Functional and structural characterization of the unique bifunctional enzyme complex involved in regulation of polyamine metabolism in *Plasmodium falciparum*

by

Lyn-Marié Birkholtz

Submitted in partial fulfilment of the requirements of the degree

Philosophiae Doctor

Department of Biochemistry
School of Biological Sciences
Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

Oct 2002

ACKNOWLEDGEMENTS

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

TABLE OF CONTENTS

LIST OF FIGURES

3D	Three-dimensional
A	Adenosine
ACD	Available chemicals directory
AdoMet	S-adenosylmethionine
AdoMetDC	S-adenosylmethionine decarboxylase
AMA	Apical membrane antigen
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
AzBE	Antizyme binding element
bp	Base pair
BCBD	N^1N^4-bis(7-chloroquinoline-4-yl)butane-1,4-diamine
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
C	Cytosine
cAMP	Cyclic adenosine monophosphate
CARP	Clustered Asp rich protein
CCD	Charge coupled devise
cDNA	Complementary DNA
CSD	Cambridge structure database
CS	Circumsporozoïte
C-terminal	Carboxy terminal
dAdoMet	Decarboxylated S-adenosylmethionine
dNTP	Deoxyribonucleotide triphosphate
DD-Poly-T	Differential display poly-T primer
DDT	Dichlorodiphenyltrichloro ethane
ddUTP	Dideoxyuridine triphosphate
DEPC	Diethyl pyrocarbonate
DHFR	Dihydrofolate reductase
DHODH	Dihydroorotate dehydrogenase
DHPS	Dihydropteroate synthetase
DIG	Digoxigenin
DMF	Dimethylformamide
DMFO	DL-α-difluoromethyl ornithine
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
ds	Double-stranded
DTT	Dithiotreitol
dUTP	Deoxyuridine triphosphate
EBA	Erythrocyte binding protein
EDTA	Ethanol diamine tetra-acetic acid
EtBr	Ethidium Bromide
FP	Ferriprotoporphyrin IX

G	Guanidine
G6-PD	Glucose-6-phosphate dehydrogenase
GM-CSF	Granulocyte macrophage colony stimulating factor
GPI	Glycophosphatidyl inositol
GRASP	Graphical representation and analyses of structural properties
GSP	Gene specific primer
GTP	Guanidine triphosphate
HABA	4-hydroxy azobenzene-2-carboxylic acid
HGPRT	Hypoxanthine-guanosine phosphoribosyltransferase
HIV	Human immunodeficiency virus
HRP	Histidine rich protein or horseradish peroxidase
HSP	Heat shock protein
I	Inosine
ICAM	Intracellular adhesion molecule
IFN	Interferon
IL	Interleukin
IMAC	Immobilised metal affinity chromatography
IMP	Inosine monophosphate
IPTG	Isopropyl-D-galactoside
IUBMB	International Union for Biochemistry and Molecular Biology
kb	Kilobase/kilobasepairs
LB	Luria Berthani
LDH	Lactate dehydrogenase
LD-PCR	Long-distance PCR
MAOPA	5'-[(3aminooxypropyl)methylamino]-5'-deoxyadenosine
MDR	Multi-drug resistance
MGBG	Methylglyoxal bis(guanylhydrazone)
MHC	Major histocompatibility complex
MI	Match index
mopp-DFB	1-methyl-3-oxo-3-phenyl difluoridoborate
MOPS	Morpholinopropanesulphonic acid
mRNA	Messenger RNA
MSA	Merozoïte surface antigen
NBT	Nitroblue tetrazolium chloride
NCBI	National Center for Biotechnology Information
NCI	National Cancer Institute (USA)
Ni-NTA	Nickel-nitrolotriacteric acid
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NOS	Nitric oxide synthase
nt	Nucleotide
N-terminal	Amino terminal
OAT	Ornithine aminotransferase
OD	Optical density

| ODC | Ornithine decarboxylase |
| ORF | Open reading frame |

PASS	Prediction of the biological activity spectra of substances
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein databank
PEG	Poly-ethylene glycol
PfAdoMetDC/ODC	*P. falciparum* S-adenosylmethionine decarboxylase/ornithine decarboxylase
Pfcrt	chloroquine resistance transporter
PfEMP	*P. falciparum*-infected erythrocyte membrane protein
PIR-PSD	Protein information resource-protein sequence database
PLP	Pyridoxal 5'-phosphate
PMSF	Phenylmethylsulfonyl fluoride
PPMP	*dl*-threo-1-phenyl-2-palmitoylamino-3-morpho-1-propanol
PPP	Pentose phosphate pathway
PPPK	Dihydroxymethylpterin pyrophosphokinase
PVM	Parasitophorous vacuolar membrane

RACE	Rapid amplification of cDNA ends
RAP1	Rhoptry-associated protein
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RNAse	Ribonuclease
RR-MAP	Methylacetylenicputrescine
RT	Reverse transcription
RT-PCR	Reverse transcription PCR

SCOP	Structural classification of proteins
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SE-FPLC	Size-exclusion fast protein liquid chromatography
SE-HPLC	Size-exclusion high-pressure liquid chromatography
SMART	Simple modular architecture research tool
STARP	Sporozoite Thr and Asp rich protein

T	Thymidine
$t_{1/2}$	Half-life
T_m	Melting temperature
TAE	Tris-acetate EDTA
TBS	Tris buffered sodium
TE	Tris EDTA buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
TIM	Triosephosphate isomerase
TMAC	Trimethylammonium chloride
TNF α	Tumor necrosis factor α
TS	Thymidylate synthetase
TVM	Tubovesicular membrane network

| UTR | Untranslated region |
| UV | Ultraviolet |

VCAM	Vascular cell adhesion molecule
WHO	World Health Organisation
X-gal	5-bromo-4-chloro-indolyl-β-D-galactoside

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

1

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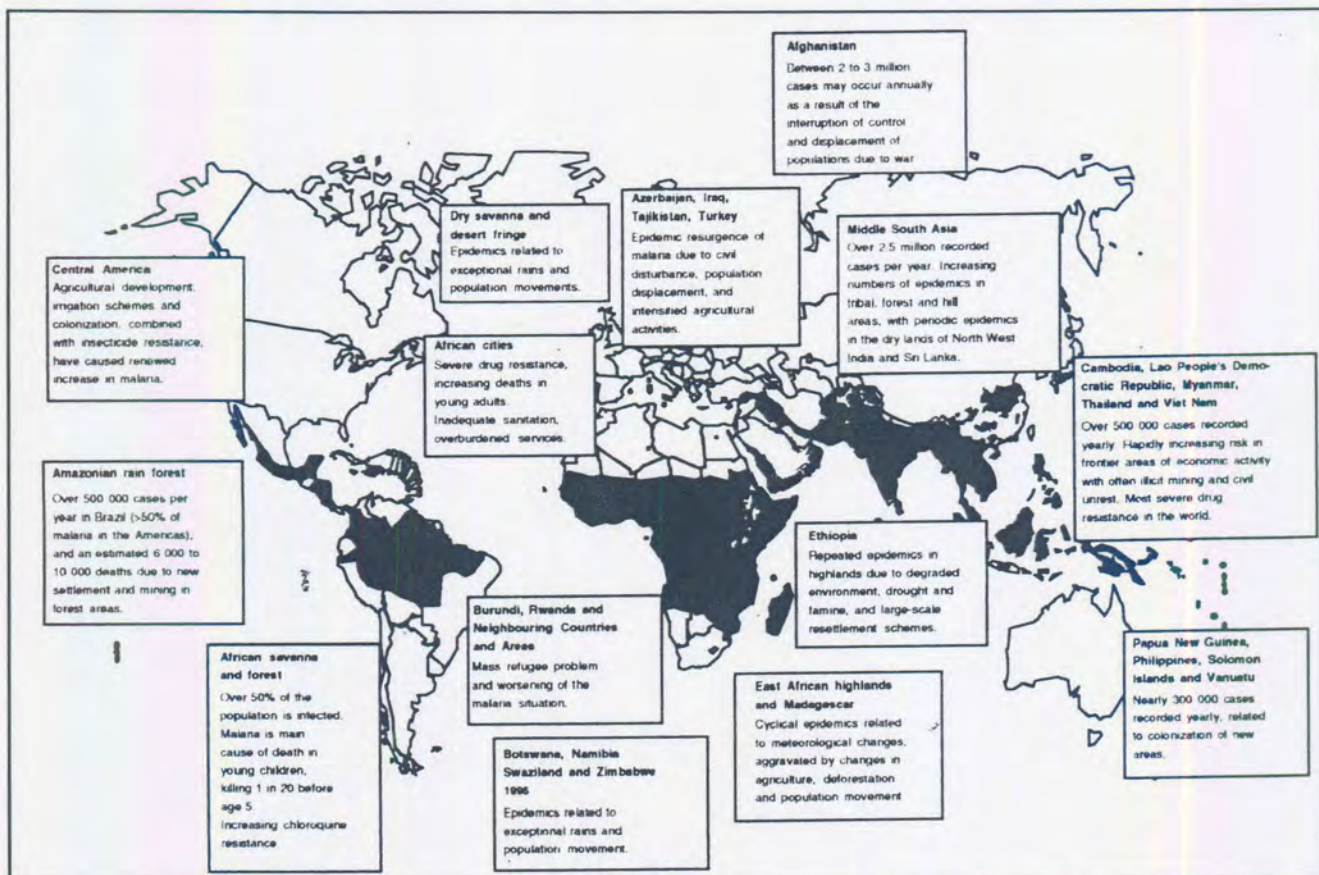
Figure 1.1: Malaria distribution and problem areas. Shading indicates the main areas where malaria transmission occurs. Data from the WHO/CTD/Health Map, 1997.

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

3

capable of developing into gametocytes, which are in turn ingested by a mosquito to complete the cycle (Cox, 1993; Fujioka and Aikawa, 1999).
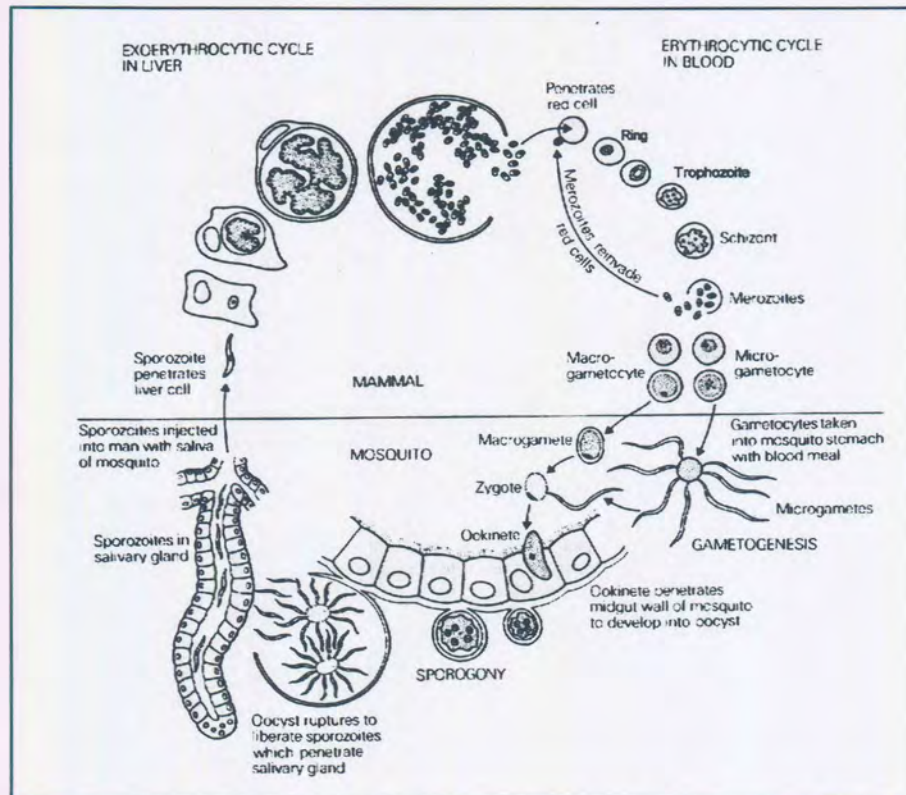
Figure 1.2: Bi-phasic life cycle of the *Plasmodium* parasite. Adapted from (Cox, 1993).

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).
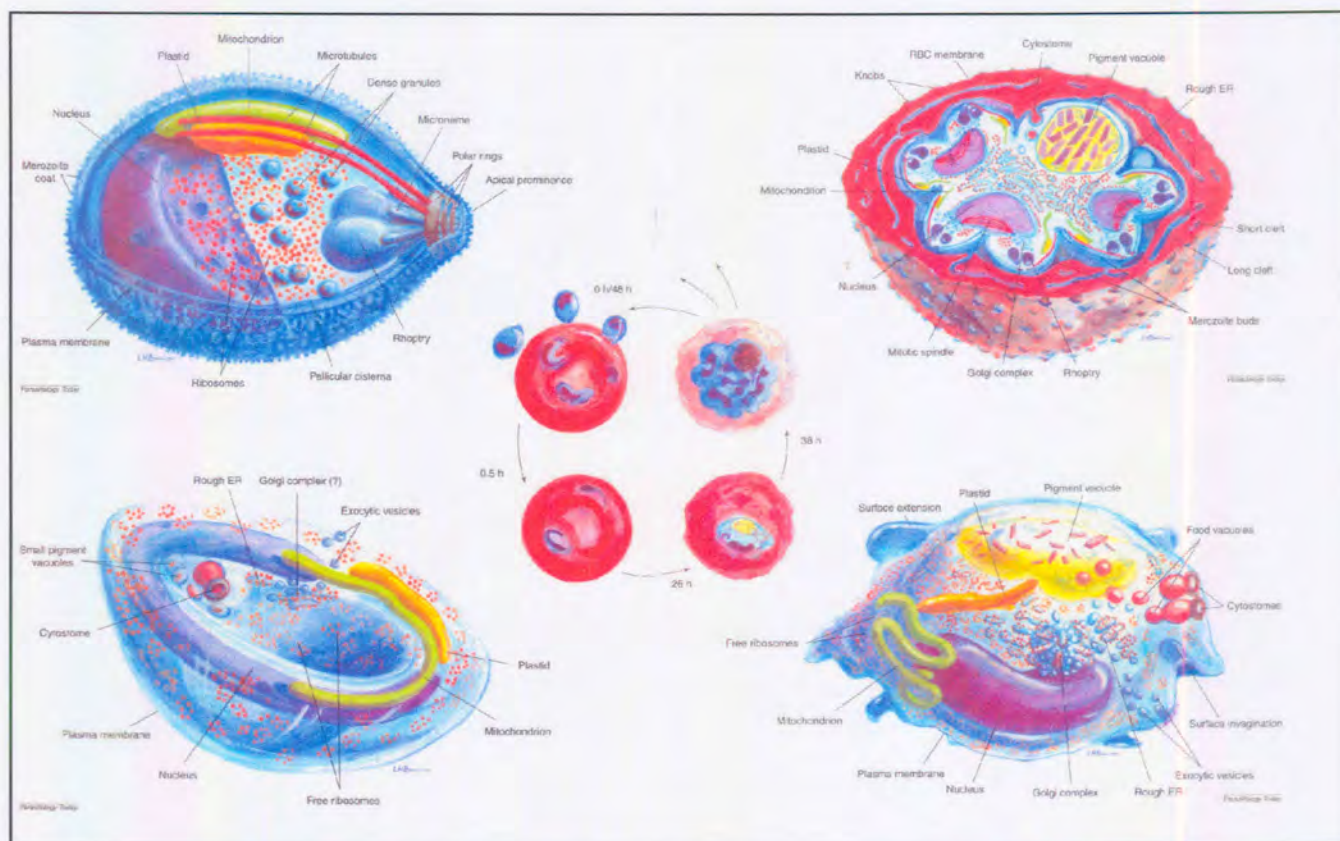
Figure 1.3: Three-dimensional representations of the ultrastructure of the different erythrocytic stages of *P. falciparum*. From top left anticlockwise: merozoite, ring, trophozoite and schizont stages. ER: endoplasmic reticulum. Insert: relative sizes of the different stages inside the erythrocyte. 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.
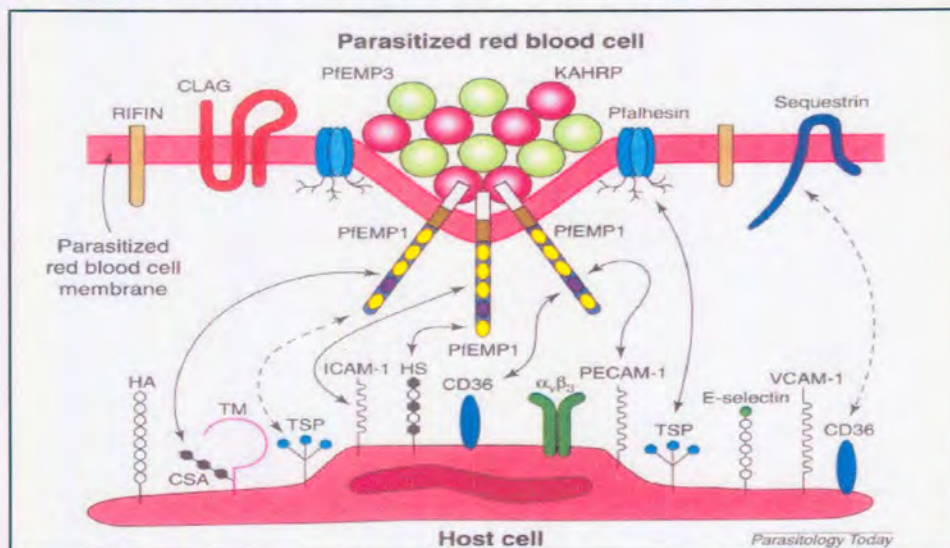
Figure 1.4: Schematic representation of the interaction at the cytoadhesive interface between a *P. falciparum* infected erythrocyte and the host vascular endothelium. CLAG: cytoadherence-linked asexual gene; CR1: complement receptor I; CSA: chondroitin sulphate A; HA: hyaluronic acid; HS: heparan sulphate; 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.
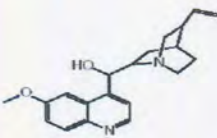
1.4.1.1) Blood schizontocytes
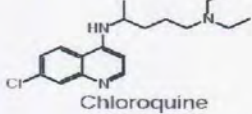
1.4.1.1.1) Quinoline-containing drugs
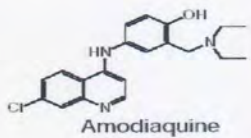
Malaria has been treated successfully for several hundred years with quinine, the active reagent in the bark of the South American Cinchona tree (*Cinchuona ledgeriana*) (WHO, 1996). Modern cinchona alkaloids or quinoline-type drugs include the type I 4-amino-quinolines (chloroquine, amodiaquine, mecaprime and sontaquine) and type II aryl-amino-alcohols (mefloquine and halofantrine) (Olliaro, 2001).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

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	Resistance
Sulphadoxine		
Dapsone		
Type II antifolates		
Pyrimethamine	Dihydrofolate reductase (DHFR) inhibition	Resistance
Proguanil		
Trimethoprim	Inhibition of	
Naphtoquinones (Atovaquone)	mitochondrial function	Compliance
Antimicrobials	Inhibition of	
Tetracycline, doxycycline, clindamycin and Azithromycin	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).
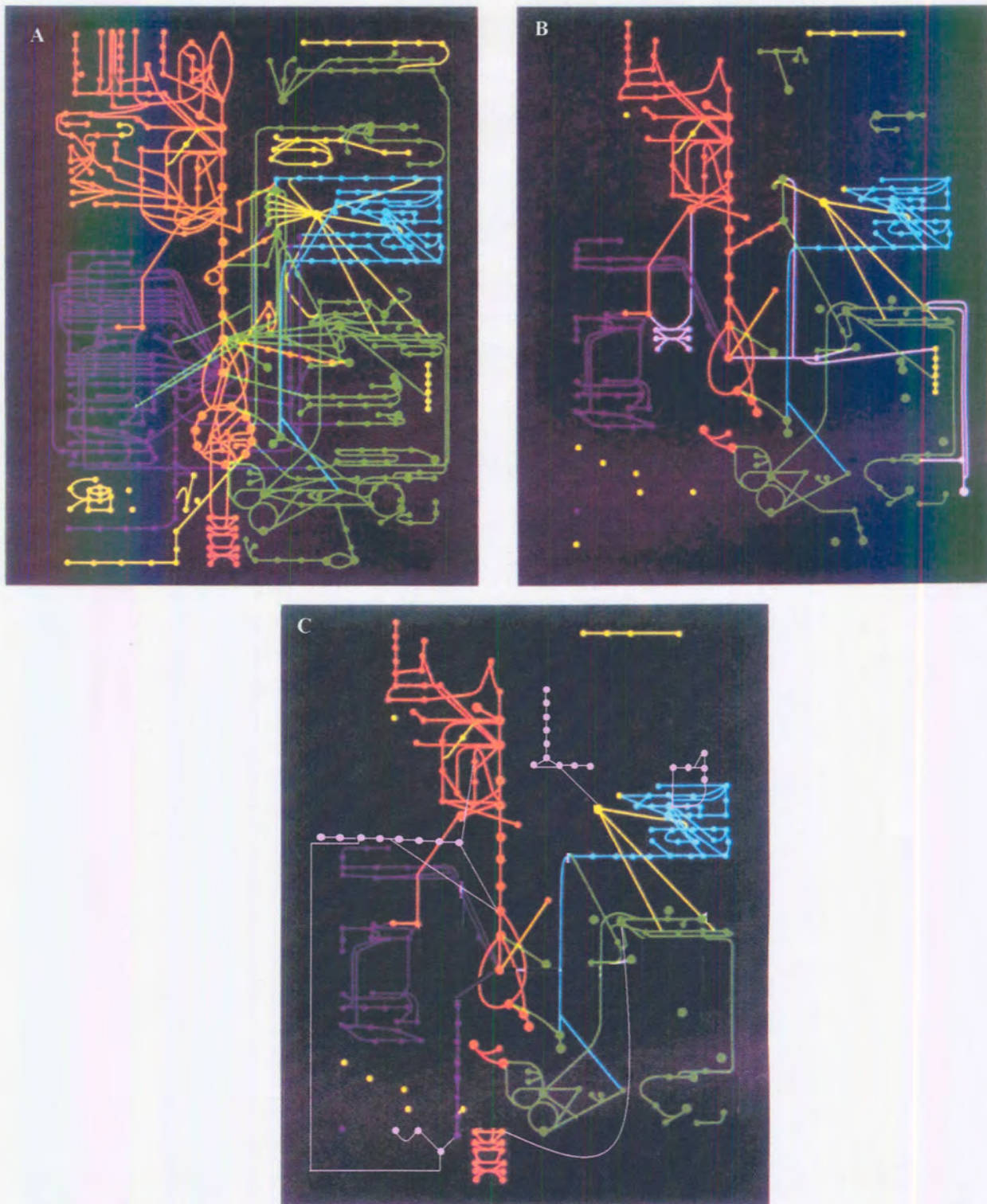
Figure 1.6: Global comparison of metabolomes. (A) A typical mammalian cell; (B) *T. brucei* and in (C) the newly constructed metabolome of *P. falciparum*. Orange indicate carbohydrate metabolism and electron transport, purple for fatty acid and sterol metabolism, green for amino acids and related metabolites, 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.
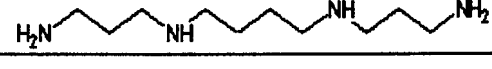
Names	Structure	Shorthand representation
Putrescine	$H_2NCH_2CH_2CH_2CH_2NH_2$	
Cadaverine	$H_2NCH_2CH_2CH_2CH_2CH_2NH_2$	
Spermidine	$H_2NCH_2CH_2CH_2CH_2NHCH_2CH_2CH_2NH_2$	
Spermine	$H_2NCH_2CH_2CH_2HNCH_2CH_2CH_2CH_2NHCH_2CH_2CH_2NH_2$	

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.
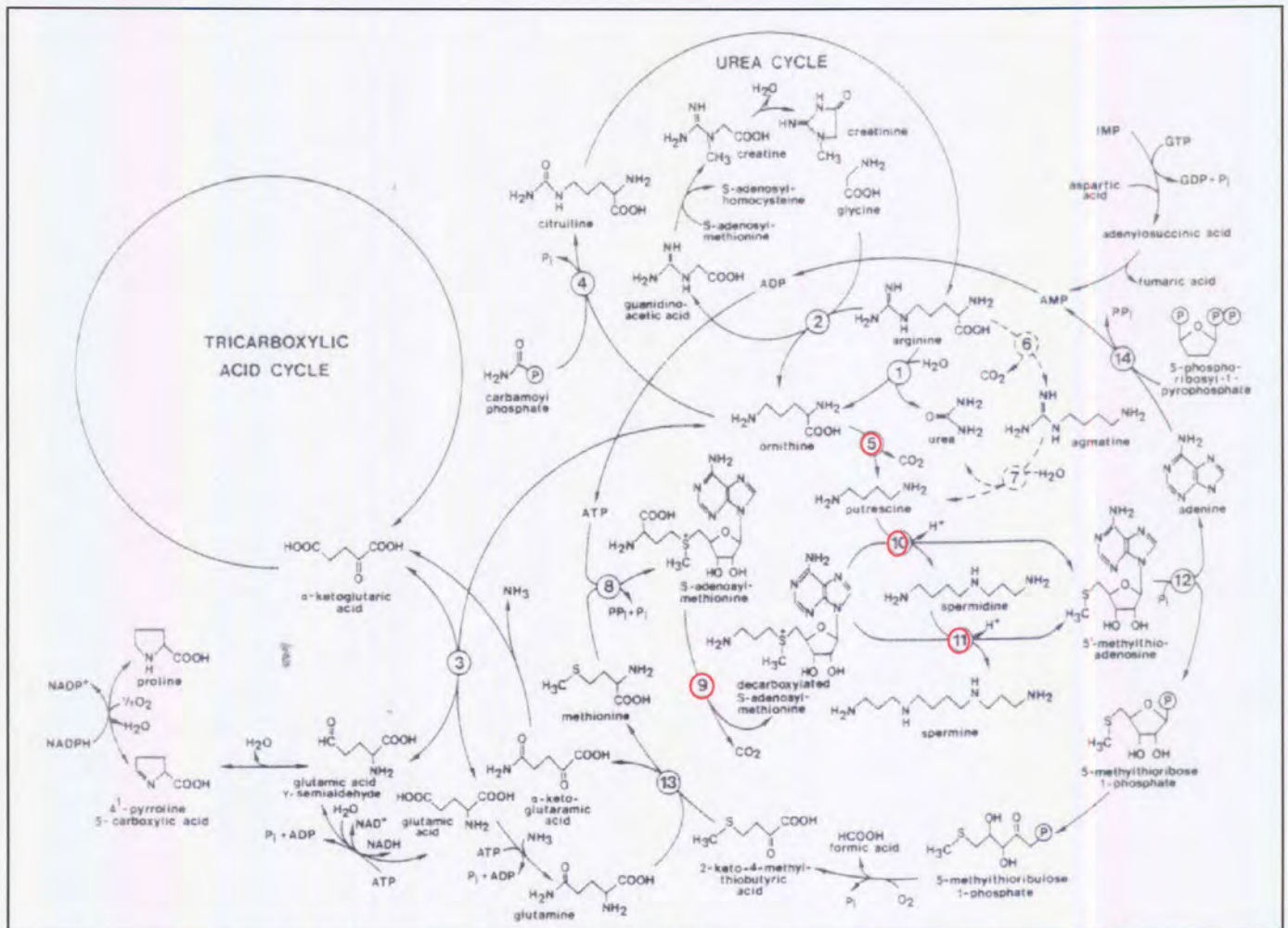
Figure 1.8: General pathway for the biosynthesis of the polyamines in pro- and eukaryotes and its linkage to the urea cycle, tricarboxylic acid cycle and methionine and adenine salvage pathways. 1) Arginase, 2) transamidinase, 3) ornithine aminotransferase, 4) ornithine carbamoyltransferase, 5) ornithine 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).
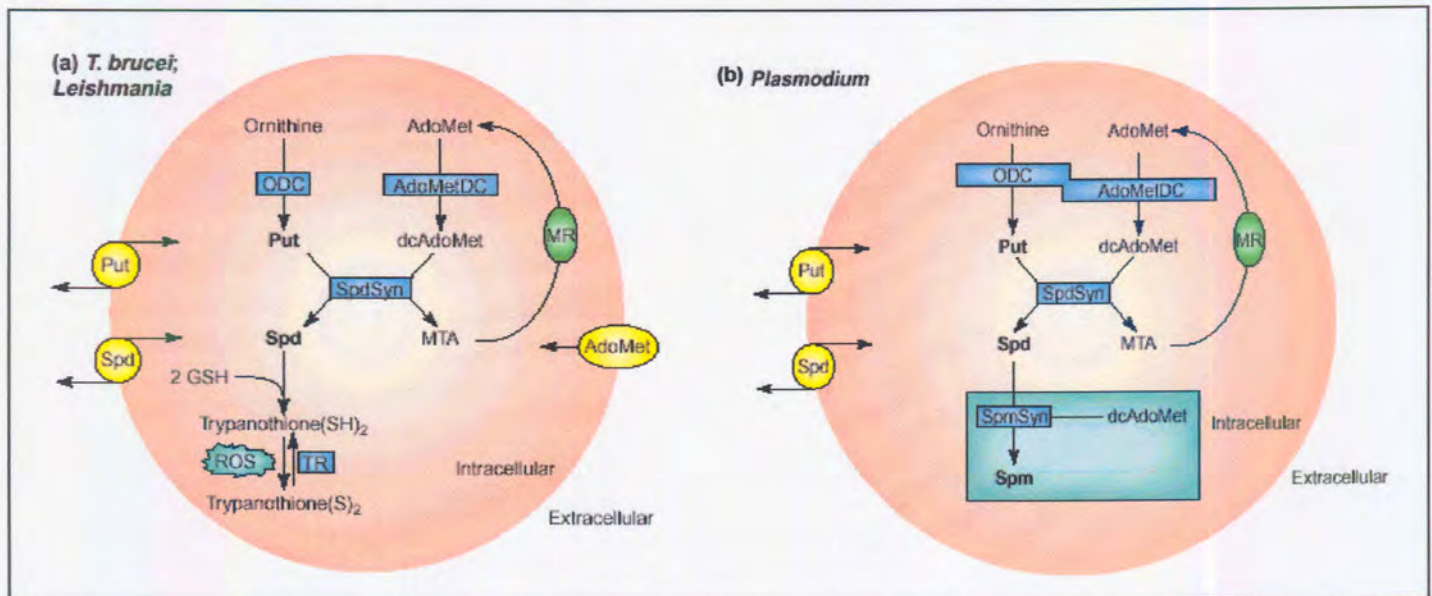
Figure 1.9: Polyamine metabolism in parasitic protozoa. (a) *T. brucei* and *Leishmania* and in (b) *Plasmodium*. ODC: ornithine decarboxylase, AdoMetDC: S-adenosylmethionine decarboxylase, SpdSyn: spermidine synthase, SpmSyn: spermine synthase, TR: trypanothione reductase, Put: putrescine, Spd: spermidine, 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

31

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:

$$AATAA\text{---}(N)_{0\text{-}23}\text{---}A\text{--------}(N)_{6\text{-}54}\text{----}G\text{+}T\text{-rich}\text{-----}$$
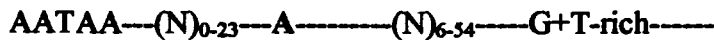
compared with the mammalian consensus sequence:

$$AATAAA\text{---}(N)_{10\text{-}23}\text{---}A\text{--------}(N)_{0\text{-}10}\text{----}G\text{+}T\text{-rich}\text{------}$$
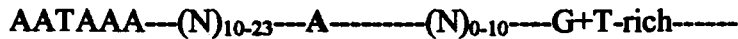
The G+T-rich sequence might therefore provide the signals for RNA processing and addition of the poly(A) tail in *Plasmodium* (Lanzer, *et al.*, 1993). An average poly-A tail is 100-200 nt long and the 3'-UTRs generally contain <400 nt (Coppel and Black, 1998; Su and Wellems, 1998).

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).
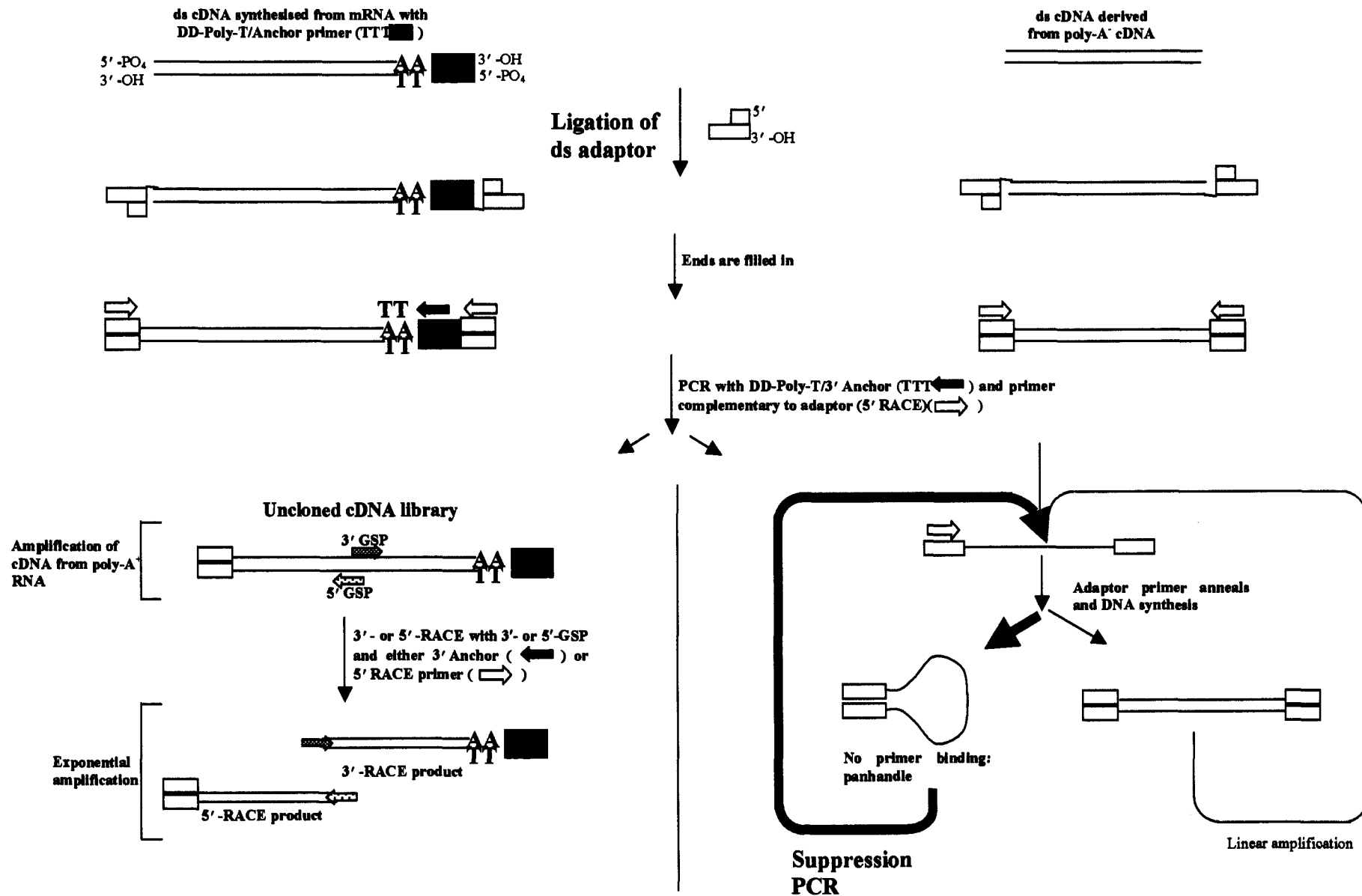
Figure 2.1: RACE protocols with double strand adaptor-ligated cDNA and suppression PCR. The adaptor is indicated with the open boxes. The DD-Poly-T and attached Anchor region are indicated with the solid boxes. The ligation to only the long strand of the adaptor is indicated with the small loop. The 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 *Eco*RI (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

47

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.

(A)

```
                                              200
Bovine         ATLKTSRLLLERAKELDIDVIGVSFHVGSGCTDPETFVQAISDARCVFDMGA
Human          ATLRTSRLLLERAKELNIDVVGVSFHVGSGCTDPETFVQAISDARCVFDMGA
Mouse          ATLRTSRLLLERAKELNIDVIGVSFHVGSGCTDPDTFVQAVSDARCVFDMAT
Shrew          ATLKTSRLLLERAKELNIDVIGVSFHVGSGCTDPETFVQAVSDARCVFDMGT
Rat            ATLKTSRLLLERAKELNIDVIGVSFHVGSGCTDPETFVQAVSDARCVFDMGT
Cricetulus     ATLRTSRLLLERAKELNIDVIGVSFHVGSGCTDPETFVQALSDARCVFDMGT
Chicken        ATLKTSRLLLERAKELDLAIVGVSFHVGSGCTDPETFVQAISDARCVFDMGA
Xenopus        ATLKTSRLLLERAKELNVDIIGVSFHVGSGCTDPQTYVQAVSDARCVFDMGA
Trypanosoma    AKVEDCRFILEQAKKLNIDVTGVSFHVGSGSTDASTFAQAISDSRFVFDMGT
Drosophila     ADVDAAALMLL-AKSLELKVTGISFHVGSGCSELQAYDRAIKKAKNLFKFGA
Ceanorhabditis DPIIAAPQLLKTASEEGINVVGISFHVGSGCNDASAYRNALQHAKNLCEIGE
Yeast          CEMENVDVLLKAIKELGLNLAGVSFHVGSGASDFTSLYKAVRDARTVFDKAA
Leishmania     APLEDVEGLLEAARQFNVTVQGVSFHVGSGNDDQSAYVSAVRDAYQVFQQAV
```

(B)

```
                                              330
Bovine         --IDATMFNPGYSMNGMKSD-GTYWTIHITPRPEFSYVSFET-------NLS
Human          --IDATMFNPGYSMNGMKSD-GTYWTIHITPRPEFSYVSFET-------NLS
Mouse          --IDATLFNPGYSMNGMKSD-GTYWTIHITPRPEFSYVSFET-------NLS
Rat            --IDATLFNPGYSMNGMKSD-GTYWTIHITPRPEFSYVSFET-------NLS
Drosophila     --IDDFLFNPGYSMNGINDK-GEYMTIHITPRNQFSYVSFET-------NVA
Dianthus       --ICDFDFEPGYSMNAIEG--PAVSTIHITPRDGFSYASFEAVG----YDLK
Nicotiana      --ICDFEFEPGYSMNSIEG--AALSTIHITPRDGFSYASFEAVG----YDMK
Arabidopsis    --ICDFEFEPGYSMNSIEG--DAISTIHVTPRDGFSYASFEAVG----YDFN
Trypanosoma    WKLHDLQFEPGYSINTIRG--AEYQTIHITPRDHCSFASYET-------NTP
Leishmania     WTLHDLQFEPGYSINAIRG--EEYQTMHITPRDHCSFASYET-------NSR
Onchocera      --VHEELFSPGYSLNGLIPHSDHYITIHVTPRPDFSYVSFET-------NQH
Yeast          FHHDAFAFTPGYSSNMILAE-KYYYTLHVTPRKGWSYASFESNIPVFDISQG
```

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.
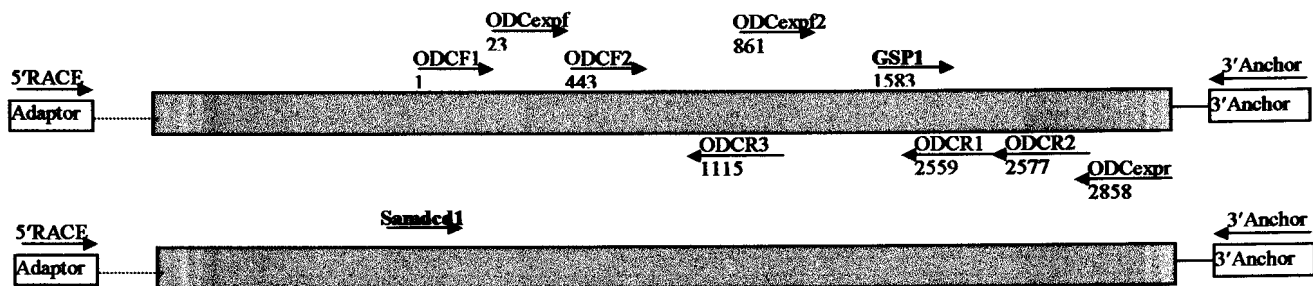
Figure 2.3: Schematic representation of the gene-specific primers used for amplification and nucleotide sequencing of the full-length *Adometdc* and *Odc* cDNAs. The proposed open reading frames (filled boxes) of unknown lengths of *Odc* (top) and *Adometdc* (bottom) are indicated. The positions and 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*	**GAA TTT** ATA CAA ACT ACT GAT G	22	51	3'-RACE product
ODCF1	3' RACE *Odc* & Primer-walking sequencing	**GAA TTT** TTA TAA TGG AAA GTA TAT G	25	52	Genbank (AF012551)
ODCF2	5' RACE *Odc*	**GTT 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).
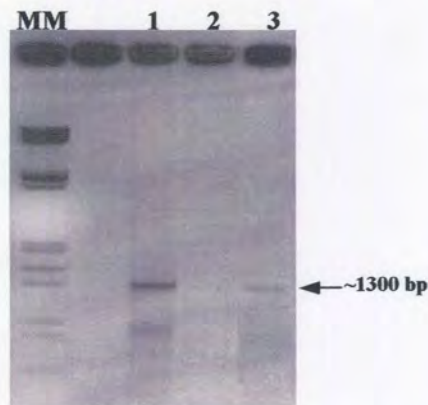
Figure 2.4: 3′-RACE PCR of the *Odc* cDNA with degenerate primer GSP1. 25 ng total RNA-cDNA equivalents of first-strand cDNA with the GSP1:3′ Anchor primer pair (lane 1), 7.86 ng mRNA-cDNA equivalents of double-stranded, adaptor-ligated amplified cDNA with GSP1:5′ RACE primer pair (lane 2) and 7.86 ng mRNA-cDNA equivalents of double-stranded, adaptor-ligated amplified cDNA with the GSP1:3′ Anchor primer pair (lane 3). **MM:** *Eco*RI-*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).
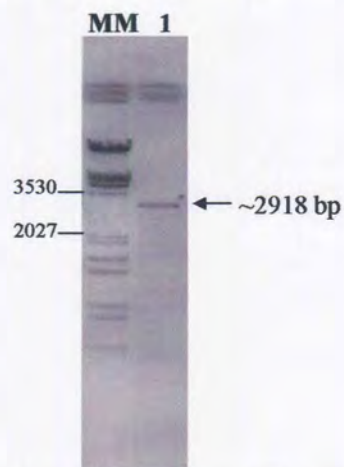
Figure 2.5: Amplification of the full-length *Odc* cDNA with 3′-RACE. The ODCF1:3′ Anchor primer pair was used in 3′-RACE on the uncloned cDNA library. The expected band of ~2918 bp was obtained. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

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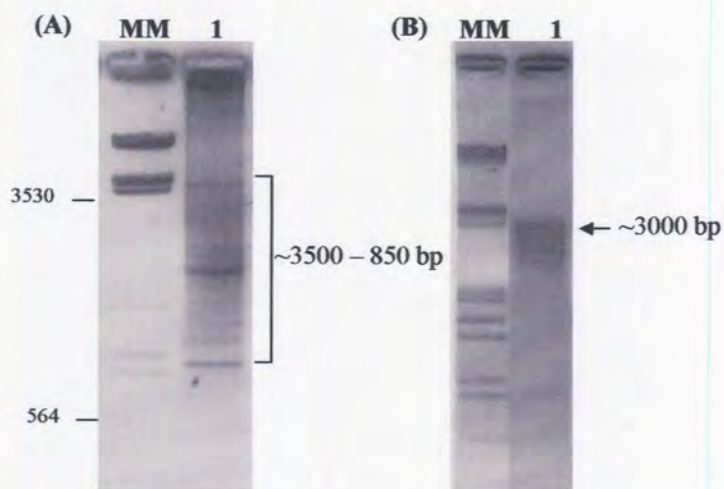
Figure 2.6: 3′-RACE of the *Adometdc* cDNA with degenerate primer Samdcd1. (A) Amplification with the Samdcd1:3′ Anchor primer pair on 7.86 ng mRNA:cDNA equivalents of the uncloned cDNA library at 72°C and in (B) at 60°C. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

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).
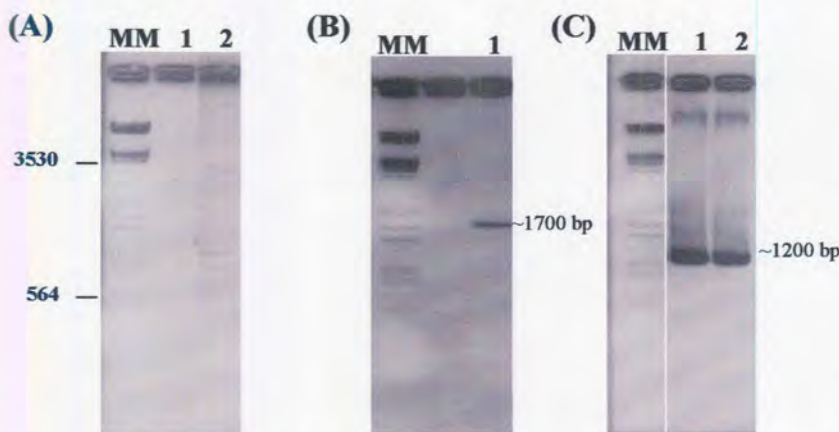
Figure 2.7: 5′-RACE of the *Odc* cDNA on the amplified, uncloned cDNA library and nested PCR strategy. (A) Several bands were observed with the ODCR1:5′RACE primer pair only with 15.72 ng mRNA-cDNA equivalents (lane 2) but not with 7.86 ng (lane 1). **(B)** The nested PCR was done with the ODCR1:ODCF1 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:** *Eco*RI-*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.
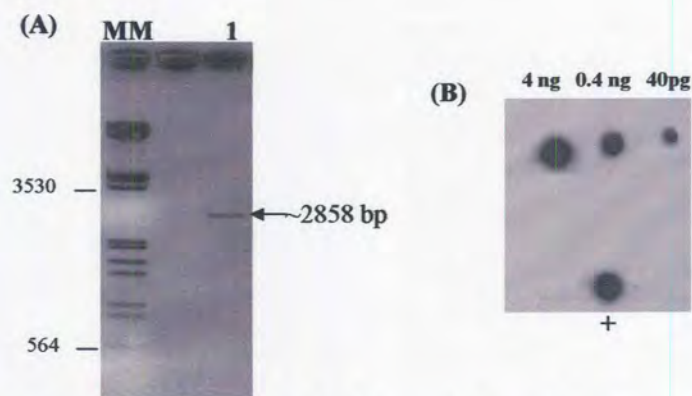
Figure 2.8: Synthesis of a DIG-labelled *Odc*-specific probe and quality analyses. (A) PCR-mediated synthesis of the 2858 bp *Odc*-specific probe. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker. **(B)** Dot-blot mock probe of the labelled *Odc*-specific probe detected 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.
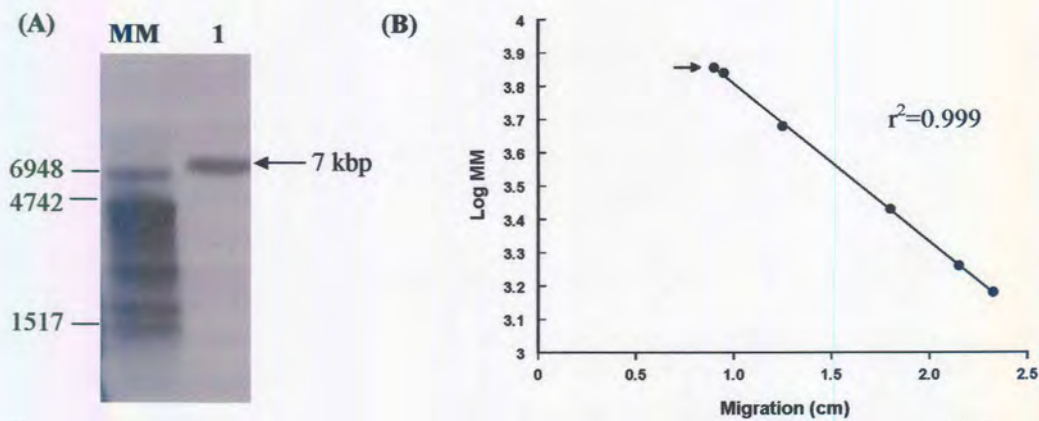
Figure 2.9: Northern blot analyses of the transcript of the bifunctional *PfAdometdc/Odc*. **(A)** Northern blot of *P. falciparum* RNA probed with a DIG-labelled *Odc*-specific probe indicating a single band (arrow). **MM**: DIG-labelled RNA molecular weight markers. **(B)** Linear regression analyses 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 MgCl₂, 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.
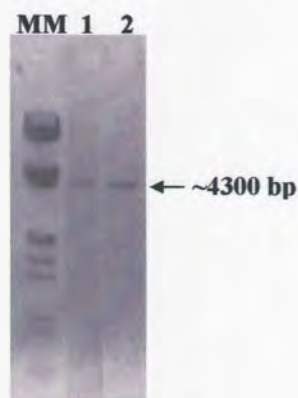
Figure 2.10: Amplification of the full-length bifunctional *PfAdometdc/Odc*. The ~4300 bp cDNA band amplified from 7.86 ng mRNA-cDNA equivalents of the uncloned cDNA library in lane 1 and in lane 2, from 50 ng *P. falciparum* genomic DNA. **MM:** *EcoRI-HindIII* digested λ-phage DNA used as high molecular 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 AATA<u>ATG</u> 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.
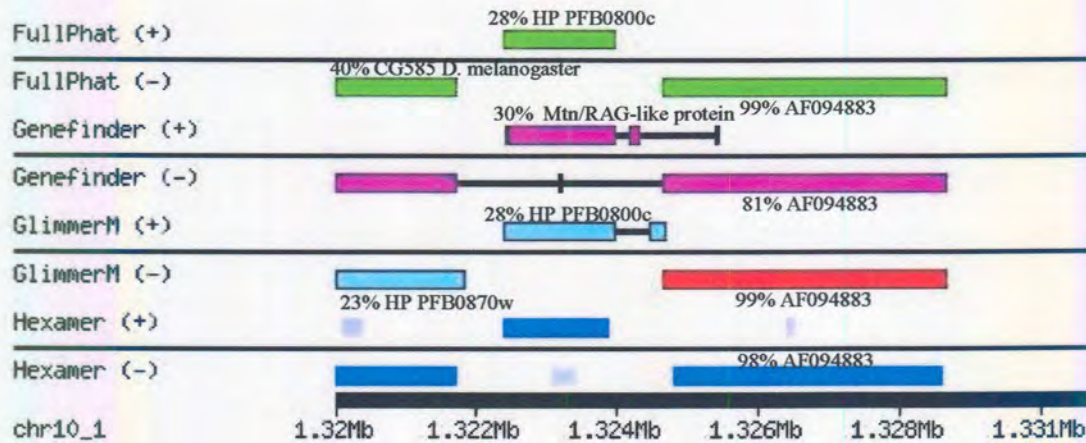
Figure 2.11: Analyses of chromosome 10 of *P. falciparum* containing the full-length ORF for the bifunctional *PfAdometdc/Odc* (red). ORF predictions were performed with different programs from PlasmoDB on chr10_1.glm_390. The location on the chromosome is indicated with the ruler at the bottom. Various 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
attttttttatttcttataaaaaagcaaatattttaaacaatattataacaaaagttatttttata
aaaaaatatatattacagtaaaaaaataaaaattttatttttatatcaaaaatatatatatatat
atatatataaaaaaaaaaaaaacgtgttcttggAttttttatataatataataaaaaaatatt
cataatacccccccaaaaaacccataaaaggaaaaaaaaacaaaaaattaaAtatatag
                                                                -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.
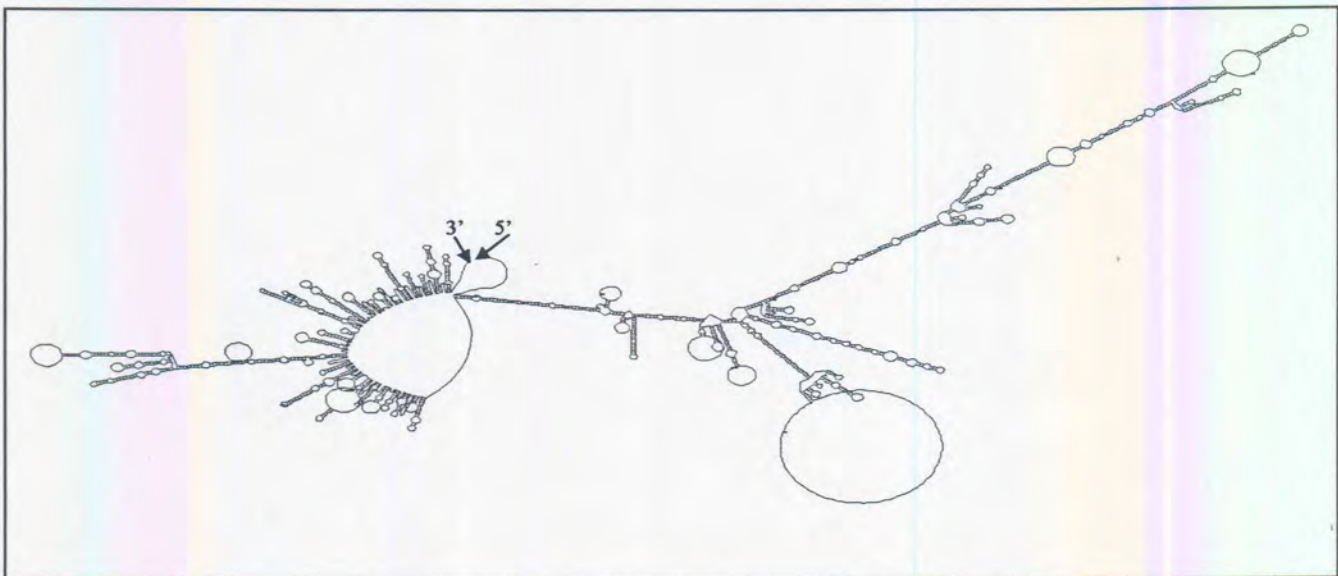
Figure 2.13: Secondary structure prediction of the ~2600 bp 5'-UTR of the bifunctional PfAdoMetDC/ODC. The 3-' and 5'-ends are indicated.

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).
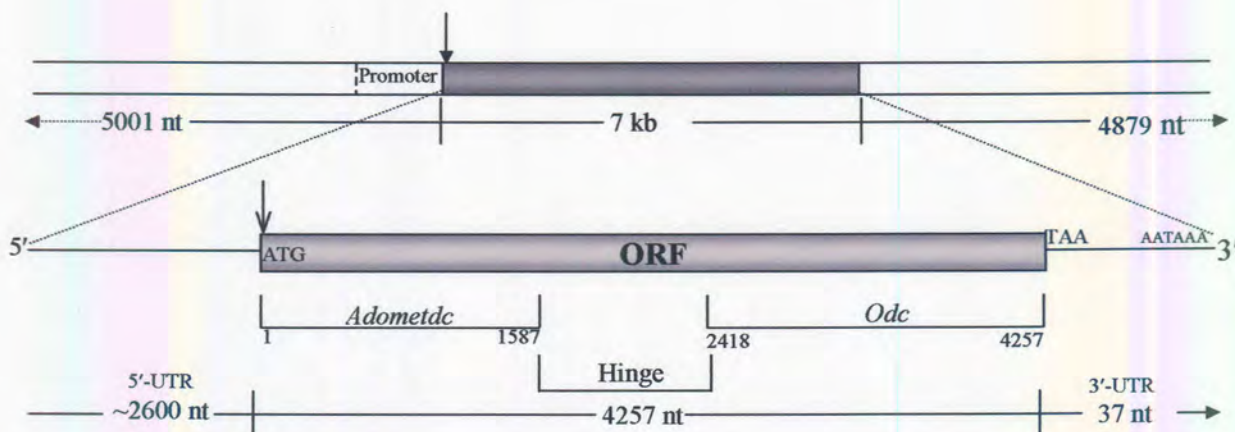
Figure 2.14: Schematic representation of the chromosomal organisation and general structures of the *PfAdometdc/Odc* gene and its corresponding mRNA. The top panel indicates the structure of the *PfAdometdc/Odc* gene without any introns and the 250 nt putative promoter region. The primary transcript 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.

CHAPTER 3

Recombinant expression and characterisation of monofunctional S-adenosylmethionine decarboxylase and ornithine decarboxylase as well as bifunctional PfAdoMetDC/ODC of *P. falciparum*.

3.1) INTRODUCTION.

3.1.1) Ornithine decarboxylase (E.C. 4.1.1.17).

The biological evidence suggesting a significant role for ODC in growth led to the characterisation of purified enzyme as one approach to understand this role. Mammalian ODC has a number of novel biochemical features including regulation by a wide variety of hormones, growth factors and other stimuli; an extremely short half-life in many species and regulation by a family of macromolecular inhibitors called antizymes (Cohen, 1998). Numerous early attempts to isolate ODC from various organisms were hampered by the low cellular content and instability of the protein (Pegg, 1989a). ODC is an extremely minor protein amounting to only 0.000002% of the total proteins in unstimulated rat tissue and depends on the presence of thiol reducing agents (5 mM dithiotreitol) and non-ionic detergents (0.02% Brij 35) for activity (Cohen, 1998; Heby, 1985). ODC has been isolated from various organisms ranging from mammals to protozoa, plants and prokaryotes and in all cases is dependent on pyridoxal-5'-phosphate (PLP) as co-enzyme (Cohen, 1998; Kaye, 1984; McCann and Pegg, 1992; Pegg, 1989a). The properties of purified ODCs are summarised in Table 3.1.

ODC has been shown to have a predominantly cytosolic localisation and undergoes limited post-translational modifications (Cohen, 1998; Heby, 1985). Multisite phosphorylation (with Ser_{303} of mammalian enzyme always involved) (Bachrach, 1984; Brown, *et al.*, 1994; Reddy, *et al.*, 1996) seems to increase the catalytic efficiency of the enzyme, whereas formation of cross-links by transamidation between two adjacent polypeptides and a diamine (like putrescine) (Bachrach, 1984; Russell, 1983) plus arginylation of the α-amino end of the enzyme (Cohen, 1998), might contribute towards the increased degradation of ODC. The only known allosteric regulator of ODC is GTP, which stimulates ODC by binding to an effector site 27Å away from the active site

(Oliveira, *et al.*, 1997). Mammalian ODC is also a lysine decarboxylase, although the Michaelis constant for L-Lysine is only 1/100th that of L-ornithine (Km: 9 mM vs. 0.09 mM for the murine enzyme) and it seems not to be a physiologically important reaction (Cohen, 1998; Pegg, 1989a).

Table 3.1 Properties of ODCs from various sources. Adapted from (Cohen, 1998; Kaye, 1984; McCann and Pegg, 1992; Pegg, 1989a). N/A: not applicable

Species	Specific activity (μmol/min/mg protein)	MM (protomer, Da)	Subunits/ molecule	Number of amino acids (deduced from cDNA)	pH optimum	Km (ornithine) (mM)	Km (PLP) (μM)
Human	12.8	51 000	2	461	--	--	--
Rat	0.01-19	51 000	2	461	7.0	~0.1	0.25
Mouse	22-70	51 000	2	461	--	0.03-0.075	0.3
Xenopus laevis	--	51 000	?	460	--	--	--
Saccharomyces cerevisiae	0.5	52 000	2	466	8	0.091	0.6
Neurospora crassa	44	53 000	2	484	--	--	--
Trypanosoma brucei	--	49 000	2	445	--	0.18	--
Leishmania	--	77 000	2	707	--	0.42-0.75	25
Physarum polycephalum	20	~50 000	--	N/A	--	--	--
Escherichia coli	130	82 000	2	732	6.9	36	--
Lactobacillus	180	85 000	2 or 12	--	--	--	--

Eukaryotic ODCs has a protomer Mr of ~ 50-54 000 depending on the species and has been proved to be an obligate homodimer of Mr ~ 100 000 (Cohen, 1998; Heby, 1985; Pegg, 1989a). Active and rapid dissociation and re-association occurs between the monomers with randomisation within five minutes (Coleman, *et al.*, 1994; Hayashi and Murakami, 1995). The dimeric form of ODC is the only active form of the enzyme, with two active sites formed at the interface between the two monomers (Osterman, *et al.*, 1994; Tobias and Kahana, 1993a). Site-directed mutagenesis studies have revealed residues essential for catalytic activity and these residues are contributed from both monomers (Tobias and Kahana, 1993a). On one monomer, Lys_{69} (mammalian enzyme numbering) forms a Schiff-base with the PLP co-factor and mutation of Lys_{169} and His_{197} leads to a major loss of activity (Cohen, 1998; Coleman, *et al.*, 1993; Lu, *et al.*, 1991; Osterman, *et al.*, 1999; Pegg, *et al.*, 1994). Cys_{360} from the second monomer completes the active site pocket (Cohen, 1998; Coleman, *et al.*, 1993; Pegg, *et al.*, 1994). Additional important residues include Gly_{387} (involved in dimerisation)(Cohen,

1998; Pegg, *et al.*, 1994), and Arg_{277} (interacting with PLP via a salt-bridge)(Osterman, *et al.*, 1997) and a Gly-rich loop (residues: LDI/VGGGF, conserved in PLP-binding enzymes and thought to form part of the co-factor binding site)(McCann and Pegg, 1992). ODC degradation is dependent on the presence of a PEST-rich region (region rich in Pro, Glu, Ser and Thr) proposed to target proteins for rapid degradation (Rogers, *et al.*, 1986). Two such areas are present, one in the C-terminal 423-449 residues and another weaker PEST-rich region in residues 298-333 (Hayashi, 1989; Hayashi and Murakami, 1995).

Multiwavelength stopped-flow spectroscopy results led to the proposal of the reaction mechanism of ornithine decarboxylation by *T. brucei* ODC (Fig. 3.1) (Brooks and Phillips, 1997; Osterman, *et al.*, 1999). The first step in catalysis is the linking of PLP to Lys_{69} via a Schiff-base and the formation of the external aldimine (second Schiff-base) with ornithine. Decarboxylation then occurs via a quinoid intermediate followed by protonation at the $C\alpha$ to form putrescine and transimination to generate free ODC and product (Brooks and Phillips, 1997; Osterman, *et al.*, 1999).
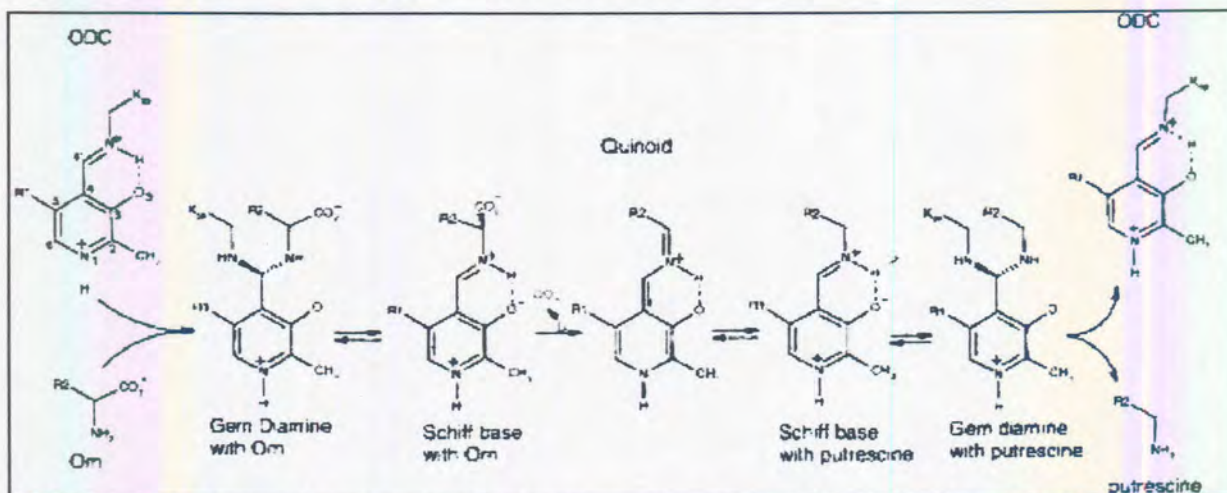
Figure 3.1: Proposed mechanism for the conversion of ornithine (Orn) to putrescine by *T. brucei* ODC. R1 is $CH_2OPO_3^{2-}$ and R2 is $(CH_2)_3NH_3^+$. Adapted from (Brooks and Phillips, 1997).

3.1.2) S-adenosylmethionine decarboxylase (E.C. 4.1.1.50).

AdoMetDC is highly regulated in mammalian cells to enable an appropriate level of decarboxylated AdoMet (dAdoMet) for the conversion of putrescine into polyamines. This rate-limiting enzyme for spermidine synthesis has been shown in all organisms not to be dependent on PLP for the decarboxylase action, but instead on a covalently bound pyruvate that is essential for activity (Tabor and Tabor, 1984a). This particularly interesting property specified a novel group of enzymes, including His decarboxylase,

Pro reductase, phosphatidylserine decarboxylase and Asp decarboxylase (Tabor and Tabor, 1984a). AdoMetDC is present in low concentrations (0.00025%) in the cytosol of animal cells (Cohen, 1998; Tabor and Tabor, 1984a). Human AdoMetDC has a Michaelis constant for AdoMet of 40 μM and 0.13 μM for putrescine and a pH optimum of 7.4-7.5 (Cohen, 1998).

AdoMetDC is synthesised as a proenzyme (Mr 38 300 for the mammalian enzyme) that undergoes an autocatalytic intramolecular cleavage at the Glu_{67}-Ser_{68} peptide bond to form the two subunits, α (Mr 30 700) and β (Mr 7 700) (Fig. 3.2 A) (Cohen, 1998; Pajunen, *et al.*, 1988; Stanley, *et al.*, 1994; Xiong, *et al.*, 1997). The cleavage generates a pyruvate at the amino terminus of the α-subunit from the Ser_{68} precursor residue through non-hydrolytic serinolysis (Ekstrom, *et al.*, 2001; Xiong, *et al.*, 1997). The side chain hydroxyl group of the Ser donates its oxygen to form the C-terminus of the β-chain through a nucleophilic attack of the Ser on the amide carbonyl group of the preceding residue. After β-elimination the β-chain is formed leaving a dehydroalanine at the amino terminus of the α-subunit. Conversion to ammonia and the pyruvoyl group via the formation of imine and cabinolamine intermediates results in the blocking of the N-terminus of the α-chain (Ekstrom, *et al.*, 1999; Ekstrom, *et al.*, 2001; Xiong, *et al.*, 1997). The result is a bipartite enzyme, a $\alpha_2\beta_2$ heterotetramer.

The decarboxylation reaction starts with Schiff-base formation between the pyruvoyl-cofactor and the α-amino group of the substrate, AdoMet (Fig. 3.2 B). The resulting electron sink facilitates removal of the α-carboxylate group from the substrate, which is replaced with a proton. Finally, the Schiff-base is hydrolysed to yield the decarboxylated product (Ekstrom, *et al.*, 1999).
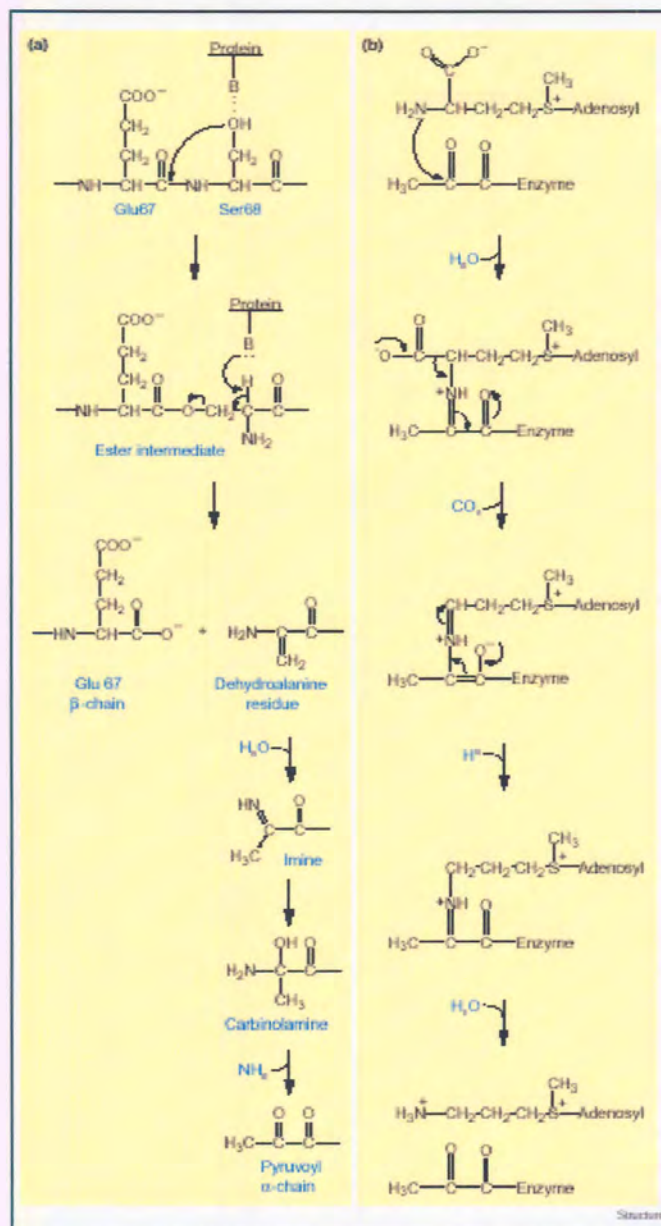
Figure 3.2: Proposed reaction mechanism for the autocatalytic intramolecular activation of AdoMetDC (A) and the decarboxylation of AdoMet (B). Adapted from (Ekstrom, *et al.*, 1999; Ekstrom, *et al.*, 2001).

The processing and activity of the mammalian, fungi, trypanosoma but not plant AdoMetDCs is enhanced by the presence of putrescine (Cohen, 1998; Pajunen, *et al.*, 1988; Stanley and Pegg, 1991; Stanley, *et al.*, 1994). AdoMetDCs fall into 2 classes, class I (microbial) with a $(\alpha\beta)_4$ structure and the cleavage sequence –DK~SHI or E~SH(L/I/V); and class II (mammalian, plants, protozoan and fungi) with a $(\alpha\beta)_2$ structure and the cleavage sequence –(L/V)L-(S/T/N)E~SS(L/M/E)(F/M)(F/I/V) (Ekstrom, *et al.*, 2001). Site-directed mutagenesis studies have indicated that putrescine needs to interact with at least four acidic residues (Glu_{11}, Glu_{178}, Asp_{174} and Glu_{256}) to accelerate post-translational processing of AdoMetDC (Xiong, *et al.*, 1997). Potato AdoMetDC processing is not affected by putrescine and this enzyme does not contain Asp_{174} (Xiong, *et al.*, 1997). Mutation of Glu_8 and Glu_{11} as well as Cys_{82} (in a

conserved area KT$\underline{C}$TG) results in inactive enzyme demonstrating an important role for the small subunit in the catalytic activity of mature AdoMetDC since the majority of these residues are present on this subunit (Cohen, 1998; Stanley and Pegg, 1991). The pyruvoyl-containing pocket includes residues Glu_8, Glu_{11}, Ser_{68}, Cys_{82}, Ser_{229} and His_{243}, which are all important for catalysis (Ekstrom, *et al.*, 1999). The putrescine-binding site is about 15-20Å removed from the active site but could exert its effects through conformational changes, electrostatic effects or both by transmitting the effects of putrescine binding to the active site residues via a chain of hydrogen bonds (Ekstrom, *et al.*, 2001).

3.1.3) AdoMetDC and ODC in *P. falciparum.*

Early studies on AdoMetDC and ODC of *P. falciparum* have shown that their activities peak in trophozoite and schizont phases of the parasite (~200 pmol CO_2 released by 10^9 parasitized erythrocytes after incubation at 37°C for 1hour) (Assaraf, *et al.*, 1984). A putrescine-dependent PfAdoMetDC (Km of 2.5 mM of putrescine) was not affected by Mg^{2+} and had an apparent Km value of 33 µM for AdoMet (Rathaur and Walter, 1987). The partial purification of PfODC on a pyridoxamine phosphate affinity column revealed two forms, a major protein band of 51 kDa and a minor 49 kDa band. The apparent Michaelis constant for ornithine was found to be 52 µM and cross-hybridisation studies with mouse ODC antibodies suggested some structural differences in comparison with the mammalian enzyme (Assaraf, *et al.*, 1988). Müller *et al.* (2000) provided conclusive data that both malarial decarboxylase activities reside on a single polypeptide. Fig. 3.3 indicates the schematic organisation of the two decarboxylase domains. The deduced amino acid sequence of the bifunctional PfAdoMetDC/ODC predicted a molecular mass of 166 kDa for the polypeptide, whereas the recombinantly expressed enzyme had a molecular mass of ~330 kDa. This suggested that the enzyme consists of a heterotetrameric structure derived from two heterodimeric proproteins (166 kDa each), post-translationally processed and cleaved within the AdoMetDC domain (157 and 9 kDa chains). The bifunctional protein revealed Michaelis constants of 58 µM and 41 µM for AdoMetDC and ODC activities, respectively (Müller, *et al.*, 2000). The two decarboxylase domains are active when expressed separately although the ODC domain loses almost a ninth of its activity (Krause, *et al.*, 2000; Wrenger, *et al.*, 2001). The Km values were 51 µM and 47 µM for the monofunctional AdoMetDC and ODC, respectively (Krause, *et al.*, 2000; Wrenger, *et al.*, 2001).

Heterodimeric polypeptide

AdoMetDC domain (1-529)   Hinge (530-804)   ODC domain (805-1419)

Heterotetrameric protein complex
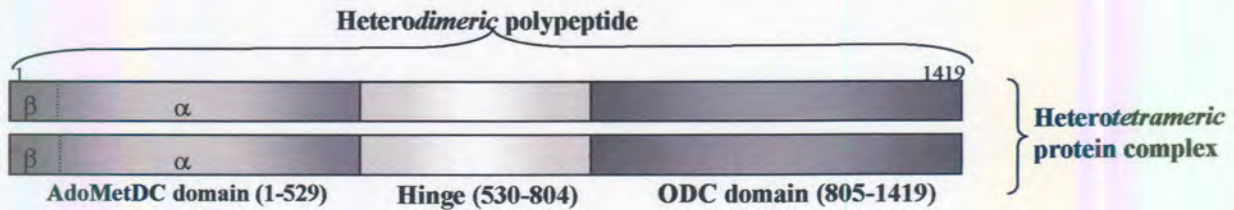
Figure 3.3: Schematic organisation of the bifunctional AdoMetDC/ODC from *P. falciparum.* AdoMetDC chains are indicated with α and β. Adapted from (Wrenger, *et al.,* 2001)

3.1.4) Recombinant protein expression and analyses.

Large amounts of pure functional proteins are required for structure-function studies of proteins. This is usually obtained by the recombinant overexpression of the proteins under investigation. Heterologous protein expression systems are based on the assumption that the basic principles of protein expression and function are similar in all organisms (Makrides, 1996). Usually, almost any protein can be recombinantly expressed in pro- or eukaryotic systems.

The choice of an expression system depends on many factors. These include cell growth characteristics, expression levels, intracellular and extracellular expression, posttranslational modifications, biological activity and regulatory issues of the protein of interest. Communication has to be established between the expression hosts, be it pro- or eukaryotic cells containing the necessary cellular machinery for protein translation, and the expression vector that contains the template from which the protein is translated. *Escherichia coli* has long been the primary prokaryotic host but the production of active eukaryotic proteins in *E. coli* is frequently hampered by the complex posttranslational modifications required, solubility problems as well as the lack of secretion and disulphide bond formation (Makrides, 1996; Weickert, *et al.*, 1996). The misfolding of heterologous proteins in *E. coli* has led to the inclusion of molecular chaperones or other folding catalysts to facilitate correct folding of proteins. In the event of inclusion body formation, the active protein might be recovered by a denaturation and refolding cycle (Thomas, *et al.*, 1997).

The isolation of the heterologously expressed protein from its counterpart in the expression host is simplified by fusing of a short peptide tag to the heterologous protein to mediate its recovery by affinity chromatography. A variety of tags are available including His-tags (with Ni^{2+}-nitrilotriacetic acid as ligand), Glutathione-S-transferase

(isolated on glutathione-Sepharose), Streptococcal protein G (albumin as ligand), Thioredoxin (thiobond resin), Avidin and streptavidin (biotin as ligand) and poly-Arg or poly-Cys (S-Sepharose or thiopropyl-Sepharose) (Makrides, 1996). Some protein tags like thioredoxin might even increase the solubility of target proteins. Most tags can also be removed posttranslationally and often does not influence protein folding or activity (Makrides, 1996).

The expression of *P. falciparum* genes in heterologous systems has often been a challenge due to the codon bias of this organism (Gardner, *et al.*, 1998; Hyde and Holloway, 1993; Pizzi and Frontali, 2001). *P. falciparum* proteins have been expressed in the yeast *Saccharomyces cerevisiae*, which also has a higher A+T-content in its genome (Bathurst, 1994). However, yeast systems produce suboptimal levels of protein and may result in early translational termination of polypeptides (Chang, 1994). Mammalian cells have been used to obtain appropriate posttranslational modifications but the yield of protein expression is mostly <1% of the dry weight (Chang, 1994; Pollack, *et al.*, 1985). The baculovirus expression system offers a higher yield and has been used for the production of *P. falciparum* antigens (Chang, 1994; Matsuoka, *et al.*, 1996). A related protozoan, *Dictyostelium discoidium* has also been exploited as a possibility for *Plasmodium* protein expression (Fasel, *et al.*, 1992; Kay and Williams, 1999). The majority of *P. falciparum* proteins have been functionally expressed in *E. coli*, but certain codons that are preferentially used by *P. falciparum* are rarely used in *E. coli* and this diminishes the expression levels of the recombinant proteins (Baca and Hol, 2000; Sayers, *et al.*, 1995). One method to overcome the codon bias is to re-engineer the *Plasmodium* gene by synthesising the gene so that it uses the preferred codons of *E. coli* (Pan, *et al.*, 1999; Prapunwattana, *et al.*, 1996). Alternatively, the co-transformation of *E. coli* with a plasmid carrying the gene for the rare codons' cognate tRNA can increase levels of heterologous protein expression. Plasmids encoding tRNAs for Arg, Ile and Gly or Arg, Ile and Leu has now been constructed for use in expression of *Plasmodium* and other parasite genes in *E. coli* (Baca and Hol, 2000).

In this chapter, the expression of active monofunctional PfAdoMetDC and PfODC as separate proteins as well as the bifunctional PfAdoMetDC/ODC in various expression systems is described. The amino acid sequence of the bifunctional PfAdoMetDC/ODC was also analysed for molecular characterisation of secondary structure elements, species-specific relationships and structural and functional motifs. The work presented

here is the result of studies performed in South Africa and during research visits to Germany.

Results of this Chapter have been presented as a paper at the BioY2K Combined Millennium Meeting (Birkholtz, 2000c) and as two posters at international and national conferences (Birkholtz, *et al.*, 2000a; Birkholtz and Louw, 2000d).

3.2) MATERIALS AND METHODS.

3.2.1) Recombinant expression of His-Tag fusion proteins.

His-Tag fusion expression systems were constructed in South Africa for the *P. falciparum* AdoMetDC and ODC monofunctional domains. Expression from the pET vector system (Novagen, Madison, USA) is obtained with T7 RNA polymerase activity in the bacteriophage DE3 lysogen cells under tight control of the *lacUV5* promoter. Addition of isopropyl-β-D-thiogalactopyranoside (IPTG) induces T7 RNA polymerase, which in turn allows the translation of the protein. Repeats of His (6x) are fused to the proteins to facilitate their isolation by immobilised-metal affinity chromatography (IMAC) with Ni-nitrolotriacetic acid (Ni-NTA) columns. His_6-Tag facilitates binding to Ni-NTA at pH 8 in its uncharged form and does not generally affect protein expression and folding. His-Tagged proteins can then be eluted under reduced pH conditions (pH 4.5-5.3) and high concentrations of competing imidazole.

3.2.1.1) Cloning of AdoMetDC and ODC open-reading frames into the pET expression vector.

The *P. falciparum* AdoMetDC and ODC domains were cloned into the pET-15b expression vector for the expression of N-terminal His-tag proteins under control of the T7*lac* promoter (*lacUV5*). Primers used in the amplification of the open-reading frames are indicated in Table 3.1 and included restriction enzyme sites to allow directional, in-frame cloning of the fragments into the expression vector. The primers used for expression of the AdoMetDC domain was based on the published sequence of the bifunctional PfAdoMetDC/ODC (Müller, *et al.*, 2000).

A short 2050 bp ODC ORF (excluding the hinge region) was amplified with primers ODCexpf2 and ODCexpr (Table 3.1) in a 25 µl PCR containing 1/10 000 dilution of a

plasmid containing the full-length PfODC of 2918 bp as template (250 ng/µl, Chapter 2), 5 pmol of each primer, 1x Takara ExTaq reaction buffer (Takara Shuzo, Japan), 0.2 mM of each dNTP, 2 mM MgCl$_2$ and 1.25 U Takara ExTaq DNA Polymerase (Takara Shuzo, Japan) in 0.2 ml thin-walled tubes in a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, California, USA). Cycling parameters consisted of denaturing at 94°C for 3 min, a 2 min pause at 80°C for the hot-start addition of polymerase followed by 30 cycles at 94°C for 30 sec, 57°C for 30 sec and 60°C for 3 min. The PCR products were analysed on a agarose gel, purified after electrophoresis, cloned into pGem-T Easy vector and the nucleotide sequence determined as described in Chapter 2, section 2.2.7-11.

The ODC cloned into pGem-T Easy vector as well as pET-15b was isolated with the High Pure Plasmid isolation Kit (Chapter 2, section 2.2.9.3) and 10 µg of each plasmid was subjected to *Cell*II (Roche Diagnostics, Mannheim, Germany) restriction digestion for 2 hours at 37°C in a reaction containing 10 U enzyme and 1x restriction enzyme buffer. Subsequently, the NaCl concentration was adjusted to 2.5 M for a further 2 hour digestion of the plasmids at 37°C with 15 U *Bam*HI (Roche) in 2 µg bovine serum albumin (BSA). The digested plasmids were again analysed by electrophoresis and purified as above. 100 ng sticky-ended ODC insert was then ligated to 100 ng digested pET-15b vector (3:1 molar ratio of insert:vector) in a reaction containing 1 U T4 DNA ligase (Promega, Wisconsin, USA) in 1x T4 DNA ligase buffer, for 16 hours at 16°C. Reactions were stopped by heat inactivation at 70°C for 15 min. Plasmids were transformed into competent DH5α and positive clones were confirmed by restriction enzyme analyses and gel electrophoresis (described in Chapter 2, section 2.2.10). Positive clones were named pET-ODC.

The AdoMetDC domain was amplified with the Sampetf1 and Sampetr1 primer pair (Table 3.1) based on the published *PfAdometdc/Odc* sequence (Genbank accession number AF0934833). The 50 µl reactions contained 1×reaction buffer, 2 mM MgCl$_2$, 0.2 mM dGTP,dCTP/0.3 mM dATP,dTTP, 10 pmol of each primer and 2.5 U Takara ExTaq DNA polymerase (Takara Shuzo, Japan) added in a hot-start protocol using 25 ng total RNA:cDNA equivalents of single-stranded cDNA as template. PCR was performed for 30 cycles at 94°C for 30 sec, and 65°C for 2.5 min. The PCR product was purified from the agarose gel and cloned into pGEM-T Easy as described in sections 2.2.7-2.2.11. The AdoMetDC insert was cut from the cloning vector with *Bam*HI and

*Cell*II as described above for the ODC domain. Sticky-end cloning was performed as for the ODC domain into pET-15b. Positive clones were named pET-ADC.

3.2.1.2) Culturing of expression cell line and induction of protein expression.

pET-ODC and pET-ADC were transformed (10 µg each) into competent BL21(DE3) *E. coli* as described in Chapter 2, section 2.2.10. Co-transformation of 10 µg of a plasmid containing the tRNAs for Arg, Ile and Gly (pRIG) was performed simultaneously to increase the levels of protein obtained (Baca and Hol, 2000). Single colonies were selected and grown overnight with shaking to saturation in LB-medium containing 50 µg/ml ampicillin and 25 µg/ml chloramphenicol (for selection of pRIG) at 37°C. 1 ml of this culture was inoculated into 1 litre LB-medium and incubated at 37°C with shaking until an OD_{600} of 0.6 was reached. Protein induction was achieved by addition of 1 mM IPTG and the cultures were grown for a further 4 hours at 30°C while shaking. Cells were subsequently harvested by centrifugation for 5 min at 3000×g in a Hermle Z252M. The pellets were dissolved in 1x TE buffer (10 mM Tris, 1 mM EDTA, pH 8); 100 µg/ml lysozyme, 1/10 the volume of 1% Triton, 50 µg/ml DNAseI, 1 µg/ml leupeptin (Roche Diagnostics), 20 µg/ml aprotinin and 0.5 mM PMSF (phenylmethylsulfonyl fluoride, Roche Diagnostics) was added and the suspension incubated for 15 min at 37°C. Leupeptin is a tripeptide protease inhibitor of serine proteases, aprotinin is a 6.5 kDa basic polypeptide inhibiting kallikrein, trypsin, chymotrypsin, plasmin and papain whereas PMSF inhibits trypsin and chymotrypsin. The cells were sonicated on ice for three cycles of 15 second intervals at an output of 40-60 W with a Branson Sonifier 250 (Branson Scientific, Danbury) each cycle followed by 15 second incubations on ice. Cell debris was collected by ultracentrifugation at 40 000×g for 20 min at 4°C in a Beckman J7-55 ultracentrifuge with a type 30 rotor (Beckman Instruments, California, USA).

3.2.1.3) Isolation of the His-Tag fusion proteins by affinity chromatography.

A column of one ml bed volume of His-bind resin (Novagen, Madison, USA) was prepared and washed with three volumes distilled water. The column was charged with five column volumes 50 mM $NiSO_4$ and equilibrated with three column volumes binding buffer (5 mM imidazole, 0.5 mM NaCl, 20 mM Tris-HCl, pH 7.9) at a flow rate of 1 ml/min. The 5 ml protein-containing supernatant from the ultracentrifugation step was loaded onto the column and non-specific proteins were removed by washing with six column volumes of wash buffer (60 mM imidazole, 0.5 mM NaCl, 20 mM Tris-HCl,

pH 7.9). The His-tagged protein was eluted in six column volumes elution buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9) and stored in 0.02% Brij-35 (Fluka, Germany) at 4°C. The column was stripped with five column volumes of 100 mM EDTA, 0.5 M NaCl and 20 mM Tris-HCl, pH 7.9 and stored at 4°C.

3.2.2) Recombinant expression of Strep-Tag fusion proteins.

The *P. falciparum* monofunctional AdoMetDC and ODC domains as well as the bifunctional PfAdoMetDC/ODC protein were expressed as Strep-Tag fusion proteins as described in (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). The engineered Strep-Tag (NH$_2$-WSHPQFEK-COOH) with highly selective binding properties for derivatised streptavidin coupled to Sepharose called Strep-Tactin can be fused to the N- or C-terminal ends of proteins (Institut für Bioanalytik, Göttingen, Germany). Proteins expression from any pASK-IBA vector occurs under tight control of the tetA promotor/operator. The tet repressor is encoded on the expression vector and is constitutively expressed under the β-lactamase promotor. Protein expression is induced with low levels of anhydrotetracycline. The tetA promotor is not leaky and not functionally coupled to any cellular regulation mechanisms or genetic background. Recombinantly expressed Strep-Tag fusion proteins are then isolated with affinity chromatography on a Strep-Tactin derivatised Sepharose column under physiological conditions using any physiological buffer. Protein elution occurs in the presence of desthiobiotin (Sigma), a competitive biotin analogue with a high affinity to Strep-Tactin. Reversibly bound desthiobiotin is removed with 4-hydroxy azobenzene-2-carboxylic acid (HABA, Sigma) to regenerate the column. The Strep-Tag does not need to be removed from the proteins since there is no interference with folding or activity as it has no ion-exchange properties and does not induce protein aggregation.

3.2.3) Size-exclusion high-pressure liquid chromatography (SE-HPLC) of the monofunctional ODC.

Monofunctional ODC expressed as a Strep-tag fusion protein was further purified with SE-HPLC. The proteins from the Strep-Tactin affinity column in buffer W (100 mM Tris-HCl, pH 8.0, 1 mM EDTA) was filtered through a 0.22 μm filter (Millex GV4, Millipore Corporation, USA) and applied to a size-exclusion column (G2000W$_{XL}$, 7.8 mm × 30 cm, 100 kDa exclusion size, TosoHaas, USA). Buffers were filtered through a 0,22 μm filter and degassed for 30 min prior to application to the column. The column was equilibrated using isocratic conditions in buffer W at a flow speed of 1 ml/min and

calibrated with low molecular weight protein markers: phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa) and α-lactoalbumin (14.5 kDa). The affinity-purified proteins were applied to the column and eluted in 1 ml fractions. Fractions were analysed for protein using SDS-PAGE (section 3.2.6). Beckman HPLC instrumentation was used consisting of two pumps (module 110B), an analog interface for control of solvent delivery by pumps (module 406), a UV detector (module 166, 280 nm) and a Beckman 340 organiser and injector with Beckman System Gold Software (Beckman Instruments, California, USA).

3.2.4) Size-exclusion fast protein liquid chromatography (SE-FPLC) of monofunctional AdoMetDC or bifunctional PfAdoMetDC/ODC.

Affinity chromatography purified monofunctional AdoMetDC or bifunctional PfAdoMetDC/ODC expressed as Strep-Tag fusion proteins were subjected to FPLC on a HiLoad 26/60 Superdex S-200 size-exclusion column (2.6 x 60 cm, exclusion limit of Mr $1.3 \cdot 10^6$ for globular proteins, separation range of 10-600 kDa, Amersham Pharmacia Biotech, Buckinghamshire, UK) according to (Müller, *et al.*, 2001; Wrenger, *et al.*, 2001). The column was calibrated with dextrane blue (2000 kDa), bovine serum albumin (~145 kDa dimeric form), alcohol dehydrogenase (150 kDa), carbonic anhydrase (29 kDa) and cytochrome C (12.4 kDa).

3.2.5) Quantitation of proteins (Bradford, 1976).

The concentrations of the purified proteins were determined according to Bradford by binding to Coomassie brilliant blue G-250 using the Pierce Coomassie Protein Assay Kit (Pierce, Illinois, USA) and monitoring the shift in absorbance of the dye from A_{465} to A_{595} due to binding of proteins via electrostatic attraction of the dyes' sulphonic groups to basic and aromatic groups in the proteins (maximally to Arg and weakly to His, Lys, Tyr, Trp and Phe). Standard protein or sample (150 µl) were pipetted into a microtitre plate well and 150 µl of the assay reagent added, shaken for 10 min and absorbance was read at 620 nm with a SLT 340 ATC scanner (SLT Labinstruments).

3.2.6) SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of proteins.

Proteins were analysed with denaturing SDS-PAGE with a 5% stacking gel and either 10% (for separation of larger proteins >100 kDa) or 12.5% separating gel (for proteins between 10-100 kDa). The stacking gel (0.625 M Tris-HCl, pH 6.8, 0.5% SDS) and separating gel (1.88 M Tris-HCl, pH 8.8, 0.5% SDS) were prepared from acrylamide

(30% acrylamide, 0.8% N',N'-methylene bisacrylamide) and electrophoresis buffer (0.02 M Tris-HCl pH 8.3, 0.06% w/v SDS, 0.1 M glycine) stock solutions. The gel solutions were degassed for 5 minutes and polymerised by addition of 30 µl of 10% ammonium persulphate and 5 µl TEMED (*N,N,N'N'*-tetramethylethylenediamine, Merck, Germany).

Protein samples were diluted 1:1 in reducing sample buffer (0.06 M Tris-HCl pH 6.8, 2% w/v SDS, 0.1% v/v glycerol, 0.05% v/v β-mercaptoethanol and 0.025% w/v bromophenol blue) and boiled at 95°C for 5 minutes. Low molecular mass markers used were phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa) and α-lactoalbumin (14.5 kDa). Benchmark 10 kDa protein ladder was used for larger proteins with a range from 10 – 200 kDa (Gibco BRL, Gaithersburg, USA). Electrophoresis was carried out in electrophoresis buffer using a Biometra electrophoresis system (Biometra GmbH, Germany) with an initial voltage of 60 V for 45 min and thereafter a voltage of 100 V until the bromophenol blue marker reached the bottom of the gel.

3.2.6.1) Staining of SDS-PAGE gels.

Low concentration (~ 50 µg/ml) of recombinantly expressed monofunctional AdoMetDC or ODC and bifunctional PfAdoMetDC/ODC protein bands were visualised with a non-ammonical, neutral silver staining procedure after SDS-PAGE, (Merril, *et al.*, 1981). The gels were fixed in 30% v/v ethanol, 10% v/v acetic acid for 30 min and then incubated in 30% v/v ethanol, 0.5 M Na-acetate, 0.02% v/v glutaraldehyde and 0.2% w/v $Na_2S_2O_3$ for 30 min. This was followed with three washes in distilled water after which a 0.1% w/v $AgNO_3$, 0.02% v/v formaldehyde solution was added and the gel soaked for 30 min. The silvernitrate solution was removed and the gel was developed in 2.5% w/v Na_2CO_3, 0.01% v/v formaldehyde until bands were visible. The reaction was stopped by the addition of 0.05 M EDTA and washing in distilled water.

3.2.7) AdoMetDC and ODC enzyme activity assays.

The *P. falciparum* monofunctional AdoMetDC and ODC domains as well as the bifunctional PfAdoMetDC/ODC were assayed for AdoMetDC as well as ODC activities as described in (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001).

3.2.8) *In silico* analyses of the predicted amino acid sequence of PfAdoMetDC/ODC.

The amino acid sequence of the bifunctional PfAdoMetDC/ODC was deduced with the Translate tool (www.expasy.ch). Amino acid sequence alignments were performed with Clustal W (Thompson, *et al.*, 1994) using the default parameters for PfAdoMetDC/ODC (Genbank Accession Number AF094833) and the corresponding enzymes from the human, mouse, *L. donovani* and *T. brucei*. Genbank accession numbers for AdoMetDCs: human: M21154, murine: D12780, *T. brucei:* U20092, *L. donovani:* LDU20091 and for ODCs: human: M31061, murine: J03733, *T. brucei:* J02771 and *L. donovani:* M81192. The amino acid composition of the bifunctional PfAdoMetDC/ODC was analysed with PHD (Rost, 1996). Possible post-translational modifications were predicted with PROSITE (Bairoch, *et al.*, 1995), and secondary-structure predictions were composed from PHD using the GOR IV algorithm. These modules are all part of the PredictProtein Server (www.dodo.cpmc.columbia.edu/pp/submit) (Rost, 1996). PEST-rich regions were predicted with the PEST Prediction Tool (www.at.embnet.org/embnet/tools/bio/PESTfind) (Rogers, *et al.*, 1986). The SMART server (Simple modular architecture research tool, smart.embl-heidelberg.de)(Letunic, *et al.*, 2002) was used to detect domains and protein sequence annotations. Analyses of low-complexity regions were performed with the SEG programme (Wootton and Federhen, 1996).

The BLOCKS database (http://www.blocks.fhcrc.org/) (Henikoff and Henikoff, 1994) was used to obtain a representation of the protein family of the aligned segments of the ODC proteins. The Pfam database (http://pfam.wustl.edu/) grouped the PfODC into existing protein families of multidomain proteins based on multiple alignments of protein domains and conserved protein regions (Bateman, *et al.*, 2002). This was also performed with the InterProScan server, which scans an amino acid sequence against BLOCKS, Pfam and PRINTS (www.ebi.ac.uk/interpro/scan.html).

Bifunctional AdoMetDC/ODC was identified in other *Plasmodium* species through analyses of the PlasmoDB database by disabling low-complexity filtering and repeat

masking. Open-reading frames were identified for *P. berghei* (berg_296a09.q1c) and *P. yoelii* (chrPyl_cpy1465) with `getorf` and `plotorf` from the Emboss package.

3.3) RESULTS.

3.3.1) Directional cloning strategy of individual ODC and AdoMetDC domains.

The monofunctional domains were amplified for expression as N-terminal His-tagged fusion proteins with primers (Table 3.2) to facilitate directional sticky-end cloning. The primer sequences were based on the nucleotide sequences of the cloned ODC and AdoMetDC domains described in Chapter 2 (Müller, *et al.*, 2000).

Table 3.2: Primers used in the cloning of the ODC and AdoMetDC domains for expression of the proteins in the pET-15b His-tag expression system.

Primer	Sequence (5' – 3')	Restriction enzyme sites	Tm (°C)[a]
ODCexpf2	GGA TCC TAT GGA AAA GAA TTA TAA AG	*Bam*HI	56
ODCexpr	GCT CAG CTT TTG TTA TTT ACC AAT G	*Cell*II	58
Sampetf1	CTG CGG GAT CCA ATG AAC GGA ATT TTT GAA G	*Bam*HI	66
Sampetr1	CTC CGG CTC AGC ATA ATC ATC ATG TAC GTC	*Cell*II	68

- [a] T_m equation=69.3+0.41(%GC)-650/length (Rychlik, *et al.*, 1990).
- Restriction sites for cloning purposes are given in bold.

3.3.2) Expression strategy of monofunctional AdoMetDC and ODC as well as bifunctional PfAdoMetDC/ODC.

Fig. 3.4 indicates the different constructs for the various expression systems used. The 1419 residue bifunctional PfAdoMetDC/ODC was expressed in pASK-IBA3 vector with a C-terminal Strep-tag, monofunctional AdoMetDC (660 residues) and ODC (758 residues) in pASK-IBA7 with N-terminal Strep-tags as described in (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). His-tag based pET-15b vectors were used for expression of the 596 residue AdoMetDC and 614 residue ODC monofunctional proteins with N-terminal His-tags.
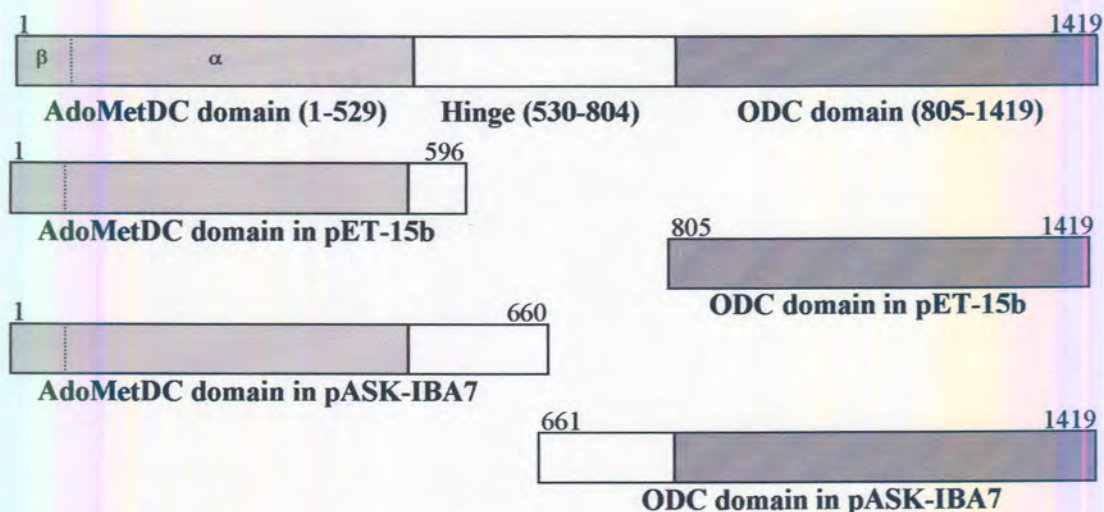
Figure 3.4: Schematic representation of the cloning strategy for expression of monofunctional AdoMetDC and ODC or bifunctional PfAdoMetDC/ODC. The top panel indicates the bifunctional PfAdoMetDC/ODC with the sizes of the domains given in number of amino acids. Monofunctional proteins were expressed as His-tag fusion proteins (pET-15b), or as Strep-tagged proteins (pASK-IBA7) with the residue sizes given. The α- and β-chains of AdoMetDC are separated with a dashed line. Parts of the hinge region (empty box) were included in some of the constructs.

3.3.3) Recombinant expression of monofunctional AdoMetDC and ODC domains.

3.3.3.1) His-tag expression system.

The monofunctional AdoMetDC (with partial hinge region) and ODC (without hinge region) were individually expressed as His-tag fusion proteins. Fig. 3.5 (A) indicates a typical elution profile after Ni^{2+}-affinity chromatography. Unbound proteins were washed from the system followed by a small amount of non-specifically bound protein that was removed with 60 mM imidazole. The recombinantly expressed AdoMetDC or ODC was eluted in a single sharp peak with 1 M imidazole in a volume of ~3 ml. The low levels of either monofunctional protein obtained after recombinant expression in *E. coli* necessitated inoculations of large-scale cell cultures. One litre of induced cells only yielded maximally ~ 150 µg protein after affinity chromatography.
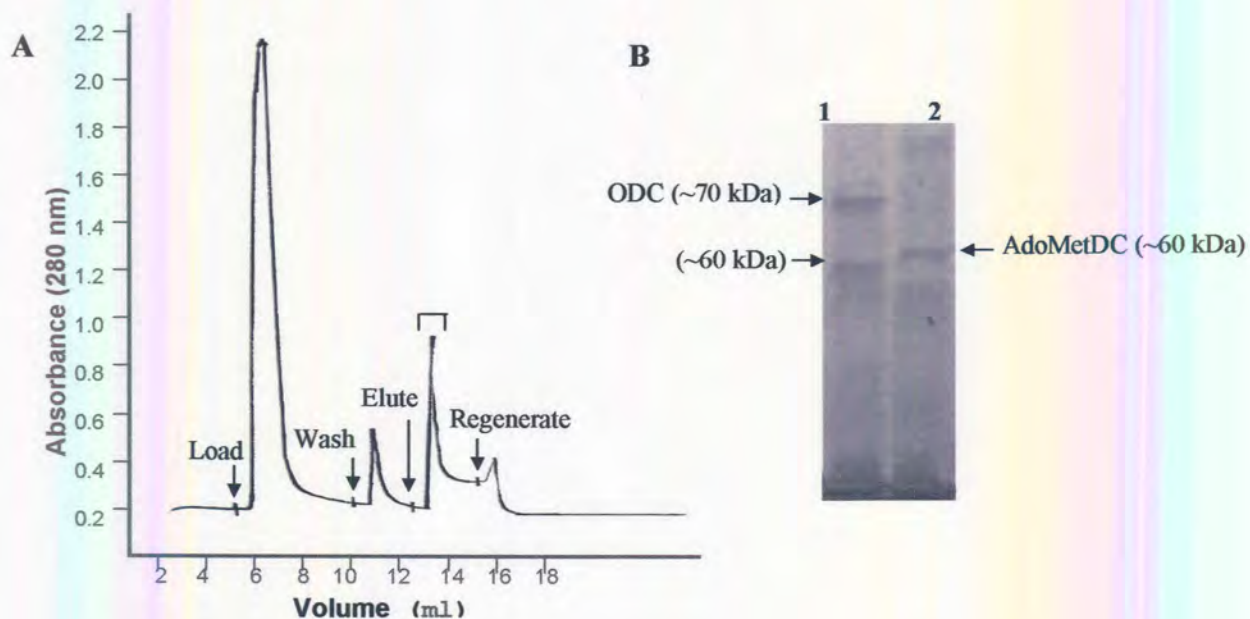
Figure 3.5: His-tag fusion protein expression of monofunctional AdoMetDC or ODC. (A) Typical elution profile of a His-tagged based purification of recombinantly expressed AdoMetDC or ODC. The elution of the recombinant proteins is indicated with the horizontal bracket. **(B)** SDS-PAGE of recombinantly expressed monofunctional AdoMetDC and ODC. Lane 1: ODC protein at ~70 kDa with a contaminating band at ~60 kDa and lane 2: AdoMetDC protein at ~60 kDa. Recombinant protein was visualised with silver staining.

The affinity chromatography purified proteins were analysed with SDS-PAGE. Expression of monofunctional AdoMetDC should result in a 68.5 kDa protein (56.6 kDa α-chain and 9 kDa β-chain with a 3 kDa His-tag). Monofunctional ODC should have a molecular weight of 70.5 kDa (67.5 kDa ODC and 3 kDa His-tag). Fig. 3.6 (B) indicates the purity of the isolated monofunctional proteins. AdoMetDC migrated at an expected ~ 60 kDa protein with the smaller 9 kDa fraction expected to be present in the migration front (not visible). However, ODC expression revealed two bands, a major band at ~70 kDa and a minor ~60 kDa band. This smaller band could be a contaminating protein or degradation product of ODC, although protease inhibitors were included during the purification procedure.

3.3.3.2) Strep-tag expression system.

The monofunctional AdoMetDC and ODC were individually expressed with N-terminal Strep-Tags. Both proteins included the full domain and part of the hinge region, as previously defined and described (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et*

al., 2001). Proteins were purified with Strep-Tactin affinity chromatography and eluted in ~ 3 ml in a single peak as indicated in Fig. 3.6. The maximal yield was ~ 250 µg per litre culture for ODC and ~ 150 µg per litre culture for AdoMetDC.
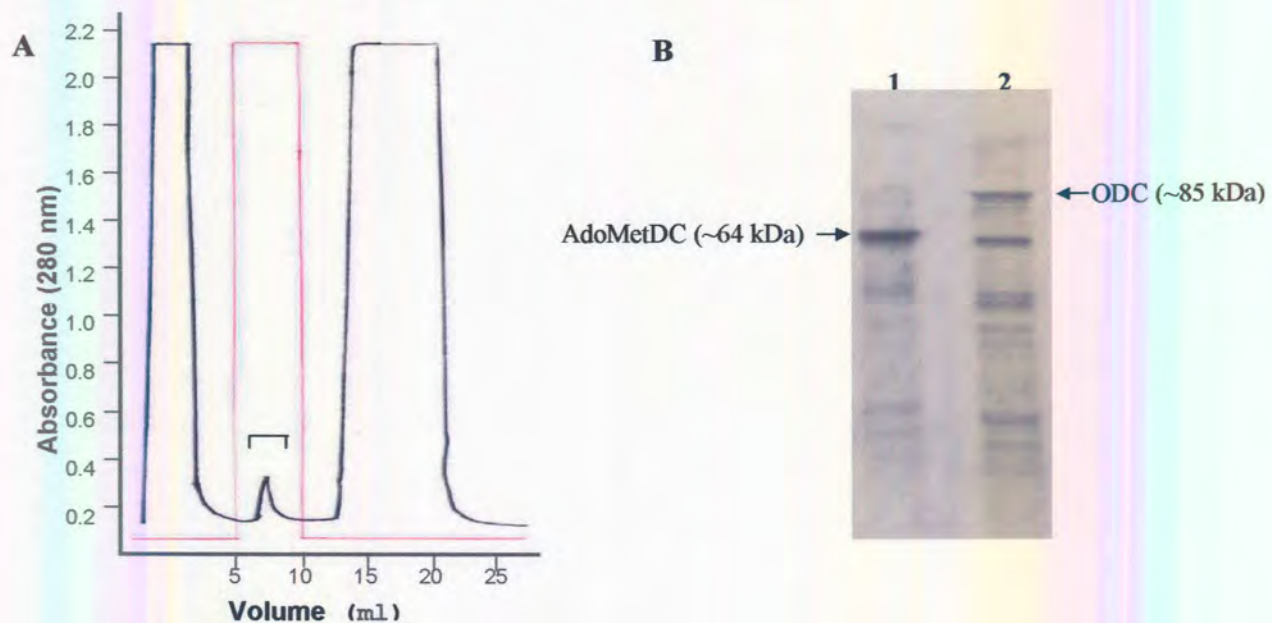
Figure 3.6: Expression of monofunctional AdoMetDC and ODC as Strep-tag proteins. (A) Typical elution profile of the Strep-tag proteins with Strep-Tactin affinity chromatography. The peak containing the eluted protein is indicated with a horizontal bracket. The red line indicates 100% elution buffer. **(B)** SDS-PAGE of the monofunctional proteins. Lane 1: AdoMetDC protein (64 kDa) and lane 2: ODC protein (~85 kDa). Proteins were visualised with silver staining.

Expression of the AdoMetDC domain (529 residues) and 131 residues of the hinge region should result in a 72.6 kDa protein (63.6 kDa and 9 kDa for the two chains) with the 1.5 kDa Strep-tag. ODC expression should yield a ~85 kDa protein including the Strep-tag and 144 residues of the hinge region. SDS-PAGE of the purified proteins indicated protein preparations with the expected sizes of the proteins observed. Fig. 3.6 (B) indicates the single ~64 kDa band for AdoMetDC (smaller 9 kDa band expected in migration front). A smaller band (60 kDa) compared with the expected band migrating at ~85 kDa was also observed during the ODC purification (Fig. 3.6, B). Both the ODC and AdoMetDC protein preparations with the Strep-Tag expression system had a higher background of non-specifically isolated proteins compared to the isolations of ODC and AdoMetDC with the His-Tag system (Fig. 3.5 B).

3.3.4) Determination of the oligomeric state of the monofunctional AdoMetDC and ODC.

The monofunctional AdoMetDC and ODC proteins isolated with affinity chromatography as Strep-Tag fusion proteins were further purified with size-exclusion chromatography to determine the oligomeric state of these proteins.

Monofunctional ODC was purified with a calibrated size-exclusion HPLC (r^2=0.967 after linear regression, Fig. 3.7 A). ODC eluted in a single peak corresponding in size to the homodimeric form of the protein of ~170 kDa (Fig. 3.7 B). Activity analyses indicated the presence of active ODC in this peak. A large peak at ~10 min (~30 kDa) was present in the ODC separations, possibly indicating the presence of the contaminating proteins from the affinity chromatography step. SDS-PAGE analyses of the ODC-containing single peak showed the presence of the ~85 kDa band and the smaller ~60 kDa band (Fig. 3.7 C). The smaller contaminating bands present after the affinity chromatography was removed by the SE-HPLC. No monomeric ODC was visible in any of the fractions corresponding in size to 85 kDa.
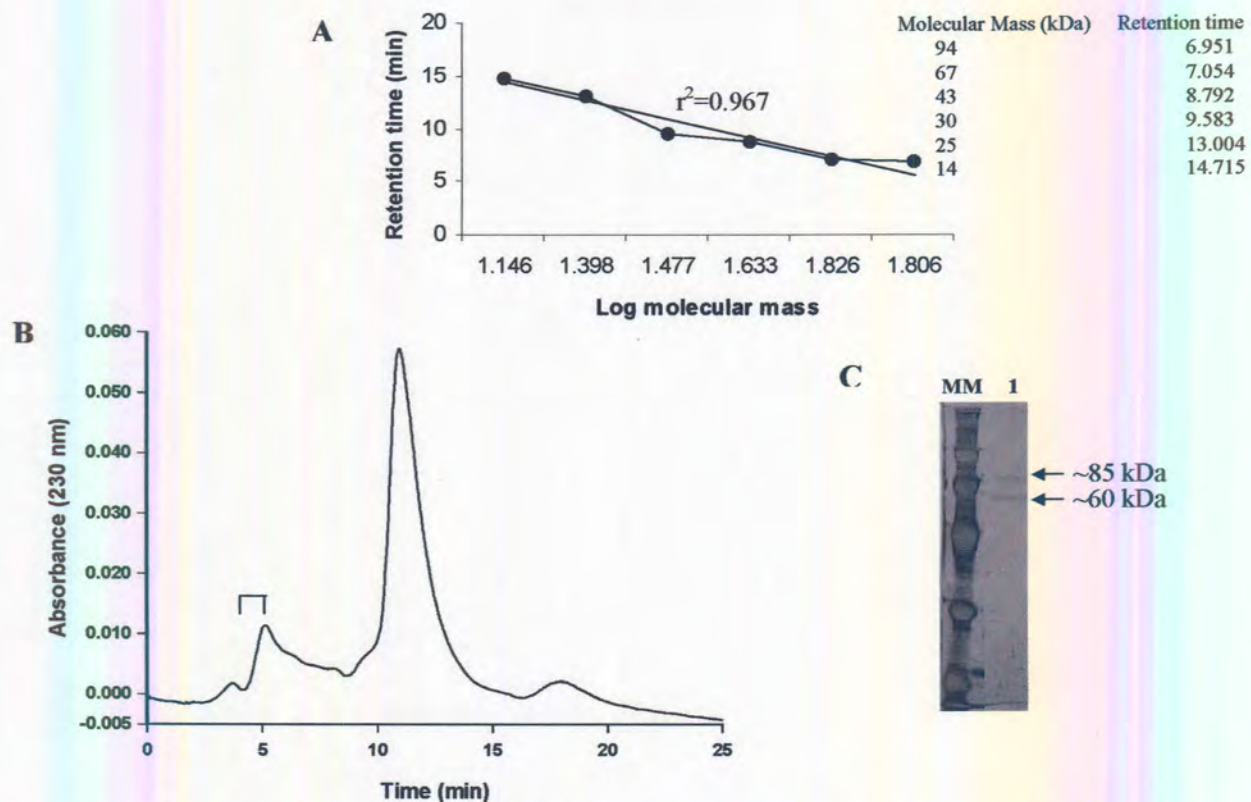
Figure 3.7: Size-exclusion HPLC of the monofunctional ODC purified with affinity chromatography. (A) Calibration curve for the size-exclusion HPLC. **(B)** Elution profile of homodimeric ODC (170 kDa). The peak containing ODC activity is indicated with a horizontal bracket. **(C)** SDS-PAGE of fractions corresponding to the ODC-specific peak indicating only two bands at ~85 and 60 kDa. MM: Low molecular weight protein markers.

AdoMetDC was further purified with SE-FPLC on a calibrated column (correlation coefficient of 0.98307 for the linear regression line derived, Fig. 3.8 A). Fig. 3.8 (B) indicates the elution profile for AdoMetDC showing peaks corresponding to heterotetrameric AdoMetDC of ~140 kDa (fraction 47) and the heterodimeric form of ~72 kDa (fraction 51). No bands were visible with SDS-PAGE. AdoMetDC activity could be detected in both fractions corresponding to the expected heterotetrameric as well as heterodimeric forms of the protein. A ~145 kDa peak would correspond to a heterotetrameric AdoMetDC of 144.6 kDa (two 63.6 kDa and two 9 kDa fragments) and the ~72 kDa peak to the heterodimeric form.
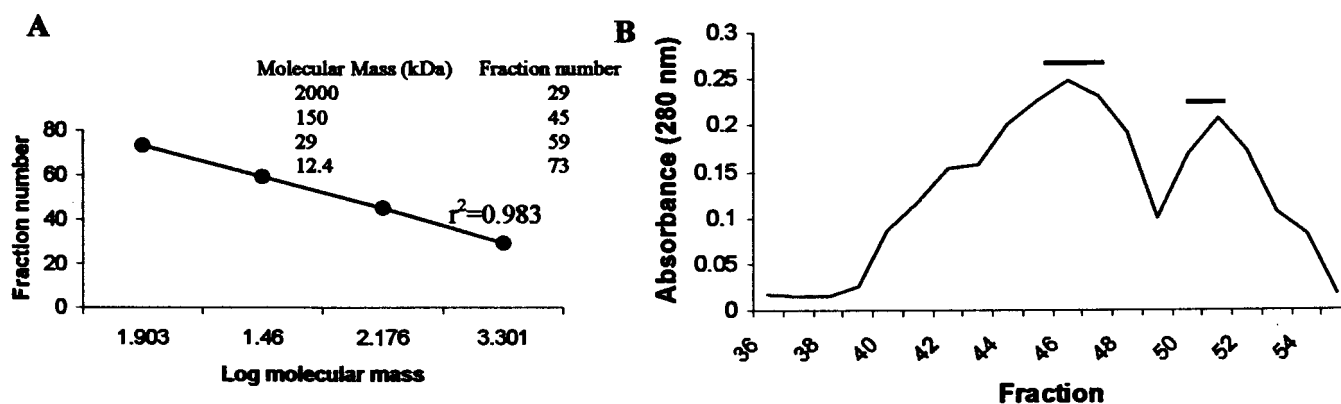
Figure 3.8: SE-FPLC curve for separation of the monofunctional AdoMetDC. (A) Calibration curve for the SE-FPLC. **(B)** Elution profile of AdoMetDC. The horizontal bars indicate AdoMetDC activity.

3.3.5) Expression and purification of the bifunctional PfAdoMetDC/ODC.

3.3.5.1) Expression and affinity chromatography of Strep-Tag PfAdoMetDC/ODC.

The full-length bifunctional PfAdoMetDC/ODC was expressed in an *E. coli* mutant lacking AdoMetDC- and ODC-activities as a C-terminally linked Strep-Tag protein, consisting of 1419 residues containing both the AdoMetDC and ODC domains (Müller, *et al.*, 2000). The protein was purified on a Strep-Tactin affinity column as described above. Expression levels obtained ranged between 100-250 µg protein per litre culture. SDS-PAGE of the protein indicated the expected protein band at ~160 kDa corresponding to the heterodimeric form of the protein (Fig. 3.9; 156 kDa consisting of two subunits of a covalently linked AdoMetDC+ODC of 147 kDa and the β-subunit of AdoMetDC of 9 kDa). Several contaminating bands of smaller molecular weight were observed after staining with silver.
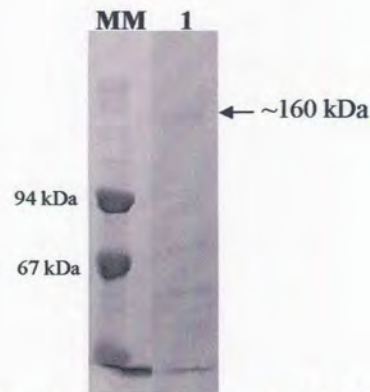
Figure 3.9: SDS-PAGE of the recombinantly expressed bifunctional PfAdoMetDC/ODC. Lane 1 indicates the heterodimeric form of the protein at ~160 kDa. MM: Low molecular weight protein molecular mass marker. Proteins were visualised with silver staining.

3.3.5.2) Size-exclusion FPLC (SE-FPLC) of the bifunctional PfAdoMetDC/ODC.

The bifunctional PfAdoMetDC/ODC isolated with affinity chromatography was further purified by SE-FLPC as described (Wrenger, *et al.*, 2001). Wild-type PfAdoMetDC/ODC elutes at a molecular mass of ~ 330 kDa (fraction 35) corresponding to a heterotetrameric complex of the PfAdoMetDC/ODC polypeptide as indicated in Fig. 3.10 A). Subsequent SDS-PAGE of the single peak indicated the presence of the PfAdoMetDC/ODC in the denatured heterodimeric form of ~160 kDa (Fig. 3.10 B).
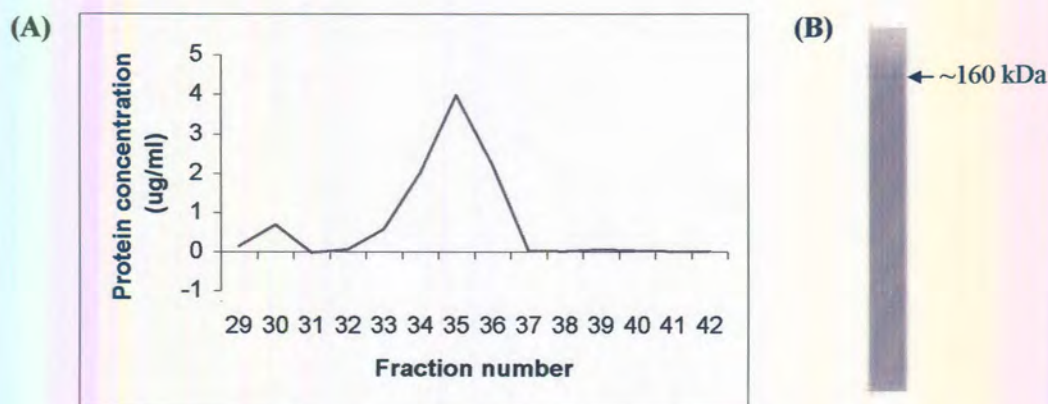
Figure 3.10: SE-FPLC purification of the bifunctional PfAdoMetDC/ODC. (A) Indicates a typical elution profile for PfAdoMetDC/ODC with the size of the protein in fraction 35 corresponding to ~330 kDa. **(B)** SDS-PAGE of fraction 35 indicating a single ~160 kDa band for denatured PfAdoMetDC/ODC.

3.3.6) Decarboxylase activities of the monofunctional AdoMetDC and ODC or bifunctional PfAdoMetDC/ODC.

ODC and AdoMetDC activities of the Strep-Tag expressed monofunctional and bifunctional proteins were assayed. Active PfODC is an obligate homodimer of ~170 kDa and is expressed containing part of the hinge region (see also Fig. 3.4). PfAdoMetDC is able to form a heterotetramer of 145 kDa (two 64 kDa and two 9 kDa fragments) but this is not essential for activity, however, the recombinantly expressed protein also contains part of the hinge region. The active form of PfAdoMetDC/ODC is a heterotetrameric structure of two 147 kDa subunits (two polypeptides containing both AdoMetDC and ODC activities) and two 9 kDa subunits (processed β-subunits of AdoMetDC) forming a ~330 kDa complex. Fig. 3.11 schematically represents the active forms of the protein.
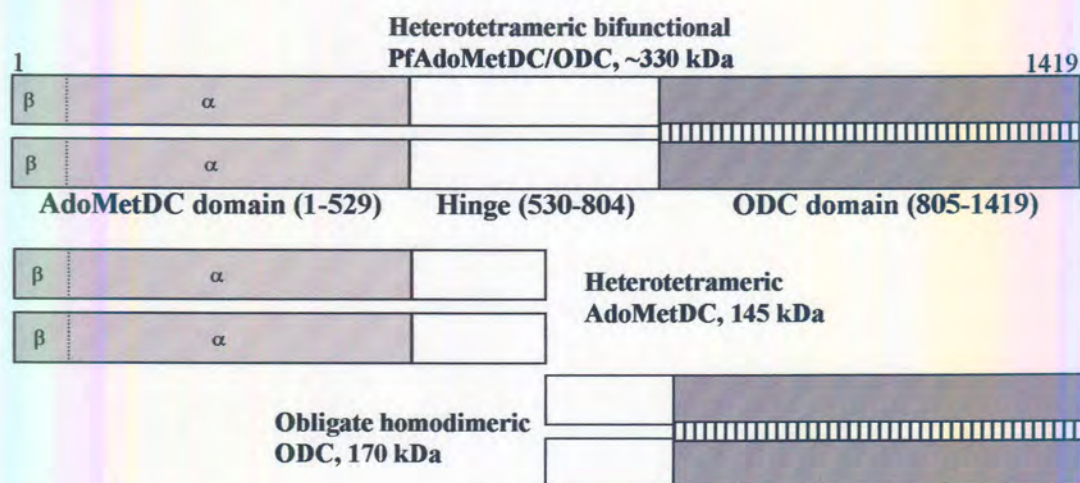
Figure 3.11: Schematic representation of the active forms of the monofunctional AdoMetDC and ODC or the bifunctional PfAdoMetDC/ODC. The dimerisation of the ODC domain is indicated with vertical lines between the monomers, and the α- and β-chains of AdoMetDC are separated with a dashed line.

Table 3.3 summarises the specific activities of the monofunctional AdoMetDC and ODC as well as for both decarboxylase domains in the bifunctional PfAdoMetDC/ODC.

Table 3.3: Decarboxylase specific activities of monofunctional AdoMetDC or ODC and bifunctional PfAdoMetDC/ODC. Results are expressed as the mean of three independent experiments performed in duplicate with standard deviations indicated. N/A: not applicable.

	AdoMetDC specific activity (nmol/min/mg)	ODC specific activity (nmol/min/mg)
Monofunctional AdoMetDC	7.38 ± 2.72	N/A
Monofunctional ODC	N/A	1.58 ± 0.46
Bifunctional PfAdoMetDC/ODC	22.19 ± 7.5	27.32 ± 14.1

The activities of the monofunctional proteins are reduced compared to the corresponding activities in the bifunctional protein. AdoMetDC is three times less active in the monofunctional form whereas ODC is 17 times less active in the monofunctional protein.

3.3.7) Analyses of the deduced amino acid sequence of the bifunctional PfAdoMetDC/ODC.

The deduced bifunctional PfAdoMetDC/ODC amino acid sequence was compared with multiple alignment to the sequences of monofunctional homologues found in other organisms (human, murine, *T. brucei* and *L. donovani*) (Fig. 3.12).

As described by Müller *et al.* (Müller, *et al.*, 2000) and evidenced in Fig. 3.12, residues 1-529 show similarity to AdoMetDCs and residues 805-1419 has homology to ODCs. However, there is no homologous sequence to the hinge region (residues 530-804) to date in the SwissProt Protein Database. The SMART server was used to define and classify the domains based on their homology with other protein families. The malarial ODC, as all the other ODCs, was classified to belong to the Group IV decarboxylase family of proteins. This family includes ornithine, arginine and diaminopimelic acid decarboxylases, probably based on a shared function and similarity between substrates. The AdoMetDC domain is grouped in the single protein family of S-adenosylmethionine decarboxylases.

The parasite-specific inserts in the bifunctional PfAdoMetDC/ODC were defined based on Fig. 3.12 as well as from previously published results (inserts of more than 10 residues)(Müller, *et al.*, 2000). The inserts in PfAdoMetDC/ODC have no sequence homologues in any other organism and are interspersed between well-conserved blocks compared to the respective homologous proteins of other organisms. Three major inserts of more than 10 residues each were identified; an area of 197 residues in the PfAdoMetDC domain (insert A_1: residues 214-410) and two areas in the PfODC domain: a 39 residue insert close to the PfAdoMetDC domain and a large 145 residue insert near the C-terminus of the protein (insert O_1: residues 1047-1085 and insert O_2: 1156-1301). The hinge region was defined by Müller *et al.* (Müller, *et al.*, 2000) as the 180 residues (H: residues 573-752) that connect the PfAdoMetDC and PfODC activities. Some discrepancy exists as to the exact boundary of this area as defined by

sequence homology to other monofunctional proteins using different multiple sequence alignments (Müller, *et al.*, 2000).

After removal of the parasite-specific regions and hinge region, the separate domains show moderate identity with the respective proteins from other organisms. The *P. falciparum* AdoMetDC domain has a 18.1% sequence identity with the human homologue, and sequence identities of 17.8%, 15% and 14.2% with the murine, *T. brucei* and *L. donovani* enzymes, respectively. The ODC domain shows a more pronounced sequence identity of 27.6%, 26.6%, 30.7% and 18.6% with the human, mouse, *T. brucei* and *L. donovani* monofunctional proteins, respectively.
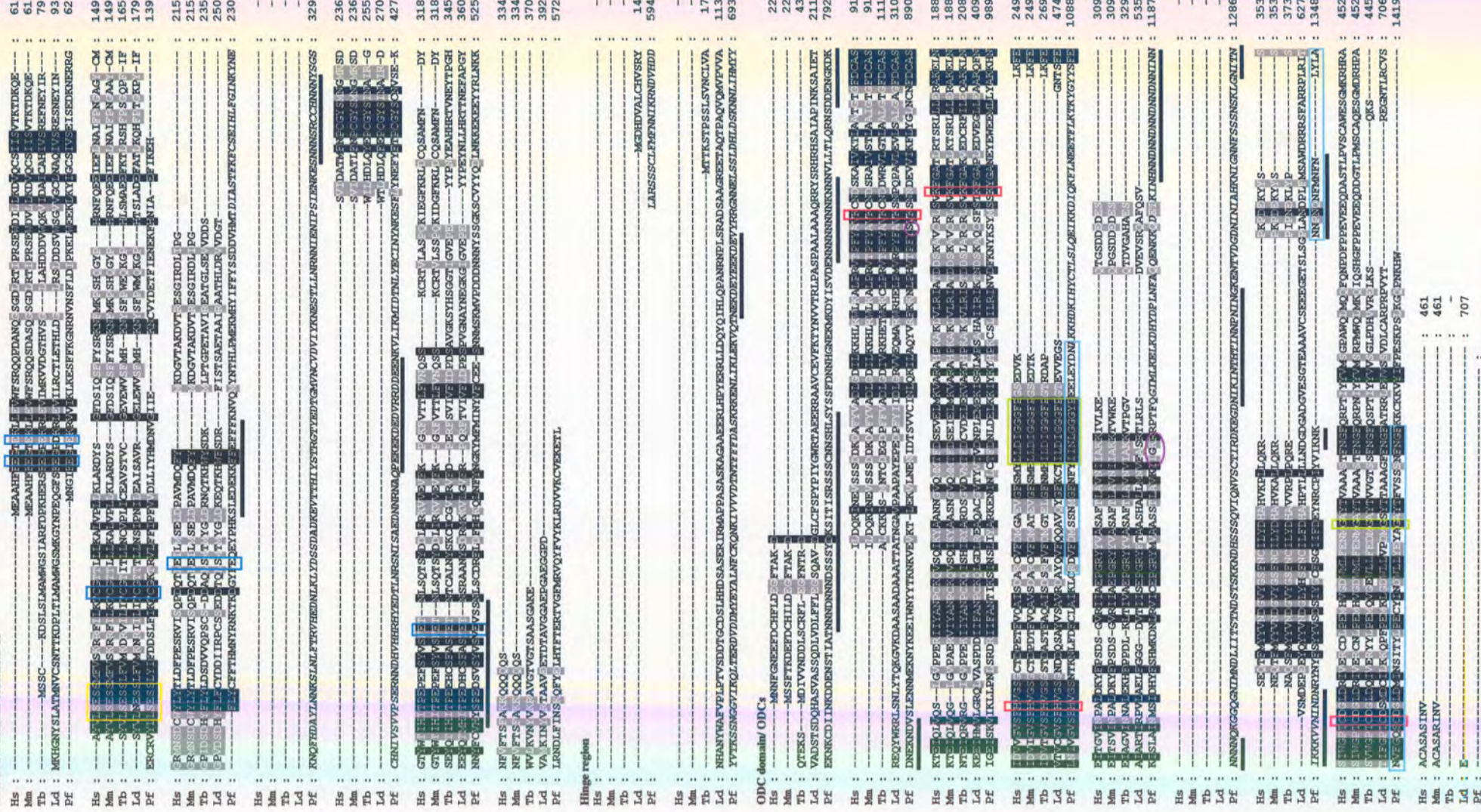
Figure 3.12: Multiple alignment of the bifunctional PfAdoMetDC/ODC amino acid sequence with homologues of the monofunctional AdoMetDC and ODC from other organisms. The deduced amino acid sequence of the PfAdoMetDC/ODC was aligned to the monofunctional sequences from *H. sapiens* (Hs*)*, *M. musculus* (Mm), *T. brucei* (TB) and *L. donovani* (Ld). The AdoMetDC and ODC domains are indicated as well as the hinge region as defined by Müller *et al.* (Müller, *et al.*, 2000). Residues with >80% conserved properties are indicated in black boxes, and >60% in grey. The processing site of AdoMetDCs is indicated in a yellow box, and residues important for activity of AdoMetDCs in blue boxes. Active site residues of ODCs are indicated with red boxes, whereas the proposed amidation site is in a purple circle. The Gly-rich area and Gly involved in dimerisation of ODC are in green boxes. Possible PEST-rich regions are in light blue boxes. The parasite-specific inserts are indicated in italics. Low-complexity areas are indicated by horizontal bars.

CHAPTER 3: Characterisation of *P. falciparum* AdoMetDC and ODC proteins

The Wootton and Federhen algorithm (SEG algorithm) was used to analyse the sequence for possible low-complexity areas. Fourteen such areas were identified (Fig. 3.12) but these were not equally dispersed throughout the sequence. The majority of the low-complexity regions (10 of the 14) were located in the parasite-specific inserts and hinge region.

Amino acid composition analyses of the 1419 residues indicated a marked prevalence for charged residues. The predominant amino acid is Asn making up 13.7% of the total residues, followed by 9.3% for Lys, 8% for Ser, 7.3% for Glu, 7.1% for Ile and 6.8% for Asp. The protein is furthermore characterised by repeats, specifically $(N)_x$- and $(NND)_x$- repeats (Fig. 3.12). These repeats are also more prevalent in the parasite-specific inserts in both decarboxylase domains as well as in the hinge region connecting the two domains.

There are some highly conserved amino acids and regions between all the proteins, most of which have been reported to be essential for catalytic activity, dimerisation and pro-enzyme processing as previously indicated (Müller, *et al.*, 2000). PROSITE analyses of the bifunctional protein indicated 14 possible sequence motifs as sites of phosphorylation by protein kinase C (Ser and Thr), as well as several sites for phosphorylation by Casein kinase II (e.g. TI/ERD or SSL/YD) and Tyrosine kinases (KKEKEEYY). The majority of the phosphorylation sites were present in either the AdoMetDC or ODC domains, and not in the hinge region or parasite-specific areas. However, the Ser_{303} of mammalian ODCs that is implicated in regulation of enzyme activity through phosphorylation is not conserved in the PfAdoMetDC/ODC (GSD vs. NEK in PfAdoMetDC/ODC). Eight different N-myristoylation sites (motif G-N-[EDRKHPFYW]-N-[STAGCN]) were identified, 3 in the AdoMetDC domain and 5 in the ODC domain. A single amidation site (motif GR/KR/K) was identified in the ODC domain at position 1131 (IGKR).

Only two poor PEST-rich regions were predicted in the ODC domain, residues 1017-1049 and in the C-terminus residues 1333-1396 (Fig. 3.12). The most C-terminal region slightly overlaps with the strong PEST-rich region found in mammalian ODCs (residues 423-449).

3.3.7.1) Secondary structure and hydrophobicity analyses of the PfAdoMetDC/ODC amino acid sequence.

Secondary structure analyses with the GOR IV algorithm indicate a preference for α-helices in the AdoMetDC domain (Fig. 3.13). The ODC domain exhibits a α-helix-β-sheet domain close to the hinge region with a strong β-sheet domain in the C-terminal of the protein (ODC domain). The hinge region and parasite-specific inserts contain unstructured areas or coils.
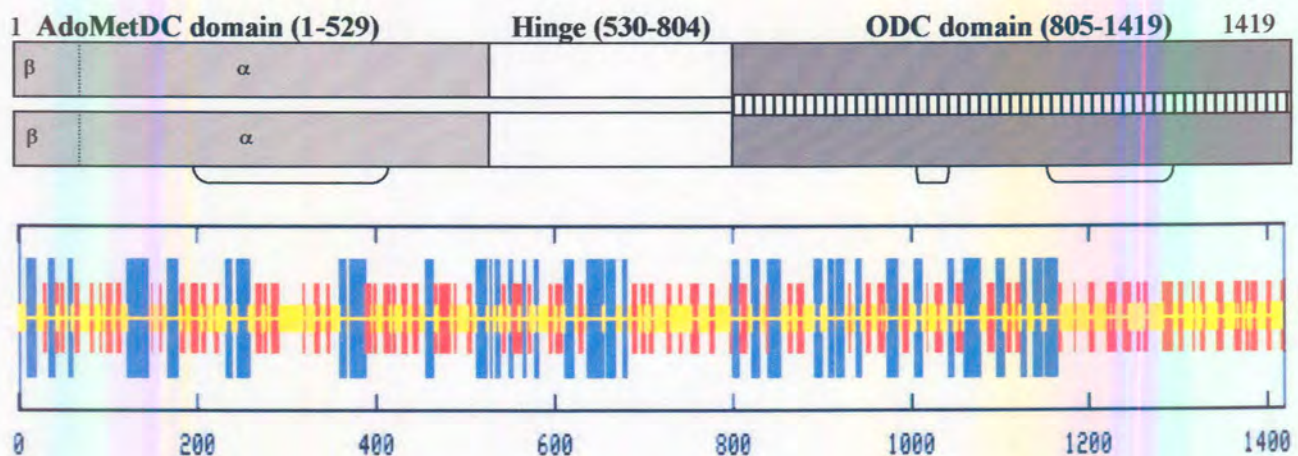
Figure 3.13: Secondary structure prediction of the PfAdoMetDC/ODC amino acid sequence. Top panel indicates the schematic representation of the bifunctional PfAdoMetDC/ODC for reference. The bottom panel indicates α-helices in blue and β-sheets in red. Unstructured areas or coils are indicated in yellow. The horizontal brackets indicate the relative positions of the parasite-specific inserts.

The Kyte and Doolittle hydropathy profile of PfAdoMetDC/ODC indicates a higher abundance of hydrophilic residues compared to hydrophobic amino acids (Fig. 3.14). A significant apolar peak is observed in the AdoMetDC domain around residue 100. This coincides with an area of β-sheets followed with an α-helix rich region (Fig. 3.13) and is also close to the processing site of the AdoMetDC domain around residue 73. Amino acid residues 900-1150 and 1300-1400 are the most apolar and corresponds to the structured α-helix-β-sheet and β-sheet regions in the ODC domain, respectively (Fig. 3.13).
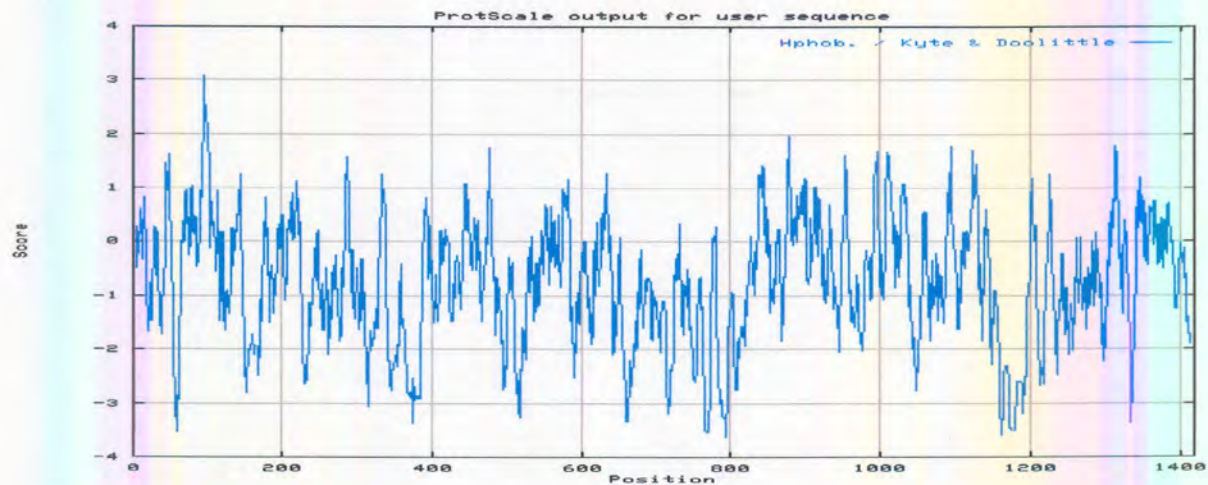
Figure 3.14: Hydrophobicity plot of the deduced PfAdoMetDC/ODC amino acid sequence. Hydrophilic residues are scored negatively.

3.3.7.2) Comparison of the relationship of PfAdoMetDC/ODC with homologues.

The bifunctional nature of PfAdoMetDC/ODC complicates phylogenetic analyses of the protein. Therefore, the PlasmoDB database was queried for similar bifunctional enzymes in other *Plasmodia*.

Partial sequences of AdoMetDC and ODC were obtained for *P. chabaudi* and *P. knowlesi* (data not shown). Full open-reading frames were identified for bifunctional AdoMetDC/ODC in the two murine *Plasmodium* species, *P. berghei* and *P. yoelii*. The deduced amino acid sequences of the bifunctional proteins from *P. falciparum*, *P. berghei* and *P. yoelii* are compared in Fig. 3.15. It appears that the bifunctional nature of these enzymes is conserved in *Plasmodia* but not in other protozoa such as *T. brucei* and *L. donovani* (Hanson, *et al.*, 1992; Phillips, *et al.*, 1987).

The *P. falciparum* amino acid sequence shows 41% identity (60% similarity) with the *P. berghei* sequence. In comparison, *P. falciparum* AdoMetDC/ODC is 42% identical (61% similar) with the *P. yoelii* sequence. The murine species share significant sequence identity (91%) and 96% sequence similarity. The parasite-specific insert A_1 in PfAdoMetDC/ODC (residues 214-410) is extended by 101 residues in the murine species. However, both the hinge region (residues 573-752) and large insert in the ODC domain (O_2: 1156-1301) is longer in the *P. falciparum* sequence. The smaller insert in the ODC domain (O_1) is better conserved between all three species. These characteristics are discussed in more detail in Chapter 4.

97

```
Pf : MNGIFEGIEKRVVIKLKESFFKGNRNVNSFLDIPKELWEEKLKYIGCSIVSEISEDKNERRGERCRVYLLSESSLYIFDDSLFIKTCGKTRVLFFIPFV : 99
Pb : MSNVFEGIEKRVVIKLRKNVFMNGNIN-SILLDISRDLWEEKLHLIGCNIVSEIKEDINAPLNERSRVYLLSESSLYIYNDMIYIKTCGKTKVLFFIPFL : 98
Py : MSNVFEGIEKRVVIKLRKNLFMNGNIN-SILLDISRDLWEEKLNLIGCNIVSEIKEDINAPLNERSRVYLLSESSLYIYNDMIYIKTCGKTKVLFFIPFL : 98

Pf : VDLLIYHMDNVGIIEKNCVYDETFIENEKPHNIA-EFIKEHFLYCFFTHMIYRNKTKDGYFEQEYPHKSLEDEKFFFDFDFKNVQMYNHLPMDKMHYI : 197
Pb : VDILIYKWKKYFVLEKPNLYNESFIQNNTKLEEILDFIQNYFEYSFFTHMIYKIVTKDGYYEQEYPHKSVDDEKVFFKFFLKSMEFISTELPMNRNHYI : 197
Py : VDILIYKWKKYFVLEKSNLYNESFIQNDTKLEEILDFIQKYFEYSFFTHMIYKNVTKDGYYEQEYPHKSVEDEKVFFKFFLKSMEFVNTELPMNRNHYI : 197

Pf : FFYSSDDVHMTDIAS---TFKFCSEIHLFGINKYNEKNQFHDAYDNNKSLNLFTRVHEDNLKLYDSSDADKEVTTHIYSTRGTYEDTGMVNCVDVIYKN : 293
Pb : FFSQVNKDATIDTSALNVFPNFCSEMHLFGIKKYSMFNQFHSEYLTYDSLNVLSKYSINIPKSEESEKIEVONSSVYSFDSVENSSKYTSNTAINNITV : 296
Py : FFSQVNKDATIDTSALNVFPNFCSEMHLFGVKKYSMPNQFHSEYLTYDSLNVLSNHSLNIPKSEESEKIEVOCSSSVYSFDSVENSSKYTNNTANTNVTV : 296

Pf : ESTLLNRNNIENIPSIENKESNNNSRGCHNNNYSGSCHNIVSVVPSERNNDHVHHRHYEDTLNRSNISAEDNNRNAQPEKEKD------------------ : 376
Pb : NVSKDLDNNSINTLYNTIDDQTVMLTCSDDKINNLNEFDSKLLLAEENSKIMDKDNTQVFD---NNNYIFISTPKAQETYSNTNSVRSNEKSTCSSN-T : 391
Py : NVSKDIDNNSINTLYNTVDDQTVILTCSDDKNNNLNESDSKLLLAGENSKNMDKDNTQVFNDNNINNYIFISNPKVQETYSNTNSVRSNEKSTCSSSSA : 395

Pf : ------------------------------------------------------------------------------------EDVRRDDEENK-- : 387
Pb : YTSLLQNDLKEFHPKNNSMDTIEDDDGKLLVEEEVISYISNEETIDAKVEQLDNLSNKNMLDNNLDKNYEVTEAAGPSLSSFTIACSDFLNKNANDEKGK : 490
Py : YTSLLQNDLKEFHPKNNAMDTIEDDGRLLVEEEDISYISNEETIDAKVEQIDNLSNKSILDNNLDKNYDT---NEPSLSSFTVACSDLINKNANDEKSE : 491

Pf : ------VLIKMIDTNLYECINYNK-DSFLYNEFYFTPCGYSCNVSEKNNYTCVHYSPEDSVSYVSVEVSSNLSCDRFLDFIHKQLNFYNGKYMFMINYV : 479
Pb : DNTSSNMGIKKIEENLYECINIQNNEKFMTNEFYFVPYGFSCNVANKNNYLCVHYSPEDGVSYVSVELSCVLKSENFMNFIKKQLNFYKAQYLYIVNYL : 589
Py : DNTSKNIRIKKNIEENLYECINTQNNDKFIYNEFYFYPVGFGFSCNVANKNNYLCVHYSPEDCVSYVSVELSCVLKSENFMNFIKKQLNFYKAQYLYIINYL : 590

Pf : FCEESNNYSKMVPDDDNNTYSSGKSCVYYQDLNKKER--EEYYRLNKKLRNDLFINSKQFYELHTPTERIVGFMPVQYPVYKLRDVVKCVEKETLJARS : 576
Pb : YDFNEDTLAQKSGDNNSDLTNRFLSVLLSSQNNKIKSINEVIEVNKKLGNNIIINNDQYYSLCAKKEQSVGFLKIQYCVYEQKNFIRNEE---RIKPY : 685
Py : YDFNENTLAQKSGDNNKDSNNRFLSVLLSSQNNKNIKSINEVVEVNKKLDNNIIINNGQYYSLCAKKEQSVGFLKIQYCVYEQKNFIRNEE---RIKSY : 686

Pf : SSCLEMFNNIKRNDVHDDYVTKSSNGGVIKOLTERVVDDMYEYALNFCKQNKIVVVDTNTFFFDASKRKENLIKLEKVQTNEKDEYEEKDEVYRRGNNE : 675
Pb : NSIYFEQPNKNAQNVN---ATKSTTP-----LSLQYVNSMYEYSYNFCKENNIHIIDI---------------------------------------KDD : 738
Py : NSICFEQSNKNAQNVN---TTKSTT------LSLQDVNSMYEYSYNFCKENNIHIIDI---------------------------------------EDD : 738

Pf : ILSSLDHLDSKNNLIHMYYEKNKCDIINKDDENSTIATNNIDNNNDSSSYDKSITTISRSSSCNNSHLSYTSFDNNIGNEKMKDYISVDENINNNNNNYKNN : 774
Pb : LNSMD-------------------------NTSNSNRNNYSISNSSDHADDIKMERNS---------SILNYAHETRKRTDHISSFLIDGOEFQSEKS : 802
Py : LNSID-------------------------NSSNSNHNIVPISNSSDNVDDIKMRKDS---------NILKESYETEKRTDPISSFLIDEQKAQNEKS : 802

Pf : NVLLTLQRNSDDENGKDKDNEKNDVSLENNMEKNVREEINYYTKNKVEVKT-LEKVLNENIDTSVVCINLQKIILAYYVRFKKNLPHVTPFYSVKSNND : 872
Pb : -------RLGSETNSNTTDDDNN----EFVLEKKHKKEMINYYRNNKAEICT-FNKMLNENIDTSVICINLQKILLQYIRFKRNLPRVTPFYAVKSNDD : 889
Py : -------RSSSETNSNITDDDNN----EFVLEKKYKKEMINYYRNNRVEIST-FNKMLNENIDTSVFCINLQKILLQYIRFKRNLPRVTPFYAVKSNDD : 889

Pf : EVVIKFLYGLNCNFDCASIGEISKVIKLLPNLSRDRIIFANTIKSINSLIYARKENINLCTFDNLDELKKIYKYHPKCSLILRINVDFKNYKSYMSSKY : 971
Pb : EVVLKFLYGLNCSFDCASIGEINKLITLLPNISIDRIIYANTIKSPASLKYAKERNINLCTFDNIDELKKICKYHPTCSLLLRINIDFKNYKSYMSSKY : 988
Py : EVVLKFLYGLNCSFDCASIGEINKLITLLPDISTDRIIYANTIKSPASLKYAKEKNINLCTFDNIDELKKICKYHPTCSLLLRINIDFKNYKSYMSSKY : 988

Pf : GANEYEWEEMLLYAKKHNLNIVGVSFHVGSNTKNLFDFCLAIKLCRDVFDMSSNMGFNFYIINLGGGYPEELEYDNAKKHDKIHYCTLSLQEIKKDIQK : 1070
Pb : GANEYEWEKILKFGKENIMNIIGVSFHVGSNTKNIFDFCQAIKLSREVFNISHRLGYKFEILNMGGGYPEELEYDFAKKNEEKHYCTLNEEELKKYIMS : 1087
Py : GANEFEWEKILKFGKENIMNIIGVSFHVGSNTKNIFDFCQAIKLSREVFDISHRLGFKFRILNMGGGYPEELEYDFAKKNEKKHYCTLNEEELKKYIMS : 1087

Pf : FLNELTFLKTKYGYYSFEKISLAINMSIDHYFSHMKDNLRVICEPGRYMVAASSTLAVKIIGKFRPTFQGIMJKELKDHYDPLNFAQQENKKQDETKIN : 1169
Pb : FLNEKSVIKKNYNFYNFEKIALAINMSIDHYFKDIKDQLRIISEPGRYMVASSTLAAKVIGKRCPTFSPIRLSULKN--------AEDNND-NVVSNE : 1177
Py : FLNEKSVIKKNYNFYNFEKIALAINMSIDHYFKDMKDQLKIISEPGRYMVASSTLAAKVIGKRCPTFSPIRLSNLKN--------AEDINDDNEVKNE : 1178

Pf : HNNDNNDNNDNNNDNNHNNNNNNQKGGOGNIMNDLIFTSTNDSTSKKNDHSSSQVIQNVSCTIRDKEGDNIKINAHTINNPNINGKENAVDEDNINIAHK : 1268
Pb : QVANNDNYQQNDDGNIQNKD------QGNCTG---AIPIAELKTEKESKHLTSEVCTNG-----------BIQTNNVYILSTNHENVNSFSPTENGKTV : 1256
Py : QVANNDNDQQNGDGNIQNKD------QGNCTG--AIPDVELRTEKESKHPSEVCTNG-----------BIQTSNVCNTNTDNENVSSFGTTTNGKTV : 1257

Pf : NIGNNESS-SNSKLGNITNIKKKVVNINDNRYNYFSYYVSDSIYGCFSGIIFDEYNRCPIYVIKNK------NNPNQNFMNFNLYLANVFGQSCDGLDM : 1360
Pb : IPGSAFNHEDNSKLGNITNIKKKVVNIKDNRYNYYSYYITDSIYGCFSGIIFDEYSRPPMYIINNNKIDSKQMILDSP-----LYLTNIFGQSCDGLDM : 1350
Py : IPGSAFNHEDNSKLGNITNIKKKVNIKDNRYNYYSYYITDSIYGCFSGIIFDEYSRPPIYINNNKMDSKKMILDSP-----LYLTNIFGQSCDGLDM : 1351

Pf : TNSITYLPECYINDWLLYEYAGAYTFVSSSNFNGFKKCKKVYIFPESKPSLKGQPNKHW : 1419
Pb : ITSSTYLPECHINDWIIYEYTGAYTFVSSSNFNGFCSCOKIYIFPRNKFPFSV------ : 1403
Py : ITSSTYLPECHINDWIIYEYTGAYTFVSSSNFNGFCPCKKIYIFPRNK----------- : 1399
```

Figure 3.15: Multiple sequence alignment of the deduced amino acid sequence of the bifunctional AdoMetDC/ODC from three *Plasmodium* species. Pf: *P. falciparum*, Pb: *P. berghei* and Py: *P. yoelii*. Residues >80% conserved are boxed in black and >60% conserved in grey.

3.4) DISCUSSION.

3.4.1) Heterologous expression of the decarboxylase proteins.

The bifunctional PfAdoMetDC/ODC was successfully expressed as a soluble protein with distinct decarboxylase activities for both the AdoMetDC and ODC domains. Furthermore, monofunctional PfAdoMetDC and PfODC could also be expressed in two different systems. However, expression levels for the different forms of the proteins were very low. Maximally 250 µg recombinantly expressed protein were obtained for either the monofunctional ODC or the bifunctional protein using the Strep-Tag expression system. However, the proteins were not homogenously isolated after affinity chromatography and the correct concentration of the proteins is most likely lower than the value obtained implying higher activities in terms of protein concentration. The expression level could not be increased using different promotor systems (i.e. T7 RNA polymerase in the His-tag system) or different protein tags. The His-tag based system was not further pursued in this study due to the initiation of collaborations using the Strep-Tag based expression system.

High-level expression of recombinant *P. falciparum* protein is generally not easily obtained. Reasons include the codon bias of this organism, the instability of the proteins in heterologous systems as well as incorrect folding of the proteins (Baca and Hol, 2000). These obstacles can in some instances be circumvented by including tRNAs or synthesising the gene to be expressed using *E. coli* codon preferences.

Co-transformation of plasmids containing the low abundance tRNAs for Arg, Ile, Leu or Gly in *E. coli* (frequencies of ~2, 4, and 3 codons used per 1000 codons) were required to achieve these protein levels. These codons are more frequently used in *P. falciparum* proteins (frequencies of 27, 33 and 15 per 1000 codons for Arg, Ile and Leu respectively, www.dna.affrc.go.jp/~nakamura/codon.html). Plasmids encoding these tRNAs have been shown to markedly increase expression of malarial proteins in heterologous systems (Baca and Hol, 2000). Recombinant proteins further needed to be stabilised by the addition of non-ionic detergents and thiol reducing agents. The total protein content for one litre *E. coli* culture (10^9 cells/ml) is usually 155 mg (Lodish, *et al.*, 1995), therefore a mere 0.062% of the total protein comprised the recombinantly expressed protein. Other expression systems were also tested including gene-complementation in ODC-deficient *S. cerevisiae*. Function could not be restored to

99

workable levels in these mutants indicating that efficient expression of the malarial ODC could not be achieved to restore the gene deficiency (Results not shown).

In both ODC expression systems, low molecular weight proteins were present after affinity chromatography, which necessitated size-exclusion chromatography. In particular, a ~60 kDa band was present in all preparations of recombinantly expressed PfODC, even after size-exclusion chromatography that removed other contaminating proteins (Fig. 3.7). Explanations for the presence of this band include inadequate separation of the proteins during size-exclusion chromatography, however, calibration of the column sufficiently separated phosphorylase b (94 kDa), BSA (64 kDa) and ovalbumin (43 kDa) (Fig. 3.7). Analyses of fractions corresponding to a size of 60 kDa did not reveal the expected contaminating band. Further explanations of the 60 kDa band include degradation products of PfODC, however a mixture of protease inhibitors was included throughout the isolation procedure and PfODC was expressed in three different fusion-protein systems (Strep-Tag and His-Tag described here and a His-Tag system described in Krause *et al.* (2000) and the contaminating band was present throughout. All three systems used N-terminal tags, therefore the product could have been cleaved in the C-terminus. This shortened PfODC will still associate with the normal PfODC and be isolated during SE-HPLC. Western-blot analyses of Strep-Tag expressed PfODC with a polyclonal antibody against PfODC also identified the ~60 kDa band and it was speculated that this is indeed a degradation product of PfODC (Krause, *et al.*, 2000). Therefore, this unresolved aspect needs further clarification. Determination of the N-terminal sequence of the 60 kDa protein will indicate its identity.

3.4.2) Multimeric states of the monofunctional and bifunctional proteins.

Analyses of the oligomeric state of the AdoMetDC indicated that the protein is active in the heterotetrameric form made up of two 64 kDa and two 9 kDa chains (Fig. 3.8). However, it is not clear if this arrangement is essential for activity since AdoMetDC activity was also found in size-exclusion fractions corresponding to the protomer of the protein, the heterodimeric form consisting of a single 64 kDa α-chain and 9 kDa β-chain. However, protein dimerisation after size-exclusion fractionation to reform the heterotetrameric AdoMetDC cannot be excluded as one explanation for activity. It is possible that the heterotetrameric and heterodimeric states may be in equilibrium with each other. AdoMetDC from mammalian species are active heterotetramers but

heterodimeric AdoMetDCs are present in plants (Ekstrom, *et al.*, 1999; Ekstrom, *et al.*, 2001).

Active monofunctional ODC is present as an obligate homodimer of ~170 kDa and migrates in its monomeric form as ~85 kDa on denaturing SDS-PAGE (Fig. 3.7). SE-HPLC revealed only the homodimeric form of the protein, no ODC was observed in fractions corresponding in size to the monomeric form of the protein after SDS-PAGE. However, ODC homodimers and monomers in other eukaryotes have been shown to be in rapid equilibrium (Coleman, *et al.*, 1994). This difference could be due to the unusual bifunctional complex of PfAdoMetDC/ODC.

The bifunctional PfAdoMetDC/ODC was expressed as a Strep-Tag fusion protein of ~330 kDa as shown by size-exclusion chromatography (Fig. 3.9 and 3.10). This corresponds to a heterotetramer consisting of two 147 kDa subunits (covalently linked α-chain of AdoMetDC with ODC) and the two 9 kDa β-chains of AdoMetDC. Two proproteins are therefore post-translationally processed and cleaved within the AdoMetDC domain and assembled into the heterotetrameric complex. These results are in good agreement with the previously described size and arrangement of the protein (Fig. 3.3) (Müller, *et al.*, 2000).

3.4.3) Decarboxylase activities of AdoMetDC and ODC in their monofunctional or bifunctional forms.

The decarboxylase activities of the monofunctional proteins are noticeably less compared to the corresponding activities of the bifunctional form of the protein (Table 3.3). AdoMetDC is three times less active in the monofunctional form compared to the bifunctional form whereas the monofunctional form of ODC seems less active than the activity found in the bifunctional PfAdoMetDC/ODC. These results are comparable to previously determined activities (Wrenger, *et al.*, 2001). This indicates that the natural bifunctional arrangement of the PfAdoMetDC/ODC and interactions between the domains have a direct impact on either the conformation of the active site centres or accessibility of the respective substrates, or both. Considering that the Michaelis constants of AdoMetDC and ODC in their monofunctional forms correspond to those of the bifunctional protein, the affinities of the active sites seems comparable in either state (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001) and the difference in activity might be due to differences in accessibility and inherent conformation of the

101

active site pockets. The PfODC activity in the bifunctional PfAdoMetDC/ODC is still two orders of magnitude lower than in mammalian proteins (McCann and Pegg, 1992; Pegg, 1989a). This is also true for the ODC activities of other parasitic organisms (Hanson, *et al.*, 1992; Müller, *et al.*, 2001; Müller, *et al.*, 2000; Phillips, *et al.*, 1987), and it is therefore possible that regulation of this activity is not as strict in these highly proliferative organisms and that some cellular mechanisms exist to rather stimulate the relatively inactive enzyme in parasitic protozoa. This might be even more complex in *P. falciparum* because of the bifunctional nature of PfAdoMetDC/ODC.

Interestingly, AdoMetDC activity is not dependent on the presence of putrescine, which normally stimulates both the processing and activity of the protein in other organisms (Müller, *et al.*, 2001). The only other instance of putrescine-independent AdoMetDCs is found in plants (S. Ealick, personal communication). However, a putrescine-dependent form of this protein from *P. falciparum* was originally described with a lower Km value for AdoMet (Rathaur and Walter, 1987). It is possible that this is a protein homologue or pseudogene product and its physiological importance needs to be evaluated. Further structure-activity analyses will be useful in determining how the putrescine-independent AdoMetDC of *P. falciparum* and potatoes are regulated, processed and activated.

3.4.4) Sequence analyses of the deduced amino acid sequence of the bifunctional PfAdoMetDC/ODC.

Analyses of the deduced amino acid sequence of PfAdoMetDC/ODC indicated that the overall homologies to similar proteins from other organisms are relatively low. However, the occurrence of homologous areas close to the proposed active site residues and other structural features in both domains strongly supports the presence of both AdoMetDC and ODC activities on a single polypeptide forming the unique bifunctional protein. Parasite-specific regions in both decarboxylase domains make a significant contribution to the low homology and increased size of the bifunctional protein compared to homologues (Fig. 3.12).

The N-terminal domain of PfAdoMetDC/ODC shows similarity with the family of S-adenosylmethionine decarboxylases. The amino acid sequence of proteins in this family is conserved but bears little similarity to any other protein currently in protein or nucleic acid databases including those of other known pyruvoyl-dependent amino acid decarboxylases. The ODC domain at the C-terminus shows homology to two other

protein families, the ornithine and arginine decarboxylase as well as the ornithine/diaminopimelic acid/arginine decarboxylase families. These proteins all belong to the group IV decarboxylases based on shared function (decarboxylation) and similarity between substrates (PLP dependent enzymes) (Pegg, *et al.*, 1994).

The entire polypeptide has a prevalence for charged residues such as Asn and Lys (Fig. 3.14). This is a characteristic of *P. falciparum* proteins probably due to the codon bias where A and T in the third base is used 3-5 times more frequently than G or C. This results in Asn, Lys, Asp and Glu accounting for 40% of the amino acids in *P. falciparum* proteins on average (Saul and Battistutta, 1988).

The protein contains three major parasite-specific inserts (Fig. 3.12). Analyses of the low-complexity segments of the protein indicated that the majority of these segments are present in these parasite-specific inserts. Pizzi *et al.* (2001) have shown that low-complexity areas found in hydrophilic regions in *P. falciparum* proteins have a good correspondence with parasite-specific, rapidly diverging insertions. The low-complexity regions found in more conserved areas of the protein make up a minor subset of prevalently hydrophobic conserved areas associated with the core structure. The parasite-specific inserts will be discussed in more detail in Chapter 4.

Furthermore, there are areas containing repeats of Nx or (NND)x in the parasite-specific inserts. Repetitive sequences are a major feature of *P. falciparum* proteins (Coppel and Black, 1998; Saul and Battistutta, 1988). These areas are usually short, tandemly repeated sequences. Although repeat sequences in proteins are not unique to malaria proteins, what appears to be different is the number of proteins that contain them and the extent of the repeats and variation of these repeats within some of these proteins (Coppel and Black, 1998; Kemp, *et al.*, 1987). The majority of proteins containing such repeats are predicted to be involved with immune evasion since it has been proposed that immune pressure may be a pivotal force for the generation of variant repeat sequences in a particular protein (Kemp, *et al.*, 1987; Schofield, 1991). Another proposed function is that repeats act as ligands or receptors for host proteins. Of interest is however the observation that the repeats can be found in proteins that would not be expected to be exposed to the immune system (Coppel and Black, 1998; Schofield, 1991). It is therefore unclear what the selective pressure is for the maintenance and

generation of these repeats. Possible functions of the parasite-specific inserts are further investigated in Chapter 4.

The AdoMetDC domain of PfAdoMetDC/ODC contains conserved residues needed for the proteolytic processing as well as enzyme activity (Fig. 3.12)(Ekstrom, *et al.*, 1999; Stanley and Pegg, 1991). Correspondingly, in the ODC domain, the necessary catalytic residues are all conserved at the amino acid level (Tobias and Kahana, 1993a). This indicates that these proteins probably use the same mechanism of catalysis as in other organisms. Post-translational modification sites including phosphorylation, N-myristoylation and trans-amidation sites were predicted and could influence the regulation of the decarboxylase activities. However, it appears that the PfAdoMetDC/ODC may not be subject to the same antizyme-mediated regulation as is observed for ODC in other organisms due to the absence of significant PEST-rich regions usually associated with this mechanism for rapid degradation. No other examples for antizyme-mediated degradation of ODC in parasitic protozoa have been described (Hanson, *et al.*, 1992; Müller, *et al.*, 2001; Müller, *et al.*, 2000; Phillips, *et al.*, 1987), including *Crithidia fasciculata* in which ODC has a rapid turnover rate (Svensson, *et al.*, 1997). Recently it was shown that ODC in this trypanosomatid is regulated via the normal ubiquitin-mediated 26S proteasome pathway (L. Persson, Personal Communication). It is therefore possible that this is a general mechanism for regulation of ODC in parasitic protozoa.

The AdoMetDC domain shows a preference for α-helices compared to two distinct structural regions in the ODC domain, a mixed α-helix/β-sheet domain close to the hinge region and a β-sheet rich domain closer to the C-terminus (Fig. 3.13). Accordingly, a mixed topology is observed for the two domains, the AdoMetDC domain can be classified into the α+β class of proteins (containing both helices and sheets in separate parts) whilst the ODC domain is a α/β class protein (proteins in which helices and sheets interact) (Sternberg, 1996). The hinge region and other parasite-specific regions are predicted to contain unstructured areas and is further discussed in Chapter 4. The bifunctional protein has an overall hydrophilic nature typical of globular proteins (Fig. 3.14).

As mentioned, the bifunctional PfAdoMetDC/ODC is grouped into protein families corresponding to the two individual decarboxylase functions. However, the bifunctional

arrangement impedes phylogenetic relationship predictions. Separate comparison of the relationship of the two decarboxylase domains did not resolve the relationships between ODC and AdoMetDCs from various organisms (Results not shown). The fact that both these proteins are predicted to belong to two different protein families could indicate that these proteins were derived from different ancestors. The bifunctional origin of the proteins might best be explained by looking at the evolution of other modular proteins, proteins containing different distinct domains due to a gene elongation or gene fusion event. The most probable mechanism for these large gene insertions to occur is through exon-shuffling. Clear-cut examples of evolution by means of exon-shuffling in *P. falciparum* come from the thrombospondin-homologue protein and a surface protein containing four epidermal growth factor-like domains (Patthy, 1999).

Plasmodium species all belong to the phylum Apicomplexa, which forms a monophyletic clade (Ayala, *et al.*, 1998). A more accurate prediction of the relationship of the bifunctional PfAdoMetDC/ODC to other organisms would therefore be to compare it to proteins of other Apicomplexa. Interestingly, bifunctional forms of the protein were discovered in two other *Plasmodia*, the murine species *P. berghei* and *P. yoelli*. Analyses of the phylogenetic relationship including the other bifunctional AdoMetDC/ODC proteins from *P. yoelii* and *P. berghei* resulted in the *Plasmodium* proteins forming a single clade removed from all the other organisms (results not shown). More than half of the *P. falciparum* AdoMetDC/ODC sequence is conserved compared with the murine sequences, which have almost identical sequences (91%). It has been shown that *P. falciparum* is more closely related to *P. reichenowi*, the chimpanzee parasite, than to any other *Plasmodium* species (Ayala, *et al.*, 1998) based on analyses of small-subunit rRNA and circumsporozoite genes. All the human malaria species are very remotely related to each other but all cluster with the primate parasites (Ayala, *et al.*, 1998). The rodent parasites do not cluster with any of the other clades and explains the low sequence identity (~41%) seen between these species and the *P. falciparum* AdoMetDC/ODC. However, the presence of the bifunctional form of AdoMetDC and ODC in other *Plasmodia*, but not in any other organism including protozoa, indicates that it is a unique property of *Plasmodia*. Therefore, to have remained, the bifunctional origin is not due to a neutral mutation but must have an evolutionary advantage since it is dispersed throughout the genus *Plasmodium*. Reasons might include adaptation to the various host environments, advantages in the regulation

of polyamine levels in these rapidly growing organisms as well as increased protein stability.

The recombinant expression and sequence analyses of the bifunctional PfAdoMetDC/ODC highlighted several unique properties of this protein. One of the most interesting is the presence of the large, parasite-specific inserts present in the protein. Chapter 4 address the structure and functional properties of these inserts.

CHAPTER 4

Functional and structural roles of parasite-specific inserts in the bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase.

4.1) INTRODUCTION.

Numerous *P. falciparum* proteins are characterised by an increased protein size relative to homologues from other organisms (Bowman, *et al.*, 1999; Gardner, *et al.*, 1998). Factors contributing towards this increased protein size include the presence of unique parasite-specific regions that intersperse conserved areas of other protein homologues as well as the peculiar bifunctional organisation of some malarial proteins, with two protein activities residing on a single polypeptide. The PfAdoMetDC/ODC presents with both of these characteristics (Chapter 3).

Different sized inserts are found in malarial protein kinases (Kappes, *et al.*, 1999), HSP 90 (Bonnefoy, *et al.*, 1994), RNA polymerases (Giesecke, *et al.*, 1991), dihydrofolate reductase-thymidylate synthase (DHFR-TS) (Bzik, *et al.*, 1987), glutathione reductase (Gilberger, *et al.*, 2000), γ-glutamylcysteine synthetase (Luersen, *et al.*, 1999) and the P-Type ATPase 3 (Rozmajzl, *et al.*, 2001). The precise function and evolutionary advantage of these inserts remain unclear. Some speculations for the functions of these inserts include possible interaction sites with as yet undefined regulatory proteins in the parasite, interaction sites with host proteins and a method to evade the host immune response (Li and Baker, 1998; Schofield, 1991). Speculations on the presence of inserts in *P. falciparum* proteins include simple evolutionary divergence that may not necessarily affect the activity and/or structure of the protein. However, strong selective pressures must exist to maintain and diversify these regions (Ramasamy, 1991).

The parasite-specific areas are normally characterised by repetitive, highly charged amino acid stretches (Chapter 3, (Pizzi and Frontali, 2001). In particular, Asn- and Asp-rich areas have been characterized in antigenic regions of membrane proteins and are speculated to play a role in evasion of the host defence mechanisms by acting as

107

antigenic smokescreens (Barale, *et al.*, 1997; Kemp, *et al.*, 1987; Reeder and Brown, 1996). *P. falciparum* has various Asn-rich proteins in particular STARP (sporozoite threonine and asparagine rich protein) (Facer and Tanner, 1997), the clustered asparagine rich protein (CARP) (Wahlgren, *et al.*, 1991) and the circumsporozoite protein (PNANP repeat) (Kwiatkowski and Marsh, 1997). These repeats are present in immunodominant domains associated with antigenic proteins on the surface of the parasite. The parasite also has other proteins rich in specific amino acids including the histidine-rich protein (Kwiatkowski and Marsh, 1997), the glutamate-rich protein (Hogh, *et al.*, 1993), a histidine-alanine rich protein (Stahl, *et al.*, 1985) and a serine repeat protein (Kwiatkowski and Marsh, 1997). These sequences probably all relate in some way to the structure-activity properties of these proteins.

Bifunctional proteins are not unusual in *P. falciparum* and indeed also in other parasitic protozoa. The malaria parasite has several bifunctional enzymes including DHFR-TS (also found in *L. donovani* and *T. brucei*)(Bzik, *et al.*, 1987; Ivanetich and Santi, 1990), dihydropteroate synthetase-dihydrohydroxymethylpterin pyrophosphokinase (DHPS-PPPK)(Triglia and Cowman, 1994), glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase (Clarke, *et al.*, 2001) and guanylyl cyclase-adenylyl cyclase (Carucci, *et al.*, 2000). Various speculations have been put forward to explain the bifunctional nature of these proteins. In the case of DHFR-TS, the two proteins catalyse consecutive reactions in the same metabolic pathway and substrate channelling has been proposed to optimise formation of products without further regulatory processes involved (Ivanetich and Santi, 1990). Other possible explanations for the bifunctional arrangements include coordinated regulation of protein concentrations/activities and intramolecular communication and interaction (Müller, *et al.*, 2000).

Obvious questions arise as to the importance of the parasite-specific inserts in the activity and/or structure of the bifunctional PfAdoMetDC/ODC. Much is known about certain key residues in PfAdoMetDC/ODC from mutagenesis results (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). Point mutations in one domain of the complex do not influence the activity of the other. Similarly, inhibition of one domain with a specific inhibitor has a singular effect on that domain. It therefore seems that the individual decarboxylase activities can function independently from each other (Wrenger, *et al.*, 2001). However, certain protein-protein interactions are expected in order to stabilize the bifunctional complex. One possible role for the inserted amino

acids and/or hinge region could be to mediate these protein-protein interactions. Clarification of the possible functions of these parasite-specific areas could contribute towards understanding the properties of PfAdoMetDC/ODC and exploitation of this knowledge in the design of selective inhibitors for antimalarial chemotherapy.

One of the most powerful developments in molecular biology has been the ability to create defined mutations in a gene and to analyse these effects on the activities of *in vitro* expressed mutated proteins. Mutants are essential in understanding the structure-function relationships of proteins and aid the rational design of proteins and their inhibitors (Lodish, *et al.*, 1995). Numerous methods are available for site-directed *in vitro* mutagenesis of genes, collectively termed protein engineering (Old and Primrose, 1994; Wilson and Walker, 2000). Mutations are designed to alter a particular codon or stretch of codons, which after translation give rise to different amino acids that may influence the properties of the protein. Cassette-mutagenesis results in the replacement of a particular DNA sequence with a synthetic DNA fragment containing the desired mutation with almost 100% efficiency (Old and Primrose, 1994). The disadvantage of this technique is the requirement of unique restriction sites flanking the region of interest. Oligonucleotide or single-primer mutagenesis (primer-extension) requires single stranded DNA (insert cloned in e.g. M13 phage vectors) from which DNA synthesis is primed with the mutant oligonucleotide. Subsequent cloning of the products produces multiple copies, half of which are mutants and half wild type, necessitating large-scale screening procedures (Old and Primrose, 1994). PCR mutagenesis relies on the incorporation of mismatches in the PCR primers into the amplified product (Wilson and Walker, 2000). PCR mutagenic methods include techniques termed overlap-extension PCR (where two primary PCRs produce two overlapping fragments that are subsequently amplified) and megaprimer PCR mutagenesis (where the products of the primary PCR are allowed to act as primers in the subsequent amplifications). Various modifications of these methods allow rapid mutagenesis at almost 100% efficiency (Old and Primrose, 1994).

This chapter describes results of studies aimed at elucidation of the role of parasite-specific inserts in interactions between the AdoMetDC and ODC domains in the bifunctional PfAdoMetDC/ODC enzyme. Our strategy was to utilise site-directed *in vitro* mutagenesis methods to gauge their effects on complex formation and activities of the two domains.

Some of the results obtained in this Chapter have been submitted for publication in the Biochemical Journal (Birkholtz, *et al.*, 2002b).

4.2) MATERIALS AND METHODS.

4.2.1) Amino acid sequence and structural analyses.

Amino acid sequence alignments were performed with Clustal W (Thompson, *et al.*, 1994) using the default parameters for PfAdoMetDC/ODC (Genbank Accession Number AF094833) and the corresponding enzymes from the human, mouse, *L. donovani* and *T. brucei*. Genbank accession numbers for AdoMetDCs: human: M21154, murine: D12780, *T. brucei:* U20092, *L. donovani:* LDU20091 and for ODCs: human: M31061, murine: J03733, *T. brucei:* J02771 and *L. donovani:* M81192. Bifunctional AdoMetDC/ODC was also identified in other *Plasmodium* species as described in Chapter 3 and these sequences were also included in the multiple alignment. Secondary structure predictions, antigenic profiles and Kyte and Doolittle hydrophobicity plots were obtained with the PredictProtein server (Rost, 1996).

4.2.2) Deletion mutagenesis (Kunkel, 1985).

Deletion mutants were created for all the major inserts present in both the PfAdoMetDC and PfODC domains, as well as for the hinge region connecting these domains in the bifunctional enzyme. Mutagenesis was based on the principle described in the QuikChange™ Site-Directed Mutagenesis Kit by Stratagene (La Jolla, California, USA). Briefly, PCR is used to introduce site-specific mutations to any double-stranded supercoiled plasmid containing the insert of interest. Two complementary mega-primers with the desired mutations are used to create mutated plasmids with staggered nicks after linear amplification. *Pfu* DNA polymerase from *Pyrococcus furiosis* is used to replicate both plasmid strands with high fidelity using its 3'-5' proofreading exonuclease activity without displacing the mutant primers. The product is then treated with *Dpn*I (target sequence: 5'-Gm^{6}ATC-3') in order to remove the methylated parental DNA template. The PCR-generated mutated plasmid is then transformed into competent *E. coli* cells where the bacterial ligase system repairs the nicks to create double stranded plasmids. This technique combines the principle of oligonucleotide mutagenesis with

110

PCR-based techniques to obtain a >80% efficiency in mutagenesis of any insert in any vector system.

Oligonucleotides used for the site-directed deletion mutagenesis are indicated in Table 4.1. A typical deletion mutagenesis reaction (50 μl final volume) contained 10 ng of the wild-type expression plasmids (as isolated in Chapter 3, section 3.2.2) with the specific inserts (pASK-IBA3 for bifunctional PfAdoMetDC/ODC; pASK-IBA7 with either the PfAdoMetDC or PfODC domains; Institut für Bioanalytik, Göttingen, Germany), 150 ng of both the mutagenic sense and antisense mega-primers (Table 4.1), 1x *Pfu* DNA polymerase reaction buffer, 2.5 mM of each dNTP and 3 U *Pfu* DNA polymerase (Promega, Wisconsin, USA). The cycling parameters were 95°C for 50 sec, 55°C for 1 min and 68°C for 12 min (bifunctional construct) or 9 min (separate domains) repeated for 18 cycles in total after an initial denaturation step of 95°C for 3 min in a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, California, USA). After 9 cycles, a further 1 U *Pfu* DNA polymerase was added to amplifications of the bifunctional constructs. The PCR products were subsequently treated with 20 U *Dpn*I (New England Biolabs, Massachusetts, USA) for 3 hours at 37°C followed by removal of the digested parental DNA templates using the standard protocols described in the High Pure PCR Product Purification Kit (Roche, Mannheim, Germany). The pure mutated constructs containing nicks were ligated at 4°C for 16 hours with 6 U T4 DNA ligase (Promega, Wisconsin, USA) to increase the transformation efficiency. The double stranded supercoiled plasmids were subsequently transformed into competent DH5α *E. coli* cells (Invitrogen, Paisley, UK) as described in Chapter 2, section 2.2.10.

4.2.3) Nucleotide sequencing of the various mutants.

The nucleotide sequences of the mutant cloned fragments were determined by automated nucleotide sequencing as described in Chapter 2, section 2.2.11. For cycle sequencing of mutant clones, primers complementary to the PfAdoMetDC/ODC nucleotide sequence was used at a site not more than 300 nucleotides removed from the mutation site. See Chapter 2, Fig. 2.3 for primer locations.

4.2.4) Recombinant expression and purification of wild type and mutant proteins.

The *P. falciparum* monofunctional AdoMetDC and ODC and bifunctional AdoMetDC/ODC were expressed as Strep-Tag fusion proteins as described in Chapter 3, section 3.2.2 (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). Mutant forms of PfAdoMetDC/ODC with individual deletion of the parasite-specific inserts, as well as single and combined insert deletion mutants in the monofunctional PfAdoMetDC and PfODC domains were isolated as for the wild type proteins. The monofunctional PfODC domain lacking the N-terminal hinge region that connects it to PfAdoMetDC was cloned into the expression plasmid pJC40 for expression as a fusion protein with an N-terminal His$_6$-tag (Krause, *et al.*, 2000). The entire coding region of the *P. falciparum* spermidine synthase was also cloned in the same plasmid (Haider *et al*, personal communication). These proteins were isolated as described in Krause *et al.* 2000. The concentrations of the purified proteins were determined with Coomassie brilliant blue G-250 (Pierce, Illinois, USA) as described by Bradford (Bradford, 1976). Purified proteins were analysed by SDS-PAGE and visualised with silver staining as described in Chapter 3, section 3.2.6.

4.2.5) Protein-protein interaction determinations.

4.2.5.1) Size-exclusion fast protein liquid chromatography (SE-FPLC) of the interacted proteins.

Protein-protein interactions were determined by SE-FPLC of wild type or mutant hinge-linked bifunctional proteins and combinations of individually expressed wild type or mutant monofunctional PfAdoMetDC and PfODC. Wild type and mutant forms of PfAdoMetDC/ODC were analysed for their ability to form heterotetrameric complexes (~330 kDa) or uncomplexed heterodimer subunits (~160 kDa). Combinations of wild type and mutant monofunctional PfAdoMetDC and PfODC were analysed for their ability to associate and form hybrid heterotetrameric complexes (~330 kDa), to remain in their monofunctional active states (heterotetrameric PfAdoMetDC of ~ 145 kDa and homodimeric PfODC of ~166 kDa) or heterodimeric ~ 64 kDa AdoMetDC and ~80 kDa monomeric ODC. Intermolecular protein-protein interactions between bifunctional PfAdoMetDC/ODC and spermidine synthase were also analysed. Separately expressed and isolated proteins were allowed to interact by co-incubation for 10 min at room temperature. Subsequently, the protein complexes were subjected to FPLC as described in Chapter 3, section 3.2.4. Protein was detected in the collected fractions with

112

Coomassie brilliant blue G-250 (Pierce, Illinois, USA), dot-blot western immunodetection or enzyme activity determinations as described in the next sections.

4.2.5.2) Dot-blot Western analyses of SE-FPLC fractions.

The collected size-exclusion chromatography fractions were transferred to nitrocellulose membranes using a BioDot apparatus (Bio-Rad) and analysed by dot blot Western. The membranes were blocked in 3% w/v low fat milk powder in 1xPBS for 16 hours at 4°C followed by incubation with a 1:4000 dilution of polyclonal Strep-tag II rabbit antiserum raised against the Strep-tag II peptide conjugated to keyhole limpet hemocyanin (Institut für Bioanalytik, Göttingen, Germany) for 1 hour at room temperature. After three washes with 0.05% Tween-20 (Merck, Germany) in 1xPBS, the membrane was incubated with a 1:2000 dilution of horseradish peroxidase (HRP) conjugated anti-rabbit donkey whole IgG (Amersham Pharmacia Biotech, UK) for 1 hour at room temperature in 1% w/v low fat milk powder in 1xPBS. The membrane was again washed three times in 0.05% Tween-20 in 1xPBS. The proteins were visualized with the ECL Plus™ Western Blotting system (Amersham Pharmacia Biotech, UK) using chemiluminescence according to the manufacturers recommendations. The detection reaction is based on the generation of an acridium ester by the enzymatic action of HRP on Lumigen PS-3 acridan substrates. The esters react with peroxide under slightly alkaline conditions to produce a high-intensity chemiluminescence with emission wavelength of 430 nm. The washed membranes were incubated in the chemiluminescent reagents for 5 min at room temperature in the dark. Excess reagents were drained off and the membrane wrapped in plastic. The membrane was placed in a X-ray film cassette and a sheet of Kodak Biomax autoradiography film (Kodak) was placed on top of the membrane and exposed from 15 sec to 30 min. The film was developed as described in Chapter 2, section 2.2.12.2.

4.2.6) Enzyme assays.

Wild type and mutant forms of the bifunctional PfAdoMetDC/ODC and monofunctional PfAdoMetDC and PfODC activities were determined as described in Chapter 3, section 3.2.7. Spermidine synthase activity was determined as described in Haider *et al.* (Personal communication). Results are the mean of three independent experiments performed in duplicate and expressed as a percentage of the normalised wild type controls.

4.3) RESULTS.

4.3.1) Explanations for the bifunctional nature of the PfAdoMetDC/ODC.

One explanation for the bifunctional nature of PfAdoMetDC/ODC is to allow substrate channelling to occur. For this to be true another enzyme, spermidine synthase, is required to use the decarboxylated products of PfAdoMetDC/ODC as substrate to produce spermidine. PfAdoMetDC/ODC was isolated and allowed to interact with separately expressed spermidine synthase. After co-incubation of the separately isolated enzymes for 30 min at 4°C, the proteins were analysed by SE-FPLC, followed by SDS-PAGE and activity analyses of various fractions. Fig. 4.1 indicates the size-exclusion elution profile of the interacting PfAdoMetDC/ODC and spermidine synthase. PfAdoMetDC/ODC protein and activity was observed at ~ 330 kDa, the size of the wild-type bifunctional protein. Spermidine synthase protein and activity did not co-elute with the decarboxylase activities but eluted at the expected size of ~75 kDa for the active dimeric form of spermidine synthase (Fig 4.1 B). None of the protein activities were present in fractions corresponding to a complex between PfAdoMetDC/ODC and spermidine synthase of ~404 kDa. It therefore seems that no interactions occur between PfAdoMetDC/ODC and spermidine synthase under the *in vitro* conditions used or that interactions are transient and not stable enough to survive size exclusion chromatography.
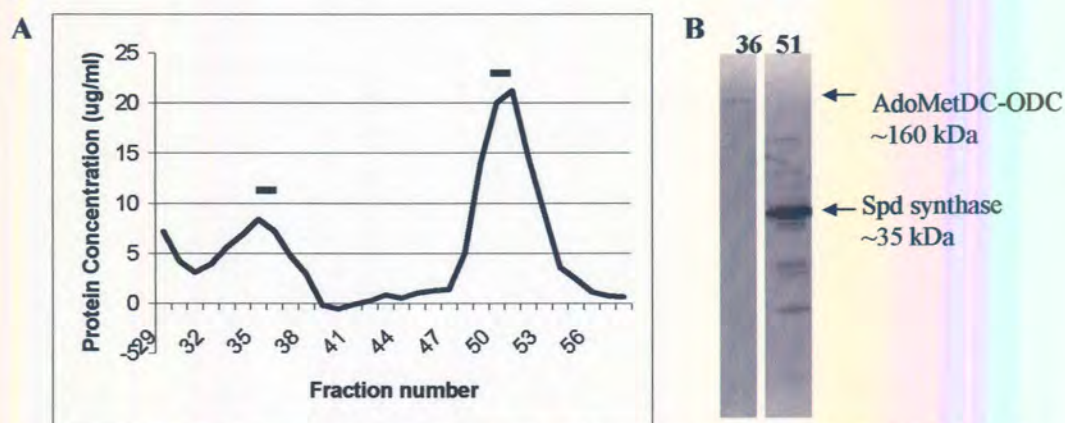
Figure 4.1: Interaction assay between the wild type bifunctional PfAdoMetDC/ODC and spermidine synthase. (A) Size exclusion elution profile of interacting PfAdoMetDC/ODC and spermidine synthase with the corresponding activities indicated in horizontal bars; PfAdoMetDC/ODC activity in fraction 36 and spermidine activity in fraction 51. **(B)** SDS-PAGE analyses of size exclusion fractions 36 and 51 corresponding to denatured protein sizes of ~160 kDa and ~35 kDa for the bifunctional, heterodimeric PfAdoMetDC/ODC and monomeric spermidine synthase, respectively.

114

4.3.2) Parasite-specific regions in PfAdoMetDC/ODC.

The parasite-specific inserts in the bifunctional PfAdoMetDC/ODC were defined based on a multiple-alignment of the PfAdoMetDC/ODC sequence with the corresponding sequences for the individual enzymes as described in Chapter 3 (Fig. 3.12). Briefly, the insert in the AdoMetDC domain includes residues 241-410 (insert A_1) and the ODC domain inserts include residues 1047-1085 (O_1) and 1156-1301 (O_2). The hinge region was defined by Müller *et al.* (2000) as residues 573-752 (H) (Fig. 4.2).

The identification of bifunctional AdoMetDC/ODC in other *Plasmodia* described in Chapter 3, provides the opportunity to better define the parasite-specific inserts. The large inserts in both domains, A_1 and O_2, show large variations in sequence composition and length between the three *Plasmodium* species (Fig. 4.2). The AdoMetDC domains of the murine parasite enzymes are ~100 residues longer than the *P. falciparum* enzyme. In contrast, the ODC domain is longer in *P. falciparum* due to a ~26 residue longer insert O_2. The amino acid compositions of both the large inserts (A_1 and O_2) seem to be conserved between the murine sequences but differ from *P. falciparum*, specifically in the distribution of Asn and $(NND)_x$-repeats in the *P. falciparum* sequence. The hinge region is also smaller in the murine bifunctional enzymes. One exception to the abovementioned characteristics is the smallest insert O_1 in the ODC domain, which is better conserved between the *Plasmodia* species both in terms of sequence composition and length. This suggests that this area might have a more defined function compared to the other larger, more variable inserts.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure 4.2: Multiple-alignment of the amino acid sequences of the bifunctional PfAdoMetDC/ODC indicating the parasite-specific areas. The putative PfAdoMetDC, hinge and PfODC domains are indicated. Amino acids shown in black boxes are >80% conserved and >60% conserved residues are shown in grey boxes. The parasite-specific inserts are in italics in the *P. falciparum* sequence (A_1: 214-410; H: 573-752; O_1: 1047-1085 and O_2: 1156-1301). Horizontal bars indicate low-complexity areas in PfAdoMetDC/ODC.

4.3.3) Sequence and structure analyses of the parasite-specific regions.

The defined parasite-specific areas of specifically the bifunctional AdoMetDC/ODC of *P. falciparum* were analysed for various sequence and structural properties to obtain information on their possible functions. The parasite-specific areas in PfAdoMetDC/ODC are rich in charged residues, predominantly Asn, Asp, Lys, Ser, Glu, Leu, and Ile (Fig. 4.2). Noticeably, both the large inserts (A_1 and O_2) and the hinge region in PfAdoMetDC/ODC are Asn-rich with insert A_1 containing 16%, the hinge region 14.8% and insert O_2, 28.7%. However, the smallest insert in the PfODC domain (O_1) is composed of 20.5% Lys residues. The inserts in PfAdoMetDC/ODC are also characterised by highly recurrent repeats of (NND)x-motifs and long stretches of Asn.

Further attempts to investigate the potential function of these parasite-specific inserts included secondary structure predictions. Kyte and Doolittle hydrophobicity analyses of the parasite-specific inserts indicated that all the inserts have a more pronounced hydrophilic nature (Fig. 4.3). Analyses of the antigenic properties of the parasite-specific regions predict several but not highly significant antigenic regions (Fig. 4.3) (Hopp-Woods equation, (Geourjon, *et al.*, 1991). The Wootton and Federhen algorithm (SEG algorithm, (Wootton and Federhen, 1996) predicted low-complexity areas in all the inserts and the hinge region except insert O_1 in the PfODC domain (Fig. 4.2 and 4.3). Furthermore, these inserts are predicted to be nonglobular with a tendency towards unstructured loops connected in the majority of cases with β-sheets (Fig. 4.3). Insert O_1 is the only area proposed to contain significant secondary structure consisting of four β-sheets arranged in an anti-parallel manner as indicated in a homology model of the PfODC domain (See Chapter 5). Thus, insert O_1 appears to be more structured without antigenic properties or low-complexity regions compared to the other parasite-specific inserts indicative of a specific role in PfAdoMetDC/ODC.
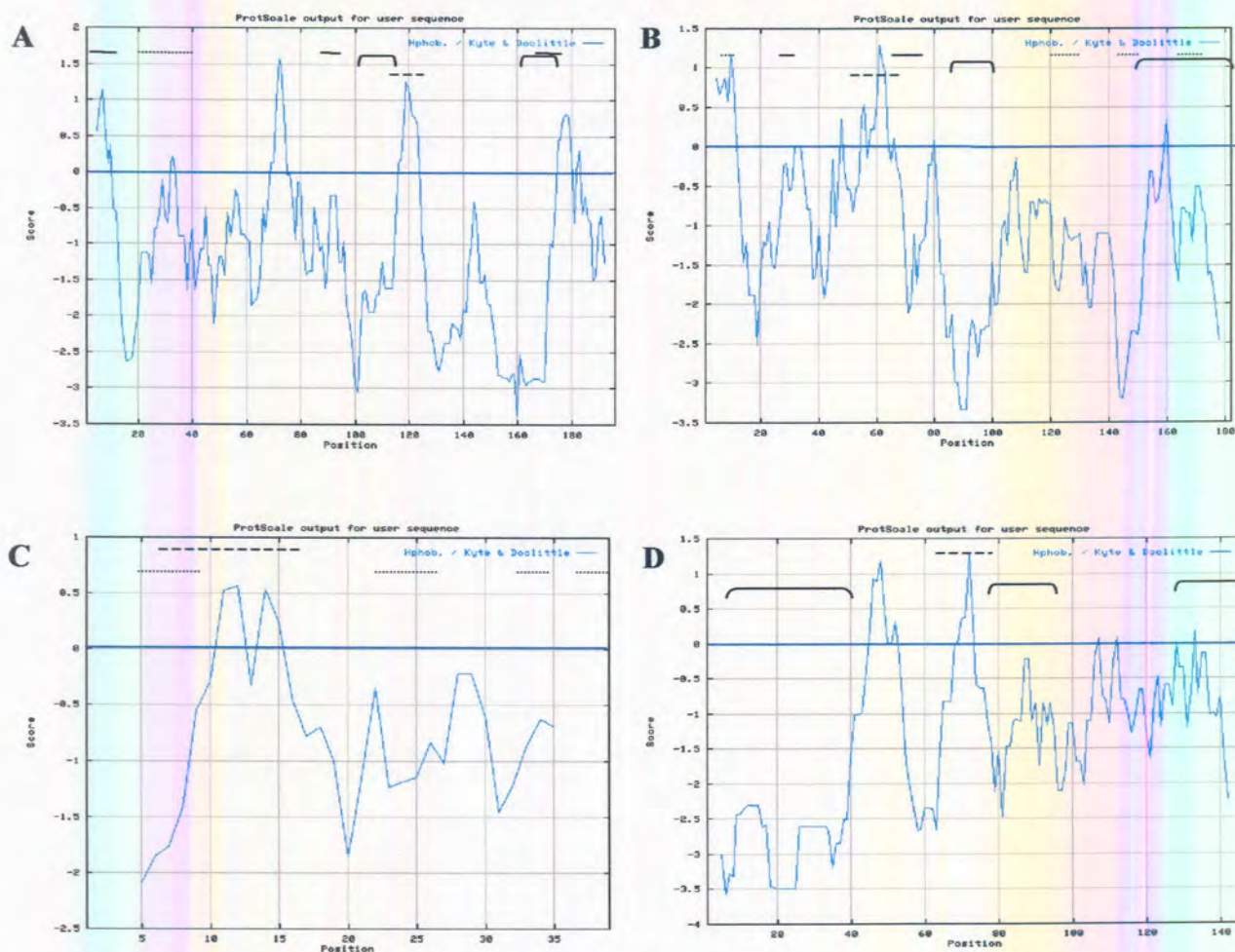
Figure 4.3: Sequence and secondary structure analyses of the parasite-specific inserts in the bifunctional PfAdoMetDC/ODC. (A) AdoMetDC insert A_1; **(B)** Hinge region; **(C)** ODC insert O_1 and **(D)** ODC insert O_2. Kyte and Doolittle hydropathy plots are shown for all four parasite-specific regions, hydrophilic residues are scored negatively. Low complexity areas are indicated with horizontal brackets. Secondary structures are predicted as α-helices (—) and β-sheets (·····) with the rest of the areas predicted to be random loops. Highest scoring antigenic regions are indicated with – – .

4.3.4) Deletion mutagenesis of parasite-specific regions in PfAdoMetDC/ODC.

To investigate what functions, if any, the parasite-specific inserts have on the activities and interactions of the bifunctional PfAdoMetDC/ODC, deletion-mutants were created for all three identified inserts as well as for the hinge region connecting the two decarboxylase activities. The mutagenic primers used to create the deletion mutants are summarised in Table 4.1.

118

Table 4.1: Mutagenic mega-primer oligonucleotides used for deletion mutagenesis of parasite-specific regions in PfAdoMetDC/ODC. Sites where deletion occurred are indicated with -. The sizes of the deletions (in nucleotides) are indicated.

Primer	Sequence 5'-3'	Length of primer	Tm	Number of nucleotides deleted
ΔA_1 sense	GAC GGA TAT AGC TTC TAC GTT T-AA TGA ATT TTA TTT TAC ACC TTG TGG	48	69.4	591
ΔA_1 antisense	CCA CAA GGT GTA AAA TAA AAT TCA TT-A AAC GTA GAA GCT ATA TCC GTC	48	69.4	591
ΔH sense	GTG TAG AAA AAG AAA CTT TG-G AAA AAA TGA AAG ATT ATA TAA GTG	45	54.6	540
ΔH antisense	CAC TTA TAT AAT CTT TCA TTT TTT C-CA AAG TTT CTT TTT CTA CAC	45	54.6	540
ΔO_1 sense	GGA GGG GGA TAT CCA GAA GAA TTA GAA TAT GAT-AGT TTT GAA AAA ATA TCA TTG GC	56	57.8	117
ΔO_1 antisense	GC CAA TGA TAT TTT TTC AAA ACT-ATC ATA TTC TAA TTC TTC TGG ATA TCC CCC TCC	56	57.8	117
ΔO_2 sense	GAC CAT TAC GAT CCT TTA AAT TTT T-TC TCA TAT TAT GTA AGC GAT AGT ATA TAT GG	56	69.4	435
ΔO_2 antisense	CCA TAT ATA CTA TCG CTT ACA TAA TAT GAG AA-A AAA TTT AAA GGA TCG TAA TGG TC	56	69.4	435

Each of the parasite-specific areas was individually removed from the bifunctional enzyme to create four deletion-mutants by using standard mega-primer deletion PCR techniques (Mutants $A-O_\Delta A_1$, $A-O_\Delta H$, $A-O_\Delta O_1$, $A-O_\Delta O_2$, Fig. 4.4).
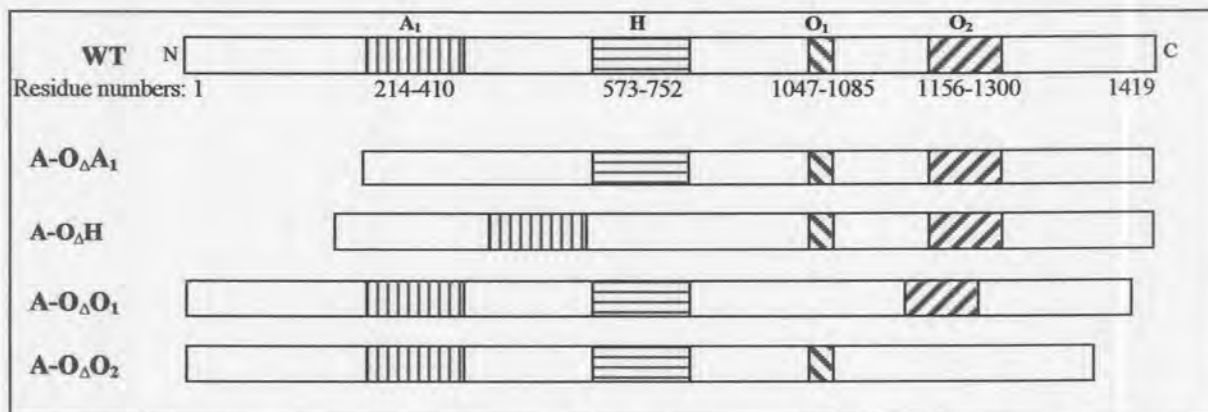
Figure 4.4: Schematic representation of the strategy used for deletion of the parasite-specific inserts and hinge region in the bifunctional PfAdoMetDC/ODC. Wild-type PfAdoMetDC/ODC is shown (top) with the positions and residue numbers of the specific inserts and the various deletion mutants indicated.

The effects of deletions on the expression of the different mutant proteins were determined by expression as for the wild-type protein followed by their isolation and analyses on SDS-PAGE (Fig. 4.5). The expressed mutant proteins had the expected decreased molecular mass. Wild-type bifunctional PfAdoMetDC/ODC migrates at the subunit size of ~160 kDa under denaturing conditions whereas mutant $A-O_\Delta A_1$ migrates at ~138 kDa corresponding to the expected size with the 21 kDa insert removed, $A-O_\Delta H$ at ~141 kDa (19.8 kDa removed), $A-O_\Delta O_1$ at ~156 kDa (4.2 kDa removed) and $A-O_\Delta O_2$ at ~144 kDa (15.9 kDa removed). The deletions did not have a major influence on the

levels of protein obtained except for the hinge deletion mutant, where expression levels dropped by 40%. This indicates that the proteins probably retained their conformations and were not expressed as misfolded proteins present in inclusion bodies in the *E. coli*.
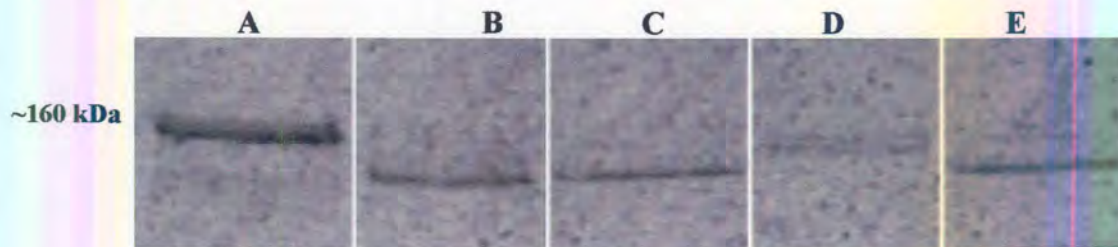
Figure 4.5: SDS-PAGE analysis of the wild-type PfAdoMetDC/ODC and the individual deletion mutants. Wild-type PfAdoMetDC/ODC (A) is compared to deletion mutants of the parasite-specific inserts (B: $A-O_\Delta A_1$; D: $A-O_\Delta O_1$ and E: $A-O_\Delta O_2$) and the hinge region (C: $A-O_\Delta H$). Proteins were revealed with silver staining.

4.3.5) Effect of deletion mutagenesis on the decarboxylase activities of the bifunctional PfAdoMetDC/ODC.

The effects of the deleted parasite-specific inserts on the decarboxylase activities were assayed in all the mutants (Fig. 4.6). AdoMetDC and ODC activities in the bifunctional enzyme are markedly reduced (>95 %) when the deletion of the specific insert occurs inside the respective domain. Interestingly, deletion of the A_1 and O_1 inserts, both closer in linear amino acid sequence to the neighbouring domain and hinge region, also influences the activity of the neighbouring domain. For instance, deletion of O_1 reduces ODC activity by 98% (1.85 % residual activity) but also decreases AdoMetDC activity by 90% (Fig. 4.6). Deletion of A_1 reduces AdoMetDC activity by 96% and ODC activity by 72%. The effect on AdoMetDC activity is not as pronounced in the O_2 deletion mutant of ODC since AdoMetDC activity was only decreased by 47%. However, this mutant shows a 98% reduction in ODC activity. Thus, deletion of the smallest O_1 insert had the most significant effect on both enzyme activities.
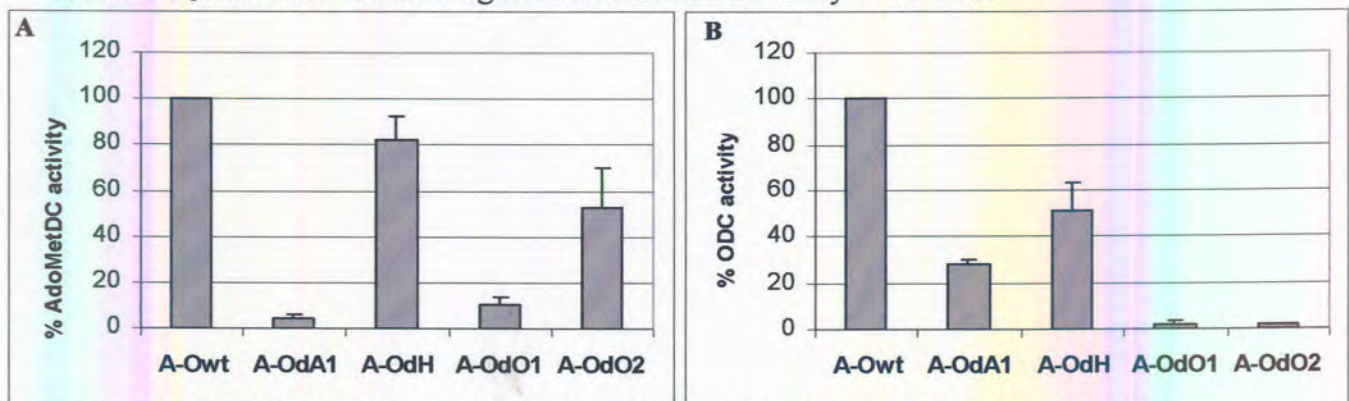
Figure 4.6: Activity analyses of wild type and mutated bifunctional PfAdoMetDC/ODC. (A) AdoMetDC activities and (B) ODC activities. Results are the mean of three independent experiments performed in duplicate and specific activities are represented as a percentage of the wild type activity with standard deviations indicated.

Deletion of the presumably flexible hinge region connecting the two domains should force the two domains into close physical proximity of each other. The PfAdoMetDC activity is only slightly decreased (18%) in this mutant but half of the PfODC activity is lost, indicating a more pronounced dependency of the ODC domain on the hinge region (Fig. 4.6). Interestingly, expression of the monofunctional PfODC domain without the hinge region also leads to a marked decrease in specific activity of the protein (Krause, et al., 2000). The results presented here indicate that the parasite-specific inserts mediate specific physical interactions between the two domains that are ultimately reflected in a decrease of both decarboxylase activities.

4.3.6) Deletion mutagenesis of the parasite-specific regions in the monofunctional PfAdoMetDC and PfODC domains.

As described in Chapter 3, the individual monofunctional PfAdoMetDC and PfODC domains can be stably expressed as a heterotetrameric protein of ~145 kDa (PfAdoMetDC) and obligate homodimeric PfODC of ~170 kDa (Krause, et al., 2000; Wrenger, et al., 2001). The direct contribution of the parasite-specific inserts to the activities of the individually expressed monofunctional domains was investigated with single and combined deletion mutants of all the inserts in the separate domains (Fig. 4.7). The ODC domain without the hinge region was previously analysed as described in Krause et al. (2000). This mutant was shown to be 50% less active with a 3.4 times decreased Km for ornithine (Krause, et al., 2000).
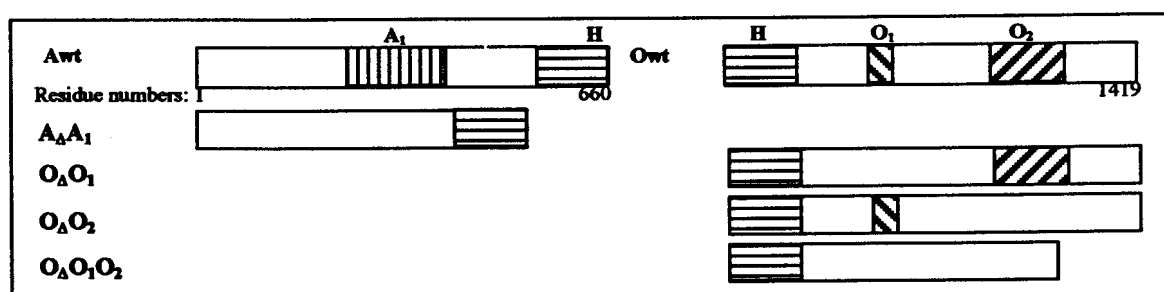
Figure 4.7: Schematic representation of the deletion mutagenesis strategy of the parasite-specific inserts in the monofunctional PfAdoMetDC and PfODC. Wild-type monofunctional PfAdoMetDC and PfODC is shown (top) with the positions of the specific inserts (A_1, H, O_1 and O_2) and the description of various deletion mutants at the bottom. Residue numbering as for the bifunctional enzyme complex.

As expected, all deletion mutants of the parasite-specific inserts did not have significant residual decarboxylase activity (Fig. 4.8). Mutant $A_\Delta A_1$ resulted in 88% decrease of AdoMetDC activity compared to 95% decrease of mutant $A-O_\Delta A_1$ in

PfAdoMetDC/ODC (Fig. 4.6). Even more pronounced activity loss was evident in the ODC domain deletion mutants: $O_\Delta O_1$ decreased ODC activity by 91%, whereas the ODC activities of $O_\Delta O_2$ and the double-deletion mutant $O_\Delta O_1 O_2$ were reduced by 97% (Fig. 4.8). Deletion of insert A_1 in the monofunctional PfAdoMetDC domain, and double deletion of both the O_1 and O_2 inserts in the monofunctional PfODC domain results in an amino acid sequence and length that closely resemble homologues for these proteins in other organisms (See also Chapter 5). The inactivity of these mutants ($A_\Delta A_1$, and $O_\Delta O_1 O_2$) implies that these areas have parasite-specific properties that are required for enzyme activity.
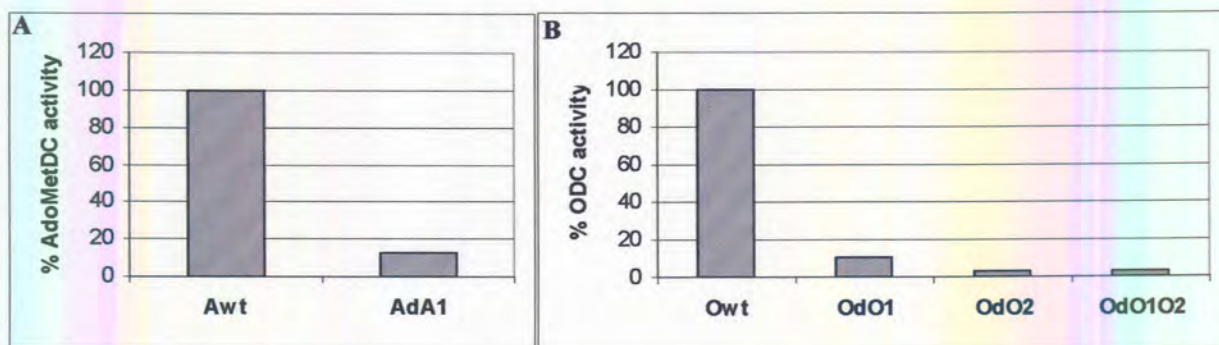
Figure 4.8: Specific activities of deletion mutants of the individual monofunctional PfAdoMetDC and PfODC domains. (A) Wild-type and deletion mutant AdoMetDC and in **(B)** Wild-type and deletion mutants activities of the ODC domain. Results are representative of duplicate experiments and specific activities are represented as a percentage of the wild type activity.

4.3.7) Oligomeric state of deletion mutants of PfAdoMetDC/ODC.

The ability of the deletion mutants to still form the heterotetrameric bifunctional complex was also investigated. The deletion mutants of the bifunctional enzyme were isolated by affinity chromatography, combined and then subjected to SE-FPLC to distinguish between heterotetrameric and heterodimeric states.

The isolated wild-type PfAdoMetDC/ODC and deletion mutants fractions were assayed for the presence of the protein (Coomassie detection and dot-blot Western) as well as for decarboxylase activity. Wild-type PfAdoMetDC/ODC elutes at a molecular mass of ~ 330 kDa (fraction 35) corresponding to a heterotetrameric complex of the AdoMetDC/ODC polypeptide as described in Chapter 3, section 3.3.5.2, Fig. 3.10 (A). Both AdoMetDC and ODC activities were observed in fractions 34-36.

Inability of the mutants to assemble in a bifunctional complex would result in the ~160 kDa heterodimeric form of the polypeptide (Fig. 4.9). Deletion of the parasite-specific

inserts in the bifunctional PfAdoMetDC/ODC did not alter the complex forming ability of the enzymes. Mutants A-O$_\Delta$A$_1$, A-O$_\Delta$O$_1$ and A-O$_\Delta$O$_2$ were still able to form heterotetrameric complexes with activities present only in the fractions corresponding to their predicted sizes of 276 kDa, 282 kDa and 312 kDa, respectively. The same is true for the hinge deletion mutant that also eluted around the expected smaller size of 288 kDa.
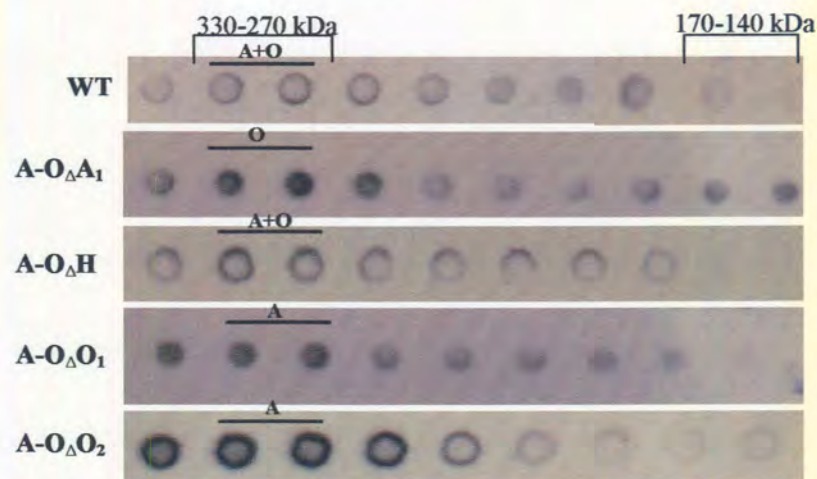
Figure 4.9: Complex forming abilities of deletion mutants of PfAdoMetDC/ODC. Protein was detected in the size-exclusion fractions with Bradford and a dot-blot Western analysis. Presence of protein and AdoMetDC and ODC activities in fractions corresponding to the heterotetrameric bifunctional complex size of 330 kDa and the uncomplexed heterodimeric form of 160 kDa is indicated. AdoMetDC and ODC activities are indicated by horizontal bars with A and O, respectively.

4.3.8) Hybrid complex forming abilities of deletion mutants of monofunctional PfAdoMetDC and PfODC.

The direct contribution of the parasite-specific inserts on hybrid bifunctional complex formation was analysed by SE-FPLC to determine the sizes of complexes after co-incubation of wild-type and mutant forms of monofunctional PfAdoMetDC and PfODC. Co-incubation of both wild-type monofunctional PfAdoMetDC and PfODC resulted in the expected formation of the ~330 kDa bifunctional heterotetrameric complex (Fig. 4.10). This is consistent with the hypothesis that the domains assembled into a heterotetramer consisting of an obligate homodimeric ODC (~170 kDa) and a heterotetrameric AdoMetDC (~145 kDa), re-establishing the natural relationship between the two domains. The new hybrid AdoMetDC + ODC complex was stable enough to survive size exclusion chromatography indicating substantial interactions between the domains.

Awt +Owt

330-270 kDa
Awt+Owt

170-140 kDa

90-70 kDa
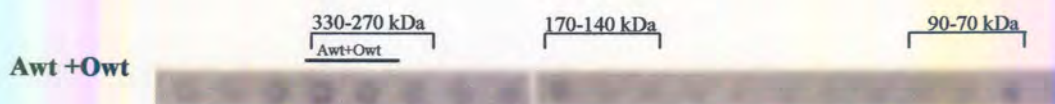
Figure 4.10: Protein-protein interactions between the separately expressed wild type AdoMetDC and ODC domains. Fractions were analysed at the different sizes corresponding to the heterotetrameric bifunctional complex size (330-270 kDa), the homodimeric size of PfODC or the heterotetrameric PfAdoMetDC (170-140 kDa) and the monomeric forms of the domains (90-70 kDa). Activities present in the different fractions are indicated in horizontal bars.

Both AdoMetDC and ODC activities were observed in the hybrid complex indicating that the association did not influence the catalytic capacity of these proteins and probably mimic the natural state of the complex (Table 4.2). Residual AdoMetDC activity was however also evident in fractions corresponding to either the heterotetrameric or heterodimeric forms of the protein.

Wild-type AdoMetDC was co-incubated with either the ODC hinge ($O_\Delta H$) or with the $O_\Delta O_1$ deletion mutants whereas wild-type ODC was co-incubated with the AdoMetDC deletion mutant $A_\Delta A_1$. Deletion of insert O_2 in the ODC domain of the bifunctional PfAdoMetDC/ODC did not seem to have a marked influence on the activity of the AdoMetDC domain and its ability to interact with AdoMetDC was not further investigated. Fig. 4.11 indicates the individual activities and dot-blot analyses of the various size exclusion fragments of the abovementioned combinations.

Awt + $O_\Delta H$

Awt + $O_\Delta O_1$

Owt + $A_\Delta A_1$

330-270 kDa

170-140 kDa
Awt

Awt

Owt

90-70 kDa
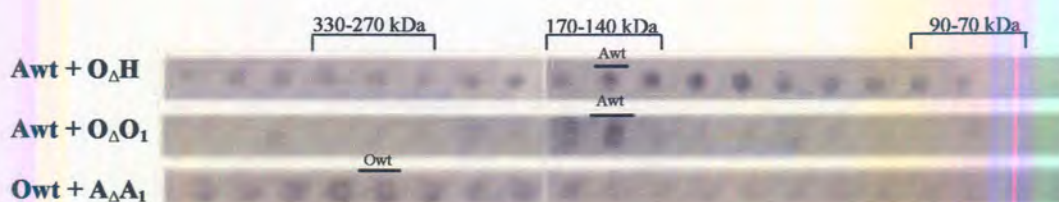
Figure 4.11: Intermolecular interaction between the wild-type and mutant forms of the monofunctional AdoMetDC and ODC. Fractions were analysed at the different sizes corresponding to the heterotetrameric bifunctional complex size (330-270 kDa), the homodimeric size of PfODC or the heterotetrameric PfAdoMetDC (170-140 kDa) and the monomeric forms of the domains (90-70 kDa). Dot blot analyses of the interaction of the wild type decarboxylase domains as well as interactions between AdoMetDCwt (Awt) and $O_\Delta H$; Awt and $O_\Delta O_1$ and ODCwt (Owt) and $A_\Delta A_1$ after size exclusion chromatography. Activities present in the different fractions are indicated in horizontal bars.

Physical association still occurred between the wild-type ODC domain and the AdoMetDC mutant $A_\Delta A_1$ to form a heterotetrameric bifunctional complex of ~330 kDa, indicating that the contribution of insert A_1 to bifunctional complex formation is

probably not that pronounced. The majority of the AdoMetDC activity after removal of the hinge region from the ODC domain (mutant $O_\Delta H$) was associated with the heterotetrameric form (~145 kDa, Table 4.2). Very little AdoMetDC activity and no ODC activity was detectable in a bifunctional complex and low levels of ODC activity was detected in fractions corresponding to the homodimeric form of the protein (~180 kDa, Table 4.2). Deletion of the much smaller O_1 insert in the ODC domain also resulted in active heterotetrameric PfAdoMetDC (~145 kDa) with no AdoMetDC activity in a bifunctional complex. Table 4.2 summarises the hybrid complex formation abilities of the mutant proteins.

Table 4.2: Hybrid complex formation abilities of mutant forms of the monofunctional AdoMetDC and ODC. Hybrid bifunctional complexes are found in the 330-270 kDa size range, monofunctional heterotetrameric AdoMetDC and homodimeric ODC in the 180-140 kDa range and monomeric proteins in the 90-70 kDa range. Results are indicated as a combination of the presence of protein (detected with the dot-blot assay) with the number of crosses indicating the relative enzyme activities. nd: Not detectable.

	330-270 kDa		180-140 kDa		90-70 kDa	
	AdoMetDC	ODC	AdoMetDC	ODC	AdoMetDC	ODC
Awt + Owt	+++	+++	++	nd	++	nd
Awt + $O_\Delta H$	+	nd	+++	+	++	nd
Awt + $O_\Delta O_1$	nd	nd	+++	nd	+	nd
Owt + $A_\Delta A_1$	nd	+++	nd	++	nd	nd

Hybrid bifunctional complex formation is only possible with the co-incubation of wild-type ODC and the AdoMetDC deletion mutant $A_\Delta A1$. No physical interactions seem to occur between the domains when the hinge region or insert O_1 is removed from the ODC domain. These areas therefore seem to enable formation of the hybrid bifunctional complex.

4.4) DISCUSSION.

4.4.1) Explanations for the bifunctional nature of the PfAdoMetDC/ODC.

Various proposals were forwarded to explain the bifunctional nature of several proteins of *P. falciparum* including substrate channelling, intramolecular communication and coordinated regulation of transcription and translation (Müller, *et al.*, 2001). In the case of DHFR-TS, substrate channelling has been proposed since the two proteins catalyse consecutive reactions in the same metabolic pathway (Miles, *et al.*, 1999). For this to be true for the bifunctional PfAdoMetDC/ODC it should interact with the subsequent

enzyme in the polyamine biosynthetic pathway, spermidine synthase, to allow the production of spermidine. Size exclusion analyses indicated that no physical association occurs between the proteins under the *in vitro* conditions used. It therefore appears unlikely that substrate channelling can be forwarded as an explanation for the bifunctional organisation of the PfAdoMetDC/ODC. More advanced techniques like the yeast two-hybrid system need to be employed to further investigate interactions with other proteins. Hypotheses advanced thus far for the bifunctional nature of PfAdoMetDC/ODC include facilitation of coordinated regulation of the protein levels and activities of both proteins or intramolecular communication and interaction.

4.4.2) Defining the parasite-specific inserts in PfAdoMetDC/ODC.

A characteristic feature of many *P. falciparum* proteins is the tendency towards large gene coding regions containing different sized inserts and increased protein sizes relative to homologues from other organisms (Bowman, *et al.*, 1999; Gardner, *et al.*, 1998). Parasite-specific inserts is usually defined with multiple amino acid sequence alignment which shows that these areas separate mutually-conserved blocks when compared to other homologous proteins (Chapter 3, Fig. 3.12 for PfAdoMetDC/ODC). Three different parasite-specific inserts were identified in PfAdoMetDC/ODC, a single large insert in the AdoMetDC domain and two inserts in the ODC domain as described in Chapter 3. The large inserts in both domains show extensive sequence composition and length variability between the different *Plasmodium* species, the AdoMetDC insert shows 18% sequence identity and the large insert in the ODC domain, O_2, shows 27% identity. However, the smallest insert in the ODC domain is more conserved (O_1, 46% identical). This suggests that the variability of this insert is constrained due to some, as yet undetermined, function.

Multiple alignment of the PfAdoMetDC/ODC amino acid sequence with homologues from other organisms as well as the bifunctional forms of the proteins in the other *Plasmodia* highlighted the inherent difficulties in predicting the correct boundaries of the hinge region between the two domains. The *L. donovani* ODC sequence has an ~200 residue extension at the N-terminus but is not part of a bifunctional protein (Hanson, *et al.*, 1992). The hinge region of the bifunctional AdoMetDC/ODC in contrast is much shorter in the sequences of the two other *Plasmodia*. Furthermore, removal of the hinge region leads to a 50% reduction in the activity of monofunctional PfODC (Krause, *et al.*, 2000). It is therefore possible that the actual hinge region in *P. falciparum* is smaller

than the currently defined 180 residues. Part of the current hinge region could therefore constitute partial sequence of ODC. This possibility is supported by the results of the deletion mutagenesis studies discussed in detail in the following sections.

4.4.3) Structural properties of the parasite-specific regions in PfAdoMetDC/ODC.

The precise function and evolutionary advantage of parasite-specific inserts in *P. falciparum* proteins are not known. Proposed functions include interaction sites with as yet undefined regulatory proteins in the parasite and sites for interaction with host proteins as a means to evade the host immune response (Li and Baker, 1998; Schofield, 1991). Analyses of the sequence and predicted structural properties of the identified parasite-specific inserts in PfAdoMetDC/ODC indicated that the inserts and the hinge region are rich in charged residues. Asp- and Asn-rich areas have been characterized in other *Plasmodial* proteins, particularly in antigenic regions of membrane proteins and as such may play a role in the diversity of the parasite population to evade the hosts defence mechanisms by acting as antigenic smokescreens (Barale, *et al.*, 1997; Kemp, *et al.*, 1987; Reeder and Brown, 1996). However, the inserts in PfAdoMetDC/ODC do not show significant antigenicity (Fig. 4.3) and it is furthermore a cytosolic protein (Müller, *et al.*, 2000). As mentioned in Chapter 3, such Asp-Asn rich areas are also found in repeat areas in proteins that would normally not be exposed to the host immune system (Coppel and Black, 1998; Schofield, 1991). Asn/Glu-rich areas have also been described as 'prion-domains' that function during protein-protein interactions and link functional domains in certain proteins (Michelitsch and Weissman, 2000; Wickner, *et al.*, 2000). Perutz has suggested that poly-Glu or poly-Asn repeats and possibly regions rich in other polar residues might behave as modular mediators of protein-protein interactions termed 'polar zippers' because of the capacity of their side chains to form hydrogen bonded networks (Perutz, *et al.*, 1994). Removal of the parasite-specific insert in glutathione reductase led to an unstable protein indicating that these areas are important in the folding and stability of the protein (Gilberger, *et al.*, 2000). Therefore, in the absence of distinctive, predicted antigenic properties of the inserts, it is possible that the parasite-specific areas containing Nx and (NND)x-repeats in PfAdoMetDC/ODC might be involved in protein-protein interactions to stabilise the heterotetrameric bifunctional complex.

In Chapter 3, the complete amino acid sequence of the bifunctional PfAdoMetDC/ODC was analysed for possible low-complexity regions, the majority of which were located

in the parasite-specific inserts and hinge region (Fig. 3.12 and 4.2). This correlates with the results of Pizzi *et al.* who showed that such areas found in hydrophilic regions in *P. falciparum* proteins co-inside with parasite-specific, rapidly diverging insertions (Pizzi and Frontali, 2001). The low-complexity regions found in more conserved areas of the protein make up a minor subset of prevalently hydrophobic areas and are proposed to be involved in the core structure of these proteins (Pizzi and Frontali, 2001).

4.4.4) Involvement of the parasite-specific inserts in the decarboxylase activities of PfAdoMetDC/ODC.

It is thus possible that the parasite-specific inserts function as interaction sites to enable formation of the bifunctional PfAdoMetDC/ODC protein or catalytic activities. These possible functions were investigated by determination of the effects of deletion of the parasite-specific inserts on bifunctional complex formation and domain activities.

Deletion mutagenesis of the parasite-specific inserts in the bifunctional PfAdoMetDC/ODC indicated that the inserts are essential for the activity (and inherent conformation) of the involved domain (Fig. 4.6). However, deletion of the parasite-specific inserts also affect the activity/conformation of the neighbouring domain. Specifically, the inserts closer to the neighbouring domain in terms of the linear amino acid sequence (A_1 and O_1) seem to have the greatest influences on the activity of the other domain. Previous point mutation studies of the active site residues indicated that the two decarboxylase activities in PfAdoMetDC/ODC are able to function independently (Wrenger, *et al.*, 2001). The results presented here indicate that the parasite-specific inserts are involved in specific communication between the two domains in the bifunctional complex.

All the deletion mutants in the covalently linked PfAdoMetDC/ODC were still able to form bifunctional complexes. It is possible that the parasite-specific inserts do not act alone in the stabilisation of the bifunctional protein complex but that other interactions also contribute to the stabilisation and conformation of the individual domain. Cumulative interactions have been proposed to stabilise the *T. brucei* ODC dimeric interface (Myers, *et al.*, 2001). It was furthermore shown that mutation of single residues far removed from the active site decreased catalytic activity mostly due to long-range energetic coupling of these residues to the active site (Myers, *et al.*, 2001). Therefore, deleting the parasite-specific inserts in PfAdoMetDC/ODC may display

similar effects on catalytic activities of the individual domains even though their properties indicate a surface location far removed from the actual active site centres.

4.4.5) Characterisation of the physical association between the decarboxylase domains.

The physical association between AdoMetDC and ODC was confirmed by the stable reassociation of the individually expressed monofunctional PfAdoMetDC and PfODC domains in a bifunctional hybrid complex that reflects the properties of the bifunctional PfAdoMetDC/ODC (Fig. 4.10). This species-specific physical contact seems to be mediated in part by a parasite-specific insert in the ODC domain. Insert O_1 is predicted as a structured hydrophilic area that does not contain low-complexity regions and shows minimal sequence variability between *Plasmodium* species (Fig. 4.2 and 4.3). This region is also more important for both decarboxylase activities in the bifunctional PfAdoMetDC/ODC (Fig. 4.6). The larger parasite-specific inserts do not seem to mediate physical interactions between the two domains (Fig. 4.11 and Table 4.2). These areas show large sequence length and property diversion between different *Plasmodial* species (Fig. 4.2 and 4.3). Structural analyses of the large parasite-specific inserts (A_1 and O_2) predicted hydrophilic, nonglobular regions containing low-complexity areas in agreement with the results of Pizzi and Frontali (Pizzi and Frontali, 2001). These authors showed that similar inserts in other *Plasmodial* proteins are also hydrophilic in nature and contain low-complexity regions. These low-complexity regions were proposed not to be involved in the correct folding of the proteins and most probably form nonglobular domains that are extruded from the protein core. This might illustrate that if large global conformation changes occurred due to the deletion of such large areas (such as A_1 and O_2) it did not influence enzyme activity as much as deletion of a smaller, more structured insert (O_1). Further investigations are required to determine the exact role of the low-complexity areas within these parasite-specific inserts in complex formation and activities of the two domains. The effect of deletions of inserts in the ODC domain on the formation of the obligate homodimer is not known at this stage. It is not unlikely that these deletions may prevent formation of the homodimeric state thereby preventing association with AdoMetDC.

Deletion of the hinge region in PfAdoMetDC/ODC had a more pronounced impact on ODC activity in the bifunctional enzyme (50% reduced) than on AdoMetDC activity (18% reduced). Monofunctional ODC lacking the hinge region is only half as active as

129

ODC expressed with part of the hinge region (Krause, *et al.*, 2000) and prevents the ODC domain to interact with the wild type AdoMetDC domain. It therefore seems that the hinge region is important for the stability/activity of the ODC domain by mediating the correct folding of the domain to ensure the active homodimeric ODC or by actually constituting part of the ODC domain and is smaller than the currently defined 180 residues.

None of the deletion mutants of the bifunctional PfAdoMetDC/ODC exhibit any effects on heterotetrameric complex formation (Fig. 4.9). However, in contrast to the results obtained with the deletion mutants of the bifunctional complex, deletions of the hinge region or insert O_1 in monofunctional ODC prevented formation of the hybrid, heterotetrameric AdoMetDC/ODC complex. However, deletion of insert A_1 in monofunctional AdoMetDC had no effect on hybrid complex formation. Taken together the results obtained with the deletion mutants of the bifunctional and monofunctional enzymes suggest that the AdoMetDC insert is not involved in heterotetrameric complex formation but only in protein-protein interactions affecting its own activity and that of the ODC domain. The roles of the ODC inserts seem to be more complex since their deletion affects not only its own activity and that of AdoMetDC but also heterotetrameric complex formation. Furthermore, when the complex formation of the mutated bifunctional proteins is compared to the hybrid complex formation of the mutant monofunctional proteins, it seems possible that the complex formation in the bifunctional proteins were due mostly to interactions between the AdoMetDC domains, with no apparent contribution of the mutant ODC. However, the ODC domain is more refractory to change to be able to interact with the AdoMetDC domain when these monofunctional proteins are not covalently linked.

The data presented here indicate that although the two decarboxylase activities can function independently of each other, physical protein-protein interactions are present in the bifunctional PfAdoMetDC/ODC that has effects on both enzyme activities and heterotetrameric complex formation. Future investigations on the role of the parasite-specific inserts in these protein-protein interactions include: 1) Deletion of only the low-complexity areas in the parasite-specific inserts to determine their contribution to protein-protein interactions and validating their proposed surface locality; 2) Mutation of only the polar residues (e.g. Asp and Asn repeats) in the parasite-specific inserts to apolar residues. This should provide evidence for the polar zipper theory of modular

mediators of protein-protein interactions and furthermore limit global conformational changes due to deletion of the large, parasite-specific inserts; 3) Determination of the oligomeric state of the mutated monofunctional proteins to determine the weight of the contribution of each domain to the protein-protein interactions via their ability or not to still form their own intramolecular interactions.

This chapter contributed to understanding the structure-functional relationships that stabilise the bifunctional heterotetrameric PfAdoMetDC/ODC. In Chapter 5, detailed structural characteristics of a three-dimensional homology model of the ODC component of the bifunctional enzyme are described.

CHAPTER 5

Comparative properties of a homology model of the ornithine decarboxylase component of the *P. falciparum* S-adenosylmethionine decarboxylase/ornithine decarboxylase.

5.1) INTRODUCTION

The understanding, modification and manipulation of protein function generally require knowledge of the three-dimensional (3D) structure of a protein at the atomic level. Detailed knowledge of the structure and function of the individual ODC and AdoMetDC enzymes of the malaria parasite is required in order to clarify and understand their arrangement and interactions in the unique bifunctional complex. Structural data is available for the murine, human, *T. brucei* and *Lactobacillus* ODC enzymes (Almrud, *et al.*, 2000; Grishin, *et al.*, 1999; Kern, *et al.*, 1999; Momany, *et al.*, 1995; Vitali, *et al.*, 1999) and for the human AdoMetDC enzyme (Ekstrom, *et al.*, 1999) but not for the malarial enzymes.

X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are the preferred methods to obtain detailed 3D structural information of proteins but both methods are time consuming and require large quantities of protein. In addition, NMR spectroscopy cannot resolve the structures of proteins larger than 200 residues or those of flexible proteins, while X-ray crystallography depends on the generation of suitable crystals (Sali, 1995). The high A+T content of the parasite genome contributes in many instances to the low or insignificant expression of protein from heterologous systems (Baca and Hol, 2000; Chang, 1994). Furthermore, crystallization of malarial proteins is problematic due to the characteristic prevalence of regions of low-complexity and/or inserted amino acids, as evidenced by the paucity of protein crystal structures.

Comparisons between various proteins have demonstrated that their tertiary structures are usually better conserved in evolution than their amino acid sequences (Blundell, *et al.*, 1987; Srinivasan, *et al.*, 1996). It is generally accepted that high sequence similarity is reflected by distinct structure similarity. The root mean square deviation (RMSD) for

132

protein α-carbon backbones sharing 50% primary sequence identity is expected to be better than 1 Å. This served as the premise for the development of knowledge-based comparative protein structure modelling methods by which a homology model for a new protein is extrapolated from the known three-dimensional structure of related proteins (Fig. 5.1) (Peitsch, 1996; Sali, 1995). This has resulted in the application of comparative or homology modelling as an additional and/or alternative method to obtain protein structural information (Srinivasan, *et al.*, 1996).
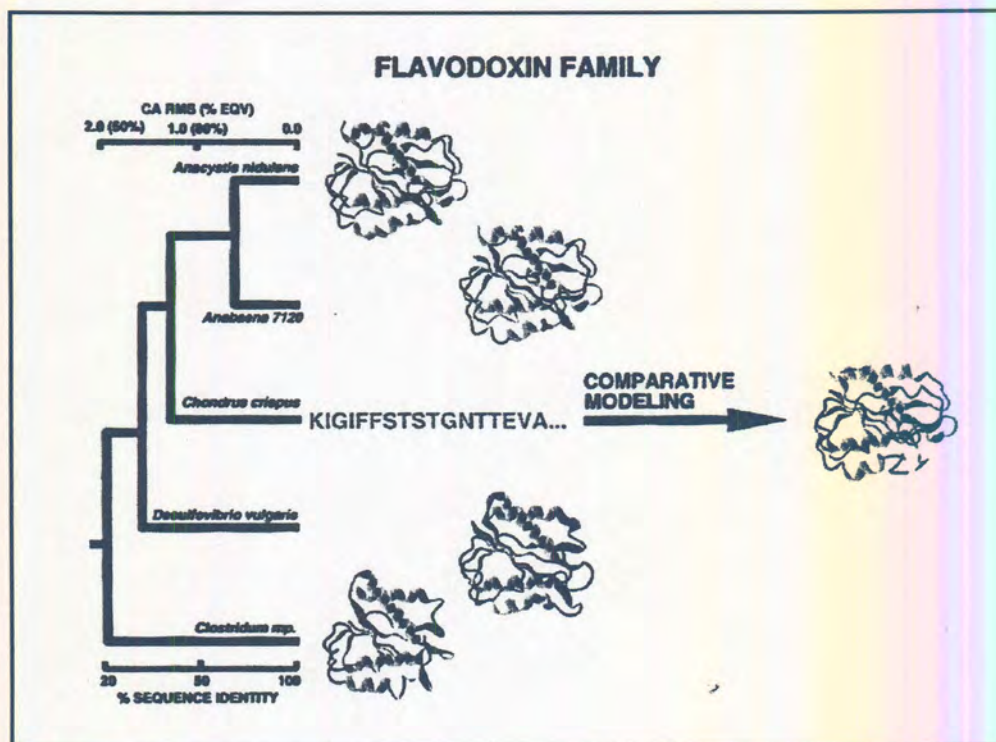
Figure 5.1: Comparative homology modelling due to the evolutionary precept that protein families have both similar sequences and 3D structures. The flavodoxin family is depicted where a protein in this family can be modelled from its sequence using the other structures in the family. The tree shows the % sequence and structural similarity. Adapted from (Sali, 1995).

A recent version of the Protein Information Resource Protein Sequence Database (PIR-PSD 71.03) contained 283 138 entries of protein sequences on 15th of February 2002. In contrast, the Protein Data Bank of experimentally determined protein structures contained only 17 304 structures on the 12th of February 2002. Since about one third of known sequences appear to be related to at least one known structure, the number of sequences that can be modelled is an order of magnitude larger than the number of experimentally determined protein structures (Oregano, *et al.*, 1994). Furthermore, the usefulness of homology modelling is increasing because the various genome projects are producing more sequences and the speed at which novel protein folds are being

determined is increasing due to the application of high-throughput methods (Sanchez, *et al.*, 2000; Taylor, 2002).

Homology modelling uses experimentally determined protein structures as templates to predict or extrapolate the conformation of another protein that has a similar amino acid sequence (>40% identity) (Sanchez and Sali, 1997). This is possible because a small change in the sequence usually results in a small change in the 3D structure. Insertions and deletions occurring during evolution are usually confined mainly to loops between secondary structures and do not alter the fold of the protein. This gives rise to families of protein folds having related structures but varying sequence identities (Jones, *et al.*, 1996). Homology modelling consists of four sequential steps beginning with the identification of the proteins with known 3D structures that are related to the target sequence and used as templates for the structure extrapolation (Srinivasan, *et al.*, 1996)(Fig. 5.2). The second and most crucial step is to align these sequences with the target sequence. This is followed by building of the model for which various calculations are possible. In the fourth step, the model is evaluated using a variety of criteria. These steps can be repeated until a satisfactory model is obtained as indicated in Fig. 5.2. The main difference between the various comparative modelling methods is in how the 3D model is calculated from the sequence alignment (Srinivasan, *et al.*, 1996). The original method of modelling used rigid-body assembly to model a protein from a few core regions, loops and sidechains obtained from related structures. The rigid bodies are assembled on a framework defined as the average of the α-carbon atoms in the conserved regions of certain folds. Another method uses modelling by segment matching that relies on the approximate positions of conserved atoms from the templates to calculate the coordinates of other atoms. The third group of methods involves modelling by satisfaction of spatial restraints with either distance geometry or optimisation techniques (Blundell, *et al.*, 1987; Sali, 1995; Srinivasan, *et al.*, 1996).
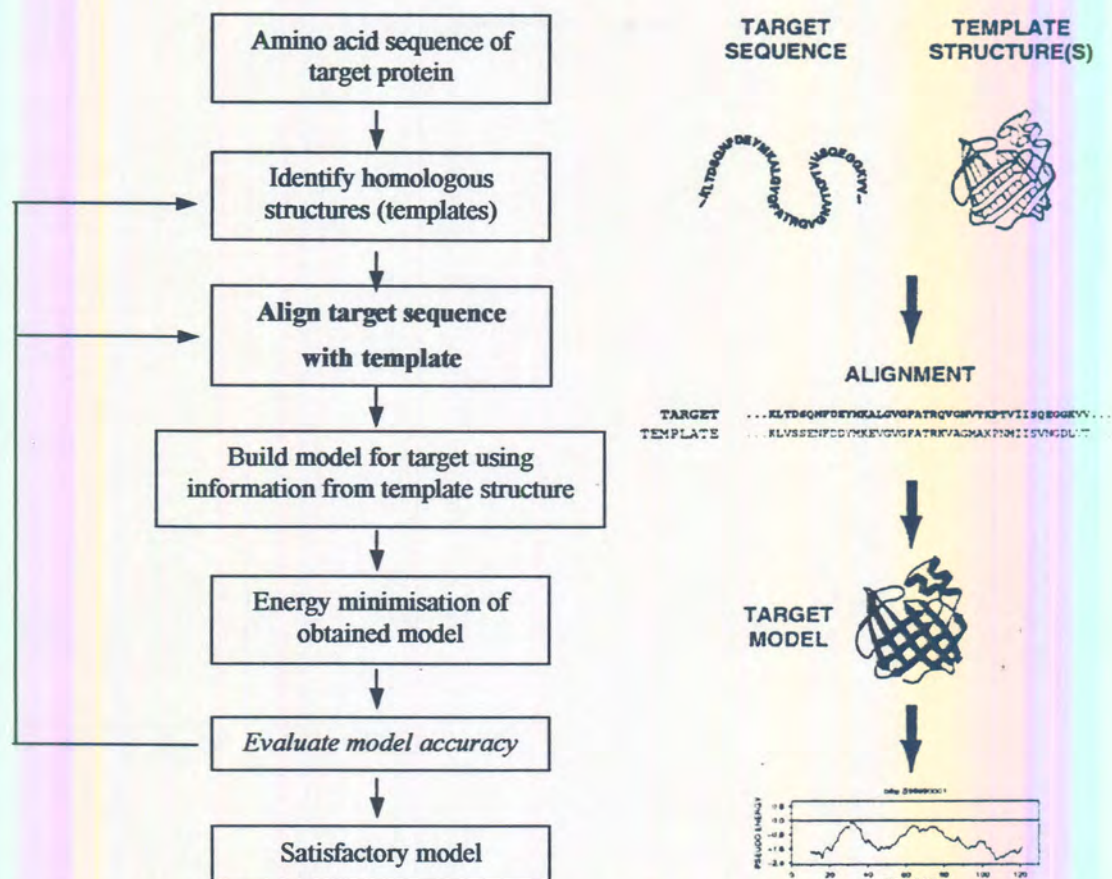
Figure 5.2: Steps in comparative protein structure modelling. Adapted from (Srinivasan, *et al.*, 1996).

No crystal structure could be obtained yet for malarial dihydrofolate reductase (DHFR) notwithstanding its widespread resistance to anti-folates and importance as a selective and validated antimalarial drug target. Malarial DHFR also occurs in a bifunctional complex with thymidylate synthase (TS) and contains inserted amino acids (Lemcke, *et al.*, 1999; Toyoda, *et al.*, 1997). The mechanism of resistance of DHFR to known antimalarials however, could be explained by a homology model and furthermore led to the discovery of lead inhibitors (Lemcke, *et al.*, 1999; Rastelli, *et al.*, 2000; Toyoda, *et al.*, 1997 Warhurst, 1998 #151).

The AdoMetDC structure for the human enzyme has been crystallised to 2.25 Å resolution (Fig. 5.3 A (Ekstrom, *et al.*, 1999). The mature protein is a dimer consisting of two α- and β-chains. The architecture of each αβ monomer is a novel four-layer α/β-sandwich fold, comprised of 2 antiparallel 8-stranded β-sheets flanked by several α- and 3_{10} helices (Ekstrom, *et al.*, 1999). The low sequence identity (<20%) between the

AdoMetDC domain of the bifunctional PfAdoMetDC/ODC (excluding the parasite-specific regions) and AdoMetDCs from other organisms (Chapter 3) complicates comparative structural modelling of this protein.
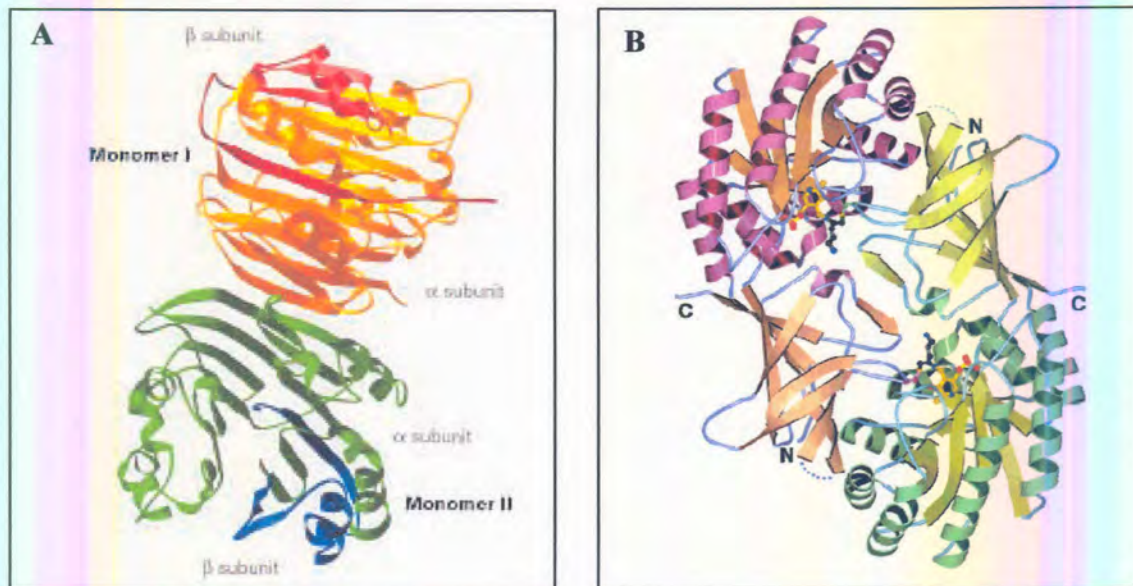
Figure 5.3: Crystal structures of mammalian AdoMetDC and protozoal ODC. (A) The AdoMetDC dimer from *H. sapiens* (Ekstrom, *et al.*, 1999) and in **(B)** the ODC structure from *T. brucei* (Grishin, *et al.*, 1999).

ODC from *T. brucei*, *H. sapiens* and *M. musculus* have been crystallised (Almrud, *et al.*, 2000; Grishin, *et al.*, 1999; Momany, *et al.*, 1995). In all the eukaryotic ODC structures, the protein is found as a dimer with each monomer consisting of two distinct domains, a N-terminal α/β-barrel and a C-terminal β-barrel domain (Fig. 5.3 B). The contacts at the dimer interface are primarily in the C-terminal domains of each monomer, whereas the active site is found in the barrel of the N-terminal domain and is closed-up by a loop from the C-terminal domain of the second monomer projecting into the cavity. The extent of sequence identity (~30%) between ODCs from a variety of species and the malarial ODC domain of the bifunctional PfAdoMetDC/ODC (excluding the parasite-specific regions, Chapter 3) suggested the feasibility of a comparative protein structure modelling approach for the generation of its three-dimensional structure.

This chapter describes the generation of a homology model of only the PfODC component of the bifunctional PfAdoMetDC/ODC and compares the putative structure to other experimentally determined ODC structures.

Some of the results obtained in this chapter have been accepted for publication in Proteins: Structure, Function and Genetics (Birkholtz, *et al.*, 2002a). The results have also been presented as papers at international meetings (Birkholtz, 2000e; 2002c; Birkholtz, *et al.*, 2000a) and as posters at international and national conferences (Birkholtz, 2000b; 2000c; Birkholtz, *et al.*, 2001b; 2001c).

5.2) MATERIALS AND METHODS

5.2.1) *In silico* analyses of predicted structural motifs of PfODC.

PfODC was structurally classified by different *in silico* techniques. A hierarchical classification of ODCs was obtained with the CATH database (Orengo, *et al.*, 1997). Protein domain structures are grouped according to a novel hierarchical classification, which clusters proteins at four major levels, Class (C), Architecture (A), Topology (T) and Homologous superfamily (H). The class is derived from secondary structure content and is automatically assigned for more than 90% of protein structures. Architecture, which describes the gross orientation of secondary structures, independent of connectivities, is currently assigned manually. The topology level clusters structures according to their topological connections and numbers of secondary structures. The homologous superfamilies cluster proteins with highly similar structures and functions. The assignments of structures to topology families and homologous superfamilies are made by sequence and structure comparisons. SCOP (Murzin, *et al.*, 1995) allows the Structural Classification of Proteins by providing a detailed and comprehensive description of the structural and evolutionary relationships between all proteins whose structures are known. As such, it provides a broad survey of all known protein folds and detailed information about the close relatives of any particular protein. Comparing 3D structures may reveal biologically interesting similarities that are not detectable by comparing sequences. The DALI server (http://www.ebi.ac.uk/dali/) (distance matrix alignment) allows the comparison of the 3D protein structure to all the other known structures of proteins in the Protein Data Bank (Holm and Sander, 1993). Low complexity regions in PfODC were identified with the SEG Program (Wootton and Federhen, 1996).

5.2.2) Comparative modelling of monomeric PfODC.

Comparative homology modelling was performed according to the method originally described by Browne in 1969 (Blundell, *et al.*, 1987; Browne, *et al.*, 1969). This is based on the assembly of a number of rigid bodies obtained from aligned protein structures followed by framework calculations, generation of mainchain atoms of the core regions, generation of loops and modelling of the sidechains based on their intrinsic conformational preferences and the conformation in the template structures. The stereochemistry of the model is then improved with energy minimisations.

Multiple pairwise alignment with the CLUSTAL W programme (Thompson, *et al.*, 1994) was used to compare the PfODC amino acid sequence as described in Chapter 4. The malaria-specific insert O_2 (residues 1139-1296) and the hinge region (residues 573-837) identified in this alignment were removed from the PfODC sequence due to the absence of the corresponding sequence in the modelling template and absence of known structural homologous. The remaining 411 residues (838-1138/1297-1406) of the ODC component was submitted to the SWISS-Model server (Automated Protein Modelling Server, Version 3.5, GlaxoWellcome Experimental Research, Geneva, Switzerland; (Guex, *et al.*, 1999; Guex and Peitsch, 1997; 1999) for comparative protein structure modelling by rigid body assembly with the following knowledge-based approach (Peitsch, 1995a; Peitsch, 1995b; Peitsch, 1996; Peitsch, *et al.*, 1996): Suitable templates on which to base the model were found by searching all similarities within the target sequence compared to sequences of known structures, using BLASTP2 searches of the ExNRL-3D database (SWISS-Model sequence database, reflecting the protein sequences of ExPDB). Templates with a sequence identity above 25% and larger than 20 residues were selected by SIM© and used to detect domains that could be modelled based on unrelated templates (Huang and Miller, 1991). ProModII was subsequently employed to generate models using ExPDB (The structure database used by the SWISS-Model is derived from the Brookhaven Protein Data Bank (PDB, BNL). Energy minimization and structure refinement was done with GROMOS96 to reduce steric overlap specifically in side-chains (default parameters using steepest gradient for 200 cycles with Gromos96 force field, BIOMOS b.v. Company).

The resulting model was validated manually with the WHAT_CHECK module of the WHAT IF program (version 19970813-1517; (Vriend, 1990) and with the PROCheck program (Laskowski, *et al.*, 1993). Molecular surfaces and potentials were created with

GRASP (Graphical representation and analyses of structural properties; Columbia University, New York; (Nicholls, *et al.*, 1991).

Models were visualized and edited with SWISS-PDB Viewer and analysed with the InsightII package (Accelrys, San Diego, USA) on a Silicon Graphics Octane workstation (Silicon Graphics, Mountainview, USA). The SWISS-PDB Viewer scenes were rendered with POV-Ray.

5.2.3) Dimerisation of monomeric PfODC.

The dimeric form of PfODC was built by superimposing the PfODC monomers on the dimeric *T. brucei* crystal structure using the Improved fit module of SWISS-PDB Viewer and merging the coordinates into one planar field. The resulting dimeric PfODC was then subjected to energy minimization with the Discover3 module of the InsightII package (cff91 force field for 10 000 iterations with a conjugate gradient) and checked for any disallowed bumps occurring between the two different chains. Interacting residues were analysed with Protein Explorer (http://www.umass.edu/microbio/chime/explorer) and LigPlot (Version4.0; (Wallace, *et al.*, 1995). The structure was analysed for accuracy with the WHAT_CHECK module of the WHAT IF program (Vriend, 1990).

5.2.4) Docking of ligands into the active site of dimeric PfODC.

Active site residues were identified as those corresponding to proven functional residues in the active site pockets of the *T. brucei* and human ODC crystal structures (Almrud, *et al.*, 2000; Coleman, *et al.*, 1993; Grishin, *et al.*, 1999; Osterman, *et al.*, 1994). These include Lys_{69}, Arg_{154}, His_{197}, $Gly_{235-237}$, Glu_{274}, Arg_{277}, Tyr_{389}, Asp_{332}, Cys_{360} and Asp_{361} (numbers according to the *T. brucei* protein). Structures for PLP and ornithine were generated with the Builder module of the InsightII package and their energies minimized as described above. Binding of PLP and ornithine requires the formation of a Schiff-base between the two ligands with ornithine then also forming a covalent bond to the S_γ atom of Cys_{360} (*T. brucei* numbering). In order to dock this transition state complex of PLP-ornithine into PfODC, a structure for the linked PLP-ornithine was created and allowed to form a covalent link with Cys_{1355}. The ligand-ODC complex was then minimized as described above. Possible interactions between the ligands and residues in PfODC were analysed with LigPlot (Wallace, *et al.*, 1995). The structures were analysed for accuracy with the WHAT IF program (Vriend, 1990).

5.2.5) Limited proteolysis studies.

Limited proteolysis is a powerful tool for probing the higher order structure of proteins by using classical biochemical methods (Hubbard, 1998). This is achieved under nondenaturing conditions by limiting the proteolytic reaction of various proteases through altering the reaction conditions such that digestion of every susceptible peptide bond is prevented and only the location of certain bonds with respect to the overall fold of the protein is obtained. Recombinantly expressed PfODC (ODC domain containing the hinge region, Chapter 3, section 3.2.2) was subjected to limited proteolysis according to a modification of the methods by Hubbard and Wilkinson (Hubbard, 1998; Wilkinson, 2000). Briefly, the expressed protein was isolated as described in Chapter 3, section 3.2.2.2 and subjected to either proteinase K (Roche, Mannheim, Germany) or trypsin (Macherey-Nagel, Duren, Germany) digestion. Enzyme:substrate ratios were optimised at between 1:50 and 1:100 and $\sim$ 750 ng protein was subjected to proteolysis at room temperature in 10 mM Tris-HCl (pH 8.0) for 0, 5, 20 and 60 min intervals. The reactions were stopped by addition of 0.1 mM PMSF and 2 x SDS-PAGE loading dye and boiling for 5 min. The digested samples were then analysed on a 12.5 % SDS-PAGE and stained with silver as described in Chapter 3, section 3.2.6. Limited proteolysis sites were predicted by analysing the obtained dimeric PfODC model with the Nickpred Server (sjh.bi.umist.ac.uk/cgi-bin/npred/nickpred) (Hubbard, 1998).

5.3) RESULTS.

5.3.1) Structural classification of PfODC.

Comparisons between homologous proteins have shown that conformations are better conserved in evolution than the corresponding amino acid sequences (Srinivasan, *et al.*, 1996). The CATH database places ODC in a hierarchical fashion with the Lyase homologous superfamily that shares topologies, consisting of a barrel-like architecture, with lyases and thrombin. These proteins are grouped into the mainly β single domain class of proteins according to evolutionary and structural groupings. SCOP places ODC in a superfamily of PLP-binding proteins that include the alanine racemase-like family, based on a triosephosphate isomerase (TIM) barrel-like fold. The α/β-barrel structure found in these proteins seems extremely well preserved even in distant homologues (Alanine racemase and TIM proteins) with diverse functions.

140

5.3.2) Modelling monomeric PfODC.

In order to apply homology modelling, the first non-trivial step is to obtain a multiple-alignment of the query amino acid sequence against sequences from other known structures. At present, there are no known homologues of the inserted or hinge region sequences. The largest insertion of 158 residues (O_2) and the hinge region were thus removed in order to arrive at a satisfactory homology model based on multiple sequence alignments used to describe and define the inserts in Chapter 4. Subsequent pair-wise alignment of the remaining 411 residues of PfODC showed the highest identity to the amino acid sequences of the *T. brucei* enzymes (41.54 %) and 39.23 % with the mouse enzyme (PDB # 7ODC, Fig. 5.4). The crystal structure of the *T. brucei* enzyme obtained with bound co-factor PLP (PDB # 1QU4 at 2.9 Å) was therefore used as template to build the PfODC homology model. Fig. 5.4 shows the alignment between the 411 PfODC residues and *T. brucei* ODC used to create the model and also indicates the secondary structural elements for each protein as predicted by the Swiss-Model server.

The resulting homology model consisted of 373 residues based on the template structure plus the *ab initio* constructed 39 residue malaria-specific insert O_1. The root mean square of deviation (RMSD) value between the model carbon α-backbone and the *T. brucei* structure was 0.816 Å as determined by PROFIT. This value increased to 6.917 Å when the *ab initio* modelled insert O_1 of 39 residues was included in the comparison.

141

```
PFODC   838           SVVCINLQKILAQYVRFKKNLPHVTPFYSVKSNNDEVVIKFLYG
TbODC    35  DEGD--PFFVADLGDIVRKHETWKKCLPRVTPFYAVKCNDDWRVLGTLAA
                      sssssshhhhhhhhhhhhhhh     ssssss        hhhhhhhhh
                      sssssshhhhhhhhhhhhhhh     ssssss        hhhhhhhhh

PFODC   882  LNCNFDCASIGEISKVIKLLPNLSRDRIIFANTIKSINSLIYARKENINL
TbODC    83  LGTGFDCASNTEIQRVRGI--GVPPEKIIYANPCKQISHIRYARDSGVDV
              h  ssss   hhhhhhhhh            sssss      hhhhhhhhhh  ss
              h  ssss   hhhhhhh              sssss      hhhhhh      ss

PFODC   932  CTFDNLDELKKIYKYHPKCSLILRINVDFKNYKSYMSSKYGANEYEWEEM
TbODC   131  MTFDCVDELEKVAKTHPKAKMVLRIST--------LSVKFGAKVEDCRFI
              sss   hhhhhhhh     ssssss                      hhhhhh
              sss   hhhhhhhh     ssssss                      hhh

PFODC   982  LLYAKKHNLNIVGVSFHVGSNTKNLFDFCLAIKLCRDVFDMSSNMGFNFY
TbODC   181  LEQAKKLNIDVTGVSFHVGSGSTDASTFAQAISDSRFVFDMGTELGFNMH
              hhhhhhh     sssss          hhhhhhhhhhhhhhhhhhhhhh    s
              hhhhhhh     sssss          hhhhhhhhhhhhhhhhhhhhhh    s

PFODC  1032  IINLGGGYPEELEYDNAKKHDKIHYCTLSLQEIKKDIQKFLNEETFLKTK
TbODC   231  ILDIGGGFPGT-----------------------------RDAPLK--
              sssss         sssss                 sssss    sss
              ssss

PFODC  1082  YGYYSFEKISLAINMSIDHYFSHMKDNLRVICEPGRYMVAASSTLAVKII
TbODC   248  -----FEEIAGVINNALEKHFPP-DLKLTIVAEPGRYYVASAFTLAVNVI
              sss   hhhhhhhhhhhhh          ssssss     ssssssss
                    hhhhhhhhhhhhh          ssssss     ssssssss

                                ↓
PFODC  1132  GKRRPTFQRYNYFSYYVSDSIYGCFSGIIFDEYNRCPIYVIKNKNNPNQN
TbODC   292  AKKVTPAQS---FMYYVNDGVYGSFNCILYDHAVVRPL--PQREPIPNEK
              sss           sssss        hhhh
              sssss   s     sssss        hhhh        s   ss

PFODC  1339  FMNFNLYLANVFGQSCDGLDMINSITYLPECYINDWLLYEYAGAYTFVSS
TbODC   350  -----LYPSSVWGPTCDGLDQIVERYYLPEMQVGEWLLFEDMGAYTVVGT
                   sssssssss     sss    sss       sssss
                   sssssssss     sss    sss       sssss

PFODC  1390  SNFNGFKKCKKVYIFPE
TbODC   395  SSFNGFQSPTIYYVVSG
                  sssss
                  sssss
```

Figure 5.4: Sequence alignment of *P. falciparum* ODC (PfODC) and the template used for homology modelling, *T. brucei* ODC (TbODC, PDB: 1QU4) obtained with SIM© using default parameters. Identical residues are shaded and the secondary structural elements are indicated: s for β-sheets and h for α-helices. The site where insert O_2 was removed to create the PfODC model is indicated with an arrow. Insert O_1 is indicated with the black bar.

5.3.3) Evaluation of the PfODC model quality and accuracy.

PROCheck analyses of the PfODC model that included insert O_1 produced a Ramachandran plot in which 84.2% of the residues were in the most favoured regions indicating the model to be sterically acceptable (Fig. 5.5). The only residues that were present in disallowed areas were Leu_{1061}, Glu_{1063} and Phe_{1077}. These residues form part of insert O_1 and predicting their exact conformational arrangement is therefore impeded by the inherent difficulty in modelling unknown loops.
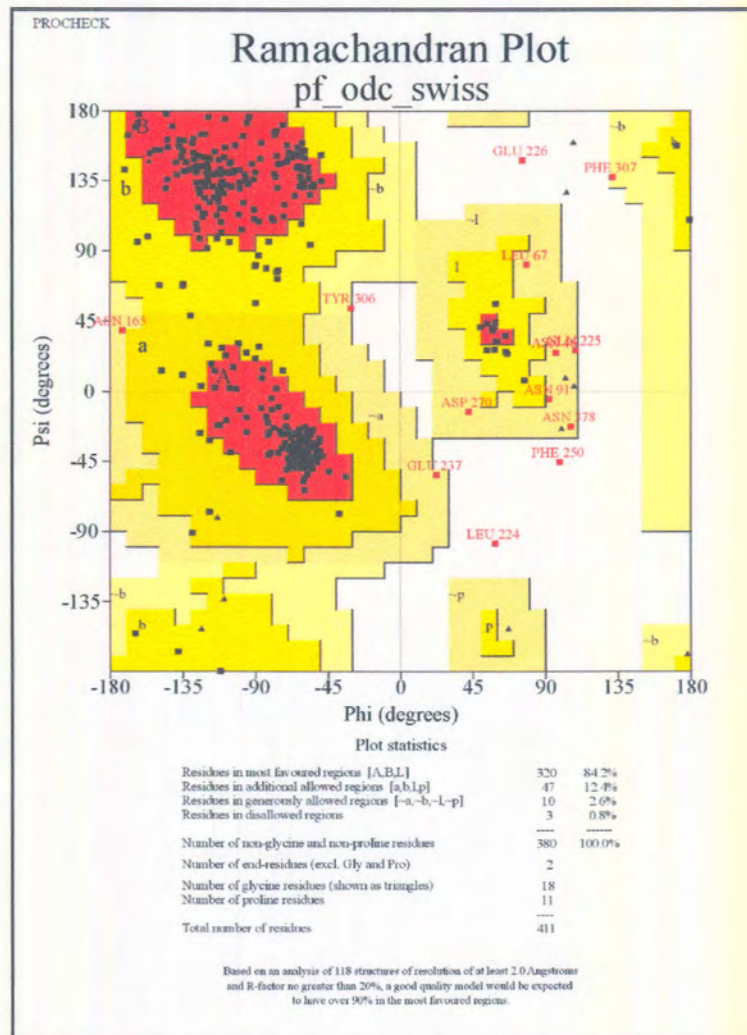
Figure 5.5: Ramachandran plot for the model of PfODC produced by PROCHECK. 84.2% of the residues are present in favourable structural areas with the exception of 3 residues in disallowed regions (Glu, Leu and Phe).

Most of the main-chain and side-chain parameters were better than typical values allowed and the rest of the parameters were within the allowed ranges (Fig. 5.6).
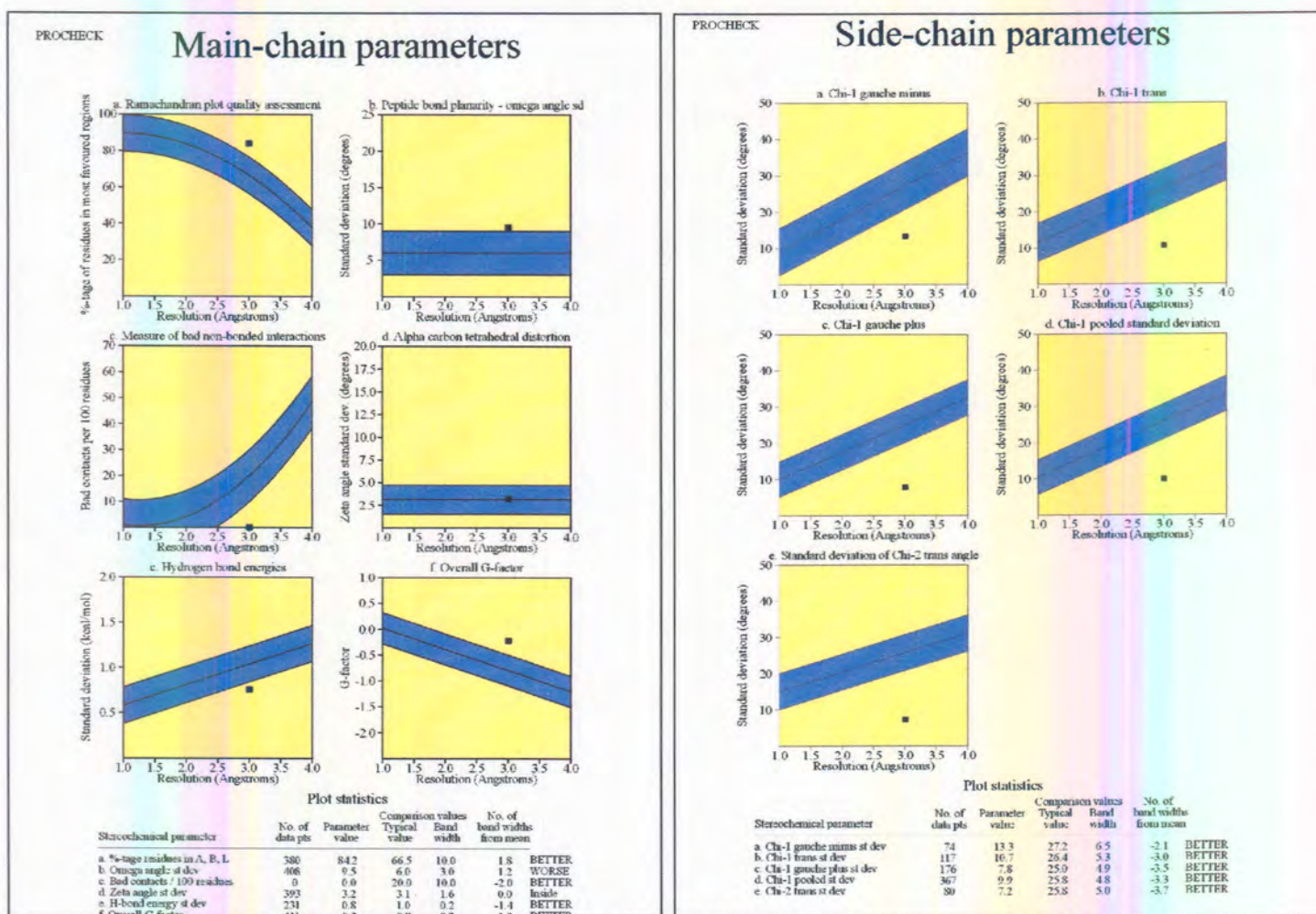
Figure 5.6: PROCHECK analyses of the main-chain and side chain parameters of the final PfODC model. Various parameters are analysed and represented in the different plots. Tables evaluate parameters into allowed ranges.

Quality analyses were also performed with the program WHAT_CHECK from the WHAT IF package and several stereochemical parameters are summarized in Table 5.1. From this Table, the quality of both the monomeric and dimeric forms of PfODC seem as good as those of the reported *T. brucei* and *H. sapiens* structures with which it is compared.

Table 5.1: Summary of WHAT IF quality assessment data. Data for the *T. brucei* (PDB # 1QU4) and *H. sapiens* (PDB # 1D7K) ODC enzymes are compared with the monomeric and dimeric form of PfODC. Values are structure Z-scores with + better than normal. * RMS Z-scores should be close to 1.0.

Structure	2nd generation packing quality	Ramachandran plot appearance	$\chi 1/\chi 2$ rotamer normality	Backbone conformation	Bond lengths*	Bond angle variability*	Omega angle restraints*
1QU4	-2.231	-2.051	-1.848	-0.052	0.612	1.674	1.527
1D7K	-1.1	-1.2	-0.3	-0.6	0.753	1.791	2.360
PfODC: Monomer	-3.658	-0.762	0.263	-1.681	1.282	1.290	1.295
Dimer	-3.978	-0.807	0.283	-1.849	1.282	1.291	1.294

5.3.4) Characterisation of monomeric PfODC.

The PfODC monomer consists of two distinct domains, a N-terminal α/β TIM barrel and a C-terminal modified Greek-key β-barrel (Fig. 5.7 A). These features correlate closely to all the other eukaryotic ODC structures characterized to date (Almrud, *et al.*, 2000; Grishin, *et al.*, 1999; Kern, *et al.*, 1999; Momany, *et al.*, 1995; Vitali, *et al.*, 1999). Evaluation of the relationships with the human ODC crystal structure (PDB # 1D7K) (Almrud, *et al.*, 2000) indicates large similarities, especially in the structural motifs of the N-terminal α/β barrel and the C-terminal β-sheet (Almrud, *et al.*, 2000) (Fig. 5.7 B). Superimposition of the PfODC monomer on the human ODC structure yields a RMSD of 0.80 Å (involving 1348 atoms) with the areas scoring the worst B-factors being the loops connecting the well-conserved structural elements of the core protein. These regions are elongated in the malaria protein compared to other ODCs. The insert O_1 is predicted to have four anti-parallel β-sheets (I1-I4) with the first two longer than the second pair and does not contain low complexity areas (Fig. 5.7 A). No firm conclusions are possible with regard to the exact orientation of the bulk of the loop in space. In the model, the insert seems to lie parallel to the rest of the protein and to bulge out towards the C-terminal domain on the same side of the monomer as the entry to the active site. No significant interactions are apparent between the protein core and the insert. The attachment site for insert O_2, which was removed in PfODC to create the model, is in the C-terminal three-quarter of the loop between C2b and C3 (Fig. 5.7 A).
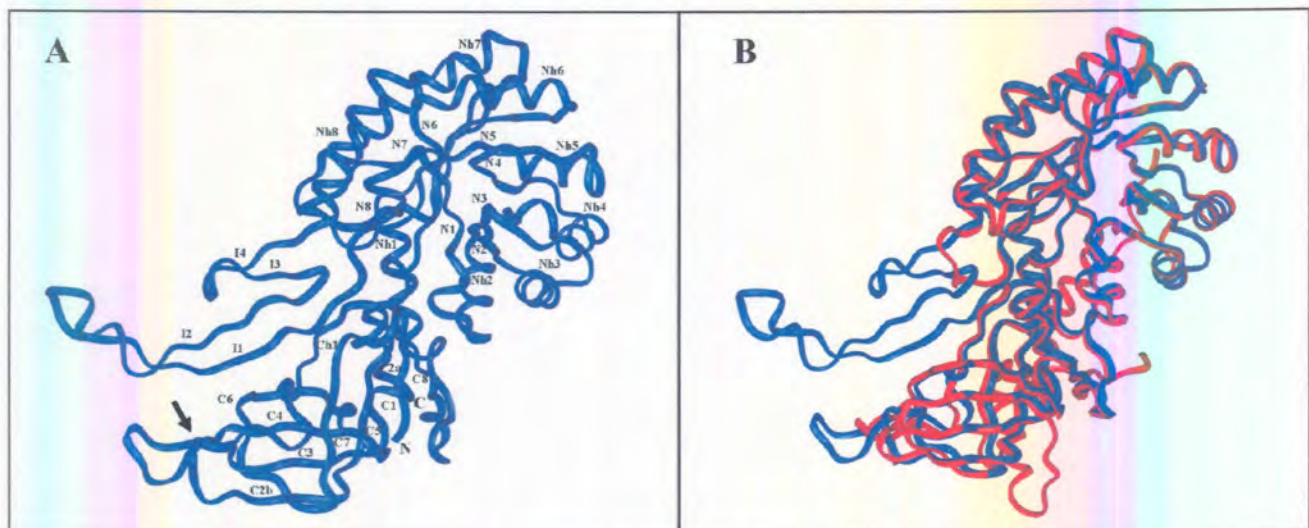
Figure 5.7: Ribbon diagram of the homology model for the PfODC monomer (A) and in (B) compared with the human enzyme. (A) The β-sheets are labelled in succession starting from the N-terminus (N1-N8, I1-I4, C1-C8) and the α-helices are specified in the N-terminal α/β-barrel domain (Nh1-Nh8). The site where insert O_2 was removed to create the PfODC model is indicated with an arrow. **(B)** PfODC in blue is superimposed on the human ODC structure in red.

5.3.5) Characterisation of dimeric PfODC.

Eukaryotic ODC is an obligate homodimeric enzyme with two active sites at the interface between the two monomers (Cohen, 1998; Pegg, *et al.*, 1994). The PfODC monomers were superimposed on the dimeric *T. brucei* crystal structure using the Improved fit module of SWISS-PdbViewer followed by energy minimization with Discover3 to create a dimeric structure of the malarial enzyme. Minimization was performed using a conjugate gradient to a maximum derivative of 0.0030 after 10 000 iterations involving 1398 atoms with RMSD of 0.37 Å and energy of −15 350 kcal/mol. Convergence to a lower derivative was not obtained, probably due to the presence of the malaria-specific areas present as unconstrained loops on the surface of the protein. The RMSD values of the structures prior to and after minimization were 2.370 Å and 3.045 Å for the backbone and side-chain atoms, respectively. Analyses of the quality of the dimer as indicated with the WHAT IF program indicates a good working structure as compared with the *T. brucei* and human structures (Table 5.1).

The dimer comprises of a head-to-tail association of the two monomers, with the C-terminal domain of one monomer vertical to the N-terminal domain of the second monomer (Fig. 5.8).
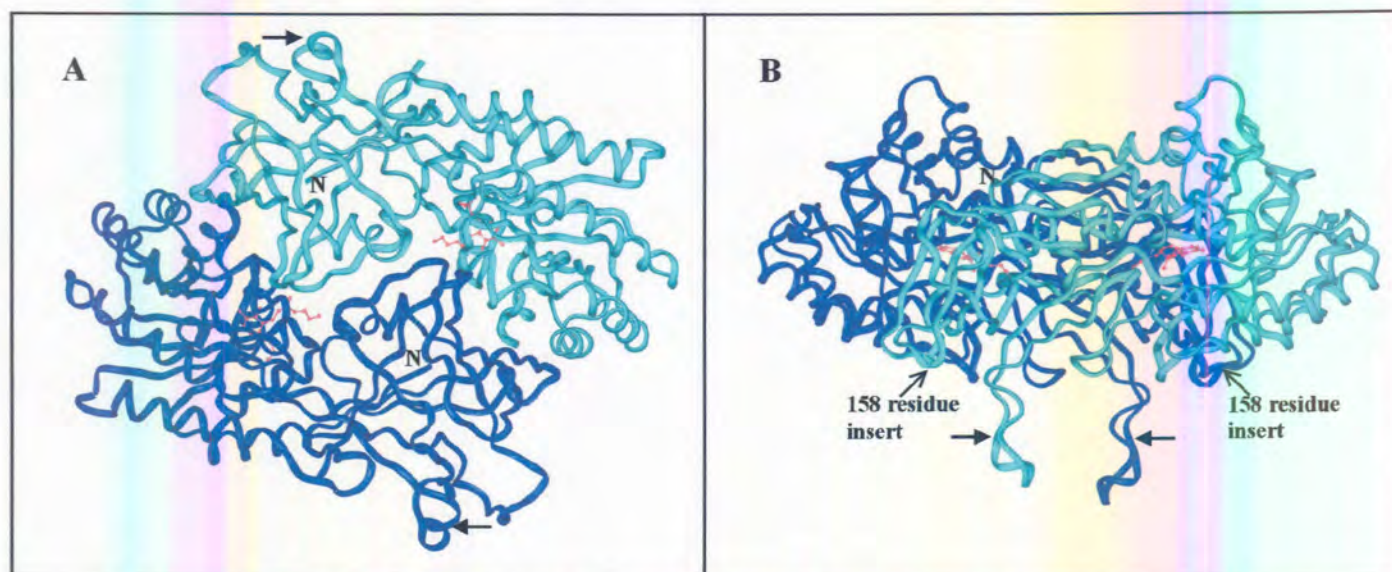
Figure 5.8: Proposed dimeric form of PfODC. The two monomers are indicated in shades of blue and the dimer is viewed from the bottom (A) and side (B). The PLP cofactor and DFMO inhibitor is indicated in red in the two active site pockets formed at the interface between the two monomers. The N-terminus in each monomer is indicated. The location of the 158 residue insert O_2 removed to create the model in the protein is shown in **(B)**. The filled arrows indicate the 39 residue insert O_1 that was modelled.

Several interactions between these two domains are apparent (Fig. 5.9). As is the case in the *T. brucei* enzyme, the dimer interface of PfODC is characterized by an aromatic amino acid zipper (Grishin, *et al.*, 1999). Phe_{1392}, Tyr_{1305} and Phe_{1319} (substituting a second Tyr residue found in *T. brucei*) are involved in hydrophobic contacts across the dimer interface forming an anti-parallel stacked interaction via their aromatic rings. A pronounced hydrophobic contact is predicted between Tyr_{1305} and Ile_{915} in an area that is well conserved in terms of sequence identity in all ODCs but not in PfODC (residues 111-115: ANPCK; PfODC residues 912-916: ANTIK). A salt bridge (1.94 Å) is predicted between Asp_{1359} and Arg_{1134} from the opposite monomer as well as between Lys_{1133} and Glu_{1369} (1.85 Å). There are several stabilizing interactions close to the active site residues. Particularly for PfODC, Lys_{970} is predicted to interact with various residues surrounding the active site Cys_{1355} donated by the opposite monomer. Lys_{970} forms a hydrogen bond with Asp_{1356} and hydrophobic contacts with Gly_{1357}, Asp_{1359} and Gly_{1352}.
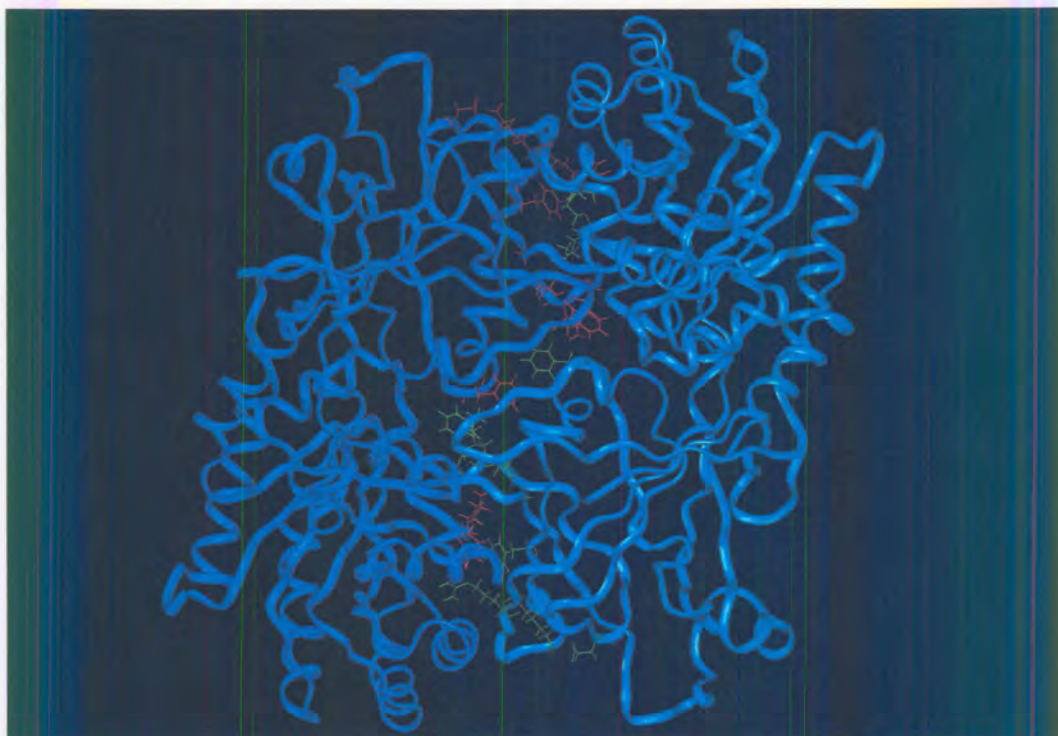
Figure 5.9: Interactions at the ODC dimer interface. The monomers are indicated in different shades of blue and the residues donated from each monomer in red and green respectively.

5.3.6) Active site pocket of dimeric PfODC.

In order to analyse the active site pocket of PfODC, the binding of PLP and ornithine was simulated. Transition state structures for PLP bound to ornithine via a Schiff-base was used and minimized into the active site pocket to an energy of –15 314 kcal/mol for

the protein-ligand complex. The predicted PLP and substrate binding site of the PfODC has a few residues within 3.0 Å to enable hydrogen bonding and charge interactions (Fig. 5.10).
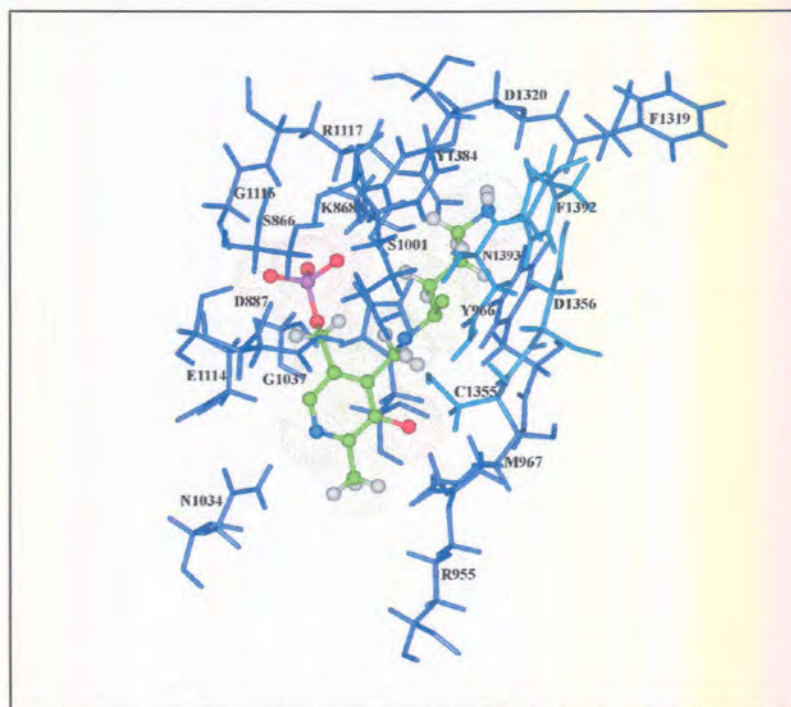
Figure 5.10: Active site residues of the PfODC indicating the interactions with PLP and ornithine. PLP and Schiff-base linked ornithine are indicated in ball-and-sticks coloured for their different atoms and with van der Waals surfaces shown. Residues are coloured in different shades of blue indicating the contribution by the two respective monomers.

Possible residues interacting with either PLP or ornithine were identified using LigPlot and are summarized in Table 5.2. The catalytic residues showed similar spatial orientations as those in the human structure. From these data it is clear that the active site residues are conserved compared to the *T. brucei* and human enzymes in terms of binding to PLP with Cys_{1355} (from the second monomer), Asp_{887}, Arg_{955}, His_{998}, Ser_{1001}, Gly_{1037}, Glu_{1114}, Gly_{1116} and Tyr_{1384}. However, residues Thr_{933}, Met_{967} and Asn_{1034} were present only in the PfODC-PLP binding site. The substrate binding site was derived from interactions with ornithine and consists of mutually conserved (compared to *T. brucei* and the human enzymes) residues Lys_{868}, Asp_{1320}, Cys_{1355}, Asp_{1356}, Tyr_{1384}, Phe_{1392} and Asn_{1393}. In the PfODC model, Cys_{1355} makes contact with Lys_{868} as well as Ala_{889} from the opposite monomer. As with the PLP-binding site, two residues (Tyr_{966} and Arg_{1117}) were only found in the PfODC substrate-binding site and not in either the *H. sapiens* or *T. brucei* binding sites. Of the five PfODC-specific residues characterising the entire active site pocket of this enzyme, only Thr_{933} and Arg_{1117} are conserved in comparison to the primary amino acid sequences of *H. sapiens, T. brucei* and *L.*

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

donovani. The identification of PfODC-specific residues Met_{967}, Asn_{1034} and Tyr_{966} could find application in the rational design of PfODC-specific lead inhibitors.

Table 5.2: Active site residues involved in interactions with ornithine as substrate and PLP as co-factor. Active site residues were identified with LigPlot v 4.0 for the *T. brucei*, *H. sapiens* and *P. falciparum* ODC enzymes. For PfODC the numbering is according to the bifunctional enzyme complex and A and B indicate which monomer contributed a residue towards the active site.

Ligand	*T. brucei* residues	*H. sapiens* residues	PfODC residues
PLP		Cys360	Cys1355A
	Arg154	Arg154	Arg955B
	Glu274	Glu274	Glu1114B
	Asp88	Asp88	Asp887B
	His197	His197	His998B
	Ala67	Ala67	Ser866B
	Tyr389	Tyr389	Tyr1384B
	Ser200		Ser1001B
	Gly276	Gly276	Gly1116B
	Arg277	Arg277	
	Gly236	Gly236	Gly1037B
	Gly237	Gly237	
			Thr933B
			Asn1034B
			Met967B
Substrate	Tyr389	Tyr389	Tyr1384B
	Cys360		Cys1355B
	Phe397	Phe397	Phe1392A
	Asp361		Asp1356A
	Tyr331		Phe1319B
	Asp332		Asp1320B
		Asn398	Asn1393A
		Val68	
		Ala67	
		Lys69	Lys868B
		Asn71	
		Glu94	
		Cys70	
		Ala392	
			Arg1117B
			Tyr966B

5.3.7) Analysis of the molecular surface of PfODC.

To obtain an indication of the surface properties of the PfODC model, the monomeric form of the model was analysed with GRASP to create the molecular surface and indicate the specific potentials of certain projections and cavities. Fig. 5.11 indicates the surfaces of PfODC (Fig. 5.11 A) compared with the surface of the *H. sapiens* ODC (Fig. 5.11 B). PfODC has overall a more positively charged surface especially in the C-

terminal β-sheet domain. Pronounced electro-negative areas are found in the α/β-barrel. One distinct difference is the electrostatic potential at the surface of the active site pockets of the PfODC model and human crystal structure. PfODC has a very pronounced positively charged ring at the entrance to the negatively charged inner pocket whereas this division is not as distinct for the human enzyme.
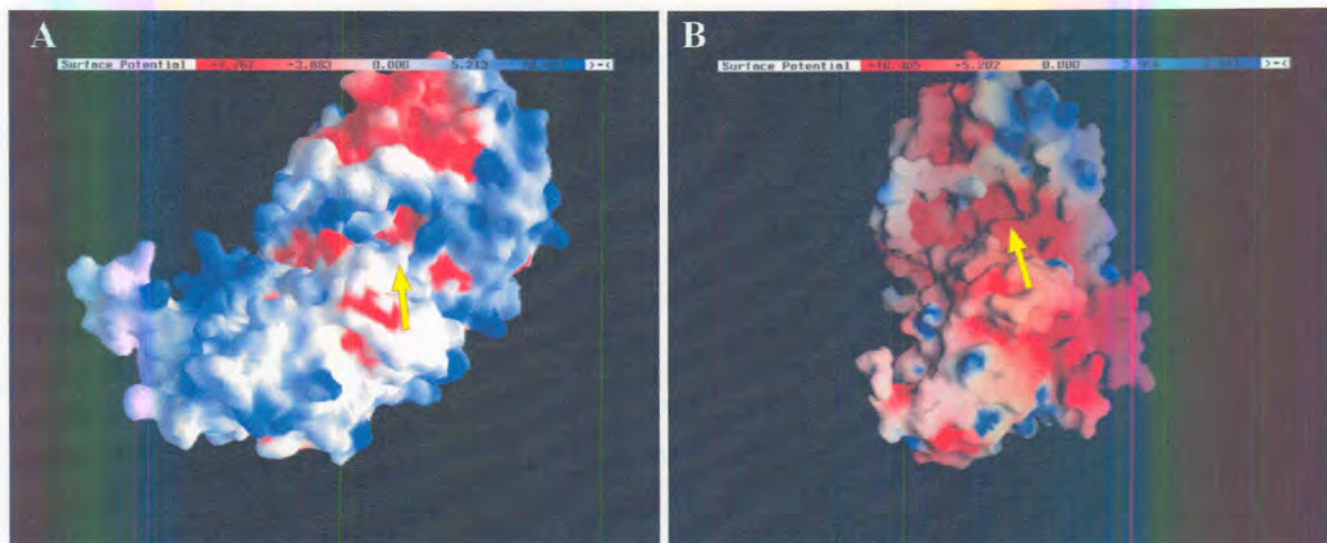
Figure 5.11: Molecular surface potentials of the monomeric PfODC (A) and human ODC (B) structures. Surfaces were created with GRASP and are coloured for electrostatic potentials with red being the most electrostatically negative and blue the most positive. Arrows indicate the view into the α/β-barrel and active site pocket.

5.3.8) Binding pocket for antizyme in PfODC.

Mammalian ODC is known to be regulated by the binding of a putrescine induced protein, antizyme. Antizyme binds to monomeric ODC in the N-terminal domain at a distinctly positively charged area called the antizyme-binding element (AzBE) and this binding is proposed to induce conformational changes in ODC to expose the C-terminal end containing a PEST region to target the degradation of ODC by the 26S proteasome (Hayashi, 1989; Hayashi and Canellakis, 1989; Hayashi and Murakami, 1995; Hayashi, *et al.*, 1996). Analyses of the corresponding area in PfODC indicate that there is only 20% sequence identity between the malarial and human ODC sequences. The α-carbon backbones of these enzymes are well conserved in this area (Fig. 5.7 B). However, the electrostatic potential at the surface of the area in PfODC corresponding to the AzBE is not as distinctly positively charged (Fig. 5.12). This is also not the case for the *T. brucei* ODC for which it was shown that antizyme does not bind the enzyme to mediate its degradation (Ghoda, *et al.*, 1990).
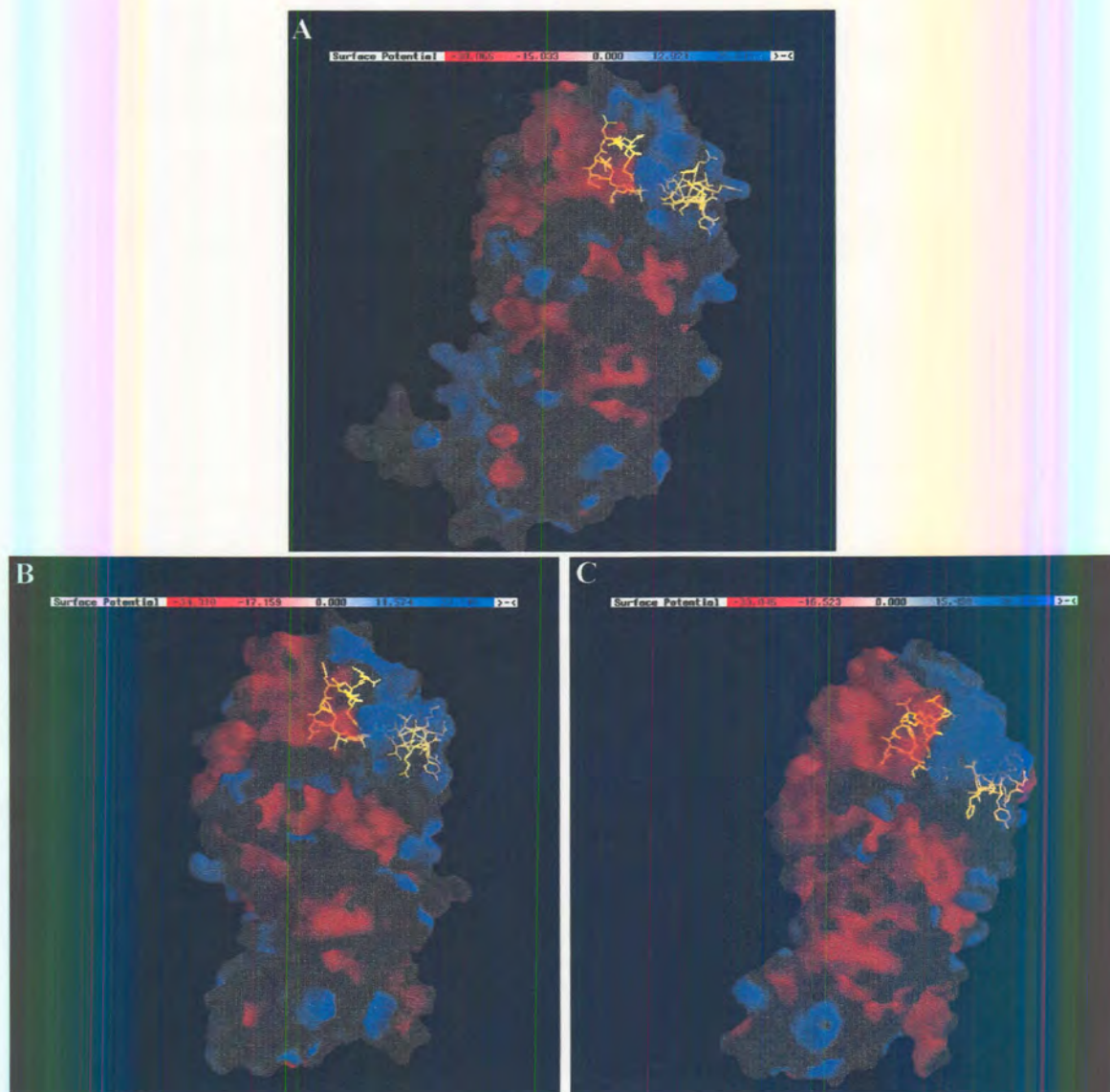
Figure 5.12: Electrostatic surface potentials for ODCs from *P. falciparum* (A), *H. sapiens* (B) and *T. brucei* (C) comparing potential antizyme binding elements. Residues forming the AzBE are indicated in yellow. Red indicates negatively charged surface areas and blue electrostatically positive areas.

5.3.9) Validation of the three-dimensional PfODC model with limited proteolysis.

Functional tests to evaluate the accuracy of the predicted PfODC model were used to delineate its predicted surface properties and domain organisation. Limited proteolysis relies on the exposure of solvent accessible areas that are selectively digested in order to reveal their organisation in the three-dimensional structure of the protein. The PfODC model was predicted to have several areas available to proteolytic splicing using the Nickpred Server. In Fig. 5.13, the model is coloured according to areas that are the most likely to be exposed to solvent and therefore be available to proteolysis (dark red areas) and areas buried in the core of the protein which is the least likely to undergo proteolysis (blue areas). This prediction shows that trypsin digestion is most likely to

151

occur in insert O_1 and secondly in a protease sensitive loop described in other eukaryotic ODCs (Osterman, *et al.*, 1995). Proteinase K digestion will most likely occur in the C-terminus or in insert O_1.
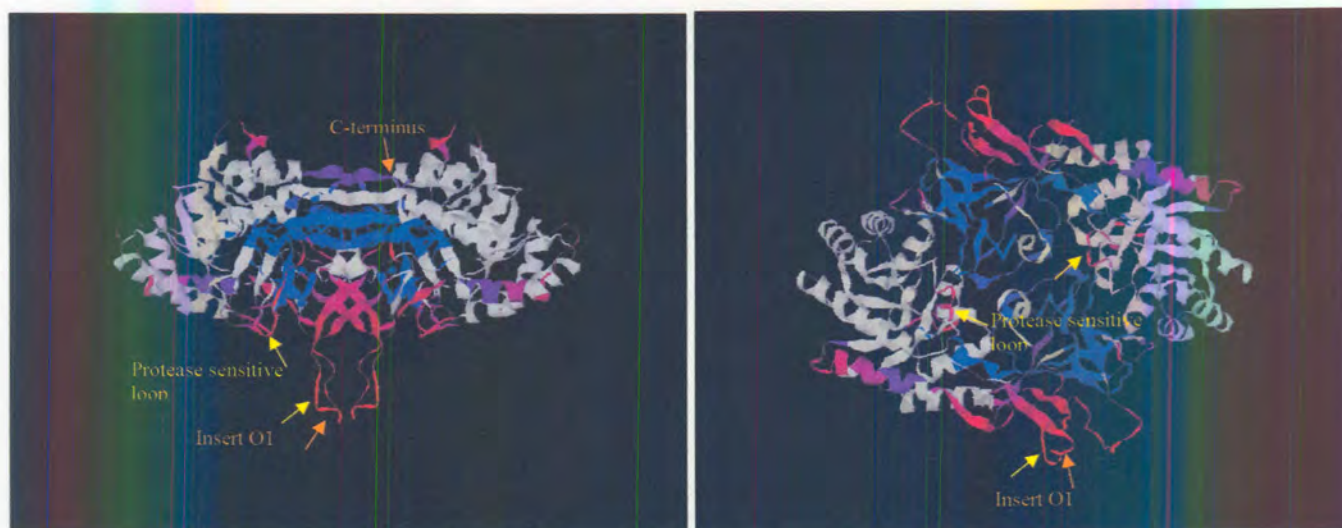
Figure 5.13: Nickpred prediction of proteolysis sites of dimeric PfODC. (A) Dimeric PfODC view from the side and in **(B)** from the top. Possible proteolytic sensitive sites are indicated in shades of red and buried areas possibly resistant to proteolysis in shades of blue. The orange arrows indicate highest scoring proteinase K prediction sites and in yellow sites for trypsin.

In order to confirm the PfODC model and the limited proteolysis predictions, the recombinantly expressed protein was subjected to diluted amounts of either trypsin or proteinase K for a short period of time. Digestion with proteinase K resulted in the wild type protein size of ~85 kDa decreasing to 82 kDa, with a 3 kDa size fragment removed (Fig. 5.14). This corresponds to the predicted cutting at the C-terminal end of PfODC where ~28 residues can be removed. Exposure of the protein to trypsin resulted in two fragments of ~45 and ~40 kDa in size. This indicates that the protein was probably cut in insert O_1. Validation of the identity of the obtained fragments was performed with peptide mass fingerprinting. However, the preliminary mass spectrometry data was inconclusive probably due to the low yield of the fragments after elution from SDS-PAGE gels.
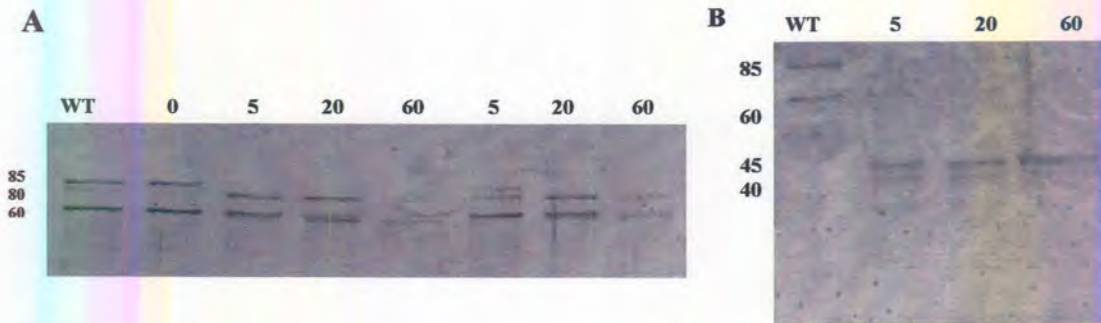
Figure 5.14: SDS-PAGE analyses of recombinantly expressed PfODC digested with either proteinase K (A) or trypsin (B). Molecular masses are indicated on the left of each figure in kDa. WT: undigested PfODC. The incubation times are indicated on top in minutes.

5.4) DISCUSSION.

The structural biology paradigm involves the determination of a protein structure to understand how the protein performs its known biological function at the molecular level (Thorton, *et al.*, 2000). Evolution has produced families of proteins whose members share the same three-dimensional architecture and frequently have detectably similar sequences. However, two structures may have very similar folds despite lacking any statistically significant sequence identity and it is now accepted that proteins having more than 30% of their sequences in common can be assumed to adopt the same folds (Jones, *et al.*, 1996). Protein structures, which are presumed to have diverged from a common ancestor in this way, are described as homologous (Jones, *et al.*, 1996). Analogous folds occur when the relationship between two structures are coincidental due to the physical limitation on protein folds. The conservation in protein folds allows a structural description of all proteins in a family even when only the structure of a single member is known (Sanchez, *et al.*, 2000). Because of the limited number of possible topologies of folds, it is a sensible approach to predict a structure by determining if a sequence could adopt one of the currently known set of protein folds (Jones, *et al.*, 1996).

5.4.1) Structural classification of PfODC.

Structural predictions are possible based on the intrinsic properties of the primary amino acid sequence of proteins. Analyses of the amino acids sequence of the PfODC with various databases and servers indicated that this decarboxylase groups into the expected family of PLP-dependent decarboxylases, with predicted α/β-barrel structure. PfODC is

grouped into the alanine racemase-like family of proteins based on sequence similarity. The proteins in this family are grouped into the TIM superfamily based on similar three-dimensional structures, which is in turn grouped into a specific protein fold, the α/β-barrel. This corresponds with the secondary structure predictions for the bifunctional PfAdoMetDC/ODC described in Chapter 3, where it was predicted that ODC contains a α-helix/β-sheet domain. All protein folds are grouped into classes of various combinations of the basic secondary structure elements of amino acids of α-helices and β-sheets. Homologous proteins sharing a known fold and having diverse functions shed light on divergent evolution. The α/β-barrel fold family is probably one of the best examples of this occurrence. The fold has been described in >20 protein families including proteins with diverse functions such as TIM, aldolase, flavocytochrome B2, tryptophan synthase, rubisco, enolase, glyoxylate oxidase and other multifold proteins including ornithine decarboxylase (Burley, 2000; Jones, et al., 1996). From the above it is apparent that characteristic structural features are conserved in evolution even between proteins with diverse primary sequences.

5.4.2) Comparative modelling of PfODC.

Homologous proteins that diverged from a common ancestor and have detectable sequence similarities (30%) commonly share folds and three-dimensional architecture. This served as the underlying principle for comparative protein structural modelling (Jones, et al., 1996; Srinivasan, et al., 1996). The first important aspect of comparative modelling is the evaluation of the predicted models since the quality of the model determines the information that can be extracted from it. The accuracy of the obtained model can be verified by determination of the correct fold of the protein, RMSD, good stereochemistry and the distribution of spatial features (Marti-Renom, et al., 2000). The results presented in this chapter indicate that the PfODC model is of a high quality that reflects the three-dimensional structure of the protein. The observed RMSD for the core of the PfODC model compared to the human ODC crystal structure compares favourably to results of the 3DCrunch project, which showed that 63% of sequences sharing 40–49% identity with their template, yield models that deviate by less than 3 Å from control X-ray crystallography structures (Guex, et al., 1999). The stereochemical parameters for the α-carbon backbone indicate that the core of the PfODC model closely resembles the human ODC crystal structure indicating a 3D-structure that resembles all known eukaryotic ODC structures characterised to date. As predicted by the protein fold family analyses, PfODC contains a α/β-barrel fold. Furthermore, it also

contains a modified Greek-key β-barrel fold. However, because of its function, it is grouped into the PLP-binding superfamily of proteins that shares the α/β-barrel fold. The accuracy of the predicted PfODC structure was experimentally shown with functional tests to delineate its predicted surface properties and domain organisation. PfODC presumably undergoes limited proteolysis in either the C-terminal 28 residues or in the smaller parasite-specific insert O_1 depending on the protease used (Fig. 5.13 and 5.14). However, these results need to be confirmed with conclusive peptide mass fingerprints. The predicted surface location of the parasite-specific insert O_1 suggests its preferential digestion and not in the protease sensitive loop as is the case for the *T. brucei* ODC which should have given rise to 34 and 50 kDa bands not observed with trypsin digestion (Fig. 5.13) (Osterman, *et al.*, 1995).

5.4.3) Structural modelling of parasite-specific inserts in PfODC.

Ab initio modelling was only possible for the smallest insert O_1 in PfODC based on the intrinsic properties of the amino acids in this area and libraries of preferred side chain conformations. The insert structure is predicted to be comprised of four anti-parallel β-sheets, corresponding to the secondary structure predictions for this insert described in Chapter 4. Since no interactions could be indicated in the model between this insert and the rest of the core structure in the PfODC model, no firm conclusions are possible with regard to the exact orientation of the bulk of the loop in space. However, the insert is bridged by flexible Gly residues ($Gly_{1036-1038}$ and Gly_{1083}), which could act as hinge regions and allow mobility of the insert. The results of Chapter 4 indicated that this area is indeed important for the activity of both decarboxylases and association between the individual domains in the bifunctional PfAdoMetDC/ODC. The analyses described here predict that this area is structured and might be flexible to mediate the proposed functions. This area might therefore directly influence the active site pockets or might act as a channel in the heterotetrameric PfAdoMetDC/ODC protein and allow the substrate and co-factor to enter the hidden active site pocket in homodimeric PfODC. Substantiating evidence from the PfODC model is that this insert lies on the same face of the protein that would allow entry to the active sites. Furthermore, as mentioned in Chapter 4, this insert might influence the activity through long range energetic coupling of residues distant from the active site pocket to the catalytic residues (Myers, *et al.*, 2001).

Removal of the large parasite-specific insert (O_2) and the hinge region was necessary to create the PfODC model. The absence of 3D structural data of these inserts pre-empt conclusions on the function of these areas. The junction region between the N- and C-terminal domains of eukaryotic ODCs (region 300-340 in the murine enzyme) varies in length from 40 residues for the mouse ODC to 115 residues for the closely related *E. coli* arginine decarboxylase enzyme (Osterman, *et al.*, 1995). The insert O_2 in PfODC occurs in the equivalent region suggesting a considerable tolerance for sequence length variations in this area and a probable species-specific property. Other studies have suggested that low-complexity regions within such inserts found in malaria proteins encode for non-globular domains that occur on the surface of proteins and are not involved in the functional folding of the proteins (Pizzi and Frontali, 2001). As indicated in Chapter 4, the insert also contains a large number of Asn and Asp residues and is unstructured and nonglobular. Investigations into the conformational characteristics of asparaginyl residues indicated a peculiar feature of its side chain to consist of a peptide plane mimic attached at its C^β atom. It is therefore the non-glycyl residue with the most potential to adopt a left-handed α-helix conformation in the (ϕ/ψ) plane. Asn is also known to prefer loops rather than structured α-helices or β-sheets (Srinivasan, *et al.*, 1994). Modelling of these areas specifically in *P. falciparum* proteins is currently not possible.

The core structure of the PfODC described here has large similarities in its α-carbon backbone with other eukaryotic ODC structures despite the absence of the large parasite-specific insert and hinge region and the presence of parasite-specific insert O_1. Homology modelling of the malarial DHFR and serine/threonine phosphatase suggested that the inserts in these proteins occur as loops pointing away from the surface of the protein, and are proposed to be separate from the catalytic site and not to affect the models in terms of active site investigations (Lemcke, *et al.*, 1999; Li and Baker, 1998). Deletion mutagenesis of the parasite-specific inserts (Chapter 4) indicated that the large insert in the ODC domain (insert O_2) is necessary for ODC activity possibly by inducing slight conformational changes in the active site centre. It is of interest to note that single point mutations on the interface of the *T. brucei* ODC homodimer resulted in significant decreases in enzyme activity (Myers, *et al.*, 2001). However, the modelling method forces the correct conformation for the active site pocket as is evidenced by the high similarity with the human and *T. brucei* structures (See next section). It is expected

that removal of the insert would have the greatest impact in the immediate vicinity from where it was removed and that the actual structure could differ from the model.

5.4.4) Structural properties of active dimeric PfODC.

Eukaryotic ODC is an obligate homodimeric enzyme with two active sites at the interface between the monomers. PfODC is predicted to dimerise in the same manner as other eukaryotes indicating the quality of the monomeric homology models in terms of the active centres obtained. The PfODC homodimer interface is characterised by several stabilising interactions at the dimer interface. Hydrophobic interactions comparable to those in the TbODC and human ODC structures are found. More importantly, electrostatic interactions and hydrogen bonds are proposed to stabilise the associating areas between the monomers surrounding the active site pockets. Several of these interactions were unique to the PfODC model. Experimental evidence to support the predicted interactions at the PfODC dimer interface was provided by site-directed point mutagenesis of Asp_{1356}, Asp_{1359} or Lys_{970} to alanine, which led to inactive PfODC (Wrenger, et al., 2001). The inability of these mutants to be catalytically active confirms the involvement of these amino acids in the dimerisation and therefore inherent activity of PfODC.

Alanine scanning mutagenesis of the *T. brucei* ODC dimer interface showed that none of the mutants caused significant weakening of the dimer interaction, suggesting that structural features contributing to dimerisation are distributed throughout the interface (Myers, et al., 2001). More importantly, all the mutations caused significant detrimental effects on enzyme activity possibly by long range energetic coupling of the interface residues to the active site. It was therefore proposed that subunit interactions in ODC are optimised for catalytic function and not for high-affinity subunit association (Myers, et al., 2001). Inhibitors that could dissociate the ODC dimer or bind in the interface and disrupt activity would have advantages over traditional active site-directed inhibitors as mentioned in Chapter 4. Since there are discrete regions predicted to be involved at the dimer interface in PfODC, this could have implications in the selective inhibition of ODC activity in the malaria parasite. However, as this activity is part of a bifunctional protein, it remains to be seen if effects predicted for the homodimeric monofunctional PfODC can be extrapolated to this protein in complex with PfAdoMetDC.

The homodimeric model of PfODC reveals two identical active sites formed at the dimer interface by contributions from both monomers. There is a large degree of residue conservation and similar spatial orientations between the active site pockets of PfODC and the ODCs from *T. brucei* and *H. sapiens*. Single substitution of conserved residues Lys_{868}, Cys_{1355} or Asp_{1356} with alanine abolished PfODC enzyme activity, providing experimental support for the role of these residues in the activity of PfODC through the predicted interactions with the substrate or co-factor as seen in the model and lends support to the accuracy of the predicted active site pocket in the model (Wrenger, *et al.*, 2001). However, parasite-specific residues (Thr_{933}, Asn_{1034}, Met_{967}, Arg_{1117} and Tyr_{966}) were identified in the PfODC model and predicted to be selectively involved in stabilising interactions with the co-factor and substrate. This allows further mapping of the PfODC active site pocket by point mutations to confirm its accuracy. Once this is confirmed, the design of parasite-specific inhibitors could be considered. Furthermore, the electrostatic potential of the PfODC active site pocket is significantly different to the corresponding human ODC. This could have positive implications to the selective entry of PfODC-specific inhibitors into the active site pocket of PfODC.

5.4.5) Potential role of antizyme in the regulation of PfODC based on structural properties.

Eukaryotic ODC is one of the most highly regulated enzymes described to date. It has an extremely short half-life and its degradation has been shown to be mediated by an inhibitory protein, antizyme. Antizyme binding not only inactivates ODC by preventing dimerisation but also induces conformational changes to expose a basal-degradation element in the C-terminus of the protein and targets it for degradation by the 26S proteasome in an ubiquitin independent manner (Almrud, *et al.*, 2000; Hayashi and Canellakis, 1989; Hayashi and Murakami, 1995; Hayashi, *et al.*, 1996; Heller, *et al.*, 1976). A distinct binding site for antizyme could not be shown in the PfODC model. Antizyme might therefore not be able to bind PfODC and regulate it in the same manner as in other organisms. The limited proteolysis of PfODC indicated that the C-terminus is exposed to solvent. However, unlike the human ODC, the PfODC does not contain a significant PEST region in this area (Chapter 3, Fig. 3.12)(Almrud, *et al.*, 2000). ODC usually has a rather long half-life in parasitic protozoa and antizyme mediated degradation of ODC has not been described for any of the parasitic protozoa as mentioned in Chapter 3.

Chapter 6 describes how the homology model for PfODC presented in this chapter is used to explain the experimental inhibition of the ODC component of PfAdoMetDC/ODC with various inhibitors and explores the rational design of novel ODC inhibitors as possible lead compounds for antimalarial chemotherapy.

CHAPTER 6

Structure-based ligand binding and discovery of novel inhibitors against *P. falciparum* ODC.

6.1) INTRODUCTION

Early observations on inhibition of spermidine synthesis in a mammalian system by known anti-tumour agents encouraged the development of other compounds potentially inhibitory to specific steps in the polyamine biosynthetic pathway (Cohen, 1998). Due to its role as a rate-limiting enzyme of polyamine biosynthesis, ODC is an obvious target for chemical intervention (Janne and Alhonen-Hongisto, 1989b). However, inhibition of the enzyme triggers a series of compensatory reactions that conserve the intracellular polyamine pools including increased uptake, release of polyamines from intracellular pools, stabilisation of ODC in the case of competitive, reversible inhibitors and the secondary induction of AdoMetDC activity (Janne, *et al.*, 1985). Therefore, AdoMetDC has to be considered as the natural second target for chemotherapy.

The ODC catalysed reaction is inhibited by structural analogues of the cofactor, substrate or product (Janne and Alhonen-Hongisto, 1989a). The earliest inhibitors were structural analogues of ornithine or putrescine. The most successful of these reversible, competitive ODC inhibitors were α-methylornithine and α-hydrazinoornithine. However, these compounds were not sufficiently active *in vivo* and α-hydrazinoornithine also non-specifically binds PLP (Janne and Alhonen-Hongisto, 1989a; Pegg, 1989b). Co-factor analogues included compounds capable of reacting with PLP, such as L-canaline (α-amino-γ-aminooxybutyric acid) that inhibit ODC but also many other PLP-binding enzymes. Synthetic analogues of the Schiff-base intermediate were designed in the hope that it would act as multisubstrate adduct inhibitors of ODC (Pegg, 1989b). One such adduct, *N*-(5'-phosphopyridoxal)ornithine was a potent inhibitor of ODC but again also inhibited other PLP-dependent enzymes (Janne and Alhonen-Hongisto, 1989a). Unphysiological diamines that act as homologues of putrescine included 1,3-diaminopropane and its hydroxylated derivative, 1,3-diamino-2-propanol. These diamines are potent inhibitors of ODC but, because of the structural

homology to putrescine, may take over some of the physiological functions of the natural polyamines (Janne and Alhonen-Hongisto, 1989a). Reversible competitive inhibitors of ODC activity included 1-amino-oxy-3-aminopropane and its derivatives with their aminooxy group being isosteric with the aminomethylene of putrescine (Cohen, 1998). These are very powerful inhibitors of ODC with IC_{50}s in the nanomolar range (Mett, *et al.*, 1993; Standek, *et al.*, 1992). Fig. 6.1 indicates the structures of some of the ODC inhibitors.
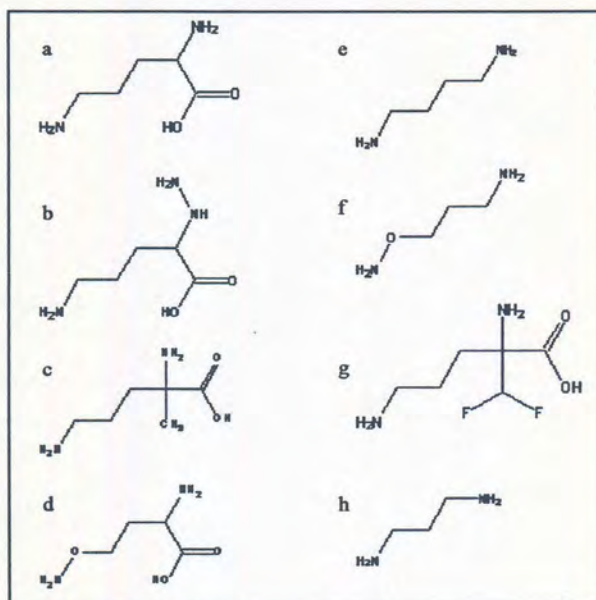
Figure 6.1: Structures of the natural substrates and reversible and irreversible inhibitors of ODC. A: ornithine, b: α-hydrazino ornithine, c: α-methyl ornithine, d: canaline, e: putrescine, f: 1-aminooxy-3-aminopropane, g: DFMO and h: 1,3-diaminopropane. Adapted from (Cohen, 1998).

It was not until the introduction of the concept of mechanism-based irreversible inhibitors, or suicide inhibitors, that more powerful and specific inhibitors of ODC were designed. Mechanism-based enzyme inactivators are unreactive compounds that bear a structural similarity to a substrate or product of a specific enzyme. Once at the active site the target enzyme converts the inactivator into a product that usually forms a covalent bond with the enzyme, generally via its normal catalytic mechanism (Silverman, 1988). The most potent of these, DL-α-difluoromethyl ornithine (DFMO), described in 1978 mimics ornithine but remains covalently bound to ODC after activation (Cohen, 1998), Fig. 6.2). A compound capable of competing favourably with DFMO is (2R,5R)-6-heptyne-2,5-diamine (or methylacetylenicputrescine, RR-MAP) which is at least ten times more potent than DFMO (Cohen, 1998; Janne and Alhonen-Hongisto, 1989a).
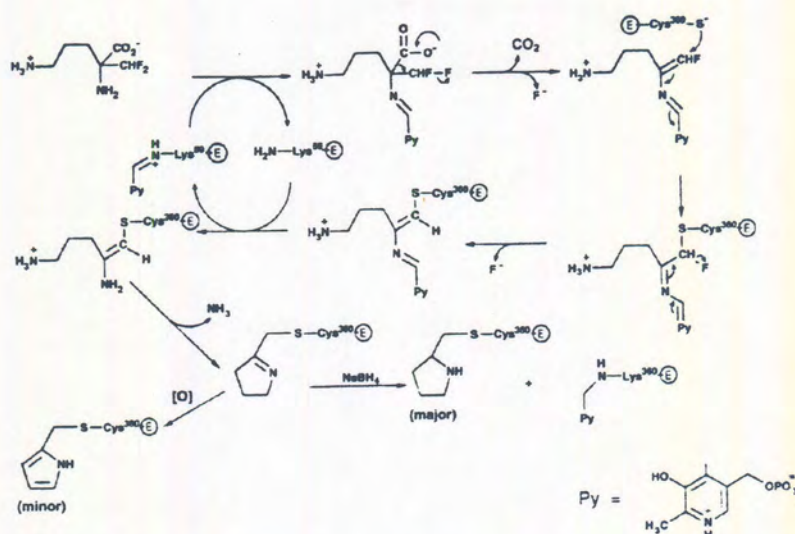
Figure 6.2: Proposed mechanism of inactivation of ODC with DFMO. Adapted from (Cohen, 1998).

The only known inhibitors of AdoMetDC of any physiological importance are derivatives of bis(guanylhydrazone) (Janne, *et al.*, 1985). The discovery that the antileukemic agent, methylglyoxal bis(guanylhydrazone) (MGBG) is an extremely potent inhibitor of eukaryotic, putrescine-activated AdoMetDC singled this compound out as a standard inhibitor for this enzyme (Janne and Alhonen-Hongisto, 1989b; Janne, *et al.*, 1985; Pegg, 1989b). Moreover, most of the bis(guanylhydrazone) derivatives, when combined with inhibitors of ODC produce a synergistic antiproliferative effect. Further alkylation at the glyoxal portion of these molecules produces even more potent inhibitors such as ethylmethylglyoxal bis(guanylhydrazone) (Cohen, 1998). Fig. 6.3 indicates the structural similarity between MGBG and spermidine as well as the substrate, adenosylmethionine (Janne, *et al.*, 1985).
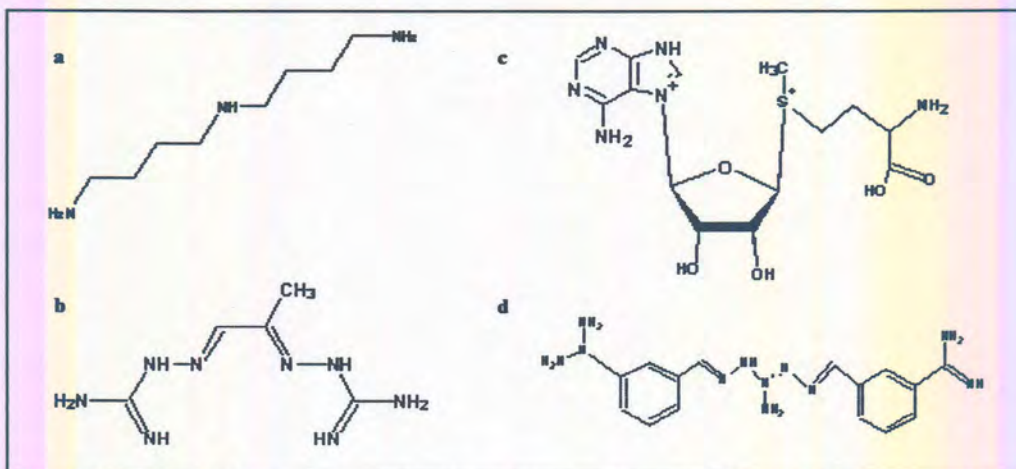
Figure 6.3: Structural similarities between spermidine, MGBG and adenosylmethionine. Adapted from (Janne, *et al.*, 1985; Müller, *et al.*, 2001). a: Spermidine, b: MGBG, c: AdoMet and d: CGP40215.

A number of other non-nucleoside substances have been described as inhibitors of AdoMetDC. Some of these, including Berenil and Pentamidine, resemble MGBG but their pharmacological actions are not well understood (Cohen, 1998; Pegg, 1989b). Other nucleoside analogues include the irreversible inhibitors 5'-[(3-aminooxypropyl)methylamino]-5'-deoxyadenosine (MAOPA) which are extremely potent inhibitors of mammalian AdoMetDC (Cohen, 1998; Pegg, 1989b).

Various ODC and AdoMetDC inhibitors have been tested for antimalarial activity. Five ODC inhibitors, DFMO (> 3 mM) and three of its α-monofluoromethyl derivatives as well as RR-MAP have been shown to inhibit *P. falciparum* schizogony *in vitro*. DFMO was also effective in limiting *P. berghei* schizogony *in vivo* (Assaraf, *et al.*, 1987a; Assaraf, *et al.*, 1986; Assaraf, *et al.*, 1987b; Bitoni, *et al.*, 1987; Whaun and Brown, 1985). Two 3-amino-oxy-1-propanamine analogues, CGP52622A and CGP54169A, were also shown to inhibit malarial ODC with K_i values in the nanomolar range (Krause, *et al.*, 2000; Wrenger, *et al.*, 2001). Irreversible inhibition of *P. falciparum* AdoMetDC with MGBG analogues prevented growth of the parasite *in vitro* (Wright, *et al.*, 1991). The most potent of these, MDL 73811, inhibited growth with an IC_{50} in the low μM range. The diamidine, CGP40215 synthesised as an anticancer compound, showed inhibition against both recombinant *P. falciparum* AdoMetDC and *P. berghei* infection in mice (Müller, *et al.*, 2001).

The relatively poor pharmacokinetics of inhibition of polyamine biosynthesis of *in vivo P. falciparum* infections prompts investigations to develop novel inhibitors. Structure-

based approaches for designing novel inhibitors of protein function have been used to generate clinically useful drugs against a number of diseases. These include antibacterial agents (DHFR as target), anticoagulants (thrombin inhibition), antiviral agents (neuramidase inhibition), anti-AIDS agents (HIV protease inhibition), anticancer agents (thymidylate synthase and purine nucleoside phosphorylase as targets), and an anti-inflammatory agent (phospholipase A_2) (Blundell, 1996; Bohm and Klebe, 1996; Whittle and Blundell, 1994). Specifically focussing on parasitic protozoa, success has been obtained in the discovery of novel lead inhibitors against trypanosomal glyceraldehyde-3-phosphate dehydrogenase (Bohm and Klebe, 1996) and the antimalarial targets, DHFR (Lemcke, *et al.*, 1999; Toyoda, *et al.*, 1997; Warhurst, 1998) and two major protease families, the serine and cysteine proteases using this approach (Ring, *et al.*, 1993).

Structure-based or rational inhibitor design methods identify favourable and unfavourable interactions between a potential ligand and the target receptor and maximise the beneficial interactions to increase binding affinity (Blundell, 1996; Whittle and Blundell, 1994). The central assumption is that good inhibitors must possess significant structural and chemical complementarity to their target receptor (Kuntz, 1992). Molecular recognition in ligand-protein complexes is responsible for the selective binding. Interactions are normally noncovalent in nature and experimentally determined inhibition constants K_i are typically in the range of 10^{-2} to 10^{-12} M (Bohm and Klebe, 1996). Hydrogen bonds are the most important stabilising force in protein interactions and define the specificity of protein-ligand interactions. Furthermore, oppositely charged functional groups of the protein and ligand are paired in electrostatic interactions. In addition, apolar groups of ligands are found in hydrophobic pockets in the active sites (Bohm and Klebe, 1996; Zubay, 1993). Structure-based approaches employ either novel ligand discovery by searching three-dimensional databases of known chemical structures (docking), modifications of existing inhibitors (linking) or attempts to *de novo* assemble compound structures from chemical fragments (building) as summarised in Fig. 6.4.
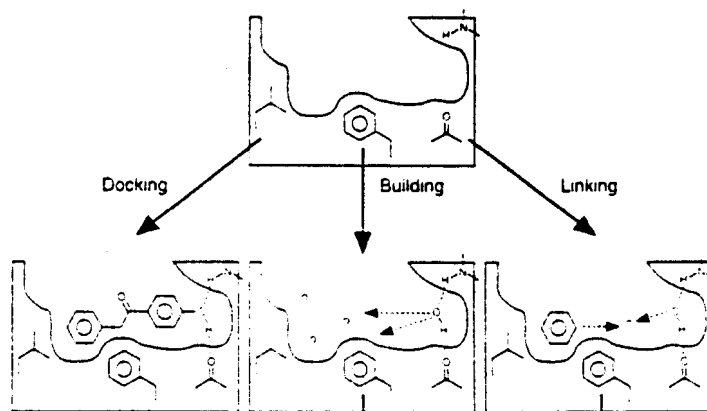
Figure 6.4: Strategies for the discovery of novel lead structures by ligand docking. Adapted from (Bohm and Klebe, 1996).

Computational methods for structure-based ligand discovery require a three-dimensional structure of the target protein. In the absence of an accurate X-ray or NMR structure, a homology model of the protein can be used if the errors of the prediction cluster away from the enzyme active site (Blundell, 1996; Kuntz, 1992; Ring, *et al.*, 1993; Walters, *et al.*, 1998). The most direct way to find a novel ligand is then to search a database of three-dimensional chemical structures. The pioneering program DOCK searches these databases based on shape and electronic complementarity (by means of a molecular mechanics force field) between the protein and ligands (Kuntz, 1992; Whittle and Blundell, 1994). The Cambridge Structure Database (CSD) contains experimentally determined crystal structures whereas other databases such as the Fine Chemicals Directory or Available Chemical Directory (ACD) can be converted to three-dimensional databases with programs such as Corina or Concord (Kuntz, 1992; Whittle and Blundell, 1994).

LUDI is a fragment-based *de novo* design program that can be used both for searching of three-dimensional databases and for the automatic construction of novel ligands (Bohm and Klebe, 1996). The program positions molecules into clefts of a protein such that hydrogen bonds can be formed and hydrophobic pockets are filled with apolar groups. The program calculates interaction sites derived from composite crystal-field environments. An important aspect of LUDI is its ability to tolerate small uncertainties in the determined protein structure. Any positioning of ligands is completely based on geometric parameters. A fast and error-tolerant empirical scoring function is used based on the assumption that noncovalent interactions are additive and the function then takes into account hydrogen bonds, ionic interactions, apolar protein-ligand contact surface,

interactions between aromatic rings, the replacement of water molecules and the number of rotatable bonds in the ligand. The capacity of the LUDI programme to discover novel ligands is supported by the identification of a novel inhibitor of human phospholipase A2, compound 23 for the immunosuppressant FK506 binding protein-12, inhibitors of trypsin, HIV-protease, purine nucleoside phosphorylase and streptavidin (Bohm and Klebe, 1996)

In this chapter, the active site of PfODC revealed by the homology model (Chapter 5) is used to explain the experimental inhibition of PfODC seen with DFMO and the CGP-series of inhibitors described above. Furthermore, novel ligands that selectively bind to PfODC are identified in a structure-based approach.

Results obtained in this chapter have been presented as a talk at the second polyamine conference: 'Polyamine metabolism as a drug target in parasitic protozoa and worms', Texas, USA (Birkholtz, 2002c).

6.2) MATERIALS AND METHODS

6.2.1) Docking of known inhibitors into the active site of dimeric PfODC

6.2.1.1) Generation of three-dimensional structures for known inhibitors.

Structures for PLP and DFMO (used as substrate analogue) were generated with the Builder module of the InsightII package and subjected to energy minimization with the Discover3 module of the InsightII package (cff91 force field for 10 000 iterations with a conjugate gradient) as described in Chapter 5. Structures for the two competitive inhibitors CGP52622A and CGP54169A from Novartis Pharma (Krause, *et al.*, 2000; Mett, *et al.*, 1993; Standek, *et al.*, 1992) were generated as described above.

6.2.1.2) Docking of the known inhibitors into the active site pocket of PfODC.

Binding of PLP and DFMO requires the formation of a Schiff-base between the two ligands with DFMO then also forming a covalent bond to the S_γ atom of Cys_{360} (*T. brucei* numbering). In order to dock this transition state complex of PLP-DFMO into PfODC, the structure for the linked PLP-DFMO was created and formed a covalent link with Cys_{1355} of the dimeric form of PfODC described in Chapter 5, section 5.2.3. The

ligand-ODC complexes were then minimized as described above. Possible interactions between the ligands and residues in PfODC were analysed with LigPlot (Wallace, *et al.*, 1995). The structures were analysed for accuracy with the WHAT IF program (Vriend, 1990).

The two competitive inhibitors, CGP52622A and CGP54169A (Novartis Pharma), are 3-amino-oxy-1-propanamine homologues of the polyamines (Krause, *et al.*, 2000; Mett, *et al.*, 1993; Standek, *et al.*, 1992). To analyse the possible interactions of these inhibitors in the active site pocket of PfODC, the structures were superimposed on the coordinates of the PLP-ornithine complex minimized in the PfODC active site pocket (Chapter 5, section 5.2.4). Energy minimization was repeated for the ligand-ODC complexes as described above. The ligands were also analysed for any bumps occurring between van der Waals radii. Possible interactions between the ligands and residues in ODC were analysed with LigPlot (Wallace, *et al.*, 1995). The structures were analysed for accuracy with the WHAT IF program (Vriend, 1990). The same procedure was followed for docking of these inhibitors in the human ODC structure to explain the selective inhibition.

6.2.2) Discovery of novel ligands for PfODC.

The three-dimensional active site pocket of the PfODC model was defined in Chapter 5, section 5.3.6. The coordinates of the defined active site pocket was used for the identification of novel ligands that could selectively bind residues involved in catalysis of PfODC. The active site was specified for the forced inclusion of residues Cys_{1355}, Asp_{1356}, Arg_{1117}, Tyr_{966}, Arg_{955} and His_{998}. The receptor atoms specified were Cys_{1355} and Lys_{868} to allow the generation of a virtual atom as the central point of the search domain. The LUDI module of the Insight II package was then used to screen a three-dimensional chemical structure library to identify novel ligands that bind to PfODC. LUDI suggests novel ligands by selecting from a library of small molecules due to the appropriate spatial orientation of hydrogen-bonding and hydrophobic contacts of the molecules with the target receptor. The LUDI parameters included a 7 Å search domain around the specified centre atom with cut-off limits on the size of the identified ligands as maximum 50 atoms and minimum 5. The USA National Cancer Institute (NCI) Open database of chemical compounds was screened. These compounds are all publicly and freely available from NCIs' Developmental Therapeutics Program and were generated originally as compounds with anticancer potential. The three-dimensional version of the

database was generated with Corina v 1.7 (Gasteiger, *et al.*, 1990) and included structures for 249 017 compounds.

The NCI Open database was screened against PfODC using LUDI. The same procedure was performed for the human ODC active site. Predicted K_i values for the compounds were obtained with the formula:

Score = -100log K_i where score is the LUDI score given for binding of each compound (Bohm and Klebe, 1996).

Compound information was obtained from the Enhanced NCI database browser. Interaction maps of the compounds with either PfODC or the human structure were obtained with LigPlot (Wallace, *et al.*, 1995). PASS predictions (Prediction of the biological activity spectra of substances) were performed for the top ten compounds identified (www.ibmh.msk.su/PASS).

6.3) RESULTS.

6.3.1) Docking of known inhibitors in the active site pocket of PfODC.

In order to explain the mechanism of competitive inhibition of PfODC (either on its own or in the bifunctional enzyme complex) by DFMO and the CGP-series of inhibitors, these compounds were docked into the proposed active site pocket.

6.3.1.1) Docking of DFMO in PfODC.

In order to simulate the binding of the irreversible competitive inhibitor DFMO, it was covalently linked to Cys_{1355} (atom S_γ) as well as coupled to PLP via a Schiff-base and minimized to an energy of $-15\ 307$ kcal/mol. Analyses of the minimized interacting site of PfODC with DFMO are shown in Fig. 6.5. The majority of the residues are also involved in binding to the substrate with the exception of Asn_{1393} and Phe_{1392} (Chapter 5, section 5.3.6). The only PfODC-specific residues involved in the binding of DFMO are Arg_{1117} and Tyr_{966} compared to the residues involved in the human enzyme. Both of these residues make contact with the ε-amino group of DFMO.
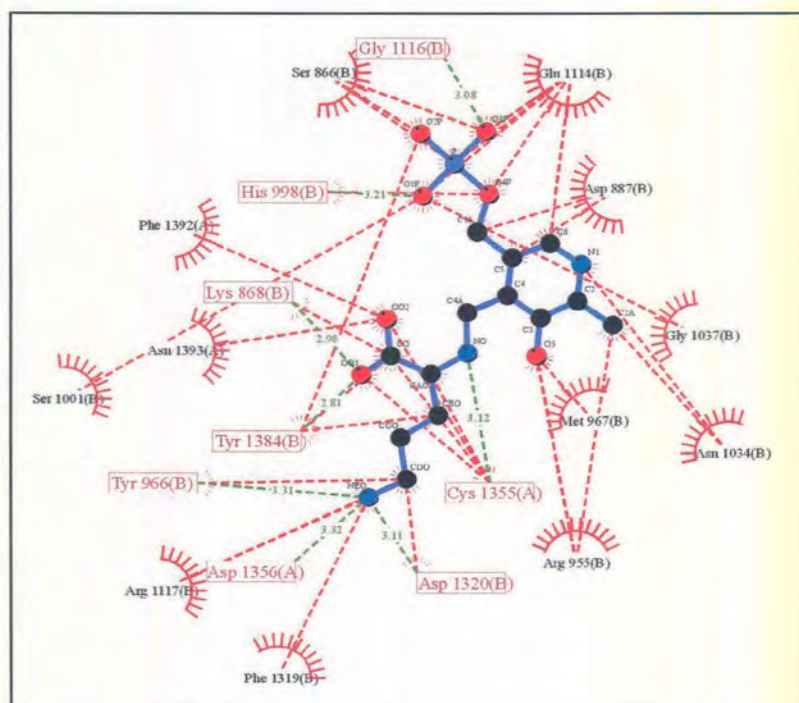
Figure 6.5: Interactions between the cofactor (PLP) and competitive inhibitor DFMO in the active site pocket of PfODC. Hydrogen bonds and their lengths are indicated in green, non-ligand residues involved in hydrophobic contacts with the ligand in red arcs with spokes.

6.3.1.2) Docking of the reversible competitive CGP inhibitors in PfODC.

To analyse the possible interactions of two reversible competitive inhibitors, CGP54169A and CGP52622A in the active site pocket of PfODC, the structures were superimposed on the coordinates of the corresponding ligands minimized in the PfODC active site pocket. Stable complexes had minimised energies of −15 343.94 kcal/mol for the CGP52622A-PfODC complex and a corresponding −15 346.68 kcal/mol for the CGP54169A-PfODC complex. Analogous procedures were followed for docking of the CGP-inhibitors against the human ODC structure and stable complexes showed minimised energies of −14 821 kcal/mol (CGP52622A-human ODC complex) and −14 807.75 kcal/mol for the CGP54169A-human ODC complex. This predicts a slightly lower ability of the human ODC active site pocket to accommodate these compounds.

Fig. 6.6 indicates the predicted Ligplot interaction sites between the CGP-inhibitors with PfODC. In the case of CGP52622A, inhibition can be explained due to the ability of this molecule to interact with essential residues involved in binding of PLP (Asp$_{887}$, Arg$_{955}$, Thr$_{993}$, Ser$_{1001}$, Asn$_{1034}$ and Cys$_{1355}$,) as well as substrate (Lys$_{868}$, Tyr$_{1384}$ and Cys$_{1355}$). Two unique interactions with Ala$_{912}$ and Phe$_{997}$ are also suggested. The calculated total free energy of interaction of this inhibitor with PfODC was −32.1

kcal/mol. As expected for an substrate analogue, CGP54169A shows major interactions with residues predicted to bind to ornithine (Arg_{1117}, Tyr_{966}, Phe_{1319}, Asp_{1320}, Asp_{1356}, Cys_{1355} and Phe_{1392}) and also to two PLP-binding residues (His_{998} and Tyr_{1384}) with a total free energy of interaction of −27.67 kcal/mol. This is lower than the binding free energy of −84.2 kcal/mol obtained for the interactions with Schiff-base linked PLP-ornithine as natural ligands. These proposed interactions of the inhibitors with the essential active site binding residues would therefore prevent or interfere with any subsequent binding of the substrate/co-factor and lead to enzyme inhibition.
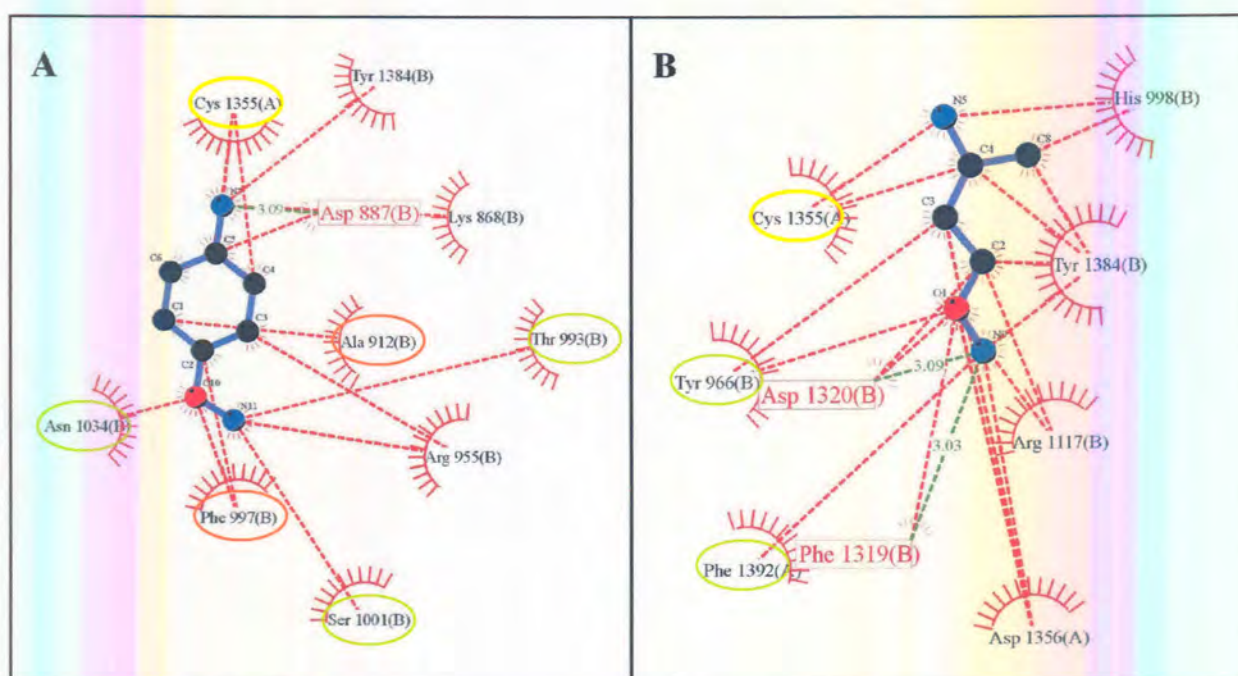
Figure 6.6: Ligplot analyses of the interactions between two competitive inhibitors and PfODC. (A) CGP52622A and (B) CGP54169A. Hydrogen bonds and their lengths are indicated in green, non-ligand residues involved in hydrophobic contacts with the ligand in red arcs with spokes. Green and yellow ellipses indicated residues in PfODC that were uniquely involved in binding to this enzyme. The catalytic Cys_{1355} is indicated in the yellow ellipse. The unique residues involved in CGP52622A binding but not in substrate or co-factor binding are shown in orange ellipses.

The CGP-series of compounds inhibit rat liver ODC with IC_{50} values in the nanomolar range and recombinantly expressed human ODC with IC_{50} values of 25 and 10 nM for CGP52622A and CGP54169A, respectively (R. Walter, personal communication). Stable complexes were minimized to energies of −14 821 kcal/mol for the CGP52622A-human ODC complex and −14 807.75 kcal/mol for the CGP54169A-human ODC complex. The predicted interaction sites of CGP52622A with the human ODC did not include Thr_{933}, Ser_{1001}, Asn_{1034}, Cys_{1355}, Tyr_{1384} and Ala_{912} (PfODC numbering) shown to be involved in its interaction with PfODC (Fig. 6.6). The total calculated free energy
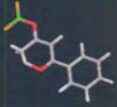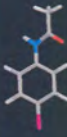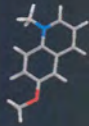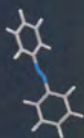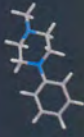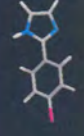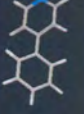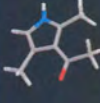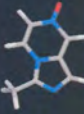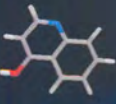
of the interaction between CGP52622A and the human ODC was −27.07 kcal/mol. Furthermore, CGP54629A-human ODC interactions also excluded residues Tyr_{966}, Cys_{1355} and Phe_{1392} of PfODC (Fig. 6.6) with a calculated free energy for the complex of −26.6 kcal/mol. Again, the interaction of the natural ligands, PLP and ornithine with human ODC has a somewhat higher free binding energy of −71.9 kcal/mol. Importantly, Cys_{1355} is essential to ODC activity and both inhibitors bind this residue in PfODC but not the equivalent residue of the human enzyme. Furthermore, Thr_{933}, Phe_{997} and Tyr_{966} uniquely interact with the inhibitors only in the malarial enzyme.

6.3.2) Discovery of novel ligands for PfODC.

The NCI Open database of three-dimensional chemical structures was screened against the PfODC structure for the identification of novel ligands that bind selectively against only the malarial ODC. Approximately 96724 compounds were screened and 694 potential hits were found. The top ten scoring compounds were analysed. Table 6.1 summarise the ten compounds identified as potential ligands or lead inhibitors against PfODC. The compounds are all iso- or heterocyclic compounds. Predicted K_i values were in the low μM range for all ten compounds with the top scoring ligand having a K_i of 9.33 μM. PASS predictions of the biological activity spectra of the compound detected antiprotozoal activity for 9 of the 10 compounds. Six of these have predicted antimalarial activity.

Table 6.1: Summary of the identified novel ligands for PfO ligands identified are listed in order of highest to lowest scores. Biological activity spectra (PASS prediction) are given in order of activity.

Compound	Structure	LUDI Score	K_i (µM)	PASS prediction
1-methyl-3-oxo-3-phenyl-1-propenyl difluoridoborate		503	9.33	Antimalarial Antitrypanosomal
4-chloroacetanilide		488	13.2	Antimalarial Antiprotozoal
6-methoxy-1-methyl-1-quinoline		483	14.8	Antiprotozoal Antimalarial
Azobenzene (1,2-diphenyldiazene)		467	17.4	Antitryponosomal Antiprotozoal
1-methyl-4-phenylpiperazine		465	22.4	Antitrypanosomal
2-(4-chlorophenyl)-1H-imidazole		446	34.7	Antiprotozoal Antimalarial Antitrypanosomal
4-phenylpyridine		419	64.5	Antimalarial Antitrypanosomal Antiprotozoal
3-acetyl-2,4-dimethylpyrrole		419	64.5	None
3-methyl-7-imidaso[1,5-a]pyrazin-7-ol		417	67.6	Antiprotozoal
4-quinolinol		416	69.2	Antiprotozoal Antimalarial

LigPlot analysis of the interacting residues of the top scoring compound, 1-methyl-3-oxo-3-phenyl-1-propenyl difluoridoborate (mopp-DFB), indicated that the predicted binding of mopp-DFB to PfODC is mediated by residues involved in binding of the PLP co-factor or ornithine substrate (Fig. 6.7). Residues Ser_{1001}, Asp_{887}, Cys_{1355} and Tyr_{1384} also show interactions with CGP52622A while the majority of residues (Cys_{1355}, Tyr_{1384}, Arp_{1117}, Tyr_{966}, Asp_{1356} and His_{998}) were also involved in interactions with CGP54169A. Two of the five residues shown to be present exclusively in the active site pocket of PfODC have interactions with the novel compound. Tyr_{966} is involved in hydrophobic contacts whereas Arg_{1117} is hydrogen-bonded to the compound. Importantly, the catalytic Cys_{1355} is again involved in interactions with the novel ligand.
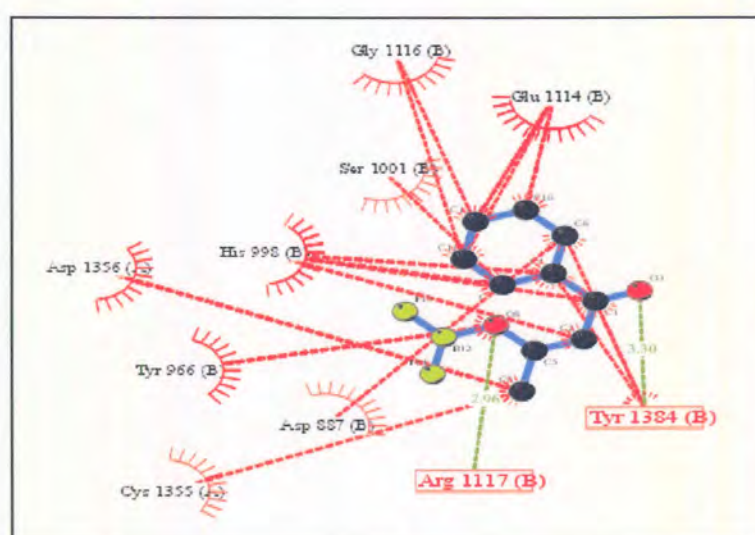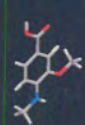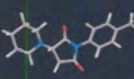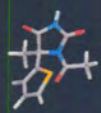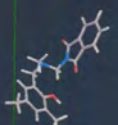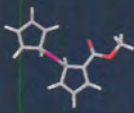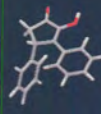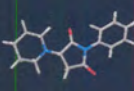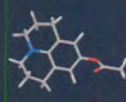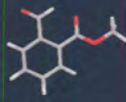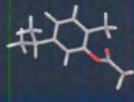
Figure 6.7: Interactions between the top scoring novel ligand and PfODC. Compound 1-methyl-3-oxo-3-phenyl-1-propenyl difluoridoborate was identified with LUDI and interactions determined with LigPlot. Hydrogen bonds and their lengths are indicated in green, non-ligand residues involved in hydrophobic contacts with the ligand in red arcs with spokes.

The specificity of the novel compounds identified as possible ligands of PfODC was validated by screening the identical NCI Open database of three-dimensional chemical structures against the human ODC structure. 1951 potential hits were found and the top ten scoring compounds are listed in Table 6.2. The predicted K_i values were lower than those found against the PfODC (between 2 and 10 μM). None of the top scoring hits identified by screening against PfODC was included in the results obtained with the human enzyme. This indicates sufficient differences between the active site pockets to allow structure-based discovery of novel ligands for PfODC.

Table 6.2: Summary of the comparative ligands of the human ODC. The top ten scoring ligands identified are listed in order of highest to lowest scores.

Compound	Structure	LUDI Score	K_i (µM)
3-methoxy-4-(methylamino)benzoic acid		564	2.29
1-(4-methylphennyl)-3-(1-piperidinyl)-2,5-pyrrolidinedione		552	3.02
1-actyl-5-methyl-5-(2-theinyl)-2,4-imidazolidinedione		551	3.09
2-(2-hydroxy-3,5-dimethylbenzyl-methyl-amino-methyl)-1H-isoindole-1,3(2H)-dione		545	3.55
2,4-cyclopentadien-1-yl(2-(methoxycarbonyl)-2,4-cyclopentadien-1-yl)iron		544	3.63
2-hydroxy-3,4-diphenyl-2-cyclopenten-1-one		530	5.01
1-phenyl-3-(1-piperidinyl)-2,5-pyrrolidinedione		502	9.55
2,3,6,7-tetrahydro-1H,5H-pyrido[3,2,1]quinolin-9-yl acetate		502	9.55
methyl 2-formylbenzoate		497	10.7
5-isopropyl-2-methylphenyl acetate		496	10.96

6.4) DISCUSSION.

6.4.1) Structural explanations for the inhibition of PfODC with known inhibitors.

The homology model created for PfODC (Chapter 5) allowed the determination of the interaction sites of DFMO and the CGP-compounds with the malarial enzyme to explain the experimental inhibition of activity. Inhibition of PfODC activity by DFMO has a K_i of 87.6±14.3 μM, almost double that of the murine enzyme (Krause, *et al.*, 2000). The majority of the residues that make contact with DFMO are usually involved in interactions with the substrate but two residues of PfODC were identified to be uniquely involved only in binding to DFMO and they are predicted to make the same contacts with DFMO as they make with ornithine. This information is useful in further mapping and validation of the PfODC active site pocket. DFMO does not show any selectivity towards experimental inhibition of ODCs from various origins (Cohen, 1998). The small differences in predicted DFMO-binding sites between PfODC and the human enzyme provide a partial explanation for the non-discriminatory inhibition. The K_i of *T. brucei* ODC for DFMO is 220 μM and can in part be explained by its long half-life (>6 hrs compared to 5-35 min of mammalian ODCs)(Ghoda, *et al.*, 1990). The slightly higher K_i of the PfODC can also be explained because this protein was also reported to be relatively stable (>2 hrs vs. 5-35 min)(Wrenger, *et al.*, 2001).

A series of potent ODC inhibitors were synthesized as analogues of 3-amino-oxy-1-propanamine (Mett, *et al.*, 1993; Standek, *et al.*, 1992). CGP52622A and CGP54169A, members of this series, were reported to inhibit rat liver ODC with an IC_{50} in the nanomolar range and are active against the growth of human T_{24} bladder carcinoma cells *in vitro* (Standek, *et al.*, 1992). They furthermore inhibit recombinantly expressed human ODC with IC_{50} values of 25 and 10 nM for CGP52622A and CGP54169A, respectively (R. Walter, personal communication) and inhibit PfODC with respective IC_{50} values of 63.5 nM and 25 nM (Krause, *et al.*, 2000). The results of Fig. 6.6 predict interaction of these compounds with essential residues for catalysis of PfODC, specifically with the catalytic Cys_{1355}, also proposed to be involved in the covalent interaction of DFMO with PfODC. The majority of the interactions were with residues normally involved in predicted interactions with the substrate (Table 5.2). This is in agreement with the original results of the activity of the CGP-compounds that indicated these inhibitors to be substrate-competitive inhibitors of mammalian ODC and do not interact non-specifically with PLP (Standek, *et al.*, 1992).

The proposed interactions of the CGP-compounds with the essential active site binding residues of PfODC would therefore prevent or interfere with any subsequent binding of the substrate/co-factor and lead to marked enzyme inhibition. Interestingly, two PfODC residues not previously implicated in ligand binding (Ala_{912} and Phe_{997}) were identified to interact with CGP52622A.

The higher free energies predicted for both CGP-human complexes compared to the energies predicted for the CGP-PfODC complexes suggest that the human ODC does not bind the CGP-inhibitors with the same affinity as PfODC. Importantly, Cys_{1355} is essential to ODC activity and both inhibitors bind this residue in PfODC but not the equivalent residue of the human enzyme. Furthermore, seven residues interacting with the PfODC does not have any corresponding residues in the CGP-human complexes. Of these parasite-specific residues Thr_{933}, Asn_{1034} and Tyr_{966}, were shown to be uniquely present in the active site pocket of PfODC (Chapter 5, Table 5.2). Taken together, the marked inhibition of PfODC seen with the CGP-compounds can be explained by their interaction with parasite-specific residues, including the essential Cys_{1355} with a higher affinity than the human ODC. Furthermore, the structures and interaction maps of these compounds could facilitate the design of even more potent, selective PfODC inhibitors in a rational drug design strategy (Bohm and Klebe, 1996).

6.4.2) Identification of novel compounds that selectively bind PfODC.

Approaches using modelled structures of malarial proteins provided promising results in other studies. Two compounds, triazinobenzimidazole and a pyridoindole were identified with experimentally determined Ki values in the low μM range (0.5-8 μM) against of DHFR (Toyoda, et al., 1997; Warhurst, 1998). Novel lead inhibitors of the malarial cysteine proteases including bis[(2-hydroxy-1-naphthylmethylene)hydrazide] were identified with experimental IC_{50} values of ~6 μM (Ring, et al., 1993). None of these compounds have been previously described as antimalarial agents and these compounds could act as scaffolds that can be modified to develop more potent inhibitors. The results obtained using the modelled active site pocket (and the predicted overall features of the substrate-binding site) of PfODC to explain experimental results obtained with various inhibitors lead to the further identification of novel and potential selective inhibitors of the enzyme by screening chemical structure libraries *in silico*.

The computational screening described in this chapter used the NCI Open database of chemical structures instead of more comprehensive sources of ligand structures like ACD for two reasons. Firstly, the former database contains compounds that were identified during drug development studies against cancer and includes comprehensive biological data on many of these compounds. It is therefore a chemical library that is focussed on a family of related targets rather than a targeted (only one single target) or more general library (broad interest targets). Since ODC inhibition has received a lot of attention in the antiproliferative field, using this database might provide a primary selection of ODC-specific ligands. Secondly, the database and all the compounds are publicly and freely available, making the synthesis of the compounds unnecessary if it is not commercially available.

The top ten scoring compounds obtained with LUDI (Table 6.1) were all iso- or heterocyclic compounds and were predicted to inhibit the enzyme with K_i values <100 µM (ranging between 9.33 and 69.2 µM). The two best compounds, 1-methyl-3-oxo-3-phenyl-1-propenyl difluoridoborate (abbreviated mopp-DFB) and 4-chloroacetanilide, were isocyclic. Analyses of databases of validated drugs revealed that only 32 frameworks are commonly found in drugs. These are all iso- or heterocyclic structures with benzene ring derivatives present in 606 out of 2548 drugs tested (Walters, et al., 1998).

Interaction maps of the top scoring compound, mopp-DFB, indicated that this compound interact exclusively with residues involved with either co-factor or substrate binding in PfODC (Fig. 6.7). The specificity of binding of mopp-DFB is defined by two hydrogen bonds with Arg_{1117} and Tyr_{1384}. Interestingly, the majority of the PfODC residues predicted to be involved in binding with this compound is also predicted in the interaction map with the most potent experimentally observed inhibitor of PfODC, CGP54169A. Parasite-specific residues Arg_{1117} and Tyr_{966} of the PfODC active site pocket were predicted to interact with mopp-DFB. The binding to Cys_{1355} and His_{998} that are important for catalysis supports the potential inhibitory capacity of mopp-DFB against PfODC. This compound does not significantly bind to the human enzyme since it was not identified as one of the top 100 scoring compounds when the same library was screened against human ODC. This furthermore indicates the potential for mopp-DFB to selectively bind to PfODC.

Biological activity spectra predictions (PASS predictions) of the compounds indicated that the majority (9 out of 10) of the compounds might have antiprotozoal activity (Table 6.1). Furthermore, six of these compounds are further predicted to have antimalarial activity. mopp-DFB is a borate derivative and boric acid is classified under the astringent and antiseptic therapeutic categories (Windholz, 1983). Compound 2 contains an acetanilide backbone with antipyretic and analgesic therapeutic applications (Windholz, 1983). Compounds 3 and 10 are derivatives of the quinoline ring structure and could explain their predicted antimalarial activity (Gillman, *et al.*, 1985). It has been speculated that the action of chloroquine could be linked to polyamine inhibition after observations of the additive inhibitory effects of chloroquine and MGBG (Das, *et al.*, 1997) and that there was a marked induction of both ODC and AdoMetDC activities in chloroquine sensitive compared to resistant *P. berghei* (Mishra, *et al.*, 1997). However, both sensitive and resistant strains had similar polyamine profiles (Mishra, *et al.*, 1997). Compounds 4 and 5 are used in the veterinary industry as an acaricide and an antihelmintic, respectively (Windholz, 1983).

A parallel screen of the same database with the human ODC crystal structure was used to identify novel ligands that will potentially bind specifically and selectively only against human ODC. None of the first 100 compounds identified in the screen against PfODC were predicted to interact with the human enzyme. Furthermore, the top 10 scoring compounds identified against the human ODC were all more complex heterocyclic compounds compared to the top 10 scoring compounds identified against PfODC (Table 6.2 and 6.1 respectively). This furthermore indicates sufficient structural variability between the active site pockets of the human and malarial ODCs. None of the compounds identified as ligands of human ODC share structural similarities to known ODC inhibitors and therefore provide novel skeletal structures that could be further investigated.

Despite the inherent limitations of computer modelled 3D structures, these structures are helpful in defining potential lead compounds. The results presented here need further investigations to evaluate mopp-DFB and the other identified compounds as potential novel inhibitors of PfODC. This include *in vitro* determination of the Ki values of these compounds against monofunctional PfODC as well as in the bifunctional PfAdoMetDC/ODC. The parasite-specific nature of these drugs also needs to be established. The predicted novel inhibitors need to be further investigated for its *in*

vivo capability in inhibiting parasite growth in culture. Furthermore, the predicted interaction sites need to be confirmed with mutagenesis studies. Once validated, these chemical compounds can be used as starting points for the process of drug development.

The predicted structure of PfODC therefore contributed towards explaining experimental results obtained with known inhibitors (DFMO and the CGP-series of compounds) as well as defining novel putative inhibitory compounds. Goodman and Gilman define therapeutic drugs in the following manner: Many drugs stimulate or depress biochemical or physiological function in man in a sufficiently reproducible manner to provide relief of symptoms or, ideally, to alter favourably the course of disease. Conversely, chemotherapeutic agents are useful in therapy because they have only minimal effects on man but can destroy or eliminate pathogenic cells or organisms' (Gillman, *et al.*, 1985). Therefore, even after the identification and validation of a novel lead inhibitor, to ultimately describe such an inhibitor as a successful therapeutic drug necessitates major investments and investigations.

The results presented in this Chapter conclude the aims set for this study. The following Chapter will highlight the relevance of the results obtained and will focus on possible future investigations of the bifunctional PfAdoMetDC/ODC and polyamine metabolism in *P. falciparum*.

CHAPTER 7

Concluding Discussion.

Malaria is still today one of the most devastating tropical infectious diseases of mankind. The emergence of multi-drug resistant parasites and the lack of a viable vaccine have stimulated the search for new chemotherapeutic targets. The traditional approach to drug discovery entails random screening of numerous compounds for parasite inhibition activity. Usually, no specific targets are identified thereby limiting understanding of the mechanisms of action of these drugs as well as the mechanisms involved in the development of resistance (as is the case for chloroquine). A more promising approach is the molecular and biochemical characterisation of a potential target protein prior to the design of specific inhibitors. This strategy of rational drug design allows the identification of particular characteristics of the target protein that could be exploited in first-line drug design (Blundell, 1996). By applying enhanced evolution techniques, the development of possible resistance mutations could be predicted and considered in the design of second-line drugs (Stemmer, 1994).

The global metabolome analyses of *P. falciparum* presented in Chapter 1 indicated several unique metabolic pathways that could be exploited as potential antimalarial drug targets (see Chapter 1, Fig. 1.6 and Table 1.2). At the start of this study in 1999, it was clear that polyamine metabolism in *P. falciparum* was unexplored and indeed that the biochemical reactions in the biosynthesis, catabolism and interconversion of these metabolites in respect to other pathways were largely unknown. However, based on the successful inhibition of polyamine metabolism in the treatment and prevention of diseases caused by other highly proliferative cells including cancer and infective diseases, the focus shifted to the elucidation of the polyamine metabolic pathway in other parasitic protozoa. The success achieved with the inhibition of polyamine biosynthesis on *T. brucei gambiense* propagation established polyamine biosynthesis as a novel antiparasitic target (Janne and Alhonen-Hongisto, 1989a; McCann and Pegg, 1992).

The study described here therefore focused on the biochemical characterisation of the polyamine metabolic pathway of *P. falciparum* in order to elucidate parasite-specific properties that can be exploited in the design of novel antimalarial therapies. The aims

set for this study included the identification and characterisation of the genes of the rate-limiting enzymes, AdoMetDC and ODC; followed by the structure and functional characterisation of the recombinantly expressed proteins and associated parasite-specific properties. The isolation and identification of the genes for *Adometdc* and *Odc*, as well as the recombinant expression of the monofunctional and bifunctional forms of the proteins are described in this thesis. The structure-activity relationships of the unique bifunctional PfAdoMetDC/ODC led to the identification of various parasite-specific properties. Ultimately, several novel and potentially specific inhibitors against the ODC component of PfAdoMetDC/ODC were identified using comparative modelling and *in silico* structure-based drug screening. The following sections highlight the significant aspects from each chapter.

The identification of genes of *P. falciparum* has always been a cumbersome task due to reasons explained in Chapter 2. An amplified, uncloned cDNA library for *P. falciparum* was constructed in our laboratory with suppression PCR technology to overcome most of the difficulties usually experienced with application of RACE procedures for this organism. We were able to use this elegant method to amplify the *Odc* cDNA as described in Chapter 2 using both 3'- and 5'-RACE techniques. This is to our knowledge one of the first examples where a *P. falciparum* cDNA was isolated and identified using degenerate primers in RACE protocols (Birkholtz, 1998c). However, this method failed to amplify the *Adometdc* cDNA of the expected size (~2900 bp). Two reasons for this result were identified. Firstly Müller *et al.* (2000) showed that *Adometdc* and *Odc* are located on the same transcript with an ORF of ~4 kDa. Secondly, the conserved amino acid sequence of AdoMetDCs of other organisms used for the design of a 3'-RACE degenerate primer was not as conserved in the *P. falciparum* sequence.

The large monocistronic transcript of 7 kbp for *Odc* and *Adometdc* was also confirmed in our laboratory using an *Odc*-specific probe on total RNA of *P. falciparum*. The 4.2 kbp cDNA was isolated by Müller *et al.* (2000) from a *P. falciparum* cDNA library that was screened with a *Odc*-specific probe amplified from genomic DNA with primers based on expressed sequence tags published by the *Plasmodium* genome project. Subsequently, the complete *PfAdometdc/Odc* cDNA was also amplified in our laboratory using the amplified, uncloned *P. falciparum* cDNA library. Preliminary predictions of the genomic structure and flanking regions of the *PfAdometdc/Odc* gene identified regulatory sequence elements flanking the gene. The transcript is proposed to

contain a large 5'-UTR of ~2600 bp although two transcription start sites were predicted (~300 bp apart) as a consequence of the high A+T content (>85%) of UTRs of *Plasmodial* genes (Lanzer, *et al.*, 1993). The size of the 5'-UTR is in agreement with the long UTR sizes observed for most malaria genes (Coppel and Black, 1998; Su and Wellems, 1998). *Adometdc* and *Odc* 5'-UTRs characterised in other organisms are also long and share certain features, including secondary ORFs and thermodynamically stable secondary structures, both of which are implicated in the regulation of translation of the proteins (Cohen, 1998; Heby and Persson, 1990). No secondary ORF was predicted for the putative 5'-UTR of *PfAdometdc/Odc* but significant secondary structures were observed. This suggests that the translation of PfAdoMetDC/ODC is not regulated in the same manner as the single proteins in other organisms through ribosome stalling (mediated by the second ORF product) but that feedback regulation through stabilisation of the secondary structures by polyamines could mediate unique translational regulatory mechanisms (Cohen, 1998).

The recombinant expression of both AdoMetDC and ODC demonstrated that the monofunctional forms of the proteins were active albeit with much lower activities compared to these activities in the natural bifunctional PfAdoMetDC/ODC state of the protein (Chapter 3 (Müller, *et al.*, 2000). All of the essential residues for catalysis of both decarboxylase functions shown for other organisms were conserved in the bifunctional PfAdoMetDC/ODC. The two decarboxylase domains of the bifunctional PfAdoMetDC/ODC were shown to group into the expected protein families of group IV decarboxylases (ODC) and S-adenosylmethionine decarboxylases (AdoMetDC). Furthermore, complete amino acid sequences were found for the bifunctional protein in two other murine *Plasmodia*, *P. berghei* and *P. yoellii* indicating an evolutionary advantage to the malaria parasites afforded by the bifunctional organisation of these enzymes. Certain parasite-specific areas were identified in the PfAdoMetDC/ODC sequence. These include the hinge region connecting the decarboxylase domains but also three additional parasite-specific insertions. Some of these areas are extremely large (>150 residues), have repetitive, polar sequences and were predicted to contain the majority of low-complexity regions present in the protein. The identification of the genes and characterisation of the recombinant proteins describes the first aims set in this study. From this, several parasite-specific properties were evident including the proposed genomic structure of the gene, the relationship of the bifunctional protein in

Plasmodia as well as parasite-specific inserts in the amino acid sequence. The structure-activity properties of the two enzymes were further investigated as described below.

In Chapter 4, the bifunctional organisation of PfAdoMetDC/ODC and the possible involvement of the parasite-specific inserts in the activities and physical interactions between the decarboxylase domains were investigated. All the inserts were shown to be important for catalytic activity of the respective domains. However, these inserts also influence the activity of the neighbouring domain to various degrees. Physical interactions between separately expressed, monofunctional AdoMetDC and ODC resulted in the reconstruction of the bifunctional complex *in vitro*. These interactions were mediated by the smallest insert in the ODC domain as well as by the hinge region. It was deduced from the results that part of the predicted hinge region also constituted a part of the ODC domain. The insert in the ODC domain was also predicted to be the most structured indicating its functional importance in the correct association of the bifunctional PfAdoMetDC/ODC complex. Protein-protein interactions between the monofunctional ODC and AdoMetDC domains seem to be more dependent on intact ODC than AdoMetDC. Unresolved aspects arising from this work include the possibility that deletion of large areas in proteins could lead to major conformational changes. The loss of activity should therefore be investigated by systematic analyses of more defined and smaller mutations. These could include removing the low-complexity areas inside the parasite-specific inserts and mutating the NND-rich areas to less polar residues to investigate the involvement of these areas in a polar zipper (Frontali, 1994; Perutz, *et al.*, 1994). However, the ODC domain seems to be more refractory to change since even single amino acid substitutions at the dimer interface result in inactive protein due to long-range coupling to the active site (Myers, *et al.*, 2001). Although the large parasite-specific inserts will be difficult to target in an antimalarial strategy, the importance of insert O_1 in mediating physical interactions in the bifunctional heterotetrameric complex could be used in strategies to prevent this association.

To exploit the bifunctional PfAdoMetDC/ODC as a potential antimalarial target, structural data is needed to aid rational drug design strategies. In Chapter 5, a homology model of the monofunctional ODC domain is described. The model predicted a global structure corresponding to those of homologous proteins in other organisms. However, certain parasite-specific properties were identified, including differences in the active site pocket and dimerisation interface. A significant degree of deviation was also

present in the solvent-accessible loops, most of which were elongated in the malarial protein. The smallest insert described in the ODC domain (Chapter 4) was predicted to be present on the surface of the protein. The predicted homology model is consistent with mutagenesis results and biochemical studies concerning some active site residues and areas involved in stabilising the dimeric state of the protein (Wrenger, *et al.*, 2001). The model described here presents the only structural data available for this protein in the absence of a crystal structure. The size of the bifunctional protein and the presence of the parasite-specific inserts as well as the low expression levels of the recombinant protein suggest that a crystal structure for this protein would be difficult to obtain. The large, nonglobular parasite-specific inserts are furthermore often removed to obtain useful crystal structures of malarial proteins (Dunn, *et al.*, 1996; Shi, *et al.*, 1999; Velanker, *et al.*, 1997). Attempts to crystalise a similar bifunctional protein, DHFR-TS, have not been successful and a similar strategy to obtain structural data using a homology model is currently the only useful alternative (Lemcke, *et al.*, 1999; Toyoda, *et al.*, 1997). Comparison of a homology model of TIM with the crystal structure revealed little deviation between the two and indicated the potential success of such an approach (Joubert, 2000; Velanker, *et al.*, 1997). Contentious aspects of the homology model include its accuracy in the absence of knowledge of the structure of the deleted parasite-specific inserts and of the AdoMetDC domain. However, the results obtained indicated a satisfactory correspondence of the core structure of the protein to those of other eukaryotic ODC enzymes. The correctness of the predicted active site is accurate in explaining the catalytic mechanism, is supported by mutagenesis results (Wrenger, *et al.*, 2001) and consistent with the experimental inhibition observed with various known ligands (Chapter 6).

Structure-based design of novel antimalarials is a strategy that is receiving widespread attention (Bohm and Klebe, 1996; Lemcke, *et al.*, 1999; Toyoda, *et al.*, 1997). This strategy was followed in Chapter 6 to explain the experimental inhibition of PfODC activity with three different known inhibitors and to identify their interaction sites with essential catalytic residues. Novel potential inhibitors were identified by the *in silico* screening of a chemical structure library. These potential inhibitors appear to be parasite-specific since they were not predicted to bind to human ODC. These compounds need further investigations e.g. inhibition of recombinantly expressed monofunctional ODC and bifunctional PfAdoMetDC/ODC as well as *in vitro* cultures of *P. falciparum*. Once confirmed, their structures could be furthermore used as

scaffolds in the design of more specific, higher affinity inhibitors. This approach has been employed to establish the essential characteristics of an inhibitor of drug-resistant *P. falciparum* DHFR (Yuthavong, 2002).

This thesis therefore described the analyses of the rate-controlling enzymes of polyamine metabolism in *P. falciparum* and resulted in the identification of parasite-specific properties. Certain of these properties were highlighted and proposed novel functions are explained. These properties include the bifunctional organisation of malarial proteins and the parasite-specific insertions. The thesis concludes with the exploitation of these parasite-specific properties in the design of potential novel antimalarials that could be developed as inhibitors of polyamine metabolism. However, these results raised important questions that are discussed in more detail in the following sections.

The bifunctional nature and origin of PfAdoMetDC/ODC.

The bifunctional nature of PfAdoMetDC/ODC is unique in several ways. Firstly, it is a property present only in *Plasmodia* and secondly, it is composed of two proteins found on different legs of the metabolic pathway. Other bifunctional proteins described in *P. falciparum* do have homologues in other protozoa (i.e. DHFR-TS found in *T. brucei* and *L. donovani*)(Bzik, *et al.*, 1987; Ivanetich and Santi, 1990). However, the majority of the described bifunctional proteins of *P. falciparum* participate in consecutive steps in the same metabolic pathways (Clarke, *et al.*, 2001; Triglia and Cowman, 1994).

Of the bifunctional nature of PfAdoMetDC/ODC there is little doubt. The bifunctional protein is translated from a single, monocistronic transcript. The deduced amino acid sequence contains homologies to protein families to which both AdoMetDC and ODC belong. Furthermore, all the necessary residues for catalysis are conserved in the primary sequence indicating that both proteins use the same mechanisms of catalysis as their eukaryotic homologues. Indeed, both decarboxylase activities were demonstrated in the recombinantly expressed PfAdoMetDC/ODC.

The origin of the bifunctional nature of the decarboxylases is unknown. An advantageous exon-shuffling event during the evolution of *Plasmodia* is one probability. Gene duplication can probably be excluded since the two decarboxylases belong to two distinct protein families. The original mutation that resulted in gene

linkage must have occurred before the divergence of the *Plasmodial* species, since the human and murine species that are more diverse than the human and primate species, both retained the bifunctional organisation of the enzymes. Therefore, the bifunctional organisation of these enzymes must confer some specific evolutionary advantage to *Plasmodia* species.

Several speculations have been forwarded to explain the advantages of the bifunctional nature of malarial proteins. These include substrate channelling, coordinated regulation of protein concentrations/activities or intramolecular communication and interaction (Ivanetich and Santi, 1990; Müller, *et al.*, 2000). Substrate channelling can in all probability be discounted as an explanation for the bifunctional nature of PfAdoMetDC/ODC since it could not be shown that this protein interacts with the subsequent enzyme in this pathway, spermidine synthase. However, alternative mechanisms must exist to regulate both the concentration and activity of the bifunctional protein in *P. falciparum* compared to the monofunctional homologues of other organisms. Firstly, both the decarboxylase activities are enhanced when present in the bifunctional complex compared to the monofunctional PfAdoMetDC and PfODC. Therefore, the bifunctional organisation appears to be beneficial for the activities probably through stabilisation of the protein conformations even though the decarboxylase activities are able to function independently from each other (Wrenger, *et al.*, 2001). Secondly, analyses of the single transcript of PfAdoMetDC/ODC indicated that translational regulation of the bifunctional protein differs from the mechanisms regulating translation of the monofunctional protein homologues of other organisms. Thirdly, the deduced amino acid sequence of PfAdoMetDC/ODC as well as the homology model of PfODC indicated that the regulation of the protein concentration via degradation is probably different to that found in other organisms. No evidence for possible antizyme-mediated regulation was observed as for other ODCs and the PfAdoMetDC is also not stimulated by putrescine as in other organisms.

Therefore, the bifunctional nature of PfAdoMetDC/ODC is proposed to allow a combination of coordinated regulation of the decarboxylase activities and protein concentrations as well as to allow beneficial intramolecular communications between the two domains.

Structure-activity properties of parasite-specific inserts in PfAdoMetDC/ODC.

One of the most intriguing and least investigated properties of a myriad of malarial proteins is the occurence of parasite-specific inserts that intersperse adjacent conserved areas shared with other protein homologues. Speculations on the functions and evolutionary advantages of these inserts include a mechanism to evade the host immune system, interaction sites with regulatory proteins in the parasite or human host or modular mediators of protein-protein interactions (Li and Baker, 1998; Perutz, *et al.*, 1994; Ramasamy, 1991; Schofield, 1991). These parasite-specific inserts in AdoMetDC/ODC are not unique to *P. falciparum* since it was shown to be also a characteristic of other *Plasmodial* species. Strong selective pressures must therefore exist to maintain and diversify these regions (Ramasamy, 1991). The conclusion drawn from the studies presented here is that the parasite-specific inserts in PfAdoMetDC/ODC are required to mediate certain protein-protein interactions in the bifunctional protein. This is supported by the following observations:

The bifunctional nature of PfAdoMetDC/ODC presented a unique opportunity to investigate the possible functions of the parasite-specific inserts in this protein. Firstly, the inserts are indeed present in the protein and must be correctly folded to allow the heterologous expression of the active form of the protein. Secondly, amino acid sequence analyses of these inserts showed a prevalence for Asn and Asp that could form polar zippers, a characteristic that have been proposed to behave as modular mediators in protein-protein interactions (Perutz, *et al.*, 1994). Thirdly, all of the parasite-specific inserts are required for effective functioning of both the decarboxylase activities. Fourthly, the large hydrophilic inserts contained low-complexity regions and, in the case of the ODC domain at least, structural analyses predicted that these areas form nonglobular domains most likely on the surface of the proteins. Finally, the hinge region as well as the structurally conserved and smallest insert in the ODC domain was essential in mediating protein-protein interactions between the two decarboxylase domains. The structured insert was predicted to be present on the surface of PfODC. It is furthermore bordered by highly flexible Gly residues, which are proposed to allow precise interaction with the PfAdoMetDC domain.

Therefore, at least in the bifunctional PfAdoMetDC/ODC, certain parasite-specific inserts mediate physical protein interactions. These results contribute towards a better understanding of the function and evolutionary advantage of the parasite-specific inserts

and is one of the first demonstrations of the essential roles of these inserts for function and protein-protein interactions.

Structural organisation of the bifunctional PfAdoMetDC/ODC.

Throughout this thesis, various structure-activity relationships were inferred for the bifunctional PfAdoMetDC/ODC from the multimeric states of the monofunctional proteins, homology modelling structural data of the PfODC component as well as mutagenesis results.

Fig. 7.1 proposes the structure of the bifunctional PfAdoMetDC/ODC and summarises the predicted protein-protein interactions. The ODC domain is depicted as a homodimer, since activity of this protein is dependent on the association between the two monomers. Furthermore, the ODC dimer has a head-to-tail association between the two monomers with the C-terminus of one monomer interacting with the N-terminus of the second as indicated by the homology model of PfODC. Both active sites are on the same side of the protein on the dimer interface.

Results from Chapter 3 showed that AdoMetDC of *P. falciparum* is a heterotetrameric protein. In the absence of structural data for the AdoMetDC domain, the association is extrapolated from the mammalian AdoMetDC in which an edge-on association of the α-chains of the proteins occurs (Ekstrom, *et al.*, 1999). All the inserts are important for the activities of the respective domains as indicated by the blue arrows in the schematic representation. However, the inserts also affect the activities of the neighbouring domains (red arrows). The large insert in the ODC domain (O_2) partly influences AdoMetDC activity (dashed red arrows) but more pronounced influences on enzyme activities are mediated between the domains by insert A_1 in the AdoMetDC domain and O_1 in the ODC domain (solid red arrows). Physical association occurs between the AdoMetDC and ODC domains as evidenced from formation of hybrid bifunctional complexes after co-incubation of separately expressed AdoMetDC and ODC. Only parasite-specific insert O_1 seems to be essential in mediating these intermolecular interactions (yellow arrows). The hinge region seems more important for ODC activity but also has, to a lesser extent, an influence on physical association with the AdoMetDC domain. Heterotetrameric complex formation of the bifunctional protein is due mostly to interactions between the AdoMetDC domains with no apparent contribution of a

mutated ODC domain. However, intramolecular interactions are more dependent on intact homodimeric ODC to associate with AdoMetDC.
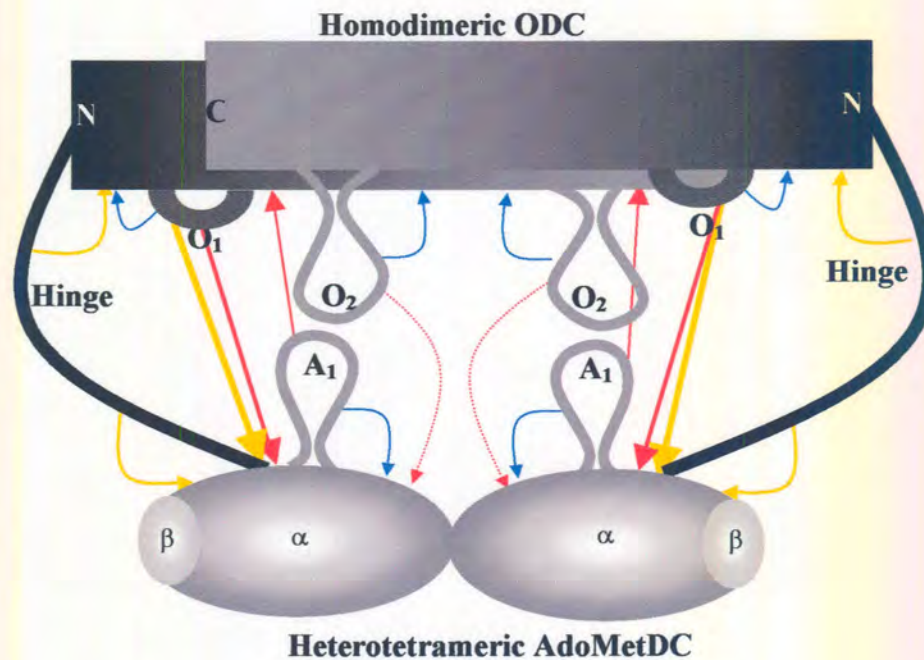
Figure 7.1: Schematic representation of the proposed structural arrangement of the bifunctional PfAdoMetDC/ODC. The obligate homodimeric ODC domain is indicated in dark blocks, and the heterotetrameric AdoMetDCs in ellipses. The hinge extends from the N-terminus of ODC to the AdoMetDC domain. The parasite-specific inserts are indicated as loops. The influences that the inserts had on the activity of the respective domains are indicated with blue arrows and influences on the neighbouring activity with red arrows. Dashed lines indicate minor roles and solid lines major influences. The contribution of the inserts to intermolecular protein-protein interactions between the domains is indicated with yellow arrows, the importance of the contribution indicated by the thickness of the arrow.

The formation of the heterotetrameric bifunctional protein is therefore due to various cumulative interactions in which the AdoMetDC domain association seems the most stable, and the ODC domain the most refractory to modifications. The parasite-specific inserts mediate some of the protein-protein interactions. In order to validate the proposed structural organisation of the bifunctional PfAdoMetDC/ODC, experimental structural data for the complete complex or, in the absence thereof, for the monofunctional domains are required. However, the crystallisation of various malarial proteins are complicated by the unstructured nature of inserted amino acid sequences containing mostly low-complexity regions. The low expression levels of the recombinant bifunctional PfAdoMetDC/ODC further impede such a strategy. Recent advances in the field of cryo-electron microscopy by which protein structures can be

determined to at least 3 Å could be employed to provide a clearer picture of the bifunctional arrangement of these proteins.

Polyamine biosynthesis of *P. falciparum* with reference to PfAdoMetDC/ODC as a viable drug target.

The ultimate aim of molecular and biochemical characterisation of metabolic processes of *P. falciparum* is to aid the design of novel antimalarial chemotherapies. However, before any metabolic pathway can be validated as a potential target for chemotherapy, it is imperative to understand the biochemistry of the processes involved. The identification of parasite-specific properties could be useful in the design of selective inhibitors.

The bifunctional nature of PfAdoMetDC/ODC and associated parasite-specific inserts present unique properties to investigate these as potential targets to inhibit polyamine metabolism. However, to date no drug could be shown to selectively inhibit either AdoMetDC or ODC activity of *P. falciparum* due to the shared conservation of the active sites of malarial and mammalian enzymes. Structural characterisation of the PfODC domain with homology modelling revealed parasite-specific residues in the active site pocket as well as at the entrance to the pocket. Additionally, the dimer interface revealed several interactions that were unique to PfODC and useful for the selective prevention of dimerisation. Novel putative inhibitors that are predicted to bind only to PfODC could be identified by *in silico* screening of chemical structure libraries.

Beyond active site targeting strategies, the bifunctional organisation of the malarial PfAdoMetDC/ODC compared to the monofunctional homologues of mammals presents another targeting strategy. A better understanding of the potential role of the parasite-specific inserts to mediate physical interactions between the domains in the bifunctional protein provides a distinct opportunity for selective chemotherapy against malaria. The greater structural variability of protein-protein interfaces, which in PfAdoMetDC/ODC also contain parasite-specific regions, suggests sufficient differentiation between such contact sites of the host and parasite enzymes. The use of small molecular weight peptidomimetics that disrupt protein-protein interactions has been explored for drug development in many systems including *Lactobacillus casei* thymidylate synthase (Prasanna, *et al.*, 1998), HIV-1 protease (Schramm, *et al.*, 1996; Zutshi, *et al.*, 1998), herpes simplex virus DNA polymerase and human glutathione reductase (Zutshi, *et al.*,

1998). Recently, synthetic peptides were also shown to prevent dimerisation and effectively inactivate *P. falciparum* TIM (Singh, *et al.*, 2001).

However, consideration of polyamine metabolism in *P. falciparum* as a viable target for chemotherapy requires demonstration of its relevance in parasite survival. Supporting evidence for validation of a metabolic pathway or specific enzyme as a chemotherapeutic target include 1) comparisons of profiles between enzyme inhibition and antiparasite action among a large number of chemical derivatives of the drug, 2) characterisation of the putative drug target from drug-resistant mutant parasites and 3) knockouts of the gene encoding the putative target accompanied by subsequent complementation of the missing gene (Wang, 1997). Of these, the latter may prove the most conclusive validation of a target and has been used to show that ODC from *T. brucei* is a *bona fide* target (Wang, 1997). Transfection is not as easily achieved in *P. falciparum* and results of knockout experiments for the bifunctional PfAdoMetDC/ODC are still lacking.

Inhibition of polyamine biosynthesis in *P. falciparum* is usually cytostatic rather than cytotoxic due to the inherent homeostatic mechanisms in place to control intracellular polyamine levels. This may include the transport of polyamines by the parasites from the extracelullar milieu or poor uptake of the drugs (Fukomoto and Byus, 1996; Müller, *et al.*, 2001). Folate metabolism is regarded as a validated chemotherapeutic target in protozoa, even in the presence of a folate transport system. The synergistic use of inhibitors against two enzymes in the pathway, DHPS and DHFR, overcomes resistance seen due to increased folate transport compared to when DHPS inhibitors are used separately. It was proposed that the DHFR inhibitor pyrimethamine counters the so-called folate effect probably through blockage of one or more steps in the uptake and utilisation of exogenous folate (Hyde, 2002). These same principles could be applied to the inhibition of polyamine biosynthesis in *P. falciparum*. Firstly, because of the bifunctional nature, inhibition of both decarboxylase activities would act synergistically in limiting polyamine biosynthesis. Furthermore, the combination with a polyamine transporter inhibitor would further contribute to foil the homeostatic mechanisms. Common features shared between the substrates/products of these enzymes and recognised by the polyamine transporter need however to be identified to exploit this route of drug design.

191

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Future directions in the study of polyamine metabolism in *P. falciparum.*

This study raised several interesting questions that need to be further investigated. At the start of the study, the genomic structure of the *PfAdometdc/Odc* gene and transcript was proposed. However, the nature of the 5'-UTR and specific contribution this region do make to regulate translation of the transcript should be investigated through analyses of the complete nucleotide sequence of the transcript as well as *in vitro* translation studies to determine the various effectors needed during formation of active PfAdoMetDC/ODC protein. This will also aid in determining how the bifunctional protein is regulated post-translationally. Structure-activity investigations reported here indicated various parasite-specific properties including those of the parasite-specific inserts and their involvement in protein-protein interactions. Finer dissection of the proposed functions attributed to these inserts should be considered, including deletion of only the low-complexity regions in these inserts and mutation of the polar repeat sequences. The structure and organisation of the bifunctional PfAdoMetDC/ODC will be clarified once structural data for the AdoMetDC domain is available. The creation of a homology model of the AdoMetDC domain will greatly aid in the understanding of the interactions between the two decarboxylase domains. Ultimately, a complete three-dimensional structure of the bifunctional protein will have immense value in explaining the parasite-specific properties as well as aiding the design of novel inhibitors.

In Chapter 1, the metabolome of *P. falciparum* was represented. However, this did not include details on the possible interacting pathways and physiological functions of the polyamines in this parasite. In order to fully understand the metabolome and the effects that inhibition of polyamine biosynthesis will have on *P. falciparum*, additional data is required. The transcriptome (mRNA profile) of *P. falciparum* after polyamine depletion will identify upregulated transcripts and the corresponding proteome will identify the mechanisms utilised by the parasite to compensate for the lack of the essential metabolites. Furthermore, enzymes from related and interacting pathways (polyamine interconversion, methionine recycling and polyamine transport) could also be identified and collectively the new data will identify new potential drug targets as well as a more complete understanding of the polyamine metabolic profile of *P. falciparum.*

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Albert L. Lehninger said the following on the challenge to Biochemists (Lehninger, 1975):

'Living things are composed of lifeless molecules. When these molecules are isolated and examined carefully, they conform to all the physical and chemical laws that describe the behaviour of inanimate matter. Yet living organisms possess extraordinary attributes not shown by collections of inanimate molecules.'

It therefore remains our responsibility as scientists not to become too absorbed in the spectacular peculiarities of single proteins or nucleic acids, but to try and unravel the mysteries of nature by completing the biochemical network from which cellular entities arises.

Summary

Title of Thesis: Functional and structural characterization of the unique bifunctional
enzyme complex involved in regulation of polyamine metabolism in
Plasmodium falciparum

Student: Lyn-Marie Birkholtz

Supervisor: Prof AI Louw

Co-supervisor: Prof RD Walter

Department: Biochemistry

Degree: *Philosophiae Doctor*

Malaria remains one of the most serious tropical infectious diseases affecting mankind. The prevention of the disease is hampered by the increasing resistance of the parasite to existing chemotherapies. The need for novel therapeutic targets and drugs is therefore of the utmost importance and detailed knowledge of the biochemistry of the parasite is imperative. This study was directed at the biochemical characterisation of the 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 antimalarials. The thesis focussed on the two rate-limiting enzymes in polyamine biosynthesis, S-adenosylmethionine decarboxylase (AdoMetDC) and ornithine decarboxylase (ODC), which occur as a unique bifunctional complex in *P. falciparum*.

The genomic structure of the bifunctional gene indicated a single, monocistronic transcript with large untranslated regions that were predicted to be involved in unique translational regulatory mechanisms. This gives rise to a bifunctional protein containing both decarboxylase activities on a single polypeptide forming a heterotetrameric complex. Activity of the decarboxylases decreases dramatically if these proteins are expressed in their monofunctional forms as homodimeric ODC and heterotetrameric AdoMetDC. The deduced amino acid sequence indicated that all the essential residues for catalysis are conserved and highlighted the presence of three parasite-specific insertions.

The parasite-specific inserts were shown to be essential for the catalytic activity of the respective domains and also to influence the activity of the neighbouring domain, indicating that intramolecular communication exists in the heterotetrameric complex. The most structured and smallest insert was also shown to mediate protein-protein interactions between the two domains and to stabilise the complex. Further structure-

functional characterisations of specifically the ODC domain were deduced from a comparative homology model. The model predicted an overall structure corresponding to those of other homologous proteins. The validity of the model is supported by mutagenesis results. However, certain parasite-specific properties were identified in the active site pocket and dimerisation interface. The former was exploited in the rational design of novel putative ODC inhibitors directed only against the *P. falciparum* protein by *in silico* screening of chemical structure libraries.

This study therefore describes the identification of certain parasite-specific properties in a unique bifunctional protein involved in regulation of polyamine metabolism of *P. falciparum*. Such discoveries are invaluable in strategies aimed at elucidating biochemical and metabolic differences between the parasite and its human host that could be exploited in the design of alternative, parasite-specific chemotherapies. Moreover, the thesis also contributed new knowledge on certain less well-understood biological phenomena characteristic of *P. falciparum*, the nature and origin of bifunctional proteins and the functional properties of parasite-specific inserts found in some proteins of the parasite.

Opsomming

Titel van Tesis: Functional and structural characterization of the unique bifunctional
enzyme complex involved in regulation of polyamine metabolism in
Plasmodium falciparum

Student: Lyn-Marie Birkholtz

Promotor: Prof AI Louw

Medepromotor: Prof RD Walter

Departement: Biochemie

Graad: *Philosophiae Doctor*

Malaria is steeds die mees kommerwekkendste tropiese infeksie wat die mensdom
teister. Voorkoming van die siekte word belemmer as gevolg van die parasiete wat
weerstandig raak teen die bestaande voorkomende en terapeutiese middels. Dit is dus
uiters noodsaaklik om nuwe terapeutiese teikens te identifiseer asook om die
biochemiese eienskappe van die parasiet beter te verstaan om sodoende nuwe
medisynes te kan ontwerp. Hierdie studie beskryf die biochemiese karakterisering van
die poli-amien metaboliese baan van die menslike malaria parasiet, *P. falciparum*, om
sodoende verskille tussen die parasiet en sy menslike gasheer te identifiseer wat gebruik
kan word in the ontwikkeling van nuwe anti-malaria middels. Twee tempo-beherende
ensieme in die baan, S-adenosielmetionien dekarbosilase (AdoMetDC) en ornitien
dekarbosilase (ODC), wat voorkom as in 'n unieke, bifunksionele kompleks in *P.
falciparum*, was bestudeer.

Die genomiese struktuur van die geen vir die bifunksionele proteïen het aangedui dat 'n
enkele, monosistroniese transkrip kodeer vir die proteïen. Die groot, ongetransleerde
gedeeltes van die transkrip word voorspel om betrokke te wees in alternatiewe
regulatoriese meganismes tydens translasie van die proteïen. Die bifunksionele proteïen
het dus beide die dekarbosilase aktiwiteite op 'n enkele polipeptied en bestaan as 'n
heterotetrameriese kompleks. Die dekarboksilase aktiwiteite verlaag dramaties indien
die proteïene uitgedruk word in hul monofunksionele vorme as heterotetrameriese
AdoMetDC en homodimeriese ODC. Analises van die afgeleide aminosuurvolgorde
van die bifunksionele protein het verskeie gekonserveerde residue aangedui wat
essensieel is vir katalitiese aktiwiteit, asook drie parasiet-spesifieke invoegsels.

Die parasiet-spesifieke invoegsels is bewys om essensieël te wees vir die aktiwiteit van
die domein waarin dit voorkom maar beïnvloed ook die aktiwiteit van die aangrensende

domein wat die bestaan van intramolekulêre kommunikasie tussen die twee domeine aandui. Die kortste en mees gestruktureerde invoegsel is ook gewys om proteïen-proteïen interaksies te bemiddel om sodoende die bifunksionele kompleks te stabiliseer. Verdere struktuur-funksie karakteriserings van spesifiek die ODC domein is verkry deur `n vergelykende homologie model. Die model voorspel `n algemene struktuur van die proteïen wat gunstig vergelyk met ander homoloë proteïene. Die akkuraatheid van die model is gestaaf deur mutagenese resultate. Verskeie parasiet-spesifieke eienskappe is egter geïdentifiseer beide in die aktiewe setel sowel as in die dimerisasie intervlak van die proteïen. Die voorspelde struktuur van die aktiewe setel is verder gebruik in die identifikasie van nuwe, vermoedelike spesifieke inhibitore van *P. falciparum* ODC deur *in silico* sifting van chemiese struktuur biblioteke.

Hierdie studie beskryf dus verskeie parasiet-spesifieke eienskappe van die unieke bifunksionele proteïen wat poli-amien metabolisme van *P. falciparum* reguleer. Sulke ontdekkings is waardevol in strategieë wat fokus op biochemiese en metaboliese verskille tussen die parasiet en sy menslike gasheer om sodoende alternatiewe, parasiet-spesifieke chemoterapeutiese middels te ontwikkel. Verder dra hierdie studie ook nuwe kennis by tot ander biologiese aspekte van *P. falciparum* insluitend die oorsprong en karakter van bifunksionele proteïene asook die funksionele bydrae van die parasiet-spesifieke invoegsels in sommige proteïene van die parasiet.

References

Adams, R. L. P., Knowler, J. T. and Leader, D. P. (1993) In: *The biochemistry of nucleic acids*. Chapman & Hall, London

Algranati, I. D. and Goldemberg, S. H. (1989) Effects of polyamines and antibiotics on the structure and function of ribosomes. In: *The physiology of the polyamines*. Bachrach, U., CRC Press, Boca Raton, 143-155Algranati, I. D. and Goldemberg, S. H.

Almrud, J. J., Oliveira, M. A., Kern, A. D., Grishin, N. V., Phillips, M. A. and Hackert, M. L. (2000) Crystal structure of human ornithine decarboxylase at 2.1 A resolution: structural insights to antizyme binding. *J Mol Biol* **295**, 7-16

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

Amador, R. and Patarroyo, M. E. (1996) Malaria vaccines. *J. Clin. Immunol.* **16**, 183-189

Anders, R. F. and Saul, A. (2000) Malaria Vaccines. *Parasitol. Today* **16**, 444-447

Assaraf, Y. G., Abu-Elheiga, L., Spira, D. T., Desser, H. and Bachrach, U. (1987a) Effect of polyamine depletion on macromolecular synthesis of the malarial parasite, *Plasmodium falcipalum*, cultured in human erythrocytes. *Biochem. J.* **242**, 221-226

Assaraf, Y. G., Golenser, J., Spira, D. T. and Bachrach, U. (1984) Polyamine levels and the activity of their biosynthetic enzymes in human erythrocytes infected with the malaria parasite, *Plasmodium falciparum*. *Biochem. J.* **222**, 815-819

Assaraf, Y. G., Golenser, J., Spira, D. T. and Bachrach, U. (1986) *Plasmodium falciparum*: Synchronization of cultures with DL-α-difluoromethylornithine, an inhibitor of polyamine biosynthesis. *Exp. Parasitol.* **61**, 229-235

Assaraf, Y. G., Golenser, J., Spira, D. T., Messer, G. and Bachrach, U. (1987b) Cytostatic effect of DL-α-difluoromethylornithine against *Plasmodium falciparum* and its reversal by diamines and spermidine. *Parasitol. Res.* **73**, 313-318

Assaraf, Y. G., Kahana, C., Spira, D. T. and Bachrach, U. (1988) *Plasmodium falciparum*: Purification, properties and immunochemical study of ornithine decarboxylase, the key enzyme in polyamine biosynthesis. *Exp. Parasitol.* **67**, 20-30

Ayala, F. J., Escalante, A. A., Lal, A. A. and Rich, S. M. (1998) Evolutionary relationships of human malaria parasites. In: *Malaria: Parasite biology, pathogenesis and protection*. Sherman, I. W., ASM Press, Washington, D. C., Ayala, F. J., Escalante, A. A., Lal, A. A. and Rich, S. M.

Baca, A. M. and Hol, W. G. (2000) Overcoming codon bias: A method for high-level over-expression of *Plasmodium* and other AT-rich parasite genes in *Eschericia coli*. *Int J Parasitol* **30**, 113-118

Bachrach, U. (1984) Physiological aspects of ornithine decarboxylase. *Cell Biochem. Funct.* **2**, 6-10

198

Bairoch, A., Bucher, P. and Hofman, K. (1995) The PROSITE database, its status in 1995. *Nucl. Acids Res.* **24**, 189-196

Balint, G. A. (2001) Artemisinin and its derivatives: an important new class of antimalarial agents. *Pharmacol. Ther.* **90**, 261-265

Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. and Mitchell, G. H. (2000) A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol. Today* **16**, 427-433

Barale, J. C., Candelle, D., Attal-Bonnefoy, G., Dehoux, P., Bonnefoy, S., Ridley, R., da Silva, L. P. and Langsley, G. (1997) Plasmodium falciparum AARP1, a giant protein containing repeated motifs rich in asparafine and aspartate residues, is associated with the infected erythrocyte membrane. *Infect. Immun.* **65**, 3003-3010

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. and Sonnhammer, E. L. L. (2002) The Pfam protein families database. *Nucl Acids Res* **30**, 276-280

Bathurst, I. C. (1994) Protein expression in yeast as an approach to production of recombinant malaria antigens. *Am. J. Trop. Med. Hyg.* **50**, 20-6

Berendt, A. R., Ferguson, D. J. P., Gardner, J., Turner, G., Rowe, A., McCormick, C., Roberts, D., Craig, A., Pinches, B., Elford, B. C. and Newbold, C. I. (1994) Molecular mechanisms of sequestration in malaria. *Parasitology* **108**, S19-S28

Berens, R. L., Krug, E. C. and Marr, J. J. (1995) In: *Purine and pyrimidine metabolism*. Ed:Marr, J. J. and Muller, M., Academic Press Ltd., London

Birkholtz, L. (1998c) Molecular characterisation of the ornithine decarboxylase gene of the human malaria parasite, *Plasmodium falciparum*. *Faculty of Biological and Agricultural Sciences, Department of Biochemistry.* University of Pretoria. 129

Birkholtz, L. (2000b) Polyamine metabolism in the human malaria parasite, *Plasmodium falciparum*. *2nd Gauteng Region Annual Biochemistry Symposium*

Birkholtz, L. (2000c) *Plasmodium falciparum* ornithine decarboxylase: Molecular characterisation and recombinant expression. *BioY2K Combined Millennium Meeting*

Birkholtz, L. (2000d) Molecular characterisation and recombinant expression of *Plasmodium falciparum* ornithine decarboxylase. *2nd Gauteng Region Annual Biochemistry Symposium*

Birkholtz, L. (2002c) Comparative properties of a three-dimensional model of Plasmodium falciparum ornithine decarboxylase. *Polyamine metabolism as a drug target in parasitic protozoa and worms*

Birkholtz, L., Joubert, F. and Louw, A. I. (2000a) *Plasmodium falciparum* ornithine decarboxylase: Molecular characterisation and recombinant expression. *Young Scientist Symposium, 18th International Conference of the IUBMB*

Birkholtz, L., Joubert, F. and Louw, A. I. (2001b) Structural characterisation of ornithine decarboxylase of *Plasmodium falciparum*. *Gordon Research Conference on Polyamines*

Birkholtz, L., Joubert, F. and Louw, A. I. (2001c) Structural characterisation of ornithine decarboxylase of *Plasmodium falciparum*. *IUBMB/SASBMB Special Meeting on the Biochemical and Molecular Basis of Disease*

Birkholtz, L., Joubert, F., Neitz, A. W. H. and Louw, A. I. (2002a) Comparative properties of a three-dimensional model of *Plasmodium falciparum* ornithine decarboxylase. *Prot: Struc. Funct. Genet.*

Birkholtz, L. and Louw, A. I. (1998b) The nucleotide sequence of a *Plasmodium falciparum* ornithine decarboxylase gene. *FASBMB/SASBMB Biochemistry in Africa*

Birkholtz, L. and Louw, A. I. (1999a) Cloning and characterisation of *Plasmodium falciparum* ornithine decarboxylase complementary DNA obtained by RACE. *Molecular Parasitology Meeting X*

Birkholtz, L. and Louw, A. I. (1999b) Molecular characterisation of the ornithine decarboxylase cDNA of the human malaria parasite, *Plasmodium falciparum*. *MIM African Malaria Conference*

Birkholtz, L. and Louw, A. I. (2000d) Molecular characterisation and recombinant expression of *Plasmodium falciparum* ornithine decarboxylase. *Molecular Aspects of Malaria Meeting*

Birkholtz, L., Visser, L., Brink, A. and Louw, A. I. (1998a) Drug-resistant and mixed-species malaria infections in Mpumalanga, South Africa. *S.A. J. Sci.* **94**, 39-43

Birkholtz, L., Wrenger, C., Joubert, F., Wells, G. A., Walter, R. D. and Louw, A. I. (2002b) Functional roles of parasite-specific inserts in the bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase of *Plasmodia*. *Biochem. J.*

Bitoni, A. J., McCann, P. P. and Sjoerdsma, A. (1987) *Plasmodium falciparum* and *Plasmodium berghei*: Effects of ornithine decarboxylase inhibitors on erythrocytic schizogony. *Exp. Parasitol.* **64**, 237-243

Bitonti, A. J., Dumont, J. A., Bush, T. L., Edwards, M. L., Stemerick, D. M., McCann, P. P. and Sjoerdsma, A. (1989) Bis(benzyl)polyamine analogs inhibit the growth of chloroquine-resistant human malaria parasites (Plasmodium falciparum) in vitro and in combination with alpha-difluoromethylornithine cure murine malaria. *Proc Natl Acad Sci U S A* **86**, 651-5.

Blundell, T. L. (1996) Structure-based drug design. *Nature* **384**, 23-26

Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E. and Thorton, J. M. (1987) Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* **326**, 347-352

Bohm, H.-J. and Klebe, G. (1996) What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs. *Angew. Chem. Int. Engl.* **35**,

Bonnefoy, S., Attal, G., Langsley, G., Tekaia, F. and Mercereau-Puijalon, O. (1994) Molecular characterization of the heat shock protein 90 gene of the human malaria parasite *Plasmodium falciparum*. *Mol Biochem Parasitol* **67**, 157-70.

Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C. M., Craig, A., Davies, R. M., Devlin, K., Feltwell, T., Gentles, S. and Gwilliam, R. (1999) The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature* **400**, 532-53

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

Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254

Brooks, H. B. and Phillips, M. A. (1997) Characterisation of the reaction mechanism of Trypanosoma brucei ornithine decarboxylase by multiwavelength stopped-flow spectroscopy. *Biochemistry* **36**, 15147-15155

Brown, A. J., Reddy, S. G. and Haddox, M. K. (1994) Multisite phosphorylation of ornithine decarboxylase increases enzyme activity and intracellular stability. *Biochem. Soc. Trans.* **22**, 859-863

Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C. and Hill, R. C. (1969) A possible three-dimensional structure of bovine alpha-lactalbumin based on that of hens egg-white lysozyme. *J Mol Biol* **42**, 65-86

Burley, S. K. (2000) An overview of structural genomics. *Nature Struc. Biol.* **Struct. Genomics Suppl.**, 932-934

Bzik, D. J., Li, W. B., Horri, T. and Inselburg, J. (1987) Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. *Proceedings of the National Academy of Science USA* **84**, 8360-8364

Carucci, D. J., Witney, A. A., Muhia, D. K., Warhurst, D. C., Schaap, P., Meima, M., Li, J.-L., Taylor, M. C., Kelly, J. M. and Baker, D. A. (2000) Guanylyl cyclase activity associated with putative bifunctional integral membrane proteins in *Plasmodium falciparum*. *J. Biol. Chem.* **275**, 22147-22156

Chang, S. P. (1994) Expression systems to best mimick the native structure. *Am J Trop Med Hyg* **50**, 20-6

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

Clarke, I. A. and Schofield, L. (2000) Pathogenesis of Malaria. *Parasitol. Today* **16**, 451-454

Clarke, J. L., Scopes, D. A., Sodeinde, O. and Mason, P. J. (2001) Glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase. A novel bifunctional enzyme in malaria parasites. *Eur J Biochem* **268**, 2013-9.

Clyde, D. F., McCarthy, V. C., Miller, R. M. and Woodward, W. E. (1975) Immunisation of man against falciparum and vivax malaria by use of attenuated sporozoites. *Am. J. Trop. Med. Hyg.* **24**, 397-401

Cohen, S. S. (1998) In: *A guide to the polyamines.* Oxford University Press, Oxford

Coleman, C. S., Stanley, B. A. and Pegg, A. E. (1993) Effects of mutations at active site residues on the activity of ornithine decarboxylase and its inhibition by active site directed irreversible inhibitors. *J. Biol. Chem* **268**, 24572-24579

Coleman, C. S., Stanley, B. A., Viswanath, R. and Pegg, A. E. (1994) Rapid exchange of subunits of mammalian ornithine decarboxylase. *J. Biol. Chem.* **269**, 3155-3158

Collins, F. H. and Paskewitz, S. M. (1995) Malaria: Current and Future Prospects of Control. *Annu. Rev. Entomol.* **40**, 195-219

Cooke, B. M., Wahlgren, M. and Coppel, R. L. (2000) Falciparum malaria: sticking up, standing out and out-standing. *Parasitol. Today* **16**, 416-420

Coppel, R. L. and Black, C. G. (1998) Malaria Parasite DNA. In: *Malaria: parasite biology, pathogenesis and protection.* Sherman, I. W., ASM Press, Washington, D. C., Coppel, R. L. and Black, C. G.

Cox, F. E. G. (1993) In: *Modern Parasitology: A textbook of parasitology.* Blackwell Scientific, London

Craig, A. and Scherf, A. (2001) Molecules on the surface of the *Plasmodium falciparum* infected erythrocyte and their role in malaria pathogenesis and immune evasion. *Mol. Biochem. Parasitol.* **115**, 129-143

Das, B., Gupta, R. and Madhubala, R. (1997) Combined action of inhibitors of S-adenosylmethionine decarboxylase with an antimalarial drug, chloroquine, on *Plasmodium falciparum.* *J. Euk. Microbiol.* **44**, 12-17

Desowitz, R. S. (1991) In: *The Malaria Capers (More tales of people, research and reality).* W.W. Norton & Company, New York

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

Dunn, C. R., Banfield, M. J., Barker, J. J., Highham, C. W., Moreton, K. M., Turgut-Balik, D., Brady, R. L. and Holbrook, J. J. (1996) The structure of lactate dehydrogenase from *Plasmodium falciparum* reveals a new target for antimalarial design. *Nat. Struct. Biol.* **3**, 912-915

Edwards, J. B. D. M., Ravassard, P., Icardi-Liepkalns, C. and Mallet, J. (1995) cDNA cloning by RT-PCR. In: *PCR2: A practical approach.* McPherson, M. J., Hames, B. D. and Taylor, G. R., IRL Press, Oxford, 89-118Edwards, J. B. D. M., Ravassard, P., Icardi-Liepkalns, C. and Mallet, J.

Edwards, M. L., Stemerick, D. M., Bitoni, A. J., Dumont, J. A., McCann, P. P., Bey, P. and Sjoerdsma, A. (1991) Antimalarial polyamine analogues. *J. Med. Chem.* **34**, 569-574

Ekstrom, J. L., Mathews, I. I., Stanley, B. A., Pegg, A. E. and Ealick, S. E. (1999) The crystal structure of human S-adenosylmethionine decarboxylase at 2.25A resolution reveals a novel fold. *Structure Fold Des.* **7**, 583-595

Ekstrom, J. L., Tolbert, W. D., Xiong, H., Pegg, A. E. and Ealick, S. E. (2001) Structure of a human S-adenosylmethionine decarboxylase self-processing ester intermediate and mechanism of putrescine stimulation of processing as revealed by the H243A mutant. *Biochemistry* **40**, 9495-9504

Facer, C. A. and Tanner, M. (1997) Clinical trails of malaria vaccines: Progress and Prospects. *Adv. Parasitol.* **39**, 1-68

Fairlamb, A. H. (1989) Novel biochemical pathways in parasitic protozoa. *Parasitol.* **S99**, S93-S112

Fairlamb, A. H. (2002) Metabolic pathway analysis in trypanosomes and malaria parasites. *Phil. Trans. R. Soc. Lond.* **357**, 101-107

Fasel, N., Begdadi-Rais, C., Bernard, M., Bron, C., Corradin, G. and Reymond, C. D. (1992) *Dictyostelium discoideum* as an expression host for the circumsporozoite protein of *Plasmodium falciparum. Gene* **111**, 157-163

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

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

Frontali, C. (1994) Genome plasticity in *Plasmodium. Genetica* **94**, 91-100

Fujioka, H. and Aikawa, M. (1999) The malaria parasite and its life-cycle. In: *Malaria: Molecular and clinical aspects.* Wahlgren, M. and Perlman, P., Harwood Academic Publishers, Amsterdam, Fujioka, H. and Aikawa, M.

Fukomoto, G. H. and Byus, C. V. (1996) A kinetic characterisation of putrescine and spermidine uptake and export in human erythrocytes. *Biochem. Biophys. Acta* **1282**, 48-56

Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Aravind, L., Koonin, E. V., Shallom, S., Mason, P. J., Yu, K., Fujii, C., Pederson, J. and Shen, K. (1998) Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum. Science* **282**, 1126-1132

Gasteiger, J., Rudolph, C. and Sadowski, J. (1990) Automatic generation of 3-D atomic coordinates for organic molecules. *Tetrahedron Comp. Method.* **3**, 537-547

Geourjon, C., Deleage, G. and Roux, B. (1991) ANTHEPROT: an interactive graphics software for analysing protein structures from sequences. *J. Mol. Graph.* **9**, 188-190

Ghoda, L., Phillips, M. A., Bass, K. E., Wang, C. C. and Coffino, P. (1990) Trypanosome ornithine decarboxylase is stable because it lacks sequences found in the carboxy terminus of the mouse enzyme which target the latter for intracellular degradation. *J. Biol. Chem.* **265**, 11823-11826

Giesecke, H., Barale, J. C., Langsley, G. and Cornelissen, A. W. (1991) The C-terminal domain of RNA polymerase II of the malaria parasite *Plasmodium berghei*. *Biochem Biophys Res Commun* **180**, 1350-5.

Gilberger, T. W., Schirmer, R. H., Walter, R. D. and Muller, S. (2000) Deletion of the parasite-specific insertions and mutation of the catalytic triad in glutathione reductase from chloroquine-sensitive *Plasmodium falciparum* 3D7. *Mol Biochem Parasitol* **107**, 169-79.

Gillman, A. G., Goodman, L. S., Rall, T. W. and Murad, F. (1985) In: *The Pharmacological basis of Therapeutics*. Macmillan Publishing Company, New York

Greenwood, B. and Mutabingwa, T. (2002) Malaria in 2002. *Nature* **415**, 670-672

Grishin, N. V., Osterman, A. L., Brooks, H. B., Phillips, M. A. and Goldsmith, E. J. (1999) X-ray structure of ornithine decarboxylase from *Trypanosoma brucei*: the native structure and the structure in complex with alpha- difluoromethylornithine. *Biochemistry* **38**, 15174-84

Guex, N., Diemand, A. and Peitsch, M. C. (1999) Protein modelling for all. *Trends Biochem Sci* **24**, 364-7

Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-23

Guex, N. and Peitsch, M. C. (1999) Molecular modelling of proteins. *Immunology News* **6**, 132-134

Hanahan, D., Jessee, J. and Bloom, F. R. (1991) Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol.* **204**, 63-114

Hanson, S., Adelman, J. and Ullman, B. (1992) Amplification and molecular cloning of the ornithine decarboxylase gene of *Leishmania donovani*. *J. Biol. Chem.* **267**, 2350-2359

Hayashi, S. (1989) Multiple mechanisms for the regulation of mammalian ornithine decarboxylase. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. Hayashi, S., Pergamon Press, Inc., Oxford, England, 35-45Hayashi, S.

Hayashi, S. and Canellakis, E. S. (1989) Ornithine decarboxylase antizymes. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. Hayashi, S., Pergamon Press, Inc., Oxford, England, 47-57Hayashi, S. and Canellakis, E. S.

Hayashi, S. and Murakami, Y. (1995) Rapid and regulated degradation of ornithine decarboxylase. *Biochem. J.* **306**, 1-10

Hayashi, S., Murakami, Y. and Matsufuji, S. (1996) Ornithine decarboxylase antizyme: a novel type of regulatory protein. *Trends Biochem. Sci.* **21**, 27-30

Heby, O. (1985) Ornithine decarboxylase as target of chemotherapy. *Adv. Enzyme Regul.* **24**, 103-124

Heby, O. (1989) Polyamines and cell differentiation. In: *The physiology of polyamines*. Bachrach, U., CRC Press, Boca Raton, 83-93Heby, O.

Heby, O. and Persson, L. (1990) Molecular genetics of polyamine synthesis in eukaryotic cells. *Trends Biochem. Sci.* **15**, 153-158

Heller, J. S., Fong, W. F. and Canellakis, E. S. (1976) Induction of a protein inhibitor to ornithine decarboxylase by the end products of its reaction. *Proc. Natl. Acad. Sci. USA* **73**, 1858-1862

Henikoff, S. and Henikoff, J. G. (1994) Protein family classification based on searching a database of blocks. *Genomics* **19**, 97-107

Hogh, B., Thompson, R., Zakiuddin, I. S., Boudin, C. and Borre, M. (1993) Glutamate rich *Plasmodium falciparum* antigen (GLURP). *Parasitologia* **35(S)**, 47-50

Hollingdale, M. R., McCann, P. P. and Sjoerdsma, A. (1985) *Plasmodium berghei*: Inhibitors of ornithine decarboxylase block exoerythrocytic schizogony. *Exp. Parasitol.* **60**, 111-117

Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138

Hommel, M. (1997) Modulation of host cell receptors: a mechanism for the survival of malaria parasites. *Parasitology* **115**, S45-S54

Horrocks, P., Dechering, K. and Lanzer, M. (1998) Control of gene expression in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **95**, 171-181

Huang, X. and Miller, W. (1991) A time-efficient linear-space local similarity algorithm. *Adv Applied Math* **12**, 337

Hubbard, S. J. (1998) The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta* **1382**, 191-206

Hyde, J. E. (2002) Mechanisms of resistance of *Plasmodium falciparum* to antimalarial drugs. *Microbes Infect.* **4**, 165-174

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

Hyde, J. E., Kelly, S. L. and Holloway, S. P. (1989) A general approach to isolating *Plasmodium falciparum* genes using non-redundant oligonucleotides inferred from protein sequences of other organisms. *Mol. Biochem. Parasitol.* **32**, 247-262

Ivanetich, K. M. and Santi, D. V. (1990) Bifunctional thymidilate synthase-dihydrofolate reductase in protozoa. *FASEB Journal* **4**, 1591-1597

Janne, J. and Alhonen-Hongisto, L. (1989a) Inhibitors of ornithine decarboxylase: Biochemistry and applications. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics.* Hayashi, S., Pergamon Press, Inc., Oxford, England, 59-85Janne, J. and Alhonen-Hongisto, L.

Janne, J. and Alhonen-Hongisto, L. (1989b) Inhibitors of polyamine biosynthesis as therapeutic targets. In: *The Physiology of polyamines.* Bachrach, U., CRC Press, Boca Raton, 251-286Janne, J. and Alhonen-Hongisto, L.

Janne, J., Alhonen-Hongisto, L., Nikula, P. and Elo, H. (1985) S-adenosylmethionine decarboxylase as target of chemotherapy. *Adv. Enzyme Regul.* **24,** 125-139

Jiang, Y. (1999) Ornithine decarboxylase gene deletion mutants of *Leishmania donovani. J. Biol. Chem.* **274,** 3781-3788

Jones, D. T., Oregano, C. A. and Thorton, J. M. (1996) Protein folds and their recognition from sequence. In: *Protein Structure Prediction.* Sternberg, M. J. E., Oxford University Press, Oxford, Jones, D. T., Oregano, C. A. and Thorton, J. M.

Joubert, F. (2000) Structural modelling of therapeutic targets and inhibitors of the malaria parasite. *Biochemistry.* University of Pretoria. 124

Kahana, C. (1989) Molecular genetics of mammalian ornithine decarboxylase. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics.* Hayashi, S., Pergamon Press, Inc., Oxford, England, 127-133Kahana, C.

Kahana, C., Berovich, Z., Erez, O., Gandre, S. and Wender, N. (2002) Regulation of intracellular polyamines, polycations that are essential for cellular viability and proliferation. *Cell Developm. Biol.* 110-111

Kahana, C. and Nathans, D. (1984) Isolation of cloned cDNA encoding mammalian ornithine decarboxylase. *Proc. Nat. Acad. Sci. USA* **81,** 3645-3649

Kappes, B., Doerig, C. D. and Graeser, R. (1999) An overview of Plasmodium protein kinases. *Parasitol Today* **15,** 449-54.

Karcher, S. J. (1995) In: *Molecular biology: A project approach.* Academic Press, Inc., San Diego

Kay, R. R. and Williams, J. G. (1999) The *Dictyostelium* genome project: an invitation to species hopping. *Trends Genet.* **15,** 294-297

Kaye, A. M. (1984) Ornithine decarboxylase: Purification and properties of ornithine decarboxylase. *Cell Biochem. Funct.* **2,** 2-5

Kemp, D. J., Coppel, R. L. and Anders, R. F. (1987) Repetitive genes and proteins in malaria. *Annu. Rev. Microbiol.* **41,** 181-208

Kemp, D. J., Cowman, A. F. and Walliker, D. (1990) Genetic diversity in *Plasmodium falciparum. Adv. Parasitol.* **29,** 75-149

Kern, A. D., Oliveira, M. A., Coffino, P. and Hackert, M. L. (1999) Structure of mammalian ornithine decarboxylase at 1.6 A resolution: stereochemical implications of PLP-dependent amino acid decarboxylases. *Structure Fold Des* **7,** 567-81

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

Krause, T., Luersen, K., Wrenger, C., Gilberger, T. W., Müller, S. and Walter, R. D. (2000) The ornithine decarboxylase domain of the bifunctional ornithine decarboxylase/S-

adenosylmethionine decarboxylase of *Plasmodium falciparum*: recombinant expression and catalytic properties of two different constructs. *Biochem J* **352** Pt 2, 287-92.

Krogstad, D. J. (1996) Malaria as a reemerging disease. *Epidem. Rev.* **18**, 77-89

Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Nat. Acad. Sci. USA* **82**, 488-492

Kuntz, I. D. (1992) Structure-based strategies for drug design and discovery. *Science* **257**, 1078-1082

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

Lang-Unnasch, N. and Murphy, A. D. (1998) Metabolic changes of the malaria parasite during the transition from the human to the mosquito host. *Annu. Rev. Microbiol.* **52**, 561-590

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

Lanzer, M., Wertheimer, S. P., De Bruin, D. and Ravetch, J. V. (1993) *Plasmodium*: Control of gene expression in malaria parasites. *Exp. Parasitol.* **77**, 121-128

Laskowski, R. A., MacAthur, M. W., Moss, D. S. and Thorton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **29**, 283-291

Latchman, D. (1995) In: *Gene regulation: a eukaryotic prespective.* Chapman and Hall, Oxford

Lehninger, A. L. (1975) In: *Principles of Biochemistry.* Worth Publishers, New York

Lemcke, T., Christensen, I. T. and Jorgensen, F. S. (1999) Towards an understanding of drug resistance in malaria: three- dimensional structure of *Plasmodium falciparum* dihydrofolate reductase by homology building. *Bioorg Med Chem* **7**, 1003-11.

Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R. R., Panting, C. P. and Bark, P. (2002) Recent improvements to the SMART domain-based sequence annotation recource. *Nucl Acid. Res.* **30**, 242-244

Li, J.-L. and Baker, D. A. (1998) A putative protein serine/threonine phosphatase from *Plasmodium falciparum* contains a large N-terminal extension and five unique inserts in the catalytic domain. *Mol. Biochem. Parasitol.* **95**, 287-295

Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Mastudaira, P. and Darnell, J. (1995) In: *Molecular cell biology.* Scientific American Books, Inc., U.S.A.

Lorenzi, E. C. and Scheffler, I. E. (1997) Co-operation of the 5' and 3' untranslated regions of ornithine decarboxylase mRNA and inhibitory role of its 3' untranslated region in regulating the translational efficiency of hybrid RNA species via cellular factors. *Biochem. J.* **326**, 361-367

Lu, L., Stanley, B. A. and Pegg, A. E. (1991) Identification of residues in ornithine decarboxylase essential for enzymatic activity and for rapid protein turnover. *Biochem. J.* **277**, 671-675

Luersen, K., Walter, R. D. and Muller, S. (1999) The putative gamma-glutamylcysteine synthetase from *Plasmodium falciparum* contains large insertions and a variable tandem repeat. *Mol Biochem Parasitol* **98**, 131-42.

Luersen, K., Walter, R. D. and Müller, S. (2000) *Plasmodium falciparum* infected red blood cells depend on a functional glutathione de novo synthesis attributable to an enhanced loss of glutathione. *Biochem. J.* **346**, 545-552

Lukyanov, K., Diatchenko, L., Chenchik, A., Nanisetti, A., Siebert, P., Usman, N., Matz, M. and Lukyanov, S. (1997) Construction of cDNA libraries from small amounts of total RNA using the suppression PCR effect. *Biochem. Biophys. Res. Comm.* **230**, 285-288

Macreadie, I., Ginsburg, H., Sirawaraporn, W. and Tilley, L. (2000) Antimalarial drug developments and new targets. *Parasitol. Today* **16**, 438-444

Makrides, S. C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **60**, 512-538

Mamroud-Kidron, E., Omer-Itsicovich, M., Bercovich, Z., Tobias, K. E., Rom, E. and Kahana, C. (1994) A unified pathway for the degradation of ornithine decarboxylase in reticulocyte lysate requires interaction with the polyamine-induced protein, ornithine decarboxylase antizyme. *Eur. J. Biochem.* **226**, 547-554

Marsh, K. (1999) Clinical features of malaria. In: *Malaria: Molecular and clinical aspects*. Wahlgren, M. and Perlman, P., Harwood Academic Publishers, Amsterdam, Marsh, K.

Marti-Renom, M. A., Stuart, A., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**, 291-325

Matsuoka, H., Kobayashi, J., Barker, G. C., Miura, K., Chizei, Y., Miyajima, S., Ishii, A. and Sinden, R. E. (1996) Induction of anti-malarial transmission blocking immunity with a recombinant ookinete surface antigen of *Plasmodium berghei* produced in silk-worm larvae using the baculovirus expression vector system. *Vaccine* **14**, 120-126

McCann, P. P. and Pegg, A. E. (1992) Ornithine decarboxylase as an enzyme target for therapy. *Pharmacology and Therapeutics* **54**, 195-215

McConkey, G. A. (1999) Targeting the shikimate pathway in the malaria parasite, *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **43**, 171-177

McKerrow, J. H., Rosenthal, P. J., Sun, E. and Bouvier, J. (1993) The proteases and pathogenicity of parasitic protozoa. *Annu. Rev. Microbiol.* **47**, 821-853

Meierjohann, S., Walter, R. D. and Müller, S. (2002) Gluthathione synthetase from *Plasmodium falciparum*. *Biochem. J.* **363**, 833-838

Mendis, K. N. and Carter, R. (1995) Clinical disease and pathogenesis in malaria. *Parasitol. Today* **11**, PTI2-PTI16

Merril, C. R., Goldman, D., Sedman, S. A. and Ebert, M. H. (1981) Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science* 211, 1437-1438

Mett, H., Standek, J., Lopez-Ballester, J. A., Janne, J., Alhonene, L., Sinervirta, R., Frei, J. and Renegrass, U. (1993) Pharmacological properties of the ornithine decarboxylase inhibitor 3-aminooxy-1-propanamine and several structural analogues. *Cancer Chemother Pharmacol* 32, 39-45

Michelitsch, M. D. and Weissman, J. S. (2000) A sensus of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions. *Proc. Nat. Acad. Sci. USA* 97, 11910-11915

Miles, E. W., Rhee, S. and Davies, R. M. (1999) The molecular basis of substrate channeling. *J. Biol. Chem.* 274, 12193-12196

Milhous, W. K. and Kyle, D. E. (1998) Introduction to the modes of action of and mechanisms of resistance to antimalarials. In: *Malaria: Parasite biology, pathogenesis and protection.* Sherman, I. W., ASM Press, Washington, D. C., Milhous, W. K. and Kyle, D. E.

Miller, L. H., Baruch, D. I., Marsh, K. and Doumbo, O. K. (2002) The pathogenic basis of malaria. *Nature* 415, 673-679

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

Mishra, M., Chandra, S., Pandey, V. C. and Tekwani, B. L. (1997) Polyamine metabolism in various tissues during pathogenesis of chloroquine-susceptible and resistant malaria. *Cell Biochem. Funct.* 15, 229-235

Momany, C., Ernst, S., Ghosh, R., Chang, N. L. and Hackert, M. L. (1995) Crystallographic structure of a PLP-dependent ornithine decarboxylase from Lactobacillus 30a to 3.0 A resolution. *J Mol Biol* 252, 643-55.

Müller, S., Coombs, G. H. and Walter, R. D. (2001) Targeting polyamines of parasitic protozoa in chemotherapy. *Trends Parasitol.* 17, 242-249

Müller, S., Da'dara, A., Luersen, K., Wrenger, C., Das Gupta, R., Madhubala, R. and Walter, R. D. (2000) In the human malaria parasite *Plasmodium falciparum*, polyamines are synthesized by a bifunctional ornithine decarboxylase, S- adenosylmethionine decarboxylase. *J Biol Chem* 275, 8097-102

Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247,

Myers, D., Jackson, L. K., Ipe, V. G., Murphy, G. E. and Phillips, M. A. (2001) Long-range interactions in the dimer interface of ornithine decarboxylase are important for enzyme function. *Biochemistry* 40, 13230-13236

Nicholls, A., Sharp, K. A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281-96

Nishimura, K., Liisanantti, M., Yasuhide, M., Kashiwagi, K., Shirahata, A., Janne, M., Kankare, K., Janne, O. A. and Igarashi, K. (1998) Structure and activity of mouse S-adenosylmethionine decarboxylase gene promoters and properties of the encoded proteins. *Biochem. J.* **332**, 651-659

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

Old, R. W. and Primrose, S. B. (1994) In: *Principles of gene manipulation.* Ed:Carr, N. G., Blackwell Science. Ltd., Oxford

Oliveira, M. A., Carroll, D., Davidson, L., Momany, C. and Hackert, M. L. (1997) The GTP effector site of ornithine decarboxylase from *Lactobacillus 30a*: Kinetic and structural characterisation. *Biochemistry* **36**, 16147-16154

Olliaro, P. and Yuthavong, Y. (1998) Chemotherapeutic targets in *Plasmodia* with potential for antimalarial drug discovery. *S. A. J. Sci.* **94**, 292-296

Olliaro, P. L. (2001) Mode of action and mechanisms of resistance for antimalarial drugs. *Pharmacol. Ther.* **89**, 207-219

Oregano, C. A., Jones, D. T. and Thorton, J. M. (1994) Protein superfamilies and domain superfolds. *Nature* **372**, 631-634

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. and Thorton, J. M. (1997) CATH: A hierarchic classification of protein domain structures. *Structure* **5**, 1093-1108

Osterman, A., Brooks, H. B., Jackson, L. K., Abbott, J. J. and Phillips, M. A. (1999) Lysine-69 plays a key role in catalysis of ornithine decarboxylase through acceleration of the Schiff base formation, decarboxylation and product relase steps. *Biochemistry* **38**, 11814-11826

Osterman, A., Brooks, H. B., Rizo, J. and Phillips, M. A. (1997) Role of Arg-277 in the binding of pyridoxal-5'-phosphate to *Trypanosoma brucei* ornithine decarboxylase. *Biochemistry* **36**, 4558-4567

Osterman, A., Grishin, N. V., Kinch, L. N. and Phillips, M. A. (1994) Formation of functional cross-species heterodimers of ornithine decarboxylase. *Biochemistry* **33**, 13662-13667

Osterman, A. L., Lueder, D. V., Quick, M., Myers, D., Canagarajah, B. J. and Phillips, M. A. (1995) Domain organization and a protease-sensitive loop in eukaryotic ornithine decarboxylase. *Biochemistry* **34**, 13431-6.

Pajunen, A., Croza, A., Janne, O. A., Ihalainen, R., Laithinen, P. H., Stanley, B. A., Madhubala, R. and Pegg, A. E. (1988) Structure and regulation of mammalian S-adenosylmethionine decarboxylase. *J. Biol. Chem* **32**, 17040-17049

Pan, W., Ravot, E. and Tolle, R. (1999) Vaccine candidate MSP-1 from *Plasmodium falciparum*: a redesigned 4917 bp polynucleotide enables synthesis and isolation fo full-lenght protein from *Escherichia coli* and mammalian cells. *Nucl. Acids Res.* **27**, 1094-1103

Patthy, L. (1999) In: *Protein Evolution*. Blackwell Science, Oxford

Pegg, A. E. (1989a) Characterisation of ornithine decarboxylase from various sources. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. Hayashi, S., Pergamon Press, Inc., Oxford, England, 21-28Pegg, A. E.

Pegg, A. E. (1989b) Inhibitors of ornithine and S-adenosylmethionine decarboxylases. In: *The Physiology of Polyamines*. Bachrach, U., CRC Press, Boca Raton, 303-313Pegg, A. E.

Pegg, A. E. and McCann, P. P. (1982) Polyamine metabolism and function. *Am. J. Physiol.* **243**, C212-C221

Pegg, A. E., Shantz, L. M. and Coleman, C. S. (1994) Ornithine decarboxylase: Structure, function and translational regulation. *Biochem. Soc. Trans.* **22**, 846-852

Peitsch, M. C. (1995a) ProMod: Automated knowledge-based protein modelling tool. *PDB Quarterly Newsletter* **72**, 4

Peitsch, M. C. (1995b) Protein modelling by E-mail. *Bio/Technology* **13**, 658-660

Peitsch, M. C. (1996) ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. *Biochem Soc Trans* **24**, 274-9

Peitsch, M. C., Herzyk, P., Wells, T. N. and Hubbard, R. E. (1996) Automated modelling of the transmembrane region of G-protein coupled receptor by Swiss-model. *Receptors Channels* **4**, 161-4

Perutz, M. F., Johnson, T., Suzuki, M. and Finch, J. T. (1994) Glutamine repeats as polar zippers: Their possible role in inherited neurodegenerative diseases. *Proc. Nat. Acad. Sci. USA* **91**, 5355-5358

Phillips, M. A., Coffino, P. and Wang, C. C. (1987) Cloning and sequencing of the ornithine decarboxylase gene from *Trypanosoma brucei*. *J. Biol. Chem.* **262**, 8721-8727

Phillips, R. S. (2001) Current status of malaria and potential for control. *Clin. Microbiol. Rev.* **14**, 208-226

Pizzi, E. and Frontali, C. (2001) Low-complexity regions in *Plasmodium falciparum* proteins. *Genome Res* **11**, 218-29.

Pollack, Y., Shemer, R., Metzger, S., Spira, D. T. and Golenser, J. (1985) *Plasmodium falciparum*: expression of the adenine phosphoribosyl transferase gene in mouse L cells. *Exp. Parasitol.* **60**, 270-275

Prapunwattana, P., Sirawaraporn, W., Yuthavong, Y. and Santi, D. V. (1996) Chemical synthesis fo the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. *Mol. Biochem. Parasitol.* **83**, 93-106

Prasanna, V., Bhattacharjya, S. and Balaram, P. (1998) Synthetic interface peptides as inactivators of multimeric enzymes: inhibitory and conformational properties of three fragments from *Lactobacillus casei* thymidylate synthase. *Biochemistry* **37**, 6883-6893

Preston, G. M. (1993) Use of degenerate oligonucleotide primers and their PCR to clone gene family members. *Methods Mol. Biol.* **15**, 317-337

Pulkka, A., Ihalainen, R., Aatsinki, J. and Pajunen, A. (1991) Structure and organization of the gene encoding rat S-adenosylmethionine decarboxylase. *FEBS Letters* **291**, 289-295

Ramasamy, R. (1991) Repeat regions in malaria parasite proteins: a review on structure and possible role in the biology of the parasite. *Indian J. Malariol.* **28**, 73-81

Ramasamy, R. (1998) Molecular basis for evasion of host immunity and pathogenesis in malaria. *Biochim. Biophys. Acta* **1406**, 10-27

Rand, K. N. (1996) Crystal Violet can be used to visualise DNA bands during electrophoresis and to improve cloning efficiency. *Elsevier Trends Journal* **T40022,**

Raney, A., Baron, A. C., Mize, G., Law, L. and Morris, D. R. (2000) In vitro translation of the upstream open reading frame in the mammaila mRNA encoding S-adenosylmethionine decarboxylase. *J. Biol. Chem* **275**, 24444-244450

Rastelli, G., Sirawaraporn, W., Sompompisut, P., Vilaivan, T., Kamchonwongpaisan, S., Ouarrel, R., Lowe, G., Thebtaranonth, Y. and Yuthavong, Y. (2000) Interaction of pyrimethamine, cycloguanil, WR99210 and their analogues with *Plasmodium falciparum* dihydrofolate reductase: Structural basis of antifolate resistance. *Bioorg Med Chem* **8**, 1117-1128

Rathaur, S. and Walter, R. D. (1987) *Plasmodium falciparum*: S-adenosylmethionine decarboxylase. *Exp. Parasitol.* **63**, 227-232

Reddy, S. G., McIlheran, S. M., Cochran, B. J., Worth, L. L., Bishop, L. A., Brown, P. J., Knutson, V. P. and Haddox, M. K. (1996) Multisite phosphorylation of ornithine decarboxylase in transformed macrophages results in increased intracellular enzyme stability and catalytic efficiency. *J. Biol. Chem.* **271**, 24945-24953

Reeder, J. C. and Brown, G. V. (1996) Antigenic variation and immune evasion in *Plasmodium falciparum*. *Immunol. Cell Biol.* **74**, 546-554

Richie, T. L. and Saul, A. (2002) Progress and challenges for malaria vaccines. *Nature* **415**, 694-701

Ridley, R. (2002) Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415**, 686-693

Ring, C. S., Sun, E., McKerrow, J. H., Lee, G. K., Rosenthal, P. J., Kuntz, I. D. and Cohen, F. D. (1993) Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc. Nat. Acad. Sci. USA* **90**, 3583-3587

Rogers, S. and Hoffman, S. L. (1999) Malaria vaccines. In: *Malaria: Molecular and clinical aspects*. Wahlgren, M. and Perlman, P., Harwood Academic Publishers, Amsterdam, Rogers, S. and Hoffman, S. L.

Rogers, S., Wells, R. and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science* **234**, 364-369

Rosenthal, P. J. and Meshnick, S. R. (1998) Hemoglobin processing and the metabolism of amino acids, heme and iron. In: *Malaria: Parasite biology, pathogenesis and protection.* Sherman, I. W., ASM Press, Washington, D. C., Rosenthal, P. J. and Meshnick, S. R.

Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.* **266**, 525-539

Rozmajzl, P. J., Kimura, M., Woodrow, C. J., Krishna, S. and Meade, J. C. (2001) Characterisation of the P-Type ATPase 3 in *Plasmodium falciparum. Mol. Biochem. Parasitol.* **116**, 117-126

Russell, D. H. (1983) Ornithine decarboxylase may be a multifunctional protein. *Adv. Enzyme Regul.* **21**, 201-222

Rychlik, W. and Rhoades, R. E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridisation, sequencing and *in vitro* amplification of DNA. *Nucl. Acid. Res.* **17**, 8543-8551

Rychlik, W., Spencer, W. J. and Rhoads, R. E. (1990) Optimisation of the annealing temperature for DNA amplification *in vitro. Nucl. Acids Res.* **18**, 6409-6412

Sali, A. (1995) Comparative protein modeling by satisfaction of spacial restraints. *Mol Med Today* **6**, 270-277

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) In: *Molecular cloning: A laboratory manual.* Cold Spring Harbour Laboratory Press, Cold Spring Harbour

Sanchez, J. C., Pieper, U., Melo, F., Eswar, N., Marti-Renom, M. A., Madhusudhan, M. S., Mirkovic, N. and Sali, A. (2000) Protein structure modeling for structural genomics. *Nature Struc. Biol.* **Struct. Genomics Suppl.**, 986-990

Sanchez, R. and Sali, A. (1997) Advances in comparative protein-structure modelling. *Current Opinion Struc Biol* **7**,

Saul, A. and Battistutta, D. (1988) Codon usage in *Plasmodium falciparum. Mol. Biochem. Parasitol.* **27**, 35-42

Sayers, J. R., Price, H. P., Fallon, P. G. and Doenhoff, M. J. (1995) AGA/AGG codon usage in parasites: Implications for gene expression in *Escherichia coli. Parasitol. Today* **11**, 345-346

Scheafer, B. C. (1995) Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. *Ann. Biochem.* **227**, 255-273

Scherf, A., Bottius, E. and Hernandes-Rivas, R. (1999) The Malaria Genome. In: *Malaria: Molecular and Clinical Aspects.* Wahlgren, M. and Perlmann, P., Harwood Academic Publishers, Amsterdam, Scherf, A., Bottius, E. and Hernandes-Rivas, R.

Schofield, L. (1991) On the function of repetitive domains in protein antigens of *Plasmodium* and other eukaryotic parasites. *Parasitol. Today* **7**, 99-105

Schramm, H. J., Boetzel, J., Buttner, J., Fritsche, E., Gohring, W., Jaeger, E., Konig, S., Thumfart, O., Wenger, T., Nager, N. E. and Schramm, W. (1996) The inhibition of human immunodeficiency virus proteases by 'interface peptides'. *Antiviral Res.* **30**, 155-170

Shantz, L. M., Viswanath, R. and Pegg, A. E. (1994) Role of the 5'-untranslated region of mRNA in the synthesis of S-adenosylmethionine decarboxylase and its regulation by spermine. *Biochem. J.* **302**, 765-772

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

Sherman, I. W. (1998a) A brief history of malaria and discovery of the parasite's life cycle. In: *Malaria: Parasite biology, pathogenesis and protection.* Sherman, I. W., ASM Press, Washington, D. C., Sherman, I. W.

Sherman, I. W. (1998b) Carbohydrate metabolism of asexual stages. In: *Malaria: Parasite biology, pathogenesis and protection.* Sherman, I. W., ASM Press, Washington, D. C., Sherman, I. W.

Shi, W., Li, C. M., Tyler, P. C., Furneux, R. H., Cahill, S. M., Girvin, M. E., Grubmeyer, C., Schramm, V. L. and Almo, S. C. (1999) The 2.0 A structure of malarial purine phosphoribosyltransferase in complex with a transition-state analogue inhibitor. *Biochemistry* **38**, 9872-9880

Sibley, C. H., Hyde, J. E., Sims, P. F. G., Plowe, C. V., Kublin, J. G., Mberu, E. K., Cowman, A. F., Winstanley, P. A., Watkins, W. M. and Nzila, A. M. (2001) Pyrimethamine-sulphadoxine resistance in *Plasmodium falciparum*: what next. *Trends. Parasitol.* **17**, 582-588

Silverman, R. B. (1988) The potential use of mechanism-based enzyme inactivators in medicine. *J. Enzyme Inhibition* **2**, 73-90

Singh, S., Puri, S. K., Singh, S. K., Srivastava, R., Gupta, R. C. and Padney, V. C. (1997) Characterisation of simian malaria parasite (*Plasmodium knowlesi*)-induced putrescine transport in rhesus monkey erythrocytes. *J. Biol. Chem.* **272**, 13506-13511

Singh, S. K., Maithal, K., Balaram, H. and Balaram, P. (2001) Synthetic peptides as inactivators of multimeric enzymes: inhibition of *Plasmodium falciparum* triosephosphate isomerase by interface peptides. *FEBS Letters* **501**, 19-23

Slater, L. A., McMonagle, F. A., Phillips, R. S. and Robins, D. J. (1998) Antimalarial activity of unsaturated putrescine derivatives. *Ann. Trop. Med. Parasitol.* **92**, 271-277

Smith, S. W., Overbeek, R., Woese, C. R., Gilbert, W. and Gillevet, P. M. (1994) The genetic data environment and expandable GUI for multiple sequence-analysis. *Comput. Appl. Biosci.* **10**, 671-675

Srinivasan, N., Anuradha, V. S., Ramakrishnan, C., Sowdhamini, R. and Balaram, P. (1994) Conformational characteristics of asparaginyl residues in proteins. *Int. J. Peptide Prot. Res.* **44**, 112-122

Srinivasan, N., Guruprasad, K. and Blundell, T. L. (1996) Comparative modelling of proteins. In: *Protein Structure Prediction*. Sternberg, M. J. E., Oxford University Press, Oxford, Srinivasan, N., Guruprasad, K. and Blundell, T. L.

Stahl, H. D., Kemp, D. J., Scalon, D. B., Woodrow, G., Brown, G. V., Bianco, A. E., Anders, R. F. and Coppel, R. L. (1985) Sequence of a cDNA encoding a small polymorphic histidine- and alanine-rich protein from *Plasmodium falciparum*. *Nucl. Acids Res.* **13**, 7837-7846

Standek, J., Frei, J., Schneider, P. and Regenass, U. (1992) 2-Substituted 3-(aminoozy)propanamines as inhibitors of ornithine decarboxylase: Synthesis and biological activity. *J Med Chem* **35**, 1339-1344

Stanley, B. A. and Pegg, A. E. (1991) Amino acid residues necessary for putrescine stimulation of human S-adenosylmethionine decarboxylase processing and catalytic activity. *J. Biol. Chem* **266**, 18502-18506

Stanley, B. A., Shantz, L. M. and Pegg, A. E. (1994) Expression of mammalian S-adenosylmethionine decarboxylase in *Escherichia coli*. *J. Biol. Chem.* **269**, 7901-7907

Stemmer, W. P. C. (1994) DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**, 10747-10751

Sternberg, M. J. E. (1996) Protein structure prediction-principles and approaches. In: *Protein structure prediction: a practical approach*. Sternberg, M. J. E., Oxford University Press, Oxford, Sternberg, M. J. E.

Su, X. and Wellems, T. E. (1997) *Plasmodium falciparum*: a rapid DNA fingerprinting method using microsatellite sequences within *var* clusters. *Exp. Parasitol.* **86**, 235-236

Su, X. and Wellems, T. E. (1998) Genome discovery and malaria research: current status and promise. In: *Malaria: Parasite biology, pathogenesis and protection*. Sherman, I. W., ASM Press, Washington, D. C., Su, X. and Wellems, T. E.

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

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

Svensson, F., Ceriani, C., Lovkvist, E., Kockum, I., Algranati, I. D., Heby, O. and Persson, L. (1997) Cloning of a trypanosomatid gene coding for an ornithine decarboxylase that is metabolically unstable even though it lacks the C-terminal degradation domain. *Proc. Natl. Acad. Sci.* **94**, 397-402

Tabor, C. W. and Tabor, H. (1984a) Methionine adenosyltransferase (S-adenosylmethionine synthetase) and S-adenosylmethionine decarboxylase. *Adv. Enzymology* **56**, 251-282

Tabor, C. W. and Tabor, T. (1984b) Polyamines. *Annu. Rev. Biochem.* **53**, 749-790

Taylor, W. R. (2002) A periodic table for protein structures. *Nature* **416**, 657-660

Thomas, J. G., Ayling, A. and Baneyx, F. (1997) Molecular Chaperones, folding catalysts and the recovery of active recombinant proteins from *E. coli*. *Appl. Biochem. Biotechnol.* **66**, 197-238

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680

Thorton, J. M., Todd, A. E., Milburn, E., Borkakoti, N. and Oregano, C. A. (2000) From structure to function: Approaches and limitations. *Nature Struc. Biol.* **Struc. Genom. Suppl.**, 991-994

Tobias, K. E. and Kahana, C. (1993a) Intersubunit location of the active site of mammalian ornithine decarboxylase as determined by hybridisation of site-directed mutants. *Biochemistry* **32**, 5842-5847

Torii, M. and Aikawa, M. (1998) Ultrastructure of asexual stages. In: *Malaria: Parasite biology, pathogenesis and protection*. Sherman, I. W., ASM Press, Washington, D. C., Torii, M. and Aikawa, M.

Toyoda, T., Brobey, R. K., Sano, G., Horii, T., Tomioka, N. and Itai, A. (1997) Lead discovery of inhibitors of the dihydrofolate reductase domain of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Biochem Biophys Res Commun* **235**, 515-9.

Trager, W. (1994) Cultivation of malaria parasites. *Methods cell biol.* **45**, 7-26

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

Trigg, P. I. and Kondrachine, A. V. (1998) The current global malaria situation. In: *Malaria: Parasite biology, pathogenesis and protection*. Sherman, I. W., ASM Press, Washington, D. C., Trigg, P. I. and Kondrachine, A. V.

Triglia, T. and Cowman, A. F. (1994) Primary structure and expression of the dihydropteroate synthase gene of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **91**, 7149-7153

Velanker, S. S., Ray, S. S., Gokhale, R. S., Balaram, S. S., Balaram, H. and Murthy, M. R. (1997) Triosephosphate isomerase from Plasmodium falciparum: the crystal structure provides insights into antimalarial drug design. *Structure* **5**, 751-761

Vitali, J., Carroll, D., Chaudhry, R. G. and Hackert, M. L. (1999) Three-dimensional structure of the Gly121Tyr dimeric form of ornithine decarboxylase from Lactobacillus 30a. *Acta Crystallogr D Biol Crystallogr* **55**, 1978-85

Vriend, G. (1990) WHAT IF: a molecular modelling and drug design program. *J. Mol. Graph.* **8**, 52-59

Wahlgren, M., Bejarano, M. T., Troye-Blomberg, M., Perlman, P., Riley, E., Greenwood, B. M., Patarroyo, M. E., Gonzales, C. I. and Martinez, A. (1991) Epitopes of the

Plasmodium falciparum clustered-asparagine-rich protein (CARP) recognised by human T-cells and antibodies. *Parasite Immunol.* **13**, 681-694

Wallace, A. C., Laskowski, R. A. and Thorton, J. M. (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot Eng* **8**, 127-134

Walter, A. E., Turner, D. H., Kim, D., Lyttle, M. H., Muller, P. and Mathews, D. H. (1994) Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Nat. Acad. Sci. USA* **91**, 9218-9222

Walters, W. P., Stahl, M. T. and Murcko, M. A. (1998) Virtual screening-an overview. *Drug Discov. Today* **3**, 160-178

Wang, C. C. (1997) Validating targets for antiparasite chemotherapy. *Parasitology* **114**, S31-S44

Wang, J., Kim, S. and Gallagher, S. (1995) Dealing with A/T content differences when using the H33258/TKO 100 DNA assay. *Hoefer news* **3**,

Warhurst, D. C. (1998) Antimalaria drug discovery: development of inhibitors of dihydrofolate reductase active in drug resistance. *Drug Discov Today* **3**, 538-546

Weickert, M. J., Doherty, D. H., Best, E. A. and Olins, P. O. (1996) Optimisation of heterologous protein production in *Escherichia coli*. *Curr. Opin. Biotechnol.* **7**, 494-499

Wellems, T. E., Panton, l. J., Gluzman, I. Y., do Rosario, R. V., Gwadz, R. W., Walker, J. A. and Krogstad, D. J. (1990) Chloroquine resistance not linked to *mdr*-like genes in a *Plasmodium falciparum* cross. *Nature* **345**, 253-255

Wellems, T. E. and Plowe, C. V. (2001) Chloroquine resistant malaria. *J. Infect. Diseases* **184**, 770-776

Whaun, J. M. and Brown, N. D. (1985) Ornithine decarboxylase inhibition and the malaria-infected red cell: a model for polyamine metabolism and growth. *J. Pharmacol. Exp. Therapeutics* **233**, 507-511

White, J. H. and Kilbey, B. J. (1996) DNA replication in the malaria parasite. *Parasitol. Today* **12**, 151-155

White, N. J. (1998) Malaria Pathophysiology. In: *Malaria: Parasite biology, pathogenesis and protection*. Sherman, I. W., ASM Press, Washington, D. C., White, N. J.

Whittle, P. J. and Blundell, T. L. (1994) Protein structure-based drug design. *Annu. Rev. Biochem. Biomol. Struct.* **23**, 349-375

WHO (1996) Malaria. *WHO information fact sheet* N94,

Wickner, R. B., Taylor, K. L., Edskes, H. K. and Maddelein, M.-L. (2000) Prions: Portable prion domains. *Current Biol.* **10**, R335-337

Wilkinson, D. (2000) Limited proteolysis provides a wealth of protein structural information. *Scientist* **14**, 21-27

Wilson, K. and Walker, J. (2000) In: *Principles and Techniques of Practical Biochemistry*. Cambridge University Press, Cambridge

Wilson, R. J. (2002) Progress with parasite plastids. *J. Mol. Biol.* **319**, 257-274

Windholz, M. (1983) In: *The Merck Index.* Merck and Co, Rahway, USA

Winstanley, P. A. (2000) Chemotherapy for falciparum malaria: the armoury, the problems and the prospects. *Parasitol. Today* **16**, 146-153

Winstanley, P. A., Ward, S. A. and Snow, R. W. (2002) Clinical status and implications of antimalarial drug resistance. *Microbes Infect.* **4**, 157-164

Woodrow, C. J., Penny, J. I. and Krishna, S. (1999) Intraerythrocytic *Plasmodium falciparum* expresses a high-affinity facilitative hexose transporter. *J. Biol. Chem* **274**, 7272-7277

Wootton, J. C. and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* **266**, 554-571

Wrenger, C., Luersen, K., Krause, T., Müller, S. and Walter, R. D. (2001) The *Plasmodium falciparum* bifunctional ornithine decarboxylase, S- adenosylmethionine decarboxylase enables a well balanced polyamine synthesis without domain-domain interaction. *J. Biol. Chem.* **276**, 29651-29656

Wright, P. S., Byers, T. L., Cross-Doersen, D. E., McCann, P. P. and Bitoni, A. J. (1991) Irreversible inhibition of S-adenosylmethionine decarboxylase in *Plasmodium falciparum*-infected erythrocytes: Growth inhibition *in vitro. Biochem. Pharmacol.* **41**, 1713-1718

Xiong, H., Stanley, B. A., Tekwani, B. L. and Pegg, A. E. (1997) Processing of mammalian and plant S-adenosylmethionine decarboxylase proenzymes. *J. Biol. Chem.* **272**, 28342-28348

Yuthavong, Y. (2002) Basis for antifolate action and resistance in malaria. *Microbes Infect.* **4**, 175-182

Zubay, G. (1993) In: *Biochemistry.* Wm. C. Brown Communications, Dubuque

Zutshi, R., Brickner, M. and Chmielewski, J. (1998) Inhibiting the assembly of protein-protein interfaces. *Curr. Opin. Chem. Biol.* **2**, 62-66

Appendix I: Multiple alignment of the genomic DNA and cDNA sequences of *PfAdometdc/Odc.*

```
PfA/O cDNA : ---------------------AAAAAAAAAAAAAAAAAAATATAGATCCATATATCGAAATTATCCCTATATCTTAACATCTCTAATAATG :   69
cDNA       : -------------------------------------------------------------------------------------ATG :    3
PfA/O gDNA : CAATCTTTACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATAGATCCATATATCGAAATTATCCCTATATCTTAACATCTCTAATAATG : 2184
gDNA       : -------------------------------------------------------------------------------------ATG :    3

PfA/O cDNA : AACGGAATTTTTGAAGGAATTGAAAAAAGGGTTGTGATCAAATTAAAGGAGAGTTTTTTCAAAGGAAATAGAAATGTGAACTCCTTTTTAG :  160
cDNA       : AACGGAATTTTTGAAGGAATTGAAAAAAGGGTTGTGATCAAATTAAAGGAGAGTTTTTTCAAAGGAAATAGAAATGTGAACTCCTTTTTAG :   94
PfA/O gDNA : AACGGAATTTTTGAAGGAATTGAAAAAAGGGTTGTGATCAAATTAAAGGAGAGTTTTTTCAAAGGAAATAGAAATGTGAACTCCTTTTTAG : 2275
gDNA       : AACGGAATTTTTGAAGGAATTGAAAAAAGGGTTGTGATCAAATTAAAGGAGAGTTTTTTCAAAGGAAATAGAAATGTGAACTCCTTTTTAG :   94
                   Sampetf1

PfA/O cDNA : ATATACCTAAAGAATTATGGGAAGAAAAATTAAAATACATTGGTTGTAGTATTGTATCGGAAATAAGTGAGGACAAGAACGAGAGAAGAGG :  251
cDNA       : ATATACCTAAAGAATTATGGGAAGAAAAATTAAAATACATTGGTTGTAGTATTGTATCGGAAATAAGTGAGGACAAGAACGAGAGAAGAGG :  185
PfA/O gDNA : ATATACCTAAAGAATTATGGGAAGAAAAATTAAAATACATTGGTTGTAGTATTGTATCGGAAATAAGTGAGGACAAGAACGAGAGAAGAGG : 2366
gDNA       : ATATACCTAAAGAATTATGGGAAGAAAAATTAAAATACATTGGTTGTAGTATTGTATCGGAAATAAGTGAGGACAAGAACGAGAGAAGAGG :  185

PfA/O cDNA : TGAACGATGTCGTGTGTATTTATTGTCAGAGAGTTCTTTATACATTTTTGATGATTCTTTATTTATTAAGACATGTGGCAAAACAAGAGTT :  342
cDNA       : TGAACGATGTCGTGTGTATTTATTGTCAGAGAGTTCTTTATACATTTTTGATGATTCTTTATTTATTAAGACATGTGGCAAAACAAGAGTT :  276
PfA/O gDNA : TGAACGATGTCGTGTGTATTTATTGTCAGAGAGTTCTTTATACATTTTTGATGATTCTTTATTTATTAAGACATGTGGCAAAACAAGAGTT : 2457
gDNA       : TGAACGATGTCGTGTGTATTTATTGTCAGAGAGTTCTTTATACATTTTTGATGATTCTTTATTTATTAAGACATGTGGCAAAACAAGAGTT :  276

PfA/O cDNA : TTATTTTTCATACCGTTTGTGGTTGATTTATTAATATATCATATGGATAATGTAGGTATAATAGAAAAGAATTGTGTATATGATGAGACGT :  433
cDNA       : TTATTTTTCATACCGTTTGTGGTTGATTTATTAATATATCATATGGATAATGTAGGTATAATAGAAAAGAATTGTGTATATGATGAGACGT :  367
PfA/O gDNA : TTATTTTTCATACCGTTTGTGGTTGATTTATTAATATATCATATGGATAATGTAGGTATAATAGAAAAGAATTGTGTATATGATGAGACGT : 2548
gDNA       : TTATTTTTCATACCGTTTGTGGTTGATTTATTAATATATCATATGGATAATGTAGGTATAATAGAAAAGAATTGTGTATATGATGAGACGT :  367

PfA/O cDNA : TTATTGAAAACGAGAAATTCCATAATATAGCTGAATTCATAAAAGAACATTTCCTTTATTGTTTTTTTACACATATGAATTACCGAAATAA :  524
cDNA       : TTATTGAAAACGAGAAATTCCATAATATAGCTGAATTCATAAAAGAACATTTCCTTTATTGTTTTTTTACACATATGAATTACCGAAATAA :  458
PfA/O gDNA : TTATTGAAAACGAGAAATTCCATAATATAGCTGAATTCATAAAAGAACATTTCCTTTATTGTTTTTTTACACATATGAATTACCGAAATAA : 2639
gDNA       : TTATTGAAAACGAGAAATTCCATAATATAGCTGAATTCATAAAAGAACATTTCCTTTATTGTTTTTTTACACATATGAATTACCGAAATAA :  458

PfA/O cDNA : AACAAAGGATGGTTATTTTGAACAGGAATATCCACACAAATCTCTTGAAGATGAAAAGAAATTTTTTGAGTTTTTTTTTTAAGAACGTACAA :  615
cDNA       : AACAAAGGATGGTTATTTTGAACAGGAATATCCACACAAATCTCTTGAAGATGAAAAGAAATTTTTTGAGTTTTTTTTTTAAGAACGTACAA :  549
PfA/O gDNA : AACAAAGGATGGTTATTTTGAACAGGAATATCCACACAAATCTCTTGAAGATGAAAAGAAATTTTTTGAGTTTTTTTTTTAAGAACGTACAA : 2730
gDNA       : AACAAAGGATGGTTATTTTGAACAGGAATATCCACACAAATCTCTTGAAGATGAAAAGAAATTTTTTGAGTTTTTTTTTTAAGAACGTACAA :  549

PfA/O cDNA : ATGTATAATACACATTTACCTATGGAGAAAATGCATTATATATTCTTCTACTCTTCTGATGATGTACATATGACGGATATAGCTTCTACGT :  706
cDNA       : ATGTATAATACACATTTACCTATGGAGAAAATGCATTATATATTCTTCTACTCTTCTGATGATGTACATATGACGGATATAGCTTCTACGT :  640
PfA/O gDNA : ATGTATAATACACATTTACCTATGGAGAAAATGCATTATATATTCTTCTACTCTTCTGATGATGTACATATGACGGATATAGCTTCTACGT : 2821
gDNA       : ATGTATAATACACATTTACCTATGGAGAAAATGCATTATATATTCTTCTACTCTTCTGATGATGTACATATGACGGATATAGCTTCTACGT :  640

PfA/O cDNA : TTAAATTCTGTTCGGAAATACATTTGTTTGGAATTAACAAATATAATGAAAAGAATCAATTCCATGACGCTTATCTGAATAACAAGTCGTT :  797
cDNA       : TTAAATTCTGTTCGGAAATACATTTGTTTGGAATTAACAAATATAATGAAAAGAATCAATTCCATGACGCTTATCTGAATAACAAGTCGTT :  731
PfA/O gDNA : TTAAATTCTGTTCGGAAATACATTTGTTTGGAATTAACAAATATAATGAAAAGAATCAATTCCATGACGCTTATCTGAATAACAAGTCGTT : 2912
gDNA       : TTAAATTCTGTTCGGAAATACATTTGTTTGGAATTAACAAATATAATGAAAAGAATCAATTCCATGACGCTTATCTGAATAACAAGTCGTT :  731

PfA/O cDNA : GAATCTATTTACAAGAGTACATGAGGATAATTTAAAGCTTTATGATAGTAGTGATGCTGATAAGGAAGTAACCACCCACATCTATAGTACC :  888
cDNA       : GAATCTATTTACAAGAGTACATGAGGATAATTTAAAGCTTTATGATAGTAGTGATGCTGATAAGGAAGTAACCACCCACATCTATAGTACC :  822
PfA/O gDNA : GAATCTATTTACAAGAGTACATGAGGATAATTTAAAGCTTTATGATAGTAGTGATGCTGATAAGGAAGTAACCACCCACATCTATAGTACC : 3003
gDNA       : GAATCTATTTACAAGAGTACATGAGGATAATTTAAAGCTTTATGATAGTAGTGATGCTGATAAGGAAGTAACCACCCACATCTATAGTACC :  822

PfA/O cDNA : AGAGGGACATATGAAGATACAGGAATGGTGAATTGTGTGGATGTGATTTATAAGAATGAAAGTACATTGTTAAATAGGAATAATATAGAAA :  979
cDNA       : AGAGGGACATATGAAGATACAGGAATGGTGAATTGTGTGGATGTGATTTATAAGAATGAAAGTACATTGTTAAATAGGAATAATATAGAAA :  913
PfA/O gDNA : AGAGGGACATATGAAGATACAGGAATGGTGAATTGTGTGGATGTGATTTATAAGAATGAAAGTACATTGTTAAATAGGAATAATATAGAAA : 3094
gDNA       : AGAGGGACATATGAAGATACAGGAATGGTGAATTGTGTGGATGTGATTTATAAGAATGAAAGTACATTGTTAAATAGGAATAATATAGAAA :  913

PfA/O cDNA : ATATTCCATCTATTGAAAATAAAGAAAGTAATAATAATAGTAGATGTTGTCATAATAATAATTATAGTGGAAGTTGTCATAATATTGTGAG : 1070
cDNA       : ATATTCCATCTATTGAAAATAAAGAAAGTAATAATAATAGTAGATGTTGTCATAATAATAATTATAGTGGAAGTTGTCATAATATTGTGAG : 1004
PfA/O gDNA : ATATTCCATCTATTGAAAATAAAGAAAGTAATAATAATAGTAGATGTTGTCATAATAATAATTATAGTGGAAGTTGTCATAATATTGTGAG : 3185
gDNA       : ATATTCCATCTATTGAAAATAAAGAAAGTAATAATAATAGTAGATGTTGTCATAATAATAATTATAGTGGAAGTTGTCATAATATTGTGAG : 1004

PfA/O cDNA : TGTTGTTCCTTCCGAAAGAAATAATGATCATGTACATCACAGACATTATGAAGATACCTTAAATCGTTCTAATATTTCTGCTGAAGATAAC : 1161
cDNA       : TGTTGTTCCTTCCGAAAGAAATAATGATCATGTACATCACAGACATTATGAAGATACCTTAAATCGTTCTAATATTTCTGCTGAAGATAAC : 1095
PfA/O gDNA : TGTTGTTCCTTCCGAAAGAAATAATGATCATGTACATCACAGACATTATGAAGATACCTTAAATCGTTCTAATATTTCTGCTGAAGATAAC : 3276
gDNA       : TGTTGTTCCTTCCGAAAGAAATAATGATCATGTACATCACAGACATTATGAAGATACCTTAAATCGTTCTAATATTTCTGCTGAAGATAAC : 1095

PfA/O cDNA : AATAGAAATGCACAACCAGAAAAAGAAAAGGACGAAGATGTAAGAAGAGATGATGAAGAAAATAAAGTTCTAATAAAAATGATAGATACGA : 1252
cDNA       : AATAGAAATGCACAACCAGAAAAAGAAAAGGACGAAGATGTAAGAAGAGATGATGAAGAAAATAAAGTTCTAATAAAAATGATAGATACGA : 1186
PfA/O gDNA : AATAGAAATGCACAACCAGAAAAAGAAAAGGACGAAGATGTAAGAAGAGATGATGAAGAAAATAAAGTTCTAATAAAAATGATAGATACGA : 3367
gDNA       : AATAGAAATGCACAACCAGAAAAAGAAAAGGACGAAGATGTAAGAAGAGATGATGAAGAAAATAAAGTTCTAATAAAAATGATAGATACGA : 1186

PfA/O cDNA : ATTTATACGAATGTATAAATTATAATAAGGAAAGCTTTTTATATAATGAATTTATTTTACACCTTGTGGTTATTCTTGTAATGTTTCTGA : 1343
cDNA       : ATTTATACGAATGTATAAATTATAATAAGGAAAGCTTTTTATATAATGAATTTATTTTACACCTTGTGGTTATTCTTGTAATGTTTCTGA : 1277
PfA/O gDNA : ATTTATACGAATGTATAAATTATAATAAGGAAAGCTTTTTATATAATGAATTTATTTTACACCTTGTGGTTATTCTTGTAATGTTTCTGA : 3458
gDNA       : ATTTATACGAATGTATAAATTATAATAAGGAAAGCTTTTTATATAATGAATTTATTTTACACCTTGTGGTTATTCTTGTAATGTTTCTGA : 1277

PfA/O cDNA : AAAAAATAATTATTTTTGTGTACATTATTCACCAGAAGATTCTGTATCGTACGTTTCTGTTGAGGTATCTTCAAATTTGTCGTGTGATCGA : 1434
cDNA       : AAAAAATAATTATTTTTGTGTACATTATTCACCAGAAGATTCTGTATCGTACGTTTCTGTTGAGGTATCTTCAAATTTGTCGTGTGATCGA : 1368
PfA/O gDNA : AAAAAATAATTATTTTTGTGTACATTATTCACCAGAAGATTCTGTATCGTACGTTTCTGTTGAGGTATCTTCAAATTTGTCGTGTGATCGA : 3549
gDNA       : AAAAAATAATTATTTTTGTGTACATTATTCACCAGAAGATTCTGTATCGTACGTTTCTGTTGAGGTATCTTCAAATTTGTCGTGTGATCGA : 1368
                                    Samdcd1

PfA/O cDNA : TTTTTAGACTTTATTCACAAGCAGTTGAATTTTTATAATGGAAAGTATATGTTCATGATAAATTATGTATTCTGTGTGGAGGAGTAACAACA : 1525
cDNA       : TTTTTAGACTTTATTCACAAGCAGTTGAATTTTTATAATGGAAAGTATATGTTCATGATAAATTATGTATTCTGTGAGGAGAGTAACAACA : 1459
PfA/O gDNA : TTTTTAGACTTTATTCACAAGCAGTTGAATTTTTATAATGGAAAGTATATGTTCATGATAAATTATGTATTCTGTGAGGAGAGTAACAACA : 3640
gDNA       : TTTTTAGACTTTATTCACAAGCAGTTGAATTTTTATAATGGAAAGTATATGTTCATGATAAATTATGTATTCTGTGAGGAGAGTAACAACA : 1459
                                            ODCf1                      ODCexpf1

PfA/O cDNA : TGTCTAAAATGGTACCTGATGATGATAATAATAATTATAGTAGTGGTAAAAGTTGCGTTTATTATCAAGATTTAAATAAGAAAGAAAAAGA : 1616
cDNA       : TGTCTAAAATGGTACCTGATGATGATAATAATAATTATAGTAGTGGTAAAAGTTGCGTTTATTATCAAGATTTAAATAAGAAAGAAAAAGA : 1550
PfA/O gDNA : TGTCTAAAATGGTACCTGATGATGATAATAATAATTATAGTAGTGGTAAAAGTTGCGTTTATTATCAAGATTTAAATAAGAAAGAAAAAGA : 3731
gDNA       : TGTCTAAAATGGTACCTGATGATGATAATAATAATTATAGTAGTGGTAAAAGTTGCGTTTATTATCAAGATTTAAATAAGAAAGAAAAAGA : 1550
```

Adometdc ←

```
PfA/O cDNA : AGAATATTATCGCTTGAACAAAAAATTAAGAAACGAGTTATTTATTAATTCGAAACAATTTTATGAATTACATACATTTACCGAACGAACG : 1707
cDNA       : AGAATATTATCGCTTGAACAAAAAATTAAGAAACGAGTTATTTATTAATTCGAAACAATTTTATGAATTACATACATTTACCGAACGAACG : 1641
PfA/O gDNA : AGAATATTATCGCTTGAACAAAAAATTAAGAAACGAGTTATTTATTAATTCGAAACAATTTTATGAATTACATACATTTACCGAACGAACG : 3822
gDNA       : AGAATATTATCGCTTGAACAAAAAATTAAGAAACGAGTTATTTATTAATTCGAAACAATTTTATGAATTACATACATTTACCGAACGAACG : 1641

PfA/O cDNA : GTTGGATTTATGAGAGTGCAATATTTTGTTTATAAATTAAGAGATGTTGTTAAATGTGTAGAAAAAGAAACTTTGCTAGCTAGGAGTTCGT : 1798
cDNA       : GTTGGATTTATGAGAGTGCAATATTTTGTTTATAAATTAAGAGATGTTGTTAAATGTGTAGAAAAAGAAACTTTGCTAGCTAGGAGTTCGT : 1732
PfA/O gDNA : GTTGGATTTATGAGAGTGCAATATTTTGTTTATAAATTAAGAGATGTTGTTAAATGTGTAGAAAAAGAAACTTTGCTAGCTAGGAGTTCGT : 3913
gDNA       : GTTGGATTTATGAGAGTGCAATATTTTGTTTATAAATTAAGAGATGTTGTTAAATGTGTAGAAAAAGAAACTTTGCTAGCTAGGAGTTCGT : 1732

PfA/O cDNA : CTTGTTTATTTATGTTTAATAATATCAAACGAAATGACGTACATGATGATTATGTAACTAAGTCGTCAAATGGTGGTGTAATAAAACAATT : 1889
cDNA       : CTTGTTTATTTATGTTTAATAATATCAAACGAAATGACGTACATGATGATTATGTAACTAAGTCGTCAAATGGTGGTGTAATAAAACAATT : 1823
PfA/O gDNA : CTTGTTTATTTATGTTTAATAATATCAAACGAAATGACGTACATGATGATTATGTAACTAAGTCGTCAAATGGTGGTGTAATAAAACAATT : 4004
gDNA       : CTTGTTTATTTATGTTTAATAATATCAAACGAAATGACGTACATGATGATTATGTAACTAAGTCGTCAAATGGTGGTGTAATAAAACAATT : 1823
                                  Sampetr1

PfA/O cDNA : AACGGAAAGAGATGTTGATGATATGTATGAGTATGCTTTAAATTTTTGTAAACAAAATAAAATAGTTGTTGTAGATACTAATACCTTTTTT : 1980
cDNA       : AACGGAAAGAGATGTTGATGATATGTATGAGTATGCTTTAAATTTTTGTAAACAAAATAAAATAGTTGTTGTAGATACTAATACCTTTTTT : 1914
PfA/O gDNA : AACGGAAAGAGATGTTGATGATATGTATGAGTATGCTTTAAATTTTTGTAAACAAAATAAAATAGTTGTTGTAGATACTAATACCTTTTTT : 4095
gDNA       : AACGGAAAGAGATGTTGATGATATGTATGAGTATGCTTTAAATTTTTGTAAACAAAATAAAATAGTTGTTGTAGATACTAATACCTTTTTT : 1914
                         ODCf2

PfA/O cDNA : TTTGATGCATCTAAAAGAAAGGAGAACTTAATAAAACTTGAAAAGGTACAAACAAATGAGAAAGATGAATATGAAGAAAAAGATGAAGTGT : 2071
cDNA       : TTTGATGCATCTAAAAGAAAGGAGAACTTAATAAAACTTGAAAAGGTACAAACAAATGAGAAAGATGAATATGAAGAAAAAGATGAAGTGT : 2005
PfA/O gDNA : TTTGATGCATCTAAAAGAAAGGAGAACTTAATAAAACTTGAAAAGGTACAAACAAATGAGAAAGATGAATATGAAGAAAAAGATGAAGTGT : 4186
gDNA       : TTTGATGCATCTAAAAGAAAGGAGAACTTAATAAAACTTGAAAAGGTACAAACAAATGAGAAAGATGAATATGAAGAAAAAGATGAAGTGT : 2005

PfA/O cDNA : ATCGAAGGGGTAATAATGAATTGAGTTCGTTGGATCATTTAGATAGTAAGAATAATTTGATTCATATGTATTATGAAAAGAACAAATGTGA : 2162
cDNA       : ATCGAAGGGGTAATAATGAATTGAGTTCGTTGGATCATTTAGATAGTAAGAATAATTTGATTCATATGTATTATGAAAAGAACAAATGTGA : 2096
PfA/O gDNA : ATCGAAGGGGTAATAATGAATTGAGTTCGTTGGATCATTTAGATAGTAAGAATAATTTGATTCATATGTATTATGAAAAGAACAAATGTGA : 4277
gDNA       : ATCGAAGGGGTAATAATGAATTGAGTTCGTTGGATCATTTAGATAGTAAGAATAATTTGATTCATATGTATTATGAAAAGAACAAATGTGA : 2096

PfA/O cDNA : TATCATAAATAAGGATGATGAGAATTCAACGATAGCGACGAATAATAATGATAATAATAATGATAGTAGTTCTTATGACAAAAGTATAACG : 2253
cDNA       : TATCATAAATAAGGATGATGAGAATTCAACGATAGCGACGAATAATAATGATAATAATAATGATAGTAGTTCTTATGACAAAAGTATAACG : 2187
PfA/O gDNA : TATCATAAATAAGGATGATGAGAATTCAACGATAGCGACGAATAATAATGATAATAATAATGATAGTAGTTCTTATGACAAAAGTATAACG : 4368
gDNA       : TATCATAAATAAGGATGATGAGAATTCAACGATAGCGACGAATAATAATGATAATAATAATGATAGTAGTTCTTATGACAAAAGTATAACG : 2187

PfA/O cDNA : ATCAGCAGAAGCAGTAGCTGTAATAATAGCCATTTGAGTTATAGTAGTTTTGATAATAATCATGGAAATGAAAAAATGAAAGATTATATAA : 2344
cDNA       : ATCAGCAGAAGCAGTAGCTGTAATAATAGCCATTTGAGTTATAGTAGTTTTGATAATAATCATGGAAATGAAAAAATGAAAGATTATATAA : 2278
PfA/O gDNA : ATCAGCAGAAGCAGTAGCTGTAATAATAGCCATTTGAGTTATAGTAGTTTTGATAATAATCATGGAAATGAAAAAATGAAAGATTATATAA : 4459
gDNA       : ATCAGCAGAAGCAGTAGCTGTAATAATAGCCATTTGAGTTATAGTAGTTTTGATAATAATCATGGAAATGAAAAAATGAAAGATTATATAA : 2278

PfA/O cDNA : GTGTTGATGAAAATAATAATAATAATAATAATAATAAAAATAATAATGTATTGTTAACTTTACAAAGGAACAGTGATGATGAAAATGGTAA : 2435
cDNA       : GTGTTGATGAAAATAATAATAATAATAATAATAATAAAAATAATAATGTATTGTTAACTTTACAAAGGAACAGTGATGATGAAAATGGTAA : 2369
PfA/O gDNA : GTGTTGATGAAAATAATAATAATAATAATAATAATAAAAATAATAATGTATTGTTAACTTTACAAAGGAACAGTGATGATGAAAATGGTAA : 4550
gDNA       : GTGTTGATGAAAATAATAATAATAATAATAATAATAAAAATAATAATGTATTGTTAACTTTACAAAGGAACAGTGATGATGAAAATGGTAA : 2369
```

Odc →

```
PfA/O cDNA : AGATAAAGATAATGAAAAAAATGACGTAAGTTTAGAAAACAATATGGAAAAGAATTATAAAGAAGAAATATGGAATTATTATACAAAAAAT : 2526
cDNA       : AGATAAAGATAATGAAAAAAATGACGTAAGTTTAGAAAACAATATGGAAAAGAATTATAAAGAAGAAATATGGAATTATTATACAAAAAAT : 2460
PfA/O gDNA : AGATAAAGATAATGAAAAAAATGACGTAAGTTTAGAAAACAATATGGAAAAGAATTATAAAGAAGAAATATGGAATTATTATACAAAAAAT : 4641
gDNA       : AGATAAAGATAATGAAAAAAATGACGTAAGTTTAGAAAACAATATGGAAAAGAATTATAAAGAAGAAATATGGAATTATTATACAAAAAAT : 2460
                         ODCexpf2

PfA/O cDNA : AAAGTGGAAGTAAAAACATTAGAAAAAGTATTAAATGAAAATATAGATACATCAGTAGTTTGTATAAATTTACAGAAAATATTAGCTCAGT : 2617
cDNA       : AAAGTGGAAGTAAAAACATTAGAAAAAGTATTAAATGAAAATATAGATACATCAGTAGTTTGTATAAATTTACAGAAAATATTAGCTCAGT : 2551
PfA/O gDNA : AAAGTGGAAGTAAAAACATTAGAAAAAGTATTAAATGAAAATATAGATACATCAGTAGTTTGTATAAATTTACAGAAAATATTAGCTCAGT : 4732
gDNA       : AAAGTGGAAGTAAAAACATTAGAAAAAGTATTAAATGAAAATATAGATACATCAGTAGTTTGTATAAATTTACAGAAAATATTAGCTCAGT : 2551
                                        ODCR3

PfA/O cDNA : ATGTTAGATTTAAAAAGAATCTTCCACATGTTACTCCATTCTATTCTGTAAAAAGTAATAATGATGAAGTTGTAATCAAATTTTTTATATGG : 2708
cDNA       : ATGTTAGATTTAAAAAGAATCTTCCACATGTTACTCCATTCTATTCTGTAAAAAGTAATAATGATGAAGTTGTAATCAAATTTTTTATATGG : 2642
PfA/O gDNA : ATGTTAGATTTAAAAAGAATCTTCCACATGTTACTCCATTCTATTCTGTAAAAAGTAATAATGATGAAGTTGTAATCAAATTTTTTATATGG : 4823
gDNA       : ATGTTAGATTTAAAAAGAATCTTCCACATGTTACTCCATTCTATTCTGTAAAAAGTAATAATGATGAAGTTGTAATCAAATTTTTTATATGG : 2642

PfA/O cDNA : ATTGAATTGTAATTTTGATTGCGCTTCGATAGGTGAAATAAGTAAAGTAATAAAATTATTACCCAAATTTATCAAGAGATAGAATAATTTTT : 2799
cDNA       : ATTGAATTGTAATTTTGATTGCGCTTCGATAGGTGAAATAAGTAAAGTAATAAAATTATTACCCAAATTTATCAAGAGATAGAATAATTTTT : 2733
PfA/O gDNA : ATTGAATTGTAATTTTGATTGCGCTTCGATAGGTGAAATAAGTAAAGTAATAAAATTATTACCCAAATTTATCAAGAGATAGAATAATTTTT : 4914
gDNA       : ATTGAATTGTAATTTTGATTGCGCTTCGATAGGTGAAATAAGTAAAGTAATAAAATTATTACCCAAATTTATCAAGAGATAGAATAATTTTT : 2733

PfA/O cDNA : GCGAATACAATTAAAAGTATTAATTCTTTAATATATGCAAGAAAGGAAAATATTAATTTATGTACTTTTGATAATTTAGATGAATTAAAAAA : 2890
cDNA       : GCGAATACAATTAAAAGTATTAATTCTTTAATATATGCAAGAAAGGAAAATATTAATTTATGTACTTTTGATAATTTAGATGAATTAAAAAA : 2824
PfA/O gDNA : GCGAATACAATTAAAAGTATTAATTCTTTAATATATGCAAGAAAGGAAAATATTAATTTATGTACTTTTGATAATTTAGATGAATTAAAAAA : 5005
gDNA       : GCGAATACAATTAAAAGTATTAATTCTTTAATATATGCAAGAAAGGAAAATATTAATTTATGTACTTTTGATAATTTAGATGAATTAAAAAA : 2824

PfA/O cDNA : AAAATATATAAATATCATCCGAAATGTTCTTTAATATTACGTATTAATGTAGATTTTAAAAATTACAAATCTTATATGTCTTCAAAATATGG : 2981
cDNA       : AAAATATATAAATATCATCCGAAATGTTCTTTAATATTACGTATTAATGTAGATTTTAAAAATTACAAATCTTATATGTCTTCAAAATATGG : 2915
PfA/O gDNA : AAAATATATAAATATCATCCGAAATGTTCTTTAATATTACGTATTAATGTAGATTTTAAAAATTACAAATCTTATATGTCTTCAAAATATGG : 5096
gDNA       : AAAATATATAAATATCATCCGAAATGTTCTTTAATATTACGTATTAATGTAGATTTTAAAAATTACAAATCTTATATGTCTTCAAAATATGG : 2915

PfA/O cDNA : AGCTAATGAATATGAATGGGAAGAAATGTTATTGTATGCAAAAAAACATAATCTAAATATTGTAGGTGTATCATTTCATGTTGGTAGTAAT : 3072
cDNA       : AGCTAATGAATATGAATGGGAAGAAATGTTATTGTATGCAAAAAAACATAATCTAAATATTGTAGGTGTATCATTTCATGTTGGTAGTAAT : 3006
PfA/O gDNA : AGCTAATGAATATGAATGGGAAGAAATGTTATTGTATGCAAAAAAACATAATCTAAATATTGTAGGTGTATCATTTCATGTTGGTAGTAAT : 5187
gDNA       : AGCTAATGAATATGAATGGGAAGAAATGTTATTGTATGCAAAAAAACATAATCTAAATATTGTAGGTGTATCATTTCATGTTGGTAGTAAT : 3006
                                                           GSP1

PfA/O cDNA : ACAAAGAATTTATTTGATTTCTGTCTAGCCATTAAATTATGTGTAGAGATGTATTCGATATGAGTAGTAATATGGGATTTAATTTTTTATATAA : 3163
cDNA       : ACAAAGAATTTATTTGATTTCTGTCTAGCCATTAAATTATGTGTAGAGATGTATTCGATATGAGTAGTAATATGGGATTTAATTTTTTATATAA : 3097
PfA/O gDNA : ACAAAGAATTTATTTGATTTCTGTCTAGCCATTAAATTATGTGTAGAGATGTATTCGATATGAGTAGTAATATGGGATTTAATTTTTTATATAA : 5278
gDNA       : ACAAAGAATTTATTTGATTTCTGTCTAGCCATTAAATTATGTGTAGAGATGTATTCGATATGAGTAGTAATATGGGATTTAATTTTTTATATAA : 3097
                                        ODCR1

PfA/O cDNA : TAAATTTAGGAGGGGGGATATCCAGAAGAATTAGAATATGATAATGCAAAGAAACATGATAAAATTCATTATTGTACTTTAAGTCTTCAAGA : 3254
cDNA       : TAAATTTAGGAGGGGGGATATCCAGAAGAATTAGAATATGATAATGCAAAGAAACATGATAAAATTCATTATTGTACTTTAAGTCTTCAAGA : 3188
PfA/O gDNA : TAAATTTAGGAGGGGGGATATCCAGAAGAATTAGAATATGATAATGCAAAGAAACATGATAAAATTCATTATTGTACTTTAAGTCTTCAAGA : 5369
gDNA       : TAAATTTAGGAGGGGGGATATCCAGAAGAATTAGAATATGATAATGCAAAGAAACATGATAAAATTCATTATTGTACTTTAAGTCTTCAAGA : 3188

PfA/O cDNA : AATTAAAAAAGATATACAAAAATTTCTTAATGAAGAAACATTTCTCAAGACGAAATATGGATACTATAGTTTTGAAAAAATATCATTGGCT : 3345
cDNA       : AATTAAAAAAGATATACAAAAATTTCTTAATGAAGAAACATTTCTCAAGACGAAATATGGATACTATAGTTTTGAAAAAATATCATTGGCT : 3279
PfA/O gDNA : AATTAAAAAAGATATACAAAAATTTCTTAATGAAGAAACATTTCTCAAGACGAAATATGGATACTATAGTTTTGAAAAAATATCATTGGCT : 5460
gDNA       : AATTAAAAAAGATATACAAAAATTTCTTAATGAAGAAACATTTCTCAAGACGAAATATGGATACTATAGTTTTGAAAAAATATCATTGGCT : 3279
```

```
PfA/O cDNA : ATTAATATGTCAATCGATCATTATTTTAGTCATATGAAAGATAATCTAAGAGTTATTTGTGAACCTGGTAGATATATGGTCGCTGCTTCGT : 3436
cDNA       : ATTAATATGTCAATCGATCATTATTTTAGTCATATGAAAGATAATCTAAGAGTTATTTGTGAACCTGGTAGATATATGGTCGCTGCTTCGT : 3370
PfA/O gDNA : ATTAATATGTCAATCGATCATTATTTTAGTCATATGAAAGATAATCTAAGAGTTATTTGTGAACCTGGTAGTAATATGGTCGCTGCTTCGT : 5551
gDNA       : ATTAATATGTCAATCGATCATTATTTTAGTCATATGAAAGATAATCTAAGAGTTATTTGTGAACCTGGTAGTAATATGGTCGCTGCTTCGT : 3370

PfA/O cDNA : CAACATTAGCTGTTAAAATTATAGGAAAGAGACGTCCAACTTTTCAGGGCATTATGTTAAAAGAATTAAAAGACCATTACGATCCTTTAAA : 3527
cDNA       : CAACATTAGCTGTTAAAATTATAGGAAAGAGACGTCCAACTTTTCAGGGCATTATGTTAAAAGAATTAAAAGACCATTACGATCCTTTAAA : 3461
PfA/O gDNA : CAACATTAGCTGTTAAAATTATAGGAAAGAGACGTCCAACTTTTCAGGGCATTATGTTAAAAGATTTAAAAGACCATTACGATCCTTTAAA : 5642
gDNA       : CAACATTAGCTGTTAAAATTATAGGAAAGAGACGTCCAACTTTTCAGGGCATTATGTTAAAAGATTTAAAAGACCATTACGATCCTTTAAA : 3461

PfA/O cDNA : TTTTGCTCAACAAGAAAATAAGAAACAAGACGAAACAAAAATAAACCACAATAATGATAATAATGATAATAATGATAATAATGATAATAAT : 3618
cDNA       : TTTTGCTCAACAAGAAAATAAGAAACAAGACGAAACAAAAATAAACCACAATAATGATAATAAagtaatgaaTaATAATAATGATAATAAT : 3552
PfA/O gDNA : TTTTGCTCCACAAGAAAATAAGAAACAAGACGAAACAAAAATAAACCACAATAATGATAATAATGATAATAATGATAATAATGATAATAAT : 5733
gDNA       : TTTTGCTCCACAAGAAAATAAGAAACAAGACGAAACAAAAATAAACCACAATAATGATAATAATGATAATAATGATAATAATGATAATAAT : 3552

PfA/O cDNA : ATTAATAATAATAATAATAATCAAAAAGGGGGCCAAGGAAATATTATGAATGATCTAATAATAACTAGCACAAATGATTCTACTAGTAAAA : 3709
cDNA       : ATTAATAATAATAATAATAATCAAAAAGGGGGCCAAGGAAATATTATGAATGATCTAATAATAACTAGCACAAATGATTCTACTAGTAAAA : 3643
PfA/O gDNA : ATTAATAATAATAATAATAATCAAAAAGGGGGCCAAGGAAATATTATGAATGATCTAATAATAACTAGCACAAATGATTCTACTAGTAAAA : 5824
gDNA       : ATTAATAATAATAATAATAATCAAAAAGGGGGCCAAGGAAATATTATGAATGATCTAATAATAACTAGCACAAATGATTCTACTAGTAAAA : 3643

PfA/O cDNA : AGAATGATCATTCTTCTAGTCAAGTTATTCAAAATGTATCGTGCACAATACGTGATAAAGAAGGAGATAATATTAAAATAAATACACATAC : 3800
cDNA       : AGAATGATCATTCTTCTAGTCAAGTTATTCAAAATGTATCGTGCACAATACGTGATAAAGAAGGAGATAATATTAAAATAAATACACATAC : 3734
PfA/O gDNA : AGAATGATCATTCTTCTAGTCAAGTTATTCAAAATGTATCGTGCACAATACGTGATAAAGAAGGAGATAATATTAAAATAAATACACATAC : 5915
gDNA       : AGAATGATCATTCTTCTAGTCAAGTTATTCAAAATGTATCGTGCACAATACGTGATAAAGAAGGAGATAATATTAAAATAAATACACATAC : 3734

PfA/O cDNA : CATAAATAATCCTAATATAAATGGAAAAGAAAATACCGTGGATGGTGATAATATTAATATTGCTCATAAAAATATTGGTAATAACTTTAGT : 3891
cDNA       : CATAAATAATCCTAATATAAATGGAAAAGAAAATACCGTGGATGGTGATAATATTAATATTGCTCATAAAAATATTGGTAATAACTTTAGT : 3825
PfA/O gDNA : CATAAATAATCCTAATATAAATGGAAAAGAAAATACCGTGGATGGTGATAATATTAATATTGCTCATAAAAATATTGGTAATAACTTTAGT : 6006
gDNA       : CATAAATAATCCTAATATAAATGGAAAAGAAAATACCGTGGATGGTGATAATATTAATATTGCTCATAAAAATATTGGTAATAACTTTAGT : 3825

PfA/O cDNA : AGTAGTAATTCAAAATTAGGCAACATAACAAATATTAAGAAAAAAGTTGTTAATATTAATGACAATAGATATAATTATTTCTCATATTATG : 3982
cDNA       : AGTAGTAATTCAAAATTAGGCAACATAACAAATATTAAGAAAAAAGTTGTTAATATTAATGACAATAGATATAATTATTTCTCATATTATG : 3916
PfA/O gDNA : AGTAGTAATTCAAAATTAGGCAACATAACAAATATTAAGAAAAAAGTTGTTAATATTAATGACAATAGATATAATTATTTCTCATATTATG : 6097
gDNA       : AGTAGTAATTCAAAATTAGGCAACATAACAAATATTAAGAAAAAAGTTGTTAATATTAATGACAATAGATATAATTATTTCTCATATTATG : 3916

PfA/O cDNA : TAAGCGATAGTATATATGGTTGTTTTAGTGGTATAATTTTTGATGAATACAATAGATGTCCTATTTATGTTATTAAAAACAAAAATAACCC : 4073
cDNA       : TAAGCGATAGTATATATGGTTGTTTTAGTGGTATAATTTTTGATGAATACAATAGATGTCCTATTTATGTTATTAAAAACAAAAATAACCC : 4007
PfA/O gDNA : TAAGCGATAGTATATATGGTTGTTTTAGTGGTATAATTTTTGATGAATACAATAGATGTCCTATTTATGTTATTAAAAACAAAAATAACCC : 6188
gDNA       : TAAGCGATAGTATATATGGTTGTTTTAGTGGTATAATTTTTGATGAATACAATAGATGTCCTATTTATGTTATTAAAAACAAAAATAACCC : 4007
                                                                                         ODCR2
PfA/O cDNA : TAATCAAAATTTTATGAATTTTAATTTGTATTTAGCTAATGTATTTGGACAATCATGTGATGGCTTGGATATGATCAATTCTATTACGTAC : 4164
cDNA       : TAATCAAAATTTTATGAATTTTAATTTGTATTTAGCTAATGTATTTGGACAATCATGTGATGGCTTGGATATGATCAATTCTATTACGTAC : 4098
PfA/O gDNA : TAATCAAAATTTTATGAATTTTAATTTGTATTTAGCTAATGTATTTGGACAATCATGTGATGGCTTGGATATGATCAATTCTATTACGTAC : 6279
gDNA       : TAATCAAAATTTTATGAATTTTAATTTGTATTTAGCTAATGTATTTGGACAATCATGTGATGGCTTGGATATGATCAATTCTATTACGTAC : 4098

PfA/O cDNA : TTACCTGAGTGTTATATTAATGATTGGCTTCTCTATGAATATGCTGGGGCATACACTTTTGTCAGCTCATCAAACTTTAATGGATTTAAGA : 4255
cDNA       : TTACCTGAGTGTTATATTAATGATTGGCTTCTCTATGAATATGCTGGGGCATACACTTTTGTCAGCTCATCAAACTTTAATGGATTTAAGA : 4189
PfA/O gDNA : TTACCTGAGTGTTATATTAATGATTGGCTTCTCTATGAATATGCTGGGGCATACACTTTTGTCAGCTCATCAAACTTTAATGGATTTAAGA : 6370
gDNA       : TTACCTGAGTGTTATATTAATGATTGGCTTCTCTATGAATATGCTGGGGCATACACTTTTGTCAGCTCATCAAACTTTAATGGATTTAAGA : 4189

PfA/O cDNA : AATGCAAGAAGGTGTATATATTCCCTGAATCGAAACCTTCCCTTAAGGGGCAACCAAACAAACATTGG[TAA]ATAACAAAATCGAAGAAAAA : 4346
cDNA       : AATGCAAGAAGGTGTATATATTCCCTGAATCGAAACCTTCCCTTAAGGGGCAACCAAACAAACATTGGTAAATAACAAAATCGAAGAAAAA : 4280
PfA/O gDNA : AATGCAAGAAGGTGTATATATTCCCTGAATCGAAACCTTCCCTTAAGGGGCAACCAAACAAACATTGG[TAA]TAACAAAATCGAAGAAAAA : 6461
gDNA       : AATGCAAGAAGGTGTATATATTCCCTGAATCGAAACCTTCCCTTAAGGGGCAACCAAACAAACATTGG--------------------- : 4257
                                                                              ODCexpr
PfA/O cDNA : GGAAATAAATAGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA--------------------------------------------------- : 4383
cDNA       : GGA[AATAA]ATAGGGAAAAAAAAAAAAAAAAAAAAAA----------------------------------------------------- : 4312
PfA/O gDNA : GGAAATAAATAGGGAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAGAAAAGAAAAAAAAAAAAAGAAAAAGAAAA : 6552
gDNA       : -------------------------------------------------------------------------------------- : -
```

Figure A.1: Multiple-alignment of the genomic (gDNA) and cDNA sequences of *PfAdometdc/Odc* **ORF.** PfA/O indicates sequences deposited in Genbank: cDNA sequence accession number AF094833, genomic DNA sequence accession number AF112367. Sequences from this study are indicated by cDNA and gDNA respectively. The start (ATG) and stop codons (TAA) are in blue boxes. The poly-adenylation signal is boxed in green (AATAA). Differences in the cDNA sequences are indicated in yellow. Primer sites used in this study are indicated with their orientation. Domain definitions are indicated in red according to (Müller, *et al.*, 2000).