CLARK, K

TOWARDS THE DEVELOPMENT OF AN RT-PCR PROTOCOL FOR MALARIA

1998

TOWARDS THE DEVELOPMENT OF AN RT-PCR PROTOCOL FOR MALARIA PARASITE GENES

by

Katherine Clark

Submitted in partial fulfilment of the requirements for the degree

of

Master of Science (Biochemistry)

in the

Faculty of Biological and Agricultural Sciences

University of Pretoria

Pretoria

November 1998

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to the following people:

Venny Nyambati, whom I worked with for two years and who taught me a great deal about scientific methods, for her friendship, enthusiasm and help;

My supervisor **Professor A.I. Louw** (Department of Biochemistry, University of Pretoria) for his enthusiasm, support, patience, insight and advice;

My co-supervisor **Professor A.W.H. Neitz** (Department of Biochemistry, University of Pretoria) for his valuable insight and advice;

My husband, **Dale** for his support, patience, encouragement and interest during the course of this degree;

Fourie Joubert and **Jaco de Ridder** the resident computer experts in the department, without whom much of this work would have been impossible;

My **parents** for their encouragement and support;

Everyone in the **malaria team** for their friendship and advice in the lab;

My **colleagues** and **friends** in the Department of Biochemistry for their friendship and help when requested;

The **Medical Research Council**, the **Foundation for Research Development** and the **University of Pretoria** for financial support.

TABLE OF CONTENTS

LIST OF FIGURES

APPENDIX B

APPENDIX C

LIST OF TABLES

APPENDIX A

LIST OF ABBREVIATIONS

ΔG	Gibbs thermodynamic value
A	Adenine
Ala	Alanine
AMA	Apical membrane antigen
Arg	Arginine
Asn	Asparagine
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
C	Cytosine
CD	Cluster of differentiation
cDNA	Complementary DNA
CIAP	Calf intestinal alkaline phosphatase
CSP	Circumsporozoite protein
Cys	Cysteine
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
DDT	Dichlorodiphenyltrichloroethane
ddUTP	Dideoxyuridine triphosphate
DEPC	Diethyl pyrocarbonate
dGTP	Deoxyguanosine triphosphate
DHFR-TS	Dihydrofolate reductase-thymidylate synthase
DHPS	Dihydropteroate synthase
DIG	Digoxygenin

DISEC-TRISEC	Di- and trinucleotide sticky end cloning
DMF	Dimethyl formamide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DTT	Dithiotreitol
dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
EBA	Erythrocyte binding antigen
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol bis (2-aminoethyl ether)-N,N,N',N'-tetraacetic acid
ELAM	Endothelial cell adhesion molecule
EM	Erythrocyte membrane
EXP	Exported protein
G	Guanine
Gln	Glutamine
Glu	Glutamic acid
GPI	Glycosyl phosphatidyl inositol
GSP	Gene specific primer
GTP	Guanosine triphosphate
HBsAg	Hepatitis B surface antigen
HEPES	(N-[2-Hydroxyethyl] piperazine-N'-[2-ethane-sulphonic acid])
HMS	Hexose monophosphate shunt
HRP	Histidine rich protein
HSP	Heat Shock Protein
ICAM	Intercellular adhesion molecule
IFN	Interferon
IL	Interleukin

Ile	Isoleucine
IPTG	Isopropyl-D-galactoside
kcal	Kilocalorie
LB Broth	Luria-Bertani Broth
Leu	Leucine
LSA	Liver stage antigen
Met	Methionine
MOPS	4-Morpholinepropanesulphonic acid
mRNA	Messenger RNA
MSP	Merozoite surface protein
NAD^+	Nicotinamide-adenine dinucleotide (oxidised form)
NBT	Nitroblue tetrazolium
NCBI	National Centre for Biotechnology Information
NIH	National Institutes of Health
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PfEMP	*Plasmodium falciparum* erythrocyte membrane protein
Pfg	Gametocyte antigen
Pfmdr	*Plasmodium falciparum* multidrug resistance protein
Pfs	Postfertilization antigen
Pfu	*Pyrococcus furiosus*
Pgh1	P-glycoprotein homologue
PPM	Parasite plasma membrane
PVM	Parasite vacuole membrane
Pwo	*Pyrococcus woesei*
RACE	Rapid amplification of complementary DNA ends

RAP	Rhoptry associated protein
RbCl	Rubidium chloride
RESA	Ring-infected erythrocyte surface antigen
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	RNase inhibitor
rRNA	Ribosomal RNA
RT-PCR	Reverse transcriptase polymerase chain reaction
RT-primer	Reverse transcriptase primer
S	Svedberg constant
SAH	S-adenosyl homocysteine
SAHH	S-adenosylhomocysteine hydrolase
SALSA	Sporozoite and liver stage antigen
SDS	Sodium dodecyl sulphate
Ser	Serine
SERA	Serine repeat antigen
STARP	Sporozoite threonine and asparagine rich protein
T	Thymine
Ta	Annealing temperature
Taq	*Thermus aquaticus*
TBS	Tris buffered saline
TEMED	N, N, N', N'- tetramethyl ethylenediamine
Tm	Melting temperature
TNF	Tumour necrosis factor
TRAP	Thrombospondin-related anonymous protein
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
tRNA	Transfer RNA
TVM	Tubovesicular membrane
U	Uracil

UTP	Uridine triphosphate
UTR	Untranslated region
VCAM	Vascular cell adhesion molecule
WHO	World Health Organisation
X-gal	5-Bromo-4-chloro-3-indolyl-galactoside
X-phosphate	5-Bromo-4-chloro-3-indolyl-phosphate

CHAPTER 1

General Introduction

1.1 Malaria

The word 'parasite' has its origins in the ancient Greek word '*parasitos*', which means one who eats at the table of another (Ackerman, 1997). Many parasites are capable of causing severe debilitation and death in humans. One such parasite causes malaria, the symptoms of which have been recognised for centuries, although it was previously thought to be caused by 'miasma' (bad air or 'mal aria') emanating from swamps.

Malaria is one of the most serious tropical diseases prevalent in Asia, South America and Africa (Collins and Paskewitz, 1995). It results in an estimated 300-500 million clinical cases and between 1.5 and 2.7 million deaths per year (MacLean *et al.*, 1997). Sub-Saharan Africa bears the brunt of the world's malaria infections, with ~90% of the malaria related deaths (le Sueur, 1996). The distribution of malaria in South Africa is predominantly in the low altitude areas of the Northern Province, Mpumalanga and the north eastern region of KwaZulu/Natal (Department of Health of South Africa, 1996).

Malaria is caused by the apicomplexan, protozoan parasite *Plasmodium*, which belongs to the order Haemosporidida (implying that it is parasitic in the blood of vertebrates) of the phylum Protozoa of the Kingdom Protista (Cox, 1993). There are four species of the genus *Plasmodium* responsible for causing human malaria, namely *Plasmodium falciparum*, *P. vivax*, *P. ovale* and *P. malariae*.

P. falciparum is the most common form of malaria, as well as the most severe. The infection is acute and often results in death (Cox, 1993). It is also responsible for more than 90% of the infections in Sub-Saharan Africa and encompasses between 90% and 95% of South African malaria infections (Department of Health of South Africa,

1996). *P. vivax*, *P. ovale* and *P. malariae* infections, though debilitating, are seldom fatal (Cox, 1993).

1.2 Life cycle of the genus *Plasmodium*

On the 20th of August 1897 Ronald Ross (1857-1932) discovered that malaria parasites are transmitted by *Anopheles* mosquitoes. For this discovery he was awarded the Nobel Prize in 1902. Prior to this, Laveran had identified malaria parasites in blood and had established the relationship between the parasites and infection. He had also reached the conclusion that there was more that one species of parasite, though he knew nothing about the mode of transmission (Hagan and Chauhan, 1997).

Today it is common knowledge that the parasite life cycle involves two hosts, namely the invertebrate mosquito (*Anopheles*) and the vertebrate host. There are at least 200 species of *Anopheles* capable of transmitting malaria, although they differ in transmission efficiency (Collins and Paskewitz, 1995). The parasitic life cycle can be divided into two distinct phases, asexual and sexual.

The asexual phase in humans commences when sporozoites are released from the salivary glands of the biting mosquito. They enter the hepatocytes and develop over several days into thousands of merozoites which, upon release, invade the erythrocytes where they develop into ring stage (immature) trophozoites, mature trophozoites and multi-nucleate schizonts. The daughter merozoites produced by the schizont-infected cells are released from the erythrocytes and can invade other erythrocytes in a number of successive cycles (Kemp *et al.*, 1987).

The life cycle of the *Plasmodium* species is summarised in Figure 1.1.

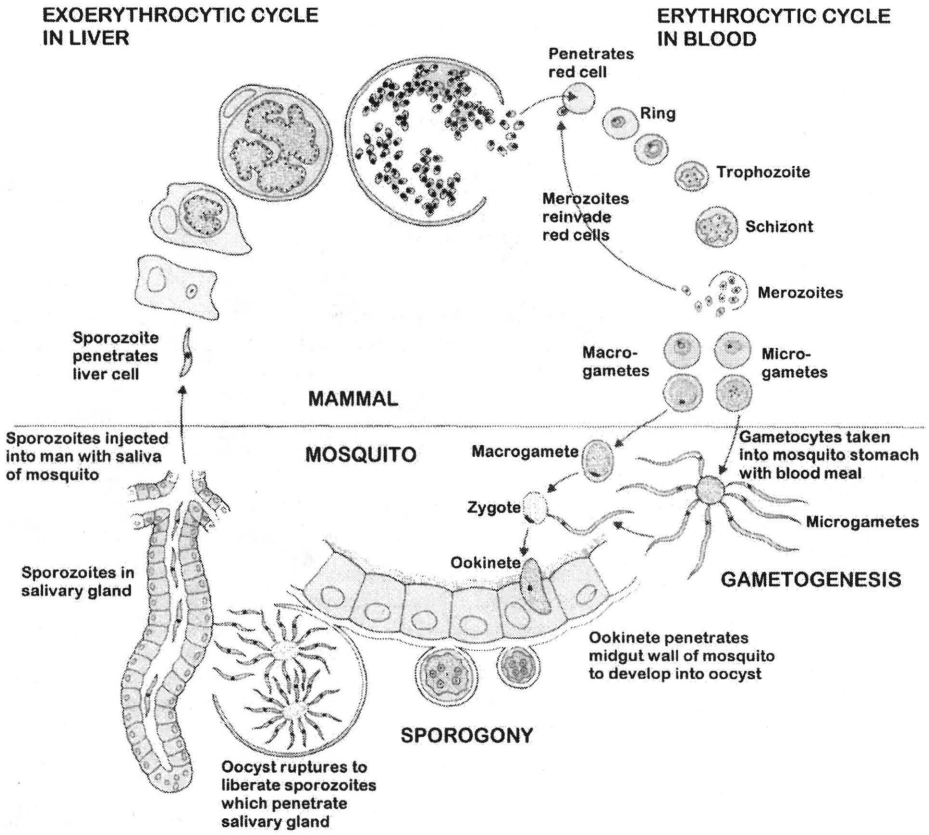


Figure 1.1 The life cycle of the *Plasmodium* species (Adapted from Cox, 1993).

After invasion some parasites differentiate into the sexual forms (gametocytes). When these gametocytes are ingested by another mosquito they initiate the sexual stage (Kemp *et al.*, 1987). After exflagellation the gametes combine to form a zygote. Sexual recombination occurs in the mosquito midgut. The ookinete subsequently penetrates the midgut where it develops into an oocyst. Thousands of new sporozoites are produced in the oocyst. The sporozoites migrate to the salivary glands, from where they can be inoculated into their next host (Doolan and Hoffman, 1997).

1.3 Pathogenesis of malaria

Pathogenesis is the interplay between various host and parasite factors that are responsible for the pathology of the disease. In malaria the interaction between the parasite and the host's immune system appears to be responsible for a great deal of the pathology. The production of cytokines is central to this interaction (Miller *et al.*,

1994). Severe and complicated malaria involves most of the tissue and organ systems of the body (Mendis and Carter, 1995).

The initial hepatocytic stage is not as severe as the erythrocytic stage, as only tens to hundreds of sporozoites are involved in the former stage (Miller *et al.*, 1994). There is however an immune reaction at this stage, which is primarily cell mediated. $CD8^+$ Tc-cells are responsible for the destruction of infected hepatocytes. They also secrete interferon IFN-γ which prevents parasite multiplication (Taverne, 1993).

The erythrocytic stage begins with the invasion of the merozoites released from the hepatocytes. In this stage they now have vastly increased numbers and are more likely to differentiate into the sexual gametocyte form. The erythrocytic stage is the most pathogenic since the parasites multiply rapidly. Fever and chills are associated with the rupture of the infected erythrocytes when the merozoites are released (Miller *et al.*, 1994).

Anaemia is characteristic of malaria but the loss of erythrocytes is greater than can be accounted for by rupture during merozoite release alone. There are three possible explanations. Firstly there is the formation of an immune complex consisting of antibodies in complex with parasite antigens. These bind uninfected erythrocytes which are cleared from the circulation by the spleen. The presence of autoantibodies (which arise from polyclonal activity) in the bloodstream also reduces the erythrocyte numbers in a similar fashion. Lastly, the formation of new red blood cells from stem cells is inhibited by the presence of TNF-α (Taverne, 1993).

Further manifestations of the disease include a nephrotic syndrome arising from deposition of immune complexes in the kidneys (mostly in *P. malariae* infection). Enlargement of the liver and the spleen is also noted and is due to an increase in the number of macrophages and lymphocytes (Taverne, 1993).

Cytokine release is thought to play a large role in the symptoms of malaria, especially fever. Macrophages and lymphocytes are stimulated to produce cytokines among

others, TNF-α and IL-1 (Taverne, 1993). TNF-α proved to be both harmful and beneficial to the host. The high fever associated with TNF-α increase has been shown to arrest or inhibit parasite growth, although there are other non-beneficial symptoms associated with cytokine release. Although the role of TNF-α has been fairly extensively studied, the role of other cytokines should not be disregarded. The phospholipid, glycosyl phosphatidyl inositol (GPI), has been recognised as the malaria 'toxin', i.e. the molecule produced by the parasite which induces the highest TNF-α production. This phospholipid is found in both a free form, as well as bound to the C-terminal of the merozoite surface proteins MSP-1 and MSP-2, where it serves to anchor the protein in the membrane (Mendis and Carter, 1995). There may also be other important toxins (Kwiatkowski and Marsh, 1997).

It is also believed that upregulation of the amount of TNF-α leads to an increase in the number of ICAM-1 (intercellular adhesion molecule) receptors (Miller *et al.*, 1994). Infected erythrocytes contain a cluster of proteins on their surface known as PfEMP-1 (*Plasmodium falciparum* infected erythrocyte membrane protein) in the form of knobs. These are parasite-expressed erythrocyte surface proteins which show substantial antigenic and adhesive phenotype variation. The PfEMP-1 proteins are involved in the adherence of the infected erythrocyte to the vascular epithelium via surface receptors such as ICAM-1, but also VCAM-1 (vascular cell adhesion molecule), ELAM-1 (endothelial cell adhesion molecule), CD36 and thrombospondin (Mendis and Carter, 1995). These PfEMP-1 proteins are encoded by the large and diverse multi-gene *var* family. All these genes contain areas homologous to an erythrocyte receptor binding domain (Smith *et al.*, 1995; Su *et al.*, 1995). This sequestration of the parasitised erythrocyte, especially in the post-capillary venules of the brain, along with the increased localised concentration of cytokines and their secondary mediators such as reactive oxygen species and nitric oxide, is thought to be responsible for cerebral malaria, although whether these factors act in isolation or in combination still requires further investigation (Berendt *et al.*, 1994; Clark and Rockett, 1994; Grau and Kossodo, 1994; Mendis and Carter, 1995).

Rosette formation is also characteristic of parasites capable of sequestration. Rosetting requires the binding of an infected erythrocyte to an uninfected erythrocyte via ligands such as carbohydrate domains, CD36 and the blood group antigens (Mendis and Carter, 1995).

Another strategy of the parasite to evade the immune system is to produce soluble antigens. These then bind antibodies in circulation and effectively distract the immune system from the parasite itself (Taverne, 1993).

People resident in areas of high endemicity, show signs of clinical immunity. This is when the people still harbour bloodstage parasites, but the frequency of clinical disease itself is greatly lowered. Children above the age of five are largely unresponsive to the presence of parasites in their blood and from the teenage years, the anti-parasite immunity becomes highly effective. This tolerance is acquired following prolonged continuous exposure to malarial infection, but in the absence of significant parasite challenge, this tolerance drops rapidly (Miller *et al.*, 1994; Mendis and Carter, 1995).

It seems that parasite and host genetics also play a large role in the pathology of the disease. Parasite genetics may include the ability to induce greater amounts of TNF, possibly through the presence of a 'toxin'. In the human host, thalassaemias and other erythrocyte abnormalities confer protection against malaria. The most notable of these is the HbS (sickle-cell) gene, which confers protection against severe *P. falciparum* malaria and homozygosity of the Duffy negative gene, which confers total refractoriness to *P. vivax* infection. It also appears that there is an association between the presence of certain MHC alleles and protection against severe anaemia and cerebral malaria (Mendis and Carter, 1995).

1.4 Impact of malaria

The heaviest burden of disease caused by *Plasmodium* infection is borne by Africa. A 1995 estimate placed childhood malaria related deaths in Africa at 800 000 per year (Mendis and Carter, 1995). Malaria is one of the most common causes of school

absenteeism (Kondrachine and Trigg, 1997). Around 40 000 survivors of cerebral malaria suffered serious, permanent neurological consequences (Mendis and Carter, 1995).

Malarial infections have economic ramifications for both the people and their country. The people suffer a loss as a result of absence from work. Loss of manpower and the cost of diagnosing and treating malaria affects the country. Productivity in the long and short term is thus affected and this then serves to reinforce poverty (Evans and Jamison, 1994).

1.5 Control of malaria

In the 1950's the WHO and other institutions made an attempt to eradicate malaria. In some areas they succeeded. The USA, Europe, most Asian areas of the then USSR and the Caribbean were malaria free. Malaria was never completely eradicated in Southeast Asia, Africa and South America and has now returned to areas such as Sri Lanka and Madagascar (Krogstad, 1996). Sub-Saharan Africa was excluded from the program due to the force of malaria transmission and a lack of infrastructure (Collins and Paskewitz, 1995).

Today the incidence of malaria is greater than it was 40 years ago (Krogstad, 1996). There are many reasons for this increase. Firstly, the WHO relied on two assumptions, namely that malaria could be treated with inexpensive drugs such as chloroquine and that its transmission could be halted by killing its vector, the *Anopheles* mosquito using residual insecticides like DDT (dichlorodiphenyltrichloroethane). Their hopes were dashed by both parasite resistance to chloroquine and insect resistance to insecticides. This prevalent drug resistance was one of the reasons why malaria was never eradicated in Southeast Asia. Logistical difficulties also made it difficult to establish an eradication campaign in Sub-Saharan Africa where transmission, morbidity and mortality were the greatest. Civil discord and easier means of travel greatly contributed to the spread of malaria from endemic to non-endemic regions in addition to the import

of drug resistant strains (Krogstad, 1996). Global warming and other environmental changes may also contribute to the spread of malaria (Greenwood, 1997).

1.5.1 Chemotherapeutic drugs and how the parasite evades them

In 1820, the alkaloid quinine was isolated from the bark of the cinchona tree (*Cinchona ledgeriana*). This compound was used for many years as an antimalarial until synthetic compounds were derived. In 1939, German chemists at Bayer synthesised chloroquine which is still being used as an antimalarial today. In 1942 Schönhofer proposed that tautomerism was required for antimalarial activity. An amino substituted pyrimidine, when opened, creates a biguanide displaying the desirable tautomeric qualities. Upon cyclisation of chloroguanide, the compound cycloguanil was obtained which showed high antimalarial activity and is also still being used today (Rost, 1989).

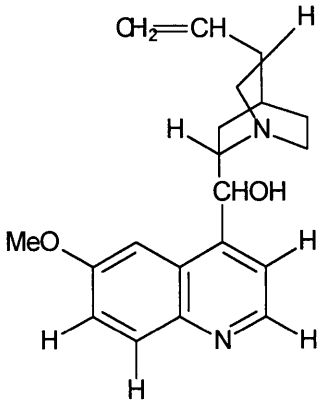
Various phases of the *Plasmodium* life cycle in humans offer potential targets for the suppression or interruption of parasite infection. Of the sporozoiticides, schizonticides and gametocytocides, the schizonticides are the most commonly used, as they are the most effective. They interfere with the growth and division of merozoites and can be divided into two groups namely, the pre-erythrocytic and the erythrocytic schizonticides. The pre-erythrocytic drugs are active against the merozoites in the hepatic tissue. These include the 8-amino quinolones such as primaquine (used in conjunction with chloroquine to treat *P. vivax* and *P. ovale* infections). The erythrocytic schizonticides are useful against the merozoites in the red blood cells. This stage is more amenable to drug therapy and the drugs include quinine and the 4-amino quinolines such as chloroquine (Roche *et al.*, 1989).

The most likely mechanism of action for chloroquine is that it interferes with the parasite's ability to polymerise ferriprotoporphyrin IX (which results from the degradation of haemoglobin), into haemozoin in food vacuoles. The unpolymerised ferriprotoporphyrin IX is toxic to the parasites (Borst and Ouellette, 1995; Collins and Paskewitz, 1995; Cowman, 1995). Since *P. falciparum* has the ability to develop

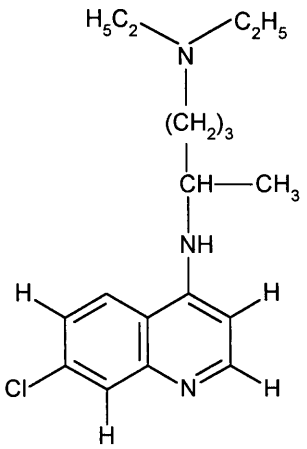
resistance to chloroquine and other molecules, compounds such as artemisinin (originally used in Chinese medicine and the only new class of antimalarial currently under development), mefloquine, halofantrine, doxycycline and sulphadoxine-pyrimethamine are used for the treatment and prophylaxis of this infection (Collins and Paskewitz, 1995; Department of Health of South Africa, 1996). Figure 1.2 shows the structures of some of the commonly used antimalarial drugs.

There are also schizonticides that are active against both life cycle stages. These are the drugs that inhibit folic acid synthesis (Roche *et al.*, 1989). DNA synthesis is then disrupted due to a limited supply of thymidylate (Basco *et al.*, 1995). They include derivatives of pyrimidine such as pyrimethamine and trimethoprim and the biguanides (chloroguanide and cycloguanil) (Roche *et al.*, 1989). Sulphonamides also inhibit DNA synthesis by inhibiting dihydropteroate synthase (DHPS). Antifolates (pyrimethamine) and sulphonamides (sulphadoxine) are often used together to deplete the folate pool (Borst and Ouellette, 1995).

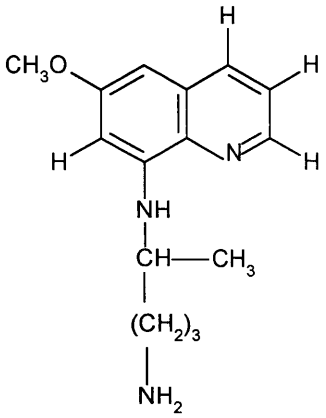
Resistance to chloroquine first arose in the 1960s in Southeast Asia and South America. It then spread over the rest of the world taking slightly longer to emerge in Africa, but by the end of the 1980s was widespread in Africa. Chloroquine resistance occurs when the parasite does not concentrate chloroquine in food vacuoles and therefore its ability to polymerise haeme is unaffected (Collins and Paskewitz, 1995). There are two potential mechanisms of chloroquine resistance. The first is the efflux of the drug and the second is alterations that decrease the uptake of the drug. These mechanisms may also act together. Chloroquine has the properties of a weak base. This means that at neutral pH it is uncharged and able to diffuse through membranes. Upon entering the acidic food vacuole (pH 5.2), the drug becomes double protonated and thus membrane impermeable. This results in the accumulation of the drug in the food vacuole. There is a P-glycoprotein homologue (Pgh 1) encoded by the *pfmdr1* gene which is mostly localised in the membrane of the food vacuole. This protein appears to be involved in the regulation of the pH in the food vacuole and may thus play a role in chloroquine accumulation.



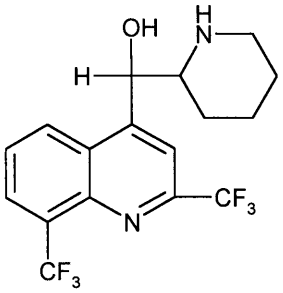
Quinine



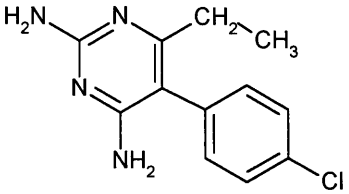
Chloroquine



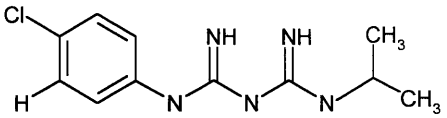
Primaquine



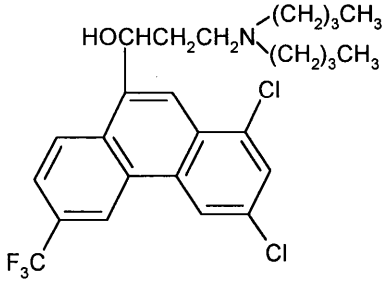
Mefloquine



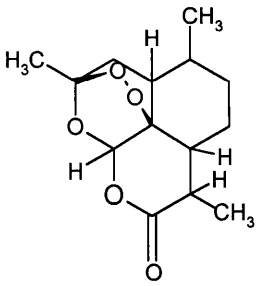
Pyrimethamine



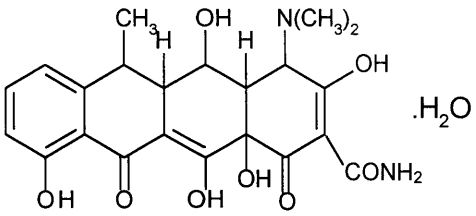
Proguanil



Halofantrine



Artemisinin



Doxycycline

Figure 1.2 Structures of the commonly used antimalarial drugs (Budavari, 1989; Roche *et al.*, 1989).

It was speculated that the product of the *pfmdr1* gene would be a multidrug transporter similar to that seen in tumour cells and would thus be involved with the efflux of the drug. This does not appear to be the case. Genetic crosses show that chloroquine resistance does not focus with the *pfmdr1* gene but rather maps to a 36kb region of chromosome 7 (Cowman, 1995). The *cg2* gene occurs in this segment and codes for an ~330kDa protein with complex polymorphisms (Su *et al.*, 1997).

Resistance to both antifolate drugs, pyrimethamine and proguanil, is widespread (Collins and Paskewitz, 1995). A recent study in Mpumalanga, South Africa showed that 43% of the single *P. falciparum* infections were parasite strains resistant to pyrimethamine (Birkholtz *et al.*, 1998). The molecular mechanism of antifolate drug resistance is due to mutations in the parasite gene for dihydrofolate reductase-thymidylate synthase (DHFR-TS). Mutations at codons 16, 51, 59, 108 or 164 result in an altered (yet still active) enzyme with a diminished affinity for the antifolate drugs. Pyrimethamine resistance is due to the following transition mutation Ser-108 (A**G**C) to Asn (A**A**C or A**C**C) The additional mutation Asn-51(A**A**T) to Ile (A**T**T) or Cys-59 (**T**GT) to Arg (**C**GT) improves the resistance. The Ala-16 (G**C**A) to Val (G**T**A) mutation can also occur and confers resistance to cycloguanil, but sensitivity to pyrimethamine, even if a S108N mutation is present. Combined mutations of S108N, N51I and C59R result in resistance to both pyrimethamine and cycloguanil. The added mutation Ile-164 (**A**TA) to Leu (**T**TA) results in the extremely high resistance generally found in South-east Asia. It is proposed that the multiple mutants arose from step wise selection from the S108N mutant (Cowman *et al.*, 1988; Petersen *et al.*, 1988; Basco *et al.*, 1995; Sirawaraporn *et al.*, 1997).

Since malaria is endemic to mostly poor Third World countries unable to afford expensive drugs, their efficacy as an avenue of prevention and treatment is dramatically reduced. This is especially so in situations requiring multiple doses of a compound (Fullick and Fullick, 1993). It is also difficult to organise the programs such that the entire country is involved at one time (Collins and Paskewitz, 1995).

As a result of this ever increasing resistance, novel targets and novel chemotherapeutic agents are being investigated (Kirby, 1997; Subbayya *et al.*, 1997). Various compounds from plants are being tested as possible chemotherapeutic agents, although potential synergism and antagonism can complicate matters (Kirby, 1997). An ideal drug target for the development of parasite specific inhibitors, is either unique to the parasite, or has properties which differ from those of the host. Enzymes of *P. falciparum* involved in glycolysis, nucleotide biosynthesis and haemoglobin catabolism are being investigated in several laboratories. Several of these are theoretically good targets and require further consideration. Proteins which are important for cytoadherence (such as PfEMP1), rosetting and invasion may be blocked and this could be another way in which parasite proliferation can be inhibited (Subbayya *et al.*, 1997).

1.5.2 Vector control

Vector control is an alternative method for reducing the incidence of malaria, although all methods of control are dependent on the behaviour and density of the mosquito vector in question (Collins and Besansky, 1994).

The *Anopheles* strains *A. gambiae*, *A. funestus* and *A. arabiensis* are the most prevalent malarial vectors present in Sub-Saharan Africa. These vectors are particularly efficient as they have a marked preference for human environments and humans as hosts. They are also capable of adapting rapidly to changes in the environment caused by human habitation and agriculture (Collins and Besansky, 1994). *A. gambiae* is the most important vector and attempts at genetic control are thus generally targeted at this species (Collins and Paskewitz, 1995).

Insecticides, larval control, biological agents, repellents and genetic control are all potential means of trying to prevent transmission of malaria through infective bites. Each method has advantages and disadvantages (Collins and Besansky, 1994).

Insecticides are a potent means of trying to rid the area of mosquitoes. The plasticity of the mosquito genome however allows resistance to develop (Collins and Besansky,

1994). Mosquitoes have developed resistance to most of the common insecticides such as DDT and dieldrin (Collins and Paskewitz, 1995). DDT has long been favoured, but this compound is toxic to more than just insects and the metabolites thereof tend to persist in the environment, with detrimental effects to many other organisms (Kelce *et al.*, 1995). Mosquitoes have also started to develop resistance to the newer insecticides such as organophosphates and carbamates. The emergence of resistance appears to be due to increased agricultural spraying (Collins and Paskewitz, 1995).

Insecticide impregnated bednets have also been tested, although opinion on their long term sustainability has been questioned. They appear to reduce the incidence of infective mosquito bites and are thus a valuable way of contributing to vector control (Collins and Paskewitz, 1995; Greenwood, 1997; Greenwood, 1997; Kwiatkowski and Marsh, 1997). However, the use of residual insecticides contributes to insecticide resistance (Collins and Paskewitz, 1995).

Synthetic and natural repellents are available. The latter include citronella products and extracts from lemon eucalyptus, although their effectiveness is fairly limited as the protective effects are non-lasting (Collins and Paskewitz, 1995).

Larval control of the *Anopheles* mosquito depends largely on the species (Collins and Paskewitz, 1995). Some such as *A. funestus* breed in relatively stable pools of water which can be fairly easily controlled with insecticides or larvicidal oils. The *A. gambiae* and *A. arabiensis* species breed in temporary pools and are thus much more difficult to eradicate (Collins and Besansky, 1994; Collins and Paskewitz, 1995). Several biological agents have been tested but only one larvivorous fish species (*Gambusia affinis*) showed any effectiveness. However, its use is disadvantageous due to competition with indigenous fish species. The endospore toxins from two *Bacillus* species are another possibility but are not really feasible due to the high costs and lack of residual activity (Collins and Paskewitz, 1995).

Genetic control is a more complex and rather controversial topic. The main concept is to replace the natural mosquito population with one that is more refractory to parasites

and thereby reduce the incidence of malaria infection (Collins and Besansky, 1994). Refractoriness of the mosquito to parasites may be obtained in various ways. One way is to manipulate the vector-parasite interrelationship such that the parasite life cycle can not be completed (Billingsley and Sinden, 1997). This can be done by exploiting the natural defences of mosquitoes. One strain of mosquito encapsulates the oocytes, while another lyses them shortly after penetration of the midgut epithelial cells (Collins and Besansky, 1994). Abnormalities in haemoglobin digestion which leads to interference with the formation of ookinetes or their penetration of the midgut can also be exploited (Curtis, 1994). Alternatively to the vector-parasite relationship, a gene for transmission blocking antibodies, could be placed in the mosquito genome (Billingsley and Sinden, 1997). Ideally, several of these components should be used together to prevent resistance. The phenotype should also be controlled by the minimum number of genes (Curtis, 1994). The requirements for mosquito-parasite specificity require further investigation, but offer vast potential for new control strategies (Billingsley and Sinden, 1997). The advantage of genetic control is that the mosquito population is not subject to rapid rebound from immigration or reproduction, which occurs after insecticide usage (Collins and Besansky, 1994). The disadvantage of genetic control is that the ability of the mosquito to serve as a host for the parasite (vector competence) probably contributes relatively little to the force of transmission of the parasite, i.e. the mosquito does not always become infected after feeding on an infected person. The ethics of the situation should be considered since mosquitoes are a pest and above the nuisance and irritation of being bitten, are capable of transmitting other parasitic and viral diseases (Spielman, 1994).

The most useful method of combating the mosquito vector would most probably require a combination of genetic control and judicious insecticide use (Collins and Besansky, 1994).

1.5.3 The elusive malaria vaccine

Attempts have been made at immunisation since 1936 and to date no vaccine has proved entirely effective (Alonso *et al.*, 1995). Over the past 30 years, considerable

research and several hundred million dollars have already gone into the search for a malaria vaccine (Collins and Paskewitz, 1995). Vaccination is the priming of the immune system prior to contact with the disease causing agent, in an attempt to provide the body with a method to combat and neutralise the parasite on infection (Fullick and Fullick, 1993). Experiments where volunteers were inoculated with irradiated sporozoites have shown that a malaria vaccine is feasible, as protection against malaria challenge was conferred on volunteers for up to 9 months (Kwiatkowski and Marsh, 1997).

There are three different life cycle stages providing targets for a malarial vaccine, namely the sporozoite (pre-erythrocytic), the asexual erythrocytic and the sexual stages (Brown and Nossal, 1986; Doolan and Hoffman, 1997).

The aim of a pre-erythrocytic vaccine would be to prevent the sporozoites from entering the hepatocytes and/or to try and inhibit their proliferation in the hepatocytes. If this could be done there would be no clinical manifestations of the disease and no transmission (Doolan and Hoffman, 1997). This type of vaccine would be ideal for non-immune temporary visitors such as tourists and military personnel (Alonso *et al.*, 1995; Collins and Paskewitz, 1995). An erythrocytic vaccine would not prevent hepatocyte infection, but would reduce the illness and death of the individual and would be of much more significant value to the permanent residents of a malaria endemic area (Alonso *et al.*, 1995; Collins and Paskewitz, 1995; Doolan and Hoffman, 1997). The pathology of the disease could also be targeted at this stage since soluble parasite derived products are released which may be used as antigens. Such a vaccine would not reduce the parasite load, but could reduce the morbidity of the disease. A sexual stage vaccine would have little benefit for the individual receiving it but would greatly affect the transmission of the disease and thus decrease morbidity and mortality (Doolan and Hoffman, 1997). Laboratory trials have shown that a vaccine based on a postfertilization protein (Pfs 25) completely blocks transmission (Collins and Paskewitz, 1995).

Since it has been shown that immunity to malaria is extremely specific (both species and strain), the choice of antigen on which to base the vaccine is crucial. It is believed that a multi-stage approach would be best. This type of vaccine would incorporate antigens from various life cycle stages (and also possibly the various parasite species) in an attempt to try and reduce the number of parasites at all stages. This is advantageous as immune evasion by the parasite would not be as significant, when more than one antigen is being targeted. Both humoral and cellular-mediated immunity would most probably be required for optimal protection (Nussenzweig and Long, 1994; Doolan and Hoffman, 1997; Kwiatkowski and Marsh, 1997).

Table 1.1 indicates some of the antigens that have been considered as vaccine targets along with the associated life cycles stages.

There are various strategies being employed in the quest for a malaria vaccine. These include DNA vaccines, protein particle vaccines and subunit vaccines (Doolan and Hoffman, 1997; Gilbert and Hill, 1997; Kwiatkowski and Marsh, 1997).

Some of the advantages of DNA vaccines are that they are easy to produce, the 'cold chain' may not be required, they are highly immunogenic and can easily incorporate more than one antigen, thus making utilisation of both the humoral and cellular facets of the immune system possible. However, there is a concern that the probable random integration into the host genome may result in the activation of oncogenes (Doolan and Hoffman, 1997). The use of DNA vaccines in mouse models has yielded some very encouraging results (Gilbert and Hill, 1997). Priming of mice with plasmid DNA (coding for a subunit vaccine of pre-erythrocytic antigens of *P.berghei)*, followed by a boost with recombinant, modified vaccinia virus Ankara, induced unprecedented, complete protection against *P.berghei* sporozoite challenge. It appears that this response is mediated by $CD8^+$ T-cells (Schneider *et al.*, 1998).

Table 1.1 Life cycle stages and target antigens for vaccine development (Doolan and Hoffman, 1997; Kwiatkowski and Marsh, 1997; Mons *et al.*, 1997).

Vaccine target	Examples of candidate antigens
Pre-erythrocytic stage	
Invasion of hepatocyte	CSP (Circumsporozoite protein)
(Antigens present on surface of sporozoite)	SSP2/TRAP (Thrombospondin-related anonymous protein)
	STARP (Sporozoite Thr & Asn rich protein)
	SALSA (Sporozoite and liver stage antigen)
Development within hepatocyte	CSP
(Peptides presented in complex with MHC recognised by $CD4^+$ and $CD8^+$ T-cells)	SSP2/TRAP
	LSA-1 (Liver stage antigen)
	SALSA
	EXP-1 (Exported protein)
	LSA-3
	MSP-1 (Merozoite surface protein)
Erythrocytic Stage	
Invasion of erythrocyte	MSP-1
(Antigens present on surface of merozoite)	MSP-2
	MSP-3
	EBA-175 (Erythrocyte binding antigen)
	AMA-1 (Apical membrane antigen)
	RAP-1 (Rhoptry associated protein)
Development within erythrocyte	PfEMP-1 (*P. falciparum* erythrocyte membrane protein)
(Antigens on or below surface of erythrocyte)	RESA (Ring-infected erythrocyte surface antigen)
	HRP2 (Histidine rich protein)
	SERA (Serine repeat antigen)
Parasite toxins	TNF inducing malarial toxins
Sexual stage	
Prefertilization	Pfg230 (Gametocyte antigen)
	Pfg48/45
	Pfg27/25
	Pfs16
Postfertilization	Pfs25 (Postfertilization antigen)
	Pfs28

Protein particle vaccines exploit the ability of viral coat proteins to assemble into particles spontaneously. They are advantageous as they are highly immunogenic. Hepatitis B surface antigen (HBsAg) and yeast derived Ty-like particles are both suitable as they retain particle-forming capability, even in the presence of long N- or C-terminal extensions that are from another organism. A vaccine consisting of 16 repeats of the circumsporozoite protein (NANP) was fused to the N-terminal of HBsAg. When used in conjunction with a new adjuvant (consisting of QS21, a saponin

from the bark of the *Quillaja saponaria* tree and 3-deacylated monophosphoryl lipid A, in an oil-in-water emulsion), good immunity was achieved (Gilbert and Hill, 1997).

One vaccine, Spf66, has reached phase III clinical trials. It is a synthetic, chimaeric molecule consisting of four epitopes, namely a repeat domain of the circumsporozoite protein (NANP) along with three blood stage antigens including MSP-1, of the Colombian malaria parasite. The vaccine afforded between 38.8% and 60.2% protection against *P. falciparum* malaria in South America. The protection in Africa varied from 30% in Tanzania to no significant protection in The Gambia (Amador and Patarroyo, 1996; Greenwood, 1997). Some researchers concluded that the vaccine confers no protection at all (Gilbert and Hill, 1997).

Field testing of vaccine candidates (especially blood stage vaccines) is rather restrictive, as it is difficult to monitor if the vaccine is reducing morbidity, since drug treatment is given as soon as parasitaemia rises. Secondly, the lack of logistical and scientific infrastructure in malaria areas is not conducive to the intensive health surveillance required for malaria field trials (Alonso *et al.*, 1995; Kwiatkowski and Marsh, 1997). There are also no immunological markers of protection in people in endemic malaria areas (Alonso *et al.*, 1995; Krishna, 1997).

The main problems involved with vaccine development are the identification of the appropriate antigens, the polymorphism of many of these antigens (antigenic variation) and the delivery of the antigens to the immune system (identification of non-toxic adjuvants) (Gilbert and Hill, 1997; Krishna, 1997).

A combination of drugs, vector control and vaccines will most probably be required for the successful control of malaria (Krogstad, 1996; Krishna, 1997). Since effective vector control and vaccines are far from reality, there is a vast requirement for alternative control methods. These may be in the form of new drugs for current targets, the modification of current drugs or the identification of entirely new targets (Olliaro and Yuthavong, 1998). Various parasite processes, such as metabolism, rosetting, cytoadherence, and erythrocyte invasion, can be targeted to inhibit parasite

proliferation. Of these, parasite metabolism offers numerous and diverse targets. In fact, most current drugs are aimed at certain aspects of parasite metabolism (Subbayya *et al.*, 1997; Olliaro and Yuthavong, 1998).

1.6 Metabolic processes of the genus *Plasmodium*

Since the advent of the technique to culture *P. falciparum* continuously *in vitro*, knowledge of many aspects of parasite development in the erythrocyte has progressed rapidly (Roth, 1990).

The metabolism of the malaria parasite requires that it modifies its host cell. Parasite metabolism includes carbohydrate, nucleotide, lipid and protein metabolism. Oxidative stress is generated and needs to be dealt with. The parasite has to induce a variety of changes in the human erythrocyte to enable it to survive. Figure 1.3 is a schematic representation of the host-parasite interrelationship (Ginsburg, 1990a).

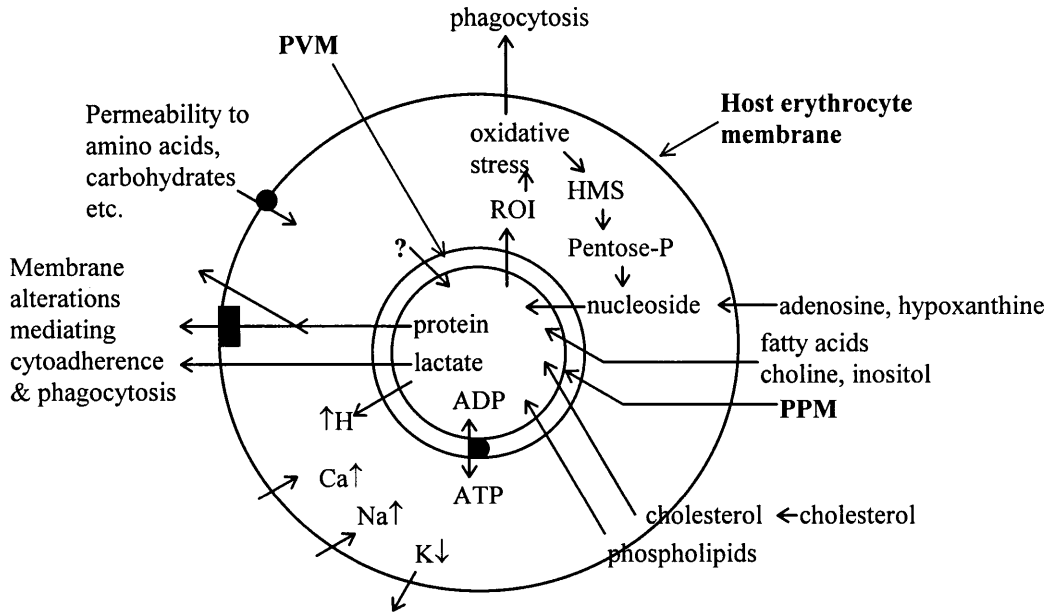


Figure 1.3 Relationship between the malaria parasite and the host erythrocyte. The question mark symbolises unknown factors which the host cell's metabolism may supply to the parasite. HMS is the hexose monophosphate shunt and ROI are reactive oxygen intermediates. PPM is the parasite plasma membrane and PVM is the parasitophorous vacuolar membrane. (Adapted from Ginsburg, 1990a).

During its 48 hour life cycle within the erythrocyte, *P. falciparum* increases its volume by as much as 25-fold and divides asexually. This rapid growth demands large

quantities of nutrients. This nutrient gathering is complicated by the relative ionic impermeability and low metabolism of the host cell. Direct passage of small nutrients across the two membranes separating the parasite from the host cytoplasm has been demonstrated for radiolabelled glucose, amino acids and purines (Desai *et al.*, 1993). *In vitro* cultured bloodstage parasites can only survive in a basal medium if the medium is supplemented with calcium pantothenate, cysteine, glutamate, glutamine, isoleucine, methionine, tryptophan and tyrosine (Divo *et al.*, 1985).

The original biochemical investigations of the metabolism of *Plasmodium* species involved the study of carbohydrates. This was not unusual as carbohydrate metabolism was the focal point of research in yeast, muscle and the liver (Sherman, 1979).

The host erythrocyte and the asexual malaria parasite have no carbohydrate reserves (glycogen, polysaccharides). As a result, both cells are entirely dependent on a continuous supply of glucose, as their entire energy (ATP) supply is derived from the metabolism of glucose through glycolysis (Sherman, 1988; Elford *et al.*, 1995). Glucose is thus a key carbon source for both cells. Another characteristic of human erythrocytes and the *P. falciparum* parasites that exploit them, is that they both lack a functional tricarboxylic acid cycle. This is the main reason why they depend entirely on the production of ATP via glycolysis (Dobeli *et al.*, 1990). The major product of glycolysis is L-lactate, although there are enzymes available for the production of D-lactate (Bryant and Behm, 1989; Vander Jagt *et al.*, 1990).

The uninfected erythrocyte consumes 1-2 mM glucose an hour per litre of cells (Roth, 1990). Glucose uptake in infected erythrocytes has been shown to be increased up to 100 times that of uninfected cells (Read *et al.*, 1994). This explains why patients with a high parasitaemia are prone to hypoglycaemia (Roth, 1990). The increase in glycolysis has been demonstrated to result from parasite derived glycolytic enzymes (Read *et al.*, 1994). At least 80% of glucose utilisation in infected blood cells is accounted for by the activity of the parasite's glycolytic enzymes. Although there is an increased demand for glucose, there is a high rate of constitutive influx, indicating that it is unlikely that glucose transport could be rate limiting (Elford *et al.*, 1995).

There is an energy connection between the parasite and its host. Both Kanaani and Ginsburg (1989) and Choi and Mikkelsen (1990) have reported the presence of an ADP/ATP transporter present on the parasite membrane. However, the authors disagree about the direction of ATP transport since Kanaani and Ginsburg suggest that the parasite provides the host cell with ATP while Choi and Mikkelsen suggest that the host supplies the parasite with ATP.

Since the parasite has an increased requirement for nutrients, transport systems for such nutrients provide potential targets. There are a variety of such nutrient transporters and these include the purine, glucose, lactate, nucleoside and amino acid transporters. The glucose transporter, an ideal target (Louw, 1998), has been identified in our laboratory by K. Nel (article submitted) and is being characterised (K. Nel, personal communication). Since Oxender and colleagues (Su *et al.*, 1995) showed that the human System L transporter (which transports leucine) at the blood brain barrier can transport the drug gabapentin, amino acid transporters seemed the next logical choice for a target capable of either antimalarial drug transport, or of being blocked to prevent nutrient access and thus parasite proliferation.

1.6.1 Parasite protein metabolism

The malaria parasite obtains amino acids by three mechanisms. Firstly, it is capable of synthesising alanine, leucine, glutamic acid and aspartic acid by fixation of CO_2 (Sherman, 1979; Kolakovich *et al.*, 1997). However, most of these amino acids are not incorporated into parasite proteins. It may be that some of the amino acids are needed for reductive metabolism, which uses glutamic acid and glutathione (Ginsburg and Stein, 1987). Secondly, it transports amino acids from the host serum. This is especially true for isoleucine which is rapidly incorporated into parasite proteins. However, the parasite obtains most of its amino acids from its degradation of haemoglobin (Sherman, 1979; Kolakovich *et al.*, 1997).

In contrast to the ring stage of the parasite, which has minimal metabolic activity, the trophozoite stage is very active. Haemoglobin is present in the erythrocyte at a

concentration of 5 mM and constitutes approximately 95% of the cytosolic protein. The host cytoplasm is ingested by means of a cytostome (invagination of the parasitophorous vacuolar membrane and the parasite plasma membrane). This cytostome fuses with the acidic digestion (food) vacuole, in which the haemoglobin is degraded by means of acidic proteases. There are two types of proteases, namely the aspartic proteases (Plasmepsins I and II) which account for between 60% and 80% of the haemoglobin degrading activity, and a cysteine protease (Falcipain) which accounts for 20% to 40%. The mechanism of haemoglobin digestion by the parasite is highly ordered. The aspartic proteases cleave in the hinge region (the domain responsible for holding the tetramer together when oxygen is bound), thus unravelling the haemoglobin and allowing further cleavages by the aspartic and cysteine proteases. The haemoglobin is broken down to smaller fragments in the food vacuole, but it is suspected that catabolism to amino acids occurs in the parasite's cytoplasm, as no exopeptidase activity has been detected in the food vacuole of *P. falciparum* (Francis *et al.*, 1997). This suggests that a peptide translocator is present in the membrane of the food vacuole, which can transport peptides into the cytoplasm (Kolakovich *et al.*, 1997). Since haemoglobin is a poor source of methionine, cysteine, glutamine, glutamic acid and contains no isoleucine, these amino acids are required exogenously in continuous culture to ensure parasite propagation. Both the amino acids that are absent and those that are present abundantly in haemoglobin are taken up exogenously. The amino acids are subsequently used in the synthesis of parasite proteins and can also be used as an energy source. Since the parasite degrades more haemoglobin than required, there may also be some diffusion of amino acids into the host serum (Francis *et al.*, 1997). The parasite polymerises the haeme to haemozoin (an inert polymer). It is speculated that the parasite's histidine rich proteins may play a role in the polymerisation of haeme. The degradation of haemoglobin is shown in Figure 1.4.

Haemoglobin degradation places oxidative stress on the digestive vacuole. This is relieved by superoxide dismutase, catalase and reduced glutathione. These molecules may possibly be obtained from the host cell on ingestion of the cytoplasm. Superoxide dismutase may be parasite derived (Francis *et al.*, 1997).

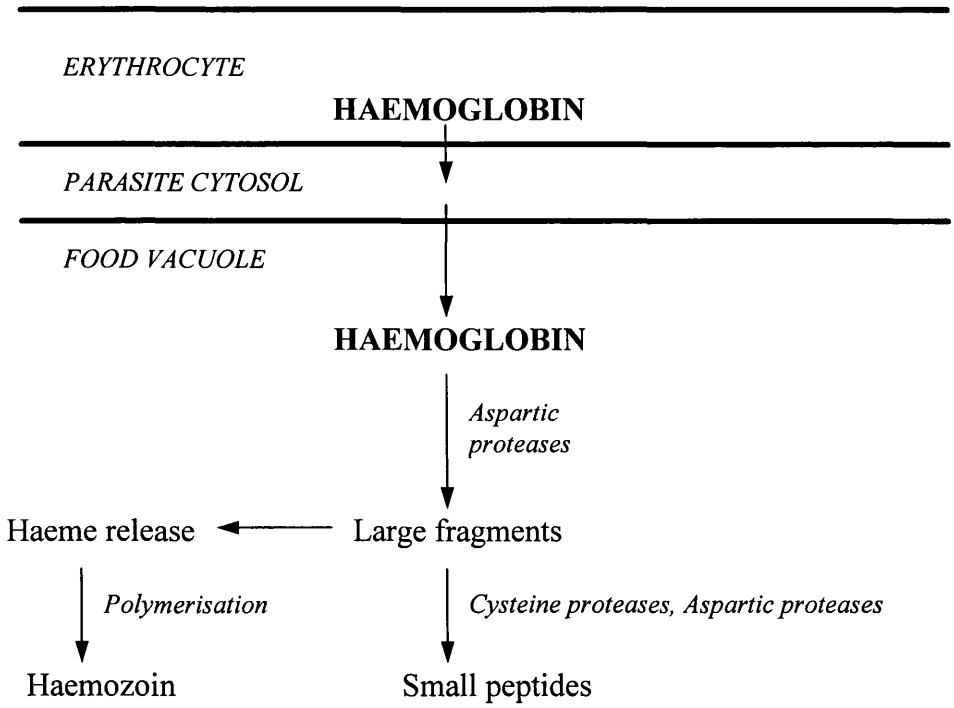


Figure 1.4 Schematic representation of haemoglobin degradation (Francis *et al.*, 1997; Subbayya *et al.*, 1997).

Haemoglobin metabolism is a current target for chemotherapeutic intervention. Compounds such as methylene blue, artemisinin, and the quinolines act on the haeme that is released after haemoglobin degradation, although their mechanisms may differ. Artemisinin appears to act by membrane damage as a result of protein alkylation, while the quinolines act by inhibiting haemozoin formation (Francis *et al.*, 1997).

Degradation may also be a target for novel therapy by using protease inhibitors to prevent haemoglobin degradation. The aspartic proteases (plasmepsins), the cysteine proteases or the processing required for plasmepsin activation, can be targeted. The cysteine protease can be inhibited by means of peptide fluoromethyl ketones such as E-64 (*L-trans*-epoxy-succinyl-leucyl-amido-(4-guanidino)butane). Peptidomimetic inhibitors are useful for the study of plasmepsins. The crystal structure of Plasmepsin II in complex with pepstatin is available which means that inhibitor identification will be much easier and more accurate (Francis *et al.*, 1997; Subbayya *et al.*, 1997).

1.7 Nutrient acquisition by *Plasmodium*

Since the plasma membrane is a selectively permeable barrier between the cell and the extracellular environment, essential molecules such as glucose, amino acids and lipids readily enter the cell, metabolic intermediates remain in the cell and waste products leave the cell. The cell can thus maintain a constant internal environment (Zubay, 1993).

Membrane transport can be divided into two main types, active and passive. Passive transport involves diffusion, either simple or facilitated. Simple diffusion would be direct diffusion of a small, relatively hydrophobic molecule across the membrane, whereas facilitated diffusion would be mediated by a protein. Active transport involves the expenditure of metabolic energy and molecules are generally transported against their concentration gradient (Zubay, 1993).

The extremely metabolically active malaria parasite has to function in a metabolically quiescent host cell, which has an outer membrane that is relatively impermeable to ions (Ginsburg, 1990b; Desai *et al.*, 1993). The membrane structures present between the parasite and the external medium are the host plasma membrane (EM), the parasite vacuolar membrane (PVM, formed during merozoite invasion) and the parasite plasma membrane (PPM) (Deitsch and Wellems, 1996; Elford *et al.*, 1997). The parasite vacuolar membrane extends towards the erythrocyte membrane to form a network known as the tubovesicular membrane (TVM) network (Deitsch and Wellems, 1996). There are three main ways that the malaria parasite in the infected erythrocyte can obtain nutrients from the extracellular environment:

- It can obtain small molecules by permeability pathways that it induces in the erythrocyte membrane;
- It gathers them directly from the extracellular environment. The way in which the parasite does this is still being debated. It may use the parasitophorous duct or it may use its TVM network in directing transport of nutrients across the erythrocyte cytosol.

- Lastly the parasite obtains macromolecules such as haemoglobin from the host cytosol by means of a cytostome (Elford *et al.*, 1997; Goodyer *et al.*, 1997).

These three methods of nutrient acquisition are explained in greater detail below.

1.7.1 Parasite membrane permeability

The parasite is responsible for increased permeation of carbohydrates, amino acids, purine nucleosides, complex lipids, anions and cations, through the three membranes to the parasite cytoplasm (Cox, 1993; Deitsch and Wellems, 1996).

Thirty years ago it was shown that parasitisation resulted in increased permeability of the erythrocyte to amino acids (Sherman, 1988). Since then numerous experiments have shown that for *P. falciparum* to grow in continuous culture, it required the addition of isoleucine, cysteine, tyrosine, glutamate, glutamine, methionine and proline (Divo *et al.*, 1985; Ginsburg and Stein, 1987; Elford *et al.*, 1995; Olliaro and Goldberg, 1995). Isoleucine is probably required as it is absent in haemoglobin, while cysteine, glutamate and glycine are needed for the reductive metabolism of the parasite which requires glutathione. Since the degradation of 1 mole of haemoglobin monomer results in 143 moles of amino acids, efflux of excess amino acids must also occur (Ginsburg and Stein, 1987). If this did not happen then the high concentration of amino acids would draw water into the compartments resulting in the osmotic lysis of these compartments (Ginsburg, 1994).

It seems that upon infection, the host cell loses its energy coupled amino acid transport systems and that free amino acids move along a concentration gradient into the erythrocyte (Sherman, 1988). It is evident that changes occur during parasite maturation which affect most of the inducible and regulated transport systems and also result in changes to the host's metabolism (Prasad, 1997). Whether the amino acids enter the cells by diffusion or through a carrier mediated system has not yet been determined (Cox, 1993).

The parasite obtains nutrients by increasing the permeability of the host cell membrane. This permeability can be induced by a couple of factors. Firstly, the parasite is incapable of synthesising cholesterol. A loss of cholesterol in the cell membrane results in increased fluidity. Secondly, the parasite incorporates polypeptides into the host membrane. It is possible that some of these may be transport proteins. Permeabilisation applies to both substrates (glucose, amino acids etc.) and waste products (lactate). It should occur without influencing the osmotic or physical integrity of the host (Ginsburg, 1990b). It seems that this permeabilisation starts six hours after infection and that it increases in parallel with the development of the parasite, with a maximum reached at the trophozoite stage (Breuer *et al.*, 1987; Ginsburg, 1990b). The transport kinetics are unaltered but the number of transport pathways increases (Ginsburg, 1990b). Two pathways appear to be present, saturable and non-saturable. One is probably the activated host pathway and the other would be the permeation pathway induced by the parasite (Ginsburg, 1994).

The characteristics of the permeation pathway are as follows. It has a broad specificity, although it is strongly selective for anions. It transports neutral molecules fairly well and weakly transports some cations (Gero and Kirk, 1994; Elford *et al.*, 1995; Goodyer *et al.*, 1997). The pathways have a high capacity, show little competition between substrates and are non-saturable (Elford *et al.*, 1995). It is size specific since monosaccharides, amino acids and ions can pass through but disaccharides cannot (Gero and Kirk, 1994; Goodyer *et al.*, 1997). The size limit appears to be less than 200 Da (Goodyer *et al.*, 1997) and the permeability pathways behave as pores in the erythrocyte membrane, with a radius of 7nm (Breuer *et al.*, 1987). The parasite derived pathways differ from those of the host erythrocyte, both pharmacologically and kinetically (Gero and Kirk, 1994). The pathways can be inhibited by various inhibitors including phloridzin (or the aglycon, phloretin), niflumic acid and cyanocinnamate derivatives (Breuer *et al.*, 1987; Ginsburg and Stein, 1987; Ginsburg, 1990b). Cations are mostly excluded from permeation and this most probably ensures the osmotic integrity of the host erythrocyte while the parasite matures (Cabantchik, 1990).

Desai *et al* (Desai *et al.*, 1993; Desai and Rosenberg, 1997) investigated a 140-pS ion channel present in the PVM. This channel is permeable to monosaccharides, amino acids and purine nucleotides, is present in high density and is primarily open. Their experiments indicated that this channel has negligible cation or anion selectivity and that permeation does not involve a binding step. The researchers suggested that the channel can be modelled as a rigid, cylindrical sieve through which soluble macromolecules of up to 1400Da can pass, if they have a size in aqueous solution of less than 23 angstrom. These molecules will probably only be limited by their diffusion coefficient in the cytosol and their size. Desai *et al* thus suggested that the function of the pore is to provide access to nutrients in the erythrocyte cytosol. It is likely that diffusion through this channel is not rate limiting and that the channel only spans the PVM and not the parasite plasma membrane as well.

It thus seems likely that the nutrients the parasite requires gain access to the erythrocyte cytoplasm via the parasite induced permeability pathways. They then traverse the PVM via the ion channel described by Desai and colleagues. Transport across the PPM is most likely the specific step (Penny *et al.*, 1998).

One of the problems encountered in the investigation of parasite derived transport mechanisms is that it is difficult to assign a transport process to a particular membrane as there are three membranes in the infected erythrocyte. Ansorge *et al* (Ansorge *et al.*, 1997) recently managed to access the PVM, using Streptolysin O, which is a bacterial pore forming protein which permeablises lipid bilayers by forming proteinaceous pores. These pores are larger than 30 nm and do not alter the membrane protein orientation. With selective permeabilisation of the erythrocyte membrane, they were able to determine the membrane topology of some of the proteins present on the PVM. The PVM is difficult to work with as the purification of the PVM or the isolation of the parasites in the vacuole, tends to disrupt the membrane.

1.7.2 Macromolecular transport

Macromolecular transport is mediated by the cytostome and/or by another, still unclarified and disputed pathway, involving the parasitophorous duct.

As mentioned previously, the cytostome is an endocytotic type of vesicle that forms from the PVM and the PPM. A double membrane section pinches off and fuses with the food vacuole where haemoglobin degradation occurs (Figure 1.5). This process is not microtubule mediated (Goodyer *et al.*, 1997).

Whether or not the parasite has direct access to the extracellular milieu and if it has, whether it is by means of a metabolic window, or a parasitophorous duct (Figure 1.5) is a controversial issue (Elford *et al.*, 1995). The parasitophorous duct could be a direct connection to the external medium formed by continuation of the erythrocyte membrane and the PVM, while the metabolic window may result from the close apposition of the PVM and PPM next to the erythrocyte membrane (Gero and Kirk, 1994).

To date there is no conclusive evidence that the parasitophorous duct exists. There is experimental data which tries to prove it but there is also data which refutes it (Goodyer *et al.*, 1997; Hibbs *et al.*, 1997). One of the theoretical problems encountered here is the fact that the duct cannot coexist with the permeability channels (Gero and Kirk, 1994; Ginsburg, 1994). If however, it could be proven that some form of direct access to the external environment does occur, it could have a significant impact on malaria control as the internal antigens present on the parasite could be used for vaccine development since the humoral arm of the immune system would have access to the parasite through this 'window', thus bypassing the erythrocyte cytosol.

Lauer and colleagues (Lauer *et al.*, 1997) have recently shown that the TVM is a transport network that allows the parasite access to nutrients (amino acids, nucleosides) in the external environment (Figure 1.5) and they speculate that this could be an efficient means of delivering antimalarial drugs directly to the parasite. The

TVM network is distinct from a duct as it is not continuous with the erythrocyte membrane and it does not allow macromolecules to diffuse into infected erythrocytes. It is likely that this means of transport occurs only later in the maturation of the parasite as the TVM is formed approximately 33 hours after infection. This suggests that it may be a means of helping the parasite to obtain the extra nutrients that it requires as it matures, which may not be provided by the other transport mechanisms (permeability pathways and cytostomes). Vesicles may bud off and these may take parasite derived proteins to the erythrocyte surface (Lauer *et al.*, 1997).

There appears to be sufficient evidence that a means of direct access to the extracellular milieu is possible, although agreement has not been reached on what form this takes. It could be a transient duct, a metabolic window or it could be via a network created by the tubovesicular membrane.

A model of probable transport pathways in *Plasmodium falciparum* infected erythrocytes is shown in Figure 1.5. It shows the non-saturable pore on the erythrocyte membrane, the 140pS channel of broad substrate specificity on the PVM and the putative specific transporters on the PPM. It also represents the possible means of macromolecular transport, namely the tubovesicular membrane network and the parasitophorous duct (Penny *et al.*, 1998).

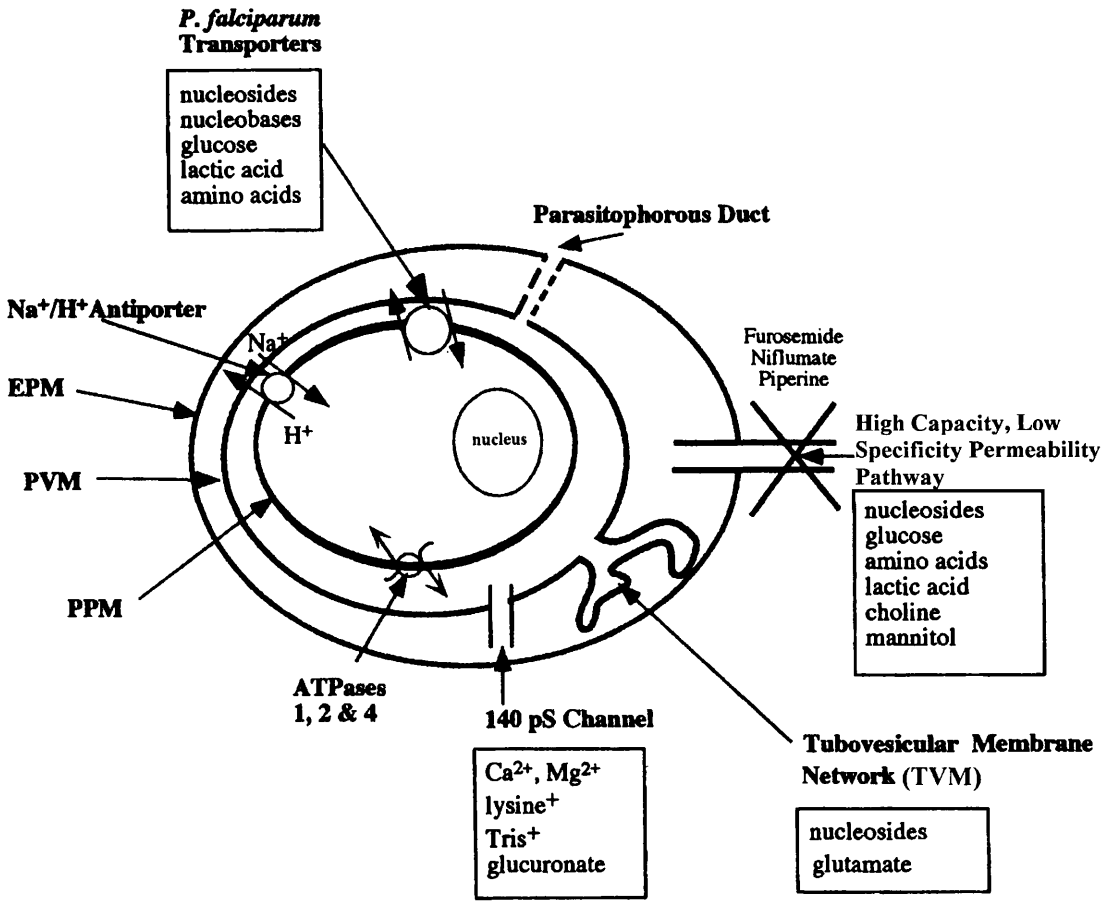


Figure 1.5 Model of transport pathways in malaria-infected erythrocytes. EPM is the erythrocyte plasma membrane, PVM is the parasitophorous vacuolar membrane and PPM is the parasite plasma membrane. (Adapted from Penny, 1998).

1.7.3 Amino acid transport

Amino acids are required by living cells to sustain a variety of metabolic activities. These include functions such as glutathione metabolism, protein synthesis, osmoregulation, donation or transport of carbon or nitrogen, as components of buffers systems, and providing a source of metabolic energy (Harvey and Ellory, 1989; Guidotti and Gazzola, 1992).

Amino acid transport systems have been a topic of study for many years and a vast quantity of research has been undertaken. Certain general characteristics of these systems have been identified. They are highly stereospecific and have a low substrate specificity, which means that one system is capable of transporting more than one amino acid. A cell will commonly also have more than one transport system and some

of these will have overlapping specificities (McGivan and Pastor-Anglada, 1994). Characterisation is determined on the basis of kinetic and physiologic studies (Guidotti and Gazzola, 1992). These transport systems fall into two categories, the sodium independent systems (uniport) and the sodium dependent systems (McGivan and Pastor-Anglada, 1994).

As erythrocytes of various organisms differ, so do the amino acid transport systems which show heterogeneity between various cell types (Harvey and Ellory, 1989; Guidotti and Gazzola, 1992).

Table 1.2 contains the five main amino acid transport systems present in the human red blood cell.

Table 1.2 The transport systems present in the human erythrocyte and some of their characteristics (Harvey and Ellory, 1989; Souba and Pacitti, 1992; McGivan and Pastor-Anglada, 1994).

System	Amino acid specificity	Sodium requirement	Characteristics
L	Neutral (branched chain and aromatic)	Independent	Highest capacity transporter in RBC
T	aromatic-Tyr, Trp, Phe	Independent	Major route for tryptophan at physiological concentrations
y+	Lys, Arg, Ornithine	Independent	Highly stereoselective, high affinity for substrate. Regulated by TNF
ASC	Ala, Ser, Cys	Dependent	Exceptional stereosensitivity. Slightly less pH sensitive. Principle route for Cys, present in RBC instead of system A. Regulated by cytokines
Gly	Gly	Dependent	More than 1 Na^+ binds per amino acid, absolute requirement for Cl^-

Since the erythrocyte is metabolically quiescent it seems unusual that it should possess these transport systems. Several reasons have been proposed for this. Firstly and most likely, is that they are functional relics of the reticulocyte stage, which has a huge requirement for amino acids. Secondly, the erythrocyte is actively synthesising glutathione and thus has a requirement for specifically the three amino acids cysteine, glutamate and glycine. A theory has also been postulated that the erythrocyte can

function as an interorgan transport system of amino acids (Harvey and Ellory, 1989). It appears that the permeability pathways induced by the parasite negate the erythrocytic transport pathways.

System L is also capable of amino acid antiport and may function physiologically for amino acid efflux rather than uptake (McGivan and Pastor-Anglada, 1994), since the degradation of haemoglobin provides the parasite with an excess of amino acids and to maintain its osmolarity the parasite also needs a means of exporting amino acids.

1.8 Conclusion

Malaria, one of the scourges of mankind, is caused by a complex and resourceful parasite. To date, eradication has been successful in only a small part of the world. In most of the tropics, however, it is omnipresent and extremely debilitating. Although there are a variety of control methods, none has proved entirely effective. The insects and parasites have developed resistance to insecticides and drugs, respectively. In addition the search for a vaccine has been hampered by the ingenuity of the parasite in evading the immune system. Through a knowledge of the genetics and biochemistry of the parasite, potential targets may be identified. A greater number of well-defined targets will increase the chance of finding an effective means to combat malaria.

Parasite metabolism and specifically nutrient acquisition provides such a target. The intra-erythrocytic parasite requires additional amino acids to those obtained from haemoglobin digestion, suggesting the presence of an amino acid transporter. Such a transporter is an ideal therapeutic target as it can be used in one of two ways. Either the transport of nutrients into the infected erythrocyte can be blocked, or the transporter can be used for the enhanced delivery of cytotoxic drugs into the parasite.

Of the known amino acid transport systems, amino acid permeases were chosen for this study, since it appears that the amino acid permeases of lower eukaryotes belong to a unique family. The characteristics of this family are as follows: they have no apparent ATP-binding sites, are of similar length (550-628 amino acid residues) and have

almost identical hydrophobicity plots indicating similar structure and function. Topology predictions of the family indicate 12 α-helical, putative membrane spanning regions with the N- and C-terminals located cytoplasmically (Sophianopoulou and Diallinas, 1995).

1.9 Aim and strategy of the study

The strategy which was followed encompasses firstly the identification of a consensus region of the gene family coding for amino acid permeases, followed by amplification with an RT-PCR protocol known as RACE (Rapid Amplification of cDNA Ends). Problems with the application of RACE techniques to *Plasmodium* genes (AT-rich genome) were anticipated, as most experimental protocols have been formulated for human or bacterial applications. Since these organisms tend to have a GC-rich genome, in marked contrast to the AT-rich *Plasmodium* genome, these protocols cannot be directly applied. This is reflected by the lack of literature available for RT-PCR, differential display and related procedures in malaria studies. However, if such techniques can be applied to malaria genes, a wealth of sequence information may be obtained, as the RT-PCR procedures are generally much less labour intensive than alternative methods, the most common of which employs the use of an oligonucleotide probe to screen a cDNA library. However, for parasite genes of unknown sequence, the PCR based approach was adopted for two primary reasons. Firstly, as mentioned previously, the amount of time required to screen a cDNA library is prohibitive. Secondly, the identified internal stretch of consensus sequence on which the primer or probe is based, is usually too short and degenerate in nature, for the specific screening of a cDNA library, but is suitable as a PCR primer.

Chapter 2 describes the selection of a consensus sequence for amino acid permeases and the design of primers, while Chapter 3 describes the use of these primers for PCR and the cloning and sequencing of the products. Chapter 3 also elaborates on the problems encountered with the RACE techniques applied to malaria genes and how these problems were circumvented. Results from this study have been presented at an

international and two local congresses (Nyambati *et al.*, 1997; Language *et al.*, 1997 a; Language *et al.*, 1997 b; Clark *et al.*, 1998).

CHAPTER 2

Primer Design

2.1 Introduction

The initial phase of the study, was concerned with the identification and selection of a consensus sequence on which to base the design of a PCR primer, for use in a RACE protocol, for the identification of an amino acid transporter gene from *P. falciparum.*

Genes can be identified in a multitude of ways. An effectual way is through the use of oligonucleotides which may either be used for the screening (hybridisation) of a library (genomic DNA or cDNA) or in a variety of PCR-based techniques. The latter method includes traditional PCR using genomic DNA, or a vast selection of RT-PCR based techniques using RNA. These techniques have been available for a number of years. The PCR approach is usually attempted first as it is more sensitive and yields unequivocal results faster. The prerequisite for the application of this technique is that a small section of protein sequence information must be available. This can be achieved in one of two ways. Firstly and ideally, a section of the protein sequence should be available. If this is not possible, a homologous section of other organisms may be used. Although this requires more conjecture, it has been used successfully and a variety of parasite genes have been identified in this way, including the DHFR-TS, phosphoglycerate kinase, RNA polymerase large subunit and α- and β-tubulin genes of *P.falciparum.* Genes for *Trypanosoma brucei* and *Schistosoma mansoni* have been identified in a similar way (Hyde and Holloway, 1993).

The selection and design of primers has a large impact on the outcome and quality of the PCR (Rychlik, 1993; Schaefer, 1995). The chance of finding the 'ideal' primer pair is greatly enhanced if more DNA sequence information is available (Rychlik, 1993). Fortunately, with the World Wide Web (WWW) gaining popularity, biological

information is readily accessible. To this effect, Gilbert (1991) wrote a prescient article describing a paradigm shift with regard to experimental procedure and theoretical conjecture. He predicted that DNA sequence data would grow exponentially and that biologists would have to become increasingly computer literate to be able to keep up to date (Gilbert, 1991).

Today, there is a vast volume of information present in various databases, located globally and which may be accessed via the WWW. Some of the most useful and comprehensive databases include GenBank, SWISS-PROT and *Entrez*. GenBank is the National Institute of Health (NIH, Bethesda, USA) database and contains all the known nucleotide sequences. GenBank is maintained by the National Centre for Biotechnology Information (NCBI) and is part of the International Nucleotide Sequence Database Collaboration, thus ensuring that a comprehensive set of sequence information is maintained (Sander, 1996). SWISS-PROT is a database of protein sequences. It is cross-referenced to several other databases (Shomer *et al.*, 1996). *Entrez* was developed by the NCBI and is a molecular biology database and retrieval system. It is particularly useful in that it presents an integrated view of biomedical data and their interrelationships as they are cross-linked by intrinsic cross-reference information and computed relationships (nearest neighbours). Biological sequences, three dimensional structures or abstracts from scientific literature can be retrieved. These documents, in turn, can serve as entry points for further exploration. Biological sequence information is drawn from various databases, primarily GenBank (Schuler *et al.*, 1996).

Homology between sequences can be determined in several ways. Two of these are the Blocks database and CLUSTAL W. The Blocks database consists of protein Blocks which are obtained by multiple alignment (without gaps) of separate sequences consisting of the most conserved regions of protein families (Henikoff and Henikoff, 1996). CLUSTAL W is a programme that performs progressive alignments of sequences, which exploits the fact that sequences are phylogenetically related. (Higgins *et al.*, 1996). The 'W' in CLUSTAL W represents 'weighting', which assigns different weights to sequences and parameters at different positions in the alignment. It seems to

be more sensitive for difficult protein alignments (Higgins *et al.*, 1996). Once these resources have been exploited and a suitable consensus region identified, the PCR primer can be designed.

Rychlik has determined the requirements for good sequencing and PCR primers. The most important factors to be considered when selecting PCR primers are the following:

- they should be as specific as possible for the target sequence;
- not form any dimers with themselves or each other;
- not form hairpin loops and should
- form a stable duplex with the target sequence.

OLIGO 4.1 (National Biosciences Inc., MN, USA) is a multi-functional computer program designed by Rychlik and based on these critical parameters. It is capable of selecting primers, be they a new primer pair, or a primer which should be compatible with a current primer. It can also determine the parameters for working with DNA and RNA, although those for RNA are less accurate. It is also capable of selecting probes and determining hybridisation conditions (Estruch, 1997).

2.2 Methods

2.2.1 Identification and selection of a consensus sequence.

Amino acid sequences for aromatic, branched chain, threonine and tryptophan/tyrosine transporters from a variety of organisms (mostly bacterial), were obtained from the databases available. These sequences were multiply aligned using CLUSTAL W. Upon careful consideration of the alignment and comparison with two *Arabidopsis* amino acid permeases, a consensus sequence became apparent. To verify that the consensus sequence was present in both eukaryotic and prokaryotic organisms, and that it is indeed conserved, approximately 45 amino acid permeases were retrieved using *Entrez* and the consensus sequence identified in most of them (~90%). The consensus sequence was subsequently confirmed as part of an amino acid permease signature motif by the Blocks database.

2.2.2 Primer design

A poly-T primer consisting of $(dT)_{18}$ with several modifications was designed for the reverse transcription reaction. The primer contained the degenerate nucleotides VN (V represents G, C or A and N is any nucleotide) on the 3′-end, which acts as a clamp by fixing the site of annealing at the start of the poly-A tail (Thomas *et al.*, 1993). It also contained an 'anchor' sequence on the 5′ end which consisted of 21 random nucleotides. The anchor was to be the reverse primer for PCR.

A permease specific primer (PSP2), to be used as a forward primer in PCR was designed around the identified amino acid permease consensus sequence. The amino acid sequence was reverse translated to mRNA using the genetic code. The complexity was subsequently reduced using the codon preference of *Plasmodium falciparum* (Saul and Battistutta, 1988). The poly-T, anchor and PSP2 primers were obtained from Genosys Products (Cambridge, England).

Subsequent to confirmation of the consensus sequence from Blocks, a second permease specific primer (PSP) was designed. The degeneracy of this primer was reduced using the codon preference in the form of the Match Index (Hyde *et al.*, 1989). Inosine was also used at position 9. The PSP primer was obtained from MWG Biotech, Germany. A third primer (TYT) was obtained from MWG Biotech, Germany. It's design was based on the same permease consensus region, but the specific amino acids chosen reflected a specificity for tyrosine and tryptophan transporters.

The primers were evaluated using OLIGO 4.1 with five types of amino acid transporter sequences as templates. They were the aromatic, threonine, tyrosine/tryptophan and branched chain transporter sequences (mostly of bacterial origin). The general amino acid permease sequence from *Saccharomyces cerevisiae* was also used. The amino acid sequences of these proteins were obtained from various databases and were reverse translated using Seqaid II™ Version 3.81 (Centre for Basic Cancer Research, KS, USA). The codons were then adjusted to fit the codon bias of *P. falciparum*. The

results obtained for the templates were similar in all cases and the results obtained using the yeast template are shown.

2.3 Results

The first permease specific primer designed (PSP2) was based on the identified consensus sequence. This sequence was identified using the CLUSTAL W aligned amino acid transporter sequences from various organisms. The relevant portion of the alignment is shown in Figure 2.1.

```
p15993    --------MMEGQQHGEQLKRGLKNRHIQLIALGGSIGTGLFLGSASVIQSAGPGIILGY
p24207    MKNASTVSEDTASNQEPTLHRGLHNRHIQLIALGGAIGTGLFLGIGPAIQMAGPAVLLGY
ecarop    --------MMEGQQHGEQLKRGLKNRHIQLIALGGSIGTGLFLGSASVIQSAGPGIILGY
p19072    ------------MTHLKGFD----LLALGFMTFALFLGAGNIIFPPSAG--MAAGEHVWS
1dbrnqgn  ---------------MKEKLTHAESLTISSMLFGLFFGAGNLIFPAYLG--EASGANLWI
p14931    -----------MTHQLKSRD----IIALAFMTFALFVGRGNIIFPPMVG--LQAGEHVWT
p11867    ----MSTSDSIVSSQTKQSSWRKSDTTWTLGLFGTAIGAGVLFFPIRAGFGGLIPILLML
p36559    ---METTQTSTIASKFSRSAWRKTDTMWMLGLYGTAIGAGVLFLPINAGVGGMIPLIIMA
heahi0289------------MKSTEKLKWNKFDATWMLNLFGTAVGAGVLFLPINAGMGGFWPLVLMA
p18199    ---------------MKNRT-----LGSVFIVAGTTIGAGMLAMPLAAAGVGFSVTLILL
heahi0528-------------MLKNKT-----FGSALIIAGTTIGAGMLAMPLTSAGMGFGYTLLLL
heahi0477----------------MNKT-----VGSTLLVAGTMIGAGMLAMPLTSAGIGFGFTLVLL
heahi0287--------------MIQQKSP--SLLGGAMIIAGTAIGAGMLANPTSTAGVWFIGSILAL
1003940   --------------MLKNKT-----FGSALIIAGTTIGAGMLAMPLTSAGMGFGYTLLLL
ECOMTRA   ----------MATLTTTQTSP--SLLGGVVIIGGTIIGAGMFSLPVVMSGAWFFWSMAAL
ehu25347  ------------MVIKKTTP--GLMSGTMLIIATVIGGGMFSLPIAMAGIWFPGASIIL
```

Figure 2.1. CLUSTAL W alignment of partial amino acid transporter sequences indicating the consensus sequence (blue).

The sequence that was used for the design of the PSP2 primer for use as a specific forward primer in the PCR was thus:

I	**G**	**A**	**G**	L	F
V		**T**		N	L
				V	I
				M	

The consensus amino acids are shown in blue.

The nucleotide sequence of the primer, obtained by reverse translation of the amino acid sequence and consideration of codon preference, is as follows:

5′ RTW GGW RCW GGW NWW DTW 3′

(W is A or T;
N is any nucleotide;
R is A or G;
D is A, G or T - See Appendix A for the complete list of degenerate nucleotides)

The degeneracy of the primer was further reduced by changing the 3′ W (underlined) to a T. The W (double underlined) at position 14 was changed to an A.

The primer that was synthesised had the following sequence:

5′ TTA RTW GGW RCW GGW NAW DTT 3′

The Blocks database subsequently confirmed that this consensus sequence was part of a signature motif for amino acid permeases. BL00218A is a partial alignment of lower eukaryote (mostly fungal and bacterial) amino acid permeases and contains their characteristic motif (Figure 2.2).

```
ANSP_SALTY ( 26) YHKAMGNRQVQMIAIGGAIGTGLFLGAGARLQMAGP  71

CAN1_CANAL ( 60) VKRDLKARHVSMIAIGGTIGTGLFISTGSLLHTTGP  51

CYCA_ECOLI ( 19) LRRNLTNRHIQLIAIGGAIGTGLFMGSGKTISLAGP  43

GABP_BACSU (  8) LKKELKTRHMTMISIAGVIGAGLFVGSGSVIHSTGP  64

GABP_ECOLI ( 10) LGGGLKSRHVTMLSIAGVIGASLFVGSSVAIAEAGP  78

GAP1_YEAST ( 86) LKHHLKNRHLQMIAIGGAIGTGLLVGSGTALRTGGP  43

HIP1_YEAST ( 86) LSKDLSVRHLLTLAVGGAIGTGLYVNTGAALSTGGP  97

INA1_TRIHA ( 61) LERPMKARHLHMIAIGGSIGAGFFVGSGGALAKGGP  57

ISP5_SCHPO ( 42) LKRTLTARHIQMIGIGGAIGTGVWVGSKNTLREGGA 100

LYSP_ECOLI ( 12) LRRELKARHLTMIAIGGSIGTGLFVASGATISQAGP  37

PUT4_YEAST (106) LKQGLQSRHVQLIALGGAIGTGLLVGTSSTLHTCGP  53

PUTX_EMENI ( 37) TKRGLSSRQLQLLAIGGCIGTGLFVGTSTVLTQTGP  64

TAT2_YEAST ( 78) LKRTLKPRHLIMIAIGGSIGTGLFVGSGKAIAEGGP  38

VAL1_YEAST ( 90) LTKSIKSRHLVMISLGTGIGTGLLVGNGQVLGTAGP  58

YBY2_YEAST ( 84) TRRKLENRHVQLIAISGVIGTALFVAIGKALYRGGP  78

YCC5_YEAST (116) LKKTIQPRHVLMIALGTGIGTGLLVGNGTALVHAGP  54

YFF5_YEAST ( 49) VKRALKNRHISLLALGGVIGPGCLVGAGNALNKGGP  64

ALP1_YEAST ( 67) VKRKLKQRHIGMIALGGTIGTGLIIGIGPPLAHAGP  50
CAN1_YEAST ( 84) VKRELKQRHIGMIALGGTIGTGLFIGLSTPLTNAGP  48
LYP1_YEAST (106) VKRALKQRHIGMIALGGTIGTGLFVGISTPLSNAGP  38

AROP_ECOLI ( 11) LKRGLKNRHIQLIALGGSIGTGLFLGSASVIQSAGP  36
PHEP_ECOLI ( 19) LHRGLHNRHIQLIALGGAIGTGLFLGIGPAIQMAGP  44

HUTM_BACSU ( 11) LKRTMKSRHLFMISLGGVIGTGLFLSTGYTLHQAGP  42
ROCC_BACSU (  8) LQRSMKSRHLFMIALGGVIGTGLFLGSGFTISQAGP  35
ROCE_BACSU ( 11) LQRTMKSRHLFMISLGGVIGTGFFLGTGFTINQAGP  43

PAP1_YEAST ( 76) LKKSMKSRHVVMMSLGTGIGTGLLVANAKGLSLAGP  55
YBR8_YEAST ( 91) LKKSMKSRHVVMMSLGTGIGTGLLVANAKGLHYGGP  57

YIFK_ECOLI (  8) LQRGLEARHIELIALGGTIGVGLFMGAASTLKWAGP  47
YIFK_SALTY (  8) LQRGLEARHIELIALGGTIGVGLFMGAASTLKWAGP  47
```

Figure 2.2 Alignment of amino acid transporter sequences present in Block BL00218A of the Blocks database. The region in red is that section of the motif that was chosen for the second permease specific primer, PSP.

A graphical representations of the motif sequence is shown in Figure 2.3

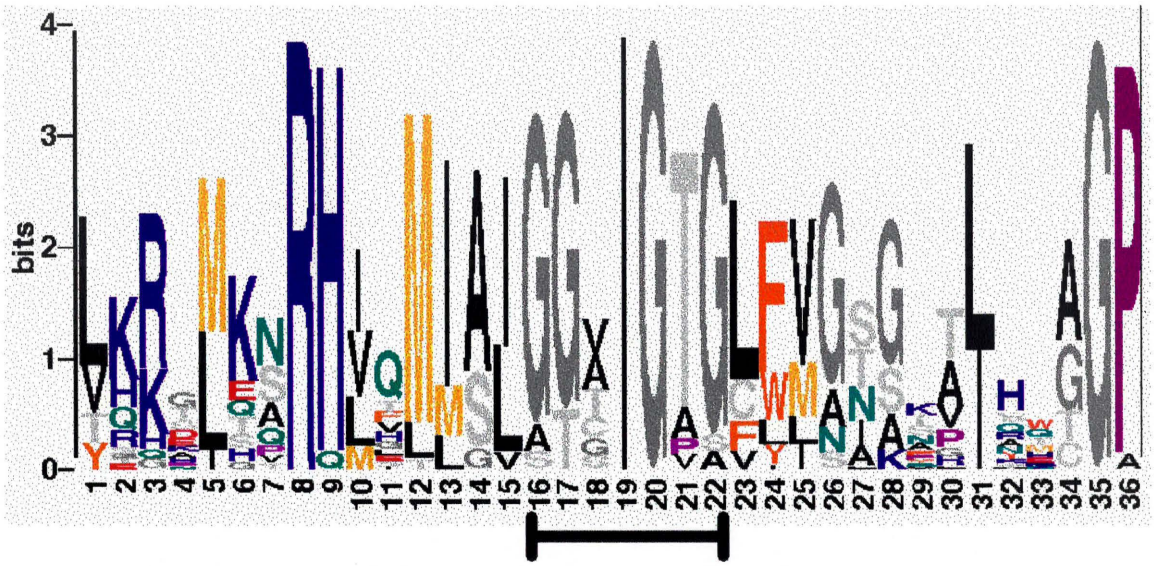


Figure 2.3 Sequence logo of BL00218A. The consensus amino acids used for the design of the primer PSP are indicated by the bar.

The same consensus sequence as was used for PSP2, was used again for PSP, except that it was shifted three amino acids towards the N-terminal, i.e.

G G V **I G T G**
A **A**
T

This was reverse translated as follows (with the Match Index form of codon preference considered).

5′ GGW GGW RYW ATW GGW RCW GGW 3′

The degeneracy of this primer was 1024. To reduce this the W at position 3 (underlined) was made a T and the W at position 9 (double underlined) was made an inosine. The primer nucleotide sequence was subsequently:

5′ GGT GGW RYI ATW GGW RCW GGW 3′

This primer has a real degeneracy of 256 but an effective degeneracy of 128 (due to the degenerate W at the 3′-end). Oligo revealed that there was a strong upper primer: upper primer association. To ameliorate this the primer was modified by changing the T at position 3 into an A (this still complies with the sequence requirements as the

original nucleotide was a W). The synthesised oligonucleotide thus had the following sequence:

5′ GGA GGW RYI ATW GGW RCW GGW 3′

The tyrosine/tryptophan amino acid permease primer, TYT was based on the following section of the consensus sequence with the strongly conserved nucleotides shown in bold.

G T T **I G A G** M
A G

Upon reverse translation and considering codon preferences, the nucleotide sequence was as follows:

5′ GGW ACW RCW ATW GGW GSW GGW ATG 3′

To alleviate the upper: upper primer association and to reduce the degeneracy, the primer was modified at positions 12 and 18 with the replacement of W with T.

5′ GGW ACW RCW ATT GGW GST GGW ATG 3′

The degeneracy was thus reduced from 512 to 128.

The suitability of all the primers for PCR was evaluated using OLIGO. The following results are those obtained using the PSP and anchor primers (as forward and reverse primers, respectively).

The energy profile (internal stability) of primer PSP is shown in Figure 2.4.

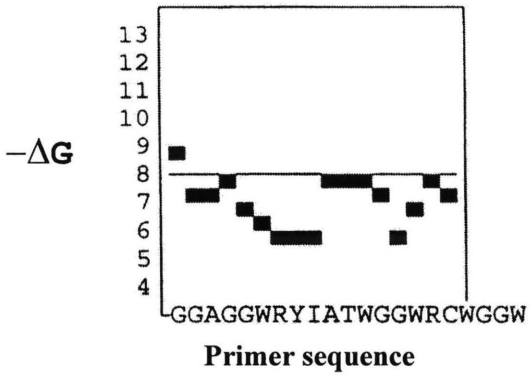


Figure 2.4 The internal stability plot of primer PSP. Each point represents the ΔG value of a pentamer.

This agrees with that recommended by Rychlik, namely that the 3′-end should be fairly unstable (less negative than -9 kcal/mol), thereby increasing the primer's specificity (Rychlik, 1993).

OLIGO calculates the Tm of the primers at various salt and formamide concentrations. These values are shown in Table 2.1

Table 2.1 DNA melting temperature (°C) in various salt and formamide concentrations. SSC consists of 150mM NaCl and 15mM sodium-citrate.

Na concentration		**Formamide concentration**		
[mM]	**(*SSC)**	**0%**	**10%**	**50%**
1	(0.006)	17.1	10.6	-15.4
10	(0.06)	33.7	27.2	1.2
50	(0.3)	45.3	38.8	12.8
165	(1)	53.9	47.4	21.4
330	(2)	58.9	52.4	26.4
500	(3)	61.9	55.4	29.4
1000	(6)	66.9	60.4	34.4

OLIGO also calculates the effect of mismatches and truncations on the Tm. These values are shown in Table 2.2.

Table 2.2 Tm (°C) of mismatched and truncated versions of primer PSP.

	Number of mismatches[a]				**1 bp short**	
Oligo	**0**	**1**	**2**	**3**	**at 5′**	**at 3′**
Upper	49.2	43.5	37.8	32.1	47.1	48.0
Lower	43.8	38.1	32.4	26.7	40.7	41.3

[a] Tm = Tm - 1.2 x (% mismatch).

Self-complementarity is a vital component of primer design. The interaction of the primers (PSP and anchor) with themselves and each other is shown below.

PSP : PSP

```
 5'        GGAGGWRYIATWGGWRCWGGW 3'
                ::::                ΔG: -5.7 kcal/mol
3' WGGWCRWGGWTAIYRWGGAGG 5'
```

Anchor : Anchor

```
5' GCTATCATTACCACAACACTC 3'
      |   ::     ||   |          ΔG: -1.9 kcal/mol
  3' CTCACAACACCATTACTATCG 5'
```

PSP : Anchor

```
5' GGAGGWRYIATWGGWRCWGGW 3'
    :::     |                    ΔG: -3.2 kcal/mol
 3' CTCACAACACCATTACTATCG 5'
```

Figure 2.5 The interaction of the PSP and anchor primers with themselves and each other. The ΔG on the right hand side is the thermodynamic value of the combined bonds indicated by colons (:).

The ΔG value of - 1.9 is relatively negligible. The values of -5.7 and - 3.5 are more negative, but are still more positive than the -ΔG of the 3′-ends (see Table 2.3) for annealing to the template and should thus not interfere with a PCR performed at the correct Tm.

OLIGO found only two potential hairpin loops. One with a loop of 3 nucleotides and the other with 5 nucleotides. Their ΔGs however, were both positive (+3.3 and +3.4 kcal/mol) and can thus be ignored.

It is useful to check the template sequence for any stretches complementary to the primer as these are potential false priming sites and may result in spurious bands during PCR. Since a *Plasmodium* amino acid permease sequence is unavailable, the yeast general amino acid permease nucleotide sequence with the *Plasmodium* codon bias was used as a template, so that an indication of the amount of false priming could

be obtained. OLIGO found the following fragments complementary to primer PSP (with a minimum length of significant duplex being 7 bp):

```
          1 GGAGGWRYIATWGGWRCWGGW 21
(1)         || |||  ||||||| |||||
       1333 CCWCCWCGWTAWCCWTGWCCW 1313

          1 GGAGGWRYIATWGGWRCWGGW 21
(2)         |    |  |||||||     |
         94 CAWRAWCCWTAWCCWATARAW 74

          1 GGAGGWRYIATWGGWRCWGGW 21
(3)                 ||||||| |   |
        670 TAWAAAATATAWCCWWGDRAW 650
```

Figure 2.6 Sites of complementarity to the PSP primer on the modified yeast general amino acid transporter template.

Site number 1 is the intended target sequence. Sites 2 and 3 are all possible false priming sites, but occurs in the middle of the primer sequence and would not interfere with PCR.

Table 2.3 contains the relevant details regarding all the primers used in this study.

Table 2.3. Summary of relevant details for all the primers used in the study.

Primer	Sequence (5' - 3')	Tm (°C) OLIGO Td[a]	Tm[b]	Tm[c]	Calculated[d] Min	Max	Degeneracy	%AT Min	Max	ΔG 3' (kcal/mol)	Ta OPT [f] (°C)
PSP2	**TTA** RTW GGW RCW GGW NAW DTT	56.2	59.1	52	48.11	55.92	1536	57.14	76.19	-7.2	47.6
PSP	GGA GGW RYI ATW GGW RCW GGW	61.9	66.9	60	55.91	61.71	256	42.86	57.14	-7.9	47.7
TYT	GGW ACW RCW ATT GGW GST GGW ATG	62.1	70.5	68	61.01	62.72	128	50	54.17	-5.7	47.0
Anchor	**GCT** ATC ATT ACC ACA ACA CTC	56.4	66.9	60	55.92		1	57.1		-6.4	--
Poly-T	Anchor+$(T)_{18}$ VN	54.8	47.8	40	36.8	40.9[e]	12	ND	ND	ND	ND

Notes:

- [a] Td is calculated by OLIGO using the nearest neighbour method.
- [b] Tm is calculated by OLIGO using the %GC method.
- [c] Tm is calculated by OLIGO using the following equation: $Tm = [2° \times (A + T) + 4° \times (G + C)]$.
- [d] These temperatures are calculated according to the following equation: $Tm = 69.3 + 0.41\ (\%GC) - {}^{650}/_{L}$ where L is the length of the primer. Min and Max are the calculations for the minimum %GC and the maximum %GC respectively.
- [e] The Tm was calculated for the $(T)_{18}$VN stretch only, as the poly-T primer is intended for first strand cDNA synthesis and the anchor does therefore not anneal.
- [f] Ta^{OPT} is calculated by OLIGO for the PSP and PSP2 primers respectively, in conjunction with the anchor primer and on the yeast template.
- ND = not determined.
- The nucleotides in bold are those that were added for the DISEC-TRISEC method of cloning.

2.4 Discussion

The consensus sequence identified using CLUSTAL W (Figure 2.1), appears to be part of a larger signature motif which is conserved throughout most of the known amino acid permeases. Sophianopoulou and Diallinas (1995) contend that the sequence M(L)I(L)AL(I)GGXIGTGLFVG, with X being any hydrophobic amino acid, is strongly conserved in bacteria and the lower eukaryotes (primarily fungi). This was confirmed by the Blocks database (Figure 2.2, positions 12-26).

Blocks can show the aligned sequences graphically in the form of a sequence logo. The characters are stacked above each other, with the height of each character proportional to its frequency. The letters that are the most common are placed at the top (Schneider and Stephens, 1990). It is easier to identify which amino acids are strongly conserved using this logo. Figure 2.3 shows the logo for the amino acid permeases. At amino acid position number 18 in this logo, there are six amino acids that are conserved in the aligned sequences, namely V, A, T, C, G and S. Amino acids V, A and T are present more frequently than the other three. These three amino acids were thus used in that position for the PSP primer. The amino acids V and A correspond to and agree with, the hydrophobic 'X' of Sophianopoulou *et al* (Sophianopoulou and Diallinas, 1995).

Ideally the primer should have the least degeneracy possible. In this instance, since the primer was rather general, being for an amino acid permease, a slight degeneracy is not disadvantageous. Successful PCRs have been performed with primers with a degeneracy of up to1024 (Rychlik, 1993).

The degeneracy of the primers was reduced using codon preferences (both the normal and Match index forms), the incorporation of inosine, the selective use of specific nucleotides instead of degenerate nucleotides and the use of 'promiscuous' T.

The match index is a value defined by Hyde *et al* (1989). It is a measure of the probable detrimental effect of choosing the wrong codon from among the group that code for a particular amino acid. Probability figures in the algorithm are derived from

the codon frequencies. They found that the most preferred codon was the same as the most frequent codon in all cases except for serine.

Inosine was incorporated at position number 9 in primer PSP, thus reducing its degeneracy by half. Several rare cDNA products have been amplified using primers containing inosine (Preston, 1993). The advantages of using inosine are that the degeneracy can be reduced as it can presumably base pair equally well with any of the four nucleotides. It creates a single bond in each case (Preston, 1993). The premise is that inosine acts as a neutral base and does not affect the stability of the DNA duplex as much as a mismatched Watson Crick base pairing (Hyde and Holloway, 1993). The disadvantages are that the presence of inosine in the primer will reduce its Tm (Preston, 1993). Inosine is much more expensive to incorporate into the primer and it is also less stable than the other phosphoramidites which may have a negative influence on the lifetime of the dissolved primer (Hyde and Holloway, 1993).

The identity of the 3′ terminal nucleotide had been speculated on by both Hyde and Holloway (1993) and Preston (1993). They are not however in agreement. Hyde is of the opinion that, depending on the amino acid sequence, the 3′ base should be a T, as a mismatched T is less likely to prevent elongation than any of the other nucleotides. Preston however believes that the 3′ nucleotide should be as specific as possible, and should thus not be N, I or T, but rather one of the strong binding nucleotides (G or C). Primer PSP2 had its degeneracy reduced by placing a T instead of the W on the 3′-end, although this T is not mismatched as W is A or T.

With regard to mismatches, it is generally accepted that primers with 15% to 20% base pair mismatch with the template, will still function effectively in PCR. However the mismatches should not be too close to the 3′-end as 2 mismatches in the last 4 base pairs will drastically reduce the product yield (Rychlik, 1993).

OLIGO was used to evaluate how suitable the primers were for PCR. It generates numerous results. According to Rychlik, the most suitable PCR or sequencing primer should have a relatively stable 5′-end, but rather unstable 3′-end (Figure 2.4). This is

because the 3′-end is the terminus driving synthesis and will thus only be extended when the primer is properly annealed, since it will only then be stable enough for the polymerase to extend (Rychlik, 1993). Primers with a 3′ terminal ΔG of between -5 kcal/mol and -10 kcal/mol are recommended (Estruch, 1997), with the more positive (below -9 kcal/mol), being more favourable (Rychlik, 1993). Figure 2.4 shows that primer PSP has a 3′ terminal stability of -7.9 kcal/mol. This type of profile also minimises false priming events. The less stable the 3′ terminal, the broader the optimal annealing temperature range, which means that performing PCR at an optimum annealing temperature is easier (Rychlik, 1993). OLIGO also evaluates any potential dimer formation. This is because if the primers form dimers with themselves, or each other, the effective primer concentration in the reaction is reduced, which could lead to an ineffective PCR with a low yield of specific product.

Self complementarity and hairpin formation are other important factors to consider. Figure 2.5 shows the self-complementarity of primers PSP and anchor with themselves and each other. The ΔG of these interactions (-5.7, -1.9 and -3.2 kcal/mol) is more positive than that of the 3′-ends (-7.9 and -6.4 kcal/mol for PSP and the anchor respectively) and should thus not be a factor in a PCR performed at the correct annealing temperature. The only hairpins formed in these primers have positive ΔGs and will thus not form under standard reaction conditions. It is important that hairpin loops be investigated as it is another way in which the primer is removed from the reaction (as the 3′-end is 'tied up') (Rychlik, 1993).

Among the valuable information generated by OLIGO is a variety of melting temperatures calculated using a variety of equations and under different conditions. Table 2.3 presents melting temperatures calculated using different equations (notes a, b and c). Table 2.1 shows the calculated melting temperatures for various salt and formamide concentrations. This information is particularly useful for reactions where the salt concentrations may differ, which is a common situation. Table 2.2 presents the effect of mismatches on the melting temperature.

Figure 2.6 shows the potential false priming sites. As previously mentioned, site number 1 is the intended target site. The remaining sites are all false priming sites and may be responsible for spurious bands in the PCR if the conditions are not sufficiently stringent.

Primer PSP is theoretically a better primer for PCR in comparison with PSP2. This is because of the following reasons: PSP has a 6 fold lower degeneracy (real degeneracy of 256, but an effective degeneracy of 128, in comparison with the 1536 of PSP2). Both PSP and PSP2 are 21-mers, but PSP has all these 21 nucleotides spread over the consensus sequence, with better placement than PSP2. PSP2 only has 18 of the 21 nucleotides placed on the consensus. This is particularly significant as the 3′ terminal drives synthesis and should thus be as specific as possible. The remaining three nucleotides of PSP2 are TTA and these were added to facilitate the DISEC-TRISEC method of cloning. The 3′-end of PSP is also situated on consensus amino acids while that of PSP2 is not. Primer PSP has a Tm higher than that of PSP2 which means that the PCR can be made more stringent. Primer TYT is not really comparable with PSP and PSP2, as it is a more specific primer. The benefits of TYT are that it has a low degeneracy and a higher calculated Tm as it is a 24-mer (Table 2.3).

It is accepted that the amino acid permeases belong to a unique family of transport proteins. They appear to all be of similar length (between 550 and 628 amino acid residues) and topology (Sophianopoulou and Diallinas, 1995). Based on this, a product of approximately 1700 nucleotides is expected from the primers. However, this should not be taken for granted as malaria parasite genes tend to have insertions compared to the genes of other species.

2.5 Conclusion

Obtaining biological sequence information is greatly facilitated by the access to the Internet, as there are many databases available containing various types of data. The use of primer analysis software is imperative for the design of primers which are to be used to amplify products larger than ~ 400 base pairs.

A stretch of consensus sequence for amino acid permeases was identified with the help of CLUSTAL W and confirmed by the Blocks database. This sequence was used for the design of forward primers PSP and PSP2. Primer TYT was based on the same region although the amino acid sequence chosen was specific for tyrosine and tryptophan permeases. These forward primers are to be used with an anchor reverse primer in PCR.

The degeneracy of the primers was reduced in various ways, namely the use of codon preferences (both normal and Match Index values) for *P. falciparum*, the incorporation of inosine and the use of 'T' mismatches. PSP is theoretically better than PSP2 as it conforms to more of Rychlik's rules for a successful primer, such as less degeneracy, better positioning over the consensus region and a higher GC-content at the 3′-end.

CHAPTER 3

3′-RACE of a Putative Amino Acid Transporter Gene of *Plasmodium falciparum*

3.1 Introduction

This chapter focuses on the application of the primers designed in Chapter 2. As previously indicated, a modified form of RT-PCR was chosen as the means to identify an amino acid transporter gene in the malaria parasite.

Since Trager and Jensen first published their method for the *in vitro* cultivation of the erythrocytic stages of *P. falciparum*, the method has become widely popular and has achieved 'classic' status. It has been applied to nearly all aspects of malaria research. These include chemotherapy; the biochemistry of the parasite and its relationship with the host erythrocyte; pathogenesis; cellular and molecular biology and immunology and vaccine development (Trager and Jensen, 1997).

The year 1983, like 1953 saw the beginning of a new age in molecular biology. The polymerase chain reaction (PCR) has revolutionised the field of molecular biology, just as the discovery of the DNA double helix started it (Watson and Crick, 1953; Mullis, 1986). Using PCR it is possible to amplify a target DNA sequence several million fold. RACE (rapid amplification of complementary ends) is a PCR based technique through which unknown 3′ or 5′-ends of a gene of interest can be amplified starting with the knowledge of only an internal consensus region of the specific gene and a 3′ or 5′ anchor primer. Synthesis of cDNAs from the mRNA is accomplished using an RT-primer that anneals to the poly(A) tail. Amplification is achieved using the anchored section of the RT-primer along with a gene specific primer. The technique is represented schematically in Figure 3.1 (Frohman, 1993)

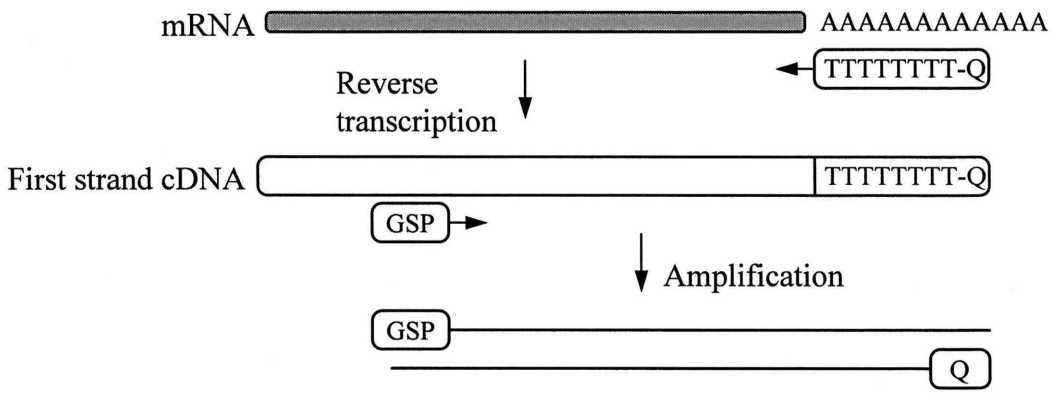


Figure 3.1 A schematic representation of 3′-RACE, a technique used if only an internal stretch of sequence is available for the gene of interest. Q is the anchor section of the RT-primer.

In this chapter we describe the isolation of RNA, the reverse transcription of the message and total RNA to cDNA, the PCRs employed and the confirmation of a novel gene sequence as of *P. falciparum* origin. Since the primary aim of this study was the application of RACE techniques to malaria genes, the optimisation of various processes, as well as the problems encountered, is elaborated on.

Optimal conditions for a PCR differ between reactions and need to be determined empirically. Various factors should be optimised, which include:

- $MgCl_2$ -. Too little Mg^{2+} renders the enzyme inactive, while an excess of Mg^{2+} results in non-specific amplification. A low concentration is desirable.
- Enzyme choice and concentration - A thermostable DNA polymerase of high fidelity is required. Since *Taq* has 5′-3′ exonuclease activity, excessive amounts of enzyme produce a smear.
- Template - the quality and the concentration of the template are important factors in the reaction. The template should thus be as pure as possible. The amount of template used in the reaction needs to be determined empirically.
- Cycling conditions - The annealing temperature is one of the factors crucial to the success of the PCR. A specific, high annealing temperature is ideal. 25 (3.4×10^7 times) - 40 ($\sim 1 \times 10^{12}$ times) cycles is generally sufficient to amplify most products.
- Primers - a well designed primer that concurs with the requirements established by Rychlik can greatly facilitate the success of a PCR (Promega, 1996).

A variety of cloning procedures are available. These include blunt end cloning, sticky end cloning and numerous variations of either process. Traditional blunt end cloning requires phosphorylation of the insert and dephosphorylation of the vector. The latter can be avoided by including the restriction enzyme for the cloning site in the ligation reaction. For sticky end ligation of PCR products, the common manner of generating sticky ends is to include restriction sites on the 5′-end of the primers. We opted for blunt end procedures (or the DISEC-TRISEC sticky end ligation method) as the restriction enzymes sites would have to be GC-rich to avoid digesting the AT-rich PCR product of unknown sequence and thus unknown restriction sites. Primers with GC-rich 5′-ends would negatively affect PCR of AT-rich genomes. The DISEC-TRISEC method of sticky end ligation is summarised in Figure 3.2.

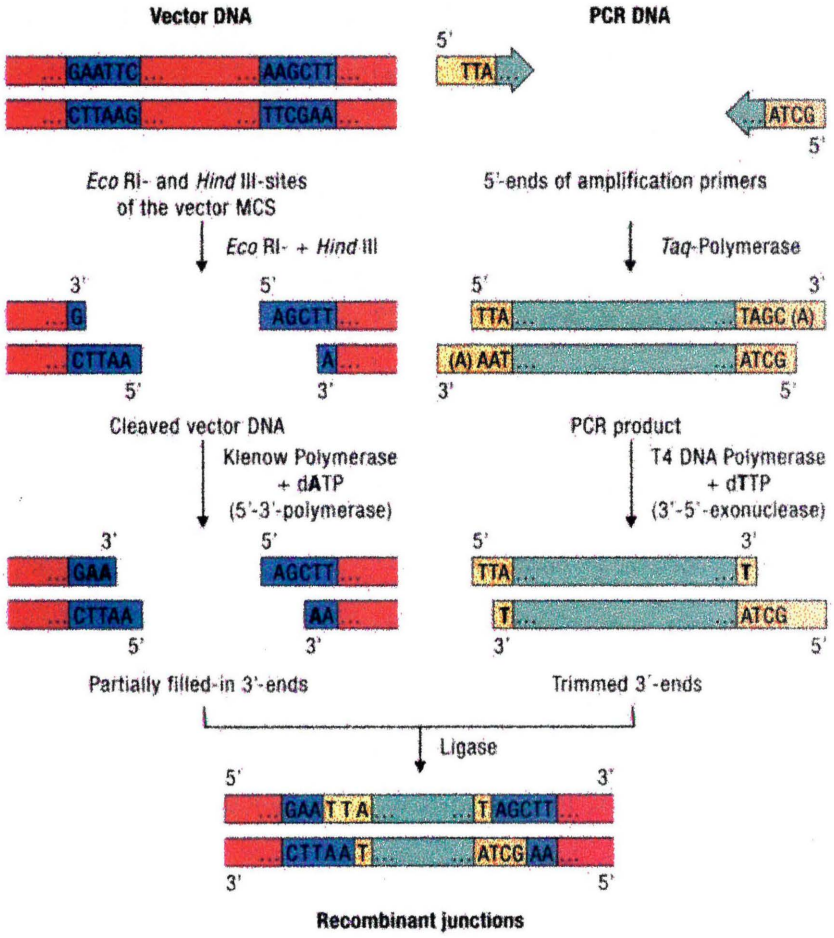


Figure 3.2 A schematic representation of the processing of vector and PCR products during the DISEC-TRISEC procedure (Dietmaier and Fabry, 1995).

This method relies on the generation of complementary di- or trinucleotide sticky end overhangs on the PCR product and digested vector. The 5′ overhangs produced by the

restriction enzymes are partially filled in using Klenow polymerase and the appropriate dNTPs. The PCR products were generated using the nucleotides, TTA and GCTA, added onto the 5′-ends of primer PSP2 and anchor, respectively. These products are then modified using the 3′-5′ exonuclease activity of T4 DNA polymerase, which is limited by the presence of the appropriate dNTP. The result is a fragment with sticky ends complementary to those on the vector (Dietmaier *et al.*, 1993).

There are a variety of non-radioactive methods for the labelling of probes. One of these employs the steroid hapten digoxygenin (DIG). This molecule is linked to a nucleotide (usually UTP) via a linker arm (Figure 3.3)

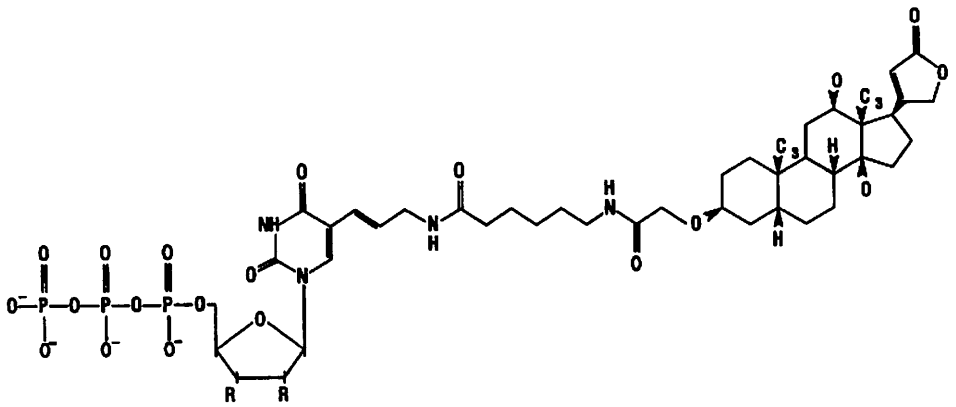


Figure 3.3 The general structure of a DIG labelled nucleotide. DIG-dUTP, R_1 = OH and R_2 = H.

The DIG system is preferable to radioactive labelling as it is vastly safer and results are obtained more rapidly (as the exposure time is shorter). An added advantage is that it is more economical, mostly due to a longer shelf life.

DIG can be incorporated into the probe in a variety of ways such as: randomly primed labelling; nick translation; incorporation during PCR; oligonucleotide tailing; oligonucleotide 3′-end labelling and 5′-end labelling (these end labelling reactions usually use DIG-ddUTP to ensure the addition of only one molecule). The detection of the DIG molecule is greatly adaptable. Detection is based on the recognition of the DIG moiety by a specific anti-DIG antibody. This antibody can be fluorescently tagged for fluorescent detection or it can be conjugated to an enzyme which would allow colourimetric or chemiluminescent detection. The enzyme most commonly used is alkaline phosphatase. Chemiluminescent detection is vastly superior to colourimetric

detection and of the chemiluminescent substrates, CDP-Star™ is the most sensitive. CDP-Star™ is a chloro-substituted 1,2-dioxetane which exhibits rapid light signal generation. This signal can be captured by X-ray film or by instrumentation. Enzymatic dephosphorylation of the molecule leads to a meta-stable dioxetane phenolate anion which decomposes and emits light at 466nm in the appropriate buffer (Figure 3.4). The signal persists for days allowing multiple exposures.

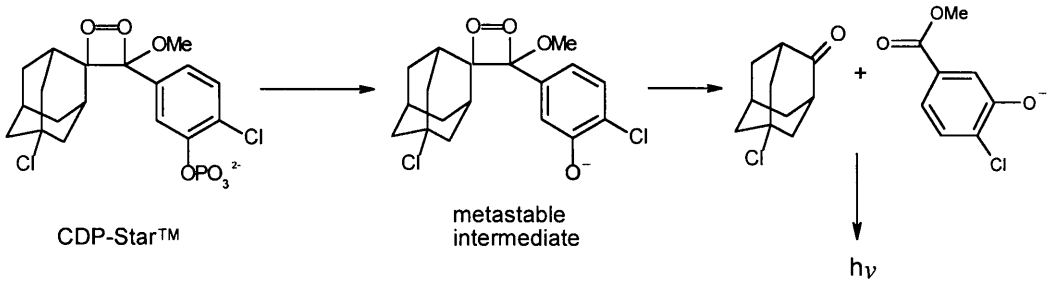


Figure 3.4 The chemical reaction of CDP-Star™ in the presence of alkaline phosphatase.

Alternatively colourimetric visualisation can be employed, though this is less sensitive than when CDP-Star™ is used. The colourimetric substrate used was Nitroblue tetrazolium (NBT) and X-phosphate (Figure 3.5).

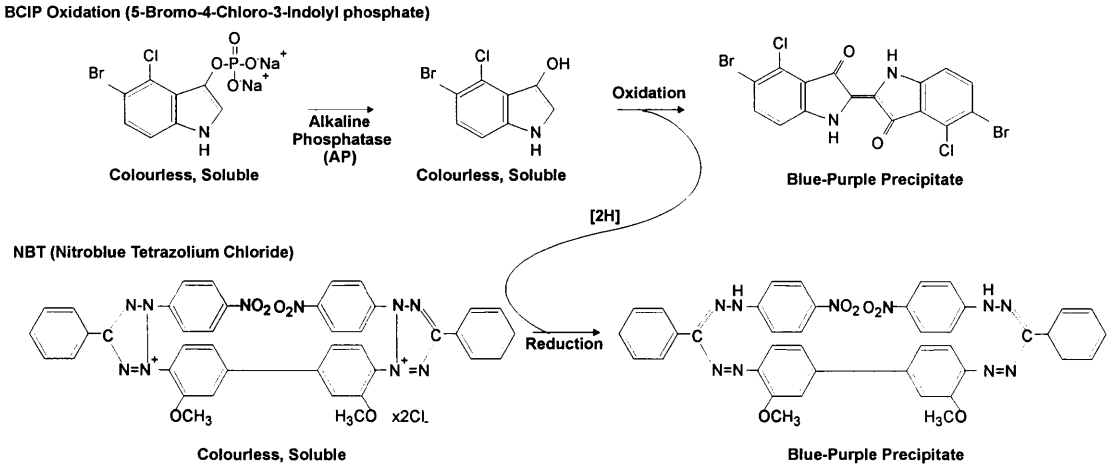


Figure 3.5 The chemical reaction of NBT and X-phosphate in the presence of alkaline phosphatase.

3.2 Materials and Methods

3.2.1 *Plasmodium falciparum* cultures (Trager and Jensen, 1976; Trager, 1994)

Maintaining *P. falciparum* in continuous culture depends on the maintenance of human erythrocytes under conditions that support intracellular development of parasites. All solutions and apparatus used in the cultivation and maintenance of parasites were sterile.

Cryopreserved *P. falciparum* isolate PfUP1, was used for these experiments. The cryotube was removed from liquid nitrogen and rapidly thawed at 37°C. The contents (between 0.5 and 1ml) were transferred to a 10ml centrifuge tube and 0.2ml of 12% (w/v) NaCl was added, and mixed well, whereafter 1.8ml of 1.6% (w/v) NaCl was added and mixed. The suspension was centrifuged (2500g, 5 minutes, room temperature) and the supernatant aspirated. The pellet was resuspended in 10 drops (~0.5ml) of fresh human erythrocytes. The cells were transferred to a culture flask ($75cm^3$) and 10ml of preheated (37°C) culture medium added. Culture medium contained 1.04% (w/v) RPMI 1640 with L-glutamine, 24.96mM HEPES, 22.22mM D-glucose, 323.27μM hypoxanthine, 4mg gentamycin solution and 21.43mM $NaHCO_3$. Human serum of the appropriate blood type (O or A) was added to a final concentration of 10% (v/v). Culture medium was filtered prior to the addition of $NaHCO_3$ and serum. The latter components were filtered independently prior to their addition. The cultures were gassed for approximately 30 seconds with a gas mixture composed of 5% O_2, 5% CO_2 and 90% N_2 and were incubated at 37°C. RPMI 1640 and gentamycin were from Highveld Biological (South Africa), HEPES from Sigma Chemical Co. (St. Louis, MO, USA) and glucose, $NaHCO_3$ and hypoxanthine from Merck (Germany).

Culture maintenance commenced with the examination of the culture on a Giemsa-stained thin smear under the light microscope. Cultures were maintained at a

parasitaemia of ~5% by the addition of fresh erythrocytes and culture medium. The medium was exchanged every 24 to 48 hours depending on parasite proliferation.

3.2.2 Quantification of nucleic acids

The concentration of nucleic acids used in the subsequent experiments was determined using one of three methods:

3.2.2.1 Spectrophotometric quantitation

The concentration of nucleic acids was determined using the Shimadzu UV-160A spectrophotometer. Absorbance values were obtained at 260nm and 280nm. The conversion values are as follows:

Double stranded DNA : 1 A_{260} is equivalent to 50µg/ml,

Single stranded DNA (oligos) : 1 A_{260} is equivalent to 33µg/ml, and

RNA : 1 A_{260} is equivalent to 40µg/ml.

The ratios obtained from A_{260}/A_{280} is an indication of the purity of the isolated DNA or RNA. A ratio of 1.8 and 2.0 for DNA and RNA respectively, indicates a pure product (Sambrook *et al.*, 1989).

We found the spectrophotometer to give reliable values for the quantification of RNA and oligonucleotides, but not for isolated plasmids and PCR products and alternative methods were used.

3.2.2.2 Fluorometric quantitation of plasmid DNA

The Hoefer Scientific Instruments (San Francisco, CA, USA) Mini-fluorometer TKO 100 was used primarily in the determination of kit-isolated plasmid concentrations, according to the manufacturer's instructions concerning calibration and use. The Hoechst 33258 dye was used at a concentration of 0.1µg/ml in 1x TNE buffer (10mM Tris-Cl, 1mM EDTA, 0.2M NaCl, pH 7.4). A 100µg/ml calf thymus DNA standard was used and the fluorometer value is thus read directly as a concentration (µg/ml). Since the H33258 dye binds to the A and T bases of double stranded DNA, the

fluorescent signal is determined by the concentration of DNA and the AT-content. As malaria contains an AT-rich genome, the standard setting requires adjustment when fluorometer readings for malaria are taken. The standard setting was adjusted as follows:

$$\text{Adjusted standard setting} = [C_{std} (0.025 (AT\%_{std} - AT\%_{sample})] + 1$$

3.2.2.3 Ethidium bromide dot quantitation assay for PCR products

For PCR fragments the fluorometer spectrophotometer values tended to be inaccurate. A grading method using ethidium bromide stained standard concentrations of calf thymus DNA was thus employed and a direct visual comparison made to estimate the DNA concentration (Sambrook *et al.*, 1989).

3.2.3 Total RNA isolation

TRI-reagent (Molecular Research Centre Inc., Cincinnati, OH), based on a mixture composed of phenol and guanidinium thiocyanate in a mono-phase solution, was used for the single step isolation of total RNA (Chomczynski, 1993). All solutions were prepared using DEPC-treated water, with DEPC obtained from Sigma Chemical Co.

A *P. falciparum* culture that was primarily in the ring stage was used for the first round of RNA isolation. The erythrocytes were pelleted from the culture medium by centrifugation at 3000g for 10 minutes. The cells were incubated in saponin at a final concentration of 0.1%, at room temperature for 5 minutes, which lysed the erythrocytes. Released parasites were pelleted by centrifugation at 3000g for 15 minutes, the supernatant aspirated, the pellet suspended in 1ml of 1x PBS (136.9mM NaCl; 2.68mM KCl; 1.76mM KH_2PO_4; 10.14mM Na_2HPO_4; pH 7.0) and transferred to microfuge tubes. Parasites were pelleted by centrifugation at 11500g for 2 minutes and washed three times with 1ml of PBS. Washed parasites were subsequently lysed by the addition of 2ml of TRI-reagent, followed by extensive homogenisation with a Wheaton homogeniser. The homogenate was incubated at room temperature for 5 minutes, followed by the addition of 400µl of chloroform. The mixture was vortexed thoroughly and incubated at room temperature for 10 minutes. Hereafter, the tubes

were subsequently centrifuged at 11500g for 15 minutes (4°C). RNA remains in the upper aqueous phase, DNA in the interphase and protein in the organic phase. RNA was processed by removing the upper phase and transferring it to a clean microfuge tube. Precipitation of RNA was achieved by adding 1ml of isopropanol for every 2 ml of TRI-reagent used. The tubes were incubated at room temperature for 10 minutes and subsequently centrifuged at 9400g for 10 minutes. After the supernatant was carefully aspirated, the pellets were washed with 75% ethanol using 1ml for every 1ml of TRI-reagent used. The tubes were centrifuged at 11500g for 5 minutes (4°C) and the supernatants aspirated. The pellets were allowed to air dry and were subsequently dissolved in 80µl of FORMAzol™. (Molecular Research Centre Inc. Cincinnati, OH). FORMAzol™ is purified and stabilised formamide, which has been shown to enhance the shelf-life of stored RNA, as it keeps the RNA stable at -20°C for longer than 12 months and also protects the RNA from degradation by RNases (Chomczynski, 1992; Chomczynski, 1993).

The entire isolation procedure was repeated 24 hours later, with a culture containing primarily trophozoites. RNA from both isolations was pooled and the concentration determined spectrophotometrically. Isolated RNA was stored at -20°C.

3.2.4 mRNA isolation

mRNA was isolated from total RNA using the Dynabead mRNA Purification Kit (Dynal A.S., Oslo, Norway) according to the manufacturers instructions. Dynabeads are superparamagnetic polystyrene beads of 2.8µm diameter with an oligo $(dT)_{25}$ covalently attached to their surface via a 5′ linker group. The beads are collected using the Dynal MPC-E-1 magnetic particle concentrator. mRNA isolated using this method is pure enough for the synthesis of cDNA libraries, translation experiments, PCR and related experiments.

Dynabeads (500µg) were washed twice with 100µl of 1x washing buffer (10mM Tris-Cl pH 7.5; 0.15M LiCl; 1mM EDTA), twice with 100µl of 1x elution buffer (2mM EDTA, pH 7.5) and once with 100µl of 2x binding buffer (20mM Tris-Cl pH 7.5;

1.0M LiCl; 2mM EDTA). The beads were subsequently suspended in 50µl of binding buffer. Approximately 40µg of RNA was made up to a volume of 50µl with DEPC-treated water. The RNA was prepared for binding by disruption of its secondary structure at 65°C for 2 minutes. The RNA was immediately added to the bead suspension and incubated at room temperature for 5 minutes to allow the RNA to bind. Hereafter the beads were washed twice with 100µl of washing buffer and suspended in 80µl of elution buffer and the poly(A^+) RNA was eluted by heating at 65°C for 2 minutes. The isolated mRNA was stored in the elution buffer at -20°C.

3.2.5 Electrophoresis of RNA (Sambrook *et al.*, 1989)

RNA samples were electrophoresed on a 1% denaturing agarose gel (1% (w/v) agarose (Promega); 1x MOPS (Sigma Chemical Co.); 1.64ml 37% (v/v) formamide (GIBCO-BRL, USA) and DEPC treated-water to 30ml, in a running buffer of 1x MOPS. Samples are prepared for electrophoresis by suspending ~5µg of RNA in a 1x MOPS solution containing 15% (v/v) formaldehyde (Sigma) and 50% (v/v) deionised formamide. The samples were heated at 55°C for 15 minutes to disrupt the secondary structure. Tracking dye (15% (w/v) Ficoll and 0.025% (w/v) Bromophenol Blue) and 2µg ethidium bromide were added to the samples prior to electrophoresis on a Minicell® EC 370M Electrophoretic Gel System (E-C Apparatus Corporation, St Petersburg, FL, USA). Gels were run at 79V (5.3V/cm) and visualised on a transilluminator (Spectroline® TC-312A, 312nm).

3.2.6 cDNA synthesis

3.2.6.1 cDNA derived from mRNA (Ausubel *et al.*, 1992)

3pmol of poly-T primer was used per 2µg of mRNA and the total volume taken to 100µl using DEPC-treated water. The mRNA and the primer were co-precipitated by adding 0.1 volumes of 3M sodium acetate and 2 volumes of ice cold ethanol, then incubating at -20°C, overnight. The precipitated nucleic acids were pelleted by centrifugation at 13600g for 15 minutes (4°C), the supernatant aspirated and the pellet washed with 200µl of 75%(v/v) ethanol. The nucleic acids were again pelleted by

centrifugation at 13600g for 15 minutes. After the removal of the supernatant, the pellet dried *in vacuo* and suspended in 12µl of DEPC-treated water with the addition of 4µl of 0.4M Tris pH 8 and 4µl of 0.4M KCl. After the reaction mixture was incubated at 40°C for 3 hours, to allow the primer to anneal to the mRNA, the following was added: 8µl of 5x first strand buffer (250mM Tris-HCL, pH 8.3; 375mM KCl; 15mM $MgCl_2$); 4µl of 0.1M DTT; 0.75mM of dGTP and dCTP and 1.125mM dATP and dTTP (in a volume of 3µl); and 300 units of Superscript II reverse transcriptase (GIBCO-BRL, USA). DEPC-treated water was added to a final volume of 40µl. The tubes were incubated at 45°C for an hour and the enzyme was subsequently inactivated by heating at 95°C for 3 minutes.

3.2.6.2 cDNA derived from total RNA

cDNA was also synthesised directly from total RNA, without co-precipitation of the RNA and poly-T primer. To destroy any secondary structure, 5µg of total RNA and 25pmol of poly-T primer were heated at 70°C for 10 minutes, then snap cooled on ice for 2 minutes. 2µl of 0.1M DTT; dNTPs in the concentrations indicated above (in a volume of 4µl); 4µl of 1x first strand buffer and 200 units of Superscript II reverse transcriptase were added. DEPC-treated water was added to a final volume of 20µl. In later experiments, a 'hot start' protocol was adopted, whereby the reaction was incubated at 45°C for 2 minutes prior to the addition of 200 units of Superscript II reverse transcriptase. Reactions were incubated at 45°C for an hour, followed by heating at 95°C for 3 minutes to inactivate the enzyme.

In the experiments using cDNA derived from total RNA, the RNA was initially precipitated out of the FORMAzol solvent as it was suspected that it may inhibit the reverse transcriptase during the cDNA synthesis reaction. This was done by adding one tenth the volume of 3M sodium acetate, pH5.4 and 2½ volumes of ethanol to 43.54µg of RNA, followed by incubation at -20°C overnight. The reaction tube was centrifuged at ~13000g for 20 minutes, the supernatant aspirated and the pellet washed with 70% ethanol, followed by centrifugation at ~13000g for 8 minutes. After aspiration of the supernatant was aspirated, the pellet was air dried, resuspended in 50µl of TE-buffer and the concentration determined spectrophotometrically. Subsequent experiments

indicated that FORMAzol had no effect on the reverse transcription reaction due to its low concentration in the reaction mixture.

The contaminating genomic DNA was removed from RNA extracts, prior to cDNA synthesis, by treatment with RNase-free DNase (Boehringer Mannheim, Germany) in the following reaction:

Table 3.1 Reaction components and concentrations required for the degradation of contaminating genomic DNA in the RNA preparation.

Reagent	Final concentration
RNA	~30µg
RNasin	40U
$MgCl_2$	10mM
DTT	1mM
RNase free DNase	1.25U
TE-buffer to	100µl

The reaction was incubated at 37°C for 15 minutes and terminated by adding 25µl of 50mM EDTA, 1.5M Na-acetate and 1% SDS. To the mixture was added premixed 37.5µl of Tris-buffered Biophenol, pH 7.8 - 8.2 (Biosolve B.V., Valkenswaard, The Netherlands) and 37.5µl of chloroform:isoamyl alcohol (24:1). After mixing by inversion and centrifugation for 2 minutes at ~13000g at 4°C, the upper aqueous phase was recovered and extracted with 75µl of chloroform:isoamyl alcohol (24:1). The tube was again mixed by inversion, centrifuged for 1 minute and the upper aqueous phase recovered. After addition of 325µl of ice-cold absolute ethanol, the tube was incubated at -20°C for 4½ hours and subsequently centrifuged for 20 minutes at ~13000g. The supernatant was aspirated and the pellet washed with 70% (v/v) ethanol, whereafter the tube was centrifuged for a further 8 minutes, the supernatant aspirated and the pellet allowed to air dry. The RNA was resuspended in 15µl of DEPC treated-water and the concentration determined spectrophotometrically (Ausubel *et al.*, 1992).

Both these treatments of total RNA (precipitation from FORMAzol and DNase treatment) were later discontinued as genomic DNA would most likely only interfere with target DNA amplification when gene specific primers are used and not when the

anchor primer is used in conjunction with a gene specific primer. cDNA was thus subsequently synthesised directly from unprecipitated, untreated total RNA.

3.2.6.3 Evaluation of cDNA preparations

cDNA derived from mRNA was first evaluated in terms of size distribution (using DIG-dUTP incorporation during synthesis) and later in terms of quality, i.e. the number of truncations present (using one gene specific primer with the anchor primer in PCR), while cDNA derived from total RNA was evaluated in terms of quality and quantity (using two gene specific primers for PCR).

- *Incorporation of DIG-dUTP into the cDNA*

The first cDNA synthesis reaction using mRNA as template, was monitored by the incorporation of DIG-dUTP (Boehringer Mannheim, Germany) into the cDNA. After annealing of the mRNA and poly-T primer as described above, 6µl was removed and added to a tube containing 100pmol DIG-dUTP. Synthesis and inactivation were completed as for normal cDNA. The DIG-incorporated cDNA was electrophoresed on a 1% agarose (Promega, Madison, USA) gel using TBE as running buffer. DIG-labelled high molecular weight marker (EcoRI/HinDIII digested lambda phage DNA) was also electrophoresed. The DNA was Southern blotted overnight onto a positively charged nylon membrane (Boehringer Mannheim, Germany) using capillary transfer (Ausubel *et al.*, 1992). The filter paper and nylon membrane were pre-wet in 20x SSC buffer (3M NaCl; 0.3M Na-citrate, SAARCHEM, Krugersdorp, South Africa), which was also the buffer used for capillary transfer. Subsequent to blotting, the membrane was blocked in 1% Elite milk powder/TBS (14mM Tris; 0.27mM KCl; 2.48mM NaCl, pH7.4) for an hour at room temperature, with gentle shaking. DIG-specific alkaline phosphatase was diluted 1 in 5000 in 0.1% Elite/TBS and added to the membrane. The membrane was incubated at room temperature for an hour, with shaking. The membrane was washed three times with 0.1% Elite/TBS (50ml, 5 minutes per wash). The substrate solution was prepared as follows: to 30ml of TBS was added, 750µl of $MgSO_4$ (final concentration of 50mM with the pH adjusted to between 9.2 and 9.5), 135µl of nitroblue tetrazolium (75mg/ml in 75% DMF, Boehringer Mannheim, Germany) and 135µl of X-phosphate (50mg/ml in water, Boehringer Mannheim,

Germany). The substrate solution was added to the membrane and the reaction allowed to proceed in the dark. Colour development was monitored visually and stopped by removing the substrate solution and rinsing the membrane with water, when sufficient colour was visible. The membrane was allowed to air dry.

- *PCR using specific primers*

The 3′-end truncated nature of the cDNA preparations was assessed by comparison of band sizes of the amplified product obtained with a forward primer, Bio1 (5′ ATCAGCTTTTGATGTTAGGGGGTATTG 3′, Tm: 60°C) specific for the 18S ribosomal RNA gene, in combination with the anchor primer. The full length product should be ~1814bp. Quantity was evaluated using Bio1 with Dig2 (5′ TAATAATAGAGGAAGCGTATT 3′, Tm: 50°C), a reverse primer specific for the 18S rRNA (Oliviera *et al.*, 1995). Both primers were obtained from Genosys Products (Cambridge, England). Bio1 and Dig 2 amplify a product of 408bp.

3.2.7 Optimisation of Polymerase Chain Reactions (PCR)

Template concentrations given for PCRs are in mRNA or total RNA equivalents, i.e. the amount of mRNA or total RNA which was used for the generation of cDNA. Theoretically, the concentration of cDNA would be between 1% and 5% of the total RNA used (Edwards *et al.*, 1995), but not all the mRNA is successfully reverse transcribed to cDNA. Total RNA derived cDNA is diluted by cDNA primed RNA fragments and non-polyadenylated RNA primed internally by the poly-T primer (CLONTECH Laboratories, 1991).

3.2.7.1 Template, primer and $MgCl_2$ concentrations

Various parameters used in PCR require optimisation. Of these, the $MgCl_2$, template and primer concentrations are the most important. These factors can initially be optimised in an orthogonal array PCR known as a Taguchi method (Cobb and Clarkson, 1994). This is an array where the effects of a number of variables and the interactions between them are investigated in a single experiment containing a few

reactions. Not all combinations are tested. Figure 3.17 shows examples of this type of optimisation.

Since the specific PCR conditions differ for almost all the experiments performed, the specific conditions are given in the figure legends. Thin-walled PCR tubes (Quality Scientific Plastics, CA, USA) were used in all PCRs. Controls included reactions with no template to confirm that no contamination had occurred. Reactions with only one primer were used to identify single primer driven products.

Conditions were optimised for primers: PSP2, PSP and TYT in several PCR experiments and typical conditions are given in Table 3.2.

Table 3.2 The typical PCR component concentrations for primers PSP2, PSP and TYT.

Reagent	**Concentration**		
	PSP2	**PSP**	**TYT**
Template	12ng[a]	200ng[b]	175ng[b]
10x Buffer	1x	1x	1x
$MgCl_2$	4mM	2mM	2mM
dNTPs	0.2mM G and C 0.3mM A and T	0.2mM G and C 0.3mM A and T	0.2mM G and C 0.3mM A and T
Gene specific primer	150pmol	60pmol	100pmol
Anchor primer	10pmol	10pmol	10pmol
Total volume	50µl	50µl	50µl
DNA polymerase	1.75U Expand™ HiFi	1.5U Takara *ExTaq*	1.5U Takara *ExTaq*

[a] The template concentration is given in mRNA equivalents.
[b] The template concentrations are given in total RNA equivalents.

PCRs were set up in the form of mastermixes to prevent contamination and to standardise the results.

3.2.7.2 DNA polymerases

Expand™ Hi Fi DNA polymerase (Boehringer Mannheim, Germany) is an enzyme which contains a mixture of *Taq* DNA polymerase and *Pwo* DNA polymerase for greater fidelity of the PCR. Takara *ExTaq* DNA polymerase (Takara Shuzo Co. Ltd. Shiga, Japan) is a similar type of formulation, but contains *Pfu* instead of *Pwo* DNA

polymerase. The cycling conditions for these reactions using the PTC-200 DNA Engine (MJ Research, Inc., Watertown, MA, USA) are given in Table 3.3.

Table 3.3 Optimal cycling conditions for each of the three primers, PSP2, PSP and TYT.

Primer	**PSP2**		**PSP**		**TYT**	
	Temp	**Time**	**Temp**	**Time**	**Temp**	**Time**
Denaturation	94°C	2'	94°C	30"	94°C	25"
Annealing	50°C	30"	55°C	15"	48.2°C	15"
Extension	72°C	1'30"	72°C	2'	72°C	2'
No of cycles	*10*		*31*		*30*	
Denaturation	94°C	20"				
Annealing	50°C	30"				
Extension	72°C	1'30" + 20"/cycle				
No of cycles	*20*					
Total no of cycles	*30*		*31*		*30*	

Hot starts were used in all instances where, after an initial denaturation period of 3 minutes, the reaction was paused at 80°C and the enzyme added (Kidd and Ruano, 1995). Normal cycling then commenced.

3.2.7.3 Annealing temperature

Stringency is an essential component of a PCR. Stringency can be increased by increasing the annealing temperature, decreasing the amount of time allocated for annealing or using a touchdown protocol (Don *et al.*, 1991; Hecker and Roux, 1996).

Annealing temperature can be calculated using various equations, namely the nearest neighbour method, the GC method or the AT method. The annealing temperature can also be determined by a computer program such as OLIGO which uses a more complex equation including the standard parameters such as GC content, primer length and the stability of the amplified template and least stable primer (See chapter 2).

3.2.8 Electrophoresis of DNA

Plasmid DNA was generally electrophoresed using 1% (w/v) agarose gel (Promega), while PCR products required 1.5% (w/v) or 2% (w/v) gels. Gels were prepared and run

in either 0.5x TBE buffer (45mM Tris-borate; 1mM EDTA) or 1x TAE buffer (40mM Tris-acetate; 1mM EDTA) using the following gel apparatus: Minicell® EC 370M and Maxicell™ EC 360M Electrophoretic Gel Systems (E-C Apparatus Corporation, St Petersburg, FL, USA) and Minnie Submarine Agarose Gel Unit, Model HE33 (Hoefer Scientific Instruments, San Francisco, CA, USA). Gels were run at 79V (Minicell®: 5.3V/cm and Minnie: 6.1V/cm) or 89V (Maxicell™: 2.8V/cm). Staining was conducted with ethidium bromide (10µg/ml) and bands were visualised on a transilluminator (312nm).

In later studies, when PCR products were to be cloned, gels contained crystal violet (10µg/ml) and were run in a buffer containing crystal violet (10µg/ml) to protect the DNA from UV damage (especially the AT-rich malaria genes) as the DNA bands were visualised on a light box. When the bands were excised and used for cloning, the cloning efficiency is improved (Rand, 1996).

3.2.9 Silica purification of nucleic acids

PCR products and plasmid DNA were purified from agarose gel fragments, or from solution, using a silica based method (Boyle and Lew, 1995). 2 volumes of 3M NaI were added to the tube containing either DNA in solution or in agarose gel. DNA-containing agarose gel fragments were first melted at 55°C for 5 minutes, whereafter 1mg of 100mg/ml silica (Sigma) in a 3M NaI solution was added per 3µg of DNA. The tube was incubated at room temperature with shaking for 30 minutes and then on ice for 15 minutes. It was subsequently microfuged at maximum speed (~13000g) for 15 seconds. The supernatant was removed and the silica pellet washed with washing buffer (50mM NaCl; 10mM Tris-Cl pH 7.5; 2.5mM EDTA; 50% (v/v) ethanol). The washing step was repeated twice and the DNA was eluted off the silica by adding water and heating at 55°C for 5 minutes. The tube was centrifuged briefly and the DNA-containing supernatant was transferred to a clean microfuge tube.

Two types of running buffers were used during the course of the study, namely TBE, which gave a weak recovery and TAE, which gave an excellent recovery of DNA from

excised gel slices (GENECLEAN *II*® Kit, Package insert, *BIO 101 Inc.*, La Jolla, CA, USA).

3.2.10 Cloning

Individual bands on an agarose gel were reamplified to obtain sufficient product for cloning. This was achieved by excising the band from the agarose gel then adding ~100µl 10mM Tris-HCl, pH 8. The gel was melted at 80°C and a fraction (between 2 and 10 µl) used as template for reamplification. A typical PCR using PSP2 and anchor would be set up along with the required controls.

Various methods of cloning were used but not all were successful.

3.2.10.1 Traditional blunt end cloning

The cloning vector pBluescript SK+ (1µg) was digested with 10 units of restriction enzyme (HincII) at 37°C, overnight. The restriction enzyme was inactivated by heating, followed by dephosphorylation with 1.5 units of calf intestinal alkaline phosphatase (CIAP, Boehringer Mannheim, Germany) for 30 minutes at 37°C. The enzyme was inactivated by the addition of 0.1 volumes of 0.2M EGTA, followed by heating at 65°C for 10 minutes. After extraction with Biophenol and chloroform:isoamyl alcohol as in section 3.2.6.2, the aqueous phase was retained and the DNA purified using silica as described above. The DNA concentration was determined using the ethidium bromide dot quantitation assay (section 3.2.2.3).

The products were phosphorylated in the supplied buffer using 20 units of polynucleotide kinase (Boehringer Mannheim, Germany), with 20mM dATP and incubation at 37°C for an hour followed by inactivation of the enzyme at 75°C for 10 minutes. The filled in PCR products were blunt-ended by the addition of 5 units of Klenow enzyme (Boehringer Mannheim, Germany), dGTP, dCTP and dTTP (4mM of each) (Kanungo and Pandey, 1993) and incubation at room temperature for 30 minutes, followed by inactivation at 75°C for 10 minutes. The modified PCR products were

purified using silica (section 3.2.9) and the concentration was determined using the ethidium bromide dot quantitation assay (section 3.2.2.3).

Blunt-end cloning was performed using plasmid and insert ratios calculated with the *j/i* ratios (Dugaiczyk *et al.*, 1975). Linear or circular products of ligation can be predicted by considering the total concentration of DNA termini *i*, and the local concentration of one terminus in the neighbourhood of the other on the same DNA molecule, *j*. With a ratio of one, intermolecular reactions, i.e. the ligation of the insert and plasmid, are promoted.

The equation is as follows:

$$j/i = \frac{j\lambda(^{MW}\lambda/_{MW})^{3/2}}{2N_0M \times 10^{-3}} = \frac{51.1}{[DNA](MW)^{½}}$$

where: MW is the molecular weight of the insert

MW_λ is the molecular weight of lambda phage and is 30.8×10^6

$j\lambda = 3.6 \times 10^{-11}$ ends/ml

N_0 is Avogadro's number

Calculated values indicated that 366ng of plasmid (~3000bp) and 848ng of insert (~560bp) were needed in a 10µl reaction volume. As the volumes needed to obtain the ratio exceeded 10µl, the plasmid and insert were co-precipitated with 0.1 volumes of 3M sodium acetate and 2 volumes of absolute ethanol, followed by incubation at –70°C for 15 minutes. The precipitated nucleic acids were pelleted by centrifugation at 14100g for 20 minutes, aspiration of the supernatant and washing of the pellet with 75% (v/v) ethanol. The nucleic acids were again pelleted by centrifugation and the pellet dried *in vacuo*.

The Rapid Link Ligation Kit (Epicentre Technologies, Madison, WI, USA) was used for ligation. rATP (0.75mM) and 3 Weiss units of T4 DNA ligase[a] were added and the reaction was incubated at room temperature for 1¾ hours. 1µl of the ligation reaction

was diluted to 10µl with water and ligated for a further 1¾ hours at room temperature. Dilution of the reaction decreases the concentration of DNA and thus promotes intramolecular (circularisation) reactions. The ligase was inactivated at 75°C for 15 minutes as it was shown that the enzyme interferes with the transformation of bacterial cells (Michelson, 1995). The ligation reaction was transformed into competent SURE *E. coli* cells.

3.2.10.2 Cloning using kits

Two types of kits were used for cloning, the pCR-Script™ Amp SK(+) Cloning Kit from Stratagene (La Jolla, CA, USA), which is a blunt end cloning kit and the pGEM® T-Easy Vector System from Promega, which is a T-vector cloning kit.

The PCR-Script™ Amp SK(+) Cloning Kit is based on the rare 8-base cutter, Srf1 (5′ GCCC↓GGGC 3′). The vector is a derivation of the pBluescript® II SK(+) phagemid in which the Sma1 site was mutated to a Srf1 site. PCR products are blunt-ended with *Pfu* DNA polymerase.

To obtain maximal results, PCR products were purified by selective precipitation prior to blunt-ending, by adding $^{1}/_{10}$ of the volume of 10x STE buffer (1M NaCl; 200mM Tris-HCl, pH 7.5; 100mM EDTA), an equal volume of 4M ammonium acetate and 2½ volumes of absolute ethanol at room temperature. Following centrifugation at 10 000g for 20 minutes (at room temperature), the supernatant was aspirated and the pellet washed with 200µl of 70% (v/v) ethanol. This was followed by centrifugation at 10 000g for 10 minutes at room temperature and the supernatant aspirated. The pellet was then dried *in vacuo* and resuspended in ~10µl of TE-buffer.

[a] 1 Weiss unit is the amount of enzyme that catalyses the exchange of 1nmol of ^{32}P from pyrophosphate into [γ, β-^{32}P] ATP in 20 minutes at 37°C.
1 Weiss unit = 0.2 units (exonuclease restriction assay) = 60 cohesive end units (New England Biolabs) (Sambrook *et al.*, 1989).

Approximately 200ng of the purified products were blunt-ended by the addition of 1µl of the supplied dNTP mix (2.5mM of each); 1.3µl of the 10x polishing buffer; 1µl of the cloned *Pfu* DNA polymerase (0.5 units) and deionised water to a total volume of 13µl. The reaction was incubated at 72°C for 30 minutes.

Ligation commenced with the addition of 1µl of 10x reaction buffer; 0.5µl of 10mM rATP; 2-4µl of the polished PCR products; 1µl of the Srf1 restriction enzyme (5 units); 1µl of T4 DNA ligase (4 Weiss units[a]) to 1µl of vector (10ng) and deionised water to 10µl. The reaction was incubated at room temperature for one hour followed by inactivation of the ligase enzyme at 65°C for 10 minutes. 2µl of the ligation reaction was transformed into competent DH5α *E. coli* cells as described below. Stratagene recommends a molar ratio of between 40:1 and 100:1 (insert:vector) as a starting point. Later studies showed that a lower ratio of insert:vector (between ~15:1 and ~25:1, depending on PCR products) was more successful.

The pGEM T-Easy Vector System is an AT cloning kit. The kit requires tailing of the PCR products using *Taq* polymerase to add an extra 'A' to the 3′ end of the product, if extensions were with proof-reading DNA polymerases. Purified products should be used for tailing and purification can be accomplished using either precipitation or silica purification. The tailing reaction was set up as follows: 1µl of *Taq* DNA polymerase reaction buffer (10x), 1µl of 25mM $MgCl_2$, dATP to a final concentration of 0.2mM, 5U of *Taq* DNA polymerase and deionised water to 10µl were added to the purified PCR product. The reaction was incubated at 70°C for between 15 and 30 minutes and 1-2µl of the tailed product was used in the ligation reaction.

Ligation was conducted with 10x ligation buffer (1µl); 1µl of pGEM T-Easy vector (50ng); 1-2µl of tailed PCR product; 1µl of T4 DNA ligase (3 Weiss units[a]) and deionised water to 10µl. The reaction was incubated at 4°C overnight, followed by inactivation of the enzyme at 70°C for 15 minutes. The ligation reaction (2µl) was transformed into competent DH5α *E. coli* cells as described below. A ratio of between 3.5 and 4:1 (insert to vector) was used. This was successful, although the optimum ratio required empirical determination.

3.2.11 Preparation of competent cells

Bacterial cells were initially grown on minimal medium agar plates such as M9 (42.25mM $Na_2HPO_4.2H_2O$; 22.04mM KH_2PO_4; 18.7mM NH_4Cl; 8.56mM NaCl; 1.5% (w/v) agar, pH 7.4. 2mM $MgSO_4$; 11mM glucose; 0.1mM $CaCl_2$; 1mM thiamine HCl were added after autoclaving and cooling of the first components to maintain the *lacZ* genotype, prior to making them competent.

Competent SURE *E. coli* or DH5α *E. coli* cells were prepared by inoculation of 1ml of overnight bacterial culture (from M9 plates) into 100ml of LB-broth (1% Tryptone (Biolab Diagnostics, Midrand, South Africa); 1% NaCl; 0.5% yeast extract (DIFCO Laboratories, Detroit, MI, USA), pH 7.5) containing 12.5µg/ml tetracycline (for SURE *E. coli* only). The culture was incubated with vigorous shaking at 37°C for 3 hours (OD between 0.3 and 0.5, i.e. log phase of growth). The cells were incubated on ice for 5 minutes and pelleted by centrifugation at 3000g for 10 minutes (4°C). They were then resuspended in 40ml of transformation buffer 1 (30mM K-acetate; 100mM RbCl (Sigma); 10mM $CaCl_2.2H_2O$; 50mM $MnCl_2.4H_2O$; 15% (v/v) glycerol; pH 5.8), incubated on ice for 5 minutes and pelleted by centrifugation at 3000g for 10 minutes. The supernatant was aspirated and the pellet resuspended in 4ml of transformation buffer 2 (10mM MOPS; 75mM $CaCl_2.2H_2O$; 10mM RbCl; 15% (v/v) glycerol; pH 6.5), incubated on ice for 15 minutes, aliquotted and frozen at -70°C.

3.2.12 Transformation of competent cells

The competent SURE or DH5α *E. coli* cells were transformed with ligation mixture by incubating 100µl (or 40µl) with $^1/_{10}$ of the ligation reaction in microfuge tubes on ice for 30 minutes. After a heat shock of 42°C for 90 seconds, incubation on ice for 2 minutes, 900µl (or 450µl) of prewarmed (37°C) LB-broth containing 20mM glucose was added and incubated at 37°C for an hour. Then, 100µl of cells were plated onto LB-agar plates (1.5% agar) containing 100µg/ml ampicillin (Boehringer Mannheim, Germany) and with 800µg of X-gal (Boehringer Mannheim, Germany) and IPTG (Boehringer Mannheim, Germany) spread on the surface (DH5α cells do not require

IPTG). After the plates were incubated at 37°C overnight, white colonies were selected and inoculated into 1ml of LB-broth containing 50µg/ml ampicillin and incubated at 37°C with vigorous shaking. After six hours, the cultures were scaled up to 10ml and incubated overnight. Positive clones were stored at -70°C in 15% glycerol.

It was subsequently found that using glass tubes (sterile Vacutainers without anticoagulants) for a heat shock of 45 seconds (reduced from the 90 seconds used previously), greatly improved transformation efficiency and this modification was thus used henceforth.

3.2.13 Plasmid isolation and restriction digestion

Plasmid DNA was isolated using a conventional mini-prep procedure (Sambrook *et al.*, 1989). The 10ml of overnight culture was centrifuged at 3000g for 10 minutes to pellet the cells, the supernatant aspirated and the cells resuspended in 200µl of Solution 1 (50mM glucose; 25mM Tris-Cl; 10mM EDTA). To this was added 200µl of Solution 2 (0.2M NaOH; 1% (w/v) SDS) and mixed by end-over-end mixing. The suspension was incubated on ice for 5 minutes, followed by the addition of 200µl of Solution 3 (3M potassium-acetate; pH 5.4). The tube was incubated on ice for a further 5 minutes, followed by centrifugation at 17300g for 15 minutes. To the supernatant was added 600µl of Biophenol (pH 7.8 - 8.2) and the mixture was vortexed for a minute, followed by centrifugation for 2 minutes at top speed (room temperature). To the upper, aqueous phase was added 600µl of chloroform: isoamylalcohol (24:1) and mixed. The tube was centrifuged briefly and 1ml of ice-cold absolute ethanol was added to the upper, aqueous phase, followed by incubation at -70°C for 20 minutes. The precipitated DNA was pelleted by centrifuging at 17300g for 15 minutes at 4°C. After aspiration of the supernatant, the pellet was washed with 1ml of 75% ethanol, followed by centrifugation at 17300g for a further 7 minutes, aspiration of the supernatant and drying of the pellet *in vacuo*. The pellet was subsequently dissolved in 40µl of TE buffer (10mM Tris; 1mM EDTA) containing 0.4µg/µl RNase.

Plasmid (pBluescript) DNA (between 1 and 1.5µg) was digested using 6 units each of KpnI and BamHI (Promega, Madison, WI) in the Multi-Core buffer supplied with the enzymes, at 37°C for 2 hours. Tracking dye was added and the plasmid electrophoresed on a 1% (w/v) agarose (Promega) /TBE gel in 1x TBE. The gel was stained in ethidium bromide and visualised at 312nm. For the clones obtained with pCR-Script, the plasmid DNA was digested as above, but with 6 units of PVuII in 1x Buffer B (Promega).

3.2.14 Manual nucleotide sequencing

Plasmid DNA for nucleotide sequencing, was isolated using the High Pure Plasmid Isolation Kit from Boehringer Mannheim (Germany); the Wizard™ *Plus* Minipreps DNA Purification System from Promega (WI, USA); or the Plasmid Midi Kit from QIAGEN (Hilden, Germany). All three kits are based on similar principles although the QIAGEN kit uses an anion exchange resin, in contrast to the silica based membranes of the others. This means that there is a difference in the adsorption and elution conditions. Figure 3.6 shows the general method of plasmid isolation.

In the High Pure and Wizard™ kits, the plasmid DNA is adsorbed under high salt conditions and eluted with low salt buffers, while in the QIAGEN kit, DNA is adsorbed under low salt conditions and eluted using high salt buffers. In the latter case, DNA thus obtained is desalted by precipitation with 0.7 volumes of isopropanol at room temperature. This is centrifuged immediately at 13600g for 30 minutes and the supernatant aspirated. The pellet is washed with 1ml of 70% (v/v) ethanol and recentrifuged at 13600g for 10 minutes. The supernatant is aspirated and the pellet allowed to air dry. The DNA pellet is resuspended in an appropriate volume of TE-buffer.

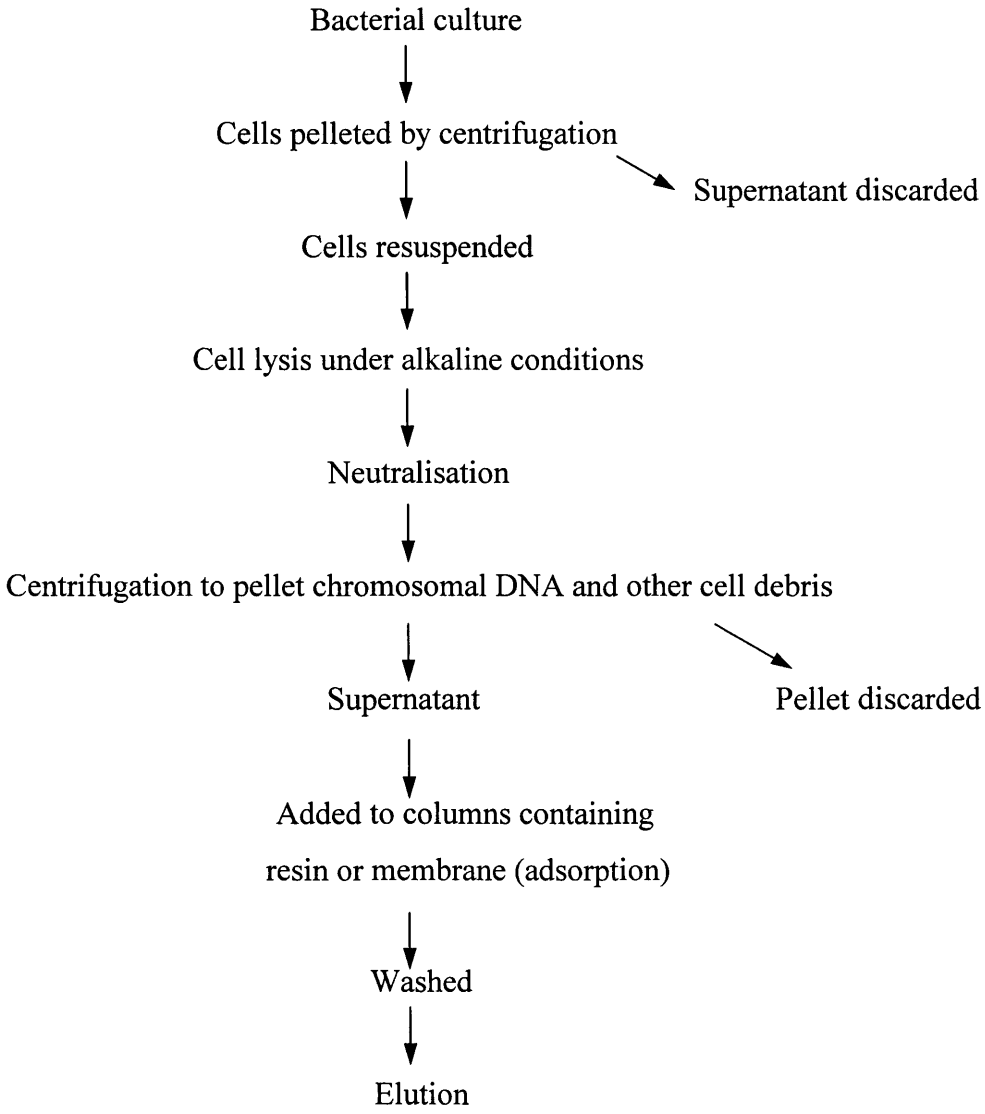


Figure 3.6 Summary of the protocol for all the plasmid isolation kits used.

DNA (5µg for each direction sequenced) was denatured by the addition of 0.1 volumes of 2M NaOH; 2mM EDTA and incubation at 37°C for 30 minutes. This was followed by the addition of 0.1 volumes of 3M Na-acetate (pH 5.5). The DNA was precipitated by the addition of 4 volumes of ice-cold absolute ethanol and incubation at -70°C for 15 minutes. The precipitated DNA was pelleted by centrifugation at 17300g for 20 minutes (4°C). The supernatant was aspirated and the pellet washed with 100µl of ice-cold 75% (v/v) ethanol, followed by centrifugation at 17300g for a further 10 minutes. The supernatant was removed and the pelleted DNA dried *in vacuo*. The pellet was dissolved in 7µl of water (per sequencing direction).

3.2.14.1 Annealing

The sequencing primer sites were located on each side of the multiple cloning site of pBluescript and were the T3 and T7 promoters, respectively. The primers were obtained from Oswel DNA Services (Edinburgh, Scotland) and had the following sequences:

- **T3** 5′ AATTAACCCTCACTAAAGGG 3′ (Tm: 50°C)
- **T7** 5′ GTAATACGACTCACTATAGGGC 3′ (Tm: 50°C)

A mixture containing pelleted DNA (2.15pmol), 4.2pmol of the appropriate primer (either T3 or T7 promoter specific in a 2:1 primer:template ratio), and 2µl of 5x reaction buffer (200mM Tris-Cl pH 7.5; 100mM $MgCl_2$; 250mM NaCl) was heated at 65°C for 2 minutes, followed by cooling to 35°C over 30 minutes and incubation on ice.

3.2.14.2 Labelling

To the annealed DNA, 1µl of 0.1M DTT was added, followed by 2µl of diluted labelling mix (1.5µM each of dGTP, dCTP, dATP and dCTP), 12.5µCi of α^{35}S-dATP (Amersham Laboratories, Buckinghamshire, England), 3.25 units of diluted Sequenase® Version 2 (United States Biochemicals, Cleveland, OH, USA) and the incubation on ice for 30 minutes.

3.2.14.3 Termination

Annealing reaction mixture (3.5µl) was transferred to each of four tubes containing 2.5µl of termination mixture (80µM each of dCTP, dGTP, dATP and dTTP; 50mM NaCl; and 8µM of the appropriate ddNTP). The termination reaction was incubated at 37°C for 5 minutes and stopped by the addition of 4µl of stop solution (95% (v/v) formamide; 20mM EDTA; 0.05% (w/v) Bromophenol Blue; 0.05% (w/v) Xylene cyanol FF).

3.2.14.4 Electrophoresis

A 6% (w/v) acrylamide, 7M urea (BDH Laboratory Supplies, Poole, England) gel was prepared using 0.8M acrylamide and 0.02M *N,N'*-methylene bis-acryalamide (BDH Laboratory Supplies, Poole, England) and 7ml of TBE buffer (0.45M Tris, 0.45M boric

acid and 0.1M EDTA). This mixture was heated briefly and made up to 70ml with deionised water. Following filtration and degassing, 40µl of TEMED and 545µl of 10% fresh ammonium peroxidisulphate was added and the gel was cast and allowed to polymerise overnight. The gel was pre-electrophoresed at 1200V for 20 minutes in 1x TBE using an Owl sequencing apparatus. Prior to loading on the gel, the termination reactions were heated to 75°C for 2 minutes and snap-cooled on ice. From each tube, 3µl was loaded onto the gel and electrophoresed at a constant power of 75W (~1600V). Following electrophoresis, the gel was removed from the glass plates using Whatman 3MM filter paper (Whatman, USA) and dried at 80°C for 2 hours under vacuum (Hoefer Scientific Instruments, Inc., San Francisco, CA, USA). Kodak Biomax X-ray film (MR, Eastman Kodak Company, Rochester, NY, USA) was exposed to the dried gel for 2 days in a cassette (Sigma, St. Louis, MO, USA). The film was developed using Polycon A (Champion photochemistry, Midrand, South Africa) and fixed with Ilford Hypam (Iso Photo, Rivonia, South Africa) for 5 minutes each.

The sequence was determined from the exposed X-ray film and evaluated by sending it to BLAST (Basic Local Alignment Search Tool), (Altschul *et al.*, 1990) for homology matching, via the Internet. An important factor to note is that the identified consensus sequence for the amino acid permeases was used for the PCR primer. This sequence is thus unavailable for identification of a potential amino acid permease gene sequence. Identification of a amino acid permease would thus have to be based on some other consensus area (there are some regions available for specific amino acid transporters), but most likely the decision would be based on the identification of suitable transmembrane regions. The AT-content also plays a role to rapidly determine if the gene is potentially of malarial origin.

3.2.15 Automated nucleotide sequencing

Automated nucleotide sequencing was originally performed using the ABI PRISM™ 310 automated nucleotide sequencer and subsequently the ABI PRISM™ 377 (PE Applied Biosystems, Foster City, CA, USA).

Fluorescent dye terminator nucleotides were incorporated into the product using PCR. Originally the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit was used. This was later substituted with the Big Dye Kit. Both kits were from PE Applied Biosystems, Foster City, CA, USA. The Big Dyes are a marked improvement over previous standard dye terminators as they are 2-3 times brighter than the standard dye terminators and have a narrower emission spectra resulting in less spectral overlap and thus conferring 4-5 fold increased sensitivity. The cycle sequencing reactions for both are identical. The PCR conditions were as follows:

Table 3.4 Typical reagent composition of a cycle sequencing PCR.

Reagent	Concentration
Terminator Ready Reaction Mix	8µl
Template	300-500ng
Primer (T3 or T7 for pBluescript vectors)	3.2 pmol
Water to	20µl.

These reaction conditions were halved and also quartered. For sequencing of the pGEM vector in the pGEM T-Easy Vector System, T7 or SP6 sites were used for the T7 and SP6 (Promega) primers, respectively. The cycling conditions were performed on the GeneAmp® PCR system 9700 from PE Applied Biosystems (Foster City, CA, USA):

Table 3.5 Cycling conditions for a cycle sequencing PCR.

Step	Temp	Time
Denaturation	96°C	10"
Annealing	50°C	5"
Extension	60°C	4'
No of cycles	*25*	

PCR extension products from the standard dye terminators were purified by precipitation with 2µl of 3M Na-acetate and 50µl of absolute ethanol, followed by incubation on ice for 10 minutes and centrifugation at ~15000g for 15-30 minutes. The supernatant was aspirated and the DNA pellet washed with 250µl of 70% (v/v) ethanol. The tube was centrifuged briefly and the supernatant was aspirated and the pellet dried *in vacuo*. Big Dye reactions were pelleted by the addition the reaction

contents to 16µl of water and 64µl of non-denatured 95% (v/v) ethanol (to a final concentration of 60 ± 3%). The tube was incubated at room temperature for 10 to 15 minutes and centrifuged at ~15000g for 15 minutes. The supernatant was aspirated and the pellet washed with 100µl of 70% (v/v) ethanol. The tube was centrifuged for a further 10 minutes, the supernatant aspirated and the pellet dried *in vacuo*.

The resuspension solution depended on the type of sequencer used. Reactions that were electrophoresed on the 310 sequencer were resuspended in 25µl of the supplied Template Suppression Reagent while samples for the 377 sequencer were resuspended in 3µl of the supplied loading buffer.

The raw data was initially analysed with the analysis software ABI PRISM™ Sequence Analysis, Version 3.0 (PE Applied Biosystems, 1989-1996) and confirmed by visual inspection. Sequences obtained were evaluated by comparison to those in the NCBI databases by searching with BLAST.

3.2.16 Screening of recombinant clones using restriction digestion

In order to be able to differentiate between recombinant clones without sequencing all of them, a few were sequenced and the restriction sites determined using DNA Strider™ Version 1.0 (Christian Marck, Commissariat a l'Energie Atomique, France).

The clones obtained with primer PSP were digested as follows:

Table 3.6 Conditions for the restriction digestion screening of recombinant clones.

Reagent	Volume
Plasmid	1µg
Buffer E (Promega)	1.5µl
HinDIII (10U/µl, Promega)	0.5µl
Water to	15µl

The plasmids were digested with HindIII at 37°C for an hour, then electrophoresed on a 1.2% agarose (Promega) gel. The restriction patterns obtained were compared and the tentative identities conferred on the unsequenced clones. All the clones identified in this manner, were subsequently sequenced to confirm the allocations conferred.

3.2.17 RNA blot of the novel ribosomal protein to determine its origin

All the reagents and apparatus used for this experiment were made RNase free, by baking of glassware, the use of DEPC-treated water and double autoclaving of appropriate reagents.

3.2.17.1 Labelling of the probe with DIG-dUTP using PCR incorporation

An asymmetric PCR was used and it was thus important to first determine which primer was required for production of the strand complementary to the RNA. This was done by reverse translation of both sides of the sequence obtained for Clone 7a using ABI PRISM™ Sequence Navigator Version 1.0.1 (PE Applied Biosystems, 1989). This showed that the strand produced using primer T3 would yield the probe in the correct orientation.

The plasmid was linearised with BamHI which cleaves an unique site in the multiple cloning site and not in the insert. The restriction site is very close to the insertion site resulting in less vector bases in the probe. 1μg of plasmid was linearised using 6 units of BamHI (Promega, Madison, WI, USA) in the appropriate buffer at 37°C for 2 hours and complete digestion was confirmed by electrophoresis.

Table 3.7 shows the PCR reaction that was set up for the generation of the DIG-labelled single stranded probe, while Table 3.8 lists the cycling conditions.

Table 3.7 Conditions used in the linear PCR used to generate a single stranded probe.

Reagent	Concentration
Plasmid	20ng
Buffer	1x
$MgCl_2$	1.5mM
dNTPs	0.2mM G and C 0.3mM A and T
DIG-dUTP	3µM
T3 primer	25pmol
Taq DNA polymerase	1U
DEPC treated-H_2O to	25µl

Table 3.8 Cycling conditions for the generation of a single stranded probe, using the GeneAmp® PCR system 9700.

Step	Temperature	Time
Denaturation	94°C	2'
Annealing	50°C	30"
Extension	72°C	1 ½'
No of cycles	*40*	

The results of the PCR were evaluated by electrophoresis.

Incorporation of DIG-dUTP into the probe was confirmed by cross-linking the probe onto positively charged nylon membranes (Boehringer Mannheim, Germany) and detection using CDP-Star™.

3.2.17.2 Hybridisation

RNA (6.22µg/µl) was serially diluted to 3µg/µl; 2µg/µl; 1µg/µl; 0.5µg/µl and 0.1µg/µl and 1µl was spotted onto the nylon membrane. The negative controls included 2µg of mouse liver RNA (a gift from Sandra van Wyngaardt); 1µg of Brome Mosaic Virus RNA (Promega); 1µl of FORMAzol and 1µl of DEPC treated-H_2O. 0.5µl of DIG-ddUTP labelled oligonucleotide was used as positive control. The nylon membrane was exposed to the transilluminator (312nm) for 3 minutes to cross-link the nucleic acids.

The membrane was prehybridised in 2ml of high SDS buffer (7% (w/v) SDS; 50% (v/v) deionised formamide; 5x SSC; 2% (w/v) blocking reagent (Boehringer

Mannheim, Germany); 50mM sodium phosphate, pH 7; 0.1% (w/v) N-lauroylsarcosine, Sigma) at 50°C for one hour.

The probe was denatured at 96°C for 10 minutes and added to the hybridisation (high SDS) buffer (25ng/ml). The membrane was placed in a clean plastic bag and the hybridisation solution added. Hybridisation was performed overnight at 50°C.

The membrane was washed twice for 15 minutes in 2x wash solution (2x SSC; 0.1% (w/v) SDS) at room temperature, followed by two washes for 15 minutes each in 0.5x wash solution (0.5x SSC; 0.1% (w/v) SDS) at 68°C. 20x SSC contained 3M NaCl and 300mM sodium citrate, pH 7.0.

The membrane was equilibrated in washing buffer (100mM maleic acid buffer; 150mM NaCl, pH 7.5) for 1 minute, blocked in blocking solution (1% (w/v) blocking reagent in maleic acid buffer) for 30 minutes and then incubated with antibody solution (anti-digoxygenin Fab fragments diluted 1:10 000 in blocking solution) for 30 minutes. The membrane was subsequently washed twice, for 15 minutes respectively, in washing buffer. After equilibration in detection buffer (100mM Tris-Cl, pH 9.5; 100mM NaCl), the membrane was incubated in CDP-Star™ (diluted 1:100 in detection buffer) for 5 minutes and exposed to Cronex® 4 X-ray film (manufactured under licence from E. I. Du Pont De Nemours and Co. Inc. by Protea, SA) for 9 minutes, followed by development and fixation of the film (3 minutes each).

3.2.18 Definitions:

In all instances, 'M' refers to the high molecular weight marker, which is λ phage DNA digested with EcoRI and HinDIII (Promega, Madison, WI, USA). 'm' refers to the low molecular length marker which is pBR322 digested with HaeIII (Boehringer Mannheim, Germany). '100bp' refers to the 100bp DNA ladder (Promega).

3.3 Results

RNA was isolated from ~650ml of unsynchronised *P. falciparum* cultures in the ring and trophozoite stages (collected 24 hours apart) respectively, at a parasitaemia of ~20%. These preparations were pooled and the concentration determined. The spectrophotometric quantitation showed an RNA concentration of 6.22µg/µl and thus a total RNA yield of 1.99mg.

Electrophoresis of the samples (Figure 3.7) shows three bands. The upper band is 28S rRNA, the middle band represents the 18S rRNA and the lower band is representative of 5S rRNA.

High integrity RNA has the following characteristics:

- the 28S band should be twice as intense as the 18S band; and
- the 5S band should be less intense than both the 28S and the 18S RNA bands (Edwards *et al.*, 1995).

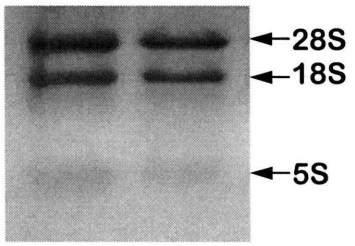


Figure 3.7 Denaturing agarose gel electrophoresis of RNA. ~5µg electrophoresed on a 1% gel (Promega).

3.3.1 cDNA derived from isolated mRNA

First strand cDNA synthesis from isolated mRNA, was monitored using the incorporation of the steroid hapten, DIG. The DIG-dUTP incorporated cDNA that was synthesised from the isolated mRNA centred between 400bp and 1500bp (Figure 3.8).

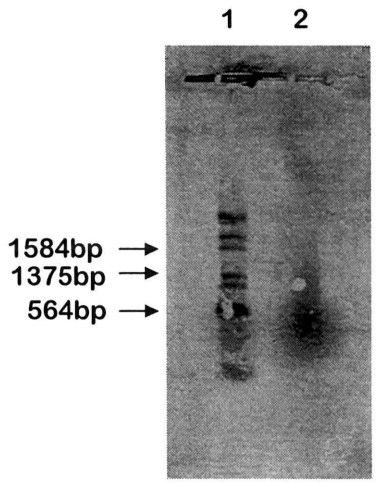


Figure 3.8 Southern blot of the DIG-dUTP incorporated cDNA, detected using anti-DIG antibodies and visualised colourimetrically using X-phosphate and NBT. Lane (1) DIG labelled high molecular weight marker; lane (2) synthesised cDNA, electrophoresed on a 1% agarose (Promega) gel.

3.3.3.1 Optimisation of PCR using PSP2

A variety of factors were optimised for primer PSP2, using cDNA derived from mRNA. These included template, $MgCl_2$, PSP2 and anchor primer concentrations and the annealing conditions (temperature and protocol). All the PCR conditions given are for 50µl reaction volumes.

- *Template concentration*

Initially, the ~1ng template concentration recommended by Frohman (Frohman, 1993) was used. No products were visible after 30 cycles (results not shown). The template concentration was subsequently tested at concentrations of ~14ng and ~100ng mRNA equivalents.

The banding pattern obtained using increased cDNA concentrations was similar to that obtained with less cDNA, though the larger ~720bp bands were more clearly defined. No new bands were observed.

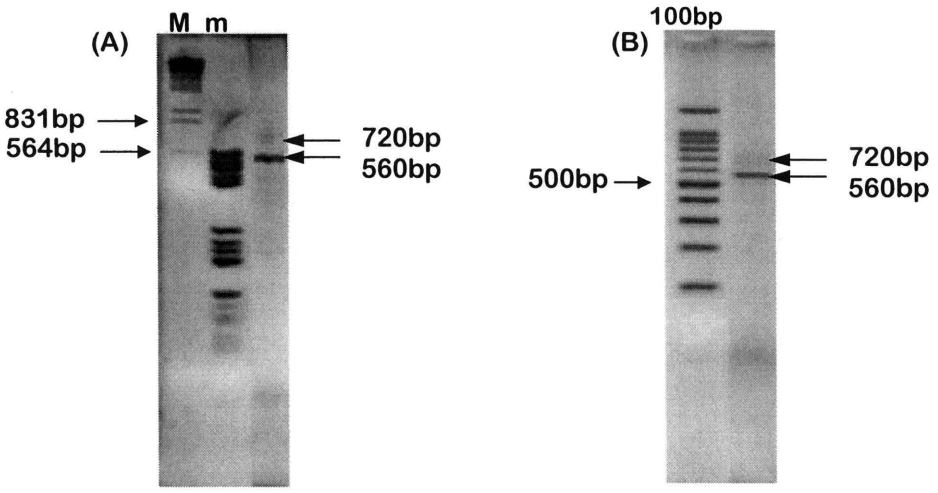


Figure 3.9 PCR to evaluate the effect of template concentration. PCR products obtained with 150pmol PSP2, 10pmol anchor **(A)** 14ng of mRNA equivalents with 4 cycles at 48°C and 26 cycles at 50°C and 2.6U of Expand™ HiFi electrophoresed on a 2% NuSieve agarose gel and **(B)** 100ng of mRNA equivalents, Ta of 50°C, 30 cycles and 1.75U of Expand™ HiFi DNA polymerase, electrophoresed using 2% NuSieve:Promega mix agarose gel.

- *$MgCl_2$*

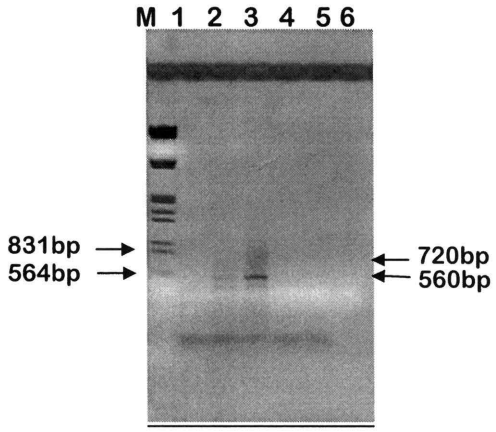


Figure 3.10 Optimisation of $MgCl_2$ requirements. PCR products obtained using 12ng of mRNA equivalents, 150pmol PSP2, 10pmol anchor, 1.75U of Expand™ HiFi and a Ta of 48°C for 2 cycles and 50°C for 28 cycles. The $MgCl_2$ was optimised over lanes 1-3 as follows: (1) 2mM; (2) 3mM and (3) 4mM. Lanes 4-6 were negative controls as follows: (4) no template; (5) no anchor and (6) no PSP2, included.

All subsequent PCRs performed with Expand™ HiFi DNA polymerase used 4mM $MgCl_2$ for optimum amplification.

- *Anchor primer concentration*

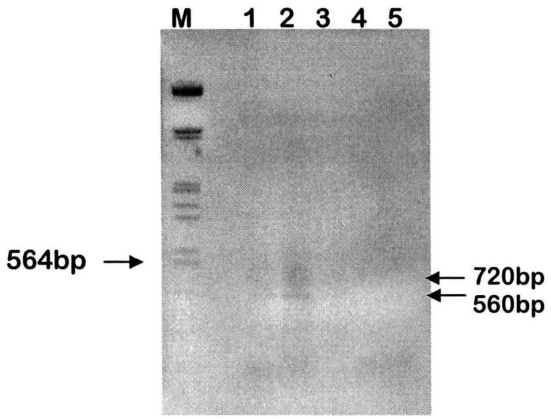


Figure 3.11 Optimisation of anchor primer concentration. PCR products obtained using primer PSP2 (150pmol) and 12ng of mRNA equivalents. Lane (1) 5pmol of anchor and lane (2) 10pmol of anchor. Lanes 3-5 were the negative controls as follows (with 10pmol of anchor): (3) no PSP2; (4) no anchor and (5) no template added.

10pmol of anchor appears to offer better amplification. The bands appear weaker than in Figures 3.8 and 3.9. This may be due to a deterioration of the Expand™ HiFi DNA polymerase mix (see later). The PSP2 primer concentrations were tested at 45pmol, 100pmol and 150pmol. Of these concentrations 150pmol of PSP2 appeared to yield the best amplification (results not shown). All subsequent PCRs used 150pmol of PSP2 and 10pmol of anchor.

- *Annealing conditions*

PCR was performed using an annealing temperature of 48°C for the first four cycles (16 times amplification of the target) as the annealing temperature recommended by OLIGO was 47.6°C. This was done to generate template and the annealing temperature was changed to 50°C after the first four cycles to increase the specificity of the reaction. There is little difference between using 48°C initially (Figure 3.12A - lane 1) or just using a constant annealing temperature of 50°C (Figure 3.12B). An annealing temperature of 50°C was therefore used henceforth. Figure 3.12A also shows two different types of annealing protocols, namely the use of a constant annealing temperature versus a touchdown protocol. No advantage was gained by using the

touchdown protocol. The slight smearing near the well in Figure 3.12A was typically obtained when excessive Expand™ HiFi DNA polymerase is used.

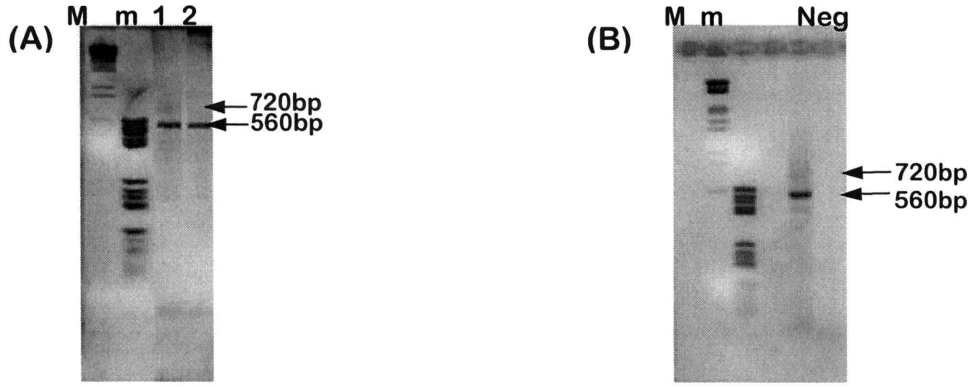


Figure 3.12 Investigation of annealing conditions for primer PSP2. **(A)** Comparison of a constant annealing temperature (4 cycles at 48°C and 26 cycles at 50°C) in Lane 1 vs. a touchdown procedure (53°C to 47°C) in Lane 2. The PCR samples were electrophoresed on a 2% NuSieve agarose gel. 14ng of mRNA equivalents and 2.6 units of Expand™ HiFi DNA polymerase were used **(B)** 12.9ng mRNA equivalents, Ta of 50°C, 40 cycles and 2.275U Expand™ HiFi on a 2% NuSieve:Promega (¾ Nusieve: ¼ Promega) mix agarose gel. 'Neg' is the negative control containing both primers, but no template.

The number of cycles used can also be compared using Figure 3.12. Gel A had 30 cycles (~1 x 10^9 times amplification) whereas Gel B had 40 (~1 x 10^{12} times). There is no noticeable difference between the bands obtained indicating that the reaction is most likely in the plateau phase. However, the 720bp band is more clearly defined after 40 cycles, but this could be due to the amplification of a secondary target after the first target (560bp) has reached the plateau phase.

3.3.1.2 Cloning and sequencing of the bands generated with primer PSP2

The 560 and 720bp bands, obtained in Figure 3.12B (Ta of 50°C), were excised together and reamplified to obtain sufficient product for cloning. The PCR products were purified from the reaction mixture using silica and subsequently cloned.

The method of cloning which proved successful was a traditional blunt end method, where the vector is dephosphorylated using CIAP, the insert is phosphorylated using polynucleotide kinase and then filled in with Klenow enzyme. Ligation was performed using the calculated *j/i* ratio of insert to vector, with pBluescript SK+ as vector (see

Appendix C, Figure 1 for the genetic map of the vector). The ligated products were transformed into competent SURE *E. coli* cells. White colonies were screened for recombinants by restriction digestion.

Sixteen clones were obtained and the plasmid DNA isolated and digested using BamHI and KpnI. Figure 3.13 shows digested plasmids electrophoresed on a 1% agarose gel. (Note: this type of restriction digestion was also always done prior to sequencing of the clones to determine the size of the inserts, although the results are not shown every time.)

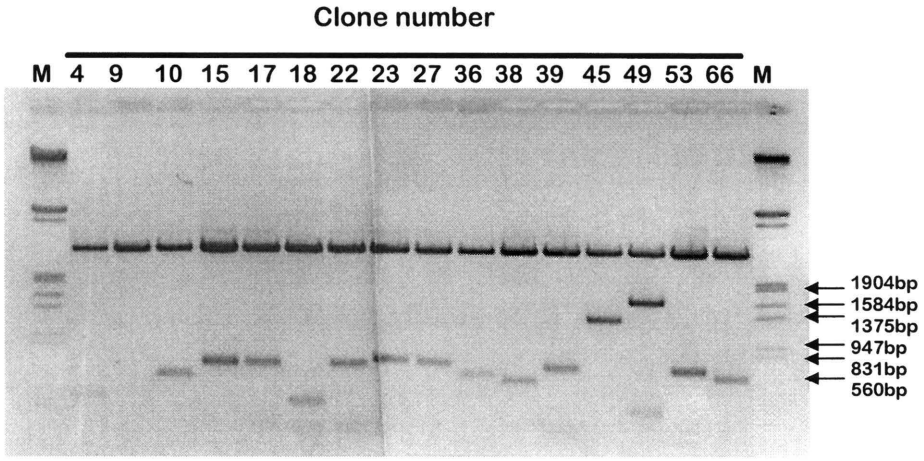


Figure 3.13 Restriction digestion of recombinant plasmids with BamHI and KpnI. The digested plasmids were electrophoresed using 1% Promega agarose.

Most of the clones showed an insert of a size corresponding to the 560bp and 720bp fragments cloned. The exceptions were clones 4 (~500bp) and 18 (~400bp), which were smaller than expected and clones 45 (~1350bp) and 49 (~1600bp) which were larger than expected. Clone 9 appears to have no insert, but the excised band may just be very faint. The insert sizes observed in clone 18, 45 and 49, were unforeseen as only the 560bp and 720bp bands were visible on the gel after the reamplification PCR. The other fragment sizes may have been present in low concentrations in the background, below the detection limit of ethidium bromide. Manual sequencing with the Sequenase® Version 2.0 kit was performed and Figure 3.14 shows part of a typical autoradiogram obtained.

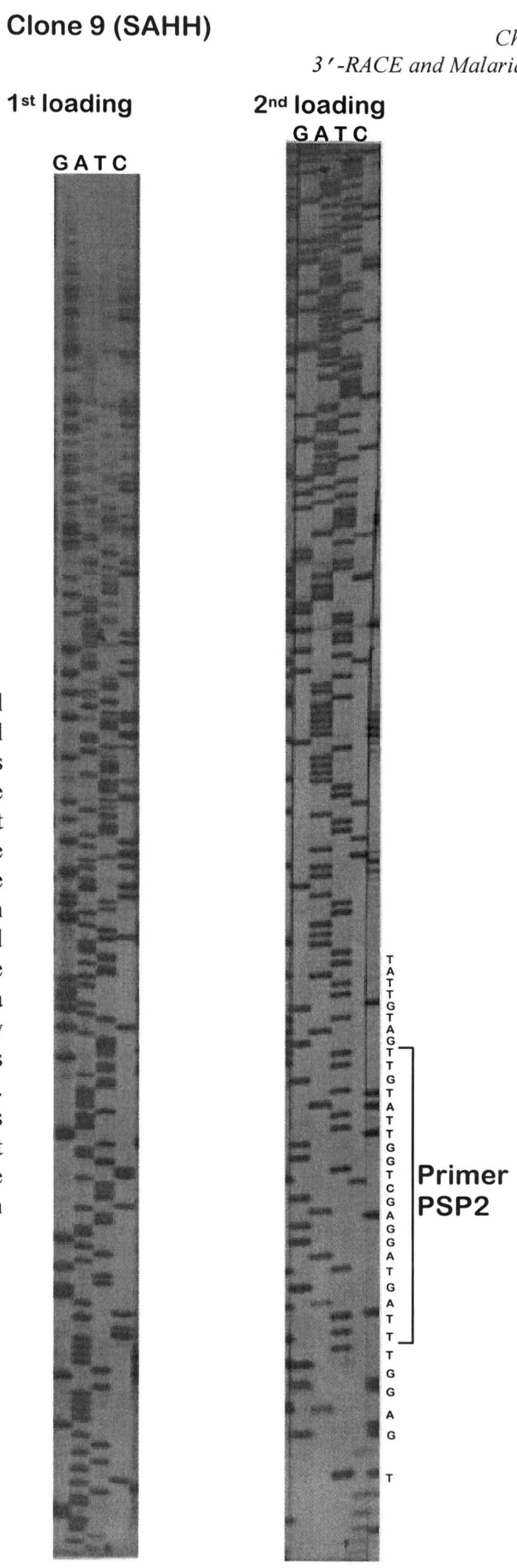


Figure 3.14 Part of a typical autoradiogram obtained from manual sequencing. Clone 9 (SAHH) was sequenced from the T7 side of the multiple cloning site. Both the first and second loadings of the termination reactions are shown. The first loading is when the termination reactions were loaded onto the gel and electrophoresed for ~2 hours, the reactions were then loaded for a second time (second loading) in new wells and all the samples electrophoresed for a further 2 hours. The purpose of loading the samples twice is so that the maximum amount of sequence information can be obtained from one set of termination reactions, on one gel.

Sequences could be read to an average of ~200bp with a maximum of ~245bp reached, per loading.

Upon determination of the sequence, it was submitted to BLAST for homology matching. The results obtained are shown in Table 3.9.

Table 3.9 Homology matches for the partial sequences obtained with manual sequencing, as determined by BLAST (alignments are given in Appendix B). The underlined clones are those where the sequence was determined using only the T7 primer. The remaining clones were sequenced from both the T3 and T7 promoter sides of the multiple cloning site.

Clone	Gene
9; 36; 66	*P. falciparum* S-adenosylhomocysteine hydrolase (SAHH)
15; 17; 22; 23; 39	*P. falciparum* nuclear GTP-binding protein homologue mRNA
10; 38	*P. falciparum* mRNA for heat shock protein 90
4	?
18	*E. coli* helicase
27	*E. coli* RecC
45	?
49	?
53	?

The obtained gene sequences are obviously not related to those expected for amino acid transporter genes. It should also be noted that bands of the same size did not necessarily have the same gene sequence. The *E. coli* genes (Clones 18, 27) could be the result of either contamination or rearrangement/recombination events. The absence of PCR primers in clones 18, 27, 45, 49 and 53 tends to support the possibility of rearrangements and/or recombinations. These probably occurred during bacterial propagation, as the PCR bands prior to cloning could be amplified by both primers.

The relevant sections of the alignments for Table 3.9 are given below (Figure 3.15), indicating the positions where the primers annealed to the respective genes. Mispriming in this instance and elsewhere, seemed to result most often with an A or T, on either the gene or the primer.

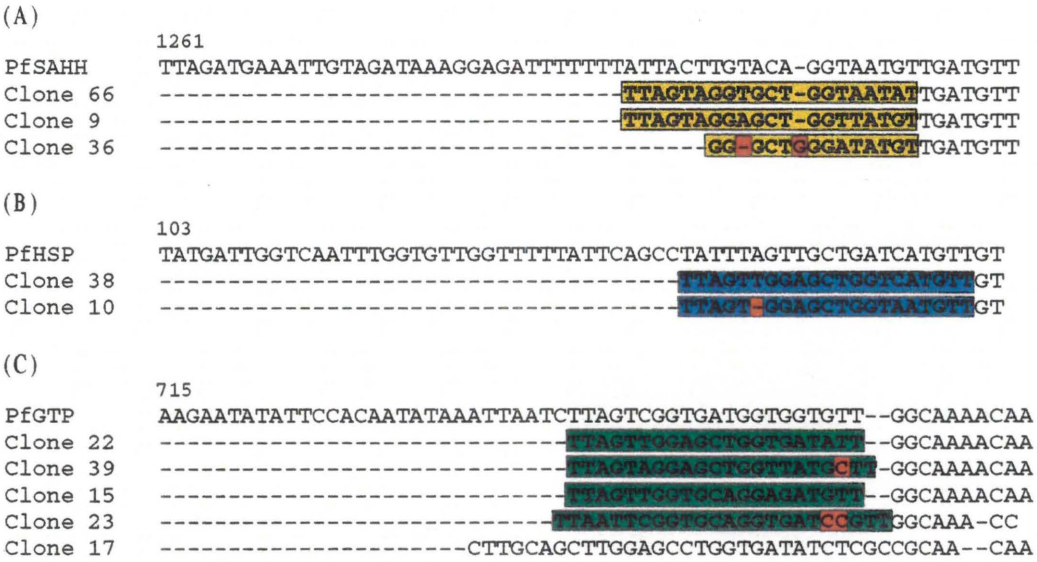


Figure 3.15 Sections of the alignments showing the positions of primer PSP2 annealing to the respective genes. **(A)** S-adenosylhomocysteine hydrolase; **(B)** Heat shock protein and **(C)** GTP binding protein. The sections that are coloured are the bases that correspond to primer PSP2. Section (C) clone 17 has no apparent primer sequence and was ignored in the homology matching. The nucleotides blocked in orange represent either insertions or deletions in the primer or may also be reading errors.

A more in depth analysis of primer annealing for SAHH is shown in Figure 3.16.

Figure 3.16 A detailed analysis of typical primer annealing using the SAHH gene. The coloured nucleotides represent bases which did not anneal correctly. The colours represent the mismatches of the primer with the template as follows: Red: G ↔ G; Blue: T ↔ T; Purple: A ↔ A; Yellow: A ↔ C and Orange: G ↔ T. The 'G' in green is either an insertion or a reading error.

However, at this early stage, we were still unaware of likely reasons for these occurrences (see discussion). Only part of the primer sequence was found in Clone 36, suggesting a truncation. Deletions of single nucleotides are also obvious.

Towards the end of the study, when the automated sequencer became available, the questionable clones were resequenced (from both sides). Since a greater number of bases could be read, identification was facilitated. The results obtained are shown in Table 3.10.

Table 3.10 Homology matches for the partial sequences obtained with automated sequencing, as determined by BLAST. The sequences were performed from both the T3 and the T7 promoter sides of the multiple cloning site.

Clone	Gene
4 T3	*P. falciparum* ribosomal phosphoprotein
4 T7	*P. falciparum* ribosomal phosphoprotein
45 T3	*P. falciparum* S-adenosylhomocysteine hydrolase
45 T7	*E. coli* gene rrnB (codes for 23S ribosomal RNA)
49 T3	*E. coli* mhp cluster for 3-hydroxy-phenylpropionic acid degradation
49 T7	*E. coli* mhp cluster for 3-hydroxy-phenylpropionic acid degradation
53 T3	?
53 T7	?

Further support for rearrangements is given by Clone 45, which has an insert consisting of a malaria gene on the T3 side and an *E. coli* gene on the T7 side of the multiple cloning site. The appearance of these apparent recombinations is surprising as SURE *E. coli* used as hosts are recombination deficient.

Clone 49 contains an inserted *E. coli* gene fragment as was also seen for Clones 18 and 27 in Table 3.10. Clone 53 showed no significant homology (on either nucleotide level, or the six reading frames of the deduced amino acid level) with any of the genes present in the NCBI databases searched by BLAST algorithm package. The insert has a AT content of ~47% which suggests that it is probably not a malaria gene.

The next round of PCRs were conducted with the second permease specific primer PSP, as well as a more specific primer, namely the tyrosine/tryptophan specific primer TYT. Both these primers were less degenerate than PSP2 (Chapter 2, Table 2.3).

Enzymes alternative to Expand™ HiFi DNA polymerase were considered at this time since it was giving smears and inconsistent results. Takara *ExTaq* DNA polymerase was used for subsequent investigations.

3.3.1.3 Optimisation of PCR with primer PSP

Figure 3.17 shows the results obtained for primer PSP using a Taguchi strategy in order to determine the optimal cDNA, $MgCl_2$ and primer concentrations. An annealing temperature of 48.2°C and 40 cycles were used.

The Taguchi method is based on the following type of matrix:

Table 3.11 Orthogonal array for two (1 and 2) variables, each at three levels (A, B, C and A', B', C').

	(1)	(2)
1	A	A'
2	A	B'
3	B	C'
4	B	A'
5	C	B'
6	C	C'

Table 3.12 The layout of the two Taguchi experiments used to optimise **(A)** the template and $MgCl_2$ concentrations (samples 1-6) and **(B)** the primer concentrations (samples 7-12).

Sample	**Gel A**		**Gel B**	
	mRNA equivalents	**[$MgCl_2$]**	**PSP**	**Anchor**
1	93.3ng	1.5mM	60pmol	10pmol
2	93.3ng	1.75mM	↓	↓
3	209.9ng	2mM		
4	209.9ng	1.5mM		
5	326.55ng	1.75mM		
6	326.55ng	2mM		
7	163.3ng	2mM	30pmol	10pmol
8	↓	↓	30pmol	20pmol
9			60pmol	30pmol
10			60pmol	10pmol
11			120pmol	20pmol
12			120pmol	30pmol

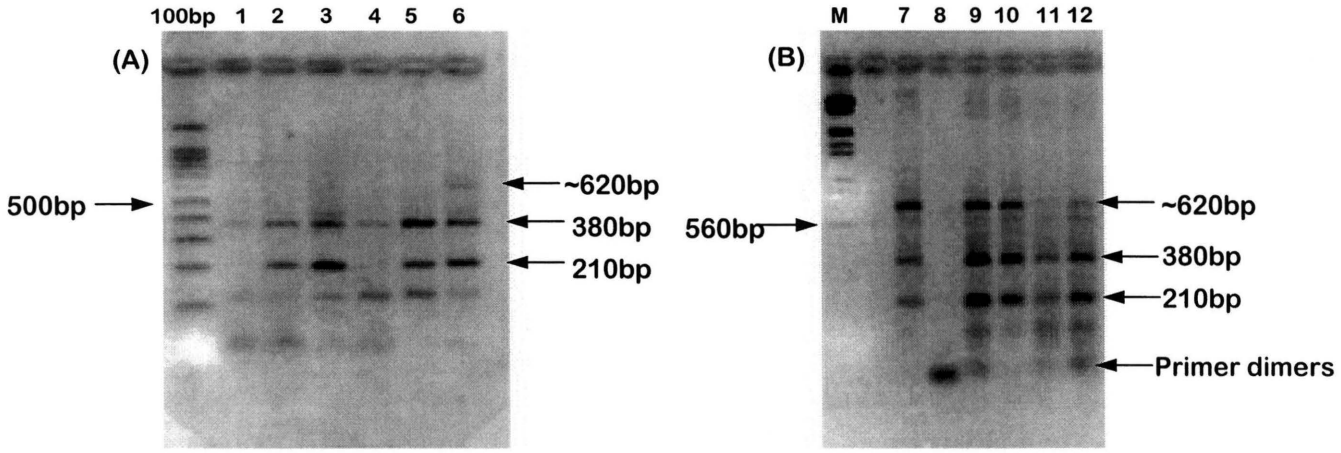


Figure 3.17 Taguchi strategy PCR, for primer PSP, to determine the optimal template, $MgCl_2$ and primer concentrations. The arrays contain **(A)** $MgCl_2$ and template concentrations as variables and **(B)** with primer concentrations as variables. The products were generated at a Ta of 48.2°C, 40 cycles and 1.5U of Takara *ExTaq* and were electrophoresed on 2% NuSieve:Promega mix agarose gel.

Sample 8 appears to be unamplified and could not be compared with the other samples. The lack of amplification may have been due to a missing reaction component, which is possible as the primers and enzyme were added individually and not as part of the mastermix.

If lanes 3, 6 and 10 are compared (they all have the same primer and $MgCl_2$ concentrations), the latter two yield similar bands, with lane 10 offering the best amplification, at a template concentration of 163ng mRNA equivalents.

Lane 2 showed improved amplification with 1.75mM $MgCl_2$, compared with the 1.5mM in lane 1 (both lanes contain the same amount of cDNA i.e. 93.3ng of mRNA equivalents). Amplification was further improved at 2mM $MgCl_2$ (if lanes 3 and 4 are compared at 209.9ng mRNA equivalents). This is confirmed at a template concentration of 326.6ng mRNA equivalents (compare lanes 5 and 6).

Optimal anchor primer concentration can be determined by comparing lane 9 with lane 10 and lane 11 with lane 12. Of lanes 9 (30pmol anchor) and 10 (10pmol anchor), lane 10 appears better as there are less primer dimers, but the 620bp band intensity is similar to that of lane 9. Lanes 11 and 12 contain 20 and 30pmol of anchor primer respectively, but the 620bp bands are less defined and there are more primer dimers

than lanes 9 and 10, which indicates that the primer concentrations (both PSP, at 120pmol and anchor) are present in excess. This would promote the amplification of secondary targets as indicated by the increased smearing between the bands. Comparing lanes 9 and 12 indicates that 60pmol rather than 120pmol of PSP appears optimal since the 620bp band is more clearly defined in lane 9. The best PCR conditions are thus 2mM $MgCl_2$ and 163.3ng of mRNA equivalents with primer concentrations of 60pmol of PSP and 10pmol of anchor.

From Figure 3.17, lanes 6 and 10, it appeared that increased cDNA concentrations yielded more bands and of higher molecular masses, in the vicinity of the predicted band size (~1200bp). The results of an experiment with 370.5ng of mRNA equivalents are shown in Figure 3.18.

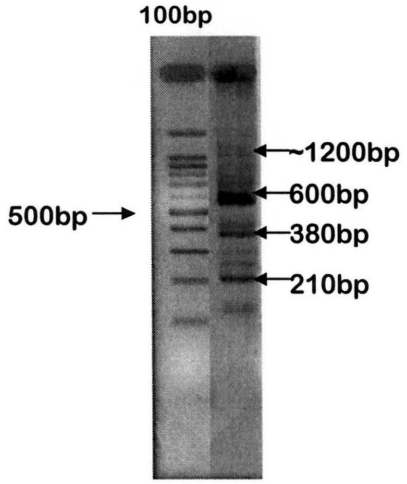


Figure 3.18 The effect of increased template concentration on PCR with primer PSP. PCR products obtained with 370.5ng of mRNA equivalents, Ta of 48.2°C, 40 cycles and 1.5U of Takara *ExTaq* electrophoresed using 2% Promega agarose gel.

Since the predicted band size of ~1200bp, for the amino acid permeases, was only emerging at vastly increased cDNA concentrations (370ng mRNA equivalents compared to the 1ng recommended by Frohman (Frohman, 1993)), it was speculated that some gene copies may be lost during mRNA isolation. This possibility was investigated using cDNA preparations synthesised from total RNA.

3.3.2 mRNA versus total RNA as a source for cDNA synthesis

The cDNA derived from mRNA and that from DNase-treated total RNA, were evaluated using primers for the known, highly abundant 18S rRNA gene (Figure 3.19). Bio1 and Dig2 (gene specific primers) evaluate quantity, while Bio1 and anchor were used to investigate the quality of the cDNA (i.e. the number of truncations present). A product of ~1814 bp should be amplified between Bio1 and the anchor primer pair. The ratio of poly-T primer to total treated RNA used for reverse transcription was also investigated. The initial PCRs were performed with cDNA synthesised using 1.5pmol of poly-T primer to 1µg of isolated mRNA. Here we show the effect of a 5:1 and a 0.5:1 ratio of poly-T primer to total RNA (this is equivalent to 100pmol to 1µg mRNA (theoretical) and 10pmol poly-T to 1µg of mRNA (theoretical), respectively). The theoretical mRNA concentration is calculated as 5% of the total RNA.

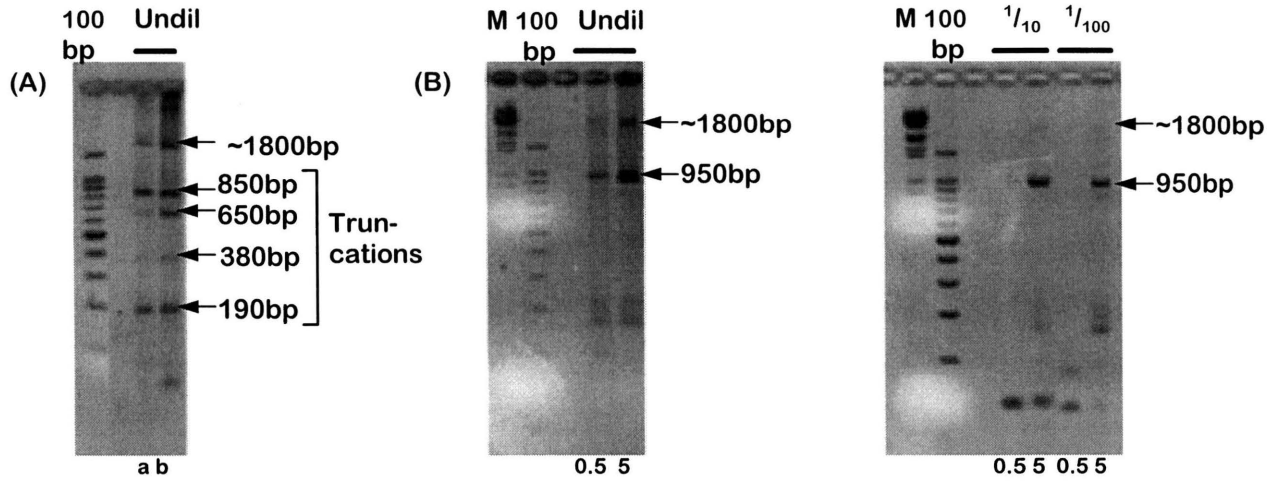


Figure 3.19 Comparison of the quality of cDNA derived from mRNA and cDNA derived from DNase-treated total RNA. PCR products (Ta of 50°C, 40 cycles and 2.25U of Takara *ExTaq*) electrophoresed on 2% NuSieve:Promega mix agarose gels. **(A)** 150ng cDNA derived from mRNA, 40pmol Bio1 and (a) 20pmol anchor or (b) 40pmol anchor and **(B)** cDNA from total RNA using 10pmol of anchor. Undiluted: 250ng; diluted $^{1}/_{10}$: 25ng and diluted $^{1}/_{100}$: 2.5ng, respectively of total RNA equivalents.

These experiments suggested that the cDNA of the 18S rRNA gene is more truncated when mRNA was used (Figure 3.19A) than when total RNA was used (Figure 3.19B). Dilution of the template by 1 in 10, to 25ng total RNA equivalents (theoretically 1.3ng mRNA) and 1 in 100, to 2.5ng total RNA equivalents (theoretically 0.13ng mRNA), resulted in a decrease in the intensity of the ~1800bp band, although the truncated band

at ~950bp did not decrease as much. It is also apparent that the 5:1 ratio of poly-T to template, is better than the 0.5:1 ratio for reverse transcription. The ~1800bp band generated with 250ng cDNA derived from total RNA synthesised with the 5:1 ratio (theoretically 12.5ng mRNA) has a similar intensity to the band generated with cDNA from 150ng of mRNA. This indicates at least a 10-fold increase in cDNA yield obtained from total RNA. The effective cDNA concentration at the $^{1}/_{10}$ dilution (theoretically ~1.3ng mRNA,) is also approaching the 1ng recommended by Frohman (Frohman, 1993).

Upon inspection of the sequence of the 18S rRNA gene, possible places where the poly-T primer could have annealed to the gene were identified (Figure 3.20).

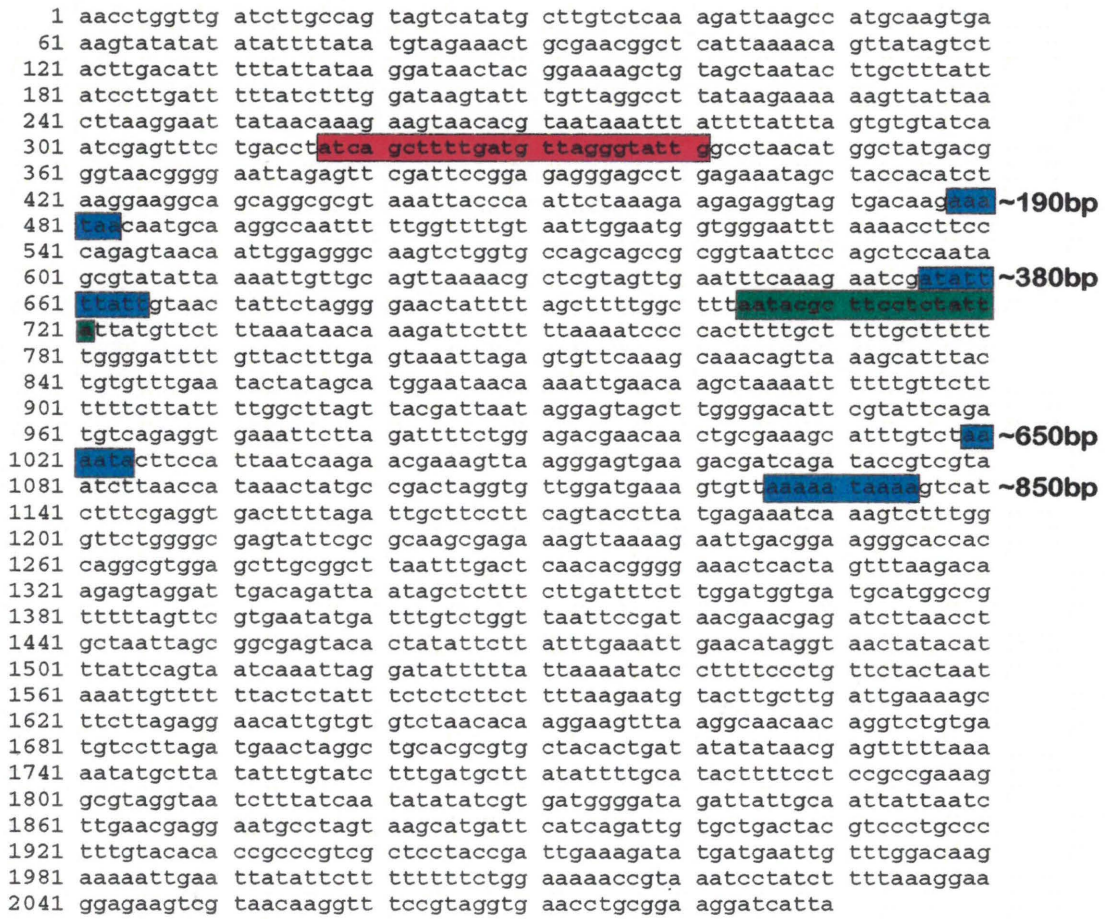


Figure 3.20 Evaluation of the possible positions of poly-T primer annealing internally in the 18S rRNA gene (four areas blocked in blue). The red block is the Bio1 primer, green is the Dig2 primer.

It is important to note that the poly-T primer appears to have annealed to A-rich areas (less than 18bp) containing some T's, with the exception of the putative position for binding to generate the 380bp band, which is T-rich. This supports the conclusion reached earlier, namely that mispriming seems to occur at the position of an A or T on either the gene or primer.

The apparent potential of the poly-T primer to drive cDNA synthesis from less than 18 internal A's in the gene led to the introduction of a 'hot start' procedure for reverse transcription in subsequent studies. This involved the equilibration of all synthesis components, excluding enzyme, at the synthesis temperature (45°C) for 2 minutes before the addition of the reverse transcriptase.

Further confirmation of the higher efficiency of a 5:1 rather than a 0.5:1 ratio of poly-T primer to total RNA, for the synthesis of cDNA, was obtained when PCR was performed using the specific forward and reverse primers for the 18S rRNA gene, namely Bio1 and Dig2 (Figure 3.21).

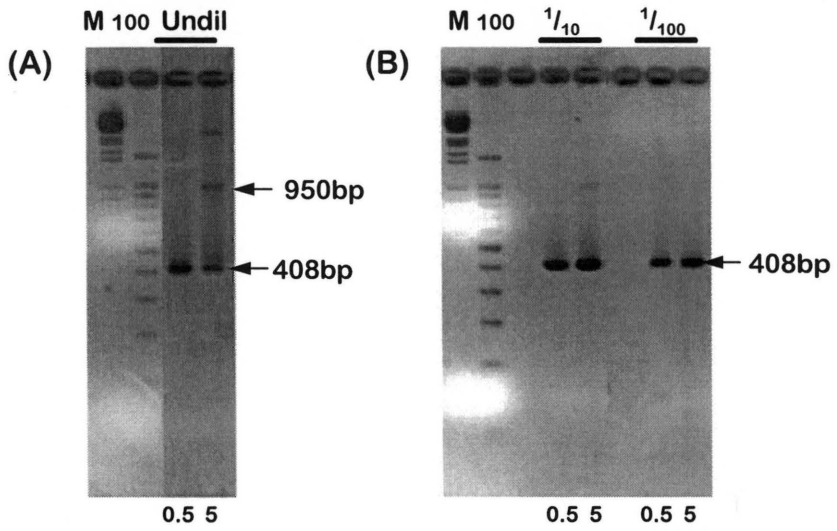


Figure 3.21 Comparison of the quantity of cDNA derived from DNase-treated total RNA. PCR products obtained using a Ta of 50°C, 30 cycles, 2.25U of enzyme and Bio1 (15pmol) with Dig2 (10pmol). **(A)** Undiluted: 250ng; **(B)** cDNA diluted $^{1}/_{10}$: 25ng and cDNA diluted $^{1}/_{100}$: 2.5ng of total RNA equivalents.

The products generated with undiluted cDNA (theoretically 12.5ng mRNA) synthesised with the 5:1 ratio of poly-T to template are overamplified. This is still the case in the $^{1}/_{10}$ dilution (theoretically 1.25ng of mRNA), while a $^{1}/_{100}$ dilution

(theoretically 0.125ng of mRNA) offers the best amplification. With the 0.5:1 ratio, only the products from undiluted cDNA are overamplified. Thus, it appears that the 5:1 ratio of poly-T: total RNA yields a higher cDNA concentration of the 18S rRNA gene, than the 0.5:1 ratio. These results are in agreement with those of Figure 3.19.

3.3.3 RNase-free DNase treatment versus non-treatment of total RNA

Initial preparations of cDNA were derived from total RNA treated with RNase-free DNase to remove traces of contaminating genomic DNA. However, loss of copies during treatment was still a concern and thus experiments were performed to determine whether DNase treatment had any effect on the results. Theoretically, since RACE protocols and thus anchor primers were to be used, minimal genomic DNA contamination of cDNA, should not interfere with amplification results. A touchdown PCR was done for PSP (65°C to 50°C over 30 cycles, with a further 10 cycles at 55°C and 1.5U of Takara *ExTaq*) using DNase-treated and untreated RNA as template for cDNA synthesis (Figure 3.22).

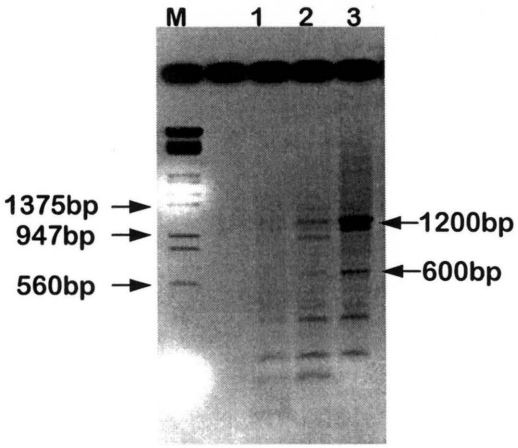


Figure 3.22 Comparison of PCR products obtained from RNase-free DNase treatment versus non-treatment of total RNA. PCR products for primer PSP (touchdown from 65°C to 50°C over 30 cycles and a further 10 cycles at 55°C). Lane (1) contains 125ng total RNA equivalents (from DNase-treated RNA), lane (2) 250ng total RNA equivalents (from DNase-treated RNA) and lane (3) 125ng total RNA equivalents (from untreated RNA) electrophoresed on 1.5% Promega agarose gel. The 'hot start' method of cDNA synthesis was used.

cDNA derived from 125ng of DNase-treated RNA (the concentration of the DNase-treated RNA was determined spectrophotometrically) resulted in almost no products whereas that derived from 125ng of untreated RNA yielded clear defined bands of

similar sizes and intensities as those produced using cDNA derived from 250ng of DNase-treated RNA. Since untreated RNA seemed to yield improved products at a lower cDNA concentration than DNase-treated RNA, it was decided to discontinue with DNase treatment of the total RNA and to minimise the potential loss of rare copies. The results also show that bands obtained are of a similar size to those obtained in Figure 3.18 (where cDNA was prepared from mRNA), except that less cDNA was used. The expected band at ~1200bp is also much more clearly defined in Figure 3.22.

Similarly, the FORMAzol in which the isolated RNA was stored was initially removed via precipitation, as it was possible that the FORMAzol could negatively influence the PCR. This was shown to be unnecessary (results not shown) and was discontinued as large quantities of RNA were lost in the precipitation steps.

3.3.4 PCR with cDNA derived from untreated total RNA

3.3.4.1 PCR, cloning and sequences of bands generated with primer PSP2

At this point we returned to primer PSP2 to investigate the products formed with the cDNA from total RNA. Figure 3.23 shows the PCR products obtained under the conditions optimised previously (see section 3.3.1.1).

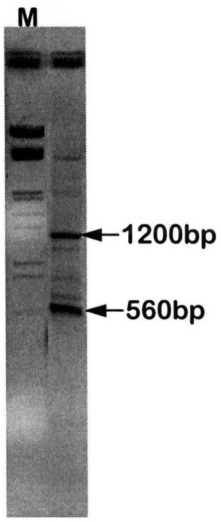


Figure 3.23 PCR products obtained with primer PSP2 using untreated RNA. A Ta of 50°C, 35 cycles, 87.5ng total RNA equivalents and 1.5U of Takara *ExTaq* were used and the products electrophoresed on 1.5% NuSieve:Promega mix agarose gel.

It is noticeable that the ~1200bp band obtained in Figure 3.22 with primer PSP, was also obtained in Figure 3.23 with primer PSP2, although the sizes of the rest of the bands appear to differ. If the results obtained in Figure 3.23 are compared with those in Figure 3.9B (both used primer PSP2, but cDNA synthesised from total RNA and mRNA, respectively), the 560bp band is present in both. However, while it is the dominant band in Figure 3.9B, the 1200bp band is of equal intensity in Figure 3.23. These two figures represent the progress made during the study as Figure 3.9B contained cDNA prepared from 100ng of mRNA, while Figure 3.23 contained cDNA derived from 87.5ng untreated total RNA (theoretically ~4.4ng mRNA). Both experiments used the PSP2 and anchor primers at the same concentrations, although the results in Figure 3.9B and Figure 3.23 were obtained with Expand™ HiFi and Takara *ExTaq* DNA polymerases, respectively.

The traditional blunt-end cloning procedure used up to this stage, required large quantities of insert, but yielded relatively few clones. It was therefore decided to use kits for subsequent cloning procedures. The 1200bp band in Figure 3.23 was cloned using the Stratagene pCR-Script blunt-end cloning kit (see Appendix C, Figure 2 for cloning vector), transformed into competent DH5α *E. coli* cells and sequenced automatically using standard dye terminators. The other bands were not cloned due to time constraints and problems encountered with cloning. The sequences were all of the 5.8S rRNA gene of *P. falciparum* (See Figure 3.26 for the partial alignment indicating the position of primer annealing and Appendix B for the complete sequence). This once again highlighted the preponderance of misprimed products. The results thus also confirmed that primer PSP2 was not an ideal primer and we reverted to the more specific primer PSP (section 2.3).

3.3.4.2 PCR, cloning and sequences of bands generated with primer PSP

Based on the mispriming that was experienced with primer PSP2, the annealing temperature for primer PSP was investigated to improve the stringency of the reaction. Figure 3.24 shows the results of a touchdown PCR with primer PSP from 65°C to 50°C and PCRs with constant annealing temperatures of 60°C, 58°C, 55°C and 50°C.

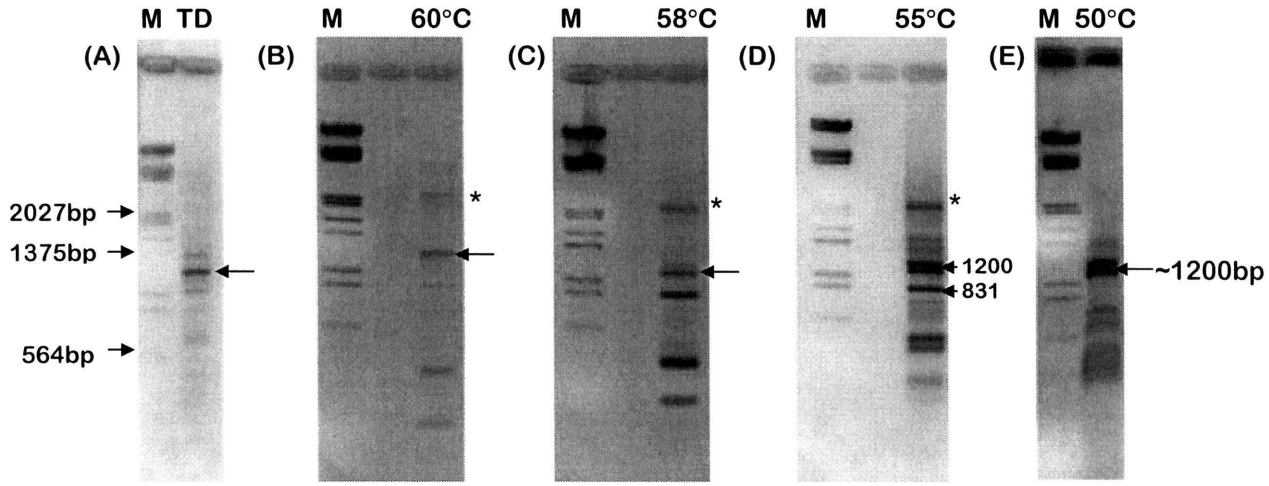


Figure 3.24 PCR products obtained with primer PSP under conditions of higher stringency. The products were electrophoresed using 1.5% NuSieve:Promega mix agarose gels (A and E) and 1.5% Promega agarose gels (B,C and D). **(A)** 87.5ng total RNA equivalents, a touchdown procedure from 65°C to 50°C over 30 cycles with a further 10 cycles at 55°C. 100ng total RNA equivalents was used for B, C and D with **(B)** Ta of 60°C; **(C)** Ta of 58°C and **(D)** Ta of 55°C. **(E)** 87.5ng of total RNA equivalents and Ta of 50°C. All experiments used 1.5U of Takara *ExTaq* and 35 cycles for B-E. The asterix (*) indicates bands that were anchor primer driven.

The intensity of the ~1200bp band (indicated by the arrow), differed at the various temperatures and became more intense with a decrease in annealing temperature. More bands also appeared with every reduction in annealing temperature. It was concluded that 55°C would be specific enough as it was 8.7°C above the temperature recommended by OLIGO (47.7°C). We also found that if the annealing temperature was too high, anchor primer driven products of high molecular weight (greater than 2000bp) resulted (results not shown).

The two bands in the vicinity of 1200bp (Figure 3.24D), were cloned together (Stratagene pCR-Script Cloning Kit or the Promega pGEM T-Easy Vector System - see Appendix C, Figure 3 for cloning vector), transformed into competent DH5α *E. coli* cells and sequenced automatically using Big-Dyes and the results are shown in Table 3.13. The band present at 831bp was also cloned, though no colonies were obtained.

The 60S ribosomal protein belongs to the family L1E and is not present in the databases for malaria. The identity was determined with homology matching using BLAST.

Table 3.13 Clones obtained using primer PSP. Clones labelled 'p' are those cloned using the Promega system. '?' indicates no homology matches found by BLAST.

Clone	Gene
2; 7a; 18a; p6; p7	60S ribosomal protein
3; 4a; 14a	*P. falciparum* 5.8S rRNA gene
7; p5	?

Clone 7 sequenced from the T3 promoter, showed no significant homology (on either nucleotide, or the six reading frames for the deduced amino acid, levels) with any sequences in the databases searched by BLAST, although the poly-T primer was identified. The AT-content was ~64%. The sequence from the T7 promoter showed extensive homology to the cloning vector used and no primer was identified on this side. This may indicate that rearrangement of the insert has occurred once again. This clone is thus not likely to be of any significant value.

Clone p5 was sequenced using the T7 primer only. The sequence obtained has an AT-content of ~70%. BLAST shows slight homology with a *Plasmodium* contig sequence (sequencing is still in progress) at nucleotide level. At amino acid level, BLAST shows fairly good homology with a kinesin-like protein (from various organisms). The best homology match found by BLAST is given below:

```
splP53086|YGW6_YEAST PUTATIVE KINESIN-LIKE PROTEIN YGL216W
Length = 805
Score = 86.2 bits (210), Expect = 4e-17
Identities = 50/140 (35%), Positives = 83/140 (58%), Gaps = 2/140 (1%)

Query: 5   LKDLFERIKMLQIMNEYKVKCSFIEIYNENICDLLNPS--NEYLDVREDPIKGVTVSNIF 178
           +++LF +I   L+    ++++   S++EIYNE I DLL P    ++ L +RED     + V+N+
Sbjct: 216 MEELFNKITDLKDEKDFEISLSYLEIYNERIRDLLKPETPSKRLVIREDTQNHIKVANLS 275

Query: 179 EVCTTSVEEIMELIHTGNRNRTQEPTDANKTSSRSHGVFK*LWKKLKRVKGYISKQKEEN 358
                +VE++M+L+  GN NRT  PT+AN+ SSRSH V +    +  ++    S+
Sbjct: 276 YHHPNTVEDVMDLVVQGNINRTTSPTEANEVSSRSHAVLQIHIMQTNKLVDLTSQHTFAT 335

Query: 359 YV*SI*LGSERAS*TNNKGMRM 424
                GSERA+ T N+G+R+
Sbjct: 336 LSIIDLAGSERAAATRNRGIRL 357
```

Figure 3.25 BLAST X (amino acid level) homology matching for Clone p5 (sequenced from the T7 promoter).

The ubiquitous 5.8S rRNA gene was amplified again (section 3.3.4.2). Figure 3.26 represents a partial alignment and indicates the region where primer PSP (from Figure 3.24D) and PSP2 (from Figure 3.23) annealed to the gene. The complete alignment is given in Appendix B, Figure 5 .

```
G924644          AGCCTATTAGCTGGTTATTTTCGAAAGA-TCTCTCA-GGATCGCTGGAGTTG
Clone 3          TTTNTTGGTGCTGGTGAAGTTCGAAAGA-TCTCTCA-GGATCGCTGGAGTTG
Clone 4a         TTAATTGGTGCTGGTGAAGTTCGAAAGA-TCTCTCA-GGNNCGCTGGAGTTG
Clone 14a        TTAATTGGTGCTGGTGAAGTTCGAAAGA-TCTCTCA-GGATCGCTGGAGTTG
Clone 1          TTAGTTGGAGCTGGTAATATTCNAAAGAATCTCTCACGGATCGCTGGANTTG
Clone 10         TTAATTGGTGCTGGTGAAGTTCGAAAGA-TCTCTCA-GGATCGCTGGAGTTG
```

Figure 3.26 The section of the alignment showing the position of primers PSP (blue) and PSP2 (red) annealing to the 5.8S rRNA gene (g924644). The nucleotide blocked in yellow is one which does not occur in the primer.

The clones arising from PCR products obtained with PSP did not contain the full primer sequence. The 5′-end of the primer may have been lost while cloning due to rearrangements occurring during propagation of the bacterial cultures (see sections 3.3.1.2, 3.3.4.4 and discussion).

3.3.4.3 Restriction digestion method for screening similar recombinant clones

Since greater numbers of clones were expected from the cloning kits, it was necessary to find a way of screening recombinant clones for differences, without sequencing all the clones. Two or three clones were sequenced and a restriction enzyme which cuts in the insert identified. This enzyme was used to digest all clones and the restriction

patterns obtained were compared. Figure 3.27 is a typical result obtained using the clones obtained after PCR with primer PSP. All the clones were subsequently sequenced to confirm the allocations conferred.

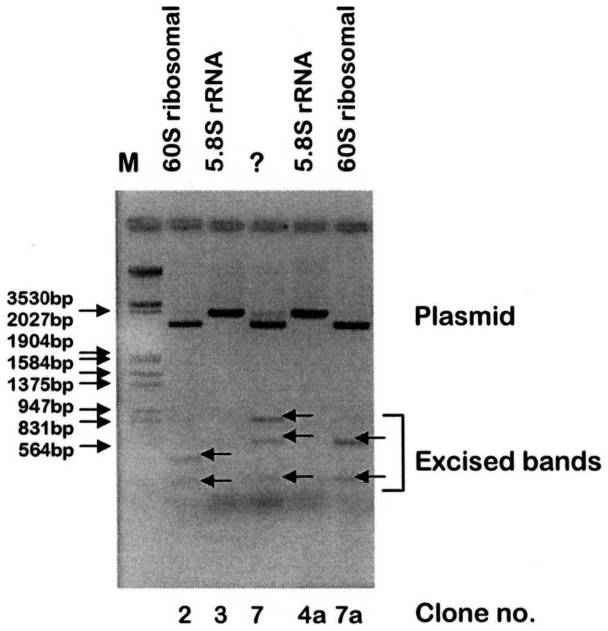


Figure 3.27 Restriction patterns obtained when the clones from primer PSP were digested with HinDIII. The '?' (clone7) represents an insert for which no homology with sequences in the databases was found.

The choice of restriction enzyme was based on the sequence of the 60S ribosomal protein. It cleaves the insert in two places and the multiple cloning site in one, which results in two bands being excised from the plasmid (clones 2 and 7a). HinDIII obviously has no site in the 5.8S rRNA gene (clones 3 and 4a) as the plasmid appears to only be linearised (the enzyme cut once in the multiple cloning site). Clone 7 appears to have three sites, other than the one in the multiple cloning site, as three bands are excised. The faint band visible above the plasmid band in Clone 7, may be partially digested plasmid. The slight shift in the excised bands of Clone 2 compared with Clone 7a may be an artefact of electrophoresis. This method appears to be a viable way of screening recombinant clones to confer a preliminary identification without sequencing.

3.3.4.4 PCR, cloning and sequences of bands generated with primer TYT

Experiments with primer TYT had been running concurrently with those of PSP and PSP2. The PCR was optimised (results not shown) and two bands obtained at a Ta of 48.2°C.

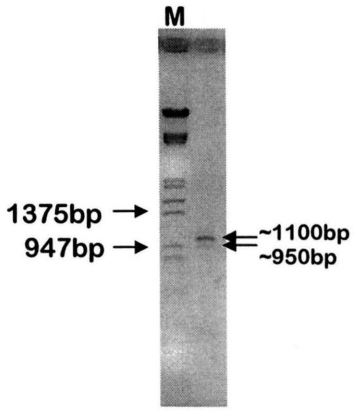


Figure 3.28 PCR products obtained with primer TYT. A Ta of 48.2°C, 30 cycles and 87.5ng of total RNA equivalents with 1.5U of Takara *ExTaq* DNA polymerase were used and the products electrophoresed on a 1.5% NuSieve:Promega mix agarose gel.

The ~1100bp band and the slightly smaller ~950bp bands were cloned together (Stratagene pCR-Script Cloning Kit) and transformed into competent DH5α *E. coli* cells. The clones were sequenced (automatically using Big-Dyes) and the results are shown in Table 3.14.

Table 3.14 Clones obtained with primer TYT.

Clone	Gene
1; 2; 3; 8; 13; 14	60S ribosomal protein
15 (5′-end)	60S ribosomal protein
15 (3′-end)	Transposon Tn10
7	*P. falciparum* ubiquitin

Clone 15 has undergone rearrangement with the bacterial genome. This clone shows that it is not a particular type of gene which is always rearranged, but occurs at random, since it was the only one of the six ribosomal protein genes amplified by primer TYT, to undergo rearrangement. It was also at this stage that the results given in Table 3.10

(section 3.3.1.2) were obtained. It was only now that it became clear that rearrangements and recombination events were occurring with the *E. coli* host genome.

Figure 3.29 shows that upon restriction digestion, the size of the excised fragment for clone 15 is the same as that for clone 14, although an additional band of larger size is also present. This band may be the result partially digested plasmid.

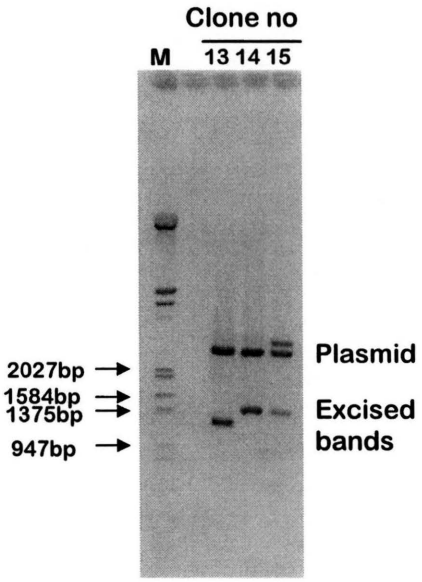


Figure 3.29 Restriction digestion of plasmids from TYT clones with PvuII. The digested plasmids were electrophoresed using 1% Promega agarose.

These results thus confirm that a band of a specific size is not necessarily all the same gene.

Figure 3.30 shows the 5′-ends of the clones from primers PSP and TYT aligned with each other, thus indicating the positions where the primers annealed.

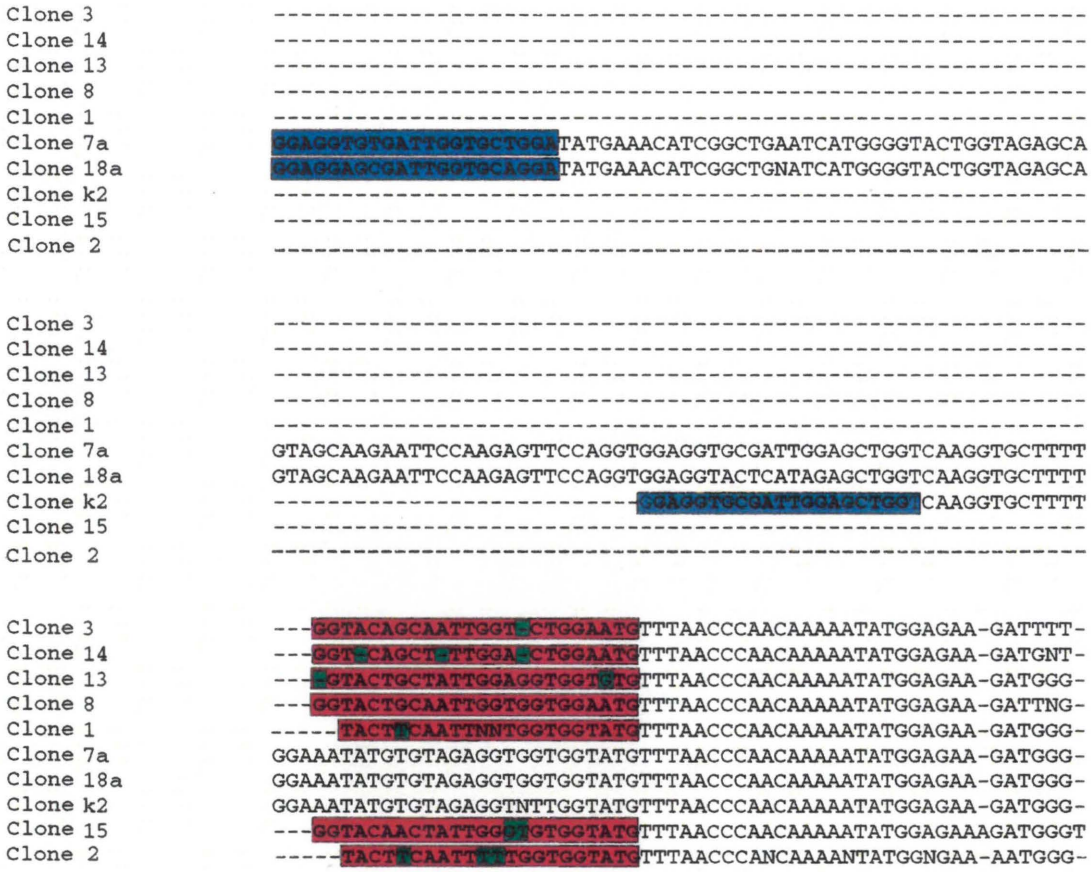


Figure 3.30 Positions of primer (PSP and TYT) annealing to the 60S ribosomal subunit protein gene. Primer TYT is outlined in blue and TYT in red, while the green blocks represent rearrangements, where the nucleotide in the sequence is not in the primer.

Primer PSP annealed at the same position in two of the clones and primed at a separate position for the third, while primer TYT appears to have primed all the clones from the same position. Primer PSP also appears to have no induced nucleotide changes, while TYT has several.

3.3.5 *P. falciparum* gene sequence for ubiquitin and the partial sequence of the large ribosomal subunit protein

Of all the clones obtained, only those for the large ribosomal subunit protein were novel. Since the 3′-UTR of ubiquitin was obtained, the sequence obtained will also be elaborated on below.

3.3.5.1 Ubiquitin

The *P. falciparum* ubiquitin protein sequence obtained consists of the same sequence repeated twice, although it is likely that there are repeats prior to primer annealing.

```
1          11         21         31         41         51         61       70
|          |          |          |          |          |          |        |
GGAACTACAN TTGGTGCTGG TATGCAAATT TTTGTCAAAA CATTAACAGG AAAAACAATA ACTCTTGACG
---------- ---------- -M--Q--I-- F--V--K--T --L--T--G- -K--T--I-- T--L--D--V

71         81         91         101        111        121        131      140
|          |          |          |          |          |          |        |
TCGAGTCATC TGATACCATT GAAAATGTTA AGGCAAAAAT TCAAGATAAA GAAGGAATTC CACCTGACCA
--E--S--S- -D--T--I-- E--N--V--K --A--K--I- -Q--D--K-- E--G--I--P --P--D--Q-

141        151        161        171        181        191        201      210
|          |          |          |          |          |          |        |
ACAAAGATTA ATATTTGCTG GAAAACAATT AGAAGATGGA AGAACTTTAT CAGATTACAA TATTCAAAAG
-Q--R--L-- I--F--A--G --K--Q--L- -E--D--G-- R--T--L--S --D--Y--N- -I--Q--K--

211        221        231        241        251        261        271      280
|          |          |          |          |          |          |        |
GAATCCACTC TACACTTAGT ATTAAGATTA AGAGGTGGTA TGCAAATTTT TGTCAAAACA TTAACAGGAA
E--S--T--L --H--L--V- -L--R--L-- R--G--G--M --Q--I--F- -V--K--T-- L--T--G--K

281        291        301        311        321        331        341      350
|          |          |          |          |          |          |        |
AAACAATAAC TCTTGACGTC GAGCCATCTG ACACCATTGA AAATGTTAAG GCAAAGATTC AAGATAAAGA
--T--I--T- -L--D--V-- E--P--S--D --T--I--E- -N--V--K-- A--K--I--Q --D--K--E-

351        361        371        381        391        401        411      420
|          |          |          |          |          |          |        |
AGGAATTCCA CCTGACCAAC AAAGATTAAT ATTTGCAGGA AAACAATTAG AAGATGGTAG AACCCTATCA
-G--I--P-- P--D--Q--Q --R--L--I- -F--A--G-- K--Q--L--E --D--G--R- -T--L--S--

421        431        441        451        461        471        481      490
|          |          |          |          |          |          |        |
GATTACAATA TTCAAAAGGA ATCCACTCTA CACTTAGGAT TAAGATTAAG AGGTGGGTTT TAATTATTTA
D--Y--N--I --Q--K--E- -S--T--L-- H--L--G--L --R--L--R- -G--G--F-- *

491        501        511        521        531        541        551      560
|          |          |          |          |          |          |        |
TTAATTTTTT TTTTTTTTCT CTTTTCCTTT TTTTCTTTAT CTTGTTTTTT TTTTTTTTTT TAAGTCTATC
---------- ---------- ---------- ---------- ---------- ---------- ----------

561        571        581        591        601        611        621      630
|          |          |          |          |          |          |        |
TTAATATAGT ATGTATATAC TCTATTATGT AATTTAATTA AGTTTTTTTA TTAAAAAATT TATTAGTTTG
---------- ---------- ---------- ---------- ---------- ---------- ----------

631        641        651        661        671        681        691      700
|          |          |          |          |          |          |        |
TGTTGTGCGC GCACATACTG TTATGAATTA AAATTTTTTT TTTTTTTTTT TTTTTTTAAA ACTTGGCCCC
---------- ---------- ---------- ---------- ---------- ---------- ----------

701        711        721        731        741        751        761      770
|          |          |          |          |          |          |        |
ATCATTTGGC ATGTGCTATA ATAATAATAT ATAGTTTTTC CTATTAAAAA ATTTACACGT ATAAAAAAAA
---------- ---------- ---------- ---------- ---------- ---------- ----------

771
|
AAAAAAAA
--------
```

Figure 3.31 The nucleotide sequence and translation of the *P. falciparum* ubiquitin protein. The asterix (*) indicates the termination codon. The sequence was reverse translated using the genetic data environment (Smith *et al.*, 1994).

One of these repeats is present in the databases (g160733). The last amino acid prior to the stop codon differs between organisms although the amino acid sequence of the protein is highly conserved. Part of the 3′ untranslated region (3′-UTR) of the gene was also obtained and is underlined. The coding area has an AT-content of 67%, while the 3′-UTR is 83% AT-rich. This conforms with the high AT-richness of areas flanking malaria genes, although up to 90% is common.

The amino acid sequence repeat of the *Plasmodium* protein aligned with that of humans is shown in Figure 3.32 illustrating the highly conserved nature of ubiquitin among unrelated organisms.

```
Human        MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAG
Plasmodium   MQIFVKTLTGKTITLDVESSDTIENVKAKIQDKEGIPPDQQRLIFAG

Human        KQLEDGRTLSDYNIQKESTLHLVLRLRGGV
Plasmodium   KQLEDGRTLSDYNIQKESTLHLVLRLRGGF
```

Figure 3.32 Amino acid sequence of one of the *P. falciparum* ubiquitin gene repeats aligned with the gene repeat from the human protein.

The underlined S in the *Plasmodium* sequence is due to the codon 'TCA' which should be 'CCA' as later in the second repeat. (Note: the sequence of the *Plasmodium* gene cannot be guaranteed as only one clone was obtained and the C to T mutation could be due to the PCR process.)

3.3.5.2 Large ribosomal subunit protein

The 60S ribosomal protein of the family L1E was obtained with both primers PSP and TYT! This family is not present in the databases for *P. falciparum*. The partial nucleotide sequence is shown in Figure 3.33. The approximate length for these genes appears to be ~1300bp which codes for a ~440 amino acid protein (see Figure 3.36).

```
1          11         21         31         41         51         61       70
 |         |          |          |          |          |          |        |
 GGAGGATTGA TTGGTGCAGG ATATGAAACA TCGGCTGAAT CATGGGGTAC TGGTAGAGCA GTAGCAAGAA
 G--G--L--I --G--A--G- -Y--E--T-- S--A--E--S --W--G--T- -G--R--A-- V--A--R--I

 71         81         91         101        111        121        131      140
 |          |          |          |          |          |          |        |
 TTCCAAGAGT TCCAGGTGGA GGTACTCATA GAGCTGGTCA AGGTGCTTTT GGAAATATGT GTAGAGGTGG
 --P--R--V- -P--G--G-- G--T--H--R --A--G--Q- -G--A--F-- G--N--M--C --R--G--G-

 141        151        161        171        181        191        201      210
 |          |          |          |          |          |          |        |
 TGGTATGTTT AACCCAACAA AAATATGGAG AAGATGGGGA AGGAAAGTTA ATTTGAAAGA AAAGAGATAT
 -G--M--F-- N--P--T--K --I--W--R- -R--W--G-- R--K--V--N --L--K--E- -K--R--Y--
```

```
211         221         231         241         251         261         271       280
|           |           |           |           |           |           |         |
GCCGTATGTT CATCTATTGC AGCTAGTGGT GTAACATCAT TAGTTTTGGC AAGAGGTCAT CGTATATCCC
A--V--C--S --S--I--A- -A--S--G-- V--T--S--L --V--L--A- -R--G--H-- R--I--S--H

281         291         301         311         321         331         341       350
|           |           |           |           |           |           |         |
ACCTTAAGGA AGTTCCATTA GTTGTTAGCA ATGATATTGA ATCTCTTTCT AAAACTAAAG AAGCTGTGAA
--L--K--E- -V--P--L-- V--V--S--N --D--I--E- -S--L--S-- K--T--K--E --A--V--N-

351         361         371         381         391         401         411       420
|           |           |           |           |           |           |         |
CTTTTTAGTT AGTCTTGGTT TAAAAGATGA AGTTAATAGA TTAGTAAAAT CGAAAAAAAT TAGAGCAGGA
-F--L--V-- S--L--G--L --K--D--E- -V--N--R-- L--V--K--S --K--K--I- -R--A--G--

421         431         441         451         461         471         481       490
|           |           |           |           |           |           |         |
AAAGGTAAAA TGAGAAACAG AAAATATAAA ATTAGAAATG GCCCTCTTAT TATTTATGAA AACGATATGG
K--G--K--M --R--N--R- -K--Y--K-- I--R--N--G --P--L--I- -I--Y--E-- N--D--M--G

491         501         511         521         531         541         551       560
|           |           |           |           |           |           |         |
GTGTAAAGAA AGCTTTTAGA AATATTCCTG GAGTTGATTT ATGTAAAGTT ACCAAATTAA ATTTATTAAA
--V--K--K- -A--F--R-- N--I--P--G --V--D--L- -C--K--V-- T--K--L--N --L--L--K-

561         571         581         591         601         611         621       630
|           |           |           |           |           |           |         |
ATTAGCTCCT GGTGGATCTA TAGGTAGATT GTGTATATGG AGTGAAAGTG CTTTTAAAAA ATTAGATGTT
-L--A--P-- G--G--S--I --G--R--L- -C--I--W-- S--E--S--A --F--K--K- -L--D--V--

631         641         651         661         671         681         691       700
|           |           |           |           |           |           |         |
ATATATGGAA AGATCCATGA GAAAAAAGTC ACTAAGAAAA ATTATATTTT ACCAAAATCT ATTGTGCACA
I--Y--G--K --I--H--E- -K--K--V-- T--K--K--N --Y--I--L- -P--K--S-- I--V--H--N

701         711         721         731         741         751         761       770
|           |           |           |           |           |           |         |
ATCCTGACAT CTATAGAATT ATACACAGCG ACAAAGTACA AGCAAGCTTA TTAGCCAAAA AGAAACCATG
--P--D--I- -Y--R--I-- I--H--S--D --K--V--Q- -A--S--L-- L--A--K--K --K--P--C-

771         781         791         801         811         821         831       840
|           |           |           |           |           |           |         |
CAAAAAGAGA TTACAAAACA AAAACTCATT AACCAACTTT GCTGTAAGAT GTAGACTTAA CCCAGCCTAC
-K--K--R-- L--Q--N--K --N--S--L- -T--N--F-- A--V--R--C --R--L--N- -P--A--Y--

841         851         861         871         881         891         901       910
|           |           |           |           |           |           |         |
AAATTATTAA GATCATTGGC TGTTCTTAGG ATGAGAAAGA GTATCCTTGA AAAATCAAAA AATAAAAAGG
K--L--L--R --S--L--A- -V--L--R-- M--R--K--S --I--L--E- -K--S--K-- N--K--K--E

911         921         931         941         951         961         971       980
|           |           |           |           |           |           |         |
AAAAGAGAGT TCAGAAACAA ATCCAAAAGA AAGAACTCCA AAAAATTAAC CATGACTACT ATAAGGGTGT
--K--R--V- -Q--K--Q-- I--Q--K--K --E--L--Q- -K--I--N-- H--D--Y--Y --K--G--V-

981         991         1001        1011        1021        1031        1041     1050
|           |           |           |           |           |           |        |
AGCTAAGGCA GTTAAGAAAA AGAAAAAGAG AGAAGAGAAA AAGGCTAAAT CTAAGAAAAC TGCTAACCAA
-A--K--A-- V--K--K--K --K--K--R- -E--E--K-- K--A--K--S --K--K--T- -A--N--Q--

1051        1061        1071        1081        1091        1101
|           |           |           |           |           |
GCTGTTATAA AAGTTGCAGC AGAAGAAAAA AAAAAAAAAG AGTGTTGTGG TAATGATAGC
A--V--I--K --V--A--A- -E--E--K-- K--K--K--E --C--C--G- -N--D--S--
```

Figure 3.33 The nucleotide sequence and translation of the 60S ribosomal subunit protein. The sequence was reverse translated using the genetic data environment (Smith *et al.*, 1994).

Since the sequence of the 60S ribosomal protein obtained with both the TYT and PSP primers was not described in the databases for malaria, it was decided to determine whether the clones were of malarial origin. This was done using a DIG-labelled single

stranded DNA probe and *P. falciparum* RNA as the target (DNA could also have been used, but RNA was available). The probe was generated by asymmetric PCR using the linearised plasmid from one of the 60S ribosomal protein clones, as a template. The probe was serially diluted to determine the detection limits (Figure 3.34) and is easily detectable to a dilution of $^{1}/_{1000}$ (~8pg).

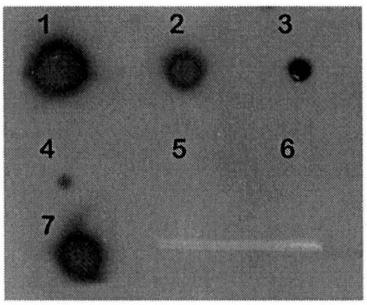


Figure 3.34 Serially diluted 60S ribosomal subunit protein probe. (1) Undiluted (~8ng); (2) Diluted $^{1}/_{10}$ (~0.8ng); (3) Diluted $^{1}/_{100}$ (~80pg); (4) Diluted $^{1}/_{1000}$ (~8pg); (5) Diluted $^{1}/_{10\,000}$ (~0.8pg); (6) Diluted $^{1}/_{100\,000}$; (7) 0.5µl of the DIG-ddUTP labelled oligonucleotide as a positive control. The probe concentration is estimated from the number of cycles and the amount of template used.

The hybridisation of the probe with the RNA on the membrane was performed at 60°C overnight, in a high SDS buffer and found to only hybridise to the RNA obtained from the parasite (Figure 3.35).

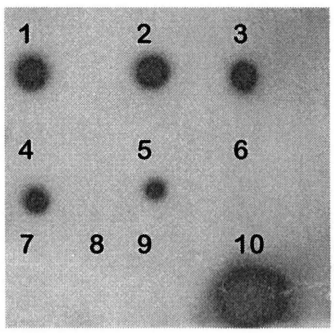


Figure 3.35 Hybridisation of the 60S ribosomal subunit protein probe to *P. falciparum* RNA. The positions on the membrane are as follows: (1-5) contain *P. falciparum* total RNA in the following concentrations (1) 3µg; (2) 2µg; (3) 1µg; (4) 0.5µg; (5) 0.1µg. Positions 6-9 are negative controls: (6) 2µg mouse liver RNA; (7) 1µg Brome Mosaic Virus RNA; (8) FORMAzol; (9) DEPC treated-water. Position (10) contains 0.5µl of DIG-ddUTP labelled oligonucleotide as positive control.

The scan shows that the *P. falciparum* RNA was detectable at 100ng which was the lowest concentration tested. It is also apparent that the blot had minimal background.

The alignment of the 60S ribosomal protein of *P. falciparum* with proteins from other organisms (Figure 3.36) is shown below.

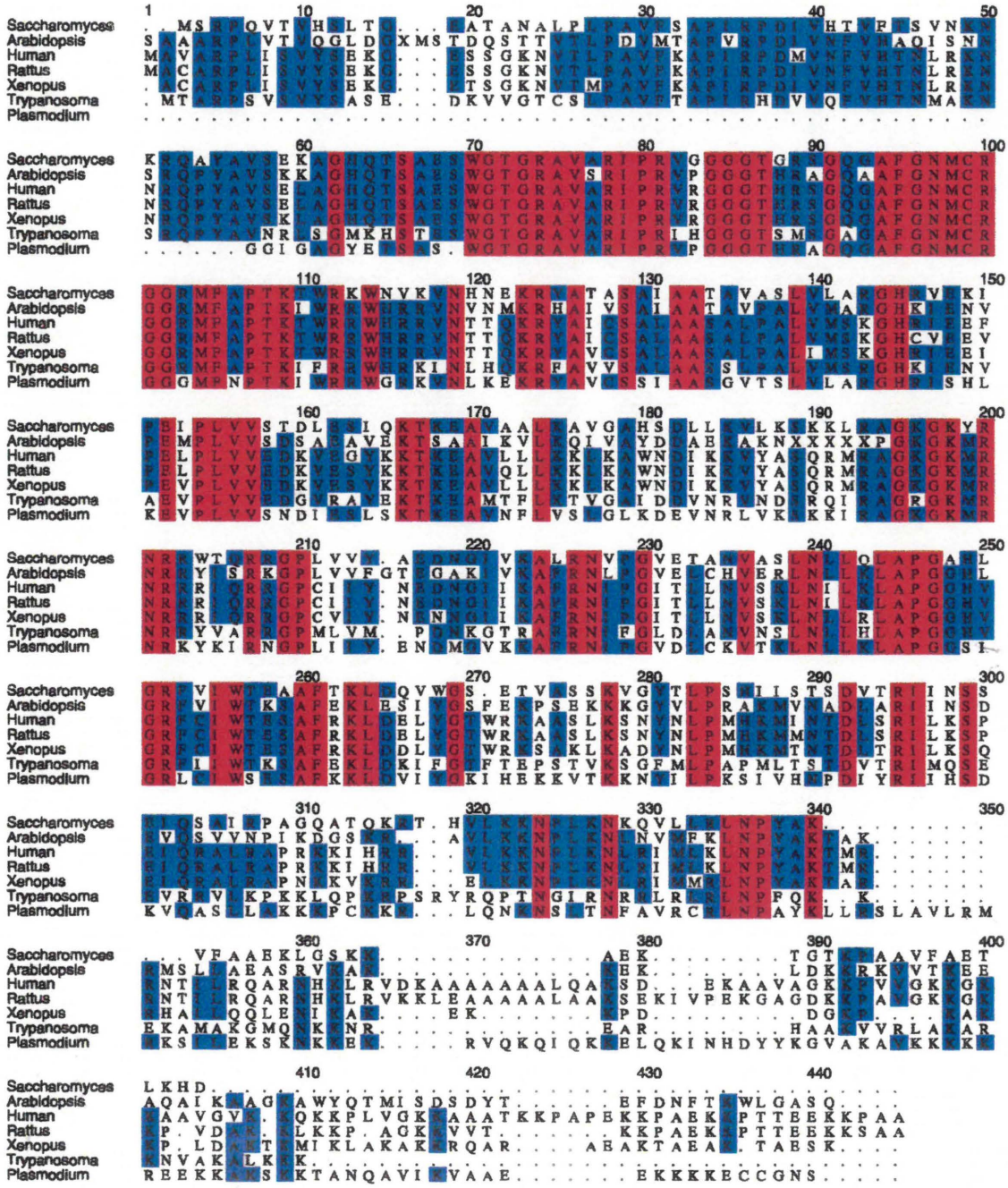


Figure 3.36 Alignment of the partial amino acid sequence of the 60S ribosomal subunit protein obtained against those of other organisms. The full length sequences are shown and red indicates 100% identity, while blue is >50% identity.

The alignment shows that the *Plasmodium* protein has fairly good homology with other organisms (an average homology of 52%) though not total identity (an average identity of 38%). The human, rat and *Xenopus* proteins are highly homologous (>75% identity), while there is 39% identity and 53% homology between the human and the *Plasmodium* proteins. The 5′-end of the *Plasmodium* sequence was not obtained as the primers annealed internally in the gene.

3.4 Discussion

This chapter presents the results of the application of 3′-RACE protocols to the AT-rich malaria genome, in an endeavour to identify a gene for an amino acid transporter. Although this objective was not reached, several obstacles were identified and have been addressed.

The study commenced with the culture of malaria parasites in an *in vitro* culture system, using the method published by Trager and Jensen (Trager and Jensen, 1976). This is a widely applied method and large quantities of asexual blood stage parasites were obtained in the ring and trophozoite stages, respectively.

Since RNA is the starting point and all further results depended on it, it was imperative that good quality RNA was obtained. RNA was isolated using TRI-reagent and the integrity was confirmed by means of denaturing agarose gel electrophoresis (Figure 3.7) (CLONTECH Laboratories, 1991; Beckler *et al.*, 1996).

cDNA was synthesised using a differential display poly-T primer, which contains the nucleotides 'VN' on the 3′-end. This serves to fix the primer at the start of the polyadenylation sequence (Thomas *et al.*, 1993), in comparison with a normal poly-T primer which may anneal anywhere on the poly-A tail, resulting in many T's before the gene itself is reached. cDNA was first synthesised from isolated mRNA, using the RNase H free Reverse Transcriptase, Superscript II (GIBCO-BRL). This enzyme is a recombinant form of the *pol* gene of Moloney Murine Leukaemia Virus (MMLV). Upon completion of the first strand, the cDNA was stored as a RNA:DNA hybrid and

used in this form for subsequent PCRs. Synthesis was monitored using DIG incorporation to determine the length of the synthesised cDNA population, which appears to be rather short (Figure 3.8). A Southern blot done subsequently with isolated mRNA (results not shown), revealed that the centring of the cDNA depends on each particular reaction and was generally slightly higher than that in Figure 3.8. The expected gene length of the amino acid permease gene is ~1700bp based on the multiple sequence alignments, but may be larger as malaria genes are prone to insertions. The band for the amino acid transporter is expected to be around ~1200bp, based on the position of primer annealing at the 5′-end of the cDNA.

The design of specific primers was addressed in Chapter 2. Procuring sequence information is relatively effortless with the WWW and the design of primers, especially those for larger templates, is expedited by the use of computer programmes such as OLIGO. If the primers designed conform to the requirements established by Rychlik (Rychlik, 1993), the PCR has a greater chance of being successful.

The optimisation of the PCR using primer PSP2 was the next important step. Barnes showed that by using a mixture of two enzymes, one highly processive, without a 3′-5′ exonuclease (proof-reading) function and a second with a proof-reading function, present in a lesser concentration (he used a 16:1 ratio), templates of up to 35kb could be amplified (Barnes, 1994). Initially Expand™ HiFi DNA polymerase was used, which is a thermostable DNA polymerase composed of two enzymes: *Taq* DNA polymerase and *Pwo* DNA polymerase. The 3′-5′ exonuclease activity confers a lower error rate (8.5×10^{-6} for *Pwo* in comparison with 2.6×10^{-5} for *Taq*) resulting in a 3-fold increase in fidelity. For a successful amplification ~10^4 template copies are required, with a primer excess to template of ~10^7 (Kidd and Ruano, 1995). Factors such as the number of cycles and the extension temperature used can influence the outcome of the PCR. The efficiency of the reaction decreases at the latter stages as the reaction components become depleted, the enzymatic activity decreases and the products accumulate. PCR products usually accumulate exponentially to a concentration of ~10^{-8}M, then linearly to 10^{-7}M (CLONTECH Laboratories, 1991).

The PCR conditions for primer PSP2 were optimised for template, primer (PSP2 and anchor) and $MgCl_2$ concentrations, along with annealing conditions (temperature and protocol). Appropriate negative controls to determine if any of the products were single primer driven or if products were formed by the primers in the absence of template, were also included (Figures 3.10 and 3.11). The negative controls indicated that the products were amplified by both primers. These typical controls were included in all new PCR reactions.

The first template concentration used (1ng of mRNA equivalents) was that recommended by Frohman (Frohman, 1993) but a higher template concentration (14ng of mRNA equivalents) was required. The observed bands (560bp and 720bp) were shorter than expected. This could have been due to: mispriming; truncated versions of the target gene being amplified; or it could be that the malaria genes are shorter than the rest. The latter is however unlikely, as malaria genes are actually prone to insertions, rather than deletions.

Initially PCR was performed at an annealing temperature of 48°C for the first four cycles in order to generate template. The annealing temperature was subsequently raised to 50°C to increase the specificity of the reaction. Little difference was observed between 48°C used initially (Figure 3.12A - lane 1) or a constant annealing temperature of 50°C used throughout (Figure 3.12B).Two different types of annealing protocols were also used, namely the use of a constant annealing temperature versus a touchdown protocol (Figure 3.12A). Touchdown PCR attempts to optimise the PCR by focusing on one variable, namely the annealing temperature, rather than analysing all the possibilities. During one amplification cycling set, a variety of temperatures surrounding the theoretical optimum, are sampled (Hecker and Roux, 1996). This is a good approach for degenerate primers where annealing temperature is not actually fixed at one temperature. Similar banding patterns (560bp and 720bp) were seen in all the optimisation experiments and little advantage was thus gained with the touchdown protocol. The 560bp and 720bp bands (generated at a Ta of 50°C - Figure 3.9B) were subsequently reamplified and cloned.

Cloning proved to be a major bottleneck in the project. The chance of mispriming with degenerate primers, is much greater and more clones than usual need to be should be screened to identify a clone containing the desired sequence. Peale *et al* (Peale *et al.*, 1998), using a multiplex display 5′-RACE protocol showed that 28% of their sequences were misprimed events, indicating that at least ten clones were required for each gene specific primer used. However, this number of clones was rarely achieved in one cloning experiment.

Since we were dealing with a gene of unknown sequence, the inclusion of restriction sites (for sticky end cloning) was undesirable, as the sites would have to be for a restriction enzyme with a GC-rich recognition site to minimise the chances that the enzyme would cleave the insert. Since restriction sites are palindromic, these GC-rich areas on the 5′-ends of the primers would anneal to each other, probably with greater stability than the AT-rich primer would anneal to the target sequence, thus removing most of the primers from the reaction. The DISEC-TRISEC method of cloning (Dietmaier *et al.*, 1993) was therefore chosen as the first method of cloning. This method failed repeatedly with no recombinant clones obtained and was particularly prone to yielding non-recombinant white clones.

Other techniques tried included: inclusion of a restriction enzyme in the ligation reaction so that dephosphorylation of the vector was not required; cycle ligation with an initial 20 cycles at 10°C for 5 minutes and 30°C for 30 minutes followed by 30°C for 30 minutes and 10°C for 2 hours for 6 cycles to improve conditions for the restriction enzyme and the ligase, respectively (Lund *et al.*, 1996). Proteinase K treatment of the PCR products prior to either DISEC-TRISEC ligation or normal blunt end ligation, to remove any contaminating proteins was also tried, as it has been shown that *Taq* DNA polymerase carryover can inhibit ligation reactions (Bennett and Molenaar, 1994; Hitti and Bertino, 1994). In all these experiments either no, or few white colonies were obtained and typically no recombinants were present.

It was at this point that we reverted to the standard blunt-ending protocol in which calculated amounts of insert and vector were required (*j/i* ratios) (Sambrook *et al.*,

1989). If *j/i* >3 then self ligation tends to occur, whereas if *j/i* <1 then intermolecular ligation is favoured. A ratio of 2 was used, which should allow intermolecular ligation (joining of insert to vector), followed by ten-fold dilution to promote intramolecular ligation (joining of the recombinant molecule) (Damak and Bullock, 1993). Calculations indicated that 848ng of insert (560bp) was required and the PCR products had to be reamplified to obtain sufficient insert. The ligation reaction was performed after co-precipitation of the insert and the vector to reduce the volume to that required in the ligation reaction. The entire reamplified PCR mix was silica purified and cloned. In addition to the expected 560bp and 720bp sizes, inserts of other sizes were also obtained. These may have been bands present in the background, but were invisible on the gels due to low concentrations.

Sequences were obtained manually (Table 3.9). The expected sequence of an amino acid transporter was not present among the clones. Some of the clones contained *P. falciparum* gene sequences, some contained *E. coli* gene sequences and some of the clones could not be identified. It was unclear at this stage whether the *E. coli* sequences could be due to contamination or possible gene rearrangements, since no primer sequences could be identified in some of the clones (including all those containing *E. coli* sequences).

The *P. falciparum* clones included genes for S-adenosylhomocysteine hydrolase (SAHH), the Ras-related Ran/TC4 GTP binding protein, heat shock protein 90. None of these clones bore any similarity to the amino acid permease genes. Of these clones, SAHH, is a potential therapeutic target as it is involved in methylation reactions and its interruption is a means to interfere with multiple metabolic pathways (Bryant and Behm, 1989; Creedon *et al.*, 1994). The GTP-binding protein is involved in chromosome condensation and the initiation of mitosis, transport of proteins into the nucleus as well as coupling the completion of cDNA synthesis with the onset of mitosis (Dontfraid and Chakrabarti, 1994), while the heat shock protein, among other functions, has chaperone function and is associated with numerous protein kinases, some nuclear receptors and cytoskeletal proteins (Bonnefoy *et al.*, 1994). An important

aspect is that all these proteins were identified by means of cDNA libraries, which is markedly different from the RT-PCR based approach adopted here.

Later studies using the automated nucleotide sequencer revealed the identity of the questionable clones (Table 3.10). The biggest advantage of the automated sequencer is that up to ~650bp can be read, compared with the ~245bp (per loading) obtained with manual sequencing. Of the clones, one was the *P. falciparum* ribosomal phosphoprotein P2 (both PSP2 and poly-T primers were identified), which is part of a family of acidic ribosomal phosphoproteins containing P0, P1 and P2 (Goswami *et al.*, 1996). Clone 45 appeared to have undergone rearrangement (see later) as the T3 side was *Plasmodium* (poly-T primer was identified) and the T7 side, *E. coli* (no primer identified). Clone 49 was an *E. coli* mhp cluster for 3-hydroxy-phenylpropionic acid degradation, from both the T3 and T7 sides and no primers were identified. Clone 53 still showed no significant homology with any of the genes available in the databases and no primers were identified in the sequence. The AT content was ~47% which suggests that it is unlikely to be of malarial origin.

The problems encountered at this time were exacerbated by the thermostable DNA polymerase, Expand™ HiFi. As the enzyme aged, PCR results were not reproducible and at times no products, or smears (such as those in Figure 3.12A), were obtained. This type of artefact has been observed in experiments using proof-reading DNA polymerase mixtures and it has been speculated that it may be one of the following: amplified contaminating DNA, complexes of primer and enzymes, primer-dimer concatamers or coagulates of buffer components (e.g. BSA) (Hengen, 1994). It is also possible that the mixture of enzymes degrade unequally, resulting in smearing. This resulted in insufficient PCR products with which to attempt cloning. Takara *ExTaq* DNA polymerase (a mixture of *Taq* and *Pfu*) was subsequently used and found to yield reproducible results.

Alignments of the position of primer PSP2 annealing (Figure 3.15) indicated that mispriming and rearrangements (e.g. Clones 17 and 36) were occurring. The rearrangements were in the form of missing primers, or nucleotide mutations, deletions

or insertions in the primers, as well as larger rearrangements where one side of the insert was replaced with a bacterial sequence. It is unlikely that these nucleotide differences may have occurred during synthesis of the primer as it only occurs in some primers and not in all of them. The primer manufacturers also supplied evidence of primer purity in the form of polyacrylamide gel electrophoresis of the primers. The larger rearrangements probably occurred during propagation of the bacteria as the bands could be amplified with both primers prior to cloning. These rearrangements are highly unusual as the clones are in SURE *E. coli,* which are supposed to be recombination deficient, although some researchers have found that rearrangements do still occur. The rearrangements also make identification of clones difficult. Mispriming may be due to:

- the stringency of the PCR was not high enough;
- the template concentration was too low or there was no template present at all;
- the primer was too degenerate; and
- the primer contained too many A's and T's (16 of the 21 nucleotides are potentially an A or T).

The number of A's and T's in the primer may affect the PCR as described for mispriming above. A more detailed analysis of the mispriming shows that an A or T mismatch is usually, but not exclusively the culprit (Figure 3.16). This is probably since 'T' is a promiscuous nucleotide that will allow effective amplification even when annealed to G, C or T (Kwok *et al.*, 1990; Hyde and Holloway, 1993). Since the malaria parasite has an AT-rich genome, this is a larger problem than with GC-rich organisms. Primer PSP2 was a badly designed primer as it contains, among others, the following type of primers in the mixture:

```
5′ TTA ATT GGT ACT GGT TAT TTT 3′
         A   A   A   A A A A
```

This primer is highly susceptible to mispriming as a result of the numerous T's and A's on the 3′-end. The fact that no clones were obtained from a primer consisting only of A and T at the 3′-end, indicates that it annealed to so many different templates that amplification of any one specific template was probably impossible.

A means of increasing the specificity of the PCR is to increase the length of the primer. Unfortunately this was not really possible for the permease specific primers as the identified consensus region was too short and extending the amino acids would increase the degeneracy substantially.

Consequently, primer PSP was designed to have more G's and C's, especially at the 3′-end (primer PSP) as depicted below:

PSP2: 5′ TTA RTW GGW RCW GGW NAW DTT 3′

PSP: 5′ GGA GGW RYI ATW GGW RCW GGW 3′

A third primer (TYT) which was specific for tyrosine and tryptophan transporters was also designed. As elaborated on in Chapter 2, primers PSP and TYT were better primers since they contained a better distribution over the consensus sequence and most importantly had a lower degeneracy. They also have an improved G and C distribution at the 3′-end.

The cDNA concentration, primer concentrations and $MgCl_2$ concentration were optimised for primer PSP using the Taguchi protocol. This method is an efficient way of optimising PCRs (Cobb and Clarkson, 1994). It is apparent that Takara *ExTaq* DNA polymerase also requires less magnesium (2mM) to amplify products than Expand™ HiFi (4mM). This may be advantageous as the specificity is increased at lower $MgCl_2$ concentrations.

Since it appeared that an increased template concentration resulted in additional bands, the concentration was increased further to 370ng of mRNA equivalents (Figure 3.18). In addition to a greater number of bands, a band of the expected size (~1200bp) was also visible for the first time, suggesting that some gene copies are now within the 10^4 copy level required for effective amplification. These results (Figure 3.18) led to the conjecture that some gene copies may have been lost during mRNA isolation, due to saturation of the oligo(dT) beads by non-polyadenylated RNA or genomic DNA, both

annealing via internal stretches of 'A'. This subsequently led to an investigation of the quality and quantity of cDNA prepared from isolated mRNA and total RNA.

The quality of the cDNA was determined using a gene specific primer for the 18S rRNA (Bio1) and the anchor on both cDNA derived from mRNA and cDNA derived from DNase-treated total RNA. Multiple bands appeared after the Bio1 and anchor primer pair PCR, using cDNA from isolated mRNA (Figure 3.19A). Most of these are likely to be truncations of the one correct 18S rRNA band of ~1800bp. These truncations are dramatically reduced when total RNA is used instead of mRNA (Figure 3.19B). It is possible that the truncations are formed by the presence of excessive amounts of poly-T primer which then anneals internally. However, a ratio of 1.5pmol poly-T primer: 1μg isolated mRNA was used, which is less than the 100pmol primer:1μg mRNA (theoretical) used for successful PCRs from total RNA. It is however likely, that since there were so many templates present in the cDNA preparation from total RNA, that the poly-T primer concentration relative to the mRNA in the total RNA, was diluted to levels below those of isolated mRNA. The 18S rRNA gene contains internal regions of poly-A where the poly-T primer may have annealed to form the truncated products (Figure 3.20). It is apparent that the primer annealed to less than the 18 A's required, possibly priming from as little as 4 A's. This observation led to the 'hot start' method of cDNA synthesis where the reaction components were equilibrated at the synthesis temperature prior to addition of the reverse transcriptase.

The quantity of cDNA was determined using the 18S rRNA specific Bio1 and Dig2 primers, with cDNA derived from DNase-treated RNA. Both 0.5:1 and 5:1 ratios of poly-T primer to total RNA were evaluated and the 5:1 ratio yielded the best results (Figure 3.21). Ideally a 10:1 ratio of poly-T to RNA should also have been tested. This has subsequently been tried in the laboratory and appears to yield improved results (L. Birkholtz, personal communication). Spurious bands were obtained in the undiluted cDNA generated with a 5:1 ratio. This may have been due to non-specific priming in the presence of excessive cDNA, as the rRNA gene amplified by Bio1 and Dig2 is of high abundance. This explains why the $^{1}/_{100}$ dilution gives the best amplification,

followed by the $^{1}/_{10}$ dilution and then the undiluted cDNA, with the non-specific priming increasing respectively. The intensity of the 408bp band is similar in all instances suggesting that a plateau phase has already been reached at a dilution of $^{1}/_{100}$.

Initially, the FORMAzol, in which the isolated total RNA is stored, was removed by precipitation, as it was possible that the FORMAzol could inhibit the PCR. This was shown to be unnecessary and the practice was discontinued as large quantities of RNA were lost in the precipitation steps. Traces of contaminating genomic DNA were removed from the total RNA using RNase-free DNase, but experiments showed that this procedure led to a reduced number of copies (Figure 3.22).

These experiments thus suggested that gene copies, of even high abundance genes, are lost during the isolation of mRNA or the DNase treatment of total RNA. Assuming that traces of genomic DNA would not affect a RACE protocol when an anchor primer is used in combinations with a gene specific primer, cDNA was thus synthesised using untreated, unprecipitated total RNA.

Primer PSP2 was tested again on the improved cDNA preparations. The 560bp band was still present, but was not the only dominant band, since a strong band was observed at the desired size of ~1200bp (Figure 3.23). Less cDNA (theoretically ~4.4ng of mRNA) was used, in comparison with where cDNA was synthesised from 100ng of mRNA equivalents (Figure 3.9B), which confirms that less gene copies were lost or not transcribed during reverse transcription. Primer PSP and PSP2 both gave the ~1200bp band on the cDNA derived from total, untreated RNA and the banding patterns between the two primers were similar (Figures 3.22 and 3.8B).

Since problems were encountered with the cloning of PCR bands, kits were used in later studies in an attempt to increase the cloning efficiency. The pGEM® T-Easy Vector System (Promega) yielded more clones than the pCR-Script™ Amp (SK+) Cloning Kit (Stratagene), but neither worked as well as expected. The 1200bp band generated with PSP2 at 50°C was cloned using the Stratagene kit and proved to be the 5.8S rRNA gene, which occurs at high copy number in the genome (see Table 3.15).

Since the results obtained with primer PSP2 confirmed that it was an unsatisfactory primer, it was decided to continue with primers PSP and TYT. The annealing temperature was optimised for primer PSP (Figure 3.24A-E). The two bands in the vicinity of 1200bp as well as the 830bp band (all generated at 55°C) were chosen for cloning, as the controls showed that anchor primer driven products of high molecular mass (above 2000bp) were more prevalent at higher temperatures (results not shown). No clones were obtained for the 830bp band and the sequences of the clones from the 1200bp bands proved to be either the ubiquitous 5.8S rRNA gene or one of the proteins present in the large (60S) subunit of the ribosome (Table 3.13). Clone p5 shows fairly good homology with a kinesin-like protein (Figure 3.25) and with an AT-content of ~70% is most likely of malarial origin. Clone 7 however, appears to have undergone rearrangement with the cloning vector on the T7 side, while the T3 side shows no significant homology with any of the genes in the NCBI databases. Of the three unknown clones (53 from PSP2; p5 and 7 from PSP), none seemed suitable for transmembrane analysis, as the likelihood that they were amino acid permeases was excluded either by AT-content or homology matching. The PCR for primer TYT was optimised (results not shown) and the 1100bp and 950bp bands generated at an annealing temperature of 48.2°C (Figure 3.28) were cloned. Their sequences also corresponded to the same 60S ribosomal subunit protein (Table 3.14) obtained using primer PSP, although their priming sites on the gene differed (Figure 3.30). One of the clones generated with primer TYT, showed a rearrangement at the 5′-end where the *Plasmodium* sequence was substituted with transposon Tn10 (see later) and another of the clones showed homology with the *P. falciparum* ubiquitin protein (Figure 3.31).

From the positions of primer PSP and PSP2 annealing to the 5.8S rRNA gene (Figure 3.26), it appears that all the PSP clones underwent some form of rearrangement (see later) as the 5′-end of all the primers was missing. Both primers PSP and PSP2 primed from approximately the same position. cDNA is probably synthesised from rRNA by the poly-T primer annealing internally in the gene, although this should be minimised by the use of the 'hot start' step during cDNA synthesis. However, mispriming was not unexpected since the rRNA genes are highly abundant (~90% of the total RNA in both eukaryotic and prokaryotic cells). Table 3.15 shows the type of RNA found in *E. coli*.

The *E. coli* information is presented, as this type of information is not that readily available for eukaryotes.

Table 3.15 The types of RNA found in *E. coli* (Zubay, 1993).

Type	Function	Number of different kinds	Number of nucleotides	% of synthesis	% of total RNA in cell
mRNA	Messenger	Thousands	500-6000	40-50	3
rRNA	Structure and function of ribosomes	3 : 23S 16S 5S	2800 1540 120	50	90
tRNA	Adapter	50-60	75-90	3	7
RNA primers	DNA replication	?	<50	<1	<1
RNA component of RNase P	Catalytic	1	377	<1	<1

Since greater numbers of clones were being obtained, a method of screening the recombinant clones for those with the same type of insert was required. Two or three of the clones were sequenced and a restriction enzyme site in the insert identified. The clones were then digested with the enzyme and the restriction patterns observed (Figure 3.27). All the clones were subsequently sequenced to confirm the allocations.

Since Clone 15 (generated with primer TYT) contained a rearranged insert (with Tn10), the size of the band excised from the plasmid by restriction digestion, was checked. The excised band had the same size as that of Clone 14 which was not rearranged (Figure 3.29).

Along with mispriming and truncations, rearrangements were an obvious problem. It appears that the AT-rich malaria genes are unstable in *E. coli* which then compensates by performing rearrangements (Triglia *et al.*, 1992). This could be the reason why, during the course of the study, various rearrangements were observed. These included the missing 5′-end of primer PSP in the 5.8S ribosomal clones, the deletions and insertions in the primer which occurred in the SAHH and GTP-binding protein clones,

the *E. coli* gene for rRNA as part of the insert in Clone 45 and the presence of transposon Tn10 on the 3′-end of one of the 60S ribosomal subunit clones. The appearance of this transposon at the 3′-end is a common occurrence in the lab (K. Nel and L. Birkholtz, personal communication). The rearrangements were obtained with both SURE (recombination deficient) and DH5α *E. coli*. SURE *E. coli* are preferable to DH5α when transforming ligation reactions and subsequently culturing the bacteria, even though DH5α have a better competency and yielded more colonies (results not shown). This is because SURE *E. coli* contain mutations that eliminate a number of independent DNA repair pathways as well as other mutations rendering them deficient for homologous recombination. Hanahan (Hanahan *et al.*, 1991) has suggested additional steps to minimise rearrangements. These include: ensuring that the culture does not grow past the mid to late logarithmic growth stage (this is when nutrients become scarce and the culture becomes stressed) and growing the cultures in a nutrient rich broth at 30°C instead of 37°C. These factors are important since rearrangements were found to occur in the SURE *E. coli* cells. Since the perception of rearrangements only became apparent with the use of the automated sequencer, limited time did not allow the repeat of the experiments under the new conditions. However, the conditions have been tested by others in the lab and appear to be successful (K. Nel and L. Birkholtz, personal communication). It has also been noted that the DNA encoding mammalian transmembrane proteins is difficult to manipulate in *E. coli*, undergoing gross rearrangements, deletions and point mutations. This has been shown for the cystic fibrosis transmembrane regulatory protein (Gregory *et al.*, 1990; Krauss *et al.*, 1992) and the Menkes (transmembrane) protein (La Fontaine *et al.*, 1998). Replacing the origin of replication that replicates to high copy number (such as those found in pGEM, pUC and pBluescript) to those that replicate to low or moderate copy number can reduce the rearrangements observed (Peden, 1992; La Fontaine *et al.*, 1998). The use of a plasmid which replicates to low or moderate copy numbers should also be tested with malaria genes, especially since Hanahan's recommendations reduce the stress on the bacteria and thus also effectively decrease the copy number of the plasmid.

A clone from the TYT primer showed extensive homology with ubiquitin. Ubiquitin is a 76 residue protein with a highly conserved sequence (Figure 3.32). It is involved in the ATP-dependent selective degradation of cellular proteins, the maintenance of chromatin structure and ribosome biogenesis (Adams *et al.*, 1992). Ubiquitin is usually synthesised as a polyubiquitin precursor, where the number of exact head-to-tail repeats depends on the organism (e.g. *Xenopus* may have up to twelve). The *Plasmodium* gene appears to contains two exact head to tail repeats. However, *Plasmodium* may have a cluster of ubiquitin genes, each containing different numbers of repeats (J. Hyde, personal communication). The final amino acid after the last repeat also differs between organisms. In *Plasmodium* this appears to be phenylalanine (Figure 3.32), while in *C.elegans* it is aspartic acid-isoleucine, in humans it is valine and in chickens, tyrosine. A single copy of ubiquitin may sometimes be attached to a ribosomal protein. The 3′-UTR of the *Plasmodium* gene is 83% AT rich. This conforms with the extreme AT-richness observed in the flanking regions of *Plasmodium* genes. No polyadenylation signal was observed, but since only one clone was obtained, this cannot be guaranteed.

It was concluded that the clones from primers PSP and TYT were a 60S ribosomal protein, based on various computer programmes. The first was BLAST N (searches for nucleotide homology) where clones showed little homology with any gene at nucleotide level. Using BLAST X (which reverse translates the nucleotide sequence into all six amino acid reading frames and then searches for homology), the clones showed homology with the 60S ribosomal protein of the L1E family, from various organisms. The nucleotide sequence was subsequently reverse translated with the genetic data environment programme (Figure 3.33) and aligned with the proteins from these other organisms using Clustal W (Figure 3.36). The protein showed extensive homology with these proteins.

It was proved that the 60S protein recombinants were of malarial origin by hybridisation of a single stranded PCR generated probe to malaria RNA. This hybridisation could be detected at an RNA concentration of 100ng which was the lowest concentration tested (Figure 3.35).

The eukaryotic ribosome is composed of a 40S (small) subunit and a 60S (large) subunit which, when combined, form the 80S ribosome. The subunits are composed of ribosomal RNA (~60%) and proteins (~40%) and are held together by magnesium. The large subunit contains 28S, 5.8S and 5S rRNA together with approximately 45 proteins, while the small subunit contains 18S rRNA with approximately 30 proteins. Each of these ribosomal proteins tends to be unique and are named L1, L2,... and S1, S2,... for the large and small subunits respectively. Most of the ribosomal proteins tend to be basic (Adams *et al.*, 1992). Ribosomes have been a classic target for antibiotic treatment as prokaryotic ribosomes differ from eukaryotic ribosomes. Figure 3.36 shows that there is some difference between the *Plasmodium* and human proteins and further investigation of these proteins as potential chemotherapeutic targets should be considered.

3.5 Conclusion

This chapter deals with the application of RACE and RT-PCR techniques to the AT-rich genome of *Plasmodium*. The applications are not limited to malaria, as many other organisms, including *Dictyostelium* and *Bacillus*, also have an AT-rich genome, although it may not be as pronounced as it is with *Plasmodium*.

Several problems were encountered during the course of the search for an amino acid permease gene. These included mispriming, truncated cDNA, loss of gene copies and rearrangements of the cloned gene fragments within the bacteria. These problems were elaborated on, and solutions presented for the synthesis of improved cDNA with less truncations and a higher yield. Suggestions for the minimisation of rearrangements were also proposed.

The amino acid permease gene was not obtained. This may be as the result of various factors namely, the template concentration may still be too low, not enough clones were obtained to identify the desired gene from among the background of misprimed products, primer error (i.e. the identified consensus sequence does not occur exactly the same in *P. falciparum)*, or more likely that the 3′ terminal is not exactly right for

amplification. The rearrangements which occur also make identification of the clones difficult, if not impossible.

Alternative methods for the identification of the amino acid permease gene are presented in Chapter 4.

CHAPTER 4

Concluding Discussion

Malaria is an age old scourge of mankind. In the 1950's it was thought that the eradication of malaria was imminent, with the discovery of cheap and effective antimalarial drugs and DDT as a potent insecticide. Unfortunately, this was unfounded and malaria is as big a problem today as it ever was, due to the increasing spread of drug resistant malaria strains. As a result, research into methods of combating malaria is occurring on a fairly large scale.

The development of a continuous culture method as well as advancements in the field of molecular biology has contributed greatly to the knowledge concerning parasite biochemistry, physiology and the host-parasite relationship. An increased appreciation of the importance of these aspects of the parasite has also accelerated research (Ginsburg and Stein, 1987; Trager and Jensen, 1997). This knowledge may lead to successful strategies to control malaria (Ginsburg and Stein, 1987).

Chapter 2 deals with the identification of a consensus sequence and the design of suitable PCR primers for the identification of a *P. falciparum* amino acid permease gene. Primer design is a factor which could influence the outcome of the PCR amplification. The satisfactory design of a suitable primer requires adequate sequence information on which to base the primer sequence, which is easily obtained using the WWW. The degeneracy should be as low as possible, with the 3′-end of the primer having a thermodynamic value between -5 and -10 kcal/mol. Degeneracy can be decreased with the use of codon preferences (both normal and Match Index), 'T' mismatches, as well as inosine if the resulting degeneracy is still too high.

A consensus sequence was originally identified using CLUSTAL W aligned amino acid transporter sequences from a variety of organisms. This consensus sequence was later confirmed using the BLOCKS database. Three primers (PSP2, PSP and TYT)

were designed based on this sequence. These primers were evaluated using OLIGO. Of the primers, PSP is better than PSP2 as it conforms to more of Rychlik's rules for a successful primer: it is less degenerate, it has a better positioning over the consensus sequence and it has more G's and C's at the 3′-end. Primer TYT was a more specific primer for tyrosine and tryptophan transporters.

Chapter 3 deals with the application of RACE and RT-PCR techniques to organisms containing an AT-rich genome. This is not just *Plasmodium* as *Dictyostelium*, *Bacillus* and some others also have this tendency, although it may not be as pronounced as it is in *Plasmodium*.

One of the main problems encountered during this study was the successful synthesis of full length cDNA as the quality of the cDNA plays an essential role in the success of the amplification. Since the malaria genome is 80% AT-rich, the likelihood of oligo(dT) bead saturation during mRNA isolation using the poly-A tail is greatly increased. The saturation is not necessarily due to poly-A tail annealing but rather through stretches of internal 'A' in either mRNA, rRNA or genomic DNA. This means that the mRNA preparation contains traces of genomic DNA and also that mRNA copies are lost. It is generally assumed that membrane proteins are encoded by genes with a rare copy number and it seems that the isolation of mRNA results in the loss of gene copies and such a loss would be amplified when rare gene copies are under investigation. Truncated cDNAs are also formed, possibly as a result of excess poly-T primer. These truncated molecules then give rise to truncated PCR products. The appearance of truncated products also led to the use of a 'hot start', where the reaction components were incubated at the synthesis temperature for 2 minutes prior to the addition of the enzyme in an attempt to allow more specific poly-T primer annealing to the mRNA.

Ideally, mRNA should thus not be used when working with an AT rich genome. This means that total RNA should be used, which results in rRNA genes being amplified during the PCR. It was also found that removal of contaminating traces of genomic DNA with a RNase-free DNase also resulted in a loss of RNA copies. This is probably

due to precipitation, as the same effect was noted with precipitation of the total RNA from FORMAzol. The use of untreated (DNAse) total RNA also appears to have minimised the number of truncated products formed. The amount of poly-T primer to total RNA should also be considered. A 5:1 ratio (pmol primer : μg total RNA) yielded improved cDNA and it has been shown that a 10:1 ratio improves the quality further (L. Birkholtz, personal communication). Too little primer results in incomplete synthesis due to internal annealing and extension.

Once cDNA of sufficiently high quality and quantity is obtained, the PCR can be performed. The PCR can be optimised by adjusting factors such as template, $MgCl_2$ and primer concentrations; the annealing temperature and time. All these factors are dependent on the gene of interest as well as the properties of the primers. The optimisation can be expedited by the use of the Taguchi method.

Numerous clones were obtained during this study, although none bore any resemblance to an amino acid transporter gene. This was mostly due to mispriming by the respective primers. The partial sequences of these clones was obtained. Most of them have already been completely sequenced and characterised, although a novel *P. falciparum* 60S ribosomal subunit protein of the L1E family was obtained. The partial sequence was obtained and the malaria origin confirmed using hybridisation of a single stranded PCR generated probe to total RNA.

Rearrangements in the inserted fragment appear to be common and necessitates the use of recombinant negative *E. coli* strains such as SURE. Further recommendations include: that the bacterial cultures should not be grown past the logarithmic growth phase, to limit nutritional stress on the bacteria, nutrient rich growth medium should be used with propagation at 30°C (Hanahan *et al.*, 1991). Although time did not allow the repetition of the experiments under these new conditions, they have been tested in the lab and appear to work satisfactorily (K. Nel and L. Birkholtz, personal communication). The use of a plasmid which replicates to low or moderate copy number should also be tried for *Plasmodium* genes.

The absence of the amino acid permease could be due to three reasons. Firstly, the template concentration may still be too low, despite the attempts to minimise the loss of copies during the manipulation required for cDNA synthesis and PCR. Secondly, the gene may be lost due to the extension temperature used. Wellems and colleagues have shown that the gene for dihydrofolate reductase is only amplified at an extension temperature of 60°C and not at the generally used 72°C, as DNA melting prevents *Taq* DNA polymerase extension at 72°C (Su *et al.*, 1996). Lastly, the identified consensus sequence may not be present in the *P. falciparum* genome and/or the amino acid transporter may not be present in *Plasmodium*.

Since the glucose transporter for *P. falciparum* was identified from cDNA synthesised with a gene specific primer, but not from the cDNA synthesised with a poly-T primer (K. Nel, personal communication), it appears that the issue of low template concentration is valid. This may be addressed in numerous ways as first strand cDNA can be synthesised using two methods:

- Firstly there is the classical method using a poly-T primer. This is only possible if the mRNA is polyadenylated. This method is preferred by Clontech as it produces fewer side products (CLONTECH Laboratories, 1991).
- Lastly, a gene specific primer can be used for the synthesis of the first strand (CLONTECH Laboratories, 1991). This method is best if two gene specific primers are available. It can however be applied if the template switching nature of Superscript II Reverse Transcriptase is exploited. Superscript tends to add 'CCC' to the end of the first strand cDNA. If an anchor primer containing 'GGG' on the 3′-end is used in the cDNA synthesis reaction, the G's and C's anneal and the reverse transcriptase switches strands and adds the complementary of the anchor to the 5′ end of the cDNA. The anchor primer can then serve as the second PCR primer in conjunction with a single gene specific primer (CLONTECH Laboratories).

We used the first method of cDNA synthesis during this study and used a differential display poly-T primer, containing the nucleotides 'VN' on the 3′-end. If this method is to be used, it is advisable to equalise the gene pool, especially for rare gene copies. This may be done using four separate poly-T primers with the 3′ nucleotide being

either A,C,G or T instead of a degenerate primer with N as the 3′ nucleotide. This will result in enriched pools of cDNA containing the appropriate nucleotide.

However, in retrospect, the third method of cDNA synthesis, using a gene specific primer and effectively performing 5′-RACE instead of 3′-RACE may be the best way to generate cDNA of the desired nature. Suppression PCR may also be incorporated to reduce non-specific background amplification. This involves the ligation of specifically designed, GC-rich, double-stranded adaptors to double stranded cDNA using T4 DNA ligase. The adaptor ligated double-stranded cDNA can then be used to amplify either 5′ or 3′ specific fragments using a combination of gene specific and adaptor specific primers. If cDNA molecules appear that have adaptor sequences at both ends (products that are not driven by the gene specific primer), the long inverted terminal repeats in the 5′ and 3′ adaptors anneal to each other creating a 'panhandle'-like structure which cannot be amplified (Chenchik *et al.*, 1996). This will effectively remove any potential products that are not driven by both primers. If after this, the amino acid permease is still not obtained, it may have to be concluded that the identified consensus sequence is not present in the *P. falciparum* genome.

Since Ronald Ross discovered the vector of the malaria parasite in 1897, much has progressed in the field of malaria research. Drug resistance to common inexpensive drugs has arisen, a vaccine candidate has undergone clinical trials, the parasite genome is in the process of being sequenced and stable transfection means that the parasite can be modified (Hagan and Chauhan, 1997). The search for therapeutic targets is an ongoing and important facet of all malaria research, both in the first and third world, although the third world has little funding for research on parasitic diseases, but is the area where the eventual applications are most urgently needed.

SUMMARY

Malaria is a life-threatening parasitic disease, which is becoming increasingly problematic in Sub-Saharan Africa due to the spread of drug-resistant parasite strains. Novel therapeutic targets and drugs are thus required to successfully treat and/or prevent malaria. The aim of this study was the identification and characterisation of an amino acid transporter gene of the human malaria parasite, *Plasmodium falciparum*, as a possible therapeutic target. Since no sequence information was available, 3'-RACE (a modified form of RT-PCR) was chosen as a method to obtain the gene sequence. Some of the obstacles encountered and the way that they were addressed are described.

The synthesis of a representative, full-length, uncloned cDNA library of the AT-rich malaria genome was one focal point of the study. The best quality cDNA was obtained when total RNA was used, with a 'hot start' protocol prior to synthesis. Three gene specific primers were designed, based on an identified consensus sequence of multiply aligned amino acid permeases from other organisms. The sequence of a transporter gene was not obtained (due mostly to mispriming), though many partial gene sequences of known therapeutic targets were obtained. Numerous rearrangements of the inserts by the bacterial host were observed and methods of minimising this were suggested. The absence of the amino acid transporter may be due to a template concentration that is too low, a slightly different consensus sequence or the absence of the corresponding gene in the malaria parasite. Possible means of continuing with further studies are proposed.

A partial sequence of one novel protein, of the large (60S) ribosomal subunit was obtained. It was confirmed, by hybridisation of a probe to total RNA, that the sequence was of malarial origin.

OPSOMMING

Malaria is `n lewensgevaarlike parasitiese siekte, wat al hoe ernstiger raak in Sub-Sahara Afrika, as gevolg van die verspreiding van parasietstamme wat weerstandbiedend teen bestaande medisyne is. Nuwe terapeutiese teikens en middels is dus nodig vir die suksesvolle behandeling en/of voorkoming van malaria. Die doel van hierdie studie was die identifikasie en karakterisering van `n aminosuur transporter geen van die menslike malariaparasiet, *Plasmodium falciparum*, as `n moontlike terapeutiese teiken. Aangesien die aminosuur transporter se volgorde onbekend is, was 3'-RACE (`n gemodifiseerde vorm van RT-PKR) gekies as `n metode om die geenvolgorde te verkry. Probleme wat ervaar was en die maniere hoe dit aangespreek is, word beskryf.

Die sintese van `n verteenwoordigende, vollengte, ongekloneerde cDNA biblioteek van die AT-ryke malaria genoom, was een fokuspunt van die studie. Die beste gehalte cDNA sintese is verkry deur totale RNA te gebruik met `n 'hot start' protokol, voordat sintese begin is. Drie geenspesifieke voorvoerders is ontwerp, gebaseer op `n konsensus volgorde wat geïdentifiseer was, na `n vergelyking van die geenvolgordes van die aminosuurpermeases van ander organismes. Die volgorde vir `n aminosuurpermease geen is nie verkry nie (meestal as gevolg van foutiewe annelering), alhoewel heelwat gedeeltelike geenvolgordes van reeds bekende terapeutiese teikens verkry is. Herrangskikkings van die invoegsels deur die bakteriële gasheer het geredelik voorgekom en metodes word voorgestel om dit te verminder. Die afwesigheid van `n aminosuur transporter mag wees as gevolg van `n templaatkonsenstrasie wat te laag is, `n effens anderse konsensus gebied, of die afwesigheid van die transportergeen in die malariaparasiet. Moontlike maniere vir die voortsetting van die studie word voorgestel.

`n Gedeeltelike volgorde van een onbeskryfde proteïen, van die groot (60S) ribosomale subeenheid, is verkry. Die oorsprong van die geenvolgorde as synde van die malariaparasiet, is bevestig deur hibridisasie van `n peiler aan totale RNA.

REFERENCES

Ackerman, J. (1997) Parasites. Looking for a free lunch. *National Geographic* **192**, 74-91.

Adams, R.L.P., Knowler, J.T. and Leader, D.P. (1992) *The Biochemistry of Nucleic Acids*. London, Chapman and Hall.

Alonso, P.L., Molyneux, M.E. and Smith, T. (1995) Design and methodology of field based intervention trials of malaria vaccines. *Parasitology Today* **11**, 197-200.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215**, 403-410.

Amador, R. and Patarroyo, M.E. (1996) Malaria vaccines. *Journal of Clinical Immunology* **16**, 183-189.

Ansorge, I., Paprotka, K., Bhakdi, S. and Lingelbach, K. (1997) Permeabilisation of the erythrocyte membrane with streptolysin O allows access to the vacuolar membrane of *Plasmodium falciparum* and a molecular analysis of membrane topology. *Molecular and Biochemical Parasitology* **84**, 259-261.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Sneidman, J.G., Smith, J.A. and Struhl, K. (1992) *Short Protocols in Molecular Biology*. New York, Greene Publishing Associates and John Wiley & Sons.

Barnes, W.M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proceedings of the National Academy of Sciences of the USA* **91**, 2216-2220.

Basco, L.K., de-Pecoulas, P.E., Wilson, C.M., Bras, J.L. and Mazabraud, A. (1995) Point mutations in the dihydrofolate reductase-thymidylate synthase gene and pyrimethamine and cycloguanil resistance in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **69**, 135-138.

Beckler, G., Bhatia, A., Brisco, P., Brondyk, W., Butler, B., DeMars, S., Doers, M., Doyle, K., Goueli, S., Harper, T., Knoche, K., Knuth, M., Lepinske, M., Lesley, S., Marcus, L. *et al.* (1996) *Promega-Protocols and Applications Guide*. Madison, Promega Corporation.

Bennett, B.L. and Molenaar, A.J. (1994) Cloning of PCR products can be inhibited by *Taq* polymerase carryover. *BioTechniques* **16**, 32-33.

Berendt, A.R., Turner, G.D.H. and Newbold, C.I. (1994) Cerebral malaria: The sequestration hypothesis. *Parasitology Today* **10**, 412-414.

Billingsley, P.F. and Sinden, R.E. (1997) Determinants of malaria-mosquito specificity. *Parasitology Today* **13**, 297-301.

Birkholtz, L., Visser, L., Brink, A. and Louw, A.I. (1998) Drug-resistant and mixed species malaria infections in Mpumalanga, South Africa. *South African Journal of Science* **94**, 39-43.

Bonnefoy, S., Attal, G., Langsley, G., Tekaia, F. and Mercereau-Puijalon, O. (1994) Molecular characterisation of the heat shock protein 90 gene of the human malaria parasite *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **67**, 157-170.

Borst, P. and Ouellette, M. (1995) New mechanisms of drug resistance in parasitic protozoa. *Annual Review of Microbiology* **49**, 427-460.

Boyle, J.S. and Lew, A.M. (1995) An inexpensive alternative to glassmilk for DNA purification. *Trends in Genetics* **11**, 8.

Breuer, W.V., Kutner, S., Sylphen, J., Ginsburg, H. and Cabantchik, Z.I. (1987) Covalent modification of the permeability pathways induced in the human erythrocyte membrane by the malarial parasite *Plasmodium falciparum*. *Journal of Cellular Physiology* **133**, 55-63.

Brown, G.V. and Nossal, G.J.V. (1986) Malaria - yesterday, today and tomorow. *Perspectives in Biology and Medicine* **30**, 65-76.

Bryant, C. and Behm, C.A. (1989) *Biochemical Adaptation in Parasites*. London, Chapman and Hall.

Budavari, S., Ed. (1989). *The Merck Index*. Rahway, Merck and Co, Inc.

Cabantchik, Z.I. (1990) Properties of permeation pathways induced in human red blood cell membranes by malaria parasites. *Blood Cells* **16**, 421-432.

Chenchik, A., Diachenko, L., Moqadam, F., Lukyanov, S. and Siebert, P.D. (1996) Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *BioTechniques* **21**, 526-534.

Choi, I. And Mikkelsen, R.B (1990) *Plasmodium falciparum*: ATP/ADP transport across the parasitophorous vacuolar and plasma membranes. *Experimental Parasitology* **71**, 452-462.

Chomczynski, P. (1992) Solubilization in formamide protects RNA from degradation. *Nucleic Acids Research* **20**, 3791-3792.

Chomczynski, P. (1993) TRI-Reagent. RNA, DNA and protein isolation reagent. Manufacturer Protocol. Molecular Research Centre Inc. Cincinnati, OH.

Clark, I.A. and Rockett, K.A. (1994) The cytokine theory of human cerebral malaria. *Parasitology Today* **10**, 410-412.

Clark, K., Nyambati, V.C.S., Neitz, A.W.H. and Louw, A.I. (1998) In Search of an Amino Acid Transporter Gene in *Plasmodium falciparum*. Biochemistry in Africa - 2nd FASBMB; 15th SASBMB, Potchefstroom, South Africa.

CLONTECH Laboratories Inc. *SMART PCR cDNA Synthesis Kit - User Manual*. Palo Alto, CA, USA, CLONTECH Laboratories Inc.

CLONTECH Laboratories Inc. (1991) *RT-PCR. Methods and Applications*. Book 1. Palo Alto, CA, USA, CLONTECH Laboratories, Inc.

Cobb, B.D. and Clarkson, J.M. (1994) A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Research* **22**, 3801-3805.

Collins, F.H. and Besansky, N.J. (1994) Vector biology and control of malaria in Africa. *Science* **264**, 1874-1875.

Collins, F.H. and Paskewitz, S.M. (1995) Malaria: Current and future prospects for control. *Annual Review of Entomology* **40**, 195-219.

Cowman, A.F. (1995) Mechanisms of drug resistance in malaria. *Australian and New Zealand Journal of Medicine* **25**, 837-844.

Cowman, A.F., Morry, M.J., Biggs, B.A., Cross, G.A.M. and Foote, S.J. (1988) Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the USA* **85**, 9109-9113.

Cox, F.E.G. (1993) *Modern Parasitology - A Textbook of Parasitology*. London, Blackwell Scientific Publications.

Creedon, K.A., Rathod, P.K. and Wellems, T.E. (1994) *Plasmodium falciparum S*-Adenosylhomocysteine hydrolase. *The Journal of Biological Chemistry* **269**, 16364-16370.

Curtis, C.F. (1994) The case for malaria control by genetic manipulation of its vectors. *Parasitology Today* **10**, 371-374.

Damak, S. and Bullock, D.W. (1993) A simple two-step method for efficient blunt-end ligation of DNA fragments. *BioTechniques* **15**, 448-452.

Deitsch, K.W. and Wellems, T.E. (1996) Membrane modifications in erythrocytes parasitized by *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **76**, 1-10.

Department of Health of South Africa (1996) *Guidelines for the prophylaxis of malaria*. Pretoria.

Department of Health of South Africa (1996) *Guidelines for the treatment of malaria*. Pretoria.

Desai, S.A., Krogstad, D.J. and McCleskey, E.W. (1993) A nutrient-permeable channel on the intra-erythrocytic malaria parasite. *Nature* **362**, 643-646.

Desai, S.A. and Rosenberg, R.L. (1997) Pore size of the malaria parasite's nutrient channel. *Proceedings of the National Academy of Sciences of the USA* **94**, 2045-2049.

Dietmaier, W. and Fabry, S. (1995). Protocol: Di-/Trinucleotide sticky end cloning. In: *PCR Applications Manual*, Boehringer Mannheim GmbH, Biochemica: 136-140.

Dietmaier, W., Fabry, S. and Schmitt, R. (1993) DISEC-TRISEC: di- and trinucleotide-sticky-end-cloning of PCR amplified DNA. *Nucleic Acids Research* **21**, 3603-3604.

Divo, A.A., Geary, T.G., Davis, N.L. and Jensen, J.B. (1985) Nutritional requirements of *Plasmodium falciparum* in culture. I. Exogenously supplied dialysable components necessary for continuous growth. *Journal of Protozoology* **32**, 59-64.

Dobeli, H., Trzeciak, A., Gillesen, D., Matile, H., Srivastava, I.K., Perrin, L.H., Jakob, P.E. and Certa, C. (1990) Expression, purification, biochemical characterisation and inhibition of recombinant *Plasmodium falciparum* aldolase. *Molecular and Biochemical Parasitology* **41**, 259-268.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* **19**, 4008.

Dontfraid, F.F. and Chakrabarti, D. (1994) Cloning and expression of a cDNA encoding the homologue of Ran/TC4 GTP-binding protein from *Plasmodium falciparum. Biochemical and Biophysical Research Communications* **201**, 423-429.

Doolan, D.L. and Hoffman, S.L. (1997) Multi-gene vaccination against malaria: A multistage, multi-immune response approach. *Parasitology Today* **13**, 171-178.

Dugaiczyk, A., Boyer, H.W. and Goodman, H.M. (1975) Ligation of EcoRI endonuclease generated DNA fragments into linear and circular structures. *Journal of Molecular Biology* **96**, 171-184.

Edwards, J.B.D.M., Ravassard, P., Icard-Liepkalns, C. and Mallet, J. (1995). cDNA cloning by RT-PCR. *PCR 2 - A Practical Approach.* M.J. McPherson, B.D. Hames and G.R. Taylor. Oxford, IRL Press: 89-118.

Elford, B.C., Cowan, G.M. and Ferguson, D.J.P. (1995) Parasite regulated membrane transport processes and metabolic control in malaria-infected erythrocytes. *Biochemical Journal* **308**, 361-374.

Elford, B.C., Cowan, G.M. and Ferguson, D.J.P. (1997) Transport and trafficking in malaria-infected erythrocytes. *Trends in Microbiology* **5**, 463-465.

Estruch, J.J. (1997) OLIGO: Primer selection. *Methods in Molecular Biology* **70**, 279-285.

Evans, D.B. and Jamison, D.T. (1994) Economics and the argument for parasitic disease control. *Science* **264**, 1866-1867.

Francis, S.E., Sullivan, D.J. and Goldberg, D.E. (1997) Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum. Annual Review of Microbiology* **51**, 97-123.

Frohman, M.A. (1993) Rapid amplification of complementary DNA ends for generation of full-length complemetary DNAs: Thermal RACE. *Methods in Enzymology* **218**, 340-356.

Fullick, A. and Fullick, P. (1993) Tropical Diseases. *New Scientist - Inside Science* **59**, 1-4.

Gero, A.M. and Kirk, K. (1994) Nutrient transport pathways in *Plasmodium*-infected erythrocytes: What and where are they? *Parasitology Today* **10**, 395-399.

Gilbert, S.C. and Hill, A.V.S. (1997) Protein particle vaccines against malaria. *Parasitology Today* **13**, 302-306.

Gilbert, W. (1991) Towards a paradigm shift in biology. *Nature* **349**, 99.

Ginsburg, H. (1990a) Some reflections concerning host erythrocyte-malarial parasite interrelationships. *Blood Cells* **16**, 225-235.

Ginsburg, H. (1990b) Alterations caused by the intraerythrocytic malaria parasite in the permeability of its host cell membrane. *Comparative Biochemistry and Physiology* **95A**, 31-39.

Ginsburg, H. (1994) Transport pathways in the malaria infected erythrocyte: Their characterisation and their use as potential targets for chemotherapy. *Biochemical Pharmacology* **48**, 1847-1856.

Ginsburg, H. and Stein, W.D. (1987) New permeability pathways induced by the malarial parasite in the membrane of the host erythrocyte: potential routes for the targeting of drugs into infected cells. *Bioscience Reports* **7**, 455-463.

Goodyer, I.D., Pouvelle, B., Schneider, T.G., Trelka, D.P. and Taraschi, T.F. (1997) Characterisation of macromolecular transport pathways in malaria-infected erythrocytes. *Molecular and Biochemical Parasitology* **87**, 13-28.

Goswami, A., Chatterjee, S. and Sharma, S. (1996) Cloning of a ribosomal phosphoprotein P0 gene homologue from *Plasmodium falciparum. Molecular and Biochemical Parasitology* **82**, 117-120.

Grau, G.E. and Kossodo, S.D. (1994) Cerebral malaria: Mediators, mechanical obstruction or more. *Parasitology Today* **10**, 408-409.

Greenwood, B.M. (1997) Malaria transmission and vector control. *Parasitology Today* **13**, 90-92.

Greenwood, B.M. (1997) What's new in malaria control? *Annals of Tropical Medicine and Parasitology* **91**, 523-531.

Gregory, R.G., Cheng, S.H., Rich, D.P., Marshall, J., Paul, S., Hehir, K., Ostegaard, L., Klinger, K.W., Welsh, M.J. and Smith, A.E. (1990) Expression and characterization of the cystic fibrosis transmembrane conductance regulator. *Nature* **347**, 382-386.

Guidotti, G.G. and Gazzola, G.C., Eds. (1992). Amino acid transporters: Systematic approach and principles of control. *Mammalian Amino Acid Transport.* New York, Plenum Press: 3-29.

Hagan, P. and Chauhan, V. (1997) Ronald Ross and the problem of malaria. *Parasitology Today* **13**, 290-295.

Hanahan, D., Jessee, J. and Bloom, F.R. (1991) Plasmid transformation of *Escherichia coli* and other bacteria. *Methods in Enzymology* **204**, 63-113.

Harvey, C.M. and Ellory, J.C. (1989) Identification of amino acid transporters in the red blood cell. *Methods in Enzymology* **173**, 122-161.

Hecker, K.H. and Roux, K.H. (1996) High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. *BioTechniques* **20**, 478-485.

Hengen, P.N. (1994) Long and accurate PCR. *Trends in Biochemical Sciences* **19**, 341.

Henikoff, J.G. and Henikoff, S. (1996) Blocks database and its applications. *Methods in Enzymology* **266**, 88-105.

Hibbs, A.R., Stenzel, D.J. and Saul, A. (1997) Macromolecular transport in malaria - does the duct exist? *European Journal of Cell Biology* **72**, 182-188.

Higgins, D.G., Thompson, J.D. and Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods in Enzymology* **266**, 383-402.

Hitti, Y.S. and Bertino, A.M. (1994) Proteinase K and T4 DNA polymerase facilitate the blunt-end sub-cloning of PCR products. *BioTechniques* **16**, 802-805.

Hyde, J.E. and Holloway, S.P. (1993) Isolation of parasite genes using synthetic oligonucleotides. *Methods in Molecular Biology* **21**, 303-318.

Hyde, J.E., Kelly, S.L., Holloway, S.P., Snewin, V.A. and Sims, P.F.G. (1989) A general approach to isolating *Plasmodium falciparum* genes using non-redundant oligonucleotides inferred from protein sequences from other organisms. *Molecular and Biochemical Parasitology* **32**, 247-262.

Kanaani, J. and Ginsburg, H. (1989) Metabolic interconnection between the human malarial parasite *Plasmodium falciparum* and its host erythrocyte. *The Journal of Biological Chemistry* **264**, 3194-3199.

Kanungo, J. and Pandey, K.N. (1993) Kinasing PCR products for efficient blunt-end cloning and linker addition. *BioTechniques* **14**, 912-913.

Kelce, W.R., Stone, C.R., Laws, S.C., Gray, L.E., Kemppainen, J.A. and Wilson, E.M. (1995) Persistent DDT metabolite p,p'-DDE is a potent androgen receptor antagonist. *Nature* **375**, 581-585.

Kemp, D.J., Koppel, R.L. and Anders, R.F. (1987) Repetitive proteins and genes of malaria. *Annual Review of Microbiology* **41**, 181-208.

Kidd, K.K. and Ruano, G. (1995). Optimizing PCR. In: *PCR2- A Practical Approach.* M.J. McPherson, B.D. Hames and G.R. Taylor. Oxford, IRL Press: 1-22.

Kirby, G.C. (1997) Plants as a source of antimalarial drugs. *Tropical doctor* **Supplement 1**, 7-11.

Kolakovich, K.A., Gluzman, I.Y., Duffin, K.L. and Goldberg, D.E. (1997) Generation of hemoglobin peptides in the acidic digestive vacuole of *Plasmodium falciparum* implicates peptides transport in amino acid production. *Molecular and Biochemical Parasitology* **87**, 123-135.

Kondrachine, A.V. and Trigg, P.I. (1997) Global overview of malaria. *Indian Journal of Medical Research* **106**, 39-52.

Krauss, R.D., Bubien, J.K., Peiper, S.C., Collins, F.S., Kirk, K.L., Frizzell, R.A. and Rado, T.A. (1992) Transfection of the wild-type CFTR into cystic fibrosis lymphocytes restores chloride conductance at G1 of the cell cycle. *EMBO Journal* **11**, 875-883.

Krishna, S. (1997) Malaria. *British Medical Journal* **315**, 730-732.

Krogstad, D.J. (1996) Malaria as a re-emerging disease. *Epidemiologic Reviews* **18**, 77-89.

Kwiatkowski, D. and Marsh, K. (1997) Development of a malaria vaccine. *The Lancet* **350**, 1696-1701.

Kwok, S., Kellog, D.E., Spasic, D., Goda, L., Levenson, C. and Sninsky, J.J. (1990) Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. *Nucleic Acids Research* **18**, 999-1005.

La Fontaine, S., Firth, S.D., Lockhart, P.J., Paynter, J.A. and Mercer, J.F.B. (1998) Eukaryotic expression vectors that replicate to low copy number in bacteria: Transient expression of the Menkes protein. *Plasmid* **39**, 245-251.

Language, K., Nyambati, V.C.S. and Louw, A.I. (1997a) In search of therapeutic targets in South African strains of the malaria parasite *Plasmodium falciparum*. South African Society of Biochemistry and Molecular Biology. Fourteenth Conference, Grahamstown, South Africa.

Language, K., Nyambati, V.C.S. and Louw, A.I. (1997b) In search of therapeutic targets in *Plasmodium falciparum*. 17th International Congress of Biochemistry and Molecular Biology in conjunction with 1997 Annual Meeting of American Society of Biochemistry and Molecular Biology, San Francisco, USA.

Lauer, S.A., Rathod, P.K., Ghori, N. and Haldar, K. (1997) A membrane network for nutrient import in red cells infected with the malaria parasite. *Science* **276**, 1122-1125.

le Sueur, D., Sharp, B.L., Gouws, E., Ngxongo, S. (1996) Malaria in South Africa. *South African Journal of Medicine* **86**, 936-939.

Louw, A.I. (1998) Molecular characterisation of faciliated transport proteins of the human malaria parasite. *South African Journal of Science* **94**, 275-277.

Lund, A.H., Duch, M. and Pedersen, F.S. (1996) Increased cloning efficiency by cycle temperature ligation. *Nucleic Acids Research* **24**, 800-801.

MacLean, M., Anderson, J. and Davies, C. (1997) Making malaria research bite. *Nature Medicine* **3**, 14-16.

McGivan, J.D. and Pastor-Anglada, M. (1994) Regulatory and molecular aspects of mammalian amino acid transport. *Biochemical Journal* **299**, 321-334.

Mendis, K.N. and Carter, R. (1995) Clinical disease and pathogensis in malaria. *Parasitology Today* **11**, PTI 1-PTI 16.

Michelson, B.K. (1995) Transformation of *Escherichia coli* increases 260-fold upon inactivation of T4 DNA ligase. *Analytical Biochemistry* **225**, 172-174.

Miller, L.H., Good, M.F. and Milon, G. (1994) Malaria pathogenesis. *Science* **264**, 1878-1883.

Mons, B., Trape, J.-F. and Hagan, P. (1997) Malaria, malaria and more malaria. *Parasitology Today* **13**, 279-284.

Mullis, K. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **1986**, 263-273.

Nussenzweig, R.S. and Long, C.A. (1994) Malaria vaccines: Multiple targets. *Science* **265**, 1381-1383.

Nyambati, V.C.S., Language, K. and Louw, A.I. (1997) Strategies for chemotherapy and vaccine development in *Plasmodium falciparum* infections. South African Society of Biochemistry and Molecular Biology. Fourteenth Conference, Grahamstown, South Africa.

Oliviera, D.A., Holloway, B.P., Durigon, E.L., Collins, W.E. and Lai, A.A. (1995) Polymerase Chain Reaction and a liquid-phase, non-isotopic hybridisation for species-specific and sensitive detection of malaria. *American Journal of Tropical Medicine and Hygiene* **52**, 139-144.

Olliaro, P. and Yuthavong, Y. (1998) Chemotherapeutic targets in Plasmodia with potential for antimalarial drug discovery. *South African Journal of Science* **94**, 292-296.

Olliaro, P.L. and Goldberg, D.E. (1995) The *Plasmodium* digestive vacuole: Metabolic headquarters and choice drug target. *Parasitology Today* **11**, 294-297.

Peale, F.V., Mason, K., Hunter, A.W. and Bothwell, M. (1998) Multiplex display polymerase chain reaction amplifies and resolves related sequences sharing a single moderately conserved domain. *Analytical Biochemistry* **256**, 158-168.

Peden, K.W. (1992) Instability of HIV sequences in high copy number plasmids. *Journal of Acquired Immune Deficiency Syndromes* **5**, 313-315.

Penny, J.I., Hall, S.T., Woodrow, C.J., Cowan, G.M., Gero, A.M. and Krishna, S. (1998) Expression of substrate-specific transporters encoded by *Plasmodium falciparum* in *Xenopus laevis* oocytes. *Molecular and Biochemical Parasitology* **93**, 81-89.

Petersen, D.S., Walliker, D. and Wellems, T.E. (1988) Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria. *Proceedings of the National Academy of Sciences of the USA* **85**, 9114-9118.

Prasad, R. (1997) Changing the transport of a cell. *Critical Reviews in Biotechnology* **17**, 171-183.

Preston, G.M. (1993) Use of degenerate oligonucleotide primers and the polymerase chain reaction to clone gene family members. *Methods in Molecular Biology* **15**, 317-337.

Promega (1996) *PCR Access. PCR Protocols and Reference Guide*. Madison, WI, USA, Promega Corporation.

Rand, K.N. (1996) Crystal violet can be used to visualise DNA bands during gel electrophoresis and to improve cloning efficiency. *Elsevier Trends Journals Technical Tips Online*.

Read, M., Hicks, K.E., Sims, P.F.G. and Hyde, J.E. (1994) Molecular characterisation of the enolase gene from the human malaria parasite *Plasmodium falciparum*. Evidence for ancestry within a photosynthetic lineage. *European Journal of Biochemistry* **220**, 513-520.

Roche, E.B., Kier, L.B. and Foye, W.O. (1989). Parasite Chemotherapy. In: *Principle of Medicinal Chemistry*. W. O. Foye. Philadelphia, Lea & Febiger: 717-730.

Rost, W.J. (1989). Chemotherapy: An introduction. In: *Principles of Medicinal Chemistry*. W. O. Foye. Philadelphia, Lea & Febiger: 629-635.

Roth, E. (1990) *Plasmodium falciparum* carbohydrate metabolism: A connection between host cell and parasite. *Blood Cells* **16**, 453-460.

Rychlik, W. (1993) Selection of primers for polymerase chain reaction. *Methods in Molecular Biology* **15**, 31-40.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning - A Laboratory Manual*. Cold Spring Harbor, Cold Spring Harbor Laboratory Press.

Sander, D.M. (1996) Molecular biology databases on the Internet. *BioTechniques* **21**, 438-440.

Saul, A. and Battistutta, D. (1988) Codon usage in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **27**, 35-42.

Schaefer, B.C. (1995) Revolutions in Rapid Amplification of cDNA Ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. *Analytical Biochemistry* **227**, 255-273.

Schneider, J., Gilbert, S.C., Blanchard, T.J., Hanke, T., Robson, K.J., Hannan, C.M., Becker, M., Sinden, R., Smith, G.L. and Hill, A.V.S. (1998) Enhanced immunogenicity for CD8+ T cell induction and complete protective efficacy of malaria DNA vaccination by boosting with modified vaccinia virus Ankara. *Nature Medicine* **4**, 397-402.

Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research* **18**, 6097-6100.

Schuler, G.D., Epstein, J.A., Ohkawa, H. and Kans, J.A. (1996) *Entrez*: Molecular biology database and retrieval system. *Methods in Enzymology* **266**, 141-162.

Sherman, I.W. (1979) Biochemistry of *Plasmodium* (Malaria parasites). *Microbiological Reviews* **43**, 453-495.

Sherman, I.W. (1988) Mechanisms of molecular trafficking in malaria. *Parasitology* **96**, S57-S81.

Shomer, B., Harper, R.A.L. and Cameron, G.N. (1996) Information services of the European Bioinformatics Institute. *Methods in Enzymology* **266**, 3-27.

Sirawaraporn, W., Sathitkul, T., Sirawaraporn, R., Yuthavong, Y. and Santi, D.V. (1997) Antifolate-resistant mutants of *Plasmodium falciparum* dihydrofolate reductase. *Proceedings of the National Academy of Sciences of the USA* **94**, 1124-1129.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I. and Miller, L.H. (1995) Switches in expression of *Plasmodium*

falciparum var genes correlates with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101-110.

Smith, S.W., Overbeek, R., Woese, C.R., Gilbert, W. and Gillevet, P.M. (1994) The genetic data environment an expandable GUI for multiple sequence analysis. *Computer Applications in the Biosciences* **10**, 671-675.

Sophianopoulou, V. and Diallinas, G. (1995) Amino acid transporters of lower eukaryotes: regulation, structure and topogenesis. *FEMS Microbiology Reviews* **16**, 53-75.

Souba, W.W. and Pacitti, A.J. (1992) How amino acids get into cells: Mechanisms, models, menus and mediators. *Journal of Parenteral and Enteral Nutrition* **16**, 569-578.

Spielman, A. (1994) Why entomological antimalaria research should not focus on transgenic mosquitoes. *Parasitology Today* **10**, 374-376.

Su, T.Z., Lunney, E., Campbell, G. and Oxender, D.L. (1995) Transport of gabapentin, a gamma-amino acid drug, by system L alpha-amino acid transporters: a comparative study in astrocytes, synaptosomes and CHO cells. *Journal of Neurochemistry* **64**, 2125-2131.

Su, X.Z., Heatwole, V., Guinet, S.P., Herrfeldt, J.A., Peterson, D.S., Ravetch, J.A. and Wellems, T.E. (1995) The large and diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89-100.

Su, X.Z., Kirkman, L.A., Fujioka, H. and Wellems, T.E. (1997) Complex polymorphisms in an ~330kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell* **91**, 593-603.

Su, X.Z., Wu, Y., Sifri, C.D. and Wellems, T.E. (1996) Reduced extension temperatures required for PCR amplification of extremely A+T-rich DNA. *Nucleic Acids Research* **24**, 1574-1575.

Subbayya, I.N.S., Ray, S.S., Balaram, P. and Balaram, H. (1997) Metabolic enzymes as potential drug targets in *Plasmodium falciparum*. *Indian Journal of Medical Research* **106**, 79-94.

Taverne, J. (1993). Immunity to protozoa and worms. In: *Immunology*. I. M. Roitt, J. Brostoff and D. K. Male. St Louis, Mosby.

Thomas, M.G., Hesse, S.A., McKie, A.T. and Farzaneh, F. (1993) Sequencing of cDNA using anchored oligo-dT primers. *Nucleic Acids Research* **21**, 3915-3916.

Trager, W. (1994) Cultivation of malaria parasites. *Methods in Cell Biology* **45**, 7-26.

Trager, W. and Jensen, J.B. (1976) Human malaria parasites in continuous culture. *Science* **193**, 673-675.

Trager, W. and Jensen, J.B. (1997) Continuous culture of *Plasmodium falciparum*: its impact on malaria research. *International Journal for Parasitology* **27**, 989-1006.

Triglia, T., Wellems, T.E. and Kemp, D.J. (1992) Towards a high-resolution map of the *Plasmodium falciparum* genome. *Parasitology Today* **8**, 225-229.

Vander Jagt, D.L., Hunsaker, L.A., Campos, N.M. and Baack, B.R. (1990) D-Lactate production in erythrocytes infected with *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **42**, 277-284.

Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids: a structure for deoxyribonucleic acid. *Nature* **4356**.

Zubay, G. (1993) *Biochemistry*. Dubuque, Wm. C. Brown Publishers.

APPENDIX A

IUBMB Degenerate Nucleotide Nomenclature

Table 1. The degenerate nucleotide alphabet.

IUBMB Codes	
A : Adenosine	**R** : A or G (puRine)
C : Cytosine	**Y** : C or T (pYrimidine)
G : Guanosine	**K** : G or T (Keto)
T : Thymidine	**M** : A or C (aMino)
B : C, G or T	**S** : G or C (Strong - 3 H-bonds)
D : A, G or T	**W** : A or T (Weak - 2 H-bonds)
H : A, C or T	**N** : aNy base
V : A, C or G	

APPENDIX B

Sequence Alignments of Clones

S-adenosyl homocysteine hydrolase

```
PfSAHH          TTTTTTTTATCATATAATACTTATTGTATATATTTATTCTTTTTTCAAGATATAATTCTT
PfSAHH          ACATTTTACAAGTCACTTCAAAATTGTATATATATATATATATATATATATATATATA
PfSAHH          TTATAACTCCATGAAGAGTATATTGAATTTGTCAATTTTTTGTGAACATATAAAAAATAA
PfSAHH          AAGCCAATAATATAATATTTTACTTTATTTACTTATTTTATTTTTATATATATATATATA
PfSAHH          TATATATACATTATTTTATATATATTTTTTTCTTCCTTTATATTAAAAGGGTAACAAGTA
PfSAHH          TCATACCCATCGGCATATAATTTTACTTCTTTTTATATATTTTTTTTTTGAGAAATGGTT
PfSAHH          GAAAATAAGAGCAAGGTCAAAGATATCAGTTTGGCCCCCTTTGGAAAAATGCAGATGGAA
PfSAHH          ATTTCTGAAAATGAAATGCCGGGATTAATGAGAATAAGAGAAGAATACGGAAAAGATCAA
PfSAHH          CCATTAAAAAATGCTAAAATTACTGGTTGTTTACATATGACTGTTGAATGTGCTTTATTA
PfSAHH          ATTGAGACTTTACAAAAATTAGGAGCTCAGATTAGGTGGTGTTCATGTAATATTTATTCA
PfSAHH          ACAGCTGATTATGCTGCAGCAGCAGTAAGTACATTAGAAAATGTAACGGTTTTTGCTTGG
PfSAHH          AAAAATGAAACTTTAGAAGAATACTGGTGGTGTGTTGAAAGTGCCCTTACGTGGGGTGAT
PfSAHH          GGAGATGATAATGGTCCAGATATGATTGTAGATGATGGGGGTGATGCAACCTTATTAGTT
PfSAHH          CATAAAGGTGTTGAGTATGAAAAATTATATGAAGAGAAAAATATATTACCTGATCCAGAA
PfSAHH          AAAGCAAAAAATGAAGAAGAACGATGTTTTCTAACTTTATTAAAAAATTCCATATTAAAA
PfSAHH          AATCCAAAGAAATGGACAAATATTGCAAAGAAAATTATAGGTGTATCTGAAGAAACTACT
PfSAHH          ACAGGAGTATTAAGATTAAAAAAAATGGACAAACAAAATGAATTATTATTTACTGCTATT
PfSAHH          AATGTAAATGATGCTGTCACTAAACAAAAATATGATAATGTGTATGGATGTAGACATTCC
PfSAHH          TATACTGACGGATTAATGAGAGCTACCGATTTTTTAATATCTGGAAAAATCGTTGTCATA
PfSAHH          TGTGGATATGGTGACGTAGGTAAAGGATGTGCTTCTTCTATGAAAGGTTTAGGTGCAAGA
PfSAHH          GTATATATAACAGAAATTGATCCCATATGTGCTATACAAGCTGTAATGGAAGGATTTAAT
```

Primer PSP2

```
PfSAHH          GTTGTTACTTTAGATGAAATTGTAGATAAAGGAGATTTTTTTATTACTTGTACAGGTAAT
Clone 66        ----------------------------------------TTAGTAGGTGCTGGTAAT
Clone 9         ----------------------------------------TTAGTAGGAGCTGGTTAT
Clone 36        ------------------------------------------------GGGC-TGGGATAT
                                                          *** **   *    *   **
```

```
PfSAHH          GTTGATGTTATTAAATTAGAACACTTACTTAAAATGAAAAATAATGCTGTTGTTGGAAAC
Clone 66        ATTGATGTTATTAAATTAGAACACTTACTTAAAATGAAAAATAATGCTGTTGTTGGAAAC
Clone 9         GTTGATGTTATTAAATTAGAACACTTACTTAAAATGAAAAATAATGCTGTTGTTGGAAAC
Clone 36        GTTGATGTTATTAAATTAGAACACTTACTTAAAATGAAAAATAATGCTGTTGTTGGAAAC
                 ************************************************************

PfSAHH          ATTGGTCATTTTGATGATGAAATACAAGTAAATGAACTTTTTAATTATAAAGGAATACAT
Clones          ATTGGTCATTTTGATGATGAAATACAAGTAAATGAACTTTTTAATTATAAAGGAATACAT
                ************************************************************

PfSAHH          ATAGAAAATGTAAAACCACAGGTTGATAGAATTACTTTACCTAATGGAAATAAAATTATT
Clones          ATAGAAAATGTAAAACCAACGGTTGATAGAATTACTTTACCTAATGGAAATAAAATTATT
                ************************************************************

PfSAHH          GTTTTAGCTAGAGGGAGATTATTAAATCTAGGATGTGCAACAGGACATCCAGCATTTGTT
Clones          GTTTTAGCTAGAGGGAGATTATTAAATCTAGGATGTCGAACAGGACATCCACGATTTGTT
                ************************************************************

PfSAHH          ATGTCCTTTTCATTTTGTAATCAAACCTTTGCTCAATTAGATTTATGGCAAAACAAGGAT
Clones          ATGTCCTTTTCATTTTGTAATCAAACCTTTGCTCAATTAGATTTATGGCAAAACAAGGAT
                ************************************************************

PfSAHH          ACAAACAAATATGAAAATAAAGTTTATTTGTTACCTAAACATCTTGATGAAAAGGTTGCT
Clones          ACAAACA-ATATGAAAATAAAGTTTATTTGTTACCTAAACATCTTGATGAAAAGGTTGCT
                ******* ************************************************

PfSAHH          CTTTATCATTTGAAAAAATTGAACGCTTCCTTGACAGAATTGGATGACAATCAATGTCAA
Clones          CTTTATCATTTGAAAAAATTGAACGCTTCCTTGACAGAATTGGATGACAATCAATGTCAA
                ************************************************************

PfSAHH          TTTTTGGGAGTCAACAAAAGTGGTCCCTTTAAGAGTAACGAATACAGATATTAAAAAAAT
Clones          TTTTTGGGAGTCAACAAAAGTGGTCCCTTTAAGAGTAACGAATACAGATATTAAAAAAAT
                ************************************************************
```

Poly-T and anchor

```
PfSAHH          AAAAGGA-------------------------------------------
Clone 9         AAAAGGAAAAAAAAAAAAAAAAAAAGAGTGTTGTGGTAATGAGACG--------
Clone 66        AAAAGGAAAAAAAAAAAAAAAAAAAGAGTGTTGTGGTAATGATAG---------
                *******
```

Figure 1. Alignment of the S-adenosylhomocysteine hydrolase gene of *Plasmodium falciparum* with the clones obtained using primer PSP2.

Nuclear GTP-Binding Protein Homologue

```
PfGTP           ------------------AAAAAAAAAAAAAATAAACAATAATTTTATAGTTGTATTATT

PfGTP           ATATATAATATAAACACCTAAATAATATATTATAAAATACAATTTAAAAAATGGATTCAC

                                                    Primer PSP2
PfGTP           AAGAATATATTCCACAATATAAATTAATCTTAGTCGGTGATGGTGGTGTTGGCAAAACAA
Clone 22        ----------------------------TTAGTTGGAGCTGGTGATATTGGCAAAACAA
Clone 39        ----------------------------TTAGTAGGAGCTGGTTATGCTTGGCAAAACAA
Clone 17        --------------CTTGCAGCTTGGAGCCTGGTGATATCTCGCCGCAACAACAACCTTC
Clone 15        ----------------------------TTAGTTGGTGCAGGAGATGTTGGCAAAACAA
Clone 23        ----------------------------TTAATTCGGTGCAGGTGATCCGTTGGCAAACC

PfGTP           CCTTTGTGAAAAGACACTTGACTGGAGAATTTGAAAAAAAATATATACCTACTCTTGGT
Clones          CCTTTGTGAAAAGACACTTGACTGGAGAATTTGAAAAAAAATATATACCTACTCTTGGT
                ************************************************************

PfGTP           GTGGAAGTTCACCCCTTAAAATTTCAAACAAACTTTGGAAAAACTCAATTTAACGTATG
Clones          GTGGAAGTTCACCCCTTAAAATTTCAAACAAACTTTGGAAAAACTCAATTTAACGTATG
                ************************************************************

PfGTP           GGATACTGCAGGACAAGAAAAGTTTGGTGGTTTAAGAGATGGATATTATATAAAAAGTG
Clones          GGATACTGCAGGACAAGAAAAGTTTGGTGGTTTAAGAGATGGATATTATATAAAAAGTG
                ************************************************************

PfGTP           ATTGTGCTATAATTATGTTTGATGTATCTTCTCGTATTACTTACAAGAACGTTCCAAAT
Clones          ATTGTGCTATAATTATGTTTGATGTATCT-CTCGTATTACTTACAAGAACGT-CCAAAT
                **************************** ********************** ******

PfGTP           TGGTATAGAGATATTACAAGAGTGTGTGAAACAATTCCTATGGTTTTAGTTGGAAACAA
Clones          TGGTATAGAGATATTACAAGAGTGTGTGAA~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                ****************************

                                ~~~ No sequence data available
PfGTP           AGTTGATGTTAAAGACAGACAAGTAAAATCAAGGCAAATTCAATTTCACAGAAAAAGGAA
Clones          ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~GGAA
                                                                        ****

PfGTP           TTTACAATACTACGATCTATCTGCAAGATCAAATTACAATTTCGAAAAACCTTTCTTATG
Clones          TTTACAATACTACGATCTATCTGCA-GATCAAATTACAATTTCCAAA--CCTTTCTTATG
                ************************* ********************  ***********

PfGTP           GTTAGCTAGGAGATTGTCCAAC-CAACCAAATCTTGTTTTCGTAGGAGAACATGCTAAAG
Clones          GTTAGCTAGGAGATTGTCCAACGCAACCAAATCTTGTTTTCGTAGGAGAACATGCTAAAG
                ********************* ************************************

PfGTP           CACCAGAGTTCCAAATTGATCTAAATATTGTAAGAGAAGCTGAAAAAGAATTAGAGCAAG
Clones          CACCAGAGTTCCAAATTGATCTAAATATTGTAAGAGAAGCTGAAAAAGAATTAGAGCAAG
                ************************************************************

PfGTP           CAGCAGCTGTAGCTATTGATGAAGAAGATATTGAAAATTAAATAAATAGAAAAAAAAAAA
Clones          CAGCAGCTGTAGCTATTGATGAAGAAGATATTGAAAATTAAATAAATAG***AAAAAAAAAA***
                ************************************************************

                    Poly-T and anchor
PfGTP           AAAAAA----------------------
Clones          ***AAAAAGAGAGTGTTGTGGTAATGATAGC***
                *****
```

Figure 2. Alignment of the GTP binding protein gene of *Plasmodium falciparum* with the clones obtained using primer PSP2.

Heat Shock Protein 90

```
PfHSP          ------AAAATAATAAAATTAAAATAAAATAAGATAATTACCAAAAATAATAAAAAAAAT

PfHSP          TAAAAAGTATTAAAAAATATATATCATATATAAATATATAAATTTTAACAACGTGAGAAA

PfHSP          TTAAGAAAAAAATTATATTTATATAAATAATATAATATTTAATATATATCTCATATAATA

PfHSP          TTATATATATATATATATATATATATATTTTTTAAAAAAAAAGAAAAAATAAGTTTAT

PfHSP          TTTTACCTTTTAATAAAAACATATTTTTAGATATACATAAAAATTTTTATATTATTTTAT

PfHSP          TGGTATATATATATATATATATATATAATAATATATATATAATAAAATATATAAATAT

PfHSP          TTACATAAATATTTTACTTTAAAAAGAAATTTTCCTTTTTTATTATATTTTTTTTCTTCC

PfHSP          CACAATTTCGAATAAAATGTCAACGGAAACATTCGCATTTAACGCCGACATCAGGCAGTT

PfHSP          GATGAGTTTGATTATCAACACTTTTTACAGTAACAAAGAAATATTTTTAAGAGAATTGAT

PfHSP          TAGTAATGCTAGTGATGCCTTAGATAAAATAAGATATGAATCAATTACAGATACTCAAAA

PfHSP          ATTATCTGCTGAGCCTGAATTTTTTATTCGTATCATTCCTGACAAAACCAACAATACATT

PfHSP          AACTATTGAAGATTCAGGTATTGGTATGACAAAAAATGATTTAATTAATAACCTTGGTAC

PfHSP          TATTGCAAGATCAGGAACCAAAGCTTTTATGGAAGCCATACAAGCCAGTGGAGATATATC

                                                                  Primer PSP2
PfHSP          TATGATTGGTCAATTTGGTGTTGGTTTTTATTCAGCCTATTTAGTTGCTGATCATGTTGT
Clone 38       ------------------------------------TTAGTTGGAGCTGGTCATGTTGT
Clone 10       ------------------------------------TTAGT-GGAGCTGGTAATGTTGT
                                                    *   * *  **** * *******

PfHSP          TGTTATCTCCAAAAATAATGATGATGAACAATATGTTTGGGAATCTGCTGCAGGAGGTTC
Clones         TGTTATCTCCAAAAATAATGATGATGAACAATATGTTTGGGAATCTGCTGCAGGAGGTTC
               ************************************************************

PfHSP          CTTCACAGTTACTAAGGATGAAACCAATGAAAAACTTGGAAGAGGTACGAAAATTATTCT
Clones         CTTCACAGTTACTAAGGATGAAACCAATGAAAAACTTGGAAGAGGTACGAAAATTATTCT
               ************************************************************

PfHSP          TCATTTAAAAGAAGATCAATTAGAATATCTTGAAGAAAAACGTATCAAAGATTTAGTTAA
Clones         TCATTTAAAAGAAGATCAATTAGAATATCTTGAAGAAAAACGTATCAAAGATTTAGTTAA
               ************************************************************

PfHSP          GAAACACTCTGAATTTATCTCTTTCCCAATCAAGTTATACTGTGAAAGGCAAAATGAAAA
Clones         GAAACACTCTGAATTTATCTCTTTCCCAATCAAGTTATACTGTGAAAGGCAAAATGAAAA
               ************************************************************

PfHSP          AGAAATCTCCGCATCTGAAGAAGAAGAAGGAGAAGGAGAAGGAGAAAGAGAAGGAGAAGA
Clones         AGAAATCACCGCATCTGAAGAAGAAGAAGGAGAAGGAGAAGGAGAAAGAGAAGGAGAAGA
               ******* ****************************************************

PfHSP          AGAAGAAGAAAAAAAAAAAAAAACAGGCGAAGATAAAAATGCTGATGAAAGTAAAGAAGA
Clones         AGAAGAAGAAAAAAAAAAAAAAACAGGCGAAGATAAAAATGCTGATGAAAGTAAAGAAGA
               ************************************************************

PfHSP          AAATGAAGATGAAGAAAAAAAAGAAGATAACGAAGAAGATGATAACAAAACTGATCATCC
Clones         AAATGAAGATGAAGAAAAAAA-GAAGATAACGAAGAAGATGATAACAAAACTGATCATCC
               ******************** ***************************************

PfHSP          AAAAGTTGAAGATGTTACCGAAGAATTAGAAAATGCTGAAAAGAAAAAAAAAGAAAAAAG
Clones         AAAAGTTGAAGATGTTACCGAAGAATTAGAAAATGCTGAAAAGAAAAAAAAAAAAAAAAG
               ************************************************************
```

Poly-T and anchor

```
PfHSP     AAAAAAAAAAATACACACAGTTGAACATGAATGGGAAGAATTAAATAAACAAAAACCATT
Clones    GAGTGTTGTGGTAATGATAGC
                  **   * **

PfHSP     ATGGATGAGAAAACCAGAAGAAGTTACAAATGAAGAATATGCAAGCTTCTATAAATCATT

PfHSP     AACAAATGATTGGGAAGACCATTTAGCTGTTAAACATTTCTCTTGTTGAAGGACAATTAG

PfHSP     AATTTAAAGCCTTATTATTTATACCAAAAAGAGCACCTTTTGATATGTTCGAAAATAGAA

PfHSP     AAAAAAGAAATAATATCAAATTATATGTAAGAAGAGTTTTTATTATGGATGATTGTGAAG

PfHSP     AAATTATTCCAGAATGGTTAAATTTTGTTAAGGGTGTTGTCGATTCAGAAGATTTACCAC

PfHSP     TTAATATTTCAAGAGAATCATTACAACAAAATAAAATACTTAAGGTTATCAAAAAAAACC

PfHSP     TTATCAAAAAATGTTTAGACATGTTCTCAGAATTAGCTGAAAATAAGGAAAACTACAAAA

PfHSP     AGTTTTATGAACAATTCAGCAAAAACTTAAAGTTGGGTATCCACGAGGATAACGCAAATC

PfHSP     GTACAAAGATCACCGAATTACTCCGATTCCAAACCTCAAAATCAGGAGACGAAATGATCG

PfHSP     GATTAAAAGAATACGTAGACAGAATGAAGGAAAACCAAAAGGATATTTACTATATCACCG

PfHSP     GTGAATCCATCAATGCTGTTTCTAATTCTCCATTCTTAGAAGCTTTGACCAAAAAAGGAT

PfHSP     TCGAAGTTATTTATATGGTTGATCCTATTGATGAATATGCAGTACAACAATTAAAAGATT

PfHSP     TTGATGGTAAGAAATTGAAATGTTGTACCAAAGAAGGTTTAGATATTGATGATTCAGAAG

PfHSP     AAGCCAAAAAAGATTTCGAAACCTTGAAAGCTGAATATGAAGGATTATGCAAAGTTATTA

PfHSP     AAGACGTATTACACGAGAAAGTTGAAAAAGTTGTTGTAGGACAAAGAATTACAGATTCTC

PfHSP     CATGTGTATTAGTCACATCAGAATTTGGATGGTCCGCAAACATGGAAAGAATTACGAAAG

PfHSP     CTCAAGCATTAAGAGATAATTCCATGACTAGCTATATGTTATCCAAAAAAATTATGGAAA

PfHSP     TCAATGCTCGTCACCCAATTATATCAGCATTAAAACAAAAAGCTGATGCAGATAAATCAG

PfHSP     ATAAAACCGTTAAAGATTTAATCTGGTTATTATTTGATACCTCTTTATTAACATCTGGTT

PfHSP     TTGCTCTTGAAGAACCAACTACCTTTTCTAAAAGAATCCACAGAATGATTAAATTAGGTT

PfHSP     TATCAATAGATGAAGAAGAAAACAATGATATCGATTTACCACCTCTTGAAGAAACTGTAG

PfHSP     ATGCAACCGATTCTAAAATGGAAGAAGTTGACTAAAGGATTTATATAATATATTTATGTA

PfHSP     CTCACAATGGGGTCTACAAAAAAAAAAAAATAAAGATATATATATATATATATATATATA

PfHSP     TATATATATATATATATATTATCTTTCATTTTATCATCACCCCAATATACATATATATAT

PfHSP     TTTTATATTTTTATATATTTATATTTTTCTTTTATTTATCCATAAAAGAAAATTAAATAT

PfHSP     AAATTTATACCTTTTATTTGAATTACCATTTGTATAAAATATAAAAAAAAA
```

Figure 3. Alignment of the heat shock protein gene of *Plasmodium falciparum* with the clones obtained using primer PSP2.

Ribosomal Phosphoprotein

```
Pfphospho       TATATATAATATACATATATATAATATACATATATGTTTATTCATTTATATATTAATTGG
Clone 4         ------------------------------------------------------------

Pfphospho       ACAATTGCACAGTTGTATATTTTCATTTGATTATTTTTTTTTATTTTTTTTTTCCTGTAG
Clone 4         ------------------------------------------------------------

Pfphospho       TTTATATATATAAAAAATGGCTATGAAATACGTTGCTGCATATCTTATGTGTGTATTGGG
Clone 4         ------------------------------------------------------------

                                                              Primer PSP2
Pfphospho       AGGAAATGAAAACCCAAGCACAAAAGAAGTTAAGAATGTGTTAGGAGCCGTAAATGCTGA
Clone 4         -----------------------------------TTAGTAGGTGCTGGAAATATTGA
                                                    *  **** ** * ****  ***

Pfphospho       TGTAGAAGATGAAGTTTTAAACAATTTTATTGATTCATTAAAAGGAAAGAGTTGCCATGA
Clone 4         TGTAGAAGATGAAGTTTTAAACAATTTTATTGATTCATTAAAAGGAAAGAGTTGCCATGA
                ************************************************************

Pfphospho       ATTAATTACTGATGGATTAAAAAAATTACAAAATATTGGAGGTGGTGTAGCTGCCGCACC
Clone 4         ATTAATTACTGATGGATTAAAAAA-TTACAAAATATTGGAGGTGGTGTAGCTGCCGCACC
                ************************ ***********************************

Pfphospho       AGCTGGTGCCGCTGCAGTAGAAACTGCTGAAGCTAAGAAAGAAGATAAGAAAGAAGAAAA
Clone 4         AGCTGGTGCCGCTGCAGTAGAAACTGCTGAAGCTAAGAAAGAAGATAAGAAAGAAGAAAA
                ************************************************************

Pfphospho       GAAAGAAGAAGAAGAGGAAGAAGAAGACGACTTAGGATTTTCCTTATTTGGTTAAATATT
Clone 4         GAAAGAAGAAGAAGAGGAAGAAGAAGACGACTTAGGATTTTCCTTATTTGGTTAAATATT
                ************************************************************

                                                          Poly-T and anchor
Pfphospho       AAATATTTTGAGGAACCAAAACAAAAACCAAAAAAAAAAAAAAAAA-ATAAAAATAAAAA
Clone 4         AAATATTTTGAGGAACCAAAACAAAAACCAAAAAAAAAAAAAAAAAGGAGTGTTGTGGTAA
                ******************************************  *      *   **

Pfphospho       TATATATTCCAGTTCAAATTTTTATATTATAATTTTTGGAAAATTTTTCCATCACTTTGA
Clone 4         TG-ATAGG----------------------------------------------------
                *  ***

Pfphospho       ATAATTTTATATAACAATATTTTA
Clone 4         ------------------------
```

Figure 4. Alignment of the phosphoprotein gene of *Plasmodium falciparum* with the clone obtained using primer PSP2.

5.8S Ribosomal gene

```
Pfribo          TTATGTTTTGAAGAAAAGATACATATAATATATATTATATATATCGTAATTCATAATATA

Pfribo          ATTAGCACAGTCTTAACGATGGATGTCTTGGTTCCTACAGCGATGAAGGCCGCAGCAAAA

Pfribo          TGCGATACGCAATGAAAATTGCAGTGACTGTGAATCATCAGAATGCTGAATGTAAACTAC

Pfribo          ACCAATTCCTTCGGGGAATAGTGGTACTCCTACAGAAAATAAAATAATACTTTGCTCTCT

Pfribo          AGGTTTCTTTTAAATATAAAATATCTTTTCATATTATATGTATAAACAAAAACACAAATA

Pfribo          GGAATTTATTCCAAACTAAAAATAATTTTTAGTTAAAGTTTCTCGCTGAATAAAATTTAT

Pfribo          TTTCATTCATTTTCTAGAGTAACAAAAAAAGAATTCTTTATTCTTTTTAACAAAGTGTTT

Pfribo          GATTTGATGTTAGGAGGGCAAATCCGCCGAATTTAAGCATATAATTAAGCGGAAGAAAAG

Pfribo          AAAATAACTATGATTCCTTTAGTAACGGCGAGTGAACAAGGACCAGCTCAAACGGTTAAT

Pfribo          CCTTCGTTTGGTGCAAACCTTTCGACGGAATTGTCGTTGAATTATTCCATGAATGCCTTT

Pfribo          AATTTGAGCAAAGTGAGCAGGAATGCTTCCCCATAGAGGGTGAAAGGCCCATAGTTCTTT

Pfribo          TTAATTTATTCATTGGAATTTTTCATGCCGAGTCGTGTTCTTTGAGATTGGAGCACTAAT

Pfribo          ACGTGTGATACATTTCACATAAAGCTAAATATGTGTAGGAGACCGATAGCAAACAAGTAC

Pfribo          CGTGAGGGAAAGATGAAATAGTACTCAGGAATGAGCAATTAAATAGTACCTGAAATCGTT

Pfribo          AAGATGGAACGGATTAAGAGAGAAAACAAGTAAAGAGGGGAATTTTTAATTTTTTTTGTT

Pfribo          ATAATTCTCTTCTTTATTAAAAGAAACATCAGTGATTAATTTAATTTCAATAAAGCAATC

Pfribo          CCCTGAAATTCAAAATTTCTTTTAATTTTCTTTTCACTTTCTCCCCGCACTAATGTGGGG

Pfribo          AAAACTGGCTTTATTTCTTCAATTATTTTTTTTGCTGAGGAATTTTAAAATTATTTGAAT

Pfribo          TTTTCTTCTTATAATTTAAGTTGTTTCTATATAGTACTTTCTTAACCCACTCGTCTTGAA

Pfribo          ACACGGACCAAGGAGTCTAGCAAATGTGCAAGTGTATATGATTCTTTAAACATTTCTCTT

Pfribo          TTTAATATACGCATAATTAATGTAATATGTTTCTTTATTGTAGATTTGTGTGGTGTTTAA

Pfribo          TTTTTATTAAATCCCCACTTTGCATACATCCGGTAAGCAATTATGCTTTATTGAGTACGA

Pfribo          GCATATTTGGTAGGACCCGAGAGGCTTTGAACTAAGCGTGATGAGATTGAAGTCAGACGA

Pfribo          AAGTCTGATGGAGGATCGAGTTGATACTGACGTGCAAATCGTTCATTTCAATCACGTTTA
```

Primer PSP2

```
Pfribo          GGGGCGAAAGACTAATCGAAAAGCCTATTAGCTGGTTATTTTCGAAAGATCTCTCAGGAT
Clone 4a        ----------------------TAATTGGTGCTNTTGAAGTTCGAAAGATCTCTCAGGNN
Clone 14a       ----------------------TAATTGNTGCTNGTGAAGTTCGAAAGATCTCTCAGGAT
Clone 3         ----------------------TTNTTGGTGCTGGTGAAGTTCGAAAGATCTCTCAGGAT
Clone 10        ----------------------TAATTGGTGCTGGTGAAGTTCGAAAGATCTCTCAGGAT
Clone 1         ----------------------TAGTTGGAGCTGGTAATATTCNAAAGATCTCTCAGGAT
                                         *    ****** *  ********************

Pfribo          CGCTGGAGTTGAATTAATATAATTTTACCAGGTAGAGATAATGATTAGAGGAATCAGGGA
clones          CGCTGGAGTTGA-TTAATATAATTTTACCAGGTAGAGATAATGATTAGAAGAATCAGGGA
                ************ ************************************ **********
```

```
Pfribo          AATCGTTTTCCTTGACCTATTCTCAAACTCCCAATAGGTAAAAAGGGCGAAATAACTCAA
clones          AATCGTTTTCCTTGACCTATTCTCAAACTCCCAATAGGTAAAAAGGGCGAAATAACTCAA
                ************************************************************

Pfribo          CTTGTATTTCACTCATATTAAATAAAATAACTCCAAGTGGGCCATTTTTGGTAAGCAGAA
clones          TTTGTATTTCACTCATATTAAATAAAATAACTCCAAGTGGGCCATTTTTGGTAAGCAGAA
                 ************************************************************

Pfribo          CTGGCGATGAGGGATGCTCCTAACGCTTGGATAAGGTGCCTAAATATTCGCTCATGAGAT
clones          CTGGCGATGAGGGATGCTCCTAACGCTTGGATAAGGTGCCTAAATATTCGCTCATGAGAT
                ************************************************************

Pfribo          CCCATAAAAGGTGTTGGTTCATAATGACAGTAGGACGGTGGTCATGGAAGTCGAAATCCG
clones          CCCATAAAAGGTGTTGGTTCATAATGACAGTAGGACGGTGGTCATGGAAGTCGAAATCCG
                ************************************************************

Pfribo          CTAAGGAGTGTGTAACAACTCACCTACCGAATGAACTAGCCCTGAAAATGGATGGCGCTA
clones          CTAAGGAGTGTGTAACAACTCACCTACCGAATGAACTAGCCCTGAAAATGGATGGCGCTA
                ************************************************************
                                                    .

Pfribo          AAGCGGATTACCGATACCAAGCCATAAAAAGAACAGAAAAAATTTATTTTTTTTTTAGAA
clones          AAGCGGATTACCGATACCAAGCCATAAAAAGAACAGAAAAAATTTATTTTTTTTT-AGAA
                ************************************************************ ****

Pfribo          TCTTTTTATGAGTAGAAAATCGTGGGGTTTGTGTTGAAGCGAAATACGTGAGTTTTCGTG
clones          TCTTTTTATGAGTAGAAAATCGTGGGGTTTGTGTTGAAGCGAAATACGTGAGTTTTCGTG
                ************************************************************

Pfribo          GAACATCTCCCTAGTGCAGATCTTGGTGGAAGTAGCAACTATTCAAAATGAGAACTTTGA
clones          GAACATCTCCCTAGTGCAGATCTTGG-GGAAGTAGCAACTATTCAAA-TGGGAACTTTGG
                ************************** ********************** ** ********

Pfribo          AGACTGAAGTGGAGAAGGGTTTCTTGTCAACTGTGATTGAACAAGAGTTAGCCGCTCCTA
clones          AGGATGAAGGGGAGA-GGGTTTNTTTTCAACTGTGATGGAACAGGAGT-ACCGGTTCCAA
                **  ***** ***** ****** ** *********** ***** **** * * * *** *

Pfribo          AGGGATAGCTG-AAAAGTGTTTAAAAGGGGGTTCCTTCCCCGTCTCAAAAGGGAAACAGG
clones          GGGGATAGCTGGAAAAGTGTTTAAAAGGGGTTTCCTTCCCCGTCTCAAAAGGGAAACAGG
                 ********** ****************** ******************************

Pfribo          TTGATATTCCTGTGCCAATAGTATTATG-AGTTTCTTAGATGGTAA-CATATATATAAAT
clones          TTGATATTCCTGTGCCAATAGTATTATGGAGTTTCTTAGATGGTAAACATATATATAAAT
                **************************** ***************** *************

Pfribo          GAACTCCTTTACATAGGCTTTACACTCGGGGTGCGTTTTCTTTGCACTTTACCTTTATAA
clones          GAACTCCTTTACATAGGCTTTACACTCGGGGTGCGTTTTCTTTGCACTTTACCTTTATAA
                ************************************************************

Pfribo          CAAACCTTGGAATCAATTTACTTGGAGAAGAGGTTCGTTGAACTCAATTCAAAAAATTTC
clones          CAAACCTTGGAATCAATTTACTTGGAGAAGAGGTTCGTTGAACTCAATTCAAAAAATTTC
                ************************************************************

Pfribo          TTGAACGGGGTTTTCGGATTCAGTTCATTTTATTTTTTGTTTTGTAGCAATAGTAATTCG
clones          TTGAACGGGGTTTTCGGATTCAGTTCATTTTATTTTTTGTTTTGTAGCAATAGTAATTCG
                ************************************************************

Pfribo          TTTTTATGAATTATCCGAAGTGGTAAGGACTATCCTTGAAAAAAGGAGGGAACGGCTTTG
clones          TTTTTATGAATTATCCGAAGTGGTAAGGACTATCCTTGAAAAAAGGAGGGAACGGCTTTG
                ************************************************************

Pfribo          TAAACTTGGTTTCTTCTGATTCCATTTTGCTTTGCTATACTCTTAATACTTGTATGAGCG
clones          TAAACTTGGTTTCTTCTGATTCCATTTTGCTTTGCTATACTCTTAATACTTGTATGAGCG
                ************************************************************

Pfribo          TACCAACAACCGCATCAGGTCTCCAAGGTTAACAGCCTCTGGTTAAATAGAAAAAAGTAG
clones          TACCAACAACCGCATCAGGTCTCCAAGGTTAACAGCCTCTGGTTAAATAGAAAAAAGTAG
                ************************************************************
```

```
Pfribo          GTAAGGGAAGTCGGCAAAAATAGATCCGTAACTTCGGGAAAAGGATTGGCTCTGAGGAC
clones          GTAAGGGAAGTCGGCAAAA-TAGATCCGTAACTTCGGGAAAAGGATTGGCTCTGAGGAC
                ****************** ************************************

                     Poly-T and anchor
Pfribo          ATTAGAAAAGAGAAGAAAAAAAGAGGGTCGAAAATAAAATTGCAGTCTTATTTGCTTTTC
clones          ATTAGAAAAAAAAAAAAAAAAGAGTGTTGTGGTAATGATAGC----------------
                ********* * ** ********** ** *     *  ** **

Pfribo          TTTCGATTTGCTTGTAATCTGCTTTTTCTTTTTCTTCTTTCTTTTTTTCTCCCCCTCTTT

Pfribo          TCGCCTTCACTTTATTGTAATTTTATTACTTTAATTTGATACCTATAATGTTAACTCAGA

Pfribo          ACTGAAACGGACAAGGGGAATCCGACTGTTTAATTAAAACATAGCATTGTGAAAAGCCAC

Pfribo          AACTGGTATTAACACAATGTGATTTCTGCCCAGTGCTTTGAATGTCAACTTGATGAAATT

Pfribo          CAATCAAGCGCAGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCT

Pfribo          CGTCATCTAATTAGTGACGCGCATGAATGGATTAACGAGATTCCCACTGTCCCTACTTGC

Pfribo          TATCTAGCGAAACCACAGCCAAGGGAACGGGCTTGGCAAAATCAGCGGGGAAAGAAGACC

Pfribo          CTGTTGAGCTTTACTCTAGTCTGGCTTTGTGAAACGACTTAAGAGGTGTAGCATAAGTGG

Pfribo          GAGTAGAAACTGAAATATGTTTTTACGACAGTGAAATACCACTACTTTTAAAGTTGTTTT

Pfribo          ACTTATCCATTTATAGGGAAATATATTATGCTTTATCCTTCGGGAAGGCATTCTGATATT

Pfribo          ATTTGAATTTTTAAGAAACTTGTTTCTTATTTTTCTCCCATTTCTATGGAGACATAGCCA

Pfribo          GGTGGGGAGTTTGACTGGGGCGGTACATCTGTTAAATATTAACGCAGATGTCCAAAGACA

Pfribo          AGCTCAAAGAGAACAGAAATCTCTTGTAGACTAAAAGGGGAAAAGCTTGTTTGATTTCTA

Pfribo          TTTTCAGAACAAGTAGAAAACGTGAAAGCGTGGCCTATCGATCCTTTATATTTGCAAAAT

Pfribo          GACGTAAATTACTTACTACTGTGCATATAGAGGTGTCTGAAAAGTTACCACAGGGATAAC

Pfribo          TGGCTTGTGGCTGCCAAGCGCTCCTAGCGACGTAGCTTTTTGATCCTTCGATGTCGGCTC

Pfribo          TTCCTATCATTGGGACGCAGAAGTCTCAAAGTGTCGGATTGTTCACCCGCTAATAGGGAA

Pfribo          CGTGAGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAATTTTAT

Pfribo          TATATTTTTTATATATACATATAGTATTGTGACAGTAATCCAACTTGGTACGAGAGGATT

Pfribo          AGTTGGTTCAGACAATTGGTACAGCAATTGGTTGACAAACCAGTGTTGCGAACGTAAGTC

Pfribo          TGTTGGATAATGGCTGAACGCCTCTTAAGCCAGAAACCATGCTGATTAAACAATACTATT

Pfribo          TCGTTCTTTTTTGTTTCCTTTTTTACTGCTAATGTAAAAGAGAAATTATATATGTACCTT

Pfribo          TTTTTGGTACATAAATAATCCTATCACCTTTTTCGTGAGCTGTGAGCTGTGAGCTAGGCA

Pfribo          ATAGTTCTTTAACAACAAAAAAGAATTGCACAAACTGTAGACGACTTTTTTGCCTCAGGG

Pfribo          TGCTGTAAACATGAAAGTAAACTTTGTTTTACGATCTGTTGAGGCTTATCCCTAGTGGTA

Pfribo          AAGTGTTT
```

Figure 5. Alignment of the 5.8S ribosomal gene of *Plasmodium falciparum* with the clones obtained using primers PSP and TYT.

APPENDIX C

Cloning Vectors

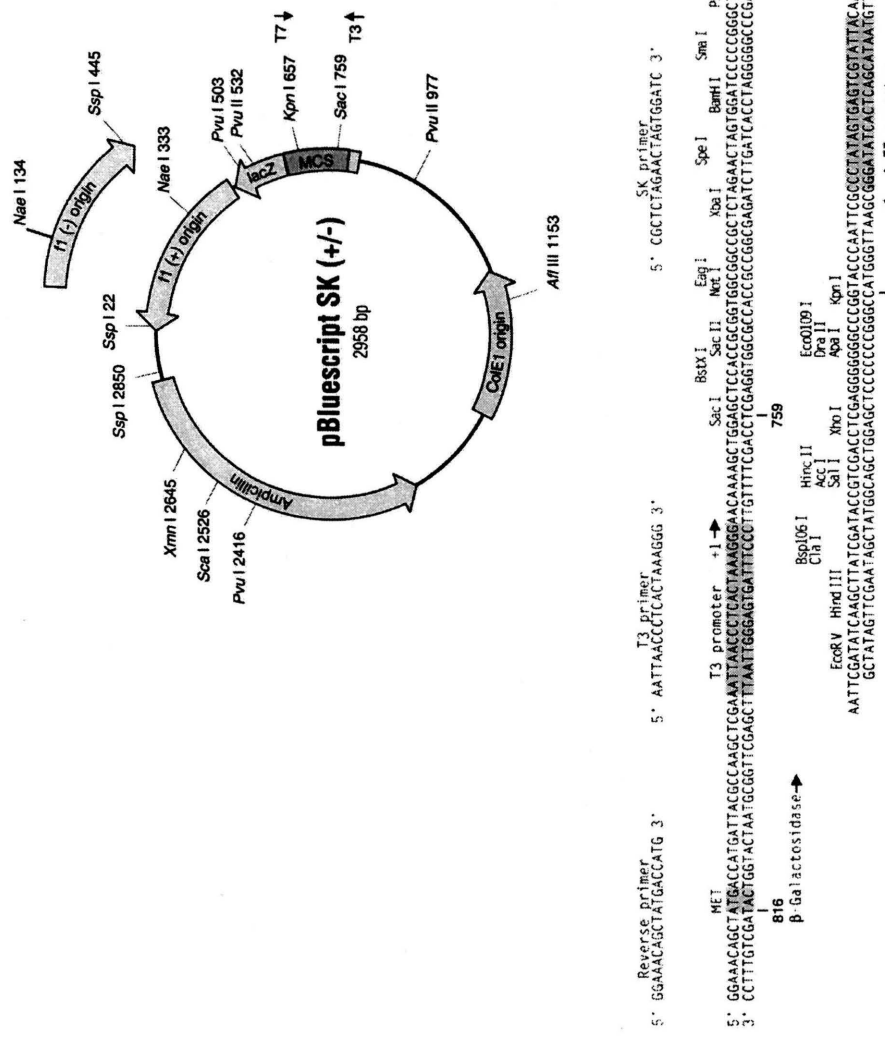


Figure 1. The genetic map of the cloning vector pBluescript SK+/- (Stratagene).

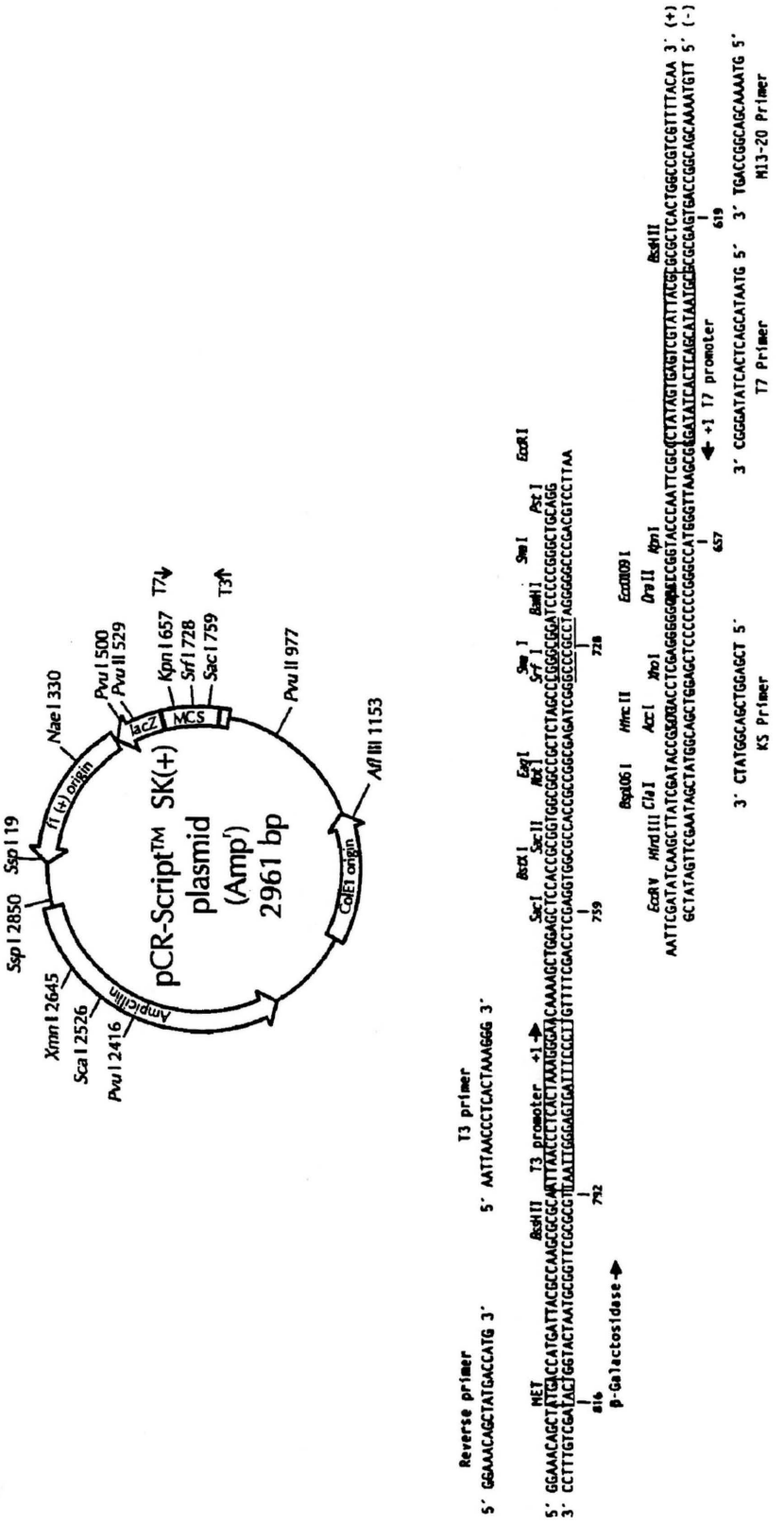


Figure 2. The genetic map of the cloning vector pCR-Script Amp (SK+) as used in the pCR-Script Amp (SK+) Blunt End Cloning Kit (Stratagene).

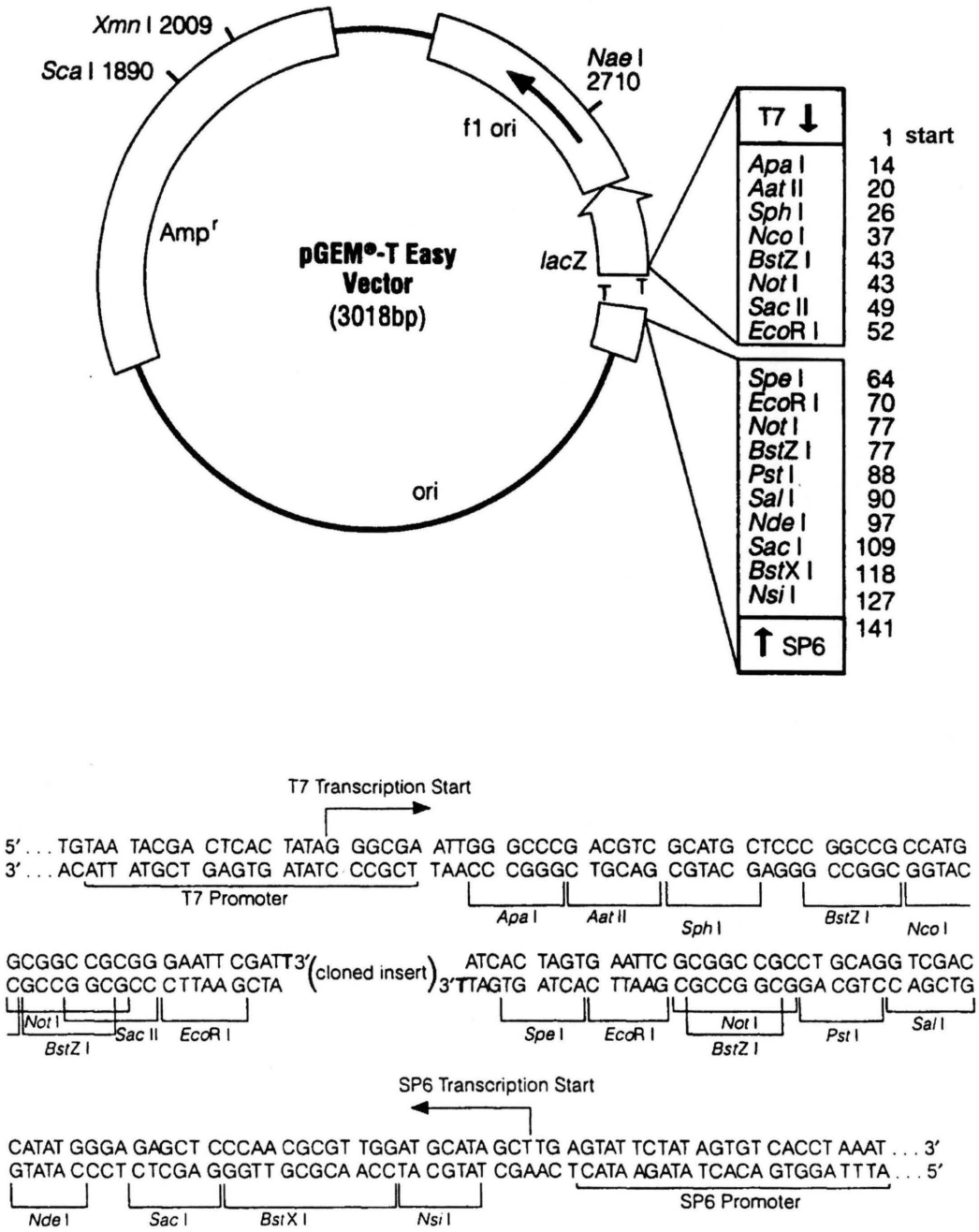


Figure 3. The genetic map of the cloning vector pGEM T-Easy as used in the pGEM T-Easy Vector System (Promega).