Characterization of a protein of the rodent malarial parasite Plasmodium chabaudi containing a novel leucine-histidine zipper

Thesis

UNRESTRICTED

Characterization of a protein of the rodent malarial parasite *Plasmodium chabaudi* containing a novel leucine-histidine zipper

Ekkehard Bruno Ernst Werner

A thesis submitted in partial fulfilment of the
requirements of the Open University
for the degree of Doctor of Philosophy

July 1995

Division of Parasitology
National Institute for Medical Research
The Ridgeway
Mill Hill, London NW7 1AA

Date of submission : July 1995
Date of award : 31st October 1995

... love hides in molecular structure ...

The Doors

meiner Mutter und meinem Vater, durch die ich das Licht der Welt erblickte

Characterization of a protein of the rodent malarial parasite *Plasmodium chabaudi* containing a novel leucine-histidine zipper.

Abstract

Clones from cDNA and genomic DNA libraries of *Plasmodium chabaudi* 96V covering the entire open reading frame for a yet uncharacterized malarial protein were isolated. Counting the first ATG as start codon the intronless gene codes for a 229 kDa protein. In the centre of the protein a 364 amino acid repeat region is located and is based on 32 11-mer repeats divided by two 6-mer repeats into three blocks.

Modelling of the repeat region led us to propose a model where each of the three units forms an α-helical coiled-coil triple-helix containing a novel leucine-histidine zipper. Each unit resembles in structure the units present in spectrin molecules. The repeat region is flanked by predicted heptad based α-helical coiled-coil regions and the 229 kDa protein has an overall character of a cytoskeletal protein.

Antisera raised against recombinant polypeptides from two different regions of the 229 kDa protein reacted in western-blotting experiments with a M_r 240 000/ 225 000 doublet present in protein extracts from *P. chabaudi* 96V. The same sera in immunofluorescence suggested a localization of the 229 kDa protein in the organelles of the apical complex, presumably in the rhoptry organelles, and an association of the 229 kDa protein with the erythrocyte membrane. Furthermore it was shown in western-blotting experiments with the recombinant polypeptides that the 229 kDa protein is a natural immunogen during infection.

We named the 229 kDa protein Repeated Organellar Protein (ROPE) and suggest that ROPE may be involved in the process of invasion, that it interacts with the erythrocyte cytoskeleton and that the leucine-histidine zipper may be involved in molecular mimicry of spectrin.

Acknowledgements

I am deeply indebted to Tony Holder for his guidance throughout the preparation of my thesis at the NIMR, where his great human sense helped me to find a solution for many difficult situations.

The modelling part of the thesis was done in collaboration with Willie Taylor from Mathematical Biology at the NIMR whose support and participation was vital.

Amongst the many people who helped me in many ways to prepare the thesis are: András Aszódi, Joseph Brock, Adel Ebrahimzadeh, Michael Green, Roger Morris, Keitan Patel, Terry Scott-Finnigan, George Snounou, William Stafford, Shabih Syed, Mrs. Way, Isabelle Wells and Julia Zola each of whom contributed with a helping hand and encouragement.

Finally I want to thank my mother, Erna Werner, whose non-material and material support was given with many sacrifices.

Index

List of Figures

ABBREVIATIONS

amp	ampicillin
A_{260}	absorbance at 260 nm
APS	ammoniumperoxydisulphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pair
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
cfu	colony-forming units
Ci	curie
CIP	calf intestinal phosphatase
cpm	counts per minute
Da	dalton
DEPC	diethyl-pyrocarbonate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
FITC	fluorescein-isothiocyanate conjugate
HEPES	N-2-hydroxyethylpiperazine-N-ethane- sulphonic acid
IFA	immunofluorescence assay
i.m.	intramuscularly
i.p.	intraperitoneally
IPTG	isopropyl-ß-D-thiogalactopyranoside
kb	kilobase
M_r	relative molecular mass
NBT	nitro blue tetrazolium
NP-40	nonidet P-40 (octylphenolethylenoxide condensate)
mRNA	messenger ribonucleic acid
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque-forming units
PMSF	phenylmethylsulphonylfluoride
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
TEMED	N,N,N',N'-tetramethylethylendiamine

tet	tetracycline
T_m	melting (or mid point) temperature
TPCK	tosyl-L-phenylalanine chloromethylketone
Triton X-100	iso-octylphenoxy polyethoxyethanol
TRIS	tris (hydroxymethyl) aminomethane
UV	ultraviolet
vol	volume
v/v	volume ratio
w/w	weight ratio
X-gal	5-bromo-4-chloro-3-indolyl-ß-D- galactopyranoside

Introduction

1.1. Malaria in the past and at present

Nowadays malaria is known in public as a tropical disease. It is not common knowledge that this disease has also had a great impact, since ancient times, on the European population and was eradicated here only after World War II by massive use of the insecticide DDT and the drug chloroquine (Resochin).

Its symptoms were already described by Hippocrates and malaria occurred not only in Mediterranean countries but also in the warmer regions of Germany and England. Since ancient times military expeditions have often been ended or seriously hampered by malaria, like most of the invasions of Italy by the German emperors or many of the more recent colonial wars.

The first antimalarial drug, quinine, was discovered by lucky chance in the year 1638, when the Spanish countess Cinon suffered in Lima (Peru) from malaria, which had been imported into the New World by the Spaniards. Her physician treated her with a powder containing quinine, which was used by the native population to treat symptoms of fever and discovered its second independent property, to be an antimalarial drug.

In 1820 quinine was isolated by Caveston. The use of synthetic quinine and the insecticide DDT led in the 1950s to attempts of worldwide eradication or containment of malaria.

In the last 25 years malaria has regained its former position as one of the major threats to human health and prosperity (Wyler, 1992). About 1 500 million people live in endemic areas, 300 million cases are reported and more than 2 million people, mostly children, die from malaria each year. In addition to mortality, severe morbidity can also result. The main reasons for the widespread revival of malaria are the lack of appropriate insecticides and the fast occurrence of resistance of the malarial parasite to

chemotherapeutic drugs. Current efforts to circumvent the resistance to present drugs are to design antimalarials with different targets than the ones of known drugs (Zhe Li *et al.*, 1994; Schirmer *et al.*, 1995) or to use synthetic and recombinant vaccines (Romero, 1992; Bathurst, 1994). These strategies are still in their infancy.

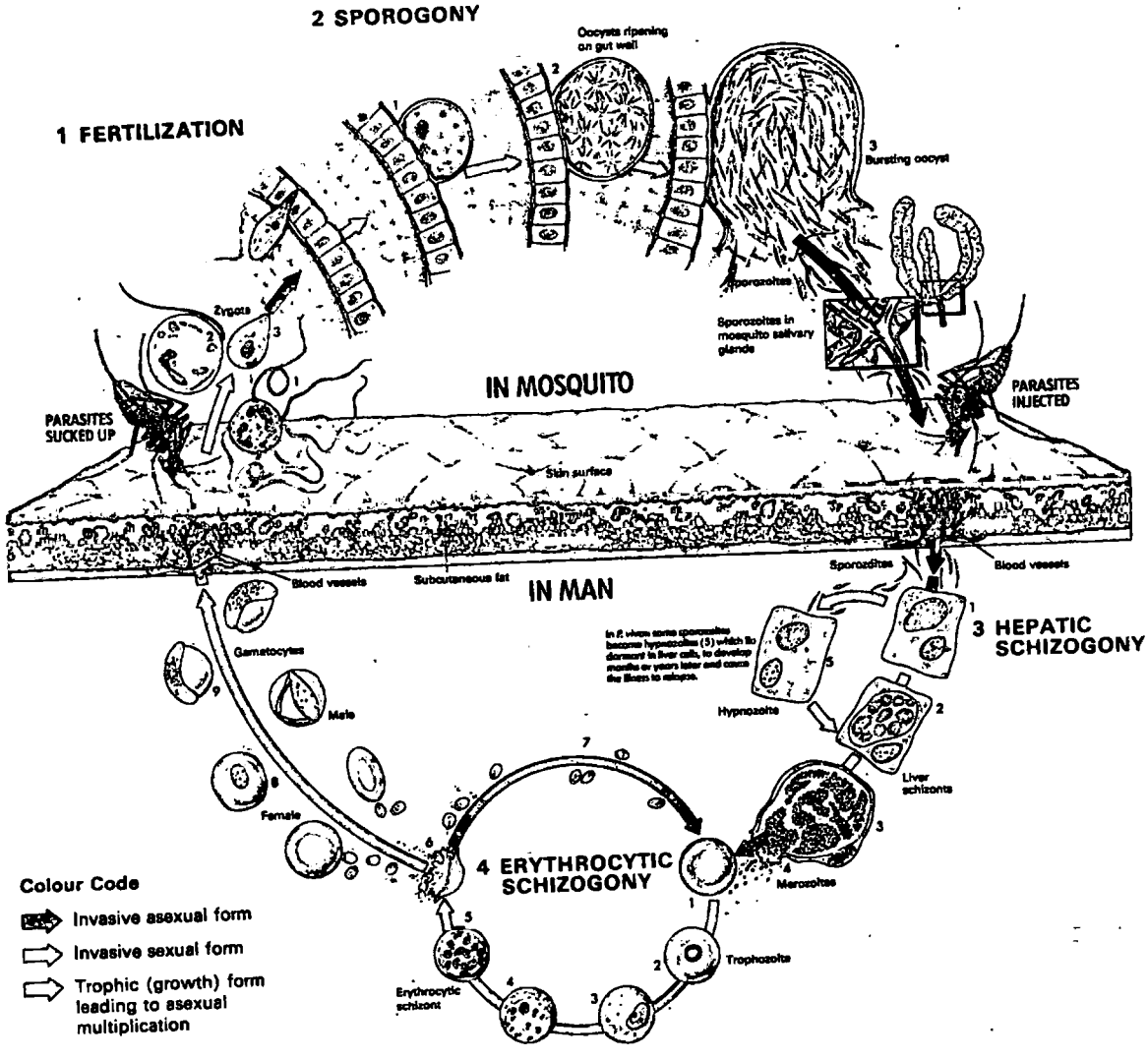
Malaria is caused by protozoan parasites of the genus *Plasmodium*. There are many different species of malaria parasites with vertebrate hosts including reptiles, birds, primates and other mammals (Garnham, 1966). Four human malaria parasite species are known but the most severe form of the disease, ending often with cerebral disease or other complications and death, is caused by *P. falciparum* (Sturchler, 1989). The life cycle of *Plasmodia* includes a vertebrate host and a mosquito of the genus *Anopheles*. In Figure 1.1 the life cycle of *P. falciparum* is depicted but its general features are the same for other species of *Plasmodia*.

Infection in the human host starts when the female mosquito takes a blood meal and injects some saliva containing sporozoite-stage parasites. The sporozoites enter the peripheral blood stream and migrate to the liver of the human host where they invade parenchymal cells, followed by differentiation and multiplication of the parasite. The exoerythrocytic stage ends with the rupture of the parenchymal cells releasing thousands of merozoites into the blood stream. The free merozoites rapidly attach to and invade an erythrocyte, beginning the blood-stage of the infection.

The merozoites first differentiate into a ring (ring-shaped morphology) which grows and develops to a trophozoite. The nucleus divides four to five times forming the multinucleated schizont and digested haemoglobin becomes visible as a dark mass (haemozoin) in the cell. Individual merozoites are formed by cell division and finally the erythrocyte ruptures and releases the new merozoites. *P. falciparum* infected erythrocytes release between 16 and 32, and *P. chabaudi* (a rodent malaria parasite) infected erythrocytes a maximum of 16 merozoites, each of which will begin another cycle of invasion and replication.

Figure 1.1: **The life cycle of *Plasmodium falciparum*.**

A schematic representation of the different stages in the life cycle of the species *Plasmodium falciparum*. At the top the stages of the life cycle in the mosquito are shown and at the bottom the stages of the life cycle in man. (Copyright by The Wellcome Trust).


2 SPOROGONY
Oocysts ripening on gut wall
1 FERTILIZATION
3 Bursting oocyst
Zygote
Sporozoites
Sporozoites in mosquito salivary glands
IN MOSQUITO
PARASITES SUCKED UP
PARASITES INJECTED
Skin surface
Blood vessels
Subcutaneous fat
IN MAN
Sporozoites
Blood vessels
3 HEPATIC SCHIZOGONY
Gametocytes
Male
Hypnozoite
Liver schizonts
Female
4 ERYTHROCYTIC SCHIZOGONY
Merozoites
Trophozoite
Erythrocytic schizont
Colour Code
Invasive asexual form
Invasive sexual form
Trophic (growth) form leading to asexual multiplication

The mature blood-stages of *P. falciparum* are mostly sequestered in capillaries causing symptoms like cerebral malaria. Some of the intraerythrocytic parasites undergo differentiation into male and female gametocytes which remain in the peripheral blood stream. If the gametocytes are ingested by a feeding mosquito they can develop into gametes and start the sexual reproductive phase which ends with the generation of new sporozoite-stage parasites.

1.2. The blood stage asexual cycle of *Plasmodium* at an ultrastructural level

In the following sections the invasion of the red blood cell by the merozoite, the differentiation and multiplication of the parasite inside the red blood cell and its final release will be discussed first at an ultrastructural level. In 1.3. some of these ultrastructural descriptions will be correlated to what is known of parasite proteins which might be involved in these processes.

1.2.1. The merozoite

The malaria parasite leaves the hepatic stage in its vertebrate host as a free merozoite of about 1-2 μm in size, surrounded by an electron-dense surface-coat which measures about 20 nm and consists of fine fibrils (Aikawa *et al.*, 1978). Merozoites have a polar organization. Two rhoptries (570 x 330 nm), a few micronemes (100 x 40 nm) and a few dense granules (140 x 120 nm) are located in the apical region and constitute the organelles of the apical complex. One mitochondrion and a nucleus are located in the posterior region (Aikawa *et al.*, 1978; Torii *et al.*, 1989). Free merozoites stained by standard electron microscopic methods show a pair of pear-like rhoptry organelles, each rhoptry appears to be connected to the merozoite exterior by a fine duct (Aikawa, 1971), these ducts may form a common duct (Bannister *et al.*, 1986). The rhoptries' interior shows a finely granular or reticular substructure and is enclosed in a

cytoplasmic membrane about 7.5 nm thick. When merozoites of *P.knowlesi* and *P. falciparum* are fixed in tannic acid-glutaraldehyde, whorls of multilamellar material, with a periodicity of 5-7 nm, become visible in the apices of the rhoptry canals and their ducts. These lamellae are not visible in micronemes or dense granules (Bannister *et al.*, 1986, Stewart *et al.*, 1986).

1.2.2. The process of invasion and the contents of the rhoptries

Free viable merozoites of *P. knowlesi* can be isolated in large quantities from infected blood by a method developed by Dennis *et al.* (1975). The following description of the invasion process is based on observations made within this species, if not mentioned otherwise. Other species should follow a similar or identical pathway of events.

When a free merozoite encounters an uninfected erythrocyte it becomes attached to the surface of the erythrocyte. Invasion itself needs a reorientation of the attached parasite so that the apical end of the merozoite adjoins the membrane of the erythrocyte to place the organelles of the apical complex next to this zone of contact. After attachment and reorientation the common rhoptry duct opens to the exterior and its contents appear to be in contact with the red cell membrane (Bannister *et al.*, 1986). In the beginning of invasion the erythrocyte membrane is slightly deformed and thickened at the site of attachment and a junction between the apex of the merozoite and the membrane of the erythrocyte is formed (Aikawa *et al.*, 1978). As invasion progresses, the parasite creates an invagination in the membrane of the erythrocyte, the thickened membrane junction moves away from the apical end of the merozoite and appears at the orifice of the invagination of the erythrocyte membrane forming a circumferential junction. This circumferential junction moves towards the posterior end of the merozoite and, when the parasite is within the erythrocyte, the orifice closes leaving the merozoite in the newly created parasitophorous vacuole (Aikawa *et al.*, 1978).

Throughout invasion the apical end of the merozoite remains in contact with the membrane of the erythrocyte and the contents of the rhoptries appear to be discharged directly onto the erythrocyte surface in a continuous membrane-lined column (Aikawa *et al.*, 1978; Bannister *et al.*, 1986). After the invasion process is finished the rhoptries appear empty (Ladda *et al.*, 1969).

The moving junction between the erythrocyte and merozoite membranes appears in freeze-fracture electron micrographs as a narrow band, about 120 nm in width, of rhomboidally arrayed particles visible on the P face of the erythrocyte membrane at the neck of the erythrocyte invagination (Aikawa *et al.*, 1981). These arrayed particles are neither present on the fracture face of the erythrocyte away from the point of invasion nor on the fracture face of the parasitophorous vacuole. The authors suggested that the arrayed particles represent rearrangements of the spectrin/actin contractile network of the erythrocyte skeleton, which had to be done in order to allow invasion. The particles could be the contact point between the contractile proteins and the transmembrane proteins.

It was suggested that components of the rhoptries could generate the early parasitophorous vacuole (Bannister *et al.*, 1986). Whether or not the parasitophorous vacuole is exclusively formed by rhoptry contents or at least partially composed of the erythrocyte membrane is unknown but theoretically the contents of the rhoptries could be sufficient to form the entire parasitophorous vacuole (Bannister *et al.*, 1989).

Freeze fracture studies on merozoites invading erythrocytes show a lower density of integral membrane proteins in the newly forming parasitophorous vacuole membrane than in the directly adjacent erythrocyte membrane, supporting the idea that the parasitophorous vacuole is not a simple invagination of the erythrocyte (Aikawa *et al.*, 1981).

The parasitophorous vacuole expands during parasite maturation and does not adhere as close to the parasite plasma membrane in later stages as it does directly after invasion.

At the segmenter stage the parasitophorous vacuolar membrane is no longer visible (Langreth *et al.*, 1978). Merozoites in late schizonts prior to release issue plumes of concentric membranes from their apices and both the lining membrane of the parasitophorous vacuole and the membrane of the red blood cell show signs of disruption (Bannister *et al.*, 1986). This suggests that the rhoptry content can expand as well as disrupt membranes and might play a role during the invasion process as well as in the release of mature merozoites (Seed *et al.*, 1976; Bannister *et al.*, 1986). Finally, though the lipid contents of the rhoptries appear at least to contribute to the formation of the parasitophorous vacuolar membrane, lipids are not the sole constituent of the interior of these organelles. Rhoptries of *P. falciparum* can be isolated by isopycnic density centrifugation. The isolated organelles have a density in sucrose of 1.16 g ml^{-1}. This relative high density indicates that proteins should be part of the rhoptry contents as well (Etzion *et al.*, 1991).

Invasion of erythrocytes by membrane invagination occurs in all species of *Plasmodia* so far studied (Langreth *et al.*, 1978). Multilamellar structures within rhoptries have not only been found in other invasive developmental stages of *Plasmodium*, e.g. within the rhoptries of *P. berghei* sporozoites (Stewart *et al.*, 1985) but also in other species of Apicomplexa, e.g. within the rhoptries of *Toxoplasma* (Nichols *et al.*, 1983). It seems that the potential of the rhoptries to generate membranes is somehow fundamental to the invasion/release mechanism used by all Apicomplexa.

1.2.3. The micronemes and dense granules

Micronemes are elongated electron-dense membrane-bound organelles present in the invasive stages of all species of the phylum Apicomplexa. Like rhoptry organelles, micronemes may store functional molecules that are released through the apex of the merozoite during the processes of attachment, junction formation and entry of the

merozoite (Hadley *et al.*, 1991).

Dense granules are commonly considered as organelles of the apical complex. They are ellipsoidal in shape, smaller in size and greater in number than the rhoptries and can be located outside of the apical region of the merozoite as well. They appear not to be involved in the process of invasion but move to the membrane of the merozoite after invasion, where their contents are released into the newly formed parasitophorous vacuolar space (Torii *et al.*, 1989).

1.2.4. *Plasmodium* induced changes of the infected erythrocyte

After invading the erythrocyte *Plasmodia* induce a variety of morphological, antigenic and functional changes to the erythrocyte and the membrane of the erythrocyte:

a) Appearance of electron-dense material, so-called knobs

During the trophozoite stage of various species of *Plasmodium*, e.g. *P. falciparum, P. malariae* and *P. brasilianum* electron-dense material, 30-50 nm in height and 90-100 nm in width, the so-called knobs, begins to appear below the membrane of the infected erythrocyte (Aikawa, 1983,1988). In *P. falciparum* appearance of the knobs correlates with sequestration of the mature stages of the parasite from the peripheral circulation (Langreth et al., 1978) and segregation to the deep vasculature of the brain and other organs.

b) Occurrence of clefts in the erythrocyte cytoplasm

Maurer's clefts, structures bounded by a unit membrane, were described in the trophozoite-stage of several species of *Plasmodium*, e.g. *P. falciparum* and *P. malariae*. Sometimes continuity of the Maurer's cleft membrane and the

parasitophorous vacuolar membrane was observed (Langreth *et al.*, 1978; Aikawa, 1988). Electron-dense material (EDM) is commonly observed in the parasitophorous vacuolar space of *P. malariae* and is sometimes associated with Maurer's clefts in erythrocytes infected with *P. malariae* and knob-forming strains of *P. falciparum*. The density and appearance of the EDM is the same as that of the material located in the knob-structures under the membrane of infected erythrocytes and suggests that the parasite derived EDM is transported from the parasite to the erythrocyte membrane via Maurer's clefts (Aikawa, 1988).

Clefts of different appearance and vesicle complexes forming a system of membranous structures within the cytoplasm of the host cell were described for example in erythrocytes infected with *P. malariae*, *P. vivax*, *P. ovale* and *P. brasilianum*. This tubularvesicular system was also suggested to be involved in protein transport (Matsumoto *et al.*, 1986; Aikawa, 1988).

c) Modifications of the erythrocyte cytoskeleton

First reports that infection with *Plasmodium* involves a change in the mechanical properties of the infected erythrocyte were made by Miller *et al.* (1971, 1972) who observed impaired flow rates through micropore filters of erythrocytes infected with *P. knowlesi* and *P. coatneyi*. Cranston *et al.* (1983), studying the behaviour of individual cells in a rheoscope, found that erythrocytes infected with ring-stage parasites were less deformable than uninfected erythrocytes and that erythrocytes infected with trophozoite- and schizont-stages did not deform. The loss of deformability as the parasite matured from the ring- to the trophozoite-stage occurred in knobless and knob-forming strains and was not associated with the formation of knobs. In this study the specific factors that reduce deformability were not investigated, so the mechanism responsible for the loss of deformability remained uncertain. Nash *et al.* (1989) used micropipette aspiration techniques to obtain information on the deformability of individual uninfected

erythrocytes or erythrocytes infected with *P. falciparum*. They showed that the loss of deformability of infected erythrocytes was partially due to a stiffening of the erythrocyte membrane.

Increasing membrane rigidity generally reflects rearrangements in the structure of the membrane skeleton (Evans *et al.*, 1980). Taylor *et al.* (1987) showed in electron microscopy studies on erythrocyte skeletons electron-dense patches of material present directly on (probably) spectrin filaments. This electron-dense material was seen in erythrocytes infected with knob-less and knob-forming strains of *P. falciparum* but was absent in uninfected erythrocytes. The authors concluded that proteins of parasite origin associate with the erythrocyte skeleton and could be responsible for some of the physical changes observed in the membranes of erythrocytes infected with *Plasmodia*.

It seems likely that at least part of the proteins of the parasite which interact with the erythrocyte skeleton are transported to the erythrocyte membrane in a similar fashion as the components of the knob structure.

1.3. The blood stage asexual cycle of *Plasmodium* at the molecular level and the proteins involved

Some of the ultrastructural findings described in 1.4.1. can be correlated with the action of certain characterized plasmodial proteins and, in some cases, be analysed down to the level of a group of amino acids, which might act as a functional domain. Because of the widespread occurrence of repeated sequences in plasmodial proteins, independent of their cellular localization, this section starts with a short description of the nature and the possible origin of these motifs.

1.3.1. Repeated sequences in proteins of *Plasmodium*

A common feature of many plasmodial protein antigens is the occurrence of tandemly

(head-to-tail) repeated sequences. Though repeated sequences in proteins are frequently encountered in nature, it is the large number of plasmodial proteins carrying these motifs that is unusual (Kemp *et al.*, 1987; Weber, 1988).

One way to 'classify' the different repeat containing plasmodial proteins is by the number of regions that are repeated and the nature of the repeat: some proteins contain a single repeat region, like the soluble-antigen (S-antigen) (Cowman *et al.*, 1985; Saint *et al.*, 1987) and the major merozoite surface protein-1 (MSP-1) (Holder *et al.*, 1985a). Some proteins contain two repeat regions, like the ring-infected erythrocyte surface antigen (RESA) (Cowman *et al.*, 1984; Favaloro *et al.*, 1986) and some contain three repeat regions, like the mature-parasite-infected erythrocyte surface antigen (MESA) (Coppel , 1992).

The repeat regions can be short, like in MSP-1, where it constitutes only a minor part of the molecule or rather extended like in the S-antigen. Tandemly repeated peptides can be highly conserved in one allele but vary considerably among different alleles, like in the S-antigens (Bickle and Coppel, 1992: Anders *et al.*, 1993) or, on the contrary, repeats may be highly degenerate along the molecule but conserved between two different isolates, like in RESA (Cowman *et al.*, 1984). In the three repeat regions of MESA exist seven blocks of different repeats, interestingly three of these repeat blocks show some sequence similarity to each other (Coppel, 1992).

Repeats generally evolve more rapidly than the remainder of the molecule (Wellems and Howard, 1986; Galinski *et al.*, 1987) and represent the more polymorphic region of a protein. Frontali and Pizzi (1991) proposed a model for the generation of repeats and showed that the distribution of regularities and defects, e.g. base substitution or the presence of related repeats of different length, along repetitive sequences conforms to some empirical rules. One of their findings was that an array of contiguous modified repeats can be recognized as a motif itself and be propagated tandemly as a new unit in a supra-repeat structure. This indicates that a general mechanism for the tandem propagation of a motif must exist.

In the case of the circumsporozoite surface protein (CS-protein), for example, protein repeats are conserved within a species and highly variable between species. This seems to exclude phenotypic selection, which acts to maintain the amino acid sequence, as a mechanism to maintain the repeats (Galinski *et al.*, 1987; Arnot *et al.*, 1988). A mechanism that acts at the DNA level was proposed by Arnot *et al.* (1988). In this model the low-thymidine content of the coding strand of the CS-protein repeat region is a direct consequence of a biased gene conversion, where each repeat unit constitutes a conversion domain. The correction of heterozygosity acts against the maintenance of thymidine and certain variant repeats have a conversion advantage, which leads to their propagation in a given allele. The mechanism of the proposed biased conversion is unknown. No current model of the generation and maintenance of repeated peptides in plasmodial proteins considers the conservation of a structure as a selection criterion for a repeated peptide.

1.3.2. Merozoite surface proteins

Merozoite surface proteins are considered to be important targets for the host's immune system and, consequently, to be possible candidates for the development of an antimalarial vaccine. Furthermore, it seems likely that some of merozoite surface proteins interact with erythrocyte receptors during invasion.

Probably the best studied merozoite surface protein at present is the major merozoite surface protein-1 (MSP-1) which was first described by Holder and Freeman (1981, 1982). MSP-1 is synthesized late in the erythrocytic cycle and varies in size from 185-205 kDa in different strains of *P. falciparum*. Functionally analogous proteins have been identified in other species of *Plasmodium* (Holder, 1988). MSP-1 is linked via a glycosylphosphatidylinositol (GPI) anchor at its C-terminus to the extracellular side of the merozoite plasma membrane. MSP-1 is processed at the end of schizogony and

some of the fragments remain on the surface of the free merozoite (Holder *et al.*, 1987; McBride and Heidrich, 1987). The GPI anchored 42 kDa C-terminal fragment is finally cleaved prior to or during invasion and a GPI anchored 19 kDa fragment, which is highly conserved among parasite lines, is carried into the newly infected erythrocyte. Interestingly the 19 kDa fragment contains cysteine residues with a spacing characteristic for two epidermal growth factor (EGF)-like domains (Blackman *et al.*, 1991).
Immunization studies with MSP-1 or the 19 kDa fragment alone showed a high degree of protection against infection with *Plasmodia* (Holder *et al.*, 1988; Daly *et al.*, 1993).

The merozoite surface antigen-2 (MSA-2), a 35-48 kDa protein in different strains of *P. falciparum*, is another GPI anchored merozoite surface protein (Ramasamy, 1987). It is less well studied than MSP-1 but it might be a protective antigen as well. Mice that were immunized with peptides from conserved parts of *P. falciparum* MSA-2 showed a substantial protection against an otherwise lethal challenge with *P. chabaudi* (Saul *et al.*, 1992).

The apical membrane antigen-1 (AMA-1) is a rhoptry protein which is membrane bound by a membrane-spanning segment. It was shown for *P. falciparum* (Peterson *et al.*, 1989; Narum and Thomas, 1994) and *P. knowlesi* (Deans *et al.*, 1984) that it is exported at around the time of schizont rupture to the merozoite surface .

Recently a novel 37-40 kDa antigen associated with *P. falciparum* merozoites, the secreted polymorphic antigen associated with merozoites (SPAM) was described (McColl *et al.*, 1994). Unlike MSP-1, MSA-2 and AMA-1 it is not an integral membrane protein. SPAM appears to be secreted into the parasitophorous vacuole, it is cleaved to a lower molecular form during merozoite maturation and becomes partially associated with the merozoite following schizont rupture. There are some indications

that SPAM is subsequently carried into newly invaded erythrocytes. SPAM contains 12 copies of a heptad repeat dispersed in three regions, which were found to be conserved in two different isolates. A 38-residue synthetic peptide comprising the first heptad region and five flanking residues at either end of the SPAM sequence was able to form a very stable α-helical structure in solution. It seems possible that the different heptad regions further form (intramolecular) coiled-coil superhelices (Mulhern *et al.*, 1995) which might reflect a specific function of this motif on the surface of the merozoite.

SPAM probably belongs to a group of malarial proteins that become associated with the merozoite surface during the development of the parasite. Some of these antigens are responsible for merozoite agglutination (immune clusters of merozoites-ICM) which occurs when parasites are grown *in vitro* in the presence of immune serum and can be found in immune complexes when these clusters are disrupted (Chulay *et al.*, 1987). The function of these antigens is uncertain. One possible role might be to assist in the release of merozoites as has been suggested for serine rich protein (SERP) and SERP H, members of a small family that are putative proteases or protease-like proteins (Knapp *et al.*, 1991).

1.3.3. Erythrocyte receptors and malarial ligands involved in invasion

Invasion of the erythrocyte by the malaria parasite requires proteins at the surface of the erythrocyte which can act as receptors for proteins present at the surface of the parasite.

a) Erythrocyte receptors for invasion by *Plasmodia*

Invasion of erythrocytes by *Plasmodia* is species specific, which suggests that merozoites recognize specific receptors on the surface of erythrocytes. *P. vivax*, a human malaria, and *P. knowlesi*, a simian malaria that infects human erythrocytes,

invade only human erythrocytes that express the Duffy blood-group antigen on their surface. Human erythrocytes missing Duffy blood group determinants are resistant to invasion by *P. vivax* and *P. knowlesi* but can be infected with *P. falciparum* (Miller *et al.*, 1975, 1976). However, it was shown that *P. knowlesi* merozoites can attach to Duffy negative erythrocytes as well but eventually the merozoites detach and interact with another erythrocyte. This suggests that at least two receptors exist, one for the attachment and a second, the Duffy blood group determinant, for the interiorization of the merozoite (Miller *et al.*, 1975).

In invasion studies with *P. falciparum* treatment of erythrocytes with neuraminidase reduces the efficiency of invasion implicating sialic acid as a ligand or part of a ligand present on the surface of erythrocytes (Miller *et al.*, 1977). Of the erythrocyte membrane proteins, glycophorins A and B, which contain about 70-80% of the erythrocyte membrane sialic acid, are extensively represented on the erythrocyte surface. Human En (a-) erythrocytes, which lack glycophorin A, are partially resistant to invasion with *P. falciparum* (Pasvol *et al.*, 1982a). S-s-U erythrocytes, which are deficient in glycophorin B, show less invasion (78%) than normal erythrocytes. In S-s-U cells treated with trypsin, which makes these cells glycophorin A deficient as well, invasion was reduced to 5% (Pasvol *et al.*, 1982b). These results suggest that glycophorins A and B have a specific role during invasion by *P. falciparum.* Mitchell *et al.* (1986) observed that the dependence on sialic acid for invasion varies between different strains of *P. falciparum.* The Thai-Tn strain invaded with 50% efficiency (compared to untreated erythrocytes) neuraminidase treated erythrocytes whereas the Camp strain invaded neuraminidase treated erythrocytes with 5% efficiency. The ability of the Thai-Tn strain to invade erythrocytes was reduced to 7% if the erythrocytes were trypsin treated. The results suggest the presence of at least two different receptors on the surface of erythrocytes for invasion by *P. falciparum*, one of which is dependent on sialic acid and a second one which is sialic acid independent but

trypsin sensitive. Furthermore these two receptors may be recognized by two different parasite proteins.

Similar results were obtained by Hadley *et al.* (1987) but they further demonstrated that part of the sialic acid used by the Camp strain was not located on glycophorin A, B or C but on an as yet uncharacterized glycolipid or a trypsin-insensitive glycoprotein.

It seems evident that different strains of *P. falciparum* can express different proteins that can recognize a variety of erythrocyte receptors and that different pathways of invasion should exist.

b) Plasmodial proteins that interact with erythrocyte receptors for invasion

The parasite proteins that interact with the Duffy blood group determinants in *P. knowlesi* (Haynes *et al.*, 1988) and *P. vivax* (Wertheimer and Barnwell, 1989) and with the sialic acid in *P. falciparum* (Camus and Hadley, 1985) were identified and subsequently cloned and sequenced (Sim *et al.*, 1990; *Fang et al.*, 1991; Adams *et al.*, 1992). There is one gene in *P. falciparum* coding for the protein that binds to sialic acid, the erythrocyte binding antigen- 175 (EBA-175), one gene in *P. vivax* coding for the protein binding to the Duffy blood group determinants and there are three genes in *P. knowlesi* coding for proteins binding to the Duffy blood group determinants. These genes have a similar intron/exon structure, where each of the exons represents a functional domain, as is often the case for other eukaryotic genes. Besides a similar gene structure, all of them show similarities in their amino acid sequence in the 5' and 3' cysteine-rich region with highly conserved cysteine and aromatic amino acid residues. Recently the 5' cysteine region was identified as the red blood cell binding domain of *P. knowlesi*, *P. vivax* (Chitnis and Miller, 1994) and *P. falciparum* (Sim *et al.*, 1994).

The similarities of the EBA-175 protein of *P. falciparum* with the Duffy binding proteins of *P. vivax* and *P. knowlesi* suggest a common origin of their genes and that

they belong to the same family of genes with different specificities.
Interestingly EBA-175 (Sim *et al.*, 1992) and the Duffy binding family of *P. knowlesi* (Adams *et al.*, 1990) were immunolocalized within the micronemes and it was shown that these proteins are not present on the surface of invasive merozoites. They may be released after free merozoites are attached to erythrocytes and apical reorientation has occurred. This would be in agreement with the observation that attachment and reorientation of *P. knowlesi* merozoites occur equally well to Duffy-negative and -positive human erythrocytes but junction formation and invasion occurs only with Duffy-positive erythrocytes (Miller *et al.*, 1975).

The reticulocyte binding protein-1 and 2 (RBP-1 and 2) from *P. vivax* were identified due to their property of binding specifically to reticulocytes, whether or not the Duffy blood group determinants were expressed on the surface of reticulocytes (Galinski *et al.*, 1992). The 280 kDa RBP-1 and 250 kDa RBP-2 are located at the apical surface of the merozoite, perhaps following translocation through the microneme organelles. It was mentioned before that after initial binding to the erythrocyte the merozoite reorientates so that its apical end adjoins the erythrocyte membrane. Before forming a junction 40-120 nm long filaments, emanating from the parasite's apical end, attach to the erythrocyte surface (Mitchell and Bannister, 1988). Following this contact junction formation can take place. Galinski and co-workers proposed that the RBP-1 and 2 might represent the long filaments seen in the electronmicroscopy studies. Once a reticulocyte is detected they could then trigger the release of other ligands, like the Duffy binding proteins, which in turn would initiate the formation of a junction in Duffy positive erythrocytes.
This model correlates with the observation that *P. knowlesi* invades primarily, if not exclusively, reticulocytes (Kitchen,1938) and can only invade Duffy positive erythrocytes. Furthermore this 'testing the water' role proposed for RBP-1 and 2 should exist in malaria parasite species that are not restricted to reticulocytes (Galinski

et al., 1992), e.g. *P. falciparum* and might be a good target to block lethal infections. The protein sequence of RBP-1 has a putative N-terminal signal sequence, a membrane spanning domain in its C-terminal region and lacks repeat motifs. It is highly hydrophilic, its predominant structures are α-helices and it can form dimers which are linked by intermolecular disulphide bonds. The two Arg- Gly- Asp (RGD) motifs, which exist in the protein, may represent adhesive domains (Ruoslahti and Pierschbacher, 1986).

Interestingly RBP-2 has a significant similarity to the Py 235 kDa rhoptry protein of *P. yoelii*, which will be described in more detail in section 1.3.4.b). Py 235 kDa might be a merozoite ligand which is released during invasion from the rhoptries and interacts with a receptor present only on the surface of mature erythrocytes (Freeman *et al.*, 1980a; Keen *et al.*, 1994).

The 42 kDa fragment of *P. falciparum* MSP-1 (see 1.3.2.) is cleaved prior to or during invasion and the remaining membrane-bound 19 kDa fragment is carried into the newly infected erythrocyte. Monoclonal antibodies directed against the 19 kDa part can block the processing step of the 42 kDa fragment and at the same time inhibit invasion *in vitro* (Blackman *et al.*, 1990, 1994). This suggests a participation of the 42/19 kDa part of MSP-1 in the process of invasion.

As was mentioned above, AMA-1 becomes associated with the merozoite surface at about the time of schizont rupture. Two rat monoclonal antibodies raised against AMA-1 of *P. knowlesi* blocked *in vitro* invasion. Furthermore *in vitro* invasion was also blocked with only the Fab fragments of these antibodies, which indicates that the mechanism is not dependent upon merozoite agglutination. The antibodies may block merozoite attachment to specific red cell receptors (Thomas *et al.*, 1984).

1.3.4. The process of invasion/merozoite release and the organelles of the apical complex

There is evidence that the organelles of the apical complex release some proteins that are involved in the process of invasion or in events that occur shortly after invasion. Some of these proteins, amongst others, may as well be involved in the processes that are responsible for the release of mature merozoites from the infected erythrocyte.

a) The proteinaceous contents of the rhoptries

It was shown by SDS-PAGE analysis that the rhoptries of *P. falciparum* contain numerous proteins with a range of about 18 - >200 kDa, some of which are characterized but with many being yet uncharacterized (Etzion *et al.*, 1991). Solubility studies with Na_2CO_3 and high-salt, and enzyme treatment with trypsin or phospholipase A_2 and C further suggest that at least some of these proteins are associated with lipids or may be membrane bound (Etzion *et al.*, 1991). This is in agreement with the localization of Rhop-H3 from the high molecular mass complex (Rhop-H1-3) and the rhoptry associated protein-1 (RAP-1) to concentric membranous whorls that appear to be associated with merozoites from the rupturing schizont or free merozoites (Sam-Yellowe *et al.*, 1988; Bushell *et al.*, 1988).

Rhoptries are formed *de novo* in each cycle of intraerythrocytic development and this appears not to be co-ordinate with nuclear division. In *P. falciparum* rhoptries isolated at the four-nucleus stage have a density of 1.12 g ml^{-1} which increases to 1.16 g ml^{-1} in the mature 16 nucleus segmenter (Jaikaria *et al.*, 1993). The start of packaging into the rhoptries is not the same for different rhoptry proteins, e.g. Rhop-H2 and 3 are contained in early rhoptries when RAP-1 is still absent.

Immunoelctronmicroscopic studies showed that Rhop-H3 is present in early immature

rhoptries that seem to be attached to a reticular network, which may be an irregular endoplasmatic reticulum. In the same study rhoptries often appeared as pairs as if they were budding from a common origin and this was suggested by the authors to be a possible mechanism of regulating the number of rhoptries per merozoite to two (Jaikaria *et al.*, 1993).

There is evidence that the Py 235 kDa rhoptry protein passes through a Golgi during synthesis and probably the sequence is: endoplasmatic reticulum → Golgi apparatus → rhoptry organelles (Ogun and Holder, 1994).

Calmodulin has been shown to be present in the cytoplasm of the parasite as well as within the ductules of rhoptries of *P. falciparum* (Scheibel *et al.*, 1987). Calmodulin serves as an intracellular receptor for Ca^{2+} and in doing so regulates a wide variety of enzymes (Means, 1988). It is possible that some proteins of the rhoptries become activated by a Ca^{2+}/calmodulin-dependent process which might trigger finally the release of lipids and proteins into the attached erythrocyte (Tanabe, 1990).

Direct evidence for the secretion of rhoptry proteins into the attached erythrocyte has been reported for the Rhop-H3 protein (Sam-Yellowe *et al.*, 1988).

b) Characterized proteins of the rhoptries

A well studied group of rhoptry proteins is a high molecular mass complex of three rhoptry proteins (Rhop-H1-3) and a low molecular mass complex of two rhoptry proteins (RAP-1 and 2).

The Rhop-H1-3 proteins were first described by Holder *et al.* (1985). A monoclonal antibody raised against *P. falciparum* schizonts immunoprecipitated a complex of three proteins of 155, 140 and 110 kDa (Rhop-H1-3) but reacted in Western blotting studies only with the 140 kDa protein. Peptide mapping showed that the 155 kDa and 140 kDa

are unrelated proteins. Other groups independently identified the same complex. Lustigman *et al.* (1988) immunoprecipitated with human antibodies affinity-purified on a cDNA clone coding for a part of Rhop-H3 (Coppel *et al.*, 1987) three proteins of 150, 135 and 110 kDa. Peptide mapping of Rhop-H1-3 indicated that the three proteins were unrelated proteins. Furthermore the monoclonal antibody from Holder et al. (1985) which was reactive with the 140 kDa protein specifically detected the 135 kDa protein of the 150, 135 and 110 kDa complex described by Lustigman *et al.* (1988) suggesting that both complexes were identical. Furthermore both complexes were located to the rhoptries and were synthesized at the late trophozoite- and schizont-stage (Holder *et al.*, 1985; Coppel *et al.*, 1987; Lustigman *et al.*, 1988). Probably the same complex was described by Cooper *et al.* (1988) who also demonstrated that the three proteins consist of discrete polypeptides.

Rhop-H3 is located in the matrix of the rhoptry organelles and on membraneous whorls secreted from the parasite. During invasion it is released from the rhoptries and spreads around the membrane of the bound erythrocyte. Another portion of Rhop-H3 appears to be associated with the membrane of the parasitophorous vacuole of the newly invaded ring-stage parasite (Sam-Yellowe *et al.*, 1988). Rhop-H1 and 2 are carried into the newly invaded erythrocyte as well (Lustigman *et al.*, 1988). It was shown that all three proteins of the Rhop-H complex can bind to erythrocyte membranes or specifically to liposomes made of phosphatidylinositol (PI) or phosphatidylserine (PS). Both PI and PS are localized preferentially on the inner leaflet of the plasma membrane of the erythrocyte and it was suggested that the proteins of the Rhop-H complex may help to disrupt the interactions of the skeleton of the erythrocyte with transmembrane proteins (Perkins and Ziefer, 1994).

The gene encoding Rhop-H3 contains seven exons, the most complex intron-exon structure of all known malarial genes, without any evidence that they represent functional domains (Brown and Coppel, 1991). The protein starts with a signal sequence which was shown to be cleaved after position 24 (Cooper *et al.*, 1989) but no

transmembrane sequence exists. The 13 cysteine residues present in the protein are all located within the N-terminal half of the molecule, similar in arrangement to some receptors and appear to form intramolecular disulphide bonds. Unlike other malarial antigens Rhop-H3 does not contain any repetitive elements and is a non-polymorphic protein (Brown and Coppel, 1991).

The rhoptry associated protein-1 and 2 (RAP-1 and 2) of *P. falciparum* are first detectable at the schizont-stage. RAP-1 is synthesized as a 84 kDa protein that is rapidly cleaved to a 80 kDa form. The 80 kDa protein then associates with RAP-2, a 42/40 kDa doublet (Clark *et al.*, 1987; Bushell *et al.*, 1988). During late schizont-stage or merozoite release some of the 80 kDa protein is further processed to a 65 kDa form, but following release the 65 kDa form is not present in free merozoites anymore (Bushell *et al.*, 1988; Harnyuttanakorn *et al.*, 1992). After invasion only the 80 kDa form of RAP-1 is carried into the parasitophorous vacuole of the newly invaded ring-stage (Schofield *et al.*, 1986).

Harnyuttanakorn *et al.* (1992) observed that monoclonal antibodies directed against epitopes adjacent to the 80/65 kDa processing site were able to inhibit *P. falciparum* multiplication *in vitro*. They suggest that binding of these monoclonal antibodies interferes with proteolytical cleavage of RAP-1 and that this cleavage is a requirement for RAP-1 to function during merozoite release. The authors propose that the 80 kDa protein performs a different function compared to its 65 kDa fragment such that the 65 kDa protein acts during merozoite release, whereas the 80 kDa protein is involved in the invasion process of an erythrocyte.

Sequencing of RAP-1 (Ridley *et al.*, 1990, 1991) showed that the protein starts with a putative N-terminal signal sequence and contains no membrane spanning domain. In the N-terminal region are dispersed five sequences related to the motif KSSSPS but tandemly repeated sequences are lacking. In the second half of the molecule is a cysteine-rich region, flanked by regions of potential amphiphilic helices which might be

able to interact with membrane structures. N-terminal sequencing of the 65 kDa protein fragment further revealed that the processing site of the 80 kDa protein falls between the region which contains the KSSSPS motif and the C-terminal part which contains the potential amphiphilic helices and the cysteine-rich region. Whether or not these two different regions contribute to selective function during merozoite release and invasion is unknown.

The sequence of RAP-2 (Ridley *et al.*, 1991; Saul *et al.*, 1992) has similarities to the sequence of RAP-1 in the sense that it starts with a N-terminal signal sequence, it has no membrane spanning domain, lacks tandemly repeated sequences, contains cysteine-residues which form disulphide bonds (Bushell *et al.*, 1988) and is non-polymorphic. RAP-1 and 2 both contain a number of moderately hydrophobic domains which probably account for their association into a complex as well as for the association with membranous material from the rhoptries (Bushell *et al.*, 1988).

The apical membrane antigen-1 (AMA-1) is a protein of 83 kDa in *P. falciparum* and 66 kDa in *P. knowlesi*. In *P. falciparum* AMA-1 is comparable in size and shows a similar processing pattern as RAP-1 but the expression of RAP-1 precedes the expression of AMA-1 slightly and AMA-1 is localized within the rhoptry neck whereas RAP-1 is localized within the body of the rhoptry (Crewther *et al.*, 1990). AMA-1 is expressed only during the last 4h of schizogony as a 83 kDa protein a portion of which undergoes rapid N-terminal processing to a 66 kDa form. Both forms remain restricted to the rhoptries until the release of merozoites occurs when a portion of the 66 kDa protein becomes associated with the circumference of free merozoites and the 80 kDa protein and the rest of the 66 kDa protein remain in the rhoptry organelles. During invasion a processed form of the 80 kDa protein (presumably the 66 kDa protein) is carried into the newly infected erythrocyte where it can be detected in the parasitophorous vacuole of the ring stage for about 3h (Narum and Thomas, 1994).

The sequence of events described here for AMA-1 bears some similarities with the

release of RAP-1. For both the first step is the processing of a portion of a precursor molecule followed by a controlled discharge where some of the precursor remains in the rhoptry organelles. However, the 66 kDa fragment of AMA-1 remains on the surface of the free merozoite whereas the processed form of RAP-1 does not, but both molecules (the processed form of AMA-1 and the 80 kDa form of RAP-1) are carried into the parasitophorous vacuole of the newly invaded erythrocyte. These events support the idea of a dual function of some rhoptry proteins. One may be to enable the parasite to leave the erythrocyte and another to render invasion of an erythrocyte possible (Peterson *et al.*, 1989; Crewther *et al.*, 1990; Narum and Thomas, 1994).
The coding region of the AMA-1 gene is intronless in *P. falciparum* (Peterson *et al.*, 1989) and *P. knowlesi* (Waters *et al.*, 1990). The protein sequence starts with a putative N-terminal signal sequence and contains at its C-terminus a hydrophobic stretch which might act as a membrane spanning domain. Repeated peptide sequences are lacking in the protein and AMA-1 shows a very low degree of polymorphism. AMA-1 of *P. falciparum* contains a unique N-terminal part, compared to AMA-1 of *P. knowlesi*, which might reflect a specific function of this part in *P. falciparum*. The rest of the molecule shows a similarity of greater than 80%, when identical and conservative substitutions are included (Waters *et al.*, 1990). The cysteine-residues are all conserved in AMA-1 of both species, they probably form intramolecular disulphide bonds and seem to be essential for the function of the protein.

The Py 235 kDa protein of *P. yoelii* is so far the only rhoptry protein of a rodent malaria parasite of known primary structure (Keen *et al.*, 1994) and is part of a multigene family of at least 11 members located on different chromosomes (Borre *et al.*, 1995). It has a continuous open reading frame and the deduced protein possesses a putative N-terminal signal sequence and a possible membrane spanning domain close to the C-terminus with a remaining cytoplasmic domain of 46 amino acids. It lacks extended repeats except for a short region of 6 repeats based on Asp, Ile, Asn (D, I, N)

close to the N-terminal side of the putative membrane spanning domain. Py 235 kDa contains 11 cysteine residues, it is extremely hydrophilic and consists almost exclusively of regions that form α-helices. Py 235 kDa shows a low but significant similarity to the reticulocyte binding protein-2 (RBP-2) and to a lesser degree to RBP-1 (Keen *et al.*, 1994). An indication that Py 235 kDa itself might be a ligand for an erythrocyte receptor(s) comes from *in vivo* passive transfer studies made with monoclonal antibodies directed against this protein (see next section) and direct binding experiments (S. Ogun, unpublished observations). *P. yoelii* YM normally infects both mature erythrocytes and reticulocytes. In the presence of monoclonal antibodies against the Py 235 kDa protein *P. yoelii* YM no longer infects mature erythrocytes and becomes restricted to reticulocytes (Freeman *et al.*, 1980a). It appears that Py 235 kDa is one of the proteins which determine host cell specificity and thus acts as parasite ligand.

To the group of rhoptry proteins of unknown primary structure belongs the p76 protein of *P. falciparum*. The p76 protein is membrane bound via a glycosyl-phosphatidylinositol (GPI) anchor and is first synthesized as a 83 kDa precursor which is rapidly processed to a 76 kDa form. Braun-Breton *et al.* (1988) showed that the membrane bound p76 becomes an active soluble serine protease only after cleavage from its GPI anchor by phosphatidylinositol-phospholipase C (PI-PLC) whereas the membrane bound p76 did not possess proteolytic activity. The inactive form of p76 is already present in mature trophozoites, but endogenous activation occurs only in late schizogony.

Interestingly a serine protease seems to be involved in the process of invasion, and invasion can be blocked by serine protease inhibitors (Dluzewski *et al.*, 1986). A monoclonal antibody against the p76 protein inhibits invasion *in vitro* as well (Perrin and Dayal, 1982) suggesting that p76 might have a function as a serine protease during invasion.

Another GPI anchored 55 kDa protein of *P. falciparum* was localized in the rhoptries by immunofluorescence microscopy (Smythe *et al.*, 1988). The authors suggest that the 55 kDa protein associates with components of the erythrocyte skeleton during the process of invasion.

A monoclonal antibody raised against culture supernatant components of *P. falciparum* identified a 240 kDa protein which was processed to a 225 kDa form and synthesized only during schizogony (Roger *et al.*, 1988). Complete processing occurred only after a period of 4h and suggests an accumulation of the precursor before the processing step. Immuno-electronmicroscopic detection of the 240 kDa protein in mature merozoites showed a localization in the neck of the rhoptry organelles.

c) Antigenicity of rhoptry proteins

First indications that the proteins located in the rhoptries are relatively immunogenic came from the observation that a high proportion of monoclonal antibodies raised against lysates of whole parasites were reactive with rhoptry proteins (Campbell *et al.*, 1984; Holder *et al.*, 1985; Schofield *et al.*, 1986).

A possible role of the rhoptry proteins as protective antigens became evident as many of the monoclonal antibodies were able to block invasion *in vitro* or were protective *in vivo*. This was first revealed by Freeman *et al.* (1980b) who showed that an Ig G2a monoclonal antibody, reactive with the Py 235 kDa protein described above, was protective upon passive transfer in mice challenged with the lethal *P. yoelii* YM strain. Monoclonal antibodies which block invasion *in vitro* were further reported, for example, for the Rhop-H complex (Cooper *et al.*, 1988), RAP-1 (Schofield *et al.*, 1986; Harnyuttanakorn *et al.*, 1992) and p76 (Perrin and Dayal, 1982) of *P. falciparum* and for AMA-1 of *P. knowlesi* (Thomas *et al.*, 1984).

Animals immunized with native rhoptry proteins were partially or completely protected against infection with different species of *Plasmodia*. Mice immunized with the Py 235 kDa protein were completely protected against challenge with *P. yoelii* YM (Holder and Freeman, 1981). The three proteins from the Rhop-H complex protected *Aotus* monkeys partially against infection with *P. falciparum* (Siddiqui *et al.*, 1987). Ridley *et al.* (1990) demonstrated that RAP-1 and 2 can protect *Saimiri* monkeys from lethal infection with *P. falciparum* and Deans *et al.* (1988) showed very strong protective responses in rhesus monkeys immunized with AMA-1 isolated from *P. knowlesi* in combination with exposure to the parasite.

Rhoptry proteins not tested in invasion inhibition or protection studies were shown to be natural antigens, e.g. the 55 kDa protein (Smythe *et al.*, 1988) and this might be a common feature of rhoptry proteins due to their character of secreted proteins.

Finally the finding that rhoptry proteins can be targets of protective immunity emphasises the important function(s) of these proteins in the development of the parasite.

d) Characterized proteins of the micronemes

Amongst the malarial proteins that were localized in the micronemes of blood-stage parasites are the Duffy binding proteins of *P. knowlesi* (Adams *et al.*, 1990), the erythrocyte binding antigen-175 (EBA-175) of *P. falciparum* (Sim *et al.*, 1992) and possibly the reticulocyte binding protein-1 and 2 (RBP-1 and 2) of *P. vivax* (Galinski *et al.*, 1992). It was suggested that these molecules function as ligands for erythrocyte receptors and that they are released prior to the secretion of rhoptry proteins (Adams *et al.*, 1990; Sim *et al.*, 1992; Galinski *et al.*, 1992).

The proposed function of micronemes to store and release parasite proteins (see 1.2.3.) that are ligands for receptors is further underlined by the finding that the circumsporozoite protein (CS protein) and the sporozoite surface protein 2 (SSP 2),

both natural antigens present on the surface of sporozoites of *Plasmodia*, are located in the micronemes as well (Fine *et al.*, 1984; Rogers *et al.*, 1992). The CS protein (Dame *et al.*, 1984; Kobayashi *et al.*, 1986; Abey *et al.*, 1986) and SSP 2 (Müller *et al.*, 1993) have many characteristics of a receptor binding protein .

e) Characterized proteins of the dense granules

The 155 kDa ring-infected erythrocyte surface antigen (RESA) of *P. falciparum* is located in the dense granules (Aikawa *et al.*, 1990). During the ring-trophozoite-stage RESA is exported into the parasitophorous vacuole, traverses the parasitophorous vacuolar membrane, is routed through the erythrocyte cytoplasm and becomes finally associated with the erythrocyte plasma membrane. RESA is apparently transported in clustered vesicles which might involve membranes (Culvenor *et al.*, 1991).
In studies with various erythrocyte membrane preparations RESA was found to bind to spectrin (Foley *et al.*, 1991; Ruangjirachuporn *et al.*, 1991) and to become phosphorylated upon binding (Foley *et al.*, 1990).
By using recombinant RESA fragments with progressive deletions of the N-terminal and C-terminal regions of the RESA molecule the putative spectrin-binding region was located within a 100 amino acid sequence (Foley et al., 1994). N-terminal to the proposed spectrin-binding region lies a module of the DnaJ heat shock protein, a member of the family of molecular chaperons (Bork *et al.*, 1992). One of the functions of DnaJ proteins is to stabilize the conformation of other proteins (Langer *et al.*, 1992). Spectrin undergoes an irreversible conformational change upon heating above 49 °C. Interestingly a recombinant RESA fragment, which contains the possible spectrin-binding region and the DnaJ region can partially abrogate the heat-induced conformational changes in spectrin in solution (Da Silva *et al.*, 1994).
It was shown in the same study that erythrocytes infected with RESA (+) parasites are less susceptible to heat-induced vesiculation than erythrocytes infected with RESA (-)

oxidative stress (Da Silva *et al.*, 1994).
RESA is invariably expressed in field isolates (Perlman *et al.*, 1987) and RESA (-) parasites exist only in culture. This demonstrates, for example, that RESA is not prerequisite for invasion.

At present RESA is the only well characterized protein that could be localized inside the dense granules. The dense granules are exocytosed after invasion and other proteins may remain in the parasitophorous vacuole or be inserted into the parasitophorous vacuolar membrane where they might alter the permeability of this membrane to the benefit of the parasite (Perkins, 1992).

In the above sections a short description was given of how proteins from the organelles of the apical complex are released at different stages of parasite invasion, multiplication or schizont rupture. One important function of some of these proteins appears to be to modify the erythrocyte skeleton at various moments in order to adapt it to the needs of the parasite. Other proteins than the ones contained in the organelles of the apical complex can contribute to these modifications as well and for some of them the interaction with the erythrocyte skeleton has been studied in more detail.
Because both groups of proteins might have common ways of interaction with the skeleton, analysis of these mechanisms in one group might further the understanding of the modes of action of the other.

1.3.5. Malarial proteins that are located outside of the organelles of the apical complex and interact with the erythrocyte cytoskeleton

On infecting the erythrocyte the malaria parasite synthesises a variety of proteins that become associated with the membrane of the infected erythrocyte. These malarial proteins can perform very different functions, e.g. change the permeability for nutrients

become associated with the membrane of the infected erythrocyte. These malarial proteins can perform very different functions, e.g. change the permeability for nutrients essential for parasite growth, like the transferrin receptor (Haldar et al., 1986; Rodriguez and Jungery, 1986), contribute to cytoadherence, like the *P. falciparum* erythrocyte membrane protein 1 (Pf EMP 1) (Howard *et al.*, 1987) or interact with the erythrocyte skeleton. The list of latter proteins is constantly growing but only for a few of them has the interaction with the skeleton been studied in some detail.

The most prominent alteration of the erythrocyte membrane occurs during the trophozoite-stage of infection with certain species of *Plasmodium*, e.g. *P. falciparum* and is the development of knobs whose appearance and morphology was described before (1.2.4.). The attachment of late stage erythrocytes infected by *P. falciparum* to capillary endothelial cells is believed to be mediated by proteins on the surface of the infected cell above the knobs (Aikawa *et al.*, 1988). Some of the major protein constituents of isolated knobs were found to be spectrin, protein 4.1 and the 48 and 52 kDa subunits of dematin. Other proteins like band 3, ankyrin, protein 4.2 or the malarial protein MESA were not identified but their presence in knobs can not be excluded (Chishti *et al.*, 1992).

The 80-108 kDa histidine-rich protein-1 (HRP-1) is a malarial protein which is localized in the electron-dense knob structure (Culvenor *et al.*, 1987; Taylor *et al.*, 1987) and it was further shown that in knobless strains of *P. falciparum* the gene encoding HRP-1 is deleted (Pologe and Ravetch, 1986). Certain regions of HRP-1 show similarities in sequence and charge distribution with the spectrin-actin-binding domain of the erythrocyte skeleton protein 4.1 (Rashid *et al.*, 1990). A recombinant part of HRP-1, which contained these regions of similarity can cross-link or oligomerize spectrin and also binds to actin in solution (Kilejian et al., 1991).

The 4.1 protein links spectrin to the transmembrane proteins band 3 and glycophorin and might be a sensitive target for rearrangements of the erythrocyte skeleton.

A second protein which might be involved in the formation of the knob structure is the 315 kDa *P. falciparum* erythrocyte membrane protein 3 (Pf EMP 3) (Pasloske *et al.*, 1993). A monoclonal antibody reacting with Pf EMP 3 and sera raised against a β-galactosidase fusion protein containing the C-terminal part of Pf EMP 3 reacted in immuno electronmicroscopy only with knob producing strains of *P. falciparum.* The monoclonal antibody reacted only with areas close to or under knobs. In contrast to the monoclonal antibody the anti fusion protein antibodies also reacted with areas between the knobs (Pasloske *et al.*, 1993). It was shown that the gene encoding Pf EMP 3 was deleted in knobless strains of *P. falciparum* and Pf EMP 3 was not expressed in these parasites, whereas the Pf EMP 3 gene product was detected in all knob-forming lines. From its partial solubility in SDS it was further suggested that Pf EMP 3 interacts with a protein(s) of the erythrocyte skeleton (Pasloske *et al.*, 1993).

Another well characterized malarial protein that binds to the erythrocyte skeleton is the mature-parasite-infected erythrocyte surface antigen (MESA, Pf EMP 2) of *P. falciparum.* MESA is polymorphic varying in size from 250-300 kDa in different isolates and shows antigenic diversity (Coppel *et al.*, 1986; Howard *et al.*, 1988a; Coppel, 1992). It is synthesized by mature stages of the parasite, transported to the inner face of the erythrocyte membrane (Howard et al., 1987) and phosphorylated mainly at its serine residues (Howard et al, 1988b). MESA is not efficiently solubilized by the non ionic detergent Triton X-100 and was shown to associate strongly with band 4.1 of the erythrocyte skeleton, the latter becomes strongly phosphorylated upon binding with MESA (Lustigman *et al.*, 1990).

Band 4.1 assists in binding the spectrin-actin skeleton to the transmembrane protein band 3 and glycophorin (Anderson and Marchesi, 1985). MESA shows sequence similarities to fibrous proteins, e.g. mouse neurofilament L and mouse myosin and long regions of predicted α-helices are present in MESA , suggesting that it is a fibrous

protein. MESA is similar in size to α- and β-spectrin, shares its ability to bind to band 4.1 and might have a similar function during the mature stage of parasite development (Howard et al., 1988b; Coppel, 1992).

A pair of 89 kDa phosphoproteins of *P. falciparum* associates with the membrane of infected erythrocytes during the entire blood-stage cycle but phosphorylation of these proteins is restricted to the trophozoite- and schizont-stage. The two proteins show very similar peptide maps and differences between them seem to be due to modifications other than phosphorylation (Suetterlin *et al.*, 1991). Three tryptic peptides of the 89 kDa proteins were shown to have identical amino acid sequences, except for a single amino acid, as parts of ankyrin, including the ankyrin repeat region (Suetterlin *et al.*, 1992). The ankyrin repeat, located on a 89 kDa fragment obtained by treatment of ankyrin with chymotrypsin, is a 33 amino acid motif which is repeated 22 times and binds to the transmembrane protein band 3 (Lux *et al.*, 1990). Ankyrin binds further to spectrin and anchors it this way to the erythrocyte membrane.

The authors propose that the 89 kDa proteins are analogues of the 89 kDa domain of ankyrin and could, upon phosphorylation, replace ankyrin from binding to spectrin and band 3; the latter loses its binding ability for spectrin and band 3 upon phosphorylation. This would occur from the trophozoite-stage onwards and might correlate to the increasing stiffening of the membrane of infected erythrocytes observed during this period (Cranston *et al.*, 1984; Nash *et al.*, 1989).

This kind of molecular mimicry of host proteins by malarial proteins was already described for the HRP-1 above where similarities in sequence and charge distribution between two regions of HRP-1 of *P. falciparum* and the spectrin-actin binding domain of band 4.1 were observed (Rashid *et al.*, 1990; Kilejian *et al.*,1991).

The evolution of motifs located on malarial proteins that are analogous to structures or binding domains of erythrocyte skeletal proteins might be an elegant and widely used strategy of *Plasmodia* to compete with host proteins in order to rearrange the host cell

skeleton to its benefit.

Finally it appears that the main target of malarial proteins which interact with the erythrocyte skeleton is spectrin, either by binding directly to it or by replacing it. This is not surprising because it is mainly the spectrin network that confers resistance against mechanical stress to the erythrocyte (Boal, 1994) and probably represents the main barrier that the parasite has to overcome either during invasion or merozoite release (Aikawa *et al.*, 1981). This is illustrated, for example, by the observation that the presence of anti-spectrin antibodies inside resealed human red cell ghosts or chemical cross-linking of spectrin molecules strongly inhibits invasion by *P. falciparum* (Olson and Kilejian, 1982; Dluzewski *et al.*, 1983).

Aim of the project

Clones from a cDNA library of *P. chabaudi* 96V were isolated which covered a 704 bp region coding for part of an uncharacterized malarial protein. The translated sequence showed a predicted α-helical coiled-coil structure, typical for cytoskeletal proteins. Little is known about this class of proteins in malaria parasites and their possible interaction with the erythrocyte skeleton.

To obtain the entire coding sequence of the protein the cDNA library constructed in λ ZAP II and a genomic DNA library constructed in pBluescript KS(+) with *Sau* 3A partially digested and size fractionated genomic DNA will be screened.

Different parts of the ORF will be expressed as recombinant proteins in vectors of the pGEX-T series. This will allow the study of whether or not the recombinant proteins react with hyper-immune sera, i.e. that the native protein is a natural antigen. Sera

raised against the recombinant proteins will be used to detect the native protein in western-blotting experiments and to localize it in erythrocytes infected with *P. chabaudi* 96V.

The secondary structure of the protein will be predicted, its hydrophilicity profile analyzed and searches for known protein motifs be done. Any structural peculiarity of the protein will be further analyzed, in order to relate it to a possible biological role of this novel protein.

MATERIALS AND METHODS

2. Materials

2.1. Reagents

Restriction endonucleases, T4 DNA ligase, Klenow fragment of DNA polymerase, T4 polynucleotide kinase, calf intestinal alkaline phosphatase, proteinase K were purchased from Boehringer-Mannheim, Bethesda Research Laboratories, New England Biolabs, Promega, Cambio and HT Biotechnology. *Taq* DNA polymerase and *Pfu* DNA polymerase were purchased from Promega and Stratagene. Ultra-pure nucleotides were purchased from Pharmacia. Radioactive isotopes used were: [γ-^{32}P]dATP (>5000 Ci mmol^{-1}), [α-^{32}P]dATP (>3000 Ci mmol^{-1}), ^{35}S-methionine (>1000 mCi mmol^{-1}), all purchased from Amersham. Nitrocellulose filters used for Western-, Southern-, Northern-blot analysis and screening of libraries were obtained from Schleicher and Schuell.

Salt-saturated phenol, phenol/chloroform /isoamyl alcohol (25:24:1) and chloroform /isoamyl alcohol were purchased from Amresco. Sea plaque low melting point agarose was purchased from FMC. General chemicals were obtained from BDH, Sigma and Pharmacia.

2.2. General buffers and media

PBS	phosphate buffered saline : 2.8 mM NaH_2PO_4, 7.2 mM Na_2HPO_4, 0.14 M NaCl pH 7.4
TE buffer	10 mM Tris /HCl pH 8.0, 1 mM EDTA pH 8.0
SM buffer	100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris/ HCl pH 7.5, 0.01 % gelatin
20 x SSC	3 M NaCl, 0.3 M Na_3citrate, pH 7.0 with 10 M NaOH
Denhardt	5 g Ficol, 5 g polyvinylpyrrolidone, 5 g BSA, H_2O to 500 ml

LB Broth	10 g bacto-tryptone, 5 g yeast extract, 5 g NaCl, H_2O to 1 litre, pH 7.5 with 1 M NaOH
Super Broth	35 g bacto-tryptone 20 g yeast extract, 5 g NaCl, H_2O to 1 litre, pH 7.5 with 1 M NaOH
SOC	20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl, 0.19 g KCl, H_2O to 1 litre, pH 7.0 with 1 M NaOH and autoclave. Add 5 ml sterile 2 M $MgCl_2$ and 20 ml sterile 1 M glucose
NZY Broth	5 g NaCl, 2 g $MgSO_4 \cdot 7 H_2O$, 5 g yeast extract, 10 g NZ amine (casein hydrolysate), H_2O to 1 litre, pH 7.5 with 1 M NaOH
2 x YT Broth	10 g NaCl, 10 g yeast-extract, 16 g bacto-tryptone, H_2O to 1 litre

2.3. Biological materials

2.3.1. Experimental animals

Balb/c and Parkes strains were obtained from the SPF colonies at the National Institute for Medical Research, London, England. Fischer rats were from the colony maintained at the Institut Pasteur de Lille, France.

2.3.2. *Plasmodium chabaudi chabaudi* 96V

In 1965 I. Landau published the discovery of a new rodent malaria parasite designated *Plasmodium chabaudi* (Landau, 1965), representing a new *Plasmodium* species. Subsequently it was shown that the original isolate of Landau was in fact a mixed infection of *P. vinckei chabaudi* and *P. vinckei petteri* (Carter & Walliker, 1974), considering *P. chabaudi* as a subspecies of *P. vinckei*. A cloned rat-adapted strain was derived from the original isolate by serial blood passages through splenectomized and non-splenectomized rats (Coombs and Gutteridge, 1975). Its

pattern of enzyme variants was typical of the subspecies *P. chabaudi.*

The parasite used in the present work was kindly provided by Dr.I. Landau, Museum d'Histoire Naturelle, Paris. It was the original rat adapted strain of Coombs and Gutteridge which was readapted to Fischer rats (Watier *et al.*, 1992), its present taxonomy is *P. chabaudi chabaudi* 96V. This strain produces a non-fatal infection in rats with a peak blood parasitaemia in the order of 50 %, and a fatal infection in BALB/c mice; mice succumb 3-14 days after infection, depending on the dose of parasites administered. The infection is highly synchronous with a 24-hour cycle. Peak schizogony occurs at about midnight and is accompanied by sequestration of many of the schizonts into the deep tissues.

2.3.3. *E. coli* strains used to propagate bacteriophage λZAPII and plasmids

BB4	supE, supF, hsdR, [F', proAB, lacIq, lacZΔM15, Tn10(tetr)] (Bullock *et al.*, 1987)
DH5αF'	F'/endA1, hsdR17($r_{K-}m_{K+}$), supE44, thi-1, recA1, gyrA96, relA1, Δ(lacZYA-argF)U169 deoR (ϕ80dlacZΔM15) (Woodcock *et al.*, 1989)
JM 109	recA1, supE44, endA1, hsdR17($r_{K-}m_{K+}$), gyrA96, relA1, thi-1, Δ (lac-proAB) F'[traD36, proAB, $lacI_q$, lacZΔM15] (Yanisch-Perron *et al.*,1985)
PLK-F'	recA, hsdR-M+, mrcA, mrcB, lacZ, supE44, galK2, galT22, metB1, hsdR2, [F', proAB, lacIq, lacZΔM15, Tn10(tetr)] (Kretz *et al.*, 1989)
SURE	e14 - (mcrA), Δ (mcrCB-hsdSMR-mrr)171, supE44, thi-1, gyrA96, endA1, relA1, lac, recB, recJ, sbcC, umuC ::Tn5(kanr), uvrC, [F', proAB, lacIq, lacZΔM15, Tn10(tetr)] (Greener *et al.*, 1990)

XL1-Blue recA1, endA1,gyrA96, thi-1, hsdR17, supE44, relA1, lac, [F', proAB, lacIq, lacZΔM15, Tn10(tetr)] (Bullock *et al.*, 1987)

3. Methods

3.1. Work with *P. chabaudi* 96V

3.1.1. Infection with *P. chabaudi* 96V

The strain was maintained by weekly serial blood passage (intraperitoneal inoculation) in a normal laboratory mouse (Balb/c, male, 20-35 g). Mice were inoculated, if not mentioned otherwise, with 10^7 parasitized erythrocytes, and if left untreated would have succumbed to the infection 8 days later.

3.1.2. Monitoring of the infection

Thin blood films from mouse tail snips were stained with Giemsa's stain. Percentage parasitaemia was calculated from the number of parasitized cells per 400 blood cells. The total number of red blood cells per ml was determined by counting in an improved Neubauer cell counting chamber.

3.1.3. Bleeding of infected animals

Mice with a parasitaemia of 30-50 % were terminally anaesthetized by ether inhalation. The thoracic cavity was opened and the aorta cut. Blood was collected with a 5 ml-syringe without needle, containing 2 ml of citric saline/ glucose-solution, enough to dilute 1 ml of blood. Citric saline/ glucose- solution is 8.5 g NaCl/ 15 g sodium citrate/ 2.25 g glucose H_2O to 1 litre. To prepare stabilates, the procedure was the same except the citric saline/ glucose-solution was replaced by one drop of heparin.

3.1.4. Preparation of stabilates

Heparinized blood was diluted with 1 vol 20 % v/v glycerol (dilution in PBS) and well mixed. The cells were equilibrated for 20 min at room temperature, then dispatched into 0.5 ml aliquots and transferred for 90 min to -70°C. Finally they were stored in liquid nitrogen. Cells kept under these conditions remain viable for many years. To thaw cells they were put directly into a 37°C waterbath. Once thawed they were kept on ice until injection into an animal.

3.1.5. Removal of lymphocytes from parasitized blood

The bottom of a suitable syringe without plunger was layered with nylon-wool. Two g CF-11 (Whatman) per mouse was swollen in citric saline/ glucose-solution and poured into the syringe. The column was washed with 1 syringe-vol of the same solution. Mice were bled as described in (3.1.3.). The blood was passed over the column and the column washed with citric saline/ glucose-solution until the red colour had disappeared. The eluted blood was centrifuged in 50 ml Falcon tubes (5 min, 1500 x g, 15°C). The fluid was discarded and the same volume PBS/ 2% glucose was added. The cells were resuspended and centrifuged under the same conditions as before. Washing was repeated one more time to remove any residual plasma proteins and haemoglobin. The cells were either used directly or kept frozen at -20°C.

3.1.6. Isolation of free parasites

After removal of the lymphocytes (3.1.5.) either a fresh or a frozen blood pellet was used. A frozen pellet was thawed in a 37°C water-bath, diluted with 5 vol ice-cold PBS in a 50 ml Falcon tube and centrifuged (5 min, 1500 x g, 15°C). The supernatant was discarded. The volume of the fresh or washed frozen pellet was estimated and 0.5 volume of 0.15% saponin in PBS (w/vol) was added. It was well mixed with a pasteur pipette and the tube was incubated for 20 min in a 37°C water-bath.The sample was removed, 10 vol ice-cold PBS were added, the solution transferred to a 30 ml Corex

tube and centrifuged (10 000 x g, 10 min, 4°C, SS-34 rotor). The supernatant was discarded, the pellet resuspended in the original volume of ice-cold PBS and centrifuged as before. The supernatant was discarded again. If saponin-lysis was complete, a two-layer pellet was visible: a lower dark brown layer of parasites and an upper greyish layer of ghosts (without any red material of unlysed erythrocytes left). The pellet was either used directly for further preparations or kept frozen. In the latter case 1 vol PBS was added to the pellet to hinder it from forming aggregates when thawed again. The pellet was then snap-frozen in a dry-ice/methanol bath and stored at -70°C, where it was kept for at least 1 year.

3.1.7. Raising of antisera

3.1.7.1. Preparation of mouse hyper-immune sera

a) Nivaquine-treatment

Mice were infected as described in (3.1.1.). when parasitaemia reached about 40 % (day 6), 1 mg of Nivaquine (SPECIA) was injected i.p. and parasites were cleared completely 5 days later.When mice were reinfected 2 weeks later no parasitaemia developed, the animals were immune against the otherwise lethal infection.Two more inoculations with the parasite were performed at 2-weeks intervals. One week after the last inoculation mice were exsanguinated as described in (3.1.3.), with the exception that no anti-coagulant was added to the syringe.

b) Preparation of sera

The blood was transferred into 1.5 ml Eppendorf tubes and left at room temperature for 1 h. The forming blood-clots were detached from the tube-walls with a sealed Pasteur-pipette and stored at 4°C overnight. The next day the sera were well separated from the blood-clots. They were transferred into fresh 1.5 ml Eppendorf tubes and

centrifuged (10 min, 10 000 rpm, 4°C). After centrifugation the sera were aliquoted and kept frozen at -20°C.

3.1.7.2. Raising of sera against recombinant proteins

Recombinant proteins were present in a solution of 50 mM TRIS pH 8.0/ 150 mM NaCl/ 2.5 mM $CaCl_2$ and were homogenized either with an equal volume of Freund's complete adjuvant (FCA) or with Freund's incomplete adjuvant (FIA). Animals were immunized either with Sj26 GST-rec 267 (3.10.4.a) fusion protein or with cleaved and purified rec 700 (3.10.4.b).

a) in Fischer rats

Female Fischer rats were immunised with either of the two recombinant proteins. Each rat received an 0.5 ml subcutaneous injection distributed at twelve different spots on their backs at day 0 (50 µg recombinant protein in FCA), week 3 (50 µg recombinant protein in FIA) and week 7 (50 µg recombinant protein in FIA) followed by bleeding at week 9 and preparation of sera (3.1.7.1.b).

b) in Parkes and Balb/c mice

Female outbred Parkes and inbred Balb/c mice were immunised with either of the two recombinant proteins. Each mouse received a 0.2 ml intraperitoneal injection at day 0 (85 mg recombinant protein in FCA), week 3 (25 mg recombinant protein in FIA) and week 7 (12 mg recombinant protein in FIA) followed by bleeding at week 9 and preparation of sera (3.1.7.1.b).

3.2. Basic methods

3.2.1. Extraction of nucleic acids with phenol, phenol/chloroform and chloroform

An equal volume of phenol was added to the nucleic acid solution in a polypropylene

tube with a plastic cup and vortexed vigorously 15 s to form an emulsion. The phases were separated by centrifugation for 2 min at top speed in a microfuge tube or for 5 min (10 000 x g, 4°C) if larger tubes were used. The upper aqueous phase was transferred to a fresh tube and extracted in the same way as before with phenol/ chloroform/ isoamyl alcohol (25:24:1 v/v) and finally with chloroform/ isoamyl alcohol (24:1 v/v) alone.

3.2.2. Precipitation of DNA

To the DNA solution 1/10 vol 3 M sodium acetate pH 5.2 and 2.5 vol cold ethanol 100 % (or 1 vol isopropanol) were added in a microfuge tube followed by brief vortexing. DNA at a concentration > 0.1 µg ml^{-1} or DNA of a size > 100 nucleotides were precipitated 30 min on ice. A lower concentration or size required precipitation for at least 1 h on ice or 30 min on dry-ice/ methanol and $MgCl_2$ at a final concentration of 10 mM was added.

The solution was centrifuged (12 000 x g, 20 min, 4°C) and the supernatant removed. To the pellet 0,5 ml cold 70 % ethanol was added and after gentle shaking centrifuged (12 000 x g, 2 min). The supernatant was discarded, the pellet dried for 20 min in a speed-vac and dissolved in the desired volume of TE pH 8.0. DNA in large aqueous volumes (0.4 to 10 ml) was centrifuged in 15- or 30-ml corex tubes (10 000 x g, 15 min, 1°C, SS-34 rotor).

The methods to precipitate RNA are given in the respective protocols.

3.3.3. Micro-dialysis of DNA

A small dish was filled with about 15 ml of distilled water and a 1.6 mm micro-disk membrane (Millipore) was laid with its shining side facing upwards onto the surface of the water. One to fifty microlitres of sample were carefully pipetted in the centre of the floating membrane and dialysis was let to proceed for 1 hour at room temperature. After this time the sample was transferred into a fresh tube.

3.2.4. Spectrophotometric quantitation of DNA, RNA and oligonucleotides

Calculation of the concentration of nucleic acid in a sample was performed by spectrophotometry at 260 nm. An absorbance of 1 corresponds to approximately:

50 μg ml^{-1} double-stranded DNA

40 μg ml^{-1} single-stranded DNA and RNA

20 μg ml^{-1} oligonucleotides

Pure preparations of DNA and RNA had A_{260}/A_{280} ratios of 1.8-2.0. If there was contamination with protein or phenol, the value was significantly less.

3.3. DNA Methods

3.3.1. Isolation of DNA

3.3.1.1. Isolation of genomic DNA

Parasites were purified as described in (3.1.6.). Frozen parasites were thawed on ice, fresh parasites were used directly. The parasite pellet was diluted in 10 volumes of proteinase K-buffer (10 mM Tris-HCl pH 7.6/ 10 mM EDTA pH 8.0/ 150 mM NaCl / 0.5 % SDS/ 100 μg ml^{-1} proteinase K) and incubated in a 50 ml Falcon tube for 3 h at 56°C. The solution was then extracted once with 1 vol phenol, twice with 1 vol phenol/ chloroform and once with 1 vol chloroform. For each extraction the phases were mixed gently on a wheel for 5 min to avoid shearing high molecular mass DNA. After each extraction phases were separated by centrifugation in Corex tubes (10 min, 10 000 x g, 4°C). The upper aqueous phase was transferred each time with a wide-bore pipette into a fresh tube. After the second phenol/ chloroform extraction the interphase was clear and free of precipitated material.

After the last extraction 0.1 vol 3 M Na-acetate pH 5.2 and 2 vol ethanol were added to the aqueous solution. The high molecular mass DNA precipitated immediately. The

solution was kept for 1 h at -20°C and then centrifuged (20 min, 4000 x g and 4°C). The pellet was washed with 2 ml 70 % ethanol and centrifuged as before. The liquid was removed completely and the pellet was air dried for 30 min. To the semi-dry pellet 0.9 ml TE-buffer was added. The tube was incubated in a 37°C water-bath overnight to dissolve the DNA.

The next day RNAse A solution was added to a final concentration of 10 µg ml^{-1} and incubated for 1 h at 37°C. The solution was transferred to an Eppendorf tube, extracted once with 1 vol phenol/ chloroform, twice with 1 vol chloroform, ethanol precipitated with 2 vol 100 % ethanol for 1 h at -20°C and centrifuged (15 min, 12 000 x g, 4°C). The pellet was washed with 1 ml 70 % ethanol, centrifuged (5 min, 12 000 x g, 4°C) and air dried for 30 min. To the semi-dry pellet 250 µl 1/10 TE was added and left in a 37°C water-bath overnight. Subsequently the DNA was stored at 4°C.

With 5 µl of the DNA-solution an UV-absorption-spectrum was performed for quantitation and purity-determination (3.2.3.).

3.3.1.2. Isolation of plasmid DNA

3.3.1.2.1. Method 1: Alkaline lysis mini-preparation

This method provides enough partially purified plasmid DNA for analytical restriction digest analysis and PCR analysis (Birnboim *et al.*,1979).

Solutions:

1. 50 mM glucose/ 25 mM Tris-HCl pH 8.0/ 10 mM EDTA pH 8.0
2. 0.2 M NaOH/ 1 % SDS
3. 60 ml 5 M potassium acetate/ 11.5 ml glacial acetic acid/ 28.5 ml H_2O

5 ml LB medium containing 50 µg ml^{-1} ampicillin was inoculated with a single bacterial colony containing a plasmid and grown overnight at 37°C with shaking. Bacteria cells in 1.5 ml were pelleted in a microfuge tube for 30 s at top speed and the pellet

resupended in 100 µl solution 1. The tube was left for 5 min at room temperature, 200 µl solution 2 was added, the tube was inverted 5 times to mix and placed on ice. After 5 min. 150 µl solution 3 was added, the inverted tube was vortexed gently 10 s and placed on ice for 5 min followed by centrifugation (5 min, 1200 x g, 4°C). The supernatant was transferred to a fresh tube and 0.9 ml ethanol (or 0.45 ml isopropanol) was added.

After 2 min at room temperature nucleic acids were pelleted (12 000 x g, 5 min, 4°C or room temperature) and the supernatant completely removed. The pellet was rinsed with 1 ml 70 % ethanol and dried 10 min in a speed-vac. It was resuspended in 20 µl TE / RNAse A (20 µg ml^{-1}) and kept for 20 min in a 37°C water-bath. The plasmid preparation was finally stored at -20°C. A typical pBluescript mini-preparation contained > 0.25 µg $µl^{-1}$ DNA.

3.3.1.2.2. Method 2: Plasmid DNA isolation with QIAGEN tips

Plasmid DNA purified with QIAGEN tips is suited for use in procedures such as transfection, sequencing and enzymatic modification. Plasmids were purified as specified by the supplier's guide.

3.3.1.3. Isolation of single-stranded pBluescript DNA

A single bacterial colony containing a pBluescript plasmid was picked from a LB/ tet (12.5 µg ml^{-1}) agar plate and grown overnight in 10 ml of LB/ amp (100 µg ml^{-1}) medium by shaking at 37°C. The following day 3 ml of Superbroth medium was inoculated with enough of the overnight culture to yield an $A_{600} = 0.1$ (approximately 1:10 dilution). When an A_{600}=0.3 (≅ 2.5 x 10^8 bacteria ml^{-1}) was reached, R408 helper phage was added at a multiplicity of infection of 20:1 and shaking was continued in a 50 ml Falcon tube for 8 hours at 37°C. The culture was heated for 15 min to 65°C, spun down in a microfuge at top speed for 2 min and the supernatant stored overnight

at 4°C. To 1.2 ml of the supernatant 300 µl 3.5 M NH_4CH_3COO, pH 7.5/ 20 % (v/v) PEG solution was added, the tube was inverted to mix, left at room temperature for 15 min and spun at top speed for 1 min. The supernatant was thoroughly drained, the centrifugation repeated and any remaining supernatant removed. The (often invisible) pellet was resuspended in 300 µl TE buffer. An equal volume of phenol/ chloroform was added. The tube was vortexed for 1 min, spun at top speed for 1 min and the supernatant transferred to a fresh tube. The extraction was repeated until there was no interphase left. Thereupon the aqueous phase was extracted with chloroform. The single stranded DNA was precipitated by adding an equal volume of 7.5 M NH_4CH_3COO, pH 7.5 and 2 vol of cold 100 % ethanol. It was spun (20 min, 11 000 x g, 4°C). The pellet was washed with 80 % ethanol, dried in a speed-vac and dissolved in 20 µl TE buffer. The amount of single stranded DNA obtained was 0.5 - 1.0 µg.

3.3.2.1. Transformation of plasmids into *E.coli* cells by electroporation

a) Preparation of electro-competent cells

A 1 litre bacterial culture was grown with shaking at 37°C to an A_{600} = 0.5 - 1.0. The flask was then chilled on ice for 30 min and the cells harvested by pelleting at 4000 x g for 15 min at 4°C. The cell pellet was resuspended in 1 litre of cold distilled water, and re-pelleted. The resuspension was repeated using 500 ml cold water and the cells re-pelleted. The cells were resuspended in 20 ml 10 % glycerol in water and re-pelleted, followed by a final resuspension in 2 ml of 10 % glycerol. The cells were split in 80 µl aliquots, snap-frozen in liquid nitrogen and stored at -70°C.

b) Electroporation

Electrocompetent cells were thawed and kept on ice. To 40 µl of cells 1-2 µl of DNA

were added, mixed and let sit on ice for 1 min. The cells were then transferred to a cold 0.1 cm electroporation cuvette, placed in the electroshock chamber and pulsed for 4.5 to 5 ms in a field strength of 12.5 kV cm^{-1}. Immediately after the electroshock 1 ml of prewarmed SOC medium was added to the cells and they were incubated at 37°C for 45 min, followed by plating out on LB agar plates with 100 µg ml^{-1} ampicillin.
The apparatus used was a Bio-Rad Gene Pulser set at 25 µF, 1.25 kV and 200 Ω.

3.3.2.2. Storage of bacterial stocks

For short term storage bacteria were kept at 4°C. For long term storage, 0.15 ml of sterile glycerol was added to 0.85 ml of bacterial culture. The culture was vortexed to ensure that the glycerol was evenly dispersed, frozen in an dry ice/ ethanol bath, and then transferred to storage at - 70°C.

3.3.3. Enzymatic reactions

3.3.3.1. Restriction enzyme digest

Restriction digests were carried in reaction buffers supplied by the manufacturers at 37°C for two to sixteen hours, unless alternative conditions were specified by the supplier's guide.

3.3.3.2. Fill-in of 5'-overhangs to create blunt-ends

5'-overhangs created by restriction digest were filled in using 10 units of Klenow fragment of *E. coli* DNA polymerase I per 10 µg of template DNA in 1 x Klenow-buffer. All four dNTPs were added to a final concentration of 0.25 mM in a 100 µl reaction and incubated 15 min at 37°C.

1 x Klenow-buffer: 10 mM Tris-HCl pH 7.2/ 50 mM NaCl/ 0.1 mM EDTA pH 8.0 / 5 mM DTT

3.3.3.3. Partial fill-in of 5'-overhangs

5'-overhangs created by restriction digest were filled in under the same conditions as to create blunt-ends, except that only the dNTPs needed to make the fill-in were added to the reaction mixture.

3.3.3.4. Dephosphorylation of DNA

30 pmol DNA termini (e.g. 10 µg digested pBluescript) were incubated 30 min at 37°C with 0.6 units of calf intestinal alkaline phosphatase (CIP) in 1 x CIP buffer.

1 x CIP buffer: 20 mM Tris-HCl pH 8.0/ 1 mM $MgCl_2$/ 1 mM $ZnCl_2$/ 0.5 mg BSA or gelatin

CIP was inactivated by heating at 65°C for 10 minutes and phenol/ chloroform extraction and ethanol precipitation of the DNA.

3.3.3.5. Ligation of DNA molecules

DNA fragments were ligated overnight at 16°C or for 2 hours at room temperature in 1 x ligation-buffer

1 x ligation-buffer: 10 mM Tris/ HCl pH 8.0/ 7.5 mM $MgCl_2$/ 1 mM dATP 10 mM DTT/ 0.2 mg ml^{-1} gelatin

1 unit of T4 DNA ligase was used for up to 1µg DNA in a final volume of 10 µl.

3.3.3.6. ^{32}P-labelling of DNA-fragments

The labelling reaction was performed with the 'Multiprime DNA labelling system' from Amersham.

In a microfuge tube 25-50 ng DNA were diluted with dH_2O to a final volume of 28 µl, boiled for 5 min and kept on ice for 5 min. The tube was spun for 10 s and 10 µl buffer/ 5 µl primer-solution/ 2 µl Klenow-fragment (1 unit μl^{-1})/ 5 µl [α-^{32}P]

dATP were added. The tube was incubated for 1 h at room temperature. 50 µl 4 M NH_4CH_3COO and 200 µl 100 % ethanol were added and the labelled DNA was precipitated for 30 min on a dry-ice/ methanol bath. The tube was centrifuged 30 min at top speed and the supernatant containing the unincorporated nucleotides was discarded. The invisible pellet was washed with 0.5 ml ethanol 70 %, dried in a speed-vac for 5 min and resuspended in 50 µl TE-buffer.

3.3.4.1. Agarose gel electrophoresis of DNA

Horizontal agarose gels (0.3-3 %) were run in 1 x TBE buffer. Ethidium bromide (0.5 µg ml^{-1}) was added to the gel. Gels were run at a constant voltage of 1-5 V cm^{-1}.

Solutions:

10 x TBE buffer : 108 g Tris base/ 55 g boric acid/ 40 ml 0.5 M EDTA pH 8.0 H_2O to 1 litre

6 x loading dye : 30 % glycerol/ 60 mM Tris-HCl pH 8.0/ 6 mM EDTA/ 0.6 % SDS/ 0.25 % Bromphenol Blue/ 0.25 % Xylene Cyanol

50 x TAE buffer : 2 M Tris-acetic acid pH 8.0/ 0.1 M EDTA pH 8.0

Prior to loading 1/5 vol loading dye was added to the samples. A standard molecular mass marker was run with the DNA samples. This was lambda DNA digested with *Eco* RI and *Hind* III from Boehringer.

Preparative gels for the isolation of DNA fragments were essentially cast and run as above using low melting point (LMP) agarose and 1 x TAE buffer.

3.3.4.1.1. Purification of DNA fragments on low melting point agarose gels

Restriction enzyme digested DNA, PCR products and partially digested genomic DNA were run on low melting point agarose (FMC) gels containing 1 x TAE buffer. The gel slice containing the DNA of interest was cut out and soaked for 1 h in a volume of 1 x GELASE buffer (Cambio), three times that of the slice's weight. The buffer was

removed, the gel melted by heating it for 20 min at 70°C and kept for 10 min at 45°C. 1 unit of GELASE(Cambio) was added for each 600 mg of gel and incubation at 45°C was continued for 1 h to digest the agarose. 1 vol 5 M ammonium acetate and 4 vol ethanol were added, the DNA pelleted (12 000 x g, 30 min room temperature), the pellet washed with 70 % ethanol, air dried and resuspended in an appropriate volume of TE pH 8.0 buffer.

3.3.5.1. Southern blotting

5µg of genomic DNA or 1 µg of plasmid DNA was digested with the appropriate restriction enzymes (3.3.3.1.) and electrophoresed in an agarose gel (3.3.4.1.). The DNA in gels containing genomic DNA was depurinated prior to denaturation by washing the gel in 0.2 M HCl for 20 min with shaking. The DNA in all gels was denaturated by washing the gel in 0.5 M NaOH/ 1.5 M NaCl for 40 min and neutralized by washing in 1 M Tris pH 7.4/ 1.5 M NaCl for 30 min with shaking. The neutralization step was repeated with a fresh solution. Subsequently the DNA was tranferred to nitrocellulose filters by capillary transfer:

A piece of Whatman 3MM paper wetted with 20 x SSC was wrapped around a plexiglass support and placed inside a large backing dish. The backing dish was filled with 20 x SSC until the level of the liquid reached almost the top of the support. The agarose gel was placed upside down on the support, followed by a layer of 1 nitrocellulose filter, 3 Whatman 3MM papers, a 8 cm stack of paper towels, 1 glass plate and a 500 g weight. The transfer of DNA was allowed to proceed overnight. The paper towels and the 3MM paper were removed and the position of the gel slots was marked on the nitrocellulose filter with a soft pencil. The nitrocellulose filter was soaked in 6 x SSC for 5 min, dried on 3MM paper for 1 hour at room temperature and the DNA was fixed to the nitrocellulose filter in a Stratalinker™ UV Crosslinker (Stratagene).

3.3.5.2. Dot blotting

To 20 ng of plasmid DNA in 1 µl TE buffer 1 µl 1 M NaOH was added and left for 5 min at room temperature. A nitrocellulose filter of an appropriate size was placed on a Whatman 3MM paper and the denatured DNA was applied in two 1 µl aliquots onto the filter to form a small spot. After drying for 30 min at room temperature the DNA was fixed to the nitrocellulose in a Stratalinker™ UV Crosslinker (Stratagene).

3.3.6.1. Hybridization of transferred DNA

This method was applied for Southern blotted DNA (3.3.5.1.), dot blotted DNA (3.3.5.2.), transferred recombinant phage plaques DNA (3.7.4.b) and transferred lysed bacterial colonies containing plasmids (3.8.4.).

a) Hybridization with ^{32}P-labelled DNA fragments

Filters were pre-hybridized in a large volume of hybridization buffer for 1 h in a 65°C water-bath with shaking.

Hybridization buffer: 6 x SSC/ 5 x Denhardt/ 0.2 % SDS/ 100 µg ml^{-1} heat denatured herring sperm DNA

Filters were hybridized in the same buffer containing the probe at >100 000 cpm ml^{-1} overnight in a 65°C water-bath with shaking. The liquid was then removed and the filters were washed (high stringency conditions)

3 x 20 min	in	1 x SSC/ 0.3 % SDS	at 60°C
3 x 20 min	in	0.2 x SSC/ 0.3 % SDS	at 60°C

b) Hybridization with ^{32}P-labelled oligonucleotides

Filters were pre-hybridized in hybridization buffer for 1 h in a 37°C water-bath with shaking. Filters were then hybridized overnight with shaking in the same buffer containing the probe at >100 000 cpm ml^{-1} in a water-bath at a temperature 12°C below

the theoretical T_m of the oligonucleotide. The liquid was removed and the filters were washed

2 x 15 min	in	5 x SSC/ 0.3 % SDS	at room temperature
2 x 20 min	in	5 x SSC/ 0.3 % SDS	at $T = T_m - 12°C$

When washing was finished (in a) and b), filters were removed from the washing solution, placed for some seconds on Whatman 3MM paper to drain and wrapped wet in Saran wrap. Filters were exposed at -80°C with a double screen (Lanex regular screen, Kodak) and Kodak X-OMAT AR film.

3.3.7. Polymerase chain reaction

PCR was typically performed in a 50 µl volume and contained in addition to the template DNA: 0.2 mM dCTP and dGTP, 0.4 mM dATP and dTTP, 1.5 mM $MgCl_2$, 15 µM of each primer, enzyme buffer to 1 x (supplied as a 10 x stock by the enzyme manufacturer) and 2.5 units of *Taq* polymerase or *Pfu* DNA polymerase. The reaction mixtures were overlaid with paraffin oil and placed in a Programmable Thermal Controller (MJ Research,INC.). The conditions employed were normally 2 min/ 94°C to denature, 1 min/ 37°C to anneal, 2 min/ 72°C to extend for a total of 40 cycles.

3.4. Synthetic oligonucleotides

3.4.1. Ethanol precipitation

All nucleotides were synthesised by the Sequencing and Synthesis Service at N.I.M.R. and were provided in 35 % ammonia solution. Oligo nucleotide was precipitated by adding 40 µl of 3 M sodium acetate (not adjusted) and 1.3 ml ethanol to 360 µl of ammonia solution. The mixture was chilled at -20°C and the oligo nucleotide was pelleted by centrifugation, washed with 80 % ethanol and dissolved in a small volume of distilled water. The concentration was determined spectrophotometrically by

measuring the absorbance at 260 nm wavelength (A_{260}).

3.4.2. Gel purification

After the precipitated oligonucleotide DNA was dissolved in distilled water 2 vol of dye mix was added, the mixture denatured for 3 min at 90°C and placed on ice for 5 min.

Dye mix:	1	ml	deionised formamide
	40	µl	0.5 M EDTA pH 8.0
	2.5	mg	Bromphenol Blue

A vertical (16 cm x 16 cm x 1.5 mm) thick denaturing polyacrylamide gel in 1 x TBE was cast.

Gel solution:	25 ml	acrylamide (38 % acrylamide, 2 % bisacrylamide)
	23 g	urea
	5 ml	10 x TBE
		H_2O to 50 ml
	120 µl	10 % APS
	120 µl	TEMED

and the sample loaded on the gel. The gel was run in 1 x TBE buffer at 400 V until the Bromphenol Blue band had migrated through 2/3 of the gel. Resolved oligonucleotide DNA products were visualized in the gel by shading on a thin layer chromatography (TLC) plate using short-wave UV light. The DNA band of interest was excised and shredded by passing through a 2 ml syringe. The DNA was eluted by addition of 3 ml of 'crush and soak' elution buffer and overnight incubation at 37°C with shaking.

Elution buffer:

0.5	M NH_4CH_3COO
10 mM	$Mg(CH_3COO)_2$
0.1 mM	EDTA pH 8.0

The polyacrylamide gel fragments were then pelleted by centrifugation and the eluate was removed and passed over a Sep-Pak® Cartridge (Millipore):
With the help of a 10 ml syringe 10 ml methanol and 10 ml H_2O were passed very slowly through the cartridge. Then the oligonucleotide DNA solution was passed through the same way. The cartridge was rinsed with 10 ml H_2O and remaining water was removed by pressing air through the cartridge with an empty syringe. The oligonucleotide DNA was eluted with 1 ml ethanol and the ethanol was evaporated in a speed-vac. The pellet was dissolved in 100 µl H_2O and the concentration was obtained by measuring the A_{260}.

3.4.3. ^{32}P-labelling of synthetic oligonucleotides

Synthetic oligonucleotides are synthesized without a phosphate group at their 5' termini and are therefore easily labelled by transfer of the γ-^{32}P from [γ-^{32}P] dATP using the enzyme bacteriophage T4 polynucleotide kinase.
In a microfuge tube were added:

1 µl	16 µM oligonucleotide DNA
2 µl	10 x forward exchange buffer (Promega)
11 µl	H_2O
5 µl	[γ-^{32}P] dATP
1 µl	polynucleotide kinase (8 U µl^{-1}) (Promega)

and incubated 1 h at 37°C.
The labelled oligonucleotides were purified using NAP-5 columns (Pharmacia): The columns were equilibrated with 10 ml of 10 mM sodium phosphate buffer pH 6.8. To the labelled oligonucleotides 480 µl of the same buffer was added to obtain a final volume of 0.5 ml. More buffer was then added to the column and 1 ml of the eluate was collected. The labelled oligonucleotides were either used directly or stored at -20°C for further use.

3.5. RNA Methods

3.5.1. RNA extraction from *P. chabaudi* 96V

This was done essentially as described by Chomczynski *et al.* (1987). To 1 vol of purified parasites (3.1.6.) or infected erythrocytes after removal of lymphocytes (3.1.5.) 5 vol of solution D was added and vortexed to mix .

Solution D:

100	g	guanidinium isothiocyanate	(4 M)
117.2	ml	H_2O	
7.04	ml	0.75 M sodium citrate pH 7.0	(25 mM)
10.6	ml	10 % sarcosyl	(0.5%)
360	µl	ß-mercaptoethanol	(23 mM)

The solution was homogenized in a glass homogenizer of an appropriate size equipped with a teflon piston and transferred to a 50 ml Falcon tube. To each 25 ml of the homogenized solution 2.5 ml 2 M sodium acetate pH 4.0, 25 ml water-saturated phenol and 5 ml chloroform/ isoamyl alcohol (24:1) was added and vortexed 5 x 20 s in a 10 min period. The tube was placed on ice for 15 min, the solution transferred to a 50 ml polyethylene tube (Nalgen) and centrifuged (12 000 x g, 20 min, 4°C, SS-34 rotor). The upper aqueous phase was transferred to a fresh 50 ml tube, 1 vol isopropanol was added and precipitated for 1 h at -20°C followed by centrifugation (12 000 rpm, 20 min, 4°C, HB-4 rotor). The pellet was dissolved in 8 ml of solution D using a small glass homogenizer, heated for 10 min at 65°C to solubilize the RNA completely and kept on ice for 15 min.

In order to remove the insoluble particles the solution was spun (12 000 rpm, 20 min, 4°C, HB-4 rotor). The supernatant was layered on a cushion of 3 ml 5.7 M CsCl, the tube filled up with solution D and centrifuged (27 000 rpm, 20 h, 20°C, SW-41 rotor).The liquid was removed and the transparent RNA pellet was washed 3 x with 0.2 ml DEPC treated water. The RNA was dissolved in 3 ml DEPC-treated water, by

adding 6 x 0.5 ml water and transferring it each time to a fresh 50 ml Falcon tube.The tube was incubated at 65°C for 10 min and placed on ice for 10 min.
With 5 µl of the RNA solution the A_{260} was measured to determine the RNA concentration in the solution. To the rest 0.1 vol sodium acetate pH 5.2 and 2 vol 100 % ethanol were added to precipitate the RNA. The RNA was stored in this form in aliquots at -70°C. To recover the RNA it was kept on dry ice/methanol, vortexed briefly and removed from the tube, spun down in a sterile microfuge tube, washed with 80 % ethanol, air dried and resuspended in the desired volume of sterile water.

3.5.2. Selection of poly(A)+ RNA

The vast majority of mRNAs of mammalian cells carry tracts of poly(A) and can therefore be separated from ribosomal RNA by affinity chromatography on oligo(dT) - cellulose (Aviv and Leder, 1972).
0.5 g of oligo(dT) - cellulose (Pharmacia) was resuspended in 0.1 M NaOH and kept at room temperature for 1 h. A column of oligo(dT) cellulose was poured in a sterile blue Gilson-tip plugged with sterile, DEPC treated glass wool and washed with 3 column volumes of sterile water. The column was washed with column-loading buffer until the pH of the effluent was less than 8.0 (about 10 column volumes)
1 x column-loading buffer:

20 mM Tris-HCl pH 7.6
0.5 M NaCl
1 mM EDTA
0.5 % SDS

The RNA was dissolved in sterile water, heated for 5 min at 65°C and quickly cooled to room temperature. One vol 2 x column-loading buffer was added, the solution applied to the column, and the eluate collected in a sterile tube. One vol column-loading buffer was added and the eluate collected. The eluate was heated as before, reapplied to the top of the column and collected. The column was washed with 10 vol 1 x column-loading

buffer and the poly(A)$^+$ RNA was eluted with 3 vol elution buffer collecting 0.5 fractions.

Elution buffer: 10 mM Tris/HCl pH 7.6
1 mM EDTA pH 8.0
0.05 % SDS

A_{260} of the different fractions was measured to determine the fractions containing the RNA. The fractions of interest were pooled. The material obtained at this stage contained equal quantities of polyadenylated and non-polyadenylated species of RNA. To purify poly(A)$^+$ further the sample was heated again, The final NaCl concentration was adjusted to 0.5 M and a second round of chromatography was carried out on the same column. To the pooled fractions of interest 0.1 vol sodium acetate and 2.5 vol cold ethanol were added. The tube was stored on ice for 1 h, spun (15 min, 10 000 x g, 4°C), the supernatant carefully discarded, the often invisible pellet washed with 70 % ethanol, air dried and redissolved in 20 µl DEPC treated water. A 2 µl aliquot was used to measure A_{260}, to the rest 3 vol ethanol was added and the poly(A)$^+$ RNA was stored in this form at -70°C until it was needed.

3.5.3. Electrophoresis of RNA through agarose gels containing formaldehyde

All solutions were made with fresh chemicals, water was treated with 0.1 % DEPC before use, glassware and electrophoresis equipment were submerged in 0.1 % DEPC overnight, combs were additionally washed in 3 % H_2O_2 for 1 h.

Solutions:

1. 10 x MOPS:	40.6	g	3-(N-morpholino)propanesulphonic acid
	800	ml	100 mM sodium acetate
			adjust pH 7.0 with 2 M NaOH
	20	ml	0.5 M EDTA
			add H_2O to 1 litre

2. RNA sample-buffer:	2,40	ml	formamide
	0.52	ml	10 x MOPS
	0.86	ml	formaldehyde
	0.66	ml	glycerol
	0.27	ml	H_2O
	0.25	ml	H_2O saturated with Bromphenol Blue

stored in 200 µl aliquots at -20°C

a) Casting the gel

2.4 g agarose, 20 ml 10 x MOPS and 125 ml H_2O were heated in a 500 ml Erlenmeyer flask until the agarose had dissolved. The solution was cooled down to 60°C, 37.8 ml formaldehyde (37 % ≡ 6.6 % final concentration) were added in a fume-hood and the mixture poured into a gel mould. The gel was left in the fume-hood until solidified. It was then placed into an electrophoresis apparatus filled with 1 x MOPS as migration buffer.

b) Running the gel

To 1 vol RNA in H_2O 5 vol RNA sample-buffer were added, heated 5 min at 65°C and put on ice for 5 min. The sample was run alongside RNA molecular weight standards (0.24 - 9.5 kb RNA ladder, BRL) at 3-4 V cm^{-1} until the Bromphenol Blue dye had migrated through 3/4 of the gel.

The gel was stained for 20 min in a solution containing 5 µg ml^{-1} ethidium bromide with shaking and destained in H_2O for 6 h with shaking.

3.5.4. <u>In vitro translation of mRNA</u>

An aliquot containing 0.1 µg poly(A)$^+$ or 3 µg total RNA was used for in vitro translation reactions. In 3 microfuge tubes were mixed:

1. Translation	35 µl	reticulocyte lysate
	1 µl	1 mM amino acids (minus methionine)
	4 µl	^{35}S-methionine (40 µCi)
	3 µl	RNA-sample
	7 µl	H_2O

2. Control without RNA

3. Control with 1 µl Brome Mosaic Virus RNA (0.5 µg $µl^{-1}$)

The tubes were incubated for 1 h at 30°C.

The alkali resistant, trichloroacetic acid (TCA) precipitable counts were assayed by following method:

23 µl H_2O_2 were mixed with 10 ml 0.33 M NaOH. To 0.75 ml of this solution 1 µl of each translation reaction was added in separate tubes and kept for 15 min at 37°C. 1 ml ice-cold 25 % TCA was added and put on ice for 15 min. The solution was filtered under vacuum onto glass fibre filters (Whatman), washed with 8 % TCA and H_2O and dried. The filters were counted in 2 ml Ready Safe (Beckman) liquid scintillation cocktail using a 1216 Rackbeta liquid scintillation counter (LKB). Aliquots containing 30 000 cpm were run on a polyacrylamide gel (3.6.1.), the gel soaked for 30 min in 250 ml Amplify (Amersham), dried under vacuum on a slab gel drier (model 1583, Bio-RAD) and exposed to Kodak X-OMAT AR films at -80°C.

3.6. **Protein Methods**

3.6.1. Polyacrylamide gel-electrophoresis (PAGE)

Separation of proteins by polyacrylamide gel-electrophoresis was essentially performed as described by Laemmli (1970). Gels of two different sizes were used: Standard gels (16 cm x 16.5 cm x 1.5 mm) were run at 35 mA and minigels (8.5 cm x 7 cm x 1.5 mm) at 20 mA. Each gel is composed of an upper stacking gel and a lower

separation gel of varying acrylamide concentration. Between them exists a pH difference of 2 units, this discontinum results in a concentration of proteins at their border.

Solutions:

1. Acrylamide-bis-solution : 30 g acrylamide/ 0.8 g methylene bis acrylamide H_2O to 1 litre
2. Separation gel buffer : 1 M Tris/HCl pH 8.8
3. Stacking gel buffer : 1 M Tris/HCl pH 6.8
4. 10 % (w/v) SDS-solution
5. 10 % (w/v) APS-solution
6. Running buffer : 25 mM (3.03 g) Tris base/ 192 mM (14.42 g) glycine/ 0.1 % (1 g) SDS/ H_2O to 1 litre
7. Sample buffer : 1.4 ml 1 M Tris pH 6.8/ 5.0 ml 10 % SDS/ 2.2 ml glycerol/ 1.4 ml 1.5 M dithiothreitol/ 10 mg Bromphenol Blue
8. Staining solution : 400 ml isopropanol/ 100 ml acetic acid/ 500 ml H_2O/ 1 g Coomassie brilliant blue R-250
9. Destaining solution : 200 ml isopropanol/ 100 ml acetic acid/ 700 ml H_2O

Separating gel [%]	5.0	7.5	10	12.5	15	17.5
Acrylamide	5.0	7.5	10	12.5	15	17.5
Tris pH 8.8	11.2	11.2	11.2	11.2	11.2	11.2
H_2O	13.7	11.2	8.7	6.2	3.7	1.2
10 % SDS	0.3	0.3	0.3	0.3	0.3	0.3
10 % APS	0.1	0.1	0.1	0.1	0.1	0.1
TEMED	0.02	0.02	0.02	0.02	0.02	0.02

Stacking gel	3.0 %	5.0 %
Acrylamide	1.00	1.67
Tris pH 6.8	1.25	1.25
H_2O	7.77	7.03
10 % SDS	0.1	0.1
10 % APS	0.05	0.05
TEMED	0.01	0.01

Samples were mixed with an appropriate volume of sample buffer, boiled for 5 min on a water-bath, spun in a microtube centrifuge for 5 min at top speed and the supernatant loaded in a slot of the stacking gel for migration. Standard molecular mass markers were run with the protein samples, these were either prestained (BRL) or not (SIGMA). When the bromphenol blue front reached the lower end of the glass plates migration was stopped. Gels were stained for 2 h in staining solution and destained overnight in destaining solution. Gels were subsequently dried using a gel-drier from NOVEX following the manufacturer's instructions. Photos were taken from dried gels using a Kodak Plus-X film.

3.6.2. Western-blotting

After separation by polyacrylamide gel electrophoresis proteins were tranferred to nitrocellulose following a modified procedure described by Towbin *et al.*(1979). In gels for western-blotting prestained molecular mass markers were used. When gel electrophoresis was finished gels were equilibrated for 10 min in transfer buffer. Protein transfer was done in a Trans-Blot Cell (Bio-Rad) at 150 mA, 4°C, overnight using transfer buffer.

Transfer buffer: 3 g Tris base/ 14.3 g glycine/ 200 ml methanol

H_2O to 1 litre (final: pH 8.8)

The nitrocellulose was removed and saturated in PBS/ 5 % (w/v) non fat milk powder

for 1 h at room temperature. The first antibody was diluted in PBS/ 5 % non fat milk powder to the desired concentration and the saturated nitrocellulose incubated with it overnight at 4°C with gentle shaking. The solution was removed, the nitrocellulose washed 3 x in PBS/ 0.05 % (v/v) Tween-20 for 20 min at room temperature and incubated with a 1:500 dilution of the second antibody labelled with alkaline phosphatase for 1 h at room temperature followed by 3 washes as before.

3.6.2.1. Detection of alkaline phosphatase

Solutions:

1. Enzyme buffer	10 ml	1 M Tris/ HCl pH 9.5
	0.6 g	NaCl
	0.5 ml	1 M $MgCl_2$
	H_2O to 100 ml	
2. NBT-solution	50 mg	NBT
	1 ml	70 % (v/v) dimethylformamide
3. BCIP-solution	50 mg	BCIP
	1 ml	dimethylformamide

To 10 ml of enzyme buffer 66 µl NBT-solution and 33 µl BCIP-solution were added and the nitrocellulose incubated with it. After 1-15 min a dark brown colour developed at places where the second antibody was bound. To avoid the development of a strong background the nitrocellulose was washed for 5 min in 10 mM EDTA pH 8.0 and dried.

3.7. Isolation of clones from a λZAPII cDNA library of *P. chabaudi* 96V coding for parts of the ROPE-protein

3.7.1. Constructing the cDNA library

Poly(A)$^+$ RNA was prepared as described in (3.5.2.). Analysis on agarose-

formaldehyde gels (3.5.3.) showed no detectable rRNA bands after 2 passages through an oligo(dT) column. Quality of the poly(A)$^+$ RNA was assessed by in vitro translation (3.5.4.), followed by polyacrylamide gel electrophoresis. The protein banding pattern obtained was comparable to those published for in vitro translated products of non degraded poly(A)$^+$ RNA from *P. chabaudi* (Da Silveira *et al.*, 1983).

The cDNA library was constructed with 6 µg of poly(A)$^+$ RNA using the ZAP-cDNA™ Synthesis Kit (Stratagene), following the manufacturer's instructions, briefly: For first strand synthesis a $(dT)_{18}$ linker-primer with an 5' adjacent *Xho* I site (underlined) followed by a "GAGA" sequence was used:

5' (GA)$_9$GAACTAGTCTCGAG(T)$_{18}$ 3'

The RNA strand was then nicked with RNAaseH and POLI "nick-translated" the RNA fragments into second strand DNA. Blunt ends were created with T4 DNA polymerase and *Eco* RI adaptors were ligated to the blunt ends:

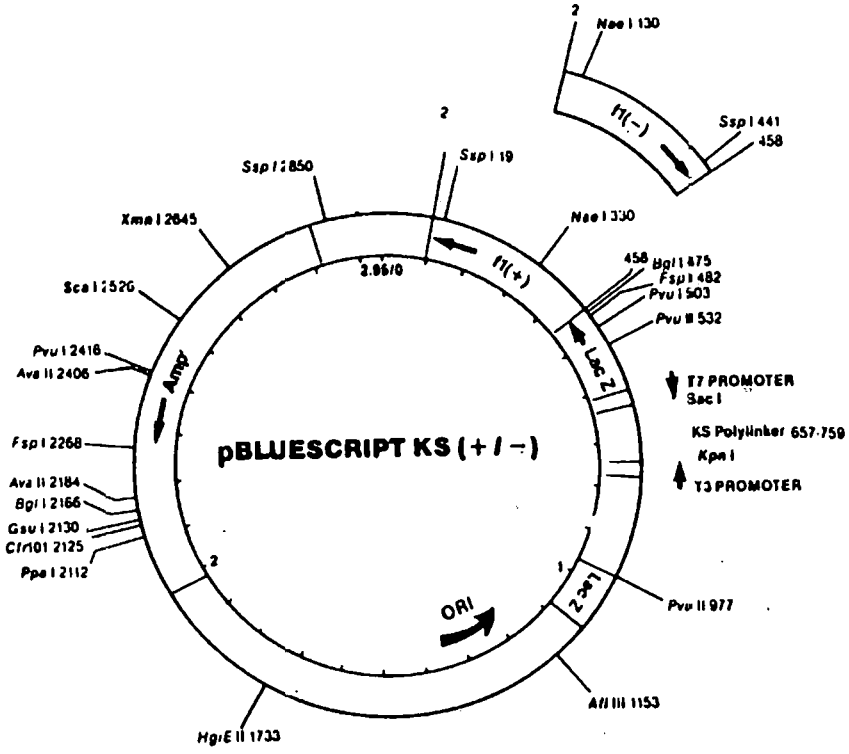
5' AATTCGGCACGAG 3'

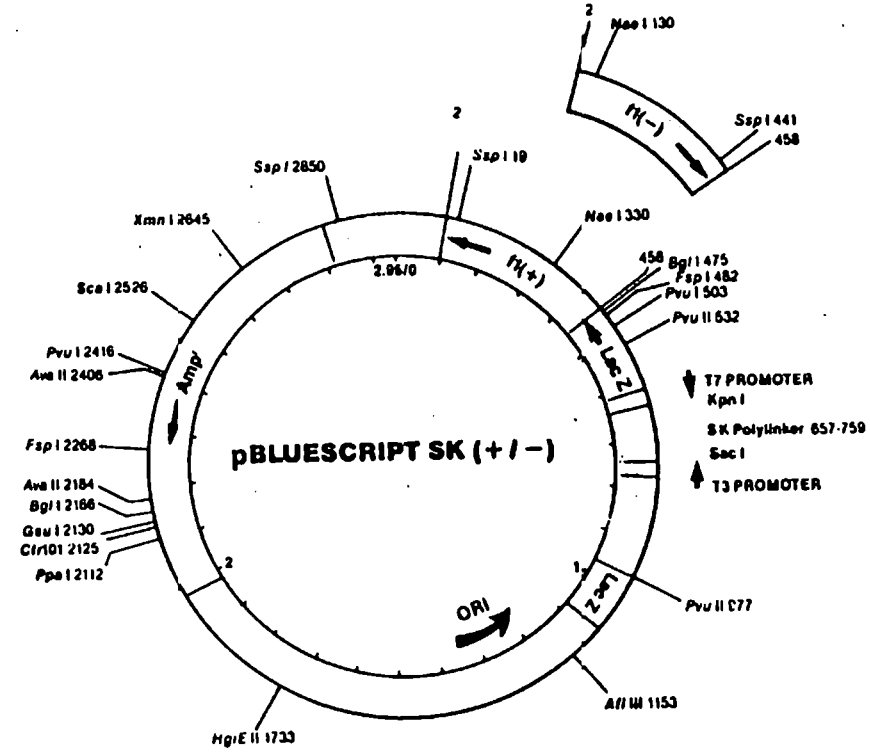
3' GCCGTGCTC 5'

The cDNAs were *Xho* I digested and size fractionated on a Sephacryl S-400 spin column. One half of the first fraction with an average size of 1.2 kb was used for ligation into the UNI-ZAP™XR vector (Lambda ZAP vector digested with *Eco* RI/ *Xho* I, CIP treated, containing pBluescript SK(-); see Figure 3.1. for pBluescript vectors). One half of the ligation products was used for in vitro packaging into Gigapack II Gold packaging extract (Stratagene). The volume was made up to 0.5 ml with SM buffer, 20 µl chloroform were added and the non-amplified library was stored in this form at 4°C

Figure 3.1 : **Structure of the pBluescript vectors.**

At the top a schematic representation the four pBluescript KS(+/-) and pBluescript SK(+/-) vectors is shown.
At the bottom the polylinker region is depicted. Some of the primer sequences present in the polylinker region were used for sequencing of recombinant plasmids from the cDNA and genomic DNA libraries.
[From: pBluescript® Exo/Mung DNA sequencing system (1990), Stratagene protocol; with the friendly permission of Dr. Sue Pearson].





Reverse Primer
5' AACAGCTATGACCATG 3'

T3 Primer
5' ATTAACCCTCACTAAAG 3'

SK Primer
5' TCTAGAACTAGTGGATC 3'

MET

T3 Promoter ←1

Sac I Bst X I Sac II Not I Eag I Xba I Spe I Bam HI Sma I Pst I Eco RI

5' GGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGG
3' CCTTTGTCGATACTGGTACTAATGCGGTTCGAGCTTTAATTGGGAGTGATTTCCCTTGTTTTCGACCTCGAGGTGGCGCCACCGCCGGCGAGATCTTGATCACCTAGGGGGCCCGACGTCCTTAA

Beta-Galactosidase

Eco RV Hind III Cla I Hinc II Acc I Sal I Xho I Apa I Dra II Kpn I

AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAA 3'
GCTATAGTTCGAATAGCTATGGCAGCTGGAGCTCCCCCCCGGGCCATGGGTTAAGCGGGATATCACTCAGCATAATGTTAAGTGACCGGCAGCAAAATGTT 5'

←1 T7 Promoter

3' GCTATGGCAGCTGGAGC 5'
KS Primer

3' GATATCACTCAGCATAA 5'
T7 Primer

3' TGACCGGCAGCAAAATG 5'
M13 -20 Primer

3.7.2. Determination of the titre of the library

First strand synthesis was made with 5-methyl-dCTP, this resulted in hemi-methylated cDNA (the *Xho* I site in the linker-primer was not methylated). Hemi-methylated cDNA was not sensitive to digestion with *Xho* I and internal *Xho* I sites were protected this way. Hemi-methylated DNA introduced into a mrcA+, mrcB+ strain would be subject to digestion by the mrcA and mrcB restriction system. Determination of the titre of the non-amplified library and amplification itself was made by infecting *E.coli* PLK-F', a mrcA-, mrcB- strain. After passing the library through *E.coli* PLK-F' the cDNA was no longer methylated and could be propagated in mrcA+, mrcB+ strains like *E.coli* BB-4, XL-1 Blue and SURE.

3.7.2.1. Preparation of competent *E.coli* bacteria for infection with bacteriophages

E.coli PLK-F' bacteria were grown on LB/ tet (12.5 µg ml^{-1}) agar plates to select for the F'-episome necessary for phage infection. A single colony was grown in 50 ml LB/ 10 mM $MgSO_4$/ amp (50 µg ml^{-1}) media overnight at 37°C with shaking. The cells were spun (1000 x g, 10 min) and resuspended in 0.5 vol 10 mM $MgSO_4$. Before use cells were diluted to $A_{600} = 0.5$ with 10 mM $MgSO_4$.

3.7.2.2. Titre before amplification

Top-agar	:	0.8	g	agarose
		100	ml	medium (NZY, LB)

A 0.1 µl aliquot of the library was incubated with 0.25 ml competent *E.coli* PLK-F' bacteria for 15 min at 37°C. 3 ml NZY top-agar, equilibrated to 48°C, was added, the mixture plated on NZY agar plates and incubated at 37°C overnight. 422 pfu were obtained corresponding to 2.11 x 10^6 pfu for the non-amplified library. The concentration of Uni-ZAP™XR vector arms in the packaging reaction was 1 µg. Vector arms ligated to insert DNA, packaged and plated should give 1 x 10^6 to 1.5 x

10^7 pfu μg^{-1}.

3.7.2.3. Titre after amplification

12 µl aliquots of the non-amplified library (approximately 50 000 pfu) were plated with 0.6 ml of competent *E.coli* PLK-F' bacteria as described in a) and incubated for 7 h at 37°C. 10 ml SM buffer were pipetted onto each plate and the plates incubated overnight with gentle shaking. The bacteriophage suspension was recovered from each plate and pooled into a 250 ml centrifuge tube (NUNC). 5 % chloroform was added and the tube incubated for 15 min at room temperature followed by centrifugation (4000 x g, 5 min). The supernatant was recovered and 0.3 % chloroform was added. The amplified library was aliquoted in sterile tubes and stored at 4°C. The titre of the library was determined as 1.6×10^{10} pfu ml^{-1}.

3.7.3. Percentage of recombinants in the amplified library

The F'-episome present in certain *E.coli* strains contains the ΔM15 mutation of the lacZ gene required for alpha-complementation of the amino-terminus of the lacZ gene present within the pBluescript part of the Uni-ZAP XR vector (Figure 3.1.). The expression of both of these partial lacZ genes is required for a functional ß-galactosidase protein. Insertion of a cDNA into the polylinker will disrupt the lacZ gene and abolish alpha-complementation (Figure 3.1.). Located on the F'-episome of some *E. coli* strains is the $lacI^q$ gene coding for the lac repressor which blocks transcription from the lac promoter P_{lac} in the absence of an inducer (e.g. IPTG). The chromogenic substrate X-gal is hydrolysed by ß-galactosidase and further oxidized by air into a blue dye.

Only non-recombinant phages can perform alpha-complementation and are able to produce blue plaques, whereas plaques from recombinant phages appear colourless. An aliquot of the amplified library containing about 1000 pfu was incubated with 0.25

ml competent *E.coli* XL1-Blue bacteria for 15 min at 37°C. 3 ml NZY top-agar, 15 µl 0.5 M IPTG (in H_2O) and 50 µl X-gal (250 mg ml^{-1} in dimethylformamide) were added, plated on a NZY agar plate and incubated overnight at 37°C. The plate was transferred for 2 h to 4°C to intensify the colour of weak blue plaques and blue/ white colonies were counted. There were 99% recombinant phages in the amplified library.

3.7.4. Screening of the library

a) Screening with antibodies

As a consequence of the way the cDNA was synthesized, the *Eco* RI site is located at the 5' end and the *Xho* I site at the 3' end with respect to the original mRNA. Ligation into the *Eco* RI/ *Xho* I digested Uni-ZAP XR vector, carrying the plasmid (phagemid) pBluescript SK(-), brings the insert under the control of the lacZ promoter P_{lac} in an orientation that allows it to be expressed as a protein, if inserted in the correct reading frame (1/3 of all clones). The resulting protein is fused at it's N-terminus to a portion of ß-galactosidase coded by the vector. Some fusion proteins are toxic to *E.coli*, therefore synthesis of the fusion-proteins was induced only after phages had already grown in the bacteria for a certain time.

Aliquots of the amplified library containing 6000 pfu were plated on numbered 80 mm NZY agar plates using competent *E.coli* BB4 as described above. Plates were incubated 3 h at 42°C to accelerate growth of the plaques. In the meantime nitrocellulose filters were numbered, soaked for 30 s in a 10 mM IPTG solution and dried for 30 min at 37°C. After incubation at 42°C the plates were removed, the plaques overlaid with the nitrocellulose filters and incubated further 4 h at 37°C. The plates were removed, filters pricked with a needle through the agar for later orientation and washed for 5 min in PBS/ 0.05 % Tween-20 with gentle shaking, saturated for 1 h in PBS/ 5 % (v/w) non fat milk powder and incubated with an 1:300 dilution of the anti-serum in PBS/ 5 %

non fat milk powder. The filters were washed 3 x 20 min in PBS/ 0.05 % Tween-20, incubated with an anti-rat IgG (whole molecule) alkaline phosphatase conjugate (SIGMA) at a 1:1000 dilution and washed under the same conditions as before. Positive clones were detected using NBT and BCIP as described for western-blotting in (3.6.2.1.). The corresponding plaques were cut out from the agar plate, transferred to 0.5 ml SM buffer/ 20 µl chloroform and kept overnight at 4°C. 200-300 pfu were plated and screened as before, a positive plaque was cut out, 50 pfu were plated and screened. A positive plaque was cored and stored in 0.5 ml SM buffer/ 20 µl chloroform. This was considered as a pure phage clone stock.

b) Screening with ^{32}P-labelled oligonucleotides

Oligonucleotides corresponding to sequences from the 5' and 3' end of inserts were synthesized to walk upstream and downstream of the coding sequence by further screening the cDNA library. 50 000 pfu were plated using *E.coli* XL1-Blue or SURE per 150 mm NZY agar plate, incubated for 10 h at 37°C and stored overnight at 4°C. Plaques were transferred for 2 min onto labelled nitrocellulose filters and the filters were pricked with a needle through the agar for later orientation. Duplicates were always made, which virtually eliminated any false positive signal. The second transfer was made for 4 min.

Spots of 5 ml 1 M NaCl/ 0.5 M NaOH solution were poured onto a plastic sheet and nitrocellulose filters were laid upon it with the transferred DNA facing upwards, to denature double stranded DNA. After 2 min the filters were removed and placed in the same way for 5 min upon 5 ml 1 M NaCl/ 0.5 M Tris-HCL pH 8.0 solution for neutralization. The filters were submerged for 30 s in 0.2 M Tris-HCl pH 7.5/ 2 x SSC, dried on 3MM Whatman paper for 30 min on the bench and fixed in a Stratalinker™ UV Crosslinker (Stratagene). Up to 20 filters were hybridized in 100 ml hybridisation solution as described in (3.3.6.1.). Positive clones were cut out and purified following the strategy described in (3.7.4.a).

3.7.5. In vivo excision of pBluescript from Uni-ZAP XR

An internal part of Uni-ZAP XR is the plasmid (phagemid) pBluescript, where pBluescript is flanked by the termination and initiation site of the f1 bacteriophage origin of replication for positive strand synthesis. By coinfection with a filamentous helper bacteriophage derived from the M13 class (e.g. R408) helper phage proteins recognize the initiator DNA sequence, nick the DNA and duplicate the DNA downstream, including any DNA inserted into the polylinker, until the termination signal is encountered. The single-stranded DNA molecule is circularized by a gene II product from the helper phage. The circularizing of the DNA molecule recreates a functional f1 intergenic (IG) region required for packaging and replication. Packaged and secreted single-stranded pBluescript can infect other *E.coli* cells where it is rapidly converted into the double-stranded replicative form and can be isolated as a plasmid pBluescript.

If a bacteria, harbouring a pBluescript plasmid, is infected with a helper phage, single-stranded copies of pBluescript are synthesized, packaged and secreted. Single-stranded DNA can be purified (3.3.1.3.) and the insert sequenced (3.9.).

The notation 'SK' signifies that the polylinker, flanked by *Sac* I and *Kpn* I sites, is oriented in a way that *Sac* I is next to the lac promoter P_{lac}, the opposite is true for the notation 'KS'. The notation (-) is respective to the orientation of the f1 intergenic region on pBluescript and thus determines which strand is synthesized in a single-stranded form. Throughout the work pBluescript SK(-) was used.

in vivo excision:

200 µl of a pure lambda ZAP clone-stock were incubated with 200 µl A_{600}=1.0 *E.coli* XL1-Blue or SURE bacteria and 1 µl R408 helper phage (1 x 10^6 pfu ml^{-1}) and incubated for 15 min at 37°C. 5 ml 2 x YT media were added followed by incubation for 3 h at 37°C with shaking. The tube was heated 20 min at 60°C to kill the bacteria

and spun (4000 x g, 5 min). 50 µl of the supernatant was mixed with 50 µl (A_{600}=1.0) *E.coli* XL1-Blue or SURE bacteria and incubated for 15 min at 37°C. The mixture was spread out onto a LB/ amp (100 µg ml^{-1}) agar plate and incubated overnight at 37°C. A single colony was streaked out on a fresh LB/ amp (100 µg ml^{-1}) agar plate and grown overnight at 37°C. Colonies from this plate were used to grow up bacteria harbouring the recombinant pBluescript SK(-) clone and for preparation of glycerol stocks.

3.8. Isolation of clones from a pBluescript KS(+) genomic DNA library of *P. chabaudi* 96V containing sequences of the ROPEgene

The genomic library was constructed in pBluescript KS(+) (Figure 3.1).

3.8.1. Constructing the genomic DNA library

a) Preparation of partially *Sau* 3A digested and partially filled-in genomic DNA of *P.chabaudi* 96V

High molecular mass genomic DNA of *chabaudi* 96V was prepared as described in (3.3.1.1.) and its analysis on a 0.6 % agarose gel showed a molecular size of greater than 25 kb. Aliquots of 10 µg genomic DNA in 100 µl 1 x Restriction Enzyme Buffer A (Boehringer) were digested for 1 hour at 37°C with 0.9, 0.6, 0.45 and 0.3 units of *Sau* 3A. The reactions were stopped by adding 0.2 vol loading dye. The samples were loaded into separate wells and run overnight at 20 V on a 0.6 % low melting point agarose gel in 1 x TAE buffer (3.3.4.1.).

Digestions under these conditions resulted in genomic DNA that was partially but not completely digested (complete digestion is indicated by an appearing banding pattern, these samples would not be used for the construction of a library) and seemed suitable for the construction of a genomic DNA library. The part of the gel containing fragments

of 2-4 kb was cut out, the agarose digested by GELASE treatment (3.3.4.1.1.), 0.2 vol loading dye was added and the sample run a second time under the same conditions. The 2-4 kb fraction was purified as before, bound to a QUIAGEN-tip 5 column followed by elution, precipitation and resuspension in 30 µl TE-buffer (3.3.1.2.2.). The *Sau* 3A ends were partially filled-in using dGTP and dATP creating a 5' CT 3' overhang (3.3.3.3.), the partially filled-in genomic DNA was precipitated and resuspended in 15 µl H_2O at a final concentration of 0.1 µg $µl^{-1}$.

b) Preparation of *Xho* I digested and partially filled-in pBluescript KS(+)

A digestion reaction of

25 µl	pBluescript KS(+)	(0.9 µg $µl^{-1}$)
20 µl	*Xho* I	(10 U $µl^{-1}$)
25 µl	10 x Buffer H	(Boehringer)
180 µl	H_2O	

was incubated at 37°C overnight and subsequently heat inactivated at 60°C for 20 min. The *Xho* I sites were partially filled-in by directly adding dTTP, dCTP and Klenow fragment of *E. coli* DNA polymerase I to the heat inactivated digestion reaction creating a 5' AG 3' overhang (3.3.3.3.). To remove possible traces of non-digested pBluescript KS(+), the sample was run on a 0.7 % low melting agarose gel in TAE buffer, the partially filled-in pBluescript KS(+) cut out of it, purified (3.3.4.1.1.) and resuspended in 50 µl H_2O at a final concentration of 0.12 µg $µl^{-1}$.

c) Cloning of partially filled-in *P. chabaudi* 96V genomic DNA into partially filled-in pBluescript KS(+) and its transformation into *E. coli* SURE

To 0.2 µg partially filled-in *P. chabaudi* 96V genomic DNA 0.12 µg partially filled-in pBluescript KS(+) was added, corresponding to a molar ratio of about 1:1.7, and

ligated for 2 hours at room temperature (3.3.3.5.). Ligation products were desalted by micro-dialysis against distilled water (3.2.3.) and subsequently 1 µl of it was transformed into electrocompetent *E. coli* SURE bacteria by electroporation (3.3.2.1.). The transformed *E. coli* SURE cells were in a final volume of 1.041 ml (3.3.2.1.).

3.8.2. Characterizing the genomic DNA library

a) Percentage and number of recombinants

The plasmid pBluescript KS(+) was the cloning vector of the genomic library and blue/ white colour selection is based on the same mechanisms as those described in (3.7.3.). After transformation into *E. coli* SURE 10 µl of the genomic library was spread out on LB agar plates containing 100 µg ml^{-1} ampicillin, 80 µg ml^{-1} X-gal and 20 mM IPTG. Colonies were grown up overnight at 37°C and kept for 2 hours at 4°C to intensify the colour of the blue colonies. The results showed that there were about 26 000 cfu ml^{-1}, and with 50 % recombinants the total amount of recombinants was 13 000 per millilitre.

b) Average insert length of recombinant plasmids

The recombinant plasmids from 10 white colonies were purified (3.3.1.2.1.), double digested with *Bam* HI-*Kpn* I (3.3.3.1.) and analyzed on a 0.8 % agarose gel. The average size of the inserts was about 3 kb.

3.8.3. Plating, growth, replication and storage of the genomic DNA library

Two different ways were used to spread out the library. In the first method described the transformed bacteria were directly spread out on plates, in the second on a nylon membrane.

Method 1: Direct plating

A total of 150 000 cfu in 6 ml SOC medium were spread out on a 24.5 x 24.5 cm LB/ amp (100 μg ml^{-1}) plate and incubated overnight at 37°C until colonies of about 0.5 mm appeared. For first replica a 22 x 22 cm nitrocellulose filter (Schleicher and Schuell) was laid on a fresh LB/ amp (100 μg ml^{-1}) plate, peeled off, inverted, replaced on the plate and peeled off again. Subsequently the moist nitrocellulose filter was placed on the plate containing the colonies. It was left for 2 min, pricked with a needle through the agar for later orientation, carefully removed and laid on a fresh LB/ amp (100 μg ml^{-1}) plate with the replica facing up. The master plate and the replica were incubated at 37°C to replenish the colonies.

To make a second replica the following assembly was made: On a 0.8 mm thick 30 x 30 cm glass plate a 25 x 25 cm Whatman 3MM paper, slightly wetted with LB medium, was placed and the first replica filter, with colonies facing up, laid on top of it. A second 22 x 22 cm nitrocellulose filter, pre-wetted on a agar plate as before, was carefully laid on the first replica filter, covered with another slightly moist Whatman 3MM paper and a second 30 x 30 cm glass plate. The assembly was put down on the ground and a person stood on top of it to bring the two nitrocellulose filters into close contact.

The first glass plate and the Whatman 3MM paper were removed, the second filter was pricked with a needle by piercing through the holes of the first replica. Both replica were carefully separated, laid, colonies facing up, on two LB/ amp (100 μg ml^{-1}) plates and incubated at 37°C until colonies were clearly visible. The master plate was kept at 4°C, where it could be used to isolate colonies for up to 1-2 months.

Method 2: Plating on Biodyne A transfer membrane

This method is a modified protocol first published by Hanahan and Meselson (1980).

a) Plating, growth and replication

On a fresh 24.5 x 24.5 cm LB/ amp (100 μg ml^{-1}) plate a 22.5 x 22.5 cm Biodyne A transfer membrane (PALL) was laid. Approximately 50 000 cfu in 6 ml SOC medium were placed in about 0.1 ml spots in a regular pattern on top of the nylon membrane. Spreading was performed with the knee of a spreader instead of its rod part. The plate was left open in a sterile hood until no more surface moisture was visible and incubated overnight at 37°C. When colonies had reached a diameter of about 0.5 mm a replica was made of the master membrane as described for first/ second replica in method 1, using another Biodyne A transfer membrane. A second nylon membrane replica was made from the first replica in the same way.

b) Freezing of the master membrane

The replenished master membrane was peeled off the LB/ amp (100 μg ml^{-1}) plate and laid on top of two layers of Whatman 3MM paper, which were soaked in LB medium containing 25 % glycerol. After 15 minutes the moist membrane was transferred on a 25 x 25 cm and 3 mm thick perspex plate, and a 24 x 24 cm graph paper grid photocopied on an acetate sheet was placed on top of it. The pattern of holes in the master filter was marked on the grid with a blue marker, a second perspex plate of the same dimensions as the first one was laid on the grid and finally the assembly was fixed with the help of four small clamps. Another acetate grid was placed precisely on top of the first one and the blue marks copied to it. The second grid was always kept together with replica 1 and 2 of the library, once their colonies were denatured and fixed (3.8.4.). The master plate assembly was stored at -80°C where it could be used to isolate colonies for many years.

3.8.4. Screening of the genomic DNA library with ^{32}P-labelled DNA fragments

Three trays, each containing a layer of two 30 x 40 cm sheets of Whatman 3MM

paper, were prepared. Enough of solution 1, 2 or 3 was poured into each tray to soak the paper but to avoid excess fluid.

Solutions :

1. 0.5 M NaOH
2. 1 M Tris pH 7.5
3. 0.5 M Tris pH 7.5 / 1.5 M NaCl

The replenished replica filters were removed from the LB/ amp (100 µg ml^{-1}) plates, denatured for 5 min on Whatman 3MM paper soaked with solution 1, colonies facing upwards, and blotted on dry Whatman 3MM paper for 5 min. The process was repeated once with neutralizing solution 2 for 2 min, followed by 2 min blotting and once with neutralizing solution 3 for 15 min, followed by 30 min blotting, finally the filters were fixed in a StratalinkerTM UV Crosslinker (Stratagene).

Up to 4 filters were hybridized in a sealed plastic bag with 50 ml hybridization solution as described in (3.3.6.1.). Subsequently the filters were placed between two sheets of Saran Wrap, mounted on Whatman 3MM paper, which was labelled with GolgosTM II Autorad Markers (Stratagene), put into X ray cassettes with double screen (HI-SPEED-X, Kodak) and exposed with a Kodak X-OMAT AR film at -80°C.

3.8.5. Isolation of positive clones

a) From the directly plated library

After autoradiography the keyholes of the filters were marked on the autoradiogram after alignment of the GolgosTM II Autorad Markers images to the originals. The autoradiogram was then aligned to the master plate, the desired colonies were picked with an inoculating loop, transferred into a 1.5 ml Eppendorf tube containing 1 ml of LB/ amp (100 µg ml^{-1}) medium and vortexed. Aliquots were spread out on LB/ amp (100 µg ml^{-1}) plates, incubated overnight at 37°C and a plate with about 100 - 500

colonies was re-screened as described for Method 1. Once an isolated positive colony was obtained, it was spread out directly on a LB/ amp (100 μg ml^{-1}) plate and used for further purification of its recombinant plasmid (3.3.1.2.2.) or to prepare glycerol stocks of the bacteria harbouring the recombinant plasmid (3.3.2.2.).

b) From the frozen master membrane

After autoradiography the keyholes of the filters were marked on the aotoradiogram as described in a). The autoradiogram was aligned to the second acetate-sheet grid (3.8.3.Method 2, b), positive signals were marked on the sheet and 5 x 5 mm squares, with the positive signal in their centre, were cut out of the sheet. The master membrane was removed from -80°C, put on dry-ice and the first perspex plate removed. The two grids were aligned to each other and same squares were cut out of the first acetate-sheet grid and the underlying nylon membrane with the spread library. The nylon membrane with the adhering colonies was transferred into a 1.5 ml Eppendorf tube containing 1 ml of LB/ amp (100 μg ml^{-1}) medium, vortexed, kept for 15 min at room temperature and vortexed again. Aliquots were plated out on LB/ amp (100 μg ml^{-1}) plates and re-screened as in a) until a pure colony could be isolated. Due to the low density plating on the nylon membrane a pure colony was normally obtained during first re-screening. The master membrane was covered with the first perspex plate and stored at -80°C until further needed.

3.9. DNA sequencing

DNA sequencing was either performed with the 'Multiwell microtitre plate DNA sequencing system T7 DNA polymerase' kit (Amersham) or the 'Sequenase (version 2.0) DNA sequencing kit' (USB) using the dideoxy chain termination method (Sanger *et al.*, 1977). Double-stranded plasmid DNA was first denatured according to the

manufacturer's instructions, single-stranded plasmid DNA was used directly. Primers corresponded either to sequences from the vectors (pBluescript SK (-), pBluescript KS(+) and pGEX-2T) or to sequences from the inserts obtained from previous sequencing. Conditions for annealing, labelling ([α-^{35}S] dATP), and extension were those given by the manufacturers. 6% polyacrylamide gels (30 cm x 40 cm x 0.4 mm) containing 7.5 M urea were cast using premade solutions (Sequagel™ Sequencing System, National Diagnostics). Gels were mounted on a S2-Sequencing Gel Electrophoresis Apparatus (BRL) and run in 1 x TBE buffer at 60 W. Gels were fixed for 20 min in 10 % acetic acid/ 20 % methanol, dried under vacuum on a slab gel drier (model 1583, Bio-RAD) and exposed to Kodak BioMAX-MR films without screen at room temperature.

3.10. Cloning and expression of sequences in p-GEX-T vectors

The pGEX expression vectors were constructed by Smith and Johnson (1988) to direct the synthesis of foreign peptides in *E.coli* as fusion proteins with the COOH-terminus of Sj 26, a 26 kDa glutathione-S-transferase encoded by the parasite helminth *Schistosoma japonicum* (Smith *et al.*, 1986). Soluble fusion proteins can be purified from crude bacterial lysates by affinity chromatography on immobilized glutathione. The presence of a protease cleavage site between the Sj 26 and the protein coded by the insert enables the expressed protein to be separated from the Sj 26 part after digestion and re-chromatography. Figure 3.2 shows a diagram of three pGEX vectors. The polycloning sites of the three vectors were designed in a way that *Eco* RI fits all three reading frames (1, 2 and 3). Each of the three vectors contain a thrombin cleavage site T. The original vector is pGEX-2T. In pGEX-1T the original pGEX-2T polylinker was replaced by the polylinker shown in Figure 3.2 (kind gift from Dr. R. Lafiatis, NIH). The plasmid pGEX-3T was made to have the third reading frame with respect to *Eco* RI.

Figure 3.2: **Structure of the pGEX-T vectors.**

Three different vectors from the pGEX-T series are shown. The three vectors contain a thrombin cleavage site between the glutathione-S-transferase (GST) encoded by the plasmids and the peptide encoded by an inserted. They vary in the linker part between the *Eco* RI and *Bam* HI sites, resulting in three different reading frames with respect to the *Eco* RI cloning site.

2T : Thrombin

LeuValProArgGlySerProGlyIleHisArgAsp

5' CTGGTTCCGCGTGGATCCCCGGGAATTCATCGTGACTGACTGACG 3'
3' GACCAAGGCGCACCTAGGGGCCCTTAAGTAGCACTGACTGACTGC 5'

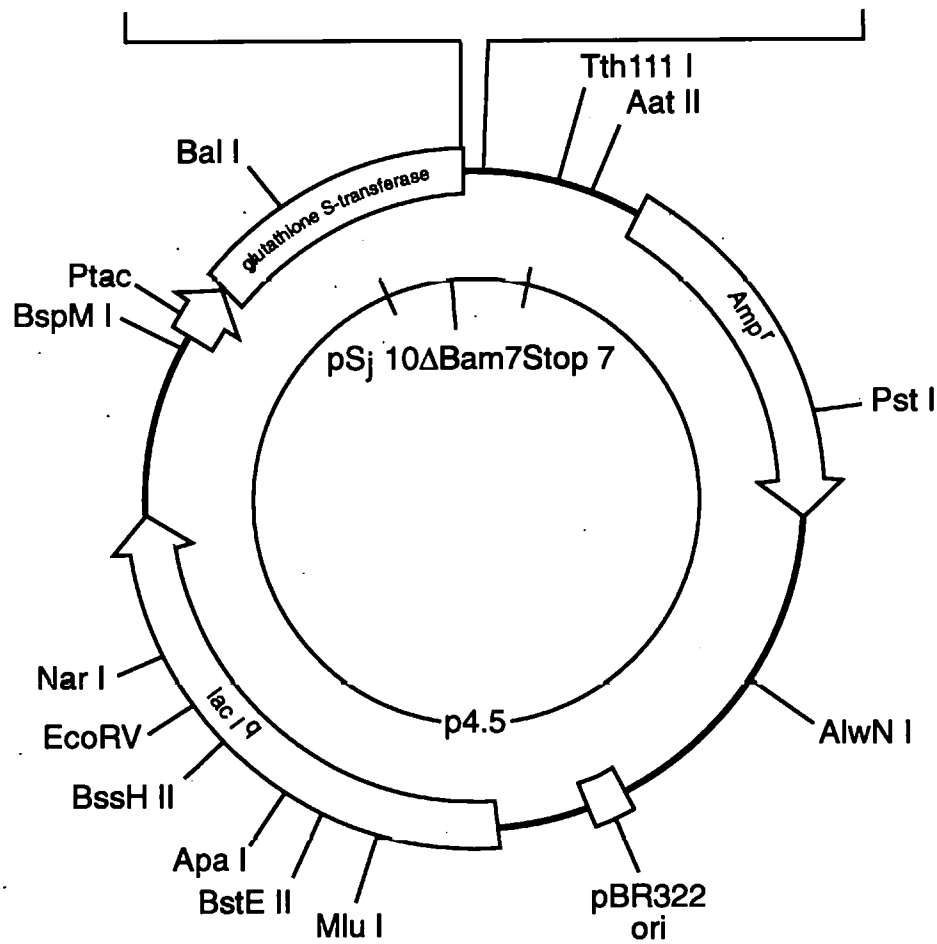
Bam HI *Sma*I *Eco*RI Stop codons

1T:

5' GATCCGGCAACAAAGGTACCATGG
3' GCCGTTGTTTCCATGGTACCTTAA
Bam HI *Eco* RI

3T:

5' GATCCGGCAACAAGGTACCATGG
3' GCCGTTGTTCCATGGTACCTTAA
Bam HI *Eco*RI



3.10.1. Construction of pGEX-3T

Two oligonucleotides, E1 and E2, were synthesized and gel-purified (3.4.2.):

```
E1 :    5'   GATCCGGCAACAAGGTACCATGG          3'
E2 :    3'        GCCGTTGTTCCATGGTACCTTAA     5'
              Bam HI                Eco RI
```

E1 and E2 were mixed in equal molar amounts by preparing 0.5 ml 20 mM E1, 20 mM E2, 10 mM $MgCl_2$, 25 mM Tris-HCl pH 8.0 in a microfuge tube. The tube was placed in a polystyrene-box filled with 2 litres of water at 95°C and left to cool down to room temperature over a period of 10 h. E1 and E2 should hybridize to form a 27 bp double-stranded DNA molecule with a sticky *Bam* HI and *Eco* RI end at either side and a theoretical $T_m = 60°C$.

Eleven µg pGEX-2T were digested with *Bam* HI and *Eco* RI (3.3.1) to cut out the polylinker, purified on a low melting point agarose gel (3.3.4.1.1.) and dephosphorylated (3.3.3.4.). 1.2 ng E1 : E2 linker were ligated (3.3.3.5.) into 50 ng of the prepared vector (5:1 molar ratio). Ligation products were introduced into *E.coli* JM 109 bacteria by electroporation (3.3.2.1.) and plated on LB/ amp (100 µg ml^{-1}) agar plates. The colonies were transferred to nitrocellulose (3.8.3.) and screened (3.8.4.) with ^{32}P-labelled E1 (3.4.3.) for pGEX-3T. Positive clones were grown up in LB/ amp (100 µg ml^{-1}) media, plasmids isolated and sequenced (3.9.). The plasmid with the new polycloning site was named pGEX-3T.

3.10.2. Subcloning into pGEX-T vectors

a) pGEX-700:

Fifteen µg of the recombinant plasmid pBluescript SK (-) 70 was digested with *Eco*

RI and *Xho* I (3.3.3.1.). The insert was purified (3.3.4.1.1.), ligated (3.3.3.5.) to an adaptor with phosphorylated blunt end :

```
5'   TCGAGGAATTC   3'
3'       CCTTAAGp  5'
   Xho I    Eco RI
```

and digested with *Eco* RI. Fifteen ng of insert carrying now a sticky *Eco* RI end at each side was ligated (3.3.3.5.) into 100 ng *Eco* RI-digested and CIP-treated (3.3.3.4.) pGEX-1T. The ligation products were introduced into *E.coli* JM 109 by electroporation (3.3.2.1.), grown on LB/ amp (100 µg ml^{-1}) agar plates, transferred to nitrocellulose (3.8.3.) and screened with an ^{32}P-labelled pBluescript SK (-) 70 insert (3.3.6.1.a). Plasmid preparations from positive clones were finally sequenced to confirm the desired insertion (3.9.).

b) pGEX-267:

Ten µg of the recombinant plasmid pBluescript SK (-) 6b was digested with *Eco* RI (3.3.3.1.) and the internal 267 bp *Eco* RI fragment gel purified (3.3.4.1.1.). 2.5 ng of this *Eco* RI fragment was ligated (3.3.3.5.) into 25 ng *Eco* RI-digested, CIP-treated (3.3.3.4.) pGEX-3T. Ligation products were transformed into *E.coli* DH5α bacteria by electroporation (3.3.2.1.). The procedure for isolating a bacterial clone harbouring the desired pGEX-267 plasmid was as described for pGEX-700. Screening was performed with the ^{32}P-labelled 267 bp *Eco* RI fragment.

3.10.3. Small-scale expression of pGEX clones

After screening for pGEX-recombinants carrying a subcloned insert expression on a small scale was performed to see if the corresponding peptide was present, before final

sequencing was performed.

A single bacterial colony harbouring a pGEX recombinant plasmid was grown in 5 ml LB/ amp (50 µg ml^{-1}) media in a 50 ml Falcon tube overnight at 37°C with shaking. One ml of the overnight culture was diluted with 9 ml LB/ amp (50 µg ml^{-1}) media and further incubated for 1.5 h at 37°C. The culture was made 0.1 mM in IPTG, then grown for 4 h under the same conditions and spun (2500 x g, 5 min). The pellet was dissolved in 0.5 ml PBS/ 1 % Triton X-100/ 2 mM PMSF/ TPCK (100 µg ml^{-1})/ 1 mM EDTA, transferred to a microfuge tube, sonicated 3 x as in a) and spun at top speed for 5 min. The supernatant was transferred to a fresh microfuge tube, 3 mg preswollen glutathione-agarose beads were added and mixed on a wheel for 1 h at room temperature. The beads were spun (1000 x g, 5 min) and washed 3 x in 0.5 ml PBS. One volume 2 x Laemmli sample buffer was added to the beads, the mixture boiled on a waterbath for 5 min and spun at top speed for 3 min. 15 µl of the supernatant were analyzed by PAGE on a mini-gel (3.6.1.).

3.10.4 Large-scale expression of pGEX clones

A single colony was grown in 100 ml LB/ amp (50 µg ml^{-1}) medium overnight at 37°C with gentle shaking. Ninehundred millilitres prewarmed (37°C) LB/ amp (50 µg ml^{-1}) media were added and incubated for 1.5 h at 37°C. Fusion protein synthesis was induced by adding 0.2 ml 500 mM IPTG (0.1 mM IPTG final concentration) and incubation was continued for 4 h at 37°C. The culture was pelleted (5000 x g, 15 min, 4°C), the pellet resuspended in 10 ml PBS/ 1% Triton X-100/ 2 mM PMSF/ TPCK (100 µg ml^{-1})/ 1 mM EDTA, sonicated 5 x 20 s on ice at 1 min intervals (Vibra Cell, microtip in position 5, Sonics & Materials INC.), centrifuged (10 000 x g, 5 min, 4°C) and the supernatant transferred to a fresh 50 ml Falcon tube. 75 mg glutathione-agarose beads (sulphur linkage, SIGMA) were preswollen for 1 h in PBS, spun (500 x g, 5 min) and washed 3 x in 10 ml PBS. The beads were added to the supernatant followed

by incubation overnight at 4°C with gentle mixing on a wheel. The mixture was spun (500 x g, 5 min) and the beads washed 3 x with PBS. At this point 2 options existed:

a) Purification as a Sj 26 fusion protein

To the bead pellet 1 ml 50 mM Tris-HCl pH 8.0/ 5 mM glutathione were added, the suspension transferred to a 2.0 ml Eppendorf tube and mixed on a wheel for 1 h at room temperature to elute the fusion protein from the beads by competition with excess glutathione. The mixture was spun (500 x g, 5 min), the supernatant transferred to a fresh tube, spun at top speed for 5 min and the supernatant transferred to dialysis tubing. After dialysis against 5 l 1 M Tris-HCl pH 8.0 overnight at 4°C, the solution of the fusion protein was analyzed by PAGE (3.6.1.) for purity, aliquoted and stored at -20°C.

b) Purification as a protein without Sj 26

The bead pellet was washed once with 30 ml thrombin buffer: 50 mM Tris-HCl pH 8.0/ 150 mM NaCl/ 2.5 mM $CaCl_2$, centrifuged (500 x g, 5 min) and the supernatant was discarded. One bead volume thrombin buffer and 60 units thrombin were added and incubated for 4 h at 25°C to cleave the unfused protein from the beads, leaving Sj 26 coupled to the beads. The mixture was spun (500 x g, 5 min), the supernatant transferred to a fresh 2.0 ml Eppendorf tube, spun at top speed for 5 min and the protein solution transferred to a new tube. Purity of the protein was assessed by PAGE (3.6.1.). If necessary traces of Sj 26 were removed by further incubation with 75 mg preswollen glutathione-agarose beads as before. Pure protein solution was finally aliquoted and stored at -20°C.

3.11. **Indirect Immunofluorescence microscopy**

When parasitaemia reached about 30 %, the animals were bled and lymphocytes

removed (3.1.5.). The erythrocytes were washed 3 x with PBS and were finally resuspended at a 1% haematocrit in PBS. Ten microlitres of the suspension were placed in each well of multiwell slides (Flow Labs) and dried in the draught of a tissue culture hood for 4 h. The erythrocytes were fixed in cold acetone (-20°C) for 30 s and dried on the bench. At this stage slides were either used directly or stored in the presence of silica gel at -20°C. Slides kept at -20°C were left at room temperature for 5 min before use. Ten microlitres of the first antibody diluted in PBS/ 1 % BSA (v/w) were added to the wells followed by incubation for 30 min at 37°C in a closed plastic box containing a wet paper towel. The slide was placed for 5 min in a slide jar filled with PBS to remove excess antibody. The washing-step was repeated 2 x with fresh PBS. After removal from the slide jar liquid between the wells was removed with a cut-off needle attached to a suction pump. Ten microlitres of an appropriate dilution of the second antibody (FITC conjugate) in PBS/ 1 % BSA was quickly added to the wells followed by incubation under the same conditions as for the first antiserum. The slide was removed, washed as before and dipped for 1 min in a 0.01 % (w/v) solution of Evans Blue in PBS before finally briefly rinsed in 200 ml PBS. Liquid between the wells was removed as before. Eight microlitres Citifluor (Citifluor Ltd.) were pipetted on each well to reduce fading of the fluorescence when illuminated with UV light and the slide was covered with a large coverslip.

Immunofluorescence microscopy was performed on a Transmitted-Light Photomicroscope III (Zeiss), fitted with an III RS Epi-Fluorescence Illuminator and a HBO 100 W/2 High Pressure Mercury Vapour Lamp. Photos were taken with Kodak Ektachrome EL 400 films.

Results

4.1. Screening of the cDNA library of *P. chabaudi* 96V

4.1.1. First screening of the cDNA library with an antiserum

A rat serum raised against material retained and eluted from a glutathione agarose column, mainly putative glutathione-S-transferase (GST), isolated from the rodent malaria parasite *Plasmodium chabaudi* 96V, was used to screen a cDNA library of the same parasite constructed in the phagemid λZAPII (3.7.). The host bacterium used for the screening procedure and further purification was *E.coli* BB4.

Screening of 5 x 10^5 pfu resulted in the detection and purification of 20 clones that gave positive signals with the serum. Recombinant pBluescript SK(-) plasmids were obtained by in vivo excision (3.7.5.). The plasmids were purified, double digested with *Eco* RI - *Xho* I and run on 1% agarose gels, where the cloned inserts showed a size of 400-900 bp. Subsequently the DNA fragments were tranferred to nitrocellulose and hybridized with different ^{32}P-labelled individual inserts of the 20 clones in order to group them. The results of the hybridization allowed the inserts to be divided into 2 groups: 19 inserts of a size of 400-700 bp that hybridized with each other and 1 insert of about 900 bp that did not hybridize with the other 19 inserts. Three clones from the group of 19 clones, c70, c64 and c100 were chosen for further sequencing (clones from the cDNA library were written with the prefix 'c'). Sequencing of the cloned inserts was performed as described in (3.9.).

The results of the sequencing are illustrated in Figures 4.1 and 4.2. Inserts c64 (position 6066-6709) and c100 (position 6244-6709) ended at the same 3'-end position, insert c70 (position 6006-6705) ended 4 bp upstream (positions are relative to the sequence shown in Figures 1 and 2). Interestingly all inserts except one of the remaining clones ended at the same 3'-end position as c70, one insert ended at the 3'-end position of c60 and c100, as demonstrated by sequencing (data not shown). All clones detected by the antiserum contained inserts located between position 6006-6709.

The aligned sequence of 704 bp was very A/T-rich, a typical feature of plasmodial sequences. Translation showed an ORF over the entire length of the sequence without a stop codon, coding for a protein fragment of 28 kDa. This indicated that the size of the native protein was most probably bigger than that of a GST, which range in size from 26-28 kDa (Hughes, 1994). A sequence homology search showed the highest similarity with coiled coil proteins like myosin, cytokeratins, spectrin and other mainly cytoskeletal proteins, suggesting that the 28 kDa protein fragment was part of a yet uncharacterized cytoskeletal protein of the parasite (see chapter 6).
Cytoskeletal proteins play a fundamental role for example in the maintenance of cell shape, motility and the fixing of organelles and they are possibly altered in erythrocytes after invasion by *Plasmodia* (Cranton *et al.*, 1983; Sherman, 1985).
Very little is known about the cytoskeleton of the parasite or about any interaction between the host's and the parasite's cytoskeleton. On the other hand there was no evidence that screening with different anti GST sera would be successful in detecting parasite GST because it might be weakly expressed or toxic in bacteria. For these reasons the decision was made to continue the sequencing and characterization of this novel protein.

4.1.2. Second screening of the cDNA library with ^{32}P -labelled oligonucleotide 70-5

Further screening of the library with radiolabeled probes was performed. Detection of cloned inserts in this way does not require that they are in frame with the ß-galactosidase sequence coded by pBluescript. It was further possible that clones coding for less antigenic parts of the protein could not be detected easily with the antiserum. From the first screening it was evident that many short clones started at position 6705 and 6709. In order not to obtain more of these short inserts, oligonucleotide 70-5 (position 6006-6026) was synthesized, representing the first 21 bp of the 5'-end of insert c70 and present only in this insert (Figures 1 and 2). It was hoped that screening

the library with this probe would result in the detection of inserts containing further upstream sequences, as well as of inserts that started downstream of the common 3'-end position of the inserts obtained from the first screening.

All screening steps were performed in *E. coli* XL-1 Blue. It was thought to be safer to propagate recombinant pBluescript plasmids containing possibly long inserts in a rec A deficient bacterial host (*E. coli* BB4 is not deficient in the rec A gene). By screening 7 x 10^5 pfu, 23 positive phage clones were detected, purified and the recombinant pBluescript SK(-) plasmids were obtained by in vivo excision. The purified plasmids containing the cloned inserts were *Eco* RI - *Xho* I double digested and separated on 0.9% agarose gels.

A common feature of the cloned inserts (18 out of 23) was that they consisted of up to 4 DNA fragments per clone. An identical pattern was obtained upon *Sma* I - *Apa* I digestion, cutting upstream and downstream of the cloned sequence, excluding the possibility that the multiple bands resulted from internal restriction sites. Digestion performed after repeated purification by another round of screening resulted again in a multiple band pattern indicating that the different bands had their origin in one plasmid. Southern blotting with ^{32}P-labelled insert c70 sometimes gave reactivity with more than 1 fragment and the stochiometry between pBluescript and its cloned inserts was often distorted. For example the DNA of two inserts of 1.3 and 1.5 kb were present in a 3 times higher amount than the DNA of the 2.9 kb plasmid pBluescript after digestion. It became evident that pBluescript carrying inserts with plasmodial sequences were prone to recombinational events when propagated in *E.coli* XL-1 Blue. In order to try to solve this problem in vivo excision was repeated in *E. coli* SURE, an *E. coli* strain that was specifically designed to suppress unwanted recombinational events in cloned sequences.

Plasmids were purified, digested and analyzed as before. Direct comparison of double digested plasmids either isolated from *E.coli* SURE or *E. coli* XL-1 Blue and run on 0.8% agarose gels showed banding patterns that were different from each other. While

most (18 out of 23) of the double digested plasmids isolated from *E. coli* XL-1 Blue showed multiple bands, only a small number (2 out of 23) did so when isolated from *E. coli* SURE.

Clone c25a contained a 1.8 kb insert, the largest of the cloned inserts, and hybridized with ^{32}P-labelled insert of clone c70 in southern blots. A restriction map of insert c25a indicated that it should contain a further 0.2 kb upstream and 1.1 kb downstream in respect to the known sequence. Insert c25a was entirely sequenced (Figure 4.1). The insert consisted of 1872 bp (position 5847-7719), including the known sequence (position 6005-6709) and represented a 5'-end and 3'-end extension of it (Figures 4.1 and 4.2). Interestingly the sequence around the common 3'-end region of the clones obtained from first screening was 5' AAAGAAAAAAATAAAA 3' (position 6706-6721). All of the clones ended either at position 6705 or 6709 and obviously the $(T)_{18}$ part of the linker primer, used to synthesize the first cDNA strand when the cDNA library was constructed, primed at this sequence accepting 1 or 2 mismatches. Insert c25a contained an ORF over its entire length, coding for a 74 kDa protein fragment, still missing the 5'- and 3'-end of the coding region.

4.1.2.1. PCR analysis of clones from the second screening with primers derived from the 5'-end and 3'-end of c25a

In order to determine if any of the other 22 clones purified during second screening contained sequences upstream and downstream of the 1872 bp ORF their inserts were analyzed by PCR. The 5'-end regions were amplified with the Bluescript reverse primer/ 70-5 pair and the 3'-regions with the Bluescript T7/ F4 primer pair. The reverse and T7 primer hybridize to pBluescript SK(-) 117 bp upstream and 48 bp downstream of the *Eco* RI - *Xho* I cloning-site respectively (see plasmid map in Figure 3.1), and primer 70-5 (position 6006-6026) and primer F4 (position 7428-7448) are derived from the ORF (Figures 4.1 and 4.2).

The PCR reactions were performed as described in 3.3.7., the conditions employed here were 2 min/ 94°C to denature, 1 min/ 37°C to anneal and 2 min/ 72°C to extend for a total of 40 cycles. Analysis of the PCR products on a 1.4% agarose gel indicated that 13 clones possessed an apparent 5'-end extension of the ORF between 50-500 bp and 1 clone , c6b, of about 1.4 kb. At the 3'-end of the ORF only 1 clone, c25b, produced a fragment during PCR amplification. This fragment was slightly larger than the one obtained under the same conditions with clone c25a, indicating that the cloned insert c25b might contain a small extension at its 3'-end compared to insert c25a.

The insert of clone c6b was completely sequenced by applying the same strategy as before, insert c25b was sequenced at its 3'-end (Figure 4.1). The entire insert of c6b (position 4393-6072) contained an ORF and showed a 225 bp overlap with the 5'-end of insert c25a (Figures 4.1. and 4.2). From position 5385 upstream to the 5'-end of insert c6b existed a repeat sequence consisting of a 33 bp element corresponding to 11 amino acids. Sequencing of this repeat was hampered by the fact that primers often matched to various parts of the repeat and especially in the beginning the sequencing procedure was based on trial and error.

The 3'-end of insert c25b extended the sequence of c25a by 14 bp, ending at position 7732 (Figures 4.1 and 4.2). The first 11 bp of the 14 bp extension were: 5' AAATAAAAAAA 3' (position 7719-7729), the region to which the linker primer of cloned insert c25a had obviously hybridized.

The aligned sequences so far obtained showed an ORF (position 4393-7732) coding for a protein fragment of 131 kDa. No stop codon could be identified and none of the other purified clones seemed to contain further upstream or downstream sequences.

4.1.3. Third screening of the cDNA library with ^{32}P-labelled oligonucleotides 6-15R and 6-16

Oligonucleotides 6-15R (position 4417-4439) and 6-16 (position 7660-7685)

(Figures 4.1 and 4.2) were synthesized, ^{32}P-labelled and used to screen 1.1 x 10^6 pfu from the cDNA library propagated in *E. coli* SURE. Each probe hybridized in southern blots only to clones containing either the 5'-end or the 3'-end of the known ORF. A total of 142 plaques hybridizing with the mixture of both probes were detected in duplicates. Thirty of them were further purified by rescreening and, after in vivo excision, their recombinant plasmids were isolated. The plasmids were run on 0.8% agarose gels, transferred to nitrocellulose and hybridized with either ^{32}P-labelled oligonucleotide 6-15R or 6-16.

Fourteen plasmids hybridized with oligonucleotide 6-15R, corresponding to the 5'-end, 16 plasmids with oligonucleotide 6-16 corresponding to the 3'-end of the ORF. To analyze 5'- and 3'-end extensions of the ORF sequence, the 5'-end of the first group of plasmids was amplified in a PCR reaction using the reverse primer/ 6-15R pair and the 3'-end of the second group with a combination of the T7/ 6-16 oligonucleotide pair. The conditions employed for the PCR reaction were 1 min/ 94°C to denature, 1 min/ 37°C to anneal and 2 min/ 72°C to extend for a total of 40 cycles. Thirteen out of 14 inserts from the 5'-end group showed extensions of about 20 - 480 bp, but only 1 out of 16 inserts from the 3'-end group showed an extension of about 100 bp.

Three cloned inserts, c111a, c112a and c121a, that gave the longest extensions at the 5'-end were partially sequenced at their 5'-end (Figure 4.1). From the 3'-end group insert c120d, which gave the 100 bp extension, was entirely sequenced. Sequence analysis showed that inserts c111a (position 3914- approx. 5650), c121a (position 3924- approx. 5450) and c112a (position 3978- approx. 5650) were overlapping with each other and the known sequence (Figures 4.1 and 4.2) (approximate positions were determined by sequencing the insert at one end and adding the insert length, determined on an agarose gel, to it). Insert c120d (position 7540-7845) contained a 113 bp 3'-end extension of the known sequence. It could be seen now that the sequence directly downstream of the 3'-end of clone c25b was 5'AAAAAAAAAAAAA 3' (position 7733-7745), providing a sequence that easily hybridized with the $(T)_{18}$ part of the

linker primer. The 33 bp repeat sequence that started at position 5385 and extended up to the 5'-end of insert c6b extended further upstream to position 4293 where it started with CCA coding for proline.

Translation of the 3931 bp so far available showed an ORF extending over the entire length coding for a protein fragment of 154 kDa.

4.1.4. Fourth screening of the cDNA library using ^{32}P-labelled oligonucleotides 6-17 and 6-18R

Despite the slow progress made in obtaining the full length of sequence of the ORF the cDNA library was screened another time. The filters used for the third screening were stripped and reprobed with ^{32}P-labelled oligonucleotides 6-17 (position 7760-7781) and 6-18R (position 3923-3947) representing sequences from the 5'-end and 3'-end of the known ORF (Figures 4.1 and 4.2).

Twenty new positive phages were detected by rescreening 1.1 x 10^6 pfu in duplicate. They were purified and, after in vivo excision in *E. coli* SURE, the plasmids isolated. From 20 inserts 17 contained 5'-end sequences and 3 contained 3'-end sequences of the known ORF as determined by dot blot hybridization of the plasmids with either ^{32}P-labelled oligonucleotide 6-17 or 6-18R. The 5'- and 3'-end extensions were measured by PCR amplification using reverse primer/ 6-18R for 5'-end extensions and 6-17/ T7 for 3'-end extensions. The PCR reactions were performed at 1 min/ 94°C to denature, 1 min/ 37°C to anneal and 2 min/ 72°C to extend for a total of 40 cycles. The PCR products were run on 2% agarose gels and analyzed.

PCR products from the 5'-end group inserts indicated possible 5'-end extensions of about 0-600 bp. Insert c120d containing the 3'-end of the known ORF and the three inserts from the 3'-end group of the fourth screening produced an identical PCR fragment, indicating that the four inserts ended at the same 3'-end position.

Four cloned inserts from the 5'-end group c119α, c122ß, c115γ and c105d were partially sequenced, the cloned inserts from the 3'-end group c102h, c119f and c116α were completely sequenced (Figure 4.1).

Alignment of clones c119α (position 3309-4124), c122ß (position 3437- approx. 4120), c115γ (position 3588-approx. 7020) and c105d (position 3648- approx. 6060) showed a further extension of 605 bp at the 5'-end of the ORF. The 3'-ends of the cloned inserts c102h (position 7436-7845), c119f (position 7617-7845) and c116α (position 7741- 7845) did not show any further extension at the 3'-end of the ORF (Figures 4.1 and 4.2).

The assembled sequencing data extended the ORF to 4537 bp, without any stop codon, coding for a protein fragment of 178 kDa.

4.1.5. Discussion

The first screening of the cDNA library was performed with a rat antiserum raised against material retained and eluted from a glutathione agarose column, mainly putative GST, isolated from *Plasmodium chabaudi* 96V and *E. coli* BB4 was used as host bacterium. Sequence analysis of the cloned inserts obtained during first screening revealed that the isolated clones were not coding for a GST.

At this stage certain aspects were carefully considered. Nineteen out of 20 of the isolated clones expressed parts of the same protein that showed homologies mainly with cytoskeletal proteins with extended coiled-coil regions, like α-spectrin, cytokeratins and myosin (see 4.3). Structure prediction indicated that the sequence so far available coded for a protein fragment that could form an extended coiled-coil structure itself (see 4.3).

Cytoskeletal proteins of *Plasmodia* are poorly characterized. So far the genes coding for α-tubulin I and II and β-tubulin have been sequenced and studied in some detail (Wesseling *et al.*, 1989a; Delves *et al.*, 1989, 1990; Holloway *et al.*, 1990; Rawlings *et al.*, 1992), as well as the genes coding for actin (Wesseling *et al.* 1989b). For some plasmodial proteins there is evidence that they interact with the cytoskeletal proteins of the infected erythrocyte, for example RESA with spectrin (Foley *et al.*, 1991, Ruangjirachuporn *et al.*, 1991) or MESA with band 4.1 (Lustigman *et al.*, 1990). Depending on the class of cytoskeletal protein to which a protein belongs its function may vary considerably. The particular coiled-coil structure, mainly predicted for the 235 aa of the 705 bp ORF, suggested that the native protein was less involved in motility, like actin and tubulin, but more in providing some mechanical support, for example like intermediate filaments (Geiger *et al.*, 1987, Steinert *et al.*, 1988) or spectrin (Bennett *et al.*, 1985, Marchesi, 1985). Considering that a characterization of this protein might give some insight into the organization of the parasite's cytoskeleton and that a successful detection of a GST in the cDNA library might be difficult, if

possible at all, the decision was taken after the first screening of the cDNA library to continue with the characterization of the novel putative cytoskeletal protein of *Plasmodium chabaudi* 96V.

The 900 bp insert of the clone which did not hybridize with the rest of the clones was also sequenced (data not shown). The insert is coding for a part of a protein which shows also similarities to cytoskeletal proteins. Its characterization was not pursued during the preparation of the thesis.

The inserts obtained during the first screening of the cDNA library with a serum could all be located inside of a 705 bp region and their size of about 400-700 bp was rather small. This might reflect the fact that they represent a highly immunogenic part of the molecule, that this part is better expressed than others or that other parts of the protein can not be maintained stably in *E. coli* BB4.The *E. coli* BB4 strain has a full complement of all the genes involved in DNA repair, like the recA gene, which are responsible for deletions and rearrangements that can occur in long cloned sequences. In order to increase the chances to obtain recombinant pBluescript plasmids with longer inserts, different strains of *E. coli* were used to propagate the cDNA library. The second screening was performed in *E. coli* XL-1 Blue, a recA deficient strain. *In vivo* excision and propagation in *E. coli* XL-1 Blue resulted in highly unstable recombinant plasmids. When the same procedures were repeated in *E. coli* SURE the stability of the recombinant plasmids was increased considerably. The *E. coli* SURE strain (Stop Unwanted Rearrangement Events) features a series of mutations that eliminate a number of independent DNA repair pathways and makes it deficient in homologous recombination. There was no evidence for deletions or rearrangements in any of the cloned inserts so far sequenced when grown in *E. coli* SURE. This was also true for the repeat sequence present in the ORF over a length of about 1092 bp. There are numerous reports on instability of cloned plasmodial repeat sequences when propagated

in *E. coli* (Stahl *et al.*, 1987; Weber, 1988; Coppel, 1992). The choice of an *E. coli* strain that is deficient in many of the genes involved in DNA repair appears to improve the situation.

After the fourth screening of the cDNA library it seemed that the strategy chosen to sequence the entire ORF would probably still be quite time-consuming, if feasible at all. Despite the efforts made to screen a large number of plaques and to characterize most of them, progress to extend the 5'-end of the ORF was slow and no further clones extending the 3'-end of the ORF were obtained during the last screening.

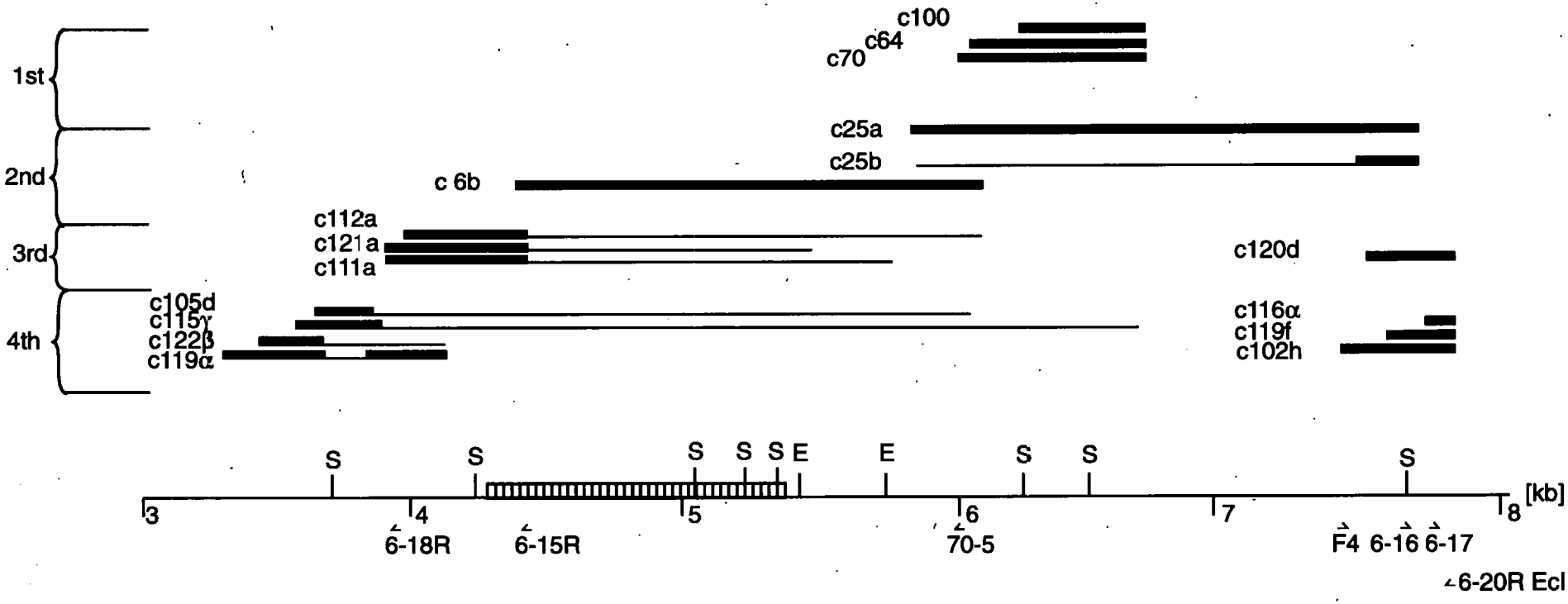
Internal runs of As to which the linker primer bound during first strand synthesis of the cDNA appeared to be a characteristic feature of the library and it seemed likely that directly downstream of the so far known ORF (position 7845) a run of As was located which was highly accessible to the $(dT)_{18}$-part of the linker primer.

To construct and screen a genomic DNA library of *Plasmodium chabaudi* 96V appeared to be a possible solution to find the missing part of the coding sequence; additionally it would give information about the existence of introns.

Figure 4.1 :

Sequence obtained from cDNA clones from the novel gene encoding the 229 kDa protein from *P. chabaudi* 96V.

1st, 2nd, 3rd, 4th : selected clones from the first, second, third and fourth screening of the cDNA library; ■ : sequenced part of a cDNA clone; - : non-sequenced part of cDNA clone; ⇀ : downstream primer; ↽ : upstream primer; ||||||||||||||||||| : repeat region; 'c' indicates that the cloned sequence is cDNA. Restriction enzyme sites: S - *Sau* 3A, E - *Eco* RI.


1st
2nd
3rd
4th
c100
c64
c70
c25a
c25b
c 6b
c112a
c121a
c111a
c120d
c105d
c115γ
c122β
c119α
c116α
c119f
c102h
S
S
S S S E
E
S
S
S
3
4
5
6
7
8
[kb]
6-18R
6-15R
70-5
F4
6-16
6-17
6-20R
Ecl

4.2. Screening of the *P. chabaudi* 96V genomic DNA library

A genomic library of P. chabaudi 96V was constructed in the plasmid pBluescript KS(+). The library was designed in a way which should allow walking along the chromosome in both directions (see discussion in 4.2.5.). Thus if an isolated clone does not contain the end of a certain sequence reprobing of the library with a new probe should result in the isolation of clones which contain extensions of the first clone.

4.2.1. First screening of the genomic DNA library with ^{32}P labelled PCR fragments

To screen the genomic library two probes, one for the 5'-end and one for the 3'-end of the known ORF were made.

5'- end probe : Plasmid c119α, isolated from the cDNA library during its 4th screening (Figure 4.1.), contained a portion of the 5'-end region (position 3309-4124) cloned into pBluescript SK(-). Using the Bluescript reverse primer/ 6-18R EcI (position 3923-3947) pair a part of the insert, c119α' (position 3309-3947) (Figures 4.1 and 4.3a), was amplified in a PCR reaction (1 min/ 94°C to denature, 1 min/ 37°C to anneal, 2 min/ 72°C to extend for a total of 40 cycles), phenol/chloroform extracted, precipitated, *Eco* RI digested and gel purified.

3'-end probe : Plasmid c102h, isolated from the cDNA during its 4th screening, contained a portion from the 3'-end region (position 7436-7845) of the known ORF (Figure 4.1.). The Bluescript reverse primer/ 6-20R EcI (position 7814-7831) pair was used for amplification of a part of the insert, c102h' (position 7436-7831) (Figures 4.1 and 4.3a), in a PCR reaction. The PCR reaction and purification steps were performed as described for the 5'-end probe.

Oligonucleotide 6-20R EcI is complementary to the coding strand and contains a T at position 3933 instead of a C. Oligonucleotide 6-18R EcI is complementary to the

coding strand and contains a C at position 7821 instead of an A. Both changes established an *Eco* RI site without changing the specificity of the PCR reaction. The amplified fragments did not contain the poly A tail of the linker primer present in the cloned inserts. This was done to minimize the risk of high background, that might rise from hybridization of the poly A tail with runs of Ts present in plasmodial sequences from non coding regions.

Fragments c119α' and c102h' were ^{32}P-labelled and used together to screen the genomic DNA library which was directly spread out on LB/amp (100 μg ml^{-1}) plates. Eleven clones gave a positive signal and were further purified, their plasmids isolated and hybridized in dot blots either with ^{32}P-labelled c119α' or c102h'. Plasmids g2, 3, 6, 8 and 10 could be assigned to the 5'-end and plasmids g1, 4, 5, 7, 9 and 11 to the 3'-end of the known ORF ('g' indicates that the respective recombinant plasmid contains an insert of genomic DNA). The size of the inserts was determined by *Apa* I-*Xba* I digestion, which did not cut the inserts, followed by analysis on 0.7% agarose gels. The recombinant plasmids and non-recombinant plasmid pBluescript KS(+) were digested with *Taq* I followed by analysis on 1.0 % agarose gels. When the library was constructed the partially filled-in *Sau* 3A and *Xho* I sites created upon ligation a new *Taq* I site. Digestion with *Taq* I would cut the cloned inserts at the cloning site and at any internal *Taq* I site. Comparison of the bands of each individual plasmid with the bands of non-recombinant pBluescript KS(+) and fragments predicted from the *Taq* I sites located on the known ORF (Figure 4.3b) resulted in restriction maps that allowed the localization of each insert in respect to the known ORF. The results showed that the cloned inserts had a length of 2.4-3.5 kb. According to their restriction map plasmids g2 and g8 showed an extension of 0.9 kb of the 5'-end, g5 a 1.1 kb and g11 a 2.1 kb extension of the 3'-end and g7 was entirely located inside of the known ORF. The other plasmids were either identical to these 5 plasmids or had shorter extensions.

Plasmids g2, 8, 5 and 11 were partially and plasmid g7 completely sequenced using either primers corresponding to sequences from pBluescript KS(+) or from the inserts (Figure 4.3a). Sequencing showed that the inserts of plasmid g2 and g8 (position 2362-5247) were identical and had a 947 bp extension of the 5'-end of the known ORF (Figure 2 and 3a). The insert of g7 (position 5074-7675) covered a region that was already sequenced from cDNA clones, plasmid g5 (position 6514-9031) had a 1186 bp extension and plasmid g11 (position 6514-9985) had a 2140 bp extension of the 3'-end of the known ORF .

Sequence translation showed that the ORF ended downstream with an UAA stop codon at position 7974, no other long ORF could be detected further downstream (Figure 4.2). Upstream the ORF extended up to position 2362, the 5'-end of insert g2 and g8, without any stop codon, indicating that the 5'-end of the ORF was not yet found. In order to obtain the entire sequence of the ORF, the genomic library was re-screened.

4.2.2. Second screening of the genomic DNA library with a ^{32}P-labelled PCR fragment

For the second screening of the genomic DNA library a part of insert g8 was amplified using the primer pair 6-29 (position 2474-2491) and 6-38R (position 3190-3208) in a PCR reaction (2 min/ 94°C to denature, 1 min/ 43°C to anneal, 2 min/ 72°C to extend, for a total of 40 cycles), resulting in fragment g8' (position 2474-3208) (Figure 4.3a). The purification steps were the same as those described for c119α' and c102h' from the first screening of the genomic library.

Fragment g8' was ^{32}P-labelled and used to screen the genomic DNA library which was spread out on a Biodyne A transfer membrane. Fourteen positive signals were detected in duplicates, the colonies cut out of the master membrane as described in (3.9.5.b) and purified by re-screening, followed by isolation of their plasmids. To determine the size of the cloned inserts plasmids were *Pvu* II digested, which did not cut the inserts, and

subsequently run on 0.8% agarose gels. Analysis showed inserts of 2.5-3.8 kb. To see the localization of the cloned inserts in respect to g8, all 14 new plasmids and g8 were digested with *Taq* I and analyzed on 1% agarose gels. The results indicated that 4 inserts ended at the same 5'-end position as inserts g2 and g8, 10 inserts contained 1-2.3 kb extensions of the 5'-end of inserts g2 and g8. Two cloned inserts, g27 and g34, which seemed to be identical and to have a 2.3 kb extension, were partially sequenced using the same strategy as for the first screening (Figure 4.3a).

The results showed that inserts g27 and g34 were identical (position 1-3718) and extended the known sequence at the 5'-end for a further 2361 bp (Figure 4.2). Translation showed that the ORF extended upstream to position 2145 at a UAA stop codon. Further upstream of position 2146 no other long ORF could be detected in any of the 3 reading frames.

The deduced amino acid sequence from the entire ORF resulted in a 229.5 kDa protein.

4.2.3 Southern-blot analysis

Malarial proteins can belong to a multigene family, like for example the Py 235 kDa protein of *P. yoelii* (Borre *et al.*, 1995) or the Duffy binding proteins of *P. knowlesi* (Adams *et al.*, 1992). Some proteins are encoded by genes that show great restriction fragment-length polymorphism in different cloned lines of *P. falciparum*, e.g. MESA (Coppel *et al.*, 1986) whereas others, like Rhop-H3 (Coppel *et al.*, 1987) are encoded by genes which are more conserved. Testing restriction enzyme digested genomic DNA of different strains and species of *Plasmodium* with ^{32}P-labelled probes from our ORF region in Southern-blots was expected to give information about the presence of this gene in these strains and species, its restriction fragment-length polymorphism and its possible copy number.

4.2.3.1. Hybridization of genomic DNA from *P. chabaudi*, *P. yoelii* and *P. falciparum* with ^{32}P-labelled probes from the 229 kDa protein of *P. chabaudi* 96V

Genomic DNA from *P.chabaudi* 96V, *P. chabaudi* CB, *P. yoelii YM* and *P. falciparum* T9/96 was digested with *Spe* I and probed with ^{32}P-labelled g7'. The fragment g7' (position 5289-7674) was obtained from the genomic clone g7 after *Taq* I digest (Figure 4.3 a) and b); note that the *Sau* 3A site at position 7674 of g7 was partially filled-in and ligated to the partially filled-in *Xho* I site of pBluescript KS (+) creating a new *Taq* I site). The digests were run on a 0.75% agarose gel; hybridization was performed, after transfer to nitrocellulose, as described in (3.3.6.1.a) at 42°C, followed by three washes of 30 min in 3 x SSC/ 0.3% SDS at 47°C (low stringency).

The g7' probe hybridized with two fragments of about 4 and 18.5 kb from both *P. chabaudi* strains one fragment of about 22 kb from *P. yoelii* and one fragment of about 29 kb from *P.falciparum* (Figure 4.6). No hybridization signals were detected with *Hind* III digested mouse genomic DNA (data not shown).
Hybridization with the ^{32}P-labelled insert from clone g2 (position 2362-5247) (Figure 4.3 a and b), which is located inside of the 2 *Spe* I sites (position 2231-6286) showed only the 4 kb band in the case of *P.chabaudi* and the 22 kb and 29 kb band for *P. yoelii* and *falciparum* respectively (data not shown).
When washing was performed at higher stringency (69°C) no hybridization signals were detectable for *P. yoelii* and *P. falciparum*.

4.2.3.2. Restriction enzyme digests of genomic DNA from *P.chabaudi* 96V and hybridization with ^{32}P-labelled insert from clone c70

Genomic DNA from *P.chabaudi* was digested with different restriction enzymes, run

on a 0.7% agarose gel and hybridized after transfer to nitrocellulose with ^{32}P-labelled insert of clone c70 (position 6006-6705) (Figure 4.1.), followed by two washes in 5 x SSC/ 0.2% SDS at 74°C (high stringency).

The results showed for all digests one major signal. In some cases minor bands appeared which are probably non-specific signals (Figure 4.7 a). Some digests allowed the construction of a simple restriction map of the genomic environment of the gene (Figure 4.7 b).

4.2.4. Discussion

The screening of the genomic DNA library resulted in the sequencing of the entire coding sequence of a novel plasmodial protein. The rational of constructing the genomic library was based on certain assumptions and observations:

1) The sequence coding for the ORF derived from the cDNA library comprised 4537 bp (position 3309-7845). Analyzing the sequence for restriction enzyme sites, using DNA-Star, it was found that 8 *Sau* 3A sites were quite evenly distributed over the sequence (Figure 4.1) with maximally 1.1 kb between two sites and it seemed possible that this pattern might continue further up- and down-stream. A partial digest with *Sau* 3A should yield a population of genomic DNA fragments that could be located anywhere between two *Sau* 3A sites if all sites were equally accessible to the enzyme. Cloning these fragments into a vector should result in a library that could be used to walk along any given sequence located on the parasite's chromosome in both directions.

2) A plasmid was preferred as cloning vector, since purified clones could be directly sequenced, making further subcloning steps unnecessary (if phages are used) or avoiding potential unstable single stranded recombinant plasmids (if phagemids are used).

3) First genomic libraries were constructed by ligating *Sau* 3A partially digested genomic DNA, which was not size fractionated, into *Bam* HI digested pBluescript KS(+). Analysis of recombinant plasmids isolated from these libraries showed that about half of the inserts had a very small size of 100-350 bp and only about 20% of the cloned inserts were longer than 2000 bp. Including a 2-4 kb size fractionation step raised the average insert length to about 3000 bp.

4) A partial fill-in of partially *Sau* 3A digested genomic DNA and *Xho* I digested pBluescript KS(+) made their ends compatible to each other and was carried out for two reasons:

I) The first libraries mentioned in 3) showed only 40% recombinant plasmids, despite dephosphorylation of pBluescript KS(+). Partial fill-in made the vector incompatible with itself and libraries made this way contained about 50-70% recombinants.

II) Partial fill-in of *Sau* 3A partially digested genomic DNA made inserts incompatible with each other and should result in single cloned inserts. Any internal *Sau* 3A site of an insert should have the same sequence environment as its corresponding chromosomal counterpart and should not be composed of two non-contiguous *Sau* 3A fragments coming from different parts of the genome. Analysis of self ligated filled-in and non filled-in partially *Sau* 3A digested genomic DNA on agarose gels could not detect any shift in size in filled-in DNA but showed a shift in size in non filled-in DNA, with the majority of DNA fragments having a size above 20 kb, supporting the expectation of a high probability of single cloned inserts (data not shown).

5) To prepare genomic libraries suitable for long term storage at -80°C they were transformed into *E. coli* SURE and spread out onto various membranes with different results. The percentage of transformed *E. coli* SURE colonies growing on them, compared to direct plating on LB/amp (100µg ml^{-1}) plates, was about: 5% for Hybond-N (Amersham), 15% for Hybond-N^+ (Amersham), 20% for nitrocellulose (Schleicher and Schuell) and 60% for Biodyne A Transfer Membrane (PALL) (data not shown). Yields were variable dependent on whether or not the membranes were washed with cold or hot distilled water, but the highest yield with the best reproducibility was obtained with the nylon membrane from PALL.

6) The genome size of *Plasmodium* is approximately 3 x 10^7 bp. If a library contains inserts with an average length of 3 kb, the number of clones required to find a particular gene can be determined by applying the formula

$$N = \frac{\ln (1 - P)}{\ln (1 - f)}$$

where P is the probability of obtaining a particular sequence and f is the ratio of the length of the insert to the entire genome (Clark and Carbon, 1976), resulting in:

Probability P	Number N of colonies
50%	6 930
90%	23 000
99%	46 000

Spreading out 25 000 - 50 000 recombinant bacteria onto a 22.2 x 22.2 cm membrane gives reasonable values for P (90-99%) at a density of colonies that allows purification of a positive colony at the first re-screening step. Combining these features a genomic DNA library of *Plasmodium* can be made that contains any particular sequence with a high probability, requires a short time to proceed from screening to sequencing, can be used for walking up- and down-stream of a sequence and allows long term storage for repeated use.

By sequencing cDNA and genomic DNA clones, which covered the ORF and parts of the flanking regions of a novel plasmodial protein, 9985 bp of genomic DNA were obtained. Most of the ORF was covered by cDNA clones and all of the 9985 bp was covered by genomic DNA clones (Figure 4.4). Both strands of the ORF were sequenced and sequencing data from each position of the ORF were compiled from at least 2 independent clones.

It seems noteworthy that using the strain *E. coli* SURE not a single cloning artefact was detected in any of the sequenced genomic clones. This includes about 12 kb of genomic sequence data which could be compared with other clones. This is true despite the fact that three of the sequenced genomic clones contained parts of the repeat sequence and that the length of the sequenced inserts was up to 3.7 kb. Long cloned plasmodial sequences are notorious for their instability in *E. coli*, especially when they

contain repeat sequences (Stahl *et al.*, 1987; Weber, 1988; Coppel, 1992). There was already a marked difference in the stability of recombinant plasmids when *E. coli* SURE was used during the screening of the cDNA library (4.1.2.). It seems to be possible that some of the reported instabilities disappear when recombinant plasmids are propagated in *E. coli* SURE (Greener, 1990).

The 9985 bp sequenced so far contain an intronless ORF at position 2146-7974, which ends with a TAA stop-codon and includes a repeat region from position 4294-5385 (Figure 4.2).

Whether or not the mRNA transcribed from the present gene is polyadenylated can not be answered because no cDNA extending the 3'-end of the ORF was found. However there are multiple potential 5' AATAAA 3' polyadenylation signals downstream of the ORF, the first at position 8163-8168 and the second at position 8550-8555 (Figure 4.2). It is possible that the polyA tail of a polyadenylated mRNA of this gene hybridized, in the conditions under which the cDNA library was constructed, to the long runs of Ts starting directly downstream of the end of the ORF. In this case the polyA tail of the mRNA would not be accessible to the linker primer any more and, as a result, only cDNA clones derived from internal priming events could be isolated. There are reports about non-polyadenylated mRNAs (Vaidya *et al.*, 1984), but the way they were isolated, based on binding to oligo(dT) cellulose, might just reflect the possible artefact mentioned here.

The plasmodial genome (especially the genome of *P. falciparum*) is known to be very A+T-rich (Weber, 1988), and it was shown that the distribution of A and T followed certain patterns (Weber 19987). The entire sequence of 9985 bp is A+T-rich (75.9%), no difference was found between the coding region (75.5%) and its flanking regions (76.5%), in contrast to most other plasmodial genes which show an increase in the A+T-content outside of the coding region (Weber 1987). There was a strong increase in

the T-content of the non coding regions and correspondingly the A/T ratio changed from 2.06 in the ORF to 0.93 in the flanking regions. The typical codon preference reported for plasmodial genes matched well with the codon usage of the present gene. Curiously this situation is quite different if one looks at the 1092 bp repeat region. This part of the ORF is less A+T-rich (61.4%) and shows a significantly higher G+C content than the rest of the ORF, although the A/T ratio of 2.5 is typical for a plasmodial coding region.

Another feature of plasmodial sequences is an increasing A+T-content from the first to the second and third position of the codons. These values are 38%, 80.8% and 65.7% respectively for the repeat region (66%, 84.5% and 83.2% for the rest of the ORF). The decrease in the A+T content at the third codon position is often reflected in a very different codon preference when compared to the rest of the ORF. For example only 2% of the glycine residues in plasmodial sequences are coded by the triplet GGG, but 62% of the glycines of the repeat (13 out of 21) are coded by GGG.

It seems interesting that the parasite uses an increased G+C content and a shifted codon preference in the repeat region. The advantage of it should be sought somehow at the level of the DNA, because this shift is not necessary for the conservation of the encoded repeat region.

Another unusual feature of the repeat region is the occurrence of a large ORF on the complementary strand, which spans the entire 1092 bp repeat region.

Outside of the coding region appears to be an underlying pattern in the base composition, because they are not randomly distributed but seem to be clustered in certain regions. There are T-rich regions (for example position 1360-1570) which are flanked by T-poor regions; or from position 0-500 the amount of G is steadily increasing and from position 500-1000 decreasing again. These patterns might result in the binding of certain proteins and/or influence the chromatin structure.

The ORF codes for 1943 aa including 44 methionines. The first methionine (met-1) located at position 5 of the amino acid sequence is proposed to be the beginning of the protein (Figure 4.2). Saul and Battistutta (1990) analyzed the sequence flanking the translational start sites of *P. falciparum* genes and found a preference for AAAAATG directly upstream of the ATG start codon. The sequence at met-1 does not fit into this consensus sequence nor does met-2 at position 72. The first match would be met-3 at position 136 or met-4 at position 193 (Figure 4.2). The determination of which ATG is the real start codon could only be made by directly sequencing the N-terminus of the native protein which is not available yet.

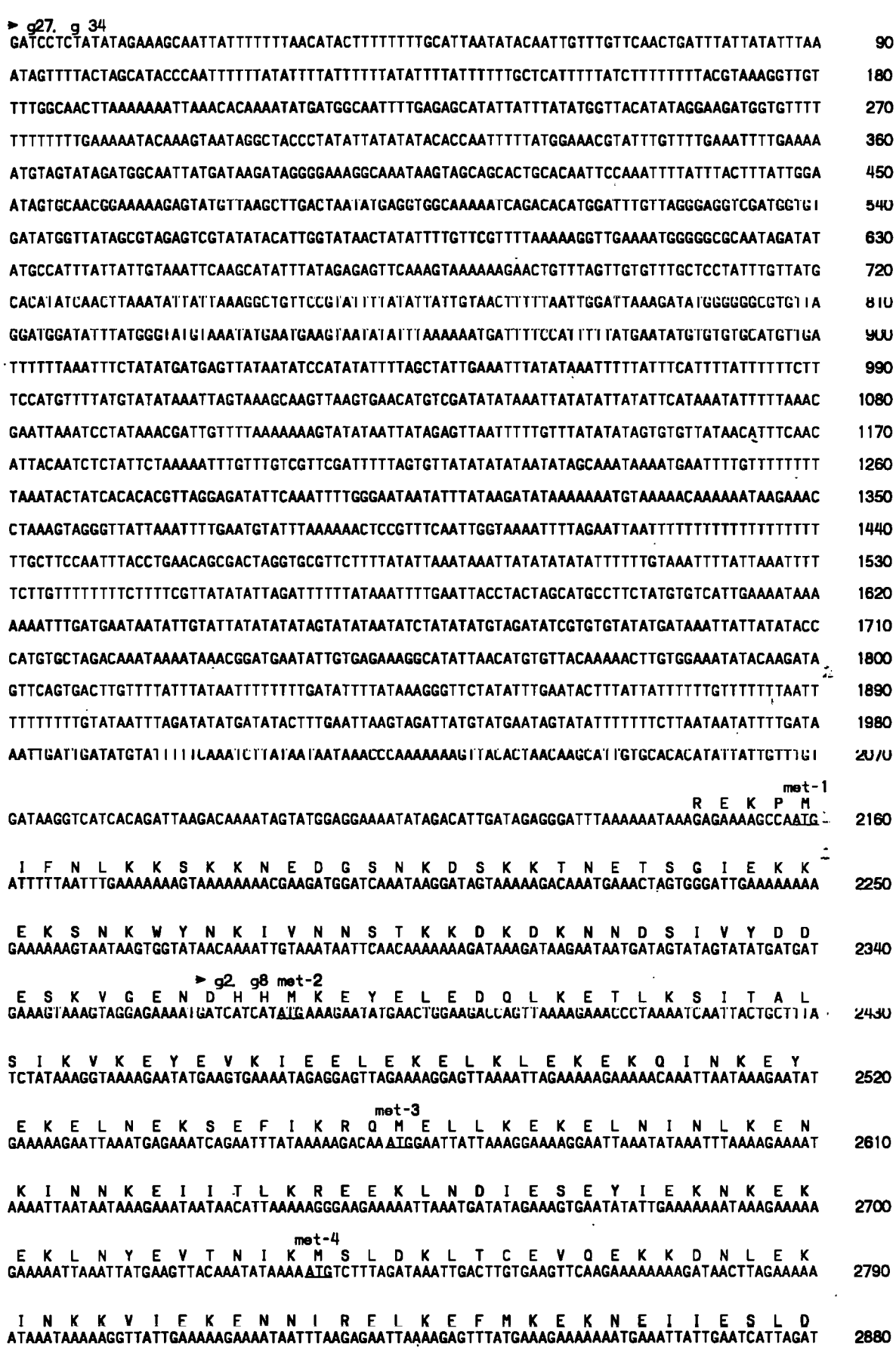
Translation from the first ATG codon of the ORF gives a 229 kDa protein composed of 1939 aa with an isoelectric point pI 5.43 (Figure 4.5). In the centre of the protein is a 364 aa repeat region (aa position 713-1076) based on an 11-mer . Three blocks of 10, 11 and 11 repeats are interrupted by a single truncated repeat composed of 6 aa.

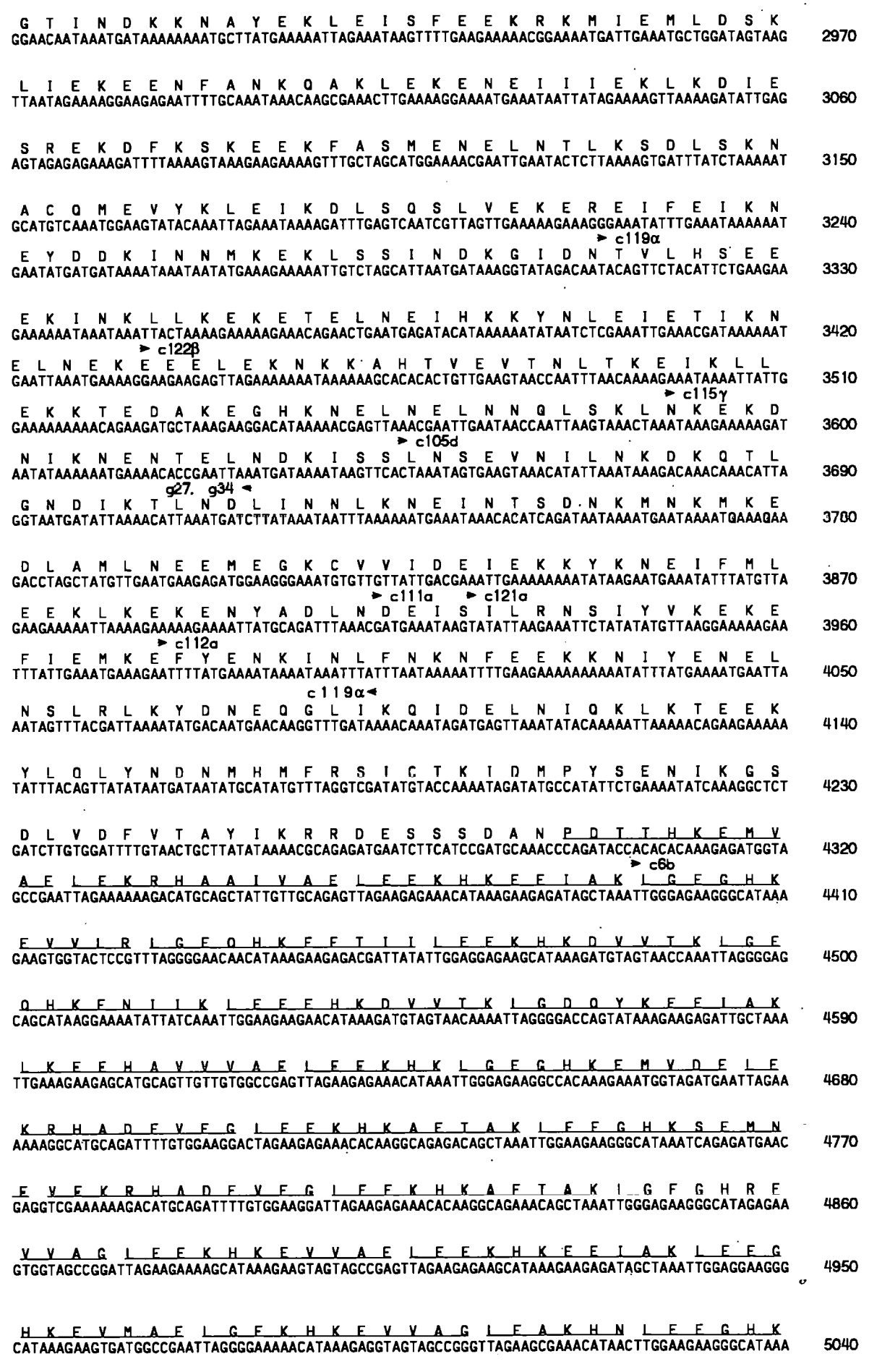
Southern-blot analysis did not reveal a restriction fragment-length polymorphism in this gene between the two different lines of *P. chabaudi*. It further showed that a homologous gene exists in *P. yoelii* and in *P. falciparum*, though their conservation at the DNA level is weak, with an apparently higher homology between the rodent malarial parasites *P. chabaudi* and *P. yoelii* than between the rodent and human parasites *P.chabaudi* and *P. falciparum*. The observation that g7', which contains an internal *Spe* I site, hybridized with a 4 kb and 18.5 kb fragment of *Spe* I digested genomic DNA of *P.chabaudi* and a 22 kb fragment of *P. yoelii* might be due to a loss of the internal *Spe* I site in the homologous gene of *P. yoelii*.

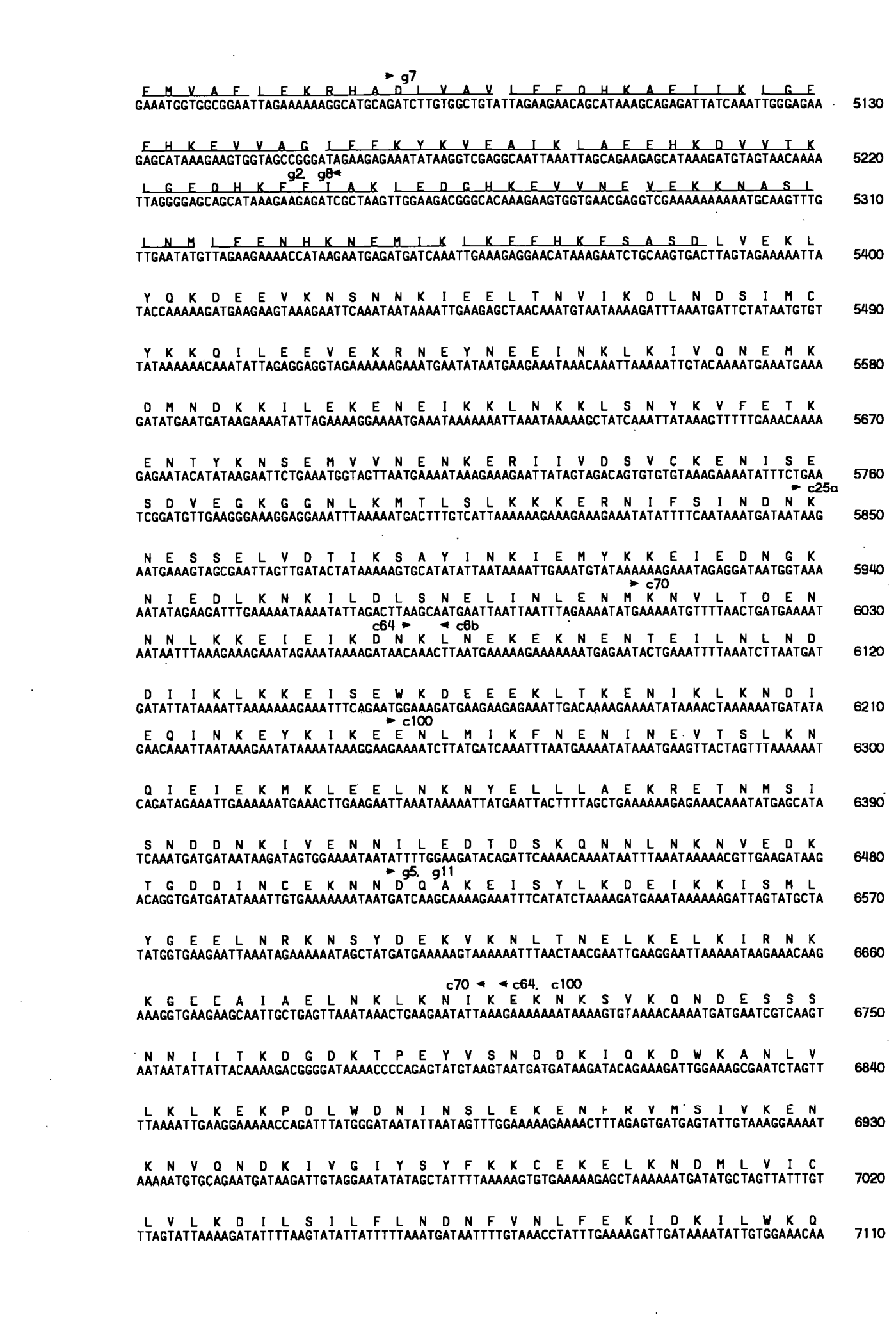
Multiple restriction enzyme digests of genomic DNA from *P. chabaudi* 96V followed by hybridization with a probe from the gene region further indicates that the gene coding for the 229 kDA protein is probably not a member of a multigene family and possibly a single copy gene. However, the presence of related genes, only detectable at low stringency, cannot be ruled out.

Figure 4.2 : **Nucleotide sequence and deduced amino acid sequence of a novel gene and its flanking regions encoding the 229 kDa protein from *P. chabaudi* 96V.**

Genomic DNA clones are indicated by a 'g'; cDNA clones are indicated by a 'c'; ▷ : start of a cloned sequence; ◁ : end of a cloned sequence; met-1,2,3,4: first, second, third and fourth ATG codons in the ORF; amino acid repeats are underlined; possible AATAAA polyadenylation signals are bold and underlined; * : TAA stop-codon of the ORF.


► g27, g 34
met-1
► g2, g8 met-2
met-3
met-4


G T I N D K K N A Y E K L E I S F E E K R K M I E M L D S K
GGAACAATAAATGATAAAAAAAATGCTTATGAAAAATTAGAAATAAGTTTTGAAGAAAAACGGAAAATGATTGAAATGCTGGATAGTAAG 2970
L I E K E E N F A N K Q A K L E K E N E I I I E K L K D I E
TTAATAGAAAAGGAAGAGAATTTTGCAAATAAACAAGCGAAACTTGAAAAGGAAAATGAAATAATTATAGAAAAGTTAAAAGATATTGAG 3060
S R E K D F K S K E E K F A S M E N E L N T L K S D L S K N
AGTAGAGAGAAAGATTTTAAAAGTAAAGAAGAAAAGTTTGCTAGCATGGAAAACGAATTGAATACTCTTAAAAGTGATTTATCTAAAAAT 3150
A C Q M E V Y K L E I K D L S Q S L V E K E R E I F E I K N
GCATGTCAAATGGAAGTATACAAATTAGAAATAAAAGATTTGAGTCAATCGTTAGTTGAAAAAGAAAGGGAAATATTTGAAATAAAAAAT 3240
► c119α
E Y D D K I N N M K E K L S S I N D K G I D N T V L H S E E
GAATATGATGATAAAATAAATAATATGAAAGAAAAATTGTCTAGCATTAATGATAAAGGTATAGACAATACAGTTCTACATTCTGAAGAA 3330
E K I N K L L K E K E T E L N E I H K K Y N L E I E T I K N
GAAAAAATAAATAAATTACTAAAAGAAAAAGAAACAGAACTGAATGAGATACATAAAAAATATAATCTCGAAATTGAAACGATAAAAAAT 3420
► c122β
E L N E K E E E L E K N K K A H T V E V T N L T K E I K L L
GAATTAAATGAAAAGGAAGAAGAGTTAGAAAAAAATAAAAAAGCACACACTGTTGAAGTAACCAATTTAACAAAAGAAATAAAATTATTG 3510
► c115γ
E K K T E D A K E G H K N E L N E L N N Q L S K L N K E K D
GAAAAAAAAACAGAAGATGCTAAAGAAGGACATAAAAACGAGTTAAACGAATTGAATAACCAATTAAGTAAACTAAATAAAGAAAAAGAT 3600
► c105d
N I K N E N T E L N D K I S S L N S E V N I L N K D K Q T L
AATATAAAAAATGAAAACACCGAATTAAATGATAAAATAAGTTCACTAAATAGTGAAGTAAACATATTAAATAAAGACAAACAAACATTA 3690
g27. g34 ◄
G N D I K T L N D L I N N L K N E I N T S D N K M N K M K E
GGTAATGATATTAAAACATTAAATGATCTTATAAATAATTTAAAAAATGAAATAAACACATCAGATAATAAAATGAATAAAATGAAAGAA 3780
D L A M L N E E M E G K C V V I D E I E K K Y K N E I F M L
GACCTAGCTATGTTGAATGAAGAGATGGAAGGGAAATGTGTTGTTATTGACGAAATTGAAAAAAAATATAAGAATGAAATATTTATGTTA 3870
► c111a ► c121a
E E K L K E K E N Y A D L N D E I S I L R N S I Y V K E K E
GAAGAAAAATTAAAAGAAAAAGAAAATTATGCAGATTTAAACGATGAAATAAGTATATTAAGAAATTCTATATATGTTAAGGAAAAAGAA 3960
► c112a
F I E M K E F Y E N K I N L F N K N F E E K K N I Y E N E L
TTTATTGAAATGAAAGAATTTTATGAAAATAAAATAAATTTATTTAATAAAAATTTTGAAGAAAAAAAAAAATATTTATGAAAATGAATTA 4050
c119α◄
N S L R L K Y D N E Q G L I K Q I D E L N I Q K L K T E E K
AATAGTTTACGATTAAAATATGACAATGAACAAGGTTTGATAAAACAAATAGATGAGTTAAATATACAAAAATTAAAAACAGAAGAAAAA 4140
Y L Q L Y N D N M H M F R S I C T K I D M P Y S E N I K G S
TATTTACAGTTATATAATGATAATATGCATATGTTTAGGTCGATATGTACCAAAATAGATATGCCATATTCTGAAAATATCAAAGGCTCT 4230
D L V D F V T A Y I K R R D E S S S D A N P D T T H K E M V
GATCTTGTGGATTTTGTAACTGCTTATATAAAACGCAGAGATGAATCTTCATCCGATGCAAACCCAGATACCACACACAAAGAGATGGTA 4320
► c6b
A E L E K R H A A I V A E L E E K H K E E I A K L G E G H K
GCCGAATTAGAAAAAAGACATGCAGCTATTGTTGCAGAGTTAGAAGAGAAACATAAAGAAGAGATAGCTAAATTGGGAGAAGGGCATAAA 4410
E V V I R L G E Q H K E E T I I L E E K H K D V V T K L G E
GAAGTGGTACTCCGTTTAGGGGAACAACATAAAGAAGAGACGATTATATTGGAGGAGAAGCATAAAGATGTAGTAACCAAATTAGGGGAG 4500
Q H K E N I I K L E E E H K D V V T K I G D Q Y K E E I A K
CAGCATAAGGAAAATATTATCAAATTGGAAGAAGAACATAAAGATGTAGTAACAAAATTAGGGGACCAGTATAAAGAAGAGATTGCTAAA 4590
L K E E H A V V V A E L E E K H K L G E G H K E M V D E L E
TTGAAAGAAGAGCATGCAGTTGTTGTGGCCGAGTTAGAAGAGAAACATAAATTGGGAGAAGGCCACAAAGAAATGGTAGATGAATTAGAA 4680
K R H A D E V E G L E E K H K A E T A K L E E G H K S E M N
AAAAGGCATGCAGATTTTGTGGAAGGACTAGAAGAGAAACACAAGGCAGAGACAGCTAAATTGGAAGAAGGGCATAAATCAGAGATGAAC 4770
E V E K R H A D E V E G L E E K H K A E T A K L G E G H R E
GAGGTCGAAAAAAGACATGCAGATTTTGTGGAAGGATTAGAAGAGAAACACAAGGCAGAAACAGCTAAATTGGGAGAAGGGCATAGAGAA 4860
V V A G L E E K H K E V V A E L E E K H K E E I A K L E E G
GTGGTAGCCGGATTAGAAGAAAAGCATAAAGAAGTAGTAGCCGAGTTAGAAGAGAAGCATAAAGAAGAGATAGCTAAATTGGAGGAAGGG 4950
H K E V M A E L G E K H K E V V A G L E A K H N L E E G H K
CATAAAGAAGTGATGGCCGAATTAGGGGAAAAACATAAAGAGGTAGTAGCCGGGTTAGAAGCGAAACATAACTTGGAAGAAGGGCATAAA 5040


g7
F M V A F L E K R H A D I V A V L F F Q H K A E I I K L G E
GAAATGGTGGCGGAATTAGAAAAAAGGCATGCAGATCTTGTGGCTGTATTAGAAGAACAGCATAAAGCAGAGATTATCAAATTGGGAGAA 5130
E H K E V V A G I F E K Y K V E A I K L A E E H K D V V T K
GAGCATAAAGAAGTGGTAGCCGGGATAGAAGAGAAATATAAGGTCGAGGCAATTAAATTAGCAGAAGAGCATAAAGATGTAGTAACAAAA 5220
g2 g8
L G E Q H K E E I A K L E D G H K E V V N E V E K K N A S L
TTAGGGGAGCAGCATAAAGAAGAGATCGCTAAGTTGGAAGACGGGCACAAAGAAGTGGTGAACGAGGTCGAAAAAAAAAAATGCAAGTTTG 5310
L N M L E E N H K N E M I K L K E E H K E S A S D L V E K L
TTGAATATGTTAGAAGAAAACCATAAGAATGAGATGATCAAATTGAAAGAGGAACATAAAGAATCTGCAAGTGACTTAGTAGAAAAATTA 5400
Y Q K D E E V K N S N N K I E E L T N V I K D L N D S I M C
TACCAAAAAGATGAAGAAGTAAAGAATTCAAATAATAAAATTGAAGAGCTAACAAATGTAATAAAAGATTTAAATGATTCTATAATGTGT 5490
Y K K Q I L E E V E K R N E Y N E E I N K L K I V Q N E M K
TATAAAAAACAAATATTAGAGGAGGTAGAAAAAAGAAATGAATATAATGAAGAAATAAACAAATTAAAAATTGTACAAAATGAAATGAAA 5580
D M N D K K I L E K E N E I K K L N K K L S N Y K V F E T K
GATATGAATGATAAGAAAATATTAGAAAAGGAAAATGAAATAAAAAAATTAAATAAAAAGCTATCAAATTATAAAGTTTTTGAAACAAAA 5670
E N T Y K N S E M V V N E N K E R I I V D S V C K E N I S E
GAGAATACATATAAGAATTCTGAAATGGTAGTTAATGAAAATAAAGAAAGAATTATAGTAGACAGTGTGTGTAAAGAAAATATTTCTGAA 5760
c25a
S D V E G K G G N L K M T L S L K K K E R N I F S I N D N K
TCGGATGTTGAAGGGAAAGGAGGAAATTTAAAAATGACTTTGTCATTAAAAAAGAAAGAAAGAAATATATTTTCAATAAATGATAATAAG 5850
N E S S E L V D T I K S A Y I N K I E M Y K K E I E D N G K
AATGAAAGTAGCGAATTAGTTGATACTATAAAAAGTGCATATATTAATAAAATTGAAATGTATAAAAAAGAAATAGAGGATAATGGTAAA 5940
c70
N I E D L K N K I L D L S N E L I N L E N M K N V L T D E N
AATATAGAAGATTTGAAAAATAAAATATTAGACTTAAGCAATGAATTAATTAATTTAGAAAATATGAAAAATGTTTTAACTGATGAAAAT 6030
c64 c6b
N N L K K E I E I K D N K L N E K E K N E N T E I L N L N D
AATAATTTAAAGAAAGAAATAGAAATAAAAGATAACAAACTTAATGAAAAAGAAAAAAATGAGAATACTGAAATTTTAAATCTTAATGAT 6120
D I I K L K K E I S E W K D E E E K L T K E N I K L K N D I
GATATTATAAAATTAAAAAAAGAAATTTCAGAATGGAAAGATGAAGAAGAGAAATTGACAAAAGAAAATATAAAACTAAAAAATGATATA 6210
c100
E Q I N K E Y K I K E E N L M I K F N E N I N E V T S L K N
GAACAAATTAATAAAGAATATAAAATAAAGGAAGAAAATCTTATGATCAAATTTAATGAAAATATAAATGAAGTTACTAGTTTAAAAAAT 6300
Q I E I E K M K L E E L N K N Y E L L L A E K R E T N M S I
CAGATAGAAATTGAAAAAATGAAACTTGAAGAATTAAATAAAAATTATGAATTACTTTTAGCTGAAAAAAGAGAAACAAATATGAGCATA 6390
S N D D N K I V E N N I L E D T D S K Q N N L N K N V E D K
TCAAATGATGATAATAAGATAGTGGAAAATAATATTTTGGAAGATACAGATTCAAAACAAAATAATTTAAATAAAAACGTTGAAGATAAG 6480
g5 g11
T G D D I N C E K N N D Q A K E I S Y L K D E I K K I S M L
ACAGGTGATGATATAAATTGTGAAAAAAATAATGATCAAGCAAAAGAAATTTCATATCTAAAAGATGAAATAAAAAAGATTAGTATGCTA 6570
Y G E E L N R K N S Y D E K V K N L T N E L K E L K I R N K
TATGGTGAAGAATTAAATAGAAAAAATAGCTATGATGAAAAAGTAAAAAATTTAACTAACGAATTGAAGGAATTAAAAATAAGAAACAAG 6660
c70 c64, c100
K G E E A I A E L N K L K N I K E K N K S V K Q N D E S S S
AAAGGTGAAGAAGCAATTGCTGAGTTAAATAAACTGAAGAATATTAAAGAAAAAAATAAAAGTGTAAAACAAAATGATGAATCGTCAAGT 6750
N N I I T K D G D K T P E Y V S N D D K I Q K D W K A N L V
AATAATATTATTACAAAAGACGGGGATAAAACCCCAGAGTATGTAAGTAATGATGATAAGATACAGAAAGATTGGAAAGCGAATCTAGTT 6840
L K L K E K P D L W D N I N S L E K E N F R V M S I V K E N
TTAAAATTGAAGGAAAAACCAGATTTATGGGATAATATTAATAGTTTGGAAAAAGAAAACTTTAGAGTGATGAGTATTGTAAAGGAAAAT 6930
K N V Q N D K I V G I Y S Y F K K C E K E L K N D M L V I C
AAAAATGTGCAGAATGATAAGATTGTAGGAATATATAGCTATTTTAAAAAGTGTGAAAAAGAGCTAAAAAATGATATGCTAGTTATTTGT 7020
L V L K D I L S I L F L N D N F V N L F E K I D K I L W K Q
TTAGTATTAAAAGATATTTTAAGTATATTATTTTTAAATGATAATTTTGTAAACCTATTTGAAAAGATTGATAAAATATTGTGGAAACAA 7110

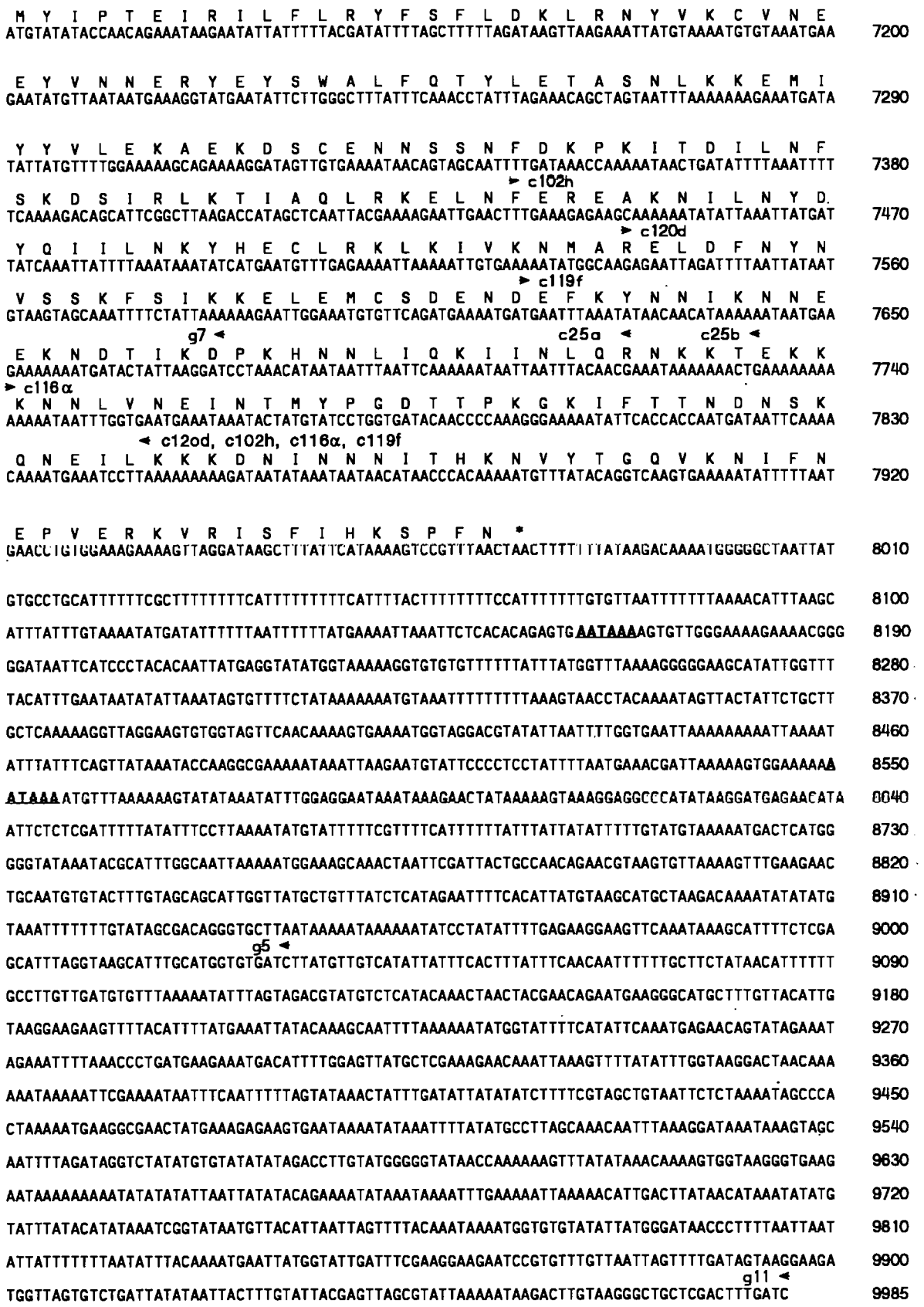

M Y I P T E I R I L F L R Y F S F L D K L R N Y V K C V N E
ATGTATATACCAACAGAAATAAGAATATTATTTTTACGATATTTTAGCTTTTTAGATAAGTTAAGAAATTATGTAAAATGTGTAAATGAA 7200
E Y V N N E R Y E Y S W A L F Q T Y L E T A S N L K K E M I
GAATATGTTAATAATGAAAGGTATGAATATTCTTGGGCTTTATTTCAAACCTATTTAGAAACAGCTAGTAATTTAAAAAAAGAAATGATA 7290
Y Y V L E K A E K D S C E N N S S N F D K P K I T D I L N F
TATTATGTTTTGGAAAAAGCAGAAAAGGATAGTTGTGAAAATAACAGTAGCAATTTTGATAAACCAAAAATAACTGATATTTTAAATTTT 7380
c102h
S K D S I R L K T I A Q L R K E L N F E R E A K N I L N Y D
TCAAAAGACAGCATTCGGCTTAAGACCATAGCTCAATTACGAAAAGAATTGAACTTTGAAAGAGAAGCAAAAAATATATTAAATTATGAT 7470
c120d
Y Q I I L N K Y H E C L R K L K I V K N M A R E L D F N Y N
TATCAAATTATTTTAAATAAATATCATGAATGTTTGAGAAAATTAAAAATTGTGAAAAATATGGCAAGAGAATTAGATTTTAATTATAAT 7560
c119f
V S S K F S I K K E L E M C S D E N D E F K Y N N I K N N E
GTAAGTAGCAAATTTTCTATTAAAAAAGAATTGGAAATGTGTTCAGATGAAAATGATGAATTTAAATATAACAACATAAAAAATAATGAA 7650
g7 c25a c25b
E K N D T I K D P K H N N L I Q K I I N L Q R N K K T E K K
GAAAAAAATGATACTATTAAGGATCCTAAACATAATAATTTAATTCAAAAAATAATTAATTTACAACGAAATAAAAAAACTGAAAAAAAA 7740
c116α
K N N L V N E I N T M Y P G D T T P K G K I F T T N D N S K
AAAAATAATTTGGTGAATGAAATAAATACTATGTATCCTGGTGATACAACCCCAAAGGGAAAAATATTCACCACCAATGATAATTCAAAA 7830
c12od, c102h, c116α, c119f
Q N E I L K K K D N I N N N N I T H K N V Y T G Q V K N I F N
CAAAATGAAATCCTTAAAAAAAAAGATAATATAAATAATAACATAACCCACAAAAATGTTTATACAGGTCAAGTGAAAAATATTTTTAAT 7920
E P V E R K V R I S F I H K S P F N *
GAACCTGTGGAAAGAAAAGTTAGGATAAGCTTTATTCATAAAAGTCCGTTTAACTAACTTTTTTTTATAAGACAAAATGGGGGCTAATTAT 8010
GTGCCTGCATTTTTTCGCTTTTTTTTCATTTTTTTTTCATTTTACTTTTTTTTCCATTTTTTTGTGTTAATTTTTTTAAAACATTTAAGC 8100
ATTTATTTGTAAAATATGATATTTTTTAATTTTTTATGAAAATTAAATTCTCACACAGAGTGAATAAAAGTGTTGGGAAAAGAAAACGGG 8190
GGATAATTCATCCCTACACAATTATGAGGTATATGGTAAAAAGGTGTGTGTTTTTTATTTATGGTTTAAAAGGGGGAAGCATATTGGTTT 8280
TACATTTGAATAATATATTAAATAGTGTTTTCTATAAAAAAAATGTAAATTTTTTTTTTAAAGTAACCTACAAAATAGTTACTATTCTGCTT 8370
GCTCAAAAAGGTTAGGAAGTGTGGTAGTTCAACAAAAGTGAAAATGGTAGGACGTATATTAATTTTGGTGAATTAAAAAAAAATTAAAAT 8460
ATTTATTTCAGTTATAAATACCAAGGCGAAAAATAAATTAAGAATGTATTCCCCTCCTATTTTAATGAAACGATTAAAAAGTGGAAAAAA 8550
ATAAAAATGTTTAAAAAAGTATATAAATATTTGGAGGAATAAATAAAGAACTATAAAAAGTAAAGGAGGCCCATATAAGGATGAGAACATA 0040
ATTCTCTCGATTTTTATATTTCCTTAAAATATGTATTTTTCGTTTTCATTTTTTATTTATTATATTTTTGTATGTAAAAATGACTCATGG 8730
GGGTATAAATACGCATTTGGCAATTAAAAATGGAAAGCAAACTAATTCGATTACTGCCAACAGAACGTAAGTGTTAAAAGTTTGAAGAAC 8820
TGCAATGTGTACTTTGTAGCAGCATTGGTTATGCTGTTTATCTCATAGAATTTTCACATTATGTAAGCATGCTAAGACAAAATATATATG 8910
TAAATTTTTTTTGTATAGCGACAGGGTGCTTAATAAAAATAAAAAATATCCTATATTTTGAGAAGGAAGTTCAAATAAAGCATTTTCTCGA 9000
g5
GCATTTAGGTAAGCATTTGCATGGTGTGATCTTATGTTGTCATATTATTTCACTTTATTTCAACAATTTTTTGCTTCTATAACATTTTTT 9090
GCCTTGTTGATGTGTTTAAAAATATTTAGTAGACGTATGTCTCATACAAACTAACTACGAACAGAATGAAGGGCATGCTTTGTTACATTG 9180
TAAGGAAGAAGTTTTACATTTTATGAAATTATACAAAGCAATTTTAAAAAATATGGTATTTTCATATTCAAATGAGAACAGTATAGAAAT 9270
AGAAATTTTAAACCCTGATGAAGAAATGACATTTTGGAGTTATGCTCGAAAGAACAAATTAAAGTTTTATATTTGGTAAGGACTAACAAA 9360
AAATAAAAATTCGAAAATAATTTCAATTTTTAGTATAAACTATTTGATATTATATATCTTTTCGTAGCTGTAATTCTCTAAAATAGCCCA 9450
CTAAAAATGAAGGCGAACTATGAAAGAGAAGTGAATAAAATATAAATTTTATATGCCTTAGCAAACAATTTAAAGGATAAATAAAGTAGC 9540
AATTTTAGATAGGTCTATATGTGTATATATAGACCTTGTATGGGGGTATAACCAAAAAAGTTTATATAAACAAAAGTGGTAAGGGTGAAG 9630
AATAAAAAAAAATATATATATTAATTATATACAGAAAATATAAATAAAATTTGAAAAATTAAAAACATTGACTTATAACATAAATATATG 9720
TATTTATACATATAAATCGGTATAATGTTACATTAATTAGTTTTACAAATAAAATGGTGTGTATATTATGGGATAACCCTTTTAATTAAT 9810
ATTATTTTTTTAATATTTACAAAATGAATTATGGTATTGATTTCGAAGGAAGAATCCGTGTTTGTTAATTAGTTTTGATAGTAAGGAAGA 9900
g11
TGGTTAGTGTCTGATTATATAATTACTTTGTATTACGAGTTAGCGTATTAAAAATAAGACTTGTAAGGGCTGCTCGACTTTGATC 9985

Figure 4.3 :

Sequence obtained from genomic clones from the novel gene and its flanking regions encoding the 229 kDa protein from *P. chabaudi* 96V.

a) 1st, 2nd : selected clones from the first and second screening of the genomic library; ■ : sequenced part of a genomic clone; - : non-sequenced part of a genomic clone; → : downstream primer; ← : upstream primer; ⊢⊣ : ^{32}P-labelled PCR fragments used as probes during 1st and 2nd screening; |||||||||||||||||||| : repeat region; ▸ cDNA ◂ : sequence covered by genomic clones as well as by cDNA clones; ▹ ORF ◃ : open reading frame region; 'g' indicates that the cloned sequence is genomic DNA.

b) Partial restriction map of the same region. Restriction enzyme sites : S - *Sau* 3A, T - *Taq* I, Sp - *Spe* I.

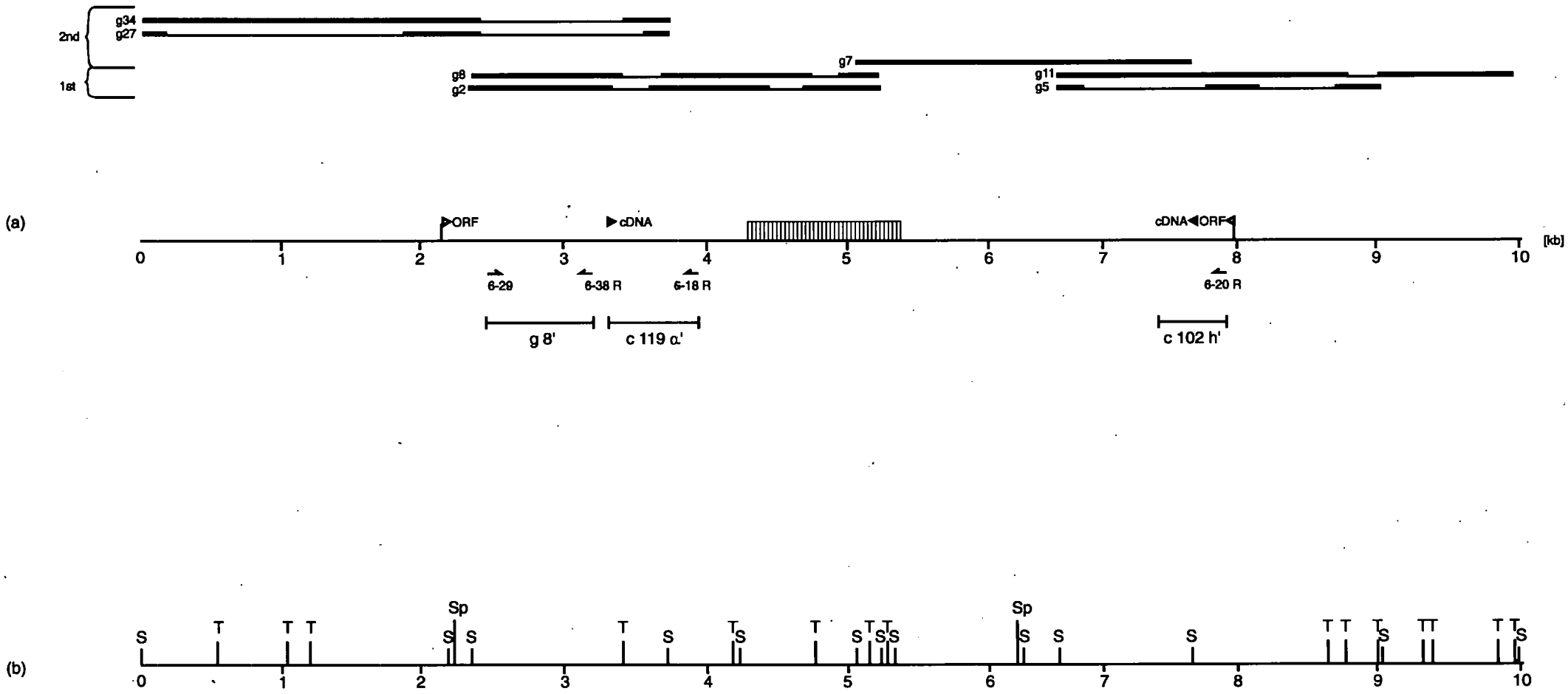

2nd
1st
g34
g27
g8
g2
g7
g11
g5
(a)
ORF
cDNA
cDNA
ORF
[kb]
0
1
2
3
4
5
6
7
8
9
10
6-29
6-38 R
6-18 R
6-20 R
g 8'
c 119 α'
c 102 h'
(b)
Sp
Sp
S
T
0
1
2
3
4
5
6
7
8
9
10

Figure 4.4 :

Sequencing strategy for the gene and its flanking region encoding the 229 kDa protein from *P. chabaudi* 96V.

Green arrows : genomic DNA clones; red arrows : cDNA clones; → : coding strand was sequenced; ← : complementary strand was sequenced; ▷ ORF ◁ : open reading frame region; ||||||||||||||||||| : repeat region.

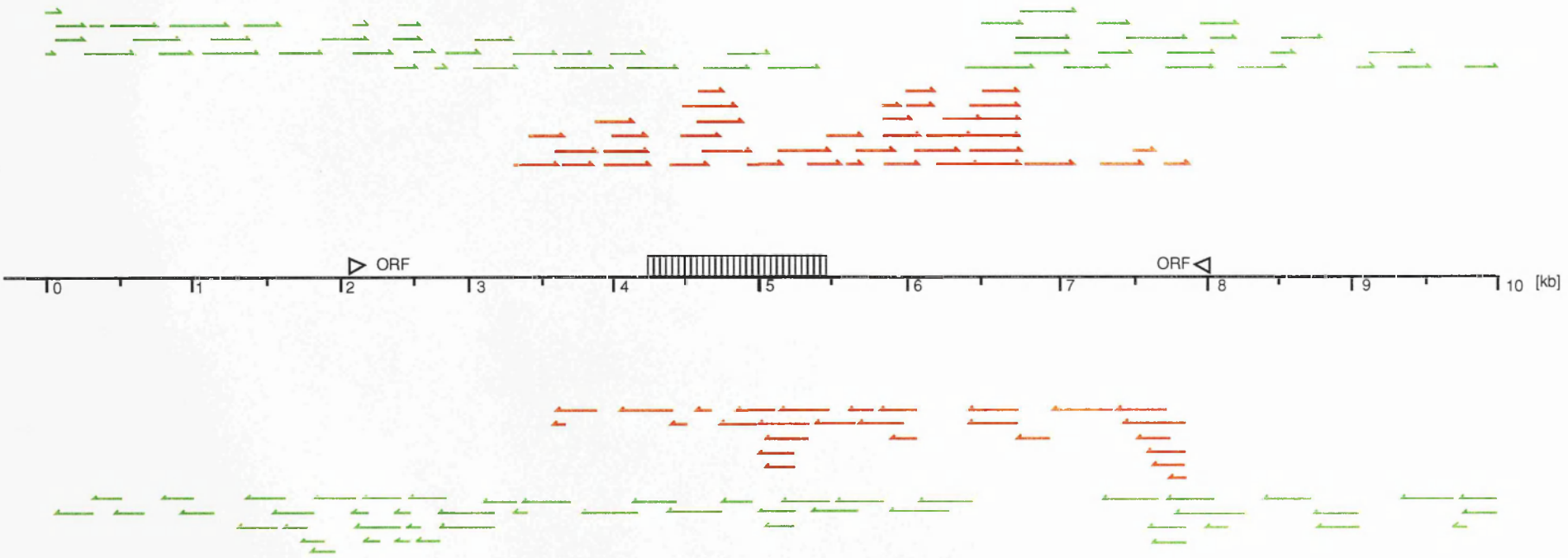

ORF
ORF
0
1
2
3
4
5
6
7
8
9
10
[kb]

Figure 4.5 : **Deduced amino acid sequence from the 229 kDa protein from P. chabaudi 96V.**

The repeat region (713-1076) is depicted in a way that the repeat units are aligned; cysteines are bold; the hydrophobic region is underlined.

MIFNLKKSKKNEDGSNKDSKKTNETSGIEKKEKSNKWYNKIVNNSTKKDKDKNNDSIVYD
DESKVGENDHHMKEYELEDQLKETLKSITALSIKVKEYEVKIEELEKELKLEKEKQINKE 60
YEKELNEKSEFIKRQMELLKEKELNINLKENKINNKEIITLKREEKLNDIESEYIEKNKE 120
KEKLNYEVTNIKMSLDKLTCEVQEKKDNLEKINKKVIEKENNLRELKEFMKEKNEIIESL 180
DGTINDKKNAYEKLEISFEEKRKMIEMLDSKLIEKEENFANKQAKLEKENEIIIEKLKDI 240
ESREKDFKSKEEKFASMENELNTLKSDLSKNACQMEVYKLEIKDLSQSLVEKEREIFEIK 300
NEYDDKINNMKEKLSSINDKGIDNTVLHSEEEKINKLLKEKETELNEIHKKYNLEIETIK 360
NELNEKEEELEKNKKAHTVEVTNLTKEIKLLEKKTEDAKEGHKNELNELNNQLSKLNKEK 420
DNIKNENTELNDKISSLNSEVNILNKDKQTLGNDIKTLNDLINNLKNEINTSDNKMNKMK 480
EDLAMLNEEMEGKCVVIDEIEKKYKNEIFMLEEKLKEKENYADLNDEISILRNSIYVKEK 540
EFIEMKEFYENKINLFNKNFEEKKNIYENELNSLRLKYDNEQGLIKQIDEL- 600
NIQKLKTEEKYLQLYNDNMHMFRSICTKIDMPYSENIKGSDLVDFVTAYIKRRDES- 660
SSDAN 712

P D T T H K E M V A E 723
L E K R H A A I V A E 734
L E E K H K E E I A K 745
L G E G H K E V V L R 756
L G E Q H K E E T I I 767
L E E K H K D V V T K 778
L G E Q H K E N I I K 789
L E E E H K D V V T K 800
L G D Q Y K E E I A K 811
L K E E H A V V V A E 822
L E E K H K 828
L G E G H K E M V D E 839
L E K R H A D F V E G 850
L E E K H K A E T A K 861
L E E G H K S E M N E 872
V E K R H A D F V E G 883
L E E K H K A E T A K 894
L G E G H R E V V A G 905
L E E K H K E V V A E 916
L E E K H K E E I A K 927
L E E G H K E V M A E 938
L G E K H K E V V A G 949
L E A K H N 955
L E E G H K E M V A E 966
L E K R H A D L V A V 977
L E E Q H K A E I I K 988
L G E E H K E V V A G 999
I E E K Y K V E A I K 1010
L A E E H K D V V T K 1021
L G E Q H K E E I A K 1032
L E D G H K E V V N E 1043
V E K K N A S L L N M 1054
L E E N H K N E M I K 1065
L K E E H K E S A S D 1076

LVEKLYQKDEEVKNSNNKIEELTNVIKDLNDSIMCYKKQILEEVEKRNEYNEEINKLKIV 1136
QNEMKDMNDKKILEKENEIKKLNKKLSNYKVFETKENTYKNSEMVVNENKERIIVDSVCK 1196
ENISESDVEGKGGNLKMTLSLKKKERNIFSINDNKNESSELVDTIKSAYINKIEMYKKEI 1256
EDNGKNIEDLKNKILDLSNELINLENMKNVLTDENNNLKKEIEIKDNKLNEKEKNENTEI 1316
LNLNDDIIKLKKEISEWKDEEEEKLTKENIKLKNDIEQINKEYKIKEENLMIKFNENINEV 1376
TSLKNQIEIEKMKLEELNKNYELLLAEKRETNMSISNDDNKIVENNILEDTDSKQNNLNK 1436
NVEDKTGDDINCEKNNDQAKEISYLKDEIKKISMLYGEELNRKNSYDEKVKNLTNELKEL 1496
KIRNKKGEEAIAELNKLKNIKEKNKSVKQNDESSSNNIITKDGDKTPEYVSNDDKIQKDW 1556
KANLVLKLKEKPDLWDNINSLEKENFRVMSIVKENKNVQNDKIVGIYSYFKKCEKELKND 1616
MLVICLVLKDILSILFLNDNFVNLFEKIDKILWKQMYIPTEIRILFLRYFSFLDKLRNYV 1676
KCVNEEYVNNERYEYSWALFQTYLETASNLKKEMIYYVLEKAEKDSCENNSSNFDKPKI 1736
TDILNFSKDSIRLKTIAQLRKELNFEREAKNILNYDYQIILNKYHECLRKLKIVKNMARE 1796
LDFNYNVSSKFSIKKELEMCSDENDEFKYNNIKNNEEKNDTIKDPKHNNLIQKIINLQRN 1856
KKTEKKKNNLVNEINTMYPGDTTPKGKIFTTNDNSKQNEILKKKDNIN- 1916
NNITHKNVYTGQVKNIFNEPVERKVRISFIHKSPFN 1939

Figure 4.6 : **Southern-blot analysis of genomic DNA from *Plasmodium* ssp. with a ^{32}P-labelled fragment from the coding region of the 229 kDa protein.**

Genomic DNA of *P. chabaudi* 96V (Pch 96V), *P. chabaudi* CB (Pch CB), *P. yoelii* (Py) and *P. falciparum* (Pf) was *Spe* I digested, separated on an agarose gel, transferred to nitrocellulose and probed with ^{32}P-labelled fragment g7'. Four µg of digested DNA from each *Plasmodium* spp. were run per lane.

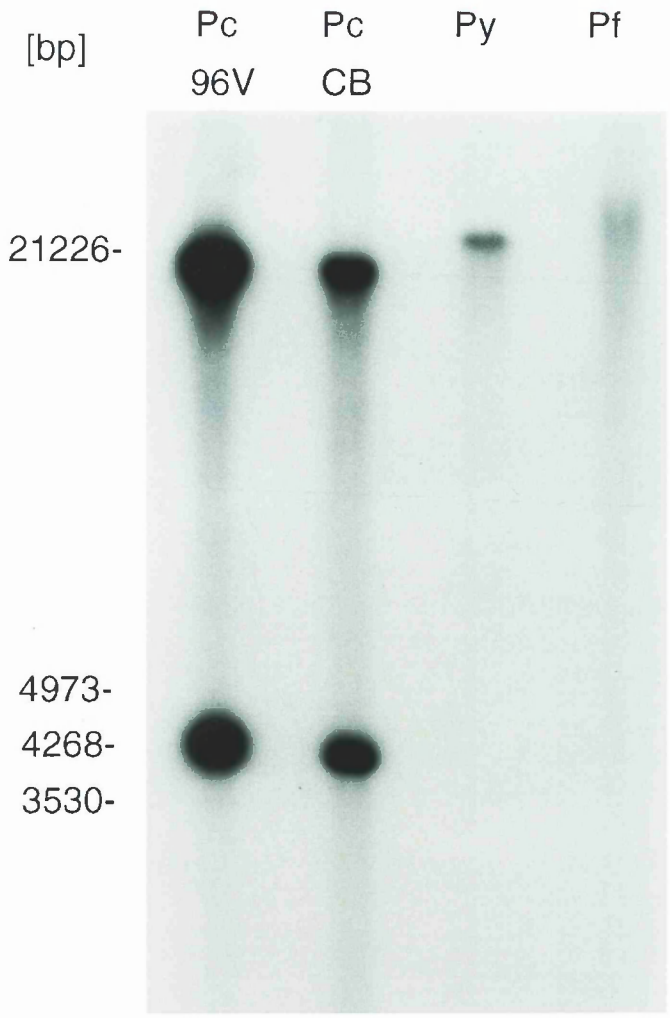
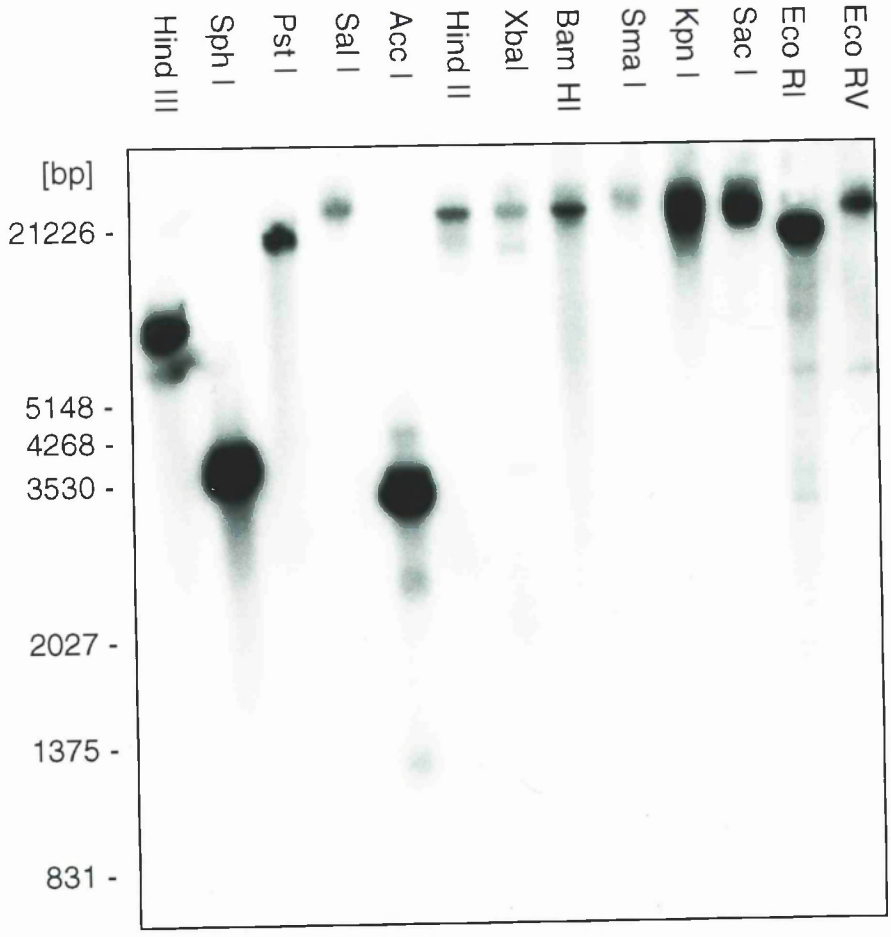

[bp]
Pc
96V
Pc
CB
Py
Pf
21226-
4973-
4268-
3530-

Figure 4.7: **Southern-blot analysis of the up- and downstream region of the gene coding for the 229 kDa protein of *P. chabaudi* 96V.**

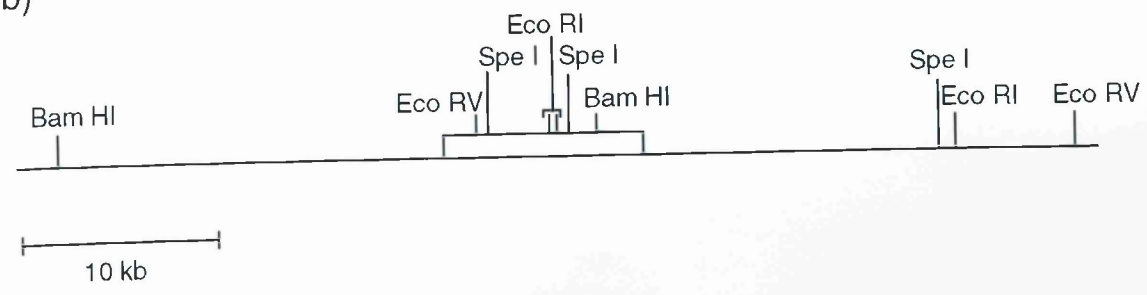
a) Genomic DNA of *P. chabaudi 96*V was digested with *Hind* III, *Sph* I, *Pst* I, *Sal* I, *Acc* I, *Hind* II, *Xba* I, *Bam* HI, *Sma* I, *Kpn* I, *Sac* I, *Eco* RI and *Eco* RV, separated on an agarose gel, transferred to nitrocellulose and probed with ^{32}P-labelled insert of clone c70. Eight µg of digested DNA were run per lane.

b) Restriction map of the chromosomal environment of the 229 kDa protein gene region of *P. chabaudi* 96V. The map was made with the restriction enzymes used in Figures 4.6 and 4.7 a), which cut inside and outside of the sequenced 9985 bp region of the gene. ▭ : sequenced region; ——— : non-sequenced region.

a)


Hind III
Sph I
Pst I
Sal I
Acc I
Hind II
Xbal
Bam HI
Sma I
Kpn I
Sac I
Eco RI
Eco RV
[bp]
21226 -
5148 -
4268 -
3530 -
2027 -
1375 -
831 -

b)


Bam HI
Eco RV
Spe I
Eco RI
Spe I
Bam HI
Spe I
Eco RI
Eco RV
10 kb

4.3. Analysis of the deduced amino acid sequence of the 229 kDa protein

4.3.1. Higher-order structure and functional sites analysis of the 229 kDa protein

a) Secondary structure predictions

Applying the Garnier, Osguthorpe and Robson method to predict α-helix, β-sheet, turn and coiled regions (Garnier *et al.*, 1978) showed that most of the protein could adopt an α-helical conformation. The probability of an α-helical conformation was particularly high for the entire 11-mer repeat region. Regions with short α-helical stretches and a high probability for β-sheet, turn and coiled conformations were found in the first 50 aa of the N-terminus, the last 400 aa of the C-terminus and about 40 aa upstream of the N-terminus of the repeat region (Figure 4.8).

b) Hydrophilicity patterns

A hydrophilicity profile was calculated using the approach of Kyte and Doolittle (1982). A window of 18 amino acid residues was chosen in order to get a precise information about possible membrane spanning segments. The overall character of the protein was highly hydrophilic. Only two adjacent stretches of 18 aa each (position 1617-1634 and 1652-1669) showed a strong hydrophobicity, of which the first was particularly hydrophobic (Figure 4.8).

c) Coiled coil arrangements

To detect coiled coil regions the approach of Lupas *et al.* (1991) was used. Figure 4.9 shows that three large coiled coil domains might exist at position 60-660, 710-1180 and 1230-1520 (domain I, II and III). Domain II might be interrupted by multiple α-helical regions that do not build coiled coil superhelices and domain III by coiled regions and by α-helical regions that do not form coiled-coil superhelices (Figures 4.8 and 4.9).

d) Potential sites for covalent modification

Seventeen N-glycosylation sites, which are characterized by the sequence N-X-(S,T) (Hubbard and Ivatt, 1981), are present along the protein (data not shown). Fifty-seven sites were found where serine, threonine or tyrosine could be phosphorylated by four different protein kinases (data not shown). Besides these motifs no other known characteristic sequence pattern were detected.

4.3.2. Repeat structures

The amino acid sequence of the 229 kDa protein contains in its centre a large repeat region based on an 11-mer repeat unit (Figure 4.5). In order to examine whether or not other, less easily detectable repeat structures, were present in the protein the Pustell protein homology matrix was used to compare the protein to itself and the resultant plot is shown in Figure 4.10.

Using a 14 aa window to highlight the coiled-coil regions and a 50% minimum score mainly the repeat region (position 713-1076) is clearly visible (data not shown). By reducing the score to 20%, regions of low homology become detectable as well (Figure 4.10) which were found to be coincident with the predicted coiled coil regions. Because the α-helical coiled coil structure is normally based on a repeating 7 residue unit (a, b, c, d, e, f, g) in which a and d are mainly hydrophobic and the other residues are often charged, these motifs can be detected when a low homology score is used.

Besides the main repeat region, based on the 11-mer unit, and the repeating 7-mer units of the α-helical coiled coil regions no other repeat structures could be detected by this method.

4.3.3. Homology searches

For sequence homology searches three segments of the 229 kDa protein were chosen:

a) Position 1-1020 : a region predicted to exist mainly in coiled-coil conformation .

b) Position 1021-1939 : a region in which the first half (position 1021-1520) was predicted to exist in coiled-coil conformation (with interruptions) and in the second half (position 1521-1939) probably no or very short coiled-coil arrangements occur but with the highest predicted degree of non α-helical conformation.

b) Position 1515-1939: about the same as the second half in b) but without the predicted coiled-coil part.

A search of the Swissprot protein sequence databank (Release No.84) was performed with the BLITZ program which uses the Smith and Waterman (1981) comparison algorithm and with the FASTA program from Pearson and Lipman (1988).
From the group of non malarial proteins the results showed a maximum similarity of about 20% mainly of the predicted α-helical coiled-coil regions of the 229 kDa protein with a variety of proteins with known α-helical coiled-coil structure, e.g. the rod region of the myosin heavy chain, paramyosin and intermediate filaments.
The similarities of malarial proteins with the 229 kDa protein were to two different parts of the 229 kDa molecule, either to regions with predicted α-helical coiled-coil structure or to the 11-mer repeat region.
The group of malarial proteins which showed similarities with the 229 kDa protein included MESA (18.4%/ 607 aa) (Coppel *et al.*, 1992), the gene 11-1 protein (26.1%/ 241 aa) (Scherf *et al.*, 1988) and MSP-1 (18.4%/ 512 aa) (Deleersnijder *et al.*, 1986, 1988) all of which showed similarities with the predicted coiled-coil regions of the 229 kDa protein.

The 11-mer repeat region of the S-antigen (Cowman *et al.*, 1985) showed a similarity of 23.8% with the entire unit II of the 11-mer repeat region from the 229 kDa protein. These similarities were based on the residues L, E, K, and the last G of the S-antigen repeat $(LEDPAKASQGG)_{13}$ relative to the 11-mer of the 229 kDa protein (Figure 4.5).

4.3.4. Discussion

In the course of the many large scale sequencing projects that are undertaken it is quite common to obtain sequences for as yet uncharacterized proteins. Significant progress has been made recently in the improvement of algorithms and methods of sequence analysis by computer (Bork *et al.*, 1994; Green, 1990, von Heijne, 1991). There are various methods to predict secondary structures, hydrophobicity patterns, transmembrane regions and coiled coil arrangements.

Another powerful set of tools are sequence database alignment searches. It is helpful to interpret a new protein sequence if it can be related to a protein of known structure and function and may give support for predicted secondary or higher ordered structures. There is an increasing number of characterized short protein sequence motifs, which represent specific binding sites, domains or sites of covalent modification and these are collected in PROSITE (Bairoch, 1993).

Applying these tools can suggest that the sequence is the one of interest and might give some insights into the structure and possible biological role of an uncharacterized protein. This helps to direct further practical approaches in order to verify predicted properties.

The results predict that large regions of the 229 kDa protein form coiled-coil superhelices. These regions can be divided into three domains, which might be further interrupted (Figure 4.9). It seems that the termini of the molecule do not form a coiled-coil superhelix and might adopt a more globular structure. This overall molecular structure resembles that of other α-fibrous proteins, e.g. myosin (Dibb *et al.*, 1989), paramyosin (Kagawa *et al.*, 1989) and especially the large multigene family of intermediate filaments. The latter show besides globular termini of variable size regular interruptions of the coiled coil rod (Parry *et al.*, 1985, North *et al.*, 1994).

The similarities found to our sequence were mainly to the same group of proteins and rather low (about 20% identity). The coiled-coil arrangements of α-fibrous proteins is based on heptad repeats (position a, b, c, d, e, f, g) which occur in an α-helix (α-helical coiled-coil superhelix). Position a and d are mainly occupied by apolar amino acids, the other positions are often occupied by charged or polar amino acids.. There is a certain preference for a subset of amino acids for each position, e.g. Leu, Ile, Val for position a and d; Asp, Glu, Lys, Arg, Gln, Asn for position b, c, e, f, g (Cohen and Parry, 1990). The observed similarities were mainly based on identities of the latter amino acids and seem just to reflect the common structures of these molecules, supporting the predicted predominant coiled-coil arrangement of our protein.

The malarial proteins that showed similarities with the α-helical coiled-coil parts of the 229 kDa protein were MESA (mature parasite infected erythrocyte surface antigen), the gene 11-1 protein and MSP-1 (major merozoite surface protein). MESA and the gene 11-1 protein have both been found to be associated with the membrane of the infected red blood cell and both exist mainly in a predicted α-helical secondary structure (Coppel *et al.*, 1992; Scherf *et al.*, 1988). MSP-1 is associated with the surface of the mature merozoite and for large regions of this molecule an α-helical secondary structure was predicted as well.

The similarities between these three molecules and the 229 kDa protein were mainly based on the same residues as those found in the non malarial proteins above suggesting that these similarities are probably based on a similar overall structure and that these three malarial proteins may form α-helical coiled-coil structures themselves. This might be noteworthy because this structure would be formed, in the case of MESA and the gene 11-1 protein, largely by repeat regions. The homologies with the 11-mer repeat region of the S-antigen were based on the residues Leu, Glu, Lys and the second Gly of the S-antigen repeat $(LEDPAKASQGG)_{13}$, a finding that will be discussed in more detail in the following chapter.

The repeat region of the 229 kDa protein is composed of 32 11-mer repeats based on LEEKHKEEVAK with some degree of variability, and 2 6-mer repeats which represent only the first 6 residues of the 11-mer repeat. These truncated repeats are located behind repeat 10 and 21. The three uninterrupted 11-mer repeat regions were named unit I, II and III respectively (Figures 4.5 and 4.11).

An intriguing feature of the sequence is the occurrence of a hydrophobic stretch of 18 aa acids located at position 1617-1634 (Figure 4.8). There is a negatively charged Glu, a positively charged Lys and a polar Ser in the centre of the hydrophobic segment (Figure 4.5). Interestingly they are all located on one face of an α-helix (data not shown) and it is possible that a number of these segments form a bundle in the cell membrane where the polar amino acids build the core and the apolar amino acid residues interact at the outside with the lipid bilayer. Transmembrane sequences that contain charged amino acids are often found in proteins that form bundles on the inside surface of a membrane, e.g. the bacterial rhodopsin (Henderson *et al.*, 1990) and the NADH:ubiquinone reductase (complex I) (Preis *et al.*, 1990) which are thought to be stabilized by the formation of ion pairs (Engelman *et al.*, 1980). A consequence of this assumption would be that this sequence is a signal-anchor sequence like the transmembrane segment found in the human transferrin receptor (Zerial *et al.*, 1986) and the human asialoglycoprotein receptor (Spiess and Lodish, 1986) because no other signal peptide can be detected in the 229 kDa protein.
The idea that the protein might be a transmembrane protein is hypothetical. The other alternative is that the hydrophobic stretch is not exposed to a polar environment but buried in a hydrophobic environment provided by interaction with a protein rather than by a lipid bilayer.

The asymmetric position of the coiled-coil arrangement led to the supposition that either two or more identical molecules build a parallel coiled-coil superhelix with about 60 aa

at the N-terminus and about 420 aa at the C-terminus forming globular structures.

The most remarkable feature of the sequence is the occurrence of the 11-mer repeat which is analyzed in greater detail in the following chapter.

Figure 4.8 :

Analysis of the secondary structure and hydrophilicity profile of the 229 kDa protein of *P. chabaudi* 96V.

The results for α-helical regions are displayed at the top as a region graph and below as a line plot graph; β-, turn- and coil-regions are displayed as region graphs and hydrophilicity as a line plot graph.

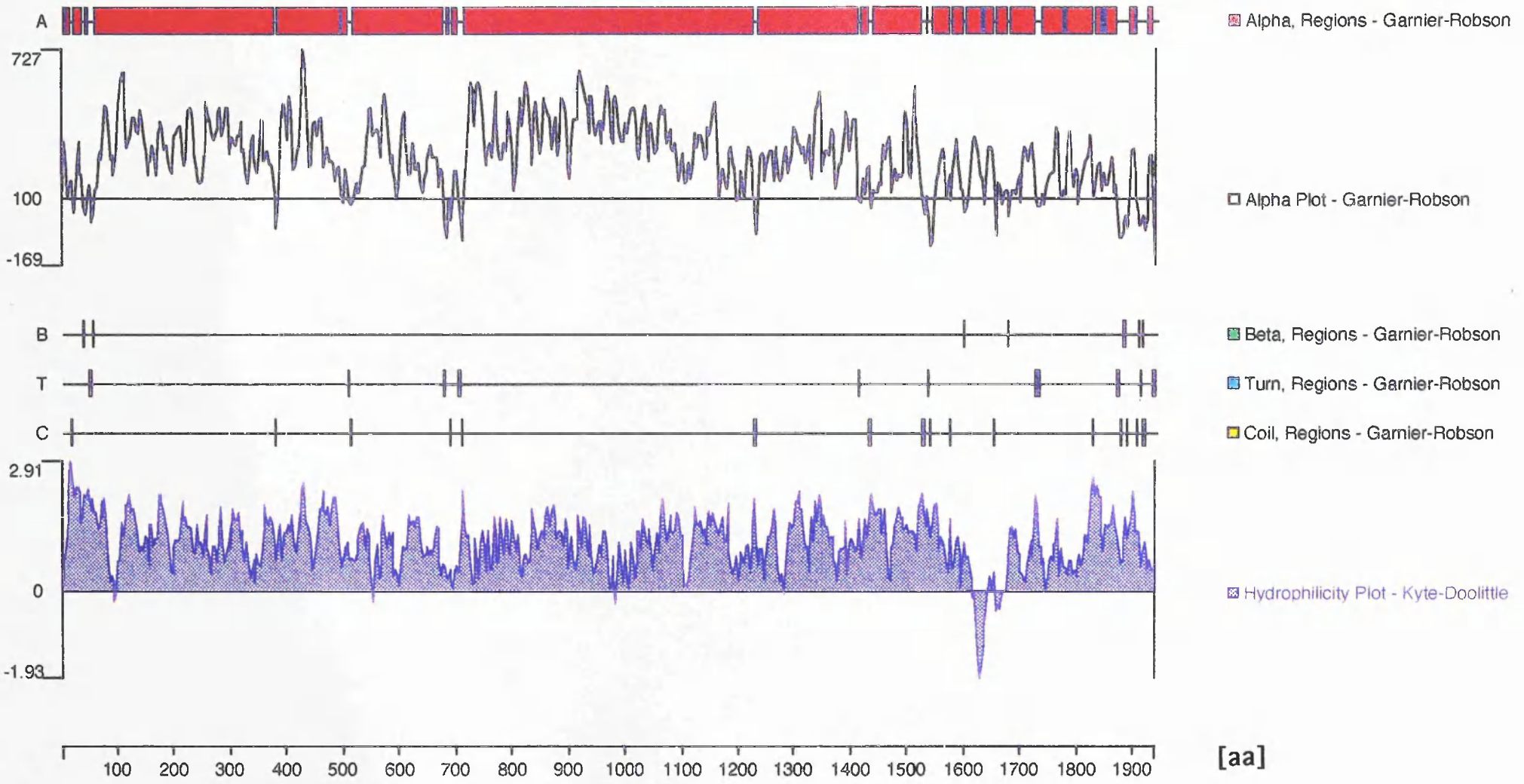

A
727
100
-169
B
T
C
2.91
0
-1.93
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
[aa]
Alpha, Regions - Garnier-Robson
Alpha Plot - Garnier-Robson
Beta, Regions - Garnier-Robson
Turn, Regions - Garnier-Robson
Coil, Regions - Garnier-Robson
Hydrophilicity Plot - Kyte-Doolittle

Figure 4.9 :

Prediction of the potential of the 229 kDa protein of *P. chabaudi* 96V to form α-helical coiled-coil superhelices.

At the top the calculated probability of the 229 kDa protein to form α-helical coiled-coil superhelices is shown (1.0 means certainty). At the bottom the 229 kDa protein is depicted as an empty bar where the shading represents the repeat region.

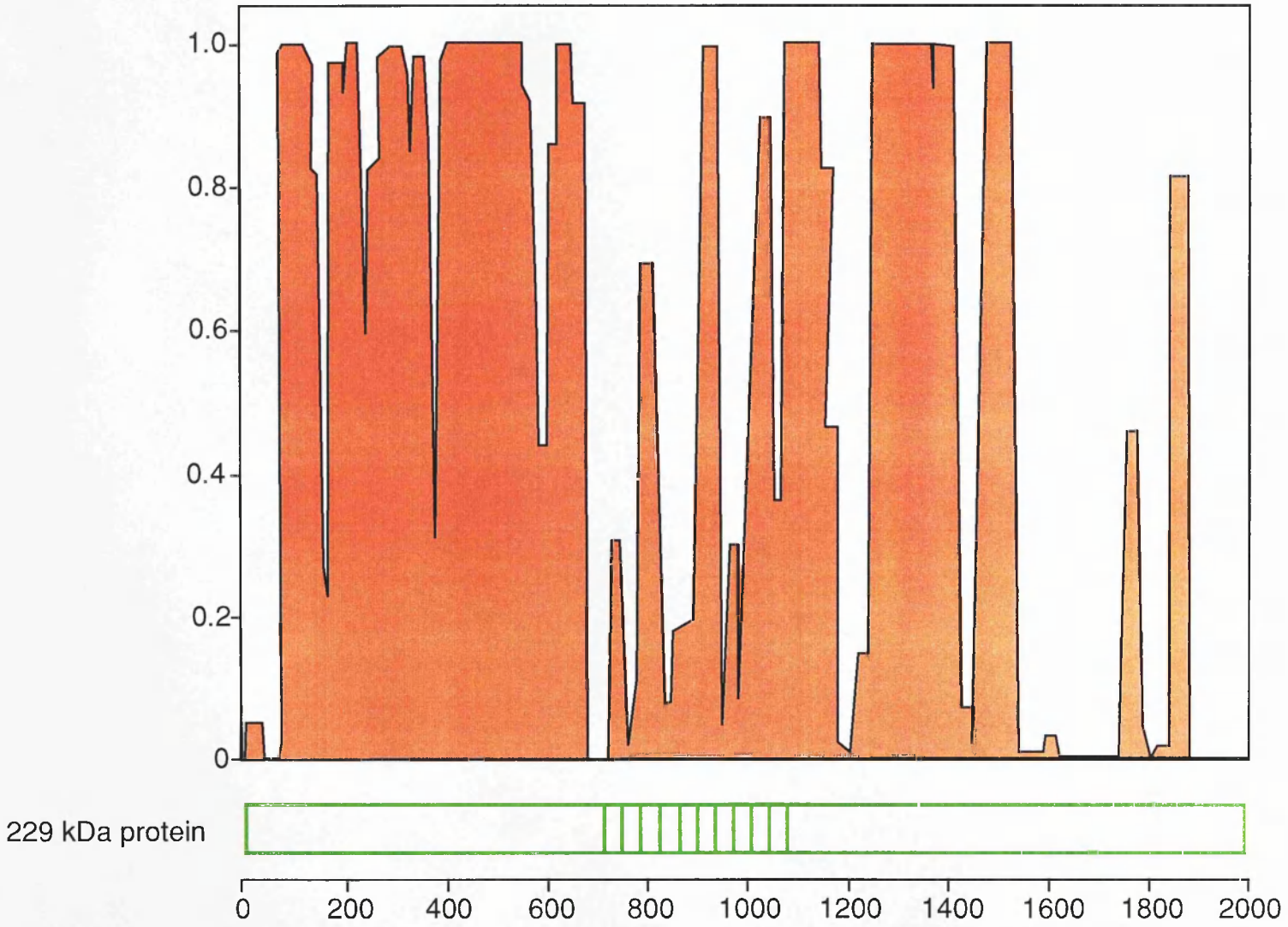
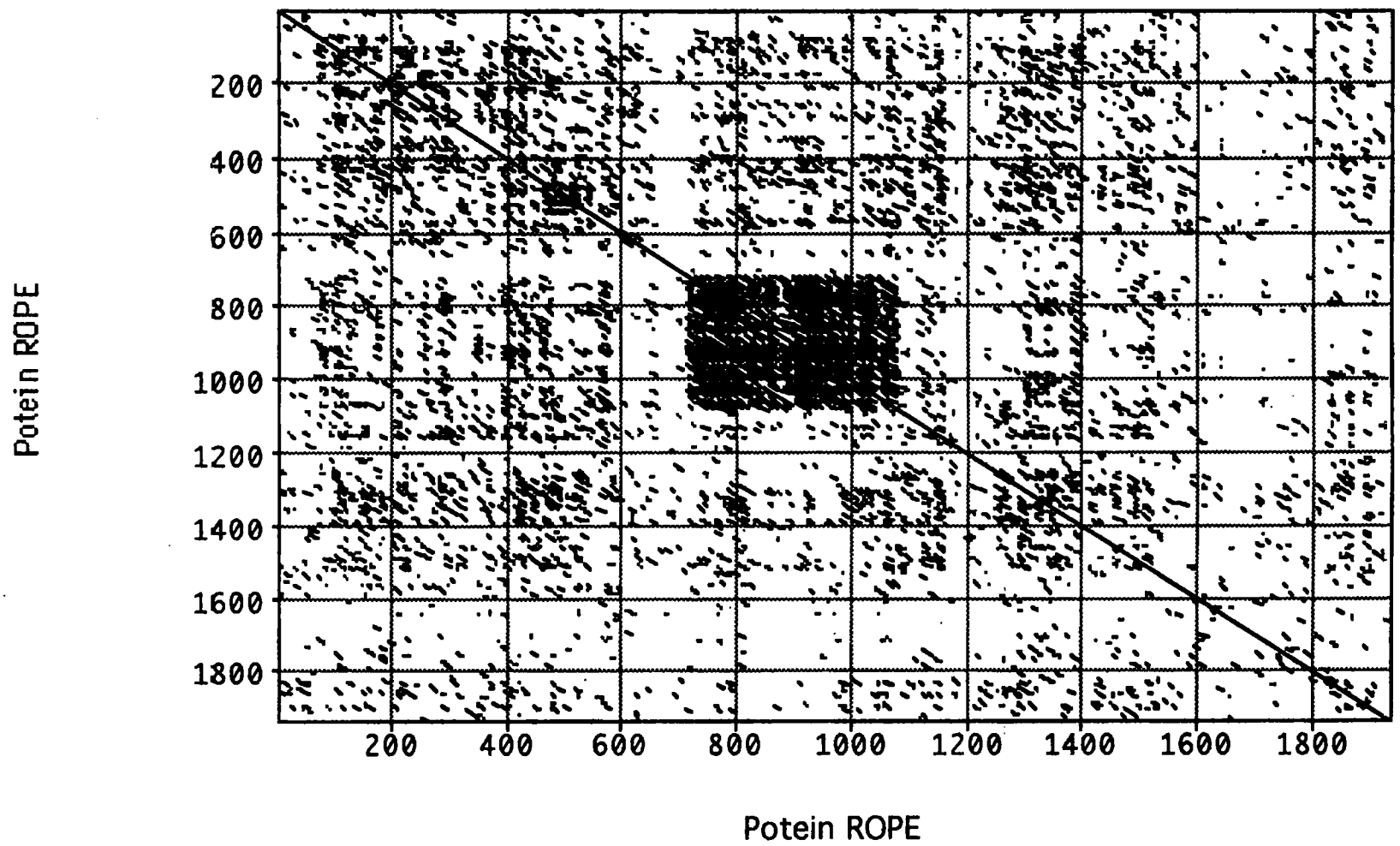

1.0
0.8
0.6
0.4
0.2
0
229 kDa protein
0
200
400
600
800
1000
1200
1400
1600
1800
2000

Figure 4.10 :

Comparison of the amino acid sequence of the 229 kDa protein of *P. chabaudi* 96V with itself in a dot plot using the Pustell protein homology matrix.

The unbroken diagonal from the upper left corner to the lower right corner indicates that the sequences are identical. Short diagonals that parallel the identity diagonal represent repeats.


Potein ROPE
200
400
600
800
1000
1200
1400
1600
1800
Potein ROPE

4.4. Modelling of the repeat region

4.4.1. Modelling of the leucine-histidine-zipper, a novel coiled-coil structure

Aligning the repeats showed the 11-mer (hendecamer) motif occurring in three blocks each separated by a truncated repeat of six residues (Figure 4.11). Of the 34 repeats, the leucine (at position **a**) is almost always present, being substituted once by isoleucine (I), twice by valine (V) and by proline (P) in the first repeat. This initial substitution may be functionally important as a 'helix-terminating' residue. The histidine position (**e**) is equally conserved with only two substitutions to tyrosine (Y) and one to asparagine (N). Both these substitutions preserve the capacity to form hydrogen-bonds while tyrosine also preserves the aromatic character of the position. Long-chain polar amino acids or amino acids with a small or no side-chain predominate in the intervening positions (**b**, **c** and **d**). Position **f** is strongly biased to lysine (K) while position **i** is exclusively hydrophobic (including threonine, T) with a strong bias towards amino acids that have side-chains branching at their β-carbon. Position **h** exhibits a clear hydrophobic trend but also admits long-chain polar amino acids. In the remaining positions (**g**, **j** and **k**) amino acids with long charged side-chains (principally, E and K) or alanine predominate. Except for the repeat length and the histidines, the amino-acid composition is typical of an α-helix in a coiled-coil structure. The leucines lie along one face of the helix with a super-twist of almost +23° per repeat based on the parameters for an ideal helix (Barlow and Thornton, 1988), however, packing forces may be sufficient to reduce or even to eliminate this twist (Pauling and Corey, 1951) (Figure 4.12). If two such helices pack as a leucine-zipper, the resulting coiled-coil would have a right-handed twist, rather than the left-handed twist observed in the standard, heptad based, coiled-coil. The absolute value of the super-twist predicted for the hendecamer (+2.1°/residue) is slightly less than for the heptad repeat (-2.6°/residue). On the helical wheel, the other two hydrophobic positions straddle the

leucine, with the more strongly conserved position (**i**) farthest (in angular measure) from the leucine. The conserved histidine position lies central in the arc between the leucine and position **i**, creating a very amphipathic helix with a conserved hydrophobic/histidine face opposing a more variable and charged face (Figure 4.12).

The conservation of residues in a periodic structure implies that each position in each repeat occupies the same structural environment. In addition, positions that exhibit conserved hydrophobicity must be shielded from solvent and highly conserved positions must interact with residues having a similar degree of conservation. These simple general constraints require that the leucines should be buried and that the histidines lie in a conserved environment, interacting with the leucines or each other.

This can be accomplished by packing two parallel helices with their leucines in contact then orienting the histidines towards each other (into the packing interface) where a hydrogen bond can form between them. A similar packing solution can also be attained with antiparallel helices where the main difference is that the histidines form a bond along the interface rather than across it. The strict steric constraints of the formation of a hydrogen-bond between histidines, combined with the limited torsional flexibility of their side-chain and the imposed symmetry of the interaction exclude any further packing solutions involving two chains.

An unsatisfactory aspect of both dimeric models is that they do not explain the strong conservation of β-branched hydrophobic side-chains at position **i** which remains relatively exposed and could easily accommodate other amino acids. From structural studies on coiled-coils from the heptad repeat, additional hydrophobic residues (especially with β-branched side-chains) are associated with higher multimer formation (Harbury *et al.*, 1993).

The simplest multimeric solutions are a triple and a four-fold helix (similar to those

found in spectrin (Yan *et al.*, 1993) and the ROP protein (Banner *et al.*, 1987), respectively). While the tetramer might simply be a dimer of the two-chain models discussed above, a triple helix would require a more radical packing solution.

Two antiparallel dimers can be arranged to shield the hydrophobic position **i** whereas with two parallel dimers, a pair of **i** positions is always exposed. Interestingly, the dimer of antiparallel chains can bring four histidines into close proximity and, with some rearrangement, these can contribute the distal nitrogens of their imidazole group to form a tetrahedral shell suitable for chelating a metallic cation such as zinc. The affinity of histidine for Zn^{++} makes this model attractive as it explains its appearance and high conservation in the repeat. However, the observed substitutions by tyrosine would not conserve this metal chelating property.

A unique role for histidine can be postulated without involving a cation. If the four chains are staggered then each histidine can form hydrogen-bonds with its two neighbours, producing an extended hydrogen-bond chain through the centre of the interaction (a histidine-zipper). In such a structure, all histidines must remain uncharged so as to be able to both accept and donate a hydrogen-bond similar to the self association of imidazole in apolar solvent (Grimmet, 1979). To maintain symmetry, the nitrogens of the histidine imidazole groups should lie in a helical arrangement. When constrained by the fixed hendecamer repeat length and the covalent/hydrogen-bonded spacing of the nitrogens, this zipper-helix is completely determined (except for chirality). The only remaining degree of freedom in the structure is the assignment of the nitrogens as proximal or distal and whether the histidines are contributed in a right- or left-handed cyclic order (Figure 4.13).

Four parallel α-helices can satisfy the hydrogen-bond zipper constraints, allowing all histidines to be connected with equal χ_1 and χ_2 torsion angles. As the resulting alignment of the proximal and distal nitrogens is almost coincident with the axis of the super-helix, the hydrogen-bonding of the histidines is virtually unaffected by the direction of the α-helices, allowing any combination of parallel/antiparallel chains to be

formed. The observed substitution of histidine by asparagine (N) is acceptable in such a model as the bifunctional side-chain of asparagine can propagate the zipper. The other substitution to tyrosine (Y), however, while sterically acceptable, would break the hydrogen-bonded helix, although bonding to the aromatic ring of tyrosine would be possible (McDonald and Thornton, 1994).

An equivalent model can also be constructed from three chains in which the central hydrogen-bond zipper is almost coincident with the axis of the α-helix bundle (Figure 4.13).

In both three and four chain models, the leucine pack towards the centre of the interaction and along with the hydrophobic position i (which lies slightly more towards the surface) shield the histidines from solvent.

4.4.2. Final model for the leucine-histidine-zipper region

The leucine-histidine-zipper (LH-zipper) region of the 229 kDa protein contains two truncated repeats of six residues. Such interruptions (or 'stutters') are common in conventional coiled-coil protein sequences (those based on the heptad sequence motif) where they can be interpreted as minor phase shifts in the hydrophobic packing face or local breaks in the α-helical structure (North *et al.*, 1994). Strangely, for the phase reversal expected across a hexa peptide, the stutters in the 229 kDa protein sequence are perfect fragments of the hendecamer repeating unit, instead of the more varied sequence normally associated with interruptions in secondary structure.

To maintain the conservation of the leucine (L) and histidine (H) residues through the stutters requires that they occupy a structural environment that is equivalent to those found in the full hendecamer repeating unit. The simplest strategy to attain this is to keep the whole sequence in the α-helix conformation and try and create the same environment with additional chains (creating an extended bundle of staggered chains

similar to that found in keratin or myosin fibres). The truncated repeat,however, introduces an reorientation of the histidines by almost 120° which leads to irreconcilable conflicts, whether the LH-zipper is constructed from three or four chains. Attempts to reset the orientation of the histidines by a segment of non-helical chain encountered similar problems as the resulting shorter gap between histidines could not be incorporated into a zipper.

Further progress came from a more detailed analysis of the sequence. Each of the segments bounded by the stutters is clearly homologous (over that expected from their repetitive nature). Referring to these as units A, B and C, each is well conserved towards their N-terminus and unit A is overall more like unit C. Within each unit there is further repetitive substructure equating histidine position 2-3-4 with 5-6-7 and (less obviously) relating 6-7 with 8-9 and possibly 10-11 (Figure 4.14).
Concentrating initially on the clearer similarity in the N-terminal part of the three units, construction of an α-helix hairpin would bring the conserved region spanning positions 1-2 together with the conserved region around position 7 and 8. This dictates the creation of a bend in the less conserved region around position 5 where both the leucine and histidine are found to vary. Extending this logic suggests a further bend between position 8 and 9 where again the conservation of both leucine and histidine is weak. The introduction of two bends implies that each unit forms a self-contained domain of three helices thus avoiding the steric conflicts that arise in any attempt to construct extended networks. This solution also allows the domains to be linked without interruption in the connecting α-helix secondary structure and the 118.4° phase shift across the stutters averts any steric conflict between neighbouring units.

In the prediction of a globular protein structure or even a conventional coiled-coil structure, little more could be added to the specificity of this model. However, the rigid structure of the LH-zipper imposes an exact registration on the helices, constraining

side-chain interactions and so allowing alternate models to be assessed.

Following a bend between helices, the next histidine can form a hydrogen-bond either on the N-terminal or C-terminal side of the previous histidine. As the third helix is forced to bridge the gap in the chain of hydrogen-bonds, there are consequently only these two possibilities (that will be referred to as the phase-1 and phase-2 respectively). In addition, viewed down the bundle axis, the helices can connect in a clockwise (right-handed super-helix) or anticlockwise (left-handed) direction giving four possible phase/hand combinations (summarised as C1, C2, A1 and A2).

Construction of the four possible structures (using the interactive computer graphic modelling program QUANTA/CHARMm), allowed the side-chain packing in each model to be assessed. These would be expected to differ because of the asymmetries of amino-acids and the α-helix itself, however, it was found that the packing in C2 equalled A1 and C1 equalled A2, giving only two distinct sets of interactions (Figure 4.15). Of these the C1/A2 set could be eliminated for a number of reasons the most compelling of which were the frequent juxtaposition of negative charges in a confined location in the absence of any countercharge and conserved leucines in a hydrophilic environment.

The sequence was then modelled on both the C2 and the A1 templates such that the less conserved 4-5- and 8-9 connections formed the turns (Figure 4.15). Maintaining the zipper and reasonable overlap of the helices, two further possibilities arise from the displacement of the third helix by a complete hendecad (designated a/b in Figure 4.15). Ideally, the less conserved histidines positions 5 and 9 would 'cap' the zipper, however, in these models (variation a) this places position 12 within the zipper and as this position is not only unconserved but also can be deleted (creating a gap in the zipper and leaving two histidines without hydrogen-bonding partners in the core of the structure), the alternative of bringing position 9 into the zipper was therefore considered preferable.

Of the two remaining possibilities, the C2b has an exposed N-terminus which would fail to account for the conservation of the initial leucine in units B and C, leaving the A1b model as the remaining possibility.

The full LH-zipper region was then modelled from the favoured A1b template. Units B and C were connected with no steric or electrostatic conflicts; however, the loss of repeat 12 in unit A resulted in the burial of negative charges in the first repeat of unit B. While some of the charges could be oriented to find counter-ions, the situation did not look favourable for the most deeply buried glutamate. This problem was avoided by shunting the register of the third helix in unit A by one repeat, maintaining the position of the terminus and shortening the preceding hairpin bend. Although this exposed the N-terminus (which was previously considered unfavourable), in the first unit, the terminus is not conserved.

With this shift, the relationship between all loops and the histidine zipper is preserved along with the junctions between units.

The LH-zipper of the 229 kDa protein is flanked by long conventional (heptad-based) coiled-coil regions which are similar to sequences with a double stranded coiled-coil structure (especially myosin). The heptad repeats begin directly after the LH-zipper region while on the N-terminal side there is a predicted uncoiled region of 40 residues (Figure 4.9.). This asymmetry, combined with the different length of the N and C terminal heptad rich regions, suggests that the 229 kDa molecule is a dimer of parallel chains. Placing the C-terminal end of the two LH-zipper models at a separation and orientation suitable to continue directly into a double-stranded coiled-coil, allowed the N-termini to be faced at an equivalent separation and created a dimer interface principally along the long (unbent) α-helix that runs between units (Figure 4.16). Any twist or kink in this helix would break the symmetry, placing the N-termini further apart.

4.4.3. Discussion

Structural proteins in animal tissue often form extended fibres consisting of long polypeptide chains in the α-helical conformation. These helices can pack in double or triple super-helices (coiled-coils) held together principally by hydrophobic interactions. The fixed period of the α-helix, combined with the twist of the super-helix is reflected in a repeated pattern of seven residues (the heptad-repeat) (Cohen and Parry, 1994). A specialised form of the repeat contains exclusively leucine at one position (leucine zipper) and crystallographic studies have shown the structure to be a coiled-coil with the leucines packed at the interface (O'Shea *et al.*, 1991).

A polar-zipper model has been proposed for the structure of polyglutamine (Perutz *et al.*, 1994), however this structure was based on a β-sheet and bears no similarity to either the conventional leucine zipper or the current model. Conserved hydrogen-bonding residues (asparagine and tyrosine) have previously been postulated to stabilize coiled-coils in a sequence containing a heptamer and hendecamer units (La Polla *et al.*, 1991). however, the complexity of these interactions did not allow a specific structure to be predicted.

From the first part of the modelling of the repeat region the hendecamer repeat motif has been predicted to form a novel coiled-coil structure (the LH-zipper) and the most probable structure for the zipper was a complex of three or four chains interacting through hydrogen-bonded histidines in the centre surrounded by packed hydrophobic residues.

Neither the three chain nor the four chain structure could be maintained beyond the truncated repeats (stutters) which bounded each segment. A more detailed analysis of the sequence considered each segment as a self-contained domain of three helices. Thus the steric conflicts that arose in any attempt to construct extended networks were avoided and a final model for the LH-zipper was proposed.

Despite being based on a different coiled-coil motif, the final model for the LH-zipper region of the 229 kDa protein exhibits many similarities to spectrin (Yan *et al.*, 1993). Both are formed from self-contained triple-helix domains of closely similar length connected in a left-handed topology. In both structures, repeated copies are linked in tandem by a continuous α-helix which in spectrin has also been modelled with a 60° phase reversal in the packing face (Yan *et al.*, 1993). In the proposed structure for the 229 kDa protein, these features were derived from structural imperatives dictated by the sequence and not as a 'homology model' based on a tentative sequence similarity.

Recently an α-helical coiled-coil structure based on a hendecamer repeat was proposed for a protein that occurs in the embryo of cotton seeds (Dure, 1993). Modelling resulted in a right-handed coiled-coil superhelix as well, but the superhelix was formed by two parallel α-helices and did not contain a zipper motif. A proposed function of this structure is not to give tensile strength to the cell, but to act with its high charge density of the solvent face as an ion carrier that prevents the crystallization of ions inside the cells of the seed tissue during a period of natural desiccation (Dure, 1993).

Proteins containing repeated sequences are widely distributed in nature. In some cases a function could be related to the repeated parts of a protein, e.g. Ca-binding in the case of the calcium-binding acidic-repeat protein precursor (ARP) of *Euglena Gracilis* (Gumpel and Smith, 1992) or building of a coiled-coil structure in collagen (Eyre, 1980). In general only a minority of proteins contain repeated sequences.

This situation is quite different for malarial proteins, where many of the characterized antigens contain repeated sequences (Kemp *et al.*, 1986), which consist, with the exception of the glycophorin binding protein (Kochan *et al.*, 1986) and FIRA (Stahl *et al.*, 1987), of rather short repeat units not exceeding 17 aa.
It is generally believed that these repeat structures are mainly maintained by the parasite

in order to evade a lethal response of the host's immune system (Anders, 1986; Schofield, 1991). This concept is based on observations like the high immunogenicity of these repeats (Kemp *et al.*, 1986), the crossreactivity that occurs sometimes between repeats of different proteins (Anders, 1986) and the extreme diversity in sequence, size and number that can occur in the same protein in different strains (Anders and Smythe, 1989).

Little attention has been drawn to the possible structure of these repeats, with some exceptions, e.g.SPAM (secreted polymorphic antigen associated with merozoites) whose heptad repeats are believed to form α-helical coiled-coil structures (McColl et al., 1994; Mulhern *et al.*, 1995).

The final model for the LH-zipper of the 229 kDa protein led us to some conclusions concerning this particular repeat region:

1) The ability of a residue in a given position to maintain a certain structure or motif is more important than its precise reiteration.

2) The two truncated repeats have a defined function of linking participating helices in spectrin-like domains. A different truncation or positioning of them would lead to a destruction of the whole superstructure.

3) Any variation in the length of the 11-mer repeat would disturb the LH-zipper as would any change in the number of repeats per unit.

4) The LH-zipper requires uncharged histidine residues, histidine has a pI of 6.5 and the LH-zipper should be stable at around pH 7. The structure would on the other hand disappear if the environment became more acidic (pH< 6).

5) The similar structure of the LH-zipper and spectrin suggests a defined biological function of the LH-zipper.

The high degree of organization that occurs in the LH-zipper implies structural constraints on this part of the 229 kDa molecule which limits its possible diversity. One

form of polymorphism of the 229 kDa protein that might occur in other plasmodial species or strains is the number of spectrin-like repeats that are present in this protein, linked by truncated repeats.

The LH-zipper part of the 229 kDa molecule shows a high charge density of the solvent face resulting in a high hydrophilicity and is expected to be water soluble. Therefore the LH-zipper is expected to be highly immunogenic. This immunogenicity might even be enhanced by the additional protection that the coiled-coil structure could give to the LH-zipper against proteolytic degradation. Therefore a certain immunodominance of the LH-zipper compared to the rest of the molecule would reflect the properties of an underlying structure but not its function as an immunological decoy.

A first step in order to describe a possible biological function of the 229 kDa protein and the LH-zipper would be to locate this protein inside the infected erythrocyte, this will be discussed in the next chapters.

Figure 4.11 : **Alignment of the hendecamer repeating motif.**

The positions of the amino acids in the hendecamer repeat are named a-k. The truncated 6-mer repeats interrupt the hendecamer repeat region twice creating three parts of tandemly repeated hendecamer repeats of comparable length, named units A,B and C.

a	b	c	d	e	f	g	h	I	j	k
P	D	T	T	H	K	E	M	V	A	E
L	E	K	R	H	A	A	I	V	A	E
L	E	E	K	H	K	E	E	I	A	K
L	G	E	G	H	K	E	V	V	L	R
L	G	E	Q	H	K	E	E	T	I	I
L	E	E	K	H	K	D	V	V	T	K
L	G	E	Q	H	K	E	N	I	I	K
L	E	E	E	H	K	D	V	V	T	K
L	G	D	Q	Y	K	E	E	I	A	K
L	K	E	E	H	A	V	V	V	A	E
L	E	E	K	H	K					
L	G	E	G	H	K	E	M	V	D	E
L	E	K	R	H	A	D	F	V	E	G
L	E	E	K	H	K	A	E	T	A	K
L	E	E	G	H	K	S	E	M	N	E
V	E	K	R	H	A	D	F	V	E	G
L	E	E	K	H	K	A	E	T	A	K
L	G	E	G	H	R	E	V	V	A	G
L	E	E	K	H	K	E	V	V	A	E
L	E	E	K	H	K	E	E	I	A	K
L	E	E	G	H	K	E	V	M	A	E
L	G	E	K	H	K	E	V	V	A	G
L	E	A	K	H	N					
L	E	E	G	H	K	E	M	V	A	E
L	E	K	R	H	A	D	L	V	A	V
L	E	E	Q	H	K	A	E	I	I	K
L	G	E	E	H	K	E	V	V	A	G
I	E	E	K	Y	K	V	E	A	I	K
L	A	E	E	H	K	D	V	V	T	K
L	G	E	Q	H	K	E	E	I	A	K
L	E	D	G	H	K	E	V	V	N	E
V	E	K	K	N	A	S	L	L	N	M
L	E	E	N	H	K	N	E	M	I	K
L	K	E	E	H	K	E	S	A	S	D

Figure 4.12 : **Helical wheel representation of the hendecamer repeating motif.**

The key features of the motif are written clockwise around points corresponding to the α-carbon positions of an α-helix with a period of 1.75 radians/residue. Two instances of the conserved leucine (**L** at position **a**) are shown along with the conserved histidine (**H** at position **e**). The black dot indicates the conserved hydrophobic position **i** and the grey dot the preceding semi-hydrophobic position **h**. The histidine lies central in this hydrophobic face, implying a buried environment.

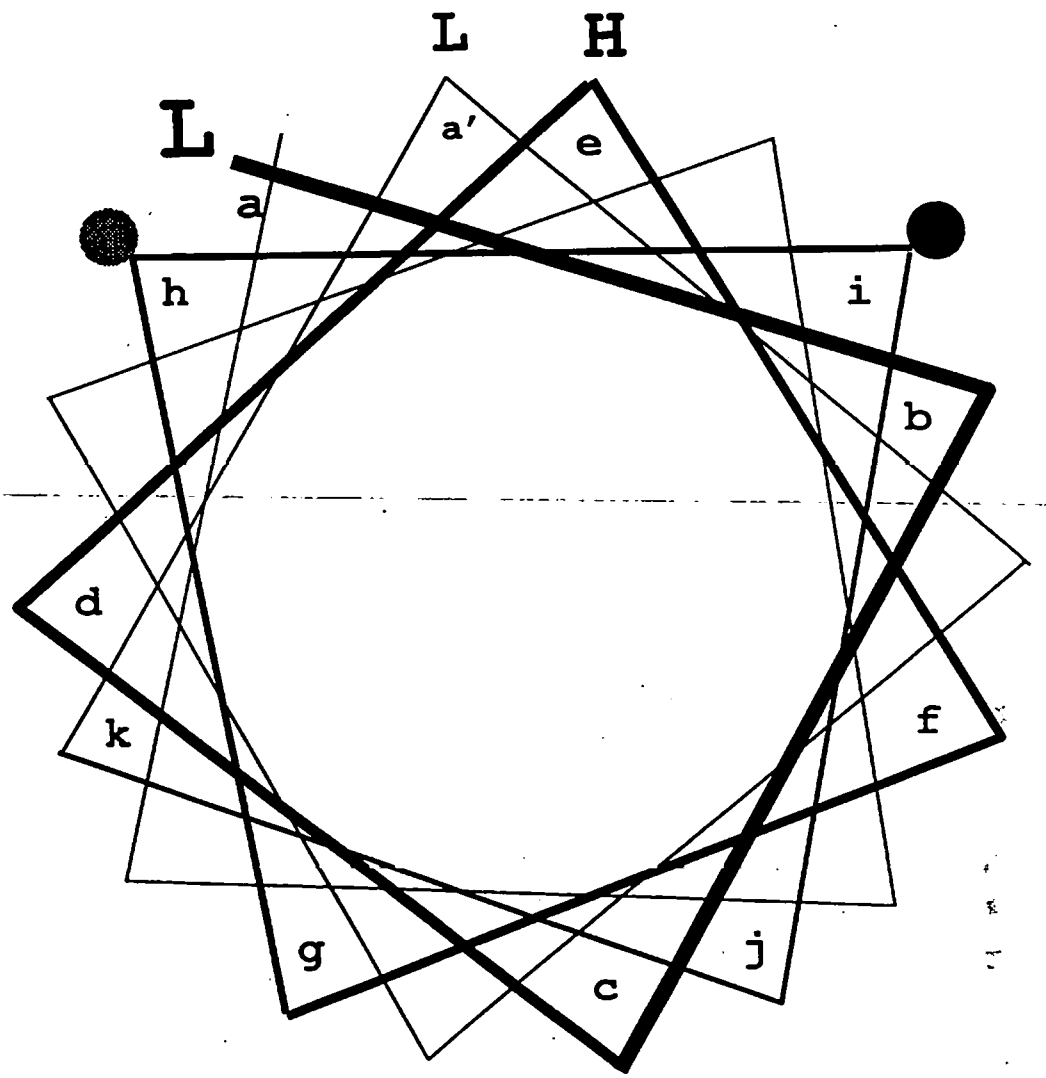

L
H
L
a'
e
a
h
i
b
d
k
f
g
c
j

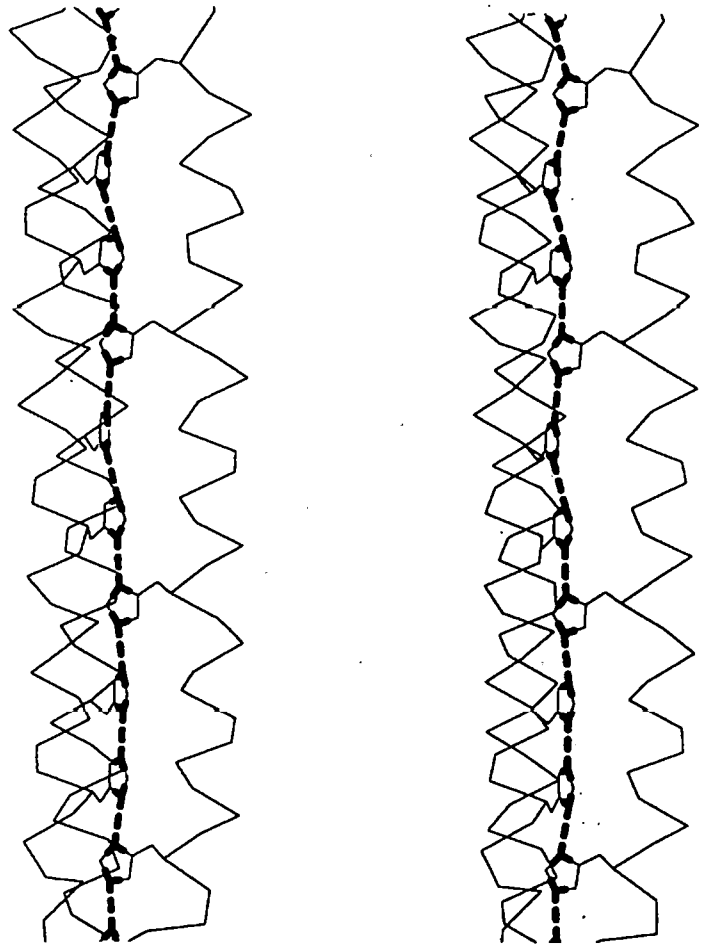
Figure 4.13 : **Molecular models of multimeric LH-zippers shown as stereopairs.**

a) A trimer composed of two parallel and one antiparallel helix.

b) A tetramer composed of two antiparallel pairs of helices.

All the non-hydrogen atoms are shown for the histidine side-chains but the remainder of the structure is represented only as a trace linking the α-carbon atoms of each residue. The histidine nitrogens are drawn bold with a dashed line connecting them to emphasize the central hydrogen-bonded zipper. The trimer had a histidine nitrogen-nitrogen separation of 2.8 Å which corresponds to the shorter end of the distribution found in well refined protein crystal structures. The tetramer can be modelled with almost any nitrogen separation depending on the spacing of the helices. 2.8 Å was again taken, giving a helix separation corresponding to that seen in the tetrameric ROP protein. The side-chain torsion angle χ_1 was slightly more than -160° in both models while χ_2 was 150° in the trimer and -66° in the tetramer (all of these angles fall within allowed regions).

a)



b)

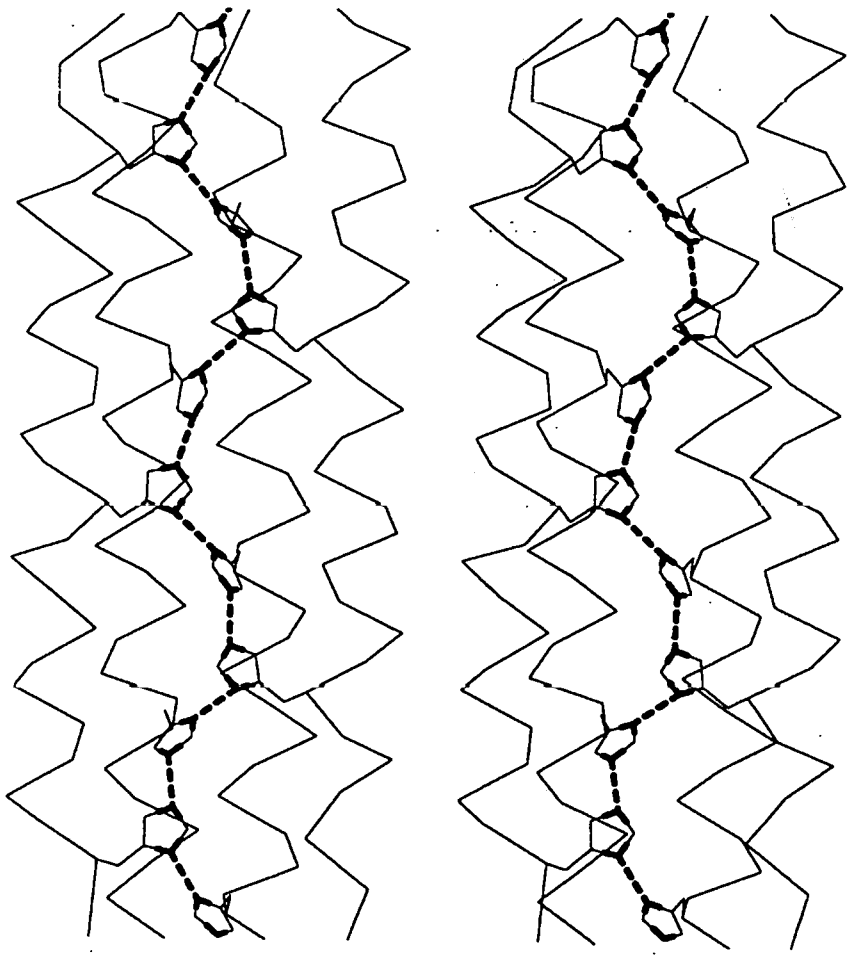


Figure 4.14 : **Internal repeats in the LH-zipper region.**

The sequence exhibits a clear triplication (units A,B and C) which are aligned following the two truncated repeats. Within these units each hendecamer motif is numbered 1-12 and the correspondence of the conserved leucine (L) and histidine (H) positions is marked by a star. As unit A is overall more similar to unit C, only the sequence similarities between A and B and B and C are indicated as identity (|) and similarity (:). Because of the inherent similarity of the hendecades, the C-terminal duplications are ambiguous and 5-6 = 7-8 = 9-10 is also plausible. The N and C terminal segments in lower-case suggest regions that may not participate in the zipper structures.

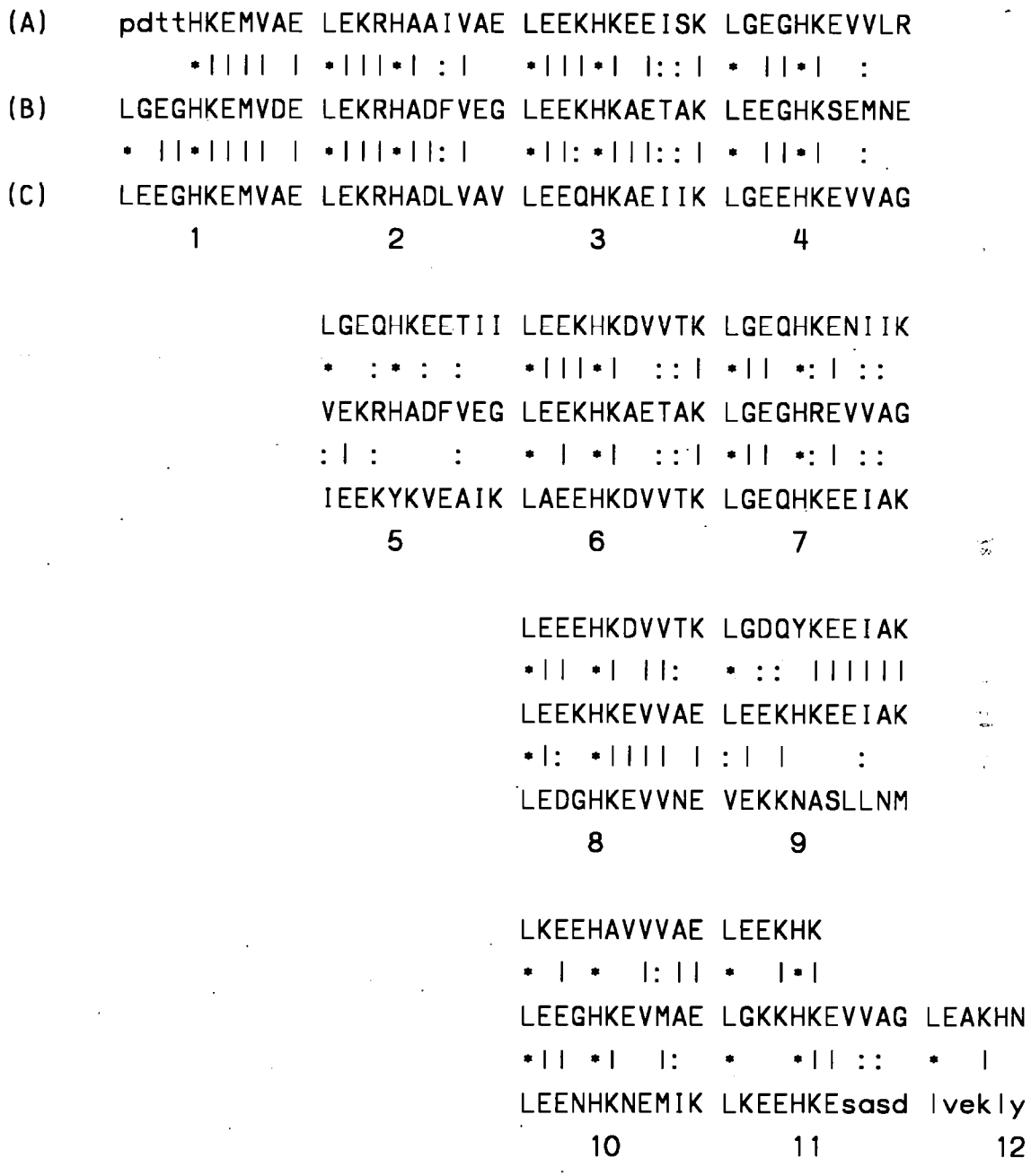
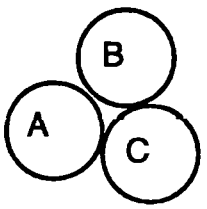

(A) pdttHKEMVAE LEKRHAAIVAE LEEKHKEEISK LGEGHKEVVLR
(B) LGEGHKEMVDE LEKRHADFVEG LEEKHKAETAK LEEGHKSEMNE
(C) LEEGHKEMVAE LEKRHADLVAV LEEQHKAEIIK LGEEHKEVVAG
1 2 3 4
LGEQHKEETII LEEKHKDVVTK LGEQHKENIIK
VEKRHADFVEG LEEKHKAETAK LGEGHREVVAG
IEEKYKVEAIK LAEEHKDVVTK LGEQHKEEIAK
5 6 7
LEEEHKDVVTK LGDQYKEEIAK
LEEKHKEVVAE LEEKHKEEIAK
LEDGHKEVVNE VEKKNASLLNM
8 9
LKEEHAVVVAE LEEKHK
LEEGHKEVMAE LGKKHKEVVAG LEAKHN
LEENHKNEMIK LKEEHKEsasd lvekly
10 11 12

Figure 4.15 : **Zipper interactions in clockwise and anticlockwise models.**

Histidine positions (1-12) in each zipper are located on a helical net in which diagonally adjacent positions make side-chain hydrogen-bonds (e.g.1-8). The net is cyclic (wraps around) so, for example, in the clockwise model histidine 1 is also bonded to 9. A desired aspect of both models is that they should form a continuous hydrogen-bonded chain. In the clockwise model the register of the first hairpin is defined to be phase-2 (histidine 5 bonds to the C-terminal side of 4) while in the anticlockwise model the register is phase-1. Although the phase of the third helix is forced (to be the opposite phase), two displacements by one repeat unit (a/b) are possible while still maintaining the intact zipper. The preferred register, based on arguments discussed in the text, is indicated by square brackets.
A, B and C written in circles represent the three helices, viewed down the bundle axis, which form the zipper.

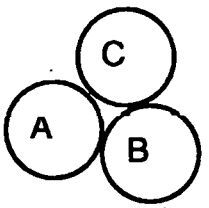
CLOCKWISE (a)/(b)




```
+     +---(9)
|     |     |
(1)   +     +
|     |     |
+     (8)   +
|     |     |
+     +   (10)/(9)
|     |     |
(2)   +     +
|     |     |
+     (7)   +
|     |     |
+     +   (11)/(10)
|     |     |
(3)   +     +
|     |     |
+     (6)   +
|     |     |
+     +   (12)/(11)
|     |     :
(4)   +     +
|     |     :
+     (5)   +
|----|      :
+     +     +  /(12)
```

phase-2

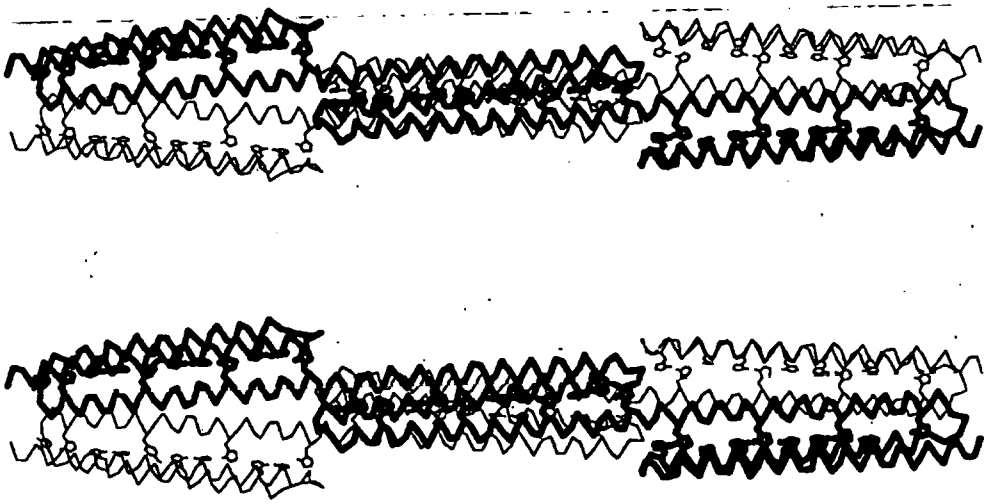
ANTICLOCKWISE (a)/(b)




```
+     +-----+
      |     |
+     +     (9)
|     |     |
+     [8]   +
|     |     |
[1]   +     +
|     |     |
+     +   (10)/[9]
|     |     |
+     [7]   +
|     |     |
[2]   +     +
|     |     |
+     +   (11)/[10]
|     |     |
+     [6]   +
|     |     |
[3]   +     +
|     |     |
+     +   (12)/[11]
|     |     :
+     [5]   +
|     |     :
[4]   +     +
 |----|     :  /[12]
```

phase-1

Figure 4.16 : **Full model of the LH-zipper region** (units A,B and C) is depicted as a stereo pair with two molecules forming a parallel dimer (one chain bold the other feint). The trace connects sequential α-carbons in the polypeptide and no side-groups are shown except those that contribute to the central zipper. The long helices connecting the three units can be seen clearly with the reversal in packing face induced by the truncated hendecamer repeat between units. The termini (carboxy on top) are positioned to continue as a conventional (heptad based) double-stranded (dimeric) coiled-coil as predicted by the flanking regions.



4.5. Expression of recombinant proteins and their reactivity with antibodies induced by parasite infection

4.5.1. Expression of rec 267 and rec 700

Different parts of the coding region of the 229 kDa protein were subcloned into pGEX vectors of the T-series which allow the expression of a fusion protein composed of the 26 kDa glutathione-S-transferase of *Schistosoma japonicum* (Sj GST) (Smith *et al.*,1986) and the polypeptide encoded by the insert, with the latter fused to the C-terminus of the Sj GST. There is a thrombin cleavage site between the Sj GST and the polypeptide encoded by the insert.
After expression in *E. coli* the fusion protein can be purified by binding to glutathione-agarose beads followed by elution with free glutathione, or the polypeptide encoded by the insert can be obtained free of Sj GST after cleavage with thrombin (Smith and Johnson, 1988).

The insert of clone c70 was cloned into pGEX-1T, the small *Eco* RI fragment of clone c6b into pGEX-3T and the big *Eco* RI fragment into pGEX-2T (Figures 4.1 and 4.17 a) as described in (3.10.2.). The recombinant polypeptides encoded by the cloned inserts were denominated rec 700, rec 267 and rec 1037 respectively.

The purified Sj GST-rec 267 protein always migrated as a 36/39 kDa double band (Figure 4.17 b), lane 1). After elution with glutathione the fusion protein was cleaved with thrombin into the 26 kDa Sj GST and two polypeptides of 12 and 15 kDa encoded by the insert (Figure 4.17 b), lane 2). When the Sj GST-rec 267 fusion protein was directly cleaved on the beads the main products were the 12/15 kDa polypeptides together with some Sj GST (Figure 4.17 b), lane 3). The M_r ratio between Sj GST and the 12/15 kDa polypeptides is about 2:1 which was not reflected in the Coomassie blue

staining intensity, since the 12/15 kDa polypeptides stained much weaker (Figure 4.17 b) lane 2). Figure 4.17 c) shows the Sj GST-rec 700 fusion protein (lane 1), the 26 kDa Sj GST that remained bound to the beads after thrombin cleavage (lane 2) and the purified rec 700 polypeptide after thrombin cleavage which migrates at 31 kDa (lane 3).

The Sj GST rec 1037 fusion protein was only expressed in trace amounts (< 0.01 mg/l; data not shown).
None of the fusion proteins was detectable in significant amounts as insoluble inclusion bodies in the sonicated bacterial pellet (data not shown).

4.5.2. Immunological reaction of rec 267 and rec 700 with mouse hyper-immune sera

The thrombin cleaved Sj GST -rec 267 fusion protein and the purified rec 700 polypeptide were separated by polyacrylamide-gel electrophoresis, electrotransferred to nitrocellulose and probed in western-blots with 5 different mouse hyper-immune sera.

The sera from all mice that were repeatedly infected with *P. chabaudi* 96 V (3.1.7.1.) clearly reacted with the rec 700 polypeptide (Figure 4.18 a), although the serum from mouse-1 (lane 1) showed a stronger reaction than the other 4 sera (lanes 2-5). None of the pre-immune sera reacted with rec 700 (lanes 6-10).
The hyper-immune sera also reacted with the rec 267 polypeptide, present here as a 12/15 kDa double band (Figure 4.18 b). The sera from mice-1 and -5 (lanes 1 and 5) showed only a very weak reaction (with the 12/15 kDa band), mice-2 and -4 (lanes 2 and 4) a weak reaction (with the 12/15 kDa band) and mouse-3 (lane 3) a strong reaction (mainly with the 15 kDa band). None of the pre-immune sera reacted with rec 267 (lanes 6-10).

4.5.3. Discussion

Three different parts from the coding region of the 229 kDa gene were subcloned into pGEX vectors of the T-series, which allow high-level expression of genes or gene fragments (Smith and Johnson, 1988).

The three constructs covered different parts of the 229 kDa protein: rec 1037 (aa position 746-1089) comprising about 90 % of the repeat region (713-1076); and rec 267 (aa position 1090-1178) and rec 700 (aa position 1284-1516), regions downstream from the repeat.

The yield of the fusion proteins increased from < 0.01 mg l^{-1} (Sj GST-rec 1037) to 0.4 mg l^{-1} (Sj GST-rec 700) and 8 mg l^{-1} (Sj GST-rec 267). Only Sj GST-rec 267 gave a yield comparable to the ones (1.5-15 mg l^{-1}) obtained by Smith and Johnson (1988). No inclusion bodies that contained any of these fusion proteins in the sonicated bacterial pellet were detectable in a significant amount, indicating that the low yield of a fusion protein was due here to low synthesis and not to low solubility. It seems noteworthy to mention that the last point is in contrast to the observations made by Smith and Johnson (1988) where low yields were generally due to low solubility. The expressed parts of the 229 kDa protein, all of which have the predicted potential to build α-helical coiled-coil superhelices, are highly water soluble.

The rec 700 polypeptide migrated at 31 kDa which is in good agreement with the calculated molecular mass of 30.7 kDa. Expression of the *Eco* RI fragment (rec 267) resulted in a 12/15 kDa double band compared to a calculated molecular mass of 11.7 kDa. This difference between Mr determined by SDS-PAGE and the calculated molecular mass has been commonly observed for malarial proteins and was attributed to their high content of charged amino acids and their resulting hydrophilic character (Anders *et al.*, 1988; Borre *et al.*, 1991; Deleersnijder *et al.*, 1992). It seems possible

that rec 267 was originally expressed as a 11.7 kDa polypeptide, which migrated at 15 kDa and that some of the molecules were cleaved by bacterial proteases resulting in a second 12 kDa band. This event may also have been responsible for the occurrence of the 36/39 kDa double band of the Sj GST-rec 267 fusion protein as well. The amount of recombinant rec 267 and rec 700 that can be produced in the pGEX expression system in *E.coli* was enough to raise sera against them in order to further characterize the native 229 kDa protein.

The immunological reaction of rec 267 and rec 700 with mouse hyper-immune sera in western-blots indicated that the native 229 kDa protein of *P.chabaudi* 96V is exposed to the immune system of the mouse during infection with the parasite, because all hyper-immune sera detected the rec 700 and rec 267 polypeptides. Interestingly the hyper-immune serum from mouse-1 gave the strongest reaction of all hyper-immune sera with rec 700 but hardly reacted with rec 267, showing that the recognition of epitopes (from the 229 kDa protein) might be very different in individual animals.
In the case of rec 267 only serum from mouse-3 reacted strongly with the 15 kDa polypeptide and weakly (like the other mice) with the 12 kDa polypeptide. Assuming that the latter is a cleavage product of the first one, the strong difference in reactivity might be due to a single epitope located on this 3 kDa (approximately 25 aa) stretch.

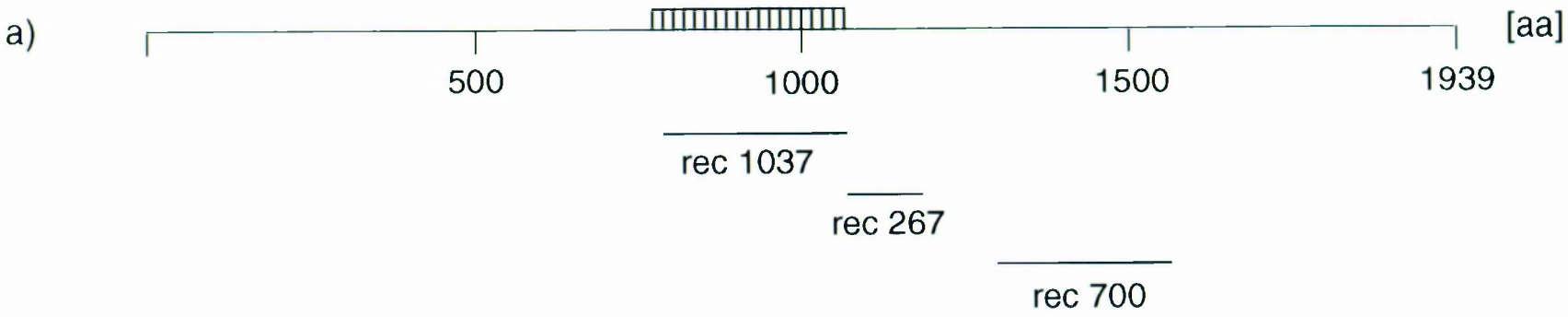
Figure 4.17:

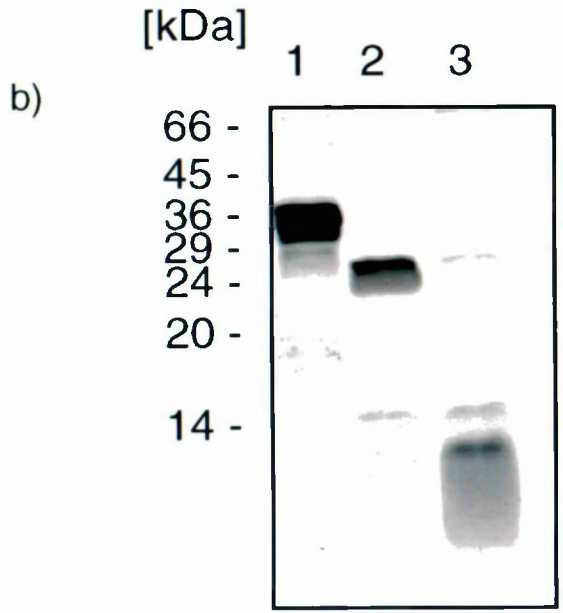
SDS-PAGE analysis with Coomassie Blue staining of recombinant polypeptides corresponding to different parts of the 229 kDa protein from *P.chabaudi* 96V.

a) Position of the sequence of the recombinant polypeptides rec 1037, rec 267 and rec 700 relative to the open reading frame of the 229 kDa protein of *P. chabaudi* 96V. |||||||||||||||| - repeat region, ———— - position of the recombinant polypeptides.

b) SDS-PAGE separation of: Sj GST-rec 267 fusion protein (lane 1), Sj GST-rec 267 cleaved with thrombin after elution (lane 2), rec 267 after thrombin cleavage of Sj GST-rec 267 bound to beads (lane 3).

c) SDS-PAGE separation of: Sj GST-rec 700 fusion protein (lane 1), Sj GST (lane 2), rec 700 (lane 3).


a)
500
1000
1500
1939
[aa]
rec 1037
rec 267
rec 700


b)
[kDa]
1 2 3
66 -
45 -
36 -
29 -
24 -
20 -
14 -

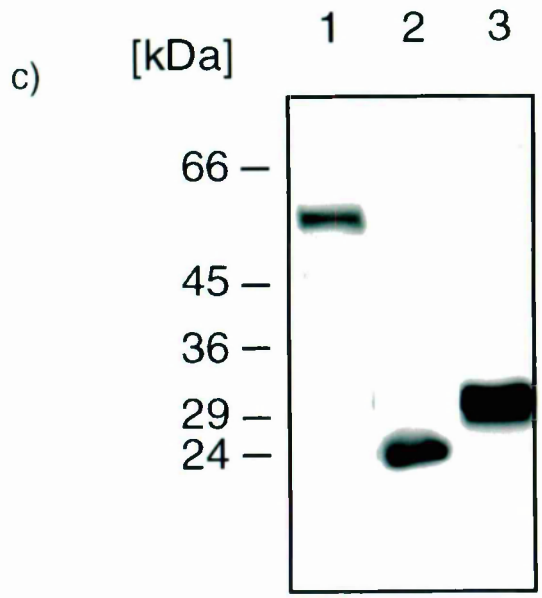
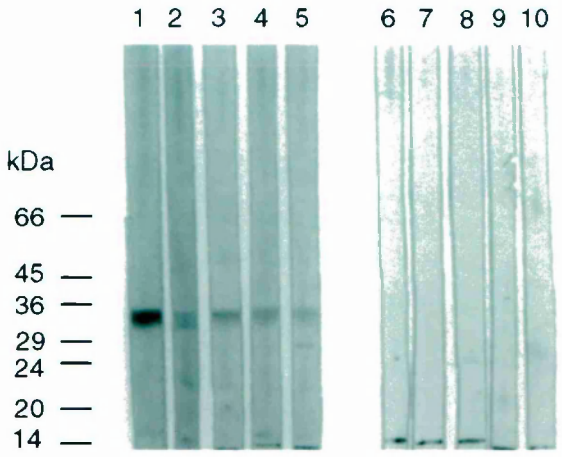

c)
[kDa]
1 2 3
66 –
45 –
36 –
29 –
24 –

Figure 4.18: **Separation of rec 700 and thrombin cleaved Sj GST-rec 267 by polyacrylamide gel-electrophoresis followed by transfer to nitrocellulose and incubation with hyper-immune sera from 5 different Balb/c mice.**

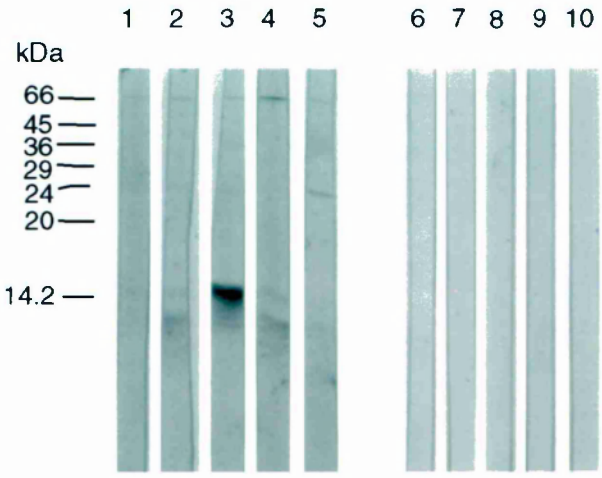
a) Reactivity of hyper-immune sera (lanes 1-5) and of pre-immune sera (lanes 6-10) from 5 different Balb/c mice with rec 700. All sera were used at 1:50 dilution.

b) Reactivity of hyper-immune sera (lanes 1-5) and of pre-immune sera (lanes 6-10) from the same Balb/c mice as in a) with thrombin cleaved Sj GST-rec 267. All sera were used at 1:50 dilution.

a)


1 2 3 4 5
6 7 8 9 10
kDa
66
45
36
29
24
20
14

b)


1 2 3 4 5
6 7 8 9 10
kDa
66
45
36
29
24
20
14.2

4.6. Reactivity of antisera raised to the recombinant proteins rec 261 and rec 700

In order to further characterize the 229 kDa protein of *P.chabaudi* 96V, antisera against rec 267 and rec 700 were produced in Fischer rats, Balb/c mice (inbred) and Parkes mice (outbred).

The different antisera were tested in Western-blots with the same preparation of recombinant protein which had been used to raise the antisera.

4.6.1. Reactivity of rec 267 and rec 700 antisera with rec 267 and rec 700

Rec 267 antisera were produced against the Sj GST-rec 267 fusion protein, consequently all sera reacted with Sj GST and with rec 267 (12/15 kDa double band) (Figure 4.19 a). Some sera detected uncleaved Sj GST-rec 267 (lanes 3-8) which was still present in small amounts in the preparation. One mouse antiserum (Parkes, lane 3) and one rat antiserum (Fischer, lane 7) reacted more strongly with rec 267 than the rest of the sera did,the mouse antiserum reacted more strongly with the 15 kDa and the rat antiserum with the 12 kDa polypeptide. Only two antisera reacted weakly with rec 267 (Parkes, lanes 1 and 2). No pre-immune sera detected Sj GST or rec 267 (lanes 10-12).

Sera against rec 700 were produced against purified rec 700 polypeptide. Figure 4.19 b) shows that all sera detected bands that correspond to Sj GST-rec 700 fusion protein and Sj GST indicating that minor amounts of these components were present in the preparation of rec 700 which was used to immunize the animals. The rec 700 polypeptide was recognized by all antisera, with sera from Balb/c (lanes 4 and 5) and Fischer rats (lanes 6 and 7) being a little more reactive than sera from Parkes mice (lanes 1-3), within the small group of 8 animals. No pre-immune serum detected Sj GST or rec 700 (lanes 8-10).

4.6.2. Detection of the native 229 kDa protein on western-blots

As it was shown that all sera that were raised against rec 267 and rec 700 recognized rec 267 or rec 700 strongly on western-blots, they were used to detect the native protein in extracts of *P.chabaudi* 96 V. For this purpose parasites were isolated as described in (3.6.1.) with the exception that the parasite pellet was resuspended in 1 volume PBS/ aprotinin (4 mg ml^{-1})/ leupeptin (4 mg ml^{-1})/ pepstatin A (2 mg ml^{-1})/ TLCK (100 mg ml^{-1})/ PMSF (4 mM)/ EDTA (2 mM), to reduce proteolysis and either snap-frozen and kept at -80°C or used directly.

The common feature of all sera was that they detected a double band of M_r225 000/ 240 000, with a higher reactivity against the 225 kDa polypeptide in most but not all sera (Figure 4.20, lanes 1-6). Most sera reacted as well with a variety of additional bands but in a less homogeneous way. Interestingly a similar 225/ 240 kDa double band reacts strongly (the 240 kDa band) with mouse hyper-immune sera (lane 7).

4.6.3. Localization of the native 229 kDa protein in *P. chabaudi* 96V infected erythrocytes by indirect immunofluorescence

So far all sera raised against rec 267 and rec 700 detected a 225/ 245 kDa protein doublet in western-blots and they were further tested in an indirect immunofluorescence assay (IFA) for their ability to detect the 229 kDa protein in erythrocytes infected with *P. chabaudi* 96 V. Mice were bled when the parasitaemia was between 20-50% with about 5-20% schizonts; under these conditions younger stages were still present. Preparation of slides and incubation with sera was performed as described in (3.11.).

The predominant result which was reproducibly obtained (with 10 out of 14 sera

tested) was the detection of dots within the intracellular parasite and sometimes these dots could be resolved into a pair of dots (Figures 4.21 a and c). The dots occurred only in schizonts where pigment was already present (b) and no infected erythrocyte with less than 4 paired dots was observed. Furthermore it seemed that the paired dots were located at the periphery of the infected erythrocyte (a) and that free merozoites contained only a single pair of dots (c).

Some sera (3 out of 14 sera tested) detected reproducibly structures that were associated with the membrane of the erythrocyte, resulting in a rim-like pattern (Figures 4.21.d and e).

4.6.3. Discussion

To produce rec 267 and rec 700 antisera animals were immunized with Sj GST-rec 267 and rec 700. Because the rec 267 polypeptide is small (11.7 kDa) it was thought to be more efficient to keep the Sj GST part as a carrier, that might contain T-cell epitopes to stimulate the antibody response against the rec 267 part. All animals immunized with these recombinant polypeptides recognized either rec 267 or rec 700.

All antisera raised either against Sj GST-rec 267 or rec 700 detected a 225/240 kDa double band in western-blots with preparations of total protein from purified *P. chabaudi* 96V parasites. The fact that sera raised against different non-overlapping parts of the 229 kDa protein react with the M_r 225 000 and M_r 240 000 protein indicates that a cross-reaction between two different proteins was not being observed, as is often the case with plasmodial proteins (Anders, 1986).

The 229 kDa polypeptide is probably synthesized as a M_r 240 000 precursor which might be proteolytically cleaved into a M_r 225 000 fragment. This remains to be proven by immunoblotting and metabolical labelling (followed by pulse-chase) experiments with parasites at different stages of their erythrocytic cycle. In addition peptide mapping studies could be used to investigate the relationship between the two proteins. Using these approaches it might also be possible to elucidate the origin of the smaller peptides observed in the western-blots presented here.

Proteolytic processing around the time of schizont rupture and invasion is characteristic for certain proteins associated with the merozoite surface, for example MSP-1 (Holder and Freeman, 1984), AMA-1 (Crewther *et al.*, 1990) and SPAM (McColl *et al.*, 1994) or proteins involved in the invasion process, e.g. RAP-1 (Clark *et al.*, 1987; Ridley *et al.*, 1990).

The calculated molecular masses of the 229 kDa protein and the M_r 240 000 determined

by SDS-PAGE are in good agreement, in contrast to the frequently large differences between the two values observed for many malaria proteins as mentioned in (4.5.3.).

In immunofluorescence studies with sera raised against rec 267 and rec 700 on glutaraldehyde fixed mouse erythrocytes invaded by *P. chabaudi* most sera gave reproducibly a punctate pattern in parasites at the schizont stage where the dots were located towards the periphery of the infected erythrocyte. These dots could sometimes be resolved into a pair of dots with free merozoites containing only a single pair.
The features of the pattern observed here resemble a characteristic pattern obtained for a number of proteins which were demonstrated to be located in the rhoptry organelles of human and rodent malaria parasites, like a 235 kDa protein from *P.yoelii* (Holder and Freeman, 1981; Oka *et al.*, 1984) and RAP-1 (Schofield *et al.*, 1986; Bushell *et al.*, 1988), AMA-1 (Peterson *et al.*, 1988; Crewther *et al.*, 1990), Rhop-H3 (Holder *et al.*, 1985; Sam-Yellow *et al.*, 1988) and a Mr 240 protein (Roger *et al.*, 1988) from *P.falciparum.*
In studies on the biogenesis of rhoptry organelles in *P.falciparum* (Jaikaria *et al.*, 1993) Rhop-H3 and Rhop-L1 (RAP-1) were first visualized as dots at the 8-nucleus stage when at least 5 pairs of rhoptries were present. The observation that the 229 kDa protein can be detected for the first time in 4 paired dots might further support its localization in the rhoptry organelles. It can not be excluded that the 229 kDa protein is located in another organelle of the apical complex, for example the microneme or dense granules and not in the rhoptry organelles. Some of the sera used in our studies reacted very strongly in immuno-fluorescence and it might be possible to use them for further analysis in immuno-electronmicroscopy, in order to determine more precisely the location of the 229 kDa protein.
In addition to a location in the rhoptries (apical complex), some of the sera gave a rim-like pattern which indicates an association of the 229 kDa protein with the membrane of the infected erythrocyte. This localization was reported for the Rhop-H3 protein as well

(Sam-Yellowe et al., 1988; Perkins and Ziefer, 1994).

One of the reasons why only three of the sera detected this location might be that they recognize stronger than the other sera an epitope that is more accessible when the protein is in this position.

Two rhoptry proteins of comparable size are known to date:

(i) A 235 kDa protein of *P. yoelii* has a sequence that differs from the 229 kDa protein and does not contain an extended repeat region (Keen *et al.*, 1994).

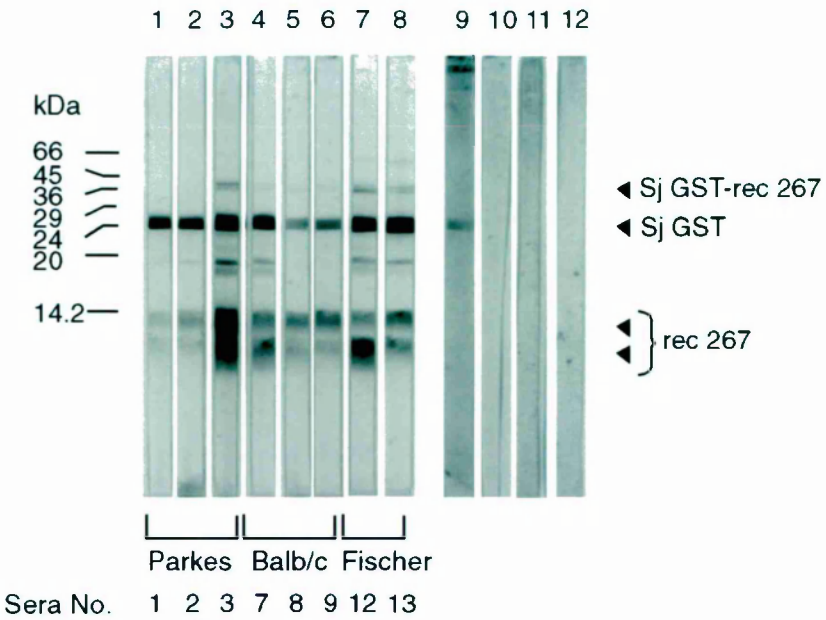
(ii) A 240 kDa protein of *P. falciparum* which is the precursor of a 225 kDa protein, (Roger *et al.*, 1988). The M_r of these proteins is identical with the 240/225 kDa doublet found in western-blots for the 229 kDa protein. There are no sequence data available for the *P. falciparum* protein and no decision can be made at present whether or not it is the homologous protein in this different species.

Figure 4.19: **Separation of rec 267 and rec 700 by polyacrylamide gel-electrophoresis, followed by transfer to nitrocellulose and incubation with sera raised against rec 267 and rec 700.**

a) The Sj GST-rec 267 protein cleaved with thrombin was electrophoresed in each lane and probed with various sera.Sera against Sj GST-rec 267 fusion protein were raised in: Parkes mice (lanes 1-3), Balb/c mice (lanes 4-6) and Fischer rats (lanes 7 and 8). Antiserum raised against Sj GST in a Balb/c mouse (lane 9). Pre-immune sera from: Parkes mouse (lane 10), Balb/c mouse (lane 11) and Fischer rat (lane 12).
All sera were used at 1:300 dilution.

b) The purified rec 700 protein was electrophoresed in each lane and probed with various sera. Sera against rec 700 were raised in: Parkes mice (lanes 1-3), Balb/c mice (lanes 4 and 5) and Fischer rats (lanes 6 and 7). Pre-immune sera from: Parkes mouse (lane 8), Balb/c mouse (lane 9) and Fischer rat (lane 10).
All sera were used at 1:1000 dilution.

a)


1 2 3 4 5 6 7 8 9 10 11 12
kDa
66
45
36
29
24
20
14.2
Sj GST-rec 267
Sj GST
rec 267
Parkes Balb/c Fischer
Sera No. 1 2 3 7 8 9 12 13

b)

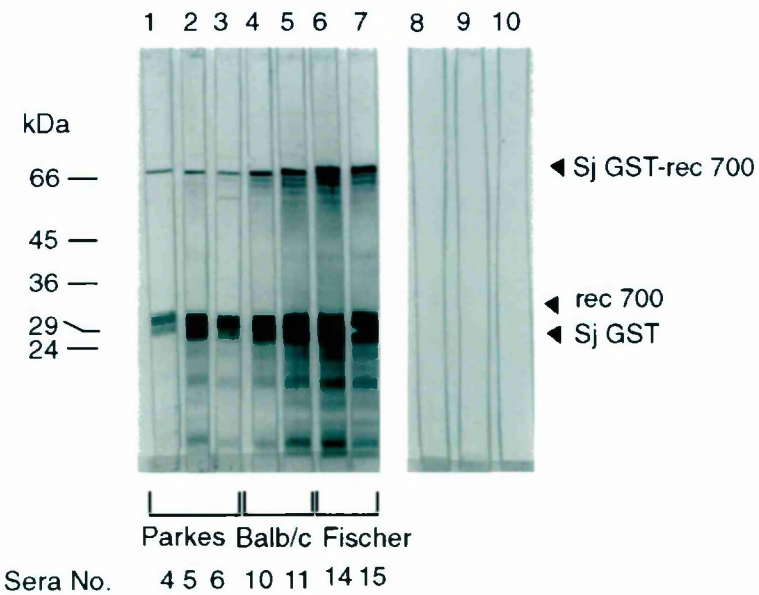

1 2 3 4 5 6 7 8 9 10
kDa
66
45
36
29
24
Sj GST-rec 700
rec 700
Sj GST
Parkes Balb/c Fischer
Sera No. 4 5 6 10 11 14 15

Figure 4.20: **Separation of total protein from *P. chabaudi* 96V infected erythrocytes by polyacrylamide gel-electrophoresis, followed by transfer to nitrocellulose and incubation with rec 267 and rec 700 antisera.**

Sera raised against rec 267: Parkes mouse (lane 1), Balb/c mouse (lane 2) and Fischer rat (lane 3). Sera raised against rec 700: Parkes mouse (lane 4), Balb/c mouse (lane 5) and Fischer rat (lane 6). Hyper-immune sera from Balb/c mouse (lane 7). Pre-immune sera: Parkes mouse (lane 8), Balb/c mouse (lane 9) and Fischer rat (lane 10) and serum raised against Sj GST in Balb/c mouse (lane 11). All sera were used at 1:100 dilution.

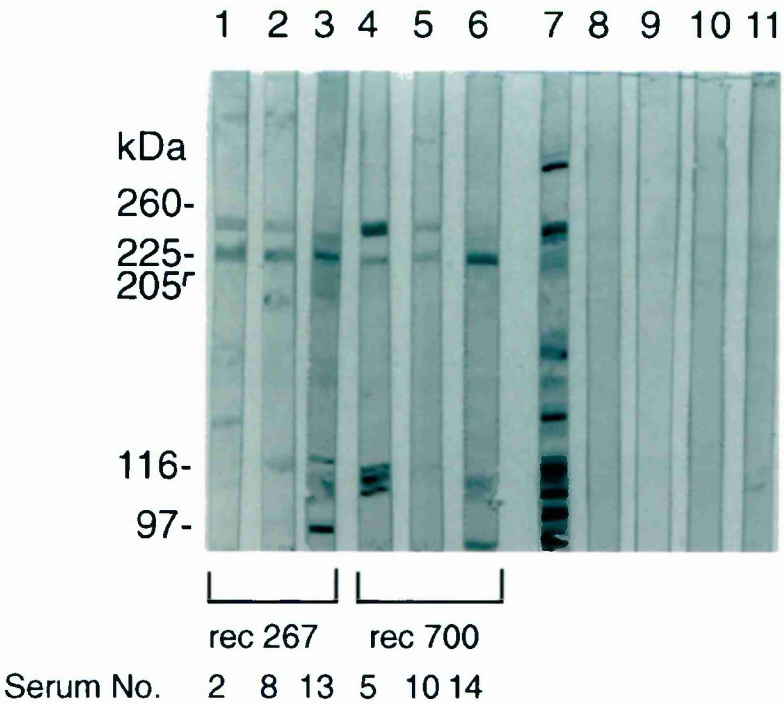

1 2 3 4 5 6 7 8 9 10 11
kDa
260-
225-
205
116-
97-
rec 267
rec 700
Serum No. 2 8 13 5 10 14

Figure 4.21: **Localization of the native 229 kDa protein of *P.chabaudi* 96V by indirect immunofluorescence.**

Serum 13:

a) dotted pattern inside of schizonts

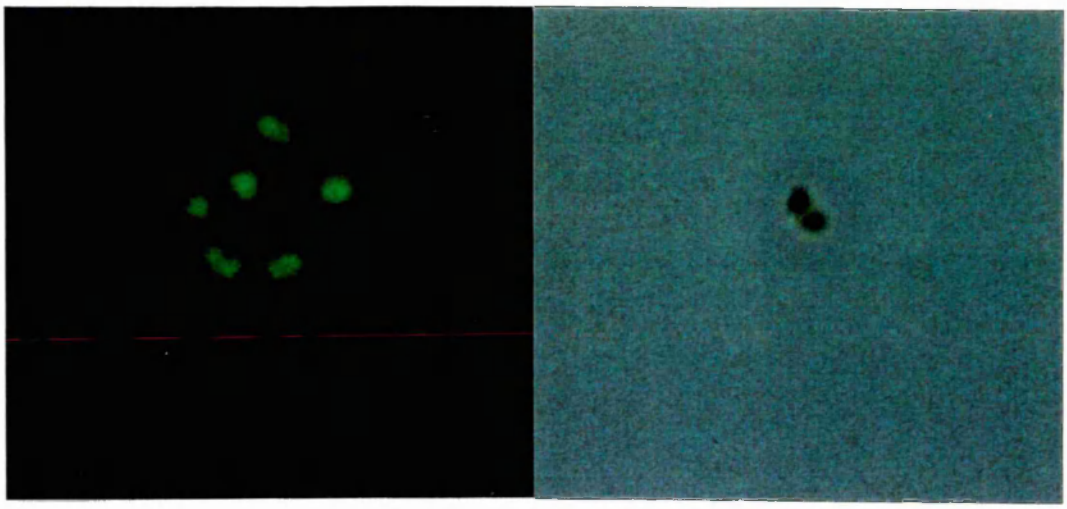
b) phase contrast image of a)

c) free merozoites each harbouring one dot that is sometimes resolved into a pair of dots.

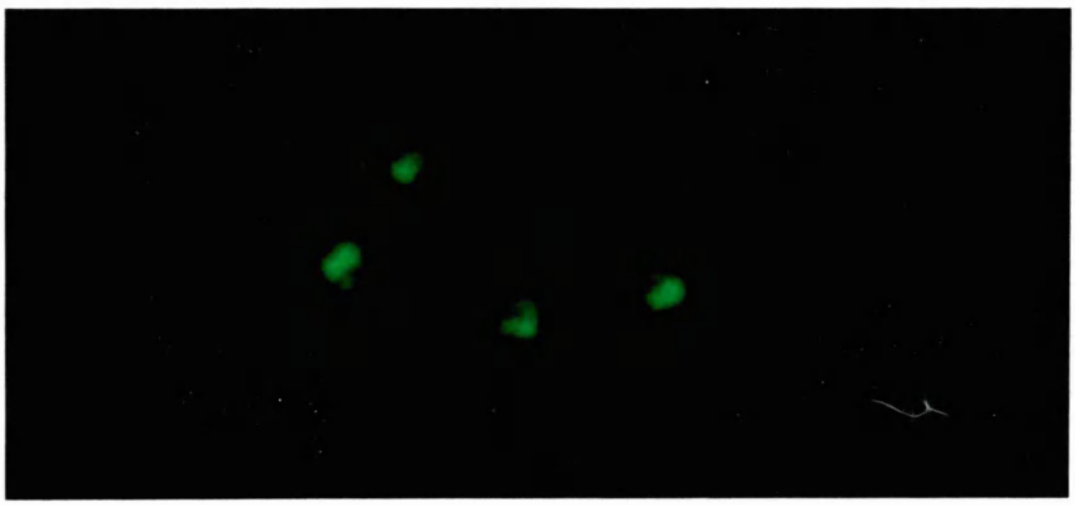
Serum 10:

d) and e) rim-like pattern associated with the membrane of the infected erythrocyte.

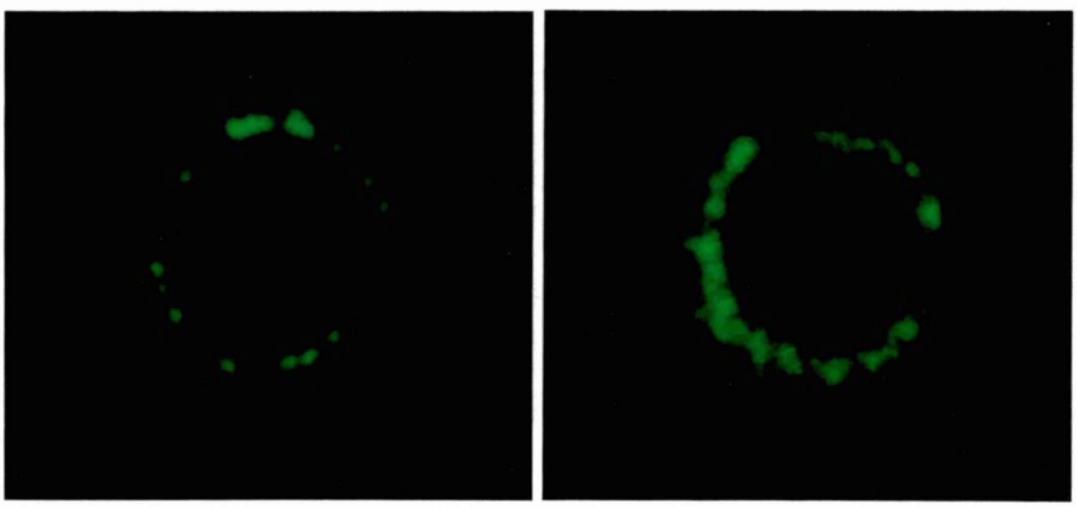
All sera were used at 1: 200 dilution.



a) b)



c)



d) e)

General discussion

5. General discussion

One of the cornerstones of the present work was the construction of cDNA and genomic DNA libraries of *P. chabaudi* 96 V. One of the critical steps in many projects is to proceed from an isolated protein to a cDNA or genomic DNA clone encoding this protein. A variety of choices is available at this stage: Sequencing of the N-terminus of the protein and reverse translation of the amino acid sequence into a nucleotide sequence. Despite the degeneracy of the genetic code sometimes oligonucleotide probes can be produced that allow isolation of clones coding for the desired protein.
Sometimes it may be more successful to sequence internal peptides of a protein which have been generated by protease treatment, this might result in less ambiguous probes and these might also be useful for PCR amplification of genomic DNA encoding the protein.
Another valuable way forward is to raise monospecific sera against the protein or to produce monoclonal antibodies against it in the hope that they specifically detect the protein in an expression library.
The latter approach was chosen in the present work. In our case a serum produced in a rat against material from *P. chabaudi* 96V infected erythrocytes that bound to a glutathione-agarose column detected quite specifically clones coding for the 229 kDa protein when a cDNA expression library of *P. chabaudi* 96 V in λZAP II was screened. It became evident that the detected protein was, with a high probability, a malarial cytoskeletal protein, which appeared interesting. No clones coding for a glutathione-S-transferase, expected to be the predominant enzyme binding specifically to glutathione agarose, were detected.
The cDNA library was screened four times and all of the sequenced clones possessed a 3'-end located inside of the open reading frame (ORF). This means that long stretches of As present in the mRNA were easily accessible to the $(dT)_{18}$-part of the linker-primer during first strand synthesis of the cDNA. This internal priming is a common

feature of malarial cDNA clones, e.g. many cDNA clones coding for malarial antigens were first published without the 5'- and 3'-end (Kemp *et al.*, 1983; Stahl *et al.*, 1985, 1987; Coppel *et al.*, 1986; Coppel, 1991). Internal priming might be avoidable by simultaneously increasing the number of dTs in a linker-primer and the annealing temperature, however, the advantage of this internal priming is that it might be possible (with some patience) to find cDNA clones that cover the entire region of a long coding sequence.

The question why no cDNA clone containing the 3'-end of the coding region was found and whether or not the mRNA coding for the protein is polyadenylated at all can't be answered here, though possible polyadenylation signals appear further downstream of the 3'-end of the ORF. Examination of the sequence directly adjacent to the 3'-end of the ORF shows a sudden appearance of long runs of Ts and it is possible that a polyA tail hybridized rapidly to these structures and was therefore no longer accessible to the linker-primer. Reports on non-polyadenylated mRNAs were based on the observation that a fraction of mRNA that failed to bind to oligo(dT) columns could be translated in vitro and the products immuno-precipitated with immune sera (Vaidya *et al.*, 1983). It seems possible that the polyA tail of these antigens had hybridized to internal Ts and thus the fact that a they failed to bind to an oligo (dT) column might not be sufficient to conclude that some plasmodial mRNAs are not polyadenylated.

The problems encountered with the cDNA library were alleviated when a genomic DNA library was used for further screening. Certain steps in the construction of the genomic library were optimized in a trial and error process resulting in a genomic library with following features:

The cloned inserts came from partially *Sau* 3A digested genomic DNA from *P. chabaudi* 96V which was 2-4 kb size fractionated and partially filled-in to make it compatible with *Xho* I digested and partially filled-in pBluescript KS(+), into which the genomic DNA was ligated. No double inserts should occur because they are not

compatible with each other after partial fill-in.

It was further observed that the choice of the *E. coli* strain used to propagate a DNA library is of primary importance. Based on observed rearrangements of pBluescript clones in certain *E. coli* strains the *E. coli* SURE (Stop Unwanted Recombinational Events) strain was found to maintain in a very stable way cloned plasmodial sequences. The average size of the cloned inserts was 3 kb and the percentage of recombinants was about 60%. Under these conditions any gene sequence can theoretically be found with a probability of P = 90% when 23 000 recombinant clones are present in the library (P = 99% for 46 000 recombinant clones). Spreading the library on a 22.2 cm x 22.2 cm Biodyne A transfer membrane gave the best growth results. Enough colonies can be grown to obtain P values between 90 - 99 % whilst still allowing a rapid purification of a positive signal. Once soaked in glycerol the library can be stored at -80°C for many years until needed again.

Theoretically inserts containing any two *Sau* 3A sites with a distance of 2-4 kb between them on a chromosome should be present in this library, thus it should be possible to walk up- and downstream from any point on a chromosome until a region is met where 2 neighbouring *Sau* 3A sites are further apart from each other than 4 kb. If a second library is made which contains longer inserts, e.g. 3-7 kb or which was partially digested with another suitable restriction enzyme followed by partial fill-in and ligation into a partial filled-in cloning site of pBluescript, probably almost any region of the genome could be sequenced by this walking strategy.

Often in (malarial) sequencing projects one or both ends of a coding sequence were not found in a library. A frequent strategy to obtain the missing part is to produce a restriction map of the adjacent region and to amplify the region outside by inverse PCR (Triglia *et al.*, 1988; Borre *et al.*, 1991; Nolte *et al.*, 1991; Knapp *et al.*, 1991). Considering the easy handling of a plasmodial genomic DNA library, due to the small size of the genome, it might be possible that these efforts are avoidable by using this

walking strategy.

The 9985 bp obtained during the sequencing procedures contained an intronless 5829 bp ORF coding for 1943 amino acids. The methionine at position 5 of the deduced amino acid sequence was counted as the start codon, resulting in a 229 kDa protein composed of 1939 aa. This decision was made arbitrarily and needs confirmation or correction by amino acid sequence data from the N-terminus of the native protein.
The up- and downstream regions adjacent to the ORF were searched for further ORFs but no long ORF (that could encode a protein of > 10 kDa) was detected. Special attention was used to identify any upstream ORF that could contain a methionine preceding a hydrophobic stretch and an exon-intron splice junction, and thus represent an N-terminal signal peptide separated by an intron from the region coding for the downstream part of the protein. No such sequence could be identified.
This search was based on the splice acceptor and donor consensus sequence for known *P. falciparum* genes (Brown and Coppel, 1991; Knapp *et al.*, 1991). It can not be excluded that there might be a variant of this consensus sequence in the 229 kDa gene or that multiple introns exist in the 5'-region of the gene as is the case for the Rhop-H3 protein (Brown and Coppel, 1991). This would hamper a precise deduction of the N-terminal sequence of the 229 kDa protein based on genomic sequence data. However likely this might be it seems necessary to confirm the 5'-end position of the ORF with sequences derived from cDNA in order to be able to exclude the existence of a potential signal sequence at the N-terminus of the 229 kDa protein.

Secondary structure predictions showed that about 90% of the 229 kDa protein could adopt an α-helical conformation and furthermore that large regions, with the exception of the first 60 aa at the N-terminus and the last 400 aa at its C-terminus, had the potential to form α-helical coiled-coil superhelices. The occurrence of globular ends and long rod-like regions of α-helical coiled-coil superhelices is typical for proteins like

intermediate filaments (Parry *et al.*, 1985; North *et* al., 1994), myosin (Dibb *et al.*, 1989) and paramyosin (Kagawa *et al.*, 1989).

Searches for similarities of the 229 kDa protein with proteins contained in the Swiss-Prot data bank showed that the highest degree of similarity (with non malarial proteins) was obtained with these proteins; the similarities were based on residues that are typical for the heptad motif of α-helical coiled-coil superhelices. These results indicated that the 229 kDa protein is most likely a cytoskeletal protein of *P. chabaudi*. Further support of this proposed role came from a detailed analysis of the repeat structure which is located in the centre of the 229 kDa protein.

On the DNA level the 1092 bp encoding the repeat region showed some features that were different from the rest of the ORF. The repeat region had a high A/T ratio typical for plasmodial coding sequences (Weber, 1987), but its G+T content was much higher (38.6%) than the one of the rest of the ORF (21.4%). Furthermore the codon preference of the repeat region was different from the codon preference summarized for plasmodial proteins (Weber, 1987), whereas the rest of the ORF showed precisely the same codon preference. Whatever the origin of the 1092 bp repeat sequence might be, it seems interesting that a non-typical codon preference is maintained in this part of the gene's coding sequence. The advantage of this conservation should be on the DNA level because the main contribution for the shift in codon preference is made by a G in the third position of a codon whereas the codon for the same amino acid in the rest of the ORF shows predominantly a different base.

Another surprising feature of the repeat region of the gene was the occurrence of another ORF over the entire repeat length on the complementary strand. Whether or not this ORF encodes a protein that is expressed during the life-cycle of *P. chabaudi* (in the mosquito or the rodent) can not be decided here. Interestingly a similar observation was made by Paloske *et al.* (1993) in the repeat region of the ORF coding for the Pf EMP 3 (*P. falciparum* erythrocyte membrane protein). Another long ORF was found on the

opposite strand and some of their results led them to the conclusion that this ORF might code for a protein that is expressed during the blood stage of *P. falciparum.*

On the protein level translation of the 1092 bp repeat region into amino acids resulted in a sequence that is composed of 32 11-mer repeats and 2 6-mer repeats which represent truncated versions of the 11-mer repeats. The truncated repeats lie between repeat 10/11 and 21/22 thus dividing the 11-mer repeat region into three parts designated units A, B and C. These hendecamers showed a strong conservation of certain positions and their presentation on a helical wheel resulted in an amphipathic helix with a conserved hydrophobic/ histidine face opposing a more variable and charged face. Based on these conservations and some general constraints concerning conserved positions in periodic structures a novel α-helical coiled-coil structure based on an 11-mer repeat containing a leucine-histidine-zipper (LH-zipper) was modelled. The high homology of each unit A,B and C led to the proposition of a final model for the LH-zipper region where each unit represents a self-contained domain.This model exhibits many similarities to spectrin (Yan *et al.*, 1993) but was not the result of a 'homology model' and rather unexpected.

The proposed model of the LH-zipper region of the 229 kDa protein suggests a function that is not very much in agreement with the current general view on the role of repeated sequences in malarial proteins. For various reasons these regions are mainly thought to act as an "immunological smoke screen" to either prevent the maturation of high-affinity antibody responses (Anders, 1986) or to cross-link Ig by the tandem array of repeated epitopes leading to a less efficient T-cell independent response (Schofield, 1991). Though each hypothesis may have validity, they were derived solely from immunological observations made on certain antigens. The modelling of the repeat region of the 229 kDa protein was led by the idea of considering it as a repeated structure and not as a repeated epitope. Thus, for example, variation of a residue in a

given position was analyzed with respect to its chemical character, e.g. charged or hydrophobic, long or short side-chain. Though variation of a residue in a given position might be variable in an immunological concept (variation in an epitope) it can be conservative from a structural point of view. The same is true for the occurrence of the truncated repeats which might be viewed as an 'accident' from an immunological perspective but have a highly specialized function from a structural point of view.

This is, as far as is known, the first time that a higher ordered structure has been proposed for a plasmodial repeat region and it is difficult to say how far this structural approach can be applied to repeat regions of other plasmodial proteins.

So far a potential to form α-helical coiled-coil superhelices was proposed for the repeats of SPAM (secreted polymorphic antigen associated with merozoites) of *P. falciparum* (McColl *et al.*, 1994; Mulhern *et al.*, 1995).

If the function of the eleven 50 aa repeats of the GBP (glycophorin binding protein) of *P. falciparum* is the binding to an erythrocyte receptor (Kochan *et al.*, 1991) then a specific 3-dimensional structure of this repeat region should exist. But there are doubts about whether this protein binds to glycophorin because only a minor fraction of GBP was found to be loosely associated with merozoites (Bianco *et al.*, 1987; Bonnefoy *et al.*, 1988).

The observed similarities of the predicted α-helical coiled-coil regions of the 229 kDa protein with the repeat regions of MESA (mature-parasite infected erythrocyte surface antigen) (Coppel, 1992) and the gene 11-1 protein (Scherf *et al.*, 1988) of *P. falciparum* might also reflect an α-helical coiled-coil superhelix conformation of these repeats. Finally the high similarity between the 11-mer repeats of the 229 kDa protein and the S-antigen (LEDPAKASQGG) (Cowman *et al.*, 1985) is mainly based on a conserved leucine (L) in position a, a glutamic acid (E) in position b and lysine (K) in position f. The view on a helical wheel showed an amphipathic helix (data not shown) and it might be possible that the repeat region of this S-antigen has a defined higher

order structure as well.

Two parts of the protein were expressed and reacted on western-blots with antibody in hyper-immune sera, indicating that the 229 kDa protein is a natural antigen. The location of rec 267 and rec 700 in the sequence of the 229 kDa protein has a predicted α-helical coiled-coil structure for these parts of the protein, with a characteristic high charge density of the solvent face giving them a high solubility in water; once in the host's blood stream they probably elicit a strong immunological response.
As mentioned before it seems possible that a certain group of proteins that are either associated with the membrane of the infected erythrocyte, e.g. MESA (Coppel *et al.*, 1986; Lustigman *et al.*, 1990) and the gene 11-1 protein (Koenen *et al.*, 1984; Scherf *et al.*, 1988) or are secreted like SPAM (McColl *et al.*, 1994) contain α-helical coiled-coil structures based on repeated sequences. These proteins necessarily encounter the host's immune system and their high immunogenicity and variable sequence might rather reflect the properties of a structure which is naturally highly immunogenic than an immunological decoy.

More details about the possible biological role of the 229 kDa protein were obtained by the use of antisera against rec 267 and rec 700 in western-blotting and immunofluorescence experiments. The sera from all animals immunized with these proteins recognized in western-blots with extracts of total protein from *P. chabaudi* a 240/225 kDa protein doublet, either both or one of them alone. Hence it is likely that the 240 kDa protein is the precursor of the 225 kDa protein rather than that they are two different proteins which cross-react. The size of some minor bands varied often with the serum that was used. Though some of the bands may represent artifacts or cross-reactions some might be products that were specifically processed inside the parasitized erythrocyte.

Detection of the 229 kDa protein by immunofluorescence with most sera gave a punctate pattern in erythrocytes infected with late stages of the parasite or in free merozoites suggesting a localization in one of the organelles of the apical complex. Taking into account that sometimes the dots were resolved into a pair and that they can be detected for the first time in 4 paired dots, it seems likely that the 229 kDa protein is a rhoptry protein (Jaikaria *et al.*, 1993). The possible localization of the 229 kDa protein in the rhoptries and the observation that it might be proteolytically cleaved would be in agreement with the frequently reported proteolytical processing of rhoptry proteins (Schofield *et al.*, 1986; Roger *et al.*, 1988; Sam-Yellowe *et al.*, 1988; Crewther *et al.*, 1990).

Some of the sera in immunofluorescence studies gave a rim-like pattern around the infected erythrocyte suggesting that one of the functions of the 229 kDa protein is to interact with the membrane of the red blood cell. This localization would be in agreement with the predicted character of the 229 kDa protein as a component of the cytoskeleton and further suggests a specific function for the spectrin-like LH-zipper in an environment that is very rich in spectrin. The observation that the 229 kDa protein can be localized at the membrane of the infected erythrocyte has also been made for other rhoptry proteins (Sam-Yellowe *et al.*, 1988; Perkins and Ziefer, 1994).

Given the probably extended, fibrous nature of the 229 kDa protein one of its functions could be to give tensile strength to the infected red blood cell and/or to maintain its (altered) shape. This function would require an interaction between the 229 kDa protein and proteins from the erythrocyte cytoskeleton. One way to establish such an interaction could be that the domains of the LH-zipper have a high affinity to spectrin binding proteins like actin, ankyrin, band 4.1 or adducin and replace spectrin (partially). This might reorganize the spectrin network, which provides the resistance to shear (Boal, 1994) and alter the physico-mechanical properties of the infected erythrocyte, an event described for erythrocytes after invasion with *Plasmodium* ssp.

(Cranston *et al.*, 1983; Taylor and Parra, 1987; Nash *et al.*, 1989).

In order to adopt a spectrin-like structure the LH-zipper would not need histidines as a zipper motif, i.e. this residue could, for this purpose, be replaced by any smaller hydrophobic residue (alanine (A) for example). The occurrence of histidine in this motif and the special nature of the histidine-zipper implies a specific function of this residue. The H-zipper part should be stable only at pH values greater than the pI of the imidazole ring (the pI of histidine is about 6.5), in a more acidic environment the imidazole-rings would become positively charged and the entire LH-zipper structure would disassemble.

It is possible that the 229 kDa protein is in an acidic environment in the rhoptries. The disassembled LH-zipper might destabilize the entire molecule and keep it in an 'inactive' form, thus not disturbing the organization of the rhoptries. When an erythrocyte is invaded and the 229 kDa protein is released into the more neutral environment of the erythrocyte's cytoplasm, under these conditions the LH-zipper would be assembled resulting in its 'active' form and induce a series of structural changes in the cytoskeleton of the erythrocyte.
These structural changes at the erythrocyte membrane might be needed only up to a certain developmental stage of the parasite. In mature stages of the parasite the cytosolic pH of the erythrocyte becomes more acidic, reaching about pH 6.5 (Tanabe,1990). Under these condition the LH-zipper might gradually disassemble, leading to a disintegration of the 229 kDa network and could represent the prelude to the parasite release process.

This scenario requires at a certain point the formation of an intact LH-zipper. Repression of the ability to form this structure might abolish the competence of the parasite for invasion/release. It might be possible to introduce molecules into the

erythrocyte that mimic parts of the histidine-zipper and therefore prevent the formation of the final spectrin-like LH-zipper structure. This might block the parasite's asexual life cycle in the blood stream. The ability of a suitable molecule to act in such a way could be easily tested in mice infected with *P. chabaudi* or in *in vitro* cultures.

The structural data discussed above are summarized in a final schematic model of the 229 kDa Protein shown in Figure 5.1:
The N-terminus of the molecule consists of a short globular domain of about 60 aa followed by a region of 600 aa of mainly coiled-coil structure which ends about 40 aa before the LH-zipper. The LH-zipper continues directly into a second coiled-coil region of about 450 aa and the molecule ends with 400 aa of mainly globular structure. The N- and C- terminal heptad rich regions are of different length suggesting that the 229 kDa molecule forms a dimer of parallel chains.Furthermore results from Southern-blot experiments suggest that, the 229 kDa protein is not member of a multigene family, thus the dimer should be formed by two 229 kDa protein molecules.
Given the structural features and the localization inside of the parasite of the 229 kDa protein we propose the name ROPE (Repeated Organellar Protein) for this yet uncharacterized malarial protein.

5.1. Perspectives

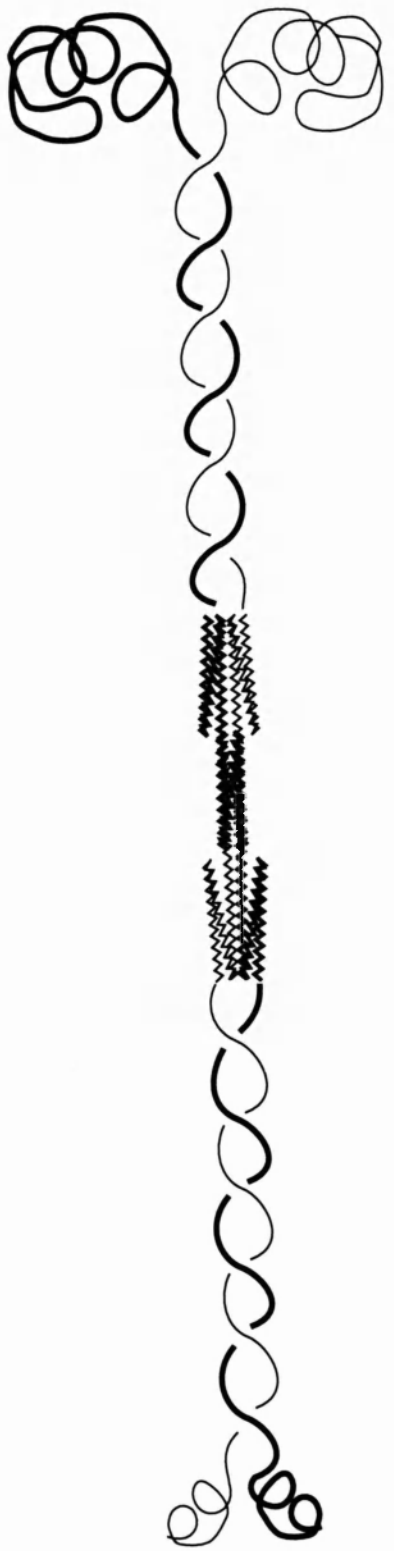
1) All plasmodial proteins, sequenced so far, that are transported out of the parasite or are located in the organelles of the apical complex contain a signal peptide in the N-terminal part of the protein (Bonnefoy and Mercereau-Puijalon, 1989; Perkins, 1992). To be compatible with this observation two solutions can be envisaged for the 229 kDa protein:

i) The part of the gene encoding a signal peptide was not identified yet. This part of the

Figure 5.1.:

Schematic presentation of the proposed higher order structure of the 229 kDa protein (ROPE).

The different regions shown in the schematic presentation are (from left to right): globular region, coiled-coil double-helices, the LH-zipper, coiled-coil double-helices and globular region. The individual regions of the schematic presentation are not drawn to scale.



gene must be separated from the rest of the coding region by at least one intron which might contain an unusual exon-intron splice junction sequence. Once a cDNA clone covering the 5'-end of the ORF is obtained and sequenced this issue should be resolved.

ii) If the ORF covers the entire coding region, then in the absence of any N-terminal signal peptide the part of the sequence that resembles most a signal peptide is the hydrophobic stretch at position 1617-1634 (Figures 4.5 and 4.8). Due to its location it would be further expected to act as an signal anchor (Zerial *et al.*, 1986; Spiess and Lodish, 1986). Though the membrane-spanning part would comprise a maximum of 18 aa, there is another rhoptry proteins which contains a putative transmembrane segment of 16 aa (Keen et al., 1994). The two residues of opposite charge would need to be buried in a transmembrane bundle composed of the hydrophobic stretch of several individual 229 kDa protein molecules. This bundle could be stabilized by interhelical salt bridges as described for charged transmembrane segments of different proteins (Engelman *et al.*, 1980; Henderson *et al.*, 1990; Preis *et al.*, 1990).
The cysteine (C) residues in the extracellular space would be in a non-reducing environment and consequently may form either intermolecular or intramolecular disulphide-bonds and might, for example, stabilize the bundle.
Even in the presence of an N-terminal signal peptide it will be of interest to test if this sequence can function as a transmembrane segment.

The potential of the hydrophobic stretch to act as a signal-anchor sequence could be investigated by *in vitro* transcription-translation followed by insertion into canine pancreatic microsomes. This could be performed with one of the clones that contain the sequence coding for the hydrophobic stretch and an appropriate "initiator" methionine because pBluescript contains T3/T7 RNA polymerase binding sites.

2)The modelling of the repeat region resulted in three self-contained spectrin-like domains comprising a leucine-histidine-zipper. In order to investigate the validity of this model a peptide representing unit B, containing repeats 11-21 and the second truncated repeat (Figure 4.11), was synthesized by chemical means (R.Ramage, Department of Chemistry, University of Edinburgh). Work is in progress to obtain crystals of this peptide in order to study the structure by X-ray diffraction analysis. This work is being done in collaboration with L. Carpenter and G. Dodson from the Department of Structural Biochemistry at the NIMR.

3) The high degree of organization in the LH-zipper implies that many constraints are acting on it and that diversity of this motif in the homologous protein of other plasmodial species and strains should be rather limited. Results from Southern-blotting showed cross-hybridization of genomic DNA from the ROPE gene of *P. chabaudi* with genomic DNA from *P. yoelii* and *P. falciparum*. Construction of a genomic library (as described for *P. chabaudi*) of these species, followed by screening with ^{32}P-labelled fragments coding for parts of the 229 kDa protein gene might allow the isolation and sequence analysis of the homologous genes.
Conservation of sequence of this protein in different species and strains would support a central role of this protein in the life of *Plasmodium* ssp.

References

Adams, J.H., Hudson, D.E., Torii, M., Ward, G.E., Wellems, T.E., Aikawa, M. and Miller, L.H. (1990). The Duffy receptor family of *Plasmodium knowlesi* is located within the micronemes of invasive malaria merozoites. Cell, 63: 141-153

Adams, J.H., Sim, B.K.L., Dolan, S.A., Fang, X., Kaslow, D.C. and Miller, L.H. (1992). A family of erythrocyte binding proteins of malaria parasites. Proc. Natl. Acad. Sci., USA, 89: 7085-7089

Aikawa, M. (1971). Parasitological reviews: the fine structure of malarial parasites. Exp. Parasitol., 30: 284-320

Aikawa, M., Miller, L.H., Johnson, J. and Rabbege, J. (1978). Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. J. Cell Biol., 77: 72-82

Aikawa, M., Miller, L.H., Rabbege, J.R. and Epstein, N. (1981). Freeze-fracture study on the erythrocyte membrane during malarial parasite invasion. J. Cell Biol., 91: 55-62

Aikawa, M., Rabbege, J.R., Udeinya, I. and Miller, L.H. (1983). Electron microscopy of knobs in *Plasmodium falciparum* - infected erythrocytes. J. Parasitol., 69: 433-437

Aikawa, M. (1988). Morphological changes in erythrocytes induced by malarial parasites. Biology of the Cell, 64: 173-181

Aikawa, M., Torii, M., Sjölander, A., Berzins, K., Perlmann, P. and Miller, L.H. (1990). Pf155/RESA antigen is localized in dense granules of *Plasmodium falciparum* merozoites. Exp. Parasitol., 71: 326-329

Aley, S.B., Bates, M.D., Tam, J.P. and Hollingdale, M.R. (1986). Synthetic peptides from the circumsporozoite proteins of *Plasmodium falciparum* and *Plasmodium knowlesi* recognize the human hepatoma cell line HepG2A16 *in vitro*. J. Exp. Med., 164: 1915-1922

Anders, R.F. (1986). Multiple cross-reactivities against antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. Parasite Immunol., 8: 529-539

Anders, R.F., Coppel, R.L., Brown, G.V. and Kemp, D.J. (1988). Antigens with repeated amino acid sequence from the asexual blood stages of *Plasmodium falciparum*. Prog. Allergy, 41: 148-172

Anders, R.F. and Smythe, J.A. (1989). Polymorphic antigens in *Plasmodiun falciparum*. Blood, 74: 1865-1875

Anders, R.F., McColl, D.J. and Coppel, R.L. (1993). Polymorphic antigens of *Plasmodium falciparum* asexual erythrocytic stages. Acta Trop., 53: 239-253

Arnot, D.E., Barnwell, J.W. and Stewart, M.J. (1988). Does biased gene conversion influence polymorphism in the circumsporozoite protein-encoding gene of *Plasmodium vivax*? Proc. Natl. Acad. Sci., USA, 85: 8102-8106

Aviv, H. and Leder, P. (1972). Purification of biologically active globine messenger RNA by chromatography on oligo-thymidylic acid-cellulose. Proc. Natl. Acad. Sci., 69: 1408-1412

Bairoch, A. (1993). The PROSITE dictionary of sites and patterns in proteins, its current status. Nucleic Acids Res., 21: 3097-3103

Banner, D.W., Kokkinidis, M. and Tsernoglou, D. (1987). Structure of the ColE1 Rop protein at a 1,7Å resolution. J. Mol. Biol., 196: 657-675

Bannister, L.H., Mitchell, G.H., Butcher, G.A. and Dennis, E.D. (1986). Lamellar membranes associated with rhoptries in erythrocytic merozoites of *Plasmodium knowlesi*: a clue to the mechanism of invasion. Parasitol., 92: 291-303

Bannister, L.H. and Mitchell, G.H. (1989). The fine structure of secretion by *Plasmodium knowlesi* merozoites during red cell invasion. J. Protozool., 36: 362-367

Barlow, D.J. and Thornton, J.M. (1988). Helix geometry in proteins. J. Mol. Biol., 201: 601-619

Bennett, V. (1985). The membrane skeleton of the human erythrocytes and its implications for more complex cells. Annu. Rev. Biochem., 54: 273-304

Bianco, A.E., Culvenor, J.G., Coppel, R.L., Crewther, P.E., McIntyre, P., Favaloro, J.M., Browm, G.V., Kemp, D.J. and Anders, R.F. (1987). Putative glycophorin-binding protein is secreted from schizonts of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 23: 91-102

Bickle, Q. and Coppel, R.L. (1992). A fourth family of the *Plasmodium falciparum* S-antigen. Mol. Biochem. Parasitol., 56: 141-150

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res., 7: 1513-1523

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: 379-382

Blackman, M.J., Ling, I.T., Nicholls, S.C. and Holder, A.A. (1991). Proteolytic processing of the Plasmodium falciparum merozoite surface protein-1 produces a membrane-bound fragment containing two epidermal growth factor-like domains. Mol. Biochem. Parasitol., 49: 29-34

Blackman, M.J., Scott-Finnigan, T.J., Shai, S. and Holder, A.A. (1994). Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein. J. Exp. Med., 180: 389-393

Boal, D.H. (1994). Computer simulation of a model network for the erythrocyte cytoskeleton. Biophysical Journal, 67: 521-529

Bonnefoy, S., Mattei, D., Dubremetz, J.-F., Guilotte, M., Jouin, H., Ozaki, L.S., Sibilli, L. and Mercereau-Puijalon, O. (1988). *Plasmodium falciparum*: molecular analysis of a putative protective antigen, the thermostable 96-kDa protein. Exp. Parasitol., 65: 69-83

Bonnefoy, S. and Mercereau-Puijalon, O. (1989). *Plasmodium falciparum*: An intervening sequence in the GBP 130/96 tR gene. Experimental Parasitology, 69: 37-43

Bork, P., Sander, C., Valencia, A. and Bukau, B. (1992). A module of the DnaJ heat shock proteins found in malaria parasites. T.I.B.S., 17: 129

Bork, P., Ouzounis, C. and Sander, C. (1994). From genome sequence to protein function. Curr. Opin. Struc. Biol., 4: 393-403

Borre, M.B., Dziegiel, M., Høgh, B., Petersen, E., Rieneck, K., Riley, E., Meis, J.F., Aikawa, M., Nakamura, K., Harada, M., Wind, A., Jakobsen, P.H., Cowland, J., Jepsen, S., Axelsen, N.H. and Vuust, J. (1991). Primary structure and localization of a conserved immunogenic *Plasmodium falciparum* glutamate rich protein (GLURP) expressed in both the preerythrocytic and erythrocytic stage of the vertebrate life cycle. Mol. Biochem. Parasitol., 49: 119-132

Borre, M.B., Owen, C.A., Keen, J.K., Sinha, K.A. and Holder, A.A. (1995). Multiple genes code for high-molecular-mass rhoptry proteins of *Plasmodium yoelii*. Mol. Biochem. Parasitol., 70: 149-155

Braun-Breton, C., Rosenberry, T.L. and Pereira da Silva, L. (1988). Induction of the proteolytic activity of a membrane protein in *Plasmodium falciparum* by phosphatidylinositol-specific phospholipase C. Nature,332: 457-459

Brown, H.J. and Coppel, R.L. (1991). Primary structure of a *Plasmodium falciparum* rhoptry antigen. Mol. Biochem. Parasitol., 49: 99-110

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: 376-379

Bushell, G.R., Ingram, L.T., Fardoulys, C.A. and Cooper, J.A. (1988). An antigenic complex in the rhoptries of *Plasmodium falciparum*. Mol. Biochem. Biol., 28: 105-112

Buthurst, I.C. (1994). Protein expression in yeast as an approach to production of recombinant malaria antigens. Am. J. Trop. Med. Hyg., 50: 20-26

Campbell, G.H., Miller, L.H., Hudson, D., Franco, E.L. and Andrysiak, P.M. (1984). Monoclonal antibody characterization of *Plasmodium falciparum* antigens. Am. J. Trop. Med. Hyg., 33: 1051-1054

Camus, D. and Hadley, T.J. (1985). A *Plasmodium falciparum* antigen that binds to host erythrocytes and merozoites. Science, 230: 553-556

Carter, R., Walliker, D. (1974). *What is Plasmodium chabaudi*? J. Protozool., 21: 449

Chishti, A.H., Andrabi, K.I., Derick, L.H., Palek, J. and Liu, S.C. (1992). Isolation of skeleton-associated knobs from human red blood cells infected with malaria parasite *Plasmodium falciparum*. Mol. Biochem. Parasitol., 52: 283-288

Chitnis, C.E. and Miller, L.H. (1994). Identification of the erythrocyte binding domains of *Plasmodium vivax* and *Plasmodium knowlesi* proteins involved in erythrocyte invasion. J. Exp. Med., 180: 497-506

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., 162: 156-159

Chulay, J.D., Lyon, J.A., Haynes, J.D., Meierovics, A.I., Atkinson, C.T. and Aikawa, M. (1987). Monoclonal antibody characterization of *Plasmodium falciparum* antigens in immune complexes formed when schizonts rupture in the presence of immune serum. J. Immunol., 139: 2768-2774

Clarke,J.T., Anand, R., Akoglu, T. and McBride, J.S. (1987). Identification and characterization of proteins associated with the rhoptry organelles of *Plasmodium falciparum* merozoites. Parasitol. Res., 73: 425-434

Clarke, L. and Carbon, J. (1976). A colony bank containing synthetic ColE1 Hybrid plasmids representative of the entire *E. coli* genome.Cell, 9: 91-99

Cohen, C. and Parry, D.A.D. (1990). α-helical coiled coils and bundles: How to design an α-helical protein. Proteins Struct. Funct. Genet., 7: 1-15

Cohen, C. and Parry, D.A.D. (1994). α-helical coiled coils: More facts and better predictions Science, 263: 488-489

Coombs, G.H. and Gutteridge, W.E. (1975). Growth in vitro and metabolism of *Plasmodium vinckei chabaudi*. J.Protozool., 22: 555-560

Cooper, J.A., Ingram, L.T., Bushell, G.R., Fardoulys, C.A., Stenzel, D., Schofield, L. and Saul, A.J. (1988). The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. Mol. Biochem. Parasitol., 29: 251-260

Cooper, J.A., Atkins, A. and Saul, A. (1989). N-terminal amino acid sequencing of the 105 kilodalton rhoptry antigen of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 33: 203-204

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 infected with mature stages of *Plasmodium falciparum*. Mol. Biochem. Parasitiol.. 20: 265-277

Coppel, R.L., Bianco, A.E., Culvenor, J.G., Crewther, P.E., Brown, G.V., Anders, R.F. and Kemp, D.J. (1987). A cDNA clone expressing a rhoptry protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 25: 73-81

Coppel, R.L. (1992). Repeat structures in a *Plasmodium falciparum* protein (MESA) that binds human erythrocyte protein 4.1. Mol. Biochem. Parasitol., 50: 335-348

Cowman, A.F., Coppel, R.L., Saint, R.B., Favaloro, J., Crewther, P.E., Stahl, H.D., Bianco, A.E., Brown, G.V., Anders, R.F. and Kemp, D.J. (1984). The ring-infected erythrocyte surface antigen (RESA) polypeptide of *Plasmodium falciparum* contains two separate blocks of tandem repeats encoding antigenic epitopes that are naturally immunogenic in man. Mol. Biol. Med., 2: 207-221

Cowman, A.F., Saint, R.B., Coppel, R.L., Brown, G.V., Anders, R.F. and Kemp, D.J. (1985). Conserved sequences flank variable tandem repeats in two S-antigen genes of *Plasmodium falciparum*. Cell, 40: 775-783

Cranston, H.A., Boylan, C.W., Carroll, G., Sutera, S.P., Williamson, J.R., Gluzman, I.Y. and Krogstad, D.J. (1984). *Plasmodium falciparum* maturation abolishes physiologic red cell deformability. Science, 223: 400-402

Crewther, P.E., Culvenor, J.G., Silva, A., Cooper, J.A. and Anders, R.F. (1990). *Plasmodium falciparum*: two antigens of similar size are located in different compartments of the rhoptry. Exp. Parasitol., 70: 193-206

Culvenor, J.G., Langford, C.J., Crewther, P.E., Saint, R.B., Coppel, R.L., Kemp, D.J., Anders, R.F. and Brown, G.V. (1987). *Plasmodium falciparum*: identification and localization of a knob protein antigen expressed by a cDNA clone. Exp. Parasitol., 63: 58-67

Culvenor, J.G., Day, K.P. and Anders, R.F. (1991). *Plasmodium falciparum* ring-infected erythrocyte surface antigen is released from merozoite dense granules after erythrocyte invasion. Infect. Immun., 59: 1183-1187

Daly, T.M. and Long, C.A. (1993). A recombinant 15-kilodalton carboxy-terminal fragment of *Plasmodium yoelii yoelii* 17XL merozoite surface protein 1 induces a protective immune response in mice. Infect. Immun., 61: 2462-2467

Dame, J.B., Williams, J.L., McCutchan, T.F., Weber, J.L., Wirtz, R.A., Hockmeyer, W.T., Maloy, W.L., Haynes, J.D., Schneider, I., Roberts, D., Sanders, G.S., Reddy, E.P., Diggs, C.L. and Miller, L.H. (1984). Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. Science, 225: 593-599

Da Silva, E., Foley, M., Dluzewski, A.R., Murray, L.J., Anders, R.F. and Tilley, L. (1994). The *Plasmodium falciparum* protein RESA interacts with the erythrocyte cytoskeleton and modifies erythrocyte thermal stability. Mol. Biochem. Parasitol., 66: 59-69

Da Silveira, J.F., Mercereau-Puijalon, O. (1983). *Plasmodium chabaudi* antigens synthesized in an mRNA-dependent cell-free translation system. Mol. Biochem. Parasitol., 7: 159-172

Deans, J.A., Thomas, A.W., Alderson, T. and Cohen, S. (1984). Biosynthesis of a putative protective *Plasmodium knowlesi* merozoite antigen. Mol. Biochem. Parasitol., 11: 189-204

Deans, J.A., Knight, A.M., Jean, W.C., Waters, A.P., Cohen, S. and Mitchell, G.H. (1988). Vaccination trials in rhesus monkeys with a minor, invariant, *Plasmodium knowlesi* 66 kDa merozoite antigen. Parasitol. Immunol., 10: 535-552

Deleersnijder, W., Hendrix, D., Bendahman, N., Hanegreefs, J., Brijs, L., Hamers-Casterman, C. and Hamers, R. (1990). Molecular cloning and sequence analysis of the gene encoding the major merozoite surface antigen of *Plasmodium chabaudi chabaudi* IP-PCI. Mol. Biochem. Parasitol., 43: 231-244

Deleersnijder, W., Prasomsitti, P., Tungpradubkul, S., Hendrix,D., Hamers-Casterman, C. and Hamers, R. (1992). Structure of a *Plasmodium chabaudi* acidic phosphoprotein that is associated with the host erythrocyte membrane. Mol. Biochem. Parasitol., 56: 59-68

Delves, C.J., Ridley, R.B., Goman, M., Holloway, S.P., Hyde, J.E. and Scaife, J.G. (1989). Cloning of a α-tubulin gene from *Plasmodium falciparum.* Mol. Microbiol., 3: 1511-1519

Delves, C.J., Alano, P., Ridley, R.G., Goman, M., Holloway, S.P., Hyde, J.E. and Scaife, J.G. (1990). Expression of α and ß tubulin genes during the asexual and sexual blood stages of *Plasmodium falciparum.* Mol. Biochem. Parasitol., 43: 271-278

Dibb, N.J., Maruyama, I.N., Krause, M. and Karn, J. (1989). Sequence analysis of the complete *Caenorhabditis elegans* myosin heavy chain gene family. J. Mol. Biol., 205: 603-613

Dluzewski, A.R., Rangachari, K., Gratzer, W.B. and Wilson, R.J.M. (1983). Inhibition of malarial invasion of red cells by chemical and immunochemical linking of spectrin molecules. Br. J. Haematol., 55: 629-637

Dluzewski, A.R., Rangachari, K., Wilson, R.J.M. and Gratzer, W.B. (1986). *Plasmodium falciparum*: protease inhibitors and inhibition of erythrocyte invasion. Exp. Parasitol., 62: 416-422

Dure, L. (1993). A repeating 11-mer amino acid motif and plant desiccation. Plant J., 3: 363-369

Engelman, D.M., Henderson, R., McLachlan, A.D. and Wallace, B.A. (1980). Path of the polypeptide in bacteriorhodopsin. Proc. Natl. Acad. Sci., USA, 77: 2023-2027

Engelman, D.M., Steitz, T.A. and Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Ann. Rev. Biophys. Chem., 15: 321-353

Enans, E.A. and Skalak,,R. (1980). Mechanics and Thermodynamics of Biomembranes. Boca Raton, FL, CRC Press, 1980, p 1

Etzion, Z., Murray, M.C. and Perkins, M.E. (1991). Isolation and characterization of rhoptries of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 47: 51-62

Eyre, D.R. (1980). Collagen: molecular diversity in the body's protein scaffold. Science, 207: 1315-1322

Fang, X., Kaslow, D.C., Adams, J.H. and Miller, L.H. (1991). Cloning of the *Plasmodium vivax* Duffy receptor. Mol. Biochem. Parasitol., 44: 125-132

Favaloro, J.M., Coppel, R.L., Corcoran, L.M., Foote, S.J., Brown, G.V., Anders.R.F. and Kemp, D.J. (1986). Structure of the RESA gene of *Plasmodium falciparum*. Nucl. Acids Res., 14: 8265-8277

Fine, E., Aikawa, M., Cochrane, A.H. and Nussenzweig, R.S. (1984). Immuno-electron microscopic observations on *Plasmodium knowlesi* sporozoites: localization of protective antigen and its precursor. Am. J. Trop. Med. Hyg., 33: 220-226

Foley, M., Murray, L.J. and Anders, R.F. (1990). The ring-infected erythrocyte surface antigen protein of *Plasmodium falciparum* is phosphorylated upon association with the host cell membrane. Mol. Biochem. Parasitol., 38: 69-76

Foley, M., Tilley, L., Sawyer, W.H. and Anders, R.F. (1991). The ring-infected erythrocyte surface antigen of *Plasmodium falciparum* binds to spectrin in the erythrocyte membrane. Mol. Biochem. Parasitol., 46: 137-148

Freeman, R.R., Holder, A.A., Trejdosiewicz, A.J. and Cross, G.A.M. (1980a). Monoclonal antibodies against the rodent malaria parasite, *Plasmodium yoelii*. in The Host-Invader Interplay (ed. Van den Bossche, H.) 121-124 (Elsevier, Amsterdam, 1980)

Freeman, R.R., Trejdosiewicz, A.J. and Cross, G.A.M. (1980b). Protective monoclonal antibodies recognising stage specific merozoite antigens of a rodent malaria parasite. Nature, 284: 366-368

Frontali, C. and Pizzi, E. (1991). Conservation and divergence of repeated structures in *Plasmodium* genomes: The molecular drift. Acta Leidensia, 60: 69-81

Galinski, M.R., Arnot, D.E., Cochrane, A.H., Barnwell, J.W., Nussenzweig, R.S. and Enea, V. (1987). The circumsporozoite gene of the *Plasmodium cynomolgi* complex. Cell, 48: 311-319

Galinski, M.R., Medina, C.C., Ingravallo, P. and Barnwell, J.W. (1992). A reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. Cell,69: 1213-1226

Garnier,J., Osguthorpe, D.J. and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol., 120: 97-120

Garnham, P.C.C. (1968). Malaria parasites and other haemosporidia. Blackwell Scientific Publications, Oxford.

Geiger, B. (1987). Intermediate filaments: looking for a function. Nature, 329: 392-393

Green, N.M. (1990). Structural deductions from receptor sequences. In Receptor Biochemistry, (ed.E.C. Hume), IRL Press, Oxford, p. 277-301

Greener, A. (1990). *E. coli* SURE: clone "unclonable" DNA. Strategies, 3: 5-6

Grimmet, M.R. (1979). Diazoles, triazoles, tetrazoles and their benzo-analogues. In P.G. Sammes, editor, *Comprehensive organic chemistry: the synthesis and reaction of organic compounds*, volume 4, heterocyclic compounds, chapter 17.3, pages 362-363. Pergamon Press, Oxford, U.K.

Gumpel, N.,J., and Smith, A.G. (1992). A novel calcium-binding protein from *Euglena gracilis*. Characterisation of a cDNA encoding a 74-kDa acidic-repeat protein targeted across the endoplasmatic reticulum. Eur. J. Biochem., 210: 721-727

Hadley, T.J., Klotz, F.W., Pasvol, G., Haynes, J.D., McGinniss, M.H., Okubo, Y. and Miller, L.H. (1987). Falciparum malaria parasites invade erythrocytes that lack glycophorin A and B ($M^k M^k$). J. Clin. Invest., 80: 1190-1193

Hadley, T.J., Miller, L.H. and Haynes, J.D. (1991). Recognition of red cells by malaria parasites: the role of erythrocyte-binding proteins. Transfusion Med. Rev., 5: 108-122

Hanahan, D. and Meselson, M. (1980). Plasmid screening at high density. Gene, 10: 63-67

Harbury, P.B., Zhang, T., Kim, P.S. and Alber, T.A. (1993). A switch between two-, three-, and four-stranded coiled-coils in GCN4 leucine zipper mutants. Science, 292: 1401-1407

Harnyuttanakorn, P., McBride, J.S., Donachie, S., Heidrich, H.G. and Ridley, R.G. (1992). Inhibitory monoclonal antibodies recognize epitopes adjacent to a proteolytic cleavage site on the RAP-1 protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 55: 177-186

Haynes, J.D., Dalton, J.P., Klotz. F.W., McGinniss, M.H., Hadley, T.J., Hudson, D.E. and Miller, L.H. (1988). Receptor-like specificity of a *Plasmodium knowlesi* malarial protein that binds to Duffy antigen ligands on erythrocytes. J. Exp. Med., 167: 1873-1881

Henderson R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckman, E. and Downing,K.H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J. Mol. Biol., 213: 899-929

Holder, A.A. and Freeman, R.R. (1981). Immunization against blood-stage rodent malaria using purified parasite antigens. Nature, 294: 361-364

Holder, A.A. and Freeman, R.R. (1982). Biosynthesis and processing of a *Plasmodium falciparum* schizont antigen recognized by immune serum and a monoclonal antibody. J. Exp. Med., 156: 1528-1538

Holder, A.A. and Freeman, R.R. (1984). The three major antigens on the surface of *Plasmodium falciparum* merozoites are derived from a single high molecular weight precursor. J. Exp. Med., 160: 624-629

Holder, A.A., Lockyer, M.J., Odink, K.G., Sandhu, J.S., Riveros-Moreno, V., Nicholls, S.C., Hillman, Y., Davey, L.S., Tizard, M.L.V., Schwarz, R.T. and Freeman, R.R. (1985a). Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. Nature, 317: 270--273

Holder, A.A., Freeman, R.R., Uni, S. and Aikawa, M. (1985b). Isolation of a *Plasmodium falciparum* rhoptry protein. Mol. Biochem. Parasitol., 14: 293-303

Holder, A.A., Sandhu, J.S., Hillman, Y., Davey, L.S., Nicholls, S.C., Cooper, H. and Lockyer, M.J. (1987). Processing of the precursor to the major merozoite antigens of *Plasmodium falciparum*. Parasitol., 94: 199-208

Holder, A.A. (1988). The precursor to major merozoite surface antigens: Structure and role in immunity. Prog. Allergy, 41: 72-97

Holloway, S.P., Gerousis, M., Delves, C.J., Sims, P.F.G., Scaife, J.G. and Hyde, J.E. (1990). The tubulin genes of the human malaria parasite *Plasmodium falciparum*, their chromosomal location and sequence analysis of the α-tubulin II gene. Mol. Biochem. Parasitol., 43: 257-270

Howard. R.J., Lyon, J.A., Uni, S., Saul, A., Aley, S.B., Klotz, F., Panton, L.J., Sherwood, J.A., Marsh, K., Aikawa, M. and Rock, E. (1987). Transport of an M_r ~ 300 000 *Plasmodium falciparum* protein (Pf EMP 2) from the intraerythrocytic asexual parasite to the cytoplasmic face of the host cell membrane. J. Cell Biol., 104: 1269-1280

Howard, R.J., Barnwell, J.W., Rock, E.P., Neequaye, J., Ofori-Adjei, D., Maloy, W.L., Lyon, J.A. and Saul, A. (1988a). Two approximately 300 kilodalton *Plasmodium falciparum* proteins at the surface membrane of infected erythrocytes. Mol. Biochem. Parasitol., 27: 207-224

Howard, R.F., Stanley, H.A. and Reese, R.T. (1988b). Characterization of a high-molecular weight phosphoprotein synthesized by the human malarial parasite *Plasmodium falciparum.* Gene, 64: 65-75

Hubbard, S.C. and Ivatt, R.J. (1981). Synthesis and processing of asparagine-linked oligosaccharides. Annu. Rev. Biochem., 50: 555-583

Hughes, A.L. (1994). Conserved proteins as immunogens: glutathione-S-transferases of *Schistosoma.* Parasitol. Today, 10: 149-151

Jaikaria, N.S., Rozario, C., Ridley, R.G. and Perkins, M.E. (1993). Biogenesis of rhoptry organelles in *Plasmodium falciparum.* Mol. Biochem. Biol., 57: 269-280

Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992). A new approach to protein fold recognition. Nature, 358: 86-89

Kagawa, H., Gengyo, K., McLachlan, A.D., Brenner, S. and Karn, J. (1989). Paramyosin gene (*unc*-15) of *Caenorhabditis elegans*. Molecular cloning, nucleotide sequence and models for thick filament structure. J. Mol. Biol., 207: 311-333

Keen, J.K., Sinha, K.A., Brown, K.N. and Holder. A.A. (1994). A gene coding for a high-molecular mass rhoptry protein of *Plasmodium yoelii*. Mol. Biochem. Parasitol., 65: 171-177

Kemp, D.J., Coppel, R.L., Cowman, A.F., Saint, R.B., Brown, G.V. and Anders, R.F. (1983). Expression of *Plasmodium falciparum* blood-stage antigens in *Escherichia coli*: Detection with antibodies from immune humans. Proc. Natl. Acad, Sci. USA, 80: 3787-3791

Kemp, D.J., Coppel, R.L., Stahl, H.D., Bianco, A.E., Corcoran, L.M., McIntyre, P., Langford, C.J., Favaloro, J.M., Crewther, P.E., Brown, G.V., Mitchell, G.F., Culvenor, J.G. and Anders, R.F. (1986). Genes for antigens of *Plasmodium falciparum.* Parasitol., 91: S83-S108

Kemp, D.J., Coppel, R.L. and Anders, R.F. (1987). Repetitive proteins and genes of malaria. Ann. Rev. Microbiol., 41: 181-208

Kitchen, S.F. (1938). The infection of reticulocytes by *Plasmodium vivax.* Am. J. Trop. Med., 18: 347-359

Kilejian, A., Rashid, M.A., Aikawa, M., Aji, T. and Yang, Y.F. (1991). Selective association of a fragment of the knob protein with spectrin, actin and the red cell membrane. Mol. Biochem Parasitol., 44: 175-182

Knapp, B., Nau, U., Hundt, E. and Küpper, H.A. (1991). A new blood stage antigen of *Plasmodium falciparum* highly homologous to the serine-stretch protein SERP. Mol. Biochem. Parasitol., 44: 1-14

Kobayashi, S., Eden-McCutchan, F., Framson, P. and Bornstein, P. (1986). Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins. Biochemistry, 25: 8418-8425

Kochan, J., Perkins, M. and Ravetch, J.V. (1986). A tandemly repeated sequence determines the binding domain for an erythrocyte receptor binding protein of *P. falciparum.* Cell, 44: 689-696

Kretz, P.L., Reid, C.H., Greener, A. and Short, J.M. (1989). Effect of lambda packaging extract *mcr* restriction activity on DNA cloning. Nucleic Acids Res., 17: 5409

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 157: 105-132

Ladda, R.L., Aikawa,M. and Sprinz, H. (1969). Penetration of erythrocytes by merozoites of mammalian and avian malarial parasites. J. Parasitol., 55: 633-644

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bateriophage T4. Nature, 227: 680-695

Landau, I. (1965). Description de *Plasmodium chabaudi* n.sp., parasite de rongeurs africain. Séanc. Acad. Sci. Paris, 260: 3756-3761

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K. and Hartl, F.U. (1992). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature, 356: 683-689

Langreth, S.G., Jensen, J.B., Reese, R.T. and Trager, W. (1978). Fine structure of human malaria *in vitro*. J. Protozool., 25: 443-452

La Polla, R.J., Haron, J.A., Kelly, C.G., Taylor, W.R., Bohart, C., Hendricks, M., Pyati, J., Graff, R.T., Ma, J.K.C. and Lehner, T. (1991). Sequence and structural analysis of surface protein antigen-I/II (SPAA) of *Streptococcus sobrinus*. Infection and Immunity, 59:2677-2685

Lupas, A., Dyke, M. and Stock. J. (1991). Predicting coiled coils from protein sequences. Science, 252: 1162-1164

Lustigman, S., Anders, R.F., Brown, G.V. and Coppel, R.L. (1988). A component of an antigenic rhoptry complex of *Plasmodium falciparum* is modified after merozoite invasion. Mol. Biochem. Parasitol., 30: 217-224

Lustigman, S., Anders, R.F., Brown, G.V. and Coppel, R. (1990). The mature-parasite-infected erythrocyte surface antigen (MESA) of *Plasmodium falciparum* associates with the erythrocyte membrane skeletal protein, band 4.1 Mol. Biochem. Parasitol., 38: 261-270

Lux, S.E., John, K.M. and Bennett, V. (1990). Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. Nature, 344: 36-42

Marchesi,V.T. (1985). Stabilizing infrastructure of cell membranes. Annu. Rev. Cell Biol., 1: 531-561

Matsumoto, Y., Matsuda, S. and Yoshida, Y. (1986). Ultrastructure of human erythrocytes infected with *Plasmodium ovale*. Am. J. Trop. Med. Hyg., 35: 697-703

Matsumoto, Y., Aikawa, M. and Barnwell, J.W. (1988). Immunoelectron microscopic localization of vivax malaria antigens to the clefts and caveola-vesicle complexes of infected erythrocytes. Am. J. Trop. Med. Hyg., 39: 317-322

McBride, J.S. and Heidrich, H.-G. (1987). Fragments of the polymorphic M_r 185, 000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites from an antigenic complex. Mol. Biochem. Parasitol., 23: 71-84

McColl, D.J., Silva, A., Foley, M., Kun, J.F.J., Favaloro, J.M., Thompson, J.K., Marshall, V.M., Coppel, R.L., Kemp, D.J., Anders, R.F. (1994). Molecular variation in a novel polymorphic antigen associated with *Plasmodium falciparum* merozoites. Mol. Biochem. Parasitol., 68: 53-67

McDonald, I.K. and Thornton, J.M. (1994). Satisfying hydrogen-bonding potential in proteins. J. Mol. Biol., 238: 777-793

Means, A.R. (1988). Molecular mechanisms of action of calmodulin. Recent Prog. Hormone Res., 44: 223-262

Miller, L.H., Usami, S. and Chien, S. (1971). Alteration in the rheologic properties of *Plasmodium knowlesi* - infected red cells. A plausible mechanism for capillary obstruction. J.Clin. Invest., 50: 1451-

Miller, L.H., Chien, S. and Usami, S. (1972). Decreased deformability of *Plasmodium coatneyi* infected red cells and its possible relation to cerebral malaria. Am. J. Trop. Med. Hyg., 21: 133

Miller, L.H., Mason, S.J., Dvorak, J.A., McGinniss, M.H. and Rothman, I.K. (1975). Erythrocyte receptors for (*Plasmodium knowlesi*) malaria: Duffy blood group determinants. Science, 189: 561-563

Miller, L.H., Mason, S.J., Clyde, D.F. and McGinniss, M.H. (1976). The resistance factor to *Plasmodium vivax* in blacks. The Duffy blood group genotype, FyFy. New Engl. J. Med., 295: 302-304

Miller, L.H., Haynes, J.D., McAuliffe, F.M., Shiroishi, T., Durocher, J.R. and McGinniss, M.H. (1977). Evidence for differences in erythrocyte surface receptors for the malarial parasite *Plasmodium falciparum* and *Plasmodium knowlesi*. J. Exp. Med., 146: 277-281

Mitchell, G.H., Hadley, T.J., McGinniss, M.H., Klotz, F.W. and Miller, L.H. (1986). Invasion of erythrocytes by *Plasmodium falciparum* malaria parasites: Evidence for receptor heterogeneity and two receptors. Blood, 67: 1519-1521

Mulhern, T.D., Howlett, G.J., Reid, G.E., Simpson, R.J., McColl, D.J., Anders, R.F. and Norton, R.S. (1995). Solution structure of a polypeptide containing four heptad repeat units from a merozoite surface antigen of *Plasmodium falciparum*. Biochemistry, 34: 3479-3491

Müller, H.-M., Reckmann, I., Hollingdale, M.R., Bujard, H., Robson, K.J.H. and Crisanti, A. (1993). Thrombospondin related anonymous protein (TRAP) of *Plasmodium falciparum* binds specifically to sulfated glycoconjugates and to HepG2 hepatoma cells suggesting a role for this molecule in sporozoite invasion of hepatocytes. EMBO J., 12: 2881-2889

Narum, D.L. and Thomas, A.W. (1994). Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of Plasmodium falciparum merozoites. Mol. Biochem. Parasitol., 67: 59-68

Nash, G.B., O'Brien, E., Gordon-Smith, E.C. and Dormandy, J.A. (1989). Abnormalities in the mechanical properties of red blood cells caused by *Plasmodium falciparum*. Blood, 74: 855-861

Nichols, B., Chiappino, M., and O'Connor, R. (1983). Secretion from the rhoptries of *Toxoplasma gondii* during host-cell invasion. J. Ultrastruct. Res., 83: 85-98

Nolte, D., Hundt, E., Langsley, G. and Knapp, B. (1991). A *Plasmodium falciparum* blood stage antigen highly homologous to the glycophorin binding protein GBP. Mol. Biochem. Parasitol., 49: 253-264

North, A.C.T., Steinert, P.M. and Parry, D.A.D. (1994). Coiled-coil stutter and link segments in keratin and other intermediate filament molecules: A computer modelling study. Prot. Struct. Funct. Genet., 20: 174-184

Ogun, S.A. and Holder, A.A. (1994). *Plasmodium yoelii*: Brefeldin A-sensitive processing of proteins targeted to the rhoptries. Exp. Parasitol., 79: 270-278

Oka, M., Aikawa, M., Freeman, R.R., Holder, A.A. and Fine, E. (1984). Ultrastructural localization of protective antigens of *Plasmodium yoelii* merozoites by the use of monoclonal antibodies and ultrathin cryomicrotomy. Am. J. Trop. Med. Hyg., 33: 342-346

Olson, J.A. and Kilejian, A. (1982). Involvement of spectrin and ATP in infection of resealed erythrocyte ghosts by the human malaria parasite, *Plasmodium falciparum*. J. Cell. Biol., 95: 757-762

Oo, M.M., Aikawa, M., Than, T., Aye, T.O.M., Myint, P.T., Igarashi, I. and Schoene, W.C. (1987). Human cerebral malaria: a pathological study. J. Neuropathol. Exp. Neurol., 46: 223-231

O'Shea, E.K., Klemm, J.D., Kim, P.S. and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper: a 2-stranded, parallel coiled-coil. Science, 254: 539-544
Parry, D.A.D. and Fraser, R.D.B. (1985). Intermediate filament structure: 1. Analysis of IF protein sequence data. Int. J. Biol. Macromol., 7: 203-213

Pasloske, B.L., Baruch, D.I., van Schravendijk, M.R., Handunnetti, S.M., Aikawa, M., Fujioka, H., Taraschi, T.F., Gormley, J.A. and Howard, R.J. (1993). Cloning and characterization of a *Plasmodium falciparum* gene encoding a novel high-molecular weight host membrane-associated protein, PfEMP3. Mol. Biochem. Parasitol., 59: 59-72

Pasvol, G., Wainscoat, J.S. and Weatherall, D.J. (1982a). Erythrocytes deficient in glycophorin resist invasion by the malarial parasite *Plasmodium falciparum*. Nature, 297: 64-66

Pasvol, G., Jungery, M., Weatherall, D.J., Parsons, S.F., Anstee, D.J. and Tanner, M.J.A. (1982b). Glycophorin as a possible receptor for *Plasmodium falciparum.* Lancet, 2: 947-950

Pauling, L. and Corey, R.B. (1951). The structure of synthetic polypeptides. Proc. Natl. Acad. Sci. USA, 37: 241-250

Pearson, W.R. and Lipman, D.J. (1988). Improved tools for biological sequence analysis. Proc. Natl. Acad. Sci., USA, 85: 2444-2448

Perkins, M.E. (1992). Rhoptry organelles of apicomplexan parasites. Parasitol. Today, 8: 28-32

Perlman, H.K., Berzins, K., Whlin, B., Udomsangpetch, R., Ruangjirachuporn, W., Wahlgren, M. and Perlmann, P.H. (1987). Absence of antigenic diversity in Pf155, a major parasite antigen in the membrane of erythrocytes infected with *Plasmodium falciparum.* J. Clin. Microbiol., 25: 2347-2354

Perrin, L.H. and Dayal, R. (1982). Immunity to asexual erythrocytic stages of *Plasmodium falciparum*: role of defined antigens in the humoral immune response. Immunol. Rev., 61: 245-269

Perutz, M.F., Johnson, T., Suzuki, M. and Finch, J.T. (1994). Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc. Natl. Acad. Sci. USA, 91: 5355-5358

Peterson, M.G., Marshall, V.M., Smythe, J.A., Crewther, P.E., Lew, A., Silva, A., Anders, R.F. and Kemp, D.J. (1989). Integral membrane protein located in the apical complex of *Plasmodium falciparum.* Mol. Cell. Biol., 9: 3151-3154

Preis, D., van der Pas, J.C., Nehls, U., Röhlen, D.A., Sackmann, U., Jahnke, U. and Weiss, H. (1990). The 49 K subunit of NADH: ubiquinone reductase (complex I) from *Neurospora crassa mitochondria*: primary structure of the gene and the protein. Curr. Genet., 18: 59-64

Ramasamy, R. (1987). Characterization of a novel merozoite surface antigen of the human malaria parasite *Plasmodium falciparum.* Immunol. Cell. Biol., 65: 419-424

Rashid, M.A., Yan, Y.F. and Kilejian, A. (1990). Expression, partial purification and immunogenicity of fragments of the knob protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 38: 49-56

Rawlings, D.J., Fujioka, H., Fried, M., Keister, D.B., Aikawa, M. and Kaslow, D.C. (1992). α-tubulin II is a male-specific protein in *Plasmodium falciparum*. Mol. Biochem. Parasitol., 56: 239-250

Ridley, R.G., Takacs, B., Lahm, H.W., Delves, C.J., Goman, M., Certa, U., Matile,H., Woollett, G.R. and Scaife, J.G. (1990a). Characterisation and sequence of a protective rhoptry antigen from *Plasmodium falciparum*. Mol. Biochem. Parasitol., 41: 125-134

Ridley, R.G., Takacs, B., Etlinger, H. and Scaife, J.G. (1990). A rhoptry antigen of *Plasmodium falciparum* is protective in *Saimiri* monkeys. Parasitology, 101: 187-192

Ridley, R.G., Lahm, H.W., Takács, B. and Scaife, J.G. (1991). Genetic and structural relationships between components of a protective rhoptry antigen complex from *Plasmodium falciparum*. Mol. Biochem. Parastol., 47:245-246

Roger, N., Dubremetz, J.F., Delplace, P., Fortier, B., Tronchin, G. and Vernes, A. (1988). Characterization of a 225 kilodalton rhoptry protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 27: 135-142

Rogers, W.O., Malik, A., Mellouk, S., Nakamura, K., Rogers, M.D., Szarfman, A., Gordon, D.M., Nussler, A.K., Aikawa, M. and Hofman, S.L. (1992). Characterization of *Plasmodium falciparum* sporozoite surface protein 2. Proc. Natl. Acad. Sci., USA, 89: 9176-9180

Romero, P. (1992). Malaria vaccines. Current Opinion in Immunology, 4: 432-441

Ruangjirachuporn, W., Udomsangpetch, R., Carlsson, J., Drenckhahn, D.D., Perlmann, P. and Berzins, K. (1991). *Plasmodium falciparum*: Analysis of the interaction of antigen Pf155/RESA with the erythrocyte membrane. Exp. Parasitol., 73: 62-72

Ruoslahti, E. and Pierschbachter, M.D. (1986). Arg-Gly-Asp: a versatile cell recognition signal. Cell, 44: 517-518

Saint, R.B., Coppel, R.L., Cowman, A.F., Brown, G.V., Shi, P.-T., Bazarga, N., Kemp, D.J. and Anders, R.F. (1987). Changes in repeat number, sequence and reading frame in S-antigen genes of *Plasmodium falciparum*. Mol. Cell. Biol., 97: 2968-2973

Sam-Yellowe, T.Y., Shio, H. and Perkins, M.E. (1988). Secretion of *Plasmodium falciparum* rhoptry protein into the plasma membrane of host erythrocytes. J. Cell Biol., 106: 1507-1513

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, 85:5463-5467

Saul, A. and Battistutta, D. (1990). Analysis of the sequences flanking the translational start sites of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 42: 55-62

Saul, A., Lord, R., Jones, G.L. and Spencer, L. (1992a). Protective immunization with invariant peptides of the *Plasmodium falciparum* antigen MSA-2. J. Immunol., 148: 208-211

Saul, A., Cooper, J., Hauquitz, D., Irving, D., Cheng, Q., Stowers, A. and Limpaiboon, T. (1992b). The 42-kilodalton rhoptry-associated protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol., 50: 139-150

Scheibel, L.W., Colombani, P.M., Hess, A.D., Aikawa, M., Atkinson, C.T. and Milhous, W.K. (1987). Calcium and calmodulin antagonists inhibit human malaria parasites (*Plasmodium falciparum*): Implications for drug design. Proc. Natl. Acad. Sci., USA, 84: 7310-7314

Scherf, A., Hilbich, C., Sieg, K., Mattei, D., Mercereau-Puijalon, O. and Müller-Hill, B. (1988). The 11-1 gene of *Plasmodium falciparum* codes for distinct fast evolving repeats. EMBO J., 7: 1129-1137

Schirmer, R.H., Müller, J.G. and Krauth-Siegel, R.L. (1995). Disulfide-reductase inhibitors as chemotherapeutic agents: The design of drugs for trypanosomiasis and malaria. Angew. Chem. Int. Ed. Engl., 34: 141-154

Schofield, L., Bushell, G.R., Cooper, J.A., Saul, A.J., Upcroft, J.A. and Kidson, C. (1986). A rhoptry antigen of *Plasmodium falciparum* contains conserved and variable epitopes recognized by inhibitory monoclonal antibodies. Mol. Biochem. Parasitol., 18: 183-195

Schofield, L. (1991). On the function of repetitive domains in protein antigens of *Plasmodium* and other eukaryotic parasites. Parasitology Today, 7: 99-105

Seed, T.M., Sterling, C.R., Aikawa, M. and Rabbege, J. (1976). *Plasmodium simium*: ultrastructure of the erythrocytic phase. Exp. Parasitol., 39: 262-276

Sherman, I.W. (1983). Membrane structure and function of malaria parasites and the infected erythrocyte. Parasitology, 91: 609-645

Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. and Kan, S. (1987). Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. Proc. Natl. Acad. Sci., USA, 84: 3014-3018

Sim, B.K., Orlandi, P.A., Haynes, J.D., Klotz, F.W., Carter, J.M., Camus, D., Zegans, M.E. and Chulay, J.D. (1990). Primary structure of the 175 K *Plasmodium falciparum* erythrocyte binding antigen and identification of a peptide which elicits antibodies that inhibit malaria merozoite invasion J. Cell Biol.,111: 1877-1884

Sim, B.K.L., Toyoshima, T., Haynes, J.D. and Aikawa, M. (1992). Localization of the 175-kilodalton erythrocyte binding antigen in micronemes of *Plasmodium falciparum* merozoites. Mol. Biochem. Parasitol., 51: 157-160

Sim, B.K.L., Chitnis, C.E., Wasniowska, K., Hadley, T.J. and Miller, L.H. (1994). Receptor and ligand domains for invasion of erythrocytes by *Plasmodium falciparum*. Science, 264: 1941-1944

Smith, D.B., Davern, K.M., Board, P.G., Tiu, W.U., Garcia, E.G., and Mitchell, G.F. (1986). M_r 26 000 antigen of *Schistosoma japonicum* recognized by resistant WEHI 129/J mice is a parasite glutathione-S-transferase. Proc. Natl. Acad. Sci. USA, 83: 8703-8707

Smith, D.B. and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-S-transferase. Gene, 67: 31-40

Smith, T.F. and Waterman, M.S. (1981). Identification of common molecular subsequences. J. Mol. Biol., 147: 195-197

Smythe, J.A., Coppel, R.L., Brown, G.V., Ramasamy, R., Kemp, D.J. and Anders, R.F. (1988). Identification of two integral membrane proteins of *Plasmodium falciparum*. Proc. Natl. Acad. Sci., USA, 85: 5195-5199

Spiess, M. and Lodish, H.F. (1986). An internal signal sequence: the asialoglycoprotein receptor membrane anchor. Cell, 44: 177-185

Spiess, M. and Handschin, C. (1987). Deletion analysis of the internal signal-anchor domain of the human asialoglycoprotein receptor H1. EMBO J., 6: 2683-2691

Steinert, P.M., Roop, D.R. Molecular and cellular biology of intermediate filaments. Ann. Rev. Biochem., 57 (1988) 593-626

Stahl, H.D., Crewther, P.E., Anders, R.F., Brown, G.V., Coppel, R.L., Bianco, A.E., Mitchell, G.F. and Kemp, D.J. Interspersed blocks of repetitive and charged amino acids in a dominant immunogen of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA., 82 (1985) 543-547

Stahl, H.D., Crewther, P.E., Anders, R.F. and Kemp, D.J. (1987). Structure of the FIRA gene of *Plasmodium falciparum*. Mol. Biol. Med., 4: 199-211

Stewart, M.J., Schulman, S. and Vanderberg, J.P. (1985). Rhoptry secretion of membranous whorls by *Plasmodium berghei* sporozoites. J. Protozool., 32: 280-283

Stewart, M.J., Schulman, S. and Vanderberg, J.P. (1986). Rhoptry secretion of membranous whorls by *Plasmodium falciparum* merozoites. Am. J. Trop. Med. Hyg., 35: 37-44

Sturchler, D. (1989). How much malaria is there worldwide? Parasitol. Today, 5: 39-40

Suetterlin, B.W., Kappes, B. and Franklin, R.M. (1991). Localization and stage specific phosphorylation of *Plasmodium falciparum* phosphoproteins during the intraerythrocytic cycle. Mol. Biochem. Parasitol., 46: 113-122

Suetterlin, B.W., Kappes, B., Jenoe, P. and Franklin, R.M. (1992). An 88-kDa protein of Plasmodium falciparum is related to the band-3-binding domain of human erythrocyte ankyrin. Eur. J. Biochem., 207: 455-461

Tanabe, K. (1990). Ion metabolism in malaria-infected erythrocytes. Blood Cells, 16: 437-449

Taylor, D.W., Parra, M. and Stearns, M.E. (1987a). *Plasmodium falciparum*: Fine structural changes in the cytoskeletons of infected erythrocytes. Experimental Parasitology, 64: 178-187

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. (19987). Localization of *Plasmodium falciparum* histidine-rich protein 1 in the erythrocyte skeleton under knobs. Mol. Biochem. Parasitol., 25: 165-174

Thomas, A.W., Jeans, J.A., Mitchell, G.H., Alderson, T. and Cohen, S. (1984). The Fab fragments of monoclonal IgG to a merozoite surface antigen inhibit *Plasmodium knowlesi* invasion of erythrocytes. Mol. Biochem. Parasitol., 13: 187-199

Torii, M., Adams, J.H., Miller, L.H. and Aikawa, M. (1989). Release of merozoite dense granules during erythrocyte invasion by Plasmodium knowlesi. Infect.Immun., 57: 3230-3233

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from acrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci., USA, 76: 4350-4354

Triglia, T., Peterson, M.G. and Kemp, D.J. (1988). A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res., 16: 8186

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

von Heijne, G. (1991). Computer analysis of DNA and protein sequences Eur. J. Biochem., 199: 253-256

Waters, A.P., Thomas, A.W., Deans, J.A., Mitchell, G.H., Hudson, D.E., Miller, L.H., McCutchan, T.F. and Cohen, S. (1990). A merozoite receptor protein from *Plasmodium knowlesi* is highly conserved and distributed throughout *Plasmodium*. J. Biol. Chem., 265: 17974-17979

Watier, H., Verwaerde, C., Landau, I., Werner, E., Fontaine, J., Capron, A., Auriault, C. (1992). T-cell dependent immunity and thrombocytopenia in rats infected with *Plasmodium chabaudi*. Infect. Immun., 60: 136-142

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. and Howard, R.J. (1986). Homologous genes encode two distinct histidine-rich proteins in a cloned isolate of *Plasmodium falciparum*. Proc. Natl. Acad. Sci., USA, 83: 6065-6069

Wertheimer, S.P. and Barnwell, J.W. (1989). *Plasmodium vivax* interaction with the human Duffy blood group glycoprotein: identification of a parasite receptor-like protein. Exp. Parasitol., 69: 340-350

Wesseling, J.G., Dirks, R., Smits, M.A. and Schoenmakers, J.G.G. (1989a). Nucleotide sequence and expression of a ß-tubulin gene from *Plasmodium falciparum*, a malarial parasite of man. Gene, 83: 301-309

Wesseling, J.G., Snijders, P.J.F., van Someren, P., Jansen, J., Smits, M.A. and Schoenmakers, J.G.G. (1989b). Stage specific expression and genomic organization of the actin genes of the malaria parasite *Plasmodium falciparum*. Mol. Biochem. Parasitol., 35: 167-176

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. (1989). Quantitative evaluation of *Escherichia coli* host stains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res., 17: 3469-3478

Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S.C. and Branton, D. (1993). Crystal structure of the repetitive segment of spectrin. Science, 262: 2027-2029

Yanisch-Perron, C., Viera, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of M13mp18 and pUC19 vectors. Gene, 33: 103

Zerial, M., Melancon, P., Schneider, C. and Garoff, H. (1986). The transmembrane segment of the human transferrin receptor functions as a signal peptide. EMBO J., 5: 1543-1550

Zhe Li, Xiaowu Chen, Davidson, E., Zwang, O., Mendis, C., Ring, C.S., Rouish, W.R., Fegley, G., Rongshi Li, Rosenthal, P.J., Lee, G.K., Kenyon, G.L., Kuntz, I.D. and Cohen, F.E. (1994). Anti-malarial drug development using models of enzyme structure. Chemistry & Biology,1: 31-37