Studies of regulation of cell cycle in *Plasmodium falciparum*

Dar-Der Ji

PhD Thesis

Institute of Cell, Animal and Population Biology

University of Edinburgh

August, 1998

DEDICATION

This thesis is dedicated to my dear father, mother and young brother. It has been their sincere and continuous love and support that have made everything possible.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisors, Dr. David E. Arnot and Prof. David Walliker. Dr. David Arnot introduced me to malaria research and shared with me his sound knowledge in molecular biology, encouraged me to work independently by watching at a distance, and showed no dismay at my continuous questions and interruptions. Prof. David Walliker, who patiently taught me for the *P. falciparum* parasite culturing and the intricacies of malaria biology, and put his tremendous laboratory facilities at my disposal with no restrictions. I really admire their esteemed scientific judgement and friendly attitude.

I especially thank Mr. William Richardson and Dr. Alison Creasy for their support and warmly friendship during this work. They also help me to inprove my English and correct my thesis. I appreciate Ms. Kay Samuel to help the mouse immunisation. I thank Mr. William Gregory and Mr. Xing-Xing Zang for many technical and scientific discussions and contributions in this study. I thank Dr. Lisa Ranford-Cartwright for advice and help with parasite cultures used in this work. I would like to thank Prof. Rick Maizels to allow me using his marvellous laboratory facilities without restrictions.

Many thanks to Drs Alister Craig, Jana McBird and Keith Gull for providing the cDNA library, anti-*Trypanosome* α-tubulin monoclonal antibody and PfRap1 monoclonal antibody used in this work. Thanks to Drs Richard Carter, Alex Loukas, Judy Alen, Mercy Sowa and Mark Blaxter for occasional different help. Also I would like to thank Mr. Chis Ward, Mr. Richard Fawcett and Mr. Ahmed Raza for excellent technical assistance at various points. Mr. Frank Johnston and Mr. Graham Brown are thanked their excellent photographic services. I also thank the doctors and staff of the Institute of Cell, Animal and Population Biology of University of Edinburgh for their support and kindness.

I am very grateful to Profs Y.T. Liu, C.C. Chen and M.F. Chiao for his encouragement and support over the years in both Taiwan and Edinburgh. I would like to thank K.T Hsia and W.L. Hu couple and other friends for their help and encouragement. The financial support for this work was provided by my parents and the British Government awarded Overseas Research Studentship.

Finally I would like to express my sincere gratitude to my parents and young brother whose their encouragement, understanding and support have been a great part of the energy that has driven me to complete this study.

ABSTRACT

Malaria is still the most serious tropical disease in the world. It kills at least one million people every year, mainly young children in Africa. Malaria transmission is increasing due to insecticide resistance of the malaria-carrying mosquito, control programme failures, drug resistance of the malaria parasite and perhaps even global warming. Therefore, a greater understanding of the genetics and molecular biology of the parasite is urgently needed in order to develop novel chemotherapeutic and vaccine based control strategies.

The research aims of this thesis were to study genes and proteins which regulate the development of the malaria parasite, *Plasmodium falciparum,* through its intra-erythrocytic life cycle. A *Plasmodium falciparum* homologue of one of the components of a Chromatin-Remodelling Complex which controls binding of transcription factors to nucleosome core particles has been cloned and characterised. This sequencing effort has been broadened into a genome data bank-based review of intons and promotor sequences in *Plasmodium falciparum.* Immunofluorescence microscopy also shows that the PfSNF2L protein is expressed throughout the asexual stage of the cell cycle and is localised within nucleus. A proteolytic process and granular pattern on the infected red blood cell membrane is revealed by using anti-PfSNF2L polyclonal antibodies on western blotting and immunoflourescence assay.

Biological experiments on the cell cycle of malaria parasite have also been carried out. Process through the cell cycle has been shown to be retarded by heat shock and blocked at certain specific cell stage by drugs known to block certain key cell cycle regulatory processes in other cells. The nature of the cell death process in cycle arrested cells has also been investigated and its role in apoptosis-like cell death in malaria cells has been investigated.

TABLE OF CONTENTS

ABBREVIATIONS

ACF	ATP-utilizing chromatin assembly and remodelling factor
Amp.	ampicillin
BAFs	BRG1-associated factors
bp	base pair(s)
Bq	Becquerels
BRG1	brm/SWI2-related gene 1
brm	braham
BSA	bovine serum albumin
^{o}C	degree centigrade
CAM	calmodulin
CaMK	Ca^{2+}/ calmodulin-dependent kinase
cdc	cell division cycle
cdc	cell division cycle gene
CDK	cyclin dependent kinase
cDNA	complementary deoxyribonucleic acid
CHRAC	chromatin assessibility complex
Ci	curie
CN	calcineurin
cpm	counts per minute
DAG	diacylglycerol
DAPI	4'-6-diamino-2-phenylindole
dATP	deoxyadenosine-5'- triphosphate
dCTP	deoxycytosine-5'- triphosphate
(d)dATP	2', 3'-dideoxyadenosine-5'-triphosphate
(d)dCTP	2', 3'-dideoxycytidine-5'-triphosphate
(d)dGTP	2', 3'-dideoxyguanosine-5'-triphosphate
dGTP	deoxyguanosine-5'- triphosphate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNAase	deoxyribonuclease
DTT	dithiothreitol
dTTP	deoxythymidine-5'- triphosphate
EDTA	ethylenediamine N,N,N',N', tetra-acetic acid
E2F	transcription factor E2F
EGTA	ethyleneglycol-bis-(β-aminoethyl ether) N,N,N',N'-tetra-acetic acid
ERK	extracellular -signal regulated kinase

FITC	fluorescein isothiocynate
G_1	gap 1, period after mitosis, before DNA synthesis
G_2	gap 2, period after DNA synthesis, before mitosis
GAP	GTPase-activating protein
GCG	Genetic Computer Group
GDP	guanosine diphosphate
$G_{\alpha,\beta,\gamma}$	trimeric G protein α,β,γ subunits
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
HEPES	N-(2-hydroxyethyl) piperazine-N'-(2-ethanesulphoic acid)
IFA	immunofluorescence assay
ISWI	initial switch
kb	kilobase
kDa	kilodalton
LB medium	Luria-Bertani medium
MAP	mitogen-activated protein
Mek	MAP kinase kinase
MPF	maturaion or mitosis promoting factor
mRNA	messenger ribonucleic acid
NP-40	NONIDET-40, nonionic detergent
NPS1	nuclear protein of *Saccharomyces*
NURF	ATP-dependent nucleosome remodelling factor
O.D.	optical density
^{32}P	β emitting isotope of phosphorous
PBS	phosphate buffer saline
PCC	premature chromosome condensation
PCR	polymerase chain reaction
PFGE	pulsed-field gradient gel electrophoresis
PI_3	inositol-3,4,5-diphosphate
PIP_2	phosphatidyl inositol-4,5-diphosphate
PIPES	piperazine-N,N'-bis (2-ethane-sulphonic acid)
PKC	protein kinase C
PLC	phospholipase C
PMSF	phenylmethyl sulfonyl fluoride
PyK2	protein tyrosine kinase 2
Ran	ras-related nuclear protein
ras	small GTPase
RBC	red blood cell

RCC1	regulator of chromosome condensation
RITC	rhodaamine isothiocyanate
RPMI	Rosewell Memorial Park Institute
RNA	ribonucleic acid
RSC	remodel the structure of chromatin
^{35}S	β emitting isotope of sulphur
SAGA	Spt-Ada-Gcn5-acetyltransferase
SANT	SWI3, ADA2, N-CoR and TFIIIB B'
SDS	sodium dodecyl sulphate
SNF	sucrose non-fermenting
SNF2L	SNF2-like protein
SPF	synthesis-promoting factor
SRE	serum response element
SRF	serum response factor
SWI	mating type switch
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TCF	ternary complex factor
TEMED	N,N,N,N,-tetramethylethylenediamine
Tris	Tris (hydroxymethyl) aminoethane
Tween 20	polxythlenesorbitan monolaurate, nonionic detergent
ts	temperature sensitive
UV	Ultra-violet
w/v	weight per volume

LIST OF FIGURES AND TABLES

Figures:

Tables:

CHAPTER 1
INTRODUCTION

1.1 Malaria: the hundred years war

It is interesting to note that exactly one hundred years ago Patrick Manson announced to a meeting in the city of Edinburgh that Major Ronald Ross of the Indian Army Medical Service had made the landmark discovery that the vector for the transmission of malaria is the *Anopheles* mosquito. Unfortunately, despite a century of research and control efforts, malaria is still the most serious tropical disease in the world and is endemic in 99 countries where 2.2 billion people (41% of the world's population) are in danger of infection. More than 120 million clinical cases and over 1 million deaths occur each year, mainly children under the age of five infected with the malaria parasite *Plasmodium falciparum* in Africa (WHO, 1992). Malaria also exerts a heavy economic burden. The total cost of malaria was estimated to be around $800 million for tropical Africa in 1987, rising to more than $1800 million in 1995 (Burgess, 1997).

By many criteria, the war against malaria is stalemated if not actually being lost. *Plasmodium* parasites, which are the disease causation agents, are resistant to a wide range of drugs and the mosquito vectors have also evolved resistance to pesticides. Despite decades of research, effective anti-malarial vaccines are not yet available although several candidates are now undergoing clinical trials. The existing rather short list of effective antimalarial drugs is constantly being threatened by the evolution of new resistant malaria strains and the anxiety this provokes in the malaria research community does not seem to have increased the rate of development of replacement therapies. A better understanding of the molecular mechanisms underlying the complex life cycle of the malaria parasite could aid the search for novel anti-malaria treatments and this hope underlies much of the work carried out in this thesis.

1.2 The life-cycle, genetics and genomics of malaria parasites

1.2.1 The life-cycle of the malaria parasites

The obligate protozoan parasites which cause malaria are restricted to the family *Plasmodiidae* within the order *Coccidiida*, sub-order *Haemosporidiidea.* There are nearly 120 known species of *Plasmodia* found in vertebrates, namely reptiles, birds, rodents, primates and other mammals. Human malaria is caused by four species, *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*, all of which are only

transmitted by female *Anopheles* mosquitoes. The life cycle of human malaria parasites essentially comprises of an exogenous cycle leading from gametic fusion to sporogony in the mosquito and an endogenous asexual multiplicative cycle in the vertebrate host. The asexual replicative cycle includes the developmental stages in the erythrocytes, erythrocytic schizogony, and exo-erythrocytic schizogony in liver hepatocytes. A comprehensive life cycle is depicted in Figure 1.1.

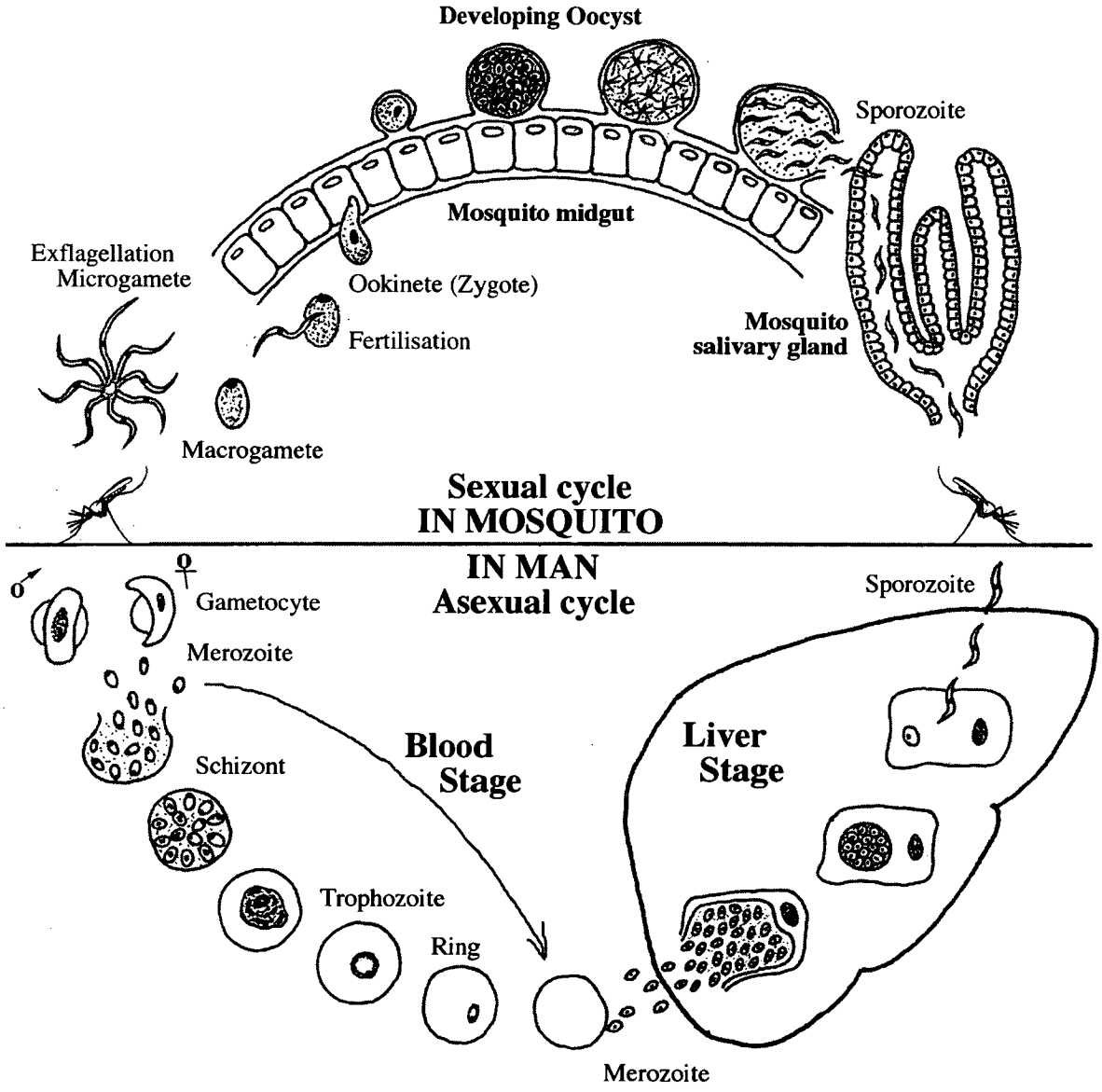
1.2.2 General genetics and genomics of malaria parasites

The malaria parasite genome is haploid for most of the life cycle, becoming diploid only briefly after fertilisation when the zygote is formed. Evidence shows that only one allele of each chromosome is found within a single cloned parasite. (Walliker *et al*, 1987). Crossing experiments between clones of the different rodent malaria parasites *P. berghei* and *P. chabaudi* and *P. falciparum* have demonstrated that blood forms of the parasites are haploid (Walliker *et al*, 1973; Walliker *et al*, 1975; Walliker *et al*, 1987). Genetic re-assortment can occur by recombination between chromosomes and by crossing-over between genes on homologous chromosomes during meiosis in the mosquito cycle (Fenton & Walliker, 1990; Sinnis & Wellems, 1988; Walker-Jonah *et al*, 1992). Intra-genic recombination within the *msp-1* gene has been detected in some of the progeny of the laboratory crosses referred to above (Kerr *et al*, 1994).

Malaria parasites have a nuclear genome and two extranuclear genomes (Creasey, 1996). *P. falciparum* has a haploid genome of around 30 megabases of DNA. The haploid genome comprises 14 chromosomes, determined both by resolution of chromosomes on pulse-field gel electrophoresis (PFGE) and by counting the number of kinetochores in electron micrographs (Kemp *et al*, 1987; Prensier & Slomianny, 1986; Wellems *et al*, 1987). The chromosome sizes range from about 650 kb (chromosome 1) to over 3 Mb (chromosome 14). In most malaria species and particularly *P.falciparum* the nuclear genome is extremely A+T-rich (82%) (McCutchan *et al*, 1984; Pollack *et al*, 1982). The mRNA coding strand is more purine rich. Overall A+T content in the triplet codon positions increases in the order 1st-2nd-3rd position and the codons used in genes are strongly biased towards using the higher A+T contents (Weber, 1987). There is no significant difference in codon usage between housekeeping and antigen genes (Musto *et al*, 1995).

A multi-copy 6kb linear molecule and a lower copy number 35kb circle of extrachromosomal DNA have been found in all malaria parasites (Wilson & Williamson, 1997; Wilson *et al*, 1991). Biochemical and genetic studies have led to the conclusion that the 6kb genome is the *Plasmodium* mitochondrial DNA, usually located in the mitochondrial fraction (Wilson *et al*, 1992), and having typical mitochondrial genes. such as cytochrome *b* and cytochrome oxidase subunit I and II

Figure 1.1 Life cycle of *Plasmodium falciparum.* The sexual cycle culminates in the production of sporozoites in mosquitos. Meiosis occurs just after zygote formation. Sporozoites infect liver hepatocytes and the parasite then reproduces asexually, in the liver cells and then in erythrocytes.


Developing Oocyst
Sporozoite
Mosquito midgut
Exflagellation
Microgamete
Ookinete (Zygote)
Fertilisation
Mosquito
salivary gland
Macrogamete
Sexual cycle
IN MOSQUITO
IN MAN
Asexual cycle
♂
♀
Gametocyte
Merozoite
Sporozoite
Blood
Stage
Liver
Stage
Schizont
Trophozoite
Ring
Merozoite

(Aldritt *et al*, 1989). It also contains at least 13 fragmentary ribosomal RNA genes (Vaidya *et al*, 1989).

The 35kb element is in some ways similar to the chloroplast and plastid genomes of plants or algae with prokaryote-like features (Gardner *et al*, 1991). It encodes about 60 genes, mainly ribosomal RNAs (rRNAs), transfer RNA (tRNAs), ribosomal proteins and the subunits of the RNA polymerase. It has now been classified as the malaria plastid-like DNA (Gardner *et al*, 1991).

Comparison of homologous chromosomes from different geographic isolates indicates that all 14 *P. falciparum* chromosomes display extensive clone-dependent size polymorphisms (of up to 15%) in both field isolates and parasites grown in culture (Bigg *et al*, 1989; Corcoran *et al*, 1986; Kemp *et al*, 1985; Van der Ploeg *et al*, 1985). While some genetic variability results from sexual recombination during meiosis, the extensive chromosomal polymorphism observed also appears to originate spontaneously during mitotic growth *in vivo* or *in vitro* (Walliker, 1989). Analysis of the chromosome 2 and 4 contig maps has revealed that polymorphisms are mainly confined to chromosome ends. The central chromosome regions are more conserved and tend to contain 'housekeeping' genes rather than antigen genes (Langsley *et al*, 1983; Lanzer *et al*, 1993; Lanzer *et al*, 1994; Lanzer *et al*, 1995; Scherf, 1996).

An abundant subtelomeric repetitive DNA sequence, termed *rep20*, has been reported by several groups (Aslund *et al*, 1985; Oquendo *et al*, 1986; Rao *et al*, 1986). This 21-bp tandemly repeated sequence is AT-rich and is not transcribed in the erythrocytic stages of *P. falciparum* parasite (Oquendo *et al*, 1986). This sequence is present in all parasite chromosomes and displays significant restriction fragment length polymorphism between different isolates of *P. falciparum* (Oquendo *et al*, 1986), possible as a result of mitotic recombination events.

1.2.3 General molecular biology of malaria parasites

Malaria chromosomes are linear and contain telomeric regions similar to other lower eukaryotes such as yeast (Dore *et al*, 1986; Vernick & McCutchan, 1988). The nuclear DNA is enclosed within a nuclear envelope, bound to histones and packaged into nucleosomes (Cary *et al*, 1994; Dore *et al*, 1986; Wunderlich, 1980). However, little is known about genome structure and organisational elements such as centromeres, replication origins and the structural elements involved in transcription.

Malaria messenger RNAs (mRNA), like other eukaryotic mRNAs appear to be capped at their 5' end (Hyde *et al*, 1984). Unlike small trypanosome and some nematode messages, there is no trans-splicing during malaria mRNA splicing (Blumenthal & Spieth, 1996; Blumenthal & Steward, 1996; Vanhamme & Pays, 1995). The 5'-untranslated regions (UTRs) of malaria messages are co-linear with

their coding sequences and these are transcribed together from genomic DNA. All of the *Plasmodium* 5'-UTRs characterised to date are much longer than those of most eukaryotes, which are usually 100 nucleotides or less. For example, in *P. falciparum* the 5'-UTRs of the *PfRan* and *PfSNF2L* genes are at least 679 and 519 bp respectively (Ji & Arnot, 1997; Sultan *et al*, 1994). Multiple potential transcription start sites are also often seen in the coding sequences of many malaria genes, although their significance is not understood(Ji *et al*, 1997; Lewis, 1990). The sequences flanking the ATG initiation codon in mRNA are believed to play a critical role in the initiation of protein translation in higher eukaryotes (Herman, 1988). From earlier studies of *Plasmodium* genes, a 5'-consensus sequence (AAAA<u>ATG</u>**A**) was found with the most conserved "A" at position –3 using the definition of Kozak where the "A" of ATG codon is position 1 and preceding base is -1 (Kozak, 1984; Kozak, 1997). This is somewhat different to the consensus sequences ((A/T)A(A/C)**A**(A/C)A <u>ATG</u>**T** and GCC(**A/G**)CC<u>ATG</u>**G**) found in yeast and vertebrates (Hamilton *et al*, 1987; Kozak, 1987).

Plasmodium introns are short and extremely AT-rich. However, the sequences of their donor and acceptor sites at splice junctions faithfully follow the GT-AG rule and closely resemble those of other eukaryotes (Prasartkaew *et al*, 1996; Vinkenoog *et al*, 1995). The 5'-splice site sequence 'AAG|**GT**AA(G/A)(A/T)' conforms to the consensus 'NAG|**GT**AAGT' from invertebrate while the 3'-sequence seems a T-rich variant of the pyrimidine enriched consensus 'T(T/C)T(T/C)T(T/C)(T/C)(T/C)(T/C)(T/C)NC**AG**|G' with some modifications (Senapathy *et al*, 1990). Most malaria genes described to date have a relatively small number of introns but many have no introns at all. Approximately 5% of the total RNA of *P. falciparum* is polyadenylated, and the poly(A) tails have, on average, 100-200 nucleotides (Vaidya *et al*, 1984). Heterogeneous polyadenylation from different sites has been observed at the 3'-end of some *Plasmodium* messages, a feature which has been proposed to play a role in stage specific gene expression (Ruvolo *et al*, 1993).

1.3 The *Plasmodium falciparum* life cycle and its relation to the cell cycle

This thesis describes work on the molecular biology of *P. falciparum* which is oriented torwards understanding how key molecular components of the parasite are integrated into its life cycle and development. A unifying theme of the work is the delineation of signaling pathways and their roles in parasite gene expression. Relevant areas of eukaryotic molecular biology are therefore reviewed here and where possible, integrated with known features of the parasite's biology.

1.3.1 Signalling pathways in eukaryotic cells

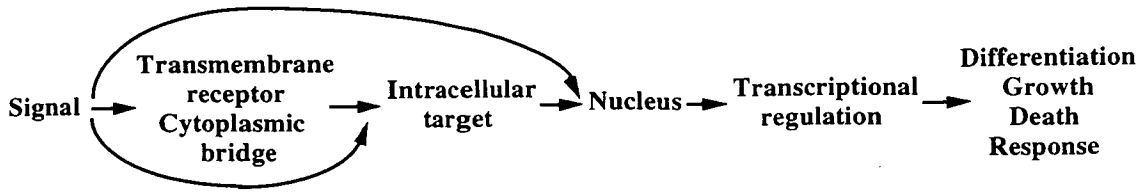
In unicellular organisms, such as bacteria, yeast or *Plasmodium*, growth is fundamentally controlled by the availibility of nutrients. Each cell is competing for these nutrients with neighbours of either the same or other species. This is different from multicellular organisms consisting of differentiated cells which integrate cellular activity in all of the basic process of life, eg. movement, metabolic activity, growth and ultimately death. However all cells must respond to the surrounding environment and control the entry and exit of molecules through the cytoplasmic membrane (Hardie, 1991). Signals such as growth factors, cytokines or calcium (Ca^{2+}), either stimulate transmembrane receptors or enter the cell cytoplasm directly and trigger receptors to modify or bind intracellular targets and thus trigger a series of signal cascades. Information may ultimately be transmitted into the nucleus to initiate gene expression (Figure 1.2A). A classic example are steroid hormones which diffuse through the plasma membrane into the cytoplasm and bind steroid receptors, before translocating to the nucleus to interact with specific promoters which regulate transcription (Mangelsdorf *et al*, 1995).

Another well-known example is the proto-oncogene *c-fos*, which is rapidly and transiently induced by a multitude of extracellular stimuli. A cascade of events is triggered by cell stimulation, as presented in Figure 1.2B. The transcription factor known as the serum response factor (SRF) and its accessory protein TCF, (ternary complex factor) bind the *c-fos* SRE (serum response element, a promoter element) co-operatively to form a ternary complex to control *c-fos* transcription (Janknecht *et al*, 1995). SRF and TCF, in turn, are tightly regulated by several signal transduction pathways, such as extracellular-signal regulated kinase (ERK), phosphatidylinositol (PtdIns) 3-kinase, PtdIns-specific phospholipase C (PLC)-γ, c-raf, trimeric G-protein and ras-GAP(see Figure 1.2B for the details) (Cahill *et al*, 1996)). An increase in *c-fos* protein can then produce a *Jun-Fos* complex (AP-1) capable of activating transcription of a variety of genes (Cantley *et al*, 1991).

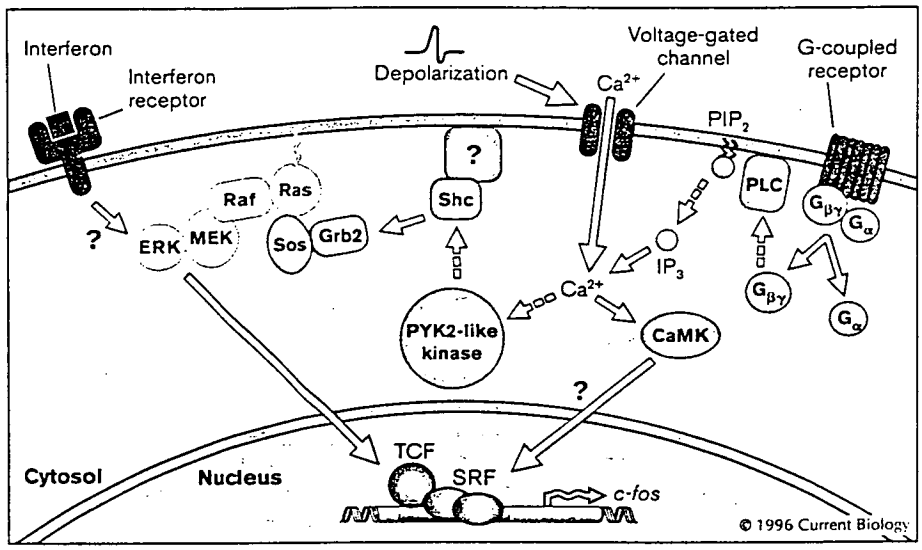
Cell cycle control is mediated by a profusion of growth factors acting on different messenger pathways (Figure 1.2C). Both G-protein linked receptors and tyrosine kinase linked receptors can stimulate proliferation. The inositol (1,4,5)-triphosphate ($InsP_3$)/calcium signalling pathway is activated by mitogenic stimuli and may contribute to changes which culminate in DNA synthesis. Activation of transcription factors which then translocate into the nucleus indicates that some of the actions of calcium occur in the cytoplasm. Calcium also can activate mitogen-activated protein (MAP) II kinase, part of the phosphorylation cascade that culminates in the activation of DNA synthesis-promoting factor (Berridge, 1993).

Figure 1.2 Signalling pathway networks in eukaryotic cell. (**A**) A comprehensive signal cascade pathway. (**B**) Intracellular signalling pathways that converge on the *c-fos* SRE. ERK activation by $G_{\beta\gamma}$ subunits, G_i-coupled receptors, voltage-gated channels and interferon receptors. Activation of CaMK, leading to SRF activation, is also shown. (**C**) Signalling pathways operating during G1 phase. Growth factors generate a number of putative mitogenic messengers which then feed into various pathways to control both early events at the G_0/G_1 boundary and later events which culminate in the onset of DNA synthesis. SRE, serum response element; SRF, serum response factor; TCF, ternary complex factor; CAM, calmodulin; CaMK, Ca^{2+}/ calmodulin-dependent kinase; PLC, phospholipase C; PIP_2, phosphatidyl inositol-4,5-diphosphate; PI_3, inositol-3,4,5-diphosphate; ERK, extracellular-signal regulated kinase; SPF,synthesis-promoting factor; GAP, GTPase-activating protein; CN, calcineurin; DAG, diacylglycerol; PKC, protein kinase C; shc, src-homology; ras, small GTPase; Grb2, adaptor protein; PyK2, protein tyrosine kinase 2; E2F, transcription factor E2F; c-fos, proto-oncogene; $G_{\alpha,\beta,\gamma}$, trimeric G protein α,β,γ subunits; MAP, mitogen-activated protein; Mek, MAP kinase kinase. (Adapted from Cahill *et al*, 1996; Berridge 1993).

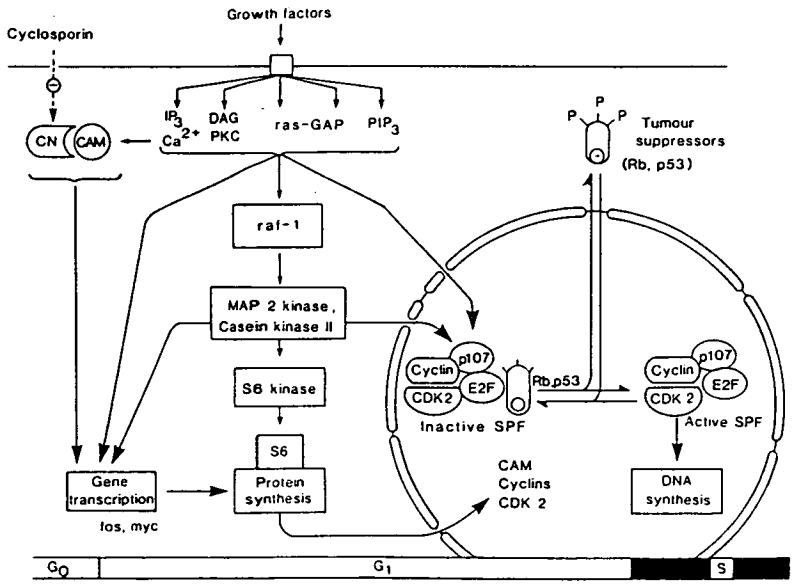
A


Signal
Transmembrane receptor Cytoplasmic bridge
Intracellular target
Nucleus
Transcriptional regulation
Differentiation
Growth
Death
Response

B


Interferon
Interferon receptor
Depolarization
Voltage-gated channel
G-coupled receptor
Ca2+
PIP2
PLC
Shc
Raf
Ras
ERK
MEK
Sos
Grb2
IP3
PYK2-like kinase
CaMK
TCF
SRF
c-fos
Cytosol
Nucleus
© 1996 Current Biology

C


Growth factors
Cyclosporin
CN
CAM
IP3
Ca2+
DAG
PKC
ras-GAP
PIP3
Tumour suppressors
(Rb, p53)
raf-1
MAP 2 kinase, Casein kinase II
S6 kinase
S6
Gene transcription
Protein synthesis
fos, myc
Cyclin
p107
E2F
CDK2
Rb,p53
Inactive SPF
Active SPF
CAM
Cyclins
CDK 2
DNA synthesis
G0
G1
S

1.3.2 DNA replication and the cell cycle in malaria parasites

The eukaryotic cell cycle is regulated both temporally and spatially. During mitosis, the nucleus, cytoskeleton, and membrane compartments undergo dramatic changes which must be tightly coordinated. Similarly, in replicating its genome the cell must ensure that each origin fires only once per cell cycle and there appears to be a set order in which particular origins initiate DNA synthesis.

The typical cycle is divided into two fundamental parts: interphase, which occupies the majority of the cell cycle and is subdivided into G_1 (gap1), S (DNA synthesis), G_2 (gap2), and mitosis (M). Interphase lasts about 30 minutes in rapidly dividing human cells. Two major 'checkpoints' regulating entry into DNA replication (the G_1/S checkpoint) and into mitosis (the G_2/M checkpoint) have been recognised (Murray & Hunt, 1993). DNA replication and synthesis of histone proteins occurs only in the S phase. Each double-stranded DNA molecule is replicated into two identical daughter double-stranded DNAs, and histone and other chromosomal proteins bind quickly to the newly synthesized DNA. S phase is preceded and followed by gap periods, G_1 and G_2 respectively, during which there is no net synthesis of DNA, although repair to damaged DNA occurs. During M phase, identical DNA molecules are condensed and distributed into each of the daughter cells through the classically observed prophase, pro-metaphase, metaphase, anaphase and telophase movement stages.

The life cycle of malaria parasites contains several cell cycles and there are at least five stages of DNA synthesis within the *P. falciparum* life cycle. In the mosquito there is microgametogenesis, meiosis-associated DNA synthesis and sporogony. Within the human there is exo-erythrocytic schizogony in the liver and erythrocytic schizogony in the RBCs. Malaria parasites are incapable of *de novo* synthesis of the purines necessary for the synthesis of nucleic acids and other metabolic functions, and operate purine salvage pathways (Homewood & Neame, 1980). However *de novo* synthesis of pyrimidines does occur involving the vitamin para-amino-benzoic acid (PABA) and folate co-factors. The key biochemical marker of cell cycle progression, DNA synthesis, starts in synchronised *P. falciparum* cultures about 28-31 h after merozoite invasion and DNA content then increases for around 8-10 h. A general model has been proposed by Arnot and Gull (1998) who interpreted the relationship between asexual parasite development and cell cycle as follows:

(1) The invasive merozoite is in a G_0-like state with condensed chromatin. The developing trophozoite is formally similar to the G_1-stage of mammalian cells or yeast and the morphologically defined 'rings' and early trophozoites are equivalent to G_1.

Since DNA synthesis begins in relatively small trophozoites ('late rings') but subdivision of the nuclear material (the morphologically defined onset of schizogony) does not appear until the trophozoite is much enlarged, growth and DNA synthesis in trophozoites must occur concurrently (Hare, 1986). Prior to DNA synthesis trophozoites have pre-accumulated high levels of RNA (Hare & Bahler, 1986). The steady-state transcripts from both PfPCNA (proliferating cell nuclear antigen) and PfDNA polymerase δ have been accumulated by the trophozoite stage (Horrocks *et al*, 1996). As parasites progress towards the stage where DNA replication occurs, there is a concomitant increase in both topoisomerase II production and activity (Cheesman *et al*, 1998). During this period, the parasite consumes host cell cytoplasm and grows to fill the erythrocyte.

(2) The S phase occurs during schizogony, but neither microscopy, precursor incorporation nor flow cytometry has been able to establish definitively whether G_2 and then M phases occur during progression from the haploid (1C) DNA content to the final genome complement of the mature segmentor (~8-30 C). Thus it is unclear whether a brief G_2 and M phase occurs after each genome doubling (6-8 hours for 4-5 rounds of mitosis). The mechanism of malaria S phase control during schizogony may be similar to the early embryonic development of *Drosophila* which *Drosophila* embryos undergo rapid rounds of DNA replication and nuclear division. In *Drosophila* embryo the first rapid rounds of replication are tightly regulated by reducing the pre-accumulated cyclin A and B protein levels (Edgar *et al*, 1994). Therefore, Leete and Rubin (1996) suggested that the overall control of the length of the schizont stage and the number of rounds of DNA replication may be controlled by the steady degradation of cyclins or some signals during the schizont stage. However, no cyclins have been identified to date from *P. falciparum* and the *P. falciparum* cdc2 homologues identified to date have not been proven encode a functional protein *in vivo* (Ross-MacDonlad *et al*, 1994; Li *et al*, 1996). Chromosomes do not condense in the nuclear divisions of *P. falciparum* and the nuclear membrane remains intact throughout genome replication. The subsequent mitotic divisions have been assumed to take place within the same nuclear envelope (Leete & Rubin, 1997; Read & Hyde, 1997). If mitotic genome segregations occur then they are colchicine insensitive and must be very rapid (Jacobberger *et al*, 1992).

(3) Segmentation occurs as an independent stage at the end of the asexual erythrocytic cycle and merozoites are formed then and not during S phase proper. Because the transitions from one stage to the next are not clearly separable biochemically or

morphologically, the relationship of schizogony to both the DNA synthetic cycle and the cell cycle is poorly defined.

Haploid (1C) merozoites can exit the erythrocytic, asexual replicative cycle and differentiate sexually into micro and macro-gametocytes. The gametocytes have generally been considered to be arrested in G_1 prior to the onset of S phase. However in both *P. berghei* and *P. falciparum* a considerable excess of DNA over the 1C value (almost reaching the 2C value in the latter case) has been measured after Feulgen staining and fluorometry (Janse *et al*, 1988; Janse *et al*, 1986; Janse *et al*, 1986; Janse *et al*, 1986). The near diploid amounts of DNA in most gametocytes make it impossible to rule out some features of an earlier model, based on the observation of gametocyte intranuclear microtubules, that gametocytes have undergone a nuclear division (Sinden, 1983; Sinden, 1982). It is, however, clear that microgametocytes have not pre-synthesised the 8C DNA compliment necessary to equip the 8 haploid microgametes released after microgametogenesis in the mosquito midgut.

1.3.3 Signal transduction and the *Plasmodium* cell and life cycles

Regulators of cell cycle progression are important targets in the signalling pathways that mediate cell growth and differentiation. In model organisms such as *Schizosaccharomyces pombe*, entry into M and S phase is regulated by the cyclin-dependent-kinase (CDK) CDC2 (Solomon *et al*, 1990), whereas in higher eukaryotes CDC2 homologues such as CDK1, control only entry into M phase, the S phase of the cell cycle being controlled by additional CDKs (Lees, 1995). Each phase of the cell cycle is initiated by accumulation of the appropriate cyclins above a threshold level, leading to activation of a cyclin-CDK complex. In general, each cyclin-CDK complex represses its predecessor and stimulates the expression of its successor, resulting in modulated waves of different cyclin expression (Lees, 1995). CDK activity is also regulated by binding of its inhibitors, like cyclin-dependent-kinase inhibitor (CKI), and by phosphorylation of regulatory sites on the CDK (Wittenberg & Reed, 1996). The relative significance of the three forms of CDK regulation (cyclin binding, phosphorylation and inhibitor binding) has not been clearly established, but cyclin accumulation appears to be the rate-limiting step in many organisms (Lees, 1995).

A putative CDC2 homologue, PfPK5, has been isolated from *P. falciparum* and can be partially activated *in vitro* by the same phosphorylation events as yeast CDC2 (Graeser *et al*, 1996; Ross-MacDonald *et al*, 1994). Studies in two recent isolated *P. falciparum* proteins, Pfmrk (a putative CAK homologue) and Pfcrk (a

cdc2-related kinase) also suggest that CDK regulation in malaria parasites may be similar to that observed in other eukaryotes (Doerig *et al*, 1995; Li *et al*, 1996).

Initiation of DNA replication in eukaryotes is regulated by 'replication licensing' to ensure that DNA replication occurs only once per cycle. Unreplicated DNA is in a replication-competent or 'licensed' state. The ability to replicate is lost in S phase and regained following passage through mitosis. The minichromosome maintenance (MCM) proteins have been indicated in determining this replication competence. Therefore, a malarial MCM homologue may be also involved in this control (Doerig, 1997; Romanowski & Madine, 1996; White & Kilbey, 1996). Other factors which may be involved in malaria cell cycle regulation include a *P. falciparum* homologue of the Ran/TC4, PfRan, a protein first characterised in this laboratory and independently in another group (Dontfraid & Chakrabarti, 1994; Sultan *et al*, 1994). Ran/TC4 is a small GTPase mainly responsible for nuclear tranport (this will be further discussed in a later part of the introduction). It has also been implicated in transmitting a signal that prevents the onset of mitosis before DNA replication is completed (Ren *et al*, 1993). Doerig further suggests that the PfRan may be involved in importing the CDK/cyclin complex into the nucleus (Doerig, 1997).

1.3.4. Signalling pathways in malaria parasites

A recent review by Doerig (1997) discusses hypotheses for the possible signal transduction pathways in malaria parasites. A brief summary of these pathways is presented below and in Figure 1.3.

cAMP-dependent pathways. cAMP is a key messeger of many developmental processes and is produced by membrane-bound adenylate cyclase, usually activated by heterotrimeric G-proteins (Tang & Gilman, 1991). Protein kinase A (PKA) can be activated by binding cAMP in order to phosphorylate a number of substrates including transcription factors which modulate gene expression. *P. falciparum* extracts have been shown to contain both adenylate cyclase and PKA activities (Read & Mikkelsen, 1990; Read & Mikkelsen, 1990). The addition of high level of cAMP has been proposed as the trigger mechanism for gametogenesis in *P. falciparum* although other groups have failed to reproduce this result (Inselburg, 1983; Kaushal *et al*, 1980). Putative gene homologues of adenylate cyclase and PKA have been isolated from different *Plasmodium* species (Doerig, 1997).

cGMP-dependent pathways. Elements of this pathway appear to be involved in gametogenesis in the vector midgut. Both cGMP and guanylate cyclase activators stimulate exflagellation in *P. falciparum* and *P. berghei.* A gene encoding a protein kinase G from *P. falciparum* has recently being reported (Doerig, 1997).

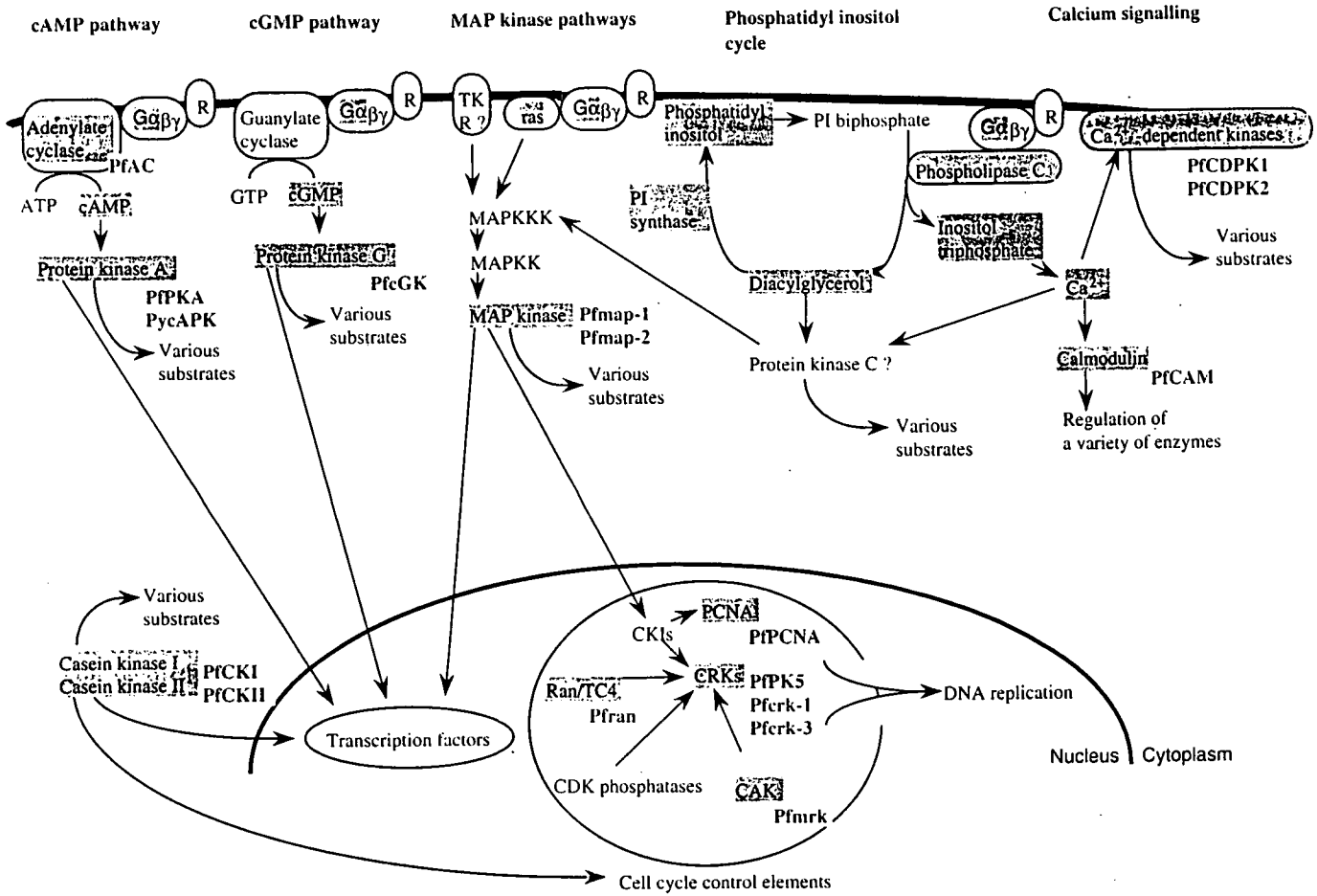


Figure 1.3 Signal transduction pathways which may operate in malaria parasites. For the detail, please see Introduction 1.3.4. Pf, *Plasmodium falciparum*; Py, *Plasmodium yeolii,* TKR, tyrosine kinase receptor; AC, adenylate cyclase, PKA, protein kinase A; PKC, protein kinase C; PK, protein kinase; KG, protein kinase G; CK, casein kinase; MAP, mitogen activated protein; CAM, calmodulin; CDPK, Ca^{2+}/calmodulin-dependent kinase; CDK, cyclin dependent kinase; CAK, CDK activating kinase; CRK, Cdc2 related kinase; mrk, MO15 related kinase, CKI, CDK inhibitor; PCNA, proliferative cell nuclear antigen. (Adapted from Doerig, 1997).

MAP kinase pathways. Several MAPKs pathways (mitogen-activated protein kinases), also called ERKs (extracellularly-regulated kinases), co-exist in eukaryotic cells and mediate responses to a wide variety of signals. MAPK substrates include transcription factors, cell cycle control elements, cytoskeleton components and downstream protein kinases. MAPK itself can be activated in a variety of ways, such as by G-protein coupled receptors via the ras protein, tyrosine kinase receptors, and protein kinase C (Cobb & Goldsmith, 1995). A MAPK homologue, Pfmap-1 (also called PfMAP or PfMRP) from *P. falciparum* has been identified and has been suggested to play a role in sexual development (Lin *et al*, 1996).

Phosphatidylinositol (PI) cycles Surface receptors stimulate phospholipase C to hydrolyse phosphatidylinositol 4,5-biphosphate (PIP_2) with the subsequent generation of the secondary messengers inositol-1,4,5-triphosphate (IP_3) and diacylglycerol (DAG) for downstream regulation. Exflagellation of *P. falciparum* microgametes is accompanied by the sequential PI degradation. A membrane bound merozoite protease can be solubilised through the action of GPI-phospholipase C in *P. falciparum* and *P. chabaudi*. The activated protease is required for merozoite invasion (Braun-Breton *et al*, 1992).

Protein phosphatases. Dephosphorylation of regulatory proteins is an important way of modulating the output of signal transduction cascades involving protein kinases (Hunter, 1995). Recently two different serine/threonine protein phosphatases (PP2A-like and PP1-like) activities have been detected and a putative protein phosphatase PP2A gene identified in *P. falciparum* (Doerig, 1997).

Calcium signalling. The importance of calcium in signal transduction is well known, and its indirect effects on the cell are widespread, including cell cycle arrest, regulation of meiosis and regulation of transcription factors. The calcium-binding protein, calmodulin, is an important mediator of calcium signalling and required for the regulation of a wide variety of Ca^{2+}-dependent enzymes (Clapham, 1995). Calcium antagonists inhibit gametogenesis in *P. falciparum* and *P. berghei* and prevent both normal morphological development and DNA synthesis (Kawamoto *et al*, 1993). Genes encoding a *P. falciparum* calmodulin homologue (Robson & Jennings, 1991), a calcium-binding protein, Pf40 (Rawlings & Kaslow, 1992), and two Ca^{2+}-dependent protein kinases, PfCDPK1 and 2 (Zhao *et al*, 1993), have been isolated. PfCDPK1 appears to be associated with the cell surface of the merozoite and is thought to play a role in RBC invasion. PfCDPK2 may be associated with a membrane or organelle fraction, as it contains a membrane-anchoring motif (Doerig, 1997).

Missing links between signalling pathways. Heterotrimeric G-proteins play an important role in activating many signalling pathways, which either directly contact their downstream receptors or act through mediators, such as the ras GTPase (Cobb *et*

al, 1995). There is some evidence that a G protein exists in asexual stage parasites (Thelu *et al*, 1994) but so far no heterotrimeric G protein gene has been reported (no sequences in GenBank and no references in Medline as of June, 1998). Surface receptors which can sense extracellular signals and initiate signal transduction cascades have not been identified and the precise signals which trigger developmental responses such as gametocytogenesis are mostly unknown. Stage-specific transcription factors are another essential group of response effectors and represent the target of many signalling pathways. Nothing is known about the machinery which regulates stage-specific transcription in malaria parasites.

1.4. GTPase mediated signal transduction pathways and nuclear transport in *P. falciparum*

1.4.1 G proteins and their structures in eukaryotic cells

Signal transduction is the process by which a receptor interacts with a ligand to transmit a signal to trigger a cascade or pathway ending with an effect within the cell. In all eukaryotic organisms, a family of heterotrimeric guanosine triphosphate-binding proteins (G-proteins) play an essential role in linking many cell surface receptors to effector proteins on the plasma membrane (Hepler & Gilman, 1992). G proteins have been found to play a central role in a broad array of cellular activities such as mating in yeast, chemical induced movement in slime mould, vision, small hormone secretion, muscle contraction and cognition in humans. Heterotrimeric G proteins are composed of α, β, and γ chains encoded by distinct genes. G_α subunits have a limited molecular weight range between 39 and 52 KDa, and 17 G_α subunits have, to date, been identified and subdivided into four main classes termed G_s, G_i, G_q and G_{12} (Hepler *et al*, 1992; Simon *et al*, 1991; Watson & Arkinstall, 1994).

The common architectural design shared by all G proteins, and the fact that they perform the same biochemical reaction, indicates their probable origin from a single progenitor. Most G proteins contain motifs of amino acid sequence that are conserved across the entire G protein superfamily. In the primary structure of each G protein these motifs appear in the same order, although spacing between them may differ. Figure 1.4 schematically depicts a G protein with conserved domains G1 to G5. Other regions of importance, for example, the 14-carbon myristic acid residues may be post-translationally added to the second glycine residue of some G proteins after cleavage of the N-terminal methionine to anchor the protein on the under-side of the plasma membrane. Cholera Toxin (CT) modifies G proteins by transfering the ADP-ribose moiety from NAD to the so-called cholera toxin site (CTx) site to block the G

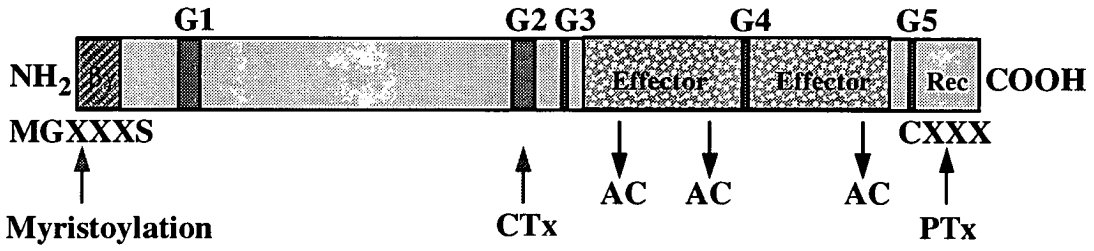


Figure 1.4 Schematic representation of generalised G_{α} protein showing domains likely to be important for interaction with receptor (Rec), effector and $\beta\gamma$ dimers. Also indicated are sites for myristoylation and ADP-ribosylation by cholera (CTx) and pertussis toxins (PTx). Shown within effector region are three zones identified within $G_{\alpha s}$ as particularly important for interaction with adenylate cyclase (AC). Conserved regions probably important for binding guanine nucleotides and GTPase activity are denoted by G1 to G5. Reproduced and modified from Watson *et al*, 1994.

protein in an active GTP-bound conformation. Pertussis toxin (PT) on G_α chains operates by ADP-ribosylating the Pertussis Toxin site (PTx) site thus uncoupling the heterotrimeric G protein from its receptors. Two effector binding sites are important for interaction with adenylate cyclase (AC): the C-terminus is important for receptor interaction and the N-terminus for interaction with $\beta\gamma$ dimers (Bourne, 1993; Watson *et al*, 1994). A generally accepted model for G protein activation is illustrated in Figure 1.5. The model is mainly based on studies performed with hormone-stimulated adenylate cyclase by $G_{\alpha s}$ protein and light-activated retinal rod cyclic GMP-phosphodiesterase (PDE) by transducin ($G_{\alpha t}$) (Bourne, 1993; Bourne *et al*, 1991; Watson *et al*, 1994).

1.4.2 Heterotrimeric G proteins in parasites

Heterotrimeric G proteins have been detected in a variety of parasites, including trypanosomes (Coulter & Hide, 1995; Oz, 1994), *Toxoplasma gondii* (Halonen *et al*, 1996), trematodes (Mansour & Mansour, 1986; Mansour & Mansour, 1989) and filarial nematodes (Grant *et al*, 1997). Studies of the liver fluke *Fasciola hepatica* found that activation of fluke adenylate cyclase by serotonin was dependent on GTP and that *F. hepatica* also contained G_α proteins of similar molecular weights which could be labelled by bacterial toxin ADP-ribosylation (Mansour *et al*, 1986). Studies in *Schistosoma mansoni* produced similar findings (Mansour *et al*, 1989) and a $G_{\alpha s}$-like gene from *S. mansoni* has been cloned which shows a 65% amino acid match with bovine $G_{\alpha s}$ (Iltzsch, 1992). Stage-specific differences in the level of expression of G_α and switches of G protein class have also been revealed by Western blot and RT-PCR experiments in *Acanthoheilonema viteae* (Grant *et al*, 1997), *Brugia pahangi*, *Onchocerca ochengi* and *Trypanosoma cruzi*. However the G proteins have not been linked to a role in regulating the adenylate cyclase signal transduction system in *T. cruzi* so far (Coulter *et al*, 1995; Oz, 1994). A recent study of G proteins in *T. gondii*, indicated the binding of radiolabelled GTP to molecules corresponding in molecular weight to G_α and the cholera toxin-induced ADP-ribosylation of a molecule of ~45 kDa (Halonen *et al*, 1996). Similar, unpublished experiments done in this laboratary indicate that heterotrimeric G-proteins may also exist in *P. falciparum* (Arnot and Shah, unpublished).

1.4.3 Other members of the GTPase family which are not part of membrane-associated complexes.

Evidence from the 3-D structure of guanine nucleotide binding domains of *ras* p21 and EF-Tu have indicated that the five stretches of amino acid sequences which define the GTPase superfamily form part of the guanine nucleotide binding pocket of

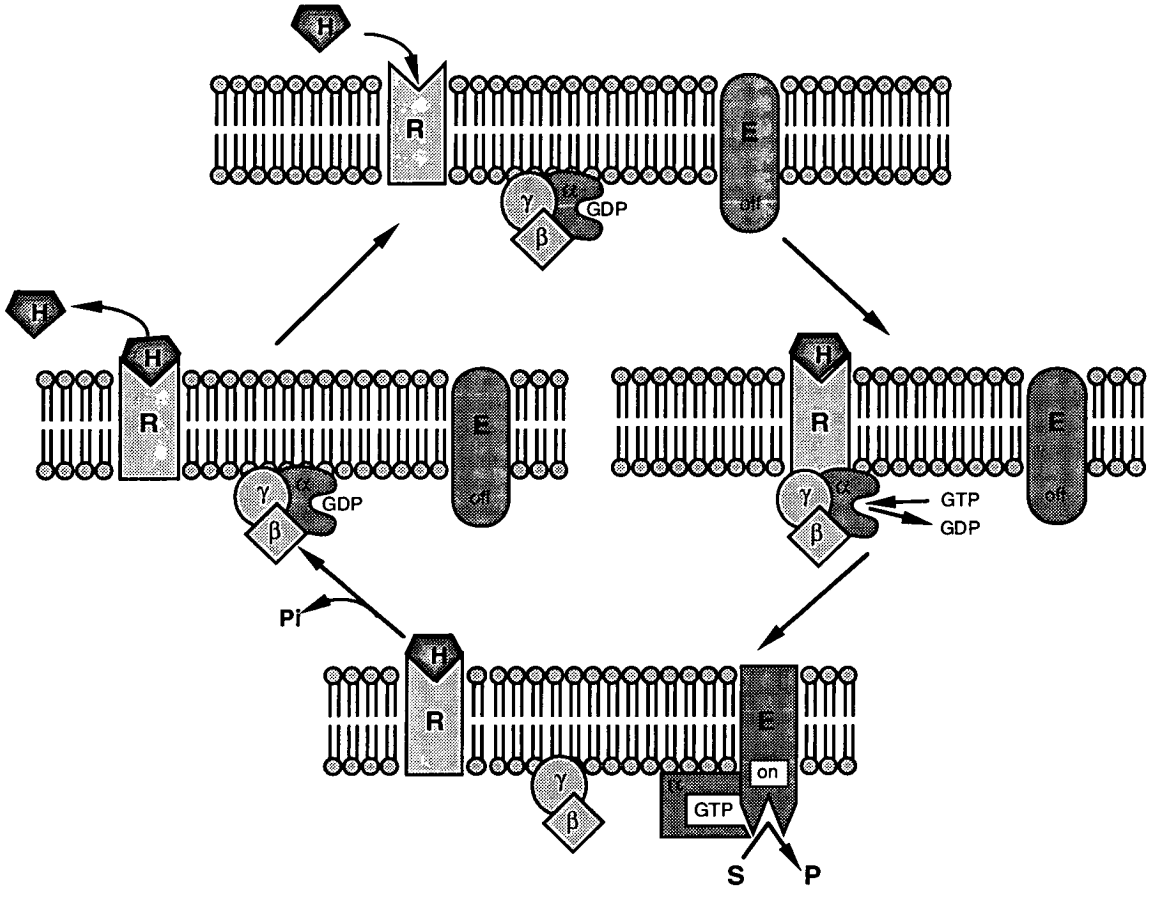


Figure 1.5 The G protein cycle, coupling hormone receptor (R) to its effector (E). Binding of hormone (H) to its receptor causes the G protein to exchange GDP for the GTP, which activates the G protein. The G protein then dissociates, after which the GTP-bound Gα (α) diffuses through and 'along' the membrane and binds to an effector thus activating it. The switch is on. Later, the Gα hydrolyses GTP to GDP and inorganic phosphate (P_i), thereby inactivating itself. The Gα will then reassemble with the Gβγ (β,γ)complex. The switch is then off. The Gαβγ complex then returns to resting state associated with the plasma membrane.

GTPase (Neer & Smith, 1996; Valencia *et al*, 1991). Although structural features of the basic switch are highly conserved, individual members of the superfamily communicate with specific proteins which may regulate the cycle and/or respond to regulation by different conformations of GTPase itself (Table 1.1). Within the families of G proteins, defined by amino acid sequence homology, the individual G proteins turn out to perform similar or closely related functions (Bourne, 1993; Hall, 1990). Structurally, members of each family resemble one another much more closely than they resemble members of other families. Thus the G_α proteins transduce signals between ligand-regulated receptors and effectors. The elongation and initiation factors (EFs & IFs) of ribosomal protein synthesis use their G protein cycle in a different way to catalyse and monitor assembly of complex macromolecular structure. The signal recognition particle family are components of the machinery that translocates polypeptides into the endoplasmic reticulum. The tubulin and cytoskeletal motor G proteins trigger the α and β tubulin polymerisation to form microtubules.

The small, monomeric subclass of G proteins are all structurally related to proteins encoded by Ras proto-oncogenes. This group has been subdivided into at least five families–Ras, Rab, Rho, Ran and the ADP-ribosylation factors (ARFs). They control a wide variety of cellular processes, serving to monitor and control much of the information flow in eukaryotic cells. Some members of this group also play an important part in control of the function of the cytoskeleton and are central to the regulation of traffic between various membrane-bound compartments in the cell. The GTP-GDP cycling is controlled by three different classes of effectors (guanine nucleotide exchange factors, GEFs, which switch the GTPase on; GTPase activating proteins, GAPs which switch it off, guanine nucleotide dissociation stimulators or inhibitors, GDSs or GDIs, which protect it from switching). A *P. falciparum* Ran/TC4 and its possible GEF, RCC1 have recently been studied in this laboratary (Ji *et al*, 1998; Sultan *et al*, 1994). Experiments on their biochemical and cellular functions in *P. falciparum* will be discussed in this thesis.

1.5 The control of gene expression in *P. falciparum*

Since this thesis is concerned with the genes and proteins involved in the control of malaria parasite development and cell cycle, it is necessary to review current knowledge on the activation of specific gene transcription in higher eukaryotic systems with a view to integrating what is known about *P. falciparum* gene activation into this body of knowledge.

1.5.1 General mechanisms of eukaryotic transcriptional activation

Table 1.1 GTPase families

Family	Example(s)
Heterotrimeric G-proteins (hormonal and sensory signalling)	G_s, the stimulatory regulator of adenylate cyclase. G_t (transducin) mediates retinal phototransduction. G_i, the inhibitory regulator of adenylate cyclase and the activator of phospholipases.
Initiation and elongation factors (ribosomalsynthesis)	EF-Tu proof-reads association between the mRNA protein codon and the anticodon of aa-tRNA.
SRP/SR family (translocates polypeptides into the ER)	*SRP54*, a GTPase component of the SRP.
Tubulins and cytoskeletal motor GTPases	α and β tubulin polymerise to form microtubules.
Monomeric (ras-related) small GTPases	
Ras family (proliferation, differentiation)	*ras* p21, oncogenic in many cells.
Rab family (vesicle traffic)	*Sec4* regulates late stage of secretion in *S. cerevisiae*.
ARF family (vesicle traffic)	*Sar1* regulates budding from the ER.
Ran family (protein and mRNA transport)	Ran/TC4 regulates protein and mRNA nuclear import and export.
Rho family (assembly and function of actin cytoskeleton)	*Rho*-A stimulates formation of actin stress fibres.

Activation of transcription requires a complex interplay between enhancer sequence-binding factors and components of the general transcription machinery. None of the genes discussed here have yet been identified in *P. falciparum* but it is likely that many functionally homologous genes exist and that the projected complete sequence of the *P. falciparum* genome will allow rapid progress to be made in this area. Co-factors are required for relaying specific activation signals to the RNA polymerase II transcription complex in order to achieve enhanced levels of mRNA synthesis. Two distinct types of DNA elements, core promoter sequences and arrayed enhancers are also important. The core promoter specifies the accurate initiation site of transcription whereas enhancers appear to direct the spatial and temporal patterns of gene-selective transcription. The core promoter is also thought to serve as an anchor point for tethering the general transcription machinery, a multiprotein complex comprising the basal transcription factors TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH and RNA PolII (reviewed in Roeder, 1996).

Three distinct classes of cofactors, upstream stimulatory activity factor (USA), mediator, and TBP-associated factors (TAFs), are required to potentiate gene-specific transcription directed by enhancer-binding factors. USA, which has the ability to bind DNA, was first identified as stimulating activator-dependent transcription *in vitro* and then subsequently found to either repress basal levels or potentiate activated levels of transcription. For instance, PC4 is able to bind single stranded DNA and repress basal transcription and enhance activated levels of transcription (Kaiser *et al*, 1995; Kretzschmar *et al*, 1994). The mediator is a multiprotein complex associated with RNA Polymerase II (RNAP II) which is able to potentiate transcriptional activation by certain activators *in vitro,* such as supressors of RNA polymerase B (SRBs), GAL11 and SIN4 (Chao *et al*, 1996; Maldonado *et al*, 1996). Interaction of GAL11 protein with the DNA-binding domain of the yeast activator GAL4 is able to stimulate transcription in yeast, by recruiting the mediator/RNA polyII complex to its promotor (Barberis *et al*, 1995; Farrell *et al*, 1996). But it is still unclear what mechanism allows the mediator to stimulate transcription.

In contrast to USA and mediators, TAFs are generally required only for activator-dependent transcription. They have been shown to directly interact with activators. *In vitro* reconstitution of TFIID complexes from recombinant TAF subunits helped to establish the functional relevance of this TAF/activator interaction. Several different classes of activators have been shown to support activation of transcription in vitro only in the presence of the TAFs targeted by these activators. The simultaneous interaction of different activators with distinct TAFs contributes to synergistic activation of transcription (Sauer *et al*, 1995; Sauer *et al*, 1995). Thus, TAFs do not work as generalized cofactors but rather as activation-domain-specific coactivators.

Interestingly, the SNF/SWI (sucrose non-fermenting / mating type switch) complex which is responsible for chromatin remodelling has also been found as a component of the mediator (Wilson *et al*, 1996). The SWI/SNF complex is highly conserved among different organisms and is itself comprised of more than 11 distinct polypeptides. The SNF2/SWI2 protein is one of the key components in the complex and appears to serve as its energy motor. A *P. falciparum* homolgue of SNF2/SWI2 has been cloned and characterised and is reported in this thesis. It constitutes one of the first tools to initiate studies in the control of *P. falciparum* transcription.

1.5.2 The role of chromatin in the control of transcription

In the 25 years since the discovery of nucleosomes (Kornberg, 1974) it has become clear that nucleosomes organised into chromatin, play the central role in transcriptional regulation (Lewin, 1994; Paranjape *et al*, 1994). In eukaryotes the nucleosome consists of 146 base pairs of DNA wrapped twice around an octamer of core histones. Although the nucleosomal organisation of chromatin is generally considered universal in eukaryotes, there is at least one report of a lower eukaryote, the dinoflagellate *Gyrodinium cohnii*, lacking nucleosomes (Bodansky *et al*, 1979). In *Entamoeba histolytica* and *Trypanosoma brucei*, nuclesome-like filaments have been reported. Significant structural differences may exist between the chromatin of these protozoan parasites and that of higher eukaryotes (Bender *et al*, 1992; Torres-Guerrero *et al*, 1991).

Little is known about the organisation of nuclear DNA structure in malaria parasites. Electron micrographs of chromatin prepared from *Plasmodium knowlesi* show a beaded fibre structure suggesting the presence of nucleosome filaments similar to higher eukaryotes (Wunderlich, 1980). However, these data are not entirely conclusive, since structures that appear similar to nucleosomes in electron micrographs can be detected even in the absence of histones (Sinha *et al*, 1982). Cary and collegues (1994) demonstrated that micrococcal nuclease digestion of chromatin in isolated *P. falciparum* nuclei yielded a ladder of DNA fragments characteristic of nucleosomal organisation. In addition, they identified a set of major proteins from *P. falciparum* nuclei that are similar in size and charge to histones of other eukaryotic cells. Furthermore, the histone H2A, H2B and H3 genes have been identified in *P. falciparum* (Bennett *et al*, 1995; Creedon *et al*, 1992). These histone genes of *P. falciparum* are closely related to histones of other lower eukaryotes such as yeast. A chromatin-associated protein also has been isloated from *P. berghei* and is thought to be involved in the dynamics of chromatin packaging (Birago *et al*, 1996).

In higher eukaryotic cell, the histone octamer is organised as a central tetramer of histones H3 and H4 flanked by two heterodimers of histones H2A and H2B. *In*

vivo and *in vitro* studies indicate that nucleosomes repress transcription by blocking the access of both the general transcription machinery and regulatory factors to their DNA sites. For example, one face of the DNA helix is completely concealed upon association with the histone core and the structure of DNA is changed dramaticly upon wrapping around the histone proteins. Several trans-acting factors, such as the yeast positive regulator PHO4, have been shown to be blocked from binding their preferred sequences when presented in a nucleosomal context.

There appear to be both replication-dependent and replication-independent pathways to cope with the nucleosome-mediated repression of transcription (Quinn & Peterson, 1997). In the replication-dependent mechanism transcription factors bind to the DNA template immediately following passage of the replication fork and prior to the assembly of the freshly replicated DNA into chromatin, thus preventing the formation of repressive chromatin structures. In the replication-independent pathway promoter regions are assembled into organised nucleosomal arrays following replication. Such repressive chromatin structures must then be directly disrupted and there is now evidence that a distinct class of transcriptional co-activators exist, whose primary role is to reconfigure chromatin structure (see below). Thus, the RNA polymerase transcription complex, associated activating factors, and chromatin are part of a single system for regulating eukaryotic gene expression.

1.5.3 Alleviating chromatin-mediated repression of transcription

A fundamental question in molecular biology is how the transcriptional machinery competes with the chromatin structure of eukaryotic DNA to activate gene expression. The specific binding of transcriptional activators disrupts nucleosomal arrays, suggesting that the primary steps leading to transcriptional initiation involve interaction between activators and chromatin. The affinity of transcription factors for nucleosomal DNA is determined by the location of recognition sequences with nucleosomes, and by the cooperative interactions of multiple proteins targeting binding sites contained within the same nucleosomes. Therefore, a network of linkages between chromatin and transcription have been proposed (Hampsey, 1997; Kadonaga, 1998) as depicted in Figure 1.6.

1.5.4 Chromatin remodeling machines

Except for the work reported in this thesis, none of the genes and proteins involved in chromatin remodelling have yet been identified in malaria parasites. However, although this literature is large and complex, it forms part of the background to the experimental work of this thesis and since it will ultimately need to be

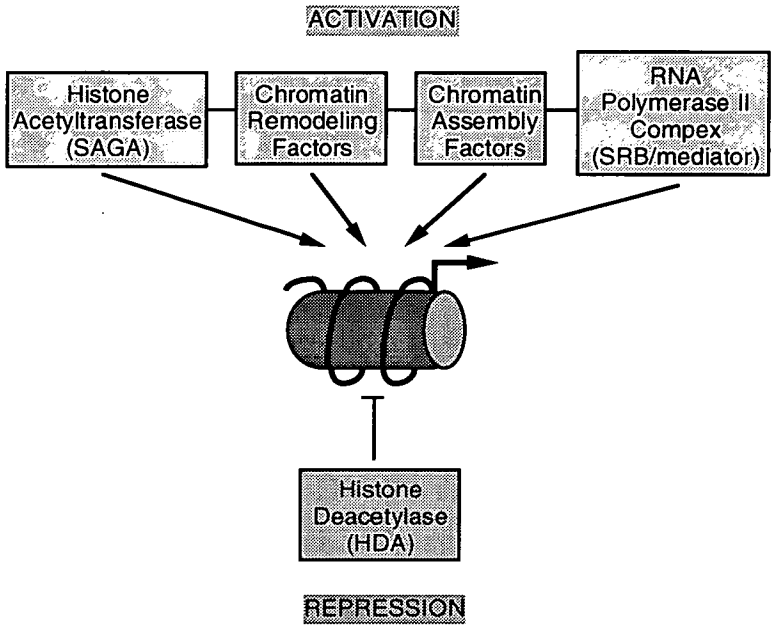


Figure 1.6 Schematic diagram of potential linkages between chromatin and chromatin modifying machines.

assimilated into our understanding of *P. falciparum* molecular biology, it is reviewed here in some detail.

Experiments in flies, yeast and human cells have uncovered a set of multi-protein chromatin remodeling machines, which are good candidates for promoter restructuring factors. These complexes include the yeast SWI/SNF complex, yeast RSC complex, *Drosophila* nucleosome remodelling factor (NURF), *Drosophila* chromatin assembly complex (CHRAC), *Drosophila* assembly chromatin factor (ACF), *Drosophila* braham (BRM) complex, as well as mammalian brm-related gene 1(BRG1) protein or the hbrm-associated complex (Cairns, 1998; Cairns *et al*, 1996; Cote *et al*, 1994; Ito *et al*, 1997; Kadonaga, 1998; Tsukiyama & Wu, 1995; Varga-Weisz *et al*, 1997; Wang *et al*, 1996). Each contain a subunit homologous to DNA-dependent ATPases (SWI2/SNF2 in the yeast SWI/SNF complex, STH1 in yeast RSC, ISWI in *Drosophila* NURF, CHRAC, and ACF, *Drosophila* BRM and mammalian BRG1 or hbrm). All complexes tested demonstrate ATP-dependent chromatin perturbation. The presence of this ATP-dependent subunit in the different factors suggests that it may carry out a related biochemical function in each complex. It has been postulated that the ATP-binding subunit might act as an ATP driven DNA-translocating motor that disrupts histone-DNA interactions. For instance, the purified yeast SWI/SNF complex was found to be able to induce perturbation of histone-DNA interactions in a mononucleosome and to facilitate binding of GAL4-AH to a mononucleosome. These transcription activating chromatin remodeling machines form part of the subject of this thesis and will be discussed further.

The RNA polymerase II (RNA pol II) holoenzyme composition controversy

As discussed previously, the RNA pol II holoenzyme has been considered to be an RNA pol II - SRB - mediator complex. An intriguing connection between this holoenzyme and the SWI/SNF complex was suggested by studies in which the SWI/SNF complex was found to be an integral and equimolar component of the holoenzyme complex (Kingston *et al*, 1996; Struhl, 1996; Wilson *et al*, 1996). But there are some arguments about this conclusion, because components of SWI/SNF complexes were not found in other preparations of mediator or holoenzyme (Cairns *et al*, 1996; Pan *et al*, 1997). In these experiments, binding of the yeast mediator/ holoenzyme to a promoter with a GAL11-PHO4Δ2 fusion protein was observed to induce nucleosome remodeling *in vivo* either in the presence or the absence of functionally active SWI/SNF complex (Gaudreau *et al*, 1997). In a recent report, an *Escherichia coli* SWI2/SNF2 homologue, RapA, has been identified which copurified with RNA pol II holoenzyme as a 1:1 complex. RapA forms a stable complex with RNA pol II and has ATPase activity which is stimulated by binding to RNA pol II

(Sukhodolets & Jin, 1998). These data indicate that the nature of the interaction between the RNA pol II holoenzyme and the SWI/SNF complex requires further clarification. The *P. falciparum* homologue of the RNA polymerase II has been cloned and characterised, could be used to further study interactions with *P. falciparum* SNF2/SWI2 homologue complexes.

Histone acetyltransferases (HATs) and deacetylases (HDAs)

The core histones, particularly H3 and H4, can be acetylated at the lysine residues in the N-terminal tails that extend outwardly from the globular core of the histone octamer. The function of this acetylation is not yet clear but it may be important on activating transcription. One theory is that charge neutralisation of the lysine residues reduces the interaction of the N-terminal tails with DNA and with nonhistone repressor proteins (Turner *et al*, 1992). A second is that the acetyl groups act as a signal for interaction of histone with other proteins, resulting in disruption of nucleosomes.

Recently, there has been progress in the identification of histone acetyltransferases (HATs) and histone deacetylases (HDAs) that can catalyze the acetylation and deacetylation of histone (reviewed in Grant *et al*, 1998; Hampsey, 1997; Kadonaga, 1998). Nuclear HATs include a basal transcription factor (TAFII250) and transcriptional co-activators (GCN5, p300/CBP, P/CAF, and SRC1). In addition, the RPD3-related HDAs appear to participate in transcriptional repression (Grunstein, 1997; Pazin & Kadonaga, 1997). These findings suggest that the HAT and HDA activities are an important component of transcriptional regulation. It remains to be determined whether it is the acetylation and deacetylation of core histone and/or other proteins that is responsible for the transcriptional effects.

The yeast GCN5 protein has been found in at least two distinct, multiprotein complexes that possess HAT function. For example, a 1.8 MDa complex, termed SAGA (SPT-ADA-GCN5-acetyltransferase), appears to contain SPT3, SPT7, SPT8, SPT20/ADA5, ADA2 ADA3 and GCN5. A combination between mutations in components of the SAGA complex and mutations in components of SWI/SNF complex is severely deleterious in yeast (Roberts & Winston, 1997). However, none of the known components of the SWI/SNF or SAGA complexes are essential for mitotic growth in yeast which implies that the complexes may function in a cooperative manner. None of these genes have yet been identified from *P. falciparum.*

Chromatin assembly factors

Chromatin assembly is a fundamental biological process that is required for the duplication and the maintenance of the genome (Ito *et al*, 1997; Kaufman, 1996; Roth

& Allis, 1996). In actively dividing cells, chromatin assembly is required to package the newly synthesised DNA into chromatin, whereas in long-lived quiesent cells such as mammalian neurons, chromatin assembly is needed to maintain the integrity of the genome upon turnover of the histones. The core chromatin assembly machinery that is required for the ATP-dependent assembly of periodic nucleosome arrays appears to comprise a protein complex termed ACF (ATP-utilising chromatin assembly and remodeling factor) (Ito *et al*, 1997). ACF was purified from *Drosophila* embryos and found to be a multiprotein factor that contains ISWI, a protein that is a member of the SWI2/SNF2 family. The finding that ACF contains an ISWI subunit and can function as a chromatin remodeling factor suggests links between chromatin assembly, chromatin remodeling, histone acetylation and deacetylation and transcription (Figure 1.7). A small subunit of CAF-1, a histone chaperone for chromatin assembly, has also been found to be identical to a protein that is associated wth a histone deacetylase and closely related to a protein that is associated with the HAT1 histone acetyltransferase (Kaufman *et al*, 1997). None have been identified in *P. falciparum*.

1.5.5 SNF/SWI complex and its related complexes

It could be argued that none of the above information is relevant to the malaria parasite which may organise its gene expression quite differently. However, data presented in the experimental chapters of this thesis indicates that this is unlikely to be the case. A *P. falciparum* homologue of one of the key components of a chromatin remodelling complex has been clearly identified (Ji *et al*, 1997). This protein is an ATPase associated with the SNF/SWI type chromatin remodelling complex.

The SWI/SNF complex is a macromolecular complex of ~2 MDa which is composed of at least 11 subunits (Cairns *et al*, 1994; Peterson *et al*, 1994). Several members of yeast SWI/SNF complex were identified for the purpose of identifying important regulators of RNA polymerase II in genetic experiments (Neigeborn & Carlson, 1984). Mutations in the SWI/SNF genes are supressed by mutations in genes coding for the nucleosome core histones and HMG1-like protein, SIN1/SPT2 (Kruger *et al*, 1995). Chromatin remodelling does not occur at SWI/SNF-dependent promoters in yeast *swi/snf* mutants (Hirshhorn *et al*, 1992). Purified yeast SWI/SNF complex mediates ATP-dependent chromatin restructuring *in vitro* (Cote *et al*, 1994). The human SWI/SNF complexes display biochemical activities similar to their yeast counterparts *in vitro* and have been linked to both transcriptional activation and cell cycle control (Kingston *et al*, 1996). In addition, a related complex from *Drosophila* has been identified which is important in the regulation of homeotic gene transcription. Furthermore, an abundant yeast complex RSC (remodels the structure of chromatin) which contains 15 subunits, and several components of RSC are strikingly similar to

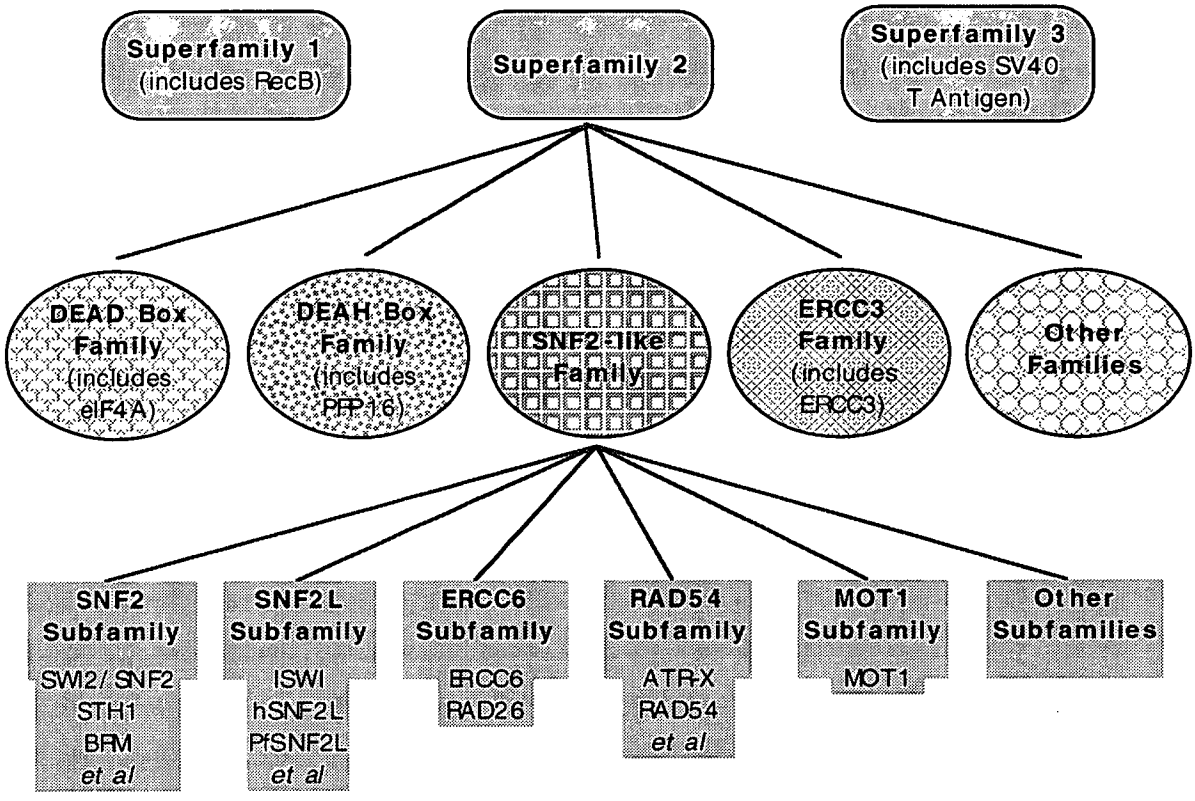


Figure 1.7 The SNF2-like family of proteins. (Adapted and modified from Kadonaga. *et al,* 1997).

components of SWI/SNF (Cairns *et al*, 1996; Treich & Carlson, 1997). Unlike SWI/SNF, RSC is essential for mitotic growth and although its genetic targets remain unidentified, the greater abundance of RSC (10^3-10^4 molecules/cell) suggests that this complex may be required for activities such as transcription or replication which involve the remodelling of large numbers of nucleosomes (Cairns *et al*, 1996; Treich *et al*, 1997).

There are three different remodelling complexes in Drosophila, NURF, CHRAC and ACF, that each contain an ISWI subunit. NURF can facilitate the ATP-dependent binding of transcription factors, e.g. GAGA factor and GAL4, to chromatin templates (Pazin *et al*, 1994; Tsukiyama *et al*, 1995). Purified NURF is composed of just four subunits and requires ATPase hydrolysis to function. Unlike SWI/SNF, the ATPase activity of NURF is stimulated more significantly by the presence of nucleosomes than by free DNA or core histones alone, suggesting that NURF is capable of detecting some unique feature of nucleosome organisation (Tsukiyama *et al*, 1994; Tsukiyama *et al*, 1995). CHRAC was identified on the basis of its ability to mobilise nucleosomes in a manner that allows enhanced access of a restriction enzyme to DNA packaged into chromatin. Extensive fractionation yielded CHRAC, a complex of five proteins. Unlike other known chromatin remodeling factors, CHRAC can also function during chromatin assembly. It uses ATP to convert irregular chromatin into a regular array of nucleosomes with even spacing. CHRAC combines enzymes that modulate nucleosome structure and DNA topology (Varga-Weisz *et al*, 1997). As mentioned previously ACF uses ATP to assemble properly spaced arrays and to facilitate activator binding to chromatin. It can mediate promoter-specific nucleosome reconfiguration by Gal4-VP16 in an ATP-dependent manner. These results suggest that ACF acts catalytically both in chromatin assembly and in the remodeling of nucleosomes that occurs during transcriptional activation (Ito *et al*, 1997). Surprisingly, NURF, CHRAC and ACF all contain the ISWI (initiation SWI2) ATPase (Tsukiyama *et al*, 1995), which was originally isolated by homology to SWI2 and encodes the ATPase component of the yeast SWI/SNF complex. ISWI and SWI2 are both members of the SWI2/SNF2 family of DNA-dependent ATPase. Interestingly, several large complexes each containing the ISWI subfamily member hSNF2L have been found in human cells (Wang *et al*, 1996; Wang *et al*, 1996). They could represent human counterparts of NURF, CHRAC and ACF and it is possible that all these large intra-nuclear complexes exist within the nucleus of all eukaryotic cells including *P. falciparum*.

1.5.6 SNF2/SWI2 and related proteins

SWI2/SNF2 proteins are a gene family whose members probably have different functions (Table 1.2). The SWI2/SNF2 proteins contain the seven conserved motifs, which could bind the nucleotide triphosphate (NTP), similar to those found in many DNA or RNA helicase protein families (Bork & Koonin, 1993). Comparative amino acid sequence analysis revealed that this group of helicases and related proteins comprises at least three superfamilies, which in turn can be further divided into families and subfamilies (Bork *et al*, 1993; Eisen *et al*, 1995; Gorbalenya & Koonin, 1993). The SNF2 family has been assigned to helicase superfamily 2, which also includes the ERCC3, RAD3, PRIA, eIF4A and PRP16 protein families. The conserved structure of these proteins may reflect a similar mechanism of action (Eisen *et al*, 1995).

Members of the SNF2-like family exhibit considerable range of biological functions. These activities include gene-specific transcriptional regulation (SNF2 subfamily), transcriptional repression (MOT1 subfamily), destabilisation of recomposed nucleosomes (SNF2 and SNF2L subfamilies), transcription-coupled repair (ERCC6 subfamily), nucleotide excision repair of non-transcribed regions of the genome (RAD16 subfamily), recombination repair (RAD54 subfamily), chromatin modification and negative regulation of cellular growth (CHD1 subfamily), developmental regulation and UV protection (ETL1 subfamily) and chromosome segregation (lodestar) (Eisen *et al*, 1995). In spite of the presence of the conserved helicase-like motif in these proteins, however, helicase activity has not yet been detected in any SNF2-like family member. The sequences within the proposed subfamilies are more related to each other than to any other characterised proteins, including other members of the SNF2-like family (Eisen *et al*, 1995). SNF2-like family members are also involved in many human diseases (Ellis, 1997). Human ERCC6/CSB mutations result in Cockayne's syndrome, which is characterised by growth deficiency, skeletal abnormalities, photosensitivity, mental degeneration, characteristic facies, cataracts and dental cavities (Troelstra *et al*, 1992). In addition, mutations in the human ATR-X gene, assigned in the RAD54 subfamily, cause a combined α-thalassemia, mental retardation syndrome, developmental delay and genital abnormalities (Gibbons *et al*, 1995).

The available evidence on SWI2/SNF2, STH1, NOT1 and ISWI suggest, as proposed earlier, that these factors may function as ATP-driven motors that translocate along DNA and destabilise protein-DNA interactions. The movement of these proteins along DNA is likely to be similar to the ATP-dependent translocation of helicases along nucleic acids. A model proposed by Pazin and Kadonaga (Pazin & Kadonaga, 1997) suggests that a DNA-translocating protein uses the energy derived from hydrolysis of ATP to traverse a nucleosome in a wave-like manner that results in only

Table 1.2 SNF2 Families

Proposed subfamily	protein	No. of amino acid	Species	Possible functions/comments
	SNF2	1703	*S.cerevisiae*	Transcription activation. DNA-dependent ATPase. Alters chromatin structure
	STH1	1359	*S.cerevisiae*	Cell cycle control. Required for normal growth
SNF2	BRM	1638	*D.melanogaster*	Transcription activation of homeotic genes
	BRG1	1022	Mouse	Binds retinoblastoma protein
	BRG1	1613	Human	Transcription co-activation with hormone receptors
	hBRM1	1586	Human	Transcription co-activation with hormone receptors
	PfSN2L	1422	*P.falciparum*	?
	SNF2L	976	Human	?
SN2L	ISWI	1027	*D.melanogaster*	DNA remodeling
	F37A4.8	971	*C.elegans*	?
	YB95	1143	*S.cerevisiae*	?
	XNP	?	Human	Causes ATR-X syndrome. Interacts with EZH2 protein
CHD1	CHD1	940	Mouse	Binds DNA, modify chromatin structure
	SYGP4	1468	*S.cerevisiae*	?
	Hrp1	1373	*S. pombe*	A negative regulator of cellular growth
ETL1	ETL1	1136	Mouse	Expressed very early in development. Concentrated in CNS and epithelium
	FUN30	1131	*S.cerevisiae*	Mutants have increased UV resistance
MOT1	MOT1	1867	*S.cerevisiae*	Transcription repression. Removes TBP from DNA
	TAF-172	1548	Human	Transcription repression. Removes TBP from DNA
ERCC6	RAD26	1085	*S.cerevisiae*	Transcription-coupled repair
	ERCC6	1493	Human	Transcription-coupled repair. Defective in Cockayne's syndrome group B
	RAD54	898	*S.cerevisiae*	Recombination repair
	DNRPPX	852	*S.pombe*	?
RAD54	YB53	958	*S.cerevisiae*	?
	NUCPRO	1298	Human	?
	NUCPRO	996	Mouse	?
	RAD16	790	*S.cerevisiae*	Nucleotide excision repair of silent genes
RAD16	RAD5	1169	*S.cerevisiae*	Post-replication repair. GT repeats more stable in mutants
	RAD8	1133	*S.pombe*	Mutants have increased sensitivity to UV and gamma irradiation
	HIP116A /Zbu1	1009	Human	DNA-dependent ATPase. Binds HIV and SPH motifs of SV40 enhancer. Aberrantly regulated in malignant cells
	P113	1017	Mouse	Activates the transcrition of the PAI-1 gene specificly
None	NPHCG42	506	*A.californica*	Viral encoded protein
None	lodestar	1061	*D.melanogaster*	Mutants have excessive chromosome breakage and tangling in mitosis
None	HepA	968	*E.coli*	Induced by DNA damage
None	RapA	990	*E.coli*	Associated with RNA Pol

a partial disruption of the nucleosome at any particular point. This sort of process could account for the ATP-dependent destabilisation of nucleosome structure and facilitated transcription factor binding that has been shown *in vitro* with SWI/SNF complex, RSC, ACF, or NURF (Pazin *et al*, 1994; Tsukiyama *et al*, 1995; Varga-Weisz *et al*, 1997). Also, the translocation of these proteins could facilitate changes in nucleosome positioning, as seen with nucleosome arrays *in vitro*. Hence, this ATP-driven DNA translocation mechanism seems to be consistent with the available data on the properties of the SNF2-like family members.

1.6 The aims of the research presented in this thesis

The interest of this laboratory in signal transduction pathways led to a series of preliminary experiments to study the possible role of the heterotrimeric G protein signalling pathways which are involved in the regulation of the cell cycle of the malaria parasite *Plasmodium falciparum*. The first task was attempt to isolate a G_α protein gene in order to initiate the study. At the same time, other work to complete the cloning of the malaria RCC1-like gene and further characterise its cellular functions was carried out.

In the course of experiments to clone genes for the G_α protein a gene for a *P. falciparum* SNF2-like protein, a putative global transcriptional regulator involved in the chromatin remodelling was identified, termed *PfSNF2L*. These transcription factors play a very important role in the regulation of cell cycle and are linked with many signalling pathways. Many reports recently published also indicate that the SNF2-like family members play a key role in the regulation of transcription and are involved in many cellular activities. Therefore, the primary aims of the project moved towards the cloning and characterise the *PfSNF2L* gene at the molecular and biochemical level. In order to provide a more biological and less purely gene-oriented basis for studying *P. falciparum* cell cycle regulation, experiments studying the effects of heat shock and drug inhibition on cell cycle progression were carried out. Finally, some computational analysis of *P. falciparum* regulating sequences was carried out, a by-product of the work on transcription and gene activation.

CHAPTER 2
MATERIALS AND METHODS

2.1 Parasite, erythrocyte and bacterial sources

Plasmodium falciparum strains 3D7A, Dd2, K1 and HB3B were used in this work. These clones are kept in the WHO Registry of Standard Strains of Malaria Parasites in the Centre for Parasite Biology, ICAPB, University of Edinburgh.

Fresh human whole blood, group O, Rh^+, in citrate phosphate dextrose-adenine pack used for the parasite culture was supplied from Edinburgh and South-East Scotland Blood Transfusion Service. The blood was washed and centrifuged at 1500g three times in incomplete RPMI medium to remove leukocytes and the preservatives. The 'buffy coat' of the white blood cells was removed from the red blood cell (RBC) pellet. The RBCs were resuspended in complete medium to a final 50% haematocrit and kept at 4°C for up to one week (Hyde, 1993).

The bacterial strains used in this study are listed in Table 2.1

Table 2.1 Bacterial strains

Strain	Genotype	Remark
BL21	*E. coli* B F^- *dcm ompT hsdS* (r_B^-, m_B^-) *gal*	For high-level expression of genes cloned into expression vectors. (Studier and Moffatt, 1986).
BL21 (DE3)	*E. coli* B F^- *dcm ompT hsdS* (r_B^-, m_B^-) *gal* λ(DE3)	For high-level expression of genes cloned into expression vectors containing bacteriophage T7 promoter. (Studier and Moffatt, 1986).
DH5α	*supE44 ΔlacU169 (ø80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	For the recombinant DNA manipulation with blue/white screening on X-gal-IPTG plate. (Hanahan, 1983).
INVαF'	F' *endA1 recA1 hsdR17* (r_k^-,m_k^+)*supE44 thi -1 gyrA96 relA1* (*ø80 lacZΔM15*) Δ(*lacZ*YA-argF)U169 *deo*R λ^-	For the recombinant DNA manipulation with blue/white screening on X-gal plate. (Invitrogen Co.).
XL1-Blue	*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac*$^-$ F' [*proAB*$^+$ *lacI*q*lacZΔM15*Tn*10(tet*^{r}*)*]	For the recombinant DNA manipulation with blue/white screening on X-gal-IPTG plate. (Bullock *et al*, 1987).

2.2 cDNA libraries

The *Plasmodium falciparum* asexual cDNA library used in this work were made from rings and late trophozoites in plasmid pJEF14 (a gift from Dr. A. Craig,

Oxford university). The RNA was prepared from ring and trophozoite stage of line ITO4 parasites.

2.3 Oligonucleotide primers used in this study

Oligonucleotide primers used in this study were synthesised by the Oswel DNA Service, Department of Chemistry, University of Southampton. The sequence of the primers are given in Table 2.2.

Table 2.2 Oligonucleotide primers used in this course of studies

Name	Sequence
Sequencing primers	
-40M13,(+)[a]	5'-GTTTTCCCAGTCACGAC-3'
M13 reverse,(-)	5'-CAGGAAACAGCTATGAC-3'
pJFE-14,(+)	5'-TCTAGAGATCCCTCGACCTC-3'
pJFE-14,(-)	5'-CGCGGACGTCCAGCGCCGGC-3'
T7 promoter	5'-TAATACGACTCACTATAGGG-3'
PGEX,(+)	5'-GGGCTGGCAAGCCACGTTTGGTG-3'
PGEX,(-)	5'-CCGGGAGCTGCATGTGTCAGAGG-3'
G-protein degenerate primers	
gp1,(+)	5'-AA(A/G)TGGAT(A/C/T)CA(C/T)TG(C/T)TT-3'
gp2,(-)	5'-(A/G)TC(A/G/T)AT(C/T)TT(A/G)TT(A/G/T/C)AA(A/G)AA-3'
gp3,(-)	5'-(A/G)TC(A/G/T)AT(C/T)TT(A/G)TT(T/C)AA(A/G)AA-3'
g1f1,(+)	5'-GG(A/T)GC(A/T/C)GG(A/T)GA(G/A)TC(A/T/C)GG(A/T)AA-3'
g1f2,(+)	5'-GG(A/T)GC(A/T/C)GG(A/T)GA(G/A)AG(T/C)GG(A/T)AA-3'
g3f1,(+)	5'-GA(T/C)G(T/C)(A/T/C)GG(A/T)GG(A/T)CA(A/G)(A/C)G-3'
g3r1,(-)	5'-C(T/G)(T/C)TG(A/T)CC(A/T)CC(A/T/G)(A/G)C(A/G)TC-3'
g4r1,(-)	5'-(A/G)TC(T/C)TT(T/C)TT(A/G)TT(T/C)AA(A/G)AA-3'
g4r2,(-)	5'-(A/G)TC(T/C)TT(T/C)TT(A/G)TT(A/G/T)AG(A/G)AA-3'
g5r1,(-)	5'-(G/A/T/C)GT(A/G)TC(G/A/T/C)GT(A/T/G)GC(A/G)CA(G/A/T/C)GT 3'
PfSNF2L primers	
S6161,(+)	5'-ATTCTTCTGCTTGTGGTTAGG-3'
S6162,(+)	5'-CGAAGTAGAGCAATCATTACCACC-3'
S6163,(-)	5'-GCTTCATCTTCTTCATTTGCATCC-3'
S6164,(-)	5'-GCTCGATCCATTGCTTGTATATCC-3'
S6165,(-)	5'-CCTTCTAAAATAGTTTCTTCCCC-3'
S6166,(+)	5'-GGATATACAAGCAATGGATCGAGC-3'
T0499,(+)	5'-GGGGAAGAAACTATTTTAGAAGG-3'
T0500,(-)	5'-GGTGGTAATGATTGCTCTACTTCG-3'

T5777,(-) 5'-GTGCTTCTATTAGTACATTG-3'

T5778,(-) 5'-ACCTAACCACAAGCAGAAGAA-3'

T6118,(+) 5'-CAATGTACTAATAGAAGCAC-3'

V1425,(+) 5'-GATGCAAATGAAGAAGATGAAG-3'

V1426,(+) 5'-AATTGTTACACACAAGTAGC-3'

V1427,(-) 5'-TGCTTTTATATCTTCCATCG-3'

V4436,(+) 5'-AACAGAAGAGGTGAAAACTG-3'

V4437,(+) 5'-ATGTTGAATATCGATGTAGG-3'

V4438,(-) 5'-AACCTACATCGATATTCAAC-3'

V4439,(-) 5'-ATTTGGTACAAAGTTATCAG-3'

PfSNF2L fusion protein primers

V8616,(+) 5'-CGTGGATCCCTCTCAGAAGGAGTTGATAG-3'

V8617,(-) 5'-CGCTCGAGCTAAGGGATAAAGTCTTGTG-3'

W7289,(+) 5'-AATTGTAGAGAGAGCAGCAAAGAAG-3'

W7290,(-) 5'-CTTCTAAAATAGTTTCTTCCCCTGA-3'

W7291,(+) 5'-ATGGAGAGCTGAAGCTAGAGGTGGT-3'

W7292,(-) 5'-TGTTGATGTTATTTAGGTAGGTAGT-3'

W7293,(+) 5'-TGTAGAAAAGGATATAGCTAGTGTA-3'

W7294,(-) 5'-CATATCTCAAATTTGATGGTAGTTG-3'

DDRT-PCR PRIMERS

M4854,(-) 5'-TTTTTTTTTTTTAG-3'

M4855,(-) 5'-TTTTTTTTTTTTAT-3'

M4856,(-) 5'-TTTTTTTTTTTTAC-3'

M4857,(-) 5'-TTTTTTTTTTTTCC-3'

M4858,(-) 5'-TTTTTTTTTTTTGG-3'

M4859,(-) 5'-TTTTTTTTTTTTGT-3'

M4860,(-) 5'-TTTTTTTTTTTTGC-3'

M4861,(-) 5'-TTTTTTTTTTTTAA-3'

M4862,(-) 5'-TTTTTTTTTTTTCA-3'

M4863,(-) 5'-TTTTTTTTTTTTCG-3'

M4864,(-) 5'-TTTTTTTTTTTTCT-3'

M4865,(-) 5'-TTTTTTTTTTTTGA-3'

N0898,(+) 5'-ACAGGTATGG-3'

N0899,(+) 5'-CCAATGTTAC-3'

N0900,(+) 5'-TACCTAAGCG-3'

N0901,(+) 5'-CTATGGTTTG-3'

---

a. (+): sense strand; , (-): anti-sense strand

2.4 Chemicals and equipments

Analytical grade standard laboratory chemicals and solvents were purchased from British Drug House (BDH) plc., and Sigma Chemical Co. Ltd., UK. Radioactive nucleotides were obtained by Amersham International plc. The sources of other materials are given in the text.

Centrifugation was either carried out in a bench top centrifuge Biofuge 13 (Heraeus, Germany) or Centrifuge 5410 (Eppendorf, Germany). Higher speed centrifugations were performed in Sorvall-RC5C-refrigerated centrifuge (DuPont, USA). The models and brands of other equipment are given in the text.

2.5 General stock solutions, buffers and media

All aqueous solutions were prepared in de-ionised glass-distilled water. pH values of solutions were measured with a pH meter model PW 9410 (Philips).

Denhart's solution (50x)	1% w/v Ficoll, 1% w/v BSA fraction V, 1% w/v polyvinylpyrrolidone.
DNA Loading Buffer	50 mM Na_2 EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol, 40% (w/v) sucrose in H_2O and stored a small aliquots at 4°C.
IPTG	Stock solution (0.1M): dissolving 0.238 g IPTG in 10 ml dH_2O, filter sterilised and stored in 1ml aliquots at -20°C.
Luria-Bertani medium (LB)	1% Bacto-tryptone, 0.5% Bacto-yeast extract, 1% NaCl adjusted to pH 7.2 with 10 N NaOH in dH_2O.
LB agar	LB medium with 1.5% agar
LB-amp.	LB supplemented with ampicillin to 50 µg/ml
M9 medium	5x M9 salts (per litre): 2.5 g NaCl, 64 g $Na_2HPO_4 \cdot 7H2O$, 15 g KH_2PO_4, 5 g NH_4Cl in dH_2O to 1 litre and sterilised by filtration through 0.22 µM Nalgene filter. Working medium : 200 ml 5x M9 salts, 20 ml 20% glucose and 780 ml sterile dH2O with or without 15 g bacto-agar.
MOPS	10x stock solution : 0.2 M 3-[N-Morpholino]- propane-

	sulphonic acid, 0.5 M Na acetate pH7.0, 0.01 M Na_2EDTA, H_2O to 1 litre. 1x working solution : (0.02 M morpholino-sulphonic acid, 0.05M Na acetate, 0.001M EDTA)
PBS	10x stock solution : 80 g NaCl, 2g KCl, 11.5 g $Na_2HPO_4 \cdot 7H2O$, 2 g KH_2PO_4 to 1 litre. working solution : 137 mM NaCl, 2.7mM KCl, 4.3mM $Na_2HPO_4 \cdot 7H_2O$, 1.4 mM KH_2PO_4, pH~7.3.
Protease inhibitors	Leupeptin 50 µg/ml (in dH_2O), Chymostatin 10 µg/ml (in acetic acid), Pepstatin 10 µg/ml in methanol, PMSF 200 mM in propanol, idoactamide 20 µg/ml in dH_2O, EDTA 500 mM, EGTA 500 mM.
RPMI medium (incomplete)	10.4 g of RPMI 1640, 5.94 g of HEPES, dissolved in 960 ml of distil dH_2O. Filtered through 0.22 µM Nalgene filter and stored at 4°C for up to four weeks.
RPMI medium (complete)	Was prepared by addition of 42 ml of Sodium bicarbonate, 50 mg/ml gentamycin, and 40 ml heat inactivated human serum to 600 ml of incomplete RPMI medium.
SDS	10% stock solution: 100 g is dissolved in 900 ml H_2O, heated to 68°C and the pH was adjusted to pH7.2 by HCl, and made up to 1 litre with dH_2O.
2x SDS-loading buffer (Studier buffer)	Core buffer: 125 mM Tris pH6.8, 4% SDS, 20% glycerol, 0.004% bromophenol blue in H_2O. Before use, 800 µl of core buffer was mixed with 200 µl of 1 M Dithiothreitol
SOC medium	2% Bacto-tryptone, 0.5% bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM Mg_2Cl, 10 mM Mg_2SO_4, 20 mM glucose in 1 litre dH_2O.
SSC	20x stock solution: 175g NaCl, 88 g Na_3 citrate$\cdot 2H_2O$, adjusted to pH 7.0 and made up to 1 litre with dH_2O. 2x working solution : (0.3 M NaCl, 0.03 M Na_3Citrate$\cdot 2H_2O$).

TAE	50x stock solution: 242 g Tris base, 37.2 g $Na_2EDTA{\cdot}2H_2O$ 57.1 ml glacial acetic acid, dH_2O to 1 litre. This solution was diluted 50 times to give 1x working solution: 40mM tris acetate, 2mM EDTA.
TBE	10x stock solution : 108g Tris base, 55g boric acid, 40ml 0.5M EDTA pH 8.0, dH_2O to 1 litre. This solution was diluted 10 times to give 1x working solution : 89mM Tris base, 89mM boric acid, 2mM EDTA.
TE buffer (1x)	10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0
Transformation Buffer I (TFB I) (amounts/litre)	30 mM potassium acetate (30 ml of a 1M stock pH7.5), 100 mM RbCl (12 g), 50 mM $MnCl_2{\cdot}2H_2O$ (9.9 g), 10 mM $CaCl_2{\cdot}2H_2O$ (1.5 g), 15% w/v glycerol (150 g), pH adjusted to 5.8 with 0.2 M acetic acid , Solution sterilised by filtration.
Transformation Buffer II (TBF II)	10 mM MOPS (20 ml of a 0.5 M stock pH6.8), 10 mM RbCl (1.2 g), 75 mM $CaCl_2{\cdot}2H_2O$ (11 g), 15% (w/v) glycerol, adjusted with KOH to pH6.8 and sterilised by filtration.
X-Gal	Fresh dry LB-amp. plates, spread with 25µl X-Gal (40mg/ml stock solution in dimethyl-formamide)

Percoll gradient:

Solution for gradient layers of 52.5%, 45% and 30% Percoll were made up in incomplete RPMI medium as follows: volume of 10x conc. RPMI medium was added to 9 volumes of stock Percoll (Sigma) to make an isotonic 90% Percoll solution. Then:-

a) to 2 volumes of 90% Percoll, 1 volume of 1xRPMI was added to make 60% Percoll.

b) to 1 volume of 60% Percoll, 1 volume of 1xRPMI was added to make 30% Percoll.

c) to 1 volume of 60% Percoll, 1 volume of the 30% Percoll was added to make 45% Percoll gradient.

d) to 1 volume of 60% Percoll, 1 volume of 45% Percoll was added to make 52.5% Percoll gradient.

2.6 *In vitro* culture of *P. falciparum*.

2.6.1 Culture of asexual parasites

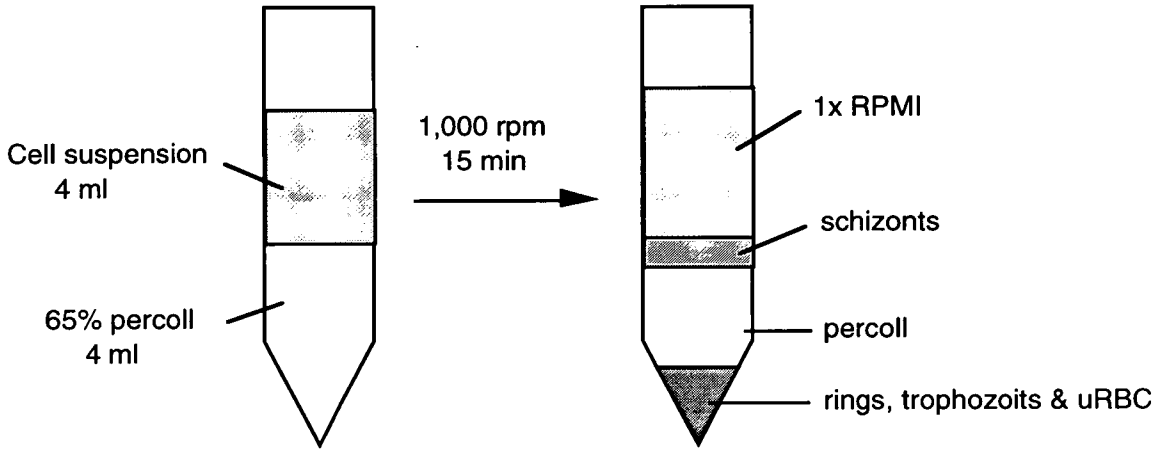
P. falciparum parasites were grown under standard conditions (Hyde, 1993; Trager & Jensen, 1976; Zolg *et al*, 1982). Parasites were maintained in 5 ml complete RPMI medium (Sigma) containing washed O^+ red blood cells (RBCs) at 5% hematocrit in 75 ml culture flasks. Cultures were typically initiated at a parasitaemia of 0.5%. The cultures were gassed with a mixture of 1% O_2, 3% CO_2, and 96% N_2 and incubated at 37°C. The medium was replaced with fresh pre-warmed medium daily. A drop of blood was taken by a Pasteur pipette and smeared on a slide. The smear was air dried, fixed by 100% methanol, stained with 5% Giemsa's solution (in Sorensen's buffer, pH7.2) for 40-45 min and the parasitaemia measured by microscopic examination. The cultures were diluted with fresh RBCs, when the parasitaemia arrived at 10-15%.

2.6.2 Synchronisation of the parasites

Three main techniques are available to obtain synchrony of different parasite stages: 1. treatment with sorbitol (Lambros & Vanderberg, 1979). 2. density gradient separation (Aley *et al*, 1986; Dulzewski *et al*, 1984). 3. exposure to high temperature (Kwiatkowski, 1989). In this study, synchronised cultures were obtained using the sorbitol treatment to destroy large parasites (trophozoites and schizonts) in erythrocytes, resulting in a synchronously growing population, starting from the remaining ring-forms.

Cultures with low parasitaemia (about 1%) and a high proportion of ring-forms were used as starting material. Cultures were centrifuged in IEC Centra-MP4R rotor 215 at 1000x g for 5 minutes, and the supernatant was discarded. The cells were then resuspended in 5% sorbitol solution in dH_2O (10% [v/v]), and allowed to stand at room temperature for 5 minutes. The mixture was then centrifuged as above and the cells resuspended in complete RPMI medium. The centrifugation was repeated once more, to remove any residual sorbitol, and the cells were finally resuspended in RPMI medium at 5% hematocrit (method 2.6.1). After 24 hours of culture, a blood smear was examined by Giemsa staining and microscopy. The culture should not contain ring-forms at this stage, but mainly trophozoites and maturing schizonts. Six hours later, another blood smear was examined and if a new generation of small rings was seen, a second sorbitol treatment was carried out as above to remove remaining trophozoites/schizonts. The culture was then monitored over 48 hours to ensure that it was synchronous. If not, more sorbitol treatments were carried out at times when the only ring-forms present were small. When cultures were fully synchronous, they were harvested for the desired experiments.

The sorbitol synchronised parasites could be further synchronised by percoll gradient. 5% mature schizonts (segmented) in 4 ml 1x incomplete RPMI medium was loaded on the top of 4 ml of 65% percoll.



The schizont layer was collected and washed 3 times with 1x incomplete RPMI medium then precipitated again. The mature schizont infected RBCs were resuspended in complete RPMI medium with fresh RBC to 2.5% hematocrit and done smear and giemsa stain again immediately. Another sorbitol treatment could be undertaken at this point to improve synchronisation of the parasites.

2.6.3 Gametocyte culture

Cultures were set up in 15 ml complete RPMI medium at 6% haematocrit and started at 0.5 to 0.7% parasitaemia in 75 mls flasks (Hyde, 1993; Ifediba & Vandenberg, 1981). Parasitised RBCs from asexual stock cultures at parasitaemia about 4% to 6% were used to ensure a high rate of commitment to sexual development. Cultures were gassed and maintained at 37°C. The medium was replaced daily with 15 ml of 37°C pre-warmed complete medium. 4 to 5 days after the start of each culturing, when a high parasitaemia had been reached, samples of the parasites were smeared, stained, and examined microscopically for morphological changes associated with gametocyte development, such as slightly triangular ring forms and early trophozoites in shape. When such signs were observed, the hematocrit was reduced to 3.6% by substituting with 25 ml of complete RPMI medium instead of 15 ml. The gametocyte cultures were then maintained with daily medium changes of 25 ml until the desired stage of maturity was reached, according to the Hawking *et al* classification (Hawking *et al*, 1971). Mature gametocytes were usually harvested 10 to 17 days after the starting of the cultures.

2.6.4 Purification of gametocytes

A high level gametocyte producing line, such as 3D7A, grown manually in 75 ml culture flask (as described in Methods 2.6.3), will produce 2-5 x 10^7 stage IV or V gametocytes. Gametocyte cultures were harvested by centrifugation at 2,300 rpm for 7 min. The cells were resuspended to a final haematocrit of about 20% in incomplete RPMI medium. Percoll gradients were prepared from the bottom up in 15 ml Corex tube (First 3 ml of 30% percoll, then 3 ml of 45% percoll, then 3 ml of 52.5% percoll, as described Methods 2.5). Finally 5 ml of harvested cells were overlaid on the top of the percoll gradient. The tube was then spun at 10,000 rpm in a 22°C Sorval centrifuge (RC's) for 10 min. Mostly mature gametocytes were recovered from the interface between the 30% and 45% percoll layers, and immature gametocytes from the interface between the 45% and 52.5% layers. The recovered cells from each interface were resuspended in at least 5 volumes incomplete RPMI medium and washed once to move percoll. The cells were condensed in 1 ml incomplete RPMI for counting gametocyte numbers and for further total RNA isolation and Immunoflourescence assay (IFA).

2.7 Preparation of parasite DNA, RNA and plasminds

Unless otherwise noted, the molecular cloning techniques used were carried out as described in Sambrook *et al.* (1989).

2.7.1 Preparation of *P. falciparum* genomic DNA

5 ml of asexual stage culture at 5-10% parasitaemia were normally used for the genomic DNA preparation. Parasitised erythrocytes were pelleted by centrifugation at 4,000 rpm for 10 minutes. The erythrocyte pellet was resuspended in 1 ml of 0.1% saponin in phosphate-buffered saline (PBS), and incubated at room temperature for 5 minutes. The erythrocyte lysate was then transferred to a microcentrifuge tube and spun in a microcentrifuge for 5 minutes (10,000 rpm). The parasite pellet was collected and washed four times with 1 ml of ice-cold PBS to remove the residual ghosts of RBCs. The parasite pellet was then lysed by addition of 600 µl of 10mM Tris-HCl pH 7.6, 50mM EDTA pH 8.0, 0.1% SDS, 1mg/ml proteinase K and incubated overnight at 50°C. The lysate was extracted twice with phenol/chloroform, then once with chloroform. DNA was then precipitated by adding 0.1 volume of 5M sodium Acetate and 1 volume of isopropanol, and left to stand at room temperature for 15 minutes. DNA was then collected by centrifugation, washed with 70% ethanol, dried and resuspended in 100 µl of Tris buffer (10mM Tris-HCl, pH8.0).

2.7.2 Plasmid DNA preparations

Mini-preparation of plasmid DNA :

Plasmid DNA was prepared either by the SDS-Alkaline Extraction method as described by Sambrook *et al.* (1989) or QIAGEN Plasmid Mini Kit or QIAprep Spin Miniprep Kit (QIAGEN Inc.), according manufacturer's instructions.

For automated sequencing, all the plasmid DNAs were isolated from *E. coli* XL1-Blue host by using the QIAGEN Plasmid Mini Kit (QIAGEN Inc.) with the following modifications. Single colony was inoculated into 5 ml of LB broth containing 50 μg/ml ampicillin and grown overnight at 37°C with shaking. The bacteria from 3-5 ml of the overnight culture were collected by centrifugation at 13,000 rpm for 2 minutes and resuspended in 300 μl of Buffer P1 (10 mM EDTA, 50 mM Tris-Cl pH 8.0, 100 μg/ml RNase A, pH 8.0). The cells were then lysed by the addition of 300 μl of Buffer P2 (0.2N NaOH, 1% SDS) and mixed gently by inversion of the tube several times. This was neutralised with 300 μl of chilled Buffer P3 (3.0 M potassium acetate pH 5.5), mixed immediately, and incubated on ice for 10 min. In order to pellet precipitated bacterial debris, the solution was centrifuged at 13,000 rpm for 15 minutes. The supernatant was then transferred to a QIAGEN-tip 20 pre-equilibrated by Buffer QBT and it was allowed to flow out by gravity. The QIAGEN-tip 20 was washed with 4 x 1 ml of Buffer QC. Then DNA was eluted with 800 μl 0f Buffer QF. The eluted DNA was precipitated with 0.7 volumes of isopropanol and centrifugated immediately at 13,000 rpm for 30 min. The DNA pellet was then washed with 70% ethanol once, air dried, resuspended in 15 μl of ddH_2O, and stored at -20°C .

Maxi-prep of plasmid DNA :

Large scale preparations of plasmid DNA were prepared using QIAfiter Plasmid Midi Kit, following the manufacturer's instructions. A single colony was picked into 250 ml of LB-amp broth and grown overnight at 37°C with vigorous shaking. The bacteria were harvested by centrifugation at 3000 rpm for 10 minutes at 4°C and the pellet resuspended in 10 ml of Buffer P1 (100 μg/ml RNase A, 50 mM Tris-HCL, 10 mM EDTA, pH 8.0). The cells were lysed by adding 10 ml of Buffer P2 (200 mM NaOH, 1% SDS). The lysate was mixed gently and incubated at room temperature for 5 minutes. After the addition of 10 ml of Buffer P3 (3 M potassium acetate pH 5.5) and mixing, the solution was then poured into the barrel of the QIAGEN Cartridge, incubated at room temperature for 10 min. During this time, a QIAGEN-tip 500 was equilibrated by applying 10 ml of Buffer QBT and the column was allowed to empty by gravity flow. The cap was removed from the QIAfilter outlet nozzle, the plunger was gently inserted into the QIAfilter Maxi Cartridge, and the cell lysate was filtered into the previously equilibrated QIAGEN-tip 500. The cleared lysate

was allowed to enter the resin by gravity flow, and then the QIAGEN-tip was washed with 2 x 30 ml of Buffer QC. The DNA was eluted with 15 ml of Buffer QF, precipitated by 10.5 ml of isopropanol, and centrifuged immediately at 15,000 x g (SS-34 Sorvall rotor) for 30 minutes at 4°C. The DNA pellet was then washed with 5 ml of 70% ethanol, air dried, resuspended in 0.25 ml of Tris buffer (10 mM Tris-Cl, pH 8.5), and stored in -20°C.

2.7.3 Purification of DNA fragments from agarose gels

DNA fragments obtained after PCR amplification and restriction digestion were purified by one of the following methods:

Phenol/chloroform extraction :

After PCR amplification and restriction digestion, the DNA was electrophoresed using the appropriate percentage low melting temperature (LMP) agarose-TAE gel and excised under long wave UV light (366 nm, IBI) in a minimum volume. The gel slice was then transferred into a 1.5 ml of Eppendorf tube and smashed to small pieces. An equal volume of phenol/chloroform was added to the mixture, vortexed for a few seconds, and kept in -70°C for 10 min. The frozen mixture was then microcentrifuged for 5 min. The aqueous upper phase was removed to a new 1.5 ml tube and extracted once with same volume of phenol/chloroform/ 3-methyl-1-butanol (50:48:2) and once with same volume of chloroform. The DNA was precipitated by addition of 0.1 volume of 3M sodium acetate and 2 volumes of absolute ethanol, and dissolved in an appropriate volume of 10 mM Tris buffer, pH 8.0 .

Magic PCR Preps™ DNA purification system :

DNA contained in LMP agarose gel slices was also purified by using Magic PCR Preps™ purification Kit (Promega), according to the manufacturers protocol. The gel slice was incubated in a 1.5 ml Eppendorf tube at 70°C until the agarose was compete melted. 1ml of Resin was added to the melted agarose and mixed immediately. The mix was pipetted into a syringe barrel attached to a mini column. Then the syringe plunger was inserted slowly to push the slurry flowing through the mini column. The mini column was washed with 2ml of 80% isopropanol and centrifuged for 20 seconds at 13,000rpm in a microcentrifuge to dry the resin. The mini column was then transferred to a new microcentrifuge tube. 50 µl of ddH_2O was applied to the column and left 3 minute. The microcentrifuge tube with the minicolumn was centrifuged for 1 minute at 13,000 rpm to elute the bound DNA fragments. The

mini column was removed and discarded. Purified DNA was stored -20°C until required.

2.7.4 Preparation of total RNA from parasite culture

Total RNA was prepared from asexual *P. falciparum* parasites using the RNA Isolator Kit (containing guanidinium thiocynate, N-lauryl sarcosine and EDTA; Genosys Biotechnologies, Inc). Cultures were centrifuged at 4,000 rpm for 5 minutes to collect the parasites. Then the parasites were released from the RBCs by lysis in 0.1% saponin in 1xPBS, centrifuged and collected prior to extraction following the manufacture's instructions. The RNA pellet was resuspended in 20 ml DEPC-treated H_2O and incubate at 55-60°C for 10 min. The total RNA was either used immediately or stored frozen at -70°C until use.

2.7.5 DNA quantification

The concentration of DNA in solution was estimated by measuring its absorbence at 260nm in spectrophotometer. It was assumed that an OD_{260} of 1.0 is equivalent to a concentration of 50 μg/ml for double stranded DNA and 35 μg/ml for oligonucleotides.

2.7.6 Restriction endonuclease digestion, electrophoresis and photography

Restriction endonuclease enzymes were obtained from Boehringer Mannheim Chemical Company and New England Biolabs. Restriction enzyme digestions were performed in the appropriate buffer supplied by the manufacturer. The reactions were carried out at 37°C for 3-6 hours (unless otherwise indicated) and 4-12 units of restriction endonuclease were used per μg of DNA. Reactions were terminated by addition of 1/10 volume of 50mM Na_2EDTA, 40% w/v sucrose, containing 0.25% bromophenol blue, 0.25% xylene cyanol.

Agarose gels were used to check the integrity of DNA, to analyse PCR products and restriction enzyme digests. 1% - 2% agarose (IBI) was dissolved in appropriate volume of 1x TAE buffer by boiling. After the solution had cooled to around 45°C, ethidium bromide (EtBr) was added at concentration of 0.5μg/ml, mixed and the gel was poured onto a pre-levelled electrophoresis gel plate (IBI) and left to polymerise at room temperature. The appropriate amount of DNA was mixed with 0.1 volumes of gel loading buffer (40% w/v sucrose , 50mM EDTA pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol) and loaded into the submerged gel. Electrophoresis was carried out in 1xTAE at 100 volts for the appropriate time depending on the size of the fragment being separated. Fluorescence from DNA

bound-ethidium bromide was visualised by short wave-length UV light and photographed.

In order to estimate the size of DNA fragments 2μg of DNA markers were loaded adjacent to the DNA samples on agarose gels. The standards, pBR328 *Bgl* I+*Hinf* I digest (Boehringer Mannheim) consisted of the following sized fragments 2176, 1766, 1230, 1033, 653, 517, 453, 394, 298, 234, 220, and 154 base pairs (bp), and *Hind* III digest of bacteriophage λ cI_{857} (Boehringer Mannheim) consisted of the following sized fragments (bp): 23130, 9416, 6562, 4368, 2322, 2027, 564 and 125.

After electrophoresis the agarose gels were stained with 10 μg/ml ethidium bromide and photographed under UV transilluminator (IBI) using short wave-length. A photograph of the gel was taken by the Polaroid camera with 667 professional film or the gel video documentation system (Video Copy Processor, Mitsubishi ™).

2.8 Polymerase Chain Reaction Techniques

2.8.1 Polymerase Chain Reaction (PCR)

The PCR is used for enzymatic synthesis of specific DNA sequences (Saiki *et al*, 1988). The single or double stranded DNA is first denatured, then annealed with specific primers corresponding to each end of the target sequence, and the sequence extended. This cycle is repeated as much as 20-40 times and has the potential of amplifying the target sequence as much as 10^7-10^9 times. Due to the high temperature required for the denaturation step, the thermostable DNA polymerase, so-called *Taq* DNA polymerase was normally used in the thermal reactions. However, the *pfu* and Deep Vent DNA polymerase could be instead of Taq to improve the PCR accuracy. PCR was performed in 0.5 ml Eppendorf tubes. Each reaction contained 1 μl of each primer at 10 μM, 1 μl of dNTP solution (contained each dGTP, dATP, dTTP, dCTP) at 750 μM, 2μl of 10x*Taq* polymerase buffer (500 mM KCl, 100 mM Tris-HCl pH 8.3, 20 mM $MgCl_2$, 0.02% gelatin), dH_2O to a total reaction volume of 20 μl and 1 μl (0.25 units) of *Taq* polymerase (Perkin Elmer Cetus Corp.). The reaction mixture was overlaid with 20 μl of mineral oil and PCR was performed using Hybaid™ Omnigene thermal reactor. The temperature have to be varied to accomodate differences in oligo composition. The products of the PCR reactions were analysed by electrophoresis on 1.5%- 2% agarose gels and ethidium bromide staining.

2.8.2 Inverse Polymerase Chain Reaction (IPCR)

This technique is used to rapidly *in vitro* amplify the DNA sequences that flank a region of known sequence. The technique used a pair of primers oriented in the

reverse direction of the usual orientation in known region. The template for the reverse primers is a restriction fragment that has been ligated upon itself to form a circle (Triglia *et al*, 1988). In such experiments, 1 μg of *P. falciparum* genomic DNA was digested to completion with EcoRI. The reaction was heat inactivated if possible, but otherwise it was phenol/chloroform extracted and alcohol precipitated. Then the DNA was resuspended in an appropriate volume of TE buffer. The product of the digestion was ligated and inverse PCR was carried out in 50 μl PCR reaction using 2 μl of the ligation mix as template DNA.

2.8.3 Reverse transcriptase (RT)- PCR

Single-stranded cDNA was reverse transcribed from total RNA of asexual or gametocytic parasites of *P. falciparum*. Each reaction contained 2 ng total RNA, 4 μl 5x MrMuLV reverse transcriptase buffer, 2 μl 0.1M DTT, 3 μl 750μM dNTP mixture, 1 μl 20 mM degenerate oligoprimer [T_{11}MN] (M=G,A or C, N=G,A,T or C) and H_2O to give 19 μl total volume in a 0.5 ml microcentrifuge tube. Then incubate the tubes at 65°C for 5 min to denature the mRNA secondary structure and incubate 10 min at 37°C to allow primer annealing. Add 1 μl 200 U/ml MrMuLV reverse transcriptase (BRL) to each tubes, mix well, and incubate at 37°C for 1.5 hours. After that, incubate at 95°C for 5 min to inactivate the reverse transcriptase and spin briefly to collect condensation. Place the tubes on ice for immediate PCR amplification or store at -20°C for later use. The PCR conditions set up as described above. The amplified PCR products were analysed by 1.5% agarose gel electrophoresis and ethidium bromide staining.

2.9 Recombinant DNA techniques

2.9.1 Subcloning and ligation of DNA

For subcloning, vector DNA (50ng) and insert DNA were cut with suitable restriction enzyme endonucleases. In order to prevent the self-ligation of the compatible ends of digested plasmid, linearised DNA was treated with 1 unit of calf intestine alkaline phosphatase (CIP) (New England BioLabs) at 37°C for 1 h which removes the 5'-phosphatase residues from DNA. Then 0.1 volume of 250 mM EDTA pH 8.0 was added and the solution incubated at 75°C for 10 minutes to inactivate the enzyme. Ligation of the insert to vector was titred at molar ratios of approximately 3:1, 5:1 and 10:1. To the mixture of vector and insert DNA was added 1μl of 10x ligase buffer (500 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 100 mM DTT, 100 mM ATP), 1 unit of T4 DNA ligase (Invitrogen Inc.) and the total reaction volume made up to 10μl with dH_2O. The reaction was incubated at 16°C overnight and then used to transform

the competent *E. coli* cells. The pBluescript KS$^+$ plasmid (see Appendix D) was used in the lab for daily subcloning works.

2.9.2 Subcloning of PCR products into pCR2.1 plasmid vector

The TA Cloning™ System (see Appendix A) from Invitrogen Inc. was used routinely during this study to provide a one-step cloning strategy for direct insertion of PCR product into a plasmid vector (Mead *et al*, 1991). The system takes advantage of the non-template dependent activity of the thermostable polymerases used in PCR that add single deoxynucleotides to the 3' end of all duplex molecules. These A-overhangs are used to insert the PCR product into specifically designed vectors providing single 3'T-overhangs ready for the insertion site.

The ligation and transformation were carried out following the manufacturer's instructions. Briefly, the appropriate amount of PCR sample was transferred to a new 0.5 ml eppendoff tube and avoid the contamination of the mineral oil. The ligations with the pCR™ vector (3.9 kb) were set up as 1:1 to 1:3 molar ratio of vector: PCR insert. For each ligation reaction, 2 µl (50 ng) of vector, 1 µl (12.5 ng) of PCR product, 1 µl 10x ligation buffer (0.5 M Tris pH 7.4, 0.1 M $MgCl_2$, 0.1 M dithiothreitol, 10 mM spermidine, 10 mM ATP, 1 mg/ml BSA), 1 µl of T4 DNA ligase and 6 µl of dH_2O. The ligation mix was incubated overnight at 14-16°C.

2.9.3 Preparation of competent bacterial cells

The method used was modified after Hanahan (1983). Cells of the appropriate *E. coli* strain from a -70°C stored master stock were plated on LB agar overnight at 37°C. Two-three colonies were picked and grown in 5 ml LB broth overnight at 37°C with shaking. The following morning the culture was diluted 50-fold in LB broth and incubation with shaking for approximately 2-4 hours until the cell density was 4-7 x 10^7 viable cells/ml (OD_{600} of 0.4-0.6). At the same time, a 100 ml of 80 mM $CaCl_2$/ 15% glycerol solution or transformation buffer I (TBI) was prepared and sterilised by filtering through a 0.22 µm filter and kept on ice bath. The culture was then chilled on ice for 15 minutes. The bacterial pellet was then collected by centrifugation at 4,500 rpm for 5 minutes at 4°C, and resuspended in the ice-cold $CaCl_2$/glycerol solution or TBI at 1/2 of the original culture volume. The cells were then incubated on ice for 5 minutes, collected again by centrifugation as above. The pellet was then resuspended in 1/2 of the original culture volume ice-cold $CaCl_2$/glycerol solution or transformation buffer II, and incubated on ice for 30 minutes. The cells were collected again and resuspended in 1/10 of the original culture volume of ice-cold $CaCl_2$/glycerol solution or transformation buffer II. The cells could then be stored at -70°C in small aliquots (100-200 µl), or could be used the same directly.

2.9.4 Transformation of *E. Coli*

Competent cells (50-100 µl) were mixed with either 5 µl of a ligation reaction or 20 ng of uncut plasmid DNA. The cells were then incubated on ice for 30 minutes, heat-shocked at 42°C for 90 seconds and incubated on ice for 2 minutes. To each transformation mixture, Then 400-450 µl of pre-warmed LB or SOC medium was added and the mixture was incubated at 37°C, with gentle shaking, for 1 hour. 25-100 µl of from each transformation were plated out on appropriate selective LB agarose plates and incubated overnight at 37°C.

2.9.5 Transformation by Electroporation

A overnight culture was set up in 10 ml of LB broth. The following morning the 10ml of overnight culture was inoculated into 1 litre of LB broth and incubated until OD_{600} to 0.5- 1.0 with shaking (about 2.5 hrs). The culture was chilled on ice for 15- 30 min and pelleted in a cold rotor at 4000g for 15 min. The pellet was collected and resuspended in 1 litre of cold water. The pellet was washed with 0.5 litre of cold water. The pellet was collected by centrifugation as above conditions and resuspended in 20 ml of cold 10% glycerol. The pellet was collected again and resuspended in 2- 3 ml of cold 10% glycerol. Cells were frozen in aliquots on dry ice, and store at -70°C for up to 6 months.

For transformation, competent cells kept at -70°C were thawed on ice and 35-40 µl of competent cells (for each reaction) were added to a 1 ng ligation mixture in a 1.5 ml tube (The DNA should be made salt- free, by placing onto a filter is floating on water to dialyse the salts out). Then the transformation mixture was placed into an electroporation cuvette (on ice), electroporated at 2.25 kV, 25mf, 200W and immediately added to 1ml of ice-cold LB broth. Different amounts of mixture were then plated out on LB agar plate with appropriate antibiotics and incubated overnight at 37°C.

2.10 DNA sequencing

2.10.1 Manual sequencing

Sequencing reaction

Sequencing of plasmid DNA was performed by the dideoxy chain termination method (Saiki *et al*, 1988), using the Sequenase version 2.0 Kit based on T7 DNA polymerase (USB), following a modification of protocol of Winship (Winship, 1989). In this protocol, 10% culture grade DMSO was added in the sequencing reactions to enhance the intensity of the signal, to reduce background and to prevent formation of secondary structure when sequencing double-stranded DNA fragments.

For sequencing plasmid and double-stranded DNA, the templates were denatured by an alkaline method. Approximately 5 µg of double stranded DNA in 10 µl of TE buffer was incubated with 1 µl of a solution of 2N NaOH and 2mM EDTA for 30 minutes at 37°C. After incubation, 1 µl of 3N Na acetate (pH5.5) and 100 µl of absolute ethanol were added and incubated at -70°C for 30 minutes. The DNA was collected by centrifugation at 13,000 rpm for 15 minutes. The pellet was washed with 500 µl 70% ethanol and air-dried. The denatured template was mixed with 2 µl of 5x reaction buffer (200mM Tris-HCl pH 7.5, 100mM $MgCl_2$, 250mM NaCl), 1µl of sequencing primer (1ng/µl), 1 µl of 10% DMSO and 7 µl of H_2O, then boiled for 5 minutes and then placed on ice. To this mixture was added 1 µl of 0.1M DTT, 1 µl of 1:5 diluted labelling mix (1.5 µM each of dCTP, dGTP, dTTP), 1 µl of [α^{35}S] dATP (1000ci mM, 10 µCi-1; Amersham SJ246), and 1 µl (2 units) Sequenase (diluted in enzyme dilution buffer of 10 mM Tris-Cl pH 7.5, 5mM DTT and 0.5mg/ml BSA) and incubated at room temperature for 5 minutes. Then 3.5 µl of the solution was aliquoted into 2.5 µl appropriate pre-warmed (37°C) termination mix (80 µM dNTP [dGTP, dATP, dCTP, dTTP]; 8 µM ddNTP [ddGTP, ddATP, ddCTP, ddTTP, respectively]; 50 mM NaCl) and 0.5 µl 10% DMSO and incubated at 37°C for 5 minutes. The reactions were terminated by adding 4 µl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol FF). Samples were boiled 3 minutes prior to loading (3 µl per well).

Sequence reactions were separated on a 380 mm x 0.3 mm, 6% polyacrylamide gel (8 M Urea, 6% acrylamide, 0.16% bis-acrylamide, 0.08% ammonium persulphate in 1x TBE pH 8.8). Electrophoresis was then carried at 55 W constant power. Following electrophoresis, the gel was fixed in 7% acetic acid and 7% methanol for 30 minutes. Gel were dried under vacuum on a gel dryer (Model 583 Gel Dryer, BIO-RAD) for 2 hours at 80°C, and exposed to X-ray film (Kodak XAR-5 or AGFA CURIX) in a cassette for about 16-24 hours. Films were developed in an automatic X-ray film processor (X-ograph X1).

2.10.2 Automatic sequencing

Automated sequencing of plasmid DNA was carried out using dye terminators with the PRISM™ cycle sequencing kit (Applied Biosystems Inc.) and the 377 automated sequencer (ABI). Protocols were achieved as described by the manufacturer (ABI PRISM dye Terminator protocol, P/N 402078). For each reaction, the reagents were aliquoted into a 0.5 ml PCR tube as follows: 8 µl Terminator Ready Reaction Premix, 5 µl dsDNA Template (0.2 µg/ml), 2 µl sequencing primer (1.6 pmol/ml), 3 µl H2O with total volume 20 µl. Cycle sequencings were performed on the HYBAID DNA thermal cycler (OmniGene) by following condition: 25 cycles (96°C for 30 sec,

50°C for 20 sec, 60°C for 4 min, respectively), then keep on 4°C. After reaction, PCR products were precipitated, washed and resuspended in 5 µl deionized formamide and 1 µl of 50 µg/ml Blue dextran in 25 mM EDTA, pH8.0. Each sample was then heated at 90°C for 5 min and 1.8 µl of each sample was loaded. 6% Denaturing gels were run at 50 W for 7 h. Chromatograms were viewed and data were edited using the Seqed™ program (ABI).

2.11 Computing analysis

Sequence analysis was performed using the GCG software package (Devereux *et al*, 1984) or through World Wide Web browsers. The cDNA sequences and translated amino acid sequences of the *SNF2L* gene were compared with the EMBL and GenBank database using the BLAST algorithm (Altschul *et al*, 1990; Altschul *et al*, 1997) as implemented in the GCG package. Local alignments of multiple sequences were constructed using these algorithms under varying parameters on individual sequence pairs, followed by manual optimisation of the alignments. The RNASPL (Solovyev *et al*, 1994) and POLYAH (Brondel *et al*, 1992) programs were used to predict intron-exon junctions and potential poly(A) sites. PSORT version 6.3 (WWW) was used to predict protein localisation sites from cDNA sequences. Statistical analysis of protein sequences was carried out using the SAPS program (Brondel *et al*, 1992). DNA and protein sequences were collected using Entrez programme in NCBI web service. Analyses of phylogenetic trees were performed using PAUP phylogenesis software package.

2.12 Pulsed field gradient gel electrophoresis

Pulsed field gradient electrophoresis (PFGE) analysis has been a important tool for fundamental genetic studies of *P. falciparum* (Foote & Kemp, 1989; Hinterberg *et al*, 1994; Schwartz & Cantor, 1984). Southern blots of PFGE separations readily serve to assign genetic markers to chromosomes. On a practical level, important features of PFGE methods include embedding cells in agarose blocks and removing protein so that the DNA remains intact but free within the gel matrix. The movement of the large chromosomal DNA molecules through the gel was by switching direction of the applied electric field at fixed time intervals.

2.12.1 Preparation of parasite DNA in agarose blocks

For preparation of chromosomal DNA in agarose blocks, parasites were first grown in 5ml asynchronous culture at 5% haematocrit to a 10% parasitaemia. Parasitised red blood cells were pelleted by centrifugation at 4000 rpm for 5 minutes at

4°C. Parasites were released from the Parasitised RBC pellet by lysis with an equal volume of 0.1% saponin/1xPBS (made up fresh). After incubation at room temperature for 10 minutes the parasites were pelleted by centrifugation as above. The parasite pellet was washed twice in excess of ice-cold 1x PBS and resuspended in 2 volumes of 1xPBS at room temperature. The suspension was fixed by addition of an equal volume of 1.5% low-melting point agarose (IBI)/1xPBS at 42°C, mixed gently and quickly pipetted into a pre-warmed mould (Bio-Rad) before the agarose had set. The mould was then placed at 4°C for 20 minutes to allow the gel to set. The agarose blocks were transferred from the mould into lysis solution (1% sarkosyl, 0.5M EDTA pH 8.0, 0.01M Tris pH 8.0, 1 mg/ml proteinase K), approximately 0.5ml per block, and incubated for 48 hours at 50°C with one change of the lysis solution. The blocks were stored in the lysis solution, without proteinase K, at 4°C until they were used. Blocks can last for many months if properly made and stored.

2.12.2 Electrophoresis conditions

Chromosome separations were performed in 0.8% PFGE-grade agarose/0.5x TBE buffer. Molten agarose was poured about 60°C. The gel was allowed to set at room temperature for about 30 minutes. After that the comb was removed and the slices of parasite DNA blocks were inserted into the wells. The blocks were sealed into the wells with low-melting agarose/1xPBS, avoiding trapped air bubbles in the wells. The temperature of the buffer was maintained at 14°C by circulating it through a cooling system. DNA separation was performed on a Bio-Rad CHEF DRII system. Table 2.1 shows the conditions used for separating DNA of various sizes using Bio-Rad CHEF DRII system. Switch time was ramped from 90 to 300 s for 22 h at 95 V followed by switch time ramped from 300 to 720 s for 30 h at 90 V.

2.13 Southern hybridisation

2.13.1 Southern blotting

Southern blotting was carried out by a modification of the method of Southern (1975). The agarose gel containing fractionated DNA was soaked in 0.25M HCl for 20 minutes to partially depurinate the DNA. The gel was then rinsed in dH_2O for 5 minutes before being soaked in 0.4 M NaOH three times for 15 minutes each. The gel was then placed onto blotting paper pads. The DNA was alkaline transfered to Hybond N^+ membrane (Amersham) with 0.4 N NaOH. After overnight transfer the membrane was removed and neutralised in 0.5M Tris-HCl for 5 minutes, air dried, short wave UV linked and wrapped in Saran wrap, (Dow Chemical Company), until it was required for the hybridisation reaction.

For transfer of DNA from mini-gels onto Hybond-N^+ membrane (Amersham) prior to probing the following method was used: After depurination, the gel was soaked in denaturation solution (0.5M NaOH, 1.5M NaCl) for 20 minutes and then transferred into neutralisation solution (1.5M NaCl, 0.5M Tris pH 7.5) for 30 minutes. The gel was then placed on a glass plate and overlaid with a piece of Hybond N^+ membrane and 3 sheets of blotting paper all pre-wetted in 10x SSC. Onto this were placed 3 sheets of dry blotting paper, a stack of paper towels, a glass plate and a weight. After transfer for 4-8 hours, the membrane was removed, rinsed in 2x SSC and air-dried. The membrane was then ready for prehybridisation and probing.

2.13.2 Probe synthesis

For maximum sensitivity in filter hybridisation, phosphorus-32 (^{32}P) is the most widely used radiolabel as it is available at high specific activity and can be detected with efficiency. Several methods are available for making probes. End-labelling with polynucleotide kinase is employed with short oligonucleotides, whereas longer DNA fragments are generally labelled by random priming labelling methods.

In this work, random primer labelling method was used which was first described by Feinberg and Vogelstein (Feinberg & Vogelstein, 1983). The approach utilises Klenow enzyme which has 5'-3'polymerase activity but lacks the 5'-3' exonuclease activity. The absence of the 5'-3' exonuclease also ensures that incorporated nucleotides are not subsequently removed as monophosphates. A mixture of random DNA hexanucleotides is used to prime the DNA and synthesise the radiolabelled complementary strand DNA from the 3'-OH termini of the primers. In this work, the Random Primed DNA Labelling Kit (Boehringer Mannheim), was used to synthesise radioactive probes for hybridisation reactions. The DNA template (25-50 ng in TE buffer) was first denatured by boiling for 10 minutes and chilled on ice to prevent reannealing. A mixture of 2 µl hexanucleotide mixture in 10x reaction buffer [900 mM HEPES, adjusted to pH 6.6 with 4 N NaOH, 100 mM $MgCl_2$], 1 µl [0.5 mM] each of dCTP, dGTP, dTTP, 5 µl [50 µCi, 1.85MBq] of [α-^{32}P] dATP (Amersham International Plc.), and 1 µl (2 units) of Klenow enzyme were then added to the denatured DNA template and incubated at 37°C for 30-60 minutes. The reaction was stopped by addition of 0.5mM EDTA. This procedure can routinely be used to obtain probes with specific activity as high as $1.8x10^9$ dpm/µg.

In general, the denatured labelled DNA can be used directly as a hybridisation probe without stopping the reaction or removing unincorporated label. Removal of unincorporated nucleotides was performed by passing the labelling reaction through a Sephadex G-50 column. A 1.5 ml Eppendorf tube was punctured by a No.26 needle, plugged by glasswool and then packed with 1 ml Sephadex G-50 prepared/TE buffer

(10mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). The column was placed on the top of a new Eppendorf tube and spun at 700x g for 5 minutes to remove excess TE buffer. After the radiolabelling reaction was completed, 5 µl of gel loading buffer was added to the DNA sample. The DNA sample was applied to the top of the column. The column was then placed on the top of another new Eppendorf tube and spun at 700x g for 5 minutes again to recover radiolabelled fractions. The radioactive fractions in the leading peak pooled and stored at -20°C until needed. The probe was boiled for 10 minutes before it was added to hybridisation solution.

2.13.3 DNA hybridisation

All hybridisations were performed in a Hybridisation oven (Hybaid™). In the course of this study all Southern blots, Northern blots and filters were probed for homologous DNA sequences and therefore hybridisation was carried out under stringent conditions throughout. In general, to maximise the rate of annealing of the probe with its target, hybridisations were usually carried out in the presence of 10% dextran sulphate in a solution of high ionic strength at a temperature that was 20-25°C below the melting temperature (T_m) (Sambrook *et al*, 1989).

The probe was added with sonicated salmon or herring sperm DNA boiled for 10 minutes before it was added to hybridisation solution. Blots were prehybridised at 65°C for 4 h using a prehybridisation buffer which was made by addition of 10 g of dextran sulphate [MW~500,000], 10ml of 10% SDS, 5.8g NaCl and 5 ml of 50x Denhart's solution to 85 ml of dH_2O and heated at 65°C for 30 minutes. For hybridisation, the prehybridisation buffer was reduced to 50 µl of hybridisation buffer per cm^2 of membrane. Just before hybridization was carried out, the probe was added 0.1 ml of 5 mg/ml sonicated herring sperm DNA and 0.9 ml of dH_2O and boiled to denature the DNAs. The boiled, denatured probe was then added and hybridisation continued overnight at 56-65°C. Non-specifically bound nucleotides were removed by washing the membrane twice with excess 2xSSC at room temperature for 10 minutes. The membrane was washed stringently with two washes in 0.5xSSC, 0.1% SDS at 56°C for 30 minutes and then two washes in 0.1x SSC for 30 minutes. Filters were then placed between two layers of Saran Wrap and autoradiographed.

2.13.4 Stripping probes from DNA blots

Blots were stripped before being used in new hybridisation reactions. DNA blots were incubated at 45°C for 30 minutes in 0.4N NaOH and then in neutralisation solution (0.1xSSC, 0.1% SDS, 0.2M Tris-Cl pH 7.5) for 30 minutes. Alternatively, blots could be boiled in 0.1x SDS for 10 minutes. The blots were then exposed to X-ray films and autordiographed to check for removal of the probe.

2.13.5 Autoradiography

This was performed using X-ray film (Kodak XAR-5 or AGFA Curix) for ^{32}P probed filters, in a cassette containing intensifying screens. The cassette containing ^{32}P probed filters were stored at -70°C while those containing ^{35}S labelled gels were stored at room temperature. Film was developed using a X-O'GRAPH X1 automatic X-ray film processor.

2.14 Colony hybridisation

The procedure used was that described by Hanahan and Meselson (1983) and Buluwela *et al* (1989) in which colony lysis, DNA denaturation and DNA fixation to the membrane are carried out in a single step.

A sterile Hybond-N^+ filter was laid on a surface of a day-old LB agar plate containing ampicillin using sterile, blunt-ended forceps (Millipore). When the filter was thoroughly wet, the bacterial colonies were directly replicated from the plate onto the surface of filter. The filters were then oriented by making a series of holes in them with an 18-gauge needle and before being peeled apart. The filters were then peeled off the plate and laid, colony side up, on 3MM paper to dry. The plate was then inverted and incubated overnight at 37°C for about 4-8 hours to allow colonies grow back.

The replica filters were removed and placed for 3 minutes, colony side up, onto filter paper pre-wetted with 10% SDS solution. The filters were then placed for 5 minutes onto a second sheet of 3MM paper that had been saturated with denaturing solution (0.5M NaOH, 1.5M NaCl). After that, filters were placed for 5 minutes onto a third sheet of 3MM paper that had been saturated with neutralising solution (1.5M NaCl, 0.5M Tris.Cl pH 8.0). Finally, the filters were rinsed briefly in 2xSSC solution and dried at room temperature for 30-60 minutes. The DNA was fixed on the filters by baking at 80°C for 2 h or cross-linking with UV for 5 minutes. The filters were either used immediately in hybridisation reactions or wrapped loosely in aluminium foil and stored under vacuum at room temperature. Positive signals were then identified and the appropriate colony on the master plate picked for further analysis.

2.15 Northern hybridisation

2.15.1 Formaldehyde gel electrophoresis

RNA was run in 1.2% w/v agarose gel, 1x MOPS buffer (20mM MOPS, 5mM NaAc pH 7.0, 1mM EDTA, pH 7.0) , 3% formaldehyde. 4.5 µl of RNA in solution was mixed with 2 µl of 10x MOPS buffer, 10 µl of formamide and 3.5 µl of

37% formaldehyde to final volume 20 µl and heated to 65°C for 10 minutes then immediately cooled on ice. After addition of 0.25 volumes of sample buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) the sample and an aliquot of RNA molecular weight marker (Boehringer Mannheim RNA marker II consisted of the following bands (kb) 7.4, 5.3, 1.9, 1.6, 1.0, 0.6, 0.4, and 0.3) were loaded. The gel was run at 70V for 1-2 hours. After the electrophoresis was complete the gel was stained and photographed.

2.15.2 Northern blotting

The RNA gels were first rinsed in 1xSSC for 10 minutes to remove formaldehyde from gels and then denatured by soaking the gels for 10 minutes in 1xSSC, 50mM NaOH solution for 10 minutes. The gels were blotted overnight onto Hybond-N^+ membrane using 10xSSC. After overnight transfer the membrane was removed and rinsed in 2xSCC. The blot was then ready for hybridisation.

2.15.3 RNA hybridisation

Membranes were prehybridised at 65°C for 4 hours in 8 ml prehybridisation buffer (50 µl / cm^2 membrane) as above. The buffer was changed once after 8 hours. The probe was random-primer labelled. Hybridisation was carried out at 56°C overnight using 3 ml of the same solution overnight. Filters were washed in 2xSSC, 0.1% SDS, then in 0.1xSSC, 0.1% SDS for 30 minutes twice before exposure to X-ray films.

2.15.4 Stripping of Northern blots

When necessary Northern blots were first stripped prior to probing. A solution of boiling 0.1% (w/v) SDS was poured over the blot and allowed to cool to room temperature. The filter was then autoradiographed to check for complete removal of the probe.

2.16 Preparation of fusion proteins

2.16.1 Construction and expression of the recombinant glutathione S-transferase (GST) fusion proteins

pGEX-4T-1 vectors were used to express polypeptides fused with glutathione-*S*-transferase (GST) in *E. coli* BL21 host. The basic protocol used for expressing, purifying and cleaving fusion proteins was that described by Smith & Corcoran (1990). For this work, the PCR fragment GST1 was amplified from the *P. falciparum* PfSNF2L clone with a BamHI site at both ends and subcloned into the *Bam*HI site of

the pGEX-4T-1 vector to produce a GST N-terminal fused fusion protein. BL21 cells were used for transformation and the cells were grown on LB-amp agar plates overnight at 37°C. Positive clones were identified by restriction enzyme digestion of purified plasmid and then further confirmed by DNA sequencing.

Single colonies were picked and put into 5 ml of LB-amp and grown overnight at 37°C with shaking. The following day the culture was diluted 1:10 in LB-amp and grown at 37°C with shaking to an OD_{600} of 0.6-0.8 (about one hour). 1ml of culture was removed as a control before induction, and to the remaining culture 100mM IPTG was added to a final concentration of 0.1mM in order to induce fusion protein expression. Cells were then grown for a further 2-3 hours. Cells were collected by centrifugation, at 5,500 rpm (Sorvall) for 10 minutes at 4°C. They were then resuspended in 1 ml of ice-cold 1x PBS. In order to check expression of the fusion protein 10 µl of the resuspension was added with 10ml of 2x loading buffer and checked by SDS-PAGE and Coomassie staining.

2.16.2 Purification of GST fusion proteins

Bacterial cells with the induced fusion protein were lysed by sonication with a 5mm diameter probe. 10% Triton X-100 was added to the lysed cell suspension to a final concentration of 1%, in order to minimise any association of fusion proteins with bacterial proteins. An aliquot of the lysed bacterial extract was then spun and a sample of both the supernatant and pellet analysed by SDS-PAGE and Coomassie staining to indicate the solubility of the GST fusion protein.

For large scale purification of a soluble fusion protein, a colony of a pGEX transformant was inoculated into 100ml LB-amp and grown overnight at 37°C with shaking. This culture was then diluted 1:10 into 500ml fresh LB-amp and grown at 37°C continuously to an OD_{600} of 0.6-0.8. The culture was induced by addition of 100 mM IPTG to a final concentration of 0.1 mM and continued the incubation was continued for an additional 3 hours. The culture was centrifuged for 10 minutes in a Beckman rotor at 5000 rpm and the supernatant discarded. The pellet was resuspended in 10 ml ice-cold 1x PBS and the cells were lysed by sonication. After sonication, 10% Triton X-100 was added to give a final concentration of 1%. This was mixed and centrifuged for 5 minutes at 10000rpm at 4°C. The supernatant was collected and 1 ml of a 50% slurry of pre-swelled S-linkage glutathione-agarose beads in PBS (Pharmacia) was added and this mixed was gently for 1 hour at room temperature. The beads were washed three times with 1xPBS. The pellet was resuspended in 1ml of ice-cold PBS and then transferred to an Eppendorf tube. The fusion protein was eluted by addition of 1ml of 50mM Tris-Cl (pH 8.0)/5 mM reduced glutathione. An aliquot of induced proteins were run in parallel to non-induced ones on SDS-PAGE.

2.16.3 Construction and expression of the recombinant histidine tagged (His-tag) fusion protein

pET29 TA vectors were used to express polypeptides fused with His6-tag in *E. coli* BL21 (DE3) host. The basic protocol used for expressing, purifying and cleaving fusion proteins was that described by Studier *et al* (Studier & Moffatt, 1986). For this work, the PCR fragment FP1, FP2 and FP3 were amplified from the *P. falciparum* PfSNF2L clone and subcloned into the A-protruding site site of the pET29 TA vector to express a His6-tag C-terminal fused fusion proetin. BL21 (DE3) cells were used for transformation and the cells were grown on LB-kam agar plates overnight at 37°C. Positive clones were identified by restriction enzyme digestion of purified plasmid and then further confirmed by DNA sequencing.

Single colonies were picked and put into into 5 ml of LB-amp and grown overnight at 37°C with shaking. The following day the culture was diluted 1:10 in LB-amp and grown at 37°C with shaking to an OD_{600} of 0.6-0.8 (about one hour). 1ml of culture was removed as a control before induction, and to the remaining culture 100mM IPTG was added to a final concentration of 0.1mM in order to induce fusion protein expression. Cells were then grown for a further 2-3 hours. Cells were collected by centrifugation at 5,500 rpm (Sorvall) for 10 minutes at 4°C. They were then resuspended in 1 ml of ice-cold 1x PBS. The cells were lysed by sonication on ice and spun by microcentrifuge. A sample of both the supernatant and pellet was analysed by SDS-PAGE and Coomassie staining to indicate the solubility of the His-tag fusion protein.

2.16.4 Purification of His tagged fusion proteins

A. Protein purification under non-denaturing conditions

The cells were harvested from 5 ml induced culture by centrifugation at 4500 rpm on 4°C for 5min. After decanting the supernatant, the cell pellet was resuspended in 0.2 ml of 1x ice-cold Binding Buffer. NP-40 can be added to 0.1% to reduce non-specific binding. With the sample in a tube on ice bath, the sample was sonicated until no longer viscous. The DNA was sheared to avoid clogging of the resin. The cell lysate was then centrifuged at 13,000 rpm for 30 min to remove debris. The post-centrifugation supernatant was filtered by passing through a 0.45 μm filter to reduce clogging of the resin (for soluble protein purification) and kept on ice before further purification.

His·Bind resin was gently mixed until completely suspended. Each 2.5 ml of settled resin can be used to purify up to 20mg of target protein. 0.125ml of settled resin (0.24 ml suspension) was aliquoted into 1.5ml tube and spun for 1min at 900

rpm. After decanting supernatant, the resin was used the following sequence of washes to charge and equilibrate (where 1 volume is equivalent to the settled bed volume): 3 vol sterile deionized water, 5 vol 1x Charge Buffer, 3 vol 1x Binding Buffer, respectively.

The charged resin was spun down and separated from the supernatant. The filtered cell extract was added into the resin tube and gently mixed to resuspend the resin, and shaken continuously for 1 hr to allow His·Tag bound to Ni^{2+} charged resin. The supernatant was removed after centrifugation and the resin was washed with 1.25 ml of 1x Binding Buffer, 0.75 ml of 1x Wash Buffer, respectively. The bound proteins were eluted with 0.75 ml of 1x Elute Buffer.

Eluted proteins were dialysed in the 1X PBS. A gradual removal of the imidazole is recommended for refolding the desired protein. After dialysis, the sample can be concentrated by sprinkling solid PEG 15,000-20,000 or Sephadex G50 on the dialysis tubing (Using dialysis tubing < 6,000 MW) or by vacuum dessication.

B. Protein purification under denaturing conditions

The cells were harvested from 5ml induced culture by centrifugation at 4500rpm for 5min. The cell pellet was then resuspended, sonicated and centrifuged as described above. The pellet was washed with 1ml of 1x Binding Buffer (without denaturant) three times to release trapped proteins. Finally, the pellet was resuspended in 0.25ml 1X Binding Buffer containing 6M urea and incubated on ice for 1 hr to completely dissolve the protein. Any remaining insoluble material was removed by centrifugation at 13,000rpm for 30min. The post-centrifugation supernatant was filtered by passing through a 0.45 μm filter to prevent clogging of the resin.

The resin is prepared as described above, except that Binding Buffer plus 6M urea should be used in the final equilibration step. The charged resin was spun down and separated from the supernatant. The filtered cell extract was added into the resin tube and gently mixed to resuspend the resin, then shaken continuously for 1 hr to allow the His·Tag to bind to the Ni^{2+} charged resin. The supernatant was removed after centrifugation and the resin was washed with 1.25 ml of 1x Binding Buffer, 0.75 ml of 20 mM imidazole buffer (combine 11ml of 1X Binding Buffer with 4.1ml of 1X Wash Buffer). The bound protein was eluted with 0.75ml of 1x Elute Buffer. Most proteins will elute with 300 mM imidazole under denaturing condition.

Eluted proteins were dialysed in the 1X PBS. A gradual removal of 6M Urea is recommended for refolding the desired protein. After dialysis, the sample may be concentrated by sprinkling solid PEG 15,000-20,000 or Sephadex G50 on the dialysis tubing (Using dialysis tubing < 6,000 MW) or vacuum pump.

2.17 Protein electrophoresis

2.17.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein extracts prepared from bacteria and parasites were fractionated by SDS-PAGE (Laemmli, 1970) with a discontinuous buffer system. Mini-gels (80mm x 70mm x 0.75mm) (Hoefer) were used in the course of this study. Protein samples were prepared by the addition of an equal volume of 2x sample loading buffer and then boiled for 5 minutes prior to loading. Gels were run in an electrode buffer (25mM Tris, 190mM glycine, 1%SDS) at 40 mA. Separated proteins were then either stained with Coomassie blue stain or electrophoretically transferred onto a PVDF membrane for Western blotting.

2.17.2 Coomassie blue staining of protein gels

Protein gels were stained with 0.1% Coomassie brilliant blue R-250 in order to visualise fractionated protein. The stain was prepared in water: methanol: glacial acetic acid 4:5:1 for 20 min to 1 h at room temperature with gentle agitation. Excess stain was removed by destaining the gel in water: methanol: glacial acetic acid 8:1:1 until the background was clear. Coomassie staining can detect approximately 0.5-1 μg of protein per band.

2.17.3 Protein molecular weight markers

Color markers (Sigma, C4337) contained Myosin (205,000; blue), β-Galactosidase (116,000; turquoise), BSA (66,000; pink), Ovalbumin (45,000; yellow), Carbonic anhydrase (29,000 orange), Trypsin inhibitor (20,100; green), α-Lysozyme (14,200; purple) and Aprotinin (6,500; blue) was used routinely in this study. The marker was prepared in 2x loading buffer and boiled for 5 minutes before being loaded. For mini-gels 2 μg of each marker protein was loaded per well, and for large gels 10 μg per lane was loaded. The gel was then dried under vacuum.

2.18 Mouse immunisation and serum preparation

Five MF1 mice were used to raise antibodies against each PfSNF2L fusion protein. Approximately 500 μg of protein was emulsified in 1ml of complete Freund's adjuvant. Each mouse was intraperitoneally (i.p.) injected 200 μl of the mixture. Prior to the first immunisation a 0.25 ml tail bleed was taken from each mouse as a sample of 'pre-immune' sera. Four weeks later the animals were boosted with protein prepared in incomplete Freund's adjuvant. Again 0.2 ml was i.p. injected into each mouse. Mice were boosted again 4 weeks later following an identical procedure. Two weeks after each boost a 0.5 ml eye bleed was taken.

To separate the serum from whole blood the blood sample was left for 1 hour at 37°C, incubated at 4°C overnight and the serum decanted into a fresh tube. In order to remove any remaining traces of blood clot the serum was centrifuged for 10 minutes at 2000 rpm. The clarified serum was then removed, aliquoted into Eppendorfs and stored at -20°C. The sera were tested by western blotting and IFA.

2.19 Western blotting

2.19.1 Preparation of total proteins from *P. falciparum*

Parasites were released from infected RBCs by saponin lysis and washed in 1xPBS as described in Method 2.7.1. The parasites were pelleted by centrifugation at 4,000 rpm for 10 minutes, washed in PBS and resuspended in SDS-PAGE loading buffer, 100 µl per ml of original infected blood. Samples were boiled for 5 minutes prior to loading, 10-15 µl per lane. 1x PBS with 25 mM EDTA, 5 mM EGTA and protease inhibitor cocktail (Boehringer Mannheim) was used instead of normal 1x PBS in order to prevent protein degradation during the sample preparation. Parasite pellets were also sometimes resuspended in 100 µl of 1x NETTIL (0.15 M NaCl, 5 mM EDTA, 50 mM Tris-HCl (pH 7.5), 0.5% Triton X-100, 0.01% Na azide, 20% glycerol, 0.004% bromophenol blue, pH7.5) and 10 µl resuspension was run in a native gel.

2.19.2 Protein blotting

Proteins separated by SDS-PAGE were transferred onto PVDF (Sigma) or Hybond-C nitrocellulose (Amersham) membrane by the electrophoretic transfer method (Renart *et al*, 1979; Towbin *et al*, 1979).

The following gel 'sandwich' was assembled from the bottom anode electrode of semi-dry blotting apparatus (Pharmacia LKB) and presoaked in transfer buffer (48mM Tris-base, 39mM glycine, 0.037% SDS, 20% methanol): four pieces of 3MM papers cut to the size of the gel, a piece of PVDF or nitrocellulose filter cut to the size, the gel, another four pieces of 3MM papers. To make sure no bubbles were trapped between each layers, a pipette was used to gently roll over each additional layer. After excess transfer buffer was removed, the upper cathode electrode was placed on the top of the stack and connected the electrodes. Transfer was conducted at 100V for 1.5 hours. Filters were then ready to be probed for bound antigen. The marker proteins can be visualised by (1) 10 min amido black stain in 0.1% amido black 10-B/45% methanol/ 10% acetic acid, and destain in 90% methanol/ 2% acetic acid/ 8% water; (2) india ink stain (i. wash blot in 0.4% Tween 20/PBS, 2x 100 ml for 5 min each; ii. Place blot in 100 µl ink/100 ml of 0.3% Tween 20/ PBS at R/T for 2 hr; iii. Destain

with several times PBS washings); (3) Ponceau S reversible stain (0.1% Ponceau S (w/v)/ 5% acetic acid (v/v), Sigma P7170).

2.19.3 Antibody binding to Western blots

After the Western transfer was completed the following method was used to detect antigens bound to PVDF or nitrocellulose filters. The filters were first incubated in blocking solution (5% fat free dried milk, 50 mM Tris pH7.5, 150 mM NaCl, 0.05% Tween 20) for 2 hour and washed with 1xPBS, 0.05% Tween 20, twice. The filters were then incubated in an antiserum in blocking solution with gentle shaking at room temperature for 2 hours. The polyclonal mouse serum was used at several dilutions. The filters were washed three times, 5-10 minutes per wash in TBST (50 mM Thris pH7.5, 150 mM NaCl, 0.05% Tween 20). The filters were then incubated for 2 hours in anti-mouse Horse Radish Peroxidase conjugated anti-IgG (Sigma) diluted 1/5000 in blocking solution. The filters were washed as before and the bound second antibody detected using Diaminobenzidine/metal ion enhancement (DAB/Metal) buffer. For 10 ml of detection solution, 9 ml of 50 mM Tris pH7.5, 1 ml of 0.3% w/v $CoCl_2$ and 10 µl of Hydrogen Peroxide were added to 6 mg of 3,3'-Diaminobenzidine tetrahydrochloride and the solution used immediately. Colour development was continued until the desired intensity was achieved. Filters were then rinsed in dH_2O and dried. All incubations and washings were performed at room temperature with shaking.

2.20 Indirect immunofluorescence assay (IFA)

2.20.1 Preparation of slides

In order to prepare slides from different developmental stages of *P. falciparum*, cultures were first synchronised according to the Method 2.6.2, and examined regularly by thin blood film, until parasitaemias of 5%-10% were obtained. Then the cultures were centrifuged at 4,000 rpm for 5 minutes, and the cells were washed three times in 1x PBS with or without 25 mM EDTA, 5 mM EGTA and protease inhibitor cocktail. The pellet was then resuspended at 30% hematocrit in 1x PBS. Then on a clean slide (Hendly, Essex, UK), 5 µl was used to make a thin smear and allowed to dry at room temperature. The slides were either used immediately or stored with desiccant (silica gel) at -20°C.

2.20.2 Staining of slides

Slides were fixed in ice-cold acetone for 5 minutes, allowed to air dry and then grided into 12 squares using nail varnish. Each square was pre-blocked with 20 µl of 1x PBS with 1% BSA and 0.02% sodium azide at room temperature for 30 minutes. The solution in each square was removed by water suction and replaced with 20 µl

appropriately diluted antibody, in PBS with 1% BSA and 0.05% sodium azide. Slides were incubated in a moist chamber at room temperature for 60 minutes, before the antibodies were then aspirated off, and washed three times in ice-cold 1x PBS, allowing 5 minutes per wash. Slides were then incubated with 20 µl of the manufacturer's recommended dilution of the second conjugated antibody per square. This was either a fluorescein isothiocyanate (FITC) conjugated anti-mouse immunoglobulin (Sigma) or a rhodamine isothiocynate (RITC) conjugate of anti-rabbit immunoglobulin (Sigma). In addition, parasite nuclei were counter-stained with DAPI (4;6 diamidino-2-phenyl-indole, Sigma). In order to do this DAPI was included in the second antibody solution at 1:40 dilution in PBS/azide (final concentration of 5 µg/ml). Slides were then incubated and washed as before, and mounted in mounting solution (90% glycerol, 100 mM Tris, pH8.0 and 2.5 mg/ml 1,4-diazabicyclo-[2,2,2] octane (DABCO) (Sigma)) to prevent fading. The slides were then examined by fluorescence microscopy (Leitz, Germany).

2.21 Differential display reverse transcriptase-PCR (DDRT-PCR)

2.21.1 Reverse transcription of *P. falciparum* RNA

This reaction was performed in 0.5 ml Eppendorf tubes. Each reaction contained 4 µl of 5X MrMuLV reverse transcriptase buffer, 2 µl of 0.1 M DTT, 1.6 µl 250 µM dNTP, 2 µl of 10 µM oligo(dT)primer: [T_{11}GA], 0.2 µg total RNA from asexual or gametocyte stages of *P. falciparum* (obtained as described previously) and H_2O to a total volume of 19 µl. Two different stage mouse embryo total RNAs were used as control. The tubes was incubated at 65°C for 5 min to denature the mRNA secondary structure and incubate at 37°C for 10 min to allow primer annealling. Add 1 µl 200 U/ml MrMuLV reverse transcriptase (BRL) to each tube, mix well, and incubate at 37°C for 1.5 hours. After that, incubate at 95°C for 5 min to inactivate the reverse transcriptase and spin briefly to collect condensation. Place the tubes on ice for immediate PCR amplification or store at -20°C for later use.

2.21.2 Performing of PCR amplification

Set up a 20 µl of reaction mixture in a 0.5 ml Eppendorf tube as follows: 8.2 µl of H_2O, 2 µl of each cDNA (as prepared above), 2 µl of 10X PCR buffer (100 mM Tris-HCl pH8.3, 500 mM KCl, 0.01% (w/v) gelatin), 1 µl of 15 mM $MgCl_2$, 1.6 µl of 25 µM dNTP (Boehringer Mannheim), 1 µl of 10 mCi/ml [α^{35}S]dATP, 2 µl of 2 µM arbitrary decamer: [AATGGATGTG], 2 µl of 10 µM oligo(dT) primer: [T_{11}GA] and 1 U AmpliTaq (Perkin-Elmer). The reaction mixture was mixed well and overlaid with 20 ml of mineral oil. Then PCR was performed using the following conditions: 40 cycles 94°C for 30 sec, 40°C for 2 min, 72°C for 30 sec, an additional one cycle

extension period at 72°C for 10 min. 3.5 ml PCR product was mixed with 2 µl of formamide loading buffer and incubated 2 min at 80°C to denature the template. The sample was loaded onto a 6% denaturing polyacrylamide gel and the gel run at 60W for about 3 hr until the Xylene cyanol ran to within 10 cm of the bottom.

2.21.3 Recovery of differential displayed amplified DNA bands

Gels were transferred to a piece of Whatman 3MM filter paper and dried by gel drier for 1 hr at 80°C. Dried gels were marked by radioactive ink to orient their position, and the film exposed for 24 to 48 hrs at room temperature. DNA bands of interest which are differentially displayed in different lanes were marked on the gel with a clean pencil and cut from the gel with a razor blade. The excised gel was placed in a 0.5 ml microcentrifuge tube and 100 µl of H_2O was added to (incubated for 10 min at room temperature) before boiling for 15 min. Spin for 2 min at 13,000rpm to pellet gel slice and paper debris. The supernatant was decanted into a clean tube and added to 10 µl of NaOAc, 5 µl of 10 mg/ ml glycogen (as a carrier) and 400 µl of 100% Ethanol. The tube was incubated at -70°C for 30 min and spin for 10 min at 13,000 rpm. The pellet was rinsed with 500 µl of 85% Ethanol, air dried and the DNA pellet was dissolved in 10 µl of H_2O.

2.21.4 Reamplification of the recovered DNA

4 ml of the recovered DNA was reamplified using the same oligo(dT) primer: [T_{11}GA] and pCR condition described as above; except 1.6 µl of 250 µM dNTP was used instead of 1.6 µl of 25 µM dNTP. The reamplified PCR products were analysed by 1.5% agarose gel electrophoresis and ethidium bromide staining. Because DNA bands are not visible after the first reamplification, 4 µl 1/100 dilution of the first reamplification sample was used for a second 40 cycles of amplification to enable visualisation of the DNA bands of interest.

2.22 Administation of heat shock to synchronous *P. falciparum* cultures

Cultures were first synchronised according to the method 2.7.2 at ring stage and examined regularly by thin blood film, until a parasitaemia of 5%-10% was obtained. Then three flasks of 5 ml cultures were set up in 5% hematocrit, 1% parasitaemia and gased. One bottle was incubated at 40°C for 3 hours and then moved to 37°C, the other two flasks were continuously incubated at 37°C. Thin blood smears were taken from these three bottle every 12 hours. One bottle was treated at 40°C after 12 hours culture (about trophozoit stage) for 3 hours as well as another bottle was

treated after 24 hour culture (about schizont stage) for 3 hours. The smears were stained with Giemsa solution and cells were counted at least 5 microscopic fields.

2.23 The effects of inhibitors and drugs on progression through the *P. falciparum* cell cycle

In order to work with synchronised developmental stages of *P. falciparum* parasites, cultures were first sorbitol treated according to the method 2.7.2 at ring stage and examined regularly by thin blood film, until parasitaemia of 5%-10% was obtained. 5 ml of synchronised cultures were set up in 5% hematocrit, 1% parasitaemia. Different inhibitors or drugs (2.5 μg/ml aphidicolin, 2.5 ng/ml pyrimethamine, 0.1 μg/ml colchicine, 30 μM verapamil and 100 ng/ml HC-toxin) were continuously added a each change of culture medium. Thin blood smears were taken fron those cultures every 12 hours, stained with Giemsa solution and counted.

2.24 Apoptosis assay with the *P. falciparum* cultures

2.24.1 DNA fragmentation assay (DNA ladder assay)

The DNA fragmentation assay was carried out using the method of Herrmann *et al* (1994) (Herrmann *et al*, 1994). After treatment, the parasitised RBCs were lysed in 0.1% w/v saponin/PBS. The parasites were pelleted, washed with 1x PBS and then lysed with 400 μl of Apoptosis buffer D (100 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 0.2 M NaCl, 0.42 mg/ ml proteinase K, 0.2 % SDS pH 8.0) at 55°C for 1 hr. Another 100 μl of Apoptosis Buffer D was added to the mixture and incubated at 37°C for at least 2 h. After incubation, 215 μl of 5 M NaCl was added to final concentration 1.5 M in the mixture. The mixture was then spun at 6,000 rpm, RT for 5 min. The supernatant was transferred to a new 1.5 ml tube and an equal volume of 100% v/v ethanol was added, and kept on ice for 10 min to precipitate the DNA. DNA pellets were washed with 70% ethanol once and then resuspended in 40 μl Apoptosis buffer E (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 200 μg/ ml DNase free RNase A) and incubate in 37°C for 2 h to digest RNA. The samples were run in 1.5% agarose gel.

2.24.2 Terminal deoxynucleotidyl transferase mediated dUTP nick-end-labeling (TUNEL) assay

The ApoAlert DNA fragmentation assay kit (Clontech) based on the TUNEL assay (Piqueras *et al*, 1996) (Piqueras *et al*, 1996) was used in these assays. After treatment, parasitised RBCs were collected by centrifugation, washed and resuspended in 1x PBS to 30% hematocrit. On a clean slide (Hendly, Essex, UK), 5 μl of the cell suspension was used to make a thin smear and allowed to dry at room temperature. To fix the cells, the slides were immersed in a Coplin jar containing fresh 4%

formaldehyde/PBS at 4°C for 25 minutes and washed with 1x PBS, twice. To permeabilise the cells, the slides were immersed in a Coplin jar containing prechilled 0.2% Triton X-100/PBS on ice for 25 minutes and twice washed with 1x PBS for 5 minutes each at each wash. Excess liquid was removed from the slides. The cells were covered in 100 µl of equilibration buffer (200 mM K cacodylate pH6.6, 25 mM Tris pH6.6, 0.2 mM DTT, 0.25 mg/ml BSA and 2 mM $CoCl_2$) and equilibrated at room temperature for 10 minutes. Terminal deoxynucleotidyl transferase (TdT) incubation buffer was prepared for the experimental samples, positive controls and negative controls and added to the cells. Then the slides were placed at 37°C for 1 hour and washed in 2x SSC at room temperature for 15 minutes. The slides were washed again in 1x PBS for 5 minutes three times. The cells were then stained with 0.5 µg/ml propidium iodide (PI)/PBS at room temperature for 10 minutes and washed in dH_2O twice. Slides were mounted in mounting solution (90% glycerol, 100 mM Tris, pH8.0 and 2.5 mg/ml 1,4-diazabicyclo-[2,2,2] octane (DABCO) (Sigma)) to prevent fading. The slides were then examined by fluorescence microscopy.

2.25 Methods for the analysis of gene organisation

The release of the GenBank (Benson *et al*, 1998), dated February 3, 1998, was used in this study. In order to analyse the intron number, size, position, splicing and initiation pattern, the sequences from *Plasmodiums* and *T. gondi*i were first chosen which have complete coding sequences. The genes from small organelle genomes were not chosen for this study since they may be different splicing pattern to those genes from nuclear genome. Then by eliminating the open reading frames (ORFs) that contain partial sequences and are repeat chosen (redundant) from different isolates, we obtained 53 genes with 100 introns, 58 gene without introns and 34 mRNA form *P. falciparum*, 19 gene with 46 introns, 37 gene without intron and 8 mRNA from other *Plasmodiums* and 15 gene with 60 introns from *T. gondii*. For each sequences in GenBank, the splice sites and initiation codons were located using the information in the annotation preceding the sequence, a FEATURES line being followed by multiple lines indicating different features present in the gene. As a consequence, there are many errors in the database: many sequence entries are either mislabelled, contaminated, incompletely or erroneously annotated, or contain sequencing errors. Each sequences was checked and corrected manually for this analysis.

To identify a consensus sequence, the criterion was that a single base at each location should represent at least 40% of the total. But the second highest nucleotide may also be chosen if it exceeds 30% or is at least twice as large as the third highest (Long *et al*, 1997; Senapathy *et al*, 1990). A quantitative measure of conservation is

the information content at the i'th site, Rs(i), which ranges from zero to two for maximum information :

$$Rs(i) = 2 + \sum f_b \log2(f_b) - e(n),$$

where f_b and e(n) are the frequency of nucleotides and a correction term for sample size (n) (Long *et al*, 1997; Rogers *et al*, 1995).

CHAPTER 3
STUDIES ON GTPASE MEDIATED PATHWAYS IN *PLASMODIUM FALCIPARUM*

3.1 Introduction

Ran is the most abundant and best conserved small GTPase in eukaryotes (Bischoff & Ponstingl, 1991; Drivas *et al*, 1990). It is also the only known member of the Ras superfamily found in the nucleus. Ran has been implicated in a variety of nuclear functions including: chromosome decondensation, assembly of a functional nucleus, initiation of S phase and import of protein with nuclear localisation signals, pre-mRNA processing, export of proteins and mRNA, and cell cycle coordination. Like all regulatory GTPases, Ran functions as a molecular switch cycling between GTP-bound "on" and GDP-bound "off". It plays an essential role in protein import into the nucleus (Moore & Blobel, 1993) and its GTP loading and hydrolysis cycle may be the major energy source for nuclear localisation signal (NLS) mediated translocation through nuclear pore complexes (Gorlich *et al*, 1996). Ran has been shown to abundantly expressed in the dividing blood stages of *Plasmodium falciparum* (Dontfraid & Chakrabarti, 1994; Sultan *et al*, 1994).

Several Ran homologues have been cloned from *Saccharomyces cerevisiae* (*GSP1/CNR1* and *GSP2/CNR2*), *Schizosaccharomyces pombe* (*Spi*), and humans (*Ran*). The human Ran protein is a 25 Kda molecule, abundant in the nucleus, (Bischoff & Ponstingl, 1991). Determination of the crystal structure of Ran-GDP·Mg^{2+} has shown that the G-domain is similar to Ras but the region involved in GDP and Mg^{2+} coordination is distinct. The N-terminus and the acidic tail are flexible in the crystal structure, suggesting that a major conformational change may occur on GTP binding (Scheffzek *et al*, 1995). Unlike other Ras family proteins, Ran lacks the C-terminal lipid modification motif for cell membrane anchoring but possesses an unusually acidic tail (DEDDDL) (Drivas *et al*, 1990), which stabilises GDP binding to Ran and is required for stable interaction with RanBP1 (Richards *et al*, 1995).

The guanidine nucleotide exchange factor (GEF) for Ran is known as RCC1 (for regulator of chromosome condensation) (Ohtsubo *et al*, 1989), a chromatin protein which stimulates guanine nucleotide loading and thus generates Ran GTP within the nucleus. Ran GTP-Ran GDP hydrolysis is catalysed by the GTPase activating protein Ran GAP1, which is located exclusively in the cell cytoplasm (Bischoff & Ponstingle, 1995). A cytoplasmic activator of GTP hydrolysis is thus separated by the nuclear membrane from a nuclear GTP generator regulating nuclear protein import by the Ran cycle. Ran GDP-associated NLS-marked protein complexes destined for nuclear import form in the cytoplasm and dissociate in the nucleus along a steep trans-

membrane Ran GTP gradient (Gorlich *et al*, 1996). Exchange of GDP for GTP on intra-nuclear RCC1/Ran complexes causes dissociation of the complex and Ran GTP export to the cytoplasm. There, Ran-GAP1 catalyses re-conversion to Ran GDP which can then once more bind and chaperone NLS-proteins to nuclear pore complexes and the nucleus. The rate of guanine nucleotide exchange on Ran is very low, but the exchange factor, RCC1, specifically stimulates the rate by about 10^5-fold (Klebe *et al*, 1995). The intrinsic rate of GTP hydrolysis by Ran is also low and requires a GTPase activating protein (RanGAP). Human RanGAP1 stimulates Ran-GTP hydrolysis by about 10^5-fold (Klebe *et al*, 1995).

RCC1 was first found as a chromatin-associated protein (Ohtsubo *et al*, 1989). RCC1 homologues are identified in *Drosophila* (BJ1) (Frasch, 1991), *Saccharomyces cerevisiae* (SRM1/ PRP20/TMR1) (Clark & Sprague, 1989), *Schizosaccharomyces pombe* (Pim-D1) (Matsumoto & Beach, 1991), and human (RCC1 and RCC1-1) (Miyabashira *et al*, 1994). All RCC1 family members have seven glycine-rich tandem repeats, 50-60 amino acids in length, an N-terminal extension which in at least one case, functions as a NLS, and a C-terminal extension whose length varies (Clark *et al*, 1989; Dasso, 1993; Frasch, 1991; Matsumoto *et al*, 1991; Nishitani *et al*, 1990; Ohtsubo *et al*, 1989). Although there are considerable sequence differences between members of this family, the conservation of key glycine residues among the repeats, both within a single species and between species is striking. RCC1 and its homologues have also been implicated in diverse cellular processes, such as the formation of nuclei, the initiation and termination of transcription and pre-mRNA splicing, and 3'-end formation (Dasso, 1993).

If Ran mediated nuclear protein import functions in intra-erythrocytic malaria parasites then a protein homologous to the Ran-GEF RCC1 must exist. The aim of the work reported in this chapter was to further analyse of a putative RCC1 gene tentatively identified in this laboratory.

3.2 Partial sequencing and structural analysis of the *P. falciparum* RCC1 gene

A 3 kb cDNA fragment encoding an *P.falciparum* RCC1-like molecule was previously identified in an asexual erythrocytic stage cDNA library in this laboratory (Sultan, 1994). Two PCR probes were amplified and used to isolate three more cDNA clones (Figure 3.1). These three clones were completely sequenced. The DNA sequence of the original 3 kb cDNA clone was confirmed and the new clones further extended the sequence on the 5' and 3' sides for additional 1.1 kb. Several sequencing errors in the original sequence have been corrected. A total of 4,161 bp of cDNA have been sequenced, encoding a single continuous reading frame without stop codons

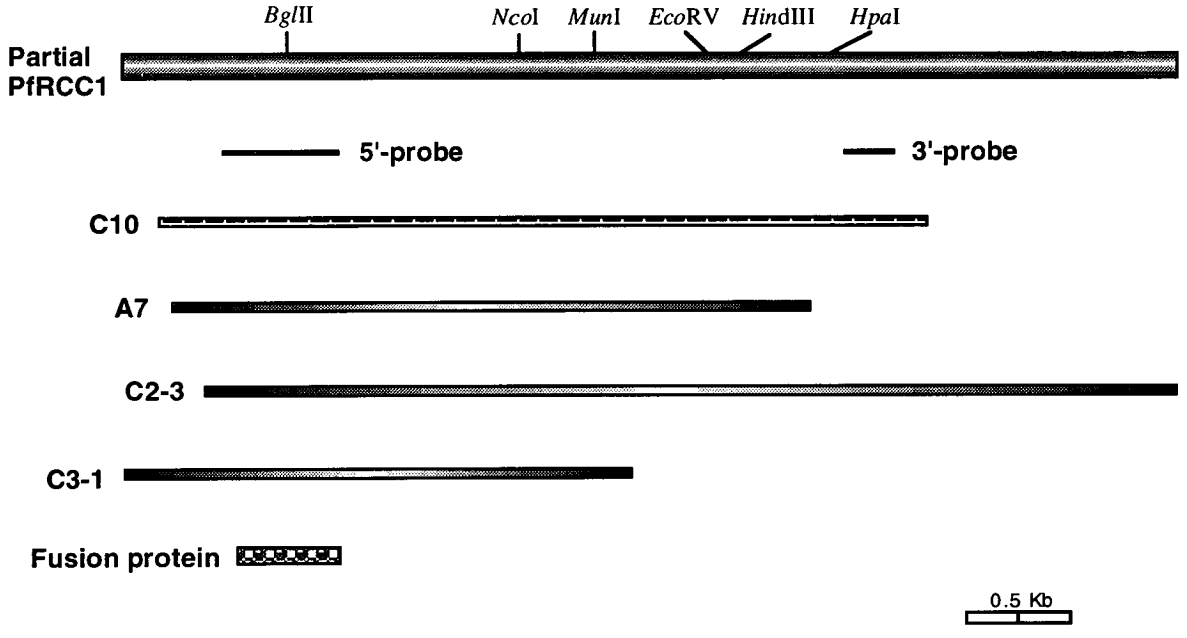


Figure 3.1 Schematic restriction map of the partial PfRCC1 gene compared to the localisation of clones A7, C2-3, C3-1, fusion protein, and 5'-and 3'-probes. Clone C10 was previously cloned by Sultan, 1994.

which has been primed from a short dA-rich 3' end sequence (Figure 3.2). This sequence has been demonstrated to be a single copy gene on chromosome 5 by restriction mapping and Southern blotting experiments (Sultan, 1994). Northern blotting experiments with RNA from synchronised cultures has also indicated that relatively low-level transcription of a single ~8 kb RNA occurs in the trophozoites and to a lesser extent in the schizonts (Sultan, 1994). This clone has been named PfRCC1 (GenBank accession number AF048836) and its size and structure compared to other RCC1 proteins are shown in Figure 3.3.

The Swissprot data base was used to compare PfRCC1 with other RCC1 proteins and the p619 Giant Protein. The human p619 giant protein has 4861 amino acid residues that contains two regions of seven internal RCC1-type repeats and other additional regions (Rosa *et al*, 1996). The alignments around the RCC1-type repeats are depicted using the Pileup programme (Figure 3.4). The alignments of the seven RCC1-type repeats from the human RCC1 protein, the *S.cerevisiae* RCC1 homologue (SRM1), and the two sets of repeats from the p619 protein confirm that PfRCC1 is a member of this family of GEFs. Internal homology of the repeats to each other is low in all of these proteins. Outside the repeat blocks there is no sequence conservation. Sequence insertions, a common feature of *P.falciparum* proteins, occur in two of the seven repeats and in repeat IV, the insertion is 140 amino acids long. Amino acid identity between each sequence in the seven RCC1 repeats is only between 11-15% although similarity in the short sequence motifs themselves rises to around 60%. The single longest block of amino acid identity between the *P. falciparum* sequence and any other is a sequence around the third tandem repeat where 11/18 amino acids (61%) are shared with the p619 Giant Protein. This observation, together with the large size of the PfRCC1 and its slightly higher homology with p619 raises the possibility that the *P.falciparum* sequence is actually a homologue of this cytoplasmic GEF rather than of the intra-nuclear RCC1 proteins. However, the extensive partial sequence obtained does not appear to leave coding space in the 8kb PfRCC1 mRNA for the additional set of RCC1 repeats that occur in p619 (encoded by a 15kb mRNA (Rosa *et al*, 1996)). Low stringency Southern blotting experiments also indicated only one copy of the PfRCC1 repeats exists in the *P.falciparum* genome (Sultan, 1994). Taken together, these experiments indicate that a RCC1-like homologue does exist in the *P. falciparum* parasites and its predicted molecular weight indicates a size in between conventional RCC1 proteins and the p619 proteins.

3.3 Further characterisation of the *P. falciparum* RCC1 protein

To establish the intra-cellular location of PfRCC1, a DNA fragment encoding 139 amino acids (comprising the whole of the first and most of the second RCC1

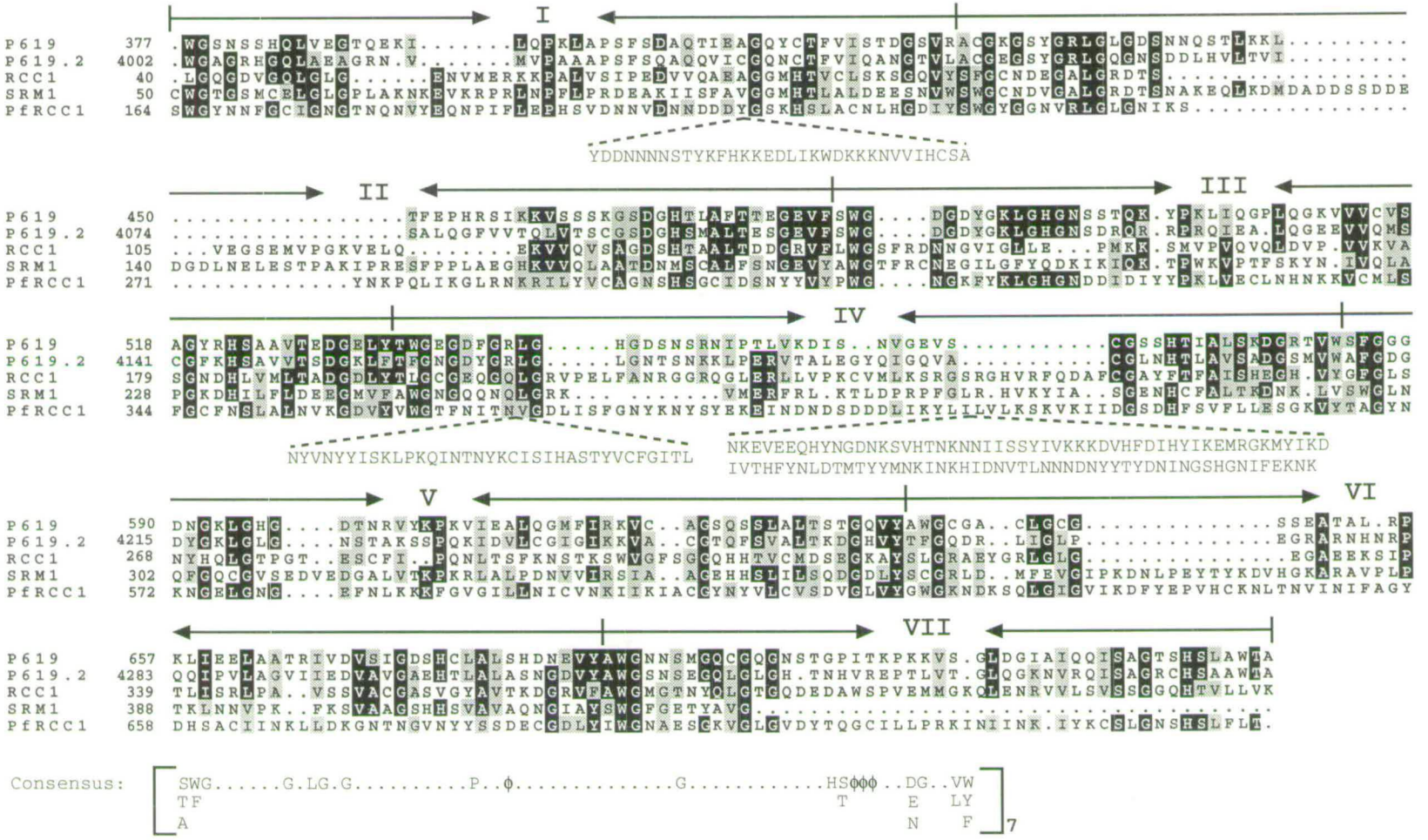

1 GGACATGAGGCAAAATATGTACAAAATAAAAAATTTTTTTACAACGAACATAATTTTAAGAACGATATAGGTCTCCCAAGCGACGACGAT
G H E A K Y V Q N K K F F Y N E H N F K N D I G L P S D D D
91 ATTTATGATGACTACAACATTTCTGGTGACGAAAGTCTTTTAAACTTTGAGGCAGATAATATTTCAAGTATACAAAATGAGGACACAAAT
I Y D D Y N I S G D E S L L N F E A D N I S S I Q N E D T N
181 ATGGATGAAGATGTATACAATAATATTGTTAATGACCATACTAATGAAGACAGACAATCGGATATTAATGAAACTATTGATAAACATAGT
M D E D V Y N N I V N D H T N E D R Q S D I N E T I D K H S
271 AATGATAATGATTATGATGATGAAAAGAAAAAGGAAAAAATGAATATATATAAAAATAAAGAAAAATTATATTCAAATATAAAAGAAAAT
N D N D Y D D E K K K E K M N I Y K N K E K L Y S N I K E N
361 GATAATAAACATGACTGGTTTACTCCATTCCCTATGAAAATAGTTTTCCCTAAAAAAAATACAAAAATAAAATTTATTTCTTGTGGTGAT
D N K H D W F T P F P M K I V F P K K N T K I K F I S C G D
I
451 ATGCACACATTGGCTATCTGCACAGGTGGTGTTTTATATTCCTGGGGTTATAACAATTTTGGTTGTATAGGAAATGGGACAAATCAAAAT
M H T L A I C T G G V L Y S W G Y N N F G C I G N G T N Q N
II
541 GTATATGAACAAAACCCGATATTTTTAGAGCCTCATTCGGTTGATAATAATGTTGATAATAATGATGATGATTATTATGATGATAATAAT
V Y E Q N P I F L E P H S V D N N V D N N D D D Y Y D D N N
631 AATAATAGTACATACAAATTTCATAAAAAGGAAGATCTCATAAAATGGGATAAAAAAAAAAATGTAGTCATACATTGCTCAGCAGGGAGT
N N S T Y K F H K K E D L I K W D K K K N V V I H C S A G S
III
721 AAGCATAGTCTTGCTTGTAACCTACACGGAGATATTTATAGTTGGGGTTATGGAGGTAATGTTAGATTGGGTCTTGGGAATATAAAAAGT
K H S L A C N L H G D I Y S W G Y G G N V R L G L G N I K S
811 TATAATAAGCCCCAATTAATTAAAGGATTAAGAAATAAAAGAATTCTTTATGTATGTGCAGGAAACTCTCATTCAGGATGTATAGATTCT
Y N K P Q L I K G L R N K R I L Y V C A G N S H S G C I D S
IV
901 AATTATTATGTATATCCGTGGGGAAATGGGAAATTTTACAAATTAGGTCATGGTAATGATGATATAGATATATATTATCCTAAATTAGTA
N Y Y V Y P W G N G K F Y K L G H G N D D I D I Y Y P K L V
991 GAATGTTTAAATCATAATAAGAAAGTATGTATGTTAAGTTTTGGGTGTTTTAATTCTTTGGCATTAAATGTTAAAGGAGATGTATATGTA
E C L N H N K K V C M L S F G C P N S L A L N V K G D V Y V
1081 TGGGGTACATTTAATATAACAAATAATTATGTAAATTATTATATTTCTAAATTACCTAAGCAAATTAATACAAATTATAAATGTATATCT
W G T F N I T N N Y V N Y Y I S K L P K Q I N T N Y K C I S
1171 ATACATGCTTCTACATATGTTTGCTTTGGTATTACGTTAGTAGGTGATTTAATTAGTTTTGGGAATTATAAAAACTATTCCTATGAAAAA
I H A S T Y V C F G I T L V G D L I S F G N Y K N Y S Y E K
1261 GAAATAAATGATAATGATTCAGATGACGATTTAATCAAATATTTAATAAATAAAGAAGTTGAAGAGCAACATTATAATGGGGATAATAAA
E I N D N D S D D D L I K Y L I N K E V E E Q H Y N G D N K
1351 AGTGTTCACACAAATAAGAATAACATCATATCATCATATATTGTAAAAAAGAAGGATGTACATTTTGATATACATTATATAAAAGAAATG
S V H T N K N N I I S S Y I V K K K D V H F D I H Y I K E M
1441 AGAGGCAAAATGTATATAAAAGATATTGTAACACATTTTTATAATTTAGATACTATGACATATTATATGAATAAGATAAACAAGCACATT
R G K M Y I K D I V T H F Y N L D T M T Y Y M N K I N K H I
1531 GATAATGTTACCTTAAATAATAATGATAATTACTATACGTATGATAATATAAATGGTTCCCATGGAAACATTTTTGAAAAGAACAAACTT
D N V T L N N N D N Y Y T Y D N I N G S H G N I F E K N K L
1621 GTCTTGAAGTCTAAAGTTAAAATAATAGATGGAAGTGATCATTTTAGTGTATTTTTATTAGAAAGCGGGAAAGTATATACTGCTGGATAT
V L K S K V K I I D G S D H F S V F L L E S G K V Y T A G Y
V
1711 AATAAGAATGGAGAATTGGGAAATGGAGAATTTAATTTAAAGAAAAAATTTGGAGTGGGTATATTATTAAACATATGTGTAAACAAGATA
N K N G E L G N G E F N L K K K F G V G I L L N I C V N K I
1801 ATAAAAATTGCATGTGGATATAATTATGTATTATGTGTAAGTGATGTGGGTTTAGTATATGGCTGGGGAAAAAATGATAAGAGCCAATTG
I K I A C G Y N Y V L C V S D V G L V Y G W G K N D K S Q L
VI
1891 GGAATTGGGGTAATAAAAGATTTTTATGAACCTGTTCATTGTAAGAATTTGACAAATGTGATAAATATATTTGCTGGATATGACCATAGT
G I G V I K D F Y E P V H C K N L T N V I N I F A G Y D H S
1981 GCTTGTATTATAAATAAGTTATTAGATAAGGGTAACACGAATGGTGTGAATTATTATAGTTCAGACGAATGTGGAGATTTATATTATTGG
A C I I N K L L D K G N T N G V N Y Y S S D E C G D L Y I W
VII
2071 GGTAATGCCGAAAGCGGAAAGGTCGGGTTAGGTGTTGATTATACCCAAGGATGTATATTATTACCTAGAAAAATTAATATAATAAATAAA
G N A E S G K V G L G V D Y T Q G C I L L P R K I N I I N K
2161 ATATATAAATGTTCTTTAGGTAATAGTCATAGTTTATTTTTAACGAATACTAACGAATTATACGTTTGTGGAAGTAATAATAATGGTAGG
I Y K C S L G N S H S L F L T N T N E L Y V C G S N N N G R
2251 TTAGGTTTATTTGAAAAATCTAAAATGGTTTGTAAATTTGTTAGTGTACCTACTAAAGTTATATTACATAATAATATATATTTAAGGAT
L G L F E K S K M V C K F V S V P T K V I L H N N I Y I K D
2341 ATCCTAGCAGGTAATACATATAGTATAATATTATCTGTTGATGGGTTTATATATATTTTTGGAGAATTCATAAAAGGTAGTAATGGTAAT
I L A G N T Y S I I L S V D G F I Y I F G E F I K G S N G N
2431 AGTAATAATATGTCATGTAGGGTCATAAGCTTATATAATGAAATTACCAATGTTAAATATATGAATGGGAAATATGAACATCTTTTTTTT
S N N M S C R V I S L Y N E I T N V K Y M N G K Y E H L F F

```
2521 TTTCTTACATATGATAATAAATTATTTGGTATAGGTAATAATAAGAATTGTCAAATATTATGTGATAATGAAAAAGAGAAGAATTATATA
     F  L  T  Y  D  N  K  L  F  G  I  G  N  N  K  N  C  Q  I  L  C  D  N  E  K  E  K  N  Y  I

2611 AAAAAACCTAAATTAATAACATATTTTTTAAAAGAAAATAATAAAATAATTGAATTCATTTTTTCTTTTCATAATGTTACATTTATTCTT
     K  K  P  K  L  I  T  Y  F  L  K  E  N  N  K  I  I  E  F  I  F  S  F  H  N  V  T  F  I  L

2701 TTCAATAATTCAGAAATATTTGTATGGGGTTATACTAATAATTATCATTTATGTATAGGAATACCTAATAATATAAAATATTTAAAACAT
     F  N  N  S  E  I  F  V  W  G  Y  T  N  N  Y  H  L  C  I  G  I  P  N  N  I  K  Y  L  K  H

2791 CCTACGAAAATTATAAAAACATGGTTAACATATGATCAGAAAGATTTAGATTATAATGGAAGTTGTAGTGATATGGAATTGAATGAATAT
     P  T  K  I  I  K  T  W  L  T  Y  D  Q  K  D  L  D  Y  N  G  S  C  S  D  M  E  L  N  E  Y

2881 AATACTACTAATAATATGTCTTATAATAATTATATAAAATATATTAATGAAATTACTCTTGTTAAAAAAAATATAAATACAACAGAACTG
     N  T  T  N  N  M  S  Y  N  N  Y  I  K  Y  I  N  E  I  T  L  V  K  K  N  I  N  T  T  E  L

2971 TTTCAAAATGAATATACTTATTTTGGTTATTATGAAGAAGAAATCGAAAATTTTATATATAATATAGTAATAATGAATAATATTATAAAT
     F  Q  N  E  Y  T  Y  F  G  Y  Y  E  E  E  I  E  N  F  I  Y  N  I  V  I  M  N  N  I  I  N

3061 TGGTGTCATTTACAAAATTTATTAAAAAAAGAAGAATATAATAATAATATGGAATATATCAAATCTTTTGAAAAAGATATAATTGATTTA
     W  C  H  L  Q  N  L  L  K  K  E  E  Y  N  N  N  M  E  Y  I  K  S  F  E  K  D  I  I  D  L

3151 TATTCTAAACATATAGAAGTTTTATTAAATCTAAATATATATGAAAAACAATTTAATTATTTATATCTTAATTATCAAAATTTTATATTA
     Y  S  K  H  I  E  V  L  L  N  L  N  I  Y  E  K  Q  F  N  Y  L  Y  L  N  Y  Q  N  F  I  L

3241 TCGAATATTGCTAATATGAAAGAACAACATCTACCAAATATTTTATATCCTTTCTCAACATATATATTTGATCAAAATAGAATAAAATTA
     S  N  I  A  N  M  K  E  Q  H  L  P  N  I  L  Y  P  F  S  T  Y  I  F  D  Q  N  R  I  K  L

3331 GAACAGTTCATATATATAATTCAACAACAGCCAGTTTATCTAATAATTTTATGTTTAATACATAATTATAAAAATGTTAAAAATGTAAAA
     E  Q  F  I  Y  I  I  Q  Q  Q  P  V  Y  L  I  I  L  C  L  I  H  N  Y  K  N  V  K  N  V  K

3421 CAACTCATGTATGGGAAAATGTCCTACAATTCTAATAAAAATACAACTCATATTTATCATAATAATATAAATGAGCAACATATAAAAAAT
     Q  L  M  Y  G  K  M  S  Y  N  S  N  K  N  T  T  H  I  Y  H  N  N  I  N  E  Q  H  I  K  N

3511 AATTATAATATTATCAACTCAAATTATATTAATCATACATATAATTATGAGCAAAATGATGCTTTAAAAAAATATAATTCAAAAGATATA
     N  Y  N  I  I  N  S  N  Y  I  N  H  T  Y  N  Y  E  Q  N  D  A  L  K  K  Y  N  S  K  D  I

3601 CAAAATAAAAAATATTCCTTTTATAAAAATAGTTCAAATGTTTTATGTTCTTTTATTTTTGATTTATATGCTGATTTTAGAAATGAAAGA
     Q  N  K  K  Y  S  F  Y  K  N  S  S  N  V  L  C  S  F  I  F  D  L  Y  A  D  F  R  N  E  R

3691 GTTCGAACTATTTTTACTATTTTCTTAATAAAATTAGGTATAGAAGAAATCAAAAATTCTTTAAACGTTCATTCTATATTTAATATTGAC
     V  R  T  I  F  T  I  F  L  I  K  L  G  I  E  E  I  K  N  S  L  N  V  H  S  I  F  N  I  D

3781 ACATCTATATTCTTTACTCTTATACAAATGCTTTTCTTCAAAAAAGAGATTCTAATAAATTTTGCTCAGTGTTTAATCGACTTGGATAAT
     T  S  I  F  F  T  L  I  Q  M  L  F  F  K  K  E  I  L  I  N  F  A  Q  C  L  I  D  L  D  N

3871 CCCAATTCTTTTGTTACTTTATTAGACCAATTATCAAGGAAAGTTCCAAAAAAAAAACATTCCCCTTAACCAACAAGGTGTTTCTAATGCA
     P  N  S  F  V  T  L  L  D  Q  L  S  R  K  V  P  K  K  N  I  P  L  N  Q  Q  G  V  S  N  A

3961 AAAAATATGCAGCATACAAAACATCTACAAACGAATACTCCTATGTATGGTTCTGCTCACCACAGGGAAAATTTATTGATAAATGATATG
     K  N  M  Q  H  T  K  H  L  Q  T  N  T  P  M  Y  G  S  A  H  H  R  E  N  L  L  I  N  D  M

4051 TATAATGATCGCATGGTAGTAGATAAAAATATTCATGTGAAGAATTCAACTCATATGGATCTCAAAAATGATGGAGCGAATAAAAAAAAA
     Y  N  D  R  M  V  V  D  K  N  I  H  V  K  N  S  T  H  M  D  L  K  N  D  G  A  N  K  K  K

4141 AAAATAAAAAAAAAAAAAAA
     K  I  K  K  K  K  K
```

Figure 3.2 The partial sequence of the PfRCC1 gene and its encoded protein. The amino acids contained within the shadowed boxes are those of the conserved RCC1-like domains. The clone contains a single contiguous reading frame of 4161 nucleotides encoding a 1387 amino acid protein and the GenBank accession number is AF048836.

Figure 3.3 Schematic diagram of the size and structure of the PfRCC1 predicted protein compared to the human p619 Giant Protein and the yeast and human RCC1 proteins. Each block indicates the seven tandem repeated RCC1 motifs.

Figure 3.4 Alignment of the conserved sequence motifs in the RCC1-type repeats. p619 is the human p619 Giant Protein (Rosa *et al,* 1996), RCC1 is the human RCC1 protein (Ohtsubo *et al*, 1989) and SRM1 is the yeast homologue of RCC1 (Clark & Sprague, 1989). Regions highlighted in black indicate sequence identity, those in grey indicate conservative amino acid substitutions. The seven conserved blocks are numbered and *P. falciparum* specific insertions are shown below Block I and IV. The consensus motif is shown below, ϕ indicating hydrophobic residue.


I
II
III
IV
V
VI
VII
P619 377 .WGSNSSHQLVEGTQEKI.....LQPKLAPSFSDAQTIEAGQYCTFVISTDGSVRACGKGSYGRLGLGDSNNQSTLKKL.....
P619.2 4002 .WGAGRHGQLAEAGRN.V.....MVPAAAPSFSQAQQVICGQNCTFVIQANGTVLACGEGSYGRLGQGNSDDLHVLTVI.....
RCC1 40 .LGQGDVGQLGLG.....ENVMERKKPALVSIPEDVVQAEAGGMHTVCLSKSGQVYSFGCNDEGALGRDTS.....
SRM1 50 CWGTGSMCELGLGPLAKNKEVKRPRLNPFLPRDEAKIISFAVGGMHTLALDEESNVWSWGCNDVGALGRDTSNAKEQLKDMDADDSSDDE
PfRCC1 164 SWGYNNFGCIGNGTNQNVYEQNPIFLEPHSVDNNVDNNDDYGSKHSLACNLHGDIYSWGYGGNVRLGLGNIKS.....
YDDNNNNSTYKFHKKEDLIKWDKKKNVVIHCSA
P619 450 .....TFEPHRSIKKVSSSKGSDGHTLAFTTEGEVFSWG....DGDYGKLGHGNSSTQK.YPKLIQGPLQGKVVVCVS
P619.2 4074 .....SALQGFVVTQLVTSCGSDGHSMALTESGEVFSWG....DGDYGKLGHGNSDRQR.RPRQIEA.LQGEEVVQMS
RCC1 105 ...VEGSEMVPGKVELQ.....EKVVQVSAGDSHTAALTDDGRVFLWGSFRDNNGVIGLLE...PMKK.SMVPVQVQLDVP.VVKVA
SRM1 140 DGDLNELESTPAKIPRESFPPLAEGHKVVQLAATDNMSCALFSNGEVYAWGTFRCNEGILGFYQDKIKIQK.TPWKVPTFSKYN.IVQLA
PfRCC1 271 .....YNKPQLIKGLRNKRILYVCAGNSHSGCIDSNYYVYPWG....NGKFYKLGHGNDDIDIYYPKLVECLNHNKKVCMLS
P619 518 AGYRHSAAVTEDGELYTWGEGDFGRLG.....HGDSNSRNIPTLVKDIS..NVGEVS.....CGSSHTIALSKDGRTVWSFGGG
P619.2 4141 CGFKHSAVVTSDGKLFTFGNGDYGRLG.....LGNTSNKKLPERVTALEGYQIGQVA.....CGLNHTLAVSADGSMVWAFGDG
RCC1 179 SGNDHLVMLTADGDLYTLGCGEQGQLGRVPELFANRGGRQGLERLLVPKCVMLKSRGSRGHVRFQDAFCGAYFTFAISHEGH.VYGFGLS
SRM1 228 PGKDHILFLDEEGMVFAWGNGQQNQLGRK... .....VMERFRL.KTLDPRPFGLR.HVKYIA..SGENHCFALTKDNK.LVSWGLN
PfRCC1 344 FGCFNSLALNVKGDVYVWGTFNITNVGDLISFGNYKNYSYEKEINDNDSDDDLIKYLILVLKSKVKIIDGSDHFSVFLLESGKVYTAGYN
NYVNYYISKLPKQINTNYKCISIHASTYVCFGITL
NKEVEEQHYNGDNKSVHTNKNNIISSYIVKKKDVHFDIHYIKEMRGKMYIKD
IVTHFYNLDTMTYYMNKINKHIDNVTLNNNDNYYTYDNINGSHGNIFEKNK
P619 590 DNGKLGHG....DTNRVYKPKVIEALQGMFIRKVC..AGSQSSLALTSTGQVYAWGCGA..CLGCG.....SSEATAL.RP
P619.2 4215 DYGKLGLG....NSTAKSSPQKIDVLCGIGIKKVA..CGTQFSVALTKDGHVYTFGQDR..LIGLP.....EGRARNHNRP
RCC1 268 NYHQLGTPGT..ESCFI..PQNLTSFKNSTKSWVGFSGCQHHTVCMDSEGKAYSLGRAEYGRLGLG.....EGAEEKSIP
SRM1 302 QFGQCGVSEDVEDGALVTKPKRLALPDNVVIRSIA..AGEHHSLILSQDGDLYSCGRLD..MFEVGIPKDNLPEYTYKDVHGKARAVPLP
PfRCC1 572 KNGELGNG....EFNLKKKFGVGILLNICVNKIIKIACGYNYVLCVSDVGLVYGWGKNDKSQLGIGVIKDFYEPVHCKNLTNVINIFAGY
P619 657 KLIEELAATRIVDVSIGDSHCLALSHDNEVYAWGNNSMGQCGQGNSTGPITKPKKVS.GLDGIAIQQISAGTSHSLAWTA
P619.2 4283 QQIPVLAGVIIEDVAVGAEHTLALASNGDVYAWGSNSEGQLGLGH.TNHVREPTLVT.GLQGKNVRQISAGRCHSAAWTA
RCC1 339 TLISRLPA..VSSVACGASVGYAVTKDGRVFAWGMGTNYQLGTGQDEDAWSPVEMMGKQLENRVVLSVSSGGQHTVLLVK
SRM1 388 TKLNNVPK..FKSVAAGSHHSVAVAQNGIAYSWGFGETYAVG.....
PfRCC1 658 DHSACIINKLLDKGNTNGVNYYSSDECGDLYIWGNAESGKVGLGVDYTQGCILLPRKINIIINK.IYKCSLGNSHSLFLT.
Consensus: SWG.....G.LG.G.....P..ϕ.....G.....HSϕϕϕ..DG..VW
TF T E LY
A N F
7

repeat) was cloned into the pET29b bacterial expression vector. The hexahistidine-tagged fusion protein was induced in *E. coli* BL21 (DE3) and purified by affinity chromatography (Yasukawa *et al*, 1995). The purified protein was injected into rabbits with Titremax adjuvant to raise a specific antiserum against the recombinant protein. Immuno-fluoresence micrographs comparing the localisation of antibodies to PfRCC1 protein in isolated *P. falciparum* nuclei, with nuclear staining with the DNA binding dye DAPI are shown in Figure 3.5. Panels a, c, and e, show the DAPI stained isolated parasite nuclei. Panels b, d and f show the corresponding parasite nuclei stained with fluor labelled pre-immune sera, mouse antiserum against the PfSNF2L chromatin remodelling ATPase (Ji & Arnot, 1997) and the rabbit anti-PfRCC1 antiserum respectively. While the pre-immune sera does not react with the isolated nuclei, antibodies against both the intra-nuclear PfSNF2L and the PfRCC1 react with these parasite nuclei.

To locate PfRCC1 in intact parasites, immuno-fluorescence micrographs with the ring, trophozoite/schizont and segmenting schizont stages of sorbitol synchronised cultures are shown in Figure 3.6. Although the resolution of internal structures within the erythrocytic parasite is at the limits of that obtainable with fluorescence microscopy, it is clear that the fluor-labelled antibody co-localises with segmented schizont nuclei stained with DAPI (panels (g) and (h)) and also with the small DAPI stained nuclei of the early ring stages (panels (c) and (d)) and the larger, probably dividing nuclei present in the late trophozoite/early schizont stages (panels (e) and (f)).

Western blots prepared from asexual stage total parsite extract were incubated with the polyclonal anti-PfRCC1 rabbit serum in order to identify the PfRCC1 protein from the total protein. However, this anti-serum did not cross-react strongly with any band on the Western blots. Multiple very faint bands were shown on the blots (data not shown). It was concluded that the polyclonal rabbit anti-sera was not of sufficient quality to use in Western blotting although it does appear to give specific nuclear staining in IFA experiments. It is possible that the major epitopes being recognised are conformational ones which are not conserved in the more denaturing Western blotting conditions.

3.4 Discussion

The conclusion drawn from these experiments is that we have identified a *P. falciparum* RCC1-type guanidine exchange factor and that antibody localisation studies indicate that PfRCC1 is located in the parasite nucleus. It is therefore likely that this gene encodes the intra-nuclear GEF involved in the Ran mediated nuclear protein import cycle. The seven RCC1-type repeats have also been found in one other protein which also appears to act as a GEF, but in this case for the small G protein ADP-

Figure 3.5 Staining of isolated *P. falciparum* nuclei with antibodies against a characterised intra-nuclear protein (PfSNF2L) and PfRCC1. Parasite nuclei were isolated by saponin lysis of infected erythrocytes followed by cell rupture by homogenisation and separation of parasite nuclei from the external membranes by centrifugation (Lanzer *et al*, 1992). Panels a, c, and e, DAPI stained isolated *P. falciparum* strain 3D7 nuclei from an asynchronous culture. Panel b, the nuclei stained panel a, co-stained with FITC labelled pre-immunisation rabbit serum. Panel d, the field corresponding to panel c, stained with a FITC labelled mouse antiserum against the intra-nuclear chromatin-remodelling complex protein PfSNF2L (Ji & Arnot, 1997). Panel e, parasites stained with DAPI in panel f, co-stained with a FITC labelled rabbit antiserum raised against a recombinant derived fragment of PfRCC1.

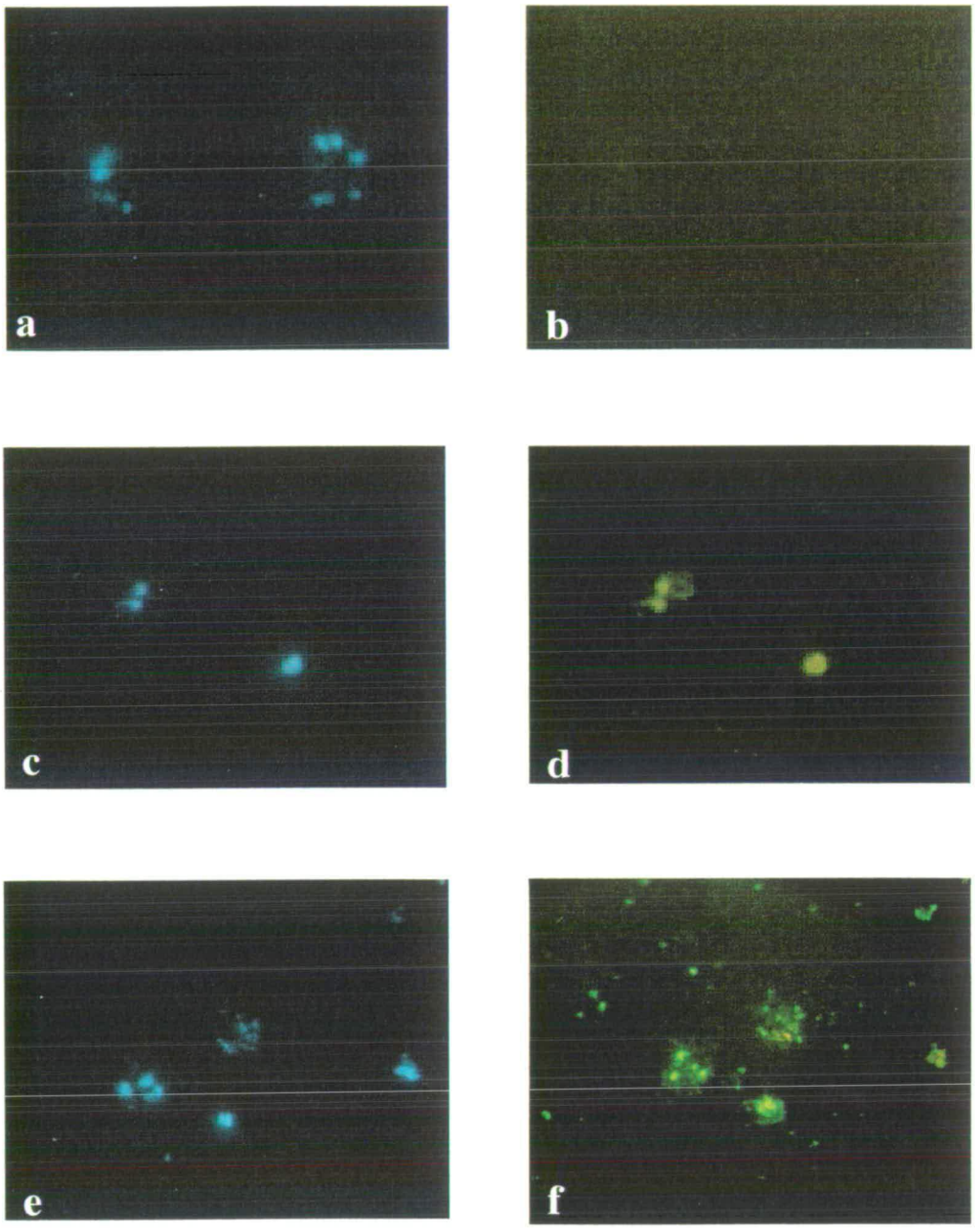
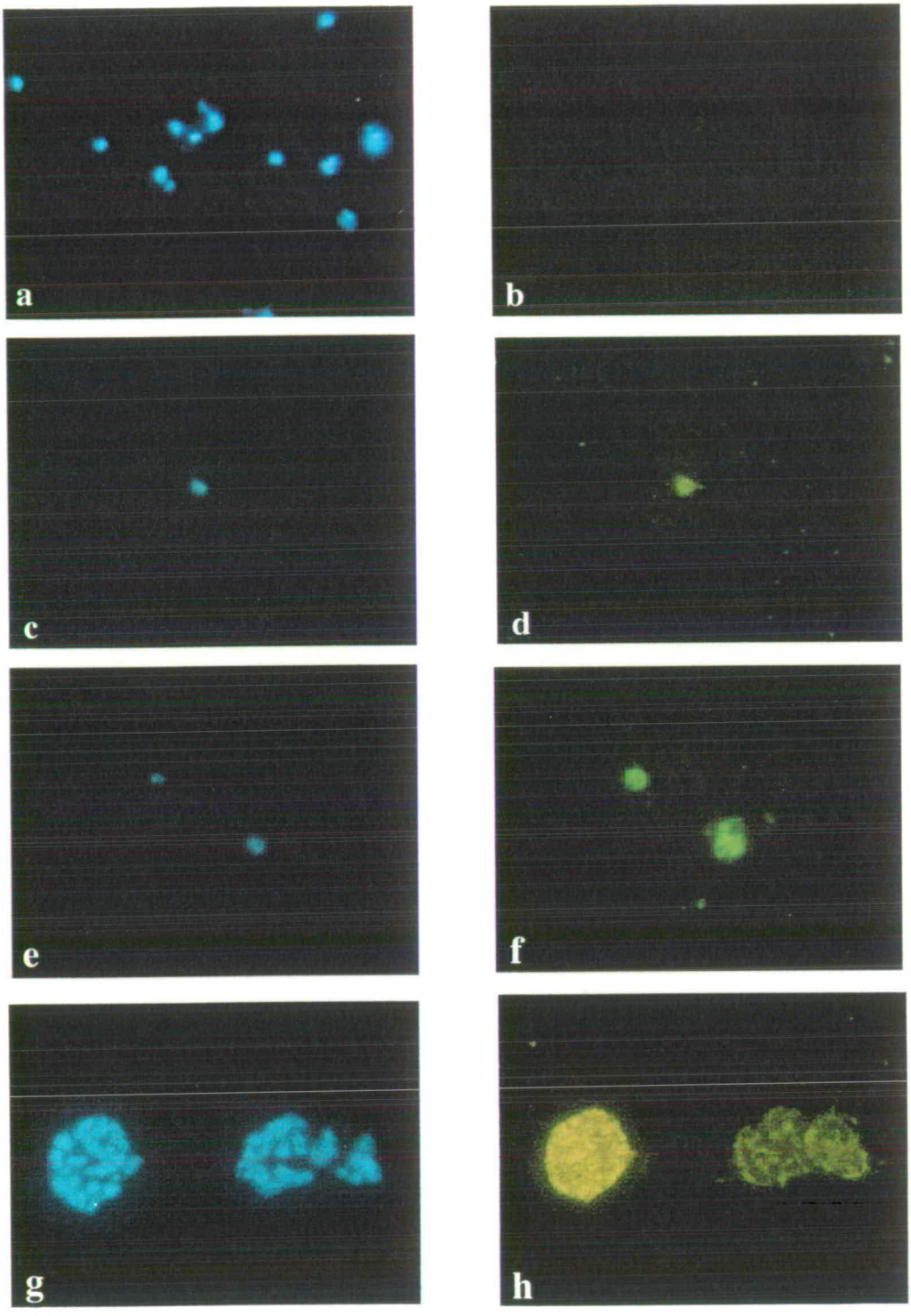

a
b
c
d
e
f

Figure 3.6 Intra-cellular localisation of PfRCC1 by immuno-fluorescence microscopy within different stages of intra-erythrocytic *P. falciparum*. Panels c, e and g. DAPI stained parasite nuclei at the ring, trophozoite/schizont and mature segmented schizont stages of development taken from separate synchronous cultures. Panels d, f and h, Immuno-fluoresence micrographs of the corresponding parasites with FITC labelled polyclonal rabbit antiserum raised against a recombinant fragment of the PfRCC1 protein. Panels a and b, DAPI stained asynchronous parasite nuclei and the nuclei co-stained with FITC labeled pre-immunisation rabbit serum.


a
b
c
d
e
f
g
h

ribosylation factors (ARF's). This extremely large, cytosolyic and Golgi-bound GEF, the p619 Giant Protein, has two sets of RCC1-type repeats, the carboxy-terminal one interacting with myristylated ARF1 in the Golgi and the amino terminal set stimulating guanine nucleotide exchange on ARF1 and Rab but not on the Ran or Ras G proteins (Rosa *et al*, 1996). The alignments of the seven RCC1-type repeats shows that four glycine and one proline residues are invariant within the internal seven repeats and in between the *P. falciparum* RCC1, the human RCC1 , the *S.cerevisiae* SRM1, and the two sets of repeats from the p619 protein. Any changes of these residues replaced by a charged or large hydrophobic group may cause mutants such as temperature sensitive which have already reported by many literatures (Aebi *et al*, 1990; Matsumoto *et al*, 1991; Ohtsubo *et al*, 1987; Renault *et al*, 1998; Sazer & Nurse, 1994). Strikingly, three large sequence insertions have been found in repeats I and IV. However, when comparing the three-dimensional structures of RCC1 proteins (Renault *et al*, 1998), none of these sequence insertions is in the β-strands which appear to constitute the framework of RCC1 architecture. The PfRCC1 is therefore concluded to be a *P. falciparum* homologue of RCC1.

What then is the function of RCC1 in *P. falciparum*? The premature mitotic phenotype of the hamster temperature sensitive (ts) BN2 cell line upon temperature shift led to the suggestion that RCC1 is part of a cell cycle checkpoint mechanism. Loss of RCC1 disrupts DNA replication from undergoing the normal mitosis pathway (Nishitani *et al*, 1991; Sazer *et al*, 1994). At restrictive temperature tsBN2 cells are unable to pass the G1/S transition if shifted during G1 phase and subject to premature chromosome condensation (PCC) and entry into mitosis if shifted during S phase (Nishitani *et al*, 1991). Using DNA-mediated gene transfer to rescue the tsBN2 cell line, the human RCC1 has been shown to regulate mitosis by preventing the activation of $p34^{cdc2}$ histone H1 kinase until DNA replication is complete. No evidence yet indicates that PfRCC1 is involved in the regulation of mitosis schizogony. With the cloned gene it may be possible to examine this question further e.g. by 'gene knock-out'.

In the fission yeast, the temperature sensitive mutant *pim* shows similar cell phenotype to its hamster $RCC1^{ts}$ homologue. The intact nuclear envelope also becomes fragmented following mitosis (Demeter *et al*, 1995). Disruption of nuclear architecture may be the initial effect of perturbation of the RanGTPase system. It is also possible that a structural component of the nucleolus is the primary target of the RanGTPase system and that the abnormal nucleolar structures observed in budding and fission yeast pim^{ts} mutants influence nuclear structure and function (Forrester *et al*, 1992). Matsumoto and Beach (Matsumoto *et al*, 1991) have isolated a *S. pombe* homologue of *Ran*, termed *spi1* (supressor of pim1). Overexpression of the wild-type

allele of spi1 suppresses the pim1 mutant phentype, although this will not rescue null mutants of *pim1*, indicating that spi1 can not completely bypass the requirement for pim1 and that both proteins are required for normal regulation of the cell cycle with respect to completion of DNA replication. From the preliminary results of these immunofluorescence experiments, the PfRCC1 protein has been located in the parasite nuclei and translocated between the parasite cytoplasm and nucleus. In some stages of malaria life cycle, the PfRCC1 protein seems be excluded from the nucleus (data not shown). The possible functions of the PfRCC1 protein in nuclear structure and nuclear transport remain to be determined.

CHAPTER 4
GENE EXPRESSION IN MALARIA PARASITES: CLONING AND CHARACTERISATION OF A *PLASMODIUM FALCIPARUM* TRANSCRIPTION FACTOR

4.1 Introduction

Plasmodium falciparum has a complex life cycle and the control of its gene expression is not understood at even a basic level. A number of preliminary observations have been made. *Plasmodium* mRNA can be translated *in vitro* and is thus inferred to be conventionally capped (Myler *et al*, 1983; Newbold *et al*, 1982). Three distinct transcripts of a 41 kDa blood stage antigen of *P. falciparum* have been found to be alternative spliced in RNA samples prepared from the schizont stage of parasites (Knapp *et al*, 1991). It therefore appears that malarial parasites, like higher eukaryotes, have some capacity to carry out mRNA processing to generate several different versions of a protein from a single transcription unit. Promoter characterisation in *P. falciparum* is difficult as the A+T rich base composition of intergenic regions (~87%) means that the 5' untranslated regions (UTRs) upstream from the initiation codon abound with sequences resembling the TATA motif of most eukaryotic transcription promoters (consensus sequence TATA/TA). However, a transcriptional start site of proliferating cell nuclear antigen (PCNA) has been mapped physically and functionally using a combination of primer extension analysis, ribonuclease protection assay and transient transfection analysis with the firefly luciferase reporter gene. This functional analysis confirms the general location of the mapped transcriptional start site and identifies a region upstream of the start site which is required for promoter efficiency (Horrocks & Kilbey, 1996).

The parasite has a TATA-binding protein and although somewhat divergent from other family members, it retains sequences essential for recognition of DNA (McAndrew *et al*, 1993). No other genes encoding transcription factors or proteins involved in the transcription process have been characterised to date. Using a pair of degenerate primers deduced from amino acid consensus sequences of alignments of the GTP binding sites of G_α protein, part of a novel malarial gene related to the SNF2 (sucrose non-fermenting)-like family of nuclear ATPases involved in transcription was recognised. In order to further characterise the transcriptional process in *P. falciparum* parasites, the whole gene encoding the *P. falciparum* homologue of the SNF2-like protein was cloned for further analysis.

4.2 Isolation and characterisation of the PfSNF2L gene cDNA clones

The *PfSNF2L* gene was identified during attempts to clone *P. falciparum* G-protein genes. The original PCR reactions were performed by using primers (GP1 and GP3) deduced from conserved GTP-binding domains of G_α proteins. A 198 bp fragment was amplified and its sequence revealed similarities to the conserved regions of the SNF2/SWI2 transcription activation factors when analysed using the BLAST search program in the NCBI net service. In attempts to clone members of the G-protein gene superfamily, we have amplified several sequences with homologies to ATPase domains. Given the importance of SNF2/SWI2 proteins in transcription and our interest in developmental gene expression in *Plasmodium*, the gene was analysed further by using the amplified fragment to screen a *P. falciparum* plasmid cDNA library constructed in plasmid vector pJFE14 (see Appendix C) from asexual stage mRNA (late rings/early trophozoites) of the *P. falciparum* clone ITO4. Each positive was repeat screened three times (see Figure 4.1 for example). An inverse PCR was also launched in order to walk through the poly(A) region in mid of the *PfSNF2L* gene (Figure 4.2). In five rounds of screening, over 65,000 colonies were screened. Ten overlapping cDNA clones were isolated and their restriction maps are shown in Figure 4.2. Each clone was sequenced on both strands using the dye dideoxynucleotide chain termination method with the ABI 377 automated sequencer. Sequences derived from these overlapping clones were combined to generate a single complete sequence of the gene which was named *pfSNF2L.*

4.3 The structure of the PfSNF2L gene

The complete sequence of *PfSNF2L* has 519 nucleotides of 5'-UTR (86.5% A+T) upstream of a putative initiation codon followed by a 4,269 nucleotide open reading frame (ORF) (74.1% A+T) and 484 nucleotides of 3'UTR (87% A+T). This open reading frame encodes a 1422 amino acid polypeptide with a molecular weight of 167 kDa (pI = 6.38). The A/T ratio of the *PfSNF2L* gene in the sense strand is 1.25; nearly all *P. falciparum* genes analysed to date have values >1.00 of A/T ratio (Weber, 1987). The exact position of the translation initiation codon is uncertain as there are four in-frame methionine codons in the region 520-595 and alignment with other members of the gene family does not resolve this problem since there is heterogeneity in the amino terminal sequences of the SNF2-like proteins. The most likely initiation site is the first of these in-frame ATG codons since *P. falciparum* 5' UTR regions as a general rule do not maintain in-frame open reading frames. Sequences around this codon are not in particularly good agreement with the 'Kozak rules' for eukaryotic translation initiation sequences, but *P. falciparum* genes in general conform poorly to this translational initiation consensus which is derived from higher eukaryotic genes (Kozak, 1984). The nucleotide and translated amino acid sequence of this protein,

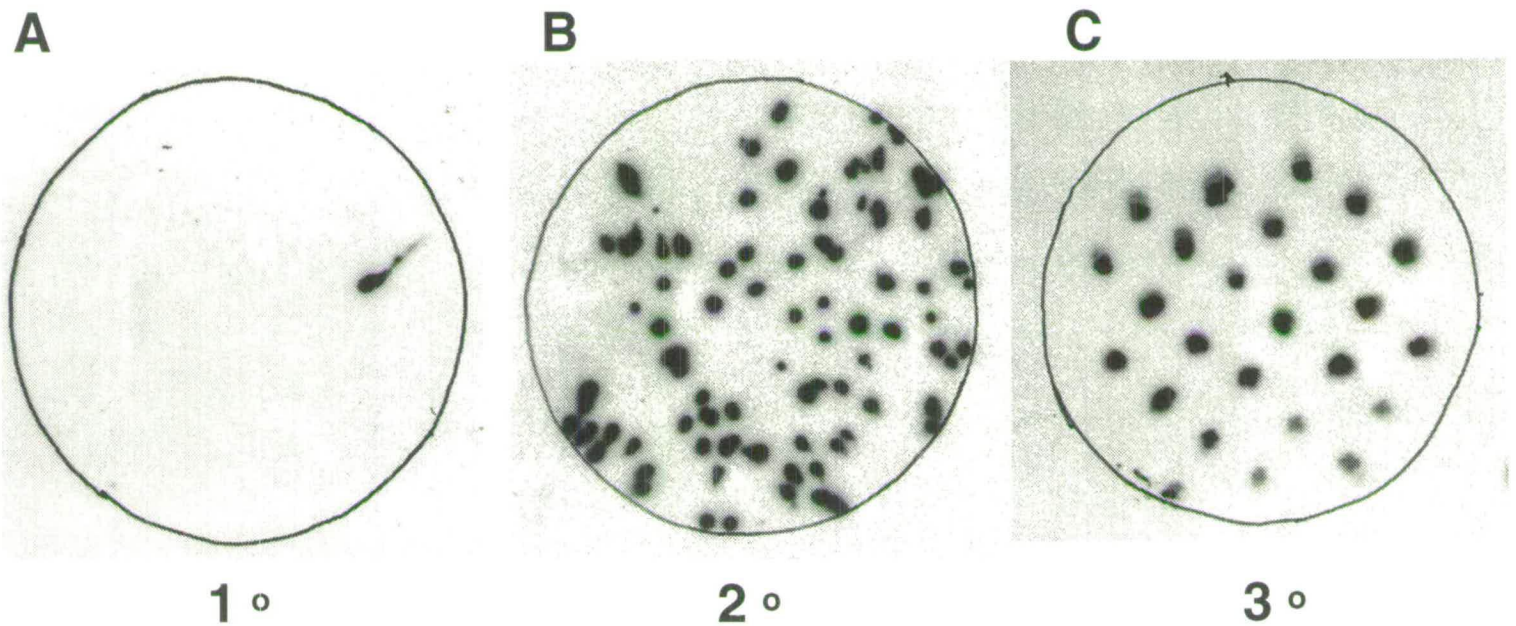


Figure 4.1 An autoradiograph showing the results of first, secondary and tertiary screening of a positive clone of *P. falciparum SNF2L* homologue isolated from the asexual cDNA library, probed with the PCR product produced by the amplifiction of clone gp7.1.

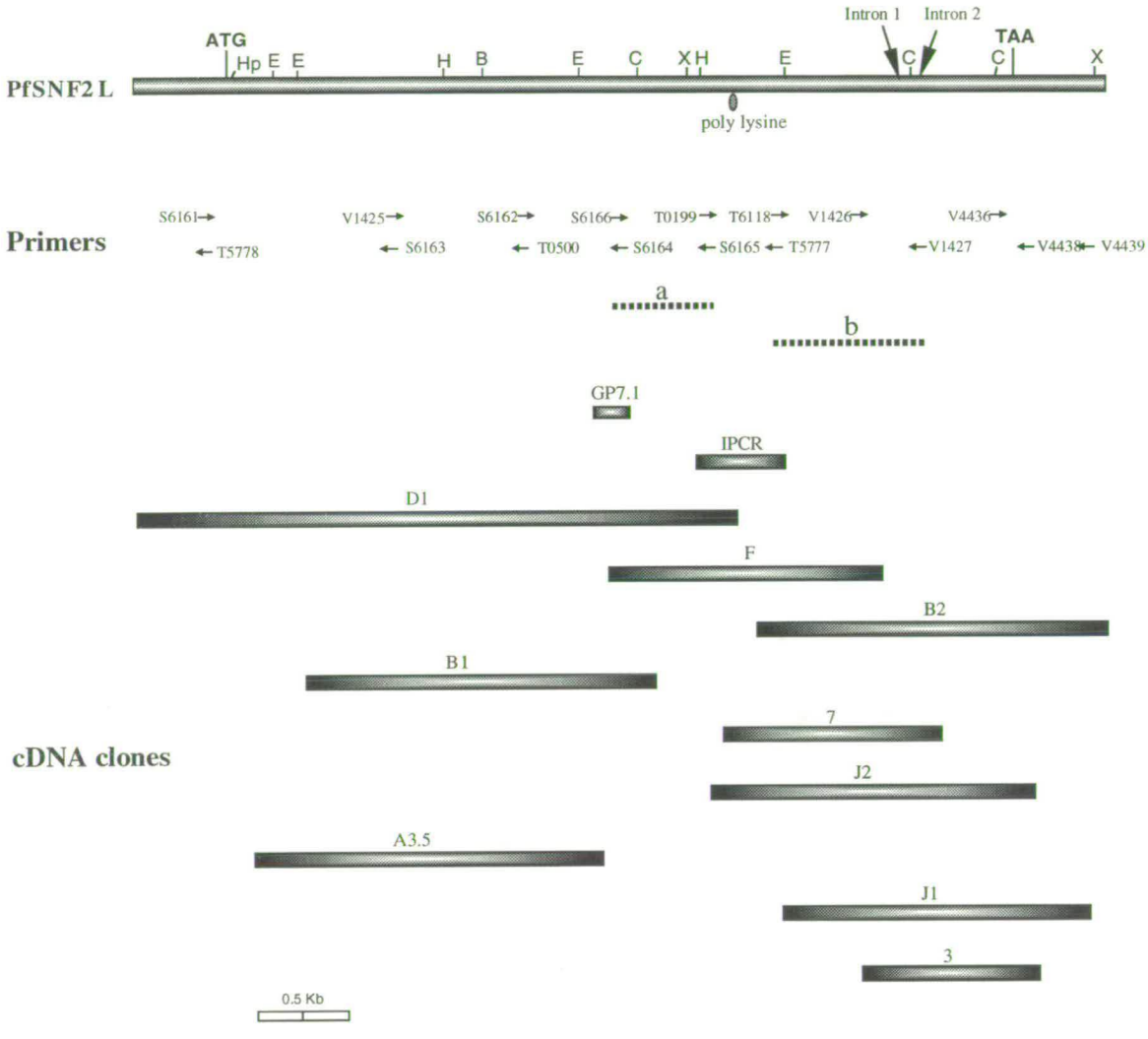


Figure 4.2 Restriction map of the *PfSNF2L* gene. The oligonucleotide primers used and their locations are underlined with arrowhead. Ten cDNA and PCR clones depicted by shaded bar are aligned together in order to indicate their locations. PCR fragments a and b depicted by dashed lines were used as the hybridisation probe for cloning and further characterisations of *PfSNF2L* gene. E. *Eco* RI; H, *Hind* III; B, *Bam* HI; C, *Cla* I; X, *Xba* I; Hp, *Hpa* I.

based on assigning the first in-frame ATG as the initiation codon, is shown in Figure 4.3.

The amino acid sequence showed that the highest levels of sequence similarity in BLAST searches were with domains conserved between members of the SNF2 and SNF2-like (SNF2L) family of transcription factors (Eisen *et al*, 1995). These proteins contain seven conserved regions which constitute the catalytic site of a DNA-dependent ATPase which also has a DNA unwinding activity. These seven regions are boxed in Figure 4.3. The alignment of the *P. falciparum* protein, which we have named SNF2L, with the *Drosophila* (Elfring *et al*, 1994), human (Okabe *et al*, 1992), *C. elegans* (unpublished, Genebank ref. P41877) and yeast (unpublished, Genebank ref. P31844) SNF2-like proteins is shown in Figure 4.4A.

Motif I, Ia and II match the consensus sequence of the bipartite nucleotide triphosphate binding motif (A box and B box motifs) found in both ATP and GTP-binding proteins (Wittinghofer & Pai, 1991) and are required for the transcriptional activation of SNF2/SWI2 protein (Laurent *et al*, 1993). The DNA-dependent ATPase/ DNA helicase domain of all the SNF2-like proteins, including *P. falciparum* sequence, is shorter than that of the corresponding domains of SNF2 proteins proper (the human and *Drosophila brm* proteins and the yeast SNF2/SWI2 protein). As in all members of the SNF2L subfamily, *P. falciparum* SNF2L has the DEAH amino acid motif, whereas the SNF2 proteins have a DEGH motif. SNF2L proteins also lack the other conserved domains found in *brm* and SNF2/SWI2, the so-called 'bromodomain', a sequence of unknown function postulated to interact with other proteins in the transcriptional activation complex.

In addition to the regions involved in ATPase activity the *P. falciparum* protein also contains a SANT domain which is characteristic of the other SNF2-like proteins, but not the SNF proteins (Figure 4.4B). The function of the SANT conserved region is unclear but it occurs in several proteins found in transcription complexes such as TFIIIB. The domain organisation of the PfSNF2L protein is compared to other SANT containing proteins and shown in Figure 4.5. Two regions of the *P. falciparum* SNF2L sequence are not found in other members of the family, although they are common in *Plasmodium* proteins. Eight tandem repeats of a NNHNDD sequence occur in the amino terminal domain and between the DNA-dependent ATPase domain and the SANT domain, the protein contains a region of 210 lysine rich, basic amino acids. Both of these insertions, which make the *P. falciparum* SNF2L substantially bigger than its higher eukaryotic homologues, are unique. Like most other *Plasmodium* proteins, *P. falciparum* SNF2L is an asparagine (12.4%) and lysine (11.5%) rich protein. The analysis of secondary structure and hydrophobicity (Figure

Figure 4.3 The complete sequence of the *PfSNF2L* gene and its encoded protein. The amino acids contained within the solid boxes are those of the conserved ATPase domains. The carboxy terminal sequences underlined by the solid bar represent the SANT domain characteristic of the SNF2-like proteins. The sequences highlighted with dashed line boxes are the asparagine-rich repeated hexapeptides unique to the *P. falciparum* gene. Sequences underlined by dashed lines indicate the position of charged residue clusters. The clone contains a single open reading frame of 4269 nucleotides encoding a 1,422 amino acid protein with an estimated 167 kDa molecular mass. The termination codon is marked by an asterisk and the GenBank accession number is AF003086.

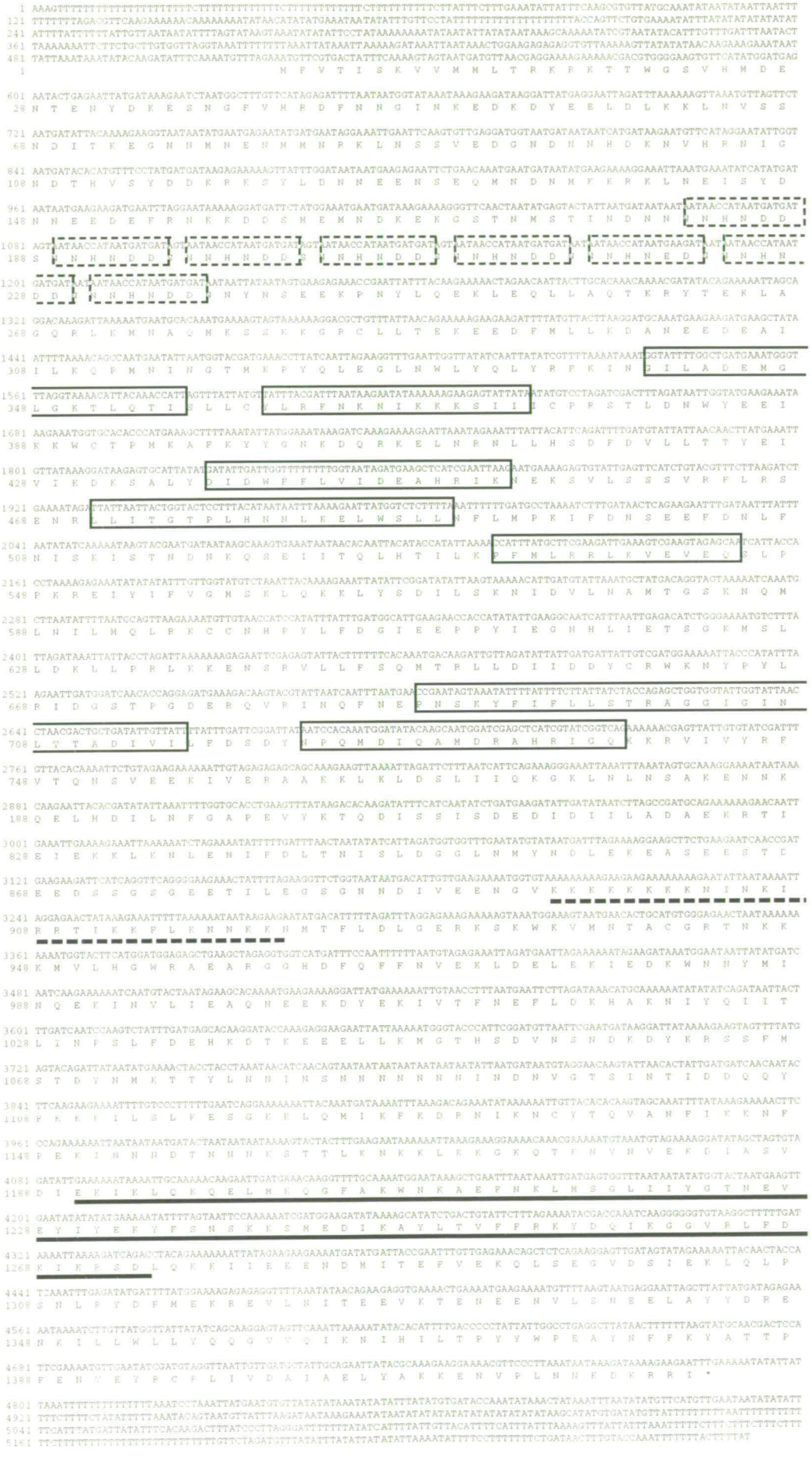
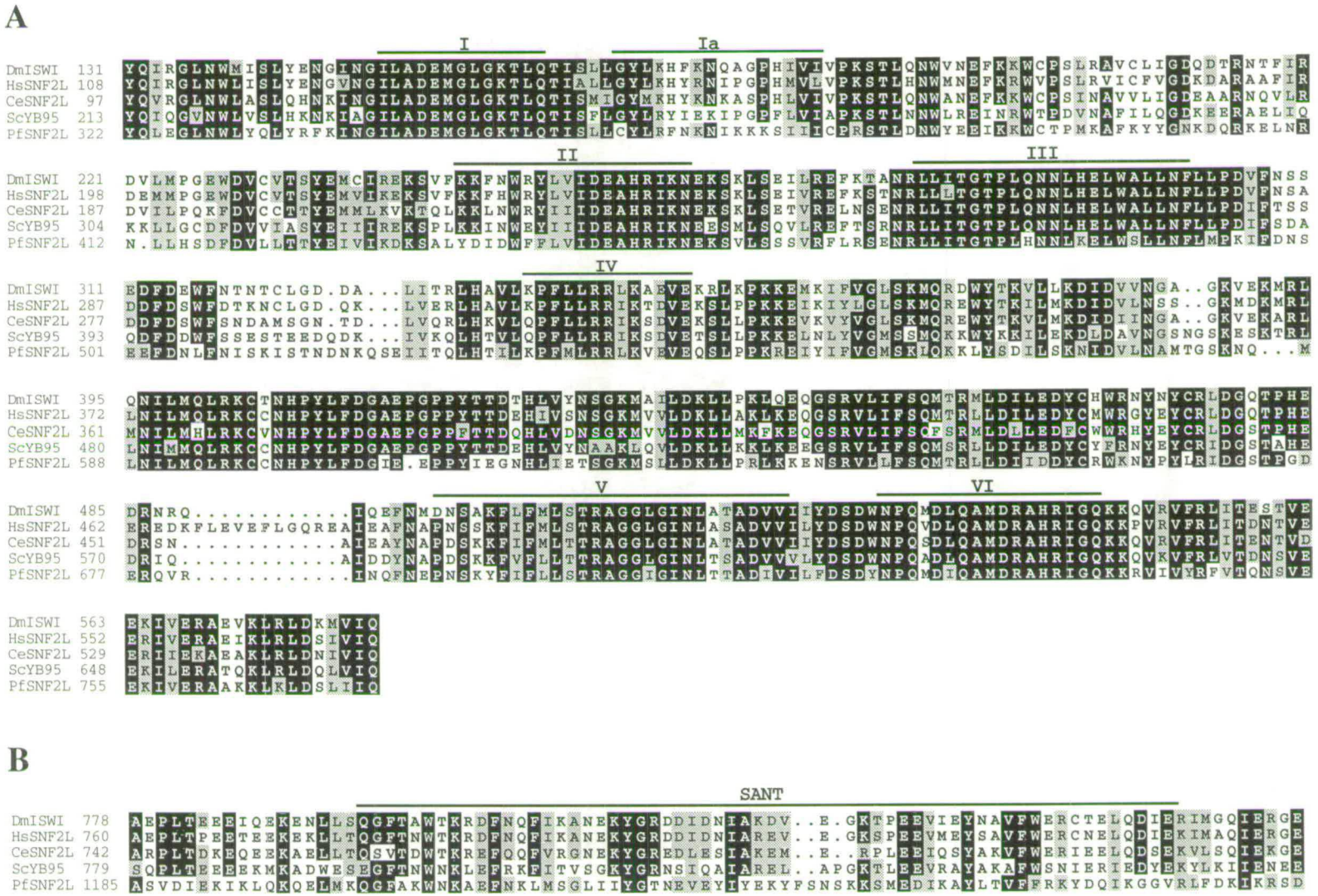


Figure 4.4 **A**. Alignment of the conserved sequences of the SNF2-like protein subfamily. ISWI is the *Drosophila* protein (Elfring *et al*, 1994), HuSNF2L is the human protein (Okabe *et al,* 1992), CeSNF2L is the *C. elegans* protein, (unpublished, Genebank ref.P41877), and ScYB95 is the budding yeast protein, (unpublished, Genebank ref. P31844). Regions highlighted in black indicate sequence identity, those highlighted in grey indicate conservative amino acid substitutions. The overlined position of the seven conserved motifs in the region encoding the framework of the DNA-dependent ATPase/DNA helicase domain is shown. **B**. The evolutionary conservation of the SANT putative DNA binding domain.


A
I
Ia
II
III
IV
V
VI
DmISWI 131
HsSNF2L 108
CeSNF2L 97
ScYB95 213
PfSNF2L 322
DmISWI 221
HsSNF2L 198
CeSNF2L 187
ScYB95 304
PfSNF2L 412
DmISWI 311
HsSNF2L 287
CeSNF2L 277
ScYB95 393
PfSNF2L 501
DmISWI 395
HsSNF2L 372
CeSNF2L 361
ScYB95 480
PfSNF2L 588
DmISWI 485
HsSNF2L 462
CeSNF2L 451
ScYB95 570
PfSNF2L 677
DmISWI 563
HsSNF2L 552
CeSNF2L 529
ScYB95 648
PfSNF2L 755
B
SANT
DmISWI 778
HsSNF2L 760
CeSNF2L 742
ScYB95 779
PfSNF2L 1185

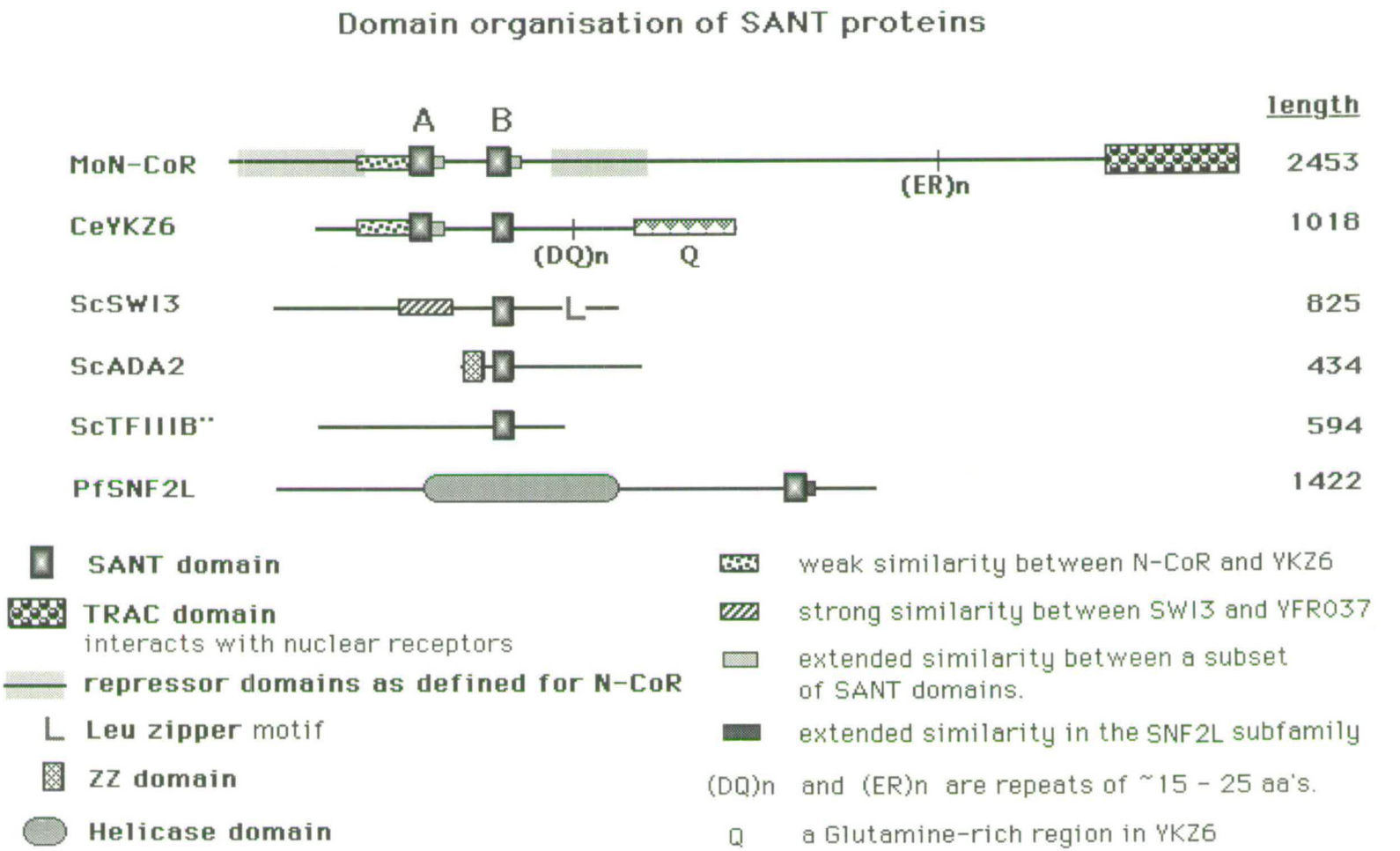


Figure 4.5 The alingment of the SANT-domain proteins and their organisation. MoN-CoR, mouse co-repressor; CeYKZ6, *Caenorhabditis elegens* predicted protein YKZ6; ScSWI3, *Saccharomyces cerevisiae* mating type switch protein 3; ScADA2, *S. cerevisiae* transcriptional factor ADA2; ScTFIIIB", *S. cerevisiae* transcription factor III B subunit; PfSNF2L, *Plasmodium falciparum* SNF2 (sucrose non-fermenting)-like protein. Adapted and modified from Aasland *et al,* 1996.

4.6) shows that the deduced PfSNF2L protein is a very hydrophilic protein except for its ATPase/helicase domain and the amino-terminal end sequences.

GenBank homology searches indicate that *P. falciparum* SNF2L is most similar to proteins of the SNF2L subfamily of the SNF2 transcription factors. *P. falciparum* SNF2L proteins are 37.6% identical and 59.8% similar to yeast YB95 SNF2L protein, 40.6% identical and 63.8% similar to the human SNF2L protein, 38.1% identical and 61.6% similar to *C. elegans* SNF2L protein, and 40.4% identical and 62.5% similar to *Drosophila* ISWI protein. Figure 4.7 compares the structure, sizes and conserved motifs of these proteins. The primary regions of homology, the seven consecutive motifs characteristic of DNA-dependent ATPase/DNA helicase and the C-terminal SANT domain were used to align the five sequences.

4.4 Phylogenetic analysis of *PfSNF2L* gene

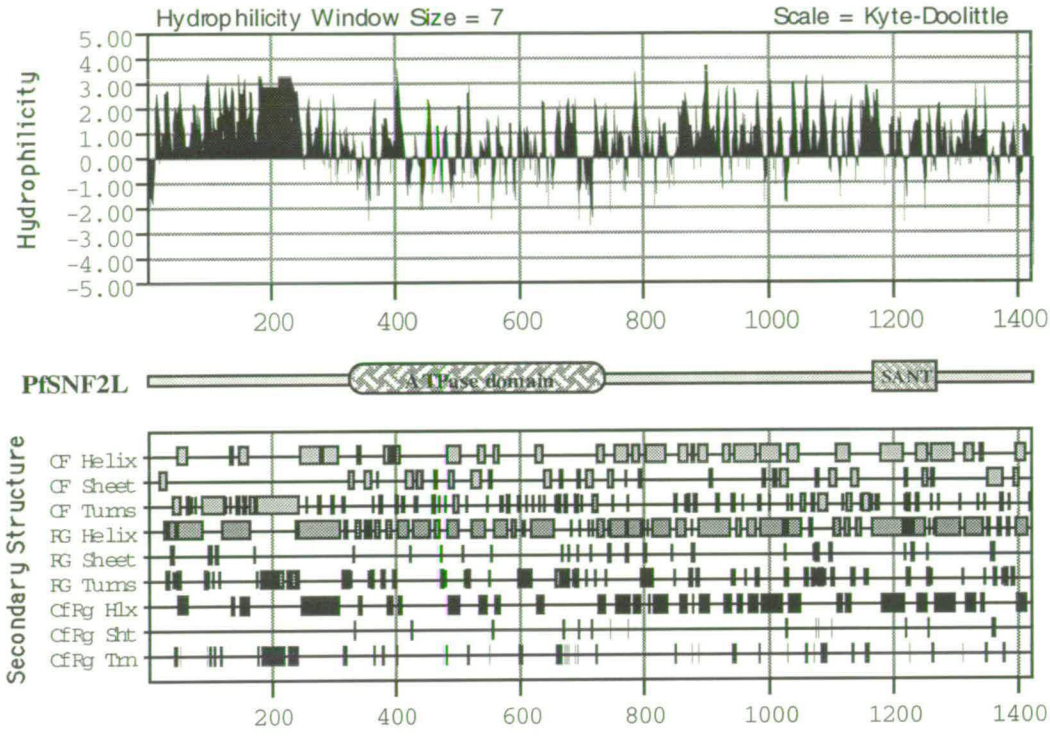
Two phylogenetic trees of *PfSNF2L* were generated depended on the ATPase and the SANT domains by using multiple Neighbour-Joining and Parsimony based methods in the PAUP computer software package. A tree of SNF2 family was created by comparisons of the ATPase regions conserved within the SNF2L subfamily members and with the typical representatives of each SNF2 subfamilies (Figure 4.8). Less conserved regions (such as the regions flanking the ATPase domain and the variable spacer regions) were not used because of problems in obtaining unambiguous alignments in these regions and because there is no established method of scoring alignment gaps in phylogenetic reconstruction. Bootstrap values for each node are shown on the tree. The bootstrap values at the divergent points of the group of SNF2L subfamily members with other subfamily members are very high, which indicate a high degree of confidence in branch positions. The PfSNF2L is obviously grouped with other members of SNF2L subfamily with a high bootstrap value and separated from other members of SNF2 families. This indicates that PfSNF2L is truly a member of SNF2L subfamily. Another tree of the SANT containing proteins was generated by comparison of only the SANT domains conserved among the SANT and MYB proteins by Neighbour-Joining method (Figure 4.9). The result also shows that PfSNF2L is grouped with the members of SNF2L subfamily and distinguished from other SANT containing proteins.

4.5 *PfSNF2L* gene copy number and chromosome localisation

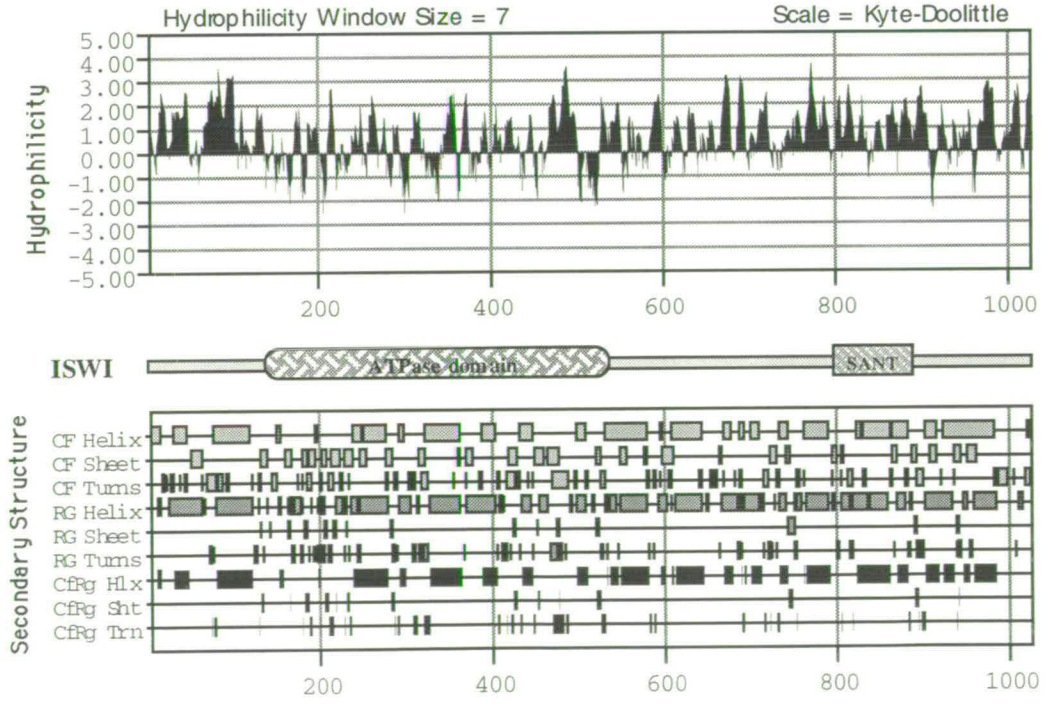
Genomic DNA was digested with restriction enzymes and analysed by Southern blotting (Figure 4.10A). All lanes containing DNA digested with restriction enzymes that do not cleave within the probe show hybridisation to a single fragment. DNA digested with enzymes which cleave within the probe yield more than one

Figure 4.6 The hydrophilicity plot and predicted secondary structure of the deduced protein encoded by (A) *PfSNF2L* and (B) *DmISWI*. Analyses of hydrophilicity and secondary structure prediction were performed using the MacVector 4.5.1 software. Secondary structure was predicted by Chou-Fasman (CF) and Robson-Garnier (RG) methods. The hydrophilicity plot was scaled by the Kyte-Doolittle method.

A


Hydrophilicity Window Size = 7
Scale = Kyte-Doolittle
Hydrophilicity
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-5.00
200
400
600
800
1000
1200
1400
PfSNF2L
ATPase domain
SANT
Secondary Structure
CF Helix
CF Sheet
CF Turns
RG Helix
RG Sheet
RG Turns
CfRg Hlx
CfRg Sht
CfRg Trn

B


Hydrophilicity Window Size = 7
Scale = Kyte-Doolittle
Hydrophilicity
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-5.00
200
400
600
800
1000
ISWI
ATPase domain
SANT
Secondary Structure
CF Helix
CF Sheet
CF Turns
RG Helix
RG Sheet
RG Turns
CfRg Hlx
CfRg Sht
CfRg Trn

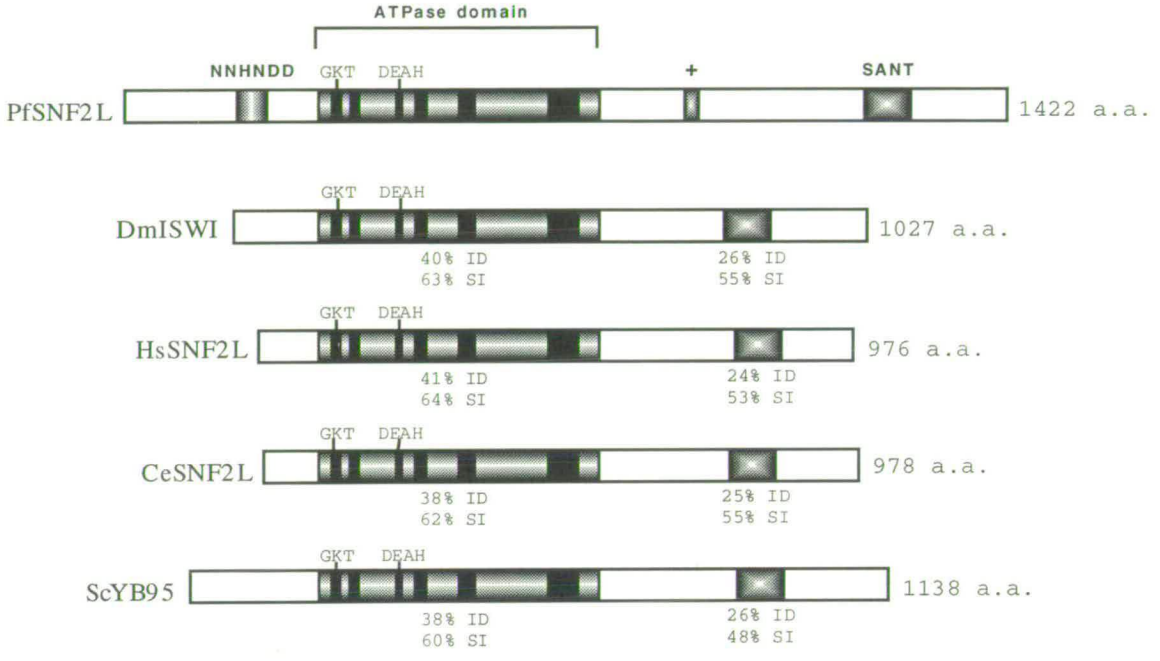


Figure 4.7 Schematic alignment of the members of the SNF2L subfamily. Different shaded boxes represent each different domain described above the alignment. Domains are indicated: NNHNDD, eight repeated NNHNDD motifs; ATPase domain, DNA-dependant ATPase domain; +, highly basic region; SANT, the SANT domain. The positions of the bipartite ATP-binding site and other highly conserved motifs within the ATPase domain are marked by black blocks. The percent identity (ID) of ATPase and SANT domains between PfSNF2L and other SNF2L members are marked under each aligned bar. SI, similarity; a.a., amino acid.

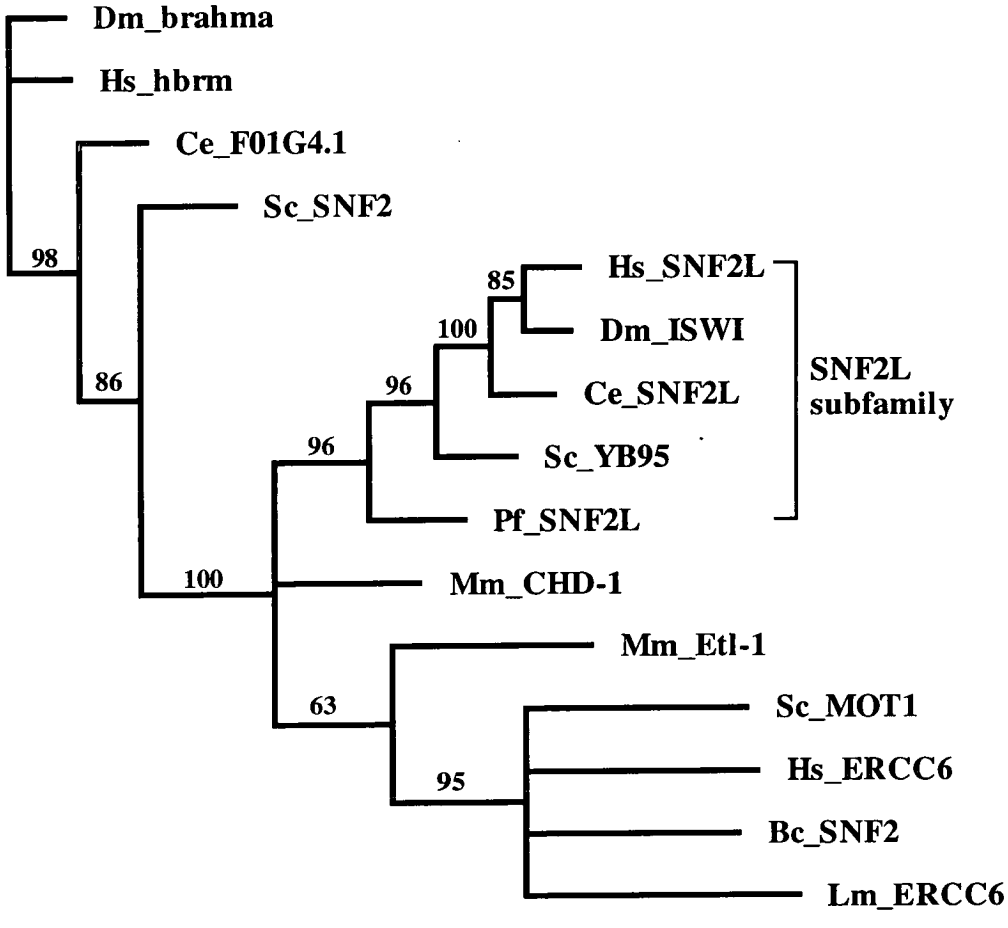


Figure 4.8 Phylogenetic tree of ATPase domains of SNF2 families. This is rooted tree generated with the program PAUP using maximum parsimony method. The bootstrap values are shown at the major points of divergence. Hs, *Homo sapiens*; Mm, *Mus musculus*; Sc, *Saccharomyces cerevisiae*, Dm, *Drosophila melanogaster*; Bc, *Bucillus cereus*; Lm, *Leishmania major*; Ce, *Caenorhabditis elegens*. Accession numbers: Dm_ brahma, M85049; Hs_hbrm, X72889; Ce_F01G4.1, Z68341; Sc_SNF2, M61703; Hs_SNF2L, M89907; Dm_ISWI, L27127; Ce_ SNF2L, P41887; Sc_YB95, Z36114; Pf_SNF2L, AF003086; Mm_CHD1, L10410; Mm_ETl-1, X69942; Sc_MOT1, M83224; Hs_ERCC6, L04791; Bc_SNF2, X98455; Lm_ERCC6, U60409.

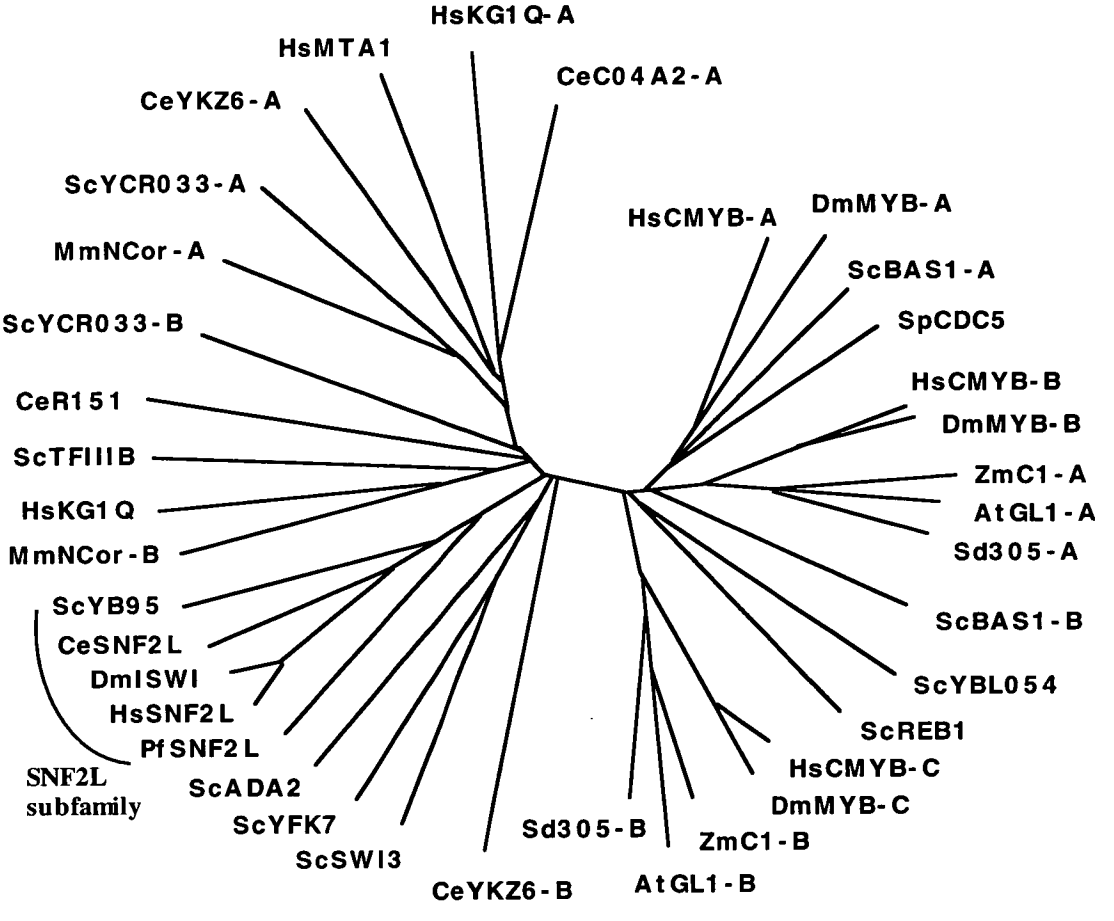
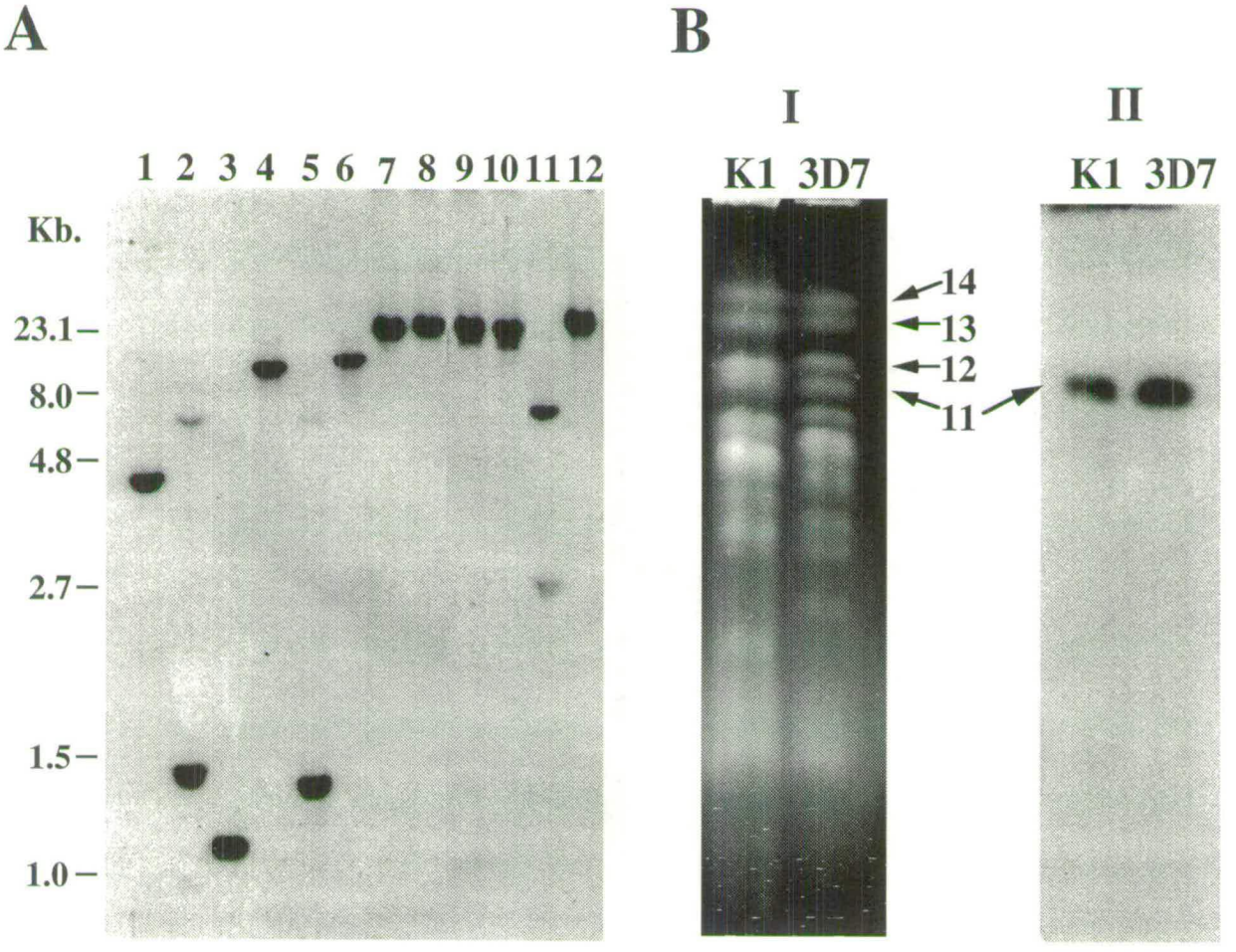


Figure 4.9 Phylogenetic tree of SANT domains and Myb domains. This is an unrooted tree generated with the program PAUP using the Neighbour-Joining method. The sequences used were (protein, database:identifier): MoNCOR, trn:MM35312_1; HuKG1Q, tr:HSORFKG1Q_1; ScSWI3, sw:SWI3_YEAST; HuMTA1, trn: HS3511310_1; CeYKZ6, sw:YKZ6_CAEEL; ScYCR033, sw:YCS3_YEAST; CeC04A2, tr:CEC04A2_2; ScTFIIIB, trn:SC31819_1; CeR151, tr:CER151_9; ScADA2, sw:ADA2_yeast; HuCMYB, sw:MYB_human (abbreviations : sw: sw issprot, tr, and trn: the translated versions of the EMBL and EMBLNEW databases . Species names are indicated as: Mm, mouse, Hs, human; Ce, *C. elegans*; Sc, *S. cerevisiae*). Adapted and modified from Aasland et al, 1996. http://www.uib.no/aasland/niph-tree.html

Figure 4.10 Copy number and genomic localisation of the *PfSNF2L* gene. **A.** Southern Blot of *P.falciparum* (clone 3D7A) DNA cut with the restriction enzymes *Bgl* II, *Cla* I, *Eco* RI, *Eco* RV, *Hin*d III, *Hpa* I, *Mlu* II, *Nar* I, *Pst* I, *Pvu* II, *Xba* I, and *Xho* I (lanes 1-12 respectively) and probed with a labelled fragmentof the cDNA clone (as shown in Figure 4.2, probe a). **BI.** Ethidium stained pulsed field gel separation of *P.falciparum* chromosomes from the K1 and 3D7 lines. **BII.** Southern blot of the same chromosome separation transferred to nylon membranes and probed with a labelled fragment of the cDNA clone.


A
1 2 3 4 5 6 7 8 9 10 11 12
Kb.
23.1
8.0
4.8
2.7
1.5
1.0
B
I
K1 3D7
14
13
12
11
II
K1 3D7

hybridising fragment, whose sizes correspond to the sizes predicted from the sequence. No further hybridising bands appeared after longer exposures with high stringently washed filters (0.1 x SSC, 0.1% SDS at 56°C), this shows that *PfSNF2L* is a single copy sequence without closely related (>80% homology) sequences. The same blot was stripped, reprobed and washed under lower stringency (0.4 x SSC, 0.1% SDS at 52°C) and again a pattern of hybridisation consistant with only a single cross-hybridising gene was observed (Data not shown).

The *P. falciparum* chromosomes were resolved on the CHEF gel system and the blots were hybridised and washed using the same probe as for Southern blotting. The PfSN2L gene was found to be located on chromosome 11 of lines K1 and 3D7A (Figure 4.10B), according to the definition of chromosome numbers proposed by Kemp *et al* (Kemp *et al*, 1987). Other genes known to reside on chromosome 11 are: PfRan, RESA2, pF118, pF332, p7-14, Ag840 and pMBA9 (Triglia *et al*, 1992).

4.6 The *PfSNF2L* intron mapping

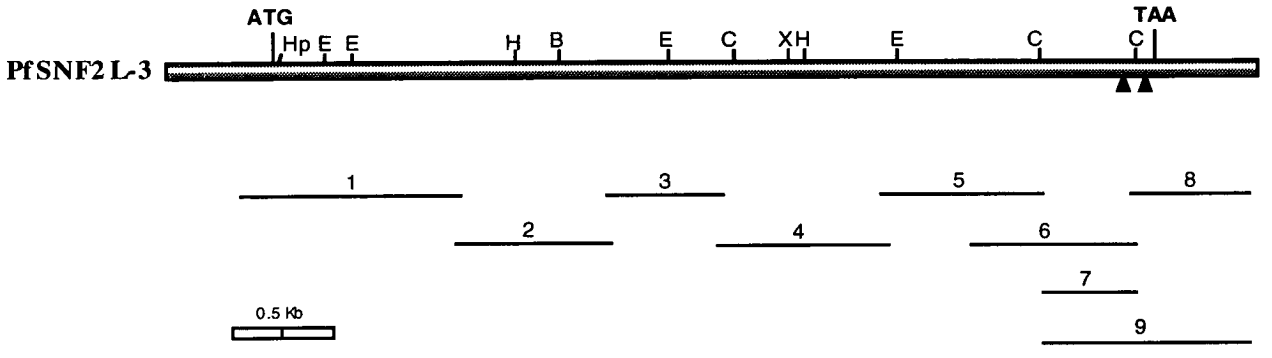
Discrepancies between the size of PCR products derived from genomic DNA and cDNA indicated that the gene has at least one intron. Differential splicing in *Plasmodium falciparum* has been noted (Knapp *et al*, 1991) and since processing of transcription factor mRNA might regulate expression, the introns of this gene were cloned, located and sequenced. Using nine pairs of PCR primers and comparing genomic DNA, the cDNA clone and the RT-PCR product, the intron-exon boundaries of the *SNF2L* gene were mapped. The intron positions and sequences are shown in Figure 4.11. Two introns were 132 and 162 bp in size with 87.1% and 87.6% A+T. The splice sites of these introns closely confirmed the mammalian consensus splice sites (GT-AG rule). No conservation of branch point sequences was observed in the two introns of *PfSNF2L* which is in agreement with observations made in the *Pfexp-1* gene (Simmons *et al*, 1987).

4.7 *PfSNF2L* transcription during development

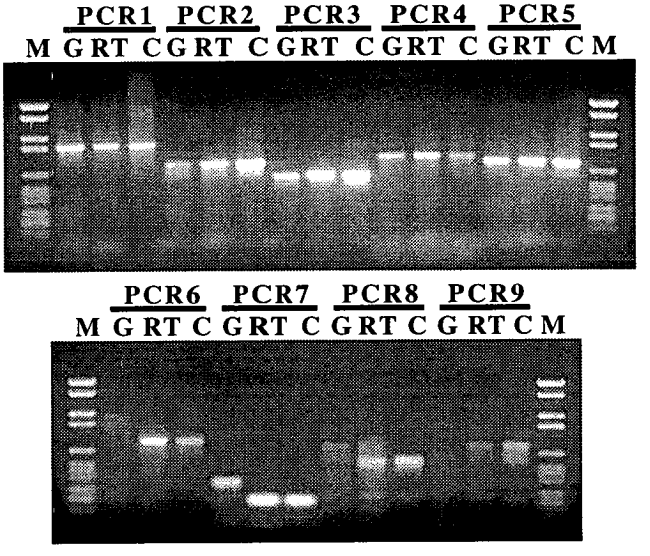
To determine mRNA size and test for the existence of any precursor or processed RNA species, total RNA was isolated from asynchronous asexual stage cultures of several *P. falciparum* strains and separated on formaldehyde gels for Northern blotting. The internal PCR fragment was used as the hybridisation probe. The *PfSNF2L* mRNA can be detected in all three laboratory lines (Dd2, 3D7A and HB3B) as a single, strong mRNA band of 6-6.5 kb (Figure 4.12A). Faint hybridisation to smaller sized fragments varies between RNA preparations and probably represents degradation rather than processing. Given that our largest cDNA clone is 5.27 kb in size, the primary transcript must be significantly longer than the

Figure 4.11 A. PCR fragments (1 to 9) which were designed to locate introns shown by solid lines. **B.** Comparison of the PCR products amplified from genomic DNA, RT-PCR product and *PfSNF2L* cDNA clone of *P. falciparum* to map the intron-exon boundaries. **C.** The sequence and position of the two carboxy-terminal introns of *PfSNF2L.* Intron 1 is 132 bp and intron 2 is 162 bp. G, genomic DNA; RT, RT-PCR product; C, PfSNF2L cDNA clone, M, DNA marker. Hp, *Hpa* I; E, *Eco*R I; H, *Hin*d III; B, *Bam*H I; C, *Cla* I; X, *Xba* I.

A



B



C

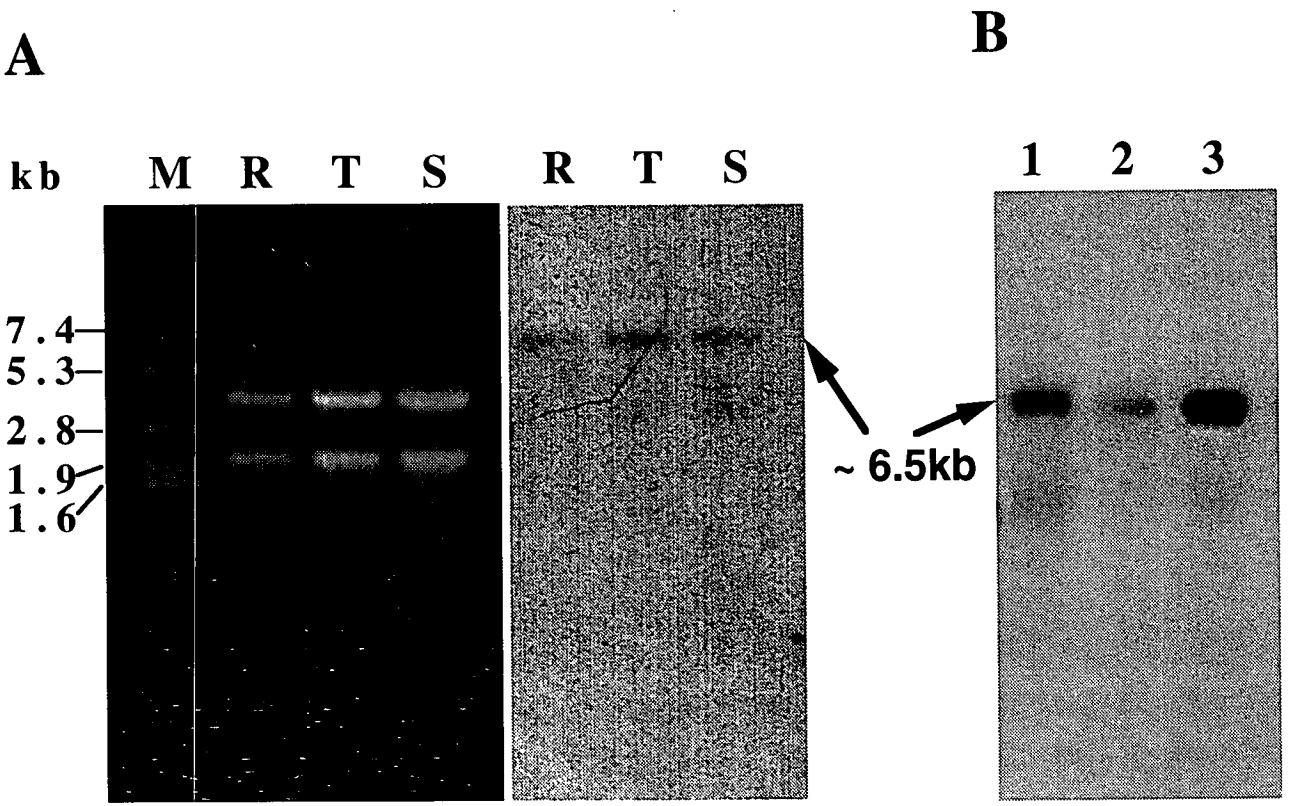
Intron1

```
          4626
           |
ATACACATTTTGgtatgtcacttgaatgtttttatattatatttatattgaacatatttttatgatattatatatatatatatatata
I  H  I  L                                                    4627
                                                               |
tttatatatgtatgtatgtaattgtttatatatttgtctatttattcatttttagACCCCCTATTATTGG
                                                               T  P  Y  Y  W
```

Intron2

```
 4746
  |
AAGgtataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagttatgtgattttgatgacctatatatatacttatatgtatatatttaaatat
K                                                                                    4747
                                                                                      |
taaattacatactaatttaacatatgtgatttattgatttattaatttatttatttttttttatccctttttattatagGAAAACGTT
                                                                                      E  N  V
```

Figure 4.12 Message size and steady state mRNA levels throughout the asexual cycle. **A.** Ethidium stained 1.2% formaldehyde/agarose gel separation of total RNA from synchronised cultures of 3D7A parasites Northern blotted onto nylon membranes and probed with a radio labelled *PfSNF2L* fragment. R, T and S represent equal amounts of total RNA from rings, trophozoites and schizonts respectively. **B.** Total *P. falciparum* RNA from three parasite clones (1. Dd2, 2. 3D7A, 3. HB3B) and the Northern blot of the same gel transferred to nylon membranes and hybridised to a radio labelled *PfSNF2L* probe a (shown in Figure 4.2).


A
B
kb
M
R
T
S
R
T
S
1
2
3
7.4
5.3
2.8
1.9
1.6
~ 6.5kb

coding sequence. 3' RACE (rapid amplification of cDNA ends) analysis using a primer at the 3' end of the *PfSNF2L* gene and an oligo dT_{15} primer indicates that there is a longer 3' UTR in the primary transcript (data not shown), and it is likely that this is also true at the 5' end of the mRNA. Low amounts of hybridising message are present in gametocytes (data not shown), indicating that the *PfSNF2L* mRNA is present in all erythrocytic stages and in low amounts in the gametocytes (Figure 4.12B).

4.8 Discussion

A *P. falciparum* homologue of the *SNF2L* genes which encode chromatin remodelling/transcription activation proteins has been identified in this study. The protein encoded by this gene may be a component of a *Plasmodium* chromatin remodelling complex. This gene joins the TATA-binding protein gene (McAndrew *et al*, 1993), the chromatin-associated protein (Birago *et al*, 1996), the histone proteins (Bennett *et al*, 1995; Cary *et al*, 1994) and two high-mobility group protein genes (Guntaka *et al*, 1992; Kun & Anders, 1995) as characterised loci encoding proteins involved in regulating *P. falciparum* transcription and chromatin remodelling. *PfSNF2L* mRNA is present in low amounts throughout the asexual cycle although it is not known whether the protein itself is active in all stages. The *P. falciparum* SNF2L protein is related to the yeast Swi/snf2 protein, the first of this class of proteins to be described, but it is more similar to a protein of a *Drosophila* nucleosome remodelling complex, ISWI (Tsukiyama *et al*, 1995). Interestingly, three different remodelling complexes in *Drosophila* that each contain an ISWI subunit have been found to date. The *Drosophila* nucleosome remodelling factor (NURF) complex appears to be involved in chromatin activation yet it is smaller (~500 kDa as opposed to 2MDa), less complex (4 as opposed to 11 major polypeptides) and much more abundant (10^5 as opposed to 10^3 copies per cell) than the SWI/SNF complex (Tsukiyama *et al*, 1995). The chromatin accessibility complex (CHRAC) was identified on the basis of its ability to mobilise nucleosomes in a manner that allows enhanced access of a restriction enzyme to DNA packaged into chromatin (Varga-Weisz *et al*, 1997). ATP-dependent chromatin assembly and remodelling factor (ACF) was purified as an ATP-utilising factor which helps chromatin resembly. Recently, a new, 15-subunit chromatin remodelling complex, (RSC - *r*emodel the *s*tructure of *c*hromatin) has been reported in yeast. This complex is 10 times as abundant as the SWI/SNF complex and also contains a novel DNA dependent ATPase subunit (Cairns *et al*, 1996). The existence of multiple chromatin remodelling complexes may indicate that they regulate different promoters and/or control different aspects of DNA metabolism such as repair or recombination. It is not known if the PfSNF2L forms part of a multi-protein complex or if it is essential for transcription. Unlike yeast and other higher eukaryotic cells

which have been found to contain more than one homologue of the SNF2/SWI2 gene, the *PfSNF2L* gene is the only one found in *P. falciparum* to date. It is unclear whether a more distantly related SNF2/SWI2 homologue exists which may form different complexes to regulate distinct cell function manners in malaria cell. The complete sequence of *P. falciparum* is being carried out at present and this may help resolve many of these question relating to the gene repertoire of *P. falciparum*.

Two substantial domains are unique to the *P. falciparum* SNF2L protein. These are the amino terminal tandem repeats of the $(NNHNDD)_8$, and the central 210 amino acid lysine rich basic domain. The role of neither of these regions is understood, although insertions of tandem repeats and basic domains and a consequent marked tendency towards larger-than-normal size are common features of *Plasmodium* proteins. The NNHNDD tandem repeat is similar to the common *Plasmodium* epitope (CPE) repeats, such as (X)KND and NKN(X'), which are found in many *Plasmodium* antigens (Cheng *et al*, 1991). These amino acid repeats are usually immunodominant compared to the other parts of the genes. It is interesting that asparagine rich repeats, basic domain insertions and the tendency towards increased size and complexity are also features of subunits of the *P.falciparum* RNA polymerases II (Li *et al*, 1989) and III (Li *et al*, 1991). The authors of these studies on RNA polymerase proposed that the enlarged variable regions contained regulatory domains for transcriptional control during the parasite's complex developmental cycle. This hypothesis remains tenable and also applicable to the SNF2L protein and its unique domains. Lysine rich sequences within nucleoproteins are common and are associated with nuclear localisation signals (Gerace, 1995). Residues 895-920 of SNF2L have a particularly high similarity to the nuclear targeting consensus sequences predicted using the PSORT protein localisation site programme (underlined in dashes in Figure 4.3). Several potential casein kinase II phosphorylation sites are also found near this region (STDE, 870-873, TDEE,871-874, SGEE, 879-882, TILE 883-886, TFLD, 928-931). Such phosphorylation sites are frequently found next to nuclear localisation sequences and can enhance the rate of nuclear transport (Rihs *et al*, 1990). In the motif II of the DNA-dependent ATPase domain, all of the members of the SNF2L subfamily have the DEAH motif, including PfSNF2L, and are thus different from the other subfamilies of the SNF2 family which use the DEGH motif. The phylogenetic study of the PfSNF2L protein shows that it is more related to the members in SNF2L subfamily than to any other characterised proteins, including other members of the SNF2 family. The reason may be that the evolutionary subdivision is paralleled by functional subdivision and therefore that particular functions are conserved within, but not between, subfamilies (Eisen *et al*, 1995).

The molecular switches which regulate the malaria parasite's responses to extra-cellular signals and define entry into developmental pathways such as endothelial sequestration or gametocytogenesis are not understood. This problem can in principle be studied by describing the components of transcriptional activation in *P. falciparum* and comparing them to better understood systems and then testing hypothetical models for *Plasmodium* developmental mechanisms. The problem with this approach is the complexity of gene expression. Thus, for RNA polymerase II (pol II) to transcribe a gene requires the recruitment of >20 proteins to a promoter (Buratowski, 1994; Tjian & Maniatis, 1994). In addition, *assembly* of the initiation complex requires the prior removal or repositioning of nucleosomes, carried out by multi-protein chromatin remodelling complexes, some of which may be promoter specific (another 20-30 proteins?) (Felsenfeld, 1996; Tsukiyama *et al*, 1994). Current models of chromatin remodelling complex function are largely based on studies of the yeast, SWI/SNF complex. Each protein in this complex is essential for transcriptional control at two sets of inducible loci, those involved in mating type switching (the *swi* mutants) and those controlling sugar fermentation (the *snf- s*ucrose *n*on-*f*ermenting mutants). However, none of the genes encoding the SWI/SNF complex proteins are essential for asexual growth and normal chromatin remodelling and transcription occurs at other promotors in *swi* and *snf* mutants.

The lack of mutant selection systems and the fragility of the culture system has inhibited progress on complex aspects of the molecular biology of the malaria parasite. The genome mapping project (Dame *et al*, 1996) and the sequencing of the *P.falciparum* genome could permit much faster identification of conserved core components of the parasite's transcription apparatus. With genetic tests for functionality such as gene 'knock outs' (Menard *et al*, 1997) it may be possible to ask whether genes such as *P.falciparum SNF2L* are essential for survival or involved in transcription of 'inessential' inducible loci such as the metabolic and mating type switches activated by the SWI/SNF chromatin remodelling complex. We have provided Dr. Kirk Deitsch of the National Institutes of Health, Bethesda, with the PfSNF2L clones to attempt the 'knock out' transfection of PfSNF2L. Stable transfectants with the gene have been obtained and the analysis of selected clones for homologous recombinants has started.

CHAPTER 5
THE EXPRESSION OF THE PFSNF2L PROTEIN DURING *PLASMODIUM FALCIPARUM* DEVELOPMENT

5.1 Introduction

Having established that PfSNF2L is a homologue of known eukaryotic transcriptional activators, the question arises of what is its role during *P. falciparum* development? The first step towards studying this question is to review what is known about the behavior of SNF2/SWI2 and SNF2L proteins in better characterised eukaryotic models. In eukaryotes, cell proliferation and differentiation is strictly regulated through the process of the cell cycle. During mitosis, the condensation of the chromatin into metaphase chromosomes is correlated with a 'global' inhibition of transcriptional activity (Johnson & Holland, 1965; Littau *et al*, 1964). Most transcriptional activators are excluded from their normal target sites although they may still retain their ability to bind DNA (Martinez-Balbas *et al*, 1995).

The human brahma (hbrm) and brm/SWI-related gene 1 (BRG-1), two human homologues of SNF2/SWI2, are phosphorylated and excluded from the condensed chromosomes during mitosis. In addition, hbrm is partially degraded. However, the two proteins are located in the nucleus and apparently expressed at approximately similar levels throughout interphase. Mitotic phosphorylation of hbrm and BRG-1 does not disrupt the association of these proteins with the hSNF5 protein (human sucrose non-fermenting gene 5), but does correlate with a decreased affinity of these proteins for the nuclear structure before pro-metaphase (Muchardt *et al*, 1996). Phosphorylation triggers the inactivation of the SNF/SWI complexes during cell division. Chromosomal exclusion or physical separation of the human SNF/SWI complex from the chromosomes at the G_2/M transition is therefore thought to be part of the mechanism leading to transcriptional arrest during mitosis (Muchardt *et al*, 1996).

To efficiently arrest cells in the G_1 phase of the cell cycle, hbrm/BRG-1 is also required to cooperate with the retinoblastoma (RB) protein through a physical interaction between RB and hbrm/BRG-1 (Dunaief *et al*, 1994; Trouche *et al*, 1997). RB is thus thought to inhibit cell proliferation through its regulation of transcription. Repressing expression of RB protein induces cell entry into S phase and may lead to malignant transformation (Bartek *et al*, 1997). The hbrm/BRG-1 is thus a co-repressor for RB and plays a potentially important role in overall cell cycle control (Trouche *et al*, 1997).

In yeast, the gene *NPS1* (nuclear protein of *Saccharomyces*), another homologue of SNF2/SWI2, is essential for mitotic growth of the *S. cerevisiae*. The

growth of cells carrying the *NPS1* null allele and the galactose-inducable *NPS1* on the plasmid is arrested under *NPS1*-repressed conditions. Development halts at the large-bud stage with a single nucleus that has a DNA content of G_2/M phase cell. When the arrested cells are further incubated under nps1-repressed conditions, re-replication of DNA occurs in some of the arrested cells without passage through mitosis (Tsuchiya *et al*, 1992).

To investigate the links between the activity of the PfSNF2L protein and the process of the parasite's intra-erythrocytic cell cycle, immunocytochemical experiments were carried out to study the expression and subcellular localisation of PfSNF2L at different phases during *P. falciparum* development.

5.2 Expression of recombinant PfSNF2L fusion proteins

The pGEX-4T-1 plasmid (see appendix F) was first chosen to express a fragment of PfSNF2L protein as a C-terminal fusion protein with glutathione-S-transferase (GST), a 26 kDa protein cloned from *Schistosoma japonicum*. A 681 bp DNA fragment (GST1) was PCR amplified from cDNA clone D1.1 (nucleotides 4399-5079) by using a pair of oligonucleotide primers modified to include BamHI restriction sites (Figure 5.1). This fragment, which contained the stop codon of *PfSNF2L*, was then inserted into pGEX-4T-1 at the BamHI site and introduced into the *E. coli* XL-1 blue strain by transformation. Positive clones were identified by restriction enzyme digestion of the purified plasmids and then confirmed by DNA sequencing. Five positive clones were used to transform *E. coli* BL21 strain in order to produce the predicted 40.3 kDa fusion protein using pGEX-4T-1 alone as a negative control. The proteins were extracted from the BL21 cells after IPTG induction at 37°C and separated on SDS-PAGE gel. As shown in Figure 5.2, none of the recombinants expressed the deduced fusion protein (40.3 kDa), although they all contained a DNA insert. However the parental pGEX-4T-1 could produce the 29 kDa GST fusion protein encoded by the pGEX. However, increasing IPTG concentration and induction time or reducing the inducing temperature failed to induce any the PfSNF2L fusion protein expression (data not shown). The pET-His tag system was therefore tried as an alternative to the pGEX system for fusion protein expression.

Three DNA fragments were PCR amplified from the PfSNF2L cDNA clone (Figure 5.1). In order to avoid potential cross-reactions with human SNF2 proteins, two unique PfSNF2L-specific regions without homology to the known homologues (nucleotides 2787-3163, FP1, and 3378-3766, FP2) were chosen. Another PCR fragment (FP3) from nucleotides 4056-4456 was used to express the SANT domain. The pET-29 T-vector (see appendix E) is designed for expression of inserts as stable fusion proteins with a C-terminal hexa histidine tag transcribed from a T7 RNA

Figure 5.1 The sequences of the *PfSNF2L* fragments used to construct fusion protein clones for *PfSNF2L* antibody production. Two oligonucleotide primers, underlined by arrowhead lines and contain BamHI sites, were used to PCR amplify the GST1 fragment from the *PfSNF2L* cDNA clone and subcloned into pGEX-4T-1 GST vector for fusion protein expression. FP1, FP2 and FP3 were also PCR amplified from the *PfSNF2L* cDNA clone but subcloned into pET29 AT-His6-tag vector in order to produce the C-terminal His6-tag fused proteins. The sites of the primers used are underlined by arrowhead lines.

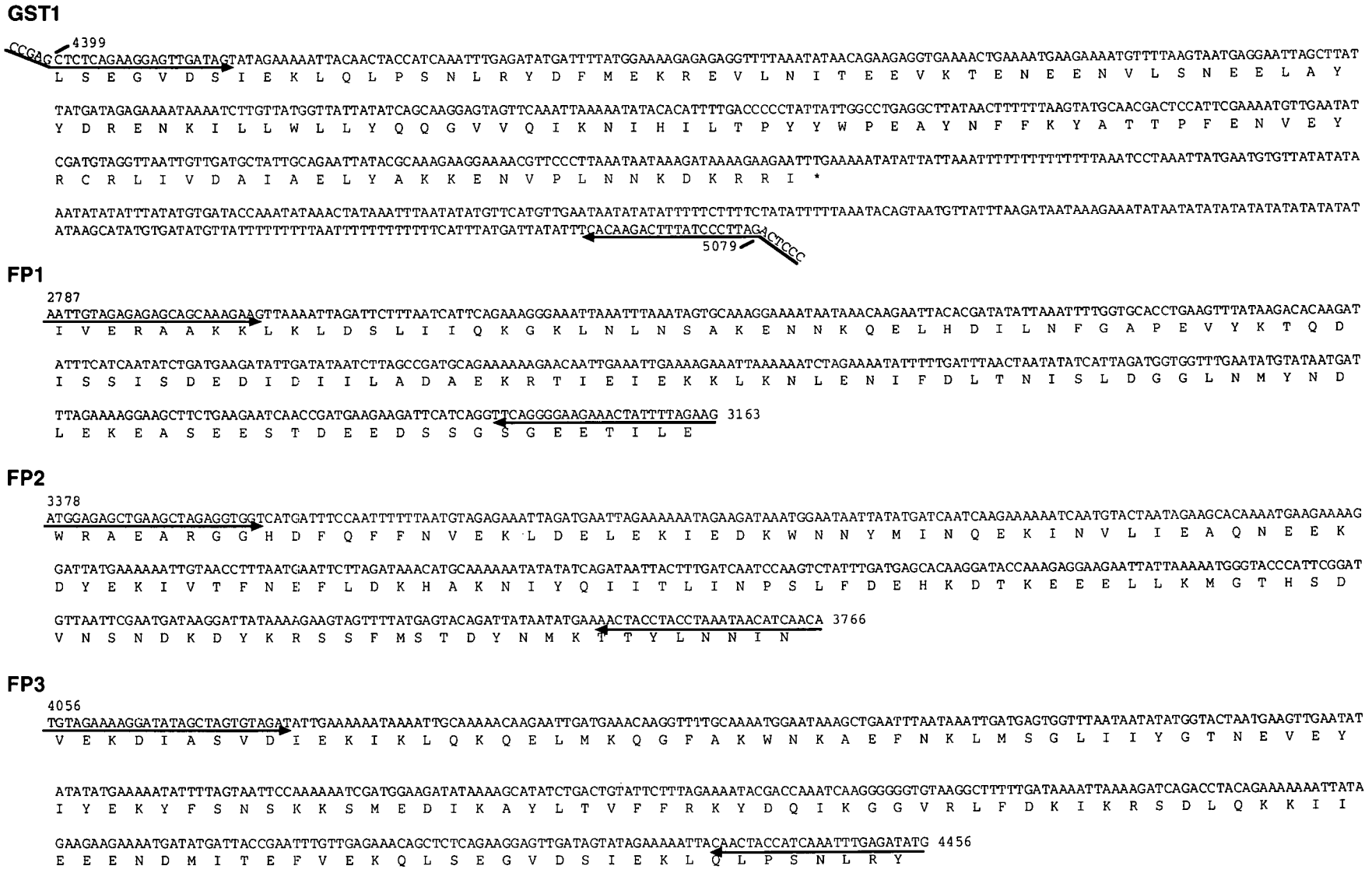
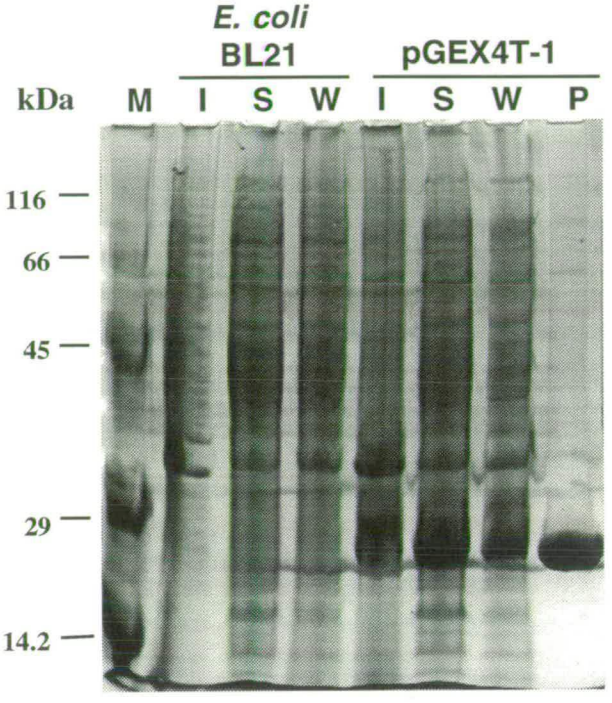
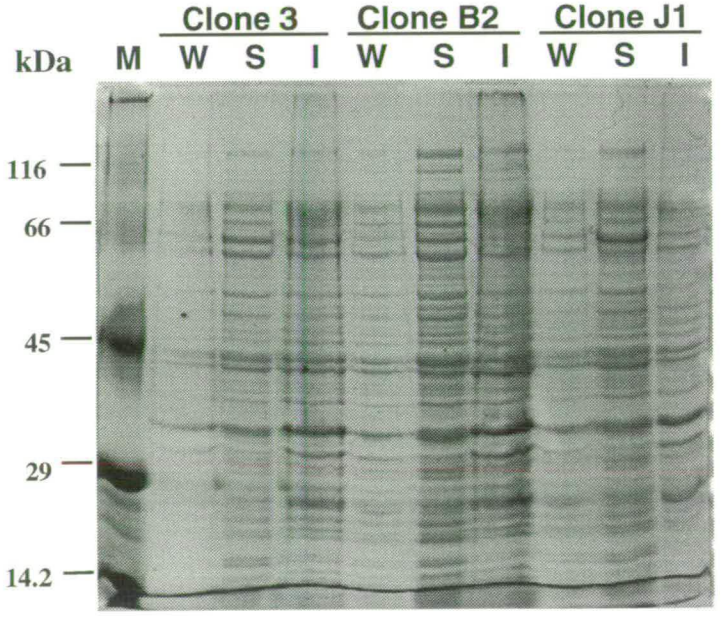

GST1
4399
CCGAGCTCTCAGAAGGAGTTGATAGTATAGAAAAATTACAACTACCATCAAATTTGAGATATGATTTTATGGAAAAGAGAGAGGTTTTAAATATAACAGAAGAGGTGAAAACTGAAAATGAAGAAAATGTTTTAAGTAATGAGGAATTAGCTTAT
L S E G V D S I E K L Q L P S N L R Y D F M E K R E V L N I T E E V K T E N E E N V L S N E E L A Y
TATGATAGAGAAAATAAAATCTTGTTATGGTTATTATATCAGCAAGGAGTAGTTCAAATTAAAAATATACACATTTTGACCCCCTATTATTGGCCTGAGGCTTATAACTTTTTTAAGTATGCAACGACTCCATTCGAAAATGTTGAATAT
Y D R E N K I L L W L L Y Q Q G V V Q I K N I H I L T P Y Y W P E A Y N F F K Y A T T P F E N V E Y
CGATGTAGGTTAATTGTTGATGCTATTGCAGAATTATACGCAAAGAAGGAAAACGTTCCCTTAAATAATAAAGATAAAAGAAGAATTTGAAAAATATATTATTAAATTTTTTTTTTTTTTAAATCCTAAATTATGAATGTGTTATATATA
R C R L I V D A I A E L Y A K K E N V P L N N K D K R R I *
AATATATATTTATATGTGATACCAAATATAAACTATAAATTTAATATATGTTCATGTTGAATAATATATATTTTTCTTTTCTATATTTTTAAATACAGTAATGTTATTTAAGATAATAAAGAAATATAATATATATATATATATATATAT
ATAAGCATATGTGATATGTTATTTTTTTTTAATTTTTTTTTTTTCATTTATGATTATATTTCACAAGACTTTATCCCTTAGACTCCC
5079
FP1
2787
AATTGTAGAGAGAGCAGCAAAGAAGTTAAAATTAGATTCTTTAATCATTCAGAAAGGGAAATTAAATTTAAATAGTGCAAAGGAAAATAATAAACAAGAATTACACGATATATTAAATTTTGGTGCACCTGAAGTTTATAAGACACAAGAT
I V E R A A K K L K L D S L I I Q K G K L N L N S A K E N N K Q E L H D I L N F G A P E V Y K T Q D
ATTTCATCAATATCTGATGAAGATATTGATATAATCTTAGCCGATGCAGAAAAAAGAACAATTGAAATTGAAAAGAAATTAAAAAATCTAGAAAATATTTTTGATTTAACTAATATATCATTAGATGGTGGTTTGAATATGTATAATGAT
I S S I S D E D I D I I L A D A E K R T I E I E K K L K N L E N I F D L T N I S L D G G L N M Y N D
TTAGAAAAGGAAGCTTCTGAAGAATCAACCGATGAAGAAGATTCATCAGGTTCAGGGGAAGAAACTATTTTAGAAG 3163
L E K E A S E E S T D E E D S S G S G E E T I L E
FP2
3378
ATGGAGAGCTGAAGCTAGAGGTGGTCATGATTTCCAATTTTTTAATGTAGAGAAATTAGATGAATTAGAAAAAATAGAAGATAAATGGAATAATTATATGATCAATCAAGAAAAAATCAATGTACTAATAGAAGCACAAAATGAAGAAAAG
W R A E A R G G H D F Q F F N V E K L D E L E K I E D K W N N Y M I N Q E K I N V L I E A Q N E E K
GATTATGAAAAAATTGTAACCTTTAATGAATTCTTAGATAAACATGCAAAAAATATATATCAGATAATTACTTTGATCAATCCAAGTCTATTTGATGAGCACAAGGATACCAAAGAGGAAGAATTATTAAAAATGGGTACCCATTCGGAT
D Y E K I V T F N E F L D K H A K N I Y Q I I T L I N P S L F D E H K D T K E E E L L K M G T H S D
GTTAATTCGAATGATAAGGATTATAAAAGAAGTAGTTTTATGAGTACAGATTATAATATGAAAACTACCTACCTAAATAACATCAACA 3766
V N S N D K D Y K R S S F M S T D Y N M K T T Y L N N I N
FP3
4056
TGTAGAAAAGGATATAGCTAGTGTAGATATTGAAAAAATAAAATTGCAAAAACAAGAATTGATGAAACAAGGTTTTGCAAAATGGAATAAAGCTGAATTTAATAAATTGATGAGTGGTTTAATAATATATGGTACTAATGAAGTTGAATAT
V E K D I A S V D I E K I K L Q K Q E L M K Q G F A K W N K A E F N K L M S G L I I Y G T N E V E Y
ATATATGAAAAATATTTTAGTAATTCCAAAAAATCGATGGAAGATATAAAAGCATATCTGACTGTATTCTTTAGAAAATACGACCAAATCAAGGGGGGTGTAAGGCTTTTTGATAAAATTAAAAGATCAGACCTACAGAAAAAAATTATA
I Y E K Y F S N S K K S M E D I K A Y L T V F F R K Y D Q I K G G V R L F D K I K R S D L Q K K I I
GAAGAAGAAAATGATATGATTACCGAATTTGTTGAGAAACAGCTCTCAGAAGGAGTTGATAGTATAGAAAAATTACAACTACCATCAAATTTGAGATATG 4456
E E E N D M I T E F V E K Q L S E G V D S I E K L Q L P S N L R Y

Figure 5.2 Comparison of the expression and solubility of the PfSNF2L-Glutathione-S-transferase (GST) fusion protein clones in pGEX-4T-1 vector. Proteins from an approximately equal number of cells induced with 1 mM IPTG were run on each lane of SDS-PAGE gels and stained with Coomassie blue. *E. coli* BL21 cell extracts alone were used as negtive control, whereas pGEX-4T-1 vector alone was given a GST protein positive control. Three PfSNF2L-GST fusion protein clones were compared. The proteins of soluble and insoluble fractions of total cell extracts and whole cell extracts was marked as S, I and W. M and P indicate protein marker and purified GST protein (The molecular weights in kDa are indicated).


E. coli
BL21
pGEX4T-1
kDa
M
I
S
W
I
S
W
P
116
66
45
29
14.2


Clone 3
Clone B2
Clone J1
kDa
M
W
S
I
W
S
I
W
S
I
116
66
45
29
14.2

polymerase promoter. The fragments were amplified using three pairs of specially designed primers and ligated to the pET-29 T-vector for in-frame expression in the pET-29 T-vector. The His6-tag facilitates the purification of PfSNF2L fusion proteins by specifically binding to immobilised nickel (Ni^{2+}) chelating resins. Positive clones from each of these three constructions were identified by restriction enzyme digestion of purified plasmid and then further confirmed by DNA sequencing.

Three positive clones of each recombinant were used to transform *E. coli* BL21 (DE3) strain in order to produce the predicted fusion protein. The proteins were extracted from the BL21 (DE3) cells after 1mM IPTG induction at 37°C and separated on SDS-PAGE gel. By this method, the growth of host cells was blocked and the bulk of the cell's protein synthesis derives from the fusion protein clones. As shown in Figure 5.3, two fusion protein FP1 clones were expressed as a 20.1 kDa soluble protein, whereas two FP2 clones were expressed as a 20.5 kDa insoluble protein. The fusion protein FP3 clones failed to produce the predicted 21 kDa protein, although a lower molecular weight protein is visible in the lane derived from an insoluble protein fraction. This clone was therefore abandoned. Solubility of the fusion protein was found to dramatically increase when the cell were induced at a lower temperature (25°C), as has previously been reported (Schein & Noteborn, 1988; Schein & Noteborn, 1988). However, reducing the induction temperature to 25°C could not increase the solubility of FP2 fusion protein (data not shown).

Before large scale purification of the fusion proteins was carried out, the plasmid stability test was carried out to examine whether each recombinant plasmid is stable in the BL21 (DE3) host. Both recombinant plasmids were transformed into the BL21 (DE3) host and then plated on the LB agar plates with or without ampicillin selection. As shown in Table 5.1, both recombinant plasmids are stable in host cells even without antibiotics selection. The FP1 fusion protein was purified under non-denaturing conditions and the FP2 fusion protein was purified under denaturing conditions in 8M urea by using the chromatography on Nickel charged resin in a batchwise fashion (as described in Method 2.16). Elution was done by stepwise gradient elution with different imidazole concentrations in PBS buffer. The FP1 fusion protein fractions of 100, 300 and 500 mM imidazole and the FP2 fractions of 40 and 60 mM imidazole were collected for further purification. No other bands appeared on the gel (see Figure 5.4). The final yield of pure FP1 and FP2 fusion proteins was 540 and 500 µg/ml, respectively.

5.3 Raising antibodies against PfSNF2L fusion proteins

Five MF1 mice were immunised with each PfSNF2L fusion protein. Each mouse was bled before the injection of the foreign protein for sera to serve as a control

Figure 5.3 Comparison of the expression and solubility of the PfSNF2L-His6-tag fusion protein clones in pET29 TA vector. Proteins from an approximately equal numbers of cells induced with 1 mM IPTG or uninduced were run on each lanes of SDS-PAGE gels and stained by Coomassie blue. *E. coli* BL21 (DE3) cell extracts alone were used as a negtive control. Two or three clones of each PfSNF2L-His6-tag fusion protein constructs FP1 and FP2 were compared. The proteins of soluble and insoluble fractions of induced total cell extracts and uninduced whole cell extracts are marked as S, I and U. M means protein marker (the molecular weights in kDa are indicated).

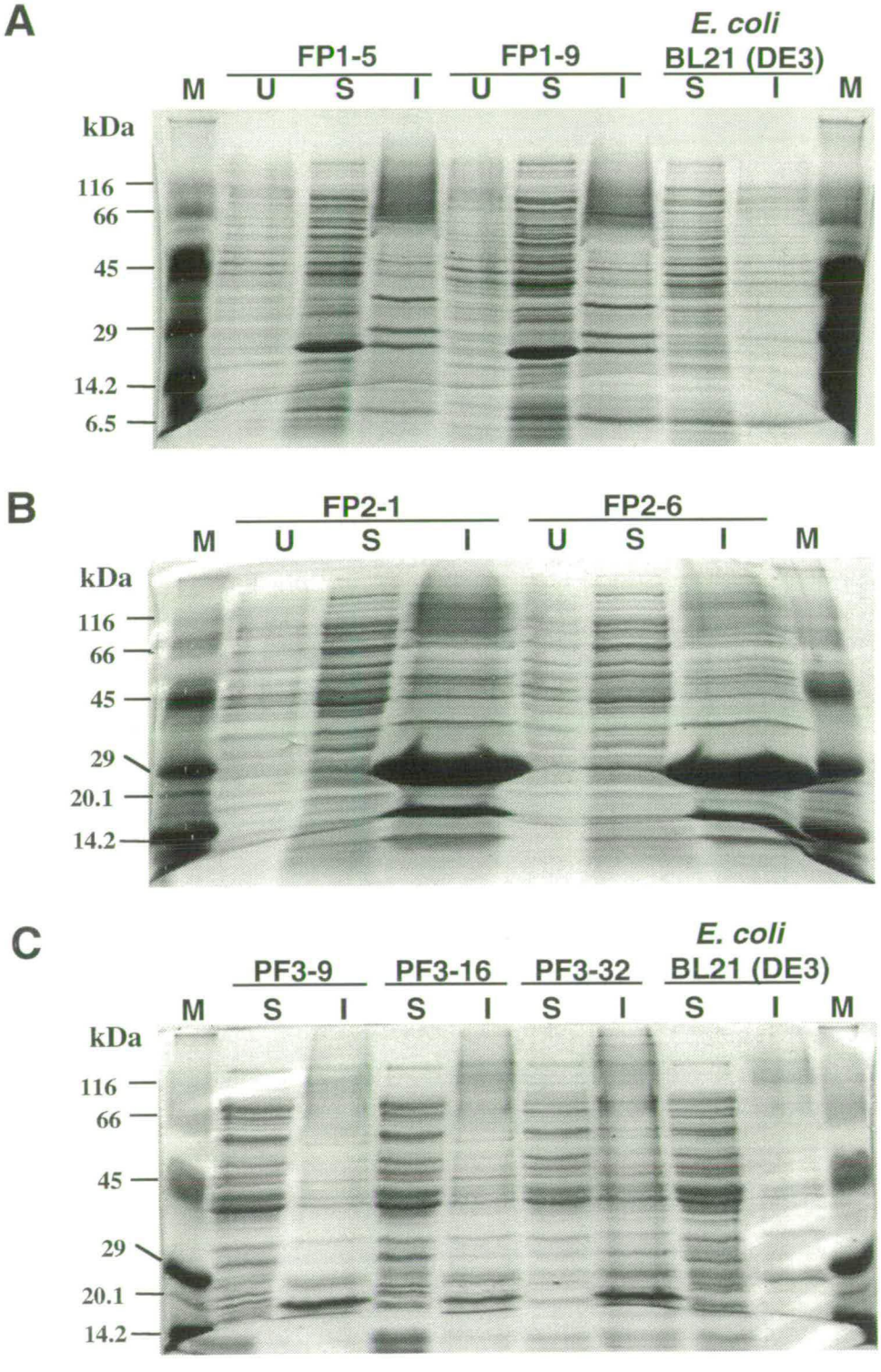

A
E. coli
FP1-5
FP1-9
BL21 (DE3)
M U S I U S I S I M
kDa
116
66
45
29
14.2
6.5
B
FP2-1
FP2-6
M U S I U S I M
kDa
116
66
45
29
20.1
14.2
C
E. coli
PF3-9
PF3-16
PF3-32
BL21 (DE3)
M S I S I S I S I M
kDa
116
66
45
29
20.1
14.2

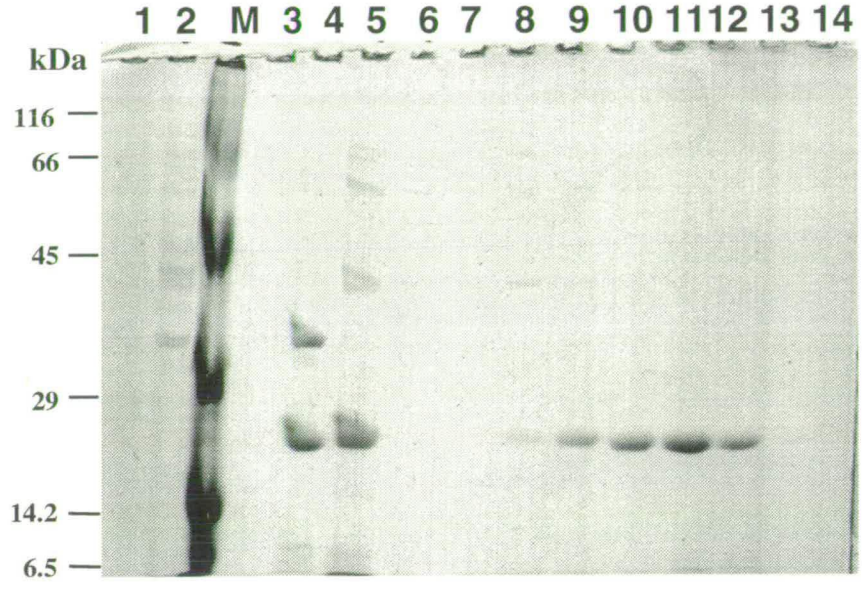
Table 5.1 Plasmid stability test of the recombinants of PfSNF2L fusion proteins

PfSNF2L fusion proteins	LB	LB + 1mM IPTG	LB + Kan	LB + Kan + 1mM IPTG
FP1-5	5×10^8[a]	0	4.7×10^8	0
FP1-9	5×10^8	0	8.2×10^8	0
FP2-1	2×10^8	0	7.2×10^8	0
FP2-6	3.5×10^8	0	6.3×10^8	0
FP3-9	2.5×10^8	0	6.3×10^8	0
FP3-16	7.5×10^8	0	7×10^8	0
FP3-32	4×10^8	0	6×10^8	0

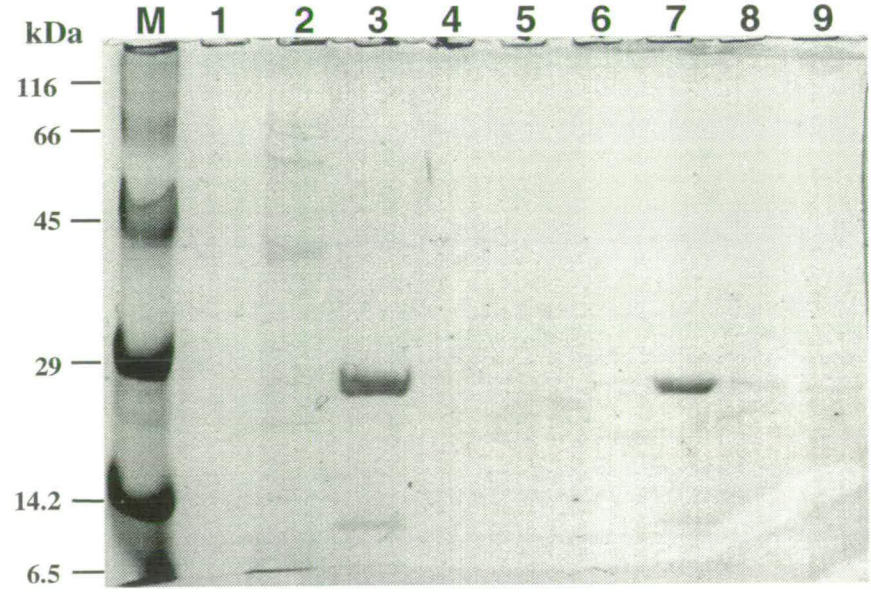
a. CFU/ml culture

Figure 5.4 Purification of the PfSNF2L-His6-tag fusion proteins from *E. coli* BL21 (DE3) cells with His-tag affinity chromatography. **A.** Protein extracts of PfSNF2L-His6-tag fusion protein clone FP1-9 were first sonicated and pelleted. The soluble fraction was purified on Nickel resin beads, washed and batchwise eluded in imidazole gradients under non-denaturing conditions. Lane: 1 and 2, IPTG uninduced and induced whole cell extracts of *E. coli* BL21 (DE3) host cell alone; 3, whole cell extract from uninduced FP1-9 clone; 4, insoluble fraction of total cell extracts from the FP1-9 clone induced by 6 hr 1 mM IPTG; 5, flow of soluble fraction of total cell extracts from the FP1-9 clone after passing through affinity chromatography; 6, elude of binding buffer wash; 7-13, eludes obtained with 20 mM, 40 mM, 60 mM, 100 mM, 300 mM, 500 mM and 1M imidazole, respectively; 14, elude of strip buffer. **B.** Protein extracts of PfSNF2L-His6-tag fusion protein clone FP2-6 were first sonicated and pelleted. The insoluble fractions was purified on Nickel resin beads, washed and batchwise eluded in imidazole gradient under a denaturing condition with 8 M urea. Lane: 1, whole cell extract from uninduced FP2-6 clone; 2, soluble fraction of total cell extracts from the FP2-6 clone induced by 6 hr 1 mM IPTG; 3, flow of resolublised insoluble fraction of the FP2-6 clone after passing through affinity chromatography; 4, elute of binding buffer wash; 5-9, eludes obtained with 10 mM, 20 mM, 40 mM, 60 mM, 100 mM imidazole, respectively; M, protein markers (the molecular weights in kDa are indicated).

A


1 2 M 3 4 5 6 7 8 9 10 11 12 13 14
kDa
116
66
45
29
14.2
6.5

B


M 1 2 3 4 5 6 7 8 9
kDa
116
66
45
29
14.2
6.5

to monitor the pre-immunisation response to malaria proteins. Approximately 100 µg of the FP1 or FP2 fusion protein in 200 µl of complete Freund's adjuvant was given to each mouse by intraperitoneal injection. The mice were boosted at four weeks, eight weeks and sixteen weeks after the initial immunisation with the fusion proteins in incomplete Freund's adjuvant. The sera were collected at three weeks, seven weeks, fifteen weeks and twenty weeks and tested by Western blotting to check for an antibody response (Figure 5.5).

5.4 Western blotting analysis of the PfSNF2L protein

After the third boost, each mouse raised a high titre anti-serum against the FP1 or the FP2 fusion proteins. The five mouse anti-sera against each FP1 or FP2 protein were therefore pooled together. Each pooled anti-serum was diluted 1:50, 1:100, 1:200 and 1:500. Three dilutions of the second antibody (horseradish peroxidase (HRP) conjugated anti-mouse IgG) were used at 1:1,000, 1:5,000 and 1:10,000. Results showed that 1:200 of each pooled anti-serum and 1:5,000 of the second antibody produced minimum background noise on the Western blots of purified FP1 and FP2 proteins (data not shown). These were therefore adopted as the optimum dilutions required to determine the specificity of the reactions.

To identify which proteins were recognised by the immune sera, Western blots prepared from asexual stage total parasite protein extracts were incubated with the pooled anti-sera. Initially, a Western blot of the total protein extracts prepared from the -20°C stored 3D7A parasites was probed with each pooled anti-sera. Strips of these blots were also probed with the pre-immune sera. Multiple bands, all of which were smaller than the expected 167 kDa PfSNF2L protein, were detected by both the anti-FP1 and the FP2 sera (Figure 5.6A, lane 4 and 5). Pre-immune sera did not recognise any *Plasmodium* proteins (lane 1), whereas the 50 kDa and 65 kDa proteins were detected by a control anti-α-tubulin monoclonal antibody and an anti-PfRap-1 sera (lane 2 and 3). Therefore either the PfSNF2L protein has been degraded during storage or the two PfSNF2L antisera may cross-react with some unknown parasite proteins.

A second Western blot was therefore prepared from freshly harvested parasites with a protease inhibitor cocktail added to the preparation to inhibit protein degradation. After using this extraction protocol, the predicted approximately 167 kDa protein did appear (lane 4 and 6, Figure 5.6B). The multiple lower molecular weight bands are however still visible and a few very faint cross-reacting bands are seen in the uninfected red blood cell extracts (lane 3 and 5). The same experiment was repeat again (Figure 5.7) and this again confirmed that a clear cross-reaction with a 167 kDa bands is present (lane 4 and 6). Very faint cross-reacting bands also appear in the lane of the *E. coli* BL21 (DE3) whole cell lysate when incubated with the FP1 anti-sera, indicating

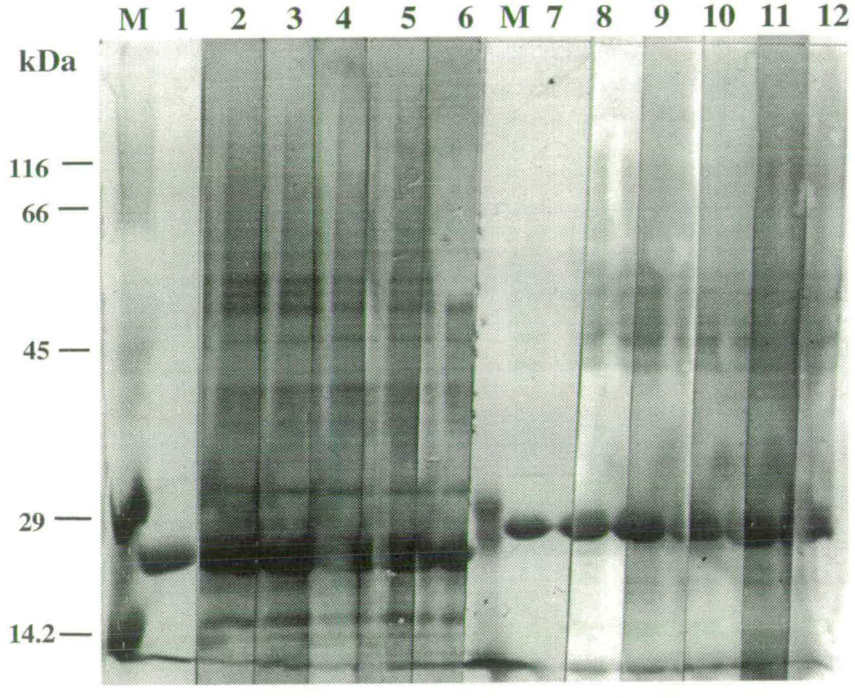
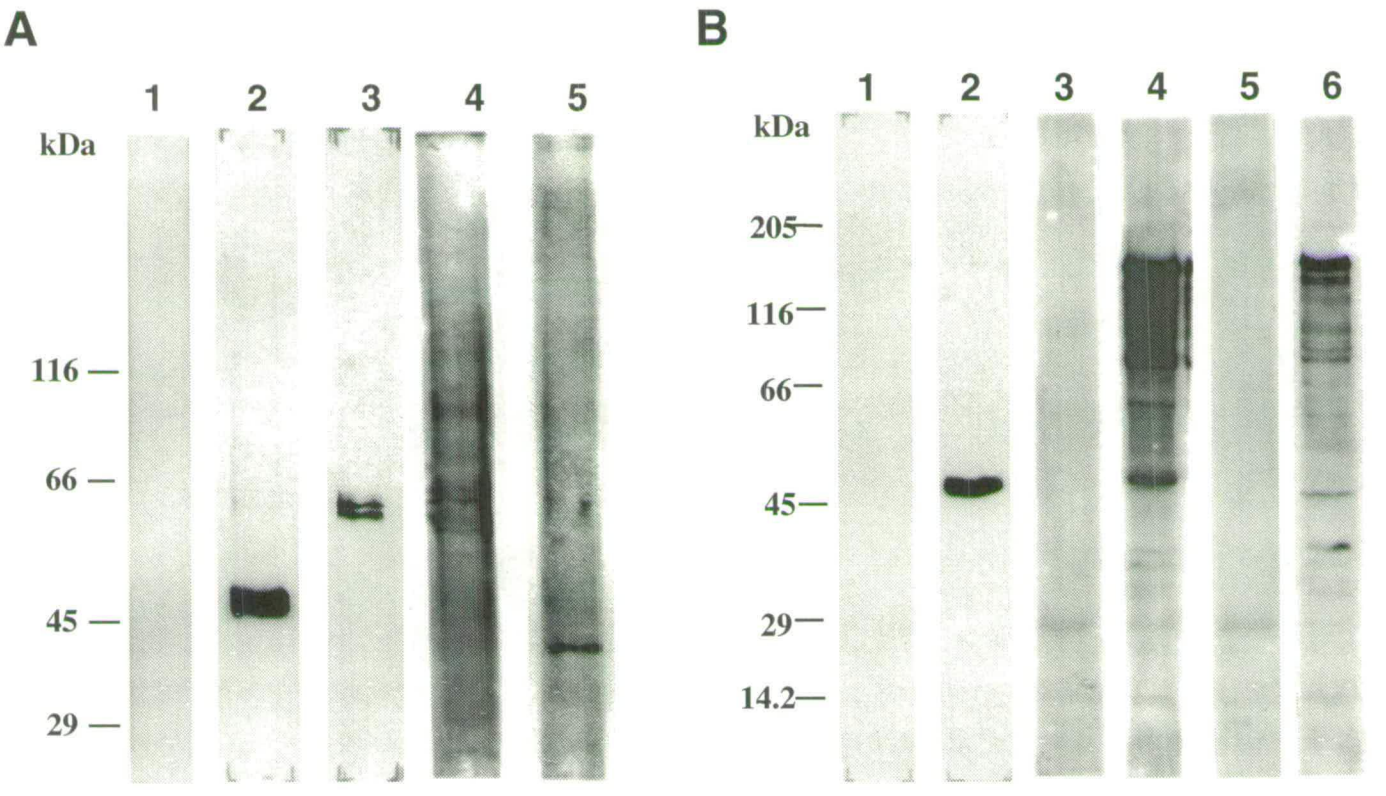


Figure 5.5 Analysis of PfSNF2L mouse sera on the Western blot of the purified PfSNF2L fusion proteins. A group of five mice were immunised by each purified PfSNF2L fusion protein FP1-9 and FP2-6. Sera were bled from individual mouse after three immunisations and tested on the Western blot of the purified PfSNF2L fusion proteins. Lane: 1, purified PfSNF2L fusion protein FP1-9 stained by amido black; 2-6, PfSNF2L fusion protein FP1-9 probed by mouse serum A to E; 7, PfSNF2L fusion protein FP2-6 stained by amido black; 8-12, PfSNF2L fusion protein FP2-6 probed by mouse serum F to J; M, protein markers (the molecular weights in kDa are indicated).

Figure 5.6 Immunoblot analysis of anti-PfSNF2L mouse sera. **A.** Immunoblots of the total cell extracts were prepared from -20°C frozen stocks of asexual stage *P. falciparum* parasites and probed with either anti-PfSNF2L mouse sera FP1, 1: 300 dilution (lane 4) or mouse antisera FP2, 1:150 dilution (lane 5). Membrane strips were then probed by peroxidase-conjugated rabbit anti-mouse immunoglobulin and detected by chemiluminescence. Pre-immunisation sera recognised no parasite protein (lane 1). The parasite α-tubulin and rhoptry proteins were recognised by anti-α-tubulin monoclonal antibody (mAb) TAT1 and anti-rhoptry mAb rap1, each used at 1: 500 dilution (lane 2 and 3). **B.** Immunoblots of the total cell extracts were prepared from freshly harvested asexual stage *P. falciparum* parasites and probed with either mouse antisera FP1 (lane 4) or antisera FP2 (lane 6). Blots of RBC extracts alone were probed with either mouse antisera FP1 (lane 3) or antisera FP2 (lane 5). Pre-immunisation sera recognised no parasite proteins (lane 1). The parasite α-tubulin were indicated by mAb TAT1 (lane 2). The conditions applied here were identical to those described above.


A
1
2
3
4
5
kDa
116
66
45
29
B
1
2
3
4
5
6
kDa
205
116
66
45
29
14.2

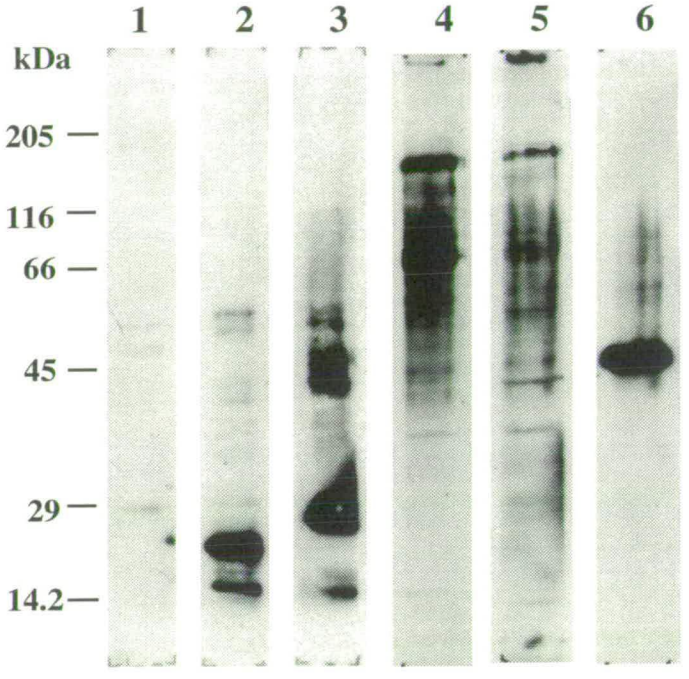


Figure 5.7 Immunoblot analysis of anti-PfSNF2L mouse sera using cell extracts prepared from freshly harvested asexual stage *P. falciparum* parasites, probed with either anti-PfSNF2L mouse sera FP1 (lane 4) or mouse antisera FP2 (lane 5). *E. coli* BL21 (DE3) cell extract was probed with antisera FP1 (lane 1), showing a faint cross-reaction with bacterial proteins. Blots of purified fusion protein FP1 and FP2 were also probed with antisera FP1 and FP2 (lane 2 and 3). The size of the parasite α-tubulin control is indicated by mAb TAT1 (lane 6). The conditions applied here were same as those described in Figure 5.6.

that the original immunisations were not significantly affected by bacterial contaminants (Figure 5.7, lane 1). In the above experiments, an anti-trypanosome α–tubulin mAb which cross-reacts with the *Plasmodium* protein was used as a positive control. Pre-immune mouse sera were the negative control. The anti-α–tubulin mAb recognised a band of 50 kDa in size, the *P. falciparum* α-tubulin. These results indicated that the predicted 167 kDa PfSNF2L protein is present and recognised by this polyclonal antiserum in the *P. falciparum* parasite. It seems that it is easily degraded by during the preparation and there may be weak cross-reactions with smaller proteins but these do not compromise the use of the antisera in immuno-localisation experiments.

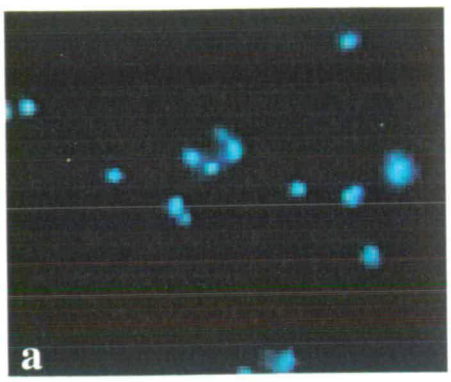
5.5 Immunolocalisation of PfSNF2L protein

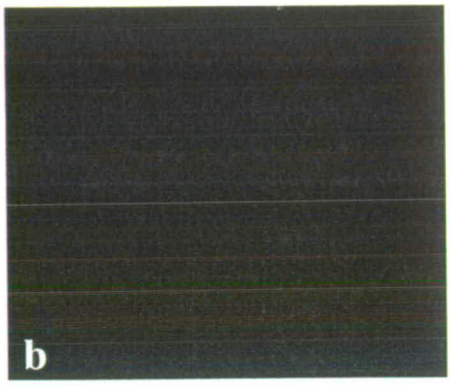
To localise the PfSNF2L protein in the parasitised red blood cell and to monitor its expression throughout the asexual life cycle of the parasite, immunofluorescence assays (IFA) were performed on air-dried and acetone fixed thin smears of parasites taken from synchronous and asynchronous cultures of *P. falciparum* strain 3D7A infected RBCs.

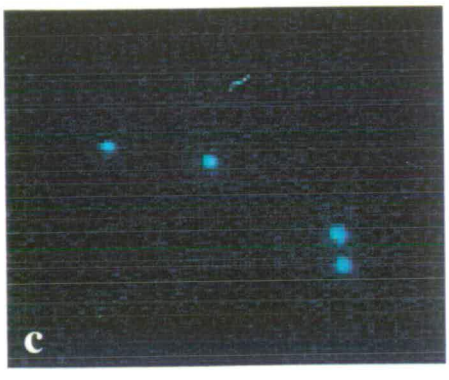
The pooled mouse polyclonal antisera made against the PfSNF2L constructs FP1 was tested at a dilution of 1:50, 1:100 and 1:200. Antibodies bound to the parasite preparation were detected using fluorescein conjugated anti-mouse immunoglobulin G (IgG) antibodies at dilutions of 1:50, 1:100 and 1:200. Bound antibodies are thus detected by a green fluorescence under U/V illumination while the parasite nuclei were simutaneously counter-stained with DAPI. Best staining conditions were achieved at 1:150 dilutions of pooled antisera and 1:100 of the second antibody.

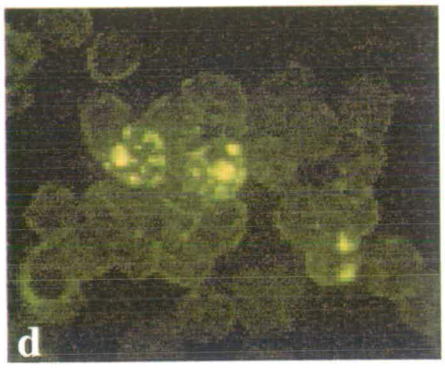
The FP1 mouse antisera were first used to examine asynchronous parasite blood smears and a series of IFA experiments were carried out. These consistently gave the same fluorescence patterns. Fluorescence was observed in all stages of the intra-erythrocytic parasites from early rings to schizonts (Figure 5.8). A granular pattern was seen in the parasitised red blood cell (RBC) cytosol although the exact location is unclear. Positive control experiments, using an anti-PfRan rabbit antisera, were performed (Figure 5.10 q and r). No fluorescence was detected when mouse pre-bleed serum (Figure 5.8 a and b) or the fluorescin-conjugated second antibody alone were used on the staining of infected or non-infected red blood cells (data not shown). When the anti-PfSNF2L antisera were used, although a diffuse pattern was seen over the whole parasitised cell throughout all stages of asexual development, an intense fluoresein spot was shown to co-localise to the DAPI stained parasite DNA (Figure 5.8) indicating that most of the PfSNF2L protein is concentrated in the parasite nucleus. Figure 5.9 shows the same pattern of spots of fluorescence co-localised with the nuclear DAPI stain when the anti-PfSNF2L antisera were used to stain the isolated

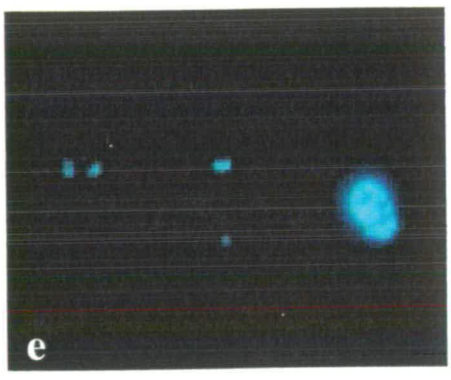
Figure 5.8 Immunofluorescence localisation of PfSNF2L within stages of intra-erythrocytic *P. falciparum*. Panels a, c and e: DAPI stained parasite nuclei taken from asynchronously growing asexual stages of *P. falciparum*. Panels d and f: Immuno-fluoresence micrographs of the corresponding parasites with FITC-labelled polyclonal mouse antisera (1:150 dilution) raised against the recombinant fragment FP1 of the PfPfSNF2L protein. Panel b: Immuno-fluoresence micrographs of the corresponding parasites with FITC-labelled pre-immune mouse serum (1:50 dilution).


a


b


c


d


e

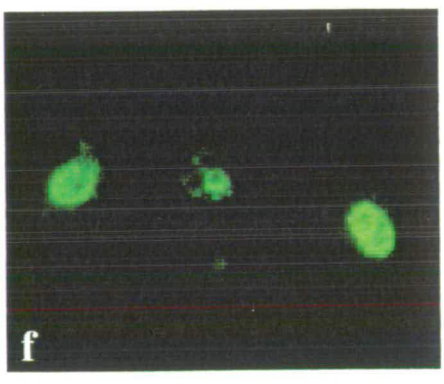

f

Figure 5.9 Staining of isolated *P. falciparum* nuclei with antibodies against a characterised PfSNF2L. Parasite nuclei were isolated by saponin lysis of infected erythrocytes followed by cell rupture by homogenisation and separation of parasite nuclei from the external membranes by centrifugation (Lanzer, 1992). Panels a, c, and e, DAPI stained isolated *P. falciparum* strain 3D7 nuclei from an asynchronous culture. Panel b, the nuclei stained in panel a, co-stained with FITC labelled pre-immunisation rabbit serum. Panels d and f, the field corresponding to panels c and e, stained with a FITC labelled mouse antiserum against a recombinant fragment FP1 of the PfSNF2L protein.

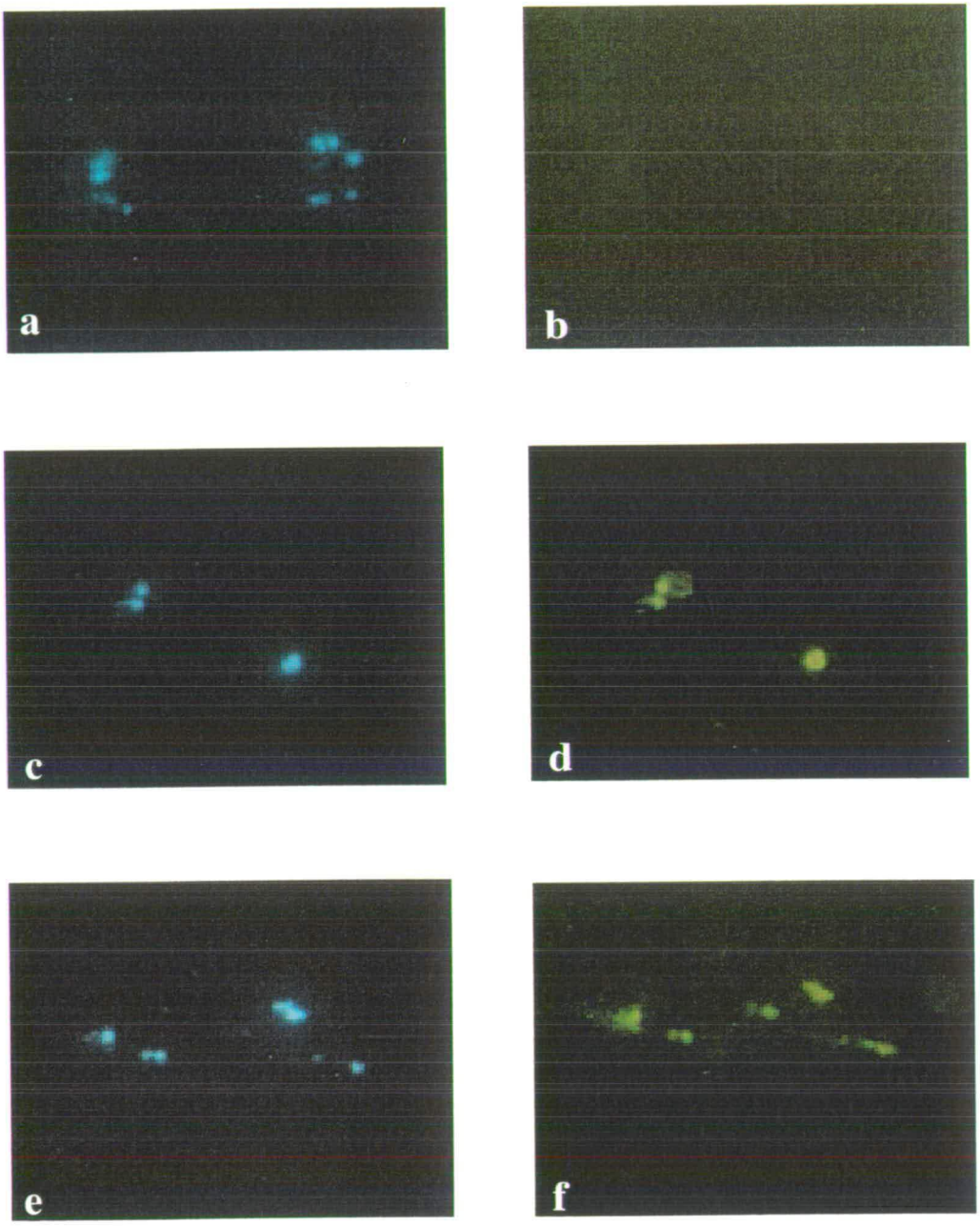
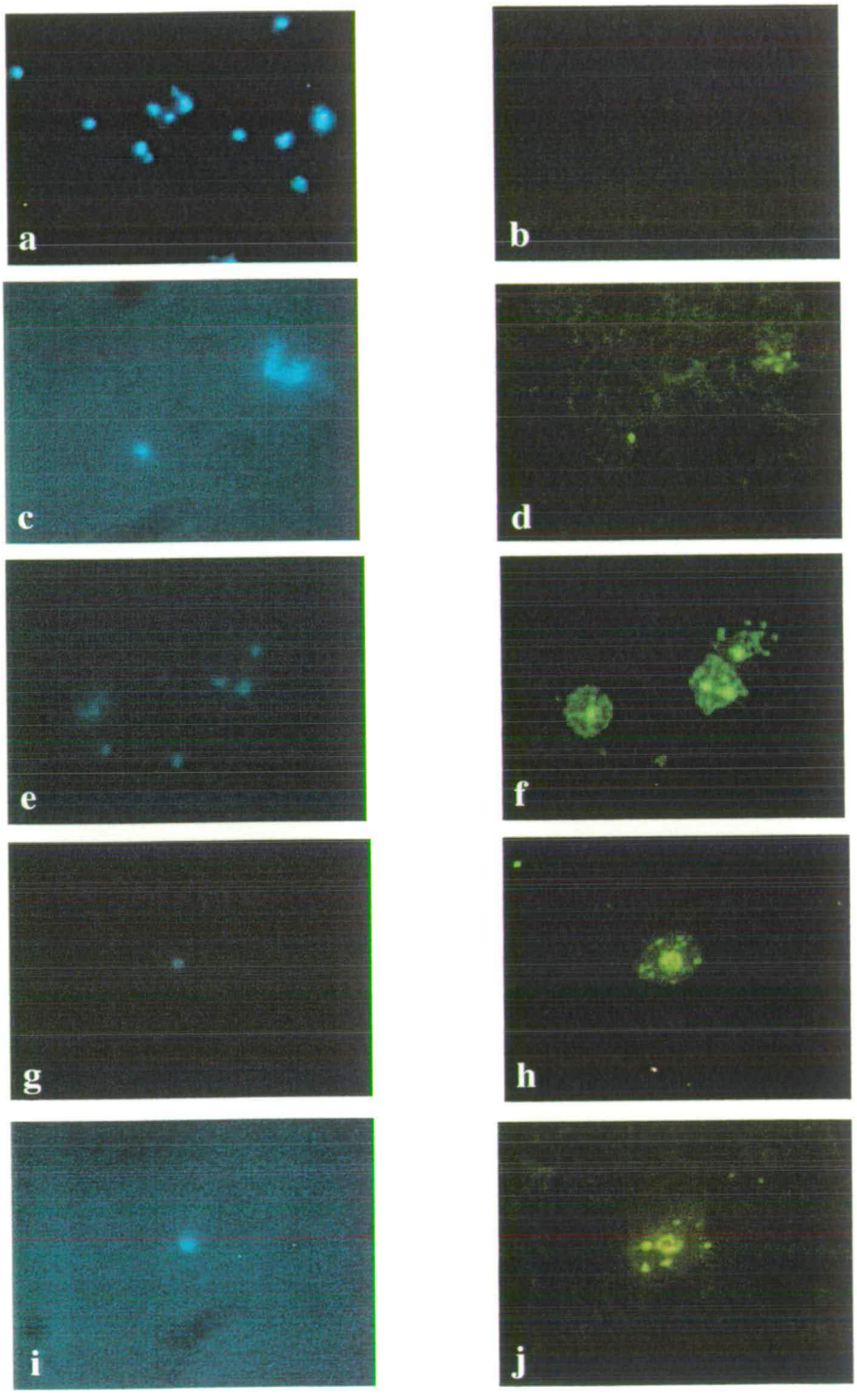
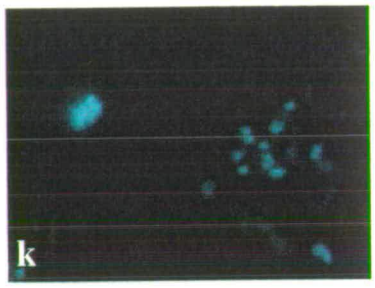
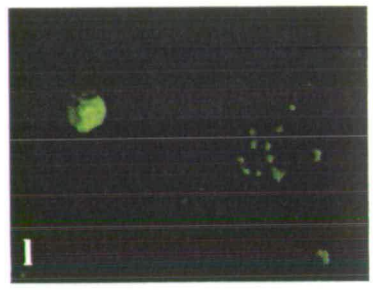
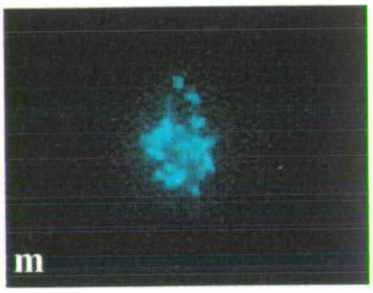

a
b
c
d
e
f

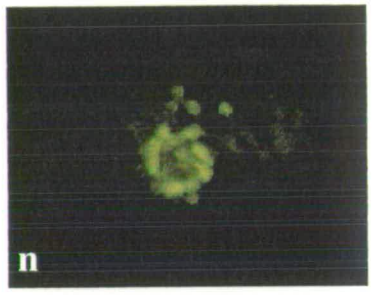
Figure 5.10 Intra-cellular localisation of PfSNF2L by immuno-fluorescence microscopy within different stages of intra-erythrocytic *P. falciparum*. Parasite nuclei were stained with DAPI at the early ring (c), ring (e), trophozoite (g and i), schizont/merozoite (k), segmenter (m) and gametocyte (o) stages of development taken from separate synchronous cultures. Immuno-fluoresence micrographs of the corresponding parasites were detected with FITC labelled polyclonal mouse antisera raised against a recombinant fragment FP1 of the PfSNF2L protein (d, f, h, j, l, n and p). *P. faciparum* parasites at segmenter stage were probed with DAPI and FITC-labelled polyclonal anti-PfRan rabbit sera as a positive control (q and r). Asexual stages parasites were also co-stained with DAPI and FITC-labelled mouse pre-immune sera as a negative control (a and b).

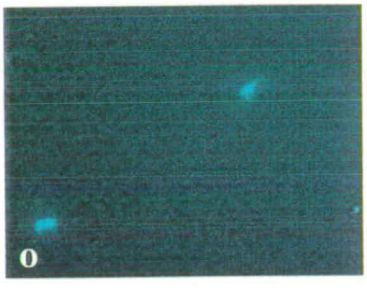

a
b
c
d
e
f
g
h
i
j

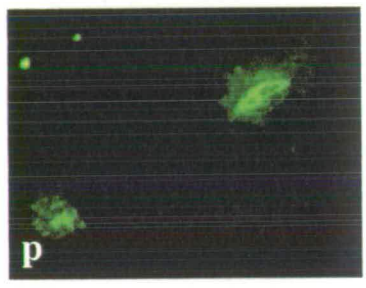

k

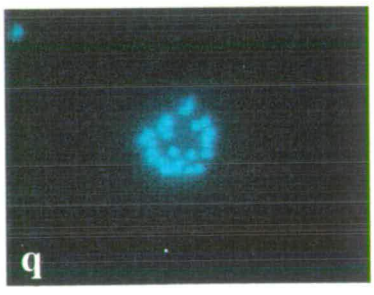

l

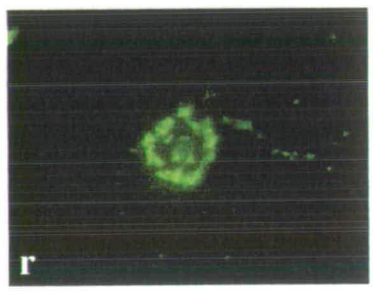

m


n


o


p


q


r

parasite nuclei. This indicates that the PfSNF2L protein, as would be expected for a transcription factor, is located in the parasite nucleus.

The changing nature of the PfSNF2L expression was further characterised throughout the asexual cell cycle of synchronous parasite cultures. PfSNF2L could be detected in young rings as a spot of fluorescence which co-localised to the nucleus with the DAPI stain (Figure 5.10 c and d). This intra-nuclear concentration then changed to more dispersed, a punctate pattern of fluorescence in late rings, although it was still predominantly concentrated around a 'bright spot' of apparently nuclear localisation (Figure 5.10 e and f). These extra-nuclear granular spots appear to be located within the erythrocyte cytoplasm of the parasitised RBCs rather than just inside the parasitophorous vacuole and some cells showed the spots apparently aligned directly under the erythrocyte membrane. The punctate spots become bigger in late trophozoites but their numbers decrease (Figure 5.10 g to j). Strikingly, the PfSNF2L seemed to be excluded from nucleus at this part (Figure 5.10 g to j). As shown in Figure 5.10 k and l, the punctate pattern also disappears during the schizont stages, and the bright fluorescein staining appears more concentrated in segmenting schizonts and merozoites. In the late schizont or segmenter stage of the parasite, the fluorescence appeared to re-localise to the nuclei (Figure 5.10 m and n). The gametocyte gave a diffuse pattern of staining with the pattern of fluorescence dispersed throughout the parasite cytoplasm and underside of the cell membrane (Figure 5.9 o and p). Positive control experiments, using an anti-PfRan rabbit antisera, were performed in Figure 5.10 q and r.

5.6 Discussion

Genetic and biochemical studies have defined the yeast SNF2/SWI2 and *Drosophila* ISWI proteins as components of a complex machinery that regulates transcription by modifying the structure of chromatin. A *Plasmodium falciparum* homologue (PfSNF2L) of the ISWI, which encodes a estimated 167 kDa protein, was cloned and characterised (Ji & Arnot, 1997). The PfSNF2L mRNA is found to be presented in all erythrocytic stages and in low amounts in the gametocytes by Northern blot studies. To investigate the expression and localisation of PfSNF2L, two mouse sera were raised against two PfSNF2L-His tag fusion proteins FP1 and FP2 (Figure 5.1).

These anti-sera were used to detect PfSNF2L protein on a Western blot in which total proteins were prepared from frozen stocks of asynchronous parasites. However, multiple bands were observed as shown in Figure 5.6A. The sizes of these bands are all smaller than 167 kDa. This indicated that either the PfSNF2L was

degraded during the protein preparation or that the anti-sera cross-reacted with unrelated *Plasmodium* proteins. However, an approximately 167 kDa protein could be detected by both FP1 and FP2 antisera on Western blots which used fresh parasite cultures, extracted with an inhibitor cocktail added during the preparation, although some smaller bands were still observed. Such protein proteolysis is often seen in other malaria proteins (Fonjungo *et al*, 1998; Holder *et al*, 1992; Wiesner *et al*, 1998). For example, during the period of schizont rupture, the 195 kDa merozoite surface protein 1 (MSP1) protein is primarily broken into four major fragments of 83, 28-30, 38 and 42 kDa. The 42 kDa fragment is then further processed into 33 and 19 kDa fragment after merozoite release (Blackman *et al*, 1990; O'Dea *et al*, 1995). It is possible that PfSNF2L degradation may be the result of protein processing *in vivo* rather than artefactual proteolytic degradation during sample preparation, although no PEST sequence is found in PfSNFL protein by the PEST-FIND program. The PEST sequences serve as signals for proteolytic degradation of many proteins such as transcription factors and cell-cycle regulators through ubiquitin-mediated proteolytic pathway (Rechsteiner, 1991).

It is also possible that the PfSNF2L antisera are cross-reacting to other *P. falciparum* proteins. Many serological cross-reactions between distinct *P. falciparum* antigens have been described, e.g. anti-Pf332 antibodies can recognise the 'PEEXXEEXX' consensus repeat of Pf11-1 antigen (Ma *et al*, 1990; Mattei *et al*, 1992; Mattei & Scherf, 1992). The common *Plasmodium* epitopes such as NKND, IKND and KKND can even be recognised by a single monoclonal antibody (mAb) M26-32 (Cheng *et al*, 1991). Comparison of PfSNF2L protein fragments FP1 and FP2 with other *Plasmodium* proteins by BLAST search of GenBank shows no significant homology to any currently known protein. Comparison of immunofluorescence assay (IFA) results of the MESA (mature erythrocyte surface antigen), PfHRP-1 (histidine-rich protein-1) and Pf332 proteins (Coppel *et al*, 1986; Crabb *et al*, 1997; Mattei *et al*, 1992; Mattei *et al*, 1992) indicate that PfSNF2L may also be associated with the classically observed structures seen in Giemsa stained parasited erythrocytes known as Maurer's clefts, a membrane structure located in the parasitised erythrocyte. However careful examination of the granular pattern of PfSNF2L antibody staining indicates that it is distinct from those described for the erythrocyte associated antigens of *P. falciparum* such as MESA (Coppel *et al*, 1986), HRP-1 (Crabb *et al*, 1997; Taylor *et al*, 1987), and Pf332 antigen (Hinterberg *et al*, 1994; Mattei *et al*, 1992). None of these proteins show this developmental switch in appearance of the granular pattern. However, the PfSNF2L polyclonal antisera raised here cannot conclusively distinguish between several hypothetical models of PfSNF2L behavior. A more specific reagent such as an anti-PfSNF2L monoclonal antibody

which could immunoprecipitate the pure protein would be very useful for the further studies of these questions.

Based on IFA results, PfSNF2L is expressed mainly in the asexual parasite with some evidence of its presence in gametocytes. As in other eukaryotic cells, many *P. falciparum* proteins are transported to specific cellular compartments or are secreted. The kinetics of PfSNF2L expression showed that it is detectable in the nuclei of young rings after merozoite invasion. When the parasite develops to late ring stage, a punctate pattern of fluorescence is observed in erythrocyte cytoplasm and under the erythrocyte membrane, although most PfSNF2L is still located in the parasite nucleus. This seems to indicate that at this stage some PfSNF2L is exported across the parasite's plasma membrane and parasitophorous vacuole membrane into the erythrocyte cytosol, although why a transcription-related protein should show this behavior is unclear. The differential localisation of PfSNF2L between nucleus and cytoplasm at different stages of parasite development suggests that its transport might be coupled with the regulation of the parasite cell cycle. However, how and why PfSNF2L, a nuclear transcriptional factor, is moved to the cytoplasm of the parasitised erythrocyte is not understood. It is possible that the protein is at all times within the parasitophorous vacuole membrane and its presence there can be explained by its being synthesised on the cytoplasmic ribosomes. Its apparent presence throughout the parasitised erythrocyte may then be merely post-fixation leakage.

In trophozoites, the PfSNF2L seems to be excluded from parasite nuclei and its granular pattern becomes more pronounced. As the schizont, segmenter and merozoite stages of parasites developed, the PfSNF2L protein is re-located back to the parasite nuclei and the punctate pattern is progressively disappears during those stages. The hbrm and BRG-1 proteins, two human homologues of SNF2/SWI2, have been reported to be phosphorylated and excluded from the condensed chromosomes during mitosis (Muchardt *et al*, 1996). The hbrm protein is further degraded in this phase of cell cycle. It has been suggested that chromosomal exclusion of the human SNF/SWI complex at the G_2/M transition could be part of the mechanism leading to transcriptional arrest during mitosis. However, very little is understood about the G_2 and M phases of the *Plasmodium* life cycle (Arnot & Gull, 1998). There is no clear definition to each phase of malaria cell cycle which can be directly compared with higher eukaryotic cells and yeast. Thus the PfSNF2L exclusion pattern cannot be directly compared to the hbrm and BRG-1. However there are indications that the expression of the PfSNF2L does correlate to the malaria cell cycle in a somewhat similar fashion to the way these proteins behave in other eukaryotic cells. Whether the PfSNF2L is phosphorylated and how it may be transported between parasite cytoplasm and nucleus, and even between parasite and erythrocytic cytoplasm remains

to be determined. There is also a possibility that the PfSNF2L antisera can recognise an unknown parasite protein which can be secreted into the erythrocytic cytoplasm. If this is the case, the PfSNF2L punctate pattern and protein exclusion from the nucleus may be separated into two independent categories. To made further progress on this question could seen to require a monoclonal antibody of defined anti-PfSNF2L specificity for electron microscope immunogold localisation studies.

CHAPTER 6
THE INTERPRETATION OF THE CELL CYCLE OF *PLASMODIUM PALCIPARUM*

This chapter is concerned with some more general and cell-biologically oriented studies of the cell cycle and development of *P. falciparum.* It involves experiments on a) gene expression switches during the parasite's progression through its cell cycle and development and b) the question of apoptosis-like symptoms in *P. falciparum.* Apoptosis has not been previously investigated in *P. falciparum* and some preliminary studies of this phenomenon have been carried out. The objective is to link studies of transcriptional regulation, signal transduction and nuclear transport and to stimulate thought on new directions for further studies.

6.1 Differential display reverse transcriptase-polymerase chain reaction (DDRT-PCR)

In some sense, changes in gene expression pattern *are* the developmental process of the parasite. Cellular and molecular studies have shown that the development of the sexual stages of malaria parasites involves the modification of basic cellular components such as ribosomes and the cytoskeleton (Alano & Carter, 1990; Sinden, 1982) as well as the expression of numerous, but uncharacterised and uncloned sexual stage-specific genes. The technique of differential display of mRNA species was therefore assessed to test its potential to identify *P. falciparum* genes whose expression is restricted to the gametocyte relative to the sexual-stage parasites.

The technique, originally described by Liang and Pardee (Liang & Pardee, 1992), is based on assumptions that every cell expresses some 15,000 genes and, in principle, every individual mRNA molecule can be reverse transcribed and amplified by the polymerase chain reaction. In order to display at least 15,000 bands on an acrylamide gel, one would therefore need about 100 tracks which could resolve 150 bands each. Liang and Pardee therefore subdivided the total reverse transcribed cDNA into 12 fractions and used a set of arbitrary primers for PCR amplification of each cDNA fraction. Denaturing polyacrylamide gel electrophoresis was then used to separate each amplified cDNA fraction.

Figure 6.1 illustrates the general strategy of differential display. An RNA sample is reverse transcribed with each of the four sets of degenerate anchored 3' oligo(dT) primers (T_{11}MN), where M can be G, A or C and N is G, A, T or C. The resulting cDNA population is PCR-amplified using the degenerate 5' primer set, an arbitrary decamer and radioactive nucleotide. The radioactivity labelled PCR products that represent a subpopulation of mRNAs defined by the given primer set are separated

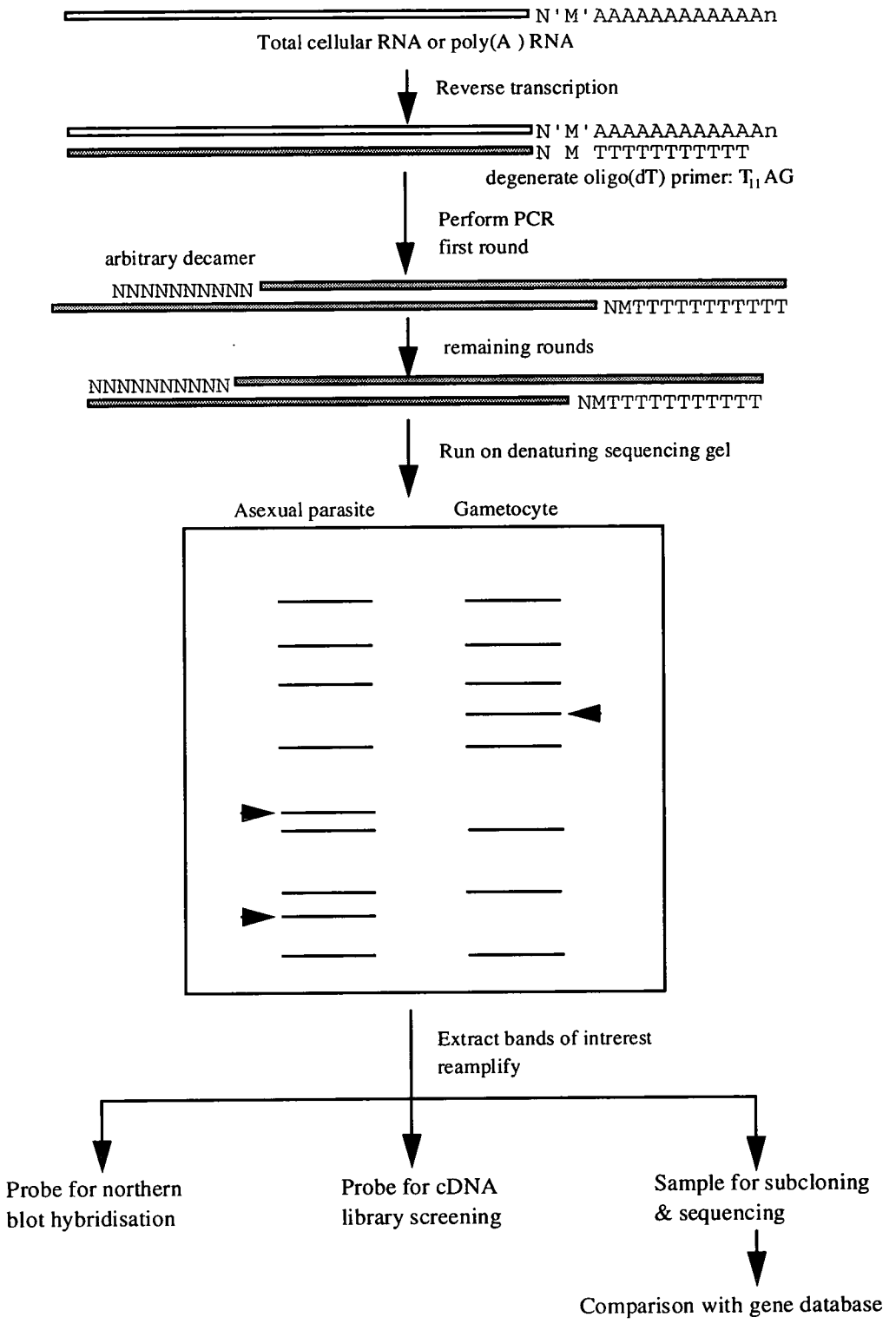


Figure 6.1 Schematic representation of Differential display. T_{11}MN, degenerate oligo(dT) primer: M indicate A,C or G; N can be A,C, G or T; ▭ : RNA; ▭ : DNA

on a denaturing polyacrylamide gel. By changing primer combinations, most of the RNA species in a cell may be represented. Side-by-side comparison of RNA samples from different cells then allows the identification and cloning of differentially expressed genes.

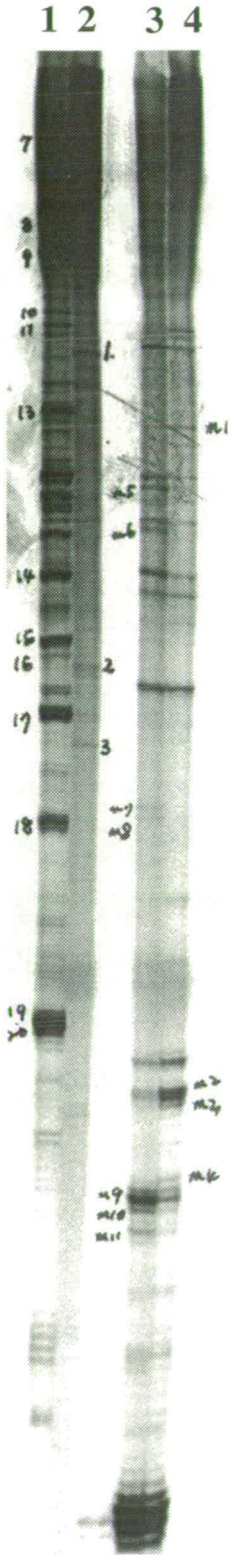
To test the potential of differential display PCR in studies of *P. falciparum* differential gene expression, total RNA from asexual or gametocyte stages of parasites was reverse transcribed with oligo(dT) primer ($T_{11}GA$) and then PCR amplified with a decamer (AATGGATGTG). The total RNA from day 9 and day 10 mouse embryo were used as a experimental control. The ^{35}S-labelled PCR products were separated on 8M urea-acrylamine sequencing gel and visualised by autoradiography. A representative differential display sequencing gel is presented in Figure 6.2. Numerous bands reflecting common gene expression can be observed along with bands which are apparently specific to in particular parasite stages.

Putative stage-specific cDNA bands were excised from the dried gel, eluted and reamplified using the same primer pair used in differential display. 13 cDNA bands were gel purified and re-amplified by PCR from asexual and gametocyte stage parasite gel. These reamplified PCR products of cDNAs were then subcloned into pCR2 vector and sequenced by the automatic DNA sequencer. In analysis of each of the 13 cDNAs, 12 clones failed to identify any significant similarity to any entries in the GenBank database. One clone Pf3-1 shows a high similarity to the yeast Sep1 (strand exchange protein 1) and mouse Dhm2 (dhp1$^+$ homologue in mouse protein 2) genes (see Appendix B), whose products are involved in many cellular processes including meiosis, cellular senescence, and telomere maintenance (Heyer *et al*, 1995; Shobuike *et al*, 1997). Strikingly, the mouse Dhm2 protein has been suggested to participate in gametogenesis in the mouse (Shobuike *et al*, 1997). If it is true that the Pf3-1 is the Sep1 and Dhm2 homologue and is involved in malaria gametogenesis it could be further determined by cloning the entire gene and characterising its protein.

The bands shown on the gametocyte lane were less numeruous than those present during the the asexual stage, indicating that gametocytes exhibit reduced expression of mRNA compared to the sexual stages. The reason is possibly that in the sexual stages, gametocytes cease growth and division, keeping a relatively small amount of protein synthesis to prolong and maintain their basic metabolism. No similar density bands were found between malaria gametocyte and asexual stage parasite lanes indicating that such messages are probably being transcribed at very different levels in sexual and asexual forms.

Since DDRT-PCR was first described, RNA fingerprinting has been used successfully to isolate differentially expressed genes in a large number of experimental systems e.g. mouse and human (McClelland *et al*, 1995). Although no publications

Figure 6.2 Differential display of mRNA from malaria gametocyte versus asexual stage parasites, and mouse embryo day 9 versus day 10 using the DDRT-PCR technique. The purified total RNA (2 mg) was reverse transcribed with [T_{11}GA]. Then the [T_{11}GA] and decamer [AATGGATGTG] were used for the PCR reaction as described in text. Lane 1. asexual stage parasites of *P. falciparum*. Lane 2. gametocyte stage parasites of *P. falciparum*. Lane 3. mouse embryo day 9. Lane 4. mouse embryo day 10. Several candidate cDNA tags that appear to be differentially expressed are marked by numbers.


1 2 3 4

using this technique have been published to date in the malaria field, this experiment indicates that it could be used to characterise and isolate some stage-specific genes of either asexual or gametocyte stages of *P. falciparum* from parasites and provide further insights into the control of sexual differentiation in malaria parasites.

6.2 Does apoptosis occur in the *P. falciparum* parasites?

Although molecular biology techniques such as DDRT-PCR can provide powerful tools for looking at the gene switches during the parasite's development, they still cannot resolve some cell biological problems in the malaria cell cycle. For instance, even in the so-called synchronous cultures, there is very little information to indicate that there is true synchrony in the cell cycle timings of individual cells within these culture populations. Malaria schizogony is also imperfectly understood in cell biological terms. Another example is the apparent mitotic asynchrony within individual parasite cells and variation in the ultimate number of schizonts produced per infected erythrocyte which has been observed during the schizogony. In fact, odd numbers of nuclei are often seen in schizont formation and there is little evidence for a 1, 2, 4, 8, 16, 32 type geometric doubling of nuclei during mitosis. Experiments of this chapter have been carried out to investigate whether apoptosis may be involved in this process.

Apoptosis is a form of programmed cell death and is a common feature of multicellular organisms. It plays an important role in development and homeotasis of metazoan organisms. During apoptosis, a cell activates an intrinsic suicide mechanism that systematically destroys the cell. An early marker of apoptosis is the exposure of phosphatidylserine (PtdSer) normally concentrated in the luminal layer of the cytoplasmic membrane on the cell surface (Martin *et al*, 1995). Nuclear DNA is cleaved between nucleosomes (Wyllie *et al*, 1980) and the chromatin condenses (Kerr *et al*, 1972). The cell finally breaks into membrane-enclosed fragments, the apoptotic bodies (Kerr *et al*, 1972). The initiation of apoptosis is a highly coordinated and regulated process controlled by many apoptotic regulators such as the Bcl-2 family proteins, IL-1β-converting enzyme (ICE) proteases (caspases), and p53.

In mammals, ageing cells are quickly destroyed and removed by apoptotic processing. Recently, cell ageing has also been reported in yeast cells to be indicated by the accumulation of extra-chromosomal rDNA circles (Guarente, 1997; Simon *et al*, 1991; Sinclair & Guarente, 1997; Sinclair *et al*, 1997), although this has not been directly linked with yeast apoptosis. This apoptosis-like phenomenon has however been reported in yeast by several research groups (Madeo *et al*, 1997; Tao *et al*, 1997; Taylor & McKeon, 1997). For example, a *Saccharomyces cerevisiae* cdc48 cell-cycle mutant has been described that exhibits a degree of chromatin condensation and fragmentation and clear exposure of PtdSer on the outer leaflet of its plasma membrane

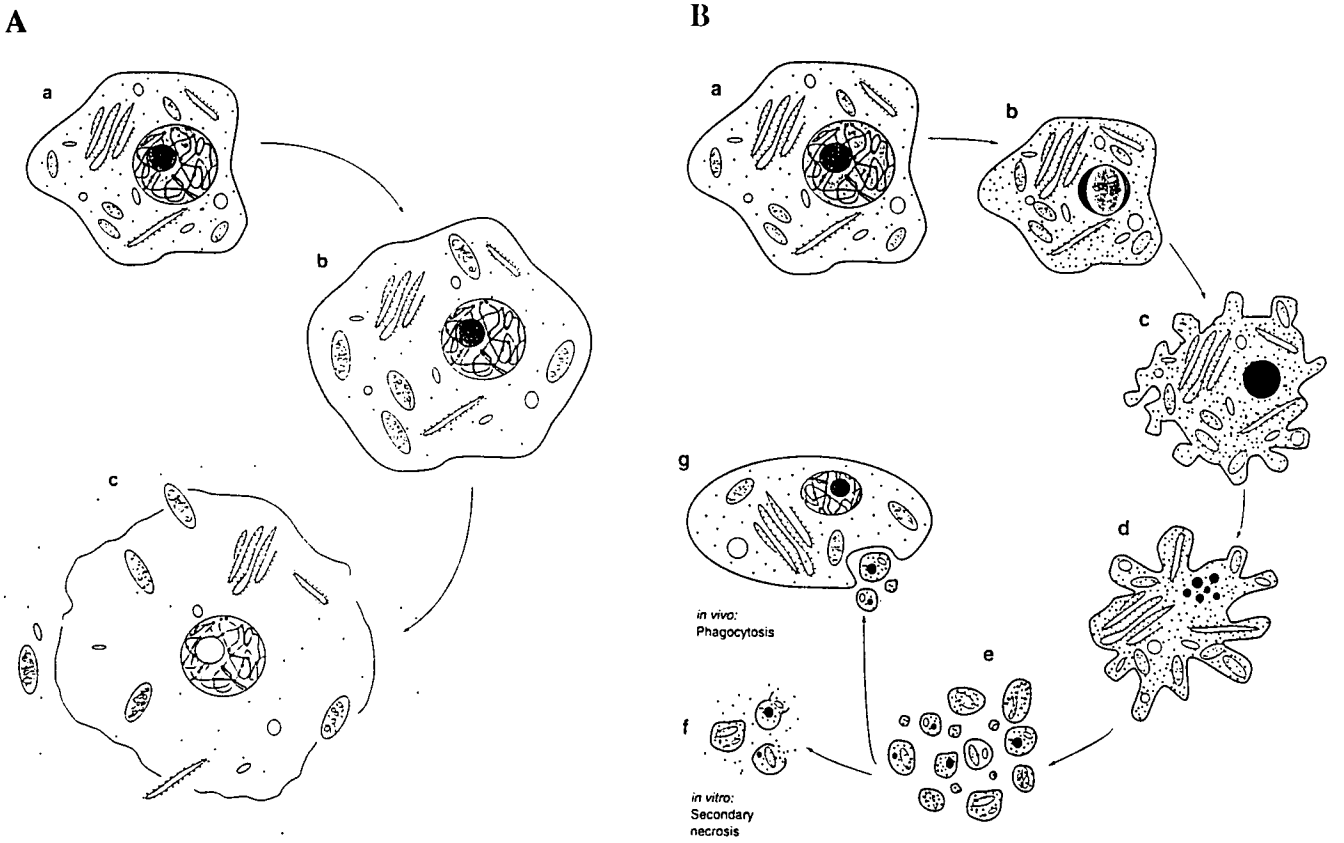


Figure 6.3 The comparison between (**A**) the necrotic cell death and (**B**) programmed cell death. Necrotic death: a normal cell (a) begins to swell (b) before losing its membrane integrity and lysing (c), spilling its contents into the surrounding area. Apoptosis: the normal cell (a) shrinks and the condensed chromatin collapses into crescents around the nuclear envolope (b). The membrane then begins to bulge and bleb (c), while the nucleus ultimately collapsed (d). The blebbing increases and the cell finally breaks apart into number of apoptotic bodies (e), which lyse in vitro (f) and are phagocytosed in vivo (g). (Adapted and modified from Mastrangelo and Betenbaugh, 1998). http://www.uib.no/aasland/SANT.html.

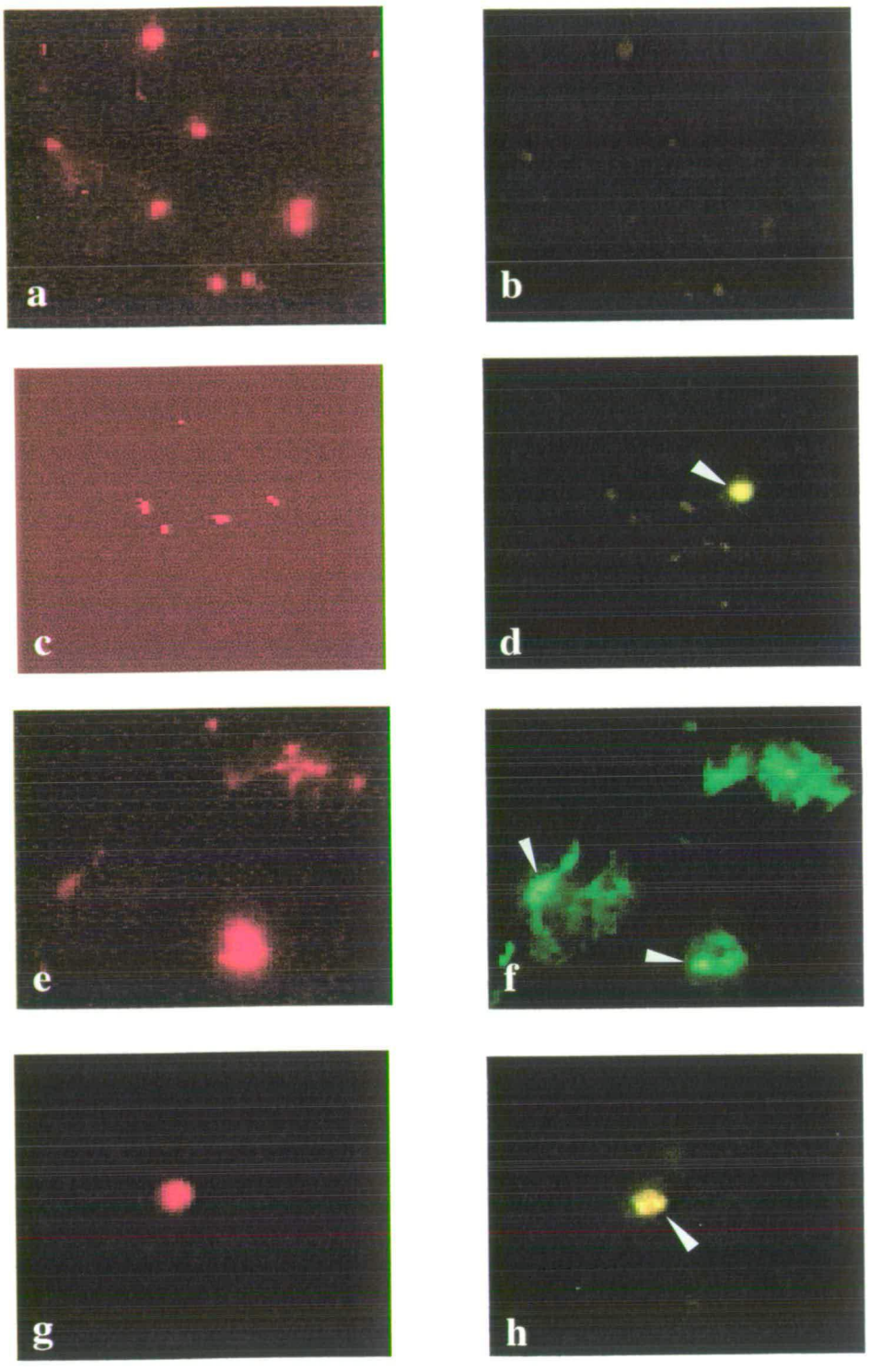
(Madeo *et al*, 1997). This protein was not however homologous to any of the apoptotic regulators that have been found nor was there evidence of nuclear DNA fragmentation in the yeast. A number of observation of apoptosis-like phenomena have also been reported in other unicellular and non-metazoan multicellular eukaryotes. The most compelling evidence of apoptosis-like cell death programmed in a unicellular organism is the developmentally regulated cell death observed in *T. cruzi*. This exhibits membrane blebbing, chromatin condensation and margination strongly reminiscent of that seen in apoptosis, and DNA fragmentation (Ameisen *et al*, 1995). The treatment of *Trypanosoma cruzi* parasites with the antibiotic drug (G418) induces the apoptotic death of epimastigotes (Billaut-Mulot *et al*, 1996). Even in bacteria, self-suicide is also thought to be induced when bacteria are treated with physical agents such as heating and freeze-thaw, chemical agents such as biocides and biological agents such as bacteriophage (Djordjevic *et al*, 1997; Dodd *et al*, 1997).

During some experiments on malaria transfection by electroporation, it was observed that malaria nuclei became condensed and appeared to cleave into dark Geimsa stained granules surrounded by cytoplasm approximately 12 hours after electric shock (data not shown). Such morphology is more similar to the apoptotic phenomenon rather than the necrotic cell death shown in Figure 6.3 (Mastrangelo & Betenbaugh, 1998). The phenomenon is similar to the apoptotic cell death described in *Trypanosoma* and yeast cells and is quite different from necrosis, a process resulting in cell lysis and release of cytoplasmic material after overwhelming cellular injury. These observations suggested that an apoptosis-like mechanism may occur in *Plasmodium* and some further experiments on the phenomenon were carried out.

Apoptosis-like processes in Plasmodium

The terminal deoxynucleotidyl transferase-mediated dUTP nick-end-labeling (TUNEL) assay was used to test for *P. falciparum* apoptosis. The technique is based on the ability of terminal deoxynucleotidyl transferase (TdT) to add several fluorescein-labeled deoxynucleotides at the 3'-OH termini of DNA fragments generated by the endonuclease digestion of genomic DNA that occurs during apoptosis. Figure 6.4 shows 3D7 parasites treated with electric shock, followed by culture for 24 h, then co-stained with TUNEL and propidium iodide (PI) which stains nuclear DNA. More than 50% of electric shocked 3D7 parasites showed an intensive green fluorescent staining with TUNEL assay (Figure 6.4, panels f & h), whereas all parasite nuclei showed red fluorescent stain with PI (Figure 6.4, panels e & g). Apoptotic-like parasites are marked by arrowheads. The parasites showed significant damage by the physical injury with morphologic damage and a large decrease in cell numbers. This is similar to the apoptotic phenomena that mammalian cells display when treated with known

Figure 6.4 DNA fragmentation assay in electric-shock-treated *P. faciparum* parasites using the TUNEL method. Synchronised ring-stage parasites were given electric shock treatment (2.25kV, 25mf, 200W) and subsequently grown in parasite culture medium for 24 h. Panels f & h: the treated parasites stained with the TUNEL assay. Panel b: a TdT-minus negative control. Panel d: the DNase I-treated positive control. Panels a, c, e, & g: IP co-stained parasite nuclei of corresponding negative control, positive control and the TUNEL assays.


a
b
c
d
e
f
g
h

apoptosis-inducers such as dexamethasone (Piqueras *et al*, 1996). The TUNEL assay results indicate that damaged parasites were undergoing DNA fragmentation. Interestingly, in some normal schizonts, some nuclei are also shown bright TUNEL staining, which indicates that these nuclei may be undergoing the apoptosis while the other nuclei in that red blood cell are not (Figure 6.4, panel f). This result might even explain why odd numbers of nuclei are often seen in Giemsa stained schizonts. During mitosis at the schizont stage, some nuclei could be undergoing apoptosis and are thus incapable of further divison. A DNase I treated positive control shows strong TUNEL fluorescence in treated parasites (Figure 6.4, panel d). Nothing is observed under the same conditions when TdT is omitted from the reaction (Figure 6.4, panel b).

Analysis of the isolated chromosomal DNA by agarose electrophoresis did not show the DNA ladder (Figure 6.5). Such results have also been reported from some high eukaryotic cells which undergo apoptosis without the occurrence of a discernable DNA ladder (Oberhammer *et al*, 1993).

Much malaria research has gone into examining ways in which parasites may be killed using drugs/antibodies etc., but little attention has been directed at examining parasite growth limitation/self-regulation. These preliminary results shown here imply that certain apoptosis-like phenomenon happen in malaria parasites. The discovery of programmed cell death in culture form *T. cruzi* suggest that the apoptosis-like process may also happen in parasitic protozoa, although no similar observations with *Plasmodium* have yet been published by other groups. As in yeast, a database search of the *Plasmodium* genome shows no potential homologues of genes involved in the triggering of apoptosis such as the caspases, p53 and bcl-2. Although it is too early to say that the apoptotic cell death does exist in *P. falciparum* parasite, this finding does provide a possible direction for the further studies.

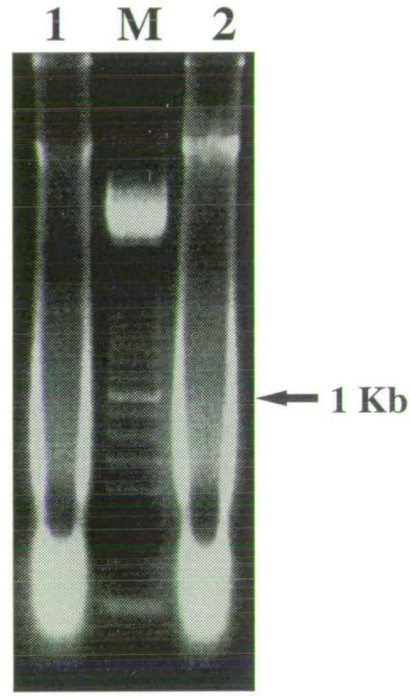


Figure 6.5 Analysis of genomic DNA after treatment of *P. falciparum* parasites without (1) or with electric shock (2). Genomic DNA was analysed in an agarose gel (15%). M is a 100-bp DNA ladder marker.

CHAPTER 7
SOME GENERAL OBSERVATIONS ON THE STRUCTURE OF *PLASMODIUM* GENES

7.1 Introduction

While carrying out computing analysis of the *PfSNF2L*, *PfRan* and *PfRCC1* genes, I found that there was no recent systematic information or general discussion of the organisation of the *Plasmodium* genes. As more and more genes from *Plasodium* and *T. gondii* have been sequenced, it is likely that a re-evaluation of these early observations on protozoan consensus sequences is necessary. Taking advantage of the increased number of sequences which now are available from the GenBank database, I report here an up-to-date analysis of the nuclear coding sequence, intron/exon junctions and the composition of the genome from *P. falciparum*, other *Plasmodium* and *T. gondii* and compare differences among them.

All members of the phylum *Apicomplexa* are parasitic and have complex, multi-stage life cycles. Most are intracellular parasites and some, including *Plasmodium* and *Toxoplasma*, reproduce sexually in one host and asexually in another. Although *Toxoplasma gondii* is normally non-pathogenic in adults, it can cause the emergence of toxoplasmosis as an opportunistic pathogen in AIDS and in congenital infection of new-born children (Sibley & Boothroyd, 1992). Because of the biomedical interest in *T. gondii*, many of its genes have been sequenced and it is the only other example from the *Apicomplexa* for which extensive sequence database exist.

The malaria parasite, during most its life cycle, possesses a haploid genome of 2-4 x10^7 bp with an extremely high A+T content of about 82% (Pollack *et al*, 1982). Coding regions in malaria parasites are somwhat less A+T-rich than flanking noncoding regions and introns (Weber, 1987). Malaria mRNAs are inferred to be capped and most of them are also polyadenylated, although some may lack poly(A) tails (Levitt, 1993; Weber, 1988). Malarial introns tend to be very short and very A+T-rich, and the sequences at their donor and acceptor splice junctions seem to closely conform to eukaryotic splice consensus sequences (Vinkenoog *et al*, 1995; Weber, 1988). The consensus sequences (often termed Kozak sequnce) flanking the translational initiation site of nuclear genes which affects the translational efficiency have been compiled for many eukaryotes (Kozak, 1996; Yamauchi, 1991). The *T. gondii*'s haploid genome contains approximately 8 x 10^7 bp (Cornelissen *et al*, 1984) and has a G+C content of approximately 55%. Its mRNA usually has a 3'-poly(A) tail

(Prince *et al*, 1985). The translation start sites of *Toxoplasma* genes also conform approximately to Kozak's criteria for initiation codons (Seeber, 1997).

7.2 Intron sizes and numbers

Introns are non-coding sequences of DNA inserted into transcribed genes. They are usually located in amino acid coding regions within eukaryotic genes, although in some genes introns are present in 5' or 3'-untranslated regions (UTR) as well. The *P. falciparum* histidine-rich protein (HRP) II and III genes, the *P. berghei* nucleosome assembly protein (PbB7) gene and the *T. gondii* hypoxanthine-xanthine-guanine phosphoribosyl transferase gene and 65 kDa cyst matrix protein gene each have one intron in their 5'- UTR. *P. falciparum* has a 15 kDa vesicule antigen gene which has a intron in its 3'-UTR. The majority of malarial genes have either a few introns or no introns at all. 53 of 111 of *P. falciparum* genes, 19 of 56 of other *Plasmodium* genes and 13 of 28 of *T. gondii* genes have introns. 40 of 53 of *P. falciparum* genes and 18 of 19 of other *Plasmdium* genes have introns near the 5' or 3' end of the coding region. However, only 3 of 15 of *T. gondii* genes have introns near the 5' or 3' end of the coding region. Simpler (sometimes called primitive) eukaryotes tend to have few or no introns, and only the most complex eukaryotes appear to have very large numbers of introns (de Souza *et al*, 1996). Figure 7.1 shows the intron numbers in *P. falciparum*, other *Plasmodium*, and *T. gondii*. Figure 7.2 displays the intron sizes in *P. falciparum*, other *Plasmodium*, and *T. gondii*. The intron positions of α and β-tubulin genes in *P. falciparum*, other *Plasmodium*, and *T. gondii* are shown in Table 7.1.

7.3 Consensus splicing signals in *Plasmodium*

Table 7.2 gives the percentage of each nucleotide at 5' splicing-site positions -10 to +10 and 3' splicing-site positions -15 to +10, for *P. falciparum*, other *Plasmodium* and *T. gondii* introns. A total of 100 introns from 53 *P. falciparum* genes, 46 introns from 19 other *Plasmodium* genes and 60 introns from *T. gondii* genes were collected from GenBank database released at February 3, 1998 (Appendix G). A consensus sequence is derived by analysis of the frequency with which individual bases occur in particular positions (Long *et al*, 1997). Thus a certain nucleotide should represent at least 40% of the total or exceed 30% and be at least twice as frequent as the third highest percentage. The consensus sequences for *Apicomplean* 5' and 3' splicing-sites are listed at the end of each row in Table 7.2. The 5' donor consensus of *P. falciparum*, other *Plasmodium* and *T. gondii* introns

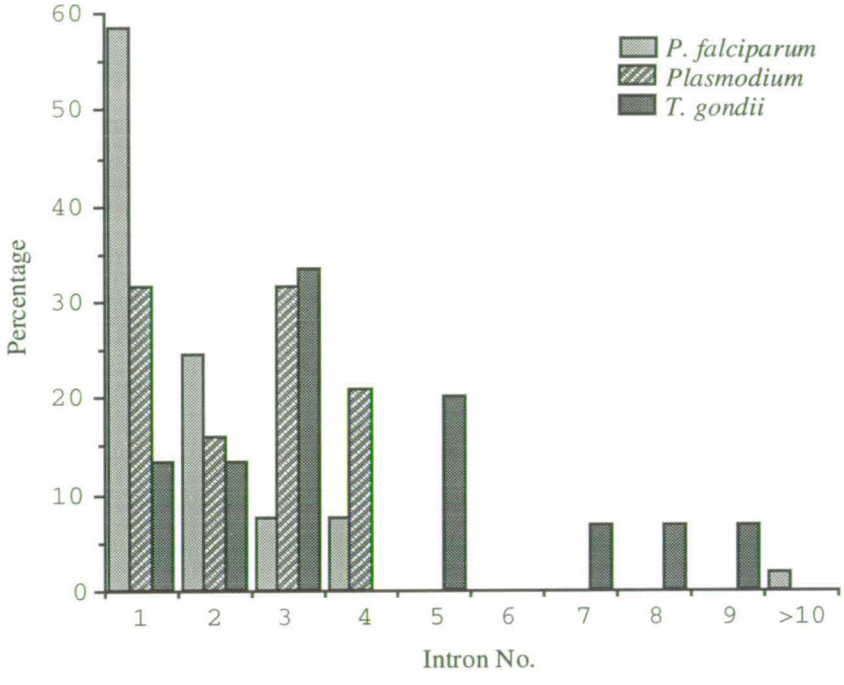


Figure 7.1 Distribution of intron number of protein-coding genes in *P. falciparum*, other *Plasmodium* and *T. gondii*. Each bar represents the percentage of intron number in each number class. The 53 *P. falciparum*, 19 other *Plasmodium* and 13 *T. gondii* protein-coding genes listed in Appendix G were used in this survey.

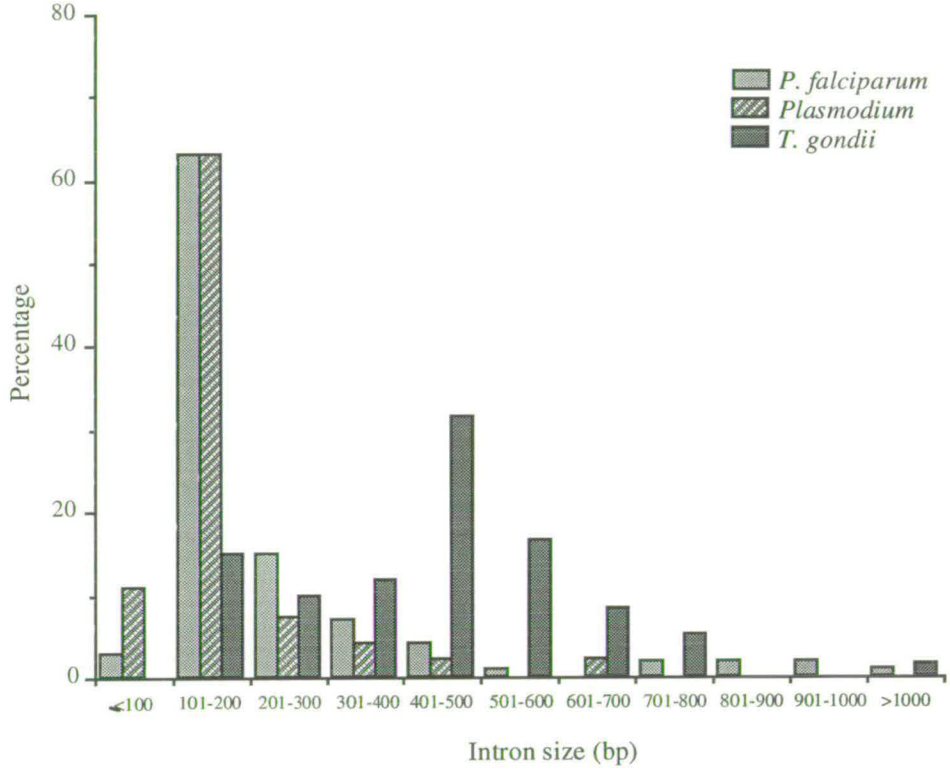


Figure 7.2 Size distribution of introns of protein-coding genes in *P. falciparum,* other *Plasmodium* and *T. gondii*. Each bar represents the number of introns in each size class. The 100 *P. falciparum*, 46 other *Plasmodium* and 60 *T. gondii* introns listed in Appendix G were used in this survey.

Table7.1 The intron locations of α and β-tubulins

Species	Intron location and size[a]			
β-tubulin gene				
P. falciparum	33 (330)	350 (174)	—	—
T. gondii	32-33 (465)	—	397 (141)	426-427 (100)
E.tenella	32-33 (541)	—	397 (95)	426-427 (94)
C. parvum	32-33 (85)	—	—	—
α-tubulin gene				
P. falciparum	21 (211)	61-62 (119)		
T. gondii	21 (511)	61-62 (150)		

a: intron located amino acid residue(s) (inserted intron size (bp))

Table 7.2 5' end intron/ exon junction

(i):	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
Plasmodium falciparum (n=100)																						
R(i):	.11	.02	.06	.25	.05	.22	.22	.30	.45	.37	1.84	1.76	1.37	.82	.35	.49	.34	.29	.46	.85	.65	.65
Percentages:																						
A	42	33	35	42	34	38	46	53	60	27	0	1	89	71	33	54	53	48	56	63	62	67
C	14	18	22	16	19	10	9	18	10	7	0	0	0	5	1	4	10	6	9	2	2	3
G	17	20	13	6	16	12	15	7	7	55	98	2	4	2	40	8	8	15	5	1	7	9
T	27	29	30	36	31	40	30	22	23	11	2	97	7	22	26	34	39	31	30	34	29	21
Consensus	A	N	N	A/T	N	A/T	A/T	A	A	G	G	T	A	A	G	A/T	A/T	A/T	A/T	A/T	A	A
Other *Plasmodia* (n=46)																						
R(i):	.01	.02	.00	-	.00	.08	.21	.24	.25	.24	1.81	1.67	1.14	.53	.08	-	.05	.13	.17	.24	.43	.23
Percentages:																						
A	24	26	17	32	17	31	48	44	52	35	2	2	85	65	13	30	39	44	43	50	61	52
C	28	20	35	20	33	15	28	22	15	9	0	0	2	4	17	24	13	17	11	9	6	9
G	33	17	22	24	28	15	9	4	24	45	98	2	9	13	37	17	22	28	33	28	13	20
T	15	37	26	24	22	39	15	30	9	11	0	96	4	18	33	29	26	11	13	13	20	19
Consensus	N	N	N	N	N	A/T	A	A	A	G	G	T	A	A	N	N	N	A	A	A	A	A
Toxoplasma gondii (n=60)																						
R(i):	.01	.02	-	.11	-	.02	.04	-	.13	1.08	2.00	2.00	.76	.68	1.07	.05	-	-	-	-	-	-
Percentages:																						
A	20	22	25	8	27	38	28	30	47	8	0	0	67	72	8	12	17	20	20	30	32	30
C	18	36	22	29	26	20	34	35	20	5	0	0	3	13	7	28	27	22	20	30	26	27
G	32	25	20	28	27	20	25	18	18	83	100	0	27	8	83	25	28	22	27	20	22	18
T	30	17	33	35	20	22	13	17	15	3	0	100	3	7	2	35	28	36	33	20	20	25
Consensus	N	N	N	N	N	N	N	N	A	G	G	T	A	A	G	N	N	N	N	N	N	N

Frequencies of nucleotides at 5' splice site locations in *P. falciparum,* other *Plasmodium* and *T. gondii.* Percentage of each of the four nucleotides at position -10 to +10 at the 5' spice site. The 100 *P. falciparum,* 46 other *Plasmodium* and 60 *T. gondii* introns listed in Appendix G were used. R(i) is the quantitive measure of conservation at the i'th site, and ranges from zero to two. The deduced 5' consensus sequences are given in this Table. N, indicates nucleotide G, A, T or C. -, indicate R(i) value < 0.00.

conformed to each other and earlier reports of the *P. falciparum* consensus (Vinkenoog *et al*, 1995). Moreover, this 5' *Aplicomplexa* splicing consensus, AAG|GTAAG, is also similar to mammalian consensus, (A/C)AG|GT(A/G)AG, where the bar "|" shows the intron/exon junction position (Mount, 1982; Senapathy *et al*, 1990). Interestingly, in *Saccharomyces cerevisiae* the 5' exon consensus is $A_{47}A_{53}A_{44}N|G$ (the numbers indicate the percentage at each nucleotide position) are quite similar to the *P. falciparum* consensus ($A_{46}A_{53}A_{60}G_{55}|G$), which is thought to interact with the U5 snRNA (small nuclear RNA) to help define 5' splice site (Long *et al*, 1997). Based on studies using constructs with a variety of point mutations, the nucleotides 5 and 6 in the highly conserved loop I sequence (GCCUUUUAC) of U5 snRNA pair with bases -2 and -3 in the 5' exon sequence of pre-mRNA. The -4 exon consensus A is also complementary to nucleotide 7 in the loop (Newman & Norman, 1992). This indicates that *Plasmodium* may share a similar 5' splicing mechanism with mammalian and yeast sequences. However, this 5' exon consensus is not well conserved in *T. gondii.*

Both the *Plasmodium* and *Toxoplasma* 3' intron consensus showed the expected pyrimidine-rich tract and 3' acceptor consensus sequences (Senapathy *et al*, 1990; Vinkenoog *et al*, 1995). In *Plasmodium* the pyrimidine tracts are usually poly(T) on the mRNA coding strand, whereas in *Toxoplasma* the pyrimidine tracts are much less pronounced. The proposed PyPyPuAPy-branch point sequences have also been looked for in the region -10 to -50 of nucleotides upstream of the 3' splicing site. However, only 28 out of 100 introns contained this sequence in the expected region in *P. falciparum*, 8 out of 46 in other *Plasmodium* and 7 out of 60 in *T. gondii.* Although Smith *et al* (1993) have demonstrated that the branch point is important for the scanning process to identify the first 3' splice-site AG, a long pyrimidine stretch was also reported to be important for efficient lariat formation. Whether long the pyrimidine stretch often seen in *Plasmodium* introns plays more a crucial role in malaria splicing than the branch point sequence is unclear. The order of 'CAG ≅ UAG > AAG > GAG' has been suggested for the nucleotide preferences at this position in 3' splice sites. Both malaria and *Toxoplasma* conform to this rule. The fact that there is not quite 100% 5'-G and T at the nucleotides +1 and +2 and 3'-A and G at the nucleotides -2 and -1 could be caused either by misassignment during database input or by certain genes using the non-consensus mRNA splice sites (Jackson, 1991).

Introns have been classified as group I, II or III (Copertino & Hallick, 1993). The eukaryotes have evolved splicing mechanism based on a multi-protein complex, the spliceosome, and spliceosome-dependent introns are thought to be evolved from group II introns. Members of this complex have introns removal via a two-step trans-esterification process (Sharp, 1987). It would be interesting to know if the group II

Table 7.3 3' end intron/exon junction

(i):	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10
Plasmodium falciparum (n=100)																									
R(i):	.56	.70	.75	.86	1.02	.68	.98	.62	.78	.71	1.15	.43	.97	1.44	1.54	.25	.13	.09	.13	.08	.10	.03	.07	.11	.13
Percentages:																									
A	26	21	16	17	11	24	12	9	20	16	10	33	7	94	3	35	29	37	39	39	31	27	38	25	41
C	7	5	9	8	6	8	8	23	7	12	7	7	14	0	0	5	11	14	11	19	13	16	16	19	10
G	5	5	3	1	2	2	1	3	2	2	0	7	1	3	93	41	18	17	18	14	18	22	17	14	21
T	62	69	72	74	81	66	79	65	71	70	83	53	78	3	4	19	42	42	32	28	38	35	29	42	28
Consensus	T	T	T	T	T	T	T	T	T	T	T	T	T	A	G	G	T	A/T	N	N	T	N	A	T	A
Other *Plasmodia* (n=46)																									
R(i):	.50	.19	.32	.77	.51	.28	.70	.36	.48	.32	.70	.08	.31	1.29	1.29	.30	.04	.21	.15	.00	.32	.13	-	-	.05
Percentages:																									
A	20	17	11	11	11	11	11	11	7	4	4	20	7	89	4	37	15	50	44	30	46	44	28	30	39
C	26	35	30	13	30	24	24	26	37	26	20	11	50	5	2	2	26	11	11	20	2	11	20	24	13
G	0	7	7	2	2	11	0	7	4	18	6	37	11	4	89	37	20	13	15	33	22	17	28	20	22
T	54	41	52	74	57	54	65	56	52	52	70	32	32	2	5	22	39	26	30	17	30	28	24	26	26
Consensus	T	T	T	T	T	T	T	T	T	T	T	N	C	A	G	N	T	A	A	N	A	A	N	N	N
Toxoplasma gondii (n=60)																									
R(i):	.14	.19	.04	.09	.13	.23	.27	.14	.37	.13	.30	.16	1.11	2.00	2.00	.17	.06	.01	.06	-	.03	.01	.09	.06	-
Percentages:																									
A	8	5	17	9	12	6	7	7	5	8	7	20	3	100	0	25	20	30	20	25	15	32	33	26	28
C	29	24	27	28	25	25	18	28	28	35	25	8	83	0	0	18	13	20	32	27	25	20	17	17	24
G	23	33	18	28	18	22	23	27	13	22	15	30	2	0	100	47	32	32	20	30	37	30	37	40	20
T	40	38	38	35	45	47	52	38	54	35	53	42	12	0	0	10	35	18	18	18	23	18	13	17	28
Consensus	T	N	N	N	T	T	T	N	T	N	T	T	C	A	G	G	T	A/T	N	N	T	N	A	T	A

Frequencies of nucleotides at 3' splice site locations in *P. falciparum*, other *Plasmodium* and *T. gondii*. Percentage of each of the four nucleotides at position -10 to +10 at the 3' spice site. The 100 *P. falciparum*, 46 other *Plasmodium* and 60 *T. gondii* introns listed in Appendix were used. R(i) is the quantitive measure of conservation at the i'th site, ranges from zero to two. The deduced 3' consensus sequences are given in this Table. N indicates nucleotides G, A, T or C. -, indicates R(i) value < 0.00.

intron-spliceosome mediated splicing machinery exists in the *Aplicomplexa* parasites and whether they have a similar biological function as in mammalian and yeast.

In conclusion, I found that the trends of the *Plasmodium* genes described previously with more limited sets of data (Weber, 1988) are still valid. This indicates that the 5' and 3' consensus sequences presented here are likely to be true for all *Plasmodium* genes.

7.4 Kozak rules and promoter sequences in *P. falciparum*

The sequence flanking potential ATG start codons has been reported to affect the efficiency of initiation of protein translation. A favourable context (often termed the Kozak sequence) is important to stop the 40S ribosomal subunit moving forward in the 3' direction on mRNA and to allow the 60S subunit joining the ribosome to translate the first codon into a peptide. Three data sets for *Plasmodium falciparum*, other *Plasmodium* and *Toxoplasma gondii* containing gene sequences from genomic or complementary DNA sequences were collected and analysed (Appendix G). Sequences were compared from position -10 relative to the ATG initiation codon up to position +10. Each consensus nucleotide was examined and tested for conformity to the rule that it should occur at >50% of all sites or constitute more than twice the frequency of the next abundant nucleotide. The second consensus nucleotide may be chosen if it exceeds 30% of all sites and is at least two times more frequent than the third highest frequency nucleotide. As shown in Table 7.4, three similar consensus sequences were obtained as follows: (A/T)A**A**AA<u>ATG</u>(**A/G**)AN, NAN**A**AA<u>ATG</u>**A**AG and GNC**A**AA<u>ATG</u>**G**NG in *Plasmodium falciparum*, other *Plasmodium* and *Toxoplasma gondii*, respectively. These consensus sequences clearly differ from the vertebrate consensus, GCC**A**CC<u>ATG</u>**G**. The A-residue is clearly dominant at position -3 in these *Apicomplexa* consensus sequences, whereas the A- or G-residues following the ATG codon are dominant at the position +4. The Kozak initiation consensus is derived from analysis of vertebrates (Kozak, 1987; Kozak, 1984), plants (Lutcke *et al*, 1987), *Drosophila* (Cavener, 1987) and yeast (Hamilton *et al*, 1987) and is taken to suggest that the purine residues at positions -3 and +4 play an important role in recognition of AUG by the ribosome in all eukaryotic mRNA. Messenger RNA species in which the first ATG codon lacks the preferred nucleotide in both of these key flanking positions are likely to initiate the translation at the second ATG codon, termed leaky scanning (Kozak, 1991; Kozak, 1996). However, several mRNAs such as *T. gondii GRA3* have also been found to produce two versions of the encoded polypeptide from both the first and second ATG codon (Bermudes *et al*, 1994; Lin *et al*, 1993). Interestingly, a *P. falciparum* gene, *PfSNF2L*, has four in-frame ATG codons within the first 20

Table 7.4 Kozak sequences

	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10
Plasmodium falciparum (n=145)																				
R(i):	.35	.41	.40	.35	.20	.36	.31	.77	.43	.56	2.00	2.00	2.00	.44	.12	.24	.18	.14	.09	.07
Percentages:																				
A	34	38	42	40	37	49	52	73	60	63	100	0	0	51	44	36	46	42	38	34
C	7	6	7	8	7	6	21	4	11	6	0	0	0	3	15	5	10	17	17	14
G	10	8	7	8	19	11	6	8	8	7	0	0	100	34	19	21	25	12	15	20
T	49	48	45	44	37	34	21	15	21	24	0	100	0	12	22	38	19	29	30	32
Consensus	A/T	A/T	A/T	A/T	N	A/T	A	A	A	A	I	A	T	G	A/G	A	N	A	A	N	N
Other *Plasmodia* (n=64)																				
R(i):	.05	.25	.31	-	.21	.31	-	.72	.68	.28	2.00	2.00	2.00	1.22	1.04	.34	.15	.11	.02	.00
Percentages:																				
A	28	18	45	35	39	49	22	84	71	56	100	0	0	83	83	28	41	44	30	39
C	18	16	6	16	12	8	23	8	11	12	0	0	0	2	5	2	11	20	26	17
G	16	12	12	27	10	10	22	4	4	17	0	0	100	15	6	47	33	13	22	13
T	38	54	37	22	39	33	33	4	14	15	0	100	0	0	6	23	15	23	22	31
Consensus	N	T	A/T	N	N	A	N	A	A	A	I	A	T	G	A	A	G	A/G	A	N	N
Toxoplasma gondii (n=15)																				
R(i):	.02	.29	.28	.28	-	-	.09	1.48	.76	.43	2.00	2.00	2.00	.05	-	.02	-	.06	.28	.13
Percentages:																				
A	13	7	0	27	20	33	20	93	74	53	100	0	0	33	20	20	20	47	27	13
C	20	13	40	40	20	20	47	7	13	7	0	0	0	7	20	13	27	27	33	20
G	20	20	27	0	40	33	26	0	13	27	0	0	100	40	33	47	27	13	40	54
T	47	60	33	33	20	13	7	0	0	13	0	100	0	20	27	20	26	13	0	13
Consensus	T	T	C	C	G	N	C	A	A	A	I	A	T	G	G	N	G	N	A	G	G

Frequencies of nucleotides around the translational starting site in the nuclear genes from *P. falciparum*, other *Plasmodium* and *T. gondii*. Percentage of each of the four nucleotides at position -10 to +10 flanking the initiation codon ATG. The 145 *P. falciparum*, 64 other *Plasmodium* and 15 *T. gondii* introns listed in Appendix were used. R(i) is the quantitive measure of conservation at the i'th site, and ranges from zero to two. The deduced translational initiation site consensus sequences are given in this Table. N indicates nucleotides G, A, T or C. -, indicates R(i) value < 0.00.

amino acids of the N-terminal. However, none of them perfectly match the *Plasmodium* most favoured context. Therefore, another rule comes into play, in-frame context which states that the first ATG is usually the best (Kozak, 1996). The description of an *Apicomplexan* consensus translational initiation sequence reported here is a useful tool for the analysis of newly cloned genes or ESTs of *Plasmodium* and *T. gondii*.

CHAPTER 8
GENERAL DISCUSSION AND FUTURE STUDIES

8.1 General discussion

The life cycle of the *Plasmodium falciparum* has been known for many years. However, very little is known about how parasite integrates the control of its multiple cell cycles. To address these tasks, I have looked at G proteins in signalling pathways, PfRCC1 in nuclear transport and PfSNF2L in transcriptional regulation, and then tried to understand more about their roles in the cell cycle. A number of questions regarding the biology and genetics of the parasite have emerged from these investigations.

Do membrane associated heterotrimeric G proteins exist in Plasmodium falciparum?

An original aim of this thesis was to identify *P. falciparum* heterotrimeric G proteins in order to study their role in the regulation of the cell cycle of the malaria parasites. However, no parasite heterotrimeric G-protein gene was successfully cloned in the course of studies. Although a total of 95 clones were selected from PCR products, and sequenced, only three human G_α protein genes, G_i, G_s and G_q, were isolated, each approximately 96% identical to published human sequences in their nucleotide or amino acid sequences (Appendix B). Although some interesting *P. falciparum* genes were found in the course of this work, none of them was a malaria G protein (Appendix B). These 'false' isolations may be due to very low level human DNA or RNA contamination from human white blood cells present when the malaria RNAs were prepared from parasites growing in human blood cultures. The choice of sequence of the 'conserved' primers may also be inappropriate for *P. falciparum* heterotrimeric G protein genes.

The question also arises, are there 'true' trimeric G-protein genes in malaria parasites? All the existing evidence indicates that G_α protein genes in *Trypanosoma*, *Toxoplasma* and *Plasmodium* are only just detectable by biochemical assays which use human or mammalian antibody reagents (Harnett & Harnett, 1998; Thelu *et al*, 1994). For example, a 46 kDa G_α protein has been shown to cross-reaction with the anti-human G_α protein monoclonal antibody by Western blotting detection (Thelu *et al*, 1994), a result which might also be criticised because of possible human G_α protein contamination of the preparations. To date there have been no G_α protein genes cloned from any parasitic protozoan. This may indicate that parasitic protozoan possess a different signalling pathway to their eukaryotic hosts, or it may simply indicate that their structures are not well conserved between distantly related phyla, a situation which always makes the identification of homologues more difficult.

The possible role of Ran and RCC1 in nuclear transport of the P. falciparum parasite

In eukaryotic cells, various cargoes are delivered between the nucleoplasm and cytoplasm through the nuclear pore complex (NPC) gateway. In a recent report (Richards *et al*, 1997), it has been proposed that the RanGTPase controls nuclear transport (see Appendix H) and that the RanGTPase is the common factor required for the transloction of cargo in both directions. Ran-GTP is the active form for translocation. The importin-α and β bind with a nuclear localisation signal (NLS)-containing protein in the cytoplasm and 'dock' at the nuclear pore complex (Gorlich *et al*, 1995; Gorlich *et al*, 1995), whereas Ran-GDP binds at or near NLS-cargo docking sites (Iovine & Wente, 1997). The conversion of Ran-GDP to Ran-GTP by an unknown cytoplasmic RanGEF triggers translocation. The Ran switch, therefore, performs a general gatekeeping function, determining when NLS cargo is admitted to the translocation channel (Iovine *et al*, 1997).

The *P. falciparum* homologue of the *Ran* gene, *PfRan*, has been cloned and characterised (Sultan, 1994). Its protein is localised in both the nucleus and cytoplasm of parasite and its mRNA is abundantly transcribed at the trophozoite and schizont stages. These results indicated that the PfRan protein also travels between the nucleus and the cytoplasm and thus appears to behave in a similar way to the better understood Ran proteins of higher eukaryotic cells. However no detailed biochemical work has yet been carried out in the *P. falciparum* system, which is not particularly well suited to such biochemical cell biology.

A Ran-GTP-free docking zone may be important in controlling NLS cargo docking (Mahajan *et al*, 1997; Saitoh *et al*, 1997). RanBP1 has been thought to shuttle between the cytoplasm and the nucleus and link the nuclear and cytoplasmic pool of Ran (Lounsbury & Macara, 1997; Mahajan *et al*, 1997; Saitoh *et al*, 1997; Zolotukhin & Felber, 1997). However, Richards *et al* have demonstrated that nuclear export sequence (NES)-directed export of RanBP1 requires Ran-GTP but not GTP hydrolysis, and that NLS-directed import and NES-directed export can be functionally uncoupled. Whereas a cytoplasmic RanGAP favors high cytoplasmic Ran-GDP levels, an abundant nuclear Ran guanidine exchange factor (RanGEF) such as RCC1 favours high nuclear Ran-GTP levels.

The PfRCC1, a *P. falciparum* homologue of the RCC1 protein, reported in this thesis is localised predominantly in the parasite nucleus. The results of immuno-fluorescence experiments also indicated that the PfRCC1 is usually co-localised with PfRan in the nucleus and that PfRCC1 is therefore the candidate nuclear pfRanGEF in malaria parasites. It is doubtful that the way forward in this area of malaria research is to try to replicate the cell biological work that has already been done in higher

eukaryotic cells which are both ' free living' and larger and more easily manipulated then malaria parasites.

The PfSNF2L structure and function and chromatin remodelling for transcription in malaria

A *P. falciparum* homologue of the SNF2-like protein gene, *PfSNF2L*, has been cloned in the course of this study (Ji & Arnot, 1997). The predicted protein sequence of PfSNF2L is 1422 amino acids long and contains the seven NTP-binding motifs of the DNA-dependent ATPase domain found in all the SNF2 superfamily of helicase II. Phylogenetic analysis of the PfSNF2L place it within the SNF2-like subfamily of these proteins.

Eukaryotic transcription initiation involves decondensation of DNA from the chromatin which is known to inhibit transcription, replication, recombination and repair by blocking access to the DNA (Koleske & Young, 1995; Peterson, 1996). Two systems that facilitate these nucleosome structure changes are the ATP-dependent chromatin remodelling complexes and the enzymatic complexes which control histone acetylation and deacetylation (Breeden & Nasmyth, 1987; Laurent *et al*, 1993; Peterson & Herskowitz, 1992; Steger & Workman, 1996; Vettese-Dadely *et al*, 1996). There are five chromatin remodelling complexes which have been described to date: SWI/SNF (switching/ sucrose non-fermenting), RSC (remodel the structure of chromatin), NURF (nucleosome remodelling factor), CHRAC (chromatin-accessibility complex) and ACF (ATP-utilising chromatin assembly and remodelling factor) (Cairns *et al*, 1994; Cairns *et al*, 1996; Ito *et al*, 1997; Tsukiyama & Wu, 1995; Varga-Weisz *et al*, 1997). All are multi-subunit complexes with molecular weights ranging from 2 MDa to 500 kDa. Strikingly, the *Drosophila* CHRAC, NURF and ACF all contain ISWI (initial switch) as an ATPase, although each complex appears to remodel chromatin structure by different mechanisms. This suggests that ISWI may function in a distinct mechanistic manner depending upon the exact protein composition of the remodelling complex. PfSNF2L is one of the ISWI homologues. It would be interesting to know whether there are more than one PfSNF2L-containing complexes in *P. falciparum*. However, the attempt to clone another *P. falciparum* PfSNF2L homologue by using PfSNF2L probe at a low stringency condition was unsuccessful and whether there is more than one SNF2 homologues in *P. falciparum* still to be found is an open question which again may be resolved by the whole genome sequencing project.

Interestingly, the p55 protein, one subunit of NURF, has also been identified as a subunit of *Drosophila* CAF1 (chromatin assembly factor 1). p55 homologues are found associated with histone acetyltransferases (HATs) and deacetylases (HDAs).

Many of the genes that show SWI/SNF dependence also demonstrate dependence upon the HATs (Pollard & Peterson, 1997) and strains containing mutations affecting both SWI/SNF function and HAT activities result in lethality or severe sickness (Roberts & Winston, 1997). Thus, many of the diverse chromatin-altering complexes may utilise a common subunit. The growth of *P. falciparum* parasite has been reported to be inhibited by low doses of histone deacetylase inhibitors such as HC-toxin and apicidin (Brosch *et al*, 1995; Darkin-Rattray *et al*, 1996). A preliminary result, shown in Appendix I, also confirmed the HC-toxin finding. This indicates that the HATs and HDAs may also exist in *P. falciparum* and co-operate with PfSNF2L chromatin remodelling complex for the regulation of chromatin dynamics.

The stage specific regulation of PfSNF2L protein

The human brahma (hbrm) and brm-related gene 1 (BRG-1) proteins are phosphorylated and excluded from the condensed chromosomes at the G_2/M transition (Muchardt *et al*, 1996). The hbrm and BRG-1 proteins are also required to co-operate with retinoblastoma (RB) protein to arrest cell in the G_1 phase (Dunaief *et al*, 1994; Trouche *et al*, 1997). Repressing expression of RB protein induces cell entry into S phase and may cause cell transformation (Bartek *et al*, 1997). In yeast, mutation in STH1, a yeast SNF2/SWI2 homologue and a subunit of RSC, has been reported to cause cell cycle arrest in the G_2/M phase of the cell cycle (Cao *et al*, 1997). Taken together, these results imply a role for hbrm, BRG-1 and STH1 chromatin remodelling activity in the progression of the cell cycle. Based on immunofluorescence results, PfSNF2L seems to be excluded from parasite nuclei during the trophozoite stage. However, it is difficult to compare the malaria cell cycle with the normal mammalian or yeast cell cycle. Nothing is known about the phosphorylation of the PfSNF2L protein throughout *P. falciparum* cell cycle. Although it may too early to say that PfSNF2L is definitely involved in the cell cycle progression, these observations do provide a new angle of attack on the problem of the malaria cell cycle.

Immuno-localisation of key P. falciparum effector proteins

In order to carry out further studies on the role of PfRan, PfRCC1 and PfSNF2L proteins in the development of the *P. falciparum* cell cycle, antibodies against these proteins have been produced in the course of the studies in this thesis and also in a previous study (Sultan, 1994). Although these three protein can all be located in the nuclei, some manifestations of their staining patterns are difficult to understand, especially the punctate pattern appearing in PfSNF2L immuno-fluorescence stains. Immuno-electron microscopy may be an essential tool to resolve this problem. However, SNF2 homologues are known to form a very high molecular weight

complex which work co-operatively in the chromatin remodelling events. It is impossible to only study SNF2 homologues without reference to any other components of these large protein complexes.

A framework to explain the appearance of the punctate fluorescent staining might be as follows:

1) During the ring/young trophozoite stages no SNF2L protein synthesis is occurring and the concentrated intra-nuclear SNF2L protein staining is due to pre-existing protein synthesised before merozoite invasion (Chapter 5, Figure 5.9, panel d bottom left).

2) As the trophozoite matures, demands on the transcriptional apparatus increase and the *SNF2L g*ene is transcribed into RNA which in turn is exported into the cytoplasm and translated into SNF2L protein. The IFA pattern observed at this stage (Chapter 5, Figure 5.9, panel c, d, e and f) is that of punctate fluorescence - an apparently cytoplasmic concentration of SNF2L protein into 3-8 small dots arranged in to a roughly circular pattern around the nuclear concentration of SNF2L protein. A possible interpretation of this pattern is that the punctate dots are specialised cytoplasmic sites for message translation into protein.

3) As shizogony is completed, SNF2L fluorescence is present in all the segmenters and has apparently been distributed to all the daughter nuclei of the original invading merozoite.

The key question appears to be, that if SNF2L protein is a part of a large multi-protein transcription activation complex, all the components of which are translated in the cytoplasm, then where is this complex assembled, in the cytoplasm or in the nucleus? In relatively long-lived eukaryotic cells, it would seem to be obvious (although probably as yet uninvestigated) that individual proteins are re-imported through nuclear pore complexes for intra-nuclear assembly since the complexes are too big to pass through the pores and into the nucleus intact. However, in *P. falciparum* it is possible to speculate that the complex problem of schizogony (1 nucleus $\rightarrow$ 8-30 nuclei in 6 hours) might be simplified if chromatin associated multi-protein complexes do not have to re-enter the S phase nucleus but are assembled in the cytoplasm and are apportioned to daughter nuclei during a phase of nuclear-cytoplasmic fusion accompanying the segmentation of replicated chromosomes into daughter merozoites.

This speculative proposal would imply that the observed staining patterns represents local intra-cytoplasmic concentration of SNF2L protein into transcription

initiation complexes which are, or will become, associated with new schizont nuclei without passing through nuclear pore complexes.

Plasmodium falciparum: a difficult organism for cell cycle studies

As in some unicellular eukaryotes, e.g. budding yeast, the nuclear envelop of *P. falciparum* cell does not break down during mitosis and microtubules are nucleated by spindle pole bodies rather than by centrosomes, which is different from multicellular eukaryotes (Arnot & Gull, 1998; Leete & Rubin, 1996; Read *et al*, 1993). In mammalian cells, the nuclear envelop breaks down during mitosis and microtubules are formed from centrosomes (Murray & Hunt, 1993). Moreover, *P. falciparum* is a rather small intracellular parasite and it is not possible to observe the G_1, S, G_2 and M phase states in its tiny nucleus by conventional light microscopy. The use of tissue culture cell fusion techniques in the studies of cell cycle transitions is also impossible in such a cell. Another difficulty for malaria cell cycle study is that there are no available techniques to select *P. falciparum* cell cycle defect mutants for further analysis.

Recently, in budding yeast, chromosome condensation has been directly observed by electron microscopy, although the chromosomes did not form a typical chromosome structure (Gottschling & Berg, 1998; Guacci *et al*, 1997; Sutani & Yanagida, 1997). Several chromosome associated proteins such as MCD1 (mitotic chromosome determinant) have been found to be important for chromsome condensation and sister chromatin cohesion, and are essential for successful completion of mitosis. However, no homologues of these proteins have been reported from *P. falciparum* to date. It is impossible to understand malaria chromosome condensation and sister chromatin cohesion during mitosis without information on the behavior of such proteins.

It seems clear that the central problem in understanding the molecular cell biology underlying the replicative phase of malaria parasites-schizogony, remains. The best way forward may be to combine malaria genome project data on candidate cell cycle and mitotic replication genes and proteins with cell biology and ultra-structural studies aimed at understanding the function of such proteins during parasite replication. Genetic techniques such as gene knock-outs can also help piece together the sequential events of this rather mysterious parasite replicative process.

REFERENCES

Aasland, R., Stewart, A. F. and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends in Biochemical Science *21*, 87-88.

Aebi, M., Clark, M. W., Vijayraghavan, U. and Abelson, J. A. (1990). A yeast mutant, PRP20, altered the mRNA metabolism and maintenance of the nuclear structure, is defective in a gene homologous to the human gene RCC1 which is involved in the cintrol of chromosome condensation. J. Mol. Gen. Genet. *224*, 72-80.

Alano, P. and Carter, R. (1990). Sexual differentiation in malaria parasites. Annual Review of Microbiology *44*, 429-429.

Aley, S. B., Sherwood, J. A., Marsh, K., Eidelman, O., and Howard, R. J. (1986). Identification of isolate-specific proteins on sorbitol-enriched *Plasmodium falciparum* infected erythrocytes from Gambian patients. Parasitology *92*, 511-525.

Aldritt, S. M., Joseph, J. T. and Wirth, D. F. (1989). Sequence identification of cytochrome *b* in *Plasmodium gallinaceum*. Mol. Cell. Biol. *9*, 3614-3620.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389-3402.

Ameisen, J. C., Idziorek, T., Billautmulot, O., Loyens, M., Tissier, J. P., Potentier, A. and Ouaissi, A. (1995). Apoptosis in a unicellular eukaryote (*Trypanosoma cruzi*) implications for the evolutionary origin and role of programmed cell death in the control of cell proliferation, differentiation and survival. Cell Death Different. *2*, 285-300.

Arnot, D. E. and Gull, K. (1998). The *Plasmodium* cell-cycle: facts and questions. Ann. Trop. Med. Parasitol, *92*, In press.

Aslund, L., Franzen, L., Weshin, G., Persson, T., Wigzell, H. and Pettersson, U. (1985). Highly reiterated non-coding sequence in the genome of *Plasmodium falciparum* is composed of 21-base pair tandem repeats. Methods Mol. Biol. *185*, 509-516.

Barberis, A., Pearlberg, J., Simkovich, N., Farrell, S., Reinagel, P., Bamdad, C., Sigal, G. and Ptashne, M. (1995). Contact with a component of the RNA polymerase II holoenzyme suffices for gene activation. Cell *81*, 359-368.

Bartek, J., Bartkova, J. and Lukas, J. (1997). The retinoblastoma protein pathway in cell cycle control and cancer. Exp Cell Res *237*, 1-6.

Bender, K., Betschart, B. and Hecker, H. (1992). Histone-DNA interactions in the chromatin of procyclic *Trypanosoma brucei brucei*. Parasitol. Res. *78*, 495-500.

Bennett, B. J., Thompson, J. and Coppel, R. L. (1995). Identification of *Plasmodium falciparum* histone 2B and histone 3 genes. Mol. Biochem. Parasitol. *70*, 231-233.

Benson, D. A., Boguski, M. S., Lipman, D. J., Ostell, J. and Francis, B. F. (1998). GenBank. Nucleic Acids Res. *26*, 1-7.

Bermudes, D., Dubremetz, J. F., Achbarou, A. and Joiner, K. A. (1994). Cloning of a cDNA encoding the dense granule protein GRA3 from *Toxoplasma gondii*. Mol. Biochem. Parasitol. *68*, 247-257.

Berridge, M. J. (1993). Inositol triphosphate and calcium signalling. Nature *361*, 315-325.

Bigg, B. A., Kemp, D. J. and Brown, G. V. (1989). Subtelomeric chromosome deletions in field isolates of *Plasmodium falciparum* and their relationship to loss of cytoadherence in vitro. Proc. Natl. Acad. Sci. USA *86*, 2428-2432.

Billaut-Mulot, O., Fernandez-Gomez, R., Loyens, M. and Ouaissi, A. (1996). *Trypanosoma cruzi* elongation factor 1-alpha: nuclear localization in parasites undergoing apoptosis. Gene *174*, 19-26.

Birago, C., Pace, T., Barca, S., Picci, L. and Ponzi, M. (1996). A chromatin-associated protein is encoded in a genomic region highly conserved in the *Plasmodium* genus. Mol. Biochem. Parasitol. *80*, 193-202.

Bischoff, F. R. and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature *354*, 80-82.

Bischoff, F. R. and Ponstingl, H. (1991). Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. Proc. Natl. Acad. Sci. USA *88*, 10830-10834.

Bischoff, F. R. and Ponstingle, H. (1995). Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran GAP1. Methods Enzymol. *257*, 135-144.

Blackman, M. J., Heidrich, H. G., Donachie, S., McBride, J. S. and Holder, A. A. (1990). A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. J. Exp. Med. *172*.

Blumenthal, T. and Spieth, J. (1996). Gene structure and organization in *Caenorhabditis elegans*. Curr. Opin. Genet. Dev. *6*, 692-698.

Blumenthal, T. and Steward, K. (1996). RNA processing and gene structure. In C. elegans II.: Cold Spring Harbor Laboratory), pp. 1-24.

Bodansky, S., Mintz, L. and Holmes, D. (1979). The mesokaryote *Gyrodinium cohnii* lack nucleosomes. Biochem. Biophys. Res. Commun. *88*, 1329-1336.

Bork, K. and Koonin, E. V. (1993). An expanding family of helicases with the DEAD/H superfamily. Nucleic Acids Res. *21*, 751-752.

Bourne, H. R. (1993). GTPase everywhere. In GTPase., B. F. Dickey and L. Birnbaumer, eds. (New York: Springer-Verlag.), pp. 3-15.

Bourne, H. R., Sanders, D. A. and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. Nature *349*, 117-127.

Braun-Breton, C., Blisnick, T., Jouin, H., Barale, J. C., Rabilloud, T., Langsley, G. and Pereira da Silva, L. H. (1992). *Plasmodium chabaudi* p68 serine protease activity required for merozoite entry into mouse erythrocytes. Proc. Natl. Acad. Sci. USA *89*, 9647-9651.

Breeden, L. and Nasmyth, K. (1987). Cell cycle control of the yeast HO gene: cis and trans-acting regulators. Cell *48*, 387-397.

Brondel, V., Bucher, P., Nourbakhsh, I., Blaisdell, B. E. and Karlin, S. (1992). Methods and algorithms for statistical analysis of protein sequences. Proc. Natl. Acad. Sci. USA *89*, 2002-2006.

Brosch, G., Ransom, R., Lechner, T., Walton, J. D. and Loidl, P. (1995). Inhibition of maize histone deacetylases by HC toxin, the host-selective toxin of *Cochliobolus carbonum*. Plant Cell *7*, 1941-1950.

Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987). XL1-Blue: a high efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. Biotechniques *5*, 376.

Buluwela, L., Forster, A., Boehm, T. and Rabbitts, T. H. (1989). A rapid procedure for colony screening using nylon filters. Nucleic Acids Res. *17*, 542.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. Cell *77*, 1-3.

Burgess, M. (1997). The hundred year's war. Biomedical News *10*, 8-9.

Cahill, M. A., Janknecht, R. and Nordheim, A. (1996). Signalling pathways: Jack of all cascades. Curr. Biol. *6*, 16-19.

Cairns, B., Kim, Y. J., Sayre, M. H., Laurent, B. C. and Kornberg, R. D. (1994). A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. Proc. Natl. Acad. Sci. USA *91*, 1950-1954.

Cairns, B. R. (1998). Chromatin remodeling machines: similar motors, ulterior motives. Trends Biochem. Sci. *23*, 20-25.

Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B. and Kornberg, R. D. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell *87*, 1249-1260.

Cantley, L. C., Auger, K. R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991). Oncogenes and signal transduction. Cell *64*, 281-302.

Cao, Y., Cairns, B. R., Kornberg, R. D. and Laurent, B. C. (1997). Sfh1p, a component of a novel chromatin-remodelling complex, is required for cell cycle progression. Mol. Cell. Biol. *17*, 3323-3334.

Cary, C., Lamont, D., Dalton, J. P. and Doerig, C. (1994). *Plasmodium falciparum* chromatin: nucleosomal organisation and histone-like proteins. Parasitol. Res. *80*, 255-258.

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. Nucleic Acids Res. *15*, 1353-1361.

Chao, D. M., Gadbois, E. L., Murray, P. J., Anderson, S. F., Sonu, M. S., Parvin, J. D. and Young, R. A. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. Nature *380*, 82-84.

Cheesman, S., Horrocks, P., Tosh, K., and Kilbey, B. (1998). Intraerythrocytic expression of topoisomerase II from *Plasmodium falciparum.* is developmently regulated. Mol. Biochem. Parasitol. *92*, 39-46.

Cheng, Q., Jones, G., Liu, E. X., Kidson, C. and Saul, A. (1991). Identification of a commone *Plasmodium* epitope (CPE) recognised by a pan-specific inhibitory monoclonal antibody. Mol. Biochem. Parasitol. *49*, 73-82.

Clapham, D. E. (1995). Calcium signaling. Cell *80*, 259-268.

Clark, K. and Sprague, G. (1989). Yeast pheromone response pathway: characterisation of a suppressor that restores mating to receptorless mutants. Mol. Cell. Biol. *9*, 2682- .

Cobb, N. H. and Goldsmith, E. J. (1995). How MAP kinases are regulated. J. Biol. Chem. *270*, 14843-14846.

Copertino, D. W. and Hallick, R. B. (1993). Group II and group III introns of twintrons: potential relationships with pre-mRNA introns. Trends Biochem. Sci. *18*, 467-471.

Coppel, R. L., Culvenor, J. G., Bianco, A. E., Crewther, P. E., Stahl, H. D., Brown, G. V., Anders, R. F. and Kemp, D. J. (1986). Variable antigen associated with the surface of erythrocytes of *P. falciparum.* Mol. Biochem. Parasitol. *20*, 265-277.

Corcoran, L. M., Forsyth, K. P., Bianco, A. E., Brown, G. V. and Kemp, D. J. (1986). Chromosome size polymorphisms in *Plasmodium falciparum* can involve deletions and are frequent in natural parasite populations. Cell *44*, 87-95.

Cornelissen, A. W. C. A., Overdulve, J. P. and ven der Ploeg, M. (1984). Determination of nuclear DNA of five Eucoccidian parasites, *Isospora (Toxoplasma) gondii, Sarcocystis cruzi, Eimeria tenella, E. acervulina,* and *Plasmodium berghei*, with special reference to gametogenesis and meiosis in *I. (T.) gondii.* Parasitology *88*, 531-553.

Cote, J., Quinn, J., Workman, J. L. and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science *265*, 53-60.

Coulter, L. J. and Hide, G. (1995). *Trypanosoma brucei*: characterisation of a life cycle stage-specific G-protein. Exp. Parasitol. *80*, 308-318.

Crabb, B. S., Cooke, B. M., Reeder, J. C., Waller, R. F., Caruana, S. R., Davern, K. M., Wickham, M. E., Brown, G. V., Coppel, R. L. and Cowman, A. F. (1997). Targeted gene disruption showns that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. Cell *89*, 287-296.

Creasey, A. M. (1996). The inheritance of extranuclear DNA in malaria parasites. (Edinburgh: University of Edinburgh).

Creedon, K. A., Kaslow, D. C., Rathod, P. K. and Wellem, T. E. (1992). Identification of a *Plasmodium falciparum* histone H2A gene. Mol. Biochem. Parasitol. *54*, 113-116.

Dame, J. B., Arnot, D. E., Bourke, P. F., Chakrabarti, D., Christodoulou, Z., Coppel, R. F., Cowman, A. F., Craig, A. G., Fischer, K., Foster, J., Goodman, N., Hinterberg, K., Holder, A. A., Holt, D. C., Kemp, D. J., Lanzer, M., Lim, A., Newbold, C. I., Ravetch, J. V., Reddy, G. R., Rubio, J., Schuster, S. M., Su, X.-Z., Thompson, J. K., Vital, F., Wellems, T. E. and Werner, E. B. (1996). Current status of the *Plasmodium falciparum* genome project. Mol. Biochem. Parasitol. *79*, 1-12.

Darkin-Rattray, S. J., Gurnett, A. M., Myers, R. W., Dulski, P. M., Crumley, T. M., Allocco, J. J., Cannova, C., Meinke, P. T., Colletti, S. L., Bednarek, M. A., Singh, S. B., Goetz, M. A., Dombrowski, A. W., Polishook, J. D. and Schmatz, D. M. (1996). Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. Proc. Natl. Acad. Sci. USA *93*, 13143-13147.

Dasso, M. (1993). RCC1 in the cell cycle: the regulator of chromosome condensation takes on new roles. Trends Biochem. Sci. *18*, 96-101.

de Souza, S. J., Long, M. and Gilbert, W. (1996). Introns and gene evolution. Genes to Cells *1*, 493-505.

Demeter, J., Morphew, M. and Sazer, S. (1995). A mutation in the RCC1-related protein pim1 results in nuclear envelope fragmentation in fission yeast. Proc. Natl. Acad. Sci. USA *92*, 1436-1440.

Devereux, J., Haederli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12*, 387-395.

Djordjevic, G. M., O'Sullivan, D. J., Walker, S. A., Conkling, M. A. and Klaenhammer, T. R. (1997). A triggered-suicide system designed as a defense against bacteriophages. J. Bacteriol. *179*, 6741-6748.

Dodd, C. E. R., Sharman, R. L., Bloomfield, S. F., Booth, I. R. and Stewart, G. S. A. B. (1997). Inimical processes: Bacterial self-destruction and sub-lethal injury. Trends Food Sci. Technol. *8*, 238-241.

Doerig, C. D. (1997). Signal transduction in malaria parasites. Parasitol. Today *13*, 307-313.

Doerig, C. D., Horrocks, P., Coyle, J., Carlton, J., Sultan, A., Arnot, D. and Carter, R. (1995). Pfcrk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum*. Mol. Biochem. Parasitol. *70*, 167-174.

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

Dore, E., Pace, T., M., P., Scotti, R. and Frontali, C. (1986). Homologous telomeric sequences are present in different species of the genus *Plasmodium*. Mol. Biochem. Parasitol. *21*, 121-127.

Drivas, G. T., Shih, A., Coutavas, E., Rush, M. G. and D'Eustachio, P. (1990). Characterisation of four novel RAS-related genes expressed in a human teratocarcinoma cell line. Mol. Cell Biol. *10*, 1793-1798.

Dulzewski, A. R., Ling, I. T., Rangachari, R., Bates, P. A., and Wilson, R. J. M. (1984). A simple method for isolating viable mature parasites of *Plasmodium falciparum* from cultures. Trans. R. Soc. Trop. Med. Hyg. *78*, 622-624.

Dunaief, J. L., Strober, B. E., Guha, S., Khavari, P. A., Alin, K., Luban, J., Begemann M, Crabtree, G. R. and Goff, S. P. (1994). The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. Cell *79*, 119-130.

Edgar, B. A., Sprenger, F., Duronio, R.J., Leopold, P. and O'Farrell, P.H. (1994). Distinct molecular mechanisms regulate cell cycle timing at successive stages of *Drosophila* embryogenesis. Genes Dev. *8*, 440-452.

Eisen, J. A., Sweder, K. S. and Hanawalt, P. C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Res. *23*, 2715-2723.

Elfring, L. K., Deuring, R., McCallum, C. M., Peterson, C. L. and Tamkun, J. W. (1994). Identification and characterization of *Drosophila* relatives of the yeast transcriptional activator SNF2/SWI2. Mol. Cell. Biol. *14*, 2225-2234.

Ellis, N. A. (1997). DNA helicases in inherited human disorders. Curr. Opin. Genet. Dev. *7*, 354-363.

Farrell, S., Simkovich, N., Wu, Y., Barberis, A. and Ptashne, M. (1996). Gene activation by recruitment of the RNA polymerase II holoenzyme. Genes Dev. *10*, 2359-2367.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analyt. Biochem. *132*, 6-13.

Felsenfeld, G. (1996). Chromatin unfolds. Cell *86*, 13-19.

Fenton, B. and Walliker, D. (1990). Genetic analysis of polymorphic proteins of the human malaria parasite *Plasomdium falciparum*. Genet. Res. Cambridge *55*, 81-86.

Fonjungo, P. N., Stuber, D. and McBride, J. S. (1998).
Antigenicity of recombinant proteins derived from rhoptry-associated protein 1 of *Plasmodium falciparum*. Infect. Immun. *66*, 1037-1044.

Foote, S. J. and Kemp, D. J. (1989). Chromosomes of malaria parasites. Trends Genet. *5*, 337-342.

Forrester, W., Stutz, F., Rosbash, M. and Wickens, M. (1992). Defects in mRNA 3'-end formation, transcription initiation, and mRNA transport associated with the yeast mutation prp20: possible coupling of mRNA processing and chromatin structure. Genes Dev. *6*, 1914-1926.

Frasch, M. (1991). The maternally expressed *Drosophila* gene encoding the chromatin-binding protein BJ1 is a homolog of the vertebrate regulator of chromosome condensation, RCC1. EMBO J. *10*, 1225-1230.

Gardner, M. J., Williamson, D. H. and Wilson, R. J. M. (1991). A circular DNA in malaria parasites encodes an RNA polymerase like that of prokaryotes and chloroplasts. Mol. Biochem. Parasitol. *44*, 115-124.

Gaudreau, L., Schmid, A., Blaschke, D., Ptashne, M. and Horz, W. (1997). RNA polymerase II holoenzyme recuitment is sufficient to remodel chromatin at the yeast PHO5 promoter.. Cell *89*, 55-62.

Gerace, L. (1995). Nuclear export signals and the fast track to the cytoplasm. Cell *82*, 341-344.

Gibbons, R. J., Picketts, D. J., Villard, L. and Higgs, D. R. (1995). Mutation in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). Cell *80*, 837-845.

Goldfarb, D. S. (1997). Whose finger is on the switch? Science *276*, 1814-1816.

Gorbalenya, A. E. and Koonin, E. V. (1993). Helicases - amino acid sequence conparisons and structure function relationships. Current Opinion Structural Biology *3*, 419-429.

Gorlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R., Hartmann, E. and Prehn, S. (1995). Two different subunits of importin cooperate to recognise nuclear localisation signals and bind them to the nuclear envelope. Curr. Biol. *5*, 383-392.

Gorlich, D., Pante, N., Kutay, U., Aebi, U. and Bischoff, F. R. (1996). Identification of different roles for Ran GDP and Ran GTP in nuclear protein import. EMBO J *15*, 5584-5594.

Gorlich, D., Vogel, F., Mills, A. D., Hartmann, E. and Laskey, R. A. (1995). Distinct functions for the two importin subunits in nuclear protein import. Nature *377*, 247-248.

Gottschling, D. E. and Berg, B. L. (1998). Chromosome dynamics: yeast pulls it apart. Curr. Biol. *8*, R76-R79.

Graeser, R., Franklin, R. M. and Kappes, B. (1996). Mechanisms of activation of the cdc2-related kinase PfPK5 from *Plasmodium falciparum*. Mol. Biochem. Parasitol. *79*, 125-127.

Grant, K., Harnett, M. and Harnett, W. (1997). Acanthocheilonema viteae: stage-specific expression of G-protein α-subunits. Exp. Parasitol. *86*, 219-227.

Grant, P. A., Sterner, D. E., Duggan, L. J., Workman, J. L. and Berger, S. L. (1998). The SAGA unfolds: convergenece of transcription regulation in chromatin-modifying complexes. Trends Cell Biol. *8*, 193-197.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. Nature *389*, 349-352.

Guacci, V., Koshland, D. and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*. Cell *91*, 47-57.

Guarente, L. (1997). Link between aging and the nucleolus. Genes Dev. *11*, 2449-2455.

Guntaka, R. V., Kandala, J. C. and Reddy, V. D. (1992). Cloning and characterisation of a highly conserved HNG-like protein (Pf16) gene from *Plasmodium falciparum*. Biochem. Biophys. Res. Commun. *182*, 412-419.

Hall, A. (1990). The cellular functions of small GTP-binding proteins. Science *249*, 635-640.

Halonen, S. K., Weidner, E. and Siebenaller, J. F. (1996). Evidence for heterotrimeric GTP-binding proteins in *Toxoplasma gondii*. J. Euk. Microbiol. *43*, 187-193.

Hamilton, R., Watanabe, C. K. and Herman, A. B. (1987). Compilation and comparison of the sequence context around the AUG start codons in *Saccharomyces cerevisiae* mRNAs. Nucleic Acids Res. *15*.

Hampsey, M. (1997). A SAGA of histone acetylation and gene expression. Trends Genet. *13*, 427-429.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. *166*, 557.

Hanahan, D. and Meselson, M. (1983). Plasmid screening at high colony density. Methods Enzymol. *100*, 333-344.

Hardie, D. G. (1991). Biochemical messengers: hormones, neurotransmitters and growth factors. (London: Chapman & Hall).

Hare, J. D. (1986). Two-color flow-cytometric analysis of the growth cycle of *Plasmodium falciparum* in vitro: identification of cell cycle compartments. J. Histochemistry and Cytochemistry *34*, 1651-1658.

Hare, J. D. and Bahler, D. W. (1986). Analysis of Plasmodium falciparum growth in culture using acridine orange and flow cytometry. J. Histochem. Cytochem. *34*, 215-220.

Harnett, W. and Harnett, M. M. (1998). Heterotrimeric Guanine nucleotide-binding proteins in eukaryotic parasites. Parasitol. Today *14*, 27-31.

Hawking, F., Wilson, M. E. and Gammage, K. (1971). Evidence for cyclic development and short lived maturity in the gametocytes of *Plasmodium falciparum*. Trans. R. Soc. Trop. Med. Hyg. *65*, 549-559.

Hepler, J. R. and Gilman, A. G. (1992). G proteins. Trends Biochem. Sci. *17*, 383-387.

Herman, R. D. (1988). Alternatives for the initiation of translation. Trends Biochem. Sci. *14*, 219-222.

Herrmann, M., Lorenz, H.-M., Voll, R., Grunke, M., Woith, W. and Kalden, J. R. (1994). A rapid and simple method for the isolation of apoptotic DNA fragments. Nucleic Acids Res. *22*, 5506-5507.

Heyer, W. D., Johnson, A. W., Reinhart, U. and Kolodner, R. D. (1995). Regulation and intracellular localization of *Saccharomyces cerevisiae* strand exchange protein 1 (Sep1/Xrn1/Kem1), a multifunctional exonuclease. Mol. Cell. Biol. *15*, 2728-2736.

Hinterberg, K., Scherf, A., Gysin, J., Toyoshima, T., Aikawa, M., Mazie, J.-C., da Silva, L. P. and Mattei, D. (1994). *Plasmodium falciparum*: the Pf332 antigen is secreted from the parasite by a brefeldin A-dependent pathway and is translocated to the erythrocyte membrane via the Maurer's clefts. Exp. Parasitol. *79*, 279-291.

Hirshhorn, J. N., Brown, S. A., Clark, C. D. and Winston, F. (1992). Evidence that SNF2/SWI2 anf SNF5 activate transcription in yeast by altering chromatin structure. Genes Dev. *6*, 2288-2298.

Holder, A. A., Blackman, M. J., Burghause, P. A., Chappel, J. A., Ling, I. T., McCallum-Deighton and Shai, S. (1992). A malaria merozoite surface protein (MSP1)-structure, surface, processing and function. Mem. Inst. Oswaldo Cruz. Rio de Janeiro *83*, 37-42 (Suppl. III).

Homewood, C. A. and Neame, K. D. (1980). Biochemistry of malarial parasites. In Malaria, J. P. Kreier, ed. (New York: Academic Press), pp. 346-406.

Horrocks, P., Jackson, M., Cheesman, S., White, J. H., and Kilbey, B. J. (1996). Stage specific expression of proliferating cell nuclear antigen and DNA polymerase d from *Plasmodium falciparum*. Mol. Biochem. Parasitol. *79*, 177-182.

Horrocks, P., and Kilbey, B. J. (1996). Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. Mol. Biochem. Parasitol. *82*, 207-215

Hunter, T. (1995). Protein kinases and phosphotases: the yin and yang of protein phosphorylation and signalling. Cell *80*, 225-236.

Hyde, J. (1993). Protocols in molecular parasitology. In Methods in molecular biology., J. M. Walker, ed. (Totowa, New Jersey, USA.: Humana Press.).

Hyde, J. E., Goman, M., Hall, R., Osland, A., Hope, I. A., Langsley, G., Zolg, J. W. and Scaife, J. G. (1984). Characterisation and translation studies of messanger RNA from the human malaria parasite *Plasmodium falciparum* and construction of a cDNA library. Mol. Biochem. Parasitol. *10*, 269-285.

Ifediba, T. and Vandenberg, J. P. (1981). Complete in vitro maturation of *Plasmodium falciparum* gametocytes. Nature *294*, 364-366.

Iltzsch, M. H. e. a. (1992). Cloning and characetrisation of a cDNA coding for the α-subunit of a stimulatory G protein from in*Schistosoma mansoni.* J. Biol. Chem. *267*, 14504-14508.

Inselburg, J. (1983). Stage-specific inhibitory effect of cyclic AMP on asexual maturation and gametocyte formation of *Plasmodium falciparum*. J. Parasitol. *69*, 592-597.

Iovine, M. K. and Wente, S. B. (1997). A nuclear export signal in Kap95p is required for both recycling the import factor and interaction with the nucleoporin GLFG repeat regions of Nup116p and Nup100p. J. Cell Biol. *137*, 797.

Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R. and Kadonaga, J. T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell *90*, 145-155.

Ito, T., Tyler, J. K. and Kadonaga, J. T. (1997). Chromatin assembly factors: a dual function in nucleosome formation and mobilisation? Genes to Cells *2*, 593-600.

Jackson, I. (1991). A reappraisal of non-consensus mRNA splice site. Nucleic Acids Res. *19*, 3795-3798.

Jacobberger, J. W., Horan, P. K. and Hare, J. D. (1992). Cell cycle analysis of asexual stages of erythrocytic malaria parasites. Cell Prolif. *25*, 431-445.

Janknecht, R., Cahill, M. A. and Nordheim, A. (1995). Signal integration at the *c-fos* promoter. Carcinogenesis *16*, 443-450.

Janse, C., Ponnudurai, T., Lensen, A. H. W., Meuwissen, J. H. E. T., Ramesar, J., Van der Ploeg, M. and Overdulve, J. P. (1988). DNA synthesis in gametocytes of *Plasmodium falciparum*.. Parasitology *96*, 1-7.

Janse, C., Van der Klooster, P. F. J., Van der Kaay, H. J., Van der Ploeg, M. and Overdulve, J. P. (1986). DNA synthesis in *P.berghei* during asexual and sexual development. Mol. Biochem. Parasitol. *20*, 173-182.

Janse, C., Van der Klooster, P. F. J., Van der Kaay, H. J., Van der Ploeg, M. and Overdulve, J. P. (1986). Mitomycin C is an unreliable inhibitor of DNA synthesis in *Plasmodium*. Mol. Biochem. Parasitol. *20*, 33-36.

Janse, C. J., Van der Klooster, P. F. J., Van der Kaay, H. J., Van der Ploeg, M. and Overdulve, J. P. (1986). Rapid repeated DNA replication during micrgametogenesis and DNA synthesis in young zygotes of Plasmodium berghei. Trans. R. Soc. Trop. Med. Hyg. *80*, 154-157.

Ji, D. D. and Arnot, D. E. (1997). A *Plasmodium falciparum* homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodeling for transcription. Mol. Biochem. Parasitol. *88*, 151-162.

Ji, D. D., Sultan, A. A. K., Chakrobarty, D., Horrocks, P., Doerig, C. and Arnot, D. E. (1998). *Plasmodium falciparum* nuclear import pathway: an RCC1-type guanidine exchange factor for the Ran G-protein in the parasite nucleus. *In preparation.*

Johnson, L. H. and Holland, J. R. (1965). Ribonucleic acid and protein synthesis in mitotic Hela cells. J. Cell Biol. *27*, 565-574.

Kadonaga, J. T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. Cell *92*, 307-313.

Kaiser, K., Stelzer, G. and Meisterernst, M. (1995). The coactivator p15 (PC4) initiates transcriptional activation during TFIIA-TFIID-promoter complex formation. EMBO J. *14*, 3520-3527.

Kaufman, P. D. (1996). Nucleosome assembly: the ACF and the HAT. Curr. Opin. Cell Biol. *8*, 369-373.

Kaufman, P. D., Kobayashi, R. and Stillman, B. (1997). Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisise* cells lacking chromatin assembly factor-1. Genes Dev. *11*, 345-357.

Kaushal, D. C., Carter, R., Miller, L. H. and Krishna, G. (1980). Gametocytogenesis by malaria parasites in continuous culture. Nature *286*, 490-492.

Kawamoto, F., Fujioka, H., Murakami, R., Syafruddin, Hagiwara, M., Ishikawa, T. and Hidaka, H. (1993). The role of Ca2+/calmodulin and cGMP-dependent pathway in gametogenesis of a rodent malaria parasite, *Plasmodium berghei.* Eur. J. Cell Biol. *60*, 101-107.

Kemp, D. J., Corcoran, L. M., Coppel, R. L., Stahl, H. D., Bianco, A. E., Brown, G. V. and Anders, R. F. (1985). Size variation in chromosomes from independent cultured isolates of *Plasmodium falciparum*. Nature *315*, 347-350.

Kemp, D. J., Thompson, J. K., Walliker, D. and Corcoran, L. M. (1987). Molecular karyotype of *Plasmodium falciparum*: conserved linkage groups and expendable histidine-rich protein genes. Proc. Natl. Acad. Sci. USA *84*, 7672-7676.

Kerr, J. F., Wyllie, A. H. and Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Bri. J. Cancer *26*, 239-257.

Kerr, P. J., Randford-Cartwright, L. C. and Walliker, D. (1994). Proof of intragenic recombination in *Plasmodium falciparum*. Mol. Biochem. Parasitol. *66*, 241-248.

Kingston, R. E., Bunker, C. A. and Imbalzano, A. N. (1996). Repression and activation by multiprotein complexes that alter chromation structure. Genes Dev. *10*, 905-920.

Klebe, C., Bischoff, F. R., Ponstingl, H. and Wittinghofer, A. (1995). Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. Biochem. *34*, 639-647.

Knapp, B., Nau, U., Hundt, E. and Kupper, H. A. (1991). Demonstration of alternative splicing in a pre-mRNA expressed in the blood stage form of *Plasmodium falciparum*. J. Biol. Chem. *266*, 7148-7154.

Koleske, A. J. and Young, R. A. (1995). The RNA polymerase II holoenzyme and its implication for gene regulation. Trends Biochem. Sci. *20*, 113-116.

Kornberg, R. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messanger RNAs. Nucleic Acids Res. *15*, 8125-8148.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: intimation of translational control. J. Cell Biol. *115*, 887-903.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. *12*, 857-872.

Kozak, M. (1996). Interpreting cDNA sequences: some insights from studies on translation. Mammalian genome *7*, 563-574.

Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J *16*, 2482-2492.

Kretzschmar, M., Kaiser, K., Lottspeich, F. and Meisterernst, M. (1994). a novel mediator of classII gene transcription with homology to viral immediate-early transcriptional regulators. Cell *78*, 525-534.

Kruger, W., Peterson, C. L., Sil, A., Coburn, C., Arents, G., Moudrianakis, E. N. and Herskowitz, I. (1995). Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. Genes Dev. *9*, 2770-2779.

Kun, J. F. J. and Anders, R. F. (1995). A *Plasmodium falciparum* gene encoding a high mobility group protein box. Mol. Biochem. Parasitol. *71*, 249-253.

Kwiatkowski, D. (1989). Febrile temperatures can synchronise the growth of *Plasmodium falciparum* in vitro. J. Exp. Med. *169*, 357-361.

Lambros, c. and Vanderberg, J. P. (1979). Synchronisation of *Plasmodium falciparum* erythrocytic stages in culture. J. Parasitol. *65*, 418-420.

Langsley, G., Hyde, J. E., Goman, M. and Scaife, J. G. (1983). Cloning and characterisation of the rRNA genes from the human malaria parasite *Plasmodium falciparum*. Nucleic Acids Res. *11*, 8703-8717.

Lanzer, M., de Bruin, D. and Ravetch, J. V. (1992). Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies a region in which transcription terminates and reinitiates. EMBO J. *11*, 1949-1955.

Lanzer, M., De Bruin, D. and Ravetch, J. V. (1993). Transcriptional differences in polymorphic and conserved domains of a complete cloned *P. falciparum* chromosome. Nature *361*, 654-657.

Lanzer, M., De Bruin, D., Wertheimer, S. P. and Ravetch, J. V. (1994). Organization of chromosomes in *Plasmodium falciparum*: a model for generating karyotypic diversity. Parasitol. Today *10*, 114-117.

Lanzer, M., Fischer, K. and Le Blancq, S. M. (1995). Parasitism and chromosome dynamics in protozoan parasites: is there a connection? Mol. Biochem. Parasitol. *70*, 1-8.

Laurent, B. C., Treich, I. and Carlson, M. (1993). The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. Genes Dev. *7*, 583-591.

Lees, E. (1995). Cyclin dependent kinase regulation. Curr. Opin. Cell Biol. *7*, 773-780.

Leete, T. H. and Rubin, H. (1996). Malaria and cell cycle. Parasitol. Today *12*, 442-444.

Levitt, A. (1993). RNA processing in malarial parasites. Parasitol. Today *9*, 465-468.

Lewin, B. (1994). Chromatin and gene expression: constant questions but changing answer. Cell *79*, 397-406.

Lewis, A. P. (1990). Sequence analysis upstream of the gene encoding the precursor to the major merozoite surface antigens of Plasmodium yoelii. Mol. Biochem. Parasitol. *39*, 285-288.

Li, J. L., Robson, K. J., Chen, J. L., Targett, G. A. and Baker, D. A. (1996). Pfmrk, a MO15-related protein kinase from *Plasmodium falciparum*. Gene cloning, sequencing, stage-specific expression and chromosome localisation. Eur. J. Biochem. *241*, 805-813.

Li, W.-B., Bzik, D. J., Gu, H., Tanaka, M., Fox, B. A. and Inselburg, J. (1989). An enlarged largest subunit of *Plasmodium falciparum* RNA polymerase II defines conserved and variable RNA polymerase domains. Nucleic Acids Res. *17*, 9621-9636.

Li, W.-B., Bzik, D. J., Tanaka, M. T., Gu, H., Fox, B. A. and Inselburg, J. (1991). Characterization of the gene encoding the largest subunit of *Plasmodium falciparum* RNA polymerse III. Mol. Biochem. Parasitol. *46*, 229-240.

Liang, P. and Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science *257*, 967-971.

Lin, D. T., Goldman, N. D. and Syin, C. (1996). Stage-specific expression of a *Plasmodium falciparum* protein related to the eukaryotic mitogen-activated protein kinases. Mol. Biochem. Parasitol. *78*, 67-77.

Lin, F.-T., MacDougald, O. A., Diehl, A. M. and Lane, M. D. (1993). A 30-kDa alternative translation product of the CCAAT/enhancer binding protein α message: translational activator lacking antimitotic activity. Proc. Natl. Acad. Sci. USA *90*, 9609-9610.

Littau, V. C., Allfrey, V. G., Frenster, J. H. and Mirsky, A. E. (1964). Active and inactive regions of nuclear chromatin as revealed by electron microscope autoradiography. Proc. Natl. Acad. Sci. USA *52*, 93-100.

Long, M., de Souza, S. J. and Gilbert, W. (1997). The yeast splice site revisited: new exon consensus from genomic analysis. Cell *91*, 739-740.

Lounsbury, K. M. and Macara, I. G. (1997). Ran-binding protein 1 (RanBP1) forms a ternary complex with Ran and karyopherin beta and reduces Ran GTPase-activating protein (RanGAP) inhibition by karyopherin beta. J. Biol. Chem. *272*, 551-.

Lutcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., Scheele, G. A., Kern, H. F. and Scheele, G. A. (1987). Selection of AUG initiation codons differs in plants and animals. EMBO J. *6*, 43-48.

Ma, H., Yanofsky, M. F. and Meyerawitz, E., M. (1990). Molecular cloning and characterisation of GPA1, a G proteina α subunit gene from *Arabidosis thaliana*. Proc. Natl. Acad. Sci. USA *87*, 3821-3825.

Madeo, F., Frohlich, E. and Frohlich, K. U. (1997). A yeast mutant showing diagnostic markers of early and late apoptosis. J. Cell Biol. *139*, 729-734.

Mahajan, R., Delphin, C., Guan, T., Gerace, L. and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell *88*, 97.

Maldonado, E., Shiekhattar, R., Sheldon, M., Cho, H., Drapkin, R., Richert, P., Lees, E., Anderson, C. W., Linn, S. and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA repair proteins. Nature *381*, 86-89.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. and Evans, R. (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835-839.

Mansour, J. M. and Mansour, T. E. (1986). GTP-binding proteins associated with serotonin-activated adenylate cyclase in *Fasciola hepatica*. Mol. Biochem. Parasitol. *21*, 139-149.

Mansour, J. M. and Mansour, T. E. (1989). Identification of GTP-binding proteins in *Fasciola hepatica* and *Schistosoma mansoni* by immunoblotting. Mol. Biochem. Parasitol. *36*, 11-18.

Martin, S. J., Reutelingsperger, C. P. M., McGahon, A. J., Rader, J. A., van Schie, R. C. A. A., LaFace, D. M. and Green, D. R. (1995). Early redistribution of plasma membrane phosphatidylserine is a general feature of apoptosis regardless of the initiating stimulus: inhibition by overexpression of Bcl-2 and Abl. J. Exp. Med. *182*, 1545-1556.

Martinez-Balbas, M. A., Dey, A., Rabindran, S. K., Ozato, K. and Wu, K. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. Cell *83*, 29-38.

Mastrangelo, A. J. and Betenbaugh, M. J. (1998). Overcoming apoptotsis: new methods for improving protein-expression systems. Trends Biotechnol. *16*, 88-95.

Matsumoto, T. and Beach, D. (1991). Premature initiation of mitosis in yeast lacking RCC1 or an interacting GTPase. Cell *66*, 247-360.

Mattei, D., Hinterberg, K. and Scherf, A. (1992). Pf11-1 and Pf332: two giant proteins synthesized in erythrocytes infected with *Plasmodium falciparum.* Parasitol. Today *8*, 426-428.

Mattei, D. and Scherf, A. (1992). The Pf332 gene of *Plasmodium falciparum* codes for a giant protein that is translocated from the parasite to the membrane of infected erythrocytes. Gene *110*, 71-79.

McAndrew, M. B., Read, M., Sims, P. F. G. and Hyde, J. E. (1993). Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T rich human malaria parasite *Plasmodium falciparum.* Gene *124*, 165-171.

McClelland, M., Mathieu-Daude, F. and Welsh, J. (1995). RNA fingerprinting and differential display using arbitrary primed PCR. Trends Genet. *11*, 242-246.

McCutchan, T. F., Dame, J. B., Miller, L. H. and Barnwell, J. (1984). Evolutionary relatedness of *Plasmodium* species as determined by the structure of the DNA. Science *225*, 808-811.

Mead, D. A., Kristy Pey, N., Herrnstadt, C., Marcil, R. A. and Smith, L. M. (1991). A universal method for the direct cloning of PCR amplified nucleic acid. Biotechniques *9*, 657-663.

Menard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S. and Nussenzweig, V. (1997). Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. Nature *385*, 336-340.

Miyabashira, J., Sekiguchi, T. and Nishimoto, T. (1994). Mammalian cells have two functional RCC1 proteins produced by alternative splicing. J. Cell Sci. *107*, 2203-2208.

Moore, M. S. and Blobel, G. (1993). The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. Nature *365*, 661-663.

Mount, S. M. (1982). A catalogue of splice junction sequences. Nucleic Acids Res. *10*, 459-472.

Muchardt, C., Reyes, J. C., Bourachot, B., Legouy, E. and Yaniv, M. (1996). The hbrm and BRG-1 proteins, components of the human SNF/SWI complex, are phosphorylated and excluded from the condensed chromosomes during mitsis. EMBO J. *15*, 3394-3402.

Murray, A. and Hunt, T. (1993). The cell cycle, an introduction. (New York: W.H. Freeman).

Musto, H., Rodriguez-Maseda, H. and Bernardi, G. (1995). Compositional properties of nuclear genes from *Plasmodium falciparum*. Gene *152*, 127-132.

Myler, P., Saul, A. and Kidson, C. (1983). The synthesis and fate of stage-specific proteins in *Plasmodium falciparum* cultures. Mol. Biochem. Parasitol. *9*, 37-45.

Neer, E. J. and Smith, T. F. (1996). G protein heterodimers: new structures propel new questions. Cell *84*, 175-178.

Neigeborn, L. and Carlson, M. (1984). Genes affecting the regulation of *SUC2* gene expression by glucose repression in *Saccharomyces cerevisiae*. Genetics *108*, 845-858.

Newbold, C. I., Boyle, D. B., Smith, C. C. and K.N., B. (1982). Stage specific protein and nucleic acid synthesis during the asexual cycle of the rodent malaria *Plasmodium chabaudi*. Mol. Biochem. Parasitol. *5*, 33-39.

Newman, A. J. and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5' and 3' splice sites. Cell *68*, 743-754.

Nishitani, H., Kobayoshi, H., Ohtsubo, M. and Nishimoto, T. (1990). Cloning of *Xenopus* RCC1 cDNA , a homologue of the human RCC1: complementation of tsBN2 mutation and identification of the product. J. Biochem. *107*, 228-235.

Nishitani, H., Ohtsubo, M., Yamashita, K., Iida, H., Pines, J., Yasudo, H., Shibata, Y., Hunter, T. and Nishimoto, T. (1991). Loss of RCC1, a nuclear DNA-binding protein, uncouples the completion of DNA replication from activation of cdc2 protein kinase and mitosis. EMBO J. *10*, 1555-1564.

O'Dea, K. P., McKean, P. G., Harris, A. and Brown, K. N. (1995). Processing of the *Plasmodium chabaudi chabaudi* AS merozoite surface protein 1 in vivo and in vitro. Mol. Biochem. Parasitol. *72*, 111-119.

Oberhammer, F., Wilson, J. W., Dive, I. D., Hickman, J. A., Wakeling, A. E., Walker, P. R. and Sikorska, M. (1993). Apoptotic death in epithelial cells: cleavage of DNA to 300 and/or 50 kb fragments prior to or in the absence of internucleosomal fragmentation. EMBO J. *12*, 3679-3684.

Ohtsubo, M., Kai, R., Furuno, N., Sekiguchi, T., Sekiguchi, M., Hayashida, H., Kuma, K., Miyata, T., Fukushige, S. and Murotsu, T. (1987). Isolation and characterization of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation. Genes Dev. *1*, 585-593.

Ohtsubo, M., Okazaki, H. and Nishimoto, T. (1989). The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. J. Cell Biol. *109*, 1389-1397.

Okabe, I., Bailey, L. C., Attree, O., Srinivasan, S., Perkel, J. M., Laurent, B. C., Carlson, M., Nelson, D. L. and Nussbaum, R. L. (1992). Cloning of human and bovine homologs of SNF2/SWI2: a global activator of transcription in yeast S. cerevisiae. Nucleic Acids Res. *20*, 4649-4655.

Oquendo, P., Goman, M., MacKay, M., Langsley, G., Walliker, D. and Scaife, J. G. (1986). Characterisation of a repetitive DNA sequence from the malaria parasite, *Plasmodium falciparum*. Mol. Biochem. Parasitol. *18*, 89-101.

Oz, H. S. e. a. (1994). Evidence for guanosine triphosphate-binding proteins in *Trypanosoma cruzi*. Am. J. Trop. Med. Hyg. *50*, 620-631.

Pan, G., Aso, T. and Greenblatt, J. (1997). Interaction of elongation factors TFIIS and elongin A with a human RNA polymerase II holoenzyme capable of promoter-specific initiation and responsive to transcriptional activators. J. Biol. Chem. *272*, 24563-24571.

Paranjape, S. M., Kamakaka, R. T. and Kadonaga, J. T. (1994). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Ann. Rev. Biochem. *63*, 265-297.

Pazin, M. J. and Kadonaga, J. T. (1997). SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein-DNA interactions? Cell *88*, 737-740.

Pazin, M. J. and Kadonaga, J. T. (1997). What's up and down with histone deacelation and transcription? Cell *89*, 325-328.

Pazin, M. J., Kamakaka, R. T. and Kadonaga, J. T. (1994). ATP-dependent nucleosome reconstitution and transcriptional activation from preassembled chromatin templates. Science *266*, 2007-2011.

Peterson, C. L. (1996). Multiple SWItches to turn on chromatin? Current. Opin. Genetics Dev. *6*, 171-175.

Peterson, C. L., Dingwall, A. and Scott, M. P. (1994). Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. Proc. Natl. Acad. Sci. USA *91*, 2905-2908.

Peterson, C. L. and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell *68*, 573-583.

Piqueras, B., Autran, B., Debre, P. and Gorochov, G. (1996). Detection of apoptosis at the single-cell level by direct incorporation of fluorescein-dUTP in DNA strand break. Biotechniques *20*, 634-640.

Pollack, Y., Katzen, A. L., Spira, D. T. and Golenzer, J. (1982). The genome of *Plasmodium falciparum*. I: DNA base composition. Nucleic Acids Res. *10*, 539-546.

Pollard, K. J. and Peterson, C. L. (1997). Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression. Mol. Cell. Biol. *17*, 6212-6222.

Prasartkaew, S., Zijlstra, N. M., Wilairat, P., Overdulve, J. P. and de Vries, E. (1996). Molecular cloning of a *Plasmodium falciparum* gene interrupted by 15 introns encoding a functional primase 53 kDa subunit as demonstrated by expression in a baculovirus system. Nucleic Acids Res. *24*, 3934-3941.

Prensier, G. and Slomianny, C. (1986). The karyotype of *Plasmodium falciparum* determined by ultrastructural serial sectioning and 3D reconstruction. J. Parasitol. *72*, 731-736.

Prince, J. B., Koven-Quinn, M. A., Remington, J. S. and Sharma, S. D. (1985). Cell free synthesis of *Toxoplasma gondii* antigens. Mol. Biochem. Parasitol. *17*, 163-170.

Quinn, J. and Peterson, C. L. (1997). Mechanisms for alleviating chromatin-mediated repression of transcription. In Chromatin structure, and gene expression., R. Van Driel and P. A. Otte, eds. (London: Oxford University Press), pp. 58-79.

Rao, A. S., Green, T. J. and Guntaka, R. V. (1986). Organisation of Plasmodium falciparum genome: II. Sequence analysis of *falci.* element. Nucleic Acids Res. *14*, 9542-9548.

Rawlings, D. J. and Kaslow, D. C. (1992). Anovel 40 kDa membrane-associated EF-hand calcium-binding protein in *Plasmodium falciparum.* J. Biol. Chem. *267*, 3976-3982.

Read, L. K. and Mikkelsen, R. B. (1990). Cyclic AMP- and Ca2+-dependent protein kinases in *Plasmodium falciparum.* Exp. Parasitol. *71*, 39-48.

Read, L. K. and Mikkelsen, R. B. (1990). *Plasmodium falciparum* infected erythrocytes contain an adenylate cyclase with properties which differ from those of the host enzyme. Mol. Biochem. Parasitol. *45*, 109-120.

Read, M., Sherwin, T., Holloway, S. P., Gull, K. and Hyde, J. E. (1993). Microtubular organisation visualised by immunofluorescence microscopy during erythrocytic schizogony in *Plasmodium falciparum* and investigation of post-transcriptional modification of parasite tubulin. Parasitology *106*, 223-232.

Rechsteiner, M. (1991). Natural substrates of the ubiquitin proteolytic pathway. Cell *66*, 615-618.

Ren, M., Drivas, G., D'Eustachio, P. and Rush, M. G. (1993). Ran/TC4: a small nuclear GTP-binding protein that regulates DNA synthesis. J. Cell Biol. *120*, 313-323.

Renart, J., Reiser, J. and Stark, G. R. (1979). Transfer proteins from gels to diazobenyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. Proc. Natl. Acad. Sci. USA *76*, 3116-3120.

Renault, L., Nassar, N., Vetter, I., Becker, J., Klebe, C., Roth, M. and Wittinghofer, A. (1998). The 1.7 A crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. Nature *392*, 97-101.

Richards, S. A., Carey, K. L. and Macara, I. G. (1997). Requirement of guanine triphosphate-bound Ran for signal-mediated nuclear protein export. Science *276*, 1842-
.

Richards, S. A., Lounsbury, K. M. and Macara, I. G. (1995). The C terminus of the nuclear Ran/TC4 GTPase stabilize the GDP-bound state and mediates interactions with RCC1, Ran-GAP and HTF9A/RanBP1. J. Biol. Chem. *270*, 14405-14411.

Rihs, H.-P., Jans, D., Fan, H. and Peters, R. (1990). The rate of nuclear cytoplasmic protein transport is determined by the casein kinase II site flanking the nuclear localisation sequence of the SV40 T-antigen. EMBO J. *10*, 633-639.

Roberts, S. M. and Winston, F. (1997). Essential functional interactions of SAGA, *Sccharomyces cerevisiae* complex of Spt, Ada, and Gcn5 proteins, with the Snf/Swi and Srb/mediatorcomplexes. Genetics *147*, 451-465.

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

Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. *21*, 327-335.

Rogers, M. J., McConkey, G. A., Li, J. and McCutchan, T. F. (1995). The ribosomal DNA loci in *Plasmodium falciparum* accumulate mutations independently. J. Mol. Biol. *254*, 881-891.

Romanowski, P. and Madine, M. A. (1996). Mechanisms restricting DNA replication to once per cell: MCM, pre-replicative complexes and kinases. TCB, 184-188.

Rosa, J. L., Casaroli-Marano, R. P., Buckler, A. J., Vilaro, S. and Barbacid, M. (1996). p619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and Rab proteins. EMBO J. *15*, 4262-4273.

Ross-MacDonald, P. B., Graeser, R., Kappes, B., Franklin, R. and Williamson, D. H. (1994). Isolation and expression of a gene specifying a *cdc2*-like protein kinase from the human malaria parasite *Plasmodium falciparum.* Europ. J. Biochem. *220*, 693-701.

Roth, S. Y. and Allis, C. D. (1996). Histone acetylation and chromatin assembly: a single escort, multiple dances? Cell *87*, 5-8.

Ruvolo, V., Altszuler, R. and Levitt, A. (1993). The transcript encoding the circumsporozoite antigen of *Plasmodium berghei* utilises heterogenous polyadenylation sites. Mol. Biochem. Parasitol. *57*, 137-150.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Elrich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science *239*, 487-491.

Saitoh, R., Pu, R., Cavenagh, M. and Dasso, M. (1997). RanBP2 associates with Ubc9p and a modified form of RanGAP1. Proc. Natl. Acad. Sci. USA *94*, 3736-.

Sambrook, J., Fritch, E. F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (New York: Cold Spring Harbor Laboratory Press).

Sauer, F., Hansen, S. and Tjian, R. (1995). DNA template and activator-coactivator requirements for transcriptional synergistic by Drosophila Bicoid. Science *270*, 1783-1788.

Sauer, F., Hansen, S. and Tjian, R. (1995). Multiple TAFIIs directing synergistic activation of transcription. Science *270*, 1783-1788.

Sazer, S. and Nurse, P. (1994). A fission yeast RCC1-related protein is required for mitosis to interphase transition. EMBO J. *13*, 606-615.

Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W. and Wittinghofer, A. (1995). Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. Nature *374*, 378-381.

Schein, C. H. and Noteborn, H. M. (1988). Low temperature induction of Glutathione-S-transferase fusion proteins. Bio/Technology *6*, 291-294.

Schein, C. H. and Noteborn, M. H. M. (1988). Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperature. Bio/Technology *6*, 291-294.

Scherf, A. (1996). Plasmodium telomeres and telomere proximal gene expression. CDB *7*, 49-57.

Schwartz, D. C. and Cantor, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell *37*, 67-75.

Seeber, F. (1997). Consensus sequence of translational initiation sites from *Toxoplasma gondii* genes. Parasitol. Res. *83*, 309-311.

Senapathy, P., Shapiro, M. B. and Harris, N. L. (1990). Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. *183*, 252-278.

Sharp, P. A. (1987). Splicing of messenger RNA precursors. Science *235*, 766-771.

Shobuike, T., Sugano, S., Yamashita, T. and Ikeda, H. (1997). Cloning and characterization of mouse *Dhm2* cDNA, a
functional homolog of budding yeast *SEP1*. Gene *191*, 161-166.

Sibley, L. D. and Boothroyd, J. C. (1992). Construction of a molecular karyotype for *Toxoplasma gondii*. Mol. Biochem. Parasitol. *51*, 291-300.

Simmons, D., Woollett, G., Bergin-Cartwright, M., Kay, D. and Scaife, J. (1987). A maralia protein exported into a new compartment within the host erythrocyte. EMBO J. *6*, 485-491.

Simon, M. I. X., Strathmann, M. P. and Gautam, N. (1991). Diversity of G protein in signal transduction. Science *252*, 802-808.

Sinclair, D. A. and Guarente, L. (1997). Extachromosomal rDNA circles-a cause of aging in yeast. Cell *91*, 1033-1042.

Sinclair, D. A., Mills, K. and Guarente, L. (1997). Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants. Science *277*, 1313-1316.

Sinden, R. E. (1983). The cell biology of sexual development in *Plasmodium*.. Parasitology *86*, 7-28.

Sinden, R. E. (1982). Gametocytogenesis in *Plasmodium falciparum* in vitro: an eletron microscopic study. Parasitol. *84*, 1-11.

Sinden, R. E. (1982). Gametocytogenesis of *Plasmodium falciparum* in vitro: an eletron microscopic study. Parasitology *84*, 1-11.

Sinha, S. N., Hellwig, R. J., Allison, D. P. and Niyogi, S. K. (1982). Conversion of simian virus 40 DNA to ordered nucleoprotein structures by extracts that direct accurate initiation by eukaryotic RNA polymerase II. Nucleic Acids Res. *10*, 5533-5552.

Sinnis, P. and Wellems, T. E. (1988). Long-range restriction maps of *Plasmodium falciparum*: crossing over and size variation among geographically distant isolates. Genomics *3*, 287-295.

Smith, C. W., Chu, T. T. and Nadal-Ginard, B. (1993). Scanning and competition between AGs are involved in 3' splice site in mammalian introns. Mol. Cell. Biol. *13*, 4939-4952.

Solomon, M. J., Glotzer, M., Lee, T. H., Philippe, M. and Kirschner, M. W. (1990). Cyclin activation of p34Cdc2. Cell *68*, 1013-1024.

Solovyev, V. V., Salamov, A. A. and Lawrence, C. B. (1994). Predicting internal exons by oligonucleotide composition and discriminant analysis of splicable open reading frames. Nucleic Acids Res. *22*, 5156-5613.

Steger, D. J. and Workman, J. L. (1996). Remodeling chromatin structures for transcription: what happens to the histones? BioEssays *18*, 875-884.

Struhl, K. (1996). Chromatin structure and RNA polymerase II connection: implications for transcription. Cell *84*, 179-182.

Studier, F. W. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned gene. J. Mol. Biol. *189*, 113.

Sukhodolets, M. V. and Jin, D. J. (1998). RapA, a novel RNA polymerase-associated protein, is a bacterial homolog of SWI2/SNF2. J. Biol. Chem. *273*, 7018-7023.

Sultan, A. A., Richardson, W., Alano, P., Arnot, D. E. and Doerig, C. (1994). Cloning and characterisation of a *Plasmodium falciparum Ras*-like GTPase involved in signal transduction for cell cycle control. Mol. Biochem. Parasitol. *65*, 531-538.

Sultan, A. A. A. (1994). Molecular studies on signal transduction in Plasmodium falciparum (Edinburgh: University of Edinburgh).

Sutani, T. and Yanagida, M. (1997). DNA renaturation activity of the SMC complex implicated in chromosome condensation. Nature *388*, 798-801.

Tang, W. J. and Gilman, A. G. (1991). Type-specific regulation of adenylate cyclase by G protein $\beta\gamma$ subunits. Science *254*, 1500-1503.

Tao, W., Kurschner, C. and Morgan, J. I. (1997). Modulation of cell death in yeast by the Bcl-2 family of proteins. J. Biol. Chem. *272*, 15547-15552.

Taylor, D. W., Parra, M., Chapman, G. B., Stearns, M. E., Rener, J., Aikawa, M., Uni, S., Aley, S. B., Panton, L. J. and Howard, R. J. (1987). Localisation of *Plasmodium falciparum* histidine-rich protein 1 in the erythrocyte skeleton under knobs. Mol. Biochem. Parasitol. *25*, 165-174.

Taylor, S. S. and McKeon, F. (1997). Kinetochore localisation of murine Bub1 is required for normal mitotic timing and checkpoint response to spindle damage. Cell *89*, 127-135.

Thelu, J., Bracchi, V., Burnod, J. and Ambroise-Thomas, P. (1994). Evidence for expression of a Ras-like and a stage-specific GTP binding homologous protein by *Plasmodium falciparum*. Cell. Signal. *6*, 777-782.

Tjian, R. and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. Cell *77*, 5-8.

Torres-Guerrero, H., Peattie, D. A. and Meza, I. (1991). Chromatin organisation in *Entamoeba histolytica*. Mol. Biochem. Parasitol. *45*, 121-130.

Towbin, H., Staehlin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some application. Proc. Natl. Acad. Sci. USA *76*, 4350-4354.

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

Treich, I. and Carlson, M. (1997). Interaction of a Swi3 homolog with Sth1 provides evidence for a Swi/Snf-related complex with an essential function in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *17*, 1768-1775.

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

Troelstra, C., Va Gool, A., De Wit, J., Vermeulen, W., Bootsma, D. and Hoeijmakers, J. H. J. (1992). ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell *71*, 939-953.

Trouche, D., Chalony, C. L., Muchardt, C. and Yaniv, M. (1997). RB and hbrm cooperate to repress the activation functions of E2F1. Proc. Natl. Acad. Sci. USA *94*, 11268-11273.

Tsuchiya, E., Uno, M., Kiguchi, A., Masuoka, K., Kanemori, Y., Okabe, S. and Mikayawa, T. (1992). The *Saccharomyces cerevisiae NPS1* gene, a novel CDC gene which encodes a 160 kDa nuclear protein involved in G_2 phase control. EMBO J. *11*, 4017-4026.

Tsukiyama, T., Becker, P. and Wu, C. (1994). ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor. Nature *367*, 525-532.

Tsukiyama, T., Daniel, C., Tamkun, J. and Wu, C. (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell *83*, 1021-1026.

Tsukiyama, T. and Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell *83*, 1011-1020.

Turner, B. M., Birley, A. J. and Lavender, J. (1992). Histine H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. Cell *69*, 375-384.

Vaidya, A. B., Akella, R. and Suplick, K. (1989). Sequences similar to genes for two mitochondrial proteins and proteins of ribosomal RNA in tandemly arrayed 6-kilobase-pair DNA of a malarial parasite. Mol. Biochem. Parasitol. *35*, 97-107.

Vaidya, A. B., Schleif, W. A., Majarian, W. R., Daly, T., M., Taylor, D. W. and Long, C. A. (1984). Analysis of mRNA coding for blood-stage antigens of a rodent malarial parasite, Plasmodium yoelii: mRNA coding for a possible protective antigen purify as poly A. J. Immunol. *132*, 3126-3130.

Valencia, A., Kjeldgaard, M., Pai, E. F. and Sander, C. (1991). GTPase domains of *ras* p21 oncogene protein and elongation factor Tu: analysis of three-dimentional structures, sequence families, and functional sites. Proc. Natl. Acad. Sci. USA *88*, 5443-5447.

Van der Ploeg, L. H. T., Smith, M., Ponnudurai, T., Vermeulen, A., Meuwissen, J. H. E. T. and Langsley, G. (1985). Chromosome-sized DNA molecules of *Plasmodium falciparum.* Science *229*, 658-661.

Vanhamme, L. and Pays, E. (1995). Control of gene expression in Trypanosomes. Microbiol. Rev. *59*, 223-240.

Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M. and Becker, P. B. (1997). Chromatin-remodeling factor CHRAC contains the ATPases ISWI and topoisomerase II. Nature *388*, 598-602.

Vernick, K. D. and McCutchan, T. F. (1988). Sequence and structure of a Plasmodium falciparum telomere. Mol. Biochem. Parasitol. *28*, 85-94.

Vettese-Dadely, M., Grant, P. A., Hebbes, T. R., Crane-Robinson, C., Allis, C. D. and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA *in vitro*. EMBO J. *15*, 2508-2518.

Vinkenoog, R., Veldhuisen, B., Speranca, M. A., del Portillo, H. A., Janse, C. and Waters, A. P. (1995). Comparison of introns in a cdc2-homologous gene within a number of *Plasmodium* species. Mol. Biochem. Parasitol. *71*, 233-241.

Walker-Jonah, A., Dolan, S. A., Gwadz, R. W., Panton, L. J. and T.E., W. (1992). An RFLP map of the *Plasmodium falciparum* genome, recombination rates and favored linkage groups in a genetic cross. Mol. Biochem. Parasitol. *51*, 313-320.

Walliker, D. (1989). Genetic recombination in malaria parasites. Exp. Parasitol. *69*, 303-309.

Walliker, D., Carter, R. and Morgan, S. (1973). Genetic recombination in *Plasmodium berghei.* Parasitology *66*, 309-320.

Walliker, D., Carter, R. and Sanderson, A. (1975). Genetic studies on *Plasmodium chabaudi*: recombination between enzyme markers. Parasitology *70*, 19-24.

Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, M. T., Corcoran, L. M., Burkot, T. R. and Carter, R. (1987). Genetic analysis of the human malaria parasite *Plasmodium falciparum.* Science *236*, 1661-1666.

Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L. and Crabtree, G. R. (1996). Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. EMBO J. *15*, 5370-5382.

Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B. R. and Crabtree, G. R. (1996). Diversity and specialization of mammalian SWI/SNF complexes. Genes Dev. *10*, 2117-2130.

Watson, S. and Arkinstall, S. (1994). The G-protein linked receptor, facts book (Oxford: Academic Press.).

Weber, J. L. (1987). Analysis of sequences from the extremely A+T rich genome of *Plasmodium falciparum*. Gene *52*, 103-109.

Weber, J. L. (1988). Molecular biology of malaria parasites. Exp. Parasitol. *66*, 143-170.

Wellems, T. E., Walliker, D., Smith, C. L., DoRosario, V. E., Maloy, W. L., Howard, R. J., Carter, R. and MacCutchan, T. F. (1987). A histidine-rich protein gene marks a linkage group favoured strongly in a genetic cross of *Plasmodium falciparum*. Cell *49*, 633-642.

White, J. H. and Kilbey, B. J. (1996). DNA replication in the malaria parasite. Parasitol. Today *12*, 151-155.

WHO (1992). Tropical Disease Research . 11th programme report. World Health Organisation, Geneva, Swizerland. .

Wiesner, J., Mattei, D., Scherf, A. and Lanzer, M. (1998). Biology of giant proteins of *Plasmodium* : resolution on polyacrylamide-agarose composite gels. Parasitol. Today *14*, 38-40.

Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E. and Young, R. A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell *84*, 235-244.

Wilson, R. J. and Williamson, D. H. (1997). Extrachromosomal DNA in the apicomplexa. Microbiology and Molecular Biology Reviews *61*, 1-16.

Wilson, R. J. M., Fry, M., Gardner, M. J., Feagin, J. E. and Williamson, D. H. (1992). Subcellular fractionation of the two orgnelle DNA of malaria parasites. Curr. Genet. *21*, 405-408.

Wilson, R. J. M., Gardner, M. J., Feagan, J. E. and Williamson, D. H. (1991). Have malaria parasites three genomes? Parasitol. Today *7*, 134-146.

Winship, R. R. (1989). An improved method for directly sequencing PCR amplified material using dimethylsulphoxide. Nucleic Acids Res. *17*, 1266-1270.

Wittenberg, C. and Reed, S. I. (1996). PLugging it in: signaling circuits and the yeast cell cycle. Curr. Opin. Cell Biol. *8*, 223-230.

Wittinghofer, A. and Pai, E. F. (1991). The structure of *Ras* protein: a model for a universal molecular switch. Trends Biochem. Sci. *16*, 382-387.

Wunderlich, F. (1980). Nucleosomal organisation of chromatin in *Plasmodium Knowlesi*. J. Parasitol. *66*, 1063-1065.

Wyllie, A. H., Kerr, J. F. R. and Currie, A. R. (1980). Cell death: the significance of apoptosis. Int. Rev. Cytol. *68*, 251-306.

Yamauchi, K. (1991). The sequence flanking translational initiation site in protozoa. Nucleic Acids Res. *19*, 2715-2720.

Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T. and Ishii, S. (1995). Increase in solubility of foreign proteins in *Escherichia coli* by co-production of the bacterial thioredoxin. J. Biol. Chem. *270*, 25328-25331.

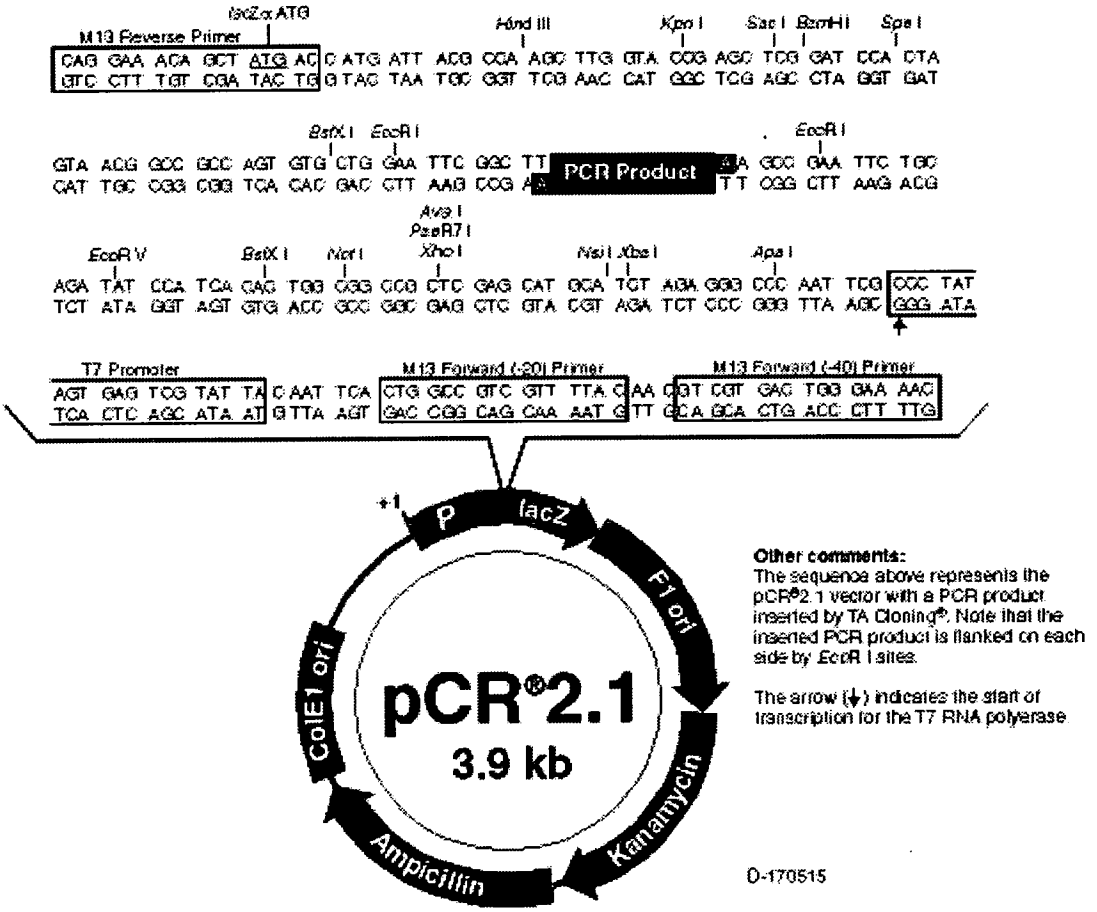
Zhao, Y., Kappes, B. and Franklin, R. M. (1993). Gene structure and expression of an unusual protein kinase from *Plasmodium falciparum* homologous at its carboxyl terminus with the EF hand calcium-binding proteins. J. Biol. Chem. *268*, 4347-4354.

Zolg, J. W., McLeod, A. J., Dickson, I. H. and Scaife, J. G. (1982). *Plasmodium falciparum*: modifications on the *in vitro* culture conditions improving parasite yields. J. Parasitol. *68*, 1072-1080.

Zolotukhin, A. S. and Felber, B. K. (1997). Mutations in the nuclear export signal of human ran-binding protein RanBP1 block the Rev-mediated posttranscriptional regulation of human immunodeficiency virus type 1. J. Biol. Chem. *272*, 11356-11360.

Appendix A

Schematic diagram of pCR2.1 vector.



LacZ alpha fragment: bases 1-587
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-355
T7 promoter/priming site: bases 362-381
M13 Forward (-20) priming site: bases 389-404
M13 Forward (-40) priming site: bases 408-424
f1 origin: bases 546-960
Kanamycin resistance ORF: bases 1294-2088
Ampicillin resistance ORF: bases 2106-2966
ColE1 origin: bases 3111-3784

Adapted from Invitrogen catalogue.

Appendix B

Sequences similarities of clones isolated from PCR fragments. The sequence smilarities of the clones isolated from PCR products during the course of this study which have a significant and/or interesting similarity to known proteins are detailed here. DNA sequences were searched with BLASTX in six reading frames through the NCBI net service (Benson *et al*, 1998).

Clone GP1-2

```
  1 GTGGATTCAG GTTTTCCAAG AATTTGTGAA GTAGATGCAG TTGGCATAGG TGCTAATAAT AATGAGTTTC
 71 TTAAACCATG TTTACGGATT TTAGCTTTTA ATTCATCCCA ATCCCAATAT TTGTTATCAA CTTTAGCATT
141 CCACATATCA AATTGTAATA TACCTTGGCT AGCTGGACTT CCTTAACAAG ATCGA
```

>> (U01323,P50648), RIR1_PLAF4 RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (RIBONUCLEOTIDE REDUCTASE), (R1 SUBUNIT), [*Plasmodium falciparum*], Length = 806

```
 Score = 307 (140.6 bits), Expect = 1.6e-34, P = 1.6e-34
 Identities = 58/60 (96%), Positives = 58/60 (96%), Frame = -2
Query:   182 GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILGKPES 3
             GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILG  ES
Sbjct:   557 GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILGNNES 616
```

>> A49412, Ribonucleoside-diphosphate reductase (EC 1.17.4.1) large chain, [*Plasmodium falciparum*], Length = 847

```
 Score = 307 (140.6 bits), Expect = 1.6e-34, P = 1.6e-34
 Identities = 58/60 (96%), Positives = 58/60 (96%), Frame = -2
Query:   182 GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILGKPES 3
             GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILG  ES
Sbjct:   598 GSPASQGILQFDMWNAKVDNKYWDWDELKAKIRKHGLRNSLLLAPMPTASTSQILGNNES 657
```

Clone GP7-1

```
  1 CGATTTTGTT TAAGAATTGA TGGATCAACA CCAGGAGATG AAAGACAAGT ACGTATTAAT CAATTTAATG
 71 AACCGAATAG TAAATATTTT ATTTTCTTAT TATCTACCAG AGCTGGTGGT ATTGGTATTA ACTTAACGAC
141 TGCTGATATT GTTATCTTAT TTGATTCGGA TTATAATCCA CAAATGGATA TACAAGCAAT GGATCCACTT
```

>> AB010882, hSNF2h, [*Homo sapiens*], Length = 1052

```
 Score = 254 (116.4 bits), Expect = 4.2e-27, P = 4.2e-27
 Identities = 46/64 (71%), Positives = 56/64 (87%), Frame = +1
Query:    13 RIDGSTPGDERQVRINQFNEPNSKYFIFLLSTRAGGIGINLTTADIVILFDSDYNPQMDIQAMD 204
             R+DG TP DERQ  IN +NEPNS  F+F+LSTRAGG+GINL TAD+VIL+DSD+NPQ+D+QAMD
Sbjct:   528 RLDGQTPHDERQDSINAYNEPNSTKFVFMLSTRAGGLGINLATADVVILYDSDWNPQVDLQAMD 591
```

>> U64598, similar to helicases of the SNF2/RAD54 family. Does not contain the bromodomain. [*Caenorhabditis elegans*], Length = 1257

```
 Score = 233 (106.7 bits), Expect = 3.3e-24, P = 3.3e-24
 Identities = 42/65 (64%), Positives = 54/65 (83%), Frame = +1
Query:    10 LRIDGSTPGDERQVRINQFNEPNSKYFIFLLSTRAGGIGINLTTADIVILFDSDYNPQMDIQAMD 204
             LR+DGST  DER   +++FN PNS+YF+F+LSTRAGG+G+NL TAD VI+FDSD+NP  D+QA D
Sbjct:   648 LRLDGSTKPDERGALLDKFNAPNSEYFLFMLSTRAGGLGLNLQTADTVIIFDSDWNPHQDMQAQD 712
```

Clone GP7-3

```
  1 TCGATTTTGT TTAAGAATAA TAATAATAAT GATATTAATG GTAATAACAA CAATAATAAT AACCACAACA
 71 ATAATAATAA CCACAACAAT AATAATAATA ATAATAATAA TAATGGTTGC TATAAAATAA TTAACGAGGG
141 TACTTTAAAC ATAAATGATG TTATGAAATA TGATGGTTTA GCATCTACAT GTTTTCCAAA TATAAACAAT
```

211 GAATCCACT

SILFKNNNNN DINGNNNNNN NHNNNNNHNN NNNNNNNNGC YKIINEGTLN INDVMKYDGL ASTCFPNINN EST

Clone GPA11-3

```
  1  AAGTGGATAC ATTGCTTCAA CGATGTGACT GCCATCATCT TCGTGGTAGC CAGCAGCAGC TACAACATGG
 71  TCATCCGGGA GGACAACCAG ACCAACCGCC TGCAGGAGGC TCTGAACCTC TTCAAGAGCA TCTGGAACAA
141  CAGATGGCTG CGCACCATCT CTGTGATCCT GTTCCTCAAT AAGATCGAC
```

>> S34421 human GTP-binding regulatory protein Gs α chain; (X56009) α subunit of Gs- GTP binding protein [Homo sapiens], Length = 419

```
Score = 315 (144.3 bits), Expect = 1.1e-35, P = 1.1e-35
 Identities = 61/63 (96%), Positives = 61/63 (96%), Frame = +1
Query:     1 KWIHCFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKID 189
             KWI CFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNK D
Sbjct:   258 KWIQCFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKQD 320
```

>> pir||JH0813 GTP-binding regulatory protein Gs α chain isoform - human (clone pSX4), Length = 335

```
Score = 315 (144.3 bits), Expect = 1.1e-35, P = 1.1e-35
 Identities = 61/63 (96%), Positives = 61/63 (96%), Frame = +1
Query:     1 KWIHCFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKID 189
             KWI CFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNK D
Sbjct:   174 KWIQCFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKQD 236
```

Clone Gpa11-5

```
  1  AAGTGGATTC ACTGCTTTGA GGGAGTGACA GCAATTATCT TCTGTGTGGC CCTCAGTGAT TATGACCTTG
 71  TTCTGGCTGA GGACGAGGAG ATGAACCGAA TGCATGAAAG CATGAAACTG TTTGACAGCA TTTGTAATAA
141  CAAATGGTTT ACAGAAACTT CAATCATTCT CTTCCTTAAC AAGATCGAC
```

>> P08754, GBAK_HUMAN G PROTEIN G(K), α SUBUNIT (G(I) α-3) pir|| RGHUI3 GTP-binding regulatory protein Gi α3 chain (adenylate cyclase -inhibiting); human (J03005) guanine nucleotide-binding regulatory protein α-inhibitory subunit [*Homo sapiens*]; (M20604) Gi3 protein α subunit; (J03198) G protein α subunit; (M27543) G protein (G) α subunit; (J03238) G protein α subunit, Length = 354

```
Score = 332 (152.1 bits), Expect = 5.0e-38, P = 5.0e-38
 Identities = 62/63 (98%), Positives = 62/63 (98%), Frame = +1

Query:     1 KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNKID 189
             KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNK D
Sbjct:   210 KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNKKD 272
```

>> M17219, G protein αi subunit, [*Homo sapiens*], Length = 349

```
Score = 329 (150.7 bits), Expect = 1.3e-37, P = 1.3e-37
 Identities = 61/63 (96%), Positives = 62/63 (98%), Frame = +1

Query:     1 KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNKID 189
             KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFT+TSIILFLNK D
Sbjct:   205 KWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKD 267
```

Clone Gpa14-2

```
  1  AAATGGATAC ACTGCTTTGA AAATGTCACC TCTATCATGT TTCTAGTAGC GCTTAGTGAA TATGATCGGG
 71  TTCTCGTGGA GTCAGACAAT GAGAACCGAA TGGAGGAAAG CAAGGCTCTC TTTAGTACAA TTATCACGTA
141  TCCCTGGTTC CAGAACTCCT CGGTTATTCT GTTCTTAAAC AAAATCGAC
```

>> AF011497, G protein α11 subunit, [*Homo sapiens*], Length = 359

Score = 311 (142.5 bits), Expect = 3.9e-35, P = 3.9e-35
Identities = 60/63 (95%), Positives = 61/63 (96%), Frame = +1

```
Query:     1 KWIHCFENVTSIMFLVALSEYDRVLVESDNENRMEESKALFSTIITYPWFQNSSVILFLNKID 189
             KWIHCFENVTSIMFLVALSEYD+VLVESDNENRMEESKALF TIITYPWFQNSSVILFLNK D
Sbjct:   215 KWIHCFENVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKD 277
```

>> AF011496, G protein αq [*Homo sapiens*], Length = 359

Score = 311 (142.5 bits), Expect = 3.9e-35, P = 3.9e-35
Identities = 60/63 (95%), Positives = 61/63 (96%), Frame = +1

```
Query:     1 KWIHCFENVTSIMFLVALSEYDRVLVESDNENRMEESKALFSTIITYPWFQNSSVILFLNKID 189
             KWIHCFENVTSIMFLVALSEYD+VLVESDNENRMEESKALF TIITYPWFQNSSVILFLNK D
Sbjct:   215 KWIHCFENVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKD 277
```

Clone Pf3-1

>> M58367 strand exchange protein 1 [Saccharomyces cerevisiae], (M36725) DNA strandtransferase β, (M90097) 5'-3' exoribonuclease , (X84705) DNA strand transferase, exoribonuclease, (Z72695) ORF YGL173c, Length = 1528

Score = 126 (57.7 bits), Expect = 6.4e-10, P = 6.4e-10
Identities = 25/44 (56%), Positives = 30/44 (68%), Frame = +2

```
Query:     5 DMAGEGEHKILKFIRENCKNDSNFKNWNHCIYGLDADLIMLSLA 136
             ++ GEGEHKI+ FIR          +N  HCIYGLDADLIML L+
Sbjct:   172 EVPGEGEHKIMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLS 215
```

>> D88026 Dhm2 protein [Mus musculus], Length = 1687

Score = 126 (57.7 bits), Expect = 6.4e-10, P = 6.4e-10
Identities = 24/40 (60%), Positives = 28/40 (70%), Frame = +2

```
Query:    14 GEGEHKILKFIRENCKNDSNFKNWNHCIYGLDADLIMLSL 133
             GEGEHKI++FIR       +  N  HC+YGLDADLIML L
Sbjct:   175 GEGEHKIMEFIRSEKAKPDHDPNTRHCLYGLDADLIMLGL 214
```

>> X91617 5'-3' exonuclease [Mus musculus], Length = 1719

Score = 126 (57.7 bits), Expect = 6.4e-10, P = 6.4e-10
Identities = 24/40 (60%), Positives = 28/40 (70%), Frame = +2

```
Query:    14 GEGEHKILKFIRENCKNDSNFKNWNHCIYGLDADLIMLSL 133
             GEGEHKI++FIR       +  N  HC+YGLDADLIML L
Sbjct:   175 GEGEHKIMEFIRSEKAKPDHDPNTRHCLYGLDADLIMLGL 214
```

>> P40848 DHP1_SCHPO DHP1 PROTEIN, S43891 dhp1 protein - fission yeast [Schizosaccharomyces pombe], (D17752 and Z69240) Dhp1 protein, Length = 991

Score = 118 (54.1 bits), Expect = 8.1e-09, P = 8.1e-09
Identities = 24/41 (58%), Positives = 28/41 (68%), Frame = +2

```
Query:    14 GEGEHKILKFIRENCKNDSNFKNWNHCIYGLDADLIMLSLA 136
             GEGEHKI++FIR          N +H +YGLDADLIML LA
Sbjct:   204 GEGEHKIMEFIRSQRVKPEYDPNTHHVVYGLDADLIMLGLA 244
```

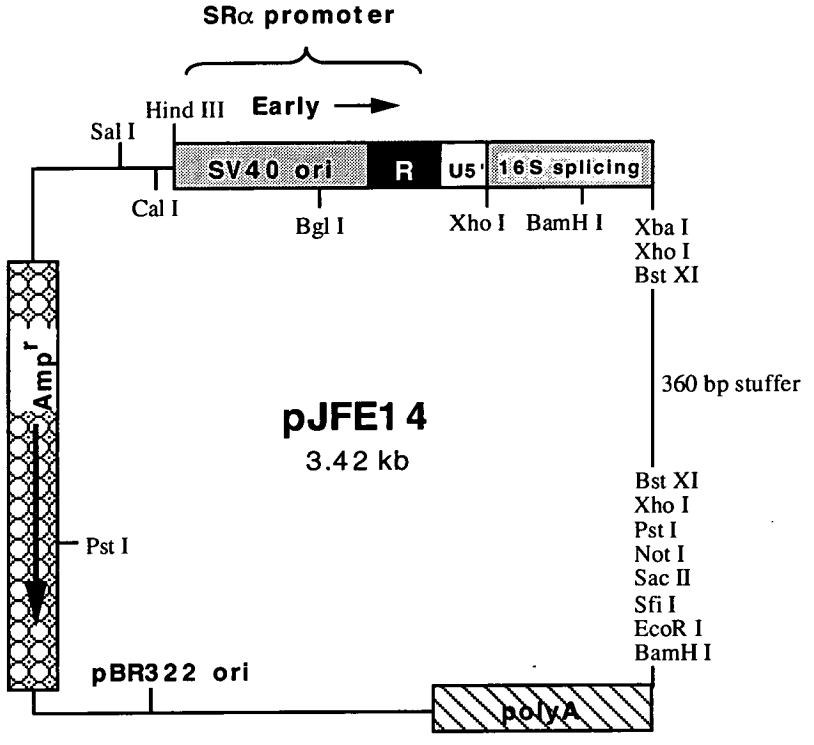
>> I49635 mouse Dhm1 protein - mouse, (D38517) Dhm1 protein [Mus musculus], Length = 947

Score = 112 (51.3 bits), Expect = 5.5e-08, P = 5.5e-08
Identities = 22/41 (53%), Positives = 27/41 (65%), Frame = +2

```
Query:    14 GEGEHKILKFIRENCKNDSNFKNWNHCIYGLDADLIMLSLA 136
             GEGEHKI+ +IR       +    N +HC+ G DADLIML LA
Sbjct:   202 GEGEHKIMDYIRRQRAQPNQDPNTHHCLCGADADLIMLGLA 242
```

Appendix C

The restriction map of the pJFE14 plasmid vector (Elliot et al, 1990) which was used to clone the *P. falciparum* asexual cDNA library for the study of this thesis. The RNA was prepared from ring and trophozoite stages of line ITO4 parasites.

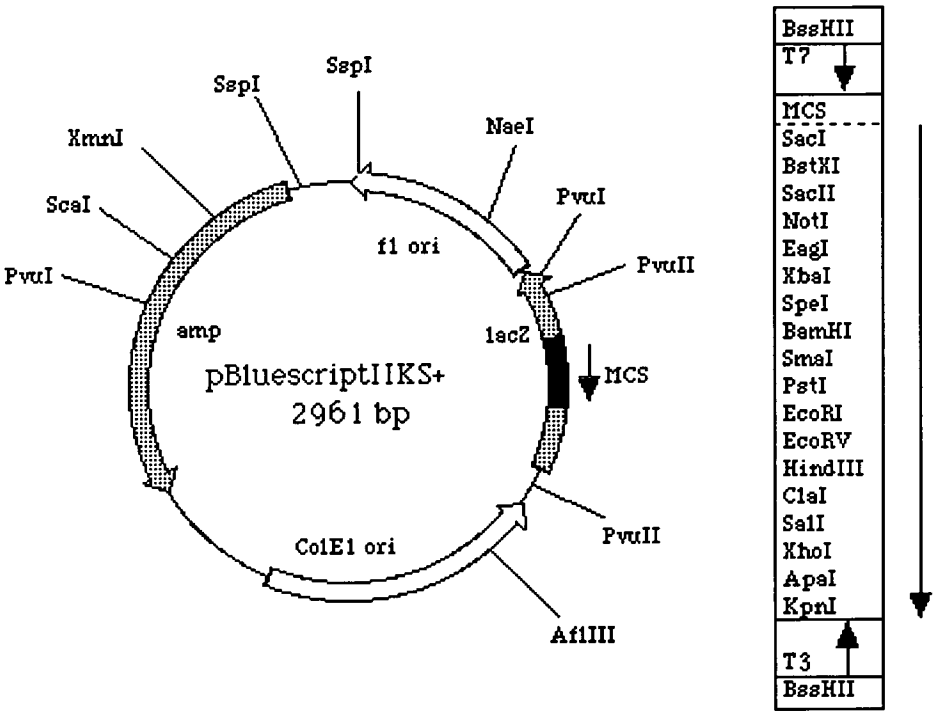


Sequence of the polylinker from pJFE14 plasmid vector:

```
   Xba I                Xho I          Bst XI                                  Bst XI
5'TCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTGGC...360bp stuffer...GCCAGCACAATGG
        S2441
   Xho I                                          Pst I       Not I Sac II   Sfi I  EcoR I
ATCTCGAGGGATCTTCCATACCTACCAGTTCTGCGCCTGCAGGTCGCGGCCGCGGGAGGCCGAATTC3'
                                                  S2442
```

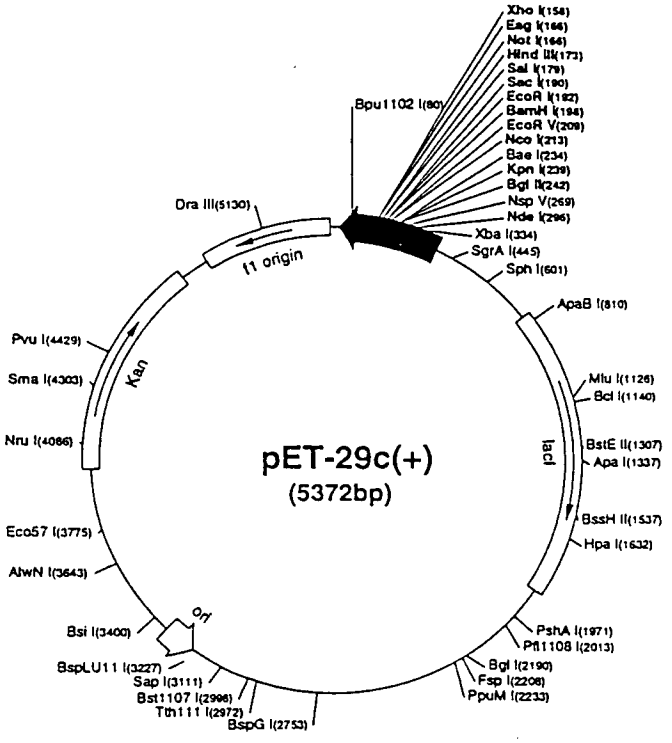
Appendix D

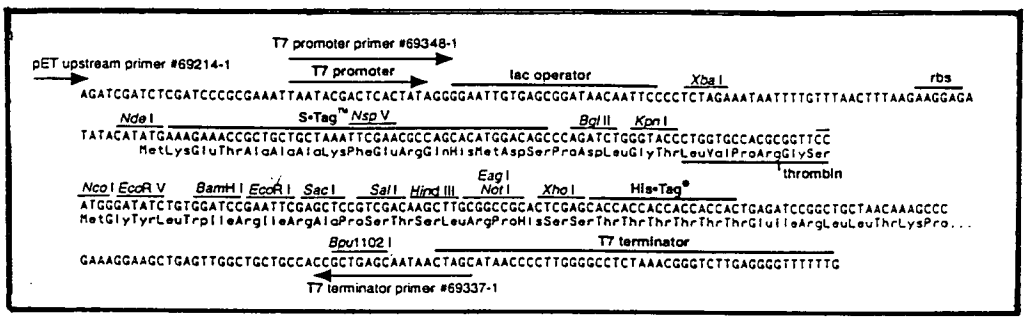
Schematic diagram of pBluescript KS+.
Adapted from Stratagene catalogue.

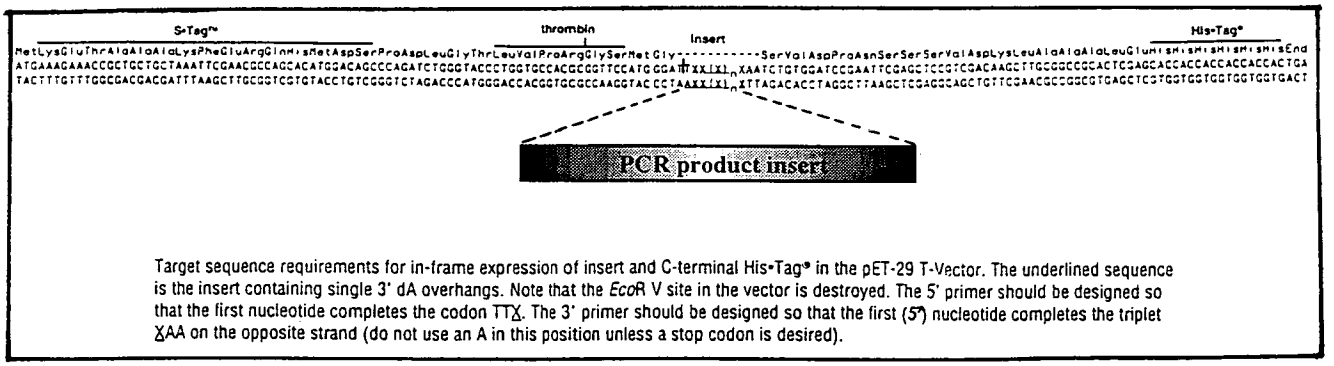


Appendix E

Schematic diagram of pET29 TA expression vector.





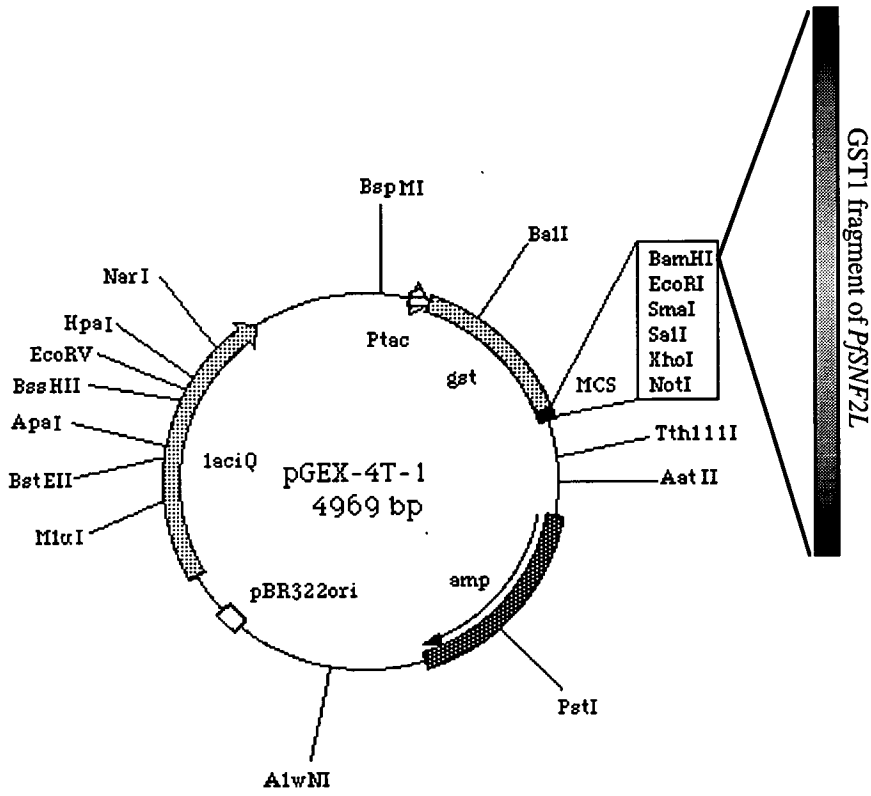


Target sequence requirements for in-frame expression of insert and C-terminal His•Tag® in the pET-29 T-Vector. The underlined sequence is the insert containing single 3' dA overhangs. Note that the *Eco*R V site in the vector is destroyed. The 5' primer should be designed so that the first nucleotide completes the codon TTX. The 3' primer should be designed so that the first (5') nucleotide completes the triplet XAA on the opposite strand (do not use an A in this position unless a stop codon is desired).

Adapted and modified from Novagen calalogue.

Appendix F

Schematic diagram of pGEX-4T-1 vector.
Adapted from Pharmacia catalogue.



Appendix G

List of the flanking sequences at the intron/exon boundary and initiation codon ATG.

	5' splice sites	3' splice sites	ATG Frankings
Z31584	P.falciparum gene for hydroxymethyl-dihydropterin pyrophosphokinase and dihydropteroate synthetase.		
	aaatatttaggtaagaaaaaag	atatacatatcactctttgattctatatatttttattttttcactttgtaggaaaaattat	tatatataatatggaaacta
	aatatatgtggtgtgtttaaaa	aatataaataatatatattcatgtatatgtatttatgtatttcttttcaggtggattagc	
X69769	P.falciparum GBPH2 gene.		
	ggccttattggtatgccgaaat	acatatctatatatatataatttttttttttttgtattatttaattttagaatgtttata	atattctttaatggatgtcc
U94594	Plasmodium falciparum heat shock protein 60 hsp60 gene, nuclear gene encoding mitochondrial protein, complete cds.		
	atgcttacaggttaatattata	tatatttatttatttattattatattttatttattaattcataaatataggttgcaataa	
Z80359	P.falciparum ARF gene.		
	ccaacaataggtatgaaaaaaa	tgtatatttatttattcattttattttatttatttatttttttttttgaaggtttcaatgt	gttattcacaatgggattat
	acaagataaggtaaaatgatta	aatgacacacgaacatatttacatattgacatgtttatttatttttatagatccgacctt	
	atagatgatggtaagaaaataa	attacatagaatgataaatgaagaagaattaaaagatgctataatttagtcttcgctaa	
	aaaggaactggtaaagaaataa	tatgtatttatatatacatatatatttttgtatccttttttttttttaggtttattcaa	
U69552	Plasmodium falciparum histidine-rich protein III (HRP III) gene, complete cds.		
	aaaattcaaggtataaaataat	tatatatatatatatatttttttattatttctttttttatattaatacagttattttaaa	ttttaaaaaaatggtttcct
	gttagataacgtaagtatttta	atatatatatgtacatttttacatttttaaatgatttttcatttttatagaataactccg	
U69551	Plasmodium falciparum histidine-rich protein II (HRP II) gene, complete cds.		
	ataattaaaagtatagaacaat	acaattatatttatacatttttgttattatttctttttctatttatatagttatttaata	tttaataaaaatggtttcct
	gttagataacgtaagcatttta	gtatatatataaatttttttcatttttaaatgctttttttatttttatatagaataattccg	
L46348	Plasmodium falciparum clone HB3 d-aminolevulinic acid synthetase gene, complete cds.		
	ccgttccaatgtaaaaaaataa	tatatatatatatatatatatttatttatttattatttttttttttttatagggaatggaa	tttttataagatgaggaaaa
X99254	P.falciparum gene encoding primase, small subunit.		
	ttttattacagtaaataaatat	cttttctttttttttttgtctgataaaaacaaattatccttttttttacagggagcttgtg	ataatactttatgaaaatgg
	tataaaaatggtataagaacaa	atatatatatatatttatgtatatacttttttatataatctttaatatagatattaaggg	
	aagatggctggtaataaagaca	ttatttatttatttatattatatctctctttctgtacttcatttatttagtcctttagta	
	agtccaaaaagtacttatcaaa	atttatttttttattataacttatatattttaatattttattttaatttaggaattaatat	
	aattgtgttggtatgactacaa	ttatatttcattatatttttttttatttatttatttgtgtatgtcatccaggaaatttttta	
	tcattgttgggtatgttaaaac	atatattaatatatcatatattaattatatttatatatttttttcattaaggtggctgatg	
	catattatcagtaataagaaaa	tatgaagataaaaaaagatatttatttttcttatccttttttgtgtttctagggttcagata	
	caatgtttgagtaaacaaaaaa	atatatatatatatatatttatttatttattcatttatccatgtttttaggagggcattt	
	ttaccttatggtatgaattaaa	tctataataaattaataatattaaaataatatgtgcatatatctttttagcttctggaaa	

```
            atttgagaaagtaataaatttt       tatatatatatatatatatttatttatttctttatttttaataattatagttcagttcta
            taaatgtcaggtaatgtattag       tatacatatatatttttttttttatatttttctttttaatttttttatagtaaagaaatt
            aacagtacaggtacggatttat       ttatatttattataaaaatatatatatatatatattatttctcaaaataggaagagtatg
            ccaacattaagtaataatttaa       gaaacatgatagtgtacatttatatatatatatatattattccctttacagaattattaag
            aattctcatagtattaataaag       tttattattatatcttaatatttttttttttttttttttttccaccttagttgaagcgga
            aaaaagaaacgtaatatatata       tctttctctatacaaaatttttcttctttacttttttttttttttttataggattgaatga
X71765 S63253  P. falciparum gene for Ca2+ - ATPase.
            acagcactaggtaataacaata       ttgtcttatacatcctattcatcaattatctttttttttaatttttttagggttcaatcc     tatataaagaatggaagagg
            aaaaattaaggtaataaatata       atgtatatatttatatatattcatttatttatatatttttttatttttaggcaagcacct
X93462  P.falciparum GR2 gene.
            gggcagcaaggttatatgaaaa       tgtgtgtgtgtgctttatatttttgtccatttctttctttttattgtaggcataacgca     aaacgcaaatatggtttacg
            tcctaaaaaagtaataccattt       gtataatttttatttatttatttttttttatttttatttttttttgtccagattatgttca
Z68200  P.falciparum glutathione peroxidase gene.
            aggattggaggtatacataata       atatatatatatatatgtataattatttatttattattcccttgctttagatcttagctt     tataaaaaatatgttcttct
            ttgtgattcggttagtaaaaaa       tattatttttttttccaaataaattatttcatatatttcatttttttccagatgcatgatg
M91672  Plasmodium falciparum RESA-2 gene, complete cds.
            cattttaaatgtaagttttatt       tgtgtagttatttaatttaatatatattaaatttttttttctttttttagtataatttta     ttgaataagtatgaaacaac
U27338  Plasmodium falciparum erythrocyte membrane protein 1, Malayan Camp variant 1 (MCvar-1 PfEMP1) gene, complete cds.
            atttttgaaggtaatatatata       catacatatacacacctacatatacatacatacctctttttattatttagaaaaaaacca     tgtcgatacaatgggggggag
X83707  P.falciparum PK1 gene.
            tggtcaataggtgtgttcaaag       atatattcatattcatttacttatatattatcattacatttttaccttaggatgcgttat     gattttcttcatgatggact
L40609 Plasmodium falciparum (strain FCR3) variant-specific surface protein (var-2, var-3) genes, complete cds's.
            atttttgaaggtaatatatata       atacctacacatacatatacacatatacatatatgttccttttttttttagaaaaaaacca     gaatacaaaaatggcacctg
            ttatctaaaggtaaatggaagt       atacctaaatacatatatacatacacatatatgttcatttttttttttagaaaaaaacca     aataacgaaaatggtgctcc
L40608 Plasmodium falciparum (strain Dd2) variant-specific surface protein (var-1) gene, complete cds.
            ttatctaaaggtaaatggaagt       atacctaaatacatatatacatacacatatatgttcatttttttttttagaaaaaaacca     acataatgtaatggtggagt
X73954 P.falciparum gene for ras-related nuclear protein.
            aaaatatatagtaagtgaaaag       tatatttttgtatttatttatttttattataatttttttttttttttttttaggctactcttg     aatttaaaaaatggattcac
M94732  Plasmodium falciparum 15 kDa vesicular-like antigen gene, exons 1 through 4.
            catcaaaatggtaagaggaaaa       tattttgtttcatttcatttaatttttttttattttttgataccttttatagagtaatgtta     tgtcatcaaaatgagtaatg
            gtcaggaaaagtaattattttt       ttatttatttatttttattttttttttttttcccaatcatattaacacaaaagatgtggcaag
            aatatgaactgttcaggcaaaa       tatataatatatttatgtttatatatatatatatatttttattttttgtagacaaaaaaaa
L01655 M98341 Plasmodium falciparum T9/106 triosephosphate isomerase (tpi) gene.
            agcaaattaggtatatatcaaa       atatatgttcaaatttgtaaaatatatatttgtatgtttttattccctagacgttgttgt     tttttacaaaatggctagaa
L34028 Plasmodium falciparum (clone HB3) heat shock protein 86 gene, complete cds.
            tcgtacaaaggtaaaaagaaga       ttacttttttacctttttactttatatttttttacttttaatttttttttttagatcaccgaat     ttcgaataaaatgtcaacgg
L34027 Plasmodium falciparum (clone Dd2) heat shock protein 86 gene, complete cds.
            tcgtacaaaggtaaaaagaaga       ttacttttttacctttttactttatatttttttacttttaatttttttttttagatcaccgaat     ttcgaataaaatgtcaacgg
Z26314 P.falciparum gene for STARP antigen.
            ttctaataaagtaagataaaaa       taatatataattatttatatttttttttttttttttaatattacttttttagaatttaaaat     acaaaaaataatgatacata
X05074 Plasmodium falciparum exp-1 gene for malaria protein exp-1.
            aaaaataaaagtaagtcatttt       tgatatatatgtgtatgatatataatattttatttttctttttgttaatagaaggatcagg     tatattcaaaatgaaaatct
```

```
            tacaccagaggtatgatttttt      tgtgtatccaaatgtctatatatttatttttatttttctttatttttatagcaaccacaag
U07706 Plasmodium falciparum NF7 dihydropteroate synthetase gene, complete cds.
            aaatatttaggtaagaaaaaag      atatacatatcactctttgattctatatatttttatttttcactttgtaggaaaaattat    tatatataatatggaaacta
X75420 P.falciparum (FCR3) cpn60 gene.
            aataaaagtggtacgttaaaaa      ctataagagctccatcatttggtgatagaagaaaagactatctaaaagatttatgtattg    ttttttttttatgagaatga
X61921 P.falciparum pfc2 gene for p34cdc2 protein kinase.
            ccatcaacaagtactacaaaaa      ttatttacataatctcctaatttttttttttctttcaatttcctaaatagcaataagaga    aagtgaaaacatggagaaat
            gtttgtgaaggtaaaaatattt      atttatttatttttaattattcttatccattttataaactacattgtataggagggttgga
            ggttagccaggtaaaatatata      tcccaataattataataaatatttctaattttataattaaatatttccagagcatttgga
            ggaatcatttgtaaagaaaaat      atataaaaatatattatttcatatatctatttttttattttgtcattctagttgaagggat
J04000 Plasmodium falciparum serine-repeat antigen protein gene (SERA), complete cds.
            ttcatattgtgtaagaatgaaa      aatataatatttttacgcatacacaaacatttgtcattatttttttttaggtgttatatt    atatatcaaaatgaagtcat
            actactgaaagtataaaaaata      attatttatattattattatttattatttatttatttttttttttagataacaaagt
            gacattccaggtaaataacaat      tatatatatatatatacatattatatatttttttaatttttttgttttagaaaaatgtga
M22718 J03988 P.falciparum actin II gene, complete cds.
            agaacaacaggtatagaatgaa      gatttaattttatctgattttatttcttggtcctaaatggttccctgtaggtatcgttct    ttttctttagatgtctgaag
L22057 Plasmodium falciparum ribonucleotide reductase R1 subunit gene, complete cds.
            tattgatcaggtaaaattaaaa      ttttgttcatacatttatattcatatatatttttttttccttttttttatagtcccaatcat    tatacaaacaatgtatgtgt
Z28975  P.falciparum (K1) Ag15 gene.
                  atggtaagaggaaaa       attttgtttcatttcatttaatttttttttattttttgataccttttatagagtaatgttaa            atggtaatgt
            gtcaggaaaagtaattattttt      ttatttatttatttttttttttttttttttcccaatcatattaacacaaaagatgtggcaag
            tgaccttaaggtaaaaaattga      ttttccttttttgttgttatatattaactattaataatataaaatatttttgtagacaaaa
L11172 Plasmodium falciparum RNA polymerase I gene, complete cds.
            gacccatacaggtataatcaaa      atatatatatgtatgtatttatttttttttttttttttttttttttcccctttaggtgatttga            atgtacgata
            aattcaattggtaagggaaaaa      tatatttttatatgtgtatatatataatatttctatttttttttttgtagattacttatc
L18785 Plasmodium falciparum DNA polymerase alpha (Pol alpha) gene.
            aaaacgatgagtgagtagataa      ttttttatattttcttatattttatttttatatttatttttcttctcagattcgatttg    ttatatatatatgtatatat
M99442 Plasmodium falciparum calmodulin gene, complete cds.
            ggagatggaagtaaatcataaa      tattttatattatattatattatattatatttttttattttttctcattagctataacaac    aaatatatcaatggcagaca
A00661 P.falciparum RESA gene for ring-infected erythrocyte surfaceantigen.
            cattttaaatgtaagttttttt      tatctttgttattatatataatttttttttttttttttttttttttcatagggtaatcttg    ctaaataattatgagacctt
X01469 Plasmodium gene fragment for histidine-rich protein.
            ttcaggaagcgtaagataaagc      ataaaaaaaaagaaaaatctttaacatttaatttttttatttatattatagaattcatgta    gaaaggaagtatgtttactt
X16075 Plasmodium falciparum beta-tubulin gene.
            aatagatccagtaagtttaaaa      tttatatatttatttatttatttttttttttttttattttattttttttagagtggtacct    ttattaaagaatgagagaaa
            acaacacaaagtaagaaggaac      attttttttttttctttttctttttttttttttttttgttttttttctttagatcaagtgtt
X15979  Plasmodium falciparum alpha-tubulin I gene.
            atgcttgctggtaaaacgataa      aatacatcttttctttattctttttataaataatttgttgaatttttttagggaattgttt    taaataaaaaatgagagaag
            aggaaaacatgtaagaaataat      tatatacatatatatatatatatatatttttttttttttttttttcctttttaggtaccacgtt
Y00060 P. falciparum SD17 gene for knob-associated histidine-rich protein.
            ctctaataacgtaagttcataa      aattttattttattttattttaatttaattttattttttttttcatatagtgcaataatg    ttattagagaatgaaaagtt
L02822 L02823 Plasmodium falciparum heat-shock protein gene, complete cds.
```

gactcaaacagtaagtattata gtaatatgtgtatatataaaaaatttgaatcttttatttttaatttttttagttgagggacc taattcaaaaatgaaccaaa

X67288 P.falciparum cpk gene.

acttaggagcgtaagaagaaag atatatatatttttttttttttttttttttttttttttttttttcacatcagtttaagaatg gtaacatgaaatggggtgtt

gaatattcaggtgaatataaaa tatatatattatatatatgtacatattatatatacattttaacttttttagagtttatttc

attagcaaatgtaaatatataa tcattatatttttttttttttttttttttttttaatgatttacctatgcagttatttggtt

agataacatggtaatattataa tgccataaaatttttaacttgaatttttttttttttttttttttttcagattgactttg

M73770 M37820 P.falciparum RNA polymerase III largest subunit gene, complete cds.

tcactttgcaggtacatacaag atatttttttttttttttattattaatttttttttttttttttttttttaggtgtggcta atgatgaaga

cgtttttcaagtaattaacatg atatatattttatgtatttgtaaatattcttatatttttatttttttcttaggttttgggta

ggcctttacgttagtttattta atttttacatcttgtcatattatttaatttttttttttttttttttttagtggagatattt

ttataaattgtaagttttgaga atatatatatatatttatatttatatttatatttttttttttttttttagtgagaaggaaa

M69183 Plasmodium falciparum (MESA) mRNA exons 1-2, complete cds.

ttttttattagtaagttgctaa aaatgaaaaaaaaagtatatttattatcaccgtaatatatttttttataggatatctata aaataatatgatggaaaatg

M19881 P.falciparum knop protein (KP) gene, complete cds.

ctctaataacgtaagttcataa attttatttttattttattttaatttaatttttatttttttttttcatatagtgcaataatg ttattagagaatgaaaagtt

J03998 P.falciparum glutamic acid-rich protein gnen, complete cds.

ttctactaaggtaacaaaaata atttacatatatatttacatatttataatatatatgtatttttttttttagtgcttttcca tttttaaaaaatgaatgtgc

M93397 Plasmodium falciparum erythrocyte binding protein gene, complete cds.

gctttttcaagtatgatattca taaaaaaaattaatttctcattaatttttttttctttttttctttttaggtatgccata tgtgcatacaatgaaatgta

aatatcaaaggttagaaaaaat tatatatatatatatatatatattgtattatatatttttttttttatagttctgaagga

tccaatatgtgtaatacaaata ttatattatcctataatatttatttttatttttatgtttttttttttttcagtcaatcaaca

M59770 P.falciparum calmodulin gene, complete cds.

ggagatggaagtaaatcataaa gttcttattttatattatattatattatttttttattttttctcattagctataacaac aaatatatcaatggcagaca

M31205 P.falciparum beta-tubulin gene (bTub), complete cds.

tagatccaaggtagtttaaaaa tattattattatatattattattattatttttttttttatttatttttttagtggtacctat ttattaaagaatgagagaaa

acaacacaaagtaagaaggaac ttttttttctttttctttttttttttttttttttgttttttttttttctttttagatcaagtgtt

M34390 P.falciparum alpha-tubulin II gene, complete cds.

atgcttgctggtaaaaaaaaaa taattaaaaatatacatatatatatatatatatatatttttttttttatagggaattgttt tttacgaacgatgagagaag

tggaaaacatgtaataataaaa tatttatttacttatacatgtgtgcttctttttgtcgttactcattttaggtaccacgtt

M28881 P.falciparum aldolase gene, complete cds.

tttgaaaatggtaagaatataa attatatatttttccttttttcttatttttatatttttatatttttttttaggctcattgca tattttgaaaatggctcatt

AF031144 Plasmodium falciparum aspartate transcarbamoylase (ATCase) gene, complete cds.

tccatcaaaggtaagaaaataa gcatatatatatatatatatatatttttatttatttatttttccttttttagaaaaatgtcg atgttttata

cgacacaaaggtaataatgtta ttaatgcatgtaaaatatatacatacatatatatatttatatatttatagatacttcacc

U89025 Plasmodium falciparum protein phosphatase-beta gene, complete cds.

ttccatatcggtatatttttat tatatatatatatatatttttctgttttttttttttttttttttttttaggaggtttacc aaagggaaggatgaaaaatt

U40232 Plasmodium falciparum exported serine/threonine protein kinase (FEST) gene, complete cds.

aaaataaaatgtataatataaa tatatatatatatatatatttatgtatttttatttttatttttttttttataggtgttcgcat ggcaaaagctatgaataata

U08113 Plasmodium falciparum serine repeat antigen (SERA3) gene, complete cds.

ttcatattgtgtaagaatgaaa aatataatatttttacgcatacacaaacatttgtcattatttttttttttaggtgttatatt atatatcaaaatgaagtcat

actactgaaagtataaaaaata attatttatattattattatttattatttatttatttttttttttagataacaaagt

gacattccaggtaaataacaat tatatatatatatatacatattatatatttttttaatttttttgttttagaaaaatgtga

M65160 P.falciparum glycophorin binding protein homology (GBPH) gene, complete cds.
agccttattggtatgataat tatatatttatatatatatttttttctgtttgcatcatttatttttagggtgtttata ctaagtaaaaatgcgtattt

AF027825 Plasmodium falciparum glutathione reductase (GR3) mRNA, complete cds. aaacgcaaatatggtttacg
AF017139 Plasmodium falciparum casein kinase 1 (CK1) mRNA, complete cds. aatttaaaaaatggaaatta
D85686 Plasmodium falciparum DNA for Pfj1, complete cds. tgattatataatgttagctt
U84403 Plasmodium falciparum adenylosuccinate lyase gene, complete cds. atttttttattatggatgtac
AF012551 Plasmodium falciparum ornithine decarboxylase (ODC) mRNA, complete cds. tggaaagtatatgttcatga
U57371 Plasmodium falciparum ADP-ribosylation factor-like protein (plARL) gene, complete cds. atgggattga
U57370 Plasmodium falciparum ADP-ribosylation factor (plARF) mRNA, complete cds. atgggtttat
AF008549 Plasmodium falciparum chorismate synthase (AroC) mRNA, complete cds. aaaaaacattatgagcacat
U38963 Plasmodium falciparum 60 kDa heat-shock protein PfHsp60 mRNA, complete cds. tgtagaaacaatgatatcaa
M77834 Plasmodium falciparum membrane-associated calcum-binding protein(Pfs40) gene, complete cds. tttttacaaaatgatgaaaa
U85963 Plasmodium falciparum asparigine-rich protein gene, complete cds. atgaataata
U70366 Plasmodium falciparum rab6 (Pfrab6) mRNA, complete cds. ttagatatatatgtatatac
U78753 Plasmodium falciparum ribosomal P2 phosphoprotein (P2) gene, complete cds. tatataaaaaatggctatga
U65407 Plasmodium falciparum apical membrane antigen 1 (AMA-1) gene, complete cds. atgagaaaat
U78291 Plasmodium falciparum cAMP-dependent protein kinase catalytic subunit (PfPKA) mRNA. acttctcgctatgcagttta
U56663 Plasmodium falciparum acidic ribosomal phosphoprotein PO (PFPO) mRNA, complete cds. atggcgaaat
U65959 Plasmodium gallinaceum circumsporozoite protein (CSP) gene, complete cds. gtacttcaaaatgaagaaat
U73195 Plasmodium falciparum MO15-related protein kinase Pfmrk (Pfmrk) gene, complete cds. ttatgttctcatggaaaata
U67764 Plasmodium falciparum thrombospondin related adhesion protein (PfTRAP) gene, complete cds. aaaataatgaatcatc
U31083 Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) gene, complete cds. acaaacaaaaatggcgcctc
D86574 Plasmodium falciparum DNA for iron-sulfur subunit of succinate dehydrogenase, complete cds. aataagaaaaatgttaaaaa
D86573 Plasmodium falciparum DNA for flavoprotein subunit of succinate dehydrogenase, complete cds. cataagatttatgcaatcaa
U36377 Plasmodium falciparum mitogen-activated protein kinase-related protein gene, complete cds. aaaaaaaaaaatgcctaaag
U49381 Plasmodium falciparum phosphate carrier (PfMPC) mRNA, nuclear gene encoding mitochondrial protein tttttttcataatgatgaaaa
U37225 Plasmodium falciparum elongation factor 3 related protein (pfgcn20) gene, complete cds. gaaaaaatatatgatggaag
U16995 Plasmodium falciparum ATPase gene, complete cds. atcaatagttatgagttctc
U18984 Plasmodium falciparum putative phosphatidylethanolamine-binding protein gene, complete cds. gttggatataatgacaatac
U54642 Plasmodium falciparum phosphoribosylpyrophosphate synthetase (prs) gene, complete cds. atggaaaatg
L04161 Plasmodium falciparum (clone Pfg377 [PfsXLX]) DNA sequence, complete cds. ttttttccaaatgaagttgc
U51645 Plasmodium falciparum cytidine triphosphate synthetase gene, complete cds. atggatagtg
U41269 Plasmodium falciparum Rab1 protein (PfRab1) gene, complete cds. gaaagaaaaaatgactgaga
U10322 Plasmodium falciparum HB3 cyclophilin (cyppf1) gene, complete cds. aatcgagaatatgaaaaata
U15994 Plasmodium falciparum histone H3 mRNA, complete cds. aacaagcaaaatggcaagaa
U34363 Plasmodium falciparum CTRP gene, complete cds. caaataacaaatgaagaaag
L42636 Plasmodium falciparum variant-specific surface protein (var-7) mRNA, complete cds. gcatacaaaaatggagccag
L31630 Plasmodium falciparum (clone: D10) high mobility group protein, complete cds. caacgctagaatgaagaata
M14632 Plasmodium falciparum GP185 mRNA encoding antigenically diverse glycoprotein, complete cds. taattcaataatgaagatca
U14735 Plasmodium falciparum histone H3 mRNA, complete cds. aacaagcaaaatggaaagaa
U14734 Plasmodium falciparum histone H2B mRNA, complete cds. atattgcaaaatggtatcaa

U14189	Plasmodium falciparum MCP1 mRNA, complete cds.	taaaaacaaaatggctcaat
U25814	Plasmodium falciparum chromodomain protein mRNA, complete cds.	agaagaaaaaatgacagggt
U16955	Plasmodium falciparum ATPase 2 gene, complete cds.	atttttaaatatgtatgaac
M80655	Plasmodium falciparum glucose-6-phosphate dehydrogenase (PFG6PD) mRNA, complete cds.	taatatatatatgaataact
L07944	Plasmodium falciparum secreted polymorphic antigen gene, complete cds.	attatacaaaatgaaaagtt
L28825	Plasmodium falciparum polymorphic antigen, complete cds.	caaatgaaaagtt
U20969	Plasmodium falciparum circumsporozoite protein (CS) gene, complete cds.	atgatgagaa
U20985	Plasmodium falciparum rhoptry associated protein (RAP-1) gene, complete cds.	atgagtttct
L13381	Plasmodium falciparum HB3\W2 transport protein (mdr2) gene, complete cds.	gtaaagaaaaatggatgtat
J03772	Plasmodium falciparum (clone HB3) dihydrofolate reductase/thymidylate synthase (DHFR-TS) gene	atgatggaac
L04162	Plasmodium falciparum (clone sY) s230 mRNA, complete cds.	aaaaatactcatgaagaaaa
U04335	Plasmodium falciparum adenine nucleotide translocase mRNA, complete cds.	atgagttctg
M59706	P.falciparum glutamate rich protein (glurp) gene, complete cds.	aaatatcaaaatgagaaacc
U08851	Plasmodium falciparum isolate K1 polymorphic antigen gene, complete cds.	caaaatgaaaagtt
U01323	Plasmodium falciparum Dd2 ribonucleotide reductase large subunit (RNR1) mRNA, complete cds.	tatacaaacaatgtatgtgt
U01322	Plasmodium falciparum Dd2 ribonucleotide reductase small subunit (RNR2) mRNA, complete cds.	tttgaatagaatggctgatg
L15631	Plasmodium falciparum (NF54) circumsporozoite-related antigen (Ag 5.1)mRNA, complete cds.	atgaaaatct
U04640	Plasmodium falciparum multi-drug resistance homolog 2 (pfmdr2) gene, complete cds.	taaagaaaaaatggatgtat
U07365	Plasmodium falciparum S-adenosylhomocysteine hydrolase mRNA, complete cds.	tttttgagaaatggttgaaa
M97214	Plasmodium falciparum S-antigen gene, complete cds.	ttataatacaatgaatagaa
L01654	Plasmodium falciparum triosephosphate isomerase (tpi) mRNA, complete cds.	tttttacaaaatggctagaa
U00152	Plasmodium falciparum enolase gene complete cds.	atggctcatg
L32150	Plasmodium falciparum carbamoyl phosphate synthetase II gene, complete cds.	atgtatattt
J03993	P.falciparum serine-repeat antigen mRNA, complete cds.	atatatcaaaatgaagtcat
J03084	P.falciparum aldolase (P41) gene, complete cds.	cactgaatatatgaatgccc
J03828	P.falciparum integral membrane protein mRNA, complete cds.	attagtcaaaatgaaggtaa
J05544	Plasmodium falciparum glucosephosphate isomerase (ggi.R1) gene, complete cds.	atatttttctatgaatatgg
M12897	Plasmodium falciparum glycophorin binding protein (GBP-130) gene, complete cds.	ttgtgtaaatatgcgacttt
J03028	P.falciparum dihydrofolate reductase-thymidylate synthase gene, complete cds.	tctccttttttatgatggaac
K02194	Plasmodium falciparum circumsporozoite (CS) protein gene, complete cds.	attacaattcatgatgagaa
M11145	Plasmodium falciparum circumsporozoite-protein related antigen(CRA) mRNA, complete cds.	tatattcaaaatgaaaatct
M22719 J03988	Plasmodium falciparum actin I gene, complete cds.	atctgtaaaaatgggagaag
J03902	Plasmodium falciparum p101/acidic basic repeat antigen (ABRA) DNA, complete cds.	tttaaaaaagatgatgaaca
J04656	Plasmodium falciparum 41-2 protein antigen gene, complete cds.	tgttttaagtatggataaat
L15426	Plasmodium falciparum ornithine aminotransferase (OAT) mRNA, complete cds.	aaaagtaaaaatggatttcg
L22058	Plasmodium falciparum ribonucleotide reductase R2 gene, complete cds.	aaagagataaatgagaagaa
L15446	Plasmodium falciparum dihydroorotate dehydrogenase homologue gene, complete cds.	tttttaaaaaatgatctcta
L08135	Plasmodium falciparum transmission blocking target antigen (Ps230) mRNA, complete cds.	cttttaataaatgaagaaaa
M80807	P.falciparum Pr86 rhoptry precursor protein (P82) gene, complete cds.	tttatatataatgagtttct
M19753	Plasmodium falciparum heat shock protein (HSP70) homologue mRNA, complete cds.	ttaagaaaaaatggctagtg
M93720	Plasmodium falciparum L-lactate dehydrogenase (LDH-P) mRNA, complete cds.	atggcaccaa
L08200	Plasmodium falciparum vacuolar proton adenosine triphosphatase(vapA) gene, complete cds.	acggagcaaaatgacaaaag

M92054 Plasmodium falciparum hexokinase gene, complete cds. taatttaaaaatgagtgagt
M64705 Plasmodium falciparum protein antigen mRNA, complete cds. tttcttcaacatgaatattc
M94013 Plasmodium falciparum sporozoite surface protein 2 gene, complete cds. atgaatcatc
M32853 P.falciparum rhoptry associated protein 1 (rop1) gene, complete cds. tttatatataatgagtttct
M59249 P.falciparum 3-phosphoglycerate kinase gene, complete cds. cttccataaaatgttaggca
M28261 P.falciparum p75 mRNA, complete cds. tggtggtgttatgaccaaat
M28890 P.falciparum membrane protein Pf7 gene, complete cds. attagtcaaaatgaaggtaa
M29154 M24322 P.falciparum multidrug resistance (MDR) gene, complete cds. gtgttgaaagatgggtaaag
M86518 Plasmodium falciparum high mobility group-like protein (PS16) gene, complete cds. accaaaagtgatgatgaaga
M37213 P.falciparum major merozoite surface antigen (gp195) gene, complete cds. atgaagatca
M64715 Plasmodium falciparum DNA polymerase-delta gene, complete cds. atggaagaac
M81341 M80590 Plasmodium falciparum cysteine proteinase gene, complete cds. taataaaaatatggttgcca
M64107 P.falciparum 17 kD sexual stage protein mRNA, complete cds. tttcttcaacatgaatattc
U03915 Plasmodium falciparum vacuolar ATPase subunit B mRNA, complete cds. aaaatatataatgagtaaag

U51723 Plasmodium vivax Serine Repeat Antigen (SERA)-Related genes
V-SERA 2 tgggccatgtgtaaggagagag tgatgtataccctttttacacgcaacccattggtgacgttctccacgcaggtgccctgat ggggaacgccatgaatccca
ccatcggacagtaatagaaa atatcccatttgctgattacatttttttttttattttttctaccccccagacaacaaagt
gacattccaagttagcacgcac acgcttatctactcatgctcacctccacacgctcgtctgctccccttcaggaacgtgcga
V-SERA 4 ttcatactaagtaagaggggaa ccattccttccactcatttttgctcgccccttgcttcccttccctgtaggcgttgcact acagaacaaaatgaagttgg
ccatcggacagtaatggaaaag taatatcccatttgctgattacatttttatttttttttttctgcccccagacaacaaggt
V-SERA 3 ctgctaataggtaacaatggga tccatcccagaggagagccccccacttctgcttcccctcaccatggggagttagatgcct atgcggcaagatgaagtccc
ggggataaaagtaaggcgggaa catcatccacataaccctcctgaccttcactcatcactacccctctgcagctgccgacag
gaaattccaggtaggttctcat agtagcgtcaacattgctccgcttcgatgacacctccctctttttgcagagaagtgcga
V-SERA 1 atgtgtgcatgtggtcgcgcct agacgttcatccctgtttatatttttctcccccctctgaatttgaacgcagattgggtctg gtttctgttaatgcttgtgc
gtggggagaagtcctaatgaag tttttccccgcacgtgcacacccccgcgcaatgacccccccctgcagcagcggacaagac
gagataccgggtaggttccatg gcaagcttaacattgcttcgcttcaatgcacacctccctctttttgcagagaagtgtga
V-SERA 5 tcttgatattgtgtaaggatga atggaagggcacccccctcattacattcccctttaatgcttttcaccgcagatatgttgct gaggaacaaaatgaagtctc
ggggagacaagtaagaggaaat catcatccacacgatcgttctgaccctcacgcatcactacccccctgcaggctccgacaa
caaattccaggtaggtttacgg gtaacgttaacattacttcgcttcaatgcacacctccctctttttgcagagaagtgtga
X84734 P.vivax GAM1 gene.
cgctgtatggtaaggtggatga ctcttccacctgcttcgtttcgcctaacccccttgccgctcccccctcaggtgcagaccac cccccacgggatgagatcgc
ccacctgtgcatggcgcacc

M61095 M37514 Plasmodium vivax Duffy receptor gene, complete cds.
gttacacaaggtatcatataag aaaaagctaataatcgtgatgtttgcaatttttttcttcatttctctcaggtaaataatg tttaaatactatgaaaggaa
ggctttaaaggtatgcagaaaa taattttaggaatgttataaacatttttgtacccaaaattctttttgcagacaagactta
tcaaaaatgagtaaccagaaaa cacatatatgtacgtatatataataaacgcacactttcttgttcgtacagttctgaagaa
atgcctaacagtaattcaaatt tgttacatcattttgcgtttttctttttttctttttttttttctttttttagatattgaaca
M88097 Plasmodium vivax retyculocyte binding protein 1 gene, complete cds.
aactatttaggtaacacacatg cttatgcgtgatggtgtaaatttgtaatttttggttatttgcgctgagtgggattatcta tgtttaaaccatgaaaaggg
X55197 P.yoelii YEL6 gene for ATPase.

actgccctcggtaagcgccgcc ttcaccccgttttcacttctcaccacttcccaccacttctcacttctcaggattcaaccc ttaatatataatggaaaata
aaaggtcaaggtaatttctact agaaaccctaacgatgtgcacacctatttgtatatctcttcatttttaggcaagtacac
D45849 Plasmodium yoelii gene for cAMP-dependent protein kinase catalytic subunit, complete cds.
cttggaactggtaagtgtgtaa tgtatatgtatatatatttttttttttttttttttttttttaaatttgaagggtcatttgg ttttgttaaaatgattcaat
tccattttgtgtaagtgaatta tgttgattttgtatatttttttgatttttttttttttttttaatggtaggttaaattat
ttgtatacaggtaggggcacct ttttacactttgcatttttacacttttatttgttcgcgctcctatgcagagacttaaag
tggaaaagcggtatgaaaagtg ggtttttttttattttattatttttttttttttttttttttattttaggtggattggt
U43538 Plasmodium yoelii hepatocyte erythrocyte protein 17 kDa gene, complete cds.
gttattaagagtaagttgattt tttatttattatttatttattatttactcacttttttaataatttgtttagaacacggaga tttcttgaaaatgaaaatca
L27591 Plasmodium chabaudi adami erythrocyte membrane antigen 1 gene,complete cds.
cacgggttgcgtaatatacgaa aaattaacataacaatatttgatatatttttaattttgttatttttttagtttggatgct gttgttgaagatgaaggcaa
L27590 Plasmodium chabaudi chabaudi erythrocyte membrane antigen 1 gene, complete cds.
cacgggttgcgtaatatacgga taataaattaacaatatttgatatgtttttaattttgttatttttgtttagtttggattcc gttgttgaagatgaaggcaa
tatataataagtaataataaat taataaataataaataaatatatatatacatatatatatttttaatgcagatgtgttatt aagaccaaatgatatttttt
AF031904 Plasmodium berghei nucleosome assembly protein (pbB7) gene, complete cds.
aaacgcacaggtaattttgaga tcgttttaatgttttttttttttttcaaaaataaaaaattttattcatagaaaataccct tttaagaaaaatgaagagag
ctaaacattcgtaagctattta atgctcacaaaattatacacttatctgtttttttccctttttctttttacagaaaaattgac
tttgaagacagtaataaatgag tttttaatttttataaacttcaaatgtttgtctgtaactttacatttgcagttcaaaagga
M90695 Plasmodium knowlesi erythrocyte binding protein (gamma) gene, complete cds.
gtcacacaaggtactgaaggac aagctaagaatcgtgattagtgtaatctgttctctttctttttctttttaggctaataatg tttaaatacaatggaaggaa
agctccacaggtatagaggaaa cattttaggaatgttataaatttttttgtaatcaatattctttttgcaggcatagttta
cctcgaatgagtaaccaaagaa gacatatatgtaccaatatataataaatgcacacttccttgttcgtacagttatgaagaa
atgcctaacggtaattcaaatt aagaagaaatattaaatcattctgcattattctttttttttttttctttagatattgagca
M90694 Plasmodium knowlesi erythrocyte binding protein (beta) gene, complete cds.
gtcacacaaggtaatgaaggac aagctaagaatcgtgatttgtgtaatcagttctttttctttttctttttaggcgaataatg tttaaatacaatggaaggaa
cacaggtatacaggaaaagatg tcattttaggaatgttataaatttttttgataatcaatatctcctttgcaggcatagttt
cctcgaatgagtaaccaaaaaa gacatatatgtacctacatataatgaatgcacacttcctagttcgtacagttatgaagag
M90466 Plasmodium knowlesi erthrocyte binding protein (alpha) gene, complete cds.
gtcacacaaggtaatgaaggaa aagctaagaatcgtgattagtataatctgttctctttctttttctttttaggctaataatg tttaaatacaatggaaggaa
agctccacaggtatacaggaaa cattttcggaatgttataaatttttttgtaatcaatattctttttgcaggcatagttta
cctcgaatgagtaaccaaaaaa agacatatatgtaccaatatataataatgcacacttccttgttcgtacagttatgaagaa
atgcctaacggtaattcaaatt atgaagaaatattacaccattctgcattattccttttatttcttctttagatattgagca
Y11396 P.cynomolgi bastianelli EBP gene.
gtcacacaaggtatcatgtaaa aaaaagctaataatcctgatgtctgctttttttttccttcatttctctcaggaaaatatg tttaaacattatgaaaggaa
agctttagaggtatgcagaaaa cattttaagaatgttatcacaatttttgtaactaatattcttttgcagacaaggcttatt
gatcaaaaatgagtaaccagaa acatatatgtacgtatatataataaatgcacacttccttgttcgtacagttctgaagaag
gatgcctaacagtaattcaaat aaagaaatattacatcattttgctttattcttttttatttcttttttagatattgaacaca
Z30339 P.reichenowi STARP gene for STARP antigen.
ttttatcaaagtaagaaagaaa tatataacatataatgatttttttttttttttattattattacttttagaatttaaaat atgatacata

U51724 Plasmodium vivax serine repeat antigen-related (V-SERA 5) mRNA, complete cds. gaggaacaaaatgaagtctc

M60807 Plasmodium vivax merozoite surface antigen 1 (MSA1), complete cds. atgaaggcgc
L26362 Plasmodium vivax cysteine proteinase mRNA, complete cds. cttatgcaaaatggcgcagg
U09737 Plasmodium vivax Gabon clone G24 circumsporozoite protein gene, complete cds attatttaacatgaagaact
M11926 J04090 M20671 Plasmodium vivax circumsporozoite protein gene, complete cds. ccagaacaagatgaagaact
U09766 Plasmodium malariae China-1 CDC circumsporozoite protein gene, complete cds. atgaagaagt
DEFINITION Plasmodium malariae (clone 255A) circumsporozoite protein (CSP) gene, complete cds. ttgctccaacatgaagaagt
U67763 Plasmodium berghei thrombospondin related adhesion protein (PbTRAP) gene, complete cds. aaagaaatatgaagctct
D88664 Plasmodium berghei DNA for ookinete surface protein, complete cds. ttatttaaaaatgaatactt
U43521 Plasmodium berghei merozoite surface protein-1 gene, complete cds. ttagttgaaaatgaaggtga
L04508 Plasmodium berghei heat shock related protein gene, complete cds. aaaagataaaatgggttatg
U45969 Plasmodium berghei apical membrane antigen-1 gene, complete cds. atgaaagaaa
U46549 Plasmodium berghei heat shock protein ClpB mRNA, complete cds. tcaattatttatgagtgatg
U12275 Plasmodium berghei ANKA dihydrofolate reductase thymidylate synthase (DHFR/TS) gene atgttacaaaatggaagact
M81793 Plasmodium berghei fructosebisphosphate aldolase (ALDO-2) gene, complete cds. aaatgcccctatgaaacttc
M28887 Plasmodium berghei (clone 236L) circumsporozoite protein (CSP) gene, complete cds. atgaagaagt
L19784 Plasmodium chabaudi adami integral membrane protein mRNA, complete cds. caattaaaaaatgaaagtcc
L22982 Plasmodium chabaudi chabaudi merozoite surface protein-1 gene, complete cds. atgaaggcga
M30834 P.chabaudi dihydrofolate reductase/thymidylate synthase (DHFR-TS) gene, complete cds. atggaagata
U80896 Plasmodium chabaudi 93 kDa protein mRNA, complete cds. tttatcaaaaatgaataaag
U49743 Plasmodium chabaudi DS apical membrane antigen 1 (AMA-1) gene, complete cds. atgaaagaaa
M25248 Plasmodium chabaudi apical membrane antigen I (AMA-1) gene, complete cds. agctataaaaatgaaagaaa
U43145 Plasmodium chabaudi repeat organellar protein gene, complete cds. agaaaagccaatgatttta
M34947 P.chabaudi major merozoite surface antigen mRNA, complete cds. ttcgttgaaaatgaaggcga
M95789 Plasmodium chabaudi acidic phosphoprotein mRNA, complete cds. gttgttgaagatgaaggcaa
D89082 Plasmodium yoelii mRNA for 28 kDa ookinete surface protein, complete cds. ttatttaaaaatgaatactt
D89081 Plasmodium yoelii mRNA for 21 kDa ookinete surface protein, complete cds. tttcacgaaaatgaatttta
U45971 Plasmodium yoelii apical membrane antigen-1 gene, complete cds. atgaaagaaa
L27838 Plasmodium yoelii rhoptry protein, complete cds. caaagatccaatgaaaaata
J04668 P.yoelii major merozoite surface antigen gene (PMMSA), complete cds. ttagttgaaaatgaaggtga
J02695 Plasmodium yoelii circumsporozoite protein (CSP) gene, complete cds. aaaatgaagaagt
M84732 Plasmodium yoelii sporozoite surface protein 2 gene , complete cds. agacagaaacatgaagctct
L33882 Plasmodium reichenowi surface antigen (Prs48/45) gene, complete cds. atgatgttat
U20986 Plasmodium reichenowi rhoptry associated protein 1 (RAP-1) gene, complete cds. atgagtttct
M75444 X61376 Plasmodium reichenowi sporozoite antigen gene, complete cds. atgaatattc
M58317 J05631 Plasmodium knowlesi merozoite recptor PK66 mRNA, complete cds. acaacacaatatgaataaaa
K00822 Plasmodium knowlesi circumsporozoite protein gene, complete cds, and 5' and 3' flanking regions. caagaacaagatgaagaact
M61097 M37854 P.knowlesi 66-kDa protein (PK66) gene, complete cds. acaacacaatatgaataaaa
M11031 Plasmodium knowlesi circumsporozoite antigen (CS) gene, complete cds. caagaacaagatgaagaact
tttttggatatgaattttt
L05069 Plasmodium simium circumsporozoite protein type 2 (CS) gene, complete cds. atgaagaact
U09765 Plasmodium simiovale circumsporozoite protein gene, complete cds. atgaagaact
M96886 Plasmodium gallinaceum ookinete stage-specific surface antigen (pgs28) gene, complete cds. atttaaaaaaatgaaaattc

M90978 Plasmodium cynomolgi heat shock protein 70 (hsp70) gene, complete cds. agttttcaaaatggccagcg
M15102 P.cynomolgi (strains Mulligan and NIH) circumsporozoite (CS) gene, complete cds. caagaacaagatgaagaact
M29898 P.fragile apical membrane antigen 1 (AMA1/AG352) gene, complete cds. acaacacaatatgaataaaa

L42007 Toxoplasma gondii protein serine/threonine kinase (PK1) gene, complete cds.
tcgctacttggtacgcgggcta cgccgtaactccgtgcaagacaactgtctgttttgagtgcttccccgcagtaattggagg tgccgaaaatatgttgatac
catgaaagctgtccaccctgaa aaccatgggtttccttaaaactctctctttttgtgtcatcgacaacaaagtggcgaaatg
cgagatccaggtaacgtattga ccacgaatctgcaccctgtgacacgttgaagcgcaattttcactggacaggcactgagag
aaggaagatggtcagtctattc ctatccagatctggatgccggcttccactttgtcatcttcgctttggcagaagcccggca
tcgtgcaccggtaagaggtctt acacaaggcatgctgtgcctggagttccttttttgtgcacaccgtcaacagggatctgaag
tgggcacttggtaaggtacagt gcaaaataccgcttctactttgtgtatttctgagccgtcgtttgctacaggagtgattct
gttcagcggcgtgaggcgtcct tgccgccatttttcttgaatagcaatttctttttggggttttggtggacaggtgcgaacga
ggcagactaggtggctgttcat aggagaaagtaccacatggtagaagtagatccttttttgtgtcttcacagtgtcaggtac
AJ001085 Toxoplasma gondii crk2 gene.
ccttcgacaggtatgtcttacg ctcacggcaaccgcattttgcttctggatctgcggcttttctggactcagcgattcgaga cttccggaaaatggagaagt
gtctgcgacggtaggtttcttt tccggattccgacgagagttgagttccactgtgggcaatttctgcgtcaggggtttgga
gactcgcacggtgcgtggggaa gaacgcgtcaggagtcgacttcttcctcgactcttctgctgtccactcagcgcctttggc
attttcgcaggtgagcttcagt tctaaaatcgtcgcggcattgcgcttttccgcaacttttttctctgcgtagaaatggtgaa
ggatcaagtggtaagagtttcg ggactcaaacttttcagacaccgcattgtgtgtgggtgtattggtgtcaggtgccaaagc
U62660 Toxoplasma gondii micronemal protein MIC2, complete cds.
gatacttcaggtaaggtgacga cccatgttgtcgctgttctttgtgggatatgtgtgtgctcgttcctgaaggcgttgaaga tgttttcacaatgagactcc
gcttgtggaggtaaggcacttg gacagtgtggttaatcgacctttgctttttgtgtgtggattgttgccagcttcctgcac
actatttgaggtaaggttatca ttagtatccgacacgatgagcctccgggtgtctgtttatctgctcttcagctccagtgtc
U81497 Toxoplasma gondii hydroxymethyldihydropterin pyrophosphokinase-dihydropteroate synthase (pppk-dhps) gene, complete cds.
agcggactgggtaagcctagcg ggcttattaccttgtagcgagcgtgagctgaggcatcccgctgccctcaggtgaacgacc cctccagaagatgagggaag
ggttatcaaggtaggcgtggaa gattgcccgttccatcacttttgaacacacttttgtgcctgcgtctccagactacgacga
gcgtcggcaggtagctgatgtt cttttggtagaaactggtgtcgtttccgactttggcgctatgttggataggtggcggatt
ccagaaaaaggtatgtcacaca cacggttgttggaccccaaaaggcgttgatttgctgcgttcatttcacagattggtcggt
tgatgtgaaggtaagcgcactg cgtggacaatgatagttatgggcctcgattttctatggctttgtgtccaggacactcggc
Z71786 T.gondii MIC1 gene.
tatggagaaggtacgtttccag cagcgacgacacgtgaactggctctctcgattctgtcggtgtttttcagcgtcgcattc atttgcaaagatgggccagg
ttacgagaaggtgaggaagaac ggaaggaatgttaggttggcattgtttgctgtcgtcgaatagctgtttaggccagcttgc
ggacgacgaggtaagttggaaa cgatggctcatccgttccccgttttcttgccggcaccccgttacctgcagggacgtcgac
Z48750 T.gondii hsp30/bag1 gene.
ctctcaacaggtacagttgaaa tcttgcggtaaacaccgatggaaattgcgatttgatgcgtggatatgcagctagactatt ttcaaaaaagatggcgccgt
ccctcccccggtaagactgaca cgccttgagtaggagggatggtcgtcccacagatgtgtgcatctgtgcagttttggggtg
cgttatcaaaggtaaatcgtct ttttttttcgtgcatgacgacggaatcgcgttgctccgtgctgccctggcaggagagaaga
Y09782 T.gondii H4 gene.
actcaatctggtgagcatggcc caggtccacccagacagttgtgtaagtgggcgtttctgtgtaaacttcagaaatgaaaga ttgaggctgtatgatttggg
tttcagcaaggtgagtcagatc ttttgtcttcagcaattttcatttttttcccattactcgattgttgcacagatggcgaaat
tctcgccgcaggtaaggttttc tcaggtaggaacaaaccgtaaaaaatcaattctgtctatctgtttgtcaggtggagtcgt

L08489 Toxoplasma gondii dihydrofolate reductase-thymidylate synthase gene, complete cds and complete genomic sequence.

cctcttccctgtgagcacacac acctccgttacgctttttttctggctttcttcttcgtctctgtttatcagcaaagaagaa gtctgggaagatgcagaaac
tttgtcgtgggtatgttgtcct gtccagactatgtcgcctgtttccccacccttctcggcttgtgctttcaggaggagcggg
tgcggaaccggtaagaggcaac acagggacgcctaccggagccccgttttctgcctttgtcgactcttgcagagcaacgcaa
gaccgaacgggtaacggcgact ttctgtgcatgtcttctgctcctcgtccttctctctttttccctgtttaggcgttggtgt
tctggaaagggtaagggcgtct aaaggagagtggaaatcattgaacagtgtcggtcgtctgttgtttcgcagggtcctcgaa
gggcgtgaaggcaagtctacgt agattctcatgcctgttgatcgccgtggtgttggtgcctgctcgtcttagatctgggaca
gactacacaggtatatgccggc tggagaagaaagacgtccagcaattcgtttctcacatgtgttctttgcagggcagggcgt
aatcctgcaggtgagccttgcg tgttttctctcaccgtgcacatttctccacgctgttgcttttgcgttcagcgctggacga
acgcatcaaggtgagcacaagt tggcccaaagagctcgcatggtgccgcgtttctatggattcgtttaccaggaaatcgacg

U09029 Toxoplasma gondii 65 kDa cyst matrix protein gene, complete cds.

atcacaccgggtaaggtcacga actcgtattcactgaagtgatcgctattttgtttgcgttactgtgcacagcaacatggat caccggcaacatggattgcg

tgagccaaaggtgagtcaagcg ttgttggtaagtgttttgcaaggcgttgtgttggcacaactggcgtgcagggtgccagag

M99392 Toxoplasma gondii antigen p28 gene, complete cds.

gagcagttgggtaagttggcaa tgtagtgtcttgagactcggtgccatgtagaattttgtgtctgcctgcagaggaaaaggt tcgcctcacaatgttcgccg

U10247 Toxoplasma gondii RH hypoxanthine-xanthine-guanine phosphoribosyl transferase gene, complete cds.

atcgcgggaggtaaagttccgc catgttgctgtaccagtggtgtttttgcgtggcttgttgtttccgtgcagaggttgctta
ccattgaagagtgagtcttggc accctcttgaaatagttcagtgttacacatgtgtataacattgttttcagatcccggtcg gtgaagcaaaatggcgtcca
cacttcaagagtaaacaccaca tgatcagaaaaaaataagataatggctcctgttctgagtgtttacctgcagctacggcaag
tgagaagttggtacgccacttg cactgtgtttggttcactgaatggtctgctgagtgcgccctctttctcaggcgtatgaca
ccatacagaagtaagcgcgcac tgttctcattgatttttgtgttgcatccgtttgtgccgatttttcacaggtacagtggt

U10246 Toxoplasma gondii RH uracil phosphoribosyl transferase gene, complete cds.

tcatcacgaggtaatccttcaa tcggaaaagttcattgtgaacgttccccttgcgtgtcatgactttatcaggtttcccaat ggcttctaaaatggcgcagg
atcattcgtggtgagtggtgac tacgaaattctcacgtgcgttttccgcctcctttttcttctgttgcttcagataaagaaac
gaagaattcggtaggaacttga gtatgttttccatctctgccccgtctatatcccctgtttgtctccttcagtcttctacgc
gctttgaacggtgagtctcaag ctagactttcaactgacgcggtgcctacgttcttctacattttactgcagaactgccgtt
ggagcctaaggtacggacttgt cttgtccgctgcacctgcttcctccgcccttgatgtttgctctgttccagctgatctacg
tcactgctgcgtaagcaaaatg gtagctttccatgtttctgcctcttgtttgctttcctcggcggtctgcagtgttgacatc
tcggtgaccggtaagaaataac cctaatgcagccagccctttatggaattgtgtgccctgctgtcttgtcaggtactttgga

M20025 T.gondii beta-tubulin gene, complete cds.

aattgatccggtgagggagaac atttcaatttgttctgcgatgtgtgtgtgtgtggtttcctctcttcagaccggtacct atcttccaaaatgagagaaa
tcttgcactggtaagttgtaca gaagtctgtgtgggcatgaatactgtttttttgcctccttggtcctataggtacacgggt
gcagtaccaggtacgctttcgc tttcgtaatcttcatcaatacttcgaatgtcttatgtgtgttctcttcaggacgccaccg

M20024 T.gondii alpha-tubulin gene, complete cds.

acgcctgctggtaagcggacac ccagttcgctgtaaccattgtcttttgtgtgttttctctccttctttttagggagctcttc tttcgacaaaatgagagagg
tggcaagcacgtgagtattgtt ggttgtctcttgtattttgtcaaacgtgttgactgcgtacgctgactcaggtgccccgat

J04018 T.gondii 28 kd antigen gene, complete cds.

gagcagttgggtaagttggcaa tgtagtgtcttgagactcggtgccatgtagaattttgtgtctgcctgcagaggaaaaggt ttgaaagagaatgttttccg

L07954 Eimeria tenella homologue of E.coli transhydrogenase (7B2) gene,complete cds.

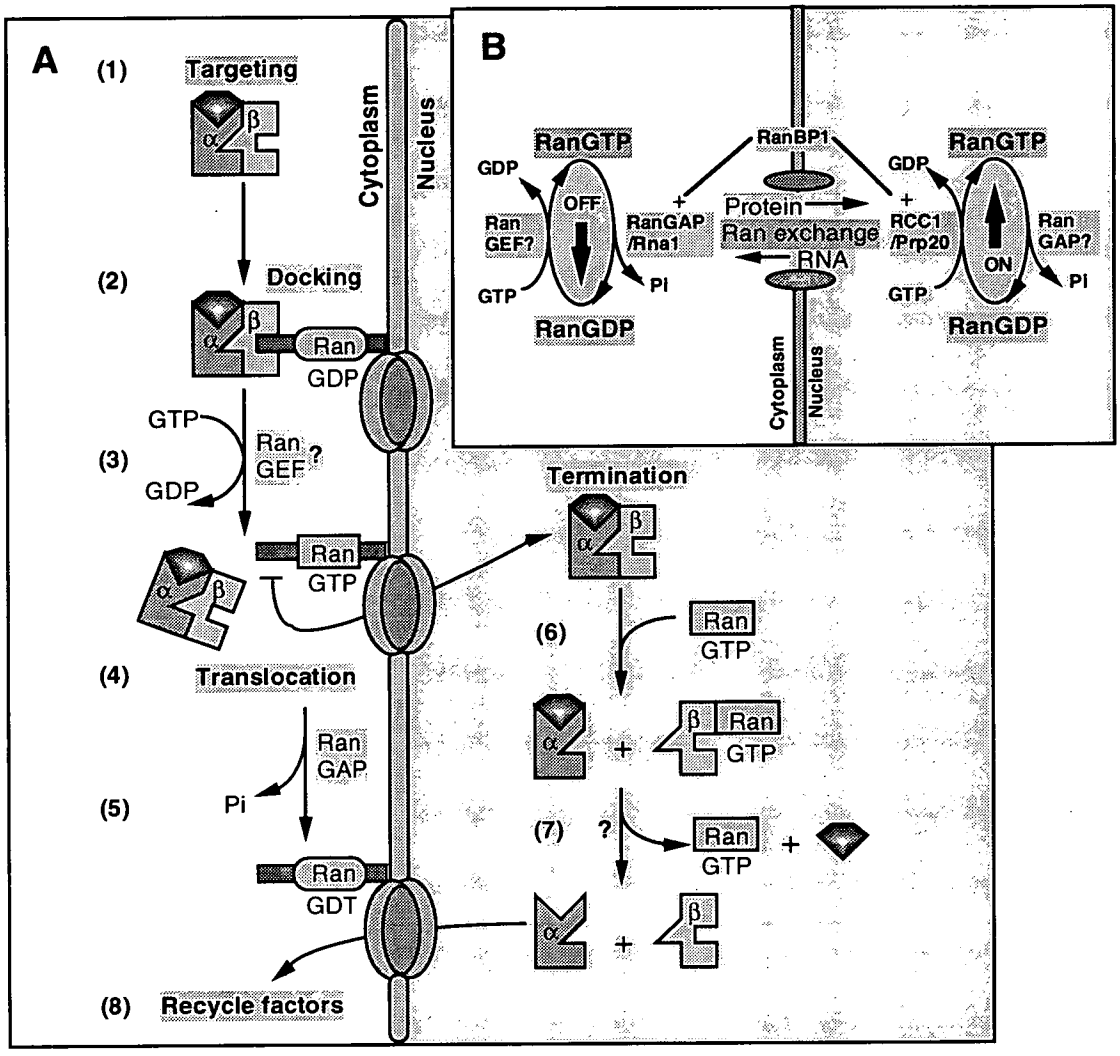
```
          acattcttgggtatgtcgtttt         atacgtgccttgcgcgaccatttcgatactcattggctttttaacagtcttgtgggc      gttctcaaatatgccttctc
M21088    E.tenella major sporozoite surface antigen gene, exons 1 to 4, complete cds.
          ccaacagcaggtgggctttttcc       aaaacgaggtgcattcacctttgcatgttgtgtgcggaaatttttatcagttacgctgga     tttgctcaaaatggctcgtc
          agtcttaggagtaagccgtcca         gaacaaggaactacactgtccttgaatttttaatcttttgttacgtacagggcggaaggt
          aaccattcgagtgagagtcagc         agcgcaatcttctaatacttgtttgtaatgtttgtaatgtttgcgtgcagcgacgagcaa
E00755    E.tenella DNA sequence coding for A4 antigen.
          ccaacagcaggtgggctttttcc       aaacgaggtgcattcaccntttgcatgttgtgtgcggaaatttttatcagttacgctgga     tttgctcaaaatggctcgtc
          atgctgtgggagacacatttgt         gaacaaggaactacactgtccttgaatttttaatcttttgttacgtacagggcggaaggt
          aaccattcgagtgagagtcagc         actaatacttgtttgtaatgtttgtaatgtttgcgtgcagcgacgagcaatggaagaaaa
U19609    Eimeria tenella beta-tubulin gene, complete cds.
          catagccccggtgaggagcccc         tgcaggcctgtgcgtcatcctgtttgctttccctttgaaatgtttcgcagaccggcacgt     tttcgcgaaaatgagagaaa
          tcctccactggtgagtttccta         gcacgcactttcaataaactccgaatggctgtgcacttcgctccgttcaggtacactgga
          gcagtaccaggtgggtttggtg         ttgcgggtagaaagaatggcttttgcctcgctttcattccattttctcaggatgcaaccg
Y12615    C.parvum beta tubulin gene.
          gatcgaccctgtaagttttgaa         aataattgacttttttttttttctttttttttctttgaaatatgaatttagactggtactg     gaaaaagagaatgagagaaa
X98824    T.parva gene encoding cdc2-related kinase.
          cctagcacaggttagt                                       cccattattctcacttttcttagccattcggga     caaaaacgaaatgaggcgat
          gcttgtgatggtaccatttgac         tctataatattgttatccaatcaaaatttgttgatataataatttgttaggtggtttgga
          gactagcaaggtactgttgaca         ttatttacattaatcttaactacatataagtattgtatataacattttagggcttttgca
          atttttgcaggtaaactcataa         ataaaattaaaaaaaaatttgcataaaattaataaaaatttaatgtttagagatgataaa
          gtcctcaattgtaagt                                     atttactgtttaatatcacattttaggtaccaaagc
U40703    Theileria parva sporozoite surface protein p67 gene, complete cds.
          agagaattaggtttg                                      ttttttgagtatatgggttttaggtaatagggc     acactgaacgatgcaaataa
M67476    Theileria parva cathepsin L-like cysteine protease gene, partial cds and p67 and ORF1 genes, complete cds.
          agagaatcaggtttgtttttttt                                    gagtatatgggttttaggcaatagggc     acactgaacgatgcaaataa
M37791    Theileria parva cysteine protease gene, complete cds.
          gaacaaaagggttagtt                                    acacacacttaatattattttataggtgatgaacc      tgtattaaccatggttagtt
U70130    Babesia bovis iron-dependent superoxide dismutase (Fe-SOD) gene, complete cds.
          aaactaaacagtaagtatta                                    ttacccatgaatataatggttcaggtctcatcaa     tgtatctacaatggccttca
          actatcgaaggttgttatag                                  gattccatgttatccatgatgccacagagttgatact
          tacattaacagtaggttgta                                     tagctcctaaatatactccacaggctggtggaa
U34076    Babesia bovis ribosomal protein L35 (bL35) gene, complete cds, and nucleoside monophosphate kinase (bNMK) gene, nuclear gene
           encoding putative mitochondrial protein, complete cds.
          tgattttctcctgtacgtaacgt      taataacgtaaaaaacaaacaaataaaaacgcaaagaagatatacttaccatttttaaa
          gtttcatcaggttcgttt                                  tgtgcattggttgttttacgcggtattaaggtgatatact      atctgttgatgattagca
          gtatagtcaggtaccatt                                    agtagccatcctgttctatatgctacagcgctgagttg
          aggccgaagggtgtgttt                                      ccatttgccactttaatttttatagcactgtgctc
          cttgctgaaggtttgtct                                    accagcaaccaccatgacttcgtttcagtgcaacggtt
          tctagaaggtgttagttt                                      tttcttagacgatttagacttaatagataccaacga
```

Appendix H

The Ran GTPase cycle.

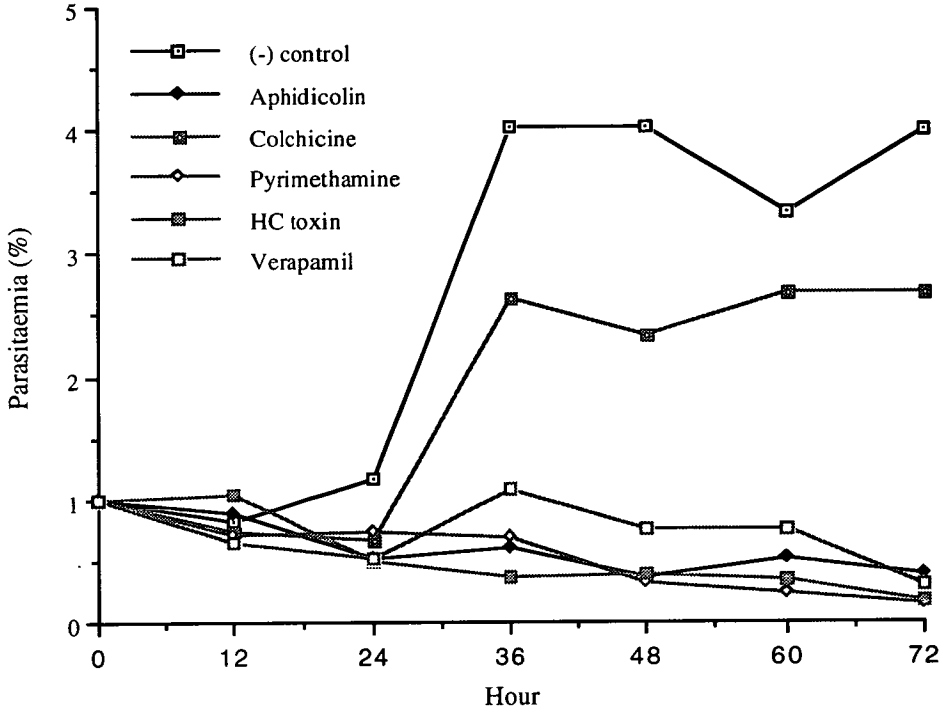
A Proposed role of Ran in the import of nuclear localisation sequence (NLS) tagged NLS cargo to nucleus. (1) The targeting complex forms when the NLS receptor a binds NLS cargo and docking factor b. (2) Docking occurs at the filamentous sites that protrude from the pore complex. Ran-GDP, presumably initially bound to the small Ran-GDP binding factor NTF2, docks independently. (3) Transfer to the translocation channel is triggered when an unidentified RanGEF converts Ran-GDP to Ran-GTP. (4) The nuclear pore complex catalyzes translocation of targeting complex. (5) Ran-GTP is recycled to Ran-GDP by docked ubiquitin-RanGAP. (6) Ran-GTP disrupts the targeting complex by binding to a site on b that overlaps with a binding site. (7) NLS cargo dissociates from a, and Ran-GTP may dissociate from b. (8) a and b factors are recycled to the cytoplam.

B The Ran translocation switch is off in the cytoplasm and on in the nucleus. Ran-GTP promotes NLS- and NES-directed translocation. Cytoplasmic Ran is enriched in Ran-GDP (OFF) by an active RanGAP and nuclear pools are in Ran-GTP (ON) by an active RanGEF. RanBP1 promotes the contrary activities of these two factors. Direct linkage of nuclear and cytoplasmic pools of Ran occurs through the nuclear pore complex by an unknown shuttling mechanism. NLS, nuclear localisation signal; NES, nuclear export signal; P_i, inorganic phosphate. (Modified and adapted from Goldfarb, 1997).


A
(1)
Targeting
Cytoplasm
Nucleus
(2)
Docking
Ran
GDP
GTP
(3)
Ran
GEF ?
GDP
Ran
GTP
(4)
Translocation
Ran
GAP
Pi
(5)
Ran
GDT
(8)
Recycle factors
Termination
(6)
Ran
GTP
(7)
?
Ran
GTP
B
RanGTP
GDP
OFF
Ran
GEF?
GTP
RanGDP
Pl
+
RanGAP
/Rna1
RanBP1
Protein
Ran exchange
RNA
Cytoplasm
Nucleus
RanGTP
GDP
+
RCC1
/Prp20
GTP
ON
Ran
GAP?
Pl
RanGDP

Appendix I

Survival of *P. faciparum* parasites in different cytotoxic drugs.



On the vertical axis are plotted the parasitaemia (%) treated with 2.5 μg/ml of aphidicolin, 0.1 μg/ml of colchicine, 2.5 ng/ml of pyrimethamine, 100 ng/ml of HC toxin, 30 μM of verapmil and without any treatment. The drugs were given in the culture from ring stage of *P. falciparum* parasites for 72 hr contineously.

Appendix J

Two publications produced from the course of this PhD study:

Ji and Arnot, 1997. Mol. Biochem. Parasitol. 88, 151-162.
Ji *et al*, 1998. Mol. Biochem. Parasitol. In press.

ELSEVIER

Molecular and Biochemical Parasitology 88 (1997) 151–162

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

A *Plasmodium falciparum* homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodelling for transcription[1]

Dar-Der Ji, David E. Arnot *

Institute of Cell, Animal and Population Biology, Ashworth Laboratory, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

Received 24 March 1997; received in revised form 27 May 1997; accepted 6 June 1997

Abstract

A *Plasmodium falciparum* homologue of one of the components of a chromatin-remodelling complex which controls binding of transcription factors to nucleosome core particles has been cloned and characterised. The gene encodes 1422 amino acids with an estimated molecular mass of 167 kDa. The protein, SNF2L, shares 60% amino acid identity in its conserved DNA-dependent ATPase domain with yeast transcription factors originally identified by characterising mating type switch mutants. It also contains sequences related to the so-called SWI3, ADA2, *N*-CoR and TFIIIB B″ or SANT DNA binding domains which are characteristic of these transcriptional activation factors. The *SNF2L* gene has two short introns in the 3′ region of the coding sequence of the gene and is transcribed into a single ~6.5 kb messenger RNA species which is present throughout the asexual stages of the cell cycle. Southern blotting and pulsed field gel electrophoresis experiments show that *SNF2L* is a single copy gene, located on *P. falciparum* chromosome 11. © 1997 Elsevier Science B.V.

Keywords: *P.falciparum* ATPase; Chromatin remodelling; Transcription factors; SNF2L; SWI/SNF

1. Introduction

Plasmodium falciparum gene expression is not understood although a number of interesting observations have been made. *Plasmodium* mRNA can be translated in vitro and is thus inferred to be conventionally capped [1,2]. These mRNAs frequently possess long, A + T rich, 5′ UTR's

Abbreviations: CRC, chromatin remodelling complex; SNF, sucrose non-fermenting; SWI, mating type switch; UTR, untranslated region.

* Corresponding author. Tel.: +44 131 6508655; fax: +44 131 6508655; e-mail: dea@holyrood.ed.ac.uk

[1] *Note*: Nucleotide sequence data reported in this paper are available in the EMBL, GenBank™ and DDJB databases under the accession number AF003086.

PII S0166-6851(97)00089-3

whose function is unclear [3–5]. Transcript mapping has indicated heterogeneous initiation at the 5′ ends of some *P. falciparum* messages [3,6] although other studies have detected single start sites in other genes [7,8]. Features characteristic of protozoan promotors such as repeated sequences upstream of the start codon have been reported for a gene transcribed during gametocytogenesis [9]. The 3′ UTR's of *P. falciparum* mRNAs are also A + T rich and multiple potential polyadenylation sites occur downstream of the stop codon [8,10,11]. As more genes are sequenced it has become clear that introns are not uncommon and probably occur in a majority of loci, although they are usually small (< 150 bp) and tend to occur near the 5′ or 3′ ends of a gene, leaving the central transcribed region uninterrupted. A similar size, frequency and distribution of introns has previously been noted in yeast [12]. *Trans*-spliced leader sequences have not been found although differential splicing has been detected in one *P. falciparum* gene with eight introns [13].

Promotor characterisation in *P. falciparum* is difficult as the A + T rich base composition of intergenic regions (~ 87%) means that the 5′ UTRs abound with sequences resembling the TATA motif of most eukaryotic transcription promoters (consensus sequence TATA/TA). However, the parasite has a TATA-binding protein and although it is somewhat divergent from other family members, it retains sequences essential for recognition of DNA [14].

Eukaryotic transcription initiation involves decondensation of DNA from chromatin which inhibits transcription, replication, recombination and repair by blocking access to the DNA [15,16]. Two types of enzymes are known to rearrange nucleosomes. Type A histone acetyltransferases increase the affinity of transcription factors for nucleosomal DNA sites by altering DNA binding via increasing the level of lysine acetylation in histone proteins [17]. The second class of known chromatin remodellers/transcriptional activators, are the SWI/SNF proteins, which catalyse ATP hydrolysis in a DNA dependent fashion [18–21]. We have identified a novel member of this important family of gene expression regulators in *P. falciparum* and characterised its structure and expression pattern.

2. Materials and methods

2.1. Parasites.

P. falciparum lines 3D7A and K1 were grown under standard conditions [22]. Synchronised cultures were obtained using the method of Lambros and Vanderberg [23] and gametocyte enriched cultures were prepared following the methods described by Carter et al. [24].

2.2. PCR cloning of the amplified fragments and cDNA library screening

Two stretches of amino acids (KWIHCF and FLNKID) that encode GTP binding residues and are conserved among members of the Gα protein family were used to design primers biased towards *P. falciparum* codon usage [25]. The primers were named as GP1 (5′AA(A/G)TGGAT(A/C/T)CA(C/T)TG(C/T)TT-3′) and GP3 (5′(A/G)TC(A/G/T)AT(C/T)TT(A/G)TT(T/C)AA(A/G)AA-3′). A total of 50ng of total asexual RNA was reverse transcribed in 40 μl of a solution containing 50 mM Tris–HCl, pH 8.3, 75 mM KCl, 10 mM dithiothreitol, 3 mM $MgCl_2$ 1 mM dNTPs, 20 U RNasin (Promega), 2.5 μM of oligo dT_{15} primer, 200 U M-MLV reverse transcriptase (BRL) at 37°C for 2 h. A total of 40 cycles of PCR were carried out with 2 μl of the cDNA product and the GP1 and GP3 primers in 50 μl of 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 10 mM dithiothreitol, 2.5 mM $MgCl_2$, 75 μM dNTPs, 1 μM primer GP1, 1 μM primer GP3, 2 U Taq polymerase (Perkin-Elmer). PCR cycling was at 94°C for 1 min, 40°C for 2 min, 72°C for 1 min, with a final 72°C extension for 10 min. The 198 bp fragment obtained was gel-purified and subcloned into the pCRII (Invitrogen) vector for sequencing. The PCR fragment was radiolabelled and used to screen a *P. falciparum* asexual stage cDNA library constructed in the plasmid vector pJFE14. Ten positive clones were selected from a screening of 50 000 colonies.

2.3. DNA isolation and southern blotting

Genomic DNA was purified from saponin-lysed cultures as described [5]. For Southern blots, genomic DNA was digested and fractionated on 0.8% agarose gels then transferred to Hybond N^+ nylon membranes (Amersham). Membranes were prehybridised for 4 h at 65°C in 1% sodium dodecyl sulphate/1 M NaCl/10% dextran sulfate/ 5 × Denhardt's solution/100 mg ml^{-1} denatured salmon sperm DNA. Hybridisation was carried out in the same buffer with radiolabelled probe at 56°C overnight. Following hybridisation, the blot was washed under conditions of increasing stringency, ending with 0.1 × SSC at room temperature.

2.4. Pulsed field gel electrophoresis

P. falciparum chromosomal DNA blocks were prepared as described [26]. Pulsed field gel electrophoresis (PFGE) was carried out using a Bio-Rad™ CHEF DRII apparatus and 1% agarose gels. The chamber buffer was 0.5 × TBE circulated at 14°C. Gels were ramped from 90 300 s pulse times at a constant 95 V for 20 h then switched to a ramp from 300–720 s pulse time at a constant 90 V for 30 h.

2.5. RNA isolation, Northern blots and intron detection

Total RNA was prepared from cultures using an RNA Isolator Kit (Genosys) after removing haemoglobin by saponin lysis of erythrocytes. For Northern blotting, 2.5 μg of total RNA was separated in 1.2% agarose gels containing 0.74% formaldehyde and transferred to Hybond N^+ membranes (Amersham) by alkaline transfer. Intron-exon detection was performed by PCR and RT-PCR amplification from genomic DNA, cDNA and the SNF2L clone [27]. PCR and RT-PCR was carried out as described above except that Taq polymerase was added to a 'hot-start' reaction at 68°C.

2.6. Computing analysis

Sequence analysis was performed using the GCG software package [28] or through World Wide Web browsers. The cDNA sequences and translated amino acid sequences of the SNF2L gene were compared with the EMBL and GenBank database using the BLAST algorithm [29] as implemented in the GCG package. Local alignments of multiple sequences were constructed using these algorithms under varying parameters on individual sequence pairs, followed by manual optimisation of the alignments. The RNASPL [30] and POLYAH [31] programs were used to predict intron-exon junctions and potential poly(A) sites. PSORT version 6.3 (WWW) was used to do the prediction of protein localisation sites from cDNA sequences. Statistical analysis of protein sequences was carried out using the SAPS program [31].

3. Results

3.1. Isolation and characterisation of the SNF2L gene cDNA clones

The *SNF2L* gene was identified during attempts to clone *P. falciparum* G-protein genes. The original PCR reactions were performed by using primers (GP1 and GP3) deduced from conserved GTP-binding domains of Gα proteins. A 198 bp fragment was amplified and its sequence revealed similarities to the conserved regions of the SNF2/SWI2 transcription activation factors. In attempts to clone members of the G-protein gene superfamily, we have amplified several sequences with homologies to ATPase domains. Given the importance of SNF2/SWI2 proteins in transcription and our interest in gene expression in *Plasmodium*, we analysed the gene further by using the amplified fragment to screen a *P. falciparum* plasmid cDNA library. Ten overlapping cDNA clones were isolated.

3.2. The structure of the SNF2L gene

The largest cDNA clone, *SNF2L-3*, has 519 nucleotides of 5'-UTR (86.5% A + T) upstream of a putative initiation codon followed by a 4,269 nucleotide open reading frame and 469 nucleotides of 3'UTR (87% A + T). The exact position of the translation initiation codon is uncertain as there are four in-frame methionine codons in the region 520–595. Alignment with other members of the gene family does not resolve this problem since there is heterogeneity in the amino terminal sequences of the SNF2-like proteins. The most likely initiation site is the first of these in-frame ATG codons since *P. falciparum* 5' UTR regions do not maintain in-frame open reading frames. Sequences around this codon are not in particularly good agreement with the 'Kozak rules' for eukaryotic translation initiation sequences, but *P. falciparum* genes in general conform poorly to this translational initiation consensus which is derived from higher eukaryotic genes [32]. The nucleotide and translated amino acid sequence of this protein, based on assigning the first in-frame ATG as the initiation codon, is shown in Fig. 1A.

The amino acid sequence showed that the highest levels of sequence similarity in BLAST searches were with domains conserved between members of the SNF2 and SNF2-like (SNF2L) family of transcription factors [33]. These proteins contain seven conserved regions which constitute the catalytic site of a DNA-dependent ATPase which also has a DNA unwinding activity. These seven regions are boxed in Fig. 1. The alignment of the *P. falciparum* protein, which we have named SNF2L, with the *Drosophila* [34], human [35], *C. elegans* (unpublished, GeneBank ref. P41877) and yeast (unpublished, GeneBank ref. P31844) SNF2-like proteins is shown in Fig. 2A.

Motif I, Ia and II match the consensus sequence of the bipartite nucleotide triphosphate binding motif found in both ATP and GTP-binding proteins [36]. The DNA-dependent ATPase/DNA helicase domain of all the SNF2-like proteins, including *P. falciparum* sequence, is shorter than that of the corresponding domains of SNF2 proteins proper (the human and *Drosophila brm* proteins and the yeast SNF2/SWI2 protein). As in all members of the SNF2L subfamily, *P. falciparum* SNF2L has the DEAH amino acid motif, whereas the SNF2 proteins have a DEGH motif. SNF2L proteins also lack the other conserved domains found in *brm* and SNF2/SWI2, the so-called 'bromodomain', a sequence of unknown function postulated to interact with other proteins in the transcriptional activation complex.

In addition to the regions involved in ATPase activity the *P. falciparum* protein also contains a SANT domain which is also characteristic of the other SNF2-like proteins, but not the SNF proteins (Fig. 2B). Again, the function of the SANT conserved sequence is unclear but it occurs in several proteins found in transcription complexes such as TFIIIB. Two regions of the *P. falciparum* SNF2L sequence are not found in other members of the family, although they are common in *Plasmodium* proteins. Eight tandem repeats of a NNHNDD sequence occur in the amino terminal domain and between the DNA-dependent ATPase domain and the SANT domain, the protein contains a region of 210 lysine rich, basic amino acids. Both of these insertions, which make the *P. falciparum* SNF2L substantially bigger than its higher eukaryotic homologues, are unique. Like most other *Plasmodium* genes, *P. falciparum* SNF2L is an asparagine (12.4%) and lysine (11.5%) rich protein.

GenBank homology searches further indicate that *P. falciparum* SNF2L is most similar to proteins of the SNF2L subfamily of the SNF2 transcription factors. *P. falciparum* SNF2L proteins are 37.6% identical and 59.8% similar to yeast YB95 SNF2L protein, 40.6% identical and 63.8% similar to the human SNF2L protein, 38.1% identical and 61.6% similar to *C. elegans* SNF2L protein and 40.4% identical and 62.5% similar to *Drosophila* ISWI protein. Fig. 2 compares the structure, sizes and conserved motifs of these proteins. The primary regions of homology, the seven consecutive motifs characteristic of DNA-dependent ATPase/DNA helicase and the C-terminal SANT domain were used to align the five sequences.

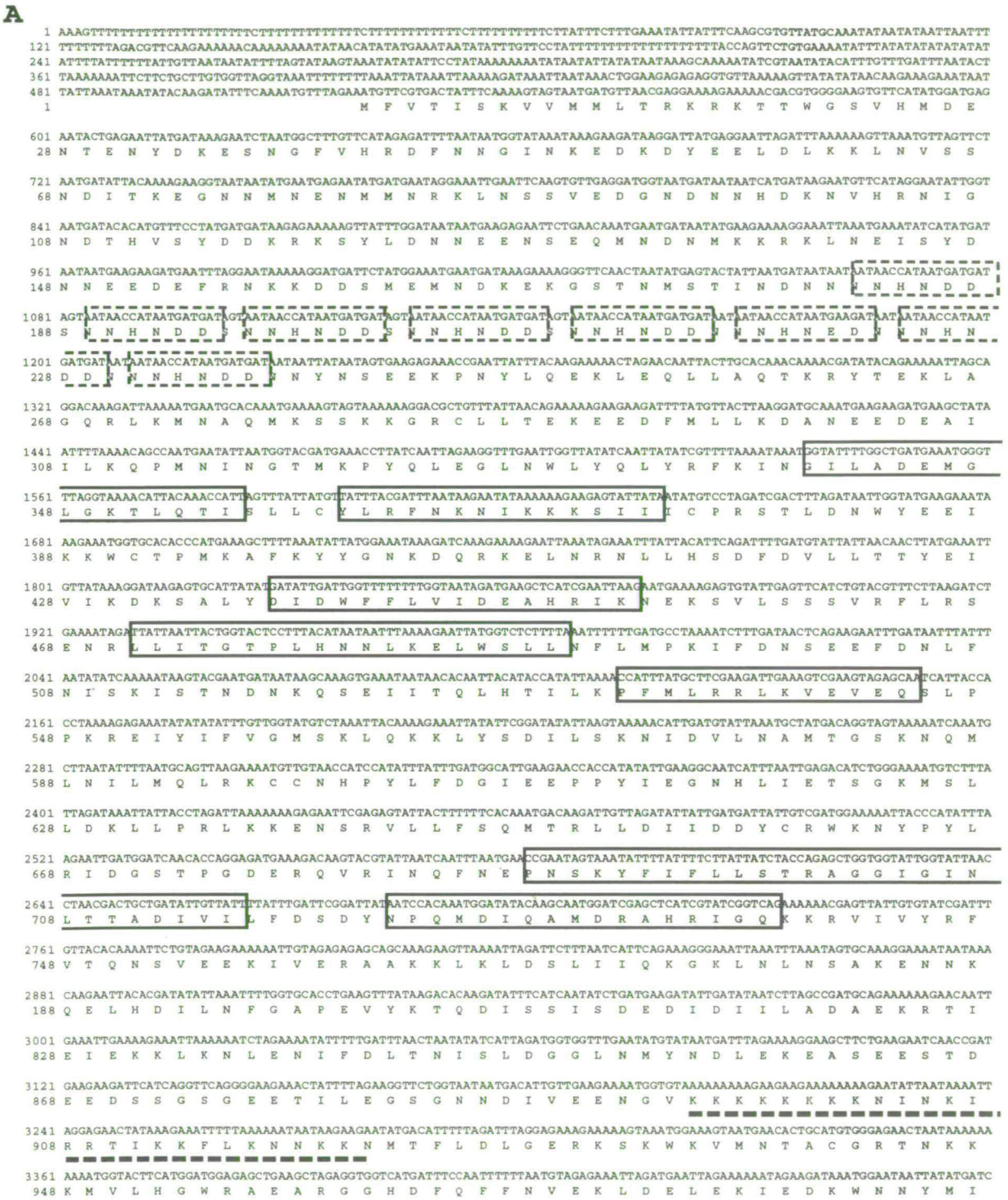
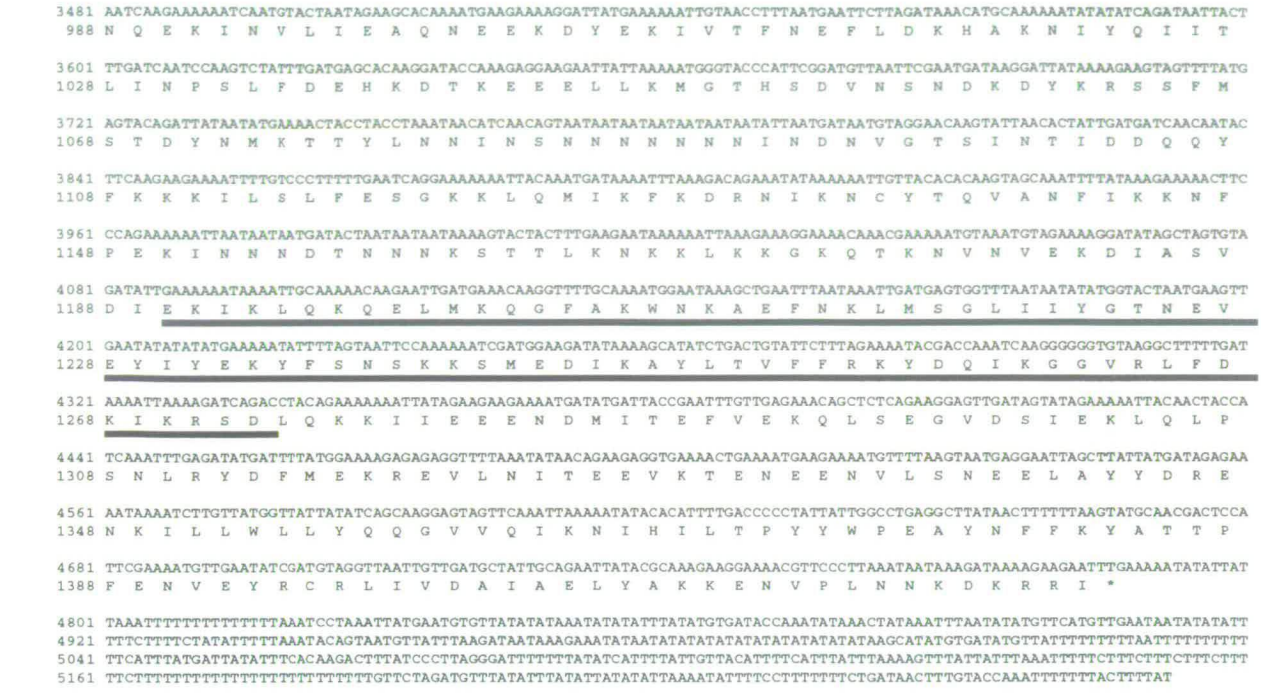


Fig. 1.



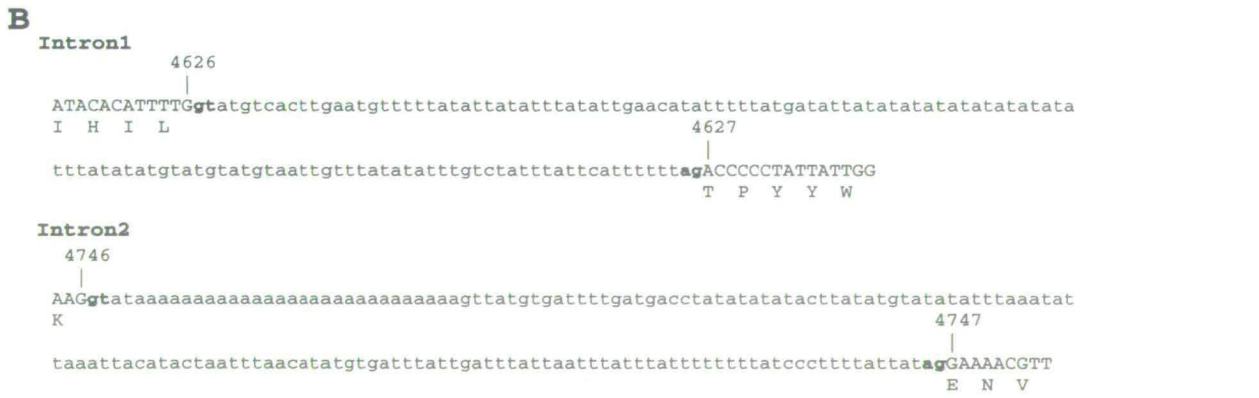


Fig. 1. (A) The complete sequence of the PfSNF2L cDNA clone and its encoded protein. The amino acids contained within the solid boxes are those of the conserved ATPase domains. The carboxy terminal sequences underlined by the solid bar represent the SANT domain characteristic of the SNF2-like proteins. The sequences highlighted with dashed line boxes are the asparagine-rich repeated hexapeptides unique to the *P. falciparum* gene. Sequences underlined by dashed lines indicate the position of charged residue clusters. The clone contains a single an open reading frame of 4269 nucleotides encoding a 1422 amino acid protein with an estimated 167 kDa molecular mass. The termination codon is marked by an asterisk and the GenBank accession number is AF003086. (B) The sequence and position of the two carboxyterminal introns of *SNF2L*. Intron1 is 132 bp and intron2 is 162 bp.

3.3. Genome localisation and intron mapping

Genomic DNA was digested with restriction enzymes and analysed by Southern blotting (Fig. 3A). This shows that SNF2L is a single copy sequence without closely related (>80% homology) sequences. Southern blotting of pulsed field gel separations of two cloned lines of *P. falciparum* maps the gene onto chromosome 11 (Fig. 3B). Discrepancies between the size of PCR prod-

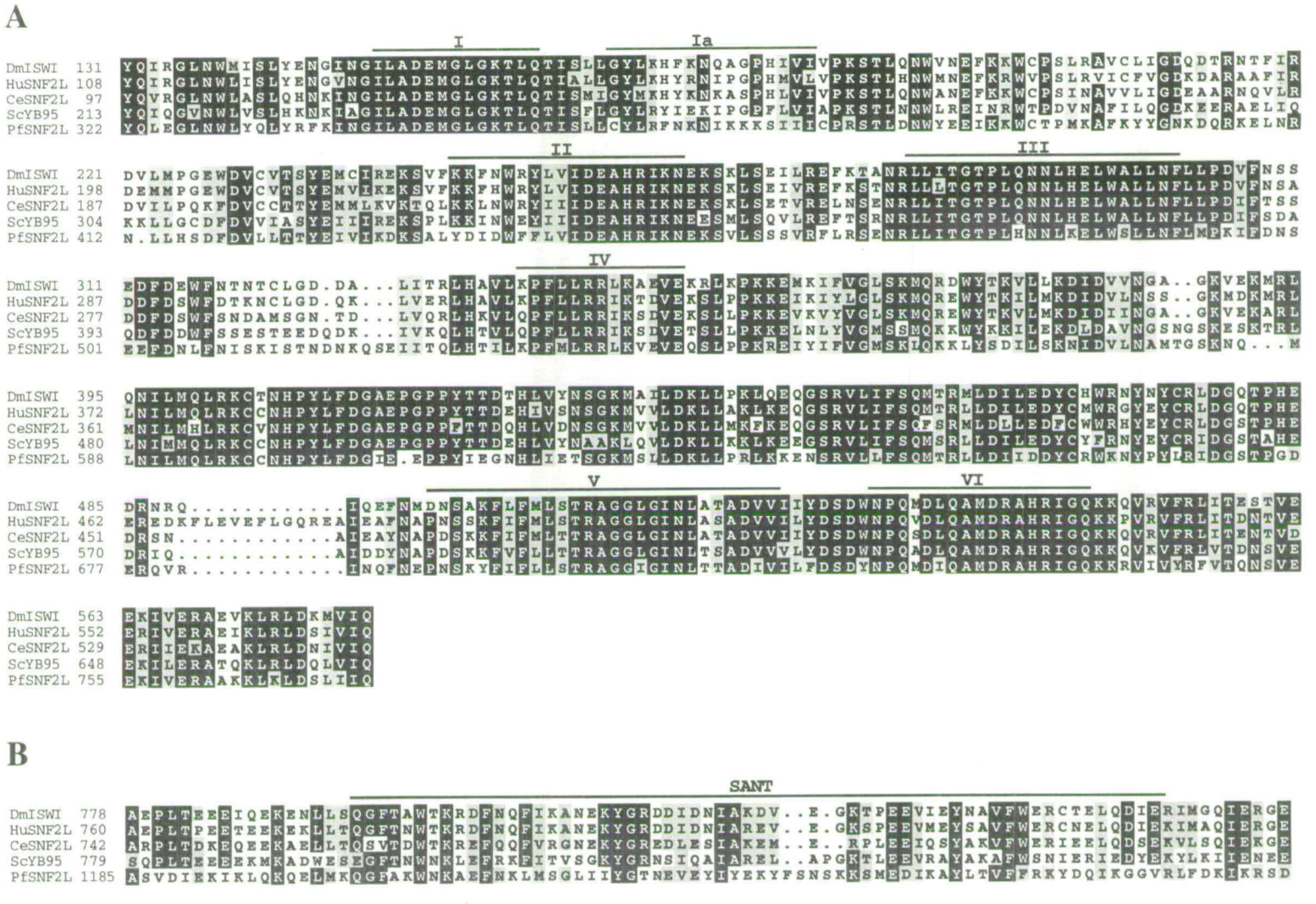


Fig. 2. (A) Alignment of the conserved sequences of the SNF2-like protein subfamily. ISWI is the *Drosophila* protein [34], HuSNF2L is the human protein [35], CeSNF2L is the *C. elegans* protein, (unpublished, GeneBank ref. P41877) and ScYB95 is the budding yeast protein, (unpublished, GeneBank ref. P31844). Regions highlighted in black indicate sequence identity, those highlighted in grey indicate conservative amino acid substitutions. The overlined position of the seven conserved motifs in the region encoding the framework of the DNA-dependent ATPase/DNA helicase domain is shown. (B) The evolutionary conservation of the SANT putative DNA binding domain.

ucts derived from genomic DNA and cDNA indicated that the gene has at least one intron. Differential splicing in *Plasmodium* has been noted [13] and since processing of transcription factor mRNA might regulate expression, the introns of this gene were located and sequenced. Using nine pairs of PCR primers and comparing genomic DNA, the cDNA clone and the RT-PCR product, the intron-exon boundaries of the SNF2L gene were mapped. The intron positions and sequences are shown in Fig. 1B.

3.4. *SNF2L transcription during development*

To determine mRNA size and the existence of any precursor or processed RNA species, total RNA was isolated from asynchronous asexual stage cultures of several *P. falciparum* strains and separated on formaldehyde gels for Northern blotting. The internal PCR fragment was used as the hybridisation probe. The SNF2L mRNA can be detected in all three laboratory lines as a single, strong mRNA band of 6–6.5 kb (Fig. 4). Faint hybridisation to smaller sized fragments varies between RNA preparations and probably represents degradation rather than processing. Given that our largest cDNA clone is 5.27 kb in size, the primary transcript must be significantly longer than the coding sequence. 3′ RACE (rapid amplification of cDNA ends) analysis using a primer at the 3′ end of the *SNF2L* gene and an oligo dT_{15} primer indicates that there is a longer 3′ UTR in the primary transcript (data not shown) and it is likely that this is also true at the 5′ end of the mRNA. Low amounts of hybridising message are present in gametocytes (data not shown), indicating that the *SNF2L* mRNA is present in all erythrocytic stages and in low amounts in the gametocytes.

4. Discussion

We have identified a *P. falciparum* homologue of the *SNF2L* genes which encode chromatin remodelling/transcription activation proteins. The protein encoded by this gene may be a component of a *Plasmodium* chromatin remodelling complex. This gene joins the TATA-binding protein gene [14] and two high-mobility group protein genes [37,38] as characterised loci encoding proteins involved in regulating *P. falciparum* transcription. *SNF2L* mRNA is present in low amounts throughout the asexual cycle although it is not known whether the protein itself is active in all stages. The *P. falciparum* SNF2L protein is related to the yeast SWI/SNF2P protein, the first of this class of proteins to be described, but it is more similar to a protein of a *Drosophila* nucleosome remodelling complex, ISWI [39]. Other higher eukaryotic genes of less well characterised function also fall in this group. These proteins form a distinct subfamily, the 'SNF2-like' proteins which lack the bromodomain motif of SNF2 proteins and instead possess a SANT type domain characteristic of several proteins involved in DNA binding and transcriptional regulation [40]. The *Drosophila* nucleosome remodelling factor (NURF) complex appears to be involved in chromatin activation, yet it is smaller ($\sim$ 500 kDa as opposed to 2 MDa), less complex (four as opposed to 11 major polypeptides) and much more abundant (10^5 as opposed to 10^3 copies per cell) than the SWI/SNF complex [39]. Recently, a new, 15-subunit chromatin remodelling complex (RSC, remodel the structure of chromatin) has been reported in yeast. This complex is ten times as abundant as the SWI/SNF complex and also contains a novel DNA dependent ATPase subunit [41]. The existence of multiple chromatin remodelling complexes may indicate that they regulate different promotors and/or control different as-pects of DNA metabolism such as repair or recombination. We do not know if the *P. falciparum* SNF2L forms part of a multi-protein complex or if it is essential for transcription.

Two substantial domains are unique to the *P. falciparum* SNF2L protein. These are the amino terminal tandem repeats of the ($8 \times$ NNHNDD) and the central 210 amino acid lysine rich basic domain. The role of neither of these regions is

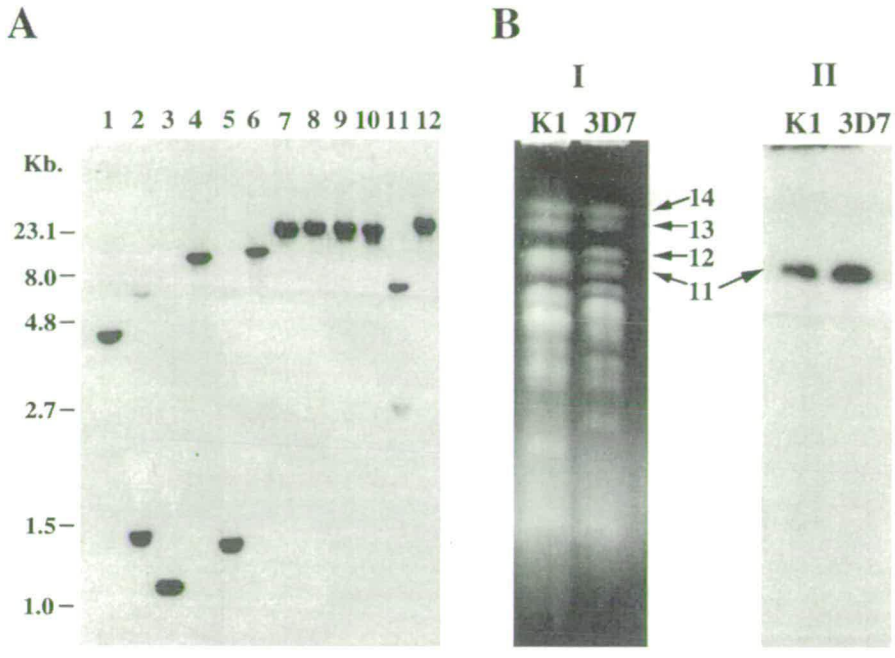


Fig. 3. Copy number and genomic localisation of the PfSNF2L gene. (A) Southern Blot of *P. falciparum* (clone 3D7A) DNA cut with the restriction enzymes *Bgl*II, *Cla*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Hpa*I, *Mlu*II, *Nar*I, *Pst*I, *Pvu*II, *Xba*I and *Xho*I (lanes 1–12 respectively) and probed with a labelled fragment of the cDNA clone. (BI) Ethidium stained pulsed field gel separation of *P. falciparum* chromosomes from the K1 and 3D7 lines. (BII) Southern blot of the same chromosome separation transferred to nylon membranes and probed with a labelled fragment of the cDNA clone.

understood, although insertions of tandem repeats and basic domains and a consequent marked tendency towards larger-than-normal size are common features of *Plasmodium* proteins. It is interesting that asparagine rich repeats, basic domain insertions and the tendency towards increased size and complexity are also features of subunits of the *P. falciparum* RNA polymerases II [42] and III [43]. The authors of these studies proposed that the enlarged variable regions contained regulatory domains for transcriptional control during the parasite's complex developmental cycle. This hypothesis remains tenable and also applicable to the SNF2L protein and its unique domains. Lysine rich sequences within nucleoproteins are common and are associated with nuclear localisation signals [44]. Residues 895–920 of SNF2L have a particularly high match to the nuclear targeting consensus sequences predicted using the PSORT protein localisation site programme (underlined in dashes in Fig. 1). Several potential casein kinase II phosphorylation sites are also found near this region (STDE, 870–873; TDEE, 871–874; SGEE, 879–882; TILE 883–886; TFLD, 928–931). Such phosphorylation sites are frequently found next to nuclear localisation sequences and can enhance the rate of nuclear transport [45].

The molecular switches which regulate the malaria parasite's responses to extra-cellular signals and define entry into developmental pathways such as endothelial sequestration or gametocytogenesis are not understood. This problem can in principle be studied by describing the components of transcriptional activation in *P. falciparum* and comparing them to better understood systems and then testing hypothetical models for *Plasmodium* developmental mechanisms. The problem with this approach is the complexity of gene expression. Thus, for RNA polymerase II

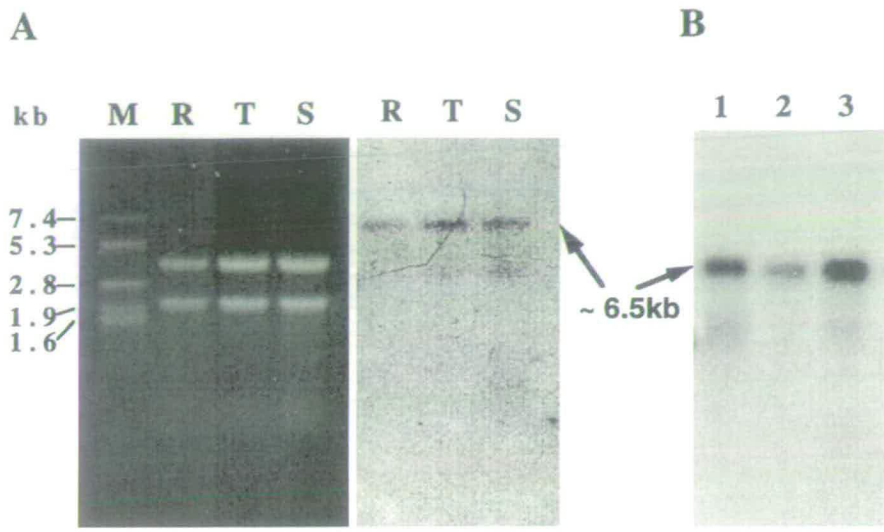


Fig. 4. Message size and steady state mRNA levels throughout the asexual cycle. (A) Ethidium stained 1.2% formaldehyde/agarose gel separation of total RNA from synchronised cultures of 3D7A parasites Northern blotted onto nylon membranes and probed with a radiolabelled *SNF2L* fragment. R, T and S represent equal amounts of total RNA from rings, trophozoites and schizonts respectively. (B) Total *P. falciparum* RNA from three clonal parasite lines (1, Dd2; 2, 3D7A; 3, HB3B) and the Northern blot of the same gel transferred to nylon membranes and hybridised to a radiolabelled *SNF2L* probe.

(pol II) to transcribe a gene requires the recruitment of >20 proteins to a promotor [46,47]. In addition, assembly of the initiation complex requires the prior removal or repositioning of nucleosomes, carried out by multi-protein chromatin remodelling complexes, some of which may be promotor specific (another 20–30 proteins?) [48,49]. Current models of chromatin remodelling complex function are largely based on studies of the yeast SWI/SNF complex. Each protein in this complex is essential for transcriptional control at two sets of inducible loci, those involved in mating type switching (the *swi* mutants) and those controlling sugar fermentation (the *snf*-sucrose non-fermenting mutants). However, none of the genes encoding the SWI/SNF complex proteins are essential for asexual growth and normal chromatin remodelling and transcription occurs at other promotors in *swi* and *snf* mutants.

The lack of mutant selection systems and the fragility of the culture system has inhibited progress on complex aspects of the molecular biology of the malaria parasite. The genome mapping project [50] and the sequencing of the *P. falciparum* genome could permit much faster identification of conserved core components of the parasite's transcription apparatus. With genetic tests for functionality such as gene 'knock outs' [51] it may be possible to ask whether genes such as *P. falciparum* SNF2L are essential for survival or involved in transcription of 'inessential' inducible loci such as the metabolic and mating type switches activated by the SWI/SNF chromatin remodelling complex.

Acknowledgements

We thank Bill Richardson for expert technical assistance and members of the Blaxter, Walliker and Maizels laboratories for helpful advice and discussions. We also thank Dr Alistair Craig for the gift of the *P. falciparum* cDNA library. This work was funded by The Wellcome Trust through a Senior Fellowship grant to D.E.A. The UK Government awarded an Overseas Research Studentship grant to D.D.J.

References

[1] Newbold CI, Boyle DB, Smith CC, Brown KN. Stage specific protein and nucleic acid synthesis during the asexual cycle of the rodent malaria *Plasmodium chabaudi*. Mol Biochem Parasitol 1982;5:33–9.

[2] Myler P, Saul A, Kidson C. The synthesis and fate of stage-specific proteins in *Plasmodium falciparum* cultures. Mol Biochem Parasitol 1983;9:37–45.

[3] Fox BA, Bzik DJ. Analysis of stage specific transcripts of the *Plasmodium falciparum* serine repeat antigen (SERA) gene and transcription from the SERA locus. Mol Biochem Parasitol 1994;68:133–44.

[4] Levitt A. RNA processing in malarial parasites. Parasitol Today 1993;9:465–8.

[5] Sultan AA, Richardson W, Alano P, Arnot DE, Doerig C. Cloning and characterisation of a *Plasmodium falciparum* *Ras*-like GTPase involved in signal transduction for cell cycle control. Mol Biochem Parasitol 1994;65:531–8.

[6] Myler PJ. Transcriptional analysis of the major merozoite surface antigen precursor (GP195) gene of *Plasmodium falciparum*. In: Agabian N, Cerami A, editors. Parasites: Molecular Biology, Drugs and Vaccine design. New York: Wiley-Liss, 1990, pp. 123–137.

[7] Lanzer M, de Bruin D, Ravetch JV. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies a region in which transcription terminates and reinitiates. EMBO J 1992;11:1949–55.

[8] Lanzer M, de Briun D, Ravetch JV. A sequence element associated with the *Plasmodium falciparum* KAHRP gene cis to the site for developmentally regulated protein-DNA interactions. Nucleic Acids Res 1992;12:3051–6.

[9] Alano P, Silvestrini F, Roca L. Structure and polymorphism of the upstream region of the *pig27/25* gene, transcriptionally regulated in gametocytogenisis of *Plasmodium falciparum*. Mol Biochem Parasitol 1996;79:207–17.

[10] Ruvolo V, Altszuler R, Levitt A. The transcript encoding the circumsporozoite antigen of *Plasmodium berghei* utilises heterogenous polyadenylation sites. Mol Biochem Parasitol 1993;57:137–50.

[11] Lanzer M, Wertheimer SP, de Bruin D, Ravetch JV. Plasmodium: Control of gene expression in malaria parasites. Exp Parasitol 1993;77:121–8.

[12] Fink GR. Pseudogenes in yeast? Cell 1987;49:5–6.

[13] Knapp B, Nau U, Hundt E, Kupper HA. Demonstration of alternative splicing in a pre-mRNA expressed in the blood stage form of *Plasmodium falciparum*. J Biol Chem 1991;266:7148–54.

[14] McAndrew MB, Read M, Sims PFG, Hyde JE. Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A + T rich human malaria parasite *Plasmodium falciparum*. Gene 1993;124:165–71.

[15] Koleske AJ, Young RA. The RNA polymerase II holoenzyme and its implication for gene regulation. Trends Biochem Sci 1995;20:113–6.

[16] Peterson CL. Multiple switches to turn on chromatin? Curr Opin Genet Dev 1996;6:171–5.

[17] Vettese-Dadely M, Grant PA, Hebbes TR, Crane-Robinson C, Allis CD, Workman JL. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J 1996;15:2508–18.

[18] Breeden L, Nasmyth K. Cell cycle control of the yeast HO gene: cis and trans-acting regulators. Cell 1987;48:389–97.

[19] Petersen CL, Herskowitz I. Characterisation of the yeast SWI1, SWI2 and SW13 genes, which encode a global activator of transcription. Cell 1992;68:573–83.

[20] Laurent BC, Treich I, Carlson M. The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. Genes Dev 1993;7:583–91.

[21] Steger DJ, Workman JL. Remodeling chromatin structures for transcription: What happens to the histones? Bioessays 1996;18:875–84.

[22] Trager W, Jensen JB. Human malaria parasites in continuous culture. Science 1976;193:673–5.

[23] Lambros C, Vanderberg JP. Synchronisation of *Plasmodium falciparum* erythrocytic stages in culture. J Parasitol 1979;65:418–20.

[24] Carter R, Ranford-Cartwright L, Alano P. The culture and preparation of gametocytes of *Plasmodium falciparum* for immunochemical, molecular and mosquito infectivity studies. Methods Mol Biol 1993;21:67–88.

[25] Weber JL. Analysis of sequences from the extremely A + T rich genome of *Plasmodium falciparum*. Gene 1987;52:103–9.

[26] Kemp DJ, Thompson JK, Walliker D, Corcoran LM. Molecular karyotype of *Plasmodium falciparum*: Conserved linkage groups and expendable histidine-rich protein genes. Proc Natl Acad Sci USA 1987;84:7672–6.

[27] Alexander H, Alexander S. Identification of introns by reverse-transcription PCR. BioTechniques 1996;20:778–80.

[28] Devereux J, Haederli P, Smithies O. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 1984;12:387–95.

[29] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10.

[30] Solovyev VV, Salamov AA, Lawrence CB. Predicting internal exons by oligonucleotide composition and discriminant analysis of splicable open reading frames. Nucleic Acids Res 1994;22:5156–613.

[31] Brondel V, Bucher P, Nourbakhsh I, Blaisdell BE, Karlin S. Methods and algorithms for statistical analysis of protein sequences. Proc Natl Acad Sci USA 1992;89:2002–6.

[32] Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res 1984;12:857–72.

[33] Eisen JA, Sweder KS, Hanawait PC. Evolution of the SNF2 family of proteins: Subfamilies with distinct sequences and functions. Nucleic Acids Res 1995;23:2715–23.

[34] Elfring LK, Deuring R, McCallum CM, Peterson CL, Tamkun JW. Identification and characterisation of *Drosophila* relatives of the yeast transcriptional activator SNF2/SWI2. Mol Cell Biol 1994;14:2225–34.

[35] Okabe I, Bailey LC, Attree O, Srinivasan S, Perkel JM, Laurent BC, Carlson M, Nelson DL, Nussbaum RL. Cloning of human and bovine homologs of SNF2/SWI2: A global activator of transcription in yeast *S. cerevisiae*. Nucleic Acids Res 1992;20:4649–55.

[36] Wittinghofer A, Pai EF. The structure of Ras protein: A model for a universal molecular switch. Trends Biochem Sci 1991;16:382–7.

[37] Guntaka RV, Kandala JC, Reddy VD. Cloning and characterisation of a highly conserved HNG-like protein (Pfl6) gene from *Plasmodium falciparum*. Biochem Biophys Res Commun 1992;182:412419.

[38] Kun JFJ, Anders RF. A *Plasmodium falciparum* gene encoding a high mobility group protein box. Mol Biochem Parasitol 1995;71:249–53.

[39] Tsukiyama T, Daniel C, Tamkun J, Wu C. ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kD subunit of the nucleosome remodelling factor. Cell 1995;83:1021–6.

[40] Aasland R, Stewart AF, Gibson T. The SANT domain: A putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends Biochem Sci 1996;21:87–8.

[41] Cairns BR, Lorch Y, Li Y, Zhang M, Lacomis L, Erdjument-Bromage H, Tempst P, Du J, Laurent B, Kornberg RD. RSC, an essential, abundant chromatin-remodeling complex. Cell 1996;87:1249–60.

[42] Li W-B, Bzik DJ, Gu H, Tanaka M, Fox BA, Inselburg J. An enlarged largest subunit of *Plasmodium falciparum* RNA polymerase II defines conserved and variable RNA polymerase domains. Nucleic Acids Res. 1989;17:9621–36.

[43] Li W-B, Bzik DJ, Tanaka M, Gu H, Fox BA, Inselburg J. Characterization of the gene encoding the largest subunit of *Plasmodium falciparum* RNA polymerase III. Mol Biochem Parasitol 1991;46:229–40.

[44] Gerace L. Nuclear export signals and the fast track to the cytoplasm. Cell 1995;82:341344.

[45] Rihs H-P, Jans D, Fan H, Peters R. The rate of nuclear cytoplasmic protein transport is determined by the casein kinase II site flanking the nuclear localisatin sequence of the SV40 T-antigen. EMBO J 1990;10:633–9.

[46] Tjian R, Maniatis T. Transcriptional activation: A complex puzzle with few easy pieces. Cell 1994;77:5–8.

[47] Buratowski S. The basics of basal transcription by RNA polymerase II. Cell 1994;77:1–3.

[48] Felsenfeld G. Chromatin unfolds. Cell 1996;86:13–9.

[49] Tsukiyama T, Becker P, Wu C. ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor. Nature 1994;367:525–32.

[50] Dame JB, Arnot DE, Bourke PF, Chakrabarti D, Christodoulou Z, Coppel RF, Cowman AF, Craig AG, Fischer K, Foster J, Goodman N, Hinterberg K, Holder AA, Holt DC, Kemp DJ, Lanzer M, Lim A, Newbold CI, Ravetch JV, Reddy GR, Rubio J, Schuster SM, Su X-Z, Thompson JK, Vital F, Wellems TE, Werner EB. Current status of the *Plasmodium falciparum* genome project. Mol Biochem Parasitol 1996;79:1–12.

[51] Menard R, Sultan AA, Cortes C, Altszuler R, van Dijk MR, Janse CJ, Waters AP, Nussenzweig RS, Nussenzweig V. Circumsporozoite protein is required for development of malaria sporozeites in mosquitoes. Nature 1997;385:336–40.

ELSEVIER

Molecular and Biochemical Parasitology 000 (1998) 000–000

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

An RCC1-type guanidine exchange factor for the Ran G protein is found in the *Plasmodium falciparum* nucleus[1]

Dar-Der Ji [a], Ali A.K. Sultan [2,a], Debopam Chakrabarti [b], Paul Horrocks [3,c], Christian Doerig [d], David E. Arnot [a,*]

[a] *Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK*
[b] *Department of Molecular Biology and Microbiology, University of Central Florida, Orlando, FL, USA*
[c] *Institute of Cell and Molecular Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK*
[d] *I.N.S.E.R.M. Unit 313, Hopital Pitie-Salpetriere, Boulevard de la Hopital, Paris, France*

Received 2 April 1998; received in revised form 6 July 1998; accepted 15 July 1998

Keywords: Ran GTP cycle; Nuclear import; Guanidine exchange factors

Ran, a small GTPase found in the cytoplasm and the cell nucleus [1,2], is abundantly expressed in the dividing blood stages of *Plasmodium falciparum* [3,4]. It plays an essential role in protein import into the nucleus [5] and its GTP loading and hydrolysis cycle may be the major energy source for nuclear localisation signal mediated translocation through nuclear pore complexes [6]. The guanidine nucleotide exchange factor for Ran is known as RCC1 (for regulator of chromosome condensation [7]), a chromatin protein which stimulates guanine nucleotide loading and thus generates Ran GTP within the nucleus. Ran GTP–Ran GDP hydrolysis is catalysed by the GTPase activating protein Ran GAP1, which is located exclusively in the cell cytoplasm [8].

A cytoplasmic activator of GTP hydrolysis separated by the nuclear membrane from a nuclear GTP generator regulate nuclear protein import by the Ran cycle. Ran GDP-associated nuclear localisation signal marked protein complexes destined for nuclear import form in the cytoplasm and dissociate in the nucleus along a steep *trans*-membrane Ran GTP gradient [6]. Exchange of GDP for GTP on intra-nuclear RCC1/Ran complexes causes dissociation of the complex and Ran GTP export to the cytoplasm. There, Ran-GAP1 catalyses re-conversion to Ran GDP which can then once more bind and chaperone nuclear localisation sequence tagged proteins to nuclear pore complexes and the nucleus.

* Corresponding author. Tel.: +44 131 6508655; fax: +44 131 6508655; e-mail: d.e.arnot@ed.ac.uk

[1] Nucleotide sequence data reported in this paper are available in the EMBL, GenBank™ and DDJB databases under the accession number AF048836.

[2] Present address: Michael Heidleberger Division of Immunology, New York University Medical School, New York, USA.

[3] Present address: University of Wurtzburg, D-97070, Wurzburg, Germany.

PII S0166-6851(98)00107-8

If a Ran-mediated nuclear import cycle functions in intra-erythrocytic malaria parasites then a protein homologous to the Ran guanidine exchange factor RCC1 must exist. In all organisms from which this guanidine exchange factor has been characterised, RCC1 is a 600–700 amino acid protein with a central region of seven 50–60 amino acid repeats each of which contains conserved sequence motifs within otherwise divergent sequences [7,9–13]. The seven RCC1-type repeats have also been found in one other protein which also appears to act as a guanidine exchange factor, but in this case for the small G protein ADP-ribosylation factors (ARF's) for trimeric G protein α chains. An extremely large, cytosolyic and Golgi-bound protein, the p619 Giant Protein, has two sets of RCC1-type repeats, the carboxy-terminal one interacting with myristylated ARF1 in the Golgi and the amino terminal set stimulating guanine nucleotide exchange on ARF1 and Rab but not on the Ran or Ras G proteins [14].

Two degenerate oligonucleotides were synthesised based on the conserved repeat sequences to try to PCR amplify a *P. falciparum* RCC1 gene fragment and characterise this protein and its role in the parasite nuclear import cycle. A 380 base pairs (bp) fragment was amplified in this way from parasite genomic DNA and this sequence encoded a reading frame with homology to the conserved motifs of the human RCC1 protein. The fragment was used to screen an asexual erythrocytic stage cDNA library from which a 3 kb cross hybridising clone was isolated, sequenced and used to isolate three more cDNA clones which extended the sequence further on the 5′ and 3′ sides. A total of 4161 bp of cDNA have been sequenced, encoding a single contiguous reading frame without stop codons which has been primed from a short dA-rich 3′ end sequence. Restriction mapping and Southern blotting experiments indicate that this is a single copy gene co-localising to chromosome 5 with the *P. falciparum* reference marker gene *Mdr*-1 on pulsed field gel blots. Northern blotting experiments with RNA from synchronised cultures indicate that low-level transcription of a single $\sim$8 kb RNA occurs in the trophozoites and to a lesser extent in the schizonts but that the gene is inactive in the young ring stages. This clone was named RCC1 and its size and structure compared with other RCC1 motif proteins are shown in Fig. 1.

The Swissprot data base was used to compare the putative *P. falciparum* RCC1 homologue with other RCC1 proteins and the p619 Giant Protein. The alignments of the seven RCC1-type repeats from the human RCC1 protein, the *Saccharomyces cerevisiae* RCC1 homologue (SRM1), and the two sets of repeats from the p619 protein confirm that the *P. falciparum* RCC1 is a member of this family of guanidine exchange factor proteins. Internal homology of the repeats to each other is low in all of these proteins. Outside the repeat blocks there is no sequence conservation. Sequence insertions, a common feature of *P. falciparum* proteins, occur in two of the seven repeats and in repeat IV the insertion is 140 amino acids long. Amino acid identity between each sequence in the seven RCC1 repeats is only between 11 and 15% although similarity in the short sequence motifs themselves rises to $\sim$60%. The single longest block of amino acid identity between the *P. falciparum* sequence and any other is a sequence around the third tandem repeat where 11/18 amino acids (61%) are shared with the p619 giant protein. This, together with the large size of the parasite RCC1 and its slightly higher homology with p619 raises the possibility that the *P. falciparum* sequence is actually a homologue of this cytoplasmic guanidine exchange factor rather than of the intra-nuclear RCC1 proteins. Arguing against this possibility is the fact the extensive partial sequence obtained does not appear to leave coding space in the 8 kb mRNA for the additional set of seven RCC1 repeats that occur in p619 (encoded by a 15 kb mRNA molecule [14]). Low stringency Southern blotting experiments also detect only one copy of the RCC1 type repeats in the *P. falciparum* genome.

To establish the intra-cellular location of RCC1 in parasitised erythrocytes, a DNA fragment encoding 139 amino acids (comprising the whole of the first and most of the second RCC1 repeat) was cloned into the pET29b bacterial expression vector. The hexahistidine-tagged fusion protein was expressed in *Escherichia coli* and purified by affinity chromatography [15]. The purified protein

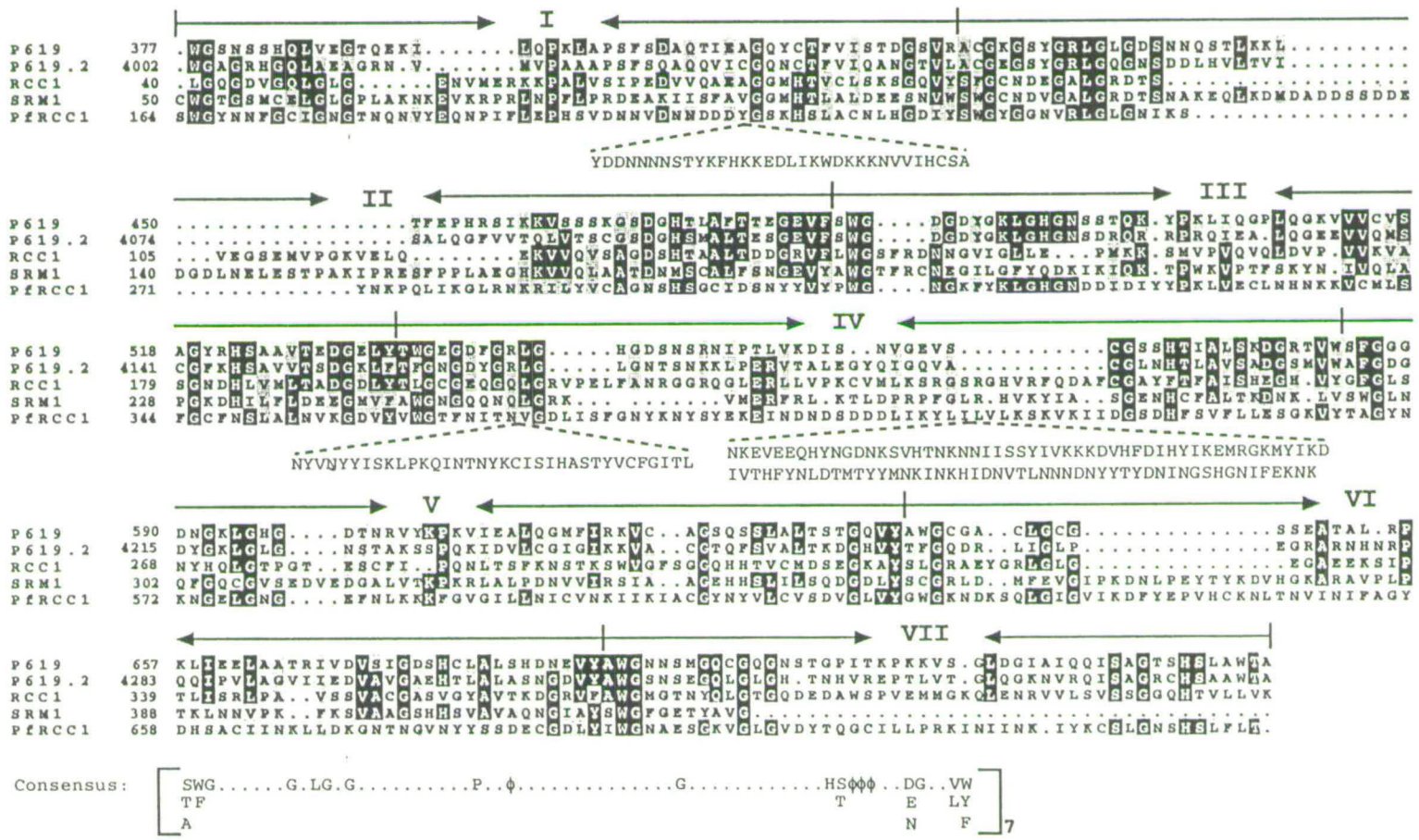


Fig. 1. Alignment of the conserved sequence motifs in the RCC1-type repeats. p619 is the human p619 giant protein [14], RCC1 is the human RCC1 protein [7] and SRM1 is the yeast homologue of RCC1 [12]. Regions highlighted in black indicate sequence identity, those in grey indicate conservative amino acid substitutions. The seven conserved blocks are numbered and *P. falciparum* specific insertions are shown below Block I and IV. The consensus motif is shown below, ϕ indicating hydrophobic residue.

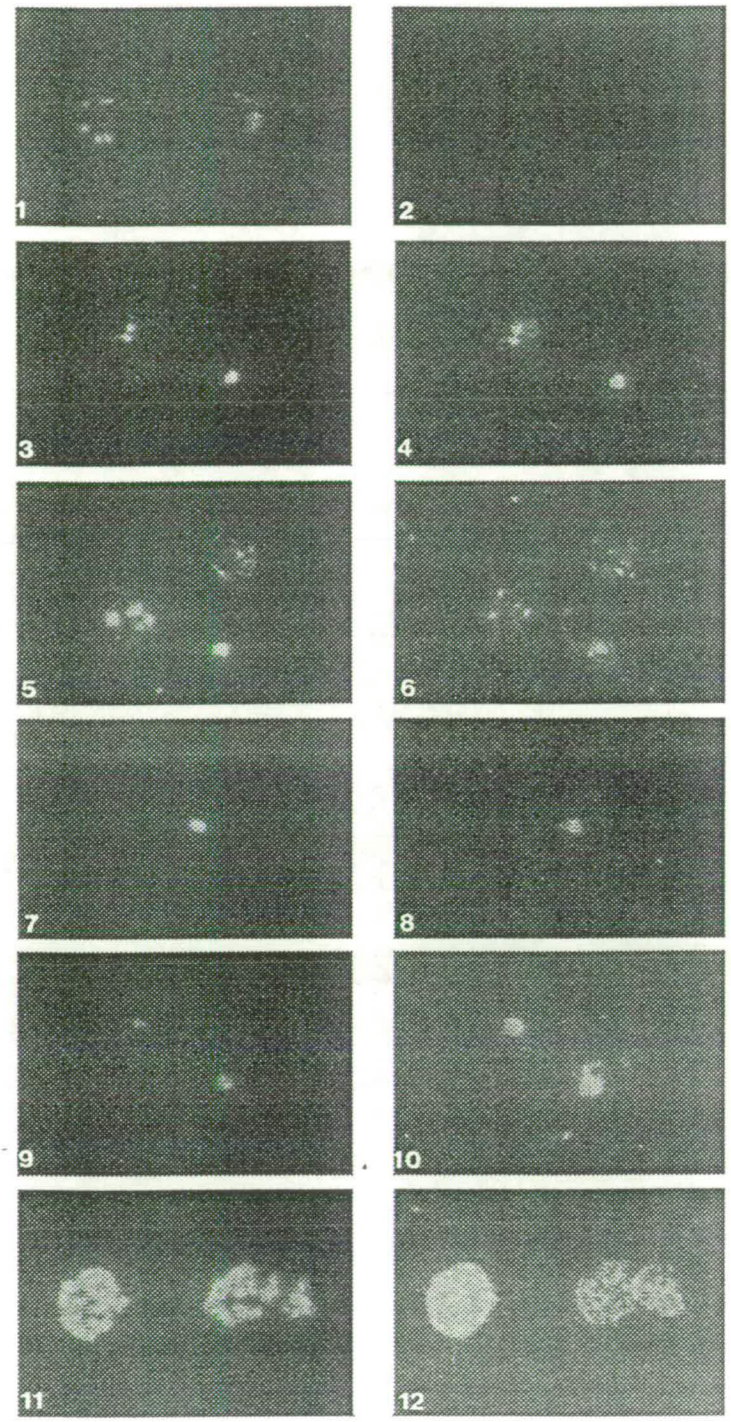


Fig. 2.

was injected into rabbits with Titremax adjuvant to raise a specific antiserum against the recombinant protein. Fig. 2 shows immunofluoresence micrographs comparing the localisation of antibodies to RCC1 in isolated *P. falciparum* nuclei, with nuclear staining using the DNA binding dye diamino-2-phenylinodle (DAPI) Plates 1, 3, and 5, show isolated parasite nuclei from mechanically disrupted parasitised erythrocytes stained with DAPI. Plates 2 shows the corresponding parasite nuclei stained with fluorescein isothiocyanate (FITC) labelled antisera from the rabbit before immunisation with recombinant RCC1 protein. Plates 4, and 6 show fluorescent staining with FITC-labelled mouse antiserum raised against the *P. falciparum* homologue of the SNF2L chromatin remodelling ATPase [16] and the rabbit anti-RCC1 antiserum respectively. While the pre-immune sera does not react with the isolated nuclei, antibodies raised against SNF2L, a chromatin associated helicase, and antibodies against RCC1 both react with the isolated nuclei. This indicates that the proteins recognised by both these antisera are located within the parasite nucleus.

To locate RCC1 in developing intra-erythrocytic parasites, immuno-fluorescence microscopy using the FITC labelled anti-RCC1 antisera with the ring, trophozoite and segmenting schizont stages of sorbitol synchronised cultures are shown in Fig. 2, plates 7–12. Although the resolution of internal structures within the erythrocytic parasite is at the limits of that obtainable with fluorescence microscopy, it is clear that the labelled antibody co-localises with segmented schizont nuclei stained with DAPI (plates 11 and 12 and also with the small DAPI stained nuclei of the early ring stages (plates 7 and 8) and the larger, probably dividing nuclei present in the late trophozoite/early schizont stages (plates 9 and 10).

The conclusion drawn from these experiments is that we have identified a *P. falciparum* RCC1-type guanidine exchange factor and that antibody localisation studies indicate that it is located in the parasite nucleus. It is therefore likely that this gene encodes the intra-nuclear guanidine exchange factor involved in the Ran mediated nuclear protein import cycle.

Acknowledgements

We wish to thank Bill Richardson for excellent technical support and Alex Loukas and Bill Gregory for helpful discussions and advice. D.E.A. is a Senior Fellow of the Wellcome Trust who supported this work.

References

[1] Drivas GT, Shih A, Coutavas E, Rush MG, D'Eustachio P. Characterisation of four novel RAS-related genes expressed in a human teratocarcinoma cell line. Mol Cell Biol 1990;10:1793–8.

[2] Bischoff FR, Ponstingl H. Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. Proc Natl Acad Sci USA 1991;88:10830–4.

[3] Sultan AA, Richardson WA, Alano P, Arnot DE, Doerig C. Cloning and characterisation of a *Plasmodium falciparum* homologue of the Ran/TC4 signal transducing GTPase involved in cell cycle control. Mol Biochem Parasitol 1994;65:331–8.

[4] Dontfraid F, Chakrabarti D. Cloning and expression of a cDNA encoding the homologue of the Ran/TC4 GTP-binding protein from *Plasmodium falciparum*. Biochem Biophys Res Commun 1994;201:423–9.

[5] Moore MS, Blobel G. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. Nature 1993;365:661–3.

[6] Gorlich D, Pante N, Kutay U, Aebi U, Bischoff FR.

Fig. 2. Immunolocalisation of the RCC1 protein. Plates 1–6 show staining of parasite nuclei which were isolated by saponin lysis of infected erythrocytes followed by cell rupture by homogenisation and separation of parasite nuclei from the external membranes by centrifugation[17]. Plates 1, 3 and 5 show DAPI stained isolated *P. falciparum* (strain 3D7) nuclei from an asynchronous culture. Plate 2 shows the nuclei stained in plate 1 co-stained with FITC labelled pre-immunisation rabbit serum. Plate 4 shows the field corresponding to plate 3, stained with FITC labelled rabbit antiserum raised against the *P. falciparum* SNF2L protein [16]. Plate 6 shows the parasites stained in plate 5, stained with a FITC labelled rabbit antiserum raised against a recombinant derived fragment of RCC1. Plates 7, 9 and 11 show DAPI stained intact rings, trophozoites and schizonts respectively. Plates 8, 10 and 12 show the same parasites co-stained with the FITC labelled rabbit anti-RCC1 antisera.

Identification of different roles for Ran GDP and Ran GTP in nuclear protein import. EMBO J 1996;15:5584–94.

[7] Ohtsubo M, Okazaki H, Nishimoto T. The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. J Cell Biol 1989;109:1389–97.

[8] Bischoff FR, Ponstingle H. Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran GAP1. Methods Enzymol 1995;257:135–44.

[9] Nishitani H, Kobayoshi H, Ohtsubo M, Nishimoto T. Cloning of Xenopus RCC1 cDNA, a homologue of the human RCC1: complementation of tsBN2 mutation and identification of the product. J Biochem 1990;107:228–35.

[10] Frasch M. The maternally expressed Drosophila gene encoding the chromatin-binding protein BJ1 is a homolog of the vertebrate regulator of chromosome condensation, RCC1. EMBO J 1991;10:1225–30.

[11] Matsumoto T, Beach D. Premature initiation of mitosis in yeast lacking RCC1 or an interacting GTPase. Cell 1991;66:247–360.

[12] Clark K, Sprague GF. Yeast pheromone response pathway: characterisation of a suppressor that restores mating to receptorless mutants. Mol Cell Biol 1989;9:2682–94.

[13] Gorlich D, Mattaj IW. Nucleocytoplasmic transport. Science 1996;271:1513–8.

[14] Rosa JL, Casaroli-Marano RP, Buckler AJ, Vilaro S, Barbacid M. p619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and Rab proteins. EMBO J 1996;15:4262–73.

[15] Yasukawa T, Kanei-Ishii C, Maekawa T, Fujimoto J, Yamamoto T, Ishii S. Increase in solubility of foreign proteins in *Escherichia coli* by co-production of the bacterial thioredoxin. J Biol Chem 1995;270:25328–31.

[16] Ji D-D, Arnot DE. A *Plasmodium falciparum* homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodelling for transcription. Mol Biochem Parasitol 1997;88:151–62.

[17] Lanzer M, de Bruin D, Ravetch JV. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. EMBO J 1992;11:1949–55.