harvesting the cells.

Effectene reagent was also obtained from Qiagen and used as recommended. The procedure was essentially the same as that above, with the exception that in place of adding SuperFect to the diluted plasmid, 8µl Enhancer solution was added followed by 25µl Effectene, in each case incubating at room temperature for 5 minutes before continuing.

2.3.3. Electroporation for transient transfection

Incomplete cytomix: 120mM KCl, 5mM $MgCl_2$, 0.15mM $CaCl_2$, 2mM EGTA,
10mM K_2H/KH_2PO_4, 25mM HEPES at pH 7.6.

Transfection was effected using a Gene Pulser II or Gene Pulser Xcell electroporator with capacitance extender module (Bio-Rad). 15µl of plasmid solution at a concentration of 3.33 mg/ml (=50µg DNA) was made up to 200µl or 400µl with incomplete cytomix. This was mixed with an equal volume of packed parasitized erythrocytes at 5-10% ring stage parasitaemia, and then transferred to a 0.2cm or 0.4cm electroporation cuvette. After pulsing at the selected conditions (cf. chapter 3), the

parasites were recovered from the cuvette using a sterile pasteur pipette, transferred into 15ml pre-warmed RPMI Alb in an 80ml tissue culture flask, and placed immediately at 37°C. Once all parasites in any given experiment had been electroporated, all flasks were gassed and returned to 37°C.

2.3.4. Electroporation of uninfected erythrocytes

Pre-loading of uninfected erythrocytes with plasmid followed the method of Deitsch et al (2001). Erythrocytes were centrifuged at 2000rpm/815g for 3 minutes to form a pellet, which was washed with incomplete cytomix and then used for electroporation exactly as described above in place of the parasitized erythrocytes. After electroporation, the erythrocytes were stored at 4°C for no more than 72 hours before use.

2.3.5. Electroporation for stable transfection

Drug preparation:

Pyrimethamine (Sigma) was dissolved to 10mM in 1% acetic acid and diluted to 200μM in PBS. WR99210 was obtained from Jacobus Pharmaceuticals, dissolved to 20mM in DMSO and diluted to 20μM in RPMI w/o. Each of these stock solutions was stored at 4°C for no more than one month.

To maximise parasite viability, a synchronous culture was maintained at <5% parasitaemia for four days prior to electroporation. On the day, fresh blood was added to reduce the parasitaemia to about 1%. Electroporation was then carried out as described above, but the pulsed parasites were cultured in 10ml of a 50/50 mixture of RPMI Alb and RPMI HS. (N.B. this gives final concentrations of 5% v/v human serum and 2.5% w/v Albumax; the '50/50' medium used by other researchers contains 5% of each.) The medium was replaced every day for the first 7 days post-electroporation, and on alternate days after that. Drug pressure was applied from 48 hours after electroporation, using a final concentration of 2μM pyrimethamine (reduced to 0.5μM after 2 days) or 10nM WR99210, and renewed every time the medium was replaced. On the sixth day post-electroporation, and every sixth day following, ¼ of the parasitized erythrocytes were discarded and replaced with the same amount of fresh erythrocytes.

2.4. Chloramphenicol acetyltransferase assay

2.4.1. Harvesting *Plasmodium falciparum* material

Saponin: 0.15% w/v saponin in PBS, sterilised by passing through a 0.22μm pore filter. Stored at 4°C.

TEN Buffer: 40mM Tris-HCl, 1mM EDTA, 150mM NaCl in H_2O, pH 7.6

A fraction containing chloramphenicol acetyltransferase (CAT) enzyme was obtained as described by Wu et al (1995). Transfected parasite cultures (15ml) were centrifuged as in section 2.1.2, the supernatant removed, washed once with PBS and centrifuged again. The resulting pellets were resuspended in 2.5 volumes of saponin solution, transferred to microfuge tubes and incubated at 37°C for 10 minutes. The tubes were then microcentrifuged at top speed for 5 minutes, the haemoglobin-rich supernatant discarded, and the pellets fast frozen on dry ice and stored at -80°C.

On the day of a CAT assay, pellets were retrieved and resuspended in 1ml TEN buffer. They were then microcentrifuged at top speed for 2 minutes, the supernatant discarded, and resuspended in 120μl 0.25M Tris-HCl, pH 7.6. Next they were subjected to three cycles of fast freeze thawing by transferring them between dry ice and a 37°C water bath. After a final incubation at 65°C for 10 minutes (to inactivate deacetylases) they were centrifuged again at top speed for 2 minutes, the supernatant being the extract containing CAT enzyme. Any extract that was not used immediately was stored at -80°C.

2.4.2. Basic assay procedure

Acetyl-CoA: Acetyl-Coenzyme A was obtained from Calbiochem, dissolved to 10mg/ml in H_2O, and stored at -20°C in 60μl aliquots.

Chloramphenicol: CAT assay grade D-*threo*-[*dichloroacetyl*-1-^{14}C]chloramphenicol (58 mCi/mmol, 25μCi/ml) was obtained from Amersham Biosciences and stored at -20°C in 24μl aliquots.

Basic assay conditions were a modified version of those described by Wu et al (1995).

To 100µl CAT enzyme extract was added 5µl acetyl-CoA solution (final concentration 500µM), and 18µl 0.25M Tris-HCl, pH 7.6. Reactions were started by adding 2µl (=50nCi) ^{14}C-labelled chloramphenicol, and incubated overnight at 37°C.

2.4.3. Analysis of products by thin layer chromatography

CAT assay products were extracted by adding 400µl ethylacetate, vortex mixing for 30s, and then spinning in a microcentrifuge at top speed for 2 minutes. The organic (upper) phases were transferred into new tubes with holes punched in the tops, and the ethylacetate evaporated off in a Savant Speed Vac Concentrator (model no. RH 40-11). Once dry, the products were redissolved in 10µl ethylacetate and spotted drop by drop onto silica coated aluminium thin layer chromatography plates (Whatman cat. no. 4420-221), allowing spots to dry between each application. The plates were then placed in a glass chromatography tank with a 19:1 mixture of chloroform:methanol, which was allowed to run up the plates for 45 minutes to 1 hour. After this the plates were removed, dried and photographed by being placed in a cassette with a sheet of Kodak Biomax MR-1 film overnight at -20°C. When densitometric analysis of the products was required, the plates were placed overnight in a phosphorimager cassette, which was then scanned using a Molecular Dynamics Storm 860 phosphorimager, the data being analysed with ImageQuant 5.0 software.

2.4.4. Analysis of products by liquid scintillation counting

At the end of the assay incubation, the tubes were placed on ice, and as far as possible all subsequent steps were carried out at 4°C. CAT assay products were extracted by adding 300µl mixed xylenes, vortex mixing for 30s, and spinning in a microcentrifuge at top speed for 3 minutes at 4°C. The organic (upper) phases were transferred to fresh tubes where they were back extracted by adding 100µl 0.25M Tris-HCl, pH 7.6, and vortexing and centrifuging as before. This was repeated. After the second back extraction, the organic phases were transferred to scintillation vials to which were added 2ml Ready Safe scintillant (Beckman). The activity of the samples was then measured on a Beckman LS6000IC scintillation counter, with a program set to count ^{14}C only for 1 minute per sample.

2.4.5. Quantitative assay procedure

Assays were set up as in section 2.4.2, but only 30µl parasite extract was used in each reaction (the volume being made up by adding another 70µl 0.25M Tris-HCl pH 7.6) and multiple reactions were set up for each parasite extract. The reactions were incubated at 37°C for a range of specific times before being stopped by xylene extraction, and the resulting time courses used to calculate the initial rate of the reaction as described in chapter 4.5.

2.5. Bioinformatics

2.5.1. Database searching

Searches for similar sequences were carried out using the program BLAST ('Basic Local Alignment Search Tool' - Altschul et al, 1990) with default search parameters on the malaria genetics and genomics server at the National Center for BioInformatics (http://www.ncbi.nlm.nih.gov/projects/Malaria/) and the online database PlasmoDB (http://www.plasmodb.org). Specific intergenic DNA sequences were downloaded from PlasmoDB, and the annotation on PlasmoDB used to locate the probable ends of the non-coding regions.

2.5.2. Sequence analysis and alignment

DNA and protein sequences were analysed using programs from DNAStar, inc. EditSeq was used to search sequences for specific motifs, build DNA constructs *in silico* and where appropriate check that they would be translated correctly. MapDraw was used to identify restriction sites and the predicted fragments to be generated by particular restriction digests.

Multiple sequence alignment of DNA sequences was carried out using ClustalX (Thompson et al, 1997).

2.6. Analysis of Nucleic Acids

2.6.1. DNA Isolation from *Plasmodium falciparum*

Parasitized erythrocytes were centrifuged at 2000rpm/815g for 3 minutes, and the supernatant discarded.

DNA for use as a template for the polymerase chain reaction was isolated using the S.N.A.P. Whole Blood DNA isolation kit (Invitrogen) according to the protocol for whole blood samples of up to 150µl volume.

DNA for Southern hybridization was extracted using phenol/chloroform as follows.

Lysis Buffer: 10mM Tris-HCl, 10mM EDTA, 1% w/v SDS in H_2O, pH 8.0

Parasitized erythrocyte pellets were washed once with PBS and saponin lysed as described in section 2.4.1. After discarding the supernatant, 10 pellet volumes of pre-warmed lysing buffer were added to each pellet. This was mixed by inversion, and then incubated at 60°C for 5 minutes. Proteinase K (Roche) was added to a final concentration of 500 µg/ml, and the samples incubated overnight at 37°C.

The next day each pellet was extracted twice using an equal volume of a 25:24:1 mixture of phenol:chloroform:isoamyl alcohol and once with chloroform, in each case shaking for 30s to mix the phases and microcentrifuging at top speed for 3 minutes to separate them again. The aqueous (upper) layer was then transferred to a fresh tube and microcentrifuged at top speed for 10 minutes to pellet any remaining solid material. To the resulting supernatant (in a fresh tube) was added 1/10 volume 3M sodium acetate and 2 volumes ethanol, and after mixing the tubes were placed at –20° for 1 hour. They were then centrifuged at top speed for 20 minutes at 4°C, and the supernatant discarded. The DNA pellet was washed once with 70% ethanol, air dried and gently redissolved in 120µl of TE buffer before storing at 4°C.

2.6.2. RNA Isolation from *Plasmodium falciparum*

Parasitized erythrocytes were centrifuged at 2000rpm/815g for 3 minutes, and the supernatant discarded. RNA was then extracted by a modification of the method of Kyes et al (2000).

Trizol reagent (Invitrogen) was pre-warmed to 37°C, and 10 pellet volumes were added to an erythrocyte pellet containing ring or trophozoite stage parasites; if the parasites were at schizont stage, 20 pellet volumes were used. This was mixed thoroughly and incubated at room temperature for 5 minutes; in some cases it was stored at 80°C at this point for up to a month, and thawed to continue. In either case 1ml aliquots of the mixture were transferred to microcentrifuge tubes.

Each tube had 200µl chloroform added, was vortexed for 15s, stood at room temperature for 3 minutes, and then microcentrifuged at top speed and 4°C for 30 minutes. The aqueous (upper) layers were then removed to fresh tubes and to them were added isopropanol to give an aqueous:isopropanol ratio of 6:5. These tubes were incubated at 4°C for at least 1 and up to 48 hours, and then microcentrifuged at top speed and 4°C for 30 minutes. The tubes had the supernatants carefully removed, were air dried for 5 minutes, and the RNA pellets redissolved in 20µl formamide. Samples were heated to 60°C for 10 minutes to drive off any remaining isopropanol, and the formamide transferred to new tubes.

RNA integrity was checked by running 1µl of each sample on a 1% w/v agarose gel (see below) made up in an electrophoresis tank which had been cleaned with RNase ERASE (ICN Biochemicals). Presence of two clear rRNA bands was taken to indicate that the RNA was intact. When required, RNA purity and concentration were determined by diluting 1µl of the sample in 1ml H_2O, and measuring absorbance at 260 and 280nm on a UNICAM UV1 spectrophotometer. A 260/280 ratio of 2.0 indicates pure RNA, and concentration was calculated on the basis that 40µg/ml single stranded RNA gives an absorbance of 1.0 at 260nm.

2.6.3. Agarose Gel Electrophoresis

DNA loading dye:	30% v/v glycerol, 10mM EDTA, 0.025% w/v of each of bromophenol blue & xylene cyanole solution in H_2O
DNA markers:	1kb or 100bp ladder DNA molecular weight markers (Invitrogen) diluted to 10ng/µl in TE mixed 4:1 with DNA loading dye

Separation of DNA fragments was by a modification of the method described by Sambrook et al (1989). Horizontal minigels were made up by dissolving Agarose MP (Roche) with heating into 50ml 0.5 x TBE buffer, typically to a concentration of 1% w/v,

and allowing the gel to set in tanks and combs obtained from Anachem-Scotlab. Larger gels (e.g. for Southern and Northern hybridisation) were of 125ml volume, made with 1 x TBE, and poured into a midigel tank (Anachem-Scotlab). DNA samples were mixed 4:1 with DNA loading dye, loaded on to the gel and electrophoresed at a constant voltage of 6-10 V/cm. applied by a EPS 500/400 DC Power Supply (Pharmacia). When sufficient separation had occurred, as displayed by the movement of the dyes, the gel was soaked in a tank containing 10μg/ml ethidium bromide for 5 minutes and visualized using a UV transilluminator.

2.6.4. Southern Hybridization

CHURCH buffer: 7% w/v SDS, 0.5M Na_2HPO_4/NaH_2PO_4 in H_2O, pH 7.6

DNA fragments were separated on a large 1% w/v agarose gel as above, and then transferred to a Hybond-N^+ (Amersham Biosciences) nylon filter as described by Sambrook et al (1989). Briefly, large DNA fragments were broken up by applying 1200J from a UV Stratalinker 1800 (Stratagene), and the gel soaked in 0.5M NaOH. It was then placed face down on top of a wick made of two large sheets of Whatman 3MM paper on a solid support, the ends of the wick being in a tank of 0.5M NaOH. A sheet of the nitrocellulose was cut to size, dampened and placed on top of the gel, followed by three similarly cut sheets of Whatman 3MM paper, a stack of paper towels, and a weight of about 500g. This was left for at least 4 hours and at most overnight, during which time the sodium hydroxide in the tank moved up into the paper towels by capillary action, transferring the DNA to the filter in the process. At the end of the transfer, the filter was neutralised by soaking in 2 x SSC buffer for 10 minutes, air dried, and cross-linked by applying 1200J from the UV cross-linker.

Probes for hybridization were prepared in two ways. Oligonucleotide probes were end-labelled by mixing 50pmol of the oligonucleotide with 150μCi Redivue γ-^{32}P-labelled ATP (Amersham Biosciences; 3000Ci/mmol, 10mCi/ml) and incubating with 20U T4 polynucleotide kinase (New England Biolabs) using the buffer supplied at 37°C for 30 minutes. Longer probes were generated using the Prime-It II radiolabel incorporation kit (Stratagene). Here 25ng of a PCR product of 300-900bp length was used as a template, to which random nonamer primers were annealed and Klenow fragment used to generate

probes which were radiolabelled by the incorporation of 30μCi Redivue α-^{32}P-labelled ATP (Amersham Biosciences; 3000Ci/mmol, 10mCi/ml). In both cases the probes were separated from unincorporated radiolabel by passing them through a NICK column (Amersham Biosciences) as described by the manufacturers.

The nylon filter was pre-hybridized by incubating with 10ml CHURCH solution with rotation in a Hybaid MINI 10 hybridization oven at 55°C for 1 hour. The probe was then added to the bottle and incubated at 55°C overnight with rotation. The next day the solution was poured off and the filter washed twice with pre-warmed 0.1% SDS in 2 x SSC at 55°C with rotation for 10 minutes each. The filter was then drained but kept moist by wrapping in cling film, and photographed by placing in a cassette with a sheet of Kodak Biomax-2 film and an intensifying screen at -80°C.

If a filter was to be re-probed it was first stripped by boiling a solution of 0.05 x SSC, 10mM EDTA in H_2O, adding SDS to a final concentration of 0.1% w/v, dropping in the filter and allowing it to cool. This was repeated with fresh buffer and the filter could then be allowed to dry.

2.6.5. Northern Hybridization

All stages of Northern hybridization were carried out similarly to those for Southern, with the following modifications:

All gel apparatus was cleaned before use with RNase ERASE (ICN Biomedicals), or by soaking in 3% v/v H_2O_2 for 10 minutes, rinsing with distilled H_2O and air drying. Guanidine thiocyanate was added to the gel solution to a final concentration of 20mM and mixed thoroughly before setting. Approximately 10-20μg RNA per lane or 2μl RNA molecular weight markers (Invitrogen) was diluted into 10μl formamide with no dye and heated to 60°C for 10 minutes before loading. 5μl DNA loading dye was loaded into an empty lane to monitor the progress of the electrophoresis.

For the transfer to nitrocellulose, 7.5mM NaOH solution was used instead of 0.5M. Probes were generated and hybridized to the filter exactly as described above.

2.6.6. Reverse Transcription Polymerase Chain Reaction

Reverse transcription polymerase chain reaction (RT-PCR) was effected using the RNase-free reagents and protocol supplied in Life Technologies' 5' RACE kit. In this method, RNA samples are first treated with DNase and split to give RT+ and RT- controls. cDNA is synthesized by annealing random hexamer primers or one end specific primer to the RNA and incubating with or without SuperScript II reverse transcriptase. The cDNA is then amplified by PCR as described in the following section.

2.7. Construction of Plasmids

2.7.1. Polymerase Chain Reaction

Amplification of specific DNA sequences was achieved using the polymerase chain reaction (PCR), described in detail by Sambrook et al (1989). Oligonucleotide primers were obtained from Oswel and gDNA or cDNA templates generated as described in sections 2.6.1 and 2.6.6.

Gene fragments for use in plasmids were amplified using Platinum *Pfx* (Life Technologies), a proof-reading, hot start polymerase, for greater accuracy. Reaction conditions were as recommended by the manufacturer with 1 x *Pfx* Buffer, a final primer concentration of 0.3μM and 1.25U Pfx used per 50μl reaction. Optimum amplification from different templates was achieved by varying the final concentrations of $MgSO_4$ (2-4mM) and dNTPs (0.6-1.2mM of each dNTP).

When determining optimum annealing temperatures, screening clones and doing RT-PCR, the polymerase used was Ampli*Taq* with 1 x PCR Buffer II (Perkin Elmer), primers at a final concentration of 1μM, $MgCl_2$ at 1.5-2.5mM, and dNTPs at 200μM. 0.25U Ampli*Taq* enzyme was used per 20μl reaction.

PCRs were carried out on a Hybaid PCR Express thermocycler, with programs similar to the following:

1. 95°C for 9min 30s (*Pfx*) or 1min 30s (Ampli*Taq*)
2. 95°C for 30s
3. (T_m-2)°C for 30s
4. 68°C (*Pfx*) or 65°C (Ampli*Taq*) for 1min/kb of target sequence
5. Repeat steps 2-4 another 34 times
6. 68 or 65°C for 10 minutes
7. Cool to 4°C

T_m is the annealing temperature of the primers, and was calculated as follows:

$$T_m = 81.5°C + 16.6 \times (\log_{10}[J^+]) - 600/n + 0.4 \times GC\%$$

where $[J^+]$ is the concentration of monovalent cations in the reaction (50mM for Ampli*Taq*, proprietary information for *Pfx*), n is the number of bases in the primer, and GC% is the percentage of those bases that are guanine or cytosine. When a template proved difficult to amplify, the temperature at step 3 above was varied by up to 10°C above and below the calculated T_m to find the optimal annealing temperature.

PCR products were separated and checked for correct size by agarose gel electrophoresis (section 2.6.3). When being used for further cloning, the DNA products were purified using the QiaQuick PCR purification kit (Qiagen). DNA was eluted from the columns in 30µl of 10mM Tris-HCl, pH 8.0.

2.7.2. Subcloning

Where PCR products were of low abundance, or from particularly AT-rich intergenic regions where sequence errors were likely, they were subcloned for amplification and sequencing. The TOPO-TA and Zero Blunt TOPO cloning kits for sequencing (Invitrogen) were used for PCR products generated with AmpliTaq and Pfx respectively, following the supplied protocol. Both kits insert the PCR sequence into the vector pCR4, with *Eco*R I sites and sequencing primer start sites flanking the insertion.

2.7.3. Restriction Digestion

Restriction endonucleases were obtained from Roche, with the following exceptions: *Bsg* I, *Bsr*G I, *Psi* I, and *Sal* I were from New England Biolabs; *Bsr*BR I was from Promega. All restriction digests were carried out according to the supplied protocol. Restriction fragments were separated from unwanted DNA by agarose gel electrophoresis (section 2.6.3), the ethidium bromide stained gel being visualized on a non-UV Dark Reader (Clare Chemical Research) and the band corresponding to the desired product excised. DNA was recovered from gel slices using the QiaQuick gel extraction kit (Qiagen) and eluted in 30µl 10mM Tris-HCl, pH 8.0.

2.7.4. Processing and Ligation of DNA fragments

10 x Nick-translation buffer: 0.5M Tris-HCl, 0.1M $MgSO_4$, 1mM DTT, 500µg/ml BSA in H_2O, pH 7.5

Where necessary, sticky ends of DNA generated by restriction digestion were blunted by the method of Sambrook et al (1989). The DNA was mixed with Nick-translation buffer to a final concentration of 1x, dNTPs to 80µM, and 1U of Klenow fragment. After incubation for 30 minutes at room temperature, the reaction was stopped by adding EDTA to a concentration of 20mM.

5'-dephosphorylation of DNA fragments was carried out using alkaline phosphatase obtained from Roche by the supplied protocol. Phosphorylation of PCR products for direct cloning used T4 polynucleotide kinase (New England Biolabs) and the supplied protocol. When multiple processing steps were required, the DNA was re-purified after each enzyme reaction using the QiaQuick PCR purification kit (Qiagen).

Ligations were carried out using T4 DNA ligase (New England Biolabs) following the supplied protocol. The concentration of each DNA fragment was estimated, and the fragments mixed to give a vector:insert molar ratio of approximately 1:3 (or 1:3:3, for triple ligations.) Incubation was overnight at 16°C for sticky ends or 4°C for blunt ends.

2.7.5. Transformation of plasmids into *E. coli*

Luria-Bertani (LB) broth: 1% w/v Bacto-Tryptone, 0.5% w/v Bacto-Yeast Extract, 170mM NaCl in H_2O, sterilized by autoclaving

LB-Agar: As above with 1.5% w/v Agar

SOB medium: 2% w/v Bacto-Tryptone, 0.5% w/v Bacto-Yeast Extract, 8.5mM NaCl, 2.5mM KCl in H_2O, sterilized by autoclaving, with $MgCl_2$ added to a final concentration of 10mM just before use.

TB Buffer: 10mM PIPES, 15mM $CaCl_2$, 250mM KCl, 55mM $MnCl_2$ in H_2O, pH 6.7, sterilized by passing through a 22μm pore filter and stored at 4°C.

TOPO cloned PCR products and ligations were transformed into INVα or TOP10 One Shot chemically competent *E. coli* (Invitrogen) following the manufacturer's protocol, and then spread on LB-Agar plates containing 100μg/ml Ampicillin.

Where long, highly AT-rich sequences were being cloned, *E. coli* strain PMC-103 (a gift from D. Baker) was used instead. To make them competent, a single colony was used to inoculate 5ml of LB broth, which was incubated overnight at 37°C with shaking at 250rpm. Forty millilitres of SOB medium was inoculated with 400μl of the overnight culture, and incubated overnight at 18°C with shaking at 220rpm. Next the culture was chilled on ice for 10 minutes, transferred to polypropylene centrifuge tubes and spun at 3000g for 10 minutes at 4°C. After discarding the supernatant, the pellet was resuspended in 12.8ml TB Buffer, held on ice for 10 minutes, spun down as before and resuspended in 3.2ml TB Buffer. To this was added 240μl DMSO dropwise with gentle mixing, and the cells were again held on ice for 10 minutes. They were then divided into aliquots in cryotubes, fast frozen in $N_{2(l)}$, and stored at -80°C.

To transform competent PMC-103 *E.coli*, 5μl of a ligation was added to 100μl thawed cells, held on ice for 1 hour, heat shocked at 42°C for 45s, and spread on an LB-Agar plate containing 100μg/ml Ampicillin.

2.7.6. Screening of transformants

Colonies generated by the transformations described above were tested to ascertain whether they contained the correct plasmid as follows: standard Ampli*Taq* PCR reactions were set up, using two primers within the DNA insert, or one in the insert and one in the vector sequence from a given ligation. A number of individual colonies from each plate of transformants were picked using a sterile loop, scraped on the inside of a PCR tube, and then transferred to a culture of 3.5ml LB broth containing 100µg/ml Ampicillin. Negative control PCR reactions containing no colonies were also set up, and the cultures were incubated overnight at 37°C, 250rpm. PCR products were analysed by agarose gel electrophoresis, and those colonies which gave no or incorrectly sized bands discarded.

The cultures from the remaining colonies had 0.5ml mixed with an equal volume of 30% v/v glycerol in LB broth, and stored at -80°C. Plasmid was purified from the other 3ml of the culture using the S.N.A.P plasmid miniprep kit (Invitrogen), from which plasmid DNA was eluted in 60µl TE. These plasmids were further checked for integrity by restriction digests and/or sequencing.

2.7.7 DNA Sequencing

Sequencing Dye: Deionized formamide mixed 5:1 with 25mM EDTA, 50mg/ml Blue dextran in H_2O.

Sections of DNA of up to 600 bases at a time were sequenced by the dye terminator method (Sanger et al, 1977). All sequencing reagents, hardware and software were supplied by PE Biosystems. Cycle sequencing reactions contained 8µl dRhodamine terminator mix, 50-500ng purified template DNA, and 3.2pmol of the sequencing primer in a total volume of 20µl. Reactions were carried out on a PTC-100 thermocycler (MJ Research) with the following program:

1. 96°C for 45s
2. 50°C for 30s
3. 60°C for 2 minutes

4. Repeat steps 1-3 24 more times
5. Cool to 4°C

After cycling the products were purified by adding 2µl 3M NaOAc and 50µl ethanol to each reaction and microcentrifuging at top speed for 25 minutes at 4°C. The supernatants were carefully removed, and the pellets washed by adding 75µl 70% ethanol and centrifuging at top speed for 10 minutes at 4°C. Once more the supernatants were carefully removed, and the pellets air dried for 5 minutes before resuspending in 2µl Sequencing Dye, and stored at -20°C if necessary.

The samples were loaded on a sequencing gel which was run in an ABI Prism 377 DNA sequencer by Irene Ling, and the data were analysed using ABI Sequence Analysis 3.4 software.

2.8. Analysis of Proteins

2.8.1. Protein isolation from *Plasmodium falciparum*

NP40 Buffer: 1% v/v Nonidet P40 (BDH Laboratory Supplies), 5mM EDTA, 5mM EGTA, 1mM PMSF, 1x Complete™ protease inhibitor cocktail (Roche), in PBS

Denaturing Buffer: 1% w/v SDS, 5mM EDTA, 50mM Tris-HCL in H_2O, pH 8.0

Cultures of parasitized erythrocytes for protein extraction were spun down as normal, the medium aspirated off, and the resulting pellets of c. 100µl packed cells stored at –80°C if necessary. After thawing the pellets were resuspended in 1ml NP40 buffer, and left on ice for 30 minutes with occasional vortexing. They were then ultracentrifuged at 55,000rpm for 45 minutes at 4°C, at the end of which the supernatants were removed and stored in 200µl aliquots at –80°C. The NP40 pellets were resuspended in 100µl Denaturing Buffer, vortexed, and heated to 100°C for 5 minutes. After cooling on ice, 900µl NP40 Buffer was added, mixed and left for 10 minutes on ice. These mixtures were ultracentrifuged as before, the supernatants and pellets being stored separately at –80°C.

2.8.2. *In vitro* transcription and translation

Genes were transcribed *in vitro* using T7 RNA polymerase, which along with all other RNase free reagents was obtained from Promega. Transcription reactions contained 20-100ng of a plasmid template including the T7 transcription start site, 60U RNasin, 2U T7 polymerase and final concentrations of 1 x T7 transcription buffer, 10mM DTT, and rNTPs at 2.5mM of each rNTP in a 50µl reaction. After incubation at 37°C for 2 hours, 1µl of each sample was run on a 1% w/v agarose gel to verify the production of RNA, which was stored at -80°C.

RNA was translated *in vitro* using rabbit reticulocyte lysate (Promega). To radiolabel the protein, 45µCi Redivue ^{35}S-labelled methionine (Amersham Biosciences; 1000Ci/mmol, 15mCi/ml) was included in the reaction along with 40U RNasin, 1µl of a 1mM mixture of all the biological amino acids except methionine, 3µl of the RNA transcript generated as described above, and 25µl rabbit reticulocyte lysate in a 50µl reaction. When it was not necessary to radiolabel, unlabelled methionine was used instead. After incubation at 30°C for 1 hour, reaction products were stored at -80°C.

2.8.3. SDS-Polyacrylamide Gel Electrophoresis

Running Buffer:	25mM Tris-HCl, 192mM Glycine, 0.1% w/v SDS in H_2O, pH 8.3, stored at room temperature
Sample Buffer:	125mM Tris-HCl, 21% v/v glycerol, 4.3% w/v SDS, 0.01% w/v Bromophenol blue in H_2O, pH 6.8. Stored wrapped in foil at room temperature, and DTT added to 0.2M just before use.
Coomassie stain:	0.1% w/v Coomassie Brilliant Blue R-250 in a Methanol:H_2O: Acetic Acid mix of volumetric ratio 45:45:10, filtered through Whatman no. 1 paper and stored at room temperature.
Destain:	20% v/v Methanol and 7% v/v Acetic Acid in H_2O.

SDS-PAGE to separate proteins was by a modification of the method of Laemmli (1970). Separating gels were made up containing final concentrations of 12.5% w/v acrylamide/methlyene bis-acrylamide (Bio-Rad), 370mM Tris-HCl pH 8.8, and 0.1% w/v SDS, with polymerization catalysed by adding ammonium persulphate to 0.05% w/v

and TEMED to 0.1% v/v. This mixture was poured into a vertical gel casting apparatus (Hoefer Scientific Instruments) to set for 1 hour, after which a stacking gel was made up similarly but containing 5% w/v acrylamide, and 125mM Tris-HCl pH 6.8. This was poured on top of the separating gel and a comb inserted to create wells before it set. Gel assemblies were then fixed to vertical electrophoresis apparatus and the reservoirs filled with Running Buffer.

Protein samples or prestained protein molecular weight markers (Bio-Rad) were mixed with the same volume of Sample Buffer, heated to 100°C for 5 minutes, and loaded into the wells. A constant current of 20-25mA was applied until the blue dye had reached the bottom of the gel, after which it was removed from the apparatus. Protein could be visualised by soaking the gel in Coomassie stain for 10 minutes and then Destain for 1 hour or more.

2.8.4. Western blotting

Transfer Buffer: 20% v/v Methanol, 150mM Glycine, 20mM Tris base in H_2O
Blocking Solution: 5% w/v BSA in PBS
Wash Solution: 0.5% v/v Tween-20 in PBS

An SDS-polyacrylamide gel was run as described above, and a piece of Hybond-C nitrocellulose (Amersham Biosciences) cut to the same size as the gel, with two pieces of Whatman 3MM paper slightly larger. The nitrocellulose and Whatman pieces were briefly soaked in Transfer Buffer before making a sandwich with the gel and the nitrocellulose between the Whatman paper pieces. This sandwich was placed in a cassette and positioned vertically inside a Hoefer Scientific Instruments TE Series electrophoresis tank, such that the nitrocellulose was on the same side of the gel as the anode. A constant voltage of 50V was applied overnight, or 75V for 4 hours.

The nitrocellulose was then removed and incubated at room temperature with Blocking Solution for 30 minutes with gentle agitation, after which the primary antibody was added to the solution at the desired concentration. This was incubated as before for 2 hours, after which the antibody solution was removed and the nitrocellulose incubated with Wash Solution for 3 x 5 minutes. Next a secondary antibody-horseradish peroxidase conjugate was added, diluted to the desired concentration in Blocking

Solution, and incubated as before for 1 hour. Finally the nitrocellulose was incubated with Wash Solution for 5 x 5 minutes, and then with PBS for 5 minutes.

Antibody binding was visualized by the ECL method, using the Western Blotting Detection Reagents (Amersham Biosciences #RPN2106) according to the manufacturer's recommendations. Kodak Biomax MR film was exposed to the freshly treated nitrocellulose for between 1 and 15 minutes.

2.8.5. Immunoprecipitation

Wash Buffer I: 50mM Tris-HCl, 5mM EDTA, 0.5% v/v Triton X-100, 1mg/ml BSA, 0.5M NaCl in H_2O, pH 8.2. Stored at 4°C.

Wash Buffer II: 50mM Tris-HCl, 5mM EDTA, 0.5% v/v Triton X-100 in H_2O, pH 8.2. Stored at 4°C.

To immunoprecipitate proteins, 5ml Protein G Sepharose 4 Fast Flow (Amersham Biosciences) was first washed three times with PBS and resuspended in 25ml Wash Buffer II. Before use this mixture was resuspended thoroughly, and a 100µl aliquot transferred to a microfuge tube for each sample being tested. Immediately before use, these tubes were microcentrifuged at top speed for 2 minutes, and aspirated dry.

A suitable quantity of radiolabelled protein was made up to a volume of 50µl with PBS. One microgram of monoclonal antibody was added, stirred with a pipette tip, and the mixture incubated for at least 2 hours at 4°C with gentle rotation. The sample was then microcentrifuged at top speed for 20 minutes at 4°C, the pellet discarded and the supernatant added to an aliquot of Protein G Sepharose (see above). This was then mixed by gentle rotation at 4°C for 2 hours, before microcentrifuging at top speed and 4°C for 1 minute. The supernatant was removed and stored at -80°C as an unbound protein control.

The sepharose pellet was then washed by adding 500µl wash buffer, resuspending, microcentrifuging at top speed and 4°C for 1 minute, and aspirating off the supernatant. This was done 3 times with Wash Buffer I and 4-6 times with Wash Buffer II. Finally the pellet was resuspended in an equal volume of SDS-PAGE sample buffer, and run on an acrylamide gel as described in section 2.8.3. The gels were not Coomassie stained,

but soaked in Amplify (Amersham Biosciences) for 10 minutes, dried, and photographed by placing in a cassette with Kodak Biomax MR-1 film overnight.

2.8.6. Immunofluorescence analysis

Parasite cultures were spun down as described in section 2.2.2, the medium aspirated off, and the pellet resuspended in PBS to 1-2% haematocrit. Aliquots of 8µl of this suspension were added to each well of 15 well multitest slides (ICN biomedicals) and allowed to air dry. The slides were fixed by immersing in a 50:50 mixture of acetone and methanol for 1 minute at room temperature, and allowed to dry again.

Both antibodies were diluted as required using 1% w/v BSA in PBS. To each well was added 8µl of primary antibody solution, which was incubated on the slide for 1hr at 37°C in a humid chamber. This was then rinsed off with PBS, and the slide washed twice with PBS for 5 minutes with gentle agitation. After air drying until damp, 8µl of secondary antibody solution was added to each well, and incubated and washed off as before. Parasite nuclei were visualized by immersing in DAPI stain (0.01 mg/ml in PBS) for 5 seconds and then washing once with PBS as before. A drop of Citifluor AF1 (Citifluor Ltd.) was added to each well, and the slide covered with a large coverslip. Slides were viewed on a Zeiss fluorescence microscope.

3. *Optimization of transfection technique*

3.1. Introduction

Transformation of *Plasmodium falciparum* is a very inefficient process, as described in section 1.3.3, with at best 0.1% of parasites receiving DNA. The homologous recombination event required to produce a stable transfectant is even rarer, occurring in no more than 1 in a million transfected parasites (B. Crabb, personal communication). Conceptually, a simple physical explanation for the problem is that, while in the transfection of most cells new DNA has to cross two membranes to reach the nucleus (the plasma membrane and the nuclear membrane of the cell), as an intracellular parasite *P. falciparum* is surrounded by two more: the parasitophorous vacuolar membrane and the erythrocyte plasma membrane. For this reason the first successful malaria transfections used extracellular parasite forms: the female gamete and fertilized zygote of *P. gallinaceum* (Goonewardene et al, 1993). Attempts to transfect the only extracellular form of the asexual blood stage parasite, the merozoite, have been confounded by the fragility and short lifespan (30 minutes or less in *P. falciparum*) of this form.

This low transfection efficiency makes many potentially valuable transient transfection experiments impossible, as phenotypic changes in those parasites that do receive DNA are masked by the vast majority of parasites which do not; additionally, plasmids do not segregate well during schizogony, so as the culture grows transfected parasites are rapidly diluted out. To date the only successful transient transfection strategies have been expression of the reporter enzymes chloramphenicol acetyltransferase (CAT - Wu et al, 1995) and luciferase (Fidock and Wellems, 1996), both of which give an activity detectable even when only a small number of parasites express them. Van Wye and Haldar (1997) reported the transient expression of Green Fluorescent Protein (GFP), but other researchers attempting to do the same have been unable to obtain expression levels high enough to detect.

The alternative approach of using a drug resistance gene to select for transfection and thus grow a 100% transfectant culture has also been used successfully and is examined in detail in chapter 6. It should be noted at this point, however, that this approach is slow

(typically requiring 3-5 weeks to obtain the transfectant culture), unpredictable, and not appropriate to every kind of experiment; for example, expressing or repressing a gene in such a way that parasite viability is adversely affected gives a negative result that is indistinguishable from a technical failure.

Clearly a substantial improvement in existing transfection techniques is required before *P. falciparum* will be truly amenable to genetic manipulation. In this study the only effective transfection technique yet characterised, electroporation, was optimised, and several other techniques which have proved useful in other systems were explored.

3.2. Lipid and Polyamidoamine based transfection reagents are ineffective when used with *P. falciparum*

In many mammalian cell systems, transfection is facilitated by first incorporating the DNA into liposomes, which fuse with the cell membrane and release their contents into the cell. An alternative approach is to package the DNA with a multiply positively charged polymer, which will neutralise the multiple negative charges of the DNA's sugar-phosphate backbone and therefore enable it to cross the uncharged lipid bilayers of cell membranes. Polyamidoamine dendrimers are a kind of molecule used in this latter approach, and Mamoun et al (1999b) reported success in using one such product, SuperFect (Qiagen), to transfect *P. falciparum* parasitized erythrocytes.

SuperFect and a liposome based reagent, Effectene (Qiagen), were compared with electroporation. The two reagents were applied as recommended by the manufacturers to pHC1-CAT plasmid DNA, which was then added to mixed cultures, or synchronised cultures of young rings, mature trophozoites, and schizonts. In no case was any CAT activity detected from parasites treated with SuperFect or Effectene, while electroporated parasites consistently showed strong activity (Fig 3.1).

The report which originally described the use of SuperFect (Mamoun et al, 1999b) commented that variation in efficiency was seen between batches of the reagent, and concluded that further investigation was required to make it a routine technique. The fact that no work published since then has used it perhaps indicates that it works neither reliably nor predictably enough to be useful in this system.

3.3 Optimization of electroporation conditions

Electroporation involves applying a short-lived but strong electric field to live cells in a suspension that also contains DNA molecules. By a poorly characterised process, the electric field causes pores transiently to open in the cell membranes through which DNA may be absorbed. This is a fairly harsh procedure and a proportion of the electroporated cells lyse. Modern electroporation devices allow variation in the characteristics of the electric field applied (Voltage, capacitance, resistance and distance between electrodes) in order to obtain the optimal balance between pore formation and cell survival.

At the start of this project, electroporation conditions for *P. falciparum* based on those used for bacterial cells had been established – 2000V, 25μF, 200Ω with a gap between electrodes of 0.4cm (Wu et al, 1995). These were used as a positive control for the following tests.

Mammalian cells are frequently electroporated using a pulse with much lower voltage, higher capacitance, and a smaller gap between electrodes (0.2cm). Over a wide range of these conditions, transfection efficiency was compared by measuring CAT activity generated following electroporation with the plasmid pHC1-CAT, and further electroporations were carried out to define tightly the optimal conditions. The results are summarised in figs 3.2 and 3.3.

Best results were obtained using 300V, 940-950mF (the exact value varied from experiment to experiment), conditions which typically gave time constants between 10 and 16ms. These values substantially agree with those published by Fidock and Wellems (1997). Voltage seemed to be the more critical of the two parameters, with a significant drop in efficiency if it varied more than 30V above or below the optimum (fig 3.2), while a wider range of capacitance had little effect on efficiency (fig 3.3a). It should be noted that the figures given for voltage and capacitance are those measured by the electroporator during the pulse itself, rather than the values to which the machine was set, which differed at times by up to 50V or 0.1 mF respectively. To duplicate the results on a different machine, the settings must be determined which give these actual values when the pulse is delivered.

The maximum CAT activity detected was slightly higher than that for the old conditions, even though the smaller gap between electrodes meant that half the volume of parasitized red blood cells and half the quantity of plasmid DNA was used in each cuvette compared to the control. Therefore a similar activity actually corresponds to a fourfold increase in transfection efficiency, supported by the observation that using half as much plasmid DNA with the control conditions gave just over half as much CAT activity (fig 3.3b). The CAT assay itself was not reaching saturation, as assays which included commercially purified CAT gave much higher values. In theory it would be possible to apply the optimal conditions to a 0.4cm cuvette by doubling the Voltage (electric field strength = Voltage/distance between electrodes). However, electroporators available today can only apply Voltages over 500V at much lower capacitances.

The increase in efficiency is significant, and so the new conditions were adopted for all subsequent transfections. However, the efficiency of electroporation remains very low; improvements in the order of 10-100 fold or more are needed to address to problems described in the introduction to this chapter. The major benefit of the optimisation described here is in reducing the quantity of parasites and DNA needed per transfection reaction, both of which are labour-intensive to produce.

3.4 Invading parasites can take up DNA from pre-loaded erythrocytes

Erythrocytes have a structurally resilient cell membrane, which can be permeabilized and resealed to form an intact enclosure as in the well established practice of producing erythrocyte 'ghosts'. If erythrocytes could be first loaded with plasmid DNA and then invaded by parasites, the first membrane barrier would already have been overcome and the efficiency of transfection – perhaps with a subsequent electroporation step – might be increased. Deitsch et al (2001) found that they could load unparasitized erythrocytes with plasmid by electroporating them using similar conditions to those for parasitized cells. On invasion, with no further electroporation, expression of the reporter gene luciferase was detected, demonstrating that during either invasion or parasite growth the parasite can take up DNA from the host erythrocyte cytoplasm.

Erythrocytes were pre-loaded by electroporation with a plasmid as described by Deitsch et al. (2001); the cells were washed with incomplete cytomix, spun down, and 300µl of

the pellet mixed with 50μg plasmid DNA in 100μl incomplete cytomix. The resulting mixture was placed in a 0.2cm. electroporation cuvette and pulsed at 300V, 950μF (the optimal conditions described above). These pre-loaded cells were resuspended in RPMI/Albumax medium and stored at 4°C prior to use.

Fresh, purified schizonts were spun down and 5μl of the pellet was added to 200μl of either pre-loaded erythrocytes, or erythrocytes which had been washed but not electroporated. These cells were resuspended in 3ml culture medium and the parasites allowed to invade by incubation at 37°C with shaking for 4 hours. After this period, the cells either had medium added to a final volume of 15ml, or were spun down, electroporated with plasmid following the protocol used in section 3.3, and resuspended in 15ml medium. In both cases the flasks were gassed and then incubated at 37°C until 48 hours following the start of invasion, after which the cells were harvested and assayed for CAT activity. All the flasks had smears taken at the end of the four hour invasion period and just before harvesting.

As shown in fig 3.4, parasites were able to take up DNA from pre-loaded parasites, as described by Deitsch et al (2001). However, in contrast to the published findings, the transfection efficiency was not as good as the optimal conditions for electroporating ring stage parasites, reaching only 20-60% of the CAT activity depending on the exact protocol followed. Interestingly, in erythrocytes which had been loaded with DNA 66 hours prior to invasion, invaded, and then subsequently electroporated again once the parasites had reached ring stage, there was much more evidence of erythrocyte damage and death on the thin blood smear than for a single electroporation. The cell pellet at the time of harvesting these samples was also about half the size of all the others. This indicates that electroporation does damage to erythrocytes which persists for at least 72 hours at 4°C, and is cumulative with later electroporation events. The additional damage resulting from a second electroporation more than outweighs any gain in transfection efficiency due to adding a second aliquot of plasmid for the second electroporation, and the extra opportunity this provides to get the DNA into the cell. Furthermore, the parasitaemia of these samples dropped over the 44 hours of culture time, whereas all the others showed a slight increase, indicating that parasitized erythrocytes were selectively killed. However, erythrocytes which had been pre-loaded with DNA 66 hours prior to invasion showed higher CAT activities than those pre-loaded just 1 hour before invasion, suggesting that there is some recovery of erythrocyte structural integrity over time.

The differences in the relative efficiency of pre-loading versus conventional electroporation observed in this study and the published material could be explained by a variety of factors. Different reporter genes and promoters were used in the two cases, and it is conceivable that the route and/or timing of transfection might favour the expression of one gene when pre-loaded, and the other when electroporated after invasion. Slightly different numbers of cells were used per transfection, which could have an impact if the efficiency of either approach was somehow dose-dependent. Finally, the parasites in this study were cultured at 1% haematocrit, whereas in the published work parasites are routinely cultured at 5% haematocrit. If, as suggested above, electroporating the parasites directly damages them more than if they invade a pre-electroporated erythrocyte, culturing them at a lower haematocrit will give them a greater supply of nutrients and thus more chance to recover.

In summary, pre-loading erythrocytes under the conditions used in this study does not appear to represent a substantial gain in transient transfection efficiency, and as it is slightly more labour intensive than conventional methods was not used routinely. However, the demonstration that parasites can take up DNA on invasion could make this approach valuable for prolonged cultures of transfectants which need new erythrocytes to be added periodically; for example, during the negative selection phase of generating a stable transfectant.

3.5 Morpholino oligonucleotides are not taken up during electroporation

Antisense technology has shown great promise in recent years as a method for investigating gene function, in principle allowing the expression of any gene to be ablated by adding a single stranded copy of the antisense gene sequence, which will then bind to the mRNA, preventing the gene from being translated. The entire mRNA sequence is not required; a short 15-25 base oligonucleotide corresponding to the translational start site on the mRNA is generally sufficient. Short, single-stranded DNAs and RNAs are, however, rapidly degraded in most cells, preventing them from reaching their target mRNA. One solution to this problem has been the development of molecules such as morpholinos, which have a structure similar to that of DNA but with a modified backbone – in the case of morpholinos the normal phosphate bond is replaced by

phosphodiamidate (fig 3.5) – rendering them resistant to biological nucleases (Summerton and Weller, 1997). In many mammalian cells, morpholinos are readily incorporated by 'scrape delivery' – simply scraping adherent cells off a surface causes pores to open through which a variety of small to medium sized molecules, including the morpholino, can enter the cell. Erythrocytes are not so amenable, however; they exist in suspension and have no need of transporting anything larger than ions and O_2 molecules across their plasma membrane in normal circumstances. However, the pores which open during electroporation might be able to transport morpholinos into the cell in addition to the much larger DNA molecules.

These molecules could be useful in genetic studies of *P. falciparum,* and owing to their small size may be easier to absorb than plasmid DNA. To test this, a morpholino oligonucleotide of the sequence TATATCCAGTGATTTTTTTCTCCAT (antisense to the translational start site of the CAT gene) was obtained from GeneTools. To check that it could inhibit CAT expression, it was first tested for function in a cell-free system. The CAT gene was cloned into pCR4-TOPO (Invitrogen), a vector containing a T7 polymerase start site, and *in vitro* transcription was used to obtain CAT mRNA. The mRNA was added to an *in vitro* translation reaction along with no morpholino, a control morpholino of irrelevant sequence (CCTCTTACCTCAGTTACAATTTATA), or the CAT antisense morpholino at a concentration of 440 nM. After incubation the *in vitro* translation products were assayed for CAT activity (fig 3.6a). No activity was seen in the absence of CAT mRNA, and while the control morpholino had no effect on the CAT activity generated in the translation, the CAT antisense morpholino greatly reduced it. Estimates of the initial rates of acetylation seen in each CAT assay suggest that the CAT antisense morpholino reduced CAT activity by greater than 50-fold.

Both morpholinos were then incorporated in electroporations of parasites with CAT expression plasmids, by adding them to the electroporation cuvettes at final concentrations of 10μM and 25μM. It was thought that this would afford the best opportunity for the morpholinos to enter the parasite. However, as shown in fig 3.6b, the CAT antisense morpholino had no discernible effect at either concentration. As the previous experiment demonstrated that the morpholino is functional, we can only conclude that the vast majority of CAT mRNA produced from the plasmid never comes into contact with the morpholino. Possibly the morpholino is unable to enter the erythrocyte during electroporation, or it enters the erythrocyte along with plasmid DNA

but whatever mechanism takes up plasmid DNA into the parasite discriminates against smaller molecules. Alternatively, the morpholino may associate non-specifically with something in the erythrocyte or parasite, removing it from solution. It is very unlikely that enough CAT mRNA is produced to swamp the morpholino, as the total measured CAT activity is much lower than in the cell-free system, and similarly the plasmid DNA should not do so, as even if it were not double-stranded, the morpholino is present in vast molar excess.

Surprisingly, at 25 μM the control morpholino gives a significant increase in measured CAT activity, despite being of an irrelevant sequence! This is clearly a sequence and concentration dependent effect. A BLAST search of the *P. falciparum* genome database gave no significant matches to the control morpholino sequence, suggesting that this is not due to the inactivation of some other gene in the parasite. The CAT morpholino has essentially similar physical characteristics to the control, but did not have the same effect, ruling out the possibility of a non-specific alteration of the electrical characteristics of the solution. Perhaps this particular morpholino forms an unusual structure in solution that improves electroporation efficiency slightly, for example by stabilising pores in the lipid bilayer. The high concentration required to obtain the effect perhaps points to a fairly abundant molecular target of this sort rather than a specific genome sequence.

3.6 Conclusions

A wide variety of different transfection approaches were examined for their applicability to *P. falciparum,* including liposome and polymer-mediated gene delivery, a range of electroporation conditions, loading erythrocytes with DNA prior to invasion, and antisense technology in the form of morpholino oligonucleotides. Of these only electroporation gave the reliability and reproducibility required of a transfection technique for general use. The pulse conditions for electroporation were varied and optimal conditions identified, which gave a significant increase in transfection efficiency. The fact that these were in agreement with earlier work (Fidock and Wellems, 1997) confirms the reproducibility of the method, demonstrating that it should work equally well with different workers and equipment.

In contrast, some other strategies which have been published either did not work at all (in the case of SuperFect) or not as well as reported (in the case of pre-loading erythrocytes) in this study, serving as a reminder that there remain a lot of unknown factors in the field of *P. falciparum* transfection, which can cause variability from lab to lab. This fact is further emphasised by the mysterious effect of a morpholino oligonucleotide of 'irrelevant' sequence which nevertheless acts to increase transfection efficiency. The identification and characterisation of these unknown factors would require a great deal more work, a lot of which would be little more than trial and error. However, if the efficiency is to increase by the orders of magnitude that would make transfection as routine as it as for some other organisms, the elucidation of these factors will certainly be required.

It was not possible to introduce an antisense oligonucleotide into parasitized erythrocytes by co-electroporation with plasmid DNA, but in theory there is no reason why this technique should not work once a more effective delivery method has been identified. In this case our lack of knowledge about the physical principles behind transfection, rather than biological, is probably the confounding factor.

Deepening our understanding of how transfection works would be facilitated by having a wider range of sensitive and quantifiable assays at our disposal for measuring the rate of transfection. In the case of reporter genes, these assays would also be valuable in a range of biological studies such as promoter characterisation. One such assay is the subject of the next chapter.

Fig 3.1: **CAT expression from parasites transfected using reagent based methods.** Typical TLC data from tests using SuperFect & Effectene with plasmid pHC1-CAT. The spots on the chromatogram are labelled as follows:

1: unacetylated chloramphenicol

2: 1-O-monoacetylated chloramphenicol

3: 3-O-monoacetylated chloramphenicol

4: 1,3-diacetylated chloramphenicol

+ve: Control of 10U commercial CAT (Promega).

-ve: No enzyme.

F1-3: Products from trophozoite cultures at 10% parasitaemia treated with Effectene under recommended conditions, in triplicate.

S1-3: Products from trophozoite cultures treated with SuperFect as recommended by the manufacturer, in triplicate.

Inset: Typical results from electroporation with pHC1-CAT.

0hr: Sample harvested immediately post-transfection.

48hr: Sample harvested after 48hr culture.

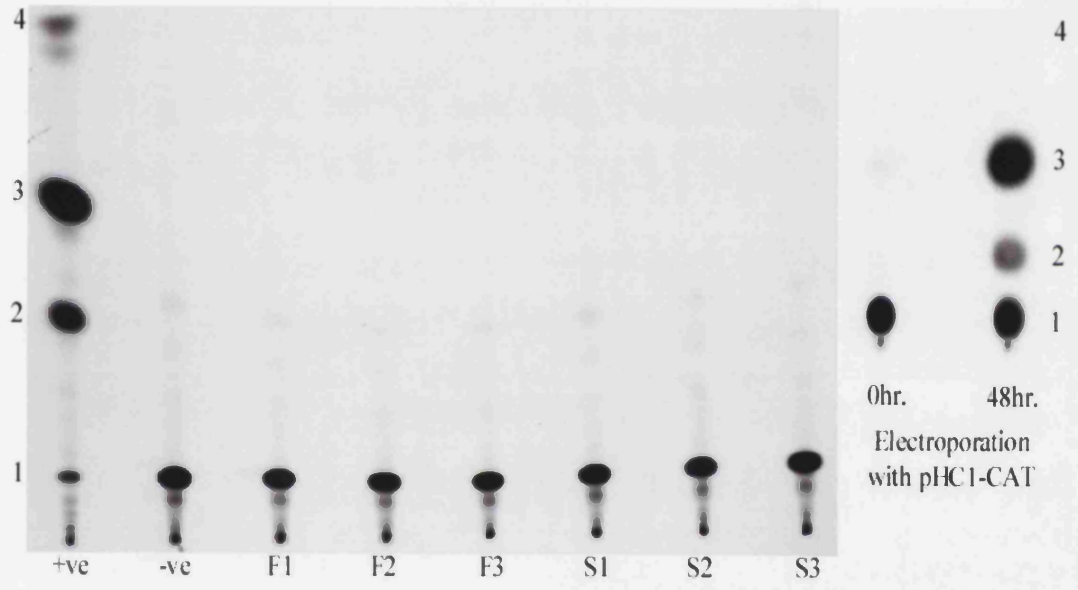

4
3
2
1
+ve
-ve
F1
F2
F3
S1
S2
S3
4
3
2
1
0hr.
48hr.
Electroporation
with pHC1-CAT

Fig 3.2: **Optimization of electroporation voltage**

Graph shows liquid scintillation count data from CAT activity assays. Aliquots of packed erythrocytes parasitized with ring stages at 10% parasitaemia were diluted to 50% haematocrit using incomplete cytomix containing pHC1-CAT, electroporated using the electrical conditions shown, and cultured for 48 hours before harvesting and CAT assay. All samples were of 200µl packed parasitized erythrocytes electroporated with 50µg plasmid in 0.2cm cuvettes at a capacitance of 0.94-0.95mF, except control (2000*). This consisted of 400µl packed parasitized erythrocytes electroporated with 100µg plasmid in 0.4cm cuvettes at 0.025mF.

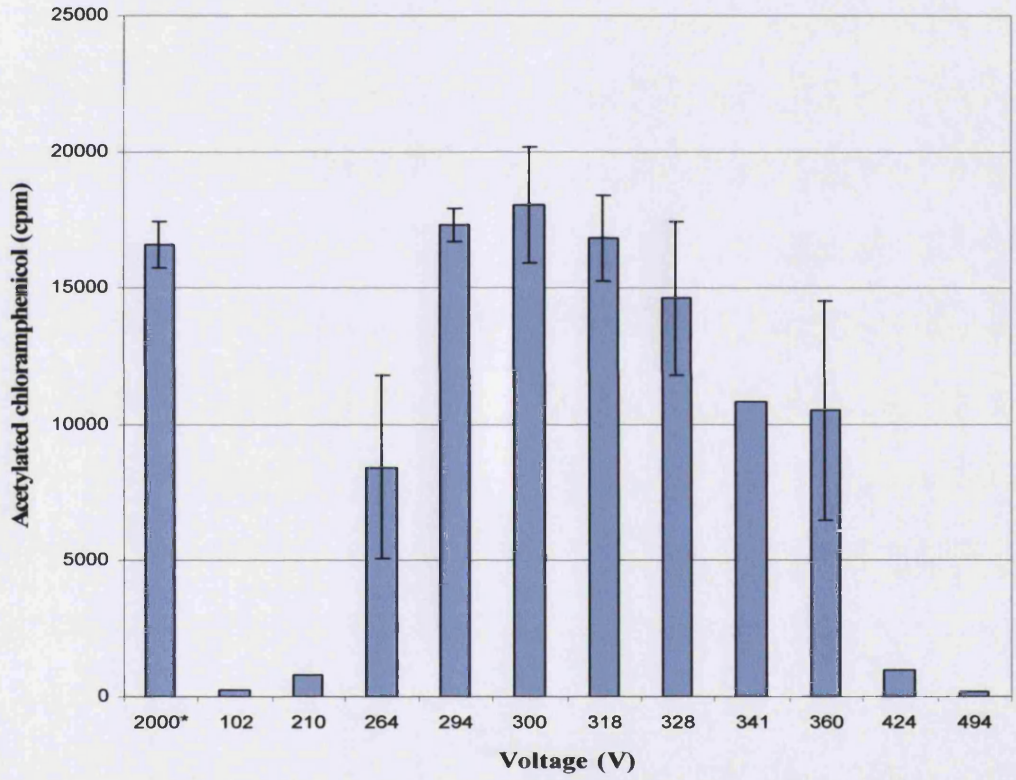

25000
20000
15000
10000
5000
0
Acetylated chloramphenicol (cpm)
2000*
102
210
264
294
300
318
328
341
360
424
494
Voltage (V)

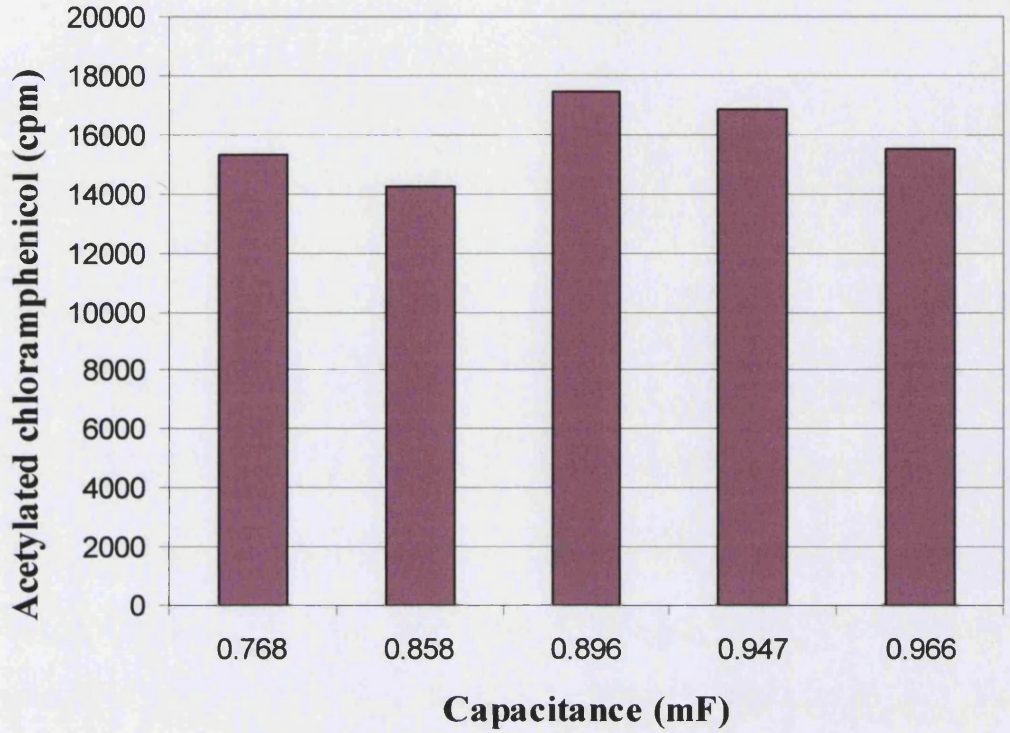
Fig 3.3: **Optimization of electroporation conditions**

Graphs show liquid scintillation count data from CAT activity assays. Aliquots of packed erythrocytes parasitized with ring stages at 10% parasitaemia were diluted to 50% haematocrit using incomplete cytomix containing pHC1-CAT, electroporated using the electrical conditions shown, and cultured for 48 hours before harvesting and CAT assay.

A: Variation of efficiency with capacitance. All samples were of 200µl packed parasitized erythrocytes electroporated with 50µg plasmid in 0.2cm cuvettes at 318V.

B: Comparison of old and new conditions. First column, 50µg plasmid added to 400µl packed parasitised erythrocytes in a 0.4cm cuvette; second column, 50µg plasmid added to 200µl packed parasitised erythrocytes in a 0.2cm cuvette. Electrical conditions were as indicated.

A.


Acetylated chloramphenicol (cpm)
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
0.768
0.858
0.896
0.947
0.966
Capacitance (mF)

B.

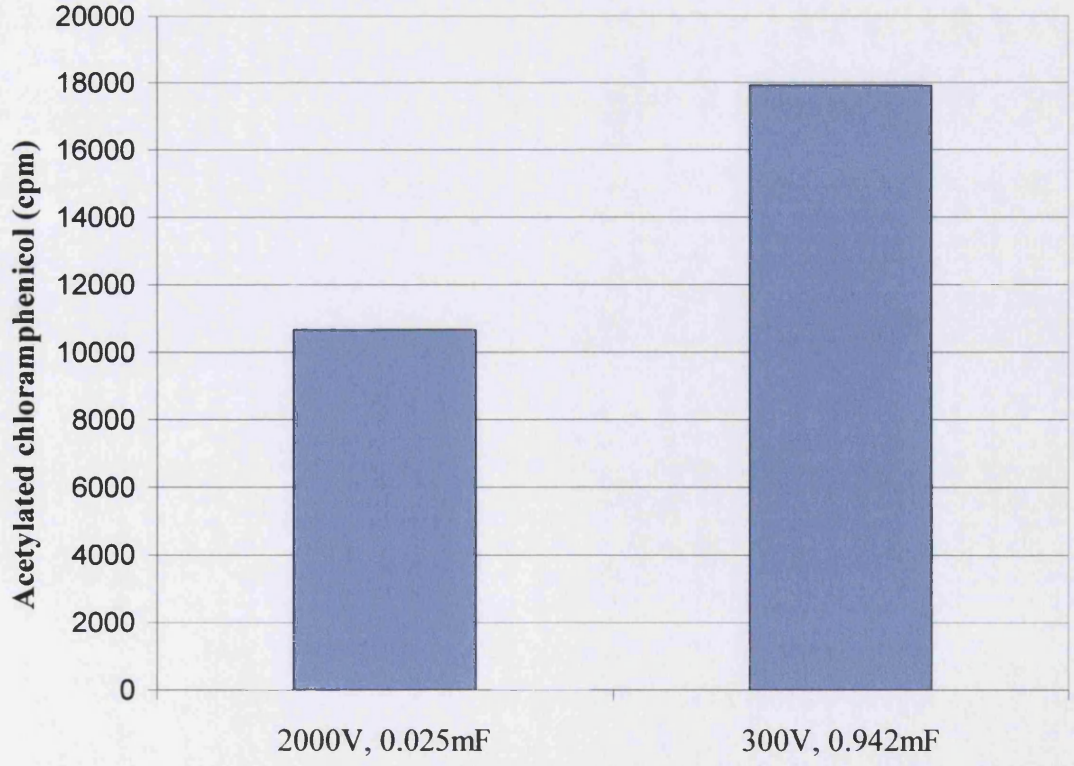

Acetylated chloramphenicol (cpm)
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
2000V, 0.025mF
300V, 0.942mF

Fig 3.4: **Transfection by pre-loading erythrocytes with plasmid.** CAT activity from transfections using 50μg pCAT.K.PcDT (see chapter 5). All samples were harvested 48 hours after the start of invasion. Blue and red columns are duplicate flasks treated with the same conditions. CAT activities were measured using quantitative protocol (see chapter 4) and error bars are the standard deviation of activity measured at three time points per sample.

-66/+4: Erythrocytes were pre-loaded with plasmid by electroporation, stored at 4°C for 66 hours, and then invaded with 5μl purified schizonts. After 4 hours incubation, the invaded erythrocytes were spun down and electroporated again with another 50μg plasmid.

-66: Erythrocytes were pre-loaded with plasmid by electroporation, stored at 4°C for 66 hours, and then invaded with 5μl schizonts.

-1: Erythrocytes were pre-loaded with plasmid by electroporation, stored at 4°C for 1 hour, and then invaded with 5μl schizonts.

+4: Erythrocytes were invaded with 5μl schizonts. After 4 hours incubation, the invaded erythrocytes were spun down and electroporated with 50μg plasmid.

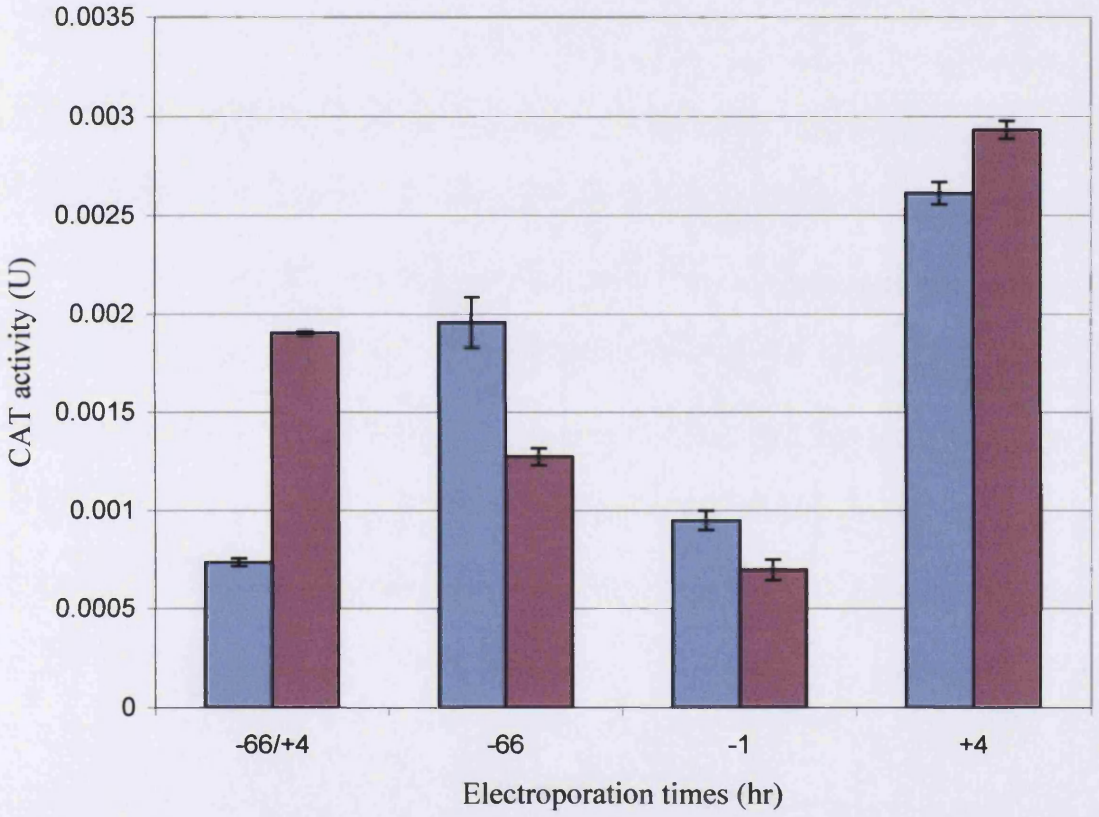
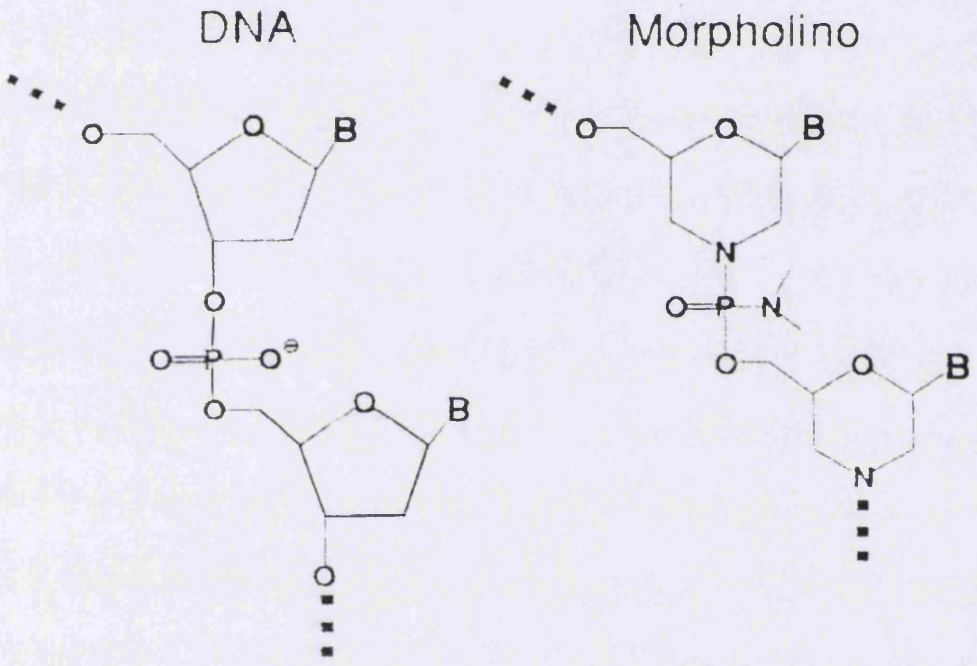

0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
CAT activity (U)
-66/+4
-66
-1
+4
Electroporation times (hr)

Fig 3.5: **Comparison of morpholino structure with DNA.** Diagram of the backbone structure of DNA and morpholino oligonucleotides, with the bases represented as 'B'. Where DNA has five-membered sugar rings linked by phosphate groups, the six-membered rings of the morpholino are connected by phosphodiamidate.



B = adenine, cytosine, guanine, thymine, uracil

Fig 3.6: **Effects of anti-CAT morpholino on expression of CAT in a cell-free system and transiently transfected *P.falciparum***

A: Acetylation of chloramphenicol with time on including 5µl *in vitro* translation products in CAT assays. CAT mRNA had been added to a rabbit reticulocyte lysate system (Promega) and translated in the following conditions:

Blue Diamonds:	no mRNA added to reaction
Purple Squares:	mRNA added, but no morpholino
Yellow Triangles:	mRNA and 440nM irrelevant morpholino added
Green Crosses:	mRNA and 440nM anti-CAT morpholino added

B: Acetylated chloramphenicol generated after overnight incubation by extracts of parasites transfected with plasmid pHC1-CAT in the presence or absence of control (irrelevant) or anti-CAT morpholino. Error bars are standard deviation of 2 samples.

-ve: Electroporation without plasmid.

+ve: Electroporation with plasmid in the absence of morpholino.

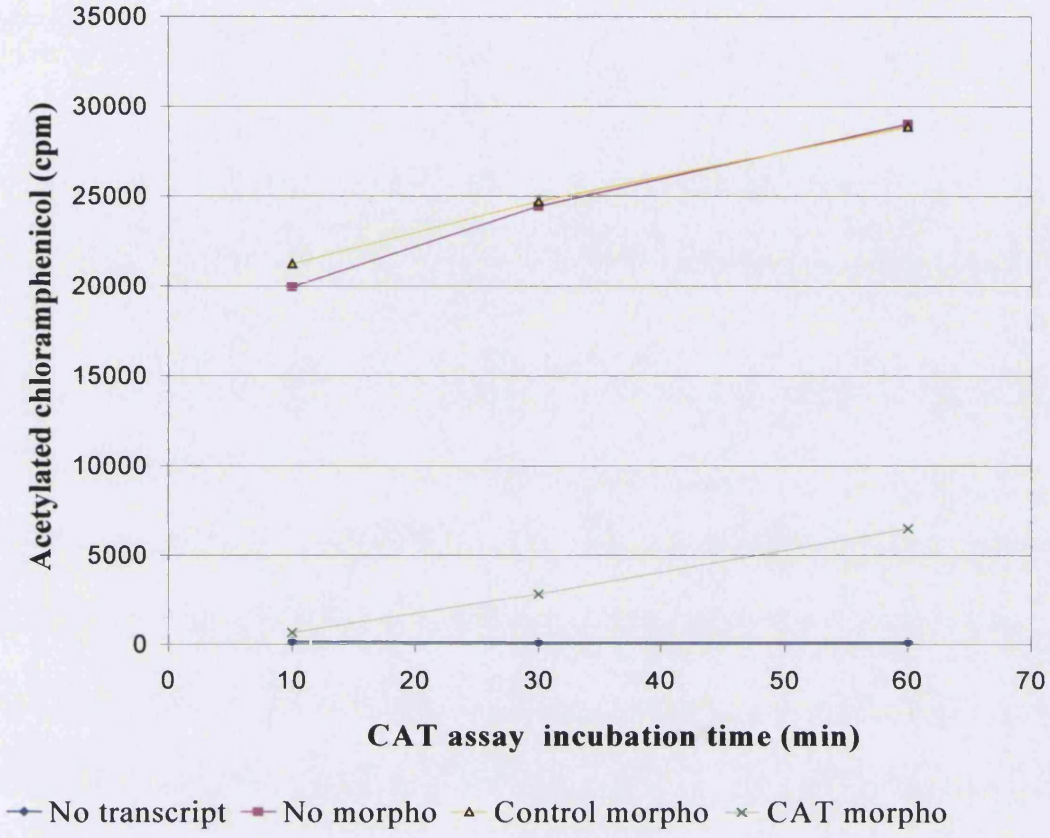
C10: Electroporation with plasmid and irrelevant morpholino at a final concentration of 10µM in the cuvette.

C25: Electroporation with plasmid and irrelevant morpholino at a final concentration of 25µM in the cuvette.

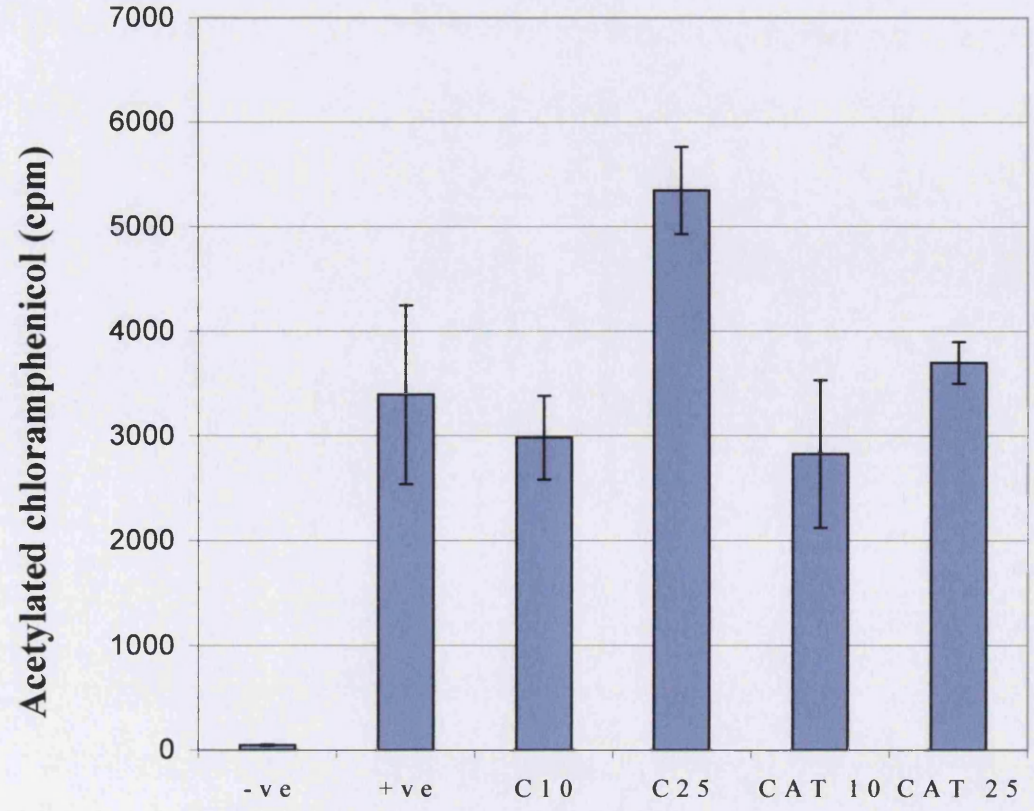
CAT10: Electroporation with plasmid and anti-CAT morpholino at a final concentration of 10µM in the cuvette.

CAT25: Electroporation with plasmid and anti-CAT morpholino at a final concentration of 25µM in the cuvette.

A.


35000
30000
25000
20000
15000
10000
5000
0
Acetylated chloramphenicol (cpm)
0
10
20
30
40
50
60
70
CAT assay incubation time (min)
No transcript
No morpho
Control morpho
CAT morpho

B.


7000
6000
5000
4000
3000
2000
1000
0
Acetylated chloramphenicol (cpm)
-ve
+ve
C10
C25
CAT 10
CAT 25

4. *Developing a quantitative CAT assay for* Plasmodium falciparum

4.1. Background

Chloramphenicol 3-O-acetyltransferase (CAT, EC 2.3.1.28) has become one of the reporter genes of choice for studying gene expression in all kinds of eukaryotic cells since it was first employed by Gorman et al. (1982). As well as being a selectable marker for the broad spectrum antibiotic chloramphenicol, its gene product is an enzyme for which very sensitive assays are available, using radioactive or chromogenic substrates. Moreover there are no analogous enzymatic activities in the vast majority of eukaryotic cells, and so no background to take into account.

CAT has also been used from the very beginning of the genetic manipulation of *P. falciparum* (Wu et al, 1995). While it has not been used effectively as a selectable marker for this organism, it has been used as a reporter gene to demonstrate successful transfection (Wu et al, 1995), compare transfection efficiencies (Crabb et al, 1997b) and assess promoter activities (Crabb and Cowman, 1996). However, the assay protocols used in these studies, and in the previous chapter, have been qualitative rather than quantitative, and relative rather than absolute. A truly quantitative enzyme assay requires that the amount of product formed varies linearly with enzyme concentration, which is usually true when the substrates are present in excess throughout the assay period. The reaction conditions described use acetyl-CoA in excess, but the more expensive radioactive substrate at a lower concentration (table 4.1) than the K_m of CAT for chloramphenicol, which has been measured at 6-7 μM (Sankaran, 1992). Therefore it is not certain that the amount of product will vary linearly with enzyme concentration, particularly as the reaction is incubated overnight at 37° C. During this time the substrate concentration may drop significantly, the product may begin to break down, and some of the enzyme may become denatured.

Reference:	Activity of chloramphenicol in reaction mix (nCi)	Specific activity (mCi/mmol)	Final concentration of chloramphenicol in 180µl reaction (µM)
Wu et al (1995)	2.5	110	0.126
Crabb and Cowman (1996)	25	58	2.39
Crabb et al (1997b)	25	58	2.39

Table 4.1: Amounts of chloramphenicol used in CAT assays of *P. falciparum* transformants

To obtain an accurate measure of promoter activity, it is important to be confident that the quantity of assay product measured is directly proportional to the amount of active enzyme produced from the promoter. In this chapter the existing CAT assay is investigated, and an improved, quantitative protocol established. The radioactive substrate used is D-*threo*-[*dichloroacetyl*-1-^{14}C]chloramphenicol (Amersham Biosciences), a form of chloramphenicol with a ^{14}C substitution three bonds away from the molecule's acetylation sites (fig 4.1); this avoids any isotope effects on the enzyme kinetics.

4.2 Liquid scintillation counting is a convenient alternative to thin layer chromatography

Previously when working with *P. falciparum*, CAT assay products have been analysed by extracting them with ethylacetate, separating the different acetylated forms from unacetylated chloramphenicol by thin layer chromatography, and then using a phosphorimager to measure the relative intensities of the spots corresponding to each chemical species. This is a fairly labour-intensive process and involves an overnight step while the phosphorimager cartridge is exposed, in addition to the assay incubation itself. In other systems, different organic solvents have been shown to separate product from substrate more efficiently than ethylacetate (e.g. Sankaran, 1992), eliminating the need for chromatography. The amount of product can then be rapidly measured by liquid scintillation counting.

A protocol using this latter technique from a CAT assay system (Promega) was adapted and compared with the established procedure. Products of a series of CAT assays were split and separated either by ethylacetate extraction followed by thin layer chromatography on silica plates with 19:1 chloroform:methanol, or by extraction with mixed xylenes and two back extractions using assay buffer. Product quantities were measured by phosphorimager analysis (fig. 4.2a) and liquid scintillation counting respectively.

Both sets of data were plotted as a percentage of a positive control of 1U commercially purified CAT (fig 4.2b). Qualitatively they compare well, with some discrepancies which could be due to experimental error in either protocol. However, the liquid scintillation data values give a slightly higher percentage of the positive control than their phosphorimager counterparts; this effect is small, but suggests that scintillation counting is slightly more sensitive, which would be valuable when measuring low CAT activities.

One disadvantage of the xylene extraction method is that it does not visually separate the substrate and different products. To ensure that no substrate was being extracted with the products, controls were carried out where no enzyme was added to the CAT assay mix (fig 4.3). In the initial xylene extraction less than 5% of the total radioactivity partitioned into the organic phase, and after two back extractions this was reduced to below 0.1%; the radioactivity at this level was indistinguishable from the background obtained using scintillation cocktail alone, showing that essentially no substrate is extracted with the acetylated products. Potentially more problematic is the inability to distinguish between mono- and di-acetylated forms of chloramphenicol. CAT is not totally specific in acetylating the 3-OH group of chloramphenicol; it can also, although less efficiently, acetylate the 1-OH group, and with less efficiency still convert either monoacetylated form to diacetylated. If a substantial proportion of the products of a CAT assay are diacetylated, the stoichiometry of the reaction is altered and enzyme activity may be underestimated. However, in all the thin layer chromatograms carried out in this study, no or minimal diacetylated products were seen when assaying extracts from *P. falciparum* (e.g. fig 4.2a); only when large quantities of pure enzyme (1 U or more) were used was diacetylation observed. Therefore, for the range of CAT activities being measured, this disadvantage of the xylene extraction method is not an issue.

Taking these factors into consideration, the xylene extraction/scintillation counting method was found to be just as reliable as thin layer chromatography for the purposes of this study, and to be substantially quicker and simpler to use. It may also give better sensitivity at the lower end of the range of CAT activities under test. Therefore this method was adopted for all subsequent CAT assays.

4.3 Long incubation times lead to assay saturation

In order to assess whether the reported CAT assay protocols give a quantitative estimate of enzyme activity, a standard curve was plotted. Purified CAT enzyme (Promega) was diluted to 1U/μl in assay buffer, and two serial 100-fold dilutions were made from this stock. From these, reactions were set up using ten different quantities of enzyme covering 4 orders of magnitude: 2, 1, 0.4, 0.1, 0.04, 0.01, 0.004, 0.001, 0.0004 and 0.0001 U enzyme in total in the 125μl reaction mix. Each assay was set up in triplicate and incubated overnight at 37°C. The following day, half of the products from each reaction (62μl) had their acetylated products measured by scintillation counting as described in section 4.2.

As shown in fig 4.4, the resulting curve is far from linear; where the total enzyme activity was 0.001U or less, no activity could be detected above background, while at higher activities the curve flattens out, suggesting that the assay is saturating. The only part of the curve which appears to be linear is from 0.01 to 0.001U (fig 4.4b), which is an extremely restrictive linear range.

The assay could be saturating because one or both of the substrates is being used up. As noted in the introduction to this chapter, while the acetyl-CoA is present in excess, chloramphenicol is not. By comparison with the total cpm of radioactive substrate at the beginning of the reactions, we can see that at least 50% of the chloramphenicol is indeed acetylated at higher enzyme activities during the overnight incubation, making this protocol unsuitable for quantitative comparisons of CAT activity.

4.4 *P. falciparum* cell extract enhances activity and stability of CAT assay

One of the other eukaryotes in which CAT has been extensively used and studied is the yeast *Saccharomyces cerevisiae.* It has been reported that the cellular material from this organism which is extracted along with CAT enzyme increases the activity measured by CAT assay in a concentration dependent manner, both for endogenously produced and for purified bacterial enzyme (Alipour et al, 1999). This was shown to be a species specific effect as neither *Schizosaccharomyces pombe* nor mammalian COS cell extracts had any such effect.

The possibility that a similar effect might occur with *P. falciparum* cell extracts was first indicated by experiments in which it was observed that assay controls using purified CAT gave highest activity measurements after 1-2 hours incubation, and that activity dropped if they were incubated overnight. In contrast, activity measurements of CAT extracted from transfected parasites continued to increase with incubation time, including overnight. To investigate this finding, *P. falciparum* cell extract was prepared from untransfected parasitized erythrocytes by the same method used to extract CAT, and verified to have no detectable activity in CAT assay. Reactions were set up including 0.1U commercially purified CAT and 0, 50 or 100µl untransfected cell extract. These were incubated at 37°C for 1, 2 or 16 hours. As shown in fig 4.5, all the reactions appeared to have reached saturation within 1 hour, and showed no more acetylation over the longer incubations. However, the reactions containing no cell extract showed a drop in acetylated product after overnight incubation, indicating that the products of the CAT reaction are not stable for long periods at 37°C.

To ascertain whether reaction rate was also affected, time courses for 0.1U, 0.03U and 0.01U of purified CAT enzyme were obtained in the presence or absence of cell extract. In all cases the cell extract was seen to increase the initial rate of the reaction, and the maximum amount of product generated (fig 4.6). Of particular note is fig 4.6b, where 0.01U enzyme was used; in the presence of 100µl cell extract in a 125 µl reaction, the measured activity increased linearly across the whole time course, whereas without cell extract the activity flattened out after about 1 hour and then began to drop on continued incubation. This surprising result was reproducible, and resembled the effect of overnight incubation on a higher enzyme concentration noted in fig 4.5.

P. falciparum cell extract, therefore, was able to stabilize the assay. When applied at a reduced concentration (10µl per 125 µl reaction) it had a partial effect, with activity continuing to climb throughout the 2 hour time course, but not increasing linearly. In order to identify the concentration which gives maximum activity, several different concentrations of cell extract were included in assays of 0.01U of commercial CAT which were incubated for 65 minutes. As a control for any non-specific effects of increased protein concentration, parallel assays were set up using the same concentration of bovine serum albumin (BSA), the total protein concentration of the extracts having been estimated using the BCA assay. As shown in fig 4.7, the effect of parasite cell extract reached its maximum by a concentration of 84µg/ml, with a further 3-fold increase in concentration conferring no additional gain in activity. In contrast BSA had a definite but limited effect, remaining well below the maximum activity even at the highest concentration used.

Clearly this effect had to be taken into account in all subsequent CAT assays. In the following sections, 30µl of transfected cell extract were included in each 125µl reaction, which corresponds to the total protein concentration of 84µg/ml noted above. In control reactions using purified enzyme, 30µl of extract from untransfected parasitized cells was added.

4.5 Generation of a standard curve for CAT assay using initial rates

It was observed in section 4.3 that long assay incubations may lead to product saturation as the chloramphenicol substrate is used up, and in section 4.4 that some loss of product occurs. The latter problem seems to be ablated in the presence of *P. falciparum* extract, however. One way to solve the former problem would be to use the substrates in vast excess; this is feasible for acetyl-CoA, but would be prohibitively expensive for the ^{14}C-labelled chloramphenicol, as well as requiring undesirably high levels of radioactivity. Alternatively, a similar quantity of radiolabelled substrate could be supplemented to excess with unlabelled chloramphenicol. This would solve the saturation problem, but with some loss of sensitivity; if hot and cold substrate are used in the ratio 1:9, then ten times as much enzyme is required to generate the same amount of labelled product.

Another approach is to measure the initial rate of the reaction, rather than the amount of product after a certain incubation period. This is widely used in enzyme kinetics because

it avoids the problems of enzyme instability and product saturation; in practice, the initial rate is usually taken as the reaction rate before 10% of the substrate has been used up.

This approach was applied to my CAT assay as follows: a wide range of CAT concentrations were freshly made in triplicate by separate serial dilutions of a commercial enzyme preparation. Each dilution was included in an assay mix as described supplemented with *P. falciparum* cell extract. Time courses at each concentration were carried out, with data shown in fig. 4.8. The initial rate of reaction for each of these curves was calculated by linear regression analysis of all the points in the linear section of the curve. These rates, in cpm/min, were plotted against CAT activity in Units (fig 4.9).

Owing to the broad range in data points, the values are plotted on a logarithmic scale. The relationship between initial rate and enzyme concentration would be predicted to be a straight line of the form:

$$V_i = M\,[E]_T$$

Where $[E]_T$ is the total enzyme concentration, V_i is the measured initial rate of reaction, and M is the gradient of the graph in fig 4.9. To calculate M from a set of data points, linear regression is frequently used, as in estimating the initial rates above. However, with a data set such as this one, covering several orders of magnitude, this is inappropriate as the data points at high values dominate the calculation. An alternative approach is to calculate $M = V_i/[E]_T$ separately for the values at each data point, and take the mean of these values. This method was applied and gave a value of M=12193 +/- 810 $cpm.min^{-1}.U^{-1}$. The best fit line computed from this value is plotted on fig. 4.9, with the faint lines either side being the standard error of the mean. This value was used in all following experiments to convert measured initial rates into equivalent enzyme activities.

4.6 Use of quantitative protocol to evaluate effect of parasitaemia on transfection efficiency

The advantages of the protocol developed above were investigated by examining the effect of parasitaemia on transient transfection efficiency. It would be expected that the level of CAT activity produced would be greater when a higher concentration of parasites is transfected, but the precise relationship between the two is unknown and might provide clues to the mechanism of transfection.

A parasite culture was synchronised and grown to ring stage at a 15% parasitaemia. This was then diluted with fresh unparasitized erythrocytes to obtain a range of parasitaemias. The final percentage of rings in each was established by counting thin blood smears; for each dilution, at least 1000 cells were counted from at least two different areas of the smear. Electroporations were then set up in duplicate using 200µl packed cells at each parasite dilution, the transfectants cultured for 44 hours, and then harvested for CAT assay.

From each transfected parasite sample, 3 assay reactions were set up using 30µl each of the cell extract, and incubated for 0, 1 and 2 hours before xylene extraction and liquid scintillation counting. Linear regression analysis of the data was used to calculate the initial reaction rate for each sample in cpm/min, which was then compared with the standard curve in fig 4.9 to estimate the CAT activity. These values are plotted in fig 4.10a.

As expected CAT activity increases with parasitaemia across most of the range, but above 10% parasitaemia it plateaus. This can be attributed to the fact that at these high parasitaemias, the health of the parasites in the culture conditions deteriorates due to overgrowth, negating any gain in number of parasites receiving plasmid.
Intriguingly, below 10% parasitaemia the relationship between CAT activity and parasitaemia is not linear, but exponential. As parasitaemia increases, the probability of a plasmid entering a parasitized erythrocyte increases by more than would be expected just from the increased number of parasites, showing that other as yet undefined factors must be involved.

The importance of the new protocol for accurate quantitation of CAT activity is illustrated by comparing this result with a similar dataset obtained using the old method with overnight incubation (fig 4.10b). From these data alone one might conclude that the increase in activity with parasitaemia levels off in a simple saturation curve. However, the shape of the curve here is very similar to fig 4.4, suggesting that what we are actually observing is the saturation of the assay masking the relationship between the variables under test. Also, the old data cannot be directly related to a known enzyme activity, but remain purely relative values. Using a non-quantitative method has practical implications; from fig 4.10b one might decide that the improvement in transfection efficiency between 5% and 10% parasitaemia is small, and therefore use the lower, and easier to obtain, parasitaemia. In fact, as shown in fig 4.10a, the increase in efficiency is about 3-fold. Moreover, only the quantitative method reveals the potentially interesting relationship between parasitaemia and transfection efficiency.

4.7 Conclusions

A truly quantitative assay protocol has been developed for the reporter enzyme CAT transfected into *P. falciparum.* This will be of considerable value as a positive control for transfection, in comparing transfection efficiencies, and in characterising promoters. Published CAT assays used for *P. falciparum* have not taken into account the exhaustion of the substrate over prolonged incubation periods, leading to saturation of the assay and potentially misleading results. The new protocol is somewhat more labour intensive than the old one, as it requires at least three time points to be measured for each sample under test, rather than just one. However, this is balanced by the shorter incubation time and simpler separation and quantification of products, meaning that the entire assay can be completed in one day rather than three.

During the development of this assay it was found that *P. falciparum* cell extract both enhances the CAT assay and stabilises its products. Considering the molecular basis for these findings raises some interesting questions. Does the addition of cell extract somehow enhance enzyme activity, or is it rather preventing a loss of activity due to instability?

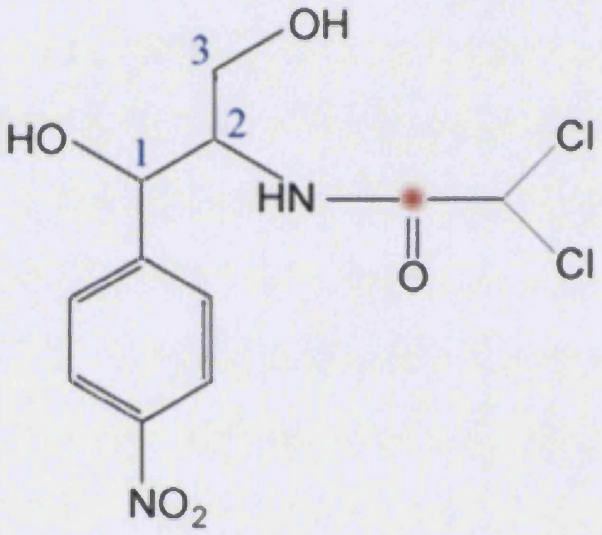
In order for the measured activity to drop in the absence of cell extract, the acetylated chloramphenicol produced in the CAT assay must be being broken down, either by chemical hydrolysis of the ester bond between the chloramphenicol and acetyl groups, or by some other chemical breakdown the products of which partition into the aqueous phase. Also, CAT activity must be dropping over time; otherwise we would expect the time course to plateau but not then fall off, reaching an equilibrium where fresh acetylated chloramphenicol is being produced at the same rate as it is being broken down, at least until the substrate in exhausted. This is significant as it shows that neither the products nor the activity of the CAT assay are stable for long periods at 37°C.

CAT activity could drop for one of three reasons: denaturation of the enzyme, reduction of substrate concentration, and product inhibition. The last case has never been reported to occur with CAT and would again cause the time course to plateau but not fall off, due to the breakdown of the product relieving the inhibition. As noted above, the chloramphenicol substrate is not present in excess, so as it is used up we might expect the reaction rate to drop. However, in fig 4.6b only about 1% of the substrate seems to have been converted to product (although if some of the products have been broken down this is an underestimate) which should not be enough to reduce the reaction rate so dramatically. The most probable explanation is inactivation of the enzyme; although it is generally thought to be stable up to 60°C, perhaps this does not hold for long incubations at low enzyme concentrations.

In any case cell extract not only stabilises the reaction but also seems to increase enzyme activity in a specific way, as an equivalent concentration of BSA had a smaller, though still significant, effect. This suggests that at least two things are occurring: a non-specific interaction which BSA can reproduce, perhaps by the higher protein concentration stabilising the enzyme or the products, and a specific enhancement of activity induced by some molecule present in *P. falciparum* cell extract, similar to that described for *S. cerevisiae.*

In any case this assay provides a valuable tool for use in transfection experiments. In addition to tests of transfection efficiency similar to those in chapter 3, another possible use is in the measurement of promoter activity, as detailed in the following chapter.

Fig 4.1: **Chemical structure of chloramphenicol.** The carbon atoms in the primary carbon chain are numbered in blue according to standard conventions. Either of the two hydroxyl groups, on the 1 and 3 carbons, may be sites for acetylation. The radiolabelled compound D-*threo*-[*dichloroacetyl*-1-^{14}C]chloramphenicol , used in this study, is identical in structure to chloramphenicol apart from the substitution of one carbon atom with ^{14}C, at the position indicated by the red spot.



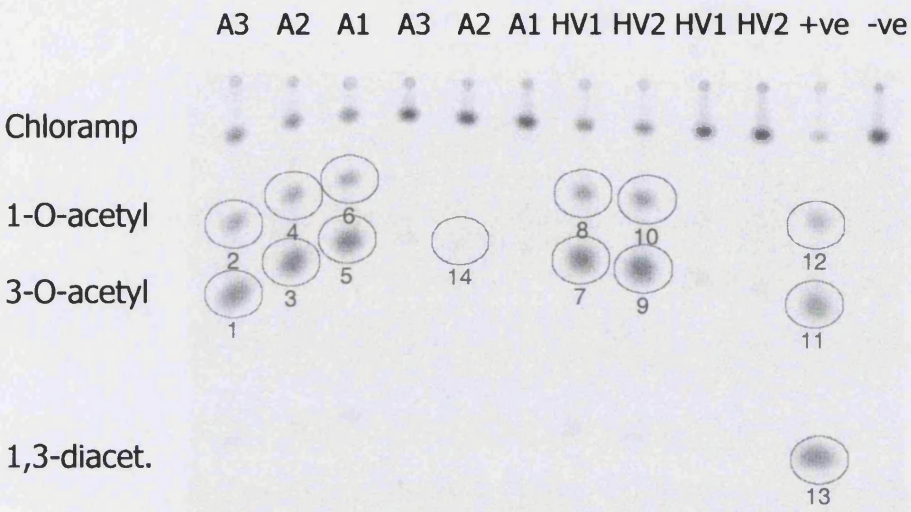
Chloramphenicol

Fig 4.2: **Comparison of data obtained by liquid scintillation counting and by phosphorimager.** Extracts from a set of transiently transfected parasites generated while optimizing electroporation conditions (see chapter 3) were used in CAT assays. After assay incubation, half of each sample was extracted with ethylacetate and separated by thin layer chromatography. The other half was extracted with mixed xylenes and used in liquid scintillation counting.

A. CAT assay products separated by thin layer chromatography were quantified using a phosphorimager. Chromatographs were inserted into a phosphorimager cartridge, which was exposed overnight before viewing. Half of the resulting scan is shown. Densitometric analysis using ImageQuant software was carried out as follows: an ellipsoid boundary was defined and placed over each acetylated CAT spot (ellipsoids 1-12) and, to measure the background, an equivalent region from a no enzyme control (ellipsoid 14). The total density in each spot was measured, the background deducted, and the values of the 1-acetyl and 3-acetyl spots for each sample (ellipsoids 1-12) added to obtain a measure of the amount of acetylated chloramphenicol. In the case of the 1U pure enzyme positive control alone, a large proportion of the substrate was diacetylated, so the density of the diacetylated spot was also measured (ellipsoid 13), doubled, and added to the monoacetylated ones.
Chloramp – unacetylated chloramphenicol
1-O-acetyl – 1-O-acetylated chloramphenicol
3-O-acetyl – 3-O-acetylated chloramphenicol
1,3-diacet. – 1,3-O,O-diacetylated chloramphenicol.

B. The values obtained by phosphorimager analysis described above were plotted as a percentage of the positive control alongside data from the same samples obtained by liquid scintillation counting.
HV1 & 2 – Parasites electroporated at 2000V, 25μF in duplicate
A1,2 & 3 – Parasites electroporated at 280V, 960μF in triplicate
B1,2 & 3 – Parasites electroporated at 250V, 960μF in triplicate
C1,2 & 3 – Parasites electroporated at 300V, 960μF in triplicate
+ve – 1U of commercially purified CAT enzyme
-ve – control reaction containing no enzyme

A.


A3 A2 A1 A3 A2 A1 HV1 HV2 HV1 HV2 +ve -ve
Chloramp
1-O-acetyl
3-O-acetyl
1,3-diacet.
1
2
3
4
5
6
14
7
8
9
10
11
12
13

B.

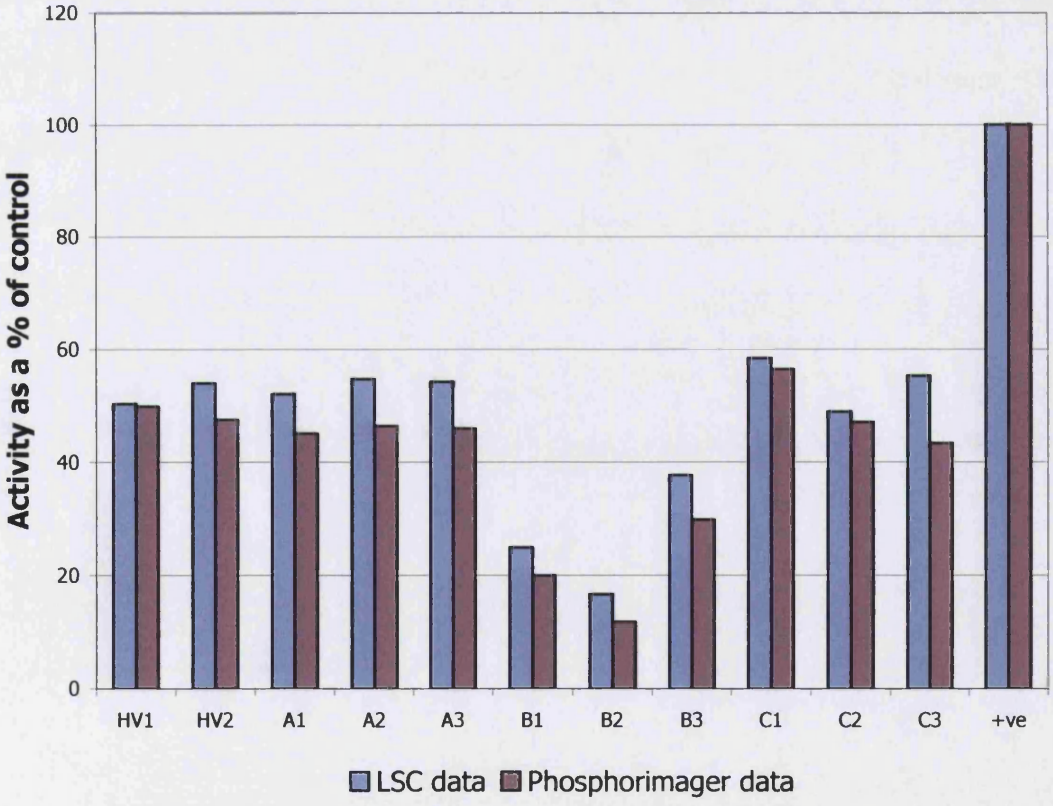

Activity as a % of control
120
100
80
60
40
20
0
HV1
HV2
A1
A2
A3
B1
B2
B3
C1
C2
C3
+ve
LSC data
Phosphorimager data

Fig 4.3: **Efficiency of extraction of unacetylated chloramphenicol by mixed xylenes.** In all samples 50nCi of D-*threo*-[*dichloroacetyl*-1-^{14}C]chloramphenicol were diluted to a final volume of 125µl in 0.25 M Tris-HCl buffer, pH 7.6.

Total counts: Entire sample was transferred directly to a scintillation vial.

Extracted: Sample was extracted with 300µl mixed xylenes, and the resulting organic phase transferred to a scintillation vial.

Extr. + Back: After extraction with 300µl mixed xylenes, the organic phase was separated from the aqueous phase and then back extracted using 100µl 0.25M Tris-HCl buffer. The organic phase resulting from this was transferred to a scintillation vial.

Extr. + 2xBack: As above, except that the back extraction of the organic phase was repeated using another 100µl 0.25M Tris-HCl buffer, before transferring it to a scintillation vial.

2ml Ready-Safe scintillant was also added to each vial. The values in the data table show the mean scintillation count of each sample in cpm. Error bars are standard deviation of duplicates.

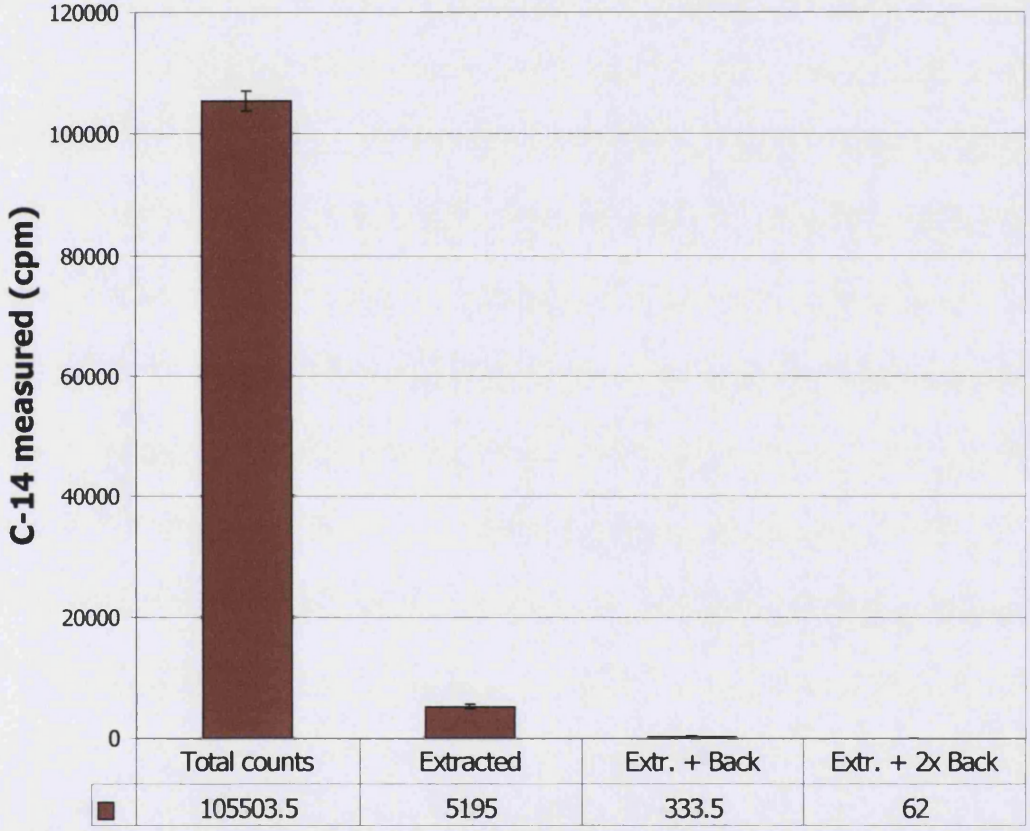
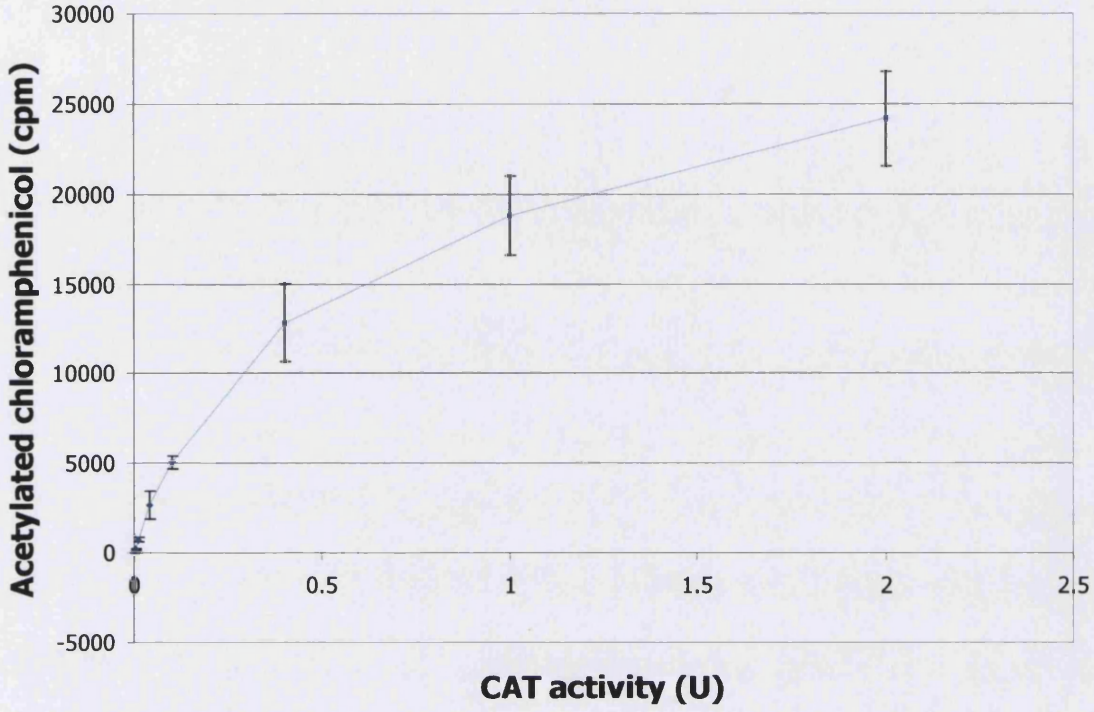

C-14 measured (cpm)
120000
100000
80000
60000
40000
20000
0
Total counts
Extracted
Extr. + Back
Extr. + 2x Back
105503.5
5195
333.5
62

Fig 4.4: **Standard curve for CAT assay after overnight incubation.** A wide range of enzyme activities were obtained by diluting commercially purified CAT and adding them to a standard CAT assay mixture. Reactions were incubated overnight at 37°C, and the acetylated chloramphenicol produced measured by xylene extraction and liquid scintillation counting.

A. Curve plotted from entire data set.
B. Close up of lower-middle range of standard curve.

Error bars are standard deviation of triplicate samples.

A.


30000
25000
20000
15000
10000
5000
0
-5000
Acetylated chloramphenicol (cpm)
0
0.5
1
1.5
2
2.5
CAT activity (U)

B.

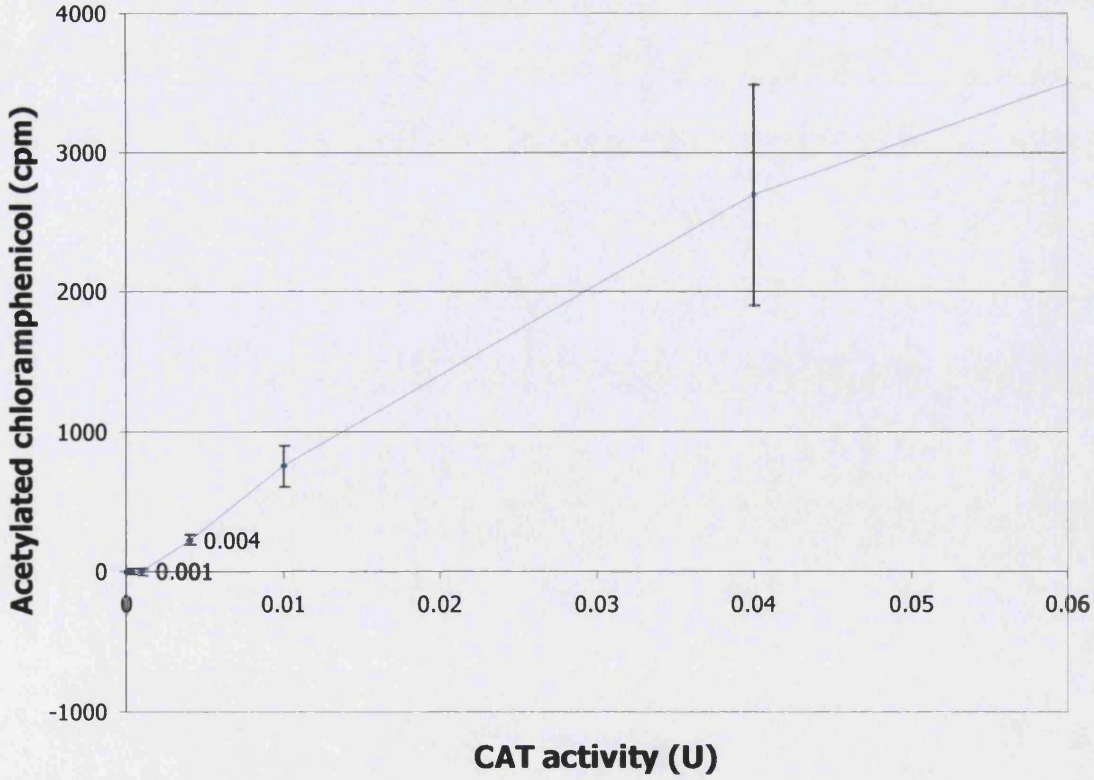

4000
3000
2000
1000
0
-1000
Acetylated chloramphenicol (cpm)
0.004
0.001
0
0.01
0.02
0.03
0.04
0.05
0.06
CAT activity (U)

Fig 4.5: **CAT enzyme in *P. falciparum* cell extract behaves differently from pure enzyme during long incubation periods.** 0.1U commercially purified CAT was added to an assay mix of a total volume of 125µl, including 0, 50 or 100µl *P. falciparum* cell extract, and incubated at 37°C for 1, 2 or 16 hours as indicated, after which the reaction was stopped by xylene extraction and the acetylated products quantified by liquid scintillation counting. Error bars are standard deviation of duplicate samples.

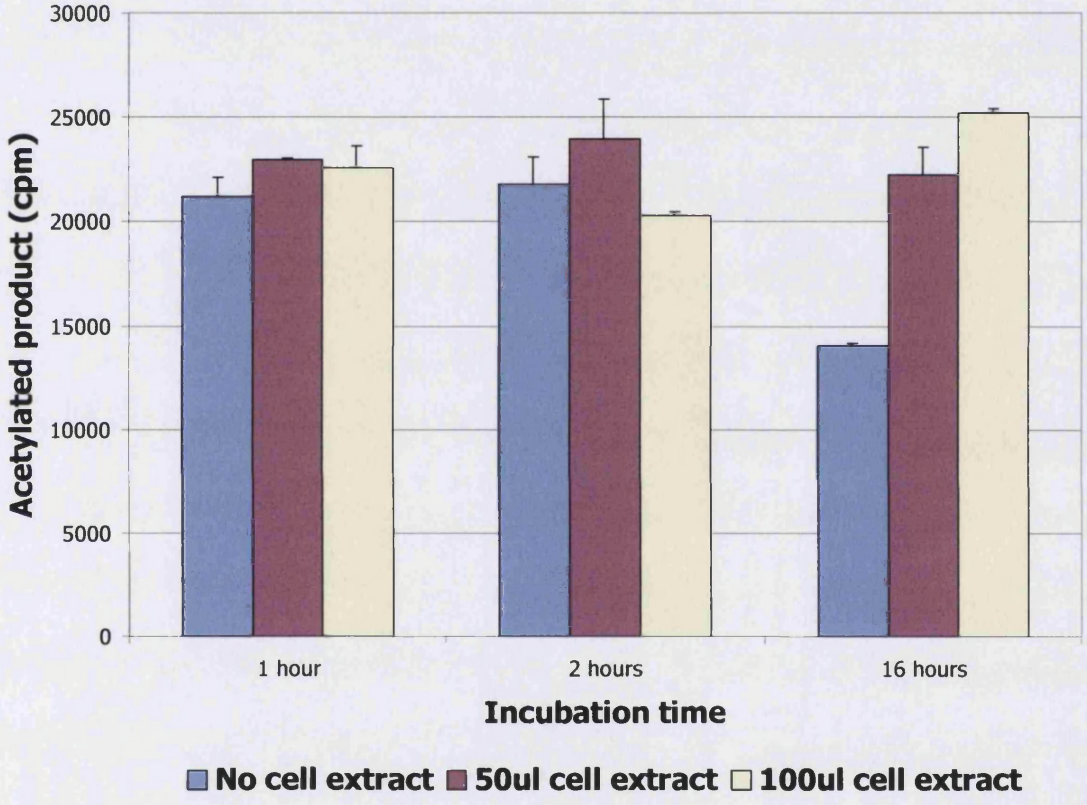

30000
25000
20000
15000
10000
5000
0
Acetylated product (cpm)
1 hour
2 hours
16 hours
Incubation time
No cell extract
50ul cell extract
100ul cell extract

Fig 4.6: **Time course of different activities of CAT enzyme in the presence or absence of *P. falciparum* cell extract.**

A. 0.1U or 0.03U commercially purified CAT enzyme were freshly diluted into assay aliquots of 125µl total volume, which were incubated at 37°C for 5, 15, 30 or 60 minutes, after which the reaction was stopped by xylene extraction and the amount of acetylated products measured using liquid scintillation counting.

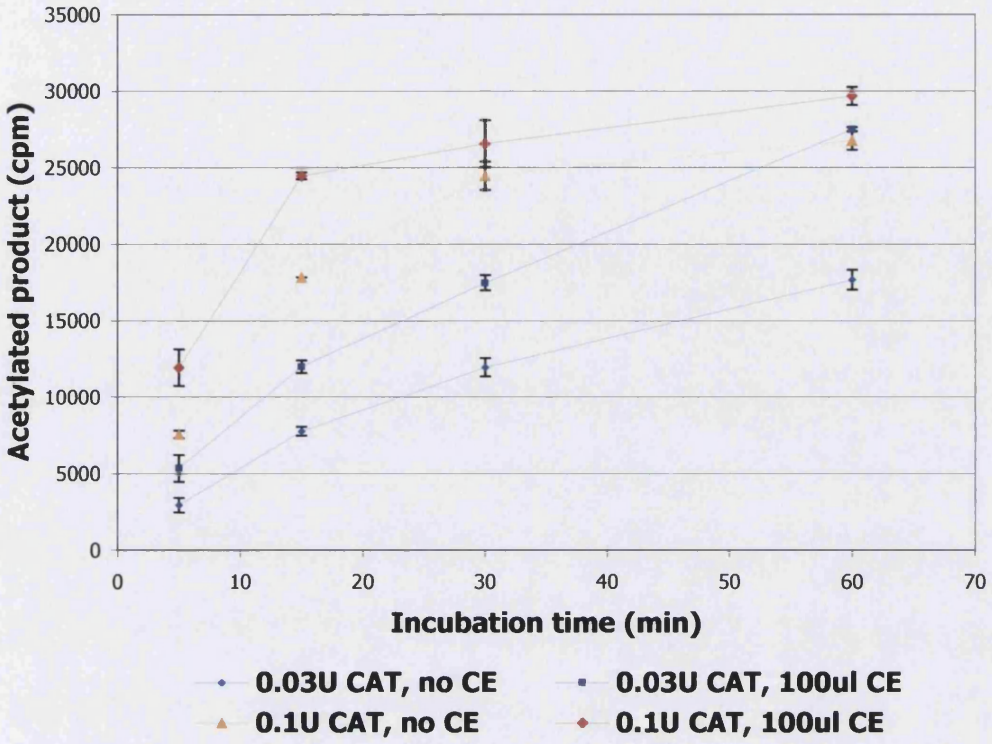
No CE: No *P. falciparum* cell extract was included in the assay mix

100µl CE: 100µl *P. falciparum* cell extract was included in the assay mix

B. 0.01U commercially purified CAT enzyme were freshly diluted into assay aliquots of 125µl total volume, including 0, 10 or 100µl *P. falciparum* cell extract. Reactions were incubated at 37°C for 5, 15, 30, 60 or 120 minutes, after which they were stopped by xylene extraction and the amount of acetylated products measured using liquid scintillation counting. A line of best fit was generated for the 100µl cell extract time course by linear regression, and is plotted on the graph.

All error bars show standard deviation of duplicate samples.

A.


35000
30000
25000
20000
15000
10000
5000
0
0
10
20
30
40
50
60
70
Acetylated product (cpm)
Incubation time (min)
0.03U CAT, no CE
0.03U CAT, 100ul CE
0.1U CAT, no CE
0.1U CAT, 100ul CE

B.

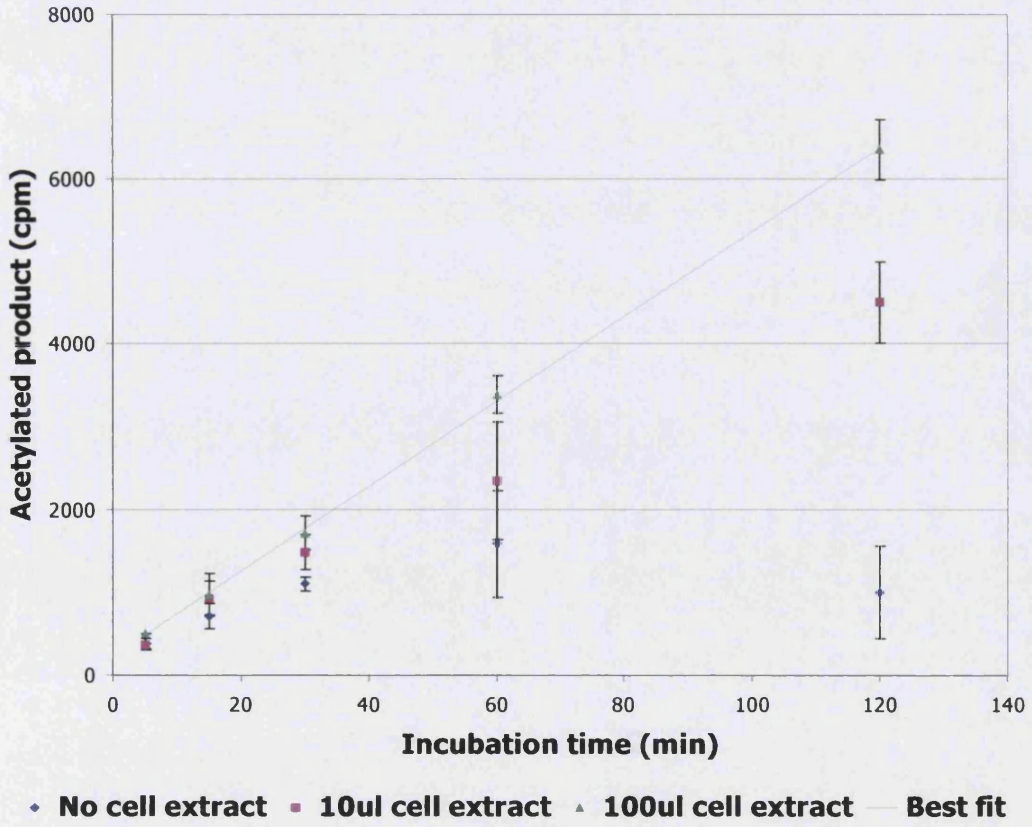

8000
6000
4000
2000
0
0
20
40
60
80
100
120
140
Acetylated product (cpm)
Incubation time (min)
No cell extract
10ul cell extract
100ul cell extract
Best fit

Fig 4.7: **Effect of *P. falciparum* cell extract is concentration dependent and specific.** *P. falciparum* cell extract was assayed for total protein concentration using the BCA method, and found to contain 350μg/ml total protein. Bovine serum albumin (Sigma) was diluted to the same concentration in 0.25M Tris-HCl, pH 7.6. 0.01U commercially purified CAT enzyme was freshly diluted into reaction aliquots of 125μl, which included 0, 10, 30, or 90μl *P. falciparum* cell extract or BSA. All reactions were incubated for 65 minutes at 37°C, and then stopped by xylene extraction followed by liquid scintillation counting of the acetylated products.

Pf extract: Samples to which *P. falciparum* cell extract was added.
BSA: Samples to which bovine serum albumin was added.

Error bars display the standard deviation of three samples.

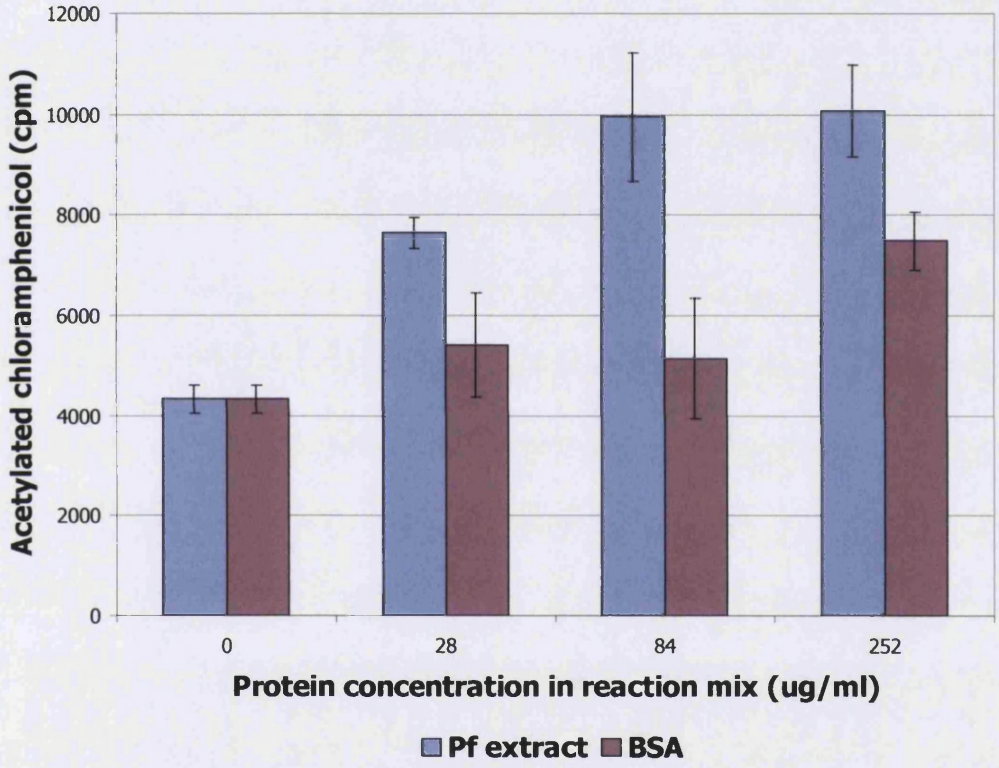
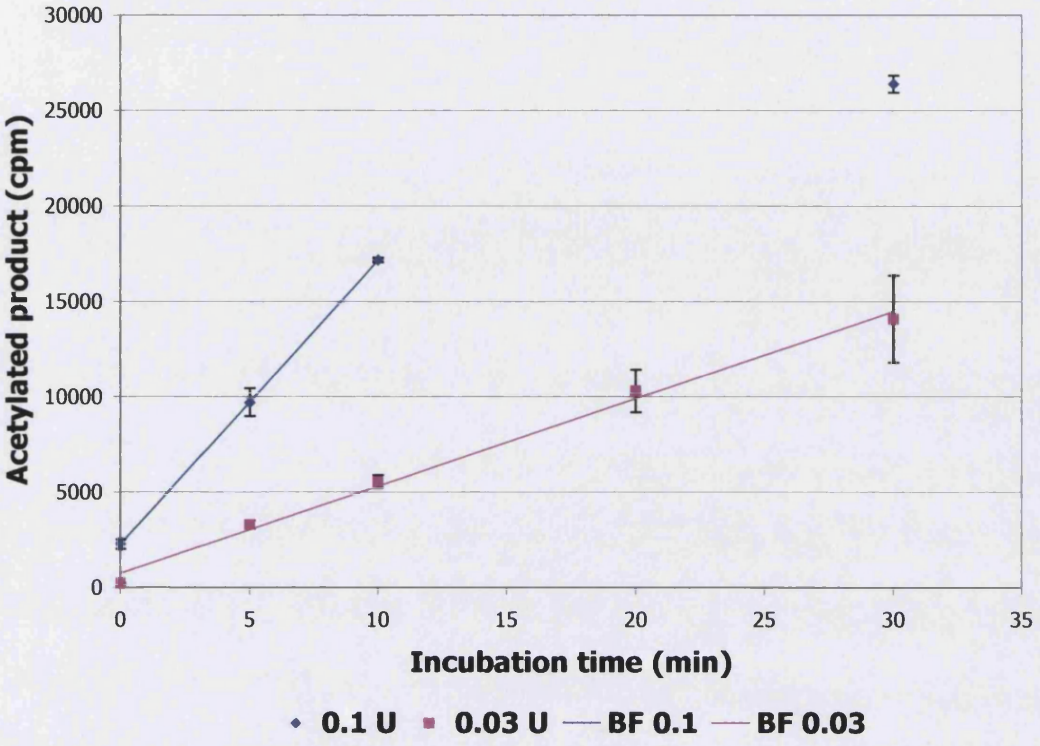

12000
10000
8000
6000
4000
2000
0
Acetylated chloramphenicol (cpm)
0
28
84
252
Protein concentration in reaction mix (ug/ml)
Pf extract
BSA

Fig 4.8: **Measurement of initial reaction rates for different quantities of CAT enzyme.** Multiple dilutions of commercially purified CAT enzyme in 0.25M Tris-HCl buffer, pH 7.6, were freshly prepared in each case. Total volume for all assays was 125µl including 30µl *P. falciparum* cell extract, and all were incubated at 37°C. Reactions were stopped at various time points by xylene extraction followed by liquid scintillation counting of the products. Lines of best fit (indicated by 'BF') were generated for the data by linear regression analysis of those points which lay in the initial linear phase of the time course.

A. Time courses for 0.1U and 0.03U CAT enzyme.
B. Time courses for 0.01U, 0.003U, and 0.001U CAT enzyme.

Error bars indicate standard error of two samples except for the 0.03U timecourse, where they show standard deviation of three samples.

A.


30000
25000
20000
15000
10000
5000
0
Acetylated product (cpm)
0
5
10
15
20
25
30
35
Incubation time (min)
0.1 U
0.03 U
BF 0.1
BF 0.03

B.

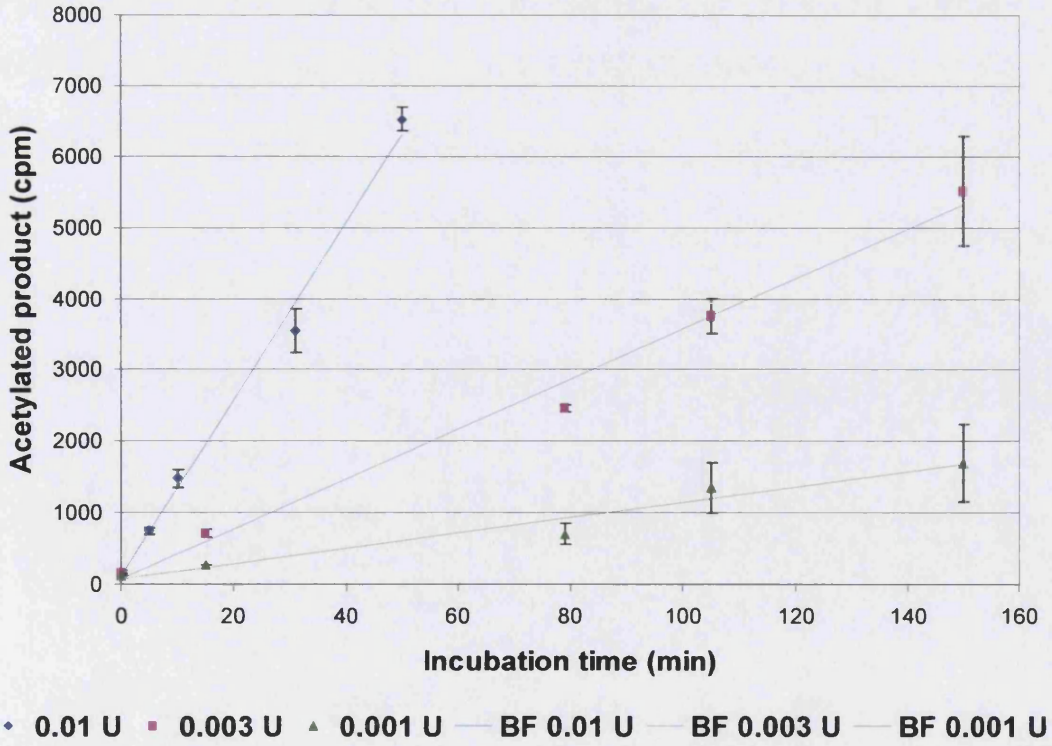

8000
7000
6000
5000
4000
3000
2000
1000
0
Acetylated product (cpm)
0
20
40
60
80
100
120
140
160
Incubation time (min)
0.01 U
0.003 U
0.001 U
BF 0.01 U
BF 0.003 U
BF 0.001 U

Fig 4.9: **Standard curve generated for quantitation of CAT assay.** The gradient of each line of best fit shown in fig 4.8 (and those for 0.0003 U and 0.0001 U CAT – data not shown) was calculated and taken to be the initial rate of reaction in cpm/min. This initial rate was plotted against enzyme activity, and is plotted on a log/log scale in order to visualise all the points. Error bars are the standard error of the gradient calculated for each line of best fit.

At each point on this graph, y/x was calculated to obtain an estimate of the gradient of the standard curve. The mean of these estimates was found and used to plot a line of best fit to the data (solid line). One standard deviation of this line of best fit is plotted above and below the line itself (dashed lines).

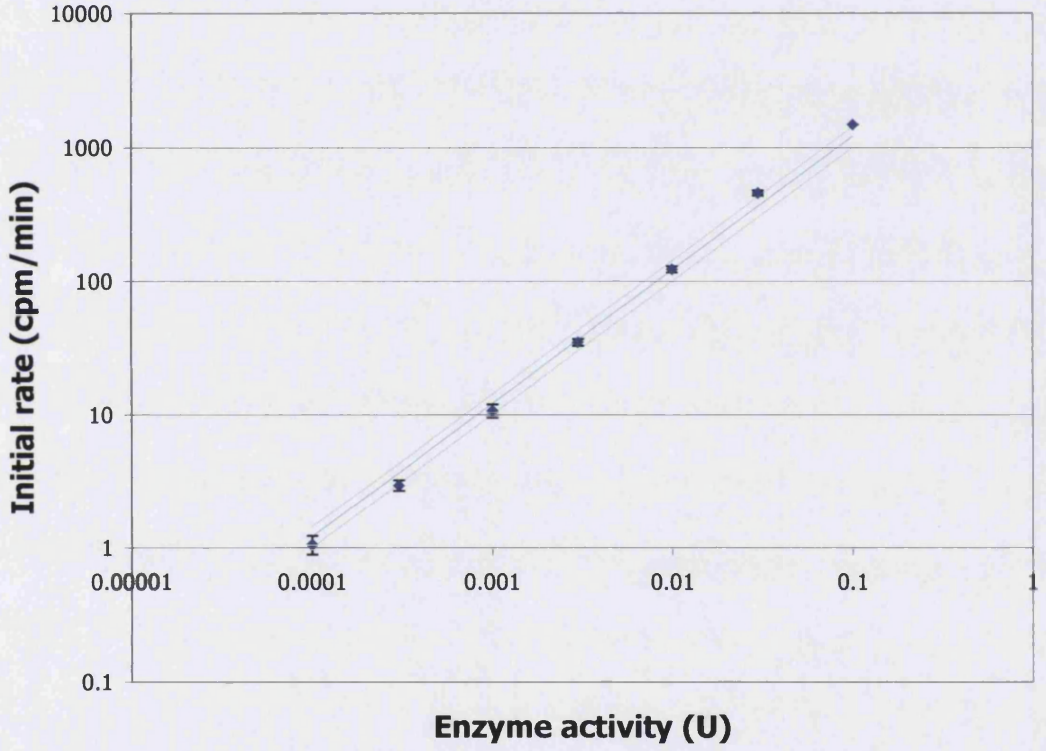

10000
1000
100
10
1
0.1
Initial rate (cpm/min)
0.00001
0.0001
0.001
0.01
0.1
1
Enzyme activity (U)

Fig 4.10: **Variation of transfection efficiency with parasitaemia.** Parasites were grown to the young ring stage at 15% parasitaemia, and diluted with unparasitized blood to the values shown; these were checked by counting thin blood smears. At each parasitaemia, duplicate aliquots of 200µl packed cells were electroporated with 50µg of a single stock of CAT expressing plasmid. After electroporation the parasites were cultured for 44 hours, harvested by saponin lysis, and the resulting pellets extracted and assayed for CAT activity.

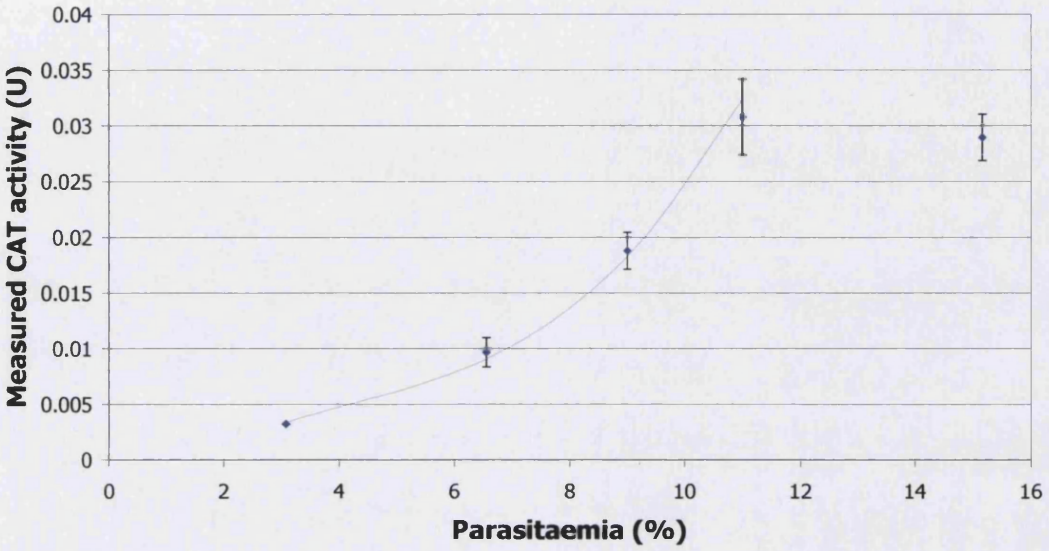
A. The plasmid used was pCATRifB^2 (see Chapter 6). From each cell extract, 30µl was added to each of three CAT assay reactions, which were incubated for 0, 1 and 2 hours before stopping with xylene extraction and liquid scintillation counting. The reaction rate was calculated by linear regression of the data, and compared with the standard curve in section 4.5 to obtain an estimate of enzyme activity in U. Errors are standard deviation of two samples. An exponential curve was fitted to the first four data points (pink line) which is described by the equation:

$$y = 0.001395*1.33^{x}$$

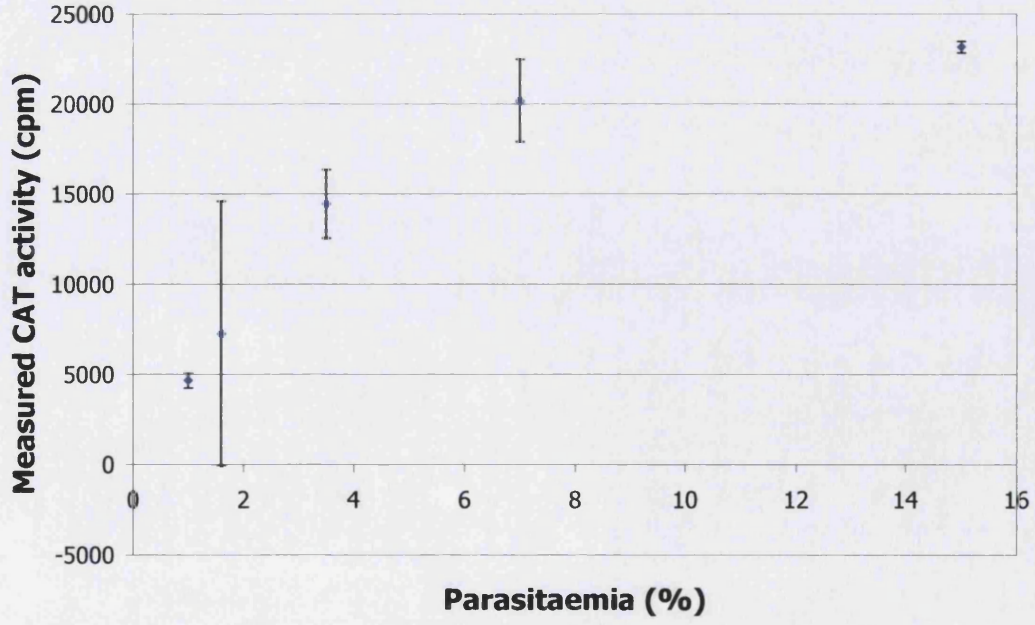
where y is the CAT activity in Units and x is the percentage parasitaemia.

B. The plasmids used were pCATRifB^2 and pCATR.B^2, which have identical CAT expression cassettes (see Chapter 6); for each pair of duplicates, one was electroporated with each plasmid. From each cell extract, 100µl was included in a CAT assay reaction which was incubated overnight before xylene extraction and liquid scintillation counting. The counts are plotted to the right, with error bars showing the standard deviation of the duplicates.

A.


Measured CAT activity (U)
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
2
4
6
8
10
12
14
16
Parasitaemia (%)

B.


Measured CAT activity (cpm)
25000
20000
15000
10000
5000
0
-5000
0
2
4
6
8
10
12
14
16
Parasitaemia (%)

5. *Studies of 5' non-coding regions of* P. falciparum *genes*

5.1 Introduction

In sections 1.4.1 and 1.4.2 we examined the relative deficiency of our understanding of the regulation of gene transcription in *P. falciparum* compared with many other organisms. Although general characteristics of the system are conserved, such as the existence of essential promoter sequences in the non-coding regions 5' of each open reading frame, it is apparently highly divergent from eukaryotes such as humans. Characterizing the system at a molecular level could therefore yield new drug targets, as in the case of reverse transcriptase; the discovery that retroviruses rely on this enzyme to replicate their genome (rather than the DNA polymerase used by human cells) and that it was unusually susceptible to inhibition by certain nucleoside analogues, was the first major breakthrough in the chemotherapy of HIV (reviewed by Mitsuya et al, 1990). Moreover, understanding transcription is technically important in any attempt to express transgenes in the parasite; only a small number of promoters have been demonstrated to be useful for *P. falciparum* transfection (Waterkeyn et al, 1999), and all are thought to be constitutive (that is, giving expression throughout the asexual erythrocytic cycle). The only detailed description of stage-specific promoters has been for those of two sexual stage genes, pfs16 and pfs25 (Dechering et al, 1999). Pfs16 was found to be specifically active in the sub-population of ring stages committed to gametocytogenesis, while pfs25 became active after transfer to the mosquito.

It would be of great value in future transfection studies to have a range of promoters specific to each stage of the asexual cycle to choose from, enabling investigators to select one most appropriate to their gene of interest. The importance of this is illustrated by the expression of *P. falciparum* AMA-1 protein in the rodent malaria *P. berghei,* which is only trafficked to the correct location when expressed at the right stage of the parasite life-cycle (Kocken et al, 1998). A great many more candidate promoters than have previously been explored are now known, thanks to the emerging genomic data. Furthermore, those promoters which have been studied were analysed independently on an *ad hoc* basis; by studying the function of a group of promoters together, it is possible that common features will emerge that might not be noticed otherwise.

To compare promoters accurately, a quantitative assay protocol is required, such as that described in the previous chapter. Here, transient transfection and the CAT reporter system were employed to analyse a group of promoters of potential technical and biological interest.

5.2 PcDT contains an enhancer

Crabb et al (1997b) observed during transient transfection studies that the activity of a shortened *P. falciparum* Calmodulin (PfCAM) promoter was increased at least eightfold when the *P. chabaudi* dihydrofolate reductase-thymidylate synthase (PcDT) promoter was adjacent to it in a head-to-head orientation. The PcDT promoter had no effect when included at another location in the plasmid. Crabb and Cowman (1996) had previously shown that a section of this latter sequence, located between positions –593 and –422 upstream of the start of the open reading frame, behaves as an enhancer in *P. falciparum;* it is required for efficient transcription from the PcDT promoter, but functions in either orientation. This suggests that the same sequence might be enhancing PfCAM promoter activity when placed sufficiently nearby.

To test the utility of my reporter system, an experiment was carried out to repeat these observations using a full length PfCAM promoter and to investigate the variation in activity across the erythrocytic cycle. Plasmid pRΔB^{2}CAT (see chapter 6) contains this same pair of promoters in head-to-head orientation with the CAT orf under the control of the PfCAM promoter. This construct was digested with *Xho* I and *Bam*H I, after which the ends were blunted and religated to form pRcpCAM, containing the PfCAM promoter alone. Additionally for comparison the two promoters were excised together from pRΔB^{2}CAT using *Sal* I and *Bam*H I, blunted, and ligated back into the plasmid in the reverse orientation to yield pRCswap, with the CAT orf under the control of the PcDT promoter. These three constructs were then used for transient transfection of *P. falciparum,* which was cultured for between 17 and 47 hours before harvesting and CAT assay.

As shown in fig 5.1a, the total CAT activity 47 hours post-transfection is fivefold higher when PcDT is included adjacent to the PfCAM promoter compared with the PfCAM promoter alone. This supports the observations of Crabb et al (1997b) and lends weight to the suggestion that the PcDT promoter contains a bi-directional enhancer. When the order of promoters was reversed, a somewhat different expression pattern is observed for the PcDT promoter in the presence of the PfCAM promoter. While the total CAT activity after 47 hours approaches a similar level to that of the PfCAM promoter, the timing of expression is altered (fig 5.1b). The CAM promoter is much more active early in the life-cycle and seems to climb to a maximum activity in the late ring stages, which is maintained with little increase until schizogony, when it drops off. In contrast the PcDT promoter appears to have only limited activity early on, but produces much more CAT during the pigmented trophozoite stage of parasite growth. Once more it loses activity once schizogony commences.

These results make sense biologically, since calmodulin is a protein involved in many regulatory pathways throughout the cell cycle, while the DHFR-TS enzyme is required for DNA synthesis, which takes place in the late trophozoite and early schizont stages. This indicates that the PfCAM promoter might be more suitable for constitutive expression, while the PcDT promoter would be useful for overexpression of transgenes during trophozoite stages.

The fact that the two promoters give different expression patterns demonstrates that the method used is sensitive enough for analyses of this nature. Also, it shows that although the enhancer element present in the PcDT sequence increases the level of transcription for both promoters, other elements must be responsible for the timing of transcription. Therefore the PcDT element might be useful as a general transcription enhancer which leaves stage specific transcription patterns intact. Of course, in any specific case it may alter the timing to some degree, and this would have to be investigated when using it.

As a caveat, the above data were obtained using the old CAT assay protocol, as the quantitative protocol described in chapter 4 had not yet been fully developed. However, they are qualitatively sound and as such support the conclusions reached above. For all following experiments, the quantitative method was employed.

5.3 Selection of promoters for analysis

In order to characterize a set of promoters of varying specificity, genes with a range of different expression patterns were required. This was facilitated by Mamoun et al (2001) who used microarrays to compare the expression patterns of 944 *P. falciparum* expressed sequence tags (ESTs) from intraerythrocytic stages. For comparison, mRNAs were purified from tightly synchronized parasites in five different time windows and normalized against 'Late Trophozoite' stage (defined as 30-36hrs post-invasion) - see fig 5.2 for full details. From these data, ESTs were identified for each stage of the life cycle that gave a >2 fold higher fluorescence yield in that stage than the others. The selected ESTs also had a high maximum fluorescence yield compared with the average for the dataset, suggesting that they are transcribed at high levels. These included part of the gene coding for MSP-1, a protein thought to be crucial in the invasion of the erythrocyte by the merozoite. Four more proteins of particular biological interest had their genes added to the selection from the literature: KAHRP, a member of the *var* gene family, PfEMP-3, and Rif2.7 (cf. Chapter 6). The *var* gene selected is known to be one of the family members which is transcribed by 3D7 parasites (P. Preiser, personal communication) and so would be expected to have an active promoter, particularly as *var* gene silencing appears to be dependent on sequences in the *var* intron (Deitsch et al. 2001b) which will not be present in the CAT reporter construct. However, much less is known about the expression of specific Rifins, so there is a risk that the Rif2.7 gene selected could have a totally inactive promoter.

A section of the expressed sequence of each EST or gene was used as a query in BLAST searches against the *P. falciparum* genome database at www.plasmodb.org. The best matches identified the location of the open reading frame of each gene. For those genes that had not previously been cloned, the open reading frame prediction programs used by PlasmoDB (hexamer, FullPhat, GeneFinder and GlimmerM) identified the putative start codon of each gene. In most cases all the programs were in agreement. The details of all the genes selected are shown in Table 5.1.

Table 5.1. Genes selected to have their promoters analyzed by their ability to drive CAT expression. Systematic gene names and chromosome locations are from PlasmoDB 4.0, with one exception. *This contig was present in an earlier version of PlasmoDB but both it and the EST used to identify it have no matches in the current version. Nevertheless, PCR products were obtained using primers designed to match its sequence.

Gene Name	Biological role	Location in genome	Expression pattern	Source
Enolase; PF10_0155	Energy metabolism (glycolysis)	Chromosome 10 637137-639010 Sense strand	High in rings and early trophozoites	Mamoun et al, 2001
Glyceraldehyde-3-phosphate dehydrogenase (GA3PDH); PF14_0598	Energy metabolism (glycolysis)	Chromosome 14 2558046-2559295 Sense strand	High in rings and early trophozoites	Mamoun et al, 2001
HSP-70 like protein	Unknown	Starts at 10165 on contig chr7_000093 Antisense strand*	High in rings, moderate in trophozoites	Mamoun et al, 2001
Knob associated histidine-rich protein (KAHRP); PFB0100c	Sub-membranous component of 'Knob' structures found in the plasma membrane of the parasitized erythrocyte	Chromosome 2 105238-105796 Antisense strand	High in rings	Vernot-Hernandez and Heidrich, 1984
Metallothionein related protein; PF10_0039	Unknown	Chromosome 10 164945-165970 Sense strand	High in early schizonts	Mamoun et al, 2001
Merozoite Surface Protein 1 (MSP-1); PFI1475w	Expressed on merozoite surface and involved in invasion	Chromosome 9 1201802-1206964 Sense strand	High in mature schizonts	Mamoun et al, 2001
var family member; PFB0010w	Codes for PfEMP-1, expressed in the plasma membrane of parasitized erythrocytes; involved in cytoadherence and antigenic variation	Chromosome 2 25232-31168 Sense strand	PfEMP-1 family members are expressed only in early rings	Kyes et al, 2000
PfEMP-3; PFB0095c	Sub-membraneous component of knobs; important for cytoadherence	Chromosome 2 91318-98838 Antisense strand	High in late rings	Waterkeyn et al, 2000
Rifin2.7	Member of multigene family expressed on parasitized erythrocyte surface and involved in antigenic variation	Chromosome 2 889660-890745 Sense strand	Rifin family members are expressed only in late rings	Kyes et al, 2000
Troponin C/ calmodulin-like; PF10_0301	In other organisms, used in cytoskeletal structure and motility	Chromosome 10 1251287-1251925 Sense strand	High in late trophozoites and early schizonts	Mamoun et al, 2001

5.4 Cloning of test constructs

It is difficult to clone large AT-rich sequences accurately, and small size is desirable in a sequence to be used in transfection vectors as it allows more of the sequence of interest to be included before the plasmid becomes unstable. Competing with this consideration is the fact that the shorter the section of DNA examined, the more likely it is that important regulatory elements will be omitted. The promoters currently used in *P. falciparum* transfection range in size from 600bp to 1.5kbp. Therefore for initial analysis of each promoter, 1.2kb of DNA sequence immediately upstream of the start codon was downloaded from PlasmoDB and PCR primers were designed that would amplify the majority of the sequence – listed in Table 5.2. The only exception is the Troponin C gene, for which the adjacent open reading frame begins only 738bp upstream of the start codon; the intergenic region alone was analyzed. In addition to the genes mentioned above, a shortened calmodulin promoter was cloned; this has been shown by Crabb and Cowman (1996) to have only 70% of the activity of the full-length promoter, but be more active than the full-length when adjacent to the PcDT element.

All the primers were designed such that the 5' end of each promoter contained an *Xho* I site and the 3' end *Bam*H I and *Sal I* sites. All of the promoters were then amplified by PCR using *Pfx* polymerase where possible, or AmpliTaq where the templates proved more difficult to amplify. The PCR products were subcloned into pCR4, PCR screened for inserts of the correct size, excised using *Xho* I and *Sal* I, and gel purified.

For initial tests it was decided that the PcDT enhancer element might prove valuable in obtaining readily measurable levels of CAT activity. Therefore the plasmid p1-4 (see section 6.4) was digested with *Xho* I and *Sal* I, the larger restriction fragment gel purified, and then ligated with the promoter restriction fragments generated above. This resulted in the promoters of interest being inserted in front of the CAT open reading frame, and adjacent to the PcDT promoter which was in the reverse orientation (fig 5.3). All the promoter fragments were successfully inserted except that of Troponin C, for which no positive clones were obtained despite repeated efforts; this sequence must be either toxic to or unusually labile in *E. coli.*

To further analyse selected promoters in the absence of the PcDT element, these constructs were digested with *Xho* I and *Sac* I, the ends blunted, and then religated, deleting the PcDT promoter. The nomenclature used for these plasmids is 'pCAT.x' or 'pCAT.x.P,' where x is the name of the promoter fragment under test, and P indicates the presence of the PcDT promoter adjacent to it.

Table 5.2. Oligonucleotides used in the cloning of promoter regions from *P. falciparum* genes. All were used as primers for PCR. Restriction sites are highlighted.

Fragment name	Fragment size (bp)	Forward primer (_5)	Reverse primer (_3)	Restriction site, 5'	Restriction site(s), 3'
Cam6	634	ATTA**CTCGAG**GT TTATATGTGATTA ATTTTATATATTA TC	CACTA**GTCGAC**TG ATATATTTCTATTA GG	*Xho* I	*Sal* I
Enol	1044	ATTA**CTCGAG**CG TAATAAATTCTCA TATATAAATACG	ATTA**GTCGACGGA TCC**GTTTATATTTT ATTTAAAATAAAAT AATATAAAAGG	*Xho* I	*Bam*H I, *Sal* I
Gpdh	1165	ATTA**CTCGAG**GA AGAAATAATTAA AAAAATATATAT CAC	ATTA**GTCGACGGA TCC**ATAATAAGAA AAGAATTAAAAAG CCG	*Xho* I	*Bam*H I, *Sal* I
H70L	1138	ATTA**CTCGAG**AA TTATAATTCAATA AAAAAAAATACA GGG	ATTA**GTCGACGGA TCC**AAAAAAGTTTT ATTTAAATGAATTA ATCTTTAC	*Xho* I	*Bam*H I, *Sal* I
KHRP	1144	ATTA**CTCGAG**CT CCGTTTGTATCTT TCTTG	ATTA**GTCGACGGA TCC**TAATAATTATG TACGAATATATAAA TTTTG	*Xho* I	*Bam*H I, *Sal* I
Met	1128	ATTA**CTCGAG**AC AATGTCGTTTTTG TGTGTAC	ATTA**GTCGACGGA TCC**ATGTTATGTTT TTTTCTTTTATTTTT TCTG	*Xho* I	*Bam*H I, *Sal* I
MSP1	1077	ATTA**CTCGAG**AT CATATATATAATA TATATAAACATGT AC	ATTA**GTCGACGGA TCC**TAAGTTAATAA TTTAAAAATATGAC TTAGC	*Xho* I	*Bam*H I, *Sal* I
PfE1	1113	ATTA**CTCGAG**TA ATGTTATTGAAAA TACGTATGCC	ATTA**GTCGACGGA TCC**GTTGCGACAA CATGATGTC	*Xho* I	*Bam*H I, *Sal* I
PfE3	1136	ATTA**CTCGAG**TA TGAATAATATTGG TACATTAGG	ATTA**GTCGACGGA TCC**GTTGTAAAATA TTTATTAACAAATT GGC	*Xho* I	*Bam*H I, *Sal* I
Rf2.7	1073	ATTA**CTCGAG**CG TAATATATATATC ATCGTGATAC	ATTA**GTCGACGGA TCC**ATTAACTAATT AATATATAAAAAT ATAATGTGC	*Xho* I	*Bam*H I, *Sal* I
TrpC	811	ATTA**CTCGAG**TT TTAAAAGATCCTT AAGATCTCAG	ATTA**GTCGACGGA TCC**GCATTTTCTAT TAATATAAATAAGC AAC	*Xho* I	*Bam*H I, *Sal* I

5.5 Erythrocytic cycle scans

An initial series of experiments aimed to ascertain which promoters were likely to have useful activities or interesting characteristics for further study. To this end, all of the pCAT.x.P constructs were transiently transfected into *P. falciparum* parasitized erythrocytes using the optimum conditions described in chapter 3. The parasites used were synchronized ring stages at 5-10% parasitaemia, and electroporation was carried out at 3-7 hours post-invasion. Separate transfectant cultures were harvested by saponin lysis at approximately 20, 28, 44 and 52 hours post-invasion, representing late rings, trophozoites, schizonts and early rings from the following life-cycle respectively. The saponin pellets were fast frozen and stored at -80°C for CAT assay at a later date. Parasitaemia and parasite stages were monitored by thin blood smear, and it was noted that even by the end of one life-cycle, synchrony had deteriorated significantly, to the extent that there would often be some rings from the next life-cycle visible in the 44hr sample, and some schizonts in the 52hr sample. It is suggested that the shock of electroporation affects some parasites more than others, slowing them down while they recover, with the effect across the population of partially dispersing the synchrony.

Levels of CAT expressed in each sample were measured using the quantitative CAT assay, with the results being presented in figs 5.4, 5.5 and 5.6. For each different parasite culture used in transfection, two aliquots were transfected with pRifCAT (see chapter 6 - contains the full-length calmodulin promoter with the PcDT element, and is here referred to as 'pCAT.Cam.P') and cultured for 44 hours before harvesting. The amount of CAT produced by this plasmid was used to normalize the results between different parasite cultures. The data is displayed in two forms; firstly, as a time course for each promoter. However, assuming the CAT enzyme generated is stable throughout the life-cycle, the promoter activity is not directly proportional to the total amount of CAT at any given time, but rather the rate at which it is being produced. Therefore the average enzyme production per hour between each pair of time points is also shown.

Only one of the plasmids showed no activity whatsoever: pCAT.PfE1.P (fig 5.4). Although reporter expression has previously been obtained using a different *var* gene promoter (Deitsch et al 1999), this example used a 2.7kb section of the upstream sequence and came from a gene shown to be the predominant family member expressed

in the parasite strain used. The 1.1kb segment I tested might lack essential promoter sequences, or be from a gene which is not expressed in 3D7. However, this serves as a useful negative control for the rest of the promoters, showing that the adjacent PcDT element alone is not sufficient to initiate transcription.

Several of the plasmids gave only very low levels of activity – producing no more than 0.001 U of CAT throughout the erythrocytic cycle. These included pCAT.MSP1.P, pCAT.PfE3.P, and pCAT.Rf2.7.P, the last of which gave barely detectable activity (fig 5.4). This may reflect the relative abundance of their biological gene products; Rifins are certainly not abundant proteins, and while PfEMP-3 and MSP-1 are both readily detected at the erythrocyte plasma membrane and in the merozoite membrane respectively, we would not expect them to be as abundant as glycolytic enzymes like enolase. The expression by pCAT.MSP1.P appeared to be only in late trophozoites and schizonts as expected, but pCAT.PfE3.P – which is expected to express early – produced most activity 28-44 hours after invasion. The levels of expression by pCAT.Rf2.7.P were too low, and the experimental error too high, to localize its stage specificity more tightly than to the second half of the life-cycle.

Some of the plasmids were much more potent, producing over 10 times as much CAT as those above (fig 5.5). Particularly powerful were pCAT.GPDH.P and pCAT.KHRP.P, both of which gave much higher expression than the pCAT.Cam.P construct used as a control. These two promoters should therefore be more effective for overexpressing proteins in *P. falciparum* than the calmodulin promoter, which is widely used in transfection constructs. Interestingly, the truncated calmodulin promoter in pCAT.C625.P seemed to be less active than the full length version in early stages, but more active later in the cycle. This could be explained if the section of the promoter upstream of the -625 position contains elements important for expression in young parasites, but removing it increases the effect of the PcDT element, perhaps by bringing it closer to the transcriptional start site.

Surprisingly, none of the promoter regions tested showed much activity in the first 20 hours post invasion, even those from genes which are known to be strongly expressed in rings. Only pCAT.Cam.P produced a level of CAT significantly higher than the background (fig 5.5b), which is in agreement with earlier results for the full length calmodulin promoter (section 5.2). However, even in this case expression was low.

Moderate expression levels were obtained from pCAT.Enol.P and pCAT.H70L.P. This last plasmid was particularly interesting in that its expression was tightly stage specific, with protein only being generated in the 28-44hr (trophozoite-early schizont) time band. This correlates with the expression pattern observed for the HSP-70 like gene by Mamoun et al (2001), and suggests that the presence of the PcDT element, while stimulating transcription, does not necessarily disrupt stage specificity.

This does not rule out the PcDT element altering timing of expression in other cases. To investigate this possibility, it was removed from several of the plasmids to give pCAT.GPDH, pCAT.KHRP, pCAT.PfE1 and pCAT.PfE3. These were used in transfections as before (fig 5.6). As expected pCAT.PfE1 gave no detectable CAT activity, but interestingly pCAT.PfE3 was also inactive. Clearly the PcDT element was important for efficient transcription from the PfEMP-3 promoter, perhaps by substituting for other genomic enhancer sequences absent from the cloned segment of this gene. However, removing the PcDT had markedly different effects on the GA3PDH and KAHRP promoters. Expression from the former in the latest time band (schizonts-rings) appeared to be greatly increased relative to that of the latter. Unfortunately technical problems with the controls for this experiment meant that the data could not be normalized relative to fig 5.5, but nevertheless it was decided that these findings warranted more detailed examination (see next section).

Overall, this series of experiments identified promoters with a wide range of levels of activity. Differences in timing of expression were more subtle, and each time course was only measured for a single culture. Therefore more data are required to support the findings described above and to characterize stage specific activity more clearly.

5.6 Activity of GA3PDH and KAHRP promoters during schizogony

Both of the strongest promoters described in section 5.5 seemed to have their activity altered by the presence of the PcDT element, an effect which was clearest in the samples measured before and after schizogony. To examine this more closely, parasites were synchronized to a less than 2hr invasion period, and then used immediately for transfection with pCAT.GPDH, pCAT.GPDH.P, pCAT.KHRP or pCAT.KHRP.P. The transfected parasites were then cultured for 2 days and duplicate aliquots harvested at 44,

48 and 52 hours post–invasion (measured from the middle of the earlier invasion period). The aim of this was to ensure that the parasites would still be synchronous enough to show any differences between these fairly close time points. Thin blood smears taken at each time point showed that despite these precautions, only about half of the schizonts visible at 44hrs post-invasion had ruptured by 52hrs, although the number of rings increased 4.5-fold over the same period.

All the samples were assayed for CAT activity (fig 5.7). The expression levels observed at the start of the experimental period (44hrs post-invasion) show the differential effect of the PcDT element on the two promoters: it has little or no effect on the activity of the GA3PDH promoter, but increases that of the KAHRP promoter 3-fold. The expression levels for GA3PDH lie between those of KAHRP with and without the PcDT element, in agreement with earlier data (fig 5.5a and 5.6a). However, in contrast with these data, none of the constructs is seen to be highly active during schizogony. The differences in CAT measured at each time point are sufficiently small that there are few significant differences across the time period (fig 5.7b). The possible exceptions to this are the apparent inactivity of both the constructs containing the PcDT element during the first time window, and the increase in activity during the second half observed for the pCAT.KHRP.P construct compared with pCAT.KHRP. If these findings are genuine, they suggest that the PcDT element acts not as an enhancer but as an inhibitor of transcription in late schizonts, while in early rings it reverts to its positive regulatory role. Sequences which regulate transcription positively or negatively depending on the developmental stage have been described in other systems, but not *P. falciparum.* However, the fact that there were still schizonts present by 52hr post-invasion makes this suggestion tenuous.

5.7 Sequences of technical interest for transfection

The advantages of stage-specific promoters for transfection experiments have already been described, but an even more flexible tool would be an inducible promoter. This would allow transcription to be turned on and off at will simply by adding an inducing agent to the culture medium. Inducible promoters are in common use in other systems and attempts have been made to adapt them for use in *P. falciparum,* notably with the Tet response element. This system has recently been successfully adapted to a related parasite, *Toxoplasma gondii* (Meissner et al, 2001); work is ongoing to try and repeat

this success in *P. falciparum,* reflecting the divergence of its transcription regulation systems from those of other eukaryotes. Given this divergence, a more productive strategy might be to adapt a response system which already exists in the parasite, rather than trying to import one from another organism.

Although no such system has been clearly defined for this species, several stimuli can be reasonably expected to specifically induce expression of certain genes. Two of these are heat shock and the increased concentration of heavy metal ions in solution; most organisms have some way of reacting to these potentially damaging conditions. The genes selected from in section 5.3 from the microarray study include one which is homologous to heat shock protein 70 and another with some sequence similarity to metallothioneins. By analogy with similar proteins found in other organisms, we might expect the promoters from these genes to give specific response elements which upregulate transcription in response to heat and metal ions respectively.

To investigate the first of these possibilities, 200µl aliquots of parasitized erythrocytes were electroporated as normal with pCAT.H70L.P. All the aliquots were cultured at 37°C until 28 hours post-invasion (trophozoite stage), at which point duplicate flasks were either left at 37°C or incubated at 39°C or 41°C for 90 minutes. After this all cultures were returned to 37°C and cultured for a further 16 hours before harvesting the cells and measuring CAT activity. Examination of blood smears taken at this point revealed that the parasitaemia had dropped by about 60% in those flasks which had been heat shocked, at either temperature. The CAT data (fig 5.8a) was normalized to take this into account; while there is possibly a slight increase in promoter activity in the heat shocked parasites, it is not large enough to be statistically significant. The high sensitivity of *P. falciparum* to heat shock displayed in this experiment suggests that this approach is unlikely to be a practical way to achieve inducible expression. Another concern would be that heat shock probably upregulates dozens of other genes as well, making it difficult to distinguish the effects of an induced transgene.

Therefore, attention was turned to the metallothionein related orf (PF10_0039). Many similar genes in other organisms feature metal response elements (MREs) in their 5' non-coding sequences, so the 5kb intergenic sequence between PF10_0039 and the preceding orf was inspected for potential MREs, both using BLAST and visually. No obvious matches to the MRE consensus were observed, although this does not rule out the

presence of divergent sequences which behave as MREs. This possibility was investigated by transfecting aliquots of parasitized erythrocytes containing c. 8% rings with 50μg pCAT.Met.P as described and culturing as normal or in medium supplemented with 1, 10 or 100μM metal ions. Different MREs respond with greater sensitivity to particular metal ions, so Cd^{3+}, Cu^{2+}, and Zn^{2+} were all added at all three concentrations, as acetate, sulphate and sulphate salts respectively. High levels (100μM) of Cu^{2+} ions induced erythrocyte lysis, but all other cultures appeared to develop normally. However, on carrying out CAT assays, none of the transfectants had any detectable enzyme activity, even though pCAT.Cam.P controls showed normal expression levels (data not shown). If PF10_0039 is a metal response gene, its key regulatory sequences lie further upstream than the 1.2kb cloned in this study.

Another sequence which is already used in some transfection constructs is the rep20 repeat, which has been demonstrated by O'Donnell et al (2002) to accelerate the generation of stable transfectants by improving the segregation of plasmids carrying it during schizogony. The mechanism which achieves this appears to be association of the plasmid with the genomic rep20 repeats (which are found sub-telomerically on all chromosomes) during sister chromatid separation. In other eukaryotes genes associated with the telomeres are frequently transcriptionally silent. Therefore if the rep20 sequence is to be used in transfection systems, it is important to ascertain whether it affects gene expression.

Therefore a rep20 sequence cloned in section 7.2 in the form of a *Hind* III restriction fragment was inserted into pCAT.KHRP.P in place of the PcDT element. This was achieved by digesting the plasmid with *Xho* I and *Bam*H I, gel purifying the larger fragment, blunting its and the rep20 fragment's sticky ends, and ligating the two together to give pCAT.KHRP.rep20. This plasmid, pCAT.KHRP and pCAT.KHRP.P were electroporated into aliquots of the same parasite culture, grown, and harvested after 45 hours (fig 5.8b). As in previous experiments pCAT.KHRP.P gave much higher expression than pCAT.KHRP, while pCAT.KHRP.rep20 gave a slight increase in activity. This shows that rep20-bearing plasmids are not impaired in gene transcription; either the rep20 sequence is not associated with the telomeres throughout the life-cycle, but only during chromosome separation, or else the association is not close enough to inhibit transcription. The slight increase in activity conferred by rep20 was unexpected; a possible explanation would be that proteins binding to the rep20 sequence alter the

DNA structure in such a way as to make it slightly more accessible to the transcription machinery.

5.8 Extended time course

In all the experiments described so far, CAT activity has been assayed during or just after the first life-cycle following transfection. As observed in section 5.5, none of the promoters under test showed much activity in the first 20 hours post-invasion, even those predicted to be most active in rings. However, in section 5.6, the construct pCAT.KHRP.P was observed to have significant activity between 48 and 52 hours post-invasion, which seems most likely to correspond to the development of early ring stages of the second life-cycle following transfection. This raises the possibility that the expression of episomal genes may be different in later life-cycles from that immediately after electroporation. The second life-cycle is accessible to study by transient assay, as even though segregation during schizogony is inefficient, a reasonable amount of plasmid persists to the next generation.

Aliquots of parasitized erythrocytes containing rings synchronized to a 6 hour invasion window at 6-7% parasitaemia were electroporated with pCAT.Enol.P or pCAT.KHRP (both promoters are predicted to have early, but slightly different, expression patterns) from 7-11 hours post-invasion. After culturing as normal, duplicate pairs of flasks were harvested at approximately 8 hour intervals from 20-100 hours post-invasion. Four control aliquots were transfected with pCAT.Cam.P, all of which were harvested 44 hours post-invasion. In order to maintain the parasite culture for the longer experimental period, at 45 hours post-invasion the medium was aspirated off from those flasks which remained, and fresh medium was added along with 800µl fresh unparisitized erythrocytes. At 77 hours post-invasion, the medium was replaced again in the flasks with the longest culture period.

Following CAT assays of extracts obtained from all the flasks, marked differences could be seen between the expression patterns of the first and second life-cycles (fig 5.9). The first life cycle develops very much as in earlier experiments, with little activity before 20 hours, then an increasing accumulation of CAT, peaking in schizont stages (36-44hr). It appears that on both occasions when the medium was changed (44-52hr and 76-83hr)

some parasites were lost, probably owing to changing the medium by aspiration. In the 44-52hr period for pCAT.Enol.P the measured CAT increased even so, showing that there must have been very high expression to outweigh any losses.

Both plasmids, however, show a drop in total CAT measured between 52-59 hours. For CAT enzyme to be lost, some of the parasites which contain it must be dying, releasing it into the medium. This can be explained if the trauma of electroporation leaves some parasites able to mature to schizonts, but unable to re-invade – a common phenotype of *P. falciparum* parasites which have been stressed. The drop occurs after 52 hours, by which time invasion was expected to be complete; however, thin blood smears showed that schizonts were still present at this time point. Stressed parasites might also be slowed in their development, and therefore form a sub-population which dies at this stage, after the healthy parasites have already re-invaded, but are expressing too little CAT to outweigh the losses. These considerations make it impossible to be sure of the actual amount of CAT produced from 44-59 hours, but the overall trends are clear, as described above.

However, there is a clear, strong burst of expression from 59-69hr with both plasmids. This corresponds to rings 11-21 hours old, which as previously mentioned showed little activity in the first life cycle; in the second life-cycle this is the period of strongest expression, correlating with the microarray data which indicates that both of these promoters are most active in rings. The enolase promoter is also expected to be active in early trophozoites, which the KAHRP promoter is not; this distinction is seen in the 69-76hr time band, although the errors make this observation less certain. From 76hr on the errors are too great to be confident of describing any of the trends, and there is little gain in overall CAT activity; examination of thin blood smears also showed that the cultures had lost much of their synchrony by this point. Presumably there is still some CAT being produced by parasites which carry plasmid, but some is also being lost due to parasite death.

The overall finding of this experiment, therefore, is that the CAT expression observed in the second life-cycle after transfection matched the pattern predicted for these two promoters much more closely than that in the first cycle. Further data are required to characterize expression in the last 20 hours of the second life-cycle.

5.9 Cloning full length intergenic regions

All the promoters examined so far in this chapter have been no more than 1.2kb in length. As was observed for the PfEMP-1 promoter, important sequences located more than 1.2kb upstream of the open reading frame may have been omitted. Therefore three promoters were selected to have the entire intergenic region between them and their preceding open reading frames cloned – the Enolase, KAHRP and MSP-1 promoters. These three were selected because they all have easily measured expression levels, have different predicted expression patterns, and are preceded by relatively short intergenic regions; 1.8kb, 3.8kb and 3.3kb respectively. The three intergenic regions were amplified from genomic DNA by PCR with *Pfx* polymerase in sections of between 800 and 1400bp, shown schematically in fig 5.10, using the primers shown in table 5.3 (the section of each gene closest to the transcriptional start had, of course, been cloned previously in each case). Each section was then subcloned into pCR4-Blunt TOPO and sequenced, before assembling the full-length intergenic regions. Fig 5.10 also shows the identity and orientation of the adjacent orf at the other end of each non-coding region; these orfs are present in a tail-to-head orientation relative to Enolase and KAHRP, so we would expect to find 3' UTR sequences for the adjacent orf within fragments Enol2 and KHRP3. However, the other orf adjacent to the MSP1 non-coding region is present in a head-to-head orientation, so additional promoter sequences should be found in MSP1.3.

Table 5.3. Oligonucleotides used in the cloning of complete intergenic regions from *P. falciparum* genes. All were used as primers for PCR. Restriction sites are highlighted.

Fragment name	Fragment size (bp)	Forward primer (_5)	Reverse primer (_3)	Restriction site, 5'	Restriction site, 3'
Enol2	848	ATTA**GGTACC**GA GAAAAACAAATA TTATTTGATGAAG	TAAA**TACGTA**TTTA TATATGAGAATTTA TTACG	*Kpn* I	*Sna*B I/ *Bsa*A I
KHRP2	1342	ATTA**GGTACC**TA TTTAGAATACAAT CTTAGCATC	ATTA**CTCGAG**TAT ATATAAAGAATAT AAAAAATATGCAT G	*Kpn* I	*Xho* I
KHRP3	1307	ATTA**GCCGGC**GA AAATAACATGAG AAATGTTAC	ATTA**GGTACC**CTT GTTTTTTTTATAAA ATTATGGTG	*Nae* I	*Kpn* I
MSP1.2	1102	ATAT**GGTACC**TT TTACTTCTTGGGT GTG	ATTA**CTCGAG**TTTT AGTAATAAACATA CATGTAATTAAAAT ATAAC	*Kpn* I	*Xho* I
MSP1.3	1140	ATAT**GAGCTC**TT TTCGTTACTCATT TTTATGTG	ATTA**GGTACC**TCA AAAGAAGAAGGAC GC	*Sac* I	*Kpn* I

The Enol2 fragment was designed to overlap with the Enol sequence cloned in section 5.4, and utilize a unique *Bsa*A I site which is found in the genomic enolase sequence. Owing to the presence of a *Bsa*A I site in the pGem3zf(+) vector backbone, the Enol2 sequence could not be directly cloned into pCAT.Enol.P; instead, a triple ligation approach was used. An old Enol PCR product was digested with *Sal* I and *Bsa*A I, Enol2 was excised from pCR4-Blunt TOPO using *BsaA* I and *Kpn* I, and p1-4 (cf. section 6.4) was digested using *Kpn* I and *Sal* I. The larger fragment from the p1-4 digest was then gel purified and ligated with the two enolase promoter fragments, with an approximate vector:insert:insert molar ration of 1:3:3. The ligation products were transformed into TOP10 cells and PCR screened using the M13(-20) primer (present in the vector sequence) and the Enol2_3 primer. Positive colonies were cultured and the presence of the entire intergenic region verified by restriction analysis and sequencing. The correctly assembled plasmid was called pCAT.Enol12.

The other two intergenic regions did not contain any restriction sites useful for cloning, so the primers were designed to make minor sequence alterations to introduce them, detailed in fig 5.10. The intergenic regions were assembled sequentially starting from their 3' ends: first pCAT.KHRP.PcDT and pCAT.MSP1.PcDT were digested with *Xho* I and *Kpn* I, while KHRP2 and MSP1.2 sequences were excised from pCR4-Blunt TOPO with the same enzymes. After gel purification of the appropriate fragments, they were ligated together, transformed into *E. coli* strain PMC-103 (which reproduces long AT-rich sequences more faithfully than standard strains) and grown to produce pCAT.KHRP12 and pCAT.MSP1.12, each of which had their integrity verified by restriction analysis and sequencing. This was followed by similarly incorporating fragment MSP1.3 into pCAT.MSP1.12 using *Kpn* I and *Sac* I, generating pCAT.MSP1.123. However, despite several attempts, KHRP3 could not be introduced into pCAT.KHRP12; after digesting these two sequences with *Kpn* I and *Nae* I, ligating them, and transforming them into PMC103 cells, few colonies were obtained and all of them contained large rearrangements or deletions in the plasmid sequence. It was concluded that the full length KHRP 5' intergenic region must be unusually unstable, or else somehow toxic to *E. coli.* Hopefully this is not too great an omission, as it is likely that all the control sequences for KAHRP were included in the plasmid pCAT.KHRP12, and the KHRP3 region mainly contains promoter sequences for the adjacent orf.

5.10 Experiments on full length intergenic regions

On the basis of the observations in section 5.8, it was decided that the second life cycle after electroporation should be examined for promoter activity from the full-length intergenic regions cloned in section 5.9. The results of section 5.8 were made more difficult to interpret by the loss of parasites when replacing medium by aspiration, and the overgrowth of the culture leading to many of the parasites dying in the latter half of the second life cycle. The experimental protocol was therefore modified to take account of these factors: additional unparasitized erythrocytes were added immediately after electroporation, reducing the parasitaemia to ensure efficient invasion at the end of the first cycle. The health of the parasites was maintained by replacing the culture medium more often and/or using a larger volume of medium. When it was necessary to replace the medium, the culture was first spun down at 2000rpm/815g to pellet all the cells, prior to the medium being gently aspirated off; this avoided the loss of parasite material observed in section 5.8. Also, in the second life cycle there are more parasites in the culture than in the first, so the volume of parasites at each time point was estimated from counting the parasitaemias of thin blood smears, and this value was used to normalize the CAT assay data.

Initially, pCAT.Enol.P was compared with pCAT.Enol12. Erythrocytes infected with rings at 15% parasitaemia and synchronized to a 2 hour time window on the day of transfection were electroporated with one of the two plasmids as normal and incubated in 15ml RPMI/Alb, which was replaced every 24 hours post-invasion. Six hundred microlitres of fresh unparasitized erythrocytes were added at 24 hours post-invasion. For each plasmid, duplicate flasks were harvested at 44, 52, 72 and 92 hours post-invasion, and the cell extracts assayed for CAT activity. As shown in fig 5.11a, both plasmids had produced similar quantities of CAT by the end of the first life cycle, and both cultures lost a small amount of activity during the first schizogony, presumably due to some CAT-bearing merozoites failing to invade. However, in the first half of the second life cycle there is a dramatic difference; pCAT.Enol.P produced essentially no new CAT, while pCAT.Enol12 was highly active. Between 72 and 92 hours post-invasion both samples lose a lot of activity. This is due probably due to cell death in the nutrient-hungry trophozoite stage, as the cultures were at a very high parasitaemia by this point.

These observations show that sequences present in the Enol2 region are necessary for high ring stage expression from the enolase promoter, for which the PcDT element cannot act as a substitute. Similarly in section 5.5 pCAT.Enol.P was not able to promote high ring stage expression in the first life-cycle after electroporation.

Next, the four constructs containing sections of the MSP1 5' intergenic region were tested. As the additional erythrocytes and medium changes in the previous experiment were not sufficient to prevent the culture from overgrowing, in this case two aliquots of 200µl parasitized erythrocytes electroporated with the same plasmid were pooled, diluted with 2ml unparasitized erythrocytes, and cultured in 50ml RPMI/Alb which was replaced daily. Parasites were harvested 80hr post-invasion, at which point they were late trophozoites in the second cycle, and not overgrown. For all the samples CAT activity was very low (fig 5.11b), as previously observed for pCAT.MSP1.P (fig 5.4). In fact, pCAT.MSP1 gave CAT activity indistinguishable from background; in this case the PcDT element was necessary to obtain detectable CAT, as previously observed with the PfEMP3 5' non-coding region (fig 5.6). However, pCAT.MSP1.12 and pCAT.MSP1.123 both gave activity which was not only detectable but higher than that of pCAT.MSP1.P. Interestingly, pCAT.MSP1.12 gave significantly better expression than pCAT.MSP1.123. We can conclude from this that important sequences for the efficient function of the MSP1 promoter lie in the MSP1.2 region. The fact that pCAT.MSP1.123 promotes less efficiently suggests that negative regulatory sequences lie in the MSP1.3 region.

Finally, pCAT.KHRP and pCAT.KHRP12 were examined. As in section 5.8, samples were taken to cover the whole of both life-cycles. Parasites were synchronized to within a 2 hour invasion window on the day of transfection. The increased culture volumes and more frequent medium changes used for the MSP-1 promoter variants above were retained, but owing to the parasitaemia being much lower at electroporation (4-5%) four aliquots of electroporated parasites were pooled in each flask. Also, rather than harvesting the entire culture at one time point, one eighth of the contents of each flask was removed using a sterile pipette at each time point, harvested, and then assayed for CAT activity. This means that only the equivalent of 100µl parasites from the initial electroporations were being harvested at each point, and this at low starting parasitaemia; however, given the strong promoter activity previously observed for pCAT.KHRP (fig 5.6) it was expected that this would still give detectable levels of CAT.

The results – shown in fig 5.12 – give a clear picture of expression pattern differences between the two life cycles, and between the two plasmids under test. The blood smears taken throughout the experiment showed some loss of synchrony with time, but no more than is usually observed in culture; this suggests that the stressful effects of electroporation and culture at high parasitaemia, which had previously been observed to lead to loss of synchrony (section 5.8) were alleviated by reducing the parasitaemia and increasing the culture volume. However, during the first life cycle CAT activity was not detectable from pCAT.KHRP due to the smaller quantity of parasite material harvested at each time point. Intriguingly, CAT activity was readily detected from pCAT.KHRP in the second cycle after electroporation, showing that it is much more active in the progeny of electroporated parasites than in those parasites themselves, even after normalizing for the increased number of parasites in the second cycle. The same is true for pCAT.KHRP12, which is more active than pCAT.KHRP throughout but also much enhanced in the second cycle relative to the first.

The timing of expression from each plasmid is shown in fig 5.12b; pCAT.KHRP12 showed strong expression during the 22-29hr time window and constant expression from 46-77hr, corresponding to the whole of the ring stages of the second erythrocytic cycle, from immediately after invasion through until development to pigmented trophozoites. This is consistent with the high expression levels observed for the genomic KAHRP promoter during ring stages. In contrast, pCAT.KHRP was only active from 46-53hr and 70-77hr. This latter time window is shifted relative to the previous experiment with this plasmid (fig 5.9) where it was active in the 59-69hr time window, but not afterwards; this may be an artifact caused by the loss of synchrony and overgrowth of the culture in the earlier experiment. The apparent drop in CAT expression per parasite during the last two time points for both plasmids is not due to a drop in total CAT activity, which in fact continued to climb at a low rate, but an increase in parasite number as the second schizogony took place – this had already begun by the 93hr time point. This highlights an interesting contrast with the first schizogony at about 46-50hr, where even though parasite number increased fourfold, CAT activity from parasites carrying either plasmid increased by an even greater amount.

These data support the earlier observations that *P. falciparum* promoter sequences included on episomes are more active and reflect their genomic stage specificity more

closely in the second erythrocytic cycle than in the cycle immediately following electroporation. Also, they show that the KHRP2 region contains important sequences for obtaining efficient expression from the KAHRP promoter.

5.11 Conclusions

In this chapter the regulation of gene expression by non-coding sequences located 5' of open reading frames was explored in *P. falciparum,* using CAT as a reporter gene. This regulation is interpreted as occurring chiefly at the level of transcription, in which case the quantitative CAT assay used relates the levels of CAT measured directly to the activity of promoters in the DNA sequences. Another important factor is mRNA stability, as an unstable RNA produced in large quantities may not persist for long enough to make as much protein as smaller quantities of a more stable RNA. Gene expression is probably also regulated at the translational and post-translational levels, but as the same open reading frame and 3' untranslated region are being transcribed from all the 5' upstream sequences in this study, it is assumed that the contribution of these systems to CAT levels is essentially constant throughout.

Two of the promoters identified – GA3PDH and KAHRP – give very high expression and may be superior to those already in general use. Although both are relatively long, not all of that length is necessarily essential to promoter activity; further studies should include truncations and internal deletions to identify the shortest segment that gives optimal expression. Putative promoters from several genes encoding proteins thought to be important in erythrocytic parasitism, MSP1, PfEMP3, and Rif2.7, showed lower expression levels, suggesting that their proteins are not as abundant as the others tested.

One aim of the experiments was to characterize promoters with stage-specific activity, and while some stage-specific differences were observed, most of what was expected was absent, at least from the first life-cycle following transfection. An exception was the H70L promoter which was only active in mature trophozoites. This promoter was also tested as a potential inducible promoter in response to heat shock; however, its response was too weak, and the effects of the shock on the parasites too great, for this to be practical. A putative metallothionein promoter was also tested for response to heavy metal ions but showed no detectable activity.

When the second life-cycle after electroporation was examined with three different plasmids, expression patterns were much closer to those expected. This raises an interesting question about the expression of episomal genes compared with those in the genome; how closely the two relate has obvious bearing on future using transfection to express genes in *trans*. One possible explanation for the data is that in the 10-20 hours immediately after electroporation, the traumatized parasites are in some form of arrest and new gene expression is halted for some time. This seems unlikely, however, given that many parasites appear to develop normally. A more interesting possibility is that the parasite cannot properly regulate expression from plasmids made in *E. coli,* but can do once it has replicated the plasmid itself, as a result of epigenetic factors. Various possible mechanisms by which this could occur are discussed in section 8.4.

The action of the PcDT element on a number of promoters supports the suggestion that it contains a bi-directional enhancer element, and therefore that *P. falciparum* promoter structure resembles that of other eukaryotes, despite the inability to detect sequence similarity. The fact that it interacts differently with different promoters (notably GA3PDH and KAHRP) implies that each promoter recruits specific transcription factors, not all of which interact in the same way. Also, the rep20 sequence was not found to have detrimental effects on gene expression, confirming its usefulness for all kinds of transfection experiments.

Promoter sequences in *P. falciparum* constitute an area where little research interest is currently invested, which could yield a great deal of valuable information about the disease and how to treat it. The findings of this chapter show that while the transcription system is divergent from other eukaryotes, it displays similar activities and is amenable to analysis.

Fig 5.1: **Differential expression patterns of the Calmodulin promoter and the Calmodulin/PcDT promoter pair.** Aliquots of erythrocytes parasitized with *P. falciparum* ring stages at 11% parasitaemia were transfected as described with plasmids containing the CAT open reading frame under the control of each promoter combination, harvested by saponin lysis at the times shown, and assayed for CAT activity.

Cam - CAT orf controlled by *P.falciparum* calmodulin promoter (only two time points measured for this construct)

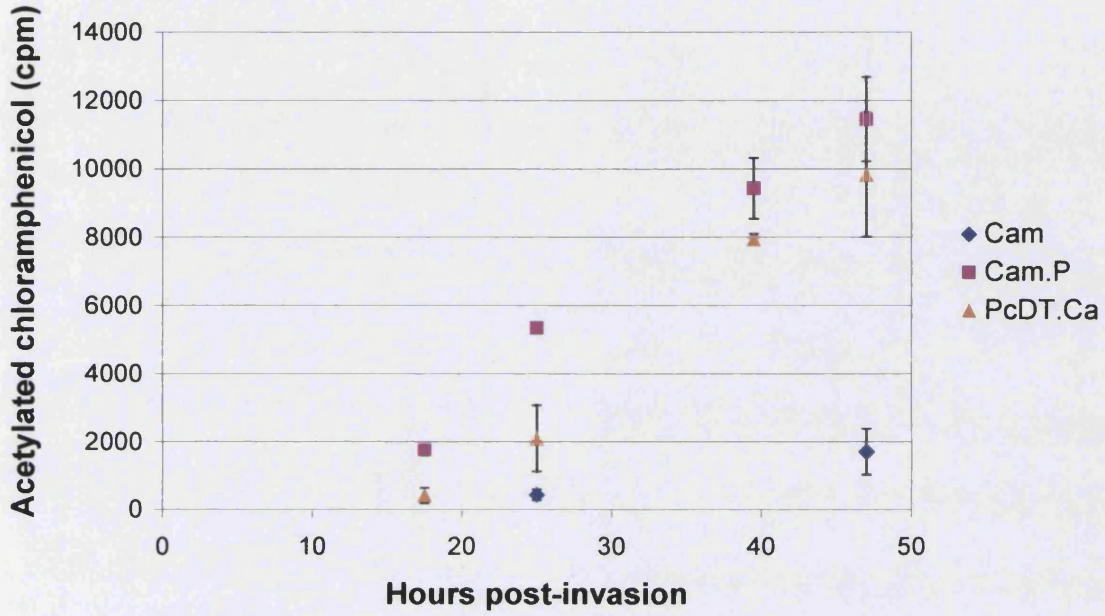
Cam.P - CAT orf controlled by calmodulin promoter with adjacent PcDT promoter in reverse orientation

PcDT.Ca - CAT orf controlled by PcDT promoter with adjacent calmodulin promoter in reverse orientation

A. Time course of total acetylation by CAT throughout one parasite life-cycle for each promoter combination. Error bars show the standard deviation of two samples.

B. Average acetylation per unit time between each pair of time points in the time course above. Error bars show the sum of the standard deviations of the end point of each time window, divided by the time period.

A.


Acetylated chloramphenicol (cpm)
14000
12000
10000
8000
6000
4000
2000
0
0
10
20
30
40
50
Hours post-invasion
Cam
Cam.P
PcDT.Ca

B.

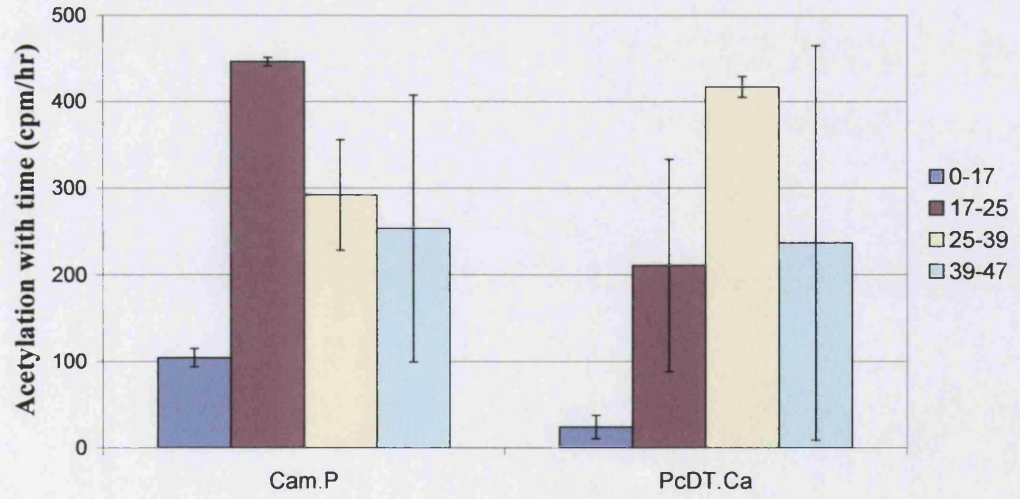
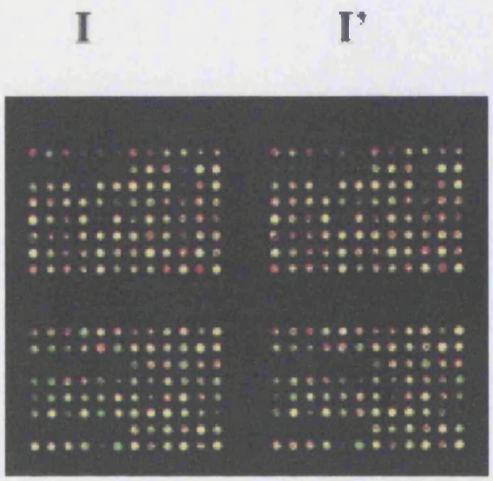

Acetylation with time (cpm/hr)
500
400
300
200
100
0
Cam.P
PcDT.Ca
0-17
17-25
25-39
39-47

Fig 5.2: **Use of microarray data to identify genes with stage-specific expression.** The upper image on the facing page depicts a small section of the microarray constructed and tested by Mamoun and colleagues. Messenger RNA was isolated from parasites synchronized to particular stages in the intraerythrocytic life-cycle: rings (0-24hr), early trophozoites (24-30hr), late trophozoites (30-36hr), early schizonts (36-42hr) and late schizonts (42-48hr). These mRNAs were labelled with the green fluorescent dye Cy5 and hybridized to microarrays containing 944 *P.falciparum* ESTs. For normalization all arrays were also hybridized with late trophozoite mRNA labelled with the red fluorescent dye Cy3. Where spots appear green, their corresponding ESTs are expressed at a higher level in the stage under test than in late trophozoites, whereas red spots indicate a lower level expression in the stage under test. Yellow spots occur when expression levels are similar in both stages. The entire array was duplicated (indicated by I and I') to control for variation induced by position on the array.

Fluorescence intensities for all array locations were quantified automatically and transferred to a spreadsheet, a small section of which is shown opposite. Each EST sequence was used in a BLAST search against the *P. falciparum* genome database to identify the gene it originates from ('NAME'). NCBI numbers are those of each EST, while 'Grid Loc.' indicates the location of each EST in the array (each of which appears twice, one for each of the duplicate arrays). The last column shown – R Cy5/3 – is the first data column, where the ratios between Cy5 and Cy3 levels for each stage (in this case rings) are calculated and displayed.



Clone	NAME	NCBI no.	Grid Loc.	R Cy5/3
313	threonyl-tRNA synthetase,cytopasmi	t18023	2-8-6	2.05
313	threonyl-tRNA synthetase,cytopasmi	t18023	2-8-6	2.34
314	glyceraldehyde 3-phosphate dehydroge	t18024	3-4-1	7.75
314	glyceraldehyde 3-phosphate dehydroge	t18024	3-4-1	6.10
315	NADH dehydrogenase subunit 1 [Ithomia	t18025	4-4-1	1.22
315	NADH dehydrogenase subunit 1 [Ithomia	t18025	4-4-1	1.43
316	circumsporozoite-related an	t18026	3-4-2	1.19
316	circumsporozoite-related an	t18026	3-4-2	1.19

Mamoun et al, 2001

Fig 5.3: **Schematic of plasmids used for promoter analysis.** Promoter sequences of interest were cloned into plasmid p1-4 (cf section 6.4) in place of the PfCAM 5' region, to generate plasmids as shown. Arrows indicate the direction of transcription from each promoter sequence.

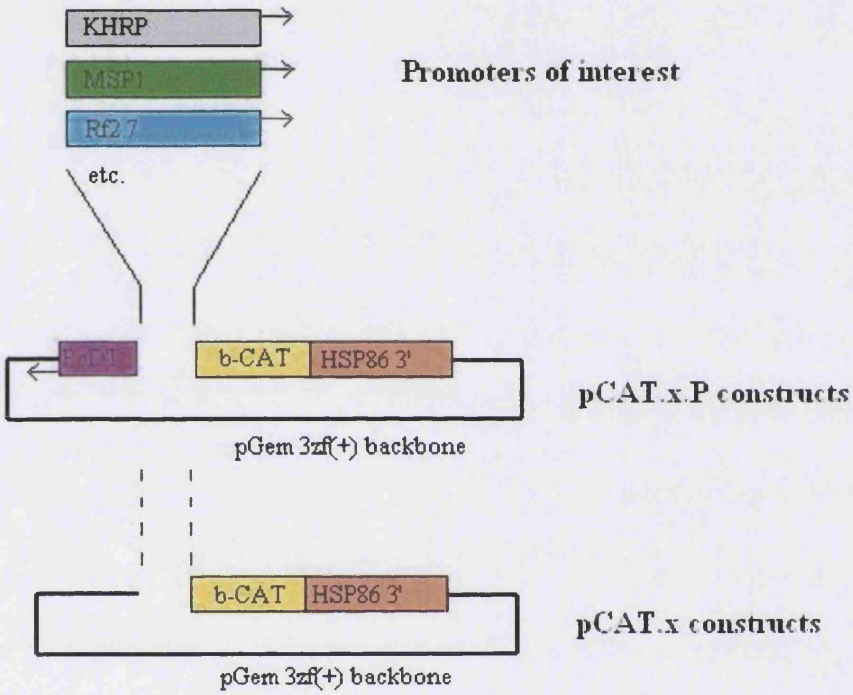

KHRP
MSP1
Rf2.7
etc.
Promoters of interest
b-CAT
HSP86 3'
pCAT.x.P constructs
pGem 3zf(+) backbone
b-CAT
HSP86 3'
pCAT.x constructs
pGem 3zf(+) backbone

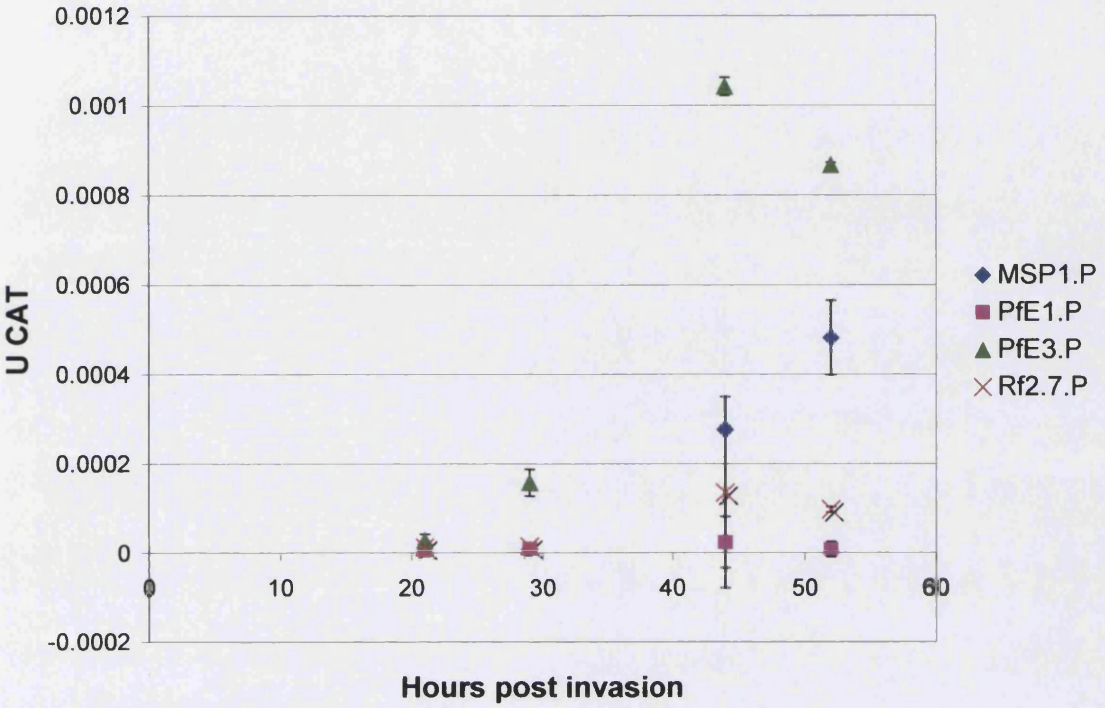
Fig 5.4: **Life-cycle scan of expression of promoters with low activity.** Erythrocytes parasitized with 7.5% ring stages were transfected at 4-6 hours post-invasion with plasmids containing the promoters indicated with adjacent PcDT promoter in reverse orientation, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay.

MSP1.P - CAT expressed by the MSP-1 promoter fragment, plus PcDT.
PfE1.P - CAT expressed by the PfEMP-1 promoter fragment, plus PcDT.
PfE3.P - CAT expressed by the PfEMP-3 promoter fragment, plus PcDT.
Rf2.7.P- CAT expressed by the Rifin2.7 promoter fragment, plus PcDT.

A. Time course of CAT activity generated by each promoter across one life-cycle. Error bars show the standard deviation of two or three CAT assay time points.

B. Average CAT activity generated per unit time between each of the time points shown in A. Error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window.

A.


0.0012
0.001
0.0008
0.0006
0.0004
0.0002
0
-0.0002
U CAT
0
10
20
30
40
50
60
Hours post invasion
MSP1.P
PfE1.P
PfE3.P
Rf2.7.P

B.

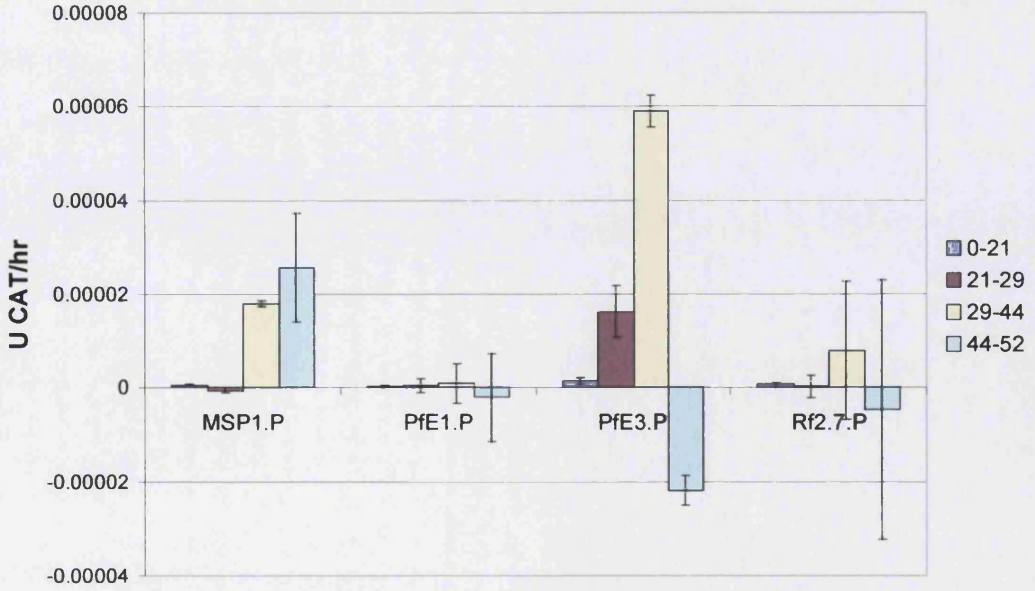

0.00008
0.00006
0.00004
0.00002
0
-0.00002
-0.00004
U CAT/hr
MSP1.P
PfE1.P
PfE3.P
Rf2.7.P
0-21
21-29
29-44
44-52

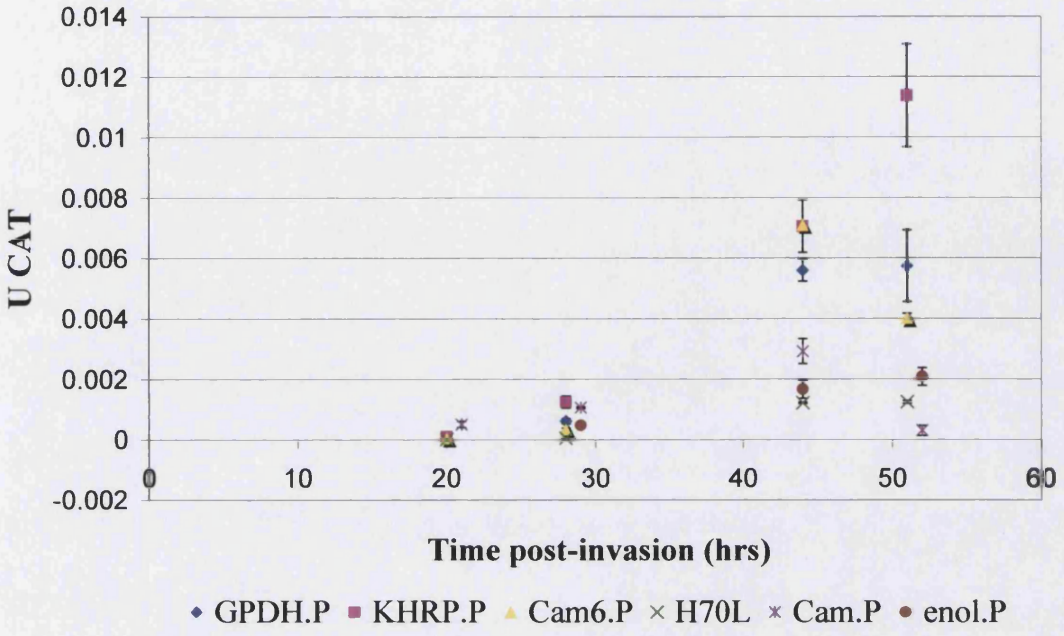
Fig 5.5: **Life-cycle scan of expression of promoters with high activity.** Erythrocytes parasitized with 5-6% ring stages were transfected at 4-6 hours post-invasion with plasmids containing the promoters indicated with adjacent PcDT promoter in reverse orientation, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay.

GPDH.P - CAT expressed by the GA3PDH promoter fragment, plus PcDT.
KHRP.P - CAT expressed by the KAHRP promoter fragment, plus PcDT.
Cam6.P - CAT expressed by the 634bp PfCAM promoter fragment, plus PcDT.
H70L.P- CAT expressed by the HSP 70-like promoter fragment, plus PcDT.
Cam.P - CAT expressed by the full-length PfCAM 5' intergenic region, plus PcDT (used as a control to normalize data from all promoters).
Enol.P - CAT expressed by the Enolase promoter fragment, plus PcDT.

A. Time course of CAT activity generated by each promoter across one life-cycle. Error bars show the standard deviation of two or three CAT assay time points.

B. Average CAT activity generated per unit time between each of the time points shown in A. Error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window.

A.


0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
-0.002
U CAT
0
10
20
30
40
50
60
Time post-invasion (hrs)
GPDH.P
KHRP.P
Cam6.P
H70L
Cam.P
enol.P

B.

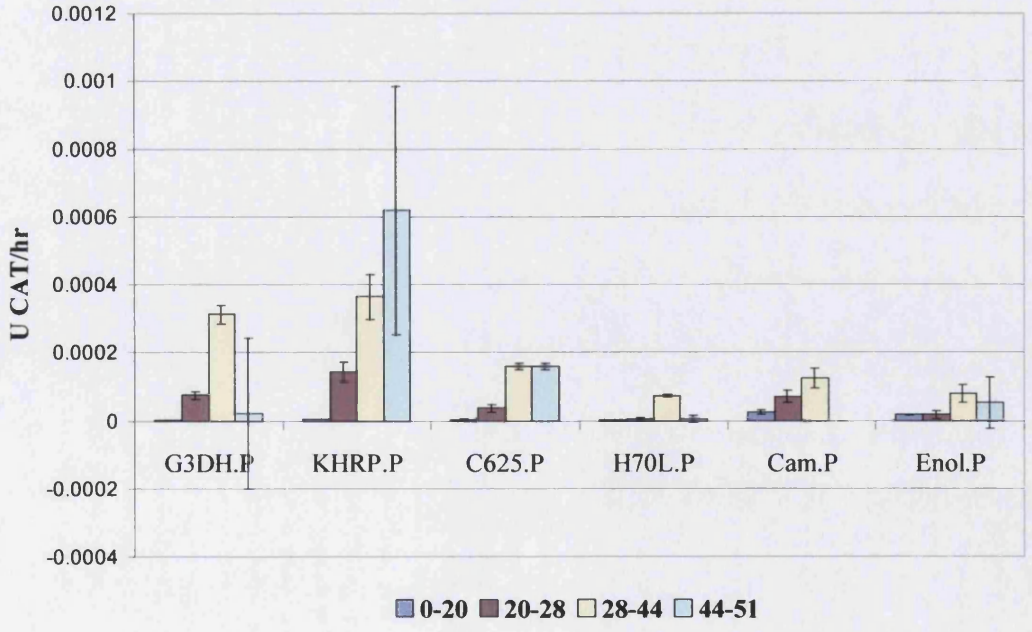

0.0012
0.001
0.0008
0.0006
0.0004
0.0002
0
-0.0002
-0.0004
U CAT/hr
G3DH.P
KHRP.P
C625.P
H70L.P
Cam.P
Enol.P
0-20
20-28
28-44
44-51

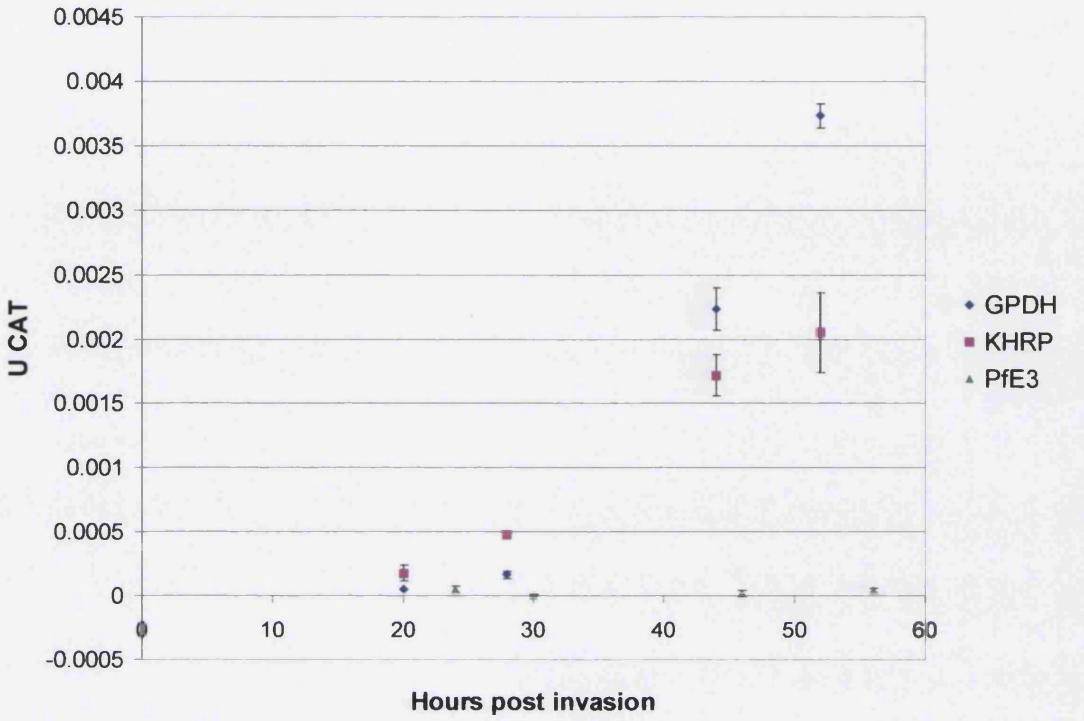
Fig 5.6: **Life-cycle scan of expression of promoters without PcDT element.** Erythrocytes parasitized with 10% ring stages were transfected at 4-6 hours post-invasion with plasmids containing the promoters indicated, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay.

GPDH - CAT expressed by the GA3PDH promoter fragment.
KHRP - CAT expressed by the KAHRP promoter fragment.
PfE3 - CAT expressed by the PfEMP-3 promoter fragment.

A. Time course of CAT activity generated by each promoter across one life-cycle. Error bars show the standard deviation of two or three CAT assay time points.

B. Average CAT activity generated per unit time between each of the time points shown in A. Error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window. No data are shown for the PfEMP-3 promoter, as the activity was zero throughout.

A.


0.0045
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
-0.0005
U CAT
0
10
20
30
40
50
60
Hours post invasion
GPDH
KHRP
PfE3

B.

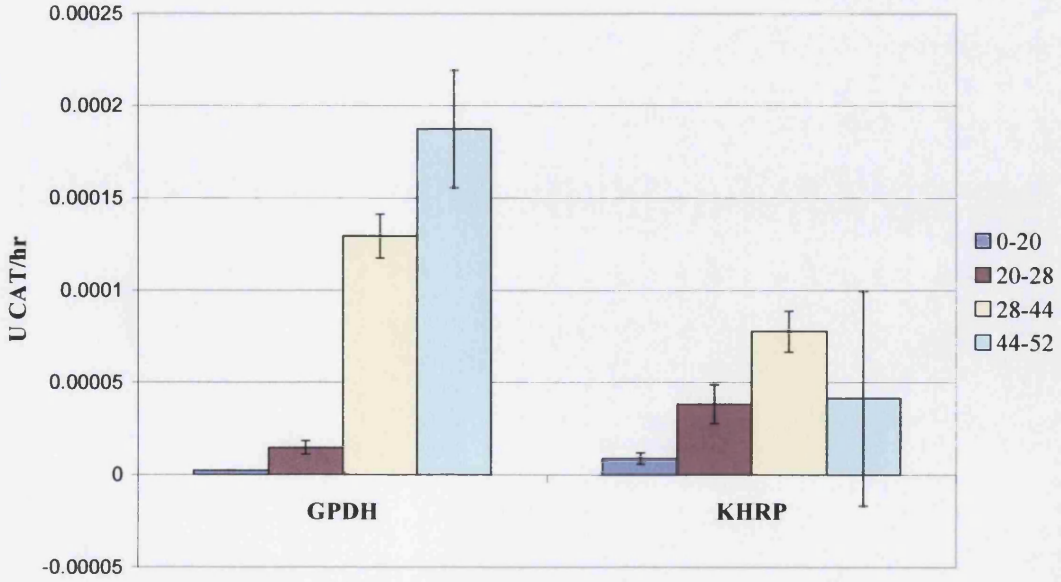

0.00025
0.0002
0.00015
0.0001
0.00005
0
-0.00005
U CAT/hr
GPDH
KHRP
0-20
20-28
28-44
44-52

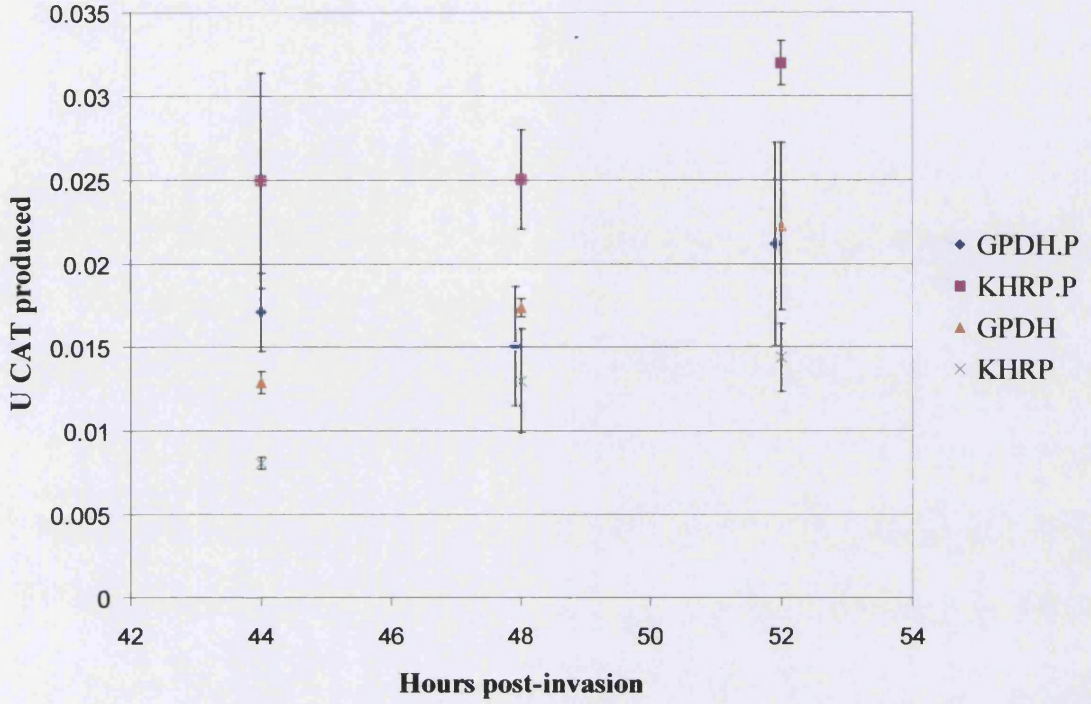
Fig 5.7: **Analysis of expression of strong promoters during schizogony.** Erythrocytes parasitized with 8% ring stages were transfected at 4-6 hours post-invasion with plasmids containing the promoters indicated with or without adjacent PcDT promoter in reverse orientation, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay.

GPDH.P - CAT expressed by the GA3PDH promoter fragment, plus PcDT.
KHRP.P - CAT expressed by the KAHRP promoter fragment, plus PcDT.
GPDH - CAT expressed by the GA3PDH promoter fragment alone.
KHRP - CAT expressed by the KAHRP promoter fragment alone.

A. Time course of CAT activity generated by each promoter across one life-cycle. Error bars show the standard deviation of four CAT assay time points (two each from duplicate transfections).

B. Average CAT activity generated per unit time between each of the time points shown in A. Error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window.

A.


0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
U CAT produced
42
44
46
48
50
52
54
Hours post-invasion
GPDH.P
KHRP.P
GPDH
KHRP

B.

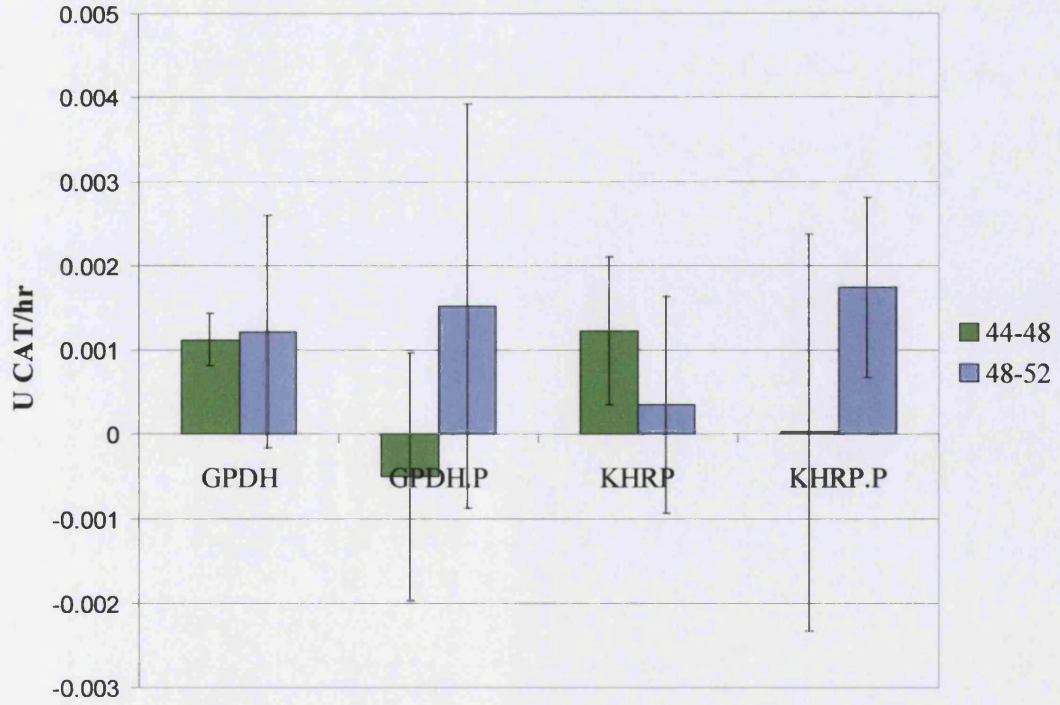

0.005
0.004
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
U CAT/hr
GPDH
GPDH.P
KHRP
KHRP.P
44-48
48-52

Fig 5.8: **Expression of CAT from plasmids containing control elements of interest.** Erythrocytes parasitized by ring stages were transfected at 4-6 hours post-invasion with plasmids, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay. Error bars show the standard deviation of six CAT assay time points (three each from duplicate transfections).

A. Activity of pCAT.H70L.P with and without heat shock. Initial parasitaemia was c. 10% rings. All samples were cultured at 37°C until 46 hours post-invasion. The net CAT activity measured is the left hand of each pair of bars, while the right hand shows the amount normalized to take into account the reduction in parasitaemia in the heat shocked cultures.

No shock - Maintained at 37°C throughout

39 - At 28 hours post-invasion, cultures were transferred to 39°C incubator for 90 minutes

41 - At 28 hours post-invasion, cultures were transferred to 41°C incubator for 90 minutes

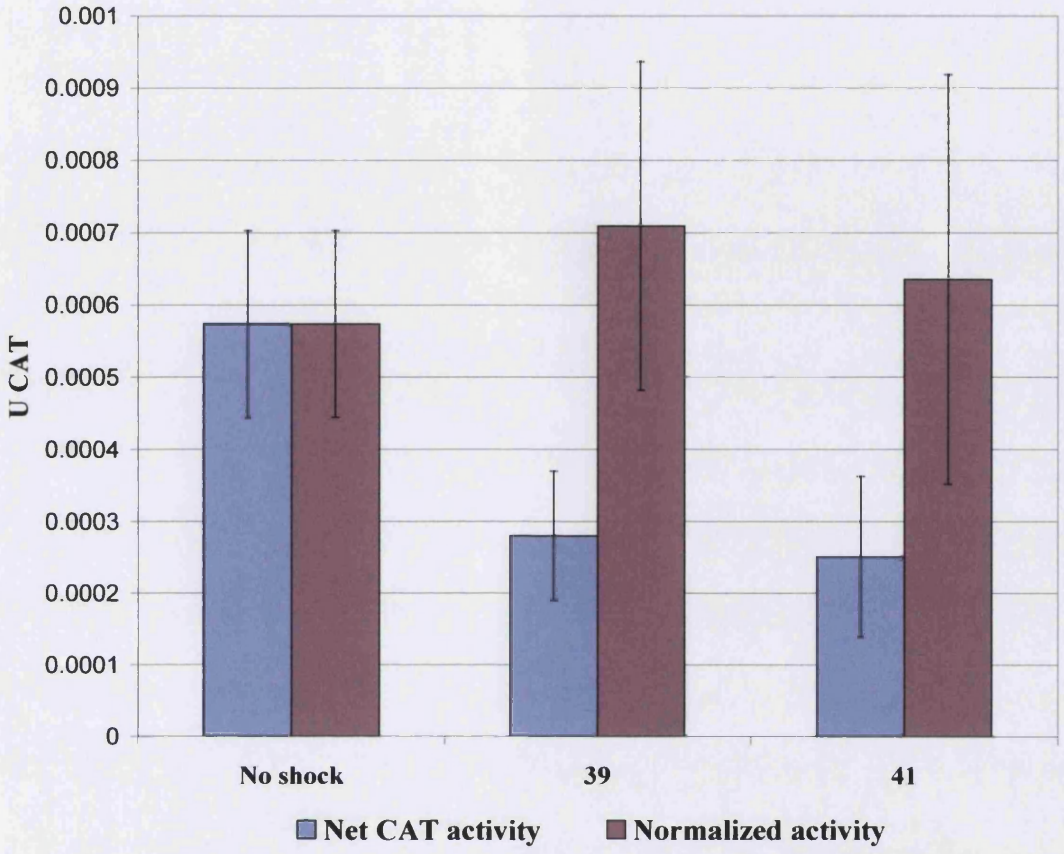
B. Activity of KHRP 5' non-coding region in the presence and absence of rep20 sequence. Samples with an initial parasitaemia of 5% were transfected and cultured for 45 hours before harvesting.

KHRP - CAT expressed by the KAHRP promoter fragment alone.

KHRP.rep - CAT expressed by the KAHRP promoter fragment, plus rep20.

KHRP.P - CAT expressed by the KAHRP promoter fragment, plus PcDT.

A.


0.001
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
U CAT
No shock
39
41
Net CAT activity
Normalized activity

B.

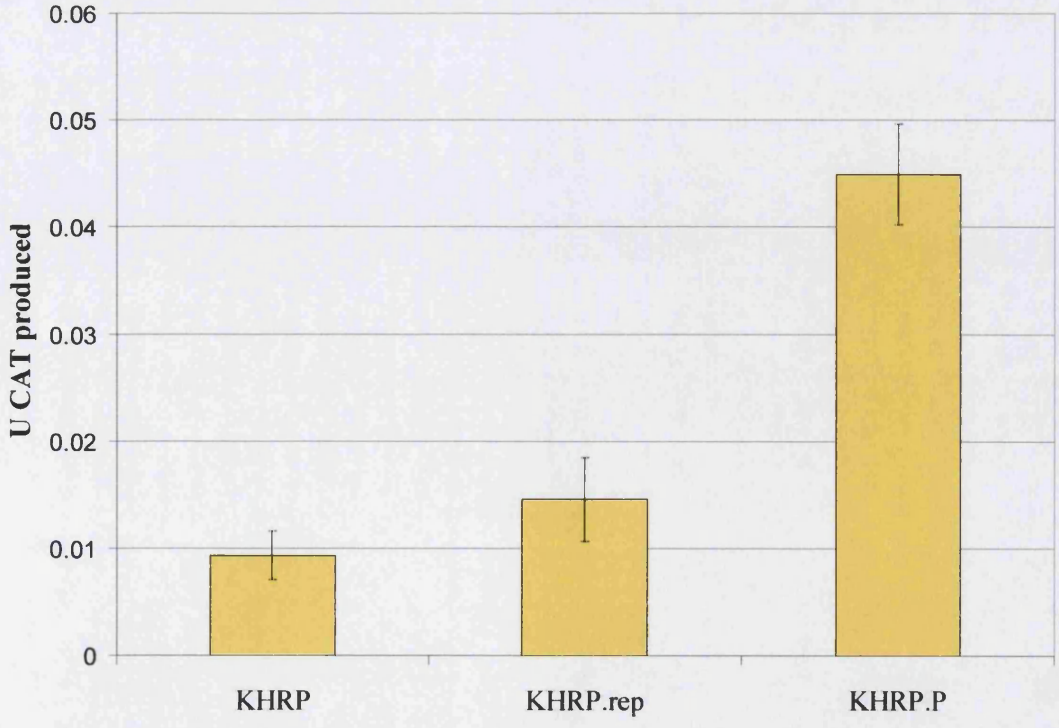

0.06
0.05
0.04
0.03
0.02
0.01
0
U CAT produced
KHRP
KHRP.rep
KHRP.P

Fig 5.9: **Two cycle time course of KAHRP and enolase promoters.** Erythrocytes parasitized with 6-7% ring stages were transfected at 7-11 hours post-invasion with plasmids containing the promoters indicated, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay.

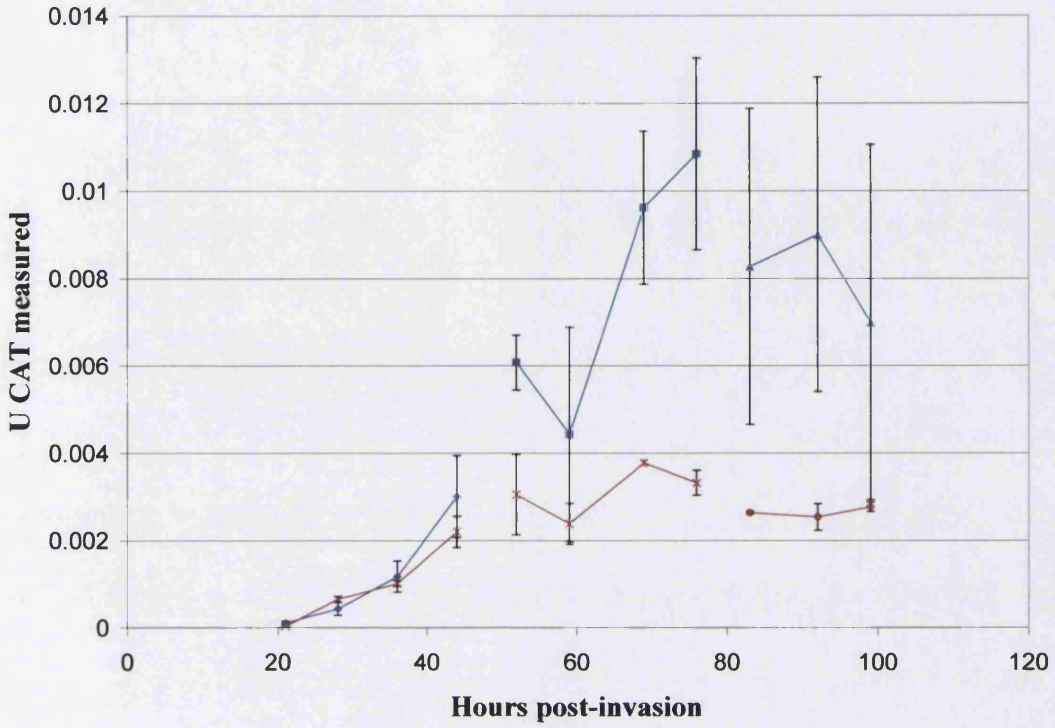
A. Time course of CAT activity generated by each promoter across two life-cycles. Blue symbols show expression from pCAT.Enol.P, and brown symbols show that from pCAT.KHRP. The gaps in the traces indicate where the medium was replaced and some loss of parasites is thought to have occurred. Error bars show the standard deviation of four or eight CAT assay time points, half each from duplicate transfections.

B. Average CAT activity generated per unit time between each of the time points shown in A. Time windows are indicated on the horizontal axis, and error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window.

Enol.P - CAT expressed by the Enolase promoter fragment, plus PcDT.

KHRP - CAT expressed by the KAHRP promoter fragment.

A.


0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
U CAT measured
0
20
40
60
80
100
120
Hours post-invasion

B.

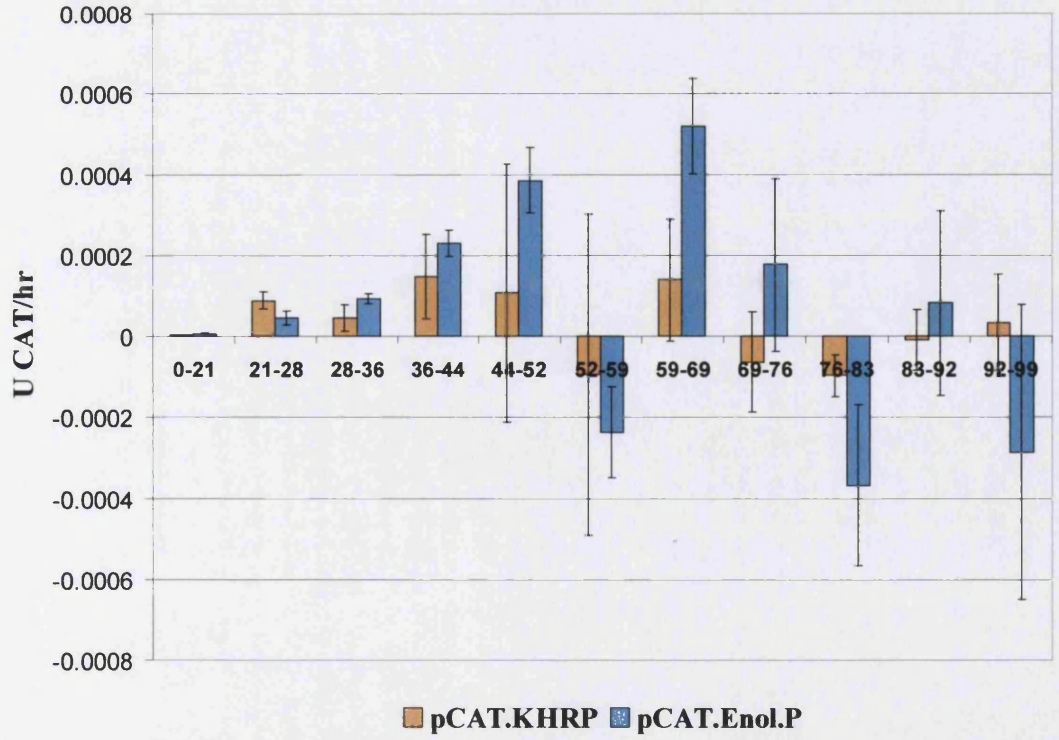
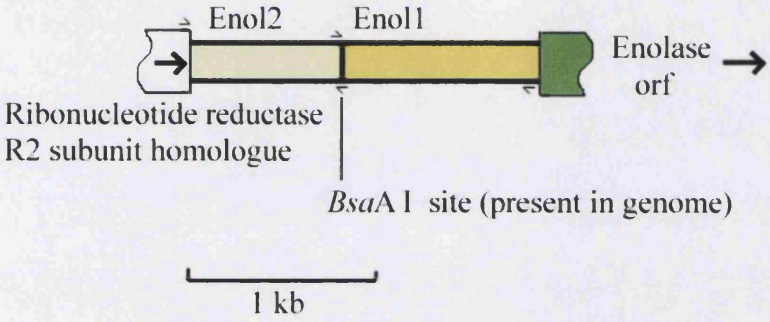
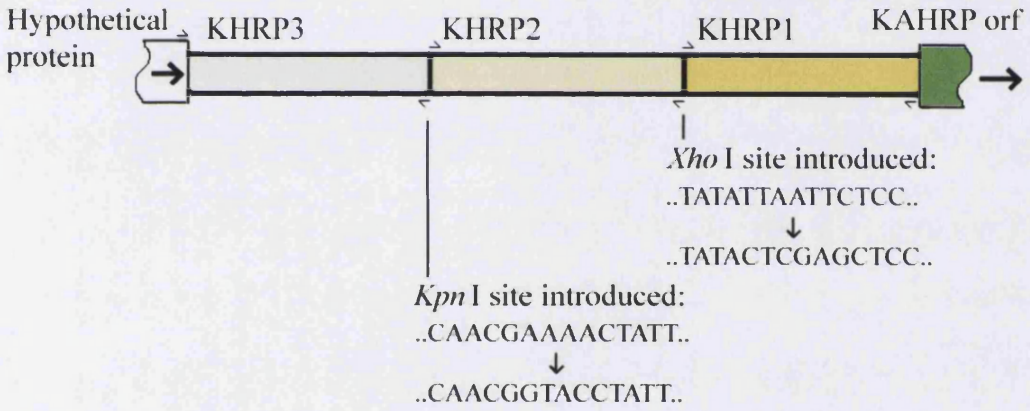

0.0008
0.0006
0.0004
0.0002
0
-0.0002
-0.0004
-0.0006
-0.0008
U CAT/hr
0-21
21-28
28-36
36-44
44-52
52-59
59-69
69-76
76-83
83-92
92-99
pCAT.KHRP
pCAT.Enol.P

Fig 5.10: **Schematic of complete intergenic regions and plan for cloning.** The intergenic regions 5' of the enolase, KAHRP and MSP-1 open reading frames are shown to scale, divided into fragments of a convenient size for cloning. The preceding open reading frames are also identified; heavy arrows show the direction of transcription. Restriction sites in the genomic sequence which were used to facilitate cloning are shown, as are those which were introduced and the sequence alterations made. The location of sequences used for PCR primers are identified by small hairpins.


Enol2
Enol1
Enolase
orf
Ribonucleotide reductase
R2 subunit homologue
*Bsa*A I site (present in genome)
1 kb


Hypothetical
protein
KHRP3
KHRP2
KHRP1
KAHRP orf
Xho I site introduced:
..TATATTAATTCTCC..
..TATACTCGAGCTCC..
Kpn I site introduced:
..CAACGAAAACTATT..
..CAACGGTACCTATT..

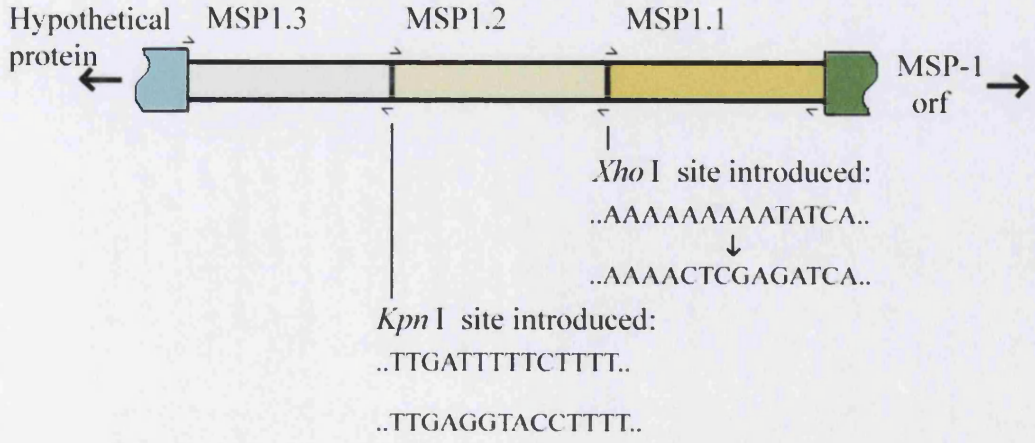

Hypothetical
protein
MSP1.3
MSP1.2
MSP1.1
MSP-1
orf
Xho I site introduced:
..AAAAAAAAATATCA..
..AAAACTCGAGATCA..
Kpn I site introduced:
..TTGATTTTTCTTTT..
..TTGAGGTACCTTTT..

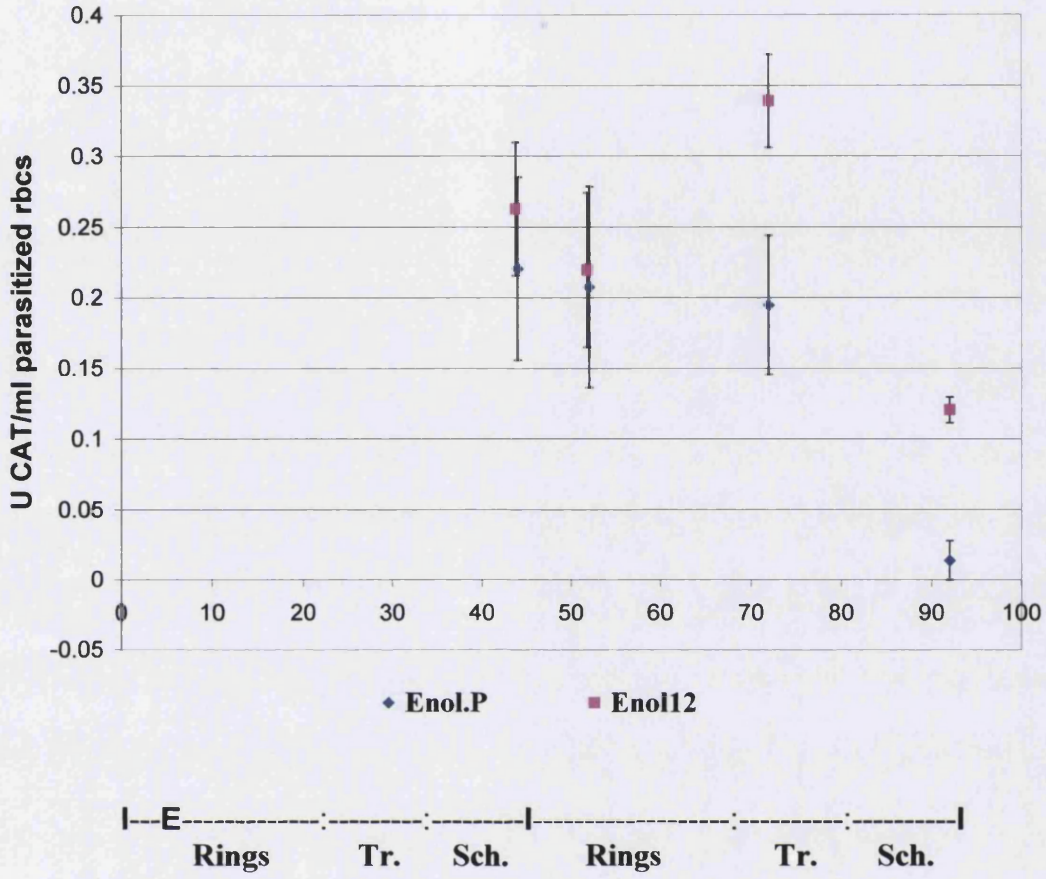
Fig 5.11: **Comparison of promoter activities from partial and complete 5' intergenic regions of PfEnolase and PfMSP1.** Erythrocytes parasitized by ring stages were transfected at 4-6 hours post-invasion with plasmids, and cultured as described. Aliquots were harvested at each time point shown and used for quantitative CAT assay. CAT assay data was normalized by dividing the measured CAT activity by the total packed cell volume of the sample, and then multiplying by the parasitaemia counted on thin blood smears.

A. Parasites were synchronized on the day of transfection to a 2 hour time window, and were at 15% parasitaemia at electroporation. At 24 hours 600µl fresh unparasitized erythrocytes were added to each flask, and the medium was changed daily. Samples were harvested at the time points shown and used for quantitative CAT assay. Error bars show the standard deviation of six CAT assay time points (three each from duplicate transfections).
Enol.P – Parasites electroporated with pCAT.Enol.P
Enol12 – Parasites electroporated with pCAT.Enol12

The line beneath the graph gives an indication of the development of the parasites during the experiment. **I** = times of invasion, **E** = time of electroporation, **Tr.** = trophozoites and **Sch.** = schizonts.

B. Parasites were synchronized to a 2 hour time window 2 cycles prior to transfection, and were all rings at 13% parasitaemia at electroporation. The cells from two electroporations for each plasmid were pooled, diluted with 2ml fresh unparasitized erythrocytes, and cultured in 50ml RPMI/Alb which was replaced 30hr and 54hr post-invasion. Half of each pooled culture was harvested at 52hr post-invasion and the other half at 80hr post-invasion; only the later samples gave measurable CAT activity. Error bars show the standard deviation of four CAT assay time points (two each from duplicate transfections).
M1P - Parasites electroporated with pCAT.MSP1.P
M1 - Parasites electroporated with pCAT.MSP1
M12 - Parasites electroporated with pCAT.MSP1.12
M123 - Parasites electroporated with pCAT.MSP1.123

A.


U CAT/ml parasitized rbcs
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
0
10
20
30
40
50
60
70
80
90
100
Enol.P
Enol12

I---E----------:---------:---------I----------:---------:---------I
Rings
Tr.
Sch.
Rings
Tr.
Sch.

B.

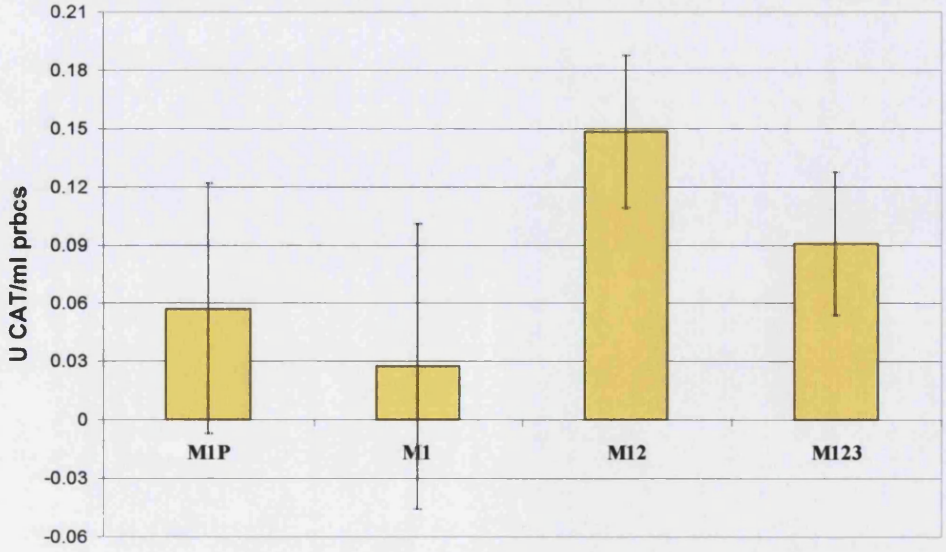

U CAT/ml prbcs
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0
-0.03
-0.06
M1P
M1
M12
M123

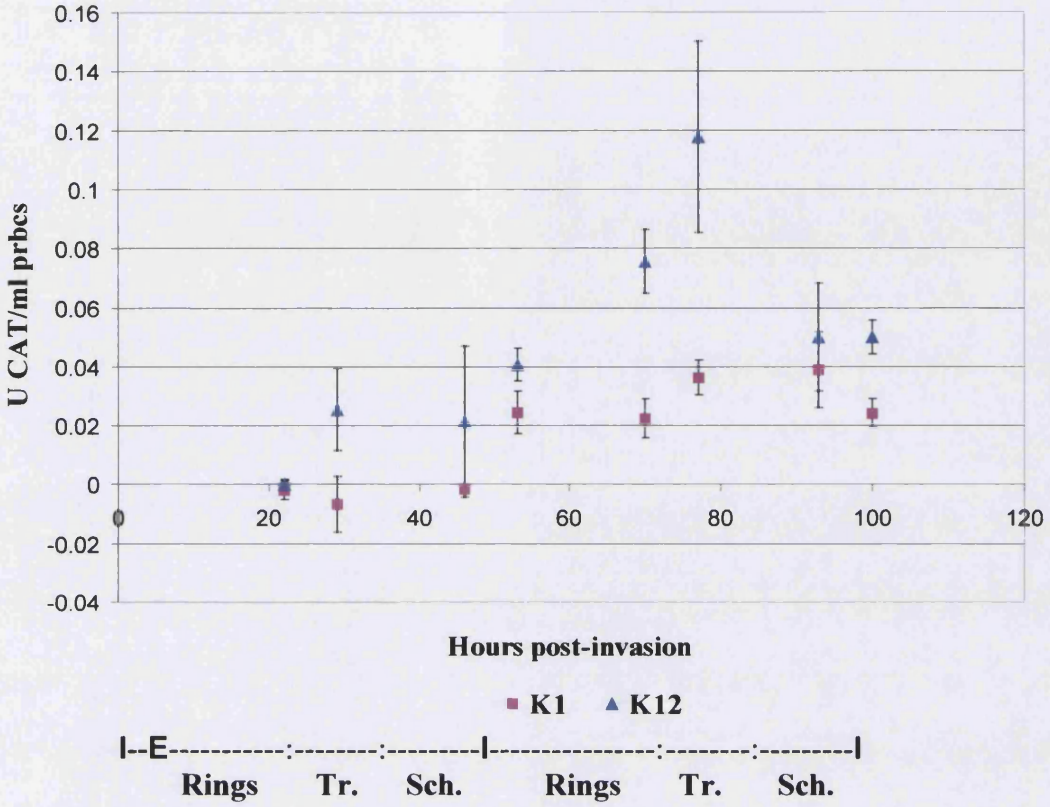
Fig 5.12: **Expression from sections of the KAHRP 5' intergenic region across two erythrocytic cycles.** Erythrocytes were parasitized with young, tightly synchronized rings at 4-5% parasitaemia. Aliquots of 200µl of these erythrocytes were electroporated as normal, and then four such aliquots were pooled in a flask along with 56ml RPMI/Alb and 2ml unparasitized erythrocytes. At this point the parasitaemia was found to be 1.2%; duplicate flasks were set up for each plasmid. At the time points indicated, the flasks were gently shaken to resuspend all the cells, and 7ml was removed from each culture for harvesting and quantitative CAT assay. Data were normalized for parasite volume as before.

K1 - Parasites electroporated with plasmid pCAT.KHRP
K12 - Parasites electroporated with plasmid pCAT.KHRP12

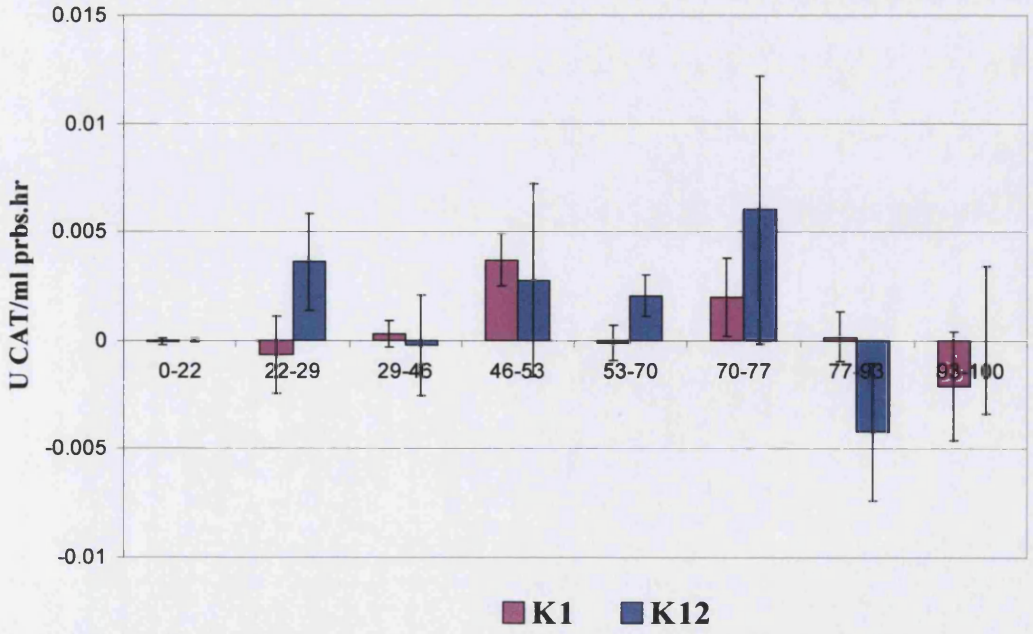
A. Time course of CAT activity generated by each promoter across two life-cycles. Error bars show the standard deviation of four CAT assay time points, two each from duplicate transfections. The line beneath the graph gives an indication of the development of the parasites during the experiment. **I** = times of invasion, **E** = time of electroporation, **Tr.** = trophozoites and **Sch.** = schizonts.

B. Average CAT activity generated per unit time between each of the time points shown in A. Time windows are indicated on the horizontal axis, and error bars are the sum of the errors at the beginning and end of each time window divided by the length of the time window.

A.


0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
U CAT/ml prbcs
0
20
40
60
80
100
120
Hours post-invasion
K1
K12
I--E----------:--------:---------I----------:--------:---------I
Rings
Tr.
Sch.
Rings
Tr.
Sch.

B.


0.015
0.01
0.005
0
-0.005
-0.01
U CAT/ml prbs.hr
0-22
22-29
29-46
46-53
53-70
70-77
77-93
93-100
K1
K12

6. *Transfecting Parasites with an Erythrocyte Surface Label*

6.1 Introduction

As has been described in earlier chapters, the low efficiency of transfection in *P. falciparum* severely limits the usefulness of the technique. Since at best 0.1% of the parasites electroporated take up plasmid, any phenotypic changes are not visible against the untransfected background – except for a phenotype which is completely absent in the wild type parasites, such as CAT activity. One way round this limitation is by stable transfection, in which the parasites which take up plasmid gain a drug resistance gene as well as the gene of interest, so that they can be purified by positive selection. However, this method is time-consuming and can be inconclusive. Typically several weeks of drug selection are required during which the level of parasitaemia is so low as to be undetectable, and if the parasites never grow up again, it is impossible to be certain whether the construct used was harmful to the parasites or whether they died due to some technical failure during the protracted selection period.

One way of sidestepping this problem would be to label the parasitized erythrocytes, which have taken up plasmid, and then separate them from the majority of unlabelled cells. This would remove the background, making it much easier to detect phenotypic changes, and allow them to be observed immediately. Even detrimental effects could be identified if the parasites could be studied within the first life-cycle after transfection.

The methods available to separate labelled cells from unlabelled were considered. If parasites were transfected with a fluorescent label, such as Green Fluorescent Protein (GFP), they could be separated by Fluorescence Activated Cell Sorting (FACS). Transient expression of GFP has been reported in *P. falciparum* (VanWye & Haldar, 1997) but other workers have been unable to repeat this result (Burghaus & Lingelbach, 2001). Stable expression of GFP from an episomal location has been achieved and used to explore protein trafficking (Waller et al, 2000; Wickham et al, 2001) but this again requires the generation of a stable transfectant. Another approach would be to use an antibody specific for a molecule found on the surface of transfected parasitized erythrocytes. The labelled cells could then be separated by physical antibody mediated

interactions (e.g. binding to a solid support such as protein G-coated magnetic beads) or if the antibody were labelled with a fluorophore, by FACS. This strategy was adopted to try and label *P. falciparum* invaded erythrocytes.

6.2 Designing a parasitized erythrocyte surface label

The protein to be used to label transfected parasites as described above must be exported to the parasitized erythrocyte surface, and modified so that it can be distinguished from protein expressed by the wild-type. As described in section 1.4.5, the only *P. falciparum* proteins clearly proven to be exposed on the outer surface of the erythrocyte plasma membrane are members of two multigene families – *var* and *rif*. Each family is found in sub-telomeric clusters at the ends of most of the parasite's chromosomes, and it is believed that the reason for having multiple copies of each gene is to allow antigenic variation; of all parasite proteins these are the most exposed to the immune system (Snounou et al, 2000). Best characterized are the *var* genes, which code for the surface antigen PfEMP-1, but these are relatively large genes and therefore difficult to manipulate in *E. coli* and to transfect with. More convenient are the smaller *rif* genes, which code for surface proteins of 35-45 kDa called Rifins (Kyes et al, 1999). All Rifins have the same gene organization, as shown in Fig 6.1a; they start with a short exon which codes for a putative membrane targeting signal sequence, separated from the major coding region by a short intron. Apart from the signal sequence, two putative transmembrane helices are present, one in the middle of the sequence (orientated from inside to outside the cell) and one towards the C-terminal end (orientated outside in). Therefore most of the C-terminal half of the sequence is predicted to be exposed to the external medium. Supporting this view, most of the sequence is highly conserved between different Rifins, but the region located between the two putative transmembrane helices is polymorphic, with a central 'hyperpolymorphic' section which varies in length and contains barely any conserved residues. It is hypothesized that this region is on the surface of the extracellular domain and is exposed to the immune system, which would explain its high variability. More recently, Abdel-Latif et al (2002) have shown that most individuals in an area endemic for *P. falciparum* carry anti-Rifin antibodies. Therefore Rifins were considered to be ideal for the kind of surface labelling being considered here; moreover, developing such a system might lead to valuable insights into the as yet unknown function of these proteins.

It was reasoned that inserting a small modification, such as an epitope tag, into the polymorphic segment of a Rifin would be likely to display it outside the erythrocyte without disrupting the structure, function or trafficking of the protein. This tag would make transfected Rifin distinct from the wild-type proteins and detectable using antibodies. Various epitope tags were considered and c-myc was selected owing to its small size (10 residues), the ready availability of antibodies against it, and its successful application in a variety of systems at either the N- or C-terminal ends of proteins. The c-myc amino acid sequence EQKLISEEDL was used as the query for a tblastn search against the *P. falciparum* genome database and produced no matches to more than six of the ten residues (and those not sequential). A DNA sequence to code for this epitope was designed taking account of *P. falciparum's* codon bias: GAACAGAAATTAATATCAGAAGAAGATTTA. In a blast search against the genome this sequence produced no hits of more than 17 residues.

Expressing an epitope tag internally in a protein sequence is an uncertain enterprise, as the protein may fold in such a way as to conceal the epitope. Therefore it was decided that the tag should be inserted in single or multiple copies at several different locations, and constructs should also be produced in which the Rifin was truncated with the tag at the new C-terminal end. Another concern was that the *rif* intron might not be correctly spliced out if the gene was expressed from a plasmid, rather than in its chromosomal context. Conversely, the intron could contain important regulatory sequences, as is the case for *var* genes (Deitsch et al, 2001). Therefore constructs should be generated with and without the intron.

The sequences of 35 known *rif* genes (from alignment by Kyes et al, 1999) were examined and the gene coding for Rifin2.7 was selected as having the most unique restriction sites in the polymorphic region (shown in fig 6.1b), and thus being most versatile for generating differently tagged proteins as outlined above. This gene is Pfb1005w, identified by the *P.falciparum* sequencing project and found on chromosome 2. It is not known how frequently this particular member of the family is expressed *in vivo.* In the 3D7 parasite line used in this study, antibodies raised against the conserved N-terminal half of Rifins immunoprecipitate a single band of about 44 kDa (*ibid,* Fernandez et al, 1999). This is larger than the predicted product of the Rifin2.7 gene,

and suggests that while 3D7 can express Rifins it does not normally express this member of the family.

6.3 Cloning of Rif2.7/c-myc and testing for antibody recognition

Genomic DNA was isolated from 3D7 parasites and used as a template for PCR with *Pfx* polymerase and 'Rif2.7' primers (table 6.1). Good PCR products were obtained by including 4mM $MgSO_4$ and 0.6mM dNTPs in the reaction. Purified DNA from these reactions was digested with *Bam*H I and *Sac* I and ligated in a 3:1 molar ratio with similarly digested pGem3zf(+) (Promega). Ligation products were transformed into *E. coli*, screened by PCR for the presence of the Rif2.7 fragment, and sequenced. Error-free clones were pooled to give the plasmid p3Rif.

Table 6.1: Oligonucleotides used in cloning Rifin2.7. All were used as primers for PCR as described in the body of the text, except for the MycTag oligos which are complementary. Restriction sites are highlighted.

Fragment name	Fragment size (bp)	Forward primer (_5)	Reverse primer (_3)	Restriction site(s), 5'	Restriction site(s), 3'
Rif2.7	1153	CGCA**GGATCC**AA TATATATGAAAAT GCATTACTCTG	AATT**GAGCTCGC TAGC**ATTAAATT GAATTTAATATC AAGAAACC	*Bam*H I	*Nhe* I, *Sac* I
RifCHOP	537	TCAT**CATATG**CG CATAGTGAGAAT AAACAATAC	CCTTTTGAG**CTG CAG**CAACCAAAG C	*Nde* I	*Pst* I
Rv	266	CCAATGAAACTT GTAAACATTGTTA ACTCTGC	GGCAATGACGGA GTAACCAATTGC ACTG	none	None
MycTag	58	TCCA**GGTACC**GA ACAGAAATTAAT ATCAGAAGAAGA AGATTT**AGCTGC AGATATC**GAT	ATC**GATATCTGC AGCT**AATCTTCTT CTGATATTAATTT CTGTTC**GGTACC** TGGA	*Asp*718 I, *Nla* IV, *Kpn* I	*Alu* I, *Pst* I, *Eco*R V

The DNA sequence coding for the c-myc epitope tag was obtained as a pair of complementary oligonucleotides ('MycTag') which were annealed to each other by mixing equal quantities, heating to 95°C for 10 minutes, and then allowing to cool. Initially, multiple restriction sites were added to the ends of these oligonucleotides to allow them to be inserted in any reading frame, but in the event it proved difficult to obtain efficient restriction digestion of such a small piece of DNA. Instead, slightly

shorter oligonucleotides were obtained and annealed to give ends ends which would have been produced by appropriate digestion of MycTag for each reading frame.

Aliquots of p3Rif were linearised with *Bsg* I, *Bsr*BR I, *Bsr*G I or *Sty* I and their ends dephosphorylated. These were mixed in a 1:10 molar ratio with appropriate phosphorylated MycTag fragments, ligated, and transformed into *E. coli.* Colonies were screened by PCR with the Rv primers (Table 6.1) which amplify across the hypervariable region of Rifin2.7, and where PCR product of increased size was seen, the clones were cultured and the DNA sequenced. The c-myc sequence was inserted successfully into the *Bsg* I site to give plasmid p3RifG, and into the *Bsr*BR I site both as a single copy and two in tandem, yielding p3RifB and p3RifB^2 respectively. Despite repeated attempts no successful insertion into the *Bsr*G I or *Sty* I sites was observed, so it was decided to proceed with the three successful constructs.

In order to check whether the internal epitope tags could be recognised by antibody, it was necessary to transcribe and translate these modified genes *in vitro.* p3RifG and p3RifB were first digested using *Bam*H I and *Sac* I, and the smaller tagged Rifin fragment gel purified and cloned into pGem7zf(+), placing it under the control of the T7 promoter. These constructs were called p7RifG and p7RifB. As it was unknown whether the intron would be correctly spliced *in vitro,* a PCR primer which bridged the intron (RifinCHOP_5) was obtained and used to amplify a segment of the gene with the intron deleted. This sequence was then cloned into p7RifG in place of the genomic sequence using the unique restriction sites *Nde* I and *Pst* I, to give p7RΔG.

p7RifG & p7RΔG were transcribed *in vitro* and then translated in a separate reaction, while p7RifB was used as a template for the TnT coupled transcription and translation system (Promega). Both methods use T7 RNA polymerase for transcription and rabbit reticulocyte lysate for translation, and ^{35}S-methionine was used to radiolabel the products in each case. These products were immunoprecipitated using the commercially available anti-c-myc monoclonal antibody 9E10 (mouse IgG1) followed by SDS-PAGE and autoradiography.

As shown in fig 6.2, in all cases a ladder of bands was specifically precipitated using 9E10 but not an isotype control. The primary band runs at around 36kDa, the predicted size of Rif2.7/c-myc. The series of bands below this are probably caused by translation initiation at some of the 11 internal methionines in the sequence. Some bands are

precipitated which are larger than expected, particularly at around 42 kDa, which may be due to read-through of the stop codon. Both p7RifG & p7RΔG produce the same bands, suggesting that the intron is correctly spliced out in the rabbit reticulocyte lysate system. The p7RifB results have a higher background, probably an artifact of the coupled transcription and translation, but importantly it can be seen that where C19 (a clone of identical sequence to p7RifB, but with the epitope tag reversed) was used, no bands are specifically precipitated by 9E10. This shows that the antibody is recognising the epitope tag and no other part of the protein.

It is likely that the *in vitro* translated form of Rif2.7/c-myc is not correctly folded, particularly as it is thought to be membrane associated, so these results do not guarantee that the c-myc tag will be accessible *in vivo*. Encouragingly, though, they demonstrate that the c-myc tag can in principle be recognised when placed internally in a sequence.

6.4 Design and construction of plasmid pRifCAT

On the strength of this an expression construct, pRifCAT, was designed for Rif2.7/c-myc (fig 6.3). This plasmid also contained a CAT expression cassette, to give a positive control for transfection. The construct was based on pHC1-CAT (Crabb & Cowman, 1997), but restriction mapping and comparison of the reported primer sequences with the relevant genomic databases revealed several restriction sites and sequences additional to those described, making pHC1-CAT unsuitable for further cloning. Therefore the various sequence elements were newly amplified by PCR, using the primers shown in Table 6.2. b-CAT and PcDT.5' were amplified from pHC1-CAT, and PfHRP2.3' and PfHSP86.3' from 3D7 genomic DNA. PfCAM.5' is extremely AT-rich and proved refractory to PCR from genomic material, so it was instead amplified from pfCAM 1561, a plasmid produced and kindly supplied by K. Robson (Robson & Jennings, 1991). Due to the >90% AT content in this element, PCR errors were expected, so 10 *Pfx* PCR reactions of this region were independently sub-cloned into pCR4-Blunt TOPO and sequenced. Despite the high-fidelity polymerase, deviations were found in all clones from the database sequence (GenBank accession no. M59770). Two long poly-(dT) stretches and a poly(dA-dT) repeat were seen to vary by up to 3 bases in length from the reported sequence. Robson (1993) reported similar polymorphisms with no effect on function in the intron of the *Pfcam* gene, so they are probably unimportant. Two single base transversions were observed (one A-T and one T-A), but both were present in all

clones and therefore were taken to be mutations in the PfCAM1561 plasmid. The clone which had sequence closest to the database sequence was used.

Table 6.2 – Oligonucleotides used in the comstruction of pRifCAT. Restriction sites are highlighted. All were used as primers for PCR.

Fragment name	Fragment size (bp)	Forward primer (_5)	Reverse primer (_3)	Restriction site, 5'	Restriction site, 3'
PfHRP2.3'	588	CTCAG**GCTAGC**C TTATTTAATAATA GATTA	CTCAG**GAATTC**T GTATAAATATGTT CTTA	*Nhe* I	*Eco*R I
b-CAT	736	CAGC**GTCGAC**GA GATTTTCAGGAGC TAAG	GAGT**AGATCT**AA GGGCACCAATAA CTGCC	*Sal* I	*Bgl* II
PfCAM.5'	1039	ATTA**CTCGAG**AT CTCTGAGCTTCTT CTTTGTTAACC	ATTA**GTCGAC**GT CTGCCAGTGATA TATTTCTATTAGG	*Xho* I	*Sal* I
PcDT.5'	612	GATA**CTCGAG**TT AGCTAATTCGCTT CTAAG	TCAA**GGATCC**TT GTAAGTTTTAGGT GTGTG	*Xho* I	*Bam*H I
PfHSP86.3'	850	CCGT**AGATCT**TT ATATAATATATTT ATGTACTCAC	TACCG**AAGCTT**T ATTTGATGAATTA ACTACAC	*Bgl* II	*Hin*d III

These DNA fragments were assembled as shown in fig 6.4. The HSP86 3' UTR and b-CAT ORF PCR products, along with pGEM 3zf(+), were digested using the appropriate enzymes, the DNA purified, and mixed together in the approximate molar ratio 1:3:3. One aliquot of the mixture was phosphorylated by T4 polynucleotide kinase and the DNA purified once more. This and an untreated aliquot were both then ligated. DNA from both triple ligations produced colonies that contained the plasmid p1-2, as confirmed by PCR screening and restriction mapping.

An attempt to do a similar triple ligation with the two 5' UTRs was unsuccessful, perhaps confounded by the AT-rich PfCAM.5' segment. Therefore the PcDT.5' PCR product was digested with *Bam*H I, before ligating into pGEM 3zf(+) which had been previously digested with *Xba* I, blunted, and then digested with *Bam*H I. The resulting plasmid, pCDT, was readily inserted with the CAM 5' UTR to produce p3-4. The non-backbone sections of p1-2 and p3-4 were each sequenced to check that all PCR products had been inserted correctly and without mutations. The HRP2 3' region PCR product was digested with *Eco*R I and *Nhe* I and ligated into similarly digested p3RifG. This plasmid, p5-6, was also sequenced. The two 5'regions were cut out of p3-4 and cloned into p1-2 to generate p1-4, and the whole non-backbone sequence of p1-4 cut out using

Hind III and *Bam*H I and cloned into p5-6 to generate pRifGCAT. The whole of this plasmid (except for backbone regions) was sequenced once more. Versions with the epitope tag in a different site and/or lacking the *rifin* intron were produced by cutting out the *rifin* sequence using *Bam*H I & *Nhe* I and replacing it with one of the alternatives, using the following nomenclature:

pRifGCAT – pRifCAT with the c-myc sequence inserted into the Rif2.7 *Bsg* I site

pRΔGCAT – as pRifGCAT with the intron deleted

pRΔB^{2}CAT – pRifCAT with two copies of the c-myc sequence inserted in tandem into the *Bsr*BR I site, and the intron deleted

6.5 pRifCAT expresses the CAT gene normally

To ensure that pRifCAT plasmids were able to express genes transiently, pRifGCAT was electroporated into 3D7 parasites following the optimized protocol described in Chapter 3. As shown in fig 6.5, the transfectants reproducibly developed transient CAT activity, being detectable 24 hours post-transfection, peaking at around 48 hours, and tailing off during the second life cycle of the transfected parasites, similar to the pattern observed for pHC1-CAT. The peak activity was 80% of that of pHC1-CAT. All the other plasmids constructed also showed CAT activity comparable to that of pHC1-CAT. This demonstrates that pRifCAT can be electroporated into *P. falciparum* with efficiency comparable to published expression constructs, and that the sequence differences in the untranslated regions of this plasmid compared to pHC1-CAT do not significantly affect gene expression. In all subsequent experiments using pRifCAT, a sample was kept for analysis by CAT assay as a positive control for transfection.

6.6 Expression of Rif2.7/c-myc is not detectable from pRifCAT

The degree and location of expression of the Rifin from parasites transfected with pRifGCAT was investigated by immunofluorescence. The fixing process of the samples in standard immunofluorescence assays permeabilizes the membranes, so unfixed, live parasites were also stained to find out whether the protein was exposed to the external medium. Anti c-myc monoclonal antibody 9E10 was incubated with the parasites at dilutions from 1:100 to 1:8000, followed by FITC-conjugated anti mouse IgG at a dilution of 1:25 or 1:50. In no case was any convincing immunofluorescence observed,

although a control antibody against MSP-1 stained the parasite as expected (data not shown). This failure suggests that either the epitope tag is obscured in the folded Rifin, or else that it is expressed at too low a level to be detected by this method.

The antibody 9E10 was also used for panning experiments, in which transfected cells were incubated with antibody and protein A sepharose 6MB beads, which rapidly settle out under gravity. The unbound fraction was removed, the beads washed with RPMI w/o, and the fractions assayed for CAT activity. As shown in fig 6.6, no CAT activity partitioned with the sepharose beads, again suggesting that insufficient Rifin is expressed on the erythrocyte surface in such a way that the c-myc tag can interact with the antibody.

This raised the question of whether the tagged Rifin was being expressed at all, so transcription of the plasmid-encoded gene was examined. RNA was isolated from parasite cultures at a range of time points post-transfection, and used to make Northern blots (fig 6.7). Probes generated against the hypervariable region of RifB showed strongly labelled low molecular weight material at a range of time points. These bands disappeared on treatment with DNase but not RNase, suggesting that they are caused by degraded plasmid material, rather than transcription from the plasmid; although if so it is puzzling that they do not appear equally at all time points. Possibly the RNA preparation process is not able to eliminate all plasmid DNA, but the amount which gets through is affected by some other factor, such as the abundance of RNA.

The same samples probed against MSP-7, known to have a transcript of about 2.5 kilobases that is expressed in late erythrocytic stages (Justin Pachebat, awaiting publication), showed that the RNA was intact. Ethidium bromide stained gels (not shown) indicated that consistent quantities of RNA were loaded in each lane, that the DNase-treated samples were slightly degraded, and that there was essentially no RNA left in the RNase-treated samples.

Various approaches were tried to eliminate the DNA contamination, but none did so without also resulting in degradation of the RNA. The same problem confounded attempts to detect transcription by RT-PCR; RT negative controls always gave a product at the end of the PCR step. These technical problems illustrate the difficulty of detecting

a relatively scarce mRNA molecule when such large quantities of plasmid are required for transfection (50μg per 200μl packed cells).

6.7 Conclusions

An attempt was made to generate transfected parasites labelled with an epitope on the parasitized erythrocyte surface. In theory, this could be a valuable research tool, although thus far technical difficulties have prevented it from being realised. Neither transcription nor surface expression of Rif2.7/c-myc could be detected, although the possibility of it being expressed at a low level cannot be ruled out. It should be noted that in previous studies using 3D7, *rif* mRNA was undetectable by Northern blot even though a Rifin was expressed on the erythrocyte surface (Kyes et al, 1999). Similarly, mRNA from the CAT gene could not be detected by this method even from parasites expressing readily detectable levels of CAT (data not shown). Antibody binding methods are also much less sensitive way for detecting protein than enzymatic assays, in which a single protein may convert thousands of substrate molecules.

In part these difficulties with detection stem from the very low efficiency of electroporation; if only a tiny fraction of parasites are transfected, we would not expect to see many positives by immunofluorescence, and the amount of specific mRNA produced will be swamped by the huge amount of plasmid required to obtain any transfection at all. Ultimately, however, it is our lack of understanding of gene regulation in *P. falciparum* which may be the confounding factor. The 5' and 3' untranslated regions used in pRifCAT were chosen owing to their having been used successfully in other constructs such as pHC1-CAT. Most transfection studies in this organism have, by this logic, used the same few proven non-coding control sequences, but relatively little is known about the precise timing and abundance of expression they induce in adjacent open reading frames. For example, the PcDT.5' promoter used in this study, when placed in front of the Dihydrofolate reductase-thymidylate synthase gene from *Toxoplasma gondii,* generates enough protein to make parasites resistant pyrimethamine (Crabb et al, 1997). This is an enzymatic activity which could be conferred by a relatively low abundance of protein; an epitope tagged protein may need to be present at a much higher level to be detectable. Furthermore, Rifins are normally expressed in the early trophozoite stages, while the stage specificity (if any) of the PcDT promoter is not clearly defined. Expression at the wrong time may prevent correct

processing and transport of the protein. These considerations highlight the importance of studies such as those in Chapter 5 which examine gene regulation in more depth, with the potential to match the ideal promoter for expression of each gene of interest.

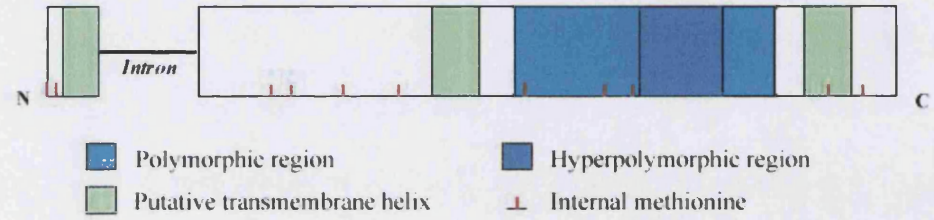
A variety of molecular events may prevent intact, tagged Rifin from being exported to the erythrocyte surface. As noted above, when removed from its chromosomal context it may not be transcribed at the right time or in sufficient abundance. A gene silencing system of some sort, analogous to that which controls the *var* family may be operating. If it is transcribed, the mRNA may not be unstable or not correctly processed, and if it is translated it may not be being transported to the right compartment. Finding out where in this process the expression of the Rifin has failed would be of great interest, but is difficult in a transient system owing to the high background of untransfected parasites. In a stable transfectant culture, however, all the parasites would contain the expression construct; the generation of such a culture is explored in the next chapter.

Fig 6.1: **Characteristics of Rifin proteins and *rif* genes.**

A. General features of Rifin protein primary structure. The small first exon codes for a trafficking signal peptide, which is proteolytically removed in the mature protein. The regions not highlighted as polymorphic are highly conserved across the family. The specific methionine residues indicated are those from Rifin2.7.

B. Schematic of the genomic DNA sequence of open reading frame Pfb1005w from chromosome 2, which codes for Rifin2.7. Unique restriction sites and PCR primers used for cloning and analysis are indicated.

A.


Intron
N
C
Polymorphic region
Hyperpolymorphic region
Putative transmembrane helix
Internal methionine

B.

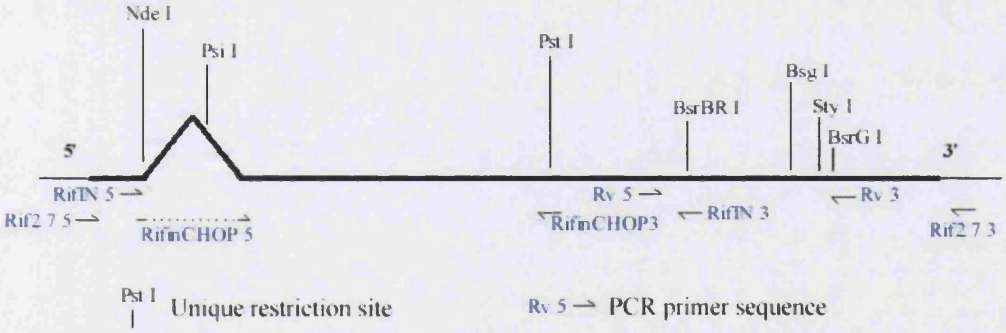

Nde I
Pst I
Pst I
Bsg I
BsrBR I
Sty I
BsrG I
5'
3'
RifIN 5
Rif2 7 5
RifinCHOP 5
Rv 5
RifinCHOP3
RifIN 3
Rv 3
Rif2 7 3
Pst I
Unique restriction site
Rv 5 PCR primer sequence

Fig 6.2: **Immunoprecipitation of c-myc tagged Rifin2.7 produced *in vitro*.** The T7 RNA polymerase system and rabbit reticulocyte lysate were used to *in vitro* transcribe and translate (incorporating ^{35}S methionine) Rifin2.7 with the c-myc epitope tag sequence inserted either in the *Bsg* I or the *Bsr*BR I site. These products were diluted 1:5 (10µl into 50µl PBS) and immunoprecipitated as described using the monoclonal anti c-myc antibody 9E10 (Sigma) and Protein G sepharose. The 'Control' lanes used the monoclonal 89.1, an antibody of the same isotype as 9E10 (mouse IgG1) raised against MSP-1. The plasmids used as templates for transcription were:

A p7RifG

B p7RΔG (identical to p7RifG, but with the intron removed)

C C19, a clone with the epitope tag inserted into the *Bsr*BR I site, but in the reverse orientation

D p7RifB

SN = supernatants removed from protein G sepharose beads after incubation with *in vitro* translation products and 9E10, prior to washing of beads (total volume 40µl). Ten microlitres of each were loaded per lane after mixing 1:1 with SDS-PAGE loading dye and heating to 95°C for 5 minutes.

IP = Material remaining bound to beads after all washes. Beads were aspirated dry (total volume c. 35µl) and mixed 1:1 with SDS-PAGE loading dye and heated to 95°C for 5 minutes. Ten microlitres of beads were loaded per lane (not including loading dye).

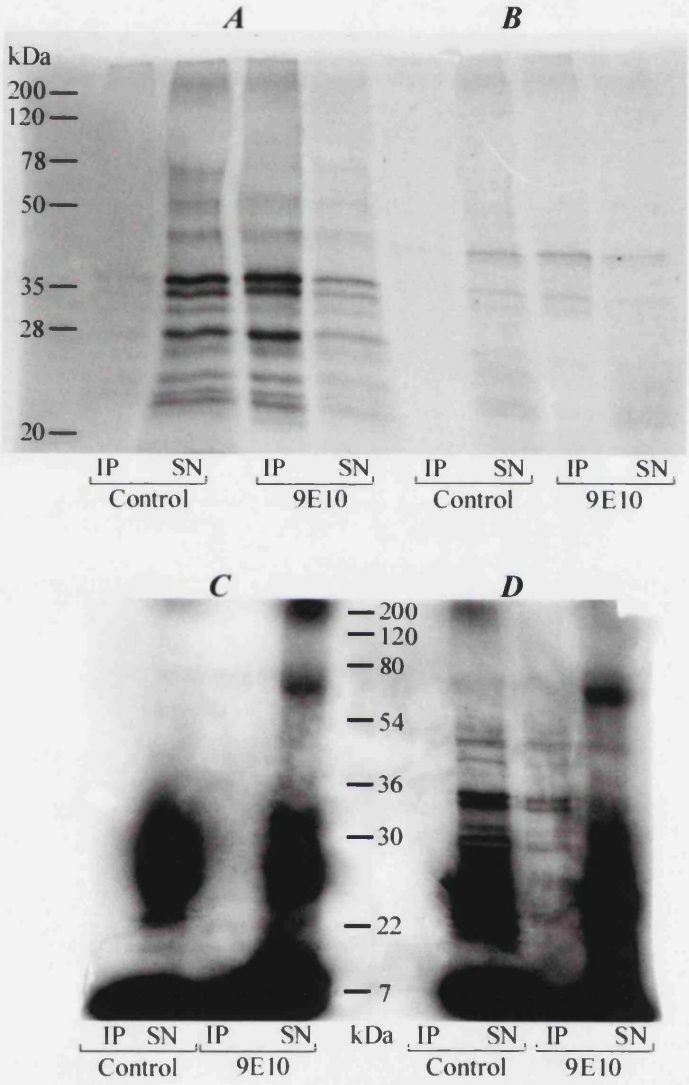

A
B
kDa
200
120
78
50
35
28
20
IP SN IP SN IP SN IP SN
Control 9E10 Control 9E10
C
D
200
120
80
54
36
30
22
7
kDa
IP SN IP SN IP SN IP SN
Control 9E10 Control 9E10

Fig 6.3: **Design for plasmid pRifCAT.** Different sequence elements, unique restriction sites used in plasmid construction, directions of transcription and the approximate location of the epitope tag sequence are all labelled. Numerals indicate the number of base pairs around the plasmid, counting clockwise from the top. Total size of plasmid = 7896bp.

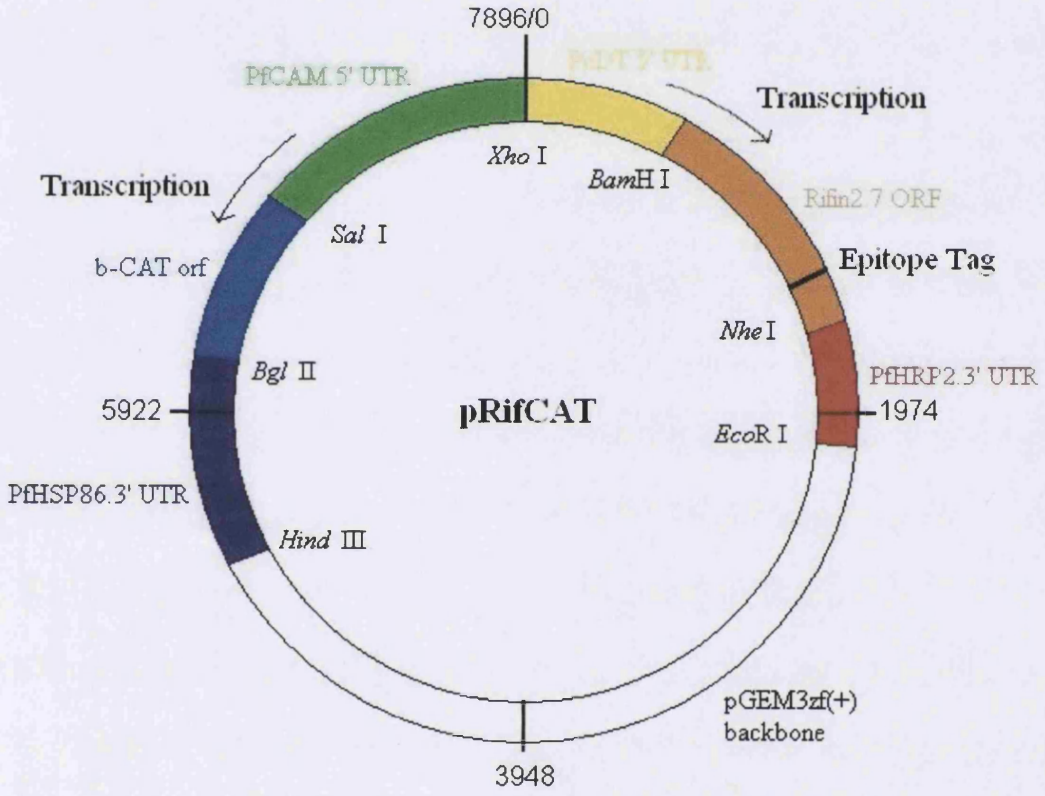

7896/0
PfCAM 5' UTR
Transcription
Transcription
Xho I
BamH I
Rifin2.7 ORF
Epitope Tag
Sal I
b-CAT orf
Nhe I
PfHRP2.3' UTR
Bgl II
5922
pRifCAT
1974
EcoR I
PfHSP86.3' UTR
Hind III
pGEM3zf(+)
backbone
3948

Fig 6.4: **Schematic of the construction of pRifCAT.** Open arrows indicate PCR reactions using the templates shown. Line arrows represent a restriction digestion, followed by ligation with another restriction fragment and transformation into TOP10 chemically competent cells. Restriction enzyme sites are abbreviated as follows: B = *Bam*H I, Bg = *Bgl* II, E = *Eco*R I, H = *Hin*d III, N = *Nhe* I, S = *Sal* I, Sc = *Sac* I, X = *Xho* I, Z = *Bsr*BR I or *Bsg* I.

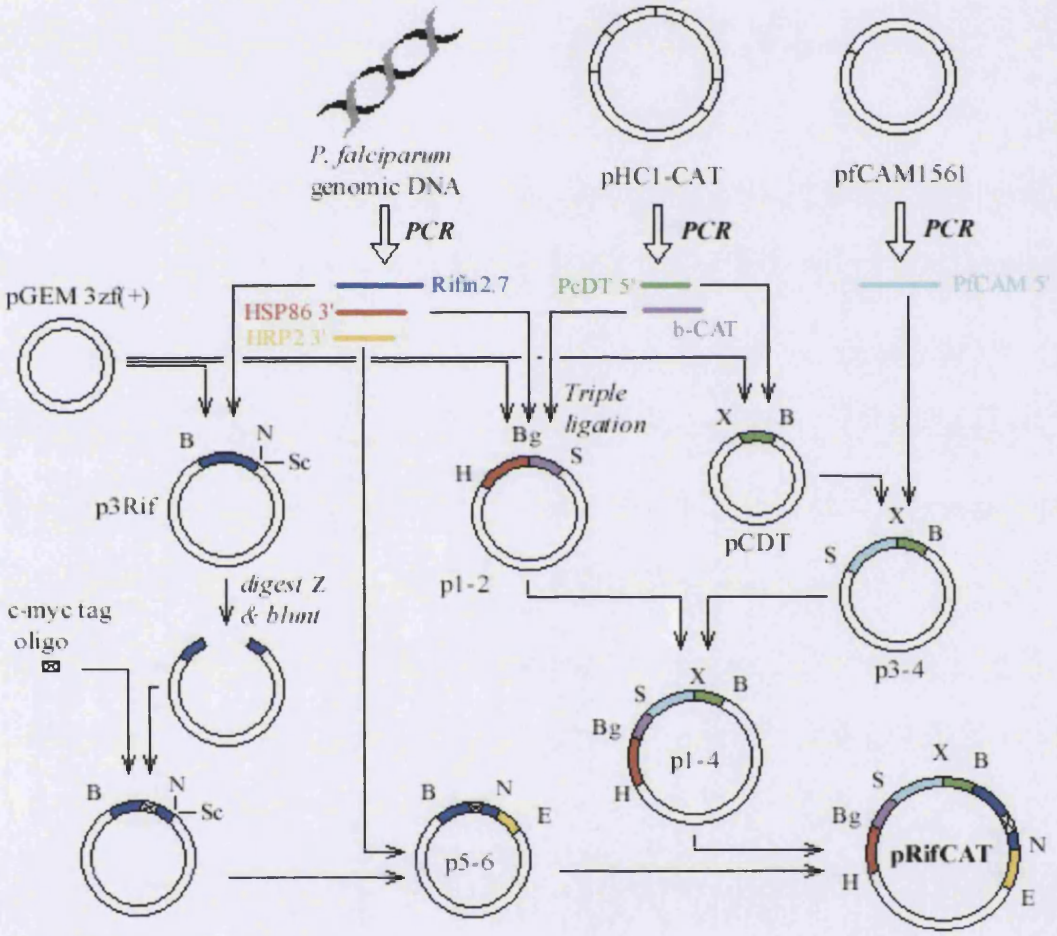

P. falciparum
genomic DNA
pHC1-CAT
pfCAM1561
PCR
PCR
PCR
pGEM 3zf(+)
Rifin2.7
HSP86 3'
HRP2 3'
PcDT 5'
b-CAT
PfCAM 5'
Triple
ligation
B
N
Sc
p3Rif
Bg
S
H
p1-2
X
B
pCDT
X
B
S
p3-4
digest Z
& blunt
c-myc tag
oligo
B
N
Sc
B
N
E
p5-6
S
X
B
Bg
H
p1-4
X
B
S
Bg
H
pRifCAT
N
E

Fig 6.5: **CAT activity of parasites transfected with pRifCAT.** Parasitized erythrocytes were electroporated with 50μg pRifGCAT and incubated for 24, 44, 52 or 72 hours prior to harvesting and assay for CAT activity. Activities are expressed as a percentage of the peak activity obtained from parasitized erythrocytes electroporated with pHC1-CAT, which were also most active at 44hrs post-transfection. Error bars indicate the standard deviation of duplicates.

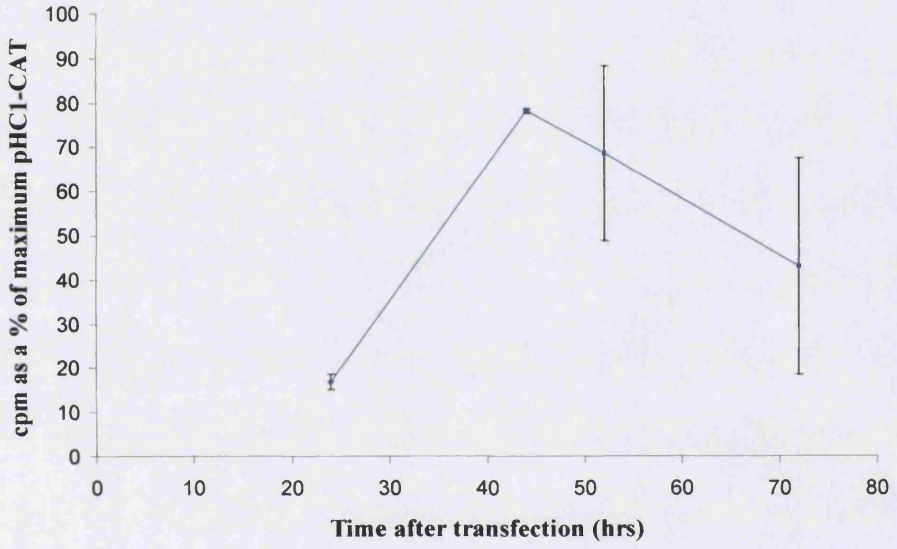

cpm as a % of maximum pHC1-CAT
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
Time after transfection (hrs)

Fig 6.6: **Immunoselection of parasites transfected with pRifCAT.** Aliquots of 200µl parasitized erythrocytes containing 10% rings were electroporated with 50µg pHC1-CAT or pRifGCAT DNA. At the time points indicated (44hr in the case of the pHC1-CAT control) half of each culture was harvested, while the other half was washed three times with RPMI without Albumax (to remove any serum antibodies in the culture) before incubating for 15 minutes at 37°C with a 1:100 dilution of anti c-myc antibody 9E10 in RPMI, adding to 100µl protein A sepharose 6MB beads, and incubating for a further 15 minutes at room temperature with constant slow rocking. The sepharose beads were then allowed to settle out under gravity, the supernatant (containing unbound parasites) removed, and the beads washed twice with RPMI, the wash then being added to the supernatant. Beads and supernatant were harvested and assayed independently for CAT activity (Bound and Unbound fraction, respectively).

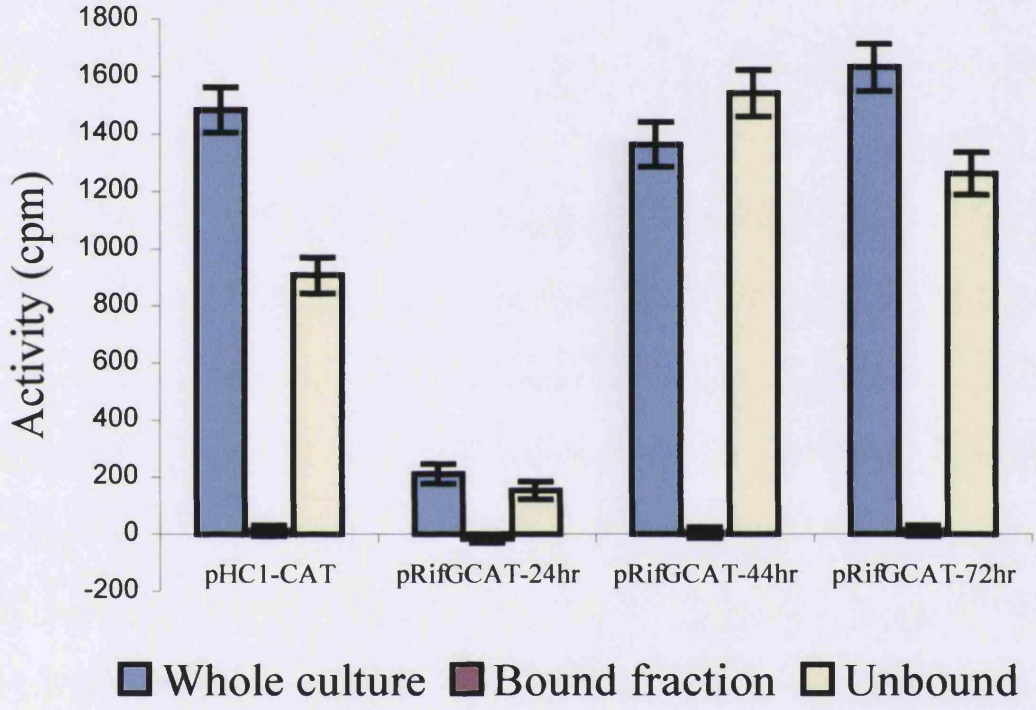

Activity (cpm)
1800
1600
1400
1200
1000
800
600
400
200
0
-200
pHC1-CAT
pRifGCAT-24hr
pRifGCAT-44hr
pRifGCAT-72hr
Whole culture
Bound fraction
Unbound

Fig 6.7: **Northern blots of RNA isolated from parasites transfected with pRifCAT.** The concentrations of aliquots of RNA purified from transfected parasites at various times were measured spectrophotometrically, and equal amounts of each were treated with RNase, DNase or nothing, and then separated on a 1% agarose gel. Two micrograms of total RNA were loaded per lane.

A – Untransfected parasites

B – 0 hours post-transfection

C – 21 hours post-transfection

D – 25 hours post-transfection

E – 28 hours post-transfection

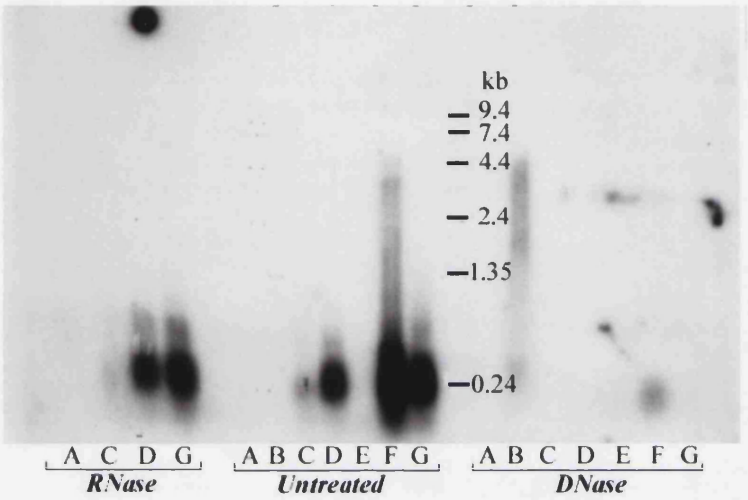
F – 45 hours post-transfection

G – Plasmid added to culture, but not electroporated; 25 hours after adding plasmid.

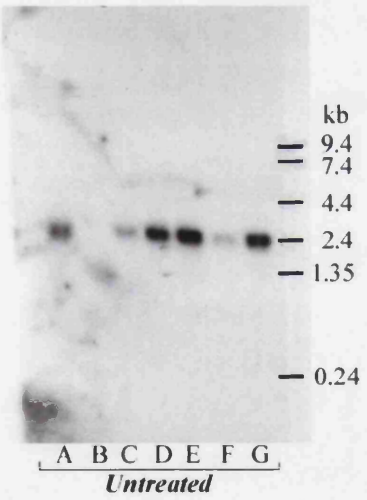
A. These samples were blotted and hybridised to a probe made using the Rv PCR product (cf. *Table 3.2*) as template.

B. Two primers, GCCGGATCCAAAGTGAAACAGATACTCAATCTAAAAATG (forward) and GCCGAATTCATTGTGTTTAGTAAATTAAATGAATATTCTAAG (reverse) were used to make a PCR product which was the template for a probe against MSP-7, which was hybridised to an identically-loaded blot of the same samples.

A.


kb
9.4
7.4
4.4
2.4
1.35
0.24
A C D G
RNase
A B C D E F G
Untreated
A B C D E F G
DNase

B.


kb
9.4
7.4
4.4
2.4
1.35
0.24
A B C D E F G
Untreated

7. *Generating Stable Transfectants in* P. falciparum

7.1 Introduction

In all previous chapters, transfection technology has been used to obtain transient transfectants, in which the change in genotype (and hence phenotype) of the parasite is temporary, lasting only as long as the introduced plasmid persists in the parasite. It has been shown that plasmids electroporated into *P. falciparum* can be replicated by the host (Wu et al, 1996), but it is thought that they do not segregate equally between the merozoite progeny of each schizont. They also confer an energy cost on the minority of parasites which carry them, with the result that the plasmid is selected out of the culture after only 2 or 3 generations.

In genetic studies on many other organisms, plasmids can be maintained if they confer a selective advantage on their host cells – such as a drug resistance gene. Wu et al (1996) and Crabb et al (1997b) have successfully applied this approach to *P. falciparum,* by including the *Toxoplasma gondii* Dihydrofolate Reductase-Thymidylate Synthase gene (TgDHFR-TS), which confers resistance to pyrimethamine. One problem with this system is that parasites can spontaneously gain resistance to pyrimethamine by mutation in the endogenous DHFR gene, creating false positives. Therefore alternative selection markers have since been described, most notably the human Dihydrofolate Reductase gene (hDHFR) which gives resistance to a drug labelled WR99210 (Fidock and Wellems, 1997). No resistance to this drug by mutation in *P. falciparum* has ever been reported.

Establishing a culture of drug-resistant, plasmid-bearing parasites takes a long time, owing to the very small proportion of parasites that take up plasmid during electroporation. On applying drug pressure, the vast majority of parasites die, leaving a parasitaemia which is too low to detect by thin blood smear. Over a period of several weeks, if any of the original transfectants survived, the parasitaemia may grow to detectable levels again. This protracted culture period, during which it is impossible to tell whether the parasites are growing healthily or not, makes establishing stable transfectants a fairly hit or miss process.

As a result, despite several published examples of stable transfection being used to answer biological questions (e.g. Crabb et al 1997a, Burghaus and Lingelbach 2001) in the majority of labs the technique is still neither routine nor reliable. In particular, at the start of this project, no stable transfectant had ever been obtained at the NIMR, despite repeated attempts.

Other recent reports have uncovered some potentially useful aspects of plasmid behaviour in *P .falciparum.* Kadekoppala et al (2000) showed that if two plasmids carrying homologous sequences are co-transfected into parasites, they frequently recombine with each other to form a larger episome. This is valuable as it allows different activities, such as drug resistance and expression of a gene of interest, to be encoded by separate plasmids and yet be selected for together. Given the instability of large plasmids in *E. coli,* which is exacerbated by AT-rich *P. falciparum* sequences, this could simplify cloning considerably.

As described in section 1.4.5, O'Donnell et al (2002) found that plasmids containing a number of copies of the 21bp sequence repeat, rep20, associate physically with chromosome ends and segregate more efficiently, meaning that they persist in culture for more generations than other plasmids. It is therefore expected that including rep20 sequence in a plasmid for stable transfection should reduce the culture time necessary to establish a drug resistant line, as the plasmid would be passed on more efficiently to the next life-cycle at each schizogony.

7.2 Generating plasmids for stable transfection

Two plasmids were used as a basis for stable transfection constructs: pHC1-CAT and pHHT-TK (fig 7.1a) both kindly supplied by A. Cowman. The former of these contains a pyrimethamine resistance cassette and a CAT expression cassette, arranged in head-to-head orientation to enhance expression, as discussed in Chapter 5. The latter has cassettes expressing genes for both WR99210 resistance and ganciclovir sensitivity (conferred by a thymidine kinase gene), allowing positive/negative selection. This has been used for targeted disruption of a non-essential *P. falciparum* gene (Duraisingh et al, 2001)

A segment of rep20 repeats, 725bp in length, was amplified from 3D7 DNA by PCR using the primers shown in table 7.1. These primers were carefully designed and their sequences searched for in the genome database using BLAST, to ensure that only the selected section of rep20 repeats would be amplified despite their abundance elsewhere in the genome. Both primers matched 3 or 4 different genomic loci, but only the chosen repeats, near the 5' end of chromosome 2, matched both primers exactly. Specific products were obtained using AmpliTaq polymerase with 2.5mM $MgCl_2$ in the reaction mix and an annealing temperature of 65°C.

Table 7.1: PCR primers used for cloning rep20 repeats from one end of chromosome 2, and checking for presence of plasmid in transfectants by PCR.

Fragment name	Fragment size (bp)	Forward primer (_5)	Reverse primer (_3)	Restriction site, 5'	Restriction site, 3'
Rep20	725	CTTG**AAGCTT**AAC AACATCAGTAAGA ACATGAC	CTGT**AAGCTT**AA CCTAAGTTAGTTA ACTTGAGGG	*Hin*d III	*Hin*d III
Cam625-hDHFR	c. 1200	ATTA**CTCGAG**GTT TATATGTGATTAAT TTTATATATTATC	AATTTC**AAGCTT** AATCATTCTTCTC ATATACTTC	*Xho* I	*Hin*d III

The resulting PCR products were subcloned into pCR4 and sequenced. The correct rep20 sequence was then excised from pCR4 using *Hin*d III, while pHC1-CAT was also digested with this enzyme followed by gel purification of the larger fragment. This was ligated with the rep20 restriction fragment to yield pTgDTrep20 (fig 7.1b). As this construct has the PcDT.5', HRP2.3' and backbone sequences in common with the pRifCAT constructs described in chapter 6, it should recombine with those plasmids on co-transfection.

7.3 Stable transfectants were obtained using pHHT-TK but not pTgDTrep20

A protocol for obtaining stable transfectants was designed with valuable advice from B. Crabb, described in detail in section 2.3.5. Briefly, 3D7 parasites were cultured for several days at <5% parasitaemia and synchronized to give a parasitaemia of 1% ring stages on the day of transfection. 200µl aliquots of packed parasitized erythrocytes were electroporated as described with 50µg each of pTgDTrep20 and one of the pRifCAT constructs, or with 100µg of pHHT-TK, and cultured at 2% haematocrit in a 50/50 mix of RPMI Alb and RPMI

HS media. Two days after electroporation, pyrimethamine was added to the cultures containing pTgDTrep20 at a final concentration of 2μM (which was reduced to 0.5μM after a further two days), and WR99210 to the pHHT-TK cultures at 10nM. Daily for seven days after electroporation and on alternate days thereafter, the cultures were spun down, and the supernatant aspirated off and replaced with fresh medium and drug. Every six days a quarter of the culture was removed and replaced with fresh erythrocytes diluted in medium.

The state of the culture was monitored regularly using thin blood smears. After electroporation, the parasites developed through one life cycle as normal until the addition of drug. At this point the parasitaemia dropped rapidly in all flasks to undetectable levels by 5 days post-electroporation, and remained so in subsequent smears. Despite repeated attempts, in all cultures transfected with pTgDTRep20 no parasites were seen by 40 days post-electroporation, at which point they were discarded.

Happily, however, parasites were observed in a flask transfected with pHHT-TK 33 days after electroporation, at about 0.1% parasitaemia. These parasites grew and developed at a similar rate in the presence of WR99210 as untransfected parasites do in its absence, and grew equally well when transferred to normal RPMI Alb medium. To verify that the drug resistance phenotype was due to the presence of pHHT-TK and not any mutation of the parasites, DNA was isolated from them and they were tested for ganciclovir sensitivity.

A culture of the transfectants at about 8% parasitaemia was divided between 3 flasks and cultured with 10nm WR99210 and 0, 1 or 10μM ganciclovir. As shown in fig 7.2, the flasks treated with ganciclovir began to show a drop in parasitaemia (particularly of early parasite stages) within 48 hours of commencing drug pressure. After 5 days of drug pressure, no parasites were visible in either of the ganciclovir treated flasks, and this drug was removed. After a further two weeks of culture in WR99210 alone, parasites reappeared in the 1μM ganciclovir but not the 10μM ganciclovir treated flask. In control flasks, transfected parasites cultured with just WR99210 and untransfected parasites cultured with just 1μM ganciclovir showed no ill effects.

Plasmid rescue was carried out by transforming TOP10 cells with an aliquot of DNA isolated from the WR99210 resistant parasites and selecting them on plates containing Ampicillin. Four of the resulting colonies were grown up and plasmid DNA purified from the resulting minicultures. All four plasmids obtained by this means were digested with combinations of *Bam*H I, *Eco*R I, and *Xho* I, and found to give an identical restriction pattern to pHHT-TK.

DNA from the parasites was also used as a template for PCR using primers for part way through the PfCAM.5' region and the 3' end of the human dihydrofolate reductase gene (table 7.1); the former is in the *P. falciparum* genome, while the latter could be present if there were any white blood cells contaminating the red blood cells used in culture, but the combination would only be present in pHHT-TK. As expected, PCR products were obtained from pHHT-TK grown in *E.coli* and from transfected parasite DNA, but not untransfected 3D7 parasite DNA (fig 7.3a). To establish whether the plasmid had integrated into the genome or was being maintained episomally, Southern blots were carried out. DNA was digested with *Eco*R I and either *Not* I or *Sca* I prior to electrophoresis, blotting and hybridization. As shown in fig 7.3b, a probe generated from the pGem3 vector backbone sequence hybridized to identical bands from pHHT-TK and transfected parasites, but did not hybridize to DNA from untransfected 3D7. These results indicate that pHHT-TK is present in the transfectants as an episome, and has not integrated into the genome.

7.4 Transfecting with Rif2.7 containing constructs

Attempts to express Rifin constructs stably by co-transfection with pTgDTRep20 in the previous section were unsuccessful, so a more conventional approach was adopted. The thymidine kinase cassette was removed from pHHT-TK by digestion with *Not* I, blunting of the resulting sticky end, and digestion with *Bgl* II. Rif2.7/c-myc expression cassettes from pRifB^{2}CAT, pRΔGCAT and pRifB^{2}ΔCAT were excised by digestion with *Eco*R I, blunting, and digestion with *Bgl* II. This allowed direction specific ligation of the Rifin cassettes to the larger pHHT-TK fragment to give the pHTKRif constructs shown in fig 7.1b. A similar cloning strategy was used to replace the thymidine kinase open reading frame and PbDT.3'

untranslated region with the CAT open reading frame and HSP86.3' untranslated region, this time using *Not* I with *Xho* I or *Sal* I (which give complementary sticky ends) on pHHT-TK and pRifGCAT respectively. The resulting construct was pHTKCAT (fig 7.1b).

These plasmids were electroporated into *P. falciparum* and cultured as described for pHHT-TK. However, owing to time constraints, they have not yet been successfully used for transfection – although they have twice been used in electroporations, due to technical problems neither set of cultures survived for more than three weeks post-electroporation.

7.5 Conclusions

A number of different constructs were tested for their utility in stable transfection, and in particular for stable expression of Rif2.7/c-myc. It was hoped that by including a Rep20 element in pTgDTRep20, it might prove easier to transfect with this construct, but in fact no transfectants were obtained. This may be because the deletion of the CAT cassette from pHC1-CAT involves removing the PfCAM.5' region, a possible enhancer of the adjacent PcDT.5' promoter which controls expression of the drug resistance gene. If the expression dropped too low, it could be insufficient to resist the drug. This might also occur if the rep20 sequence, included adjacent to the PcDT.5' region, associates with chromosome ends in such a way as to block the transcription machinery from reaching the plasmid. The silencing of genes inserted near the telomeres of chromosomes is a well known phenomenon, and could extend to genes on episomes associated with the chromosome ends; however, no such effect was observed for the rep20 plasmids tested in section 5.7.

As reported by Duraisingh et al (2001), stable transfectants were generated using pHHT-TK. Although the plasmid does contain some *P. falciparum* sequences, and therefore could theoretically recombine with the genome, it was found to be maintained episomally. This is a desirable characteristic for a construct which might be used in gene targeting experiments.

The efficiency of transfection is clearly very low; drug resistant parasites were not observed until 16 generations after the start of drug pressure. We might expect the proliferation of

drug resistant parasites to be slower than usual, as only progeny which receive the plasmid would survive, but even if we assume that on average only 2 merozoites survive per schizont 16 generations is a huge amplification. Taking the number of drug resistant parasites present when they were first observed (0.1% parasitaemia in 200µl packed cells = approximately 2 x 10^6 parasitized erythrocytes) and dividing by 2^{16} gives an estimate of just 30 viable resistant parasites at the start of drug treatment! It is to be expected that some of the parasites which take up plasmid might prove unable to re-invade after the damage done to them during electroporation, but this figure is still extremely low. Furthermore, once drug resistant parasites had grown up they seemed to proliferate at a rate close to normal for 3D7 *in vitro*. Perhaps as well selecting for parasites which took up the plasmid in the first place, there is also a period of selection for parasites which are better at maintaining episomes; during this time the number of viable parasites would increase very little with each generation, but at its end they would begin to proliferate more rapidly. This would help to explain the protracted culture period and the exquisite sensitivity of the transfectants to adverse conditions, particularly soon after transfection.

It is this protracted culture period and sensitivity which makes stable transfection a demanding technique in *P. falciparum*, and explains why many laboratories have found it difficult to do reliably, if at all. In this context it is a significant step to obtain a stable transfectant for the first time, as local conditions frequently have been observed to affect the success of the method. The next step is to express the Rifin2.7/c-myc stably, a goal which should in theory be readily attainable. As this project shows, transferring theory to practice is not trivial, and the long culture time requires prolonged close attention on the part of the worker. For this reason, development of strategies to reduce the culture time should be a high priority.

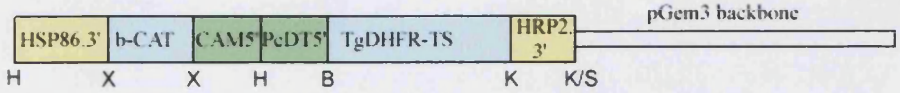
Fig 7.1: **Schematics of plasmids constructed for stable transfection.** Constructs are shown linearized. Specific sequence elements and restriction sites are indicated. Colour coding: green = promoter regions, cyan = open reading frames, yellow = 3' untranslated region. Restriction sites: B = *Bam*H I, Bg = *Bgl* II, E = *Eco*R I, H = *Hin*d III, Hp = *Hpa* I, K = *Kpn* I, Nh = *Nhe* I, N = *Not* I, S = *Sac* I, S = *Sac* II, X = *Xho* I.

A. Transfection constructs kindly provided by Alan Cowman. 'TgDHFR-TS' is the *T. gondii* dihydrofolate reductase-thymidylate synthase ORF. 'hDHFR' is the human dihydrofolate reductase ORF. 'TK' is the thymidine kinase ORF of *Herpes simplex* virus-2.

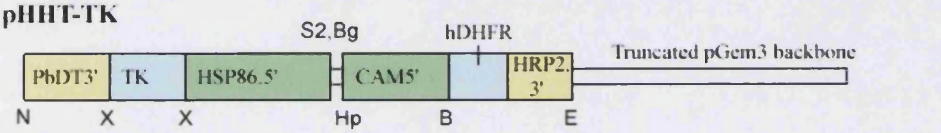
B. Constructs produced in this project for stable transfection. 'rep20' is a 725bp long section of rep20 repeats cloned from the end of chromosome 2. For a detailed description of $RifB^2$, RΔG and $RifB^2\Delta$ see section 6.3.

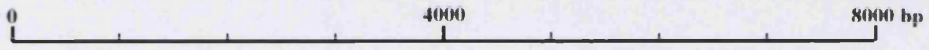
A.

pHC1-CAT



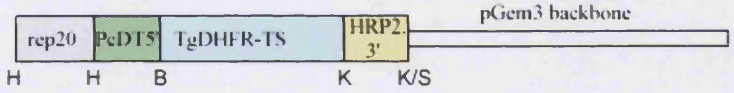
pHHT-TK



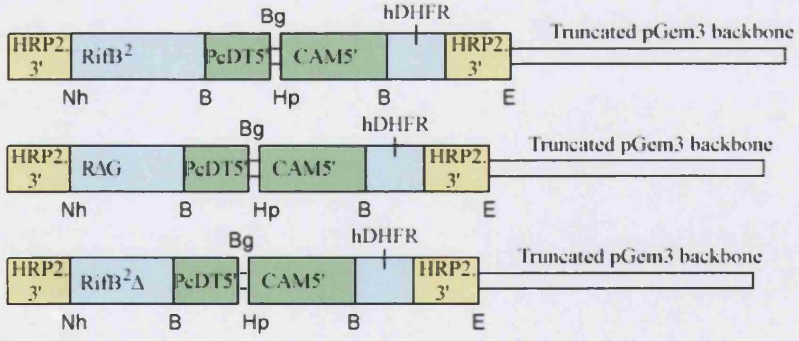


B.

pTgDTrep20



phDRif variants



phDCAT

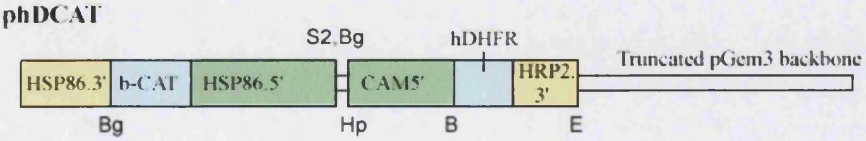


Fig 7.2: **Effect of ganciclovir on parasites transfected with pHHT-TK.** A culture of parasites which had been electroporated with pHHT-TK and selected using 10nm WR99210 was split, and aliquots were cultured in the presence of 10nm WR99210 and 0, 1 or 10μM ganciclovir. Thin blood smears were taken at the time intervals indicated and the number and stages of the parasites present were counted. 'Rings' = the ring stage of the trophozoite, 1-22hrs post-invasion; 'Trophozoites' = mature trophozoites, 23-34hrs post-invasion; 'Schizonts' = parasites with discernible multiple nuclei, 35-48hrs post-invasion. Error bars show standard deviation of duplicates.

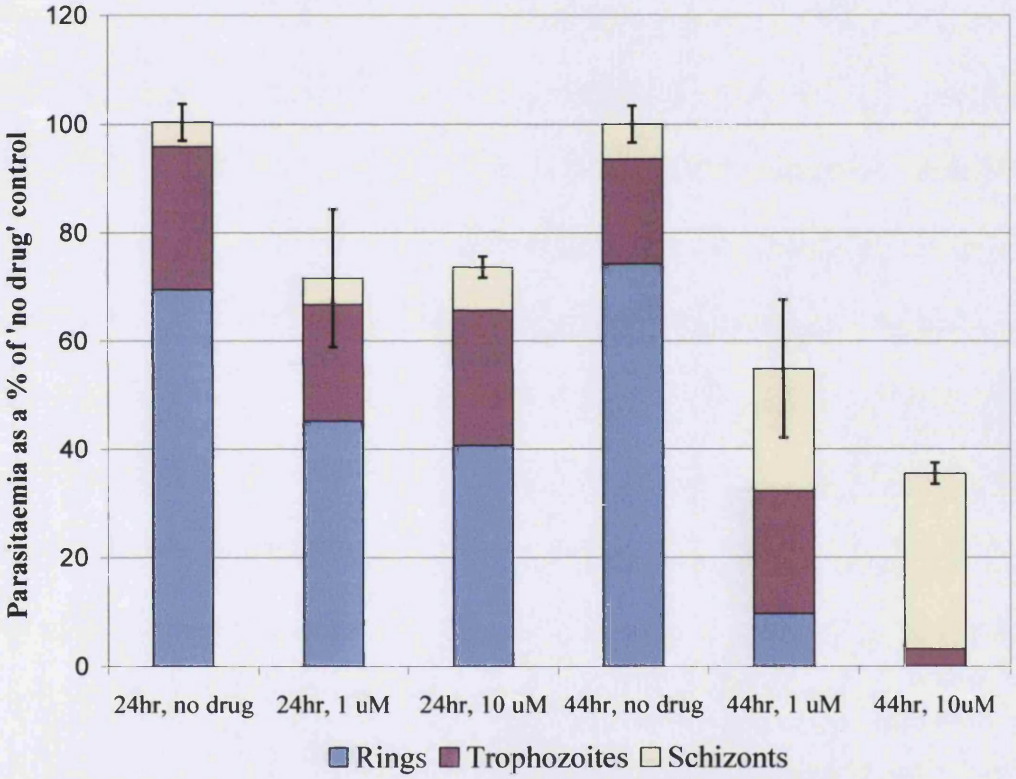
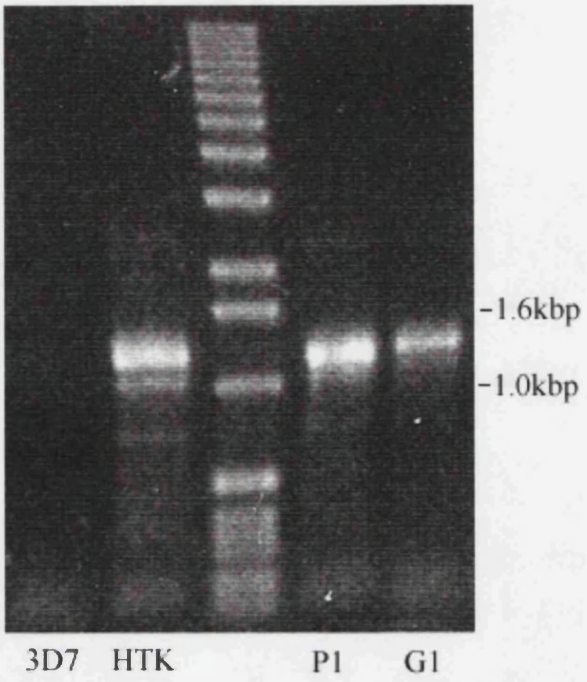

120
100
80
60
40
20
0
Parasitaemia as a % of 'no drug' control
24hr, no drug
24hr, 1 uM
24hr, 10 uM
44hr, no drug
44hr, 1 uM
44hr, 10uM
Rings
Trophozoites
Schizonts

Fig 7.3: **WR99210 resistant parasites contain the plasmid pHHT-TK.** A variety of DNA preparations were used as templates for PCR and material for Southern hybridization. Molecular weight markers are indicated in each case. HTK = purified pHHT-TK plasmid (grown in *E. coli*), 3D7 = genomic DNA purified from untransfected parasites, P1 = DNA purified from WR99210 resistant parasites using the S.N.A.P. plasmid miniprep kit, G1/G2 = genomic DNA purified from WR99210 resistant parasites.

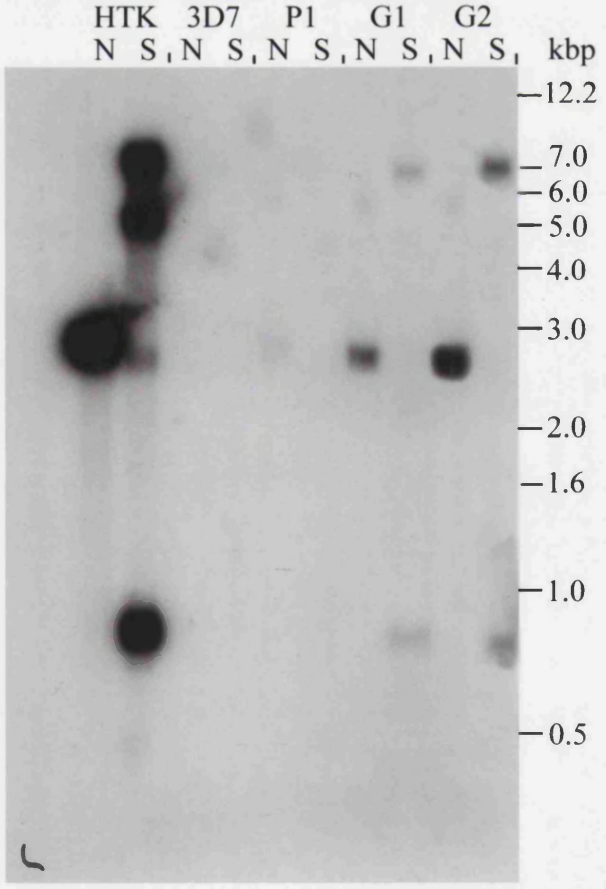
A. Products of PCR using primers against the middle of the CAM5' region and the 3' end of the hDHFR open reading frame. Predicted size of correct PCR product is 1.2 kb.

B. Southern hydridization using a probe against the pGem3zf(+) plasmid backbone. Tracks labelled 'N' contain DNA digested with *Eco*R I and *Not* I, while those labelled 'S' were digested with *Eco*R I and *Sca* I.

A.


-1.6kbp
-1.0kbp
3D7
HTK
P1
G1

B.


HTK
3D7
P1
G1
G2
N
S
kbp
12.2
7.0
6.0
5.0
4.0
3.0
2.0
1.6
1.0
0.5

8. Discussion

8.1 Project Summary

Plasmodium falciparum is the parasitic protozoan responsible for the most severe forms of human malaria, with annual mortality in the order of millions and huge associated morbidity. With over a quarter of the world's population exposed to the disease and current front line treatments becoming ineffective due to drug resistance, identifying new potential drug and vaccine targets in this species is a priority.

The asexual erythrocytic stages of the *P. falciparum* life cycle are both of primary importance pathologically, and particularly amenable to molecular biology as they can be cultured *in vitro*. One disadvantage of working with these stages is that they have a haploid genome, making classical genetic methods unusable. Therefore transgenic techniques become an essential tool for molecular genetics, which by gene knockouts, heterologous and orthologous gene expression, and gene replacements provide the means to identify and characterize the potential drug and vaccine targets mentioned above.

This project aimed to develop the emerging *P. falciparum* transfection technology (reviewed in Waterkeyn et al, 1999) by improving the efficiency at which plasmid DNA is taken up into erythrocytic parasites, and establishing new transfection-based approaches to the molecular biology of this species. These methods were then to be used to investigate important aspects of parasite biology, such as the control of gene expression and the function of particular genes.

In Chapter 3 a wide variety of transfection protocols were tested, and while many were unsuccessful the efficiency of electroporation was significantly improved; however, further improvement of at least an order of magnitude is still desirable. The development of a quantitative CAT reporter and assay system to assist with this and other investigations is detailed in Chapter 4. These methods are put to investigative use in Chapter 5, where the control of gene expression by a range of *P. falciparum* 5' non-coding sequences is explored. In this chapter several such regions which would be suitable for mediating stage-specific or high level expression of transgenes are characterised, and important promoter sequences for some of these are localized.

Chapter 6 described a novel approach to circumventing the problem of low transfection efficiency using a transgene to label the surface of transfected parasitized erythrocytes. The transgene used coded for a c-myc tagged Rifin (Rif2.7/c-myc). The protein was not shown to be expressed on the erythrocyte surface; however, the attempt revealed some of the important issues in *trans* expression, and raised some interesting questions about Rifin biology. Finally, Chapter 7 dealt with the production of a stable transfectant using the human DHFR gene as a drug selection marker.

In what follows, the implications of these findings in the light of other contemporary research and their impact on the field is discussed, culminating in proposals for future work in this area. Throughout, the ultimate goal of identifying new treatment strategies must be kept in mind.

8.2 Transfection efficiency: a major hurdle

At every stage of this investigation, the problem of low transfection efficiency proved to be a critical factor. Optimizing electroporation conditions – the best being 300V, 950μF in a 0.2cm cuvette - gave a four-fold improvement in the efficiency of transfection by this method (section 3.3) in line with that observed by Fidock and Wellems (1997), but still fell far short of that seen in other organisms. The fact that so few parasites take up transgenes makes any kind of signal from a transgene much weaker, as shown in section 5.5 when some of the less active promoters examined produced too little CAT to be characterized. Obviously if more parasites contained the transgene these would have been easier to detect. Similarly, the approach tested in Chapter 6 (which itself was aimed to circumvent the efficiency problem) was confounded by the inability to detect mRNA or protein from the small fraction of parasites containing the transgene (section 6.6), even though protein could be readily obtained by *in vitro* transcription & translation. Stable transfection, the only method currently available to generate a culture in which all parasites contain the transgene, is itself rendered much more difficult and laborious by the low transfection efficiency. Again, if more parasites took up the plasmid initially, such protracted culture periods (a month in section 7.3) to obtain transfectants would not be required.

Two physical explanations have been mentioned which contribute to this problem: the fact that electroporation destroys 50-70% of the cells exposed to it (section 1.3.3) and the need to get the plasmid across four membranes to the parasite nucleus rather than two (section 3.1). We would expect the pre-loaded erythrocyte strategy pioneered by Deitsch et al (2001) and here tested in section 3.4 to avoid the former of these two problems by not electroporating the parasites themselves, but in fact this method gave lower levels of CAT expression than direct electroporation under the conditions used. This may simply be because the indirect route is less efficient at getting DNA into the parasite, but it was also observed that the erythrocytes which had been electroporated once to pre-load (and then washed and left at 4°C to recover) were much more fragile than fresh red blood cells. Higher rates of lysis were observed when these cells were electroporated a second time after parasite invasion than in those which were only electroporated once, whether before or after invasion. Therefore even though pre-loading erythrocytes does not harm the parasites as directly as electroporating them would, it may damage their erythrocyte hosts sufficiently that a similar loss in viability occurs. A more gentle technique than electroporation is needed to bypass this problem, but unfortunately lipid and polymer based transfection reagents tested in section 3.2 were ineffective at achieving transfection.

The latter problem – of getting the DNA across four membranes – begs the question, when plasmids do get into the parasite nucleus, what route do they take? The details of the mechanism of electroporation (such as what, if any, host proteins are involved) are not well understood for any organism, despite its widespread use. Pores are induced in a cell's plasma membrane by electroporation which are sufficiently large and stable for DNA to pass through, but they are very short-lived. It therefore seems highly unlikely that in this case electroporation perforates all four membrane barriers for long enough for the plasmid to find its way across and between them. We also know from controls that the parasites cannot take up plasmid DNA at all from the external medium in the absence of electroporation.

The fact that pre-loading erythrocytes works at all shows that *P. falciparum* can transport plasmid DNA intact from the erythrocyte cytoplasm to the parasite nucleus, although whether this occurs only on invasion or spontaneously throughout intra-erythrocytic growth is unclear. It was observed in Chapter 5 that no plasmids gave much CAT activity soon after electroporation, but most seemed to give increasing CAT expression

as the parasites developed to schizonts, regardless of their expected stage specificity. One explanation for this would be if the parasites continued to take up more plasmid from the erythrocyte cytoplasm throughout their development, perhaps at an increasing rate in mature trophozoites and schizonts, when ingestion of material from the host cell also increases. If so, electroporation would merely serve to get the plasmid through the erythrocyte plasma membrane, and the parasite would do the rest. How the parasite would do this is unclear; only one structure, the cytostome, is known to traffic material from the erythrocyte cytoplasm across the parasitophorous vacuole. However, haemoglobin, the major substance ingested by this route, is targeted to the food vacuole and degradation, and it is not obvious how a plasmid could escape this fate. Even if there are no DNases to degrade the plasmid in the food vacuole, there is not thought to be any traffic between it and the nucleus.

The other possibility – that parasites can take up plasmid from pre-loaded erythrocyte during invasion – is no more likely, as at no stage are the host cell and parasite cytoplasm in contact, and an additional mechanism would be required to explain how direct electroporation works. Therefore a two stage process similar to that described above seems most probable. Of course, the central fact that electroporation is so inefficient suggests that the mechanism by which it works is a fairly improbable one, and more to do with an occasional failure of function (such as an ingested vesicle not going to the food vacuole) than one behaving normally.

To date little research effort has been put into increasing the efficiency at which DNA enters the parasite, instead focusing on facilitating the generation of stable transfectants once the plasmid is in the nucleus. This project's findings suggest that in order to obtain the dramatic increases in transfection efficiency that are desirable, we should examine the ways in which *P. falciparum* takes up other molecules from the erythrocyte cytoplasm or the external medium. As described above, the role of the cytostome in haemoglobin ingestion is well established, but other endocytic pathways may also operate. Furthermore, we know that parasites take up nutrients from outside the erythrocyte, because they can only be cultured successfully in medium enriched with human serum or a substitute such as Albumax. In section 1.3.1 the tubo-vesicular network seen in the parasitized erythrocyte was mentioned as a structure allowing controlled transport of material between the parasite and the external medium; it has even been proposed that a 'parasitophorous duct' exists linking the parasitophorous

vacuole directly to the outside, although the observations on which this is based are now thought to be artefactual (Taraschi, 2000). Elucidating the function of the TVN could suggest alternative ways of getting DNA into the parasite.

These considerations are of limited value if the molecules naturally absorbed by intra-erythrocytic parasites are all relatively small, such as amino acids, lipids or even globular proteins; plasmids are much larger than all of these. However, H. Imrie et al ('Human haptoglobin-related protein kills *Plasmodium falciparum*' – 12th British Society for Parasitology Malaria Meeting, 10th-12th September 2001) found that one essential component of human serum required by *P. falciparum* in culture is high density lipoprotein (HDL), a complex of lipids and proteins hundreds of kiloDaltons in size. It is not known whether this is ingested whole by the parasite or not, but perhaps if a plasmid could be packaged to resemble an HDL particle it would be more efficiently transported to the parasite cytoplasm.

Another transfection technology which has recently been reported to function in *P. falciparum* is RNAi, in which double-stranded RNA molecules inhibit the expression of genes of the same sequence by targeting their mRNA for degradation. In two recent papers double-stranded RNAs were found to inhibit the expression of their cognate mRNAs, in one case by electroporating them into the parasite and the other simply by incorporating them in the culture medium (McRobert & McConkey, 2002; Malhotra et al, 2002) The use of the morpholino oligonucleotides tested in section 3.5 is a similar technique, although it works by a different mechanism. In this study morpholinos were not readily taken up by the parasites, and other workers using RNAi methods have been unable to repeat the findings reported above (M. Blackman, T. Surentheran, personal communications). This emphasizes once again the large unknowns in malaria transfection which cause such variation between laboratories. Even so, methods which target parasite mRNA do have the potential to be very effective in this system, and morpholino oligonucleotides have the added advantage of being relatively small, which might make them more suitable for exploiting the parasite's existing transport mechanisms.

8.3 Reporter genes and assays in *P. falciparum*

The last few years since transfection was first achieved have seen a range of reporter genes and selectable markers described for *P. falciparum.* In addition to CAT (described in section 4.1) and TgDHFR-TS & hDHFR (section 7.1) three other drug selectable markers have been reported to be effective in this system: neomycin phosphotransferase, blasticidin S deaminase (Mamoun et al, 1999a) and puromycin N-acetyltransferase (de Koning-Ward et al, 2001). The next reporter gene to be used after CAT was firefly luciferase, which had already been shown to function in the avian parasite *P. gallinaceum* (Goonewardene et al, 1993). This was followed by green fluorescent protein (GFP - Van Wye and Haldar, 1997) and more recently a different luciferase from *Renilla regiformis* (Militello & Wirth, 2003). Kadekoppala et al (2000) achieved something of the best of both worlds with a DHFR-GFP chimera.

All of these genes have a role to play in future transgenic research in *P. falciparum;* the drug selection markers are essential for obtaining stable transfectants, while reporter genes allow all aspects of gene expression and protein trafficking to be investigated, as well as being used as a control for successful transfection. GFP is currently unique in that it allows the visualization of parasite gene expression *in situ*, whereas the other reporter genes require the parasites to be killed to extract the enzyme. Its expression can also be quantitated at a population level by fluorescence activated cell sorting (FACS). However, some workers have been unable to express detectable levels of GFP in transient transfections (Burghaus and Lingelbach, 2001), which is perhaps unsurprising when only one in a thousand parasites may be carrying the gene. For this reason GFP and GFP fusions have only been used productively in stable transfection studies.

The two luciferase genes mentioned above and CAT are the genes of choice for transient transfections. All encode enzymes which can transform many thousands of substrate molecules to product per copy of the protein, making them exquisitely sensitive; therefore they can be detected even at the low levels of transfection efficiency currently obtained in *P. falciparum.* Both firefly luciferase and CAT have been used successfully by many different workers in a range of laboratories, demonstrating the robustness of these methods. However, while the luciferase assay is relatively quick and easy to carry out, the published CAT assays are fairly slow and laborious. The quantitative CAT assay developed in Chapter 4 is a significant improvement on those which have been

published: it is quicker, less laborious, and more reproducible. It is also the first fully quantitative assay protocol described for transgenic use in this species, with a broad linear range of three orders of magnitude. This permits more accurate measurement of biological phenomena, such as the effect of parasitaemia on transfection efficiency measured in section 4.6. In this project the assay was used to measure transfection efficiencies (sections 4.6 and 6.5) and promoter activities (Chapter 5). It could also be coupled to new sub-cellular fractionation techniques (*ibid.*) to localize expression from different genes, particularly those expressed at levels low enough to make standard protein detection techniques difficult. One advantage of CAT over luciferase in some situations is that the former enzyme is the more stable of the two, making it more suitable for some of the time course and total expression experiments in Chapter 5; for example, luciferase produced from a promoter with a fairly small expression window might decay before the next measurement.

To date the relative sensitivities of these two reporter genes have not been directly compared in *P. falciparum.* Regardless of the relative advantages and disadvantages of each, having both of them available opens up a range of powerful experiments that are not possible with one alone; for example, dual labelling by fusing them to two different proteins, or using one to measure expression while using the other as an internal control.

When measuring promoter activity, an important alternative to a reporter gene encoding an enzyme is to measure mRNA levels. Transiently expressed transgenes in *P. falciparum,* even from strong promoters, produce insufficient mRNA to detect on a Northern blot (e.g. section 6.6), but this still leaves RT-PCR as a possible approach. Horrocks and Lanzer (1999b) developed an RT-PCR method to measure promoter activity after first discounting the CAT assay on two grounds: that CAT accumulated steadily throughout the erythrocytic cycle, and that the pattern of accumulation seemed to be independent of the promoter used. Using the improved CAT assay protocol described here, and measuring the increase in CAT between different time points rather than the total amount, different promoters were clearly seen to have different expression patterns. Furthermore, RT-PCR is not necessarily quantitative, although using real time techniques quantitative RT-PCR has now been applied to *P. falciparum* (Blair et al, 2002). Measuring steady state mRNA levels does not take account of any differences in the efficiency of translation initiation by different 5' UTRs; conversely, the CAT assay

gives no information about mRNA stability. It would be a valuable test to compare the data obtained by CAT assay and by RT-PCR for the same promoter.

These considerations, and the unexpected stabilizing effect of parasite cell extract on the CAT assay described in section 4.4, illustrate the importance of thoroughly developing and testing an assay before use. The work described here will allow future investigations to use CAT as a reporter with greater confidence and accuracy.

8.4 Unravelling transcriptional regulation

Transcription is a relatively poorly understood area of *P. falciparum* molecular biology, despite being of great importance both technically and clinically. On the technical side, an appropriate promoter is required for correct timing of expression and trafficking of proteins, as shown by heterologous expression studies on Apical Membrane Antigen 1 (AMA-1). Both PfAMA-1 transfected into *P. berghei* (Kocken et al, 1998) and *P. chabaudi* AMA-1 transfected into *P. falciparum* (Triglia et al, 2000) were only correctly expressed when placed under the control of the transfected parasite species' AMA-1 promoter. Clinically, transcriptional regulation is integral both to the development of the parasite, typified by the transcription of different rRNA genes in different developmental stages (Waters et al, 1989), and to immune evasion, such as during *var* gene switching (Deitsch et al, 2001).

In this project a wide range of possible promoter sequences from *P. falciparum* were analysed, several of which proved to be of great technical value. The GA3PDH 5' non-coding region consistently gave much higher CAT expression than the calmodulin promoter (section 5.5), which is one of those currently widely used in transfection experiments. Not only is it therefore more likely to produce detectable gene expression after transient transfection, but it could also facilitate stable transfections by increasing the expression of drug resistance genes. The H70L 5' region had activity restricted to mature trophozoites (section 5.5), and would therefore be useful for expressing transgenes specifically in this developmental stage. Similarly, the KAHRP 5' region, discussed in more detail below, gave high levels of expression in ring stages after the transfectants had passed through schizogony (section 5.10). These two make a substantial step towards having a range of promoters available for transfection which are specifically active in every parasite stage. Just as valuable would be an inducible

promoter, and two potential such promoters were tested in section 5.7. Neither gave sufficient response to the predicted stimulus to be useful, but the strategy of harnessing existing stimulus-response mechanisms in the parasite should be taken further. Given the apparent divergence of *P. falciparum* transcription from other eukaryotes, inducible promoter systems established in other organisms are unlikely to function. One possibility would be to use a microarray analysis of transcription similar to that described by Mamoun et al (2001) to identify genes which are activated in response to specific stimuli, and then investigate them in detail. This would have the added benefit of showing how many genes apart from the transgene would also be activated on induction.

The promoters from three genes of interest, *Pfmsp1, Pfemp3* and *rif2.7* all gave very low levels of expression. While Rifin proteins are found at relatively low levels *in vivo,* the other two are generally readily detectable – particularly MSP1, which is the most abundant protein in the merozoite membrane. One possible explanation for this discrepancy would be if the orf or 3' UTR of *Pfmsp1* confers greater stability on its mRNA than the orf and 3' UTR used here, leading to higher protein levels; mRNA stability is often determined by sequences in the 3' UTR (Mitchell & Tollervey, 2000). Another possible explanation is that the CAT activities observed do actually correspond to sufficient expression of MSP1 *in vivo,* and that it is in fact critical that not too much of this protein is expressed. MSP1 is transported to the merozoite surface, where it stays until the merozoite invades an erythrocyte, and it is possible that an excess of MSP1 might clog up the transport system. Interestingly, most of the 5' non-coding sequences which showed high promoter activity were from genes coding for cytosolic enzymes, which we would expect to be present in large quantities and undergoing a degree of protein turnover.

The exception to this rule is the *kahrp* gene; large amounts of CAT were expressed from the putative promoter of this gene. KAHRP is also an exported protein, this time to the inner surface of the erythrocyte plasma membrane, and is probably not undergoing turnover. However, it is by far the most abundant parasite protein found at the erythrocyte surface, being the first such protein identified (Vernot-Hernandez & Heidrich, 1984) and present in much greater quantities than, for example, PfEMP3. CAT expression from a 2.6kb length of the *kahrp* 5' non-coding region was higher and closer in timing to the *in vivo* production of KAHRP (*ibid.)* than when a shorter, 1.2kb length of the region was used. Lanzer et al (1992) had previously identified a sequence which

binds a nuclear factor and promotes transcription from this gene 1040bp upstream of the ATG, and therefore this sequence was included even in the shorter construct. Clearly additional sequences further upstream of the ATG are also an important part of this promoter.

Similarly, longer sections of the upstream regions from both the *enolase* gene and *Pfmsp1* gave higher expression levels than the 1.1kb segments first tested (section 5.10). The former gene has another orf present in head-to-head orientation just 1.9kb upstream of its ATG, suggesting that this longer sequence may contain bi-directional enhancer elements which increase the transcription of both orfs. Expression from the other end of the region should also be measured; if it is high enough, this 1.9kb region would be a relatively compact way of transcribing two orfs from the same plasmid. We can also conclude that an enhancer of the *Pfmsp1* gene lies in the MSP1.2 region (1.1 to 2.2 kb upstream of the ATG). In the absence of this element expression is below detectable levels, although a different enhancer (the PcDT element) restores partial expression. The PcDT element is not sufficient to promote detectable expression by itself, as shown by the inactivity of a construct containing this element adjacent to a non-functional section of a *Pfemp1* promoter, showing that the MSP1.1 region (the first 1.1kb upstream of the ATG) also contains vital promoter elements.

The fact that all three of the complete intergenic regions examined gave different activities from shortened segments of the same region shows that *P. falciparum* promoters have a bipartite structure, similar to that of other eukaryotes. This is particularly clear for *Pfmsp1,* where no expression is detected unless both sections of the promoter are present, or else MSP1.2 is substituted with the PcDT element. However, the increased activity of all three intergenic regions also suggests that *P. falciparum* promoters are longer than is typical for other eukaryotes. This would be in keeping with the findings of Watanabe et al (2002) who observed from a library of transcribed sequences that the 5' UTRs of *P. falciparum* genes are on average three times longer than in humans. This extension may be enforced by the extreme AT-richness of the non-coding sequences, meaning that longer sequences are needed to contain the same amount of information required for specifying promoter activity. In the same study it was observed that all the genes represented had multiple transcriptional start sites, which is also in contrast to most eukaryotic systems. The results described here are consistent

with the theory that transcription in *P. falciparum* has similar overall features but is relatively divergent in the details from that of other organisms characterized to date.

One of the most interesting observations of this project was that, of three different plasmids tested, all behaved differently in terms of stage specificity of expression in the second erythrocytic cycle after electroporation compared with the first (section 5.10). As discussed in section 5.11, this could be due to an effect of electroporation on the parasites, but is more likely to be related to the epigenetic status of the plasmid. Horrocks and Lanzer (1999b) had previously found that the GBP130 promoter loses its stage specific regulation when introduced transiently on a plasmid, and shown by limited micrococcal nuclease digestion that the episomes were not assembled on to nucleosomes, even though this occurs spontaneously when other eukaryotic cells are transfected (Jeong & Stein, 1994). In the results presented here, none of the plasmids tested showed the predicted pattern of expression immediately after transfection, but after schizogony all reflected the *in vivo* expression patterns of the genes from which the promoters came. A probable explanation for this transition is that when the parasite replicates the plasmid DNA it alters its epigenetic status, allowing stage-specific regulation. This hypothesis is supported by the observations of Deitsch et al (2001b), who found that silencing of an episomal *var* gene promoter transfected transiently into *P. falciparum* was only efficient after the parasites had passed through S-phase (i.e. replicated their DNA). Such an epigenetic change might occur by a number of mechanisms, such as DNA methylation or assembly of the episomes on to nucleosomes; supporting the latter view, Horrocks et al (2002) found that the aforementioned GBP130 promoter expressed on a stable episome (and therefore having undergone several rounds of DNA replication) did show stage specificity, and was correctly assembled on to nucleosomes. Another difference noted between plasmids electroporated into *P. falciparum* and those detected after several erythrocytic cycles is that the latter form multimers (Kadekoppala et al, 2001). We can therefore suggest a general mechanism for episome maintenance: plasmid produced in *E. coli* is electroporated into the parasite, where it is poorly regulated until the onset of DNA replication. At this point the plasmids are replicated, possibly undergoing homologous recombination into multimers during the replication process, after which they are loaded on to nucleosomes where they are correctly regulated.

These discoveries are crucial to future transient transfection experiments, showing that specific regulation of transgenes will only occur once their episomes have been replicated by the parasite and incorporated into chromatin. Therefore all future studies

into promoter activity or transient protein expression should look for results in the second life cycle following electroporation.

8.5 Rifins and *trans* expression

Rifins are the protein products of the largest multigene family in *P. falciparum,* yet their function remains a mystery. One of the major obstacles to carrying out functional studies on these proteins since their discovery has been their low level of expression *in vivo* (Kyes et al, 1999) and the inability to express them in a heterologous system. Another approach to the question would be to express a unique Rifin from an episomal location in *P. falciparum,* and this was one of the potential benefits of the epitope tagged Rifin construct developed in Chapter 6. Additionally, as described in section 6.1, if the epitope-tagged Rifin was transported to the erythrocyte surface, it could be used to identify and select transfectants from the background of untransfected cells, by-passing many of the problems of low transfection efficiency. Unfortunately the Rifin was not expressed successfully, but in the attempt several interesting issues concerning the *trans* expression of *P. falciparum* proteins generally, and Rifins in particular, arose.

First of all, no protein was detected in transient transfection experiments using pRifCAT, even though the c-myc tagged Rifin could be readily recognized by antibody when translated *in vitro* (section 6.3). This might occur if the protein folds differently *in vivo* such that the c-myc epitope is concealed, although it was hoped that the insertion of the tag at two different sites would avoid this. Another potential confounding factor when dealing with a multigene family like *rif* is that it is not known whether the *rif2.7* gene is ever expressed in erythrocytic stages. Proteomic analysis by Florens et al (2002) suggests that more Rifins are expressed in sporozoites than in erythrocytic forms of *P. falciparum,* and the *rif2.7* promoter was found to express only minimal CAT activity (section 5.5). Furthermore, the *rif* family undergoes antigenic variation, so by analogy to *var* genes any given member of the family may be silenced under normal conditions. Addressing these concerns, the gene encoding Rifin2.7/c-myc was placed under the control of an asexual stage promoter; all reported examples of stage-specific expression to date are controlled at the level of transcription (e.g. for the sexual stage protein Pfs16, in Dechering et al, 1997), as is *var* silencing (Deitsch et al, 2001) and these mechanisms would have been circumvented. Moreover, at least some *rif* transcripts have been

detected in 3D7 parasites by Cheng et al (1998). Therefore perhaps the correct explanation for non-detection of Rifin2.7/c-myc is the simplest one – that it is expressed at a level too low to detect. In section 8.3 it was commented that only enzymatic reporter genes are readily detected in transient transfections, owing to the low efficiency of electroporation.

If this is the case, steps can be taken to boost the expression levels. The GA3PDH promoter described in Chapter 5 should give much greater expression of Rifin2.7/c-myc, possibly taking it into the detectable range. Also, higher levels of expression should be obtained in the second life cycle following electroporation, as was the case for all the plasmids tested in sections 5.8 and 5.10.

Given these difficulties in detecting transiently expressed proteins, it would be valuable to demonstrate that they are transcribed successfully. The difficulty here is that such a large quantity of plasmid DNA is used in the electroporation that it tends to mask the mRNA produced from the transgene. The two species could potentially be resolved by their different sizes in a Northern blot, but in section 6.6 this technique was not found to be sensitive enough to detect the mRNA. RT-PCR is more sensitive and has been used to detect expression from transient transfections (Horrocks & Lanzer, 1999b); in this study, however, left over plasmid could not be eliminated from the RNA preparations despite DNase treatment.

All these considerations suggest that the best way to obtain *trans* expression of proteins is to generate a stable transfectant, in which all the parasites contain the transgene and so the levels of both mRNA and protein will be much higher. The process of achieving this for Rif2.7/c-myc was begun in Chapter 7, and the important step was taken of identifying a vector which could be used successfully to select for a stable transfectant. Although stable transfectants containing the transgene encoding Rif2.7/c-myc have not yet been generated, it should now be possible to do so.

8.6 Future work

A wide range of areas of research into the molecular biology of *P. falciparum* have been developed in the course of this project, opening up many interesting avenues for future work.

The highest priority must be to obtain stable transfectants expressing Rif2.7/c-myc. This is only a small step beyond what has already been achieved, and will potentially be very valuable. It will enable the question of the elusive function of Rifins to be addressed; also, by providing a system in which episomal Rif2.7/c-myc can be characterized, allow the development of selection methods to facilitate transient transfection experiments. To facilitate this, the gene encoding Rif2.7/c-myc should be placed under the control of the powerful GA3PDH promoter and further effort put into detecting its mRNA by RT-PCR, and its protein, in both stable and transient transfections.

Additionally, further experiments on the full-length *kahrp* and *msp1* promoter regions, such as nested deletions and site-directed mutagenesis, should be carried out to pinpoint the location of the novel control sequences detected in this study. This might lead to important information about the regulation of these genes and the transcriptional mechanisms of *P. falciparum* in general.

The hypothesis concerning the changes undergone by plasmids during their replication by *P. falciparum* should be confirmed by isolating DNA from parasites in the second cycle following electroporation, and testing it for i) the presence of nucleosomes by limited micrococcal nuclease digestion, and ii) the oligomerization of the plasmid by restriction digestion and Southern hybridisation.

These approaches should ultimately lead to several valuable spheres of research: the evaluation of Rifins as potential vaccine candidates, the identification of drug targets in the *P. falciparum* transcription mechanism, and the facilitation of transgenic studies to further elucidate the biology and future treatment of malaria.

Bibliography

Abdel-Latif, M. S., Khattab, A., Lindenthal, C., Kremsner, P. G. and Klinkert, M. Q. (2002). "Recognition of variant Rifin antigens by human antibodies induced during natural Plasmodium falciparum infections." Infect Immun **70**(12): 7013-21

Adisa, A., Rug, M., Klonis, N., Foley, M., Cowman, A. F. and Tilley, L. (2003). "The signal sequence of exported protein-1 directs the green fluorescent protein to the parasitophorous vacuole of transfected malaria parasites." J Biol Chem **278**(8): 6532-42

Akompong, T., Kadekoppala, M., Harrison, T., Oksman, A., Goldberg, D. E., Fujioka, H., Samuel, B. U., Sullivan, D. and Haldar, K. (2002). "Trans expression of a Plasmodium falciparum histidine-rich protein II (HRPII) reveals sorting of soluble proteins in the periphery of the host erythrocyte and disrupts transport to the malarial food vacuole." J Biol Chem **277**(32): 28923-33

Alipour, H., Eriksson, P., Norbeck, J. and Blomberg, A. (1999). "Quantitative aspects of the use of bacterial chloramphenicol acetyltransferase as a reporter system in the yeast Saccharomyces cerevisiae." Anal Biochem **270**(1): 153-8

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10

Bahl, A., Brunk, B., Coppel, R. L., Crabtree, J., Diskin, S. J., Fraunholz, M. J., Grant, G. R., Gupta, D., Huestis, R. L., Kissinger, J. C., Labo, P., Li, L., McWeeney, S. K., Milgram, A. J., Roos, D. S., Schug, J. and Stoeckert, C. J., Jr. (2002). "PlasmoDB: the Plasmodium genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished)." Nucleic Acids Res **30**(1): 87-90

Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. and Mitchell, G. H. (2000). "A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages." Parasitol Today **16**(10): 427-33

Baruch, D. I., Gormely, J. A., Ma, C., Howard, R. J. and Pasloske, B. L. (1996). "Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1." Proc Natl Acad Sci U S A **93**(8): 3497-502

Behari, R., and Haldar, K. (1994). "Plasmodium falciparum: protein localization along a novel, lipid-rich tubovesicular membrane network in infected erythrocytes." Exp Parasitol **79**(3): 250-9.

Berendt, A. R., Simmons, D. L., Tansey, J., Newbold, C. I., and Marsh, K. (1989). "Intercellular adhesion molecule-1 is an endothelial cell adhesion receptor for Plasmodium falciparum." Nature **341**(6237): 57-9.

Bischoff, E., Guillotte, M., Mercereau-Puijalon, O. and Bonnefoy, S. (2000). "A member of the Plasmodium falciparum Pf60 multigene family codes for a nuclear protein

expressed by readthrough of an internal stop codon." Mol Microbiol **35**(5): 1005-16

Blackman, M. J., Scott-Finnigan, T. J., Shai, S. and Holder, A. A. (1994). "Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein." J Exp Med **180**(1): 389-93

Blair, P. L., Witney, A., Haynes, J. D., Moch, J. K., Carucci, D. J. and Adams, J. H. (2002). "Transcripts of developmentally regulated Plasmodium falciparum genes quantified by real-time RT-PCR." Nucleic Acids Res **30**(10): 2224-31

Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C. M., Craig, A., Davies, R. M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Barrell, B. G. and et al. (1999). "The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum." Nature **400**(6744): 532-8

Brown, K. N., and Brown, I. N. (1965). "Immunity to malaria: antigenic variation in chronic infections of Plasmodium knowlesi." Nature **208**(17): 1286-8.

Bruce, M. C., Alano, P., Duthie, S., and Carter, R. (1990). "Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development." Parasitology **100**(2): 191-200.

Buffet, P. A., Gamain, B., Scheidig, C., Baruch, D., Smith, J. D., Hernandez-Rivas, R., Pouvelle, B., Oishi, S., Fujii, N., Fusai, T., Parzy, D., Miller, L. H., Gysin, J. and Scherf, A. (1999). "Plasmodium falciparum domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection." Proc Natl Acad Sci U S A **96**(22): 12743-8

Burghaus, P. A. and Lingelbach, K. (2001). "Luciferase when fused to an N-terminal signal peptide is secreted from transfected Plasmodium falciparum and transported to the cytosol of infected erythrocytes." J Biol Chem **276**(29): 26838-45

Chen, Q., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P., and Wahlgren, M. (1998). "Developmental selection of var gene expression in Plasmodium falciparum." Nature **394**(6691): 392-5.

Cheng, Q., Cloonan, N., Fischer, K., Thompson, J., Waine, G., Lanzer, M. and Saul, A. (1998). "stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens." Mol Biochem Parasitol **97**(1-2): 161-76

Crabb, B. S. and Cowman, A. F. (1996). "Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum." Proc Natl Acad Sci U S A **93**(14): 7289-94

Crabb, B. S., Cooke, B. M., Reeder, J. C., Waller, R. F., Caruana, S. R., Davern, K. M., Wickham, M. E., Brown, G. V., Coppel, R. L. and Cowman, A. F. (1997a). "Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress." Cell **89**(2): 287-96

Crabb, B. S., Triglia, T., Waterkeyn, J. G. and Cowman, A. F. (1997b). "Stable transgene expression in Plasmodium falciparum." Mol Biochem Parasitol **90**(1): 131-44

de Koning-Ward, T. F., Thomas, A. W., Waters, A. P., and Janse, C. J. (1998). "Stable expression of green fluorescent protein in blood and mosquito stages of Plasmodium berghei." Mol Biochem Parasitol **97**(1-2): 247-52.

de Koning-Ward, T. F., Waters, A. P. and Crabb, B. S. (2001). "Puromycin-N-acetyltransferase as a selectable marker for use in Plasmodium falciparum." Mol Biochem Parasitol **117**(2): 155-60

Dechering, K. J., Thompson, J., Dodemont, H. J., Eling, W. and Konings, R. N. (1997). "Developmentally regulated expression of pfs16, a marker for sexual differentiation of the human malaria parasite Plasmodium falciparum." Mol Biochem Parasitol **89**(2): 235-44

Dechering, K. J., Kaan, A. M., Mbacham, W., Wirth, D. F., Eling, W., Konings, R. N., and Stunnenberg, H. G. (1999). "Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite Plasmodium falciparum." Mol Cell Biol **19**(2): 967-78.

Deitsch, K., Driskill, C. and Wellems, T. (2001a). "Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes." Nucleic Acids Res **29**(3): 850-3

Deitsch, K. W., Calderwood, M. S. and Wellems, T. E. (2001b). "Malaria. Cooperative silencing elements in var genes." Nature **412**(6850): 875-6

Duraisingh, M. T., Triglia, T. and Cowman, A. F. (2002). "Negative selection of Plasmodium falciparum reveals targeted gene deletion by double crossover recombination." Int J Parasitol **32**(1): 81-9

Egan, A. F., Morris, J., Barnish, G., Allen, S., Greenwood, B. M., Kaslow, D. C., Holder, A. A. and Riley, E. M. (1996). "Clinical immunity to Plasmodium falciparum malaria is associated with serum antibodies to the 19-kDa C-terminal fragment of the merozoite surface antigen, PfMSP-1." J Infect Dis **173**(3): 765-9

Elmendorf, H. G., and Haldar, K. (1994). "Plasmodium falciparum exports the Golgi marker sphingomyelin synthase into a tubovesicular network in the cytoplasm of mature erythrocytes." J Cell Biol **124**(4): 449-62.

Fernandez, V., Hommel, M., Chen, Q., Hagblom, P. and Wahlgren, M. (1999). "Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses." J Exp Med **190**(10): 1393-404

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

Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D., Moch, J. K., Muster, N., Sacci, J. B., Tabb, D. L., Witney, A. A., Wolters, D.,

Wu, Y., Gardner, M. J., Holder, A. A., Sinden, R. E., Yates, J. R. and Carucci, D. J. (2002). "A proteomic view of the Plasmodium falciparum life cycle." Nature **419**(6906): 520-6

Foth, B. J., Ralph, S. A., Tonkin, C. J., Struck, N. S., Fraunholz, M., Roos, D. S., Cowman, A. F., and McFadden, G. I. (2003). "Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum." Science **299**(5607): 705-8.

Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Smith, H. O., Fraser, C. M., Venter, J. C. and Hoffman, S. L. (1999). "The malaria genome sequencing project: complete sequence of Plasmodium falciparum chromosome 2." Parassitologia **41**(1-3): 69-75

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M. and Barrell, B. (2002). "Genome sequence of the human malaria parasite Plasmodium falciparum." Nature **419**(6906): 498-511

Goonewardene, R., Daily, J., Kaslow, D., Sullivan, T. J., Duffy, P., Carter, R., Mendis, K. and Wirth, D. (1993). "Transfection of the malaria parasite and expression of firefly luciferase." Proc Natl Acad Sci U S A **90**(11): 5234-6

Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982). "Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells." Mol Cell Biol **2**(9): 1044-51

Haynes, J. D. and Moch, J. K. (2002). "Automated synchronization of Plasmodium falciparum parasites by culture in a temperature-cycling incubator." Methods Mol Med **72**: 489-97

Horrocks, P., Dechering, K. and Lanzer, M. (1998). "Control of gene expression in Plasmodium falciparum." Mol Biochem Parasitol **95**(2): 171-81

Horrocks, P., and Lanzer, M. (1999a). "Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in Plasmodium falciparum." Mol Biochem Parasitol **99**(1): 77-87.

Horrocks, P. and Lanzer, M. (1999b). "Differences in nucleosome organization over episomally located plasmids coincides with aberrant promoter activity in P. falciparum." Parasitol Int **48**(1): 55-61

Horrocks, P., Pinches, R., Kriek, N. and Newbold, C. (2002). "Stage-specific promoter activity from stably maintained episomes in Plasmodium falciparum." Int J Parasitol **32**(10): 1203-6

Jeong, S. and Stein, A. (1994). "Micrococcal nuclease digestion of nuclei reveals

extended nucleosome ladders having anomalous DNA lengths for chromatin assembled on non-replicating plasmids in transfected cells." Nucleic Acids Res **22**(3):270-5

Kadekoppala, M., Kline, K., Akompong, T. and Haldar, K. (2000). "Stable expression of a new chimeric fluorescent reporter in the human malaria parasite Plasmodium falciparum." Infect Immun **68**(4): 2328-32

Kadekoppala, M., Cheresh, P., Catron, D., Ji, D., Deitsch, K., Wellems, T. E., Seifert, H. S. and Haldar, K. (2001). "Rapid recombination among transfected plasmids, chimeric episome formation and trans gene expression in Plasmodium falciparum." Mol Biochem Parasitol **112**(2): 211-8

Kaviratne, M., Khan, S. M., Jarra, W. and Preiser, P. R. (2002). "Small Variant STEVOR Antigen Is Uniquely Located within Maurer's Clefts in Plasmodium falciparum-Infected Red Blood Cells." Eukaryot Cell **1**(6): 926-35

Kirk, K. (2001). "Membrane transport in the malaria-infected erythrocyte." Physiol Rev **81**(2): 495-537.

Kocken, C. H., van der Wel, A. M., Dubbeld, M. A., Narum, D. L., van de Rijke, F. M., van Gemert, G. J., van der Linde, X., Bannister, L. H., Janse, C., Waters, A. P. and Thomas, A. W. (1998). "Precise timing of expression of a Plasmodium falciparum-derived transgene in Plasmodium berghei is a critical determinant of subsequent subcellular localization." J Biol Chem **273**(24): 15119-24

Kocken, C. H., van der Wel, A., and Thomas, A. W. (1999). "Plasmodium cynomolgi: transfection of blood-stage parasites using heterologous DNA constructs." Exp Parasitol **93**(1): 58-60.

Kocken, C. H., Ozwara, H., van der Wel, A., Beetsma, A. L., Mwenda, J. M., and Thomas, A. W. (2002). "Plasmodium knowlesi provides a rapid in vitro and in vivo transfection system that enables double-crossover gene knockout studies." Infect Immun **70**(2): 655-60.

Kwiatkowski, D. (1989). "Febrile temperatures can synchronize the growth of Plasmodium falciparum in vitro." J Exp Med **169**(1): 357-61

Kyes, S., Pinches, R. and Newbold, C. (2000). "A simple RNA analysis method shows var and rif multigene family expression patterns in Plasmodium falciparum." Mol Biochem Parasitol **105**(2): 311-5

Kyes, S. A., Rowe, J. A., Kriek, N. and Newbold, C. I. (1999). "Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum." Proc Natl Acad Sci U S A **96**(16): 9333-8

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(259): 680-5

Lanzer, M., de Bruin, D. and Ravetch, J. V. (1992). "A sequence element associated with the Plasmodium falciparum KAHRP gene is the site of developmentally regulated protein-DNA interactions." Nucleic Acids Res **20**(12): 3051-6

Malhotra, P., Dasaradhi, P. V., Kumar, A., Mohmmed, A., Agrawal, N., Bhatnagar, R. K. and Chauhan, V. S. (2002). "Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of Plasmodium falciparum." Mol Microbiol **45**(5): 1245-54

Mamoun, C. B., Gluzman, I. Y., Goyard, S., Beverley, S. M. and Goldberg, D. E. (1999a). "A set of independent selectable markers for transfection of the human malaria parasite Plasmodium falciparum [published erratum appears in Proc Natl Acad Sci U S A 1999 Sep 14;96(19):10944]." Proc Natl Acad Sci U S A **96**(15): 8716-20

Mamoun, C. B., Truong, R., Gluzman, I., Akopyants, N. S., Oksman, A. and Goldberg, D. E. (1999b). "Transfer of genes into Plasmodium falciparum by polyamidoamine dendrimers." Mol Biochem Parasitol **103**(1): 117-21

Mamoun, C. B., Gluzman, I. Y., Hott, C., MacMillan, S. K., Amarakone, A. S., Anderson, D. L., Carlton, J. M., Dame, J. B., Chakrabarti, D., Martin, R. K., Brownstein, B. H. and Goldberg, D. E. (2001). "Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite Plasmodium falciparum revealed by microarray analysis." Mol Microbiol **39**(1): 26-36

McAndrew, M. B., Read, M., Sims, P. F. and Hyde, J. E. (1993). "Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite Plasmodium falciparum." Gene **124**(2): 165-71

McRobert, L. and McConkey, G. A. (2002). "RNA interference (RNAi) inhibits growth of Plasmodium falciparum." Mol Biochem Parasitol **119**(2): 273-8

Meissner, M., Brecht, S., Bujard, H., and Soldati, D. (2001). "Modulation of myosin A expression by a newly established tetracycline repressor-based inducible system in Toxoplasma gondii." Nucleic Acids Res **29**(22): E115.

Menard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S., and Nussenzweig, V. (1997). "Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes." Nature **385**(6614): 336-40.

Militello, K. T. and Wirth, D. F. (2003). "A new reporter gene for transient transfection of Plasmodium falciparum." Parasitol Res **89**(2): 154-7

Miller, S. K., Good, R. T., Drew, D. R., Delorenzi, M., Sanders, P. R., Hodder, A. N., Speed, T. P., Cowman, A. F., de Koning-Ward, T. F., and Crabb, B. S. (2002). "A subset of Plasmodium falciparum SERA genes are expressed and appear to play an important role in the erythrocytic cycle." J Biol Chem **277**(49): 47524-32.

Mitsuya, H., Yarchoan, R., and Broder, S. (1990). "Molecular targets for AIDS therapy." Science **249**(4976): 1533-44.

Molineaux, L. (1997). "Malaria and mortality: some epidemiological considerations." Ann Trop Med Parasitol **91**(7): 811-25

Mota, M. M., Thathy, V., Nussenzweig, R. S., and Nussenzweig, V. (2001). "Gene targeting in the rodent malaria parasite Plasmodium yoelii." Mol Biochem Parasitol **113**(2): 271-8.

O'Donnell, R. A., de Koning-Ward, T. F., Burt, R. A., Bockarie, M., Reeder, J. C., Cowman, A. F., and Crabb, B. S. (2001a). "Antibodies against merozoite surface protein (MSP)-1(19) are a major component of the invasion-inhibitory response in individuals immune to malaria." J Exp Med **193**(12): 1403-12.

O'Donnell, R., Preiser, P. R., Williamson, D. H., Moore, P. W., Cowman, A. F. and Crabb, B. S. (2001b). "An alteration in concatameric structure is associated with efficient segregation of plasmids in transfected Plasmodium falciparum parasites." Nucleic Acids Res **29**(3): 716-24

O'Donnell, R. A., Freitas-Junior, L. H., Preiser, P. R., Williamson, D. H., Duraisingh, M., McElwain, T. F., Scherf, A., Cowman, A. F. and Crabb, B. S. (2002). "A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes." Embo J **21**(5): 1231-9

Oh, S. S., Voigt, S., Fisher, D., Yi, S. J., LeRoy, P. J., Derick, L. H., Liu, S. and Chishti, A. H. (2000). "Plasmodium falciparum erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton." Mol Biochem Parasitol **108**(2): 237-47

Perkins, S. L. and Schall, J. J. (2002). "A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences." J Parasitol **88**(5): 972-8

Phillips, R. S., Wilson, R. J. and Pasvol, G. (1978). "Differentiation of gametocytes in microcultures of human blood infected with Plasmodium falciparum." J Protozool **25**(3 Pt 2): 394-8

Robson, K. J. and Jennings, M. W. (1991). "The structure of the calmodulin gene of Plasmodium falciparum." Mol Biochem Parasitol **46**(1): 19-34

Ross, R. (1897). "On some peculiar pigmented cells found in two mosquitoes fed on malarial blood." British Medical Journal **1897**(II): 1786-8

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) "Molecular Cloning - A Laboratory Manual" 2nd edition, publ. Cold Spring Harbour Laboratory Press

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-7

Sankaran, L. (1992). "A simple quantitative assay for chloramphenicol acetyltransferase by direct extraction of the labeled product into scintillation cocktail." Anal Biochem **200**(1): 180-6

Sherman, I. W. (1979). "Biochemistry of Plasmodium (malarial parasites)." Microbiol Rev **43**(4): 453-95

Shirley, M. W., Biggs, B. A., Forsyth, K. P., Brown, H. J., Thompson, J. K., Brown, G. V. and Kemp, D. J. (1990). "Chromosome 9 from independent clones and isolates

of Plasmodium falciparum undergoes subtelomeric deletions with similar breakpoints in vitro." Mol Biochem Parasitol **40**(1): 137-45

Sidhu, A. B., Verdier-Pinard, D. and Fidock, D. A. (2002). "Chloroquine resistance in Plasmodium falciparum malaria parasites conferred by pfcrt mutations." Science **298**(5591): 210-3

Snounou, G., Jarra, W. and Preiser, P. R. (2000). "Malaria multigene families: the price of chronicity." Parasitol Today **16**(1): 28-30

Snow, R. W., Omumbo, J. A., Lowe, B., Molyneux, C. S., Obiero, J. O., Palmer, A., Weber, M. W., Pinder, M., Nahlen, B., Obonyo, C., Newbold, C., Gupta, S. and Marsh, K. (1997). "Relation between severe malaria morbidity in children and level of Plasmodium falciparum transmission in Africa." Lancet **349**(9066): 1650-4

Subbayya, I. N., Ray, S. S., Balaram, P. and Balaram, H. (1997). "Metabolic enzymes as potential drug targets in Plasmodium falciparum." Indian J Med Res **106**: 79-94

Sultan, A. A., Thathy, V., Frevert, U., Robson, K. J., Crisanti, A., Nussenzweig, V., Nussenzweig, R. S., and Menard, R. (1997). "TRAP is necessary for gliding motility and infectivity of plasmodium sporozoites." Cell **90**(3): 511-22.

Sultan, A. A., Thathy, V., de Koning-Ward, T. F., and Nussenzweig, V. (2001). "Complementation of Plasmodium berghei TRAP knockout parasites using human dihydrofolate reductase gene as a selectable marker." Mol Biochem Parasitol **113**(1): 151-6.

Summerton, J. and Weller, D. (1997). "Morpholino antisense oligomers: design, preparation, and properties." Antisense Nucleic Acid Drug Dev **7**(3): 187-95

Taraschi, T. F. (1999). "Macromolecular transport in malaria-infected erythrocytes." Novartis Found Symp **226:** 114-20

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Res **25**(24): 4876-82

Trager, W. and Jensen, J. B. (1976). "Human malaria parasites in continuous culture." Science **193**(4254): 673-5

Trape, J. F., Pison, G., Spiegel, A., Enel, C. and Rogier, C. (2002). "Combating malaria in Africa." Trends Parasitol **18**(5): 224-30

Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S. and Cowman, A. F. (2000). "Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species." Mol Microbiol **38**(4): 706-18

Turner, R. B. and Woodward, R. B. (1953). "The Alkaloids" publ. New York Academic Press.

van der Wel, A. M., Tomas, A. M., Kocken, C. H., Malhotra, P., Janse, C. J., Waters, A. P., and Thomas, A. W. (1997). "Transfection of the primate malaria parasite

Plasmodium knowlesi using entirely heterologous constructs." J Exp Med **185**(8): 1499-503.

van Dijk, M. R., Waters, A. P., and Janse, C. J. (1995). "Stable transfection of malaria parasite blood stages." Science **268**(5215): 1358-62.

van Dijk, M. R., Janse, C. J., and Waters, A. P. (1996). "Expression of a Plasmodium gene introduced into subtelomeric regions of Plasmodium berghei chromosomes." Science **271**(5249): 662-5.

van Dijk, M. R., Vinkenoog, R., Ramesar, J., Vervenne, R. A., Waters, A. P., and Janse, C. J. (1997). "Replication, expression and segregation of plasmid-borne DNA in genetically transformed malaria parasites." Mol Biochem Parasitol **86**(2): 155-62.

VanWye, J. D. and Haldar, K. (1997). "Expression of green fluorescent protein in Plasmodium falciparum." Mol Biochem Parasitol **87**(2): 225-9

Vernot-Hernandez, J. P. and Heidrich, H. G. (1984). "Time-course of synthesis, transport and incorporation of a protein identified in purified membranes of host erythrocytes infected with a knob-forming strain of Plasmodium falciparum." Mol Biochem Parasitol **12**(3): 337-50

Waller, R. F., Reed, M. B., Cowman, A. F., and McFadden, G. I. (2000). "Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway." Embo J **19**(8): 1794-802.

Wang, P., Wang, Q., Sims, P. F. and Hyde, J. E. (2002). "Rapid positive selection of stable integrants following transfection of Plasmodium falciparum." Mol Biochem Parasitol **123**(1): 1-10

Watanabe, J., Sasaki, M., Suzuki, Y. and Sugano, S. (2002). "Analysis of transcriptomes of human malaria parasite Plasmodium falciparum using full-length enriched library: identification of novel genes and diverse transcription start sites of messenger RNAs." Gene **291**(1-2): 105-13

Waterkeyn, J. G., Crabb, B. S. and Cowman, A. F. (1999). "Transfection of the human malaria parasite Plasmodium falciparum." Int J Parasitol **29**(6): 945-55

Waterkeyn, J. G., Wickham, M. E., Davern, K. M., Cooke, B. M., Coppel, R. L., Reeder, J. C., Culvenor, J. G., Waller, R. F., and Cowman, A. F. (2000). "Targeted mutagenesis of Plasmodium falciparum erythrocyte membrane protein 3 (PfEMP3) disrupts cytoadherence of malaria-infected red blood cells." Embo J **19**(12): 2813-23.

Waters, A. P., Syin, C. and McCutchan, T. F. (1989). "Developmental regulation of stage-specific ribosome populations in Plasmodium." Nature **342**(6248): 438-40

Wickham, M. E., Rug, M., Ralph, S. A., Klonis, N., McFadden, G. I., Tilley, L., and Cowman, A. F. (2001). "Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes." Embo J **20**(20): 5636-49.

Wu, Y., Sifri, C. D., Lei, H. H., Su, X. Z. and Wellems, T. E. (1995). "Transfection of Plasmodium falciparum within human red blood cells." Proc Natl Acad Sci U S A **92**(4): 973-7

Wu, Y., Kirkman, L. A. and Wellems, T. E. (1996). "Transformation of Plasmodium falciparum malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine." Proc Natl Acad Sci U S A **93**(3): 1130-4