CHAPTER 1

Literature Overview

1.1) Malaria: The disease.

'This day designing God
Hath put into my hand
A wondrous thing. And God
Be praised. At his command
I have found thy secret deed
Oh million murdering Death.
I know that this little thing
A million men will save
Oh death where is thy sting? Thy victory oh grave?'

These were the words written on 20 August 1897 by Ronald Ross upon his definitive discovery of the mosquito transmission of malaria (Sherman, 1998a). However, of the antiquity of malaria there is no doubt and descriptions of the disease have been present since the beginning of the written word. The Ebers Papyrus (ca. 1570 B.C.) as well as clay tablets inscribed with cuneiform writing found in the library of Ashurbanipal (ca. 2000 B.C.) mention deadly periodic fevers (Sherman, 1998a). A clear discussion of quartan (every fourth day) and tertian (every third day) fevers by the physician Hippocrates (470-370 B.C.) in his '*Book of Epidemics*' leaves little doubt that by the fifth century B.C., the disease was present in Europe. Eventually, the disease described as the Roman fever gave rise to the Italian word *mal'aria*, meaning 'bad air' to describe the cause of the sickness superstitiously thought to be due to air emanating from swamps. It was only in 1880 that Charles Louis Alphonse Laveran discovered crescent shaped organisms in a patient suffering from malaria. Laveran received the Nobel Prize for this discovery of *Plasmodia* in 1907 (Desowitz, 1991; Sherman, 1998a).

Currently, malaria is the most important and devastating of all infectious diseases. In any given year, nearly ten percent of the global population (92 countries) will suffer from malaria with 500 million known clinical cases. Estimated mortalities range between 1.5 to 2.7 million deaths annually resulting in a death from malaria every 30 seconds (Collins and Paskewitz, 1995; Greenwood and Mutabingwa, 2002). Africa is particularly affected by the disease, which kills one child in twenty before the age of

five. Ninety percent of the world's infections occur in sub-Saharan Africa due to the persistence of the mosquito vector, *Anopheles gambiae*. The burden of malaria is increasing because of drug and insecticide resistance, war and civil disturbances, environmental and climatic changes, population increase and migration (Greenwood and Mutabingwa, 2002). The monetary cost of malaria worldwide is around US $ 2.200 million per annum (Trigg and Kondrachine, 1998). In recent years large-scale malaria epidemics have occurred on practically all continents (Fig. 1.1).

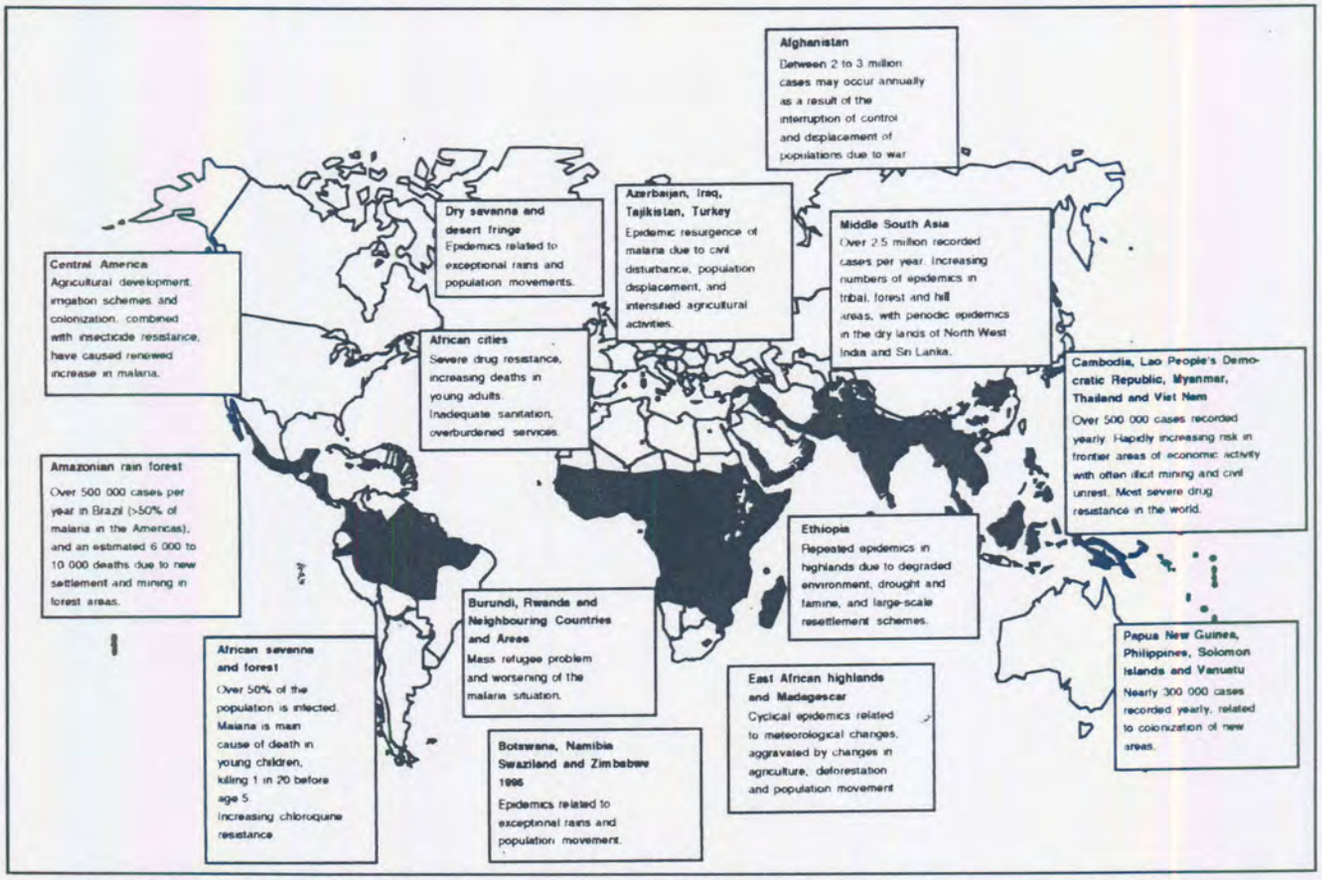


Figure 1.1: Malaria distribution and problem areas. Shading indicates the main areas where malaria transmission occurs. Data from the WHO/CTD/Health Map, 1997.

This literature review will first introduce the causative agents of the disease and their biochemical properties. The pathological manifestations of the disease and efficacy of different control strategies are then examined. In order to understand the biochemistry of the malaria parasite, some differential metabolic pathways are considered which could be interfered with in order to contain the disease. One such target, polyamine metabolism and its unique bifunctional regulatory enzyme complex are then discussed in more detail.

1.2) The etiologic agents of malaria.

The malaria parasites of mammals are all transmitted by blood feeding female mosquitoes belonging to the genus *Anopheles*, including the three species *A. gambiae*, *A. arabiensis* and *A. funestus*, of which *A. gambiae* is the most important (Cox, 1993).

The malaria parasite is a protozoan (unicellular eukaryotic organism) of the kingdom Protista. It is further distinguished in the phylum *Apicomplexa*, in the class *Hematozoa* and order *Haemosporidida* (parasitic in the blood of vertebrates) (Ayala, *et al.*, 1998). The genus *Plasmodium* consists of nearly 200 known species that parasitize reptiles, birds and mammals (Ayala, *et al.*, 1998). Four species of morphologically distinct parasites infect humans: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae* (Ayala, *et al.*, 1998; Cox, 1993). The human malaria parasites form a clade with chimpanzee *Plasmodia*, predicting a time of divergence between 6 to 8 million years ago, which is consistent with the time of divergence between the host species (Ayala, *et al.*, 1998). *P. falciparum* infection can be lethal while *P. vivax* and *P. ovale* cause relapsing malaria by remaining dormant in the liver (hypnozoite). *P. malariae* is unique in its ability to persist in an infected host for decades at very low parasitaemias (Collins and Paskewitz, 1995; Fujioka and Aikawa, 1999). Because of the severity associated with *P. falciparum* infections, this thesis will focus on the many aspects of falciparum malaria.

1.2.1) Life cycle of the human malaria parasites.

Human malaria parasites all undergo the same bi-phasic life cycle. Fig. 1.2 illustrates the sexual (in the insect host) and asexual (in the human host) phases of the parasite. After gametogenesis and sporogony in the mosquito, the sporozoites are inoculated into the human peripheral circulation and invade hepatocytes within 30 minutes. After a 5-15 day period of asexual exoerythrocytic schizogony, as many as 30 000 uninucleate merozoites are produced. During this period there are no clinical symptoms of malaria. The released merozoites invade erythrocytes within about 30 seconds to commence erythrocytic schizogony consisting of four stages: merozoites develop into the ring stage (immature trophozoites visible in peripheral blood of the infected patient), which in turn mature into trophozoites and multinucleate schizonts; schizonts erupt after 48-72 hours and release as many as 36 daughter merozoites per schizont which initiate the next round of invasion and multiplication in erythrocytes. The rupture of the schizonts is typically associated with bouts of fever. Alternatively, some of the merozoites are

capable of developing into gametocytes, which are in turn ingested by a mosquito to complete the cycle (Cox, 1993; Fujioka and Aikawa, 1999).

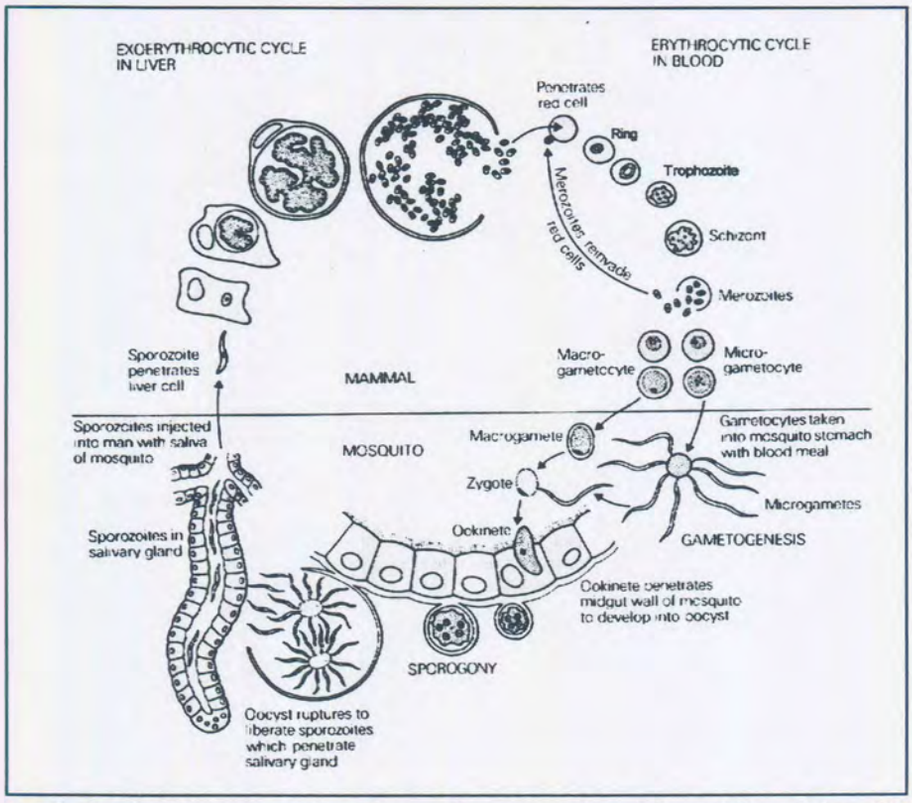


Figure 1.2: Bi-phasic life cycle of the *Plasmodium* parasite. Adapted from (Cox, 1993).

1.2.2) Ultrastructure of the erythrocytic stages of *P. falciparum*.

The erythrocytic stages of the malarial parasites share morphological features and also induce structural and functional changes in the infected erythrocytes. These alterations appear to relate to the capacity of the parasites to alter the properties of the erythrocyte and its membrane to allow survival but are also involved in the development of malaria-related complications in the host (Torii and Aikawa, 1998).

The merozoite is the only extracellular form, has an elliptical shape and measures 1.5 μm in length and 1 μm in diameter (Bannister, *et al.*, 2000; Fujioka and Aikawa, 1999) (Fig. 1.3). The merozoite is surrounded by a trilaminar pellicle composed of a plasma membrane and two closely aligned inner membranes. Below this lies a row of subpellicular microtubules thought to allow motility and underneath this lies the single mitochondrion and a characteristic organelle of the apicomplexa, the plastid. The apical end is truncated and cone-shaped and contain two electron-dense rhoptries, dense granules and micronemes (Fujioka and Aikawa, 1999; Torii and Aikawa, 1998). Together with other features these make up the apical complex, the term that gave rise

to the name of the phylum Apicomplexa (Ayala, *et al.*, 1998). Proteins on the surface of the merozoite are thought to play a role in the invasion of erythrocytes. These are localised in the apical complex and include the polymorphic family of merozoite surface antigens (MSA), rhoptry-associated protein 1 (RAP-1), apical membrane antigen 1 (AMA 1), erythrocyte binding protein (EBA-175) and Duffy-binding ligands in *P. vivax* (Bannister, *et al.*, 2000; Fujioka and Aikawa, 1999; Torii and Aikawa, 1998). The parasite induces a vacuole derived from the erythrocyte plasma membrane and enters the vacuole by a moving junction, therefore enveloping it in a third membrane layer called the parasitophorous vacuolar membrane (PVM) (Miller, *et al.*, 2002).

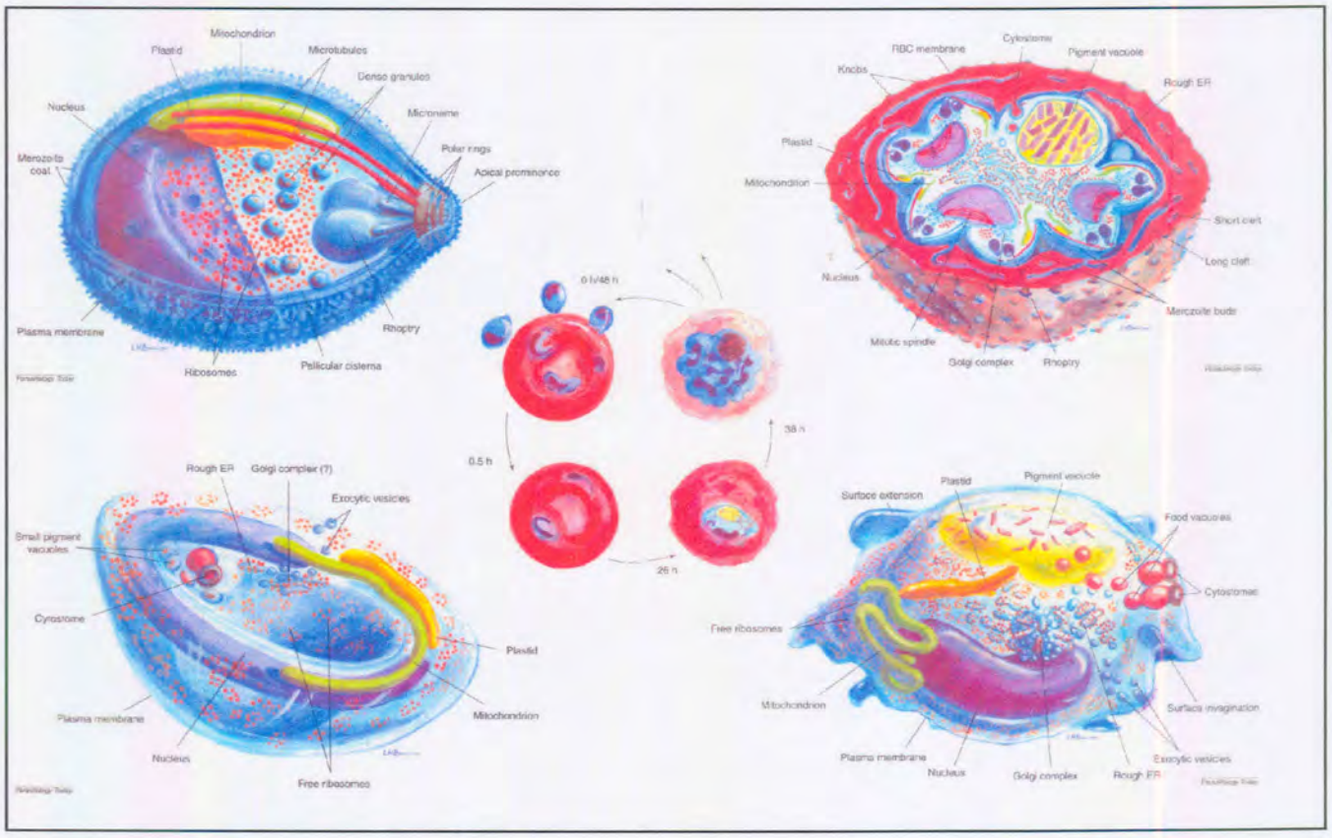


Figure 1.3: Three-dimensional representations of the ultrastructure of the different erythrocytic stages of *P. falciparum*. From top left anticlockwise: merozoite, ring, trophozoite and schizont stages. ER: endoplasmic reticulum. Insert: relative sizes of the different stages inside the erythrocyte. Adapted from (Bannister, *et al.*, 2000).

After invasion the parasite flattens into a thin, discoidal flat or cup shaped ring-form of the trophozoite stage (Fig. 1.3) (Bannister, *et al.*, 2000). The nucleus shape varies from sausage-like to a disc. The parasite survives intracellularly by ingesting host cell cytoplasm through a circular structure named the cytostome, a small, dense ring at the surface of the parasite (Bannister, *et al.*, 2000; Fujioka and Aikawa, 1999). Approximately 70 to 80% of the erythrocyte haemoglobin is degraded during

schizogony and converted to inert brown haemozoin crystals that accumulate within the pigment vacuoles (Fujioka and Aikawa, 1999).

The ring changes shape to a more rounded or irregular trophozoite due to the area around the parasite and the PVM increasing in size and extending finger-like projections into the erythrocyte. During this stage, the PVM hugs the parasite membrane closely but also forms membranous stacks penetrating deep into the erythrocyte thought to be in some way involved in trafficking of substances (Bannister, *et al.*, 2000). The parasite gradually alters the erythrocyte membrane producing five structural modifications: knobs, caveolae, caveola-vesicle complexes, cytoplasmic clefts and electron dense materials (Fujioka and Aikawa, 1999; Torii and Aikawa, 1998). Knobs are involved in adhesion of parasitized erythrocytes during sequestration and cytoadherence, the pathogenic obstruction of vasculature (Bannister, *et al.*, 2000).

During the schizont stage the single parasite undergoes repetitive nuclear division. The nucleus divides about four times in an endomitotic manner to produce 16 nuclei. This division is accompanied by numerous cytoplasmic changes including proliferation of rough endoplasmic reticulum, multiplication of mitochondria and plastids and accumulation of large lipid vacuoles (Bannister, *et al.*, 2000). A series of centres of merozoite formation is then observed beginning with the apical organelles in a highly ordered sequence of steps. A constriction ring then separates each merozoite, which is released to invade new erythrocytes.

1.3) Pathogenic basis and clinical features of malaria.

The pathogenic process of malaria occurs during the asexual erythrocytic cycle where clinical symptoms including nausea, headache and chills, accompany the characteristic bouts of fever. In the untreated patient, severe complications of *P. falciparum* malaria infections manifest as cerebral malaria (with convulsions and coma), anaemia, hypoglycaemia, renal failure, metabolic acidosis, severe liver failure, respiratory distress, circulatory collapse, raised intracranial pressure and non-cardiac pulmonary oedema often resulting in death (Marsh, 1999; Mendis and Carter, 1995; Ramasamy, 1998; White, 1998).

The abovementioned morbid conditions associated with the disease are ascribed to various host-parasite interactions. Several substances are released by the intra-

erythrocytic parasite including malarial mitogens, toxic proteins, prostaglandins (D_2, E_2 and $F_{2\alpha}$) and polar lipids, particularly glycophosphatidyl inositol (GPI; covalently bound to the merozoite surface antigens, MSA1 and MSA2) that has been speculatively named malaria toxins. The proposed main consequence of these bio-active molecules is to direct the systemic release of several pro-inflammatory cytokines, in particular tumour necrosis factor α (TNFα), interferon γ (IFN γ), interleukin 1 (IL1), IL6 and IL10 (Clarke and Schofield, 2000; Hommel, 1997; Miller, *et al.*, 2002; Miller, *et al.*, 1994; Ramasamy, 1998; White, 1998). A major role for the pro-inflammatory cytokines is to generate the inducible form of nitric oxide synthase (iNOS) to produce a continuous release of the nitric oxide (NO) mediator (Clarke and Schofield, 2000). This could be clinically important in some reversible cerebral symptoms, immunosuppression and weight loss seen in malaria (Clarke and Schofield, 2000; Miller, *et al.*, 2002).

One of the most distinctive pathophysiological characteristics, which evolved for the survival of *P. falciparum*, is to effectively modify the ultrastructure of the erythrocyte in which it resides. Several parasite-specific proteins are produced on the surface to allow the unique, parasite-induced ability of infected erythrocytes to adhere to post-capillary microvascular endothelial cells (a process termed sequestration), and the *in vitro* adherence to uninfected erythrocytes (rosetting), other infected erythrocytes (auto-agglutination or clumping) and to platelets, monocytes and lymphocytes (Berendt, *et al.*, 1994; Miller, *et al.*, 2002; White, 1998). Collectively, cytoadherence enables the parasite to avoid destruction by the reticulo-endothelial system in the spleen and has as consequence a decreased peripheral parasitaemia with only ring stage parasites visible in peripheral blood. Sequestration occurs non-homogeneously in various organs including heart, lung, brain, liver, kidney, subcutaneous tissues and placenta resulting in considerable obstruction to tissue perfusion with systemic or local production of cytokines as mentioned above (Miller, *et al.*, 2002).

The most potent modification is the expression of a family of *P. falciparum*-infected erythrocyte membrane proteins called PfEMP1 (200-300 kDa) that have been implicated as the main mediators of adhesion in both sequestration and possibly also in rosetting (Miller, *et al.*, 2002). This protein is encoded by a large and diverse family of *var* genes that are involved in clonal antigenic variation (Miller, *et al.*, 2002). PfEMP1 is continually exposed to the host immune system but *P. falciparum* escapes the onslaught of an antibody response by varying the antigenic and adhesive characteristics

of the protein family. Antigenic variation is achieved by switching the gene transcription of the 50-150 *var* genes, leading to an altered antigenic phenotype of the VAR protein family (PfEMP1 family) with an associated change in the cytoadherent properties of a clonal population. This spontaneous variation is estimated to occur at a rate of more than 2% per generation. This predicts that a clonally derived organism will rapidly become phenotypically heterogeneous by eliminating a phenotype recognised by the host immune system and predominantly expressing another (Berendt, *et al.*, 1994; Hommel, 1997; Miller, *et al.*, 2002; Ramasamy, 1998). A second, and by far the largest, multigene family called the *rif* genes (and their subfamily *stevor*) encode the RIFINS, a group of polymorphic antigens that may be co-expressed on the surface of an erythrocyte infected with a single clonal parasite (Cooke, *et al.*, 2000; Craig and Scherf, 2001).

Other parasite-specific proteins and host-cell receptors on the vascular endothelium that are thought to mediate cytoadherence are summarised in Fig. 1.4.

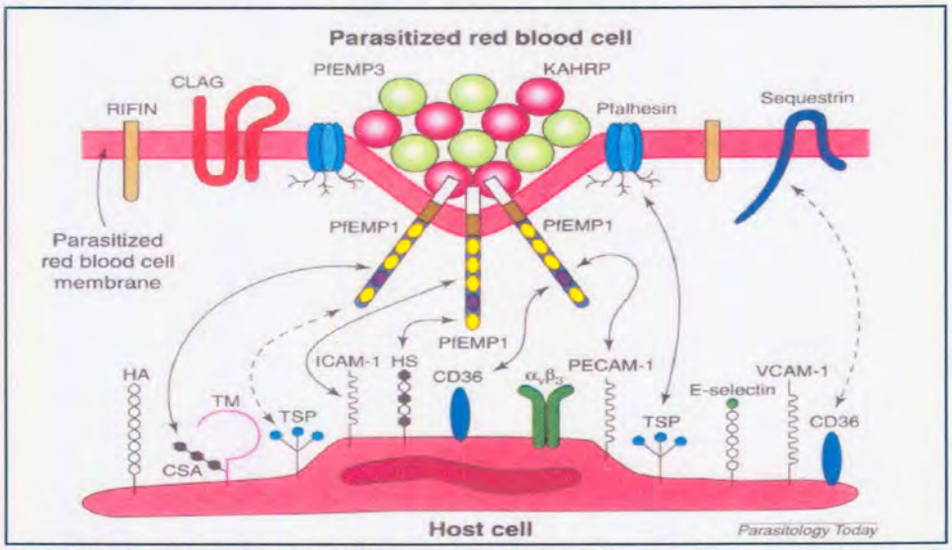


Figure 1.4: Schematic representation of the interaction at the cytoadhesive interface between a *P. falciparum* infected erythrocyte and the host vascular endothelium. CLAG: cytoadherence-linked asexual gene; CR1: complement receptor I; CSA: chondroitin sulphate A; HA: hyaluronic acid; HS: heparan sulphate; ICAM: intercellular adhesion molecule I; KAHRP: knob-associated His-rich protein; PECAM-1: platelet-endothelial cell adhesion molecule I; Pfalhesin: modified form of band 3; TM: thrombomodulin; TSP: thrombospondin; VCAM-1: vascular cell adhesion molecule 1. Adapted from (Cooke, *et al.*, 2000)

1.4) Global control strategies of malaria.

In order to control a disease with such severe pathological manifestations as malaria, several innovative strategies must be devised. The most effective public health tools in dealing with malaria have been antiparasitic drugs and insecticides (Collins and Paskewitz, 1995). Unfortunately, resistant parasites and vectors compromise the

efficacy of these strategies. The global malaria eradication campaign was launched in 1955 and resulted in the subsequent elimination of malaria from Europe, most of the Asian regions of Russia, the United States of America and most of the Caribbean (Krogstad, 1996). However, the eradication programme has been abandoned since it became obvious that malaria cannot be dealt with as a single and uniform worldwide problem susceptible to one global control strategy. There are a variety of factors to take into account including 1) the biological and cultural characteristics of the population, 2) the intensity and periodicity of malaria transmission, 3) the species of malaria and their sensitivity to antimalarial drugs, 4) the mosquito vector, 5) the presence of social and ecological changes and 6) the characteristics of the existing health services (Phillips, 2001). Attention is now focused on creating alternative strategies consisting of chemotherapy/prophylaxis, vector control and vaccines. Several international partnerships have been created with the ultimate goal to eradicate the disease in the foreseeable future. These include the Roll Back Malaria campaign (www.rbm.who.int), the Medicines for Malaria Venture (www.mmv.org, based on the expertise of the Special Program for Research and Training in Tropical Diseases of the World Health Organisation) and the Multilateral Initiative against Malaria (mim.nih.gov). The continued and sustainable improvement and discovery of antimalarials through focussed research and rational drug design programmes are essential in combating the disease.

1.4.1) Chemotherapy and –prophylaxis.

Chemotherapy remains the only viable and practical tool to control falciparum malaria. Fig. 1.5 lists the current antimalarials, their mechanisms of action and limits of use and highlights the structures of some of these compounds.

1.4.1.1) Blood schizontocytes

1.4.1.1.1) Quinoline-containing drugs

Malaria has been treated successfully for several hundred years with quinine, the active reagent in the bark of the South American Cinchona tree (*Cinchuona ledgeriana*) (WHO, 1996). Modern cinchona alkaloids or quinoline-type drugs include the type I 4-amino-quinolines (chloroquine, amodiaquine, mecaprime and sontaquine) and type II aryl-amino-alcohols (mefloquine and halofantrine) (Olliaro, 2001).

Antimalarial	Proposed function	Limits of efficacy
Blood schizontocytes		
Quinolines		
Quinine	Inhibition of haem detoxification	Compliance, safety, resistance
Chloroquine		Resistance
Amodiaquine		Safety, resistance
Mefloquine		Cost, psychosis, compliance, (resistance)
Halofantrine		Safety, resistance, cost, cardiotoxicity
Artemisinins		
Dihydroartemisinin		
Arteether	Oxidative stress, free radical formation	Compliance, safety, cost, recrudescence
Artemether		
Artesunate		
Nucleic acid metabolism antagonists		
Type I antifolates	Dihydropteroate synthetase (DHPS) inhibition	
Sulphadoxine		Resistance
Dapsone		
Type II antifolates		
Pyrimethamine	Dihydrofolate reductase (DHFR) inhibition	
Proguanil		Resistance
Trimethoprim		
Naphtoquinones (Atovaquone)	Inhibition of mitochondrial function	Compliance
Antimicrobials		
Tetracycline, doxycycline, clindamycin and Azithromycin	Inhibition of prokaryote-like protein synthesis in apicoplast	Limited usefulness

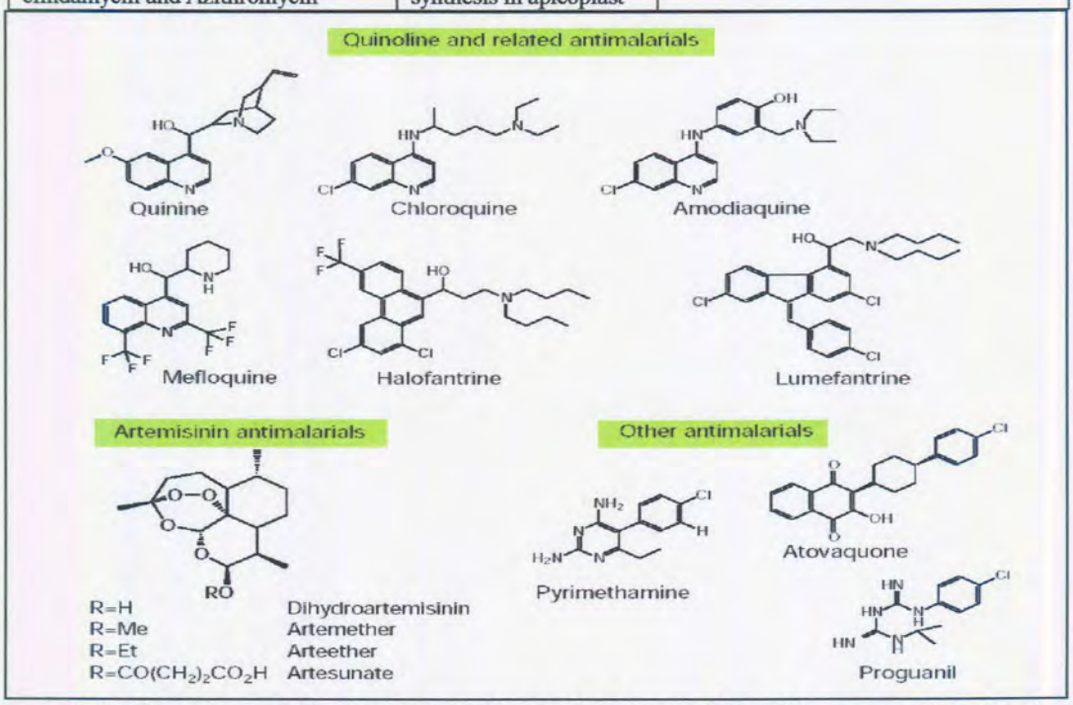


Figure 1.5: Overview of the current antimalarial drugs. The table summarises the drugs, their targets and limits of efficacy while structures of some of the drugs are given in the bottom panel. Compiled from (Hyde, 2002; Macreadie, *et al.*, 2000; Milhous and Kyle, 1998; Olliaro, 2001; Ridley, 2002; Winstanley, 2000).

The quinoline-type antimalarials are thought to disrupt or prevent effective formation of haemozoin by binding to haem through π-π stacking of their planar aromatic structures (Ridley, 2002). Normally, haemoglobin is digested in the lysosomal food vacuole into usable short peptides and the released haem is sequestered as ferriprotoporphyrin IX (FP). Free FP is toxic to the parasite and is polymerised to haemozoin, which is released during rupture of the schizonts (Olliaro, 2001).

In the early 1960s there were two simultaneous reports of chloroquine resistant *P. falciparum* from South America and Southeast Asia (Collins and Paskewitz, 1995). Resistance has gradually spread throughout the globe, with Africa the last to succumb in 1978 (Krogstad, 1996). Cross-resistance is also apparent between chloroquine, mefloquine and halofantrine (Hyde, 2002). Currently, it is accepted that resistance to chloroquine is probably conferred by multiple gene mutations. A chloroquine resistance determinant on chromosome 7 has been named *pfcrt* (chloroquine resistance transporter), a protein that likely functions as a transporter on the food vacuolar membrane. A single mutation in this gene is linked to resistance (Hyde, 2002; Olliaro, 2001; Wellems and Plowe, 2001). A second gene encoding a P-glycoprotein, *mdr1*, has been linked to the altered susceptibility of the parasites to mefloquine and halofantrine (Ridley, 2002; Wellems, *et al.*, 1990).

1.4.1.1.2) Artemisinins.

Several synthetic derivatives of artemisinin, the active ingredient of the Chinese herb qinghao (*Artemisia annua*, wormwood), have been increasingly used in the past two decades (Balint, 2001; Ridley, 2002). These drugs achieve the quickest and highest reduction rates of parasitaemia per cycle compared to any other known drug and also act on gametocytes (Olliaro, 2001). Although the mechanisms of action is not wholly understood, the prevailing hypothesis is that oxidoreductive cleavage of the intact peroxide of the drug occurs in the food vacuole probably through interaction with iron as Fe(II). This generates fatal free radicals, which in turn would alkylate biomolecules (Olliaro, 2001; Ridley, 2002). To date, no clinical cases of resistance against the artemisinin derivatives have been documented. However, monotherapy leads to recrudescence probably due to suboptimal levels of the drug (Hyde, 2002).

1.4.1.2) Antifolate Antimalarials.

The antifolates are a group of antimalarials designed to limit the availability of folate in the parasite, which prevents the synthesis of thymidylate and disrupts DNA synthesis. The two major enzymatic targets in the folate pathway are dihydrofolate reductase (DHFR) and dihydropteroate synthetase (DHPS) (Hyde, 2002). The most significant antifolates to treat malaria are the combination of pyrimethamine and sulphadoxine (Fansidar), synergistically blocking both actions of DHFR and DHPS (Ridley, 2002). Unfortunately, resistance developed rapidly when this combination was used

extensively. At this stage, antifolate resistance is common in South America, Southeast Asia and also in Africa (Birkholtz, *et al.*, 1998a; Krogstad, 1996). Resistance has been mapped to single or combined point mutations in the active sites of both enzymes leading to decreased capability of these competitive inhibitors to effectively bind and inhibit the enzymes (Hyde, 2002; Sibley, *et al.*, 2001; Yuthavong, 2002).

1.4.1.3) Combination therapy against malaria.

After the unsuccessful use of single compound therapies, malariologists followed protocols for the treatment of cancers and other infectious diseases including those caused by *Mycobacterium tuberculosis* and the human immunodeficiency virus. Combinations of different drugs that synergistically act to treat a disease have the added effect of preventing the rapid development of resistance. Certain old antimalarials now have new uses when used in synergistic combinations. These include Atovaquone-proguanil (Malarone), lumefantrin-arthemeter, mefloquine-artemisinin, chlorproguanil-dapsone (LapDap) and other artemisinin combination therapies (Winstanley, 2000; Winstanley, *et al.*, 2002).

1.4.1.4) New Antimalarial drugs and potential drug targets.

Very few successful new drugs are in clinical trails or are undergoing preclinical development. One compounding factor is the lack of economic incentive for the pharmaceutical industry to develop new antimalarials. However, chloroquine resistance reversers such as verapamil (Ca-antagonist) are receiving attention (Milhous and Kyle, 1998). Pyronaridine may also have utility for multiresistant malaria (Winstanley, 2000). A series of compounds developed by the Walter Reed Army Institute of Research such as WR238605 (Etaquine) have promising activity (Milhous and Kyle, 1998). However, rational drug design strategies based on genomic, bioinformatic and protein homology modelling coupled with virtual docking of inhibitors will aid in identification of more effective drugs with known targets.

1.4.2) Strategies for vector control.

The failure of the malaria eradication program and emergence of insecticide-resistant mosquitoes has shifted the focus to other strategies for the control of human-vector contact. Chemicals continue to be the mainstay of mosquito control and broadly fall into five groups: Petroleum oils (larvicide), copper acetoarsenite (larvicide), the pyrethroids, organochlorines including dichlorodiphenyltrichloroethane (DDT), the

organophosphates and carbamates (Phillips, 2001). Bednets or curtains impregnated with pyrethroid insecticides such as permethrin or deltamethrin, have been shown to reduce the number of mosquito bites by ≥ 95% and the prevalence of infection by ≥ 40-45% in low intensity transmission areas (Krogstad, 1996). Biological control agents such as larvivorous fish and bacterial endospore toxins from *Bacillus sphaericus* and *Bacillus thuringiensis* have been investigated (Phillips, 2001). Although successful, it is too expensive for common application. The most promising, although still speculative development, has been the possible introduction of transgenic vectors that are unable to host the malaria parasite, into natural populations (Collins and Paskewitz, 1995; Phillips, 2001).

1.4.3) Malaria Vaccines.

The feasibility of malaria vaccines was first indicated in the early 1970s when complete protection against malaria was obtained in humans after vaccination with irradiation-attenuated sporozoites (Clyde, *et al.*, 1975). Further support for a vaccine strategy includes acquired clinical immunity after long-term exposure. Immunity is obtained after repeated infections during which the host is exposed to numerous parasite-induced erythrocyte surface proteins which elicit variant-specific antibodies to inhibit cytoadherence of an increased number of different parasite variants (Richie and Saul, 2002). This immunity is dependent on continuous exposure and after a short break in antigenic stimulation, individuals lose the essential diversity in their immune response. Unfortunately, after more than 20 years since the first efforts, a reliable, effective vaccine is still not available.

Vaccine design is complicated by the many stages of the *Plasmodium* parasite as well as by antigenic variations and polymorphisms in a whole range of *P. falciparum* proteins. Two main strategies in malaria vaccine development have been investigated: 1) blocking infection and transmission at the pre-erythrocytic (sporozoite and hepatic) stage and 2) inhibition of the blood stage (asexual erythrocytic stage) where both parasite growth and pathogenesis is blocked (Phillips, 2001; Rogers and Hoffman, 1999). The first step in any of these strategies is the identification of suitable parasite antigens to be targeted. Several candidate antigens of the different parasite stages have been identified, and are listed in Table 1.1. Secondly, antigens from different stages should be combined (multistage vaccine). Thirdly, several antigens from a single stage (multivalent) are needed and lastly, the vaccines need to be simple and elicit the correct

type of immune response (Richie and Saul, 2002). An effective vaccine is envisaged as a cocktail of a multistage, -valent, -immune response vaccine.

The most successful asexual blood-stage vaccine to date has been the multicomponent Spf66 developed by Patarroyo and colleagues (Amador and Patarroyo, 1996). One epitope is derived from the repeat domain (PNANP) of the circumsporozoite antigen (CS), two peptides (35.1 and 55.1) are based on unidentified *P. falciparum* molecules and a third peptide (83.1) corresponds to the highly conserved residues 45-53 of the major surface antigen (MSA-1). Results of phase III trials between different study groups were contradictory but gave a good indication of what could be achieved with a multicomponent vaccine. Two other vaccines that have recently undergone field trials are RTS,S (a pre-erythrocytic stage vaccine) as well as a combination of three blood-stage antigens (parts from MSA-1, MSA-2 and RESA, see Table 1.1) showed promise but did not induce long-lasting protection (Anders and Saul, 2000; Richie and Saul, 2002; Rogers and Hoffman, 1999). Other vaccines in or close to clinical trials include TRAP/SSP2, $MSA1_{42}$-3D7 and AMA-1-3D7 (Richie and Saul, 2002).

Alternative technologies such as recombinant DNA vaccines may allow the combination of many DNA sequences of many different antigens (multivalent) from one or more stages (multistage) to broaden the immune response. One such vaccine, NYVAC-Pf7 has been developed and employs a live attenuated vaccinia virus vector containing genes for seven malaria antigens (CS, SSP2 [TRAP], LSA-1, MSA-1, AMA-1, SERA, and Pfs25; see Table 1.1 for abbreviations)(Facer and Tanner, 1997). There are however safety concerns with DNA as immunogen such as possible integration into the host genome or germ line cells, elicitation of anti-DNA antibodies, autoimmunity and a possible induction of tolerance.

Table 1.1: Synopsis of the candidate *Plasmodium* antigens for malaria vaccine development. The table was compiled from the following references: (Anders and Saul, 2000; Doolan and Hoffman, 1997; Facer and Tanner, 1997; Kwiatkowski and Marsh, 1997; Nussenzweig and Long, 1994; Richie and Saul, 2002; Rogers and Hoffman, 1999).

Vaccine target stage	Antigen	Abbreviation	Function
Pre-erythrocytic stage	Circumsporozoite antigen	CS	Invasion of erythrocyte
	Thrombospondin related anonymous protein/Sporozoite surface protein	TRAP/SSP 2	Development in hepatocyte
	Liver stage antigen	LSA-1 and -3	Development in hepatocyte
	Pfs 16	N/A	?
	Sporozoite surface threonine and aspartic acid rich protein	STARP	?
	Sporozoite and liver stage antigen	SALSA	Hepatocyte invasion
	PfEXP-1	N/A	Development in hepatocyte
Asexual blood-stage	Merozoite surface antigen	MSA-1/2/3/4 or MSP	Red cell invasion
Merozoite proteins	Ring infected surface antigen	RESA (pf155)	Stabilisation of erythrocyte spectrin
	Erythrocyte binding antigen	EBA-175	Red cell invasion
	Apical membrane antigen	AMA-1	Rhoptry organelle/ Red cell invasion
	Rhoptry associated protein	RAP-1/2 (QF3)	Rhoptry organelle/ Red cell invasion
	Glutamate-rich protein	GLURP	Red cell invasion
Infected erythrocyte membrane proteins	*P. falciparum* infected erythrocyte membrane protein	PfEMP-1	Cytoadherence
	P. falciparum histidine rich protein	PfHRP-2	Development in erythrocyte
	Rosettin	N/A	Rosetting process
	Ag332 (Pf332)	N/A	Asexual growth
	RIFINS	N/A	Cytoadherence ?
Soluble antigens	*P. falciparum* histidine rich protein	PfHRP-2	Malarial toxin
	Ag-7	N/A	Malarial toxin
	Ag-2/ Serine repeat protein	SERA(SERP)	Malarial toxin
	Glycosylphosphatidylinositol	GPI	Red cell rapture, malarial toxin
Sexual transmission stage	Post-fertilisation antigen	Pfs 25	Sexual development
	Post-fertilisation antigen	Pfs 28	Sexual development
	Gametocyte antigen	Pfg 230	Sexual development
	Gametocyte antigen	Pfg 48/45	Sexual development

This thesis is based on the premise that a better understanding of the metabolic processes of the *Plasmodium* parasite would assist in the identification of suitable therapeutic targets, which could be exploited in the development of a novel intervention strategy. Once a target protein has been identified and characterised it can be evaluated in terms of the design of novel antimalarials. The following sections deal with the metabolic processes of *P. falciparum*, and primarily focus on polyamine biosynthesis as a possible antimalarial targets.

1.5) Biochemistry and metabolic pathways of *Plasmodium.*

Defining the metabolome of an organism allows deeper insights into its biology. The integration of the metabolome in sequential reactions in metabolic pathways and networks represents a convenient method of defining specific processes at the biological level. However, rather than specifying each metabolite and enzyme by name, Alberts *et al.* (1983) devised a dramatic way of illustrating the plethora of metabolic reactions by representing these components with dots and lines, respectively. This approach has been successfully applied by A. Fairlamb in metabolic pathway analysis of trypanosomes to indicate unique features of this parasite (Fairlamb, 1989; 2002). Using this elegant concept, I have produced a metabolome for *P. falciparum* with the aid of the *Plasmodium* electronic metabolic pathway database (sites.huji.ac.il/malaria) as well as the standard chart of metabolic pathways (IUBMB, 21st edition, designed by Dr. D. E. Nicholson). Global comparisons are made with this new metabolic pathway analysis of *P. falciparum* compared to those constructed for a typical mammalian cell and *Trypanosoma brucei* (Fig. 1.6).

Using the global metabolome comparison, it is apparent that there is an absence of many core metabolic pathways in both parasites. Intra-erythrocytic *P. falciparum* does not store glycogen and therefore needs a constant supply of glucose. Glucose metabolism through anaerobic glycolysis increases as much as 50–100 fold over uninfected erythrocytes (Sherman, 1998b). The general pathway of carbohydrate metabolism is very similar to that of other eukaryotic cells, but as the parasites are microaerophilic homolactate fermenters, almost all the glucose passes through the Emden-Meyerhoff-Parnas pathway to produce lactate and two ATP molecules (Sherman, 1998b). A unique glucose transporter was recently characterised for *P. falciparum* (Woodrow, *et al.*, 1999).

The parasite does have a functional pentose phosphate pathway (Sherman, 1998b) but erythrocytic stage *P. falciparum* lacks a functional citric acid cycle (Lang-Unnasch and Murphy, 1998). Despite this, there is substantial evidence of a functional electron transport chain due to the presence of cytochromes a, a_3 , b, c, and c_1 (Lang-Unnasch and Murphy, 1998). Furthermore, a plant-like alternative electron transfer occurs directly from ubiquinone to oxygen in *P. falciparum* (Fig. 1.6)(Lang-Unnasch and Murphy, 1998).

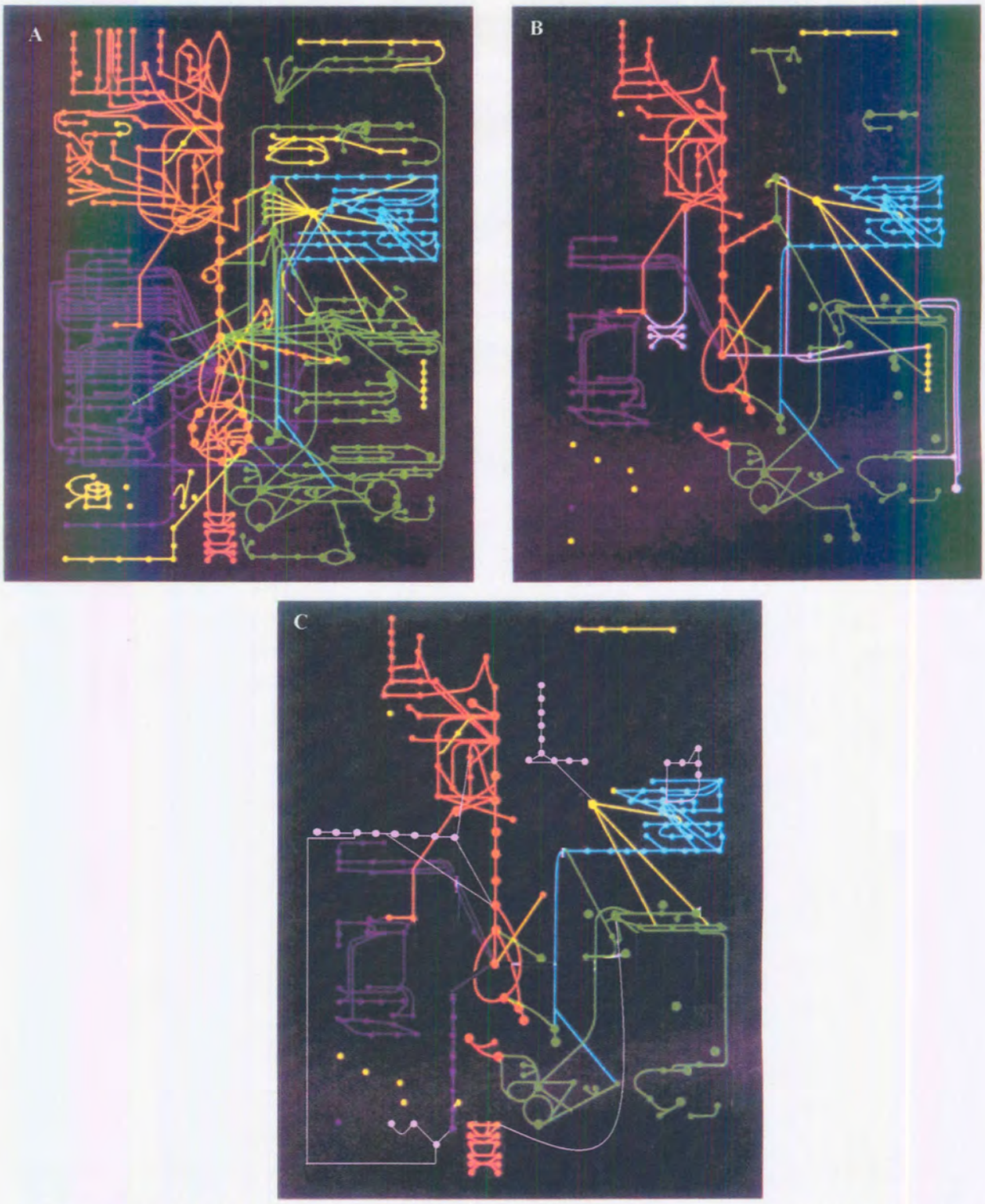


Figure 1.6: Global comparison of metabolomes. (A) A typical mammalian cell; (B) *T. brucei* and in (C) the newly constructed metabolome of *P. falciparum*. Orange indicate carbohydrate metabolism and electron transport, purple for fatty acid and sterol metabolism, green for amino acids and related metabolites, blue for purine, pyrimidine and nucleic acid biosynthesis, yellow for co-enzyme biosynthesis. Unique pathways in the parasites are shown in pink. (A) and (B) taken from (Fairlamb, 2002).

Virtually all protozoan parasites lack the enzymes for *de novo* synthesis of purines and they scavenge these nucleotides from their hosts. The enzymes needed for the salvage pathway are therefore potential targets for interruption of parasite growth and development. The major purine source for the *P. falciparum* is hypoxanthine (the base of inosine nucleosides), which is normally generated from ATP in the erythrocyte (Fairlamb, 2002; Sherman, 1979). Once in the parasite, hypoxanthine is converted to both adenosine and guanosine via IMP by the key enzyme, hypoxanthine-guanosine phosphoribosyltransferase (HGPRT) (Fairlamb, 2002). Furthermore, the conversion of IMP to AMP involves the enzymes adenylosuccinate lyase and –synthase, which appear to be unique to the parasite, and thus ideal antiparasitic targets.

Plasmodia depend solely on *de novo* synthesis for their supply of pyrimidine nucleotides. Human erythrocytes do not contain significant levels of pyrimidines and this pathway is another potential antimalarial target. The role of the electron transport system in *P. falciparum* is for the *de novo* synthesis of pyrimidines via the enzymatic link dihydroorotate dehydrogenase (DHODH) (Berens, *et al.*, 1995; Olliaro and Yuthavong, 1998).

De novo synthesised folate is essential in *P. falciparum* whereas mammalian cells salvage these co-factors from the diet (Lang-Unnasch and Murphy, 1998). The two bi-functional enzymes in this pathway, dihydropteroate synthetase-pyrophosphokinase (DHPS-PPPK) and dihydrofolate reductase-thymidylate synthase (DHFR-TS) have been thoroughly studied and are the targets of the antifolates as was discussed in section 1.4.1.2.

The primary source for amino acids in *P. falciparum* is haemoglobin, which is degraded first into peptides and, after export from the food vacuole to the cytoplasm, into amino acids by exopeptidases. Haemoglobin is a poor source of Met, Ile, Cys, Glu/Gln, Leu and His. Malaria parasites need to obtain these amino acids from the external environment (Lang-Unnasch and Murphy, 1998; Rosenthal and Meshnick, 1998). The only amino acids that the parasite synthesises are Glu, Asp, Ala and Leu (Rosenthal and Meshnick, 1998). Malarial proteases are required for processing of parasite proteins, degradation of haemoglobin and erythrocyte invasion and rupture. One cysteine (analogous to cathepsin L) and two aspartic proteases (analogous to cathepsin D) have

been implicated in haemoglobin degradation. Two serine proteases, three cysteine, five aspartic and two aminopeptidase proteases have been described for *P. falciparum* (McKerrow, *et al.*, 1993).

The shikimate pathway, conserved in plants, algae, fungi and bacteria have recently been described in the Apicomplexa (McConkey, 1999). Seven enzymes in this pathway generate chorismate, which is metabolised to para-amino benzoic acid, ubiquinone and the aromatic amino acids. The absence of this pathway in mammals makes it an attractive target for antimalarials. The distinguishing plastid organelle of the Apicomplexa also has compartmentalised metabolic pathways that are targets for the selective inhibition of malaria. Type II fatty acid synthesis and isoprenoid synthesis are present in this organelle in *P. falciparum* and seems to indicate a bacterial or plant-like origin to these pathways (Wilson, 2002).

Most cells contain high concentrations of two important classes of metabolites: the polyamines (discussed in the following sections) and the thiols. *P. falciparum* possess a functional redox system including glutathione reductase, peroxidase and synthase (Luersen, *et al.*, 2000; Meierjohann, *et al.*, 2002).

A fundamental reason for studying the biochemistry of malaria parasites is to uncover metabolic differences between the host and the parasite that might be exploited in the design of drugs specifically targeted to *Plasmodium*, as well as to provide an understanding of the mode of action of existing antimalarials. Table 1.2 summarises the identified potential targets for intervening with *P. falciparum* metabolism. Promising new targets for antimalarial therapy include glycolytic enzymes (lactate dehydrogenase), nucleotide biosynthesis inhibitors, plastid DNA replication and transcription, Type II fatty acid biosynthesis, non-mevalonate isoprenyl biosynthesis, the shikimate pathway, plasmepsin aspartic proteases and falcipain cysteine proteases, nutrient transporters (including glucose and Na^+/H^+ transporters), mitochondrial enzymes and cell signalling pathways (Ridley, 2002).

Table 1.2: Summary of the major metabolic target proteins in *P. falciparum*. This table was compiled from references: (Olliaro and Yuthavong, 1998; Ridley, 2002; Subbayya, *et al.*, 1997)

Target protein	Function/Metabolic pathway
Hexokinase (HK)	First enzyme in glycolysis
Aldolase	Glycolysis
Lactate dehydrogenase (LDH)	Last enzyme of glycolysis
Triosephosphate isomerase (TIM/TPI)	Glycolysis
Hypoxanthine-guanosine phosphoribosyl transferase (HGPRT)	Key enzyme in purine salvage pathway.
Adenylosuccinate lyase	Formation of AMP from IMP
Adenylosuccinate synthase (ASS)	Formation of AMP from IMP
Inosine monophosphate dehydrogenase (IMPDH)	Formation of GMP from IMP
Guanosine monophosphate synthase	Formation of GMP from IMP
Dihydroorotase	*De novo* pyrimidine synthesis
Dihydroorotase dehydrogenase (DHODH)	*De novo* pyrimidine synthesis
Carbamoyl phosphate synthetase	*De novo* pyrimidine synthesis
Orotoate phosphoribosyl transferase	*De novo* pyrimidine synthesis
Orotidine-5′ -monophosphate decarboxylase	*De novo* pyrimidine synthesis
Dihydrofolate reductase-thymidylate synthase (DHFR-TS)	*De novo* pyrimidine synthesis/ Folate synthesis
Dihydropteroate synthetase-pyrophosphokinase (DHPS-PPPK)	Folate synthesis
Haeme polymerase ?	Degradation of haemoglobin
Phosphatidyl choline cytidyl transferase (CTP)	Phospholipid synthesis
Sphingomyelin synthase	Tubovesicular membrane (TVM) synthesis
Falcipain (Cysteine protease)	Degradation of haemoglobin
Plasmepsin I (Aspartic protease)	Degradation of haemoglobin
Plasmepsin II (Aspartic protease)	Degradation of haemoglobin
Serine protease	Involved in MSA-1 processing
Ribonucleotide reductase	Cell Cycle regulated enzyme
DNA topoisomerase I and II	DNA replication
RNA polymerase (plastid encoded)	Transcription
DNA polymerase α	DNA replication
LSU rRNA (plastid encoded)	Protein synthesis
Tubulin (α, β subunits)	Cytoskeleton
Na+/H+ transporter	Nutrient transport
Hexose transporter	Nutrient transport
Ornithine decarboxylase (ODC)	Polyamine biosynthesis
S-adenosylmethionine decarboxylase	Polyamine biosynthesis

1.6) Polyamine metabolism.

'...Among the various classes of organic substances, there is perhaps none of which, from an early period, chemists have so constantly endeavoured to attain a general conception as the group of compounds, which have received the name of organic bases.'

A. W. Hofmann (1850)

The polyamines are a discrete class of biological substances that are low-molecular aliphatic, nonprotein, nitrogenous bases found ubiquitously in all cells, both pro- and eukaryotic, at relatively high concentrations. This large group of compounds contains four most important molecules, the diamines putrescine (1,4-diaminopropane) and cadaverine (prokaryotic 1,5-diaminopropane) and the triamine and tetraamine putrescine derivatives spermidine (N-(3-aminopropyl)-1,4-diaminobutane) and spermine (N,N^1-bis(3-aminopropyl)-1,4-butanediamine) (Fig. 1.7 indicates the structures of these compounds) (Cohen, 1998). Their polycationic character gives them a high affinity for acidic cell constituents including nucleic acids, acidic proteins and phospholipids and as such are involved in a myriad of essential cellular processes integral to macromolecular synthesis, cell proliferation and differentiation (Algranati and Goldemberg, 1989; Heby, 1989). Rapidly growing cells contain higher levels of polyamines and their biosynthetic enzymes than quiescent cells.

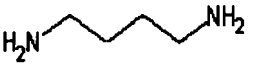
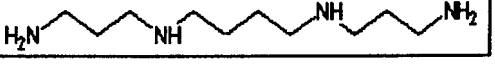
Names	Structure	Shorthand representation
Putrescine	$H_2NCH_2CH_2CH_2CH_2NH_2$	
Cadaverine	$H_2NCH_2CH_2CH_2CH_2CH_2NH_2$	
Spermidine	$H_2NCH_2CH_2CH_2CH_2NHCH_2CH_2CH_2NH_2$	
Spermine	$H_2NCH_2CH_2CH_2HNCH_2CH_2CH_2CH_2NHCH_2CH_2CH_2NH_2$	

Figure 1.7: Structures of the most important polyamines. Adapted from (Cohen, 1998).

The polyamines are formed from ornithine as by-products of the urea cycle in eukaryotes. Arginine is converted to ornithine and this is the only route for the synthesis of putrescine via the enzyme ornithine decarboxylase (ODC) (shown in Fig. 1.8). Putrescine is the precursor for all the other polyamines and is converted to spermidine by the addition of an aminopropyl group donated by methionine after decarboxylation of S-adenosylmethionine (AdoMet) by S-adenosylmethionine decarboxylase (AdoMetDC). Spermine is formed in a similar aminopropyl transfer reaction. The interconversion of the different polyamines occurs via the transfer of aminopropyl groups and oxidation (Cohen, 1998; Pegg and McCann, 1982; Tabor and Tabor, 1984b). An active transport system for the uptake of polyamines is present in all cells. The system is distinct from amino acid transport and is stimulated by polyamine depletion (Pegg and McCann, 1982). Polyamine biosynthesis is also linked with the citric acid

cycle, methionine recycling and adenine nucleotide salvage pathways (Heby, 1985). Prokaryotes are capable of synthesising putrescine from agmatine using arginine as starting point.

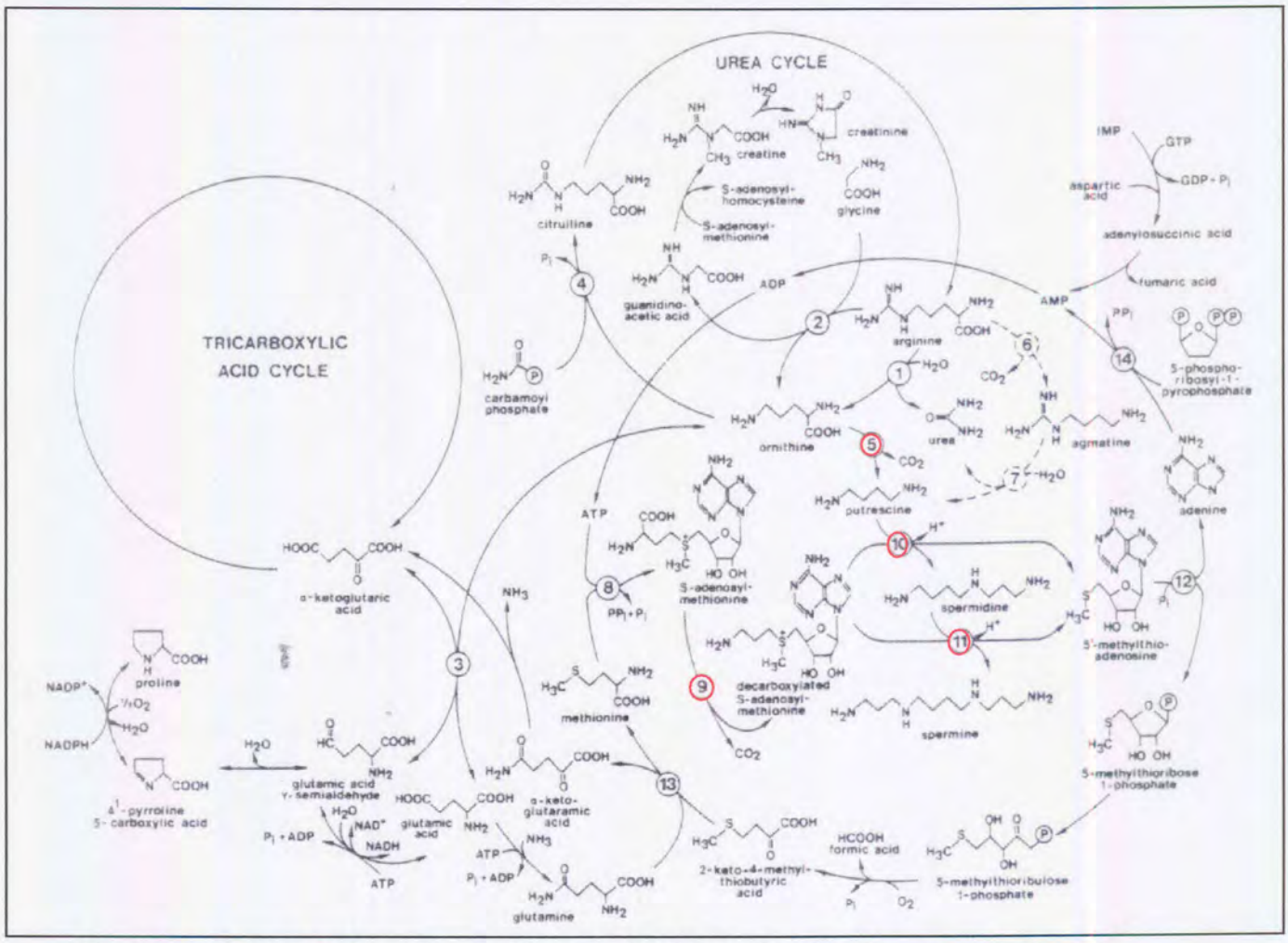


Figure 1.8: General pathway for the biosynthesis of the polyamines in pro- and eukaryotes and its linkage to the urea cycle, tricarboxylic acid cycle and methionine and adenine salvage pathways. 1) Arginase, 2) transamidinase, 3) ornithine aminotransferase, 4) ornithine carbamoyltransferase, 5) ornithine decarboxylase, 6) arginine decarboxylase, 7) agmatine ureohydrolase, 8) S-adenosylmethionine synthetase, 9) S-adenosylmethionine decarboxylase, 10) spermidine synthase, 11) spermine synthase, 12) 5'-methylthioadenosine phosphorylase, 13) glutamine transaminase, 14) adenine phosphoribosyltransferase. Adapted from (Heby, 1985). The four main enzymes in polyamine biosynthesis are highlighted in red.

Due to the critical role of polyamines in important cellular functions, multiple pathways such as biosynthesis, catabolism, uptake, interconversion and excretion tightly regulate their intracellular concentration (Kahana, *et al.*, 2002). The two decarboxylase enzymes, ODC and AdoMetDC, are most important since polyamine synthesis is totally dependent on their activities and they function as the rate-limiting enzymes in the pathway (Heby, 1985). Both these enzymes are highly regulated and are subject to feedback control by cellular polyamines. ODC is one of the most rapidly degraded

proteins in eukaryotic cells with a half-life between 5-35 min (Hayashi, 1989; Hayashi and Murakami, 1995; Heby, 1985). Its degradation mechanism is furthermore unique as it is ubiquitin independent. Instead, ODC is marked for degradation by the 26 S proteasome by interaction with a unique polyamine-induced inhibitory protein named antizyme (Hayashi and Canellakis, 1989; Hayashi and Murakami, 1995). Antizyme is synthesised in a polyamine-dependent translational frameshift mechanism (Hayashi and Canellakis, 1989). It then binds to monomeric ODC, inactivates the protein and induces a conformational change to expose a C-terminal area rich in Pro, Glu, Ser and Thr (PEST-rich region) and targets the protein for rapid degradation (Mamroud-Kidron, *et al.*, 1994; Rogers, *et al.*, 1986). Antizyme is itself subject to regulation by another inhibitory protein, antizyme inhibitor (Hayashi, *et al.*, 1996). AdoMetDC in turn is dependent on putrescine for both activation and enzyme activity (Tabor and Tabor, 1984a).

1.6.1) Polyamine metabolism in parasitic protozoa.

The complete characterisation of the biochemical properties of the enzymes involved in the synthesis of the polyamines in parasites was based on the assumption that parasite-specific properties found would contribute towards the antiparasitic potential of polyamine inhibition (Cohen, 1998). The synthetic pathway in trypanosomatids (*Trypanosoma* and *Leishmania*) is well characterised but very simple compared to those of mammals. These organisms synthesise putrescine from the decarboxylation of ornithine via ODC and then produces spermidine with the action of AdoMetDC and spermidine synthase (Fig. 1.9)(Cohen, 1998). However, no spermine is synthesised in these trypanosomatids and back conversion via the interconversion pathway has not been demonstrated (Müller, *et al.*, 2001). One unique and novel deviation is the synthesis of a bis-glutathionyl derivative of spermidine called trypanothione with trypanothione reductase the essential enzyme in keeping the metabolite in its reduced form. This co-factor is required to maintain the redox balance in the cells (Fairlamb, 1989; Müller, *et al.*, 2001). *Leishmania* contains distinct transporters for putrescine and spermidine, a property that is absent from *T. brucei* (Cohen, 1998). *T. cruzi* does not synthesise putrescine but scavenges this polyamine from the host (Müller, *et al.*, 2001).

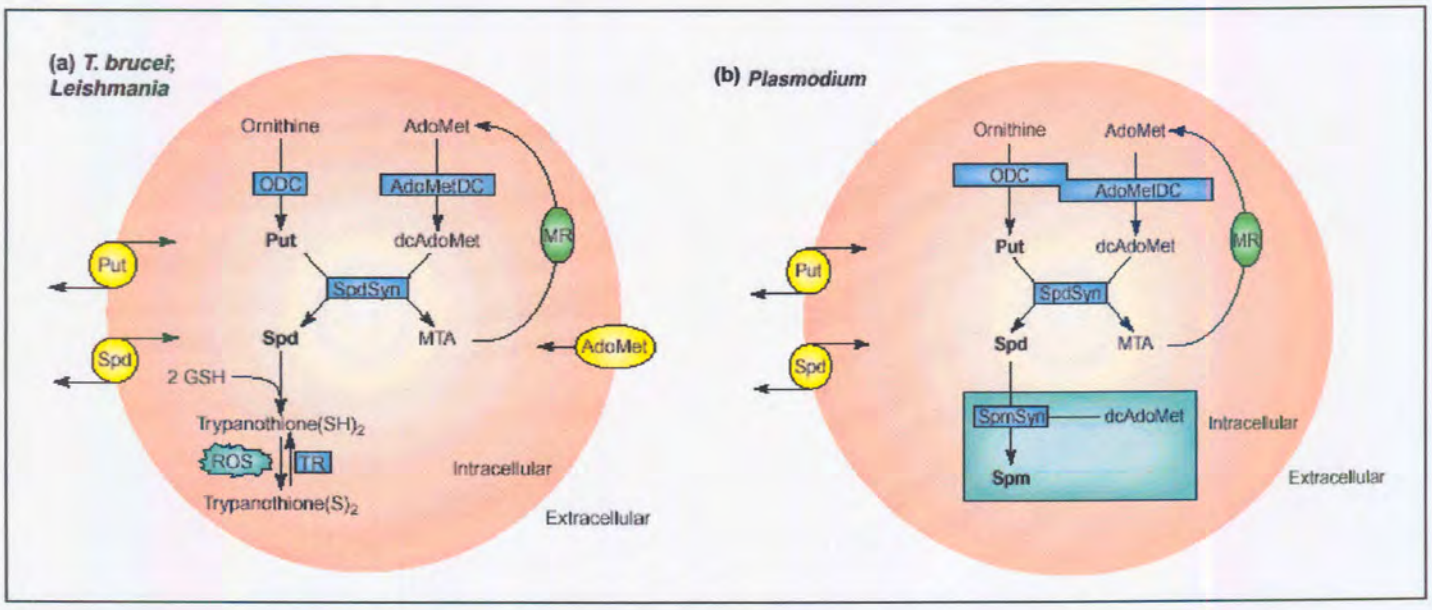


Figure 1.9: Polyamine metabolism in parasitic protozoa. (a) *T. brucei* and *Leishmania* and in (b) *Plasmodium*. ODC: ornithine decarboxylase, AdoMetDC: S-adenosylmethionine decarboxylase, SpdSyn: spermidine synthase, SpmSyn: spermine synthase, TR: trypanothione reductase, Put: putrescine, Spd: spermidine, AdoMet: S-adenosylmethionine, Spm: spermine, ROS: reactive oxygen species, MTA: methylthioadenosine, MR: methionine recycling pathway. Adapted from (Müller, *et al.*, 2001).

At the start of this study in 1999, information regarding polyamine biosynthesis in *P. falciparum* lacked the research status of the other parasitic protozoa and was limited primarily to *in vitro* monitoring of polyamine levels. Polyamine biosynthesis and the activity of the biosynthetic enzymes peaked at the early trophozoite stage (Assaraf, *et al.*, 1984). Putrescine levels were increased 95-fold in parasite-infected red blood cells followed with a 51 times increase in spermidine levels and only a slight increase in spermine levels. *P. falciparum* contains 9 pmol putrescine, 33 pmol spermidine and 8 pmol spermine per 10^6 parasitized erythrocytes (Tabor and Tabor, 1984b). The parasites are capable of transporting putrescine and spermidine but it seems that the interconversion pathway is absent (Müller, *et al.*, 2001). The activities of ODC, AdoMetDC and ornithine aminotransferase (OAT) have been described in the malarial pathway and ODC and OAT have been partially isolated (Assaraf, *et al.*, 1988; Rathaur and Walter, 1987).

It was shown in 2000 that both the rate-limiting decarboxylase activities of ODC and AdoMetDC are found on a single polypeptide in *P. falciparum* (Müller, *et al.*, 2000). This is an extraordinary deviation from the mammalian enzyme organisation and could

lead to novel selective inhibition of polyamine biosynthesis in *P. falciparum*. Spermidine synthase has also been isolated (Müller, *et al.*, 2001).

1.6.2) Polyamine metabolism as an antiprotozoal target.

The observation that the polyamines and their biosynthetic enzymes are present in increased concentrations in proliferating cells, including malignant cells and parasitic organisms, makes inhibition of polyamine biosynthesis a logical approach for chemotherapy and antiparasitic drugs (Janne and Alhonen-Hongisto, 1989b). The antiproliferative effects of polyamine depletion have been exploited in cancer treatment and chemoprevention, in psoriasis and in infectious diseases caused by viruses, bacteria, fungi and parasitic protozoa (Janne and Alhonen-Hongisto, 1989b). Interest in polyamine metabolism in the parasitic protozoa (including *Plasmodia*, *Leishmania* and *Trypanosoma*) increased when it was shown that inhibition of polyamine biosynthesis in these highly proliferative organisms could be curative of disease (Janne and Alhonen-Hongisto, 1989a).

Polyamine analogues have been used with much success and their functions vary between ODC inhibition and replacing the normal polyamine pool in cells and thereby preventing their action as growth factors (Cohen, 1998). This discussion will focus on the enzyme-based inhibition of polyamine biosynthesis.

Due to their roles as the rate-limiting enzymes in polyamine biosynthesis, ODC and AdoMetDC serve as obvious targets for chemical intervention aimed at depletion of the polyamine pool. Inhibition of ODC leads to decreased putrescine and spermidine levels (Janne and Alhonen-Hongisto, 1989a) whereas inhibition of AdoMetDC leads to large increases in putrescine and a decline in spermidine and spermine due to the absence of dAdoMet (Pegg, 1989b). The effect on cell proliferation is usually cytostatic rather than cytocidal. A successful and useful ODC inhibitor should therefore be able to overcome the homeostatic maintenance of cellular polyamine pools, which is controlled by uptake, synthesis, interconversion and catabolic reactions (Cohen, 1998). This probably contributes to reasons for the apparent inability of ODC inhibitors to be clinically useful against human cancers with a high tumour burden (Janne and Alhonen-Hongisto, 1989a).

A landmark in ODC inhibition is the synthesis of the substrate analogue, DL-α-difluoromethyl ornithine (DFMO), which is enzymatically decarboxylated and generates an irreversible alkylation of the enzyme at or near the active site to cause a rapid loss of enzyme activity (Cohen, 1998; Heby, 1985; Janne and Alhonen-Hongisto, 1989a). The only physiologically meaningful inhibitor of AdoMetDC was synthesised as an antileukemic agent, methylglyoxal bis(guanylhydrazone) (MGBG) (Cohen, 1998; Janne, *et al.*, 1985).

Regarding clinical antiprotozoal activity of ODC inhibitors, the results are just the opposite of that found in cancer. DFMO exert a dramatic and rapid therapeutic effect in African sleeping sickness caused by *T. brucei gambiense* (Janne and Alhonen-Hongisto, 1989a). In 1990, DFMO was approved as the first drug in forty years for the treatment of this disease (McCann and Pegg, 1992). Reasons for the sensitivity of these parasites to DFMO include an increased half-life of ODC (> 7 hrs), inability of these parasites to take up polyamines, the presence of trypanothione and altered levels of AdoMetDC (Janne and Alhonen-Hongisto, 1989a; Pegg, 1989b; Wrenger, *et al.*, 2001). The polyamine pathway, with reference to ODC, has also been fully validated as a chemotherapeutic target in *T. brucei* since genetic elimination of ODC rendered the parasites auxotrophic for polyamines (Wang, 1997). This is in contrast to *L. donovani* ODC knock-out mutants where the auxotrophy could be circumvented by the addition of a series of polyamines (Jiang, 1999). Prolonged use of DFMO resulted in the development of resistant cell lines. In *L. donovani* the ODC gene is amplified 10-20 fold in the resistant strains (Hanson, *et al.*, 1992) and this therefore needs to be considered in the design of new inhibitors of ODC. In addition the multiple, intricate mechanisms regulating ODC activity, polyamine transport and the development of resistance need careful consideration. AdoMetDC inhibitors were also found to be trypanocidal and include MGBG and some derivatives called Berenil and Pentamidine that were found to be curative of infections in mice (Cohen, 1998).

Inhibiting polyamine synthesis in *P. falciparum* with DFMO blocks transformation of trophozoites to mature schizonts and decreases the levels of putrescine and spermidine but not spermine (Assaraf, *et al.*, 1984; Assaraf, *et al.*, 1986). The synthesis of spermine in *P. falciparum* is questionable and the absence of an interconversion pathway might explain the small effect seen on spermine levels (Müller, *et al.*, 2001). The use of DFMO prevented parasite proliferation and decreased the parasitaemia *in vivo* by 42-

70%, indicating the possibility of preventing disease (McCann and Pegg, 1992). Both erythrocytic as well as exoerythrocytic schizogony seem to be limited in *P. berghei in vivo* (Bitoni, *et al.*, 1987; Hollingdale, *et al.*, 1985). The inhibitory concentration of DFMO appeared quite high at about 5-6 mM (Cohen, 1998). *In vitro* growth inhibition is also observed when AdoMetDC is inhibited with MGBG (Wright, *et al.*, 1991). Inhibition of polyamine synthesis inhibited synthesis of selected proteins, caused a partial inhibition of RNA synthesis and completely blocked DNA synthesis. This led to the suggestion that the polyamines are required for the synthesis of malaria parasite proteins involved in DNA synthesis (Assaraf, *et al.*, 1987a). Prolonged treatment of the parasites with DFMO induced massive accumulation of pigment followed by death (Assaraf, *et al.*, 1987b). Unfortunately, DFMO inhibition of ODC seems to be circumvented by an exogenous supply of polyamines (Assaraf, *et al.*, 1987b) and explains the *in vivo* refractoriness of DFMO (Singh, *et al.*, 1997). The intracellular polyamine pool is simply maintained by decreased polyamine excretion and/or increased polyamine uptake. Furthermore, DFMO might not be transported well across the three membrane layers of intraerythrocytic *P. falciparum* (Müller, *et al.*, 2001).

A combination of ODC and polyamine transporter inhibitors like N^1,N^4-bis(7-chloroquinoline-4-yl)butane-1,4-diamine [BCBD], appears to be required in order to deplete *P. falciparum* of intracellular polyamines (Singh, *et al.*, 1997). The uptake of bisbenzyl analogues occurs via a system quite distinct from the polyamine transport system and, in combination with DFMO, cured mice infected with *P. berghei* (Bitonti, *et al.*, 1989). Unsaturated putrescine derivatives and novel tetraamines furthermore showed antimalarial activity against *P. falciparum in vitro* (Edwards, *et al.*, 1991; Slater, *et al.*, 1998).

Despite the success of polyamine inhibition in curing sleeping sickness, this strategy is currently still ineffective against *Leishmania* and *Plasmodium*. The main reasons for this might be the poor uptake of drugs by these parasites, the low affinity of the current inhibitors to the target proteins as well as the circumvention of the inhibition by polyamine uptake systems. However, there are still several reasons for optimism that polyamine metabolism of parasitic protozoa can be exploited by chemotherapy including: 1) Biochemical elucidation of the metabolic pathway in the parasites indicated marked differences with the mammalian host pathways, in particular the bifunctional decarboxylase in *P. falciparum* and the presence of trypanothione in the

Trypanosomatids; 2) Structural and inhibitor studies suggest that there is scope for obtaining more potent inhibitors than DFMO (Müller, *et al.*, 2001); 3) Stage-specific inhibition seems to lead to more pronounced effects including the sensitivity of *P. falciparum* asexual stages to DFMO (Müller, *et al.*, 2001); 4) A myriad of compounds have been developed in the anticancer field that can be investigated as antiparasite agents and 5) the combination of polyamine synthesis inhibitors with polyamine transport inhibitors or polyamine analogues seems promising in curing parasitic protozoal infections.

1.7) Research Objectives.

This study was aimed at the biochemical characterisation of the complete polyamine metabolic pathway of *P. falciparum* in order to elucidate differences between the parasite and its human host that can be exploited in the design of novel antimalarial therapies. The approach followed was to identify and characterise the genes and recombinantly expressed proteins of the key enzymes, ODC and AdoMetDC, followed by various structure-functional characterisations of these proteins to infer parasite-specific properties. Ultimately, these properties are used in the design of novel, potential antimalarial agents.

While this study was in progress, the characterisation of the bifunctional form of the *P. falciparum* ODC and AdoMetDC was revealed (PfAdoMetDC/ODC) (Müller, *et al.*, 2000). A collaboration with the research group of Prof. R. D. Walter (Biochemistry Department, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany) was established in 2000. The methodology and results of the study presented in this thesis are therefore a combination of work performed in South Africa and during research visits to Germany.

Chapter 2 describes the sequences and characterisation of the *Odc* and *Adometdc* genes obtained in South Africa by a novel protocol in which an uncloned, amplified cDNA library was constructed for *P. falciparum* and used for rapid amplification of cDNA ends (RACE) procedures. The full-length combined *Adometdc/Odc* gene was amplified based on the published sequence (Müller, *et al.*, 2000) using our in-house cDNA library. Molecular genetic analyses were independently performed on the obtained nucleotide sequence.

Chapter 3 presents results of the recombinant expression and determination of biochemical properties of the enzymes in different expression systems. One expression system for the recombinant monofunctional enzymes was developed and optimised in South Africa. Further studies of the monofunctional as well as bifunctional proteins were conducted using the expression systems described in Müller *et al.* (2000). Bioinformatic analyses of the deduced amino acid sequence and inherent properties of the bifunctional PfAdoMetDC/ODC are also described.

Chapter 4 describes determinations of the inherent structure-functional relationship of the enzymes using mutagenesis techniques to determine loss-of-function in mutants lacking certain parasite-specific structural elements. The results were obtained during a research visit to Germany.

In Chapter 5, the bioinformatics field is explored to obtain a three-dimensional model of the ODC protein structure and to infer some functional properties from this structure.

Finally, the identification of potential novel inhibitors of PfODC enzyme function is presented in Chapter 6.

Chapter 7 presents a concluding discussion of the knowledge obtained during this study, specifically highlighting structure and functional characteristics of the bifunctional PfAdoMetDC/ODC.

CHAPTER 2

Molecular genetic analyses of *P. falciparum* S-adenosylmethionine decarboxylase (*Adometdc*), ornithine decarboxylase (*Odc*) and the bifunctional *Adometdc/Odc* genes.

2.1) INTRODUCTION.

2.1.1) Genetic analyses of *Plasmodia.*

Genetic analyses of the *Plasmodium* genome has been tedious mostly because the organism has various non-classical genetic features to allow it to adapt to changing environments during complex and varied host interactions: *P. falciparum* is the organism with the highest A+T content in its genome (82%), *in vitro* genetic crossing is extremely difficult, the chromosomes are not condensed during meiosis and show size variation between wild type isolates and the organism is capable of extreme genetic diversity through extensive genome plasticity involving loss of dispensable functions under non-selective conditions (Coppel and Black, 1998; Foote and Kemp, 1989; Frontali, 1994; Kemp, *et al.*, 1990; Lanzer, *et al.*, 1995; Scherf, *et al.*, 1999). The parasite genome is haploid during its life cycle in the vertebrate host with the only diploid stage being the zygote produced by fertilisation of gametes in the mosquito (Coppel and Black, 1998; White and Kilbey, 1996). The 14 chromosomes range in size from 630 kbp (chromosome 1) to 3.4 megabases (chromosome 14) with the estimated haploid genome size, 2-3 x 10^7 bp (Foote and Kemp, 1989; Gardner, *et al.*, 1998; Scherf, *et al.*, 1999). Other DNA elements in the cytosol include a multi-copy 6 kb DNA associated with the mitochondria and a maternally inherited 35 kb circular DNA thought to be present within the plastid-like organelle (Su and Wellems, 1998).

One of the most interesting features of the *P. falciparum* genome is its base content. The distribution of the base composition is not uniform with genes having a lower A+T content (60-70%) whereas introns and flanking regions tend to be very A+T-rich (>85%) (Coppel and Black, 1998; Scherf, *et al.*, 1999; Su and Wellems, 1998). One of the most striking consequences is the highly biased codon usage, with I>II>III the order of usage of G/C among the three codon positions (Coppel and Black, 1998; Hyde and Holloway, 1993; Scherf, *et al.*, 1999). The compositional effects on genome function

are still largely unknown and a biologically feasible concept to explain the differences found in DNA composition between different organisms is still lacking. The flexibility of the genetic code might tolerate a certain degree of deviation without introducing significant selective pressure on the translation apparatus (Scherf, *et al.*, 1999). Speculations for the underlying mechanism of extreme DNA composition bias include mutations in the DNA polymerase that might alter the way by which a DNA mismatch is repaired. In *P. falciparum*, DNA repair might be biased and preferentially use the base A or T (Scherf, *et al.*, 1999).

Eukaryotic genes transcribed by RNA polymerase II frequently contain transcription signals upstream of the RNA initiation site such as the highly conserved TATA-box (consensus sequence 5′ TATAt/aAt/a 3′)(Adams, *et al.*, 1993; Latchman, 1995). Other more unique upstream promoter elements include the CAAT-box and a GC-rich area that bind specific transcriptional regulators (Adams, *et al.*, 1993). Initial attempts to identify *Plasmodial* promoters have been hampered by the instability of the noncoding *Plasmodial* sequences as a consequence of the high A+T content (Lanzer, *et al.*, 1993). The identification of transcriptional signals upstream of the transcription start sites is usually limited to structural sequence analyses although recent successes with transfection of *P. falciparum* genes have facilitated functional analyses of transcriptional processes (Horrocks, *et al.*, 1998). Indications are that the *P. falciparum* promoters conform to the characteristic bipartite structure of eukaryotic promoters. The upstream regions of *Plasmodial* genes contain at least the minimal necessary elements for transcription initiation including TATA-boxes, CAAT and SP1 binding elements (Lanzer, *et al.*, 1993; Su and Wellems, 1998).

All of the *Plasmodium* mRNAs investigated to date are co-linear. *Trans*-splicing that occurs in other protozoa has not been observed. Post-transcriptional processing however occurs through *cis*-splicing and polyadenylation (Lanzer, *et al.*, 1993). Most *Plasmodial* genes have none or only one intron with splice sites conforming to those of eukaryotes (Gardner, *et al.*, 1998; Lanzer, *et al.*, 1993). The presence of unusually long 5′-untranslated regions (5'-UTR) appears to be a common feature of many malarial mRNA transcripts, with sizes between 500 and 1000 nt in length (Coppel and Black, 1998; Su and Wellems, 1998). Analyses of the sequences around the start codon indicated a consensus of AAAA/ATG (Coppel and Black, 1998). All *Plasmodial* 3′-end polyadenylation sites of transcripts are flanked by conserved sequences, the

polyadenylation signal (AATAA) is found immediately or up to 30 nt upstream of the poly-A tail downstream of which, within 50 nt, a diffuse G+T-rich sequence is found. This makes up the following conserved structural organisation:

AATAA---$(N)_{0-23}$---A--------$(N)_{6-54}$-----G+T-rich------

compared with the mammalian consensus sequence:

AATAAA---$(N)_{10-23}$---A--------$(N)_{0-10}$----G+T-rich------

The G+T-rich sequence might therefore provide the signals for RNA processing and addition of the poly(A) tail in *Plasmodium* (Lanzer, *et al.*, 1993). An average poly-A tail is 100-200 nt long and the 3′-UTRs generally contain <400 nt (Coppel and Black, 1998; Su and Wellems, 1998).

2.1.2) Molecular characteristics of the *Adometdc* and *Odc* genes.

The observed increase in polyamine levels and the activities of the biosynthetic enzymes under conditions that promote cell growth has spurred investigations of the molecular aspects of the relevant genes ever since the isolation of the cDNA for *Odc* in 1984 (Kahana and Nathans, 1984). *Odc* mRNA levels are largely increased in mouse kidney cells in response to androgens, possibly due to the stabilisation of the message. Furthermore, DFMO-resistant cell lines are usually caused by amplification of the functional *Odc* gene (Heby and Persson, 1990). Analyses of the *Odc* gene from murine species indicated that a family of up to twelve *Odc* related genes are present in these genomes (Kahana, 1989). It is uncertain however if these are functional genes or pseudogenes with simply a high homology to *Odc* (Cohen, 1998). In contrast, the haploid human genome has only two *Odc* genes on two different chromosomes, which may both transcribe mRNA (Heby and Persson, 1990). The genomic sequences of the two mammalian *Odcs* contain 11 introns interspersing 12 exons, in contrast with the yeast and *Trypanosoma Odcs*, which lack any introns (Heby and Persson, 1990). Furthermore, the mammalian 5′-untranscribed flanking region contains several promotor/enchancer elements, a TATA-box, a putative CAAT-box, a GC-box and a consensus sequence of a cAMP responsive element making it as strong as the Rous sarcoma virus long terminal repeat promoter (Heby and Persson, 1990). The *Adometdc* gene of murine genomes is also present as multiple copies although it is uncertain if these are part of a multigene family or pseudogenes (Heby and Persson, 1990; Pulkka, *et al.*, 1991). The human *Adometdc* gene consists of 9 exons with 8 introns transcribing one active and one pseudogene (Nishimura, *et al.*, 1998). The 5′-untranscribed flanking region of *Adometdc* is similar between mammalian sequences with 60% identity in the

proximal 500 nt. Several features that characterise housekeeping gene promoters were identified including a TATA-box, Sp1, CAAT, AP2 and EB1 elements but no GC-rich octamer (Nishimura, *et al.*, 1998; Pulkka, *et al.*, 1991).

Both *Adometdc* and *Odc* mammalian transcripts belong to a very small class of mRNAs containing a large leader sequence (>200 nt 5′-UTR). Two mRNA species have been described as *Odc* transcripts in various organisms, a predominant transcript of 2.0-2.4 kbp and a minor 2.6-2.7 kbp band (Heby and Persson, 1990). The size heterogeneity has been ascribed to differing lengths of the 3′-UTR due to the presence of two polyadenylation signals (AATAAA) in the gene. The biological relevance of such alternative polyadenylation sites is presently unclear (Kahana, 1989). Analyses of mammalian *Odc* mRNAs have revealed an extremely GC-rich 5′-UTR that may form secondary structures including stable hairpin loops with very high free energies of stabilisation (Heby and Persson, 1990). It is postulated that these structures might be involved in the translational regulation of ODC and that polyamines might influence the secondary structure formation in a feedback regulation mechanism (Cohen, 1998). The 5′-UTR also contains a second small open reading frame (ORF) but it is unlikely that functional protein is produced from this area. Both the 3′- and 5′-UTRs has been further implicated in the regulation of *Odc* through co-operative interactions to inhibit translation (Lorenzi and Scheffler, 1997).

The *Adometdc* transcript in mammals has two species, a 1.7-2.4 kbp band and a 3.0-3.6 kbp band also produced through the use of alternative polyadenylation signals 1272 bp apart (Heby and Persson, 1990; Pulkka, *et al.*, 1991). The mammalian 5′-UTR is ~320 nt long and appears to also be GC-rich with strong secondary structures of –289 kJ/mol (Shantz, *et al.*, 1994). A definite second small ORF is present and has been shown to translate a peptide thought to be involved in ribosome stalling as a mechanism of polyamine mediated regulation of AdoMetDC translation (Raney, *et al.*, 2000). The synthesis of mammalian *Adometdc* mRNA is also inhibited by high levels of spermidine and spermine (Heby and Persson, 1990; Shantz, *et al.*, 1994).

2.1.3) The molecular characterisation of genes and their mRNAs.

The characterisation of proteins in terms of their genes and RNA has been made possible by the application of molecular biology techniques. The methods available for characterising cDNA include screening a cDNA library and reverse-transcriptase PCR

(RT-PCR). The latter is an extremely sensitive and rapid method and is therefore an attractive approach for obtaining a specific cDNA sequence (Karcher, 1995; Kidd and Ruano, 1995). However, in some instances, RT-PCR alone is ineffective in obtaining complete cDNAs when only a limited consensus protein sequence is available. Rapid Amplification of cDNA Ends (RACE) is a PCR-based technique that facilitates the cloning of cDNA 5′- and 3′-ends after amplification with a gene-specific primer (GSP) based on a consensus sequence, and a 5′- or 3′-end anchor primer, respectively (Frohman, 1993; Karcher, 1995; Scheafer, 1995). A modified RACE procedure has been developed by Chenchik *et al.* (1996). The principle of the method is the ligation of a double-stranded (ds) adaptor oligonucleotide to both ends of ds cDNA. This is followed by 5′- or 3′-RACE with a gene-specific primer and adaptor-specific primer as indicated in Fig. 2.1. First-strand cDNA is generated with a differential-display Poly-T (DD-Poly-T) primer. This primer contains a 3′-clamp to position the primer at the edge of the poly-A tail of the mRNA and is normally linked to an anchor sequence at its 5′-end to facilitate RACE.

The ds adaptor-ligated RACE method has however, one disadvantage. Ligation of the ds adaptor to both ends of the ds cDNA allows background amplification of cDNA derived from all RNA species (including poly-A$^-$ RNA) if total RNA is used. This is due to self-priming and priming by short RNA or DNA fragments present in the RNA sample. A derivation of this method employs a novel suppression PCR technology to inhibit amplification of cDNA originating from poly-A$^-$ RNA (Lukyanov, *et al.*, 1997). Double-strand cDNA is synthesised and the ds adaptor is then ligated only by its long strand since none of the oligonucleotides are phosphorylated. After the ends of the duplexes are filled in, the cDNA derived from mRNA is then effectively flanked by the DD-Poly-T/Anchor sequence and the adaptor sequence whereas cDNA derived from the poly-A$^-$ fraction will be flanked only by adaptor sequences (Fig. 2.1) (Lukyanov, *et al.*, 1997). The mRNA-derived cDNA is then amplified with the DD-Poly-T primer and the adaptor primer. The amplification of the cDNA derived from poly-A$^-$ RNA is inhibited by the suppression PCR effect since the long inverted sequences of the adaptor at both ends of the cDNAs will hybridise to form a panhandle structure (Fig. 2.1). In addition, the adaptor sequence is GC-rich, and has a high negative ΔG and a Tm at least 10°C higher than the Tm of the primers. Intramolecular annealing events are thus facilitated and effectively enrich the cDNA population derived from mRNA to establish an uncloned and representative cDNA library.

In this chapter an in-house designed kit, employing the PCR suppression technology to create an uncloned cDNA library as described above, was used to obtain the full-length *P. falciparum Odc* and *Adometdc* cDNAs. This chapter is a continuation of work done for a M. Sc. on the cloning of the *Odc* cDNA by L. Birkholtz at the University of Pretoria as part of a larger project using a molecular approach in elucidating the complete polyamine metabolic pathway in *P. falciparum* (Birkholtz, 1998c). The M. Sc. study described the synthesis and control-analyses of the in-house designed uncloned cDNA library as well as the optimisation of RACE procedures using this library. Furthermore, part of the *P. falciparum Odc* cDNA was identified and cloned in 1997-1998 (Genbank accession number AF139900).

Our strategy was continued in 1999 at the start of this PhD study and included the investigation of the *Adometdc* cDNA. However, early in 2000, Müller *et al.* (2000) published results showing that both the *Odc* and *Adometdc* cDNAs are derived from a single transcript (*PfAdometdc/Odc*). A collaboration was established with Prof. R. D. Walter of Germany in July 2000. Therefore, this chapter is divided into two parts: Part one (p. 37) describes the characterisation in South Africa of partial *Odc* and *Adometdc* cDNAs with RACE using the uncloned cDNA library as well as the identification of a large mRNA transcript using the *Odc* cDNA as probe. Part two (p. 57) describes the amplification of the full-length *Odc* and *Adometdc* cDNAs as well as amplification of the full-length *PfAdometdc/Odc* cDNA based on the published sequence data from Müller *et al.* (2000)(Genbank accession number AF0934833). Novel sequence analyses of the full-length *PfAdometdc/Odc* gene and cDNA are also described.

Some of the results obtained in this Chapter have been presented as papers at the Young Scientist Symposium, 18th International Conference of the IUBMB, Birmingham, UK (2000) (Birkholtz, *et al.*, 2000a), 2nd Gauteng Region Annual Biochemistry Symposium, Pretoria, South Africa (2000) (Birkholtz, 2000b) and at the BioY2K Combined Millennium Meeting, Grahamstown, South Africa (2000) (Birkholtz, 2000c). Furthermore, the results were presented as posters at six international conferences (Birkholtz, 2000c; Birkholtz, *et al.*, 2000a; Birkholtz and Louw, 1998b; 1999a; 1999b; 2000d) and one national conference (Birkholtz, 2000d).

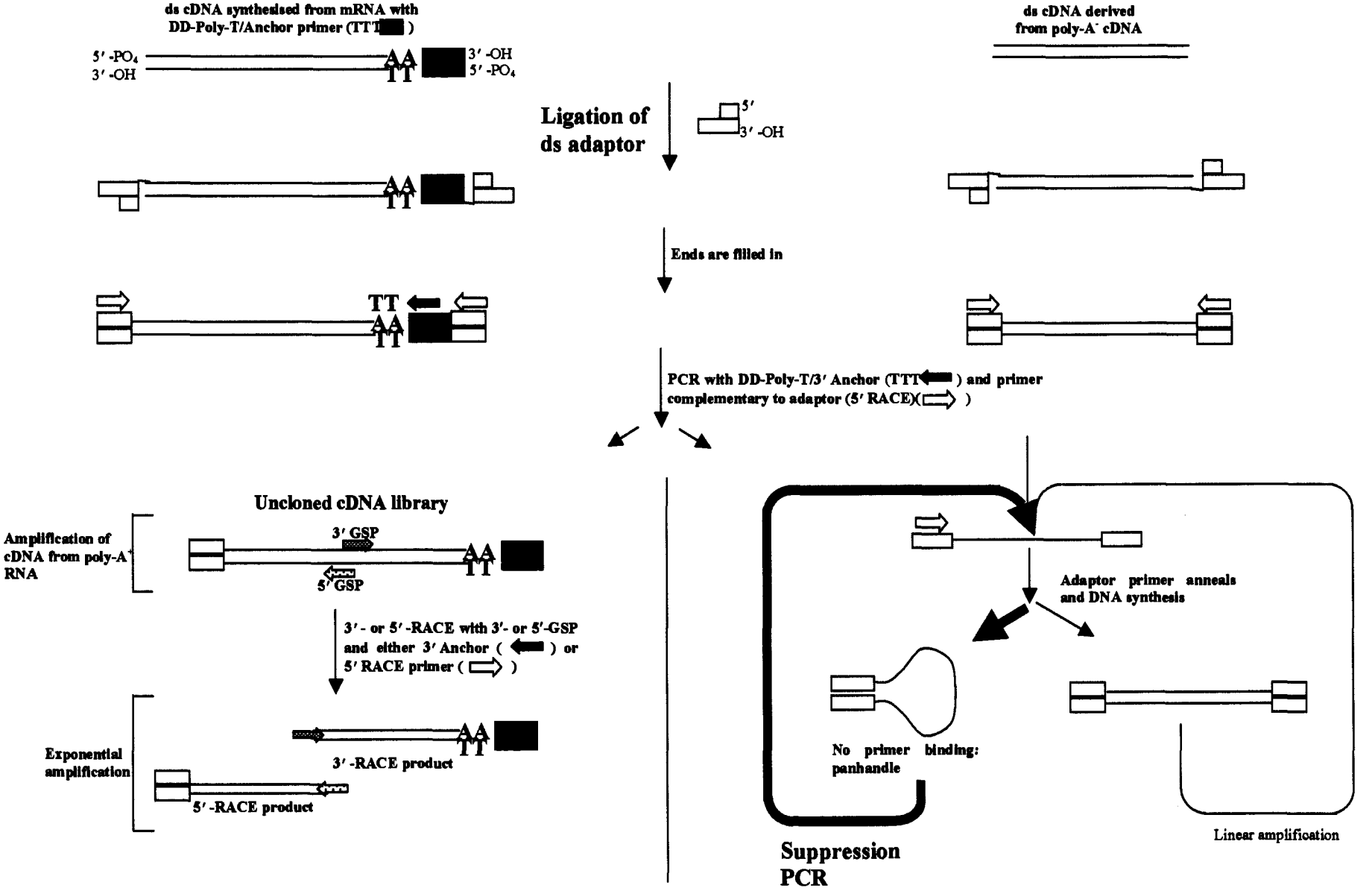


Figure 2.1: RACE protocols with double strand adaptor-ligated cDNA and suppression PCR. The adaptor is indicated with the open boxes. The DD-Poly-T and attached Anchor region are indicated with the solid boxes. The ligation to only the long strand of the adaptor is indicated with the small loop. The suppression PCR effect is shown on the right hand side of the scheme and the formation of the uncloned cDNA library to the left. Adapted from (Lukyanov, *et al.*, 1997).

PART I: Identification of *Adometdc* and *Odc* cDNAs with RACE.

2.2) MATERIALS AND METHODS.

2.2.1) *In vitro* cultivation of malaria parasites.

A continuous culture of the isolate, *P. falciparum* UP1 (PfUP1), was maintained under conditions that support intracellular development of the parasites (Trager, 1994; Trager and Jensen, 1976). 500 µl PfUP1-infected erythrocytes, cryopreserved at -70°C, were thawed rapidly at 37°C and 200 µl of a 12% (w/v) NaCl solution was added. After incubation for 10-20 sec, 1.8 ml of 1.6% (w/v) NaCl was added and the parasites collected by centrifugation at 2500×g for 15 min at room temperature. The supernatant was aspirated and 500 µl fresh human erythrocytes (Blood type O^+) were added to obtain a hematocrit of 2-4%. The culture was established in a 75 cm^3 culture flask (Corning, New York, USA) with 10 ml culture medium [1 % (w/v) RPMI-1640 (Highveld Biologicals, South Africa); 25 mM HEPES (BDH); 22 mM D-glucose (Merck, Germany); 323 µM hypoxanthine (Merck, Germany); 4 mg gentamycin (Highveld Biologicals, South Africa) in 900 ml sterile Milli Q H_2O (0.22 µm filter sterilised) and 21 mM $NaHCO_3$ and sterile, heat-inactivated human serum (A^+ or O^+ blood) added]. The culture was then gassed for 30 sec with a 5% O_2, 5% CO_2, and 90% N_2 mixture and incubated at 37°C. Growth was monitored daily on a Giemsa-stained thin blood smear by light microscopy. Medium was aspirated daily and replaced with 10 ml fresh culture medium and 500 µl erythrocytes if needed. If the parasitaemia increased to 5-10%, cultures were either subdivided or the volume scaled up in larger culture flasks.

2.2.2) Nucleic acid isolation from *P. falciparum* cultures.

2.2.2.1) Isolation of genomic DNA from *P. falciparum*.

P. falciparum parasites were harvested from one millilitre of unsynchronised (>50% ring stages) PfUP1 cultures at a parasitaemia of 5 % by the addition of 1/100 volume of 10% saponin in 10 mM EDTA and incubation at room temperature for 5 min to lyse the erythrocytes. The parasites were collected by centrifugation at 3000×g for 15 min. The supernatant and erythrocyte ghosts were aspirated and the parasite pellet resuspended in 1 ml 1×PBS (137 mM NaCl, 1.5 mM KH_2PO_4, 2.7 mM KCl, 10 mM Na_2HPO_4, pH 7). The pellet was washed four times in total with 1×PBS and boiled for 10 min in 1 ml

1×PBS according to the method of Su and Wellems (Su and Wellems, 1997). The genomic DNA was further purified from the solution by treatment with 100µg/ml DNAse-free RNAse and Proteinase K (Roche Diagnostics, Mannheim, Germany) for 2 hours at 50°C. The Proteinase K was heat-inactivated at 80°C for 5 min. The proteins and cell debris were removed with an equal volume cold Tris-saturated phenol (pH 8.0) and vortexed for 1 min. The phases were separated by centrifugation for 2 min at 13000×g. The upper aqueous phase was transferred to a clean microfuge tube and 600 µl chloroform:isoamyl alcohol (24:1) was added to remove traces of phenol. The tube was briefly centrifuged and the upper, DNA-containing aqueous phase precipitated with 1/10th volume 3M Na-acetate (pH 8.0) and three volumes cold absolute ethanol for 30 min at -70°C. The DNA was sedimented by centrifugation at 13000×g for 30 min (4°C), and the pellet washed with 70% ethanol. The pellet was dried *in vacuo* and redissolved in 60 µl TE buffer with 0.5 µg/ml RNAse (Sambrook, *et al.*, 1989).

2.2.2.2) RNA isolation from *P. falciparum* cultures.

RNA was isolated according to the Chomczynski method (Chomczynski and Sacchi, 1987), which is based on disruption of non-covalent interactions between proteins by guanidinium thiocyanate and subsequent phenol-chloroform extraction of the proteins. The use of TRI Reagent (Molecular Research Centre, Ohio, USA), which contains phenol and guanidinium thiocyanate in a mono-phase solution, allows a single step separation of RNA, DNA and proteins. Samples homogenised in TRI Reagent are separated into aqueous (containing RNA), interphase (DNA) and organic phases (proteins) by the addition of chloroform and subsequent centrifugations.

Four-hundred millilitres of unsynchronised (>50% ring stages) PfUP1 cultures at a parasitaemia of 15 % were harvested as described above. The pellet was resuspended in 1 ml TRI Reagent/5-10 × 10^6 cells. Parasites were homogenised on ice and incubated for a further 5 min at room temperature to dissociate the nucleo-protein complexes after which 200 µl chloroform was added. After a brief vortex for 15 sec, the reaction was incubated for 15 min at room temperature followed by centrifugation at 12000×g for 15 min (4°C) for phase separation. The RNA was extracted from the upper aqueous phase with 500 µl isopropanol and incubation at room temperature for 10 min. The RNA was pelleted by centrifugation at 12000×g for 10 min (4°C), washed with 1 ml 75% ethanol and airdried. The RNA pellet was resuspended in FORMAzol (Molecular Research

Centre, Ohio, USA; 100 µl/40-400 µg RNA), which is stabilised formamide to protect the RNA from RNAse degradation. The RNA was stored at –20°C. The whole procedure was repeated 24 hours later with a >50% trophozoite culture with a parasitaemia of 18% in order to obtain a representative RNA sample from the different development phases.

2.2.3) Nucleic acid quantification.

2.2.3.1) Spectrophotometry.

The concentrations of DNA, RNA and oligonucleotides were determined on a Shimadzu UV 160 A spectrophotometer. The absorbancy at 260 nm estimates the yield (double stranded DNA: one unit=50 µg/ml DNA; single stranded DNA: one unit=33 µg/ml DNA and for RNA: one unit=40 µg/ml) and the ratio of absorbances at 260 nm to 280 nm was used as an indication of the purity of the sample. Pure double stranded DNA should have a ratio $A_{260/280}$ of 1.7-1.9 while RNA should have a ratio between 1.8 and 2 $A_{260/280}$ units (Sambrook, *et al.*, 1989).

2.2.3.2) Fluorometry.

To obtain a reliable quantification of isolated plasmid concentrations, fluorometry was used. A Hoefer TK100 Mini-Fluorometer (Hoefer Scientific Instruments, USA) was calibrated with 1×TNE buffer (10 mM Tris-HCl, 1 mM EDTA; 0.2 M NaCl, pH 7.4) including one tenth the volume Hoechst 33258 DNA binding dye and calf-thymus DNA (10 µg/ml) as standard. Hoechst dye primarily binds to A and T nucleotides in double stranded DNA and the readings were multiplied with 0.4 to compensate for the A/T rich *Plasmodium* genome (~81%) with the following equation (Wang, *et al.*, 1995):

$$\text{Compensation factor} = C_{std}[0.025(AT_{\%std}-AT_{\%sample})]+1$$

where $AT_{\%std}$=58% for the calf-thymus DNA, $AT_{\%sample}$=81% for *Plasmodium* and C_{std}=100 µg/ml.

2.2.4) Primer design for RT-PCR.

The National Centre for Biotechnology Information (NCBI) SwissProt database was searched for known ODC and AdoMetDC amino acid sequences. Multiple alignments of the homologous amino acid sequences of related organisms with the CLUSTALW Program were used for the identification of conserved areas (Thompson, *et al.*, 1994). The sequences for ODC were from *Bos taurus, Homo sapiens, Mus musculus, Mus*

pahari, Rattus norvegicus, Cricetulus griseus, Gallus gallus, Xenopus laevis, Trypanosoma brucei, Drosophila melanogaster, Ceanorhabditis elegans, Saccharomyces cerevisiae and *Leishmania donovani* compared to the 12 AdoMetDC sequences from *R. norvegicus, M. musculus, B. taurus, H. sapiens, D. melanogaster, Dianthus, Nicotiana tabacum, Arabidopsis taliana, T. brucei, L. donovani, Onchocera volvulus*, and *S. cerevisiae*. These alignments were verified with the BLOCKS database (Henikoff and Henikoff, 1994) based on the protein family representations of blocks of aligned segments of the most highly conserved regions. The identified conserved areas were verified using the Basic Local Alignment Search Tool (BLAST) Program (Altschul, *et al.*, 1990) and the Blosum62 matrix and disabling gap filtering for homology testing against the SwissProt database.

The amino acid sequences of the human cDNA sequence homologues were used as templates to evaluate the identified primer sequences with the Oligo Version 4.0 Program (National Biosciences, Hamel, USA) (Rychlik and Rhoades, 1989). The codon preferences of *P. falciparum* were used in the translation of the human protein sequence with the Seqaid II Program, Version 3.81 (Centre for Basic Cancer Research, Kansas, USA). The codon preferences used in the primer design were based on the MI values (Match Index: the measure of the probable detrimental effect of choosing the wrong codon) for each codon (Hyde and Holloway, 1993; Preston, 1993), and inosine was included at positions of high redundancy. Two primers were designed and designated GSP1 (for the *Odc* cDNA) and Samdcd1 (*Adometdc*), each with 64-fold degeneracy.

Non-degenerate gene-specific primers used in RACE as well as in the primer walking strategy to sequence the entire *Odc* and *Adometdc* cDNAs were based on the obtained nucleotide sequences of the 3′-RACE fragments. However, a 2961 bp *Odc* cDNA fragment deposited in Genbank (Accession number AF012551, in 1997) allowed the synthesis of primer ODCF1 that was used in amplification of a longer *Odc* cDNA with 3′-RACE as well as in nested PCRs during 5′-RACE of the *Odc* cDNA.

2.2.5) 3′-RACE of *Odc* and *Adometdc* cDNAs.

The precise concentration of cDNA synthesised from total RNA is usually unknown and was defined as RNA equivalents. cDNA derived from total RNA is referred to as total RNA-cDNA equivalents. The concentration of the cDNA derived from total RNA was calculated from spectrophotometrical determinations of the total RNA

concentration. mRNA usually constitutes 1-5% of the total RNA. The concentration of the cDNA should therefore be between 1-5% of the total RNA used, assuming quantitative conversion (Edwards, *et al.*, 1995). cDNA derived from mRNA is referred to as mRNA-cDNA equivalents.

All the 3′-RACE reactions on cDNA as template utilised a reverse primer with a sequence identical to the 5′-anchor region of the DD-Poly-T primer used for cDNA synthesis (3′Anchor, see Table 2.1) as well as a gene-specific degenerate primer. 3′-RACE was performed on the in-house amplified uncloned cDNA library using the GSP1:3′Anchor and Samdcd1: 3′Anchor primer pairs. Non-degenerate primer ODCF1 was used to amplify a longer *Odc* cDNA fragment with 3′-RACE (ODCF1:3′ Anchor primer pair). The 50 µl reactions contained 1×reaction buffer, 2 mM $MgCl_2$, 0.2 mM dGTP,dCTP/0.3 mM dATP,dTTP, 40 pmol GSP1 or Samdcd1 and 10 pmol 3′ Anchor and ODCF1, and 2.5 U Takara ExTaq DNA polymerase (Takara Shuzo, Japan) added in a hot-start protocol. As template, either of the following was used: 50 ng total RNA-cDNA equivalents of first-strand cDNA, 100 ng total RNA-cDNA equivalents of ds cDNA, 100 ng total RNA-cDNA equivalents of adaptor-ligated ds cDNA or 7.86 ng mRNA-cDNA equivalents of the amplified uncloned cDNA library. The PCR consisted of 30 cycles at 94°C for 30 sec, 50°C for 30 sec and 68°C for 2 min.

All the PCRs were performed on a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, USA) in 0.2 ml thin-walled tubes.

2.2.6) 5′-RACE of the *P. falciparum Odc* cDNA.

The in-house amplified uncloned cDNA library was used as template for the 5′-RACE with the ODCR1:5′RACE primer pair (5′RACE sequence identical to 5′-end of Adapter, see Table 2.1). The 50 µl reaction contained 10 pmol of each primer and 15.72 ng mRNA-cDNA equivalents of the uncloned cDNA library, 1×reaction buffer, 2 mM $MgCl_2$ and 0.2 mM dGTP,dCTP/ 0.3 mM dATP,dTTP. 2.5 U Takara ExTaq (Takara Shuzo, Japan) was added in a hot-start protocol followed by 30 cycles at 94°C for 30 sec, 50°C for 30 sec and 68°C for 2 min. Specific amplification was confirmed with a nested PCR with the ODCR1:ODCF1 primer pair (Table 2.1 and Fig 2.3) to obtain the predicted ~1700 bp full-length 5′-fragment of *Odc*. The reaction contained 10 pmol each of the ODC-specific primers ODCR1 and ODCF1, 1 µl of the ODCR1:5′RACE

products as template and the rest of the components as above. The PCR was performed for 25 cycles at 94°C for 30 sec, 50°C for 30 sec and 68°C for 2 min. A nested PCR was conducted on these PCR products (1 µl of 50- and 100-fold dilutions) with the ODCR3:ODCF1 primer pair under the same reaction conditions as above.

2.2.7) Agarose gel electrophoresis of PCR products.

All PCR reactions were analysed on 1.5% (w/v) agarose (Promega, Wisconsin, USA)/TAE (0.04 M Tris-acetate, 1 mM EDTA) gels by electrophoresis in 1×TAE at 78 V (5.2 V/cm) in a Minicell EC370M electrophoretic system (E-C Apparatus Corporation, USA). The DNA was loaded in 6×loading dye (30% (v/v) glycerol, 0.025% (w/v) bromophenol blue). The gels were subsequently stained in a 10 µg/ml EtBr solution and the DNA bands visualised on a Spectroline TC-312 A UV transilluminator at 312 nm. Images were captured with a charged-coupled devise (CCD) camera linked to a computer system.

2.2.8) Purification of agarose-electrophoresed DNA fragments.

2.2.8.1) Crystal Violet visualisation of DNA bands during electrophoresis (Rand, 1996).

Direct visualisation of PCR products with crystal violet under white light was used for cloning purposes to prevent exposure of the A+T-rich DNA to the damaging effects of UV light. Normal 1.5% (w/v) agarose (Promega)/TAE gels were prepared and crystal violet (20 mg/ml) added to a final concentration of 10 µg/ml just before casting. The gel was subsequently run in 1×TAE with the same concentration of dye added. Samples were prepared by addition of glycerol to a final concentration of 5% (v/v) and electrophoresed at 6.4 V/cm.

2.2.8.2) Silica purification procedure.

PCR products and plasmid DNA in agarose gel fragments were purified using a modified silica-based method (Boyle and Lew, 1995). This method is based on the binding of DNA to a silica matrix in the presence of chaotropic salts. The DNA-containing bands were excised from the agarose gels and dissolved in three volumes 6 M NaI at 55°C. 1 mg of a 100 mg/ml silica solution (Sigma) in 3 M NaI was added to the dissolved agarose fragments and vortexed (maximum binding capacity of 3 µg DNA/mg silica). The tubes were left for 30 min at room temperature while shaking and

then incubated for another 15 min on ice with mixing every 2 min to allow the DNA to bind to the silica matrix. The bound DNA-silica was sedimented at high speed for 30 sec and the supernatant removed. The pellet was washed twice with wash buffer (10 mM Tris-HCl, pH 7.5; 50 mM NaCl; 2.5 mM EDTA; 50% v/v ethanol) before the elution of DNA at 55°C for 5 min in the minimum volume (10-30 µl) of Milli Q H_2O. The average recovery was ~50-100 ng DNA/µl.

2.2.9) Cloning protocols.

2.2.9.1) Preparation of competent cells.

Competent SURE (Stratagene, La Jolla, CA, USA) and DH5α *E. coli* cells (Gibco BRL, Life Technologies, USA) were prepared according to a calcium/manganese-based method (Hanahan, *et al.*, 1991). Bacteria streaked on M9 minimal medium agar plates (0.05 M Na_2HPO_4-$2H_2O$; 0.02 M KH_2PO_4; 8 mM NaCl; 0.02 M NH_4Cl; 2 mM $MgSO_4$; 0.01 M D-glucose; 0.1 mM $CaCl_2$; 1 mM thiamine hydrochloride; 1.5 % agar (w/v), pH 7.4) were picked and streaked on a LB plate (1% tryptone, 0.5% yeast extract, 1% NaCl pH 7.5, 1.5% (w/v) noble agar) containing the appropriate antibiotic (12.5 µg/ml tetracycline for SURE cells) and grown overnight at 30°C. Several colonies were dispersed into 1 ml SOB medium (2% Tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, pH 6.8-7.2) by vortexing and then inoculated into 50 ml SOB medium. The cells were incubated with shaking at 250 rpm at 30°C until an OD_{600} of 0.3 was reached at which the cells were in the early exponential phase of growth. The cells were transferred to 50 ml centrifuge tubes and incubated on ice for 10 min and collected by centrifugation at 1000×g for 15 min (4°C). The supernatant was removed and the pellet resuspended in one third the volume of CCMB 80 medium (80 mM $CaCl_2$-$2H_2O$, 20 mM $MnCl_2$-$4H_2O$, 10 mM $MgCl_2$-$6H_2O$, 10 mM K-acetate, 10% glycerol, pH ~6.4) and incubated on ice for 20 min. The cells were pelleted at 1000×g for 10 min (4°C) and resuspended in a twelfth of the original volume of CCMB 80. The competent cells were aliquotted on ice, flash frozen in liquid nitrogen and stored at –70°C.

2.2.9.2) Conventional miniprep plasmid isolation (Sambrook, *et al.*, 1989).

This protocol was used when plasmid DNA (recombinant or non-recombinant) was isolated from a 5 ml overnight culture (30°C) of the transformed cells in LB-Broth (1% Tryptone, 0.5% yeast extract and 1% NaCl, pH 7.5) with the appropriate antibiotic (e.g. 50 µg/ml ampicillin for pBluescript based vectors). The cells were harvested by

centrifugation at 3000×g for 10 min (4°C) and the pellet resuspended in 200 µl Solution I (50 mM D-Glucose; 25 mM Tris-HCl, pH 8; 10 mM EDTA). 200 µl fresh Solution II (0.2 M NaOH, 1% SDS) was added followed by incubation on ice for 5 min. Ice-cold 200 µl Solution III (3 M K-acetate) was added and the reaction incubated on ice for 15 min. The tubes were centrifuged for 15 min at 13000×g (4°C) and the supernatant added to 600 µl cold Tris-saturated phenol (pH 8.0) and vortexed for 1 min. The phases were separated by centrifugation for 2 min at 13000×g. The upper aqueous phase was transferred to a clean microfuge tube and 600 µl chloroform:isoamyl alcohol (24:1) was added to remove traces of phenol. The tube was briefly centrifuged and the upper, DNA-containing aqueous phase precipitated with 1 ml cold absolute ethanol for 30 min at -70°C. The plasmid DNA was sedimented by centrifugation at 13000×g for 15 min (4°C), and the pellet washed with 70% ethanol. The pellet was dried *in vacuo* and redissolved in 40 µl TE buffer with 0.5 µg/ml RNAse.

2.2.9.3) High-Pure Plasmid Isolation Kit (Roche Diagnostics, Mannheim, Germany).

This protocol was used for the isolation of plasmid DNA for nucleotide sequencing. It entails the conventional alkaline lysis of cells followed by binding of the plasmid DNA to glass fibres in the presence of chaotropic salts. DNA is eluted with a low salt buffer or Milli Q water. 10 ml culture in LB-Broth was grown overnight at 30°C to a OD_{600} of ~ 2 units and the cells were collected by centrifugation for 10 min at 3000×g. The manufacturers protocol was followed henceforth. Briefly, the pellet was suspended in 250 µl suspension buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8) to which 250 µl lysis buffer (0.2 M NaOH; 1% SDS) was added. Incubation for 5 min at room temperature was followed by the addition of 350 µl binding buffer (4 M guanidine hydrochloride; 0.5 M K-acetate, pH 4.2). After a 5 min incubation on ice and centrifugation for 10 min at 13000×g (4°C), the clear supernatant was transferred to the filter-tube and centrifuged for 45 sec at high speed. The filter was then washed with wash buffer I (5 M guanidinium hydrochloride; 20 mM Tris-HCl, pH 6.6) and once with wash buffer II (20 mM NaCl, 2 mM Tris-HCl, pH 7.5). The DNA was eluted in Milli Q H_2O and the concentration determined fluorometrically (section 2.2.3.2).

2.2.10) A/T cloning strategies.

pGEM-T Easy cloning vector (Promega, Wisconsin, USA) containing 3′-terminal thymidine at both ends in the multiple cloning site was used. PCR products were A-tailed during amplification with ExTaq DNA polymerase. The PCR products were then ligated to the pGEM-T Easy vector in a 3:1 molar ratio of insert to vector. The reaction containing 1×T4 DNA ligase buffer, 50 ng vector and 3 U of T4 DNA ligase (Promega) was incubated at 4°C overnight and the ligase inactivated at 70°C for 15 min. The ligation mixture was stored at –20°C until needed.

Transformation of competent cells with the recombinant plasmids was performed according to the conventional heat-shock method (Sambrook, *et al.*, 1989). Competent DH5α or SURE *E. coli* cells were thawed on ice and 5 µl of the ligation mixture was added to 100 µl of the cells in a glass Vacutainer tube (Vacutest, USA). Five ng non-recombinant pBluescript SK$^-$ was transformed in 100 µl cells as a positive control. The negative control consisted of 100 µl cells without any DNA added. Incubation of cells and plasmids on ice for 30 min was followed by a heat shock at 42°C for 45 sec and 2 min on ice. 900 µl preheated SOC (SOB with 50 mM D-glucose) was added and the tubes incubated at 30°C for 1 hour while shaking. 8 mg X-gal (5-bromo-4-chloro-indolyl-β-D-galactoside, 20 mg/ml in dimethylformamide, DMF) and 0.4 mM IPTG (Isopropyl-D-galactoside, for SURE *E. coli* cells only) was plated on LB-Agar plates containing 100 µg/ml ampicillin, and left at room temperature to allow absorption of DMF. 100 µl of the transformation mixtures was plated and the plates incubated at 30°C for no longer than 16 hours.

White colonies were picked for screening of inserts. The colonies were grown overnight at 30°C in LB-Broth with 50 µg/ml ampicillin and plasmid DNA was isolated according to the miniprep plasmid isolation method (see section 2.2.9.2). 500 ng plasmid DNA was digested with 10 U of *EcoR*I (Promega, Wisconsin, USA) in the appropriate 1×buffer for four hours or overnight at 37°C. The digested plasmids were analysed on a 1% (w/v) agarose (Promega)/TAE gel in 1×TAE at 7.8 V/cm and visualised by staining with EtBr. Positive clones were stored in 15% (v/v) glycerol/LB broth at –70°C.

2.2.11) Automated nucleotide sequencing.

The nucleotide sequences of the cloned fragments were determined with an automated ABI Prism 377 DNA Sequencer (PE Applied Biosystems, California, USA). For cycle sequencing of cloned PCR fragments in pGEM-T Easy, primers complementary to the T7 and SP6 promoters in the vector were used. The nucleotide sequences for the respective primers are as follows (all with a Tm of ~ 50°C):

T7: 5′ GTA ATA CGA CTC ACT ATA GGG C 3′
SP6: 5′ ATT TAG GTG ACA CTA TAG AAT AC 3′

The cycle-sequencing reaction contained 200-500 ng double stranded DNA template, 3.2 pmol of the respective primer and 2 µl terminator ready reaction mix from the Big Dye Sequencing Kit (version 2.0, PE Applied Biosystems) in a final reaction volume of 5 µl. The cycle-sequencing was performed in a Perkin Elmer GeneAmp PCR system 9700 with 25 cycles of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min. The labelled extension products were purified by the addition of one-tenth the volume of 3 M sodium acetate (pH 5.4) and 2.5 volumes of cold absolute ethanol. After vortexing and standing on ice for 10 minutes, the labelled products were collected by centrifugation at 13000×g for 25 minutes (4°C). The supernatant was removed and the pellet was washed with 70% ethanol and dried *in vacuo*. The pellet was resuspended in 3 µl loading dye (5:1 ratio of deionised formamide to 25 mM EDTA, pH 8 and blue dextran, 30 mg/ml). The samples were denatured at 95°C for 2 minutes and snapcooled on ice and were analysed on a 36 cm gel according to the ABI PRISM 377 Genetic Analyser Users Manual. The raw data was analysed with the Analysis software (ABI Prism Sequencing Analysis Version 3.0, PE) and preliminary alignments done with the Navigator software (ABI Prism Sequencing Navigator, Version 1.0.1, PE). Results were confirmed by visual inspection of the electropherograms obtained.

Direct PCR product automated fluorescent cycle sequencing was used to sequence the *Odc* 5′-RACE products. The RACE product was purified from agarose with a silica-based method (section 2.2.8.2) and 62.5 ng was used in a reaction containing 3.2 pmol of the respective primers and 2 µl terminator ready reaction mix in a final volume of 5 µl. The reaction conditions and purification of the labelled products were as above.

The sequences obtained by the T7 and SP6 directional sequencing of the *Odc* and *Adometdc* were used for the design of non-degenerate gene-specific primers to sequence

the remaining internal fragments in a primer-walking strategy. The T7 and SP6 directional sequencing was repeated twice for three positive clones and the internal sequence once for three clones. The sequences were aligned to obtain a consensus sequence. The alignments were done with the CLUSTAL W Program (Thompson, *et al.*, 1994) and the consensus sequences deduced with the Genetic Data Environment program (Smith, *et al.*, 1994).

All nucleotide sequences were submitted to BLAST (Altschul, *et al.*, 1990) using the Blosum62 matrix and disabling gap filtering to identify the various clones based on homology testing against the known nucleotide and deduced amino acid sequences in the Genbank and SwissProt databases.

2.2.12) Northern blot analyses of *P. falciparum* total RNA with a *Odc*-specific probe (Sambrook, *et al.*, 1989).

2.2.12.1) Synthesis of DIG-labelled *Odc*-specific DNA probe.

Non-radioactive labelling was performed by incorporation of a digoxigenin-labelled nucleotide. Digoxigenin (DIG), a cardenolide steroid isolated from *Digitalis* plants, is detected using an enzyme-linked immunoassay with an anti-DIG antibody conjugated to alkaline phosphatase (Roche Diagnostics, Mannheim, Germany). The entire *Odc* fragment (2918 bp) was used as template and the predicted ORF of 2858 bp was labelled by the incorporation of DIG-labelled dUTP in a PCR. The 25 µl reaction contained 10 pg plasmid DNA, 10 pmol each of ODCexpf and ODCexpr (see Table 2.1), 1× rTaq DNA polymerase reaction buffer, 2 mM $MgCl_2$, 0.2 mM dGTP,dCTP and dATP, 0.17 mM dTTP, 0.03 mM DIG-11 dUTP (Roche Diagnostics, Mannheim, Germany) and 2.5 U Takara rTaq DNA polymerase (Takara Shuzo, Japan). The exact reaction was repeated in tandem but without DIG-labelled dUTP and analysed with agarose gel electrophoresis to determine the concentration of the probe by comparison with the intensities of the molecular mass markers of known concentration. The DIG-labelled probes were stored at -20°C in silanised tubes.

2.2.12.2) Dot-blot mock probe of the *Odc*-specific probe.

One in ten dilutions (ranging from 4 ng/µl to 0.4 pg/µl) were made and 1 µl dot-blots of each dilution was spotted onto positively charged nylon membranes (Roche Diagnostics, Mannheim, Germany). 2.5 pmol of a DIG-labelled control oligonucleotide

was spotted as control (Roche Diagnostics, Mannheim, Germany). The DNA was crosslinked to the membrane for 3 min on an UV transilluminator (312 nm wavelength). The membrane was equilibrated in wash buffer (3% (v/v) Tween-20 (Merck) in maleic acid buffer: 0.1 M maleic acid, 0.15 M NaCl, pH 7.5) for 1 min and incubated for 30 min in 1×blocking solution (1% (v/v) blocking reagent in maleic acid buffer, Roche Diagnostics) while shaking, followed by incubation in a 1:10 000-fold dilution of sheep anti-DIG alkaline phosphatase Fab fragment (Roche Diagnostics, Mannheim, Germany) at room temperature for 30 min while shaking. The membrane was washed twice with wash buffer for 15 min each and equilibrated in detection buffer (100 mM Tris-HCl, pH 9.5; 1 mM NaCl) for 2 min.

Detection was performed with CDP-Star (Roche Diagnostics), an ultra-sensitive chloro-substituted 1,2 dioxetane chemiluminescent substrate for alkaline phosphatase, which emits extremely rapid light signals detected and recorded on film. Enzymatic dephosphorylation of CDP-Star forms the meta-stable dioxetane phenolate anion, which decomposes and emits light at 466 nm. 1:100 dilution of CDP-Star was added to the membrane in the detection buffer and incubated at room temperature for 5 min. The membrane was sealed in a clean plastic bag and exposed to Konica Medical Film (AX; X-Ray Imaging Services, SA) for 3 min in the dark. The chemiluminescence was detected by developing the film for 3 min in developer (PolyCon A Variable Contrast X-ray developer, Champion Photochemistry, SA), washing in distilled water and fixing for 3 min in fixer (Perfix High Speed X-ray Fixer, Champion Photochemistry, SA).

2.2.12.3) Northern blot of *P. falciparum* total RNA with *Odc*-specific probe.

All reactions were performed under RNAse-free conditions by using buffers made up in diethyl pyrocarbonate (DEPC)-treated H_2O (1 ml DEPC in 1L Milli Q H_2O, autoclaved twice) and using plastic and glassware treated with 0.5 M NaOH. 30 µg of total RNA from *P. falciparum* was electrophoresed on a denaturing agarose gel along with DIG-labelled RNA molecular weight marker (0.3-0.69 kb, Roche Diagnostics, Mannheim, Germany). A 1% denaturing agarose gel (w/v) was prepared by adding 3 ml 10×MOPS (morpholinopropanesulphonic acid), 1.64 ml 37% formamide and 7 ml DEPC-treated H_2O to a 30 ml agarose gel solution (Promega, Wisconsin, USA). The gel was run in a 1×MOPS buffer. Samples were denatured in 2 µl 10×MOPS, 3 µl formaldehyde and 10 µl 37% formamide at 55°C for 15 minutes and loaded in 6×tracking dye (15% (w/v) Ficoll, 0.025% (w/v) bromophenol blue) with EtBr (76 µg/µl final concentration). The

gel was electrophoresed at 4.5 V/cm in a Minicell EC370M electrophoretic system (E-C Apparatus Corporation, USA) and visualised on a Spectroline TC-312 A UV transilluminator (312 nm wavelength). The gel was washed three times in RNAse-free 20×SSC (3 M NaCl, 0.3 M Na-citrate, pH 7) to remove the EtBr and formaldehyde from the gel. The RNA was transferred from the gel to a positively charged nylon membrane by capillary action (Roche Diagnostics, Mannheim, Germany) by placing the gel on filter paper hanging on both sides in 20×SSC and covering the gel with nylon membrane pre-wet in 20×SSC and cut to size. The capillary action was started by covering the nylon membrane with 5 pieces filter paper pre-wet in 20×SSC and 30 layers of dry paper towel. A 500g weight was placed on the top and transfer was allowed to proceed for 16 hours at 22°C. The membrane was subsequently equilibrated in 5×SSC for 30 sec and the RNA cross-linked to the membrane by exposure to UV light for 3 min.

The membrane was equilibrated in a high-SDS hybridisation buffer (Church buffer: 7% w/v SDS, 50% v/v formamide, 5×SSC, 2% v/v blocking reagent, 0.1% N-laurylsarcosine in 50mM $NaPO_4$) for 1 hour at 50°C to limit background. 50 ng of the 2858 bp DIG-labelled *Odc*-specific probe was heat-denatured at 94°C for 10 min and snap-cooled on ice. The 50 ng probe was hybridised to the *P. falciparum* RNA on the membrane by incubation in high-SDS hybridisation buffer for 16 hours at 50°C. The membrane was washed twice for 15 min each in low-stringency 2× wash solution (2×SSC, 0.1% w/v SDS) at room temperature while shaking and then twice for 15 min in a high-stringency 0.5× wash solution (0.5×SSC, 0.1% w/v SDS) at 68°C to remove any non-specifically bound probe and leave only fully homologous bound probe. The detection of the hybridised probe was performed as for the dot-blot mock probe described in section 2.2.12.2.

2.3) RESULTS.

2.3.1) Primer design.

The primers used in 3′-RACE were designed based on consensus amino acid sequences identified by multiple-homologous alignment of ODC and AdoMetDC amino acid sequences from different organisms as depicted in Fig. 2.2. A search of the SwissProt databank revealed several partially or fully characterised sequences of which 13 full

ODC sequences and 12 AdoMetDC sequences where chosen based on the relationship between *P. falciparum* and the higher eukaryotes (Hyde and Holloway, 1993). The boxed areas indicate the consensus regions used for primer design with low degeneracy using match-index values (the measure of the probable detrimental effect of choosing the wrong codon) for the different codons used by *P. falciparum* (Hyde and Holloway, 1993). BLOCKS analyses of the conserved area for ODC grouped this area into proteins of the group IV decarboxylases i.e. the family of PLP dependent decarboxylases including ornithine, arginine and diaminopimelic acid decarboxylases. The identified AdoMetDC conserved area was grouped into BLOCKS for decarboxylase proteins involved in the biosynthesis of S-adenosylmethionine. A BLASTP search of the SWISSProt database with the conserved areas as search term revealed significant alignments only with ODC or AdoMetDC sequences, respectively.

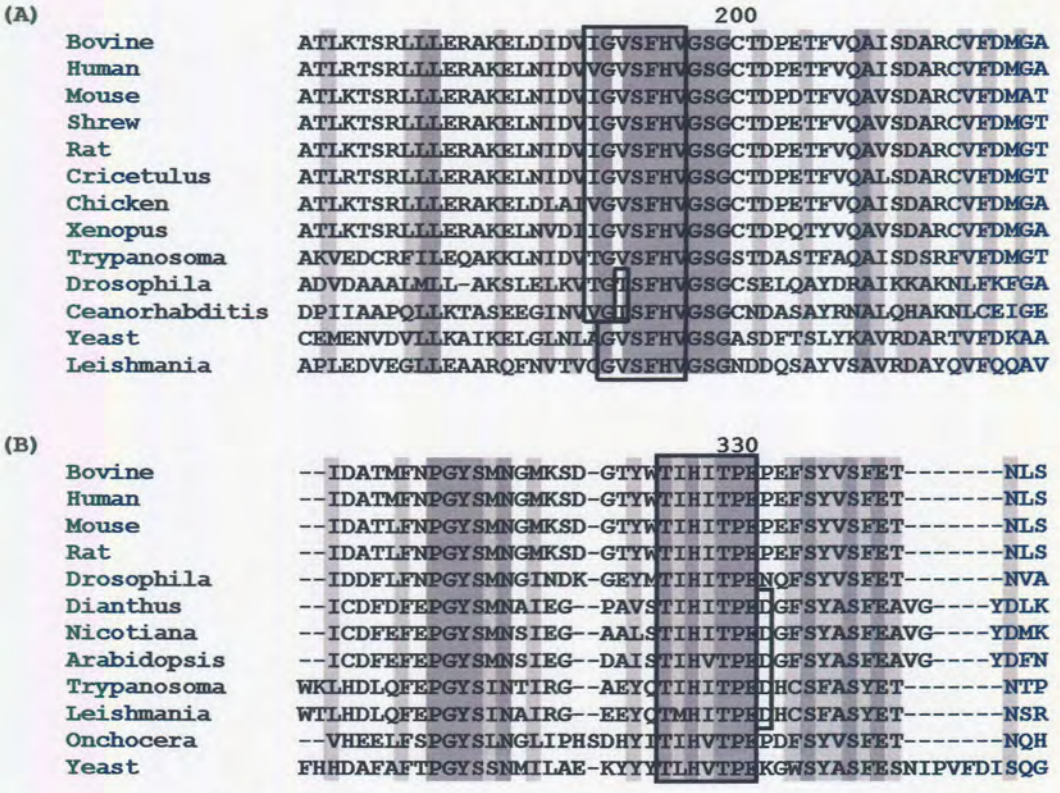


Figure 2.2: Partial multiple-alignment of ODC (A) and AdoMetDC (B) amino acid sequences from different organisms. Relevant selections of the total sequences are indicated by the residue numbers of the human sequence. Dark grey blocks indicate identical amino acids and light grey indicates residues conserved in terms of properties. The boxed residues indicate those used in the primer design, the area in (A) for GSP1 and the one in (B) for Samdcd1. See section 2.2.4 for the organisms used.

Non-degenerate gene-specific primers used in RACE to amplify *Odc* cDNA as well as in the primer walking strategy to sequence the entire *Odc* and *Adometdc* cDNAs were based on the obtained nucleotide sequences of the 3′-RACE fragments. Positions of the primers are indicated in Fig 2.3 and their characteristics are listed in Table 2.1. Primer ODCF1 was based on a *Odc* cDNA sequence of 2961 bp deposited in Genbank (Accession number AF012551). This allowed the amplification of a longer *Odc* fragment with 3′-RACE as well as in applying a nested-PCR protocol during 5′-RACE. In Fig. 2.3, the numbering for the *Odc* cDNA is represented such that primer ODCF1 anneals at position 1.

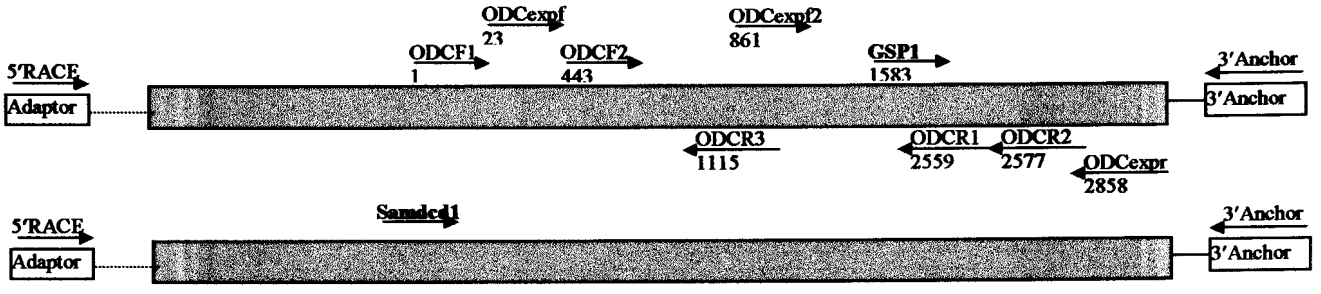


Figure 2.3: Schematic representation of the gene-specific primers used for amplification and nucleotide sequencing of the full-length *Adometdc* and *Odc* cDNAs. The proposed open reading frames (filled boxes) of unknown lengths of *Odc* (top) and *Adometdc* (bottom) are indicated. The positions and directions of annealing of the different gene-specific primers are indicated. The positions of the degenerate Samdcd1 and GSP1 primers are included in bold. The RACE primer sites introduced during cDNA synthesis are boxed.

The characteristics and sequences of the primers are summarised in Table 2.1.

Table 2.1: Summary of the characteristics of the various primers used in PCR. The specific use of the primers are indicated.

Primer	Utility	Sequence (5' – 3')	Length	Tm (°C) [a]	Origin
DD-Poly-T	1st strand cDNA synthesis	GCT ATC ATT ACC ACA ACA CTC $(T)_{18}$VN	41	max:41[b] min:37	Original design
3' Anchor	3′ RACE	GCT ATC ATT ACC ACA ACA CTC	21	56	Original design
GSP1	3′ RACE *Odc*	TTW GGW GTW WCI TTT CAY GTW	21	max:52 min:50	Multiple alignment
Samdcd1	3′ RACE *Adometdc*	ACW ATW CAT RTW ACW CCA GAA G	22	max:55 min:53	Multiple alignment
5′ RACE	5′ RACE	TTA CAG GAC CAC ATC AAC TAT CGG G	25	63	Original design
ODCR1	5′ RACE *Odc*	**GCT** ACT CAT ATC GAA TAC ATC TCT AC	26	60	3′-RACE product
ODCR2	Primer-walking sequencing	GTA ATA ACA TAA ATA GGA CAT C	22	49	3′-RACE product
ODCR3	5′ RACE *Odc*	**G**AA TTT ATA CAA ACT ACT GAT G	22	51	3′-RACE product
ODCF1	3′ RACE *Odc* & Primer-walking sequencing	**G**AA TTT TTA TAA TGG AAA GTA TAT G	25	52	Genbank (AF012551)
ODCF2	5′ RACE *Odc*	**G**TT GAT GAT ATG TAT GAG TAT G	25	53	3′-RACE product
ODCexpf	*Odc* domain for expression	**CTC GAG** TTC ATG ATA AAT TAT GTA TTC	27	57	3′-RACE product
ODCexpr	*Odc* domain for expression	**GCT CAG** CTT TTC TTA TTT ACC AAT G	25	58	3′-RACE product

- [a] T_m equation=69.3+0.41(%GC)-650/length (Rychlik, *et al.*, 1990). Min and max are the calculations for the minimum and maximum %GC, respectively.
- [b] The Tm was calculated for the T(18)VN stretch on its own, as this is the only area involved in first-strand cDNA synthesis.
- W: A or T, R: A or G and Y: C or T according to the IUBMB degenerate nucleotide nomenclature.
- The nucleotides in bold were included for cloning strategies.

2.3.2) 3′-RACE of the *P. falciparum Odc* and *Adometdc* cDNAs from the uncloned cDNA library.

2.3.2.1) Amplification of *Odc* cDNA.

3′-RACE for the *Odc* cDNA was performed on the amplified, uncloned cDNA library with the degenerate primer GSP1 and the reverse 3′ Anchor primer. 3′-RACE with GSP1:3′ Anchor revealed the expected 1300 bp band for 25 ng first-strand (total RNA-cDNA equivalents) although a lower molecular mass band was present, possibly a truncation product of first-strand cDNA synthesis (Fig. 2.4). After amplification of the uncloned cDNA library, the 1300 bp band was sharper when 7.86 ng mRNA-cDNA equivalents were used (lane 3). The truncated product was not present indicating that the library resulted in cleaner, full-length cDNA. No band was visible after amplification

with the GSP1:5′ RACE primer pair as expected for successful suppression PCR (lane 2).

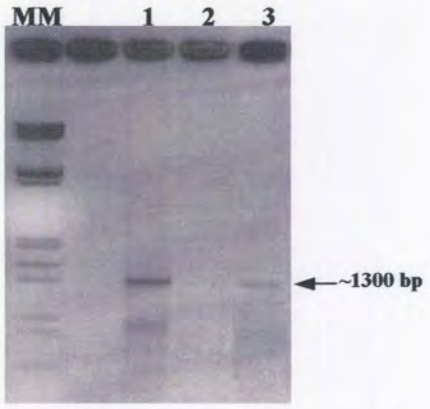


Figure 2.4: 3′-RACE PCR of the *Odc* cDNA with degenerate primer GSP1. 25 ng total RNA-cDNA equivalents of first-strand cDNA with the GSP1:3′ Anchor primer pair (lane 1), 7.86 ng mRNA-cDNA equivalents of double-stranded, adaptor-ligated amplified cDNA with GSP1:5′ RACE primer pair (lane 2) and 7.86 ng mRNA-cDNA equivalents of double-stranded, adaptor-ligated amplified cDNA with the GSP1:3′ Anchor primer pair (lane 3). **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

A longer *Odc* cDNA fragment of 2918 bp was amplified with the ODCF1:3′ Anchor primer pair using 7.86 ng of the uncloned cDNA library as template in 3′-RACE (Fig. 2.5). The ODCF1 primer sequence was based on the Genbank sequence of *Odc* (Accession number AF012551). The amplification of the longer *Odc* fragment was only successful when a low extension temperature was used (60-68°C). This might be due to the long A+T-rich product undergoing premature denaturation or truncation and derailing of the DNA polymerase at these sites at higher temperatures (Su, *et al.*, 1996).

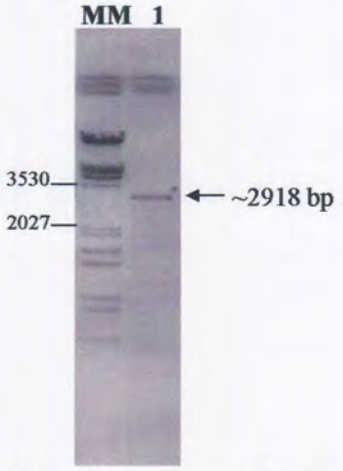


Figure 2.5: Amplification of the full-length *Odc* cDNA with 3′-RACE. The ODCF1:3′ Anchor primer pair was used in 3′-RACE on the uncloned cDNA library. The expected band of ~2918 bp was obtained. **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

2.3.2.2) 3′-RACE of the *Adometdc* cDNA.

3′-RACE for the *Adometdc* cDNA was performed on the amplified, uncloned cDNA library with the degenerate primer Samdcd1 and the reverse 3′ Anchor primer. Several bands ranging in size from ~3500 bp to ~850 bp were observed with a background smear (Fig. 2.6 A). The expected size for the single *Adometdc* domain was at this stage ~500 bp. Amplification at a lower extension temperature of 60°C decreased the number of smaller bands but the large ~3500 bp band was still not resolved well (Fig. 2.6 B).

The 850 bp band was cloned into pGEM T-Easy and the nucleotide sequence determined. However, this band was identified as a *P. falciparum* gene already characterised as the highly-abundant 5.8S rRNA gene (89% identical). This product was therefore non-specifically amplified due to the degeneracy of the Samdcd1 primer used. Further investigations were not performed on the 3500 bp band at that stage because of its extreme size, however, as will be shown in section 2.5.1, because of the bifunctional nature of *PfAdometdc/Odc*, a 2963 bp band with the Samdcd1:3′ Anchor primer pair should have been expected. The ~3500 bp band seen in Fig. 2.6 is therefore in the correct size range and probably contains the *Adometdc* and *Odc* cDNAs.

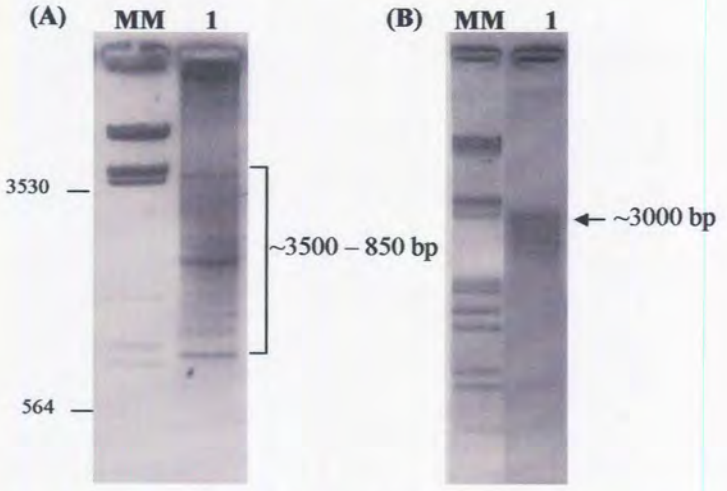


Figure 2.6: 3′-RACE of the *Adometdc* cDNA with degenerate primer Samdcd1. (A) Amplification with the Samdcd1:3′ Anchor primer pair on 7.86 ng mRNA:cDNA equivalents of the uncloned cDNA library at 72°C and in **(B)** at 60°C. **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

2.2.3) 5′-RACE of *Odc* cDNA.

5′-RACE was performed with the ODCR1:5′ RACE primer pair on the amplified, uncloned cDNA library. Multiple faint bands ranging in size from ~800 bp to ~3000 bp were observed (Fig. 2.7 A) when 15.72 ng mRNA-cDNA equivalents were used (lane 2) but not when 7.86 ng mRNA-cDNA equivalents were used (lane 1). The multiple bands are a common occurrence in 5′-RACE protocols since the adaptor is ligated equally to full-length and 3′-truncated cDNA species. A nested PCR was performed with the ODCR1:ODCF1 primer pair on the 5′-RACE products. The expected 1700 bp band was amplified (Fig. 2.7 B) and was distinct since no other amplification products were observed. The authenticity of the 1700 bp band was confirmed when the expected 1200 bp band was obtained with the ODCR3:ODCF1 nested primer pair (Fig. 2.7 C).

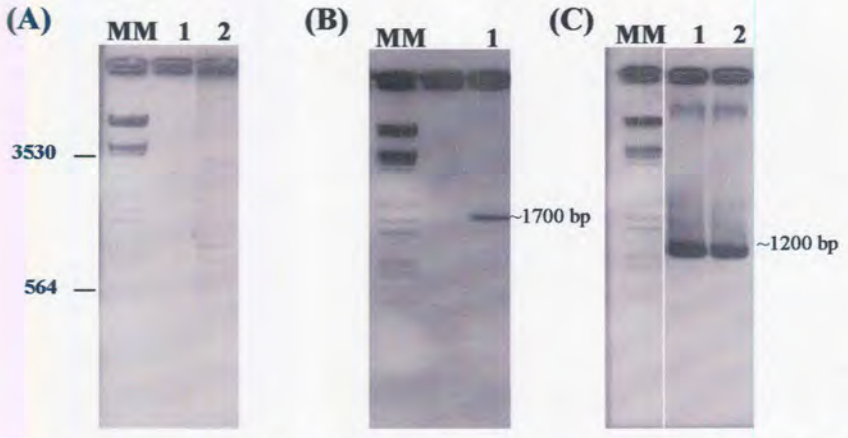


Figure 2.7: 5′-RACE of the *Odc* cDNA on the amplified, uncloned cDNA library and nested PCR strategy. (A) Several bands were observed with the ODCR1:5′ RACE primer pair only with 15.72 ng mRNA-cDNA equivalents (lane 2) but not with 7.86 ng (lane 1). **(B)** The nested PCR was done with the ODCR1:ODCF1 primer pair on the products from (A). **(C)** The second nested PCR with ODCR3:ODCF1 on the products obtained in (B) with 50-fold (lane 1) and 100-fold dilutions (lane 2). **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

All the abovementioned *Odc* fragments (the 3′-RACE 1300 bp and 2918 bp bands as well as the 1700 bp 5′-RACE band) were cloned into pGEM T-Easy and the nucleotide sequences determined. The fragments all had high homologies to known *Odc* genes as well as 92% identity to a partial sequence for the *P. falciparum* ODC gene submitted to Genbank (Genbank accession number AF012551). The full consensus sequence obtained here was compiled and deposited in Genbank (Accession number AF139900). Based on this sequence, an open reading frame of 2858 bp was predicted as the longest in-frame read starting from the first ATG codon. This information was used in the design of primers ODCexpf and ODCexpr (used in the following section).

2.3.4) Northern blot analyses of *P. falciparum* total RNA with an *Odc*-specific probe.

2.3.4.1) Synthesis of a DIG-labelled *Odc*-specific probe.

The transcript size of *Odc* cDNA was analysed with a Northern blot. Total RNA from *P. falciparum* was probed with a DIG-labelled probe corresponding to the entire *Odc* predicted ORF of 2858 bp. Fig. 2.8 indicates the amplified probe (A) and the results of a dot-blot mock probe experiment to test the efficiency of DIG-labelling of the probe (Fig. 2.8 B). The PCR was sufficient in labelling the probe such that picogram amounts of the probe were still detectable.

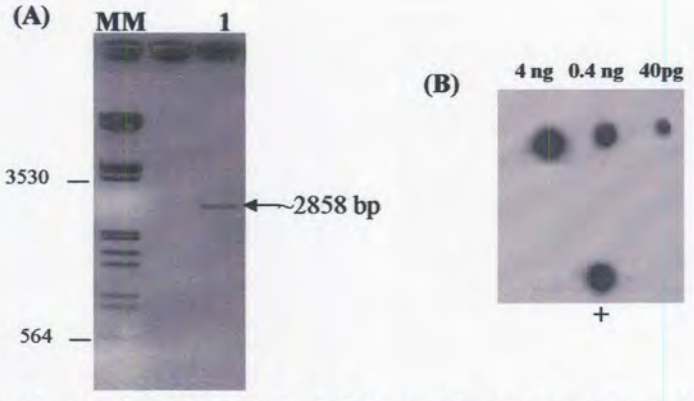


Figure 2.8: Synthesis of a DIG-labelled *Odc*-specific probe and quality analyses. (A) PCR-mediated synthesis of the 2858 bp *Odc*-specific probe. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker. **(B)** Dot-blot mock probe of the labelled *Odc*-specific probe detected with chemiluminescence. 10 times dilutions are indicated at the top. + indicates the DIG-labelled oligonucleotide as control.

2.3.4.2) Northern blot of *P. falciparum* total RNA with the DIG-labelled *Odc*-specific probe.

The DIG-labelled *Odc*-specific probe was used to probe total RNA isolated from *P. falciparum* cultures and electrophoresed under denaturing conditions. The DIG-labelled *Odc*-specific probe hybridised to mRNA corresponding to 7 kbp (Fig. 2.9 A). The molecular mass of the mRNA was determined compared to the migration profile of known DIG-labelled RNA standards. The Rf values of the RNA standards and the concluding regression line (r^2=0.999) are indicated in Fig. 2.9 (B). The size of the single distinct band obtained was unexpected, implying a considerable contribution of untranslated regions.

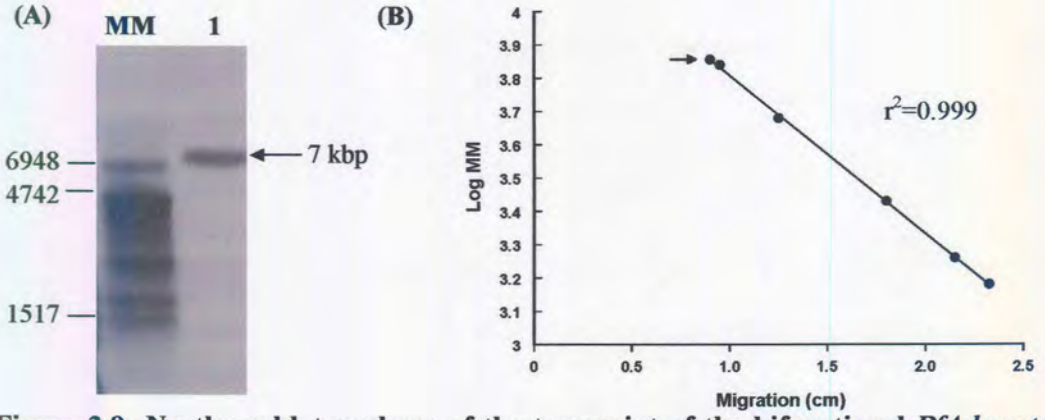


Figure 2.9: Northern blot analyses of the transcript of the bifunctional *PfAdometdc/Odc*. **(A)** Northern blot of *P. falciparum* RNA probed with a DIG-labelled *Odc*-specific probe indicating a single band (arrow). **MM**: DIG-labelled RNA molecular weight markers. **(B)** Linear regression analyses of the Rf values for the RNA molecular weight markers. Arrow indicates the migration of the *Odc*-probed *P. falciparum* RNA band.

PART II: Molecular genetics of the full-length *PfAdometdc/Odc.*

2.4) MATERIALS AND METHODS.

2.4.1) Long-distance PCR (LD-PCR) of the full-length bifunctional *PfAdometdc/Odc*.

2.4.1.1) 3′-RACE of the full-length *PfAdometdc/Odc* cDNA.

The full-length *PfAdometdc/Odc* was amplified with LD-PCR using the in-house designed, uncloned *P. falciparum* cDNA library. 3′-RACE was performed using the Sampetf1:3′ Anchor primer pair. Sampetf1 (5' CTG CGG GAT CCA **ATG** AAC GGA ATT TTT GAA G 3') was designed based on the nucleotide sequence of the full-length *PfAdometdc/Odc* (Müller, *et al.*, 2000) and overlapped the start codon indicated in bold. The 50 µl reactions contained 1×reaction buffer, 2 mM $MgCl_2$, 0.2 mM dGTP,dCTP/0.3 mM dATP,dTTP, 10pmol Sampetf1 and 10 pmol 3′ Anchor, and 2.5 U Takara ExTaq DNA polymerase (Takara Shuzo, Japan) added in a hot-start protocol. As template, 7.86 ng mRNA-cDNA equivalents of the amplified uncloned cDNA library was used. The PCR consisted of 30 cycles at 94°C for 30 sec, 58°C for 30 sec and 65°C for 4 min. PCR products were cloned and the nucleotide sequences determined as described in sections 2.2.8-2.2.11.

2.4.1.2) Amplification of the full-length *PfAdometdc/Odc* gene from *P. falciparum* genomic DNA.

The full-length *PfAdometdc/Odc* gene was amplified from 50 ng *P. falciparum* genomic DNA in a 50 µl reaction containing 1×reaction buffer, 2 mM $MgCl_2$, 0.2 mM dGTP,dCTP/0.3 mM dATP,dTTP and 2.5 U Takara ExTaq DNA polymerase (Takara Shuzo, Japan) added in a hot-start protocol. The gene-specific primers Sampetf1 and ODCexpr were used at 10 pmol each. The PCR conditions were the same as above. PCR products were cloned and the nucleotide sequences determined as described in sections 2.2.8-2.2.11.

2.4.2) *In silico* nucleotide sequence analyses of the *PfAdometdc/Odc* gene

The full-length nucleotide sequence of the *PfAdometdc/Odc* gene and cDNA was compared with the published sequences submitted to Genbank (accession numbers AF093488 and AF112367 for cDNA and genomic DNA, respectively). Multiple nucleotide sequence alignments were performed with Clustal W (Thompson, *et al.*, 1994). The *Plasmodium* genome project database (www.plasmoDB.org) was searched for the chromosomal localisation of the gene as well as to obtain sequences for the untranslated and flanking regions. Promoter predictions were performed with the ProScan (bimas.dcrt.nih.gov/molbiol/proscan) and the Neural Network Promoter Prediction Server (www.fruitfly.org/seq-tools/promoter.html). Secondary structure analyses of the 2600 bp of the genomic DNA upstream of the start codon was performed with MFOLD (Walter, *et al.*, 1994)

2.5) RESULTS.

2.5.1) Amplification of the full-length cDNA of the bifunctional *PfAdometdc/Odc.*

Based on the nucleotide sequence described in Müller *et al.* (2000), we independently preformed long-distance PCR of the full-length *PfAdometdc/Odc* cDNA with the in-house designed, uncloned *P. falciparum* cDNA library as well as from *P. falciparum* genomic DNA.

The full-length cDNA of 4333 bp of the bifunctional *PfAdometdc/Odc* was amplified with 3'-RACE from the uncloned *P. falciparum* cDNA library with the Sampetf1:3' Anchor primer pair. This consists of the open reading frame of 4260 bp (including the

start and stop codons) with the 3′ UTR of 34 bp, 18 bp poly-A tail and 21 bp 3′ Anchor. A single distinct band was obtained after performing the LD-PCR using lowered extension temperatures (65°C, Fig. 2.10, lane 1). The bifunctional *PfAdometdc/Odc* gene was also amplified with LD-PCR with the Sampetf1:ODCexpr primer pair on genomic DNA isolated from *P. falciparum* cultures (Fig. 2.10, lane 2). The 4260 bp single, distinct band was obtained indicating that there are no introns present in the gene and that it is possible that there is only one genomic copy for this gene.

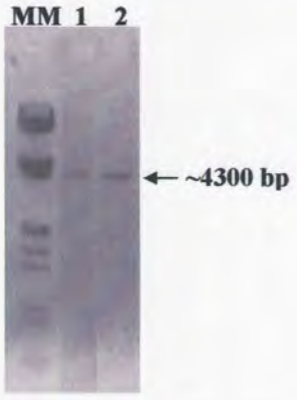


Figure 2.10: Amplification of the full-length bifunctional *PfAdometdc/Odc*. The ~4300 bp cDNA band amplified from 7.86 ng mRNA-cDNA equivalents of the uncloned cDNA library in lane 1 and in lane 2, from 50 ng *P. falciparum* genomic DNA. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

2.5.2) Analyses of the nucleotide sequence of the full-length *PfAdometdc/Odc* gene and cDNA.

The *PfAdometdc/Odc* gene amplified from genomic DNA and the full-length *PfAdometdc/Odc* cDNA was cloned into the pGEM-T Easy vector and the nucleotide sequences determined using a primer-walking strategy using the primers described in Fig. 2.3. Three independent clones of each were sequenced twice.

Appendix I indicates an alignment of the nucleotide sequences of *PfAdometdc/Odc* derived from cDNA and genomic DNA compared to the Genbank sequence submitted by Müller *et al.* (Müller, *et al.*, 2000). The consensus nucleotide cDNA sequence is almost identical to the published sequence with only a few point mutations in the 3′-end of the *Odc* domain including a single transition and three A or T nucleotide insertions in a stretch of AAT-repeats without causing a frameshift. Virtually no difference was observed in the 5′-end of our consensus sequence compared with the sequence of Müller *et al.* (Müller, *et al.*, 2000). The *Adometdc* domain stretches from 1-1587 bp, the hinge region was defined by Müller *et al.* as the next 822 bp that does not show similarity to

other genes/proteins, and the *Odc* domain is the 1842 bp at the 3′-end (2418-4257 excluding the stop codon). The entire ORF of 4257 bp contains 75.8% A and T nucleotides. The transcript is therefore monocistronic and contains a single ORF encoding two decarboxylase functions. A single translation start site defined the bifunctional ORF with a sequence of AATAATG compared to consensus sequences of a quartet of A nt found before the ATG in other *Plasmodium* genes (Coppel and Black, 1998).

The 3′-UTR consists of 34 nt excluding the stop codon (TAA, residues 4257-5260) and poly-adenylation signal (AATAAA, residues 4284-4289, Appendix I). The poly-A signal is only 5 nucleotides removed from the start of the poly-A tail, which is characteristic of *Plasmodia* whereas this site is typically found 15-25 residues upstream of the poly-A tail in other mRNAs (Lanzer, *et al.*, 1993). The identification of the poly-A site indicates that the DD-Poly-T primer did not anneal internally in the cDNA, but at the correct position at the 3′-end. No obvious alternative poly-adenylation signal was apparent.

There was no observable difference in the nucleotide sequences obtained from genomic DNA compared to cDNA. This supports the fact that there are no introns present in the genomic DNA for the bifunctional *PfAdometdc/Odc* and that these two genes are found in *cis* in the genome on a single chromosome. Analyses of the *Plasmodium* Genome Sequencing Project (www.plasmoDB.org) data indicated 98-99% identity between the deduced amino acid sequences of the *PfAdometdc/Odc* full-length cDNA and a 4305 bp segment of chromosome 10 (chr10_1.glm_390, accession code 8584428). This is in an area of 14 350 nt of the chromosome (Fig. 2.11). Fragments of the cDNA sequence also had low levels of identity (<30%) with chromosomal segments from chromosomes 11, 14 and 13 but the areas showing homology was mostly in the AAT-repeats and poly-A or –T stretches. Various other open reading frames were identified in the vicinity of the *PfAdometdc/Odc* ORF, the majority of which were identified as hypothetical *P. falciparum* proteins.

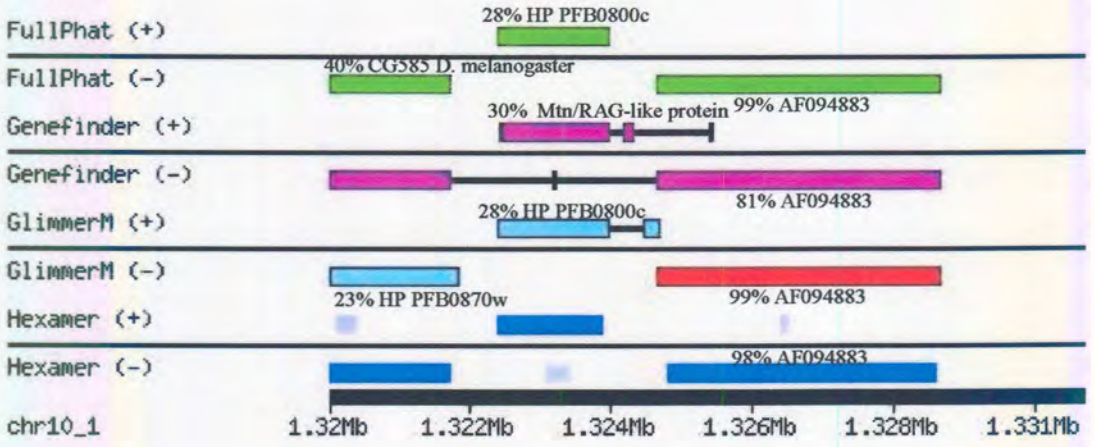


Figure 2.11: Analyses of chromosome 10 of *P. falciparum* containing the full-length ORF for the bifunctional *PfAdometdc/Odc* (red). ORF predictions were performed with different programs from PlasmoDB on chr10_1.glm_390. The location on the chromosome is indicated with the ruler at the bottom. Various other ORFs are present in the chromosomal segment and their homologies to known genes are indicated. HP: hypothetical protein, PF: *P. falciparum.*

Analyses of the genome sequence upstream of the *PfAdometdc/Odc* ORF from PlasmoDB for possible transcription start sites and promoter areas were performed with two different programs. The predicted promoter areas and transcription start sites obtained with both programs indicated an 250 bp area between residues –2808 and –2558 upstream of the ORF as indicated in Fig. 2.12.

-2808
attttttatttcttataaaaaagcaaatattttaaacaatattataacaaaagttattttata
aaaaaatatatattacagtaaaaaaataaaaattttattttatatcaaaaatatatatatatat
atatatatataaaaaaaaaaaaacgtgttcttgg**A**ttttttatataatataataaaaaaatatt
cataataccccccaaaaaaccc**ataaa**ggaaaaaaaaacaaaaaattaa**A**tatatag
-2558

Figure 2.12: Predicted promoter area (250 bp) for *PfAdometdc/Odc*. The region between residues –2808 and –2556 are indicated. Underlined areas indicated putative TATA-boxes and the predicted transcription start sites are indicated in bold (ProScan and Neural Network Promoter Prediction Server).

The transcription start site was predicted at either position –2645 (Neural Network Promoter prediction) or –2565 (Proscan). The predicted promoter region is extremely A+T rich (88%), with long stretches of AT-tracks. Putative TATA-boxes were predicted; the first TATA-box 24 nt upstream of the first predicted transcription start site, and the second 22 nt removed from the second start site.

The identification of a putative transcription start site indicates a 5′-UTR of ~2600 nt. This could result in a transcript size of 7022 nt including the ORF of 4257 nt, the 37 nt

3′-UTR (34 nt and the stop codon), the poly-A tail and the predicted maximum length of the 5′-UTR of 2645 nt. This corresponds to the 7 kbp band observed with the Northern blot (Fig. 2.9). However, the 5′ UTR could not successfully be amplified and sequenced from cDNA using our procedure (results not shown). Analyses of 2600 nt of the predicted 5′-UTR of the *PfAdometdc/Odc* ORF in the genomic DNA with MFOLD indicated marked secondary structures present in this area that would make up part of the 5′-UTR of the *PfAdometdc/Odc* mRNA (Fig. 2.13). The ΔG of the structure was –1595.2 kJ/mol. No secondary ORF was predicted in the 5′-UTR.

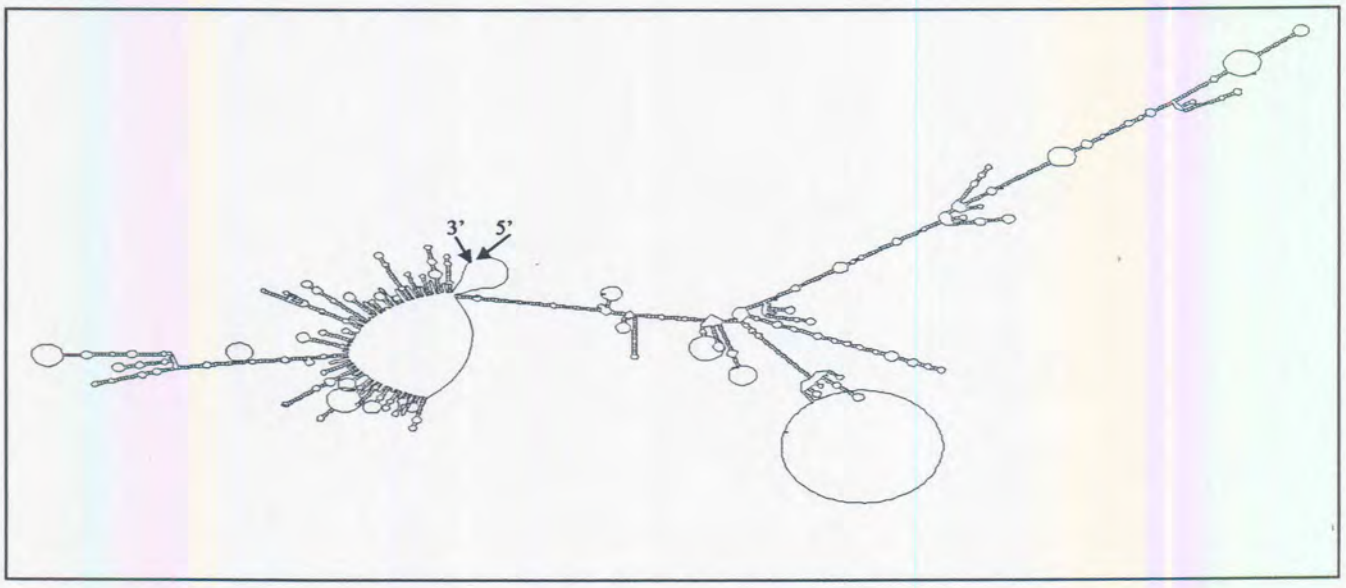


Figure 2.13: Secondary structure prediction of the ~2600 bp 5′-UTR of the bifunctional PfAdoMetDC/ODC. The 3-' and 5'-ends are indicated.

2.5) DISCUSSION.

This discussion is based on the results presented in both Parts I and II.

2.5.1) Design of *Adometdc* and *Odc*-specific degenerate primers for 3′-RACE.

The construction and design of the *Adometdc* and *Odc*-specific primers used in the 3′-RACE were based on sequence data for the proteins characterised in related organisms. The number of organisms chosen in the multiple alignments for the identification of the conserved areas was limited to only 13 for ODC and 12 for AdoMetDC based on the close resemblance of *P. falciparum* to higher rather than lower eukaryotes (Hyde, *et al.*, 1989). The primer sequences were based on two identified conserved areas (Fig. 2.5). The degeneracies of the primers were limited to <128 by using the MI values of the codon preferences of *P. falciparum* and by the inclusion of inosine (Hyde, *et al.*, 1989).

2.5.2) Identification of the *Odc* and *Adometdc* cDNAs with 3'-RACE.

In the absence of any sequence data for the *Odc* and *Adometdc* cDNAs when this study was initiated, 3'-RACE was preformed using degenerate primers as discussed above. The use of the in-house designed uncloned cDNA library led to improved amplification of the 3'-RACE products as evidenced from the *Odc* amplifications (Fig. 2.4). These products were identified as *Odc* cDNAs based on the homology of their nucleotide sequences to a *P. falciparum Odc* cDNA fragment sequence deposited in Genbank. However, the *Adometdc* cDNA was not amplified using this strategy. Multiple 3'-RACE products ranging in size from 3500 bp to 850 bp were obtained that could not be resolved after multiple optimisations (Fig. 2.6), even when amplification was performed at a lower extension temperature of 60°C to amplify sequences that are normally refractory to PCR (Su, *et al.*, 1996). Retrospective analyses of the conserved area used in the design of the degenerate primer Samdcd1 indicated that this area is not as highly conserved in *P. falciparum*: TIHITPE in the other organisms were found as CVHYSPE in *P. falciparum* (conserved residues in bold). This could have contributed towards the non-specific priming seen with this primer as was evident after sequencing the 850 bp band (in the expected size range). The identity of the larger bands was not pursued but it is conceivable that the ~3500 bp band could contain both the *Adometdc* and *Odc* cDNAs.

2.5.3) Analyses of the mRNA transcript of *Odc*.

The transcript size for *Odc* was analysed with a Northern blot and indicated a single 7 kbp RNA fragment (Fig. 2.9). Only this single, distinct band was obtained in contrast to *Odc* transcripts from other organisms where two copies of the mRNA is present due to alternative use of a polyadenylation signal (Heby and Persson, 1990). However, the transcript is extremely large, almost three times the size (7 kbp vs. ~ 2 kbp) of other eukaryotic *Odc* mRNAs (Heby and Persson, 1990). The large size of the transcript was shown by Prof Walters' group in Germany to contain a single open reading frame for both *Adometdc* and *Odc* (Müller, *et al.*, 2000). However, this still results in a ~2600 bp 5'-UTR, a size corresponding to the predictions of the regulatory elements in the transcript (section 2.5.2) but more than double the size of other malarial transcripts, which are usually characterised by 5'-UTRs between 500 and 1000 bp long (Coppel and Black, 1998).

2.5.4) 5′-RACE of *Adometdc* and *Odc*.

The uncloned cDNA library was used to amplify the upstream regions of the *Odc* and *Adometdc* cDNAs in 5′-RACE strategies. No product could be obtained with 5′-RACE using the uncloned cDNA library and reverse primers directed against either *Adometdc* or *Odc*, even after extensive optimisations of various amplification parameters including primer:template ratios and cycle parameters (results not shown). Multiple faint bands were observed for the *Odc* 5′-RACE products, which could fortuitously be identified as *Odc* cDNA fragments with nested-PCR protocols (Fig. 2.7). Unfortunately, the same strategy could not be applied for the *Adometdc* cDNA. Analyses of the predicted 5′-UTR for the bifunctional PfAdoMetDC/ODC transcript (Fig. 2.12) indicated extensive secondary structures spread along the whole ~2600 bp. It is known that *Adometdc* and *Odc* mRNAs from other organisms contain a >200 nt long, GC-rich 5′-UTR with thermodymically stable secondary structures (-289 kJ/mol)(Heby and Persson, 1990). The higher ΔG observed for the *PfAdometdc/Odc* transcript (-1595.2 kJ/mol) could be due to its remarkable length (13 times longer 5′-UTR than that of mammalian mRNAs). It is therefore conceivable that these secondary structures in the 5′-UTR of the *PfAdometdc/odc* transcript could reduce the efficacy of the reverse transcription reaction resulting in a reduction of full-length cDNA (i.e. ~7 kbp) and consequently, the relative abundance of the transcript. Probing the uncloned cDNA library with e.g. the *Odc*-specific probe would have indicated if full-length cDNAs for this transcript were synthesised. Further manipulation of the conditions of 5′-RACE, including extended elongation times (from 2 min to around 7 min) to accommodate for the length of the transcript, could in retrospect have resulted in products, the size of which would have been much larger than expected at the start of the study.

2.5.5) Amplification of the full-length bifunctional *PfAdometdc/Odc* cDNA.

The full-length bifunctional cDNA for *PfAdometdc/Odc* was amplified from the uncloned cDNA library as well as from genomic DNA isolated from *P. falciparum*. The equivalent size products were obtained in both instances and sequence analyses of these products indicated a single ORF encoding both decarboxylase functions with a base content of 75.8% A and T nucleotides, slightly higher than the mean 60-70% A+T content of *P. falciparum* genes (Scherf, *et al.*, 1999). Furthermore, the gene does not contain any introns. This is not unusual as the majority of *P. falciparum* genes contain no or only one intron (Gardner, *et al.*, 1998; Lanzer, *et al.*, 1993) and furthermore, the *Trypanosoma* and yeast *Odcs* also have no introns (Heby and Persson, 1990).

The obtained PCR fragments for *Odc* and *Adometdc* as well as the full-length PfAdoMetDC/ODC was cloned and the nucleotide sequences determined. Various precautions were used during the amplification of the long cDNA fragments to compensate for the A+T-richness of the *P. falciparum* genome. These included lowering of the extension temperatures during amplification, the avoidance of UV light during cloning procedures, using recombinant negative *E. coli* cell lines grown at 30°C only to the exponential phase to limit rearrangements and where possible using direct PCR product nucleotide sequencing, thereby eliminating problem-prone cloning strategies (Adams, *et al.*, 1993; Hanahan, *et al.*, 1991). Similar observations have been included in texts on cloning and amplification of malaria genes (Coppel and Black, 1998; Su and Wellems, 1998).

2.5.6) Genomic structure of *PfAdometdc/ODC* gene and structure of the single transcript.

The unique single *Adometdc/Odc* transcript of *P. falciparum* is monocistronic and there is no evidence for multiple copies of genes. All *P. falciparum* transcripts are monocistronic implying the presence of regulatory sequence elements flanking the coding regions (Horrocks, *et al.*, 1998). The enormous amount of data now available from the *Plasmodium* genome sequencing project as well as the monocistronic arrangement of *P. falciparum* genes allowed preliminary predictions on the gene structure and flanking regions. Putative promoter regions were identified for *PfAdometdc/Odc* and contained two TATA-boxes, but no other significant transcription initiation elements could be predicted. The ambiguous assignment of the TATA-boxes and presence of two possible transcription elements is complicated by the extraordinary high A+T content of these regions, also found in other upstream regions of *Plasmodial* genes (Lanzer, *et al.*, 1993). However, the promoter contains long tracks of these nucleotides, a characteristic common to other *P. falciparum* promoter sequences. These areas can adopt a unique rigid structure of a propeller twist, which may stimulate transcriptional activity (Horrocks, *et al.*, 1998).

The identification of putative transcription start sites allowed the identification of a possible 5'-UTR of ~2600 bp for the transcript. The size of this UTR is in agreement with both the long UTR sizes of *Odc* and *Adometdc* transcripts found in other organisms (>200 nt), as well as the fact that most malaria genes contain long 5'-UTRs (500-1000

nt)(Coppel and Black, 1998; Su and Wellems, 1998). Possible explanations for the length include a role in transcript stability (Su and Wellems, 1998), or in regulation of expression by interactions with regulatory proteins or ribosomes (Coppel and Black, 1998). The predicted 5′-UTR has an extremely high A+T content of 88%, characteristic of all *Plasmodium* transcripts (Scherf, *et al.*, 1999). No secondary ORF was predicted in this area using various prediction programs. Secondary ORFs are a characteristic of *Adometdc* transcripts in other organisms that mediate translational regulation through ribosome stalling (Raney, *et al.*, 2000). The 5′-UTR of the *PfAdometdc/Odc* does however contain significant secondary structures, which is a feature of all other *Odc* and *Adometdc* transcripts and might facilitate translational regulation mediated by polyamine stabilisation (Cohen, 1998). The thermodynamically stable secondary structure could explain the inability to sufficiently amplify any full-length 5′-RACE products for *PfAdometdc/Odc* as mentioned in the previous section. Collectively, the characteristics of the predicted 5′-UTR suggests that the translation of *PfAdometdc/Odc* is not regulated in the same manner as the single *Adometdc* or *Odc* transcripts in other organisms via ribosome stalling but that feedback regulation of polyamines through stabilisation of the secondary structures could mediate unique translational regulatory mechanisms. These predictions will be validated once the complete nucleotide sequence of the 5′-UTR has been conclusively established. The transcription start site can also be determined and validated by S1 nuclease mapping, primer extension analyses and RNAse protection assays (Horrocks, *et al.*, 1998).

The 3′-UTR of the transcript contains all the required signature sequences, including the poly-adenylation signal only 5 residues removed from the start of the poly-A tail. This short interspersing region is a characteristic of *Plasmodia* compared to the longer 15-25 residues in mRNAs from other eukaryotes (Lanzer, *et al.*, 1993). *Odc* and *Adometdc* from eukaryotes both have two transcripts of varying lengths, due to the use of alternative polyadenylation sites (Kahana, 1989; Pulkka, *et al.*, 1991). However, no alternative polyadenylation site was obvious for the *PfAdometdc/Odc*. The physiological advantage of the alternative polyadenylation has not been determined, it is therefore unclear if the lack of this characteristic in *PfAdometdc/Odc* will have any biological relevance.

From the results presented in this chapter, a structural organisation for the *PfAdometdc/Odc* gene and its mRNA is proposed and summarised in Fig. 2.14. From

the *Plasmodium* genome sequencing project, the chromosomal location and ~9880 nt flanking both sides of the gene (5001 nt upstream and 4879 nt downstream) was identified. A putative promoter region of 250 nt was indicated upstream of the transcription start site. The primary transcript (~ 7 kb) contains the single ORF of 4257 nt encoding both the *Adometdc* and *Odc* functions divided by a hinge region. Furthermore, the transcript contains a long 5′-UTR of ~ 2600 bp and downstream, a short 37 nt 3′-UTR containing the stop (TAA) and poly-adenylation signal (AATAAA).

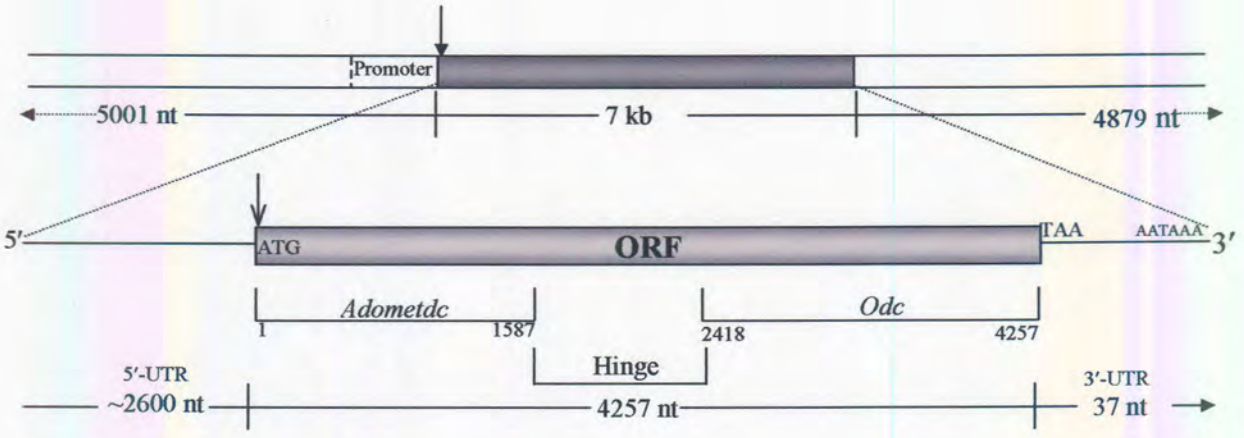


Figure 2.14: Schematic representation of the chromosomal organisation and general structures of the *PfAdometdc/Odc* gene and its corresponding mRNA. The top panel indicates the structure of the *PfAdometdc/Odc* gene without any introns and the 250 nt putative promoter region. The primary transcript is indicated in the bottom panel and contains long 5′- and 3′-UTRs and an ORF of 4257 nt. ATG: start codon, TAA: stop codon, AATAAA: poly-adenylation signal. The putative transcription start site is indicated with the filled arrow, and the translation start site with the open arrow.

Chapter 3 describes the heterologous expression of the full-length PfAdoMetDC/ODC protein as well as the monofunctional PfAdoMetDC and PfODC domains based on the cDNA sequence described in this Chapter.