MOLECULAR CHARACTERISATION OF THE ORNITHINE DECARBOXYLASE GENE OF THE HUMAN MALARIA PARASITE, *PLASMODIUM FALCIPARUM*

by

Lyn-Marié Birkholtz

Submitted in partial fulfilment of the requirements for the degree of Master of Science (Biochemistry)

in the
Faculty of Biological and Agricultural Sciences

University of Pretoria
Pretoria
December, 1998

ACKNOWLEDGEMENTS

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

- Prof. A. I. Louw, Department of Biochemistry, University of Pretoria, as supervisor of this project, for his guidance and encouragement, helpful suggestions and insightful ideas, moral support and willingness to allow discussions on many a subject;

- Prof. L. Visser, Department of Biochemistry, University of Pretoria, Co-supervisor, for his valuable criticism of experimental results and emphasis on correct scientific procedure;

- Karen Nel, Fourie Joubert, Katherine Clark and the other members of the malaria research team for their guidance, friendship and advice about laboratory techniques;

- Fourie Joubert and Jaco de Ridder, for their assistance with computers;

- My fellow students and friends, for helping me retain a balanced outlook on life during the course of the degree;

- My parents and family, for always being interested in my studies, and for their continual love and support;

- My husband, Franz Birkholtz, for his endless love, patience and understanding, encouragement and unfailing support and belief in me;

- The Foundation for Research Development, Medical Research Council and the University of Pretoria for financial assistance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	Adenosine
AMA	Apical membrane antigen
AMP	Adenosine monophosphate
AMV	Avian myeloblastosis virus
ASS	Adenylosuccinate synthase
ATP	Adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BCBD	N^1,N^4-bis(7-chloroquinoline-4-yl)butane-1,4-diamine
BLAST	Basic local alignment search tool
bp	Base pair
C	Cytosine
°C	Degrees Celsius
CARP	Clustered asparigine rich protein
CCD	Charge coupled device
cDNA	Complementary DNA
CD	Cluster of differentiation
C-terminal	Carboxy terminal
CS	Circumsporozoite
CTP	Phosphatidylcholine cytidyl transferase
DAP	Diaminopimelic acid
DDT	Dichlorodiphenyltrichloroethane
ddNTP	Dideoxyribonucleotide triphosphate
ddUTP	Dideoxyuridine triphosphate
DD-Poly-T	Differential display poly-T primer
DEPC	Diethyl pyrocarbonate
DFMO	Difluoromethyl ornithine
DHFR	Dihydrofolate reductase
DHPS	Dihydropteroate synthetase
DHOH	Dihydroorotase dehydrogenase
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
DMF	Dimethylformamide
DMSO	Dimethyl sulphoxide
ds	Double-stranded
DTT	Dithiotreitol
dUTP	Deoxyuridine triphosphate
EBA	Erythrocyte binding antigen
EDTA	Ethylenediaminetetraacetic acid
EM	Erythrocyte membrane
EtBr	Ethidium Bromide
FP	Ferriprotoporphyrin IX

G	Guanidine
GM-CSF	Granulocyte macrophage colony stimulating factor
G6-PD	Glucose-6-phosphate dehydrogenase
GSP	Gene-specific primer
GPI	Glycophosphatidyl inositol
HGPRT	Hypoxanthine-guanosine phosphoribosyltransferase
HK	Hexokinase
HRP	Histidine rich protein
I	Inosine
ICAM	Intracellular adhesion molecule
IL	Interleukin
IMPDH	Inosine monophosphate dehydrogenase
IPTG	Isopropyl-D-galactoside
IUBMB	International Union of Biochemistry and Molecular Biology
IFN	Interferon
IMP	Inosine monophosphate
kDa	kiloDalton
LA-PCR	Ligation-anchored PCR
LAMA	Local alignments of multiple alignments
LB	Luria Bertani
LDH	Lactate dehydrogenase
LSA-1	Liver stage antigen
M	Molar
mM	Millimolar
µM	Micromolar
MM	Molecular mass
mRNA	Messenger RNA
MDR	Multi-drug resistance
MHC	Major histocompatibility complex
MI	Match index
MMLV	Murine moloney leukaemia virus
MOPS	Morpholinopropanesulphonic acid
MSA	Merozoite surface antigen
MSP	Merozoite surface protein
ng	Nanogram
nt	Nucleotide
NBT	Nitroblue tetrazolium chloride
NCBI	National Center for Biotechnology Information
NTP	Nucleotide triphosphate
N-terminal	Amino terminal
NO	Nitric oxide
OD	Optical density
ODC	Ornithine decarboxylase
ORF	Open reading frame

PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Poly-ethylene glycol
PEST	Pro, Glu, Ser, Thr rich region
PfEMP	*P. falciparum*-infected erythrocyte membrane protein
PfHRP	*P. falciparum* histidine rich protein
Pfu	*Pyrococcus furiosis*
Pgh	P-glycoprotein homologue
PLP	Pyridoxal-5'-phosphate
PMV	Parasite vacuole membrane
PPM	Parasite plasma membrane
PPMP	*dl*-threo-1-phenyl-2-palmitoylamino-3-morpho-1-propanol
PPP	Pentose phosphate pathway
PPPK	Pyrophosphokinase
PSSM	Position specific sorting matrix
R	A or G nucleotides
RACE	Rapid amplification of cDNA ends
RAP-1	Rhoptry associated protein
RESA	Ring infected surface antigen
RLM-RACE	RNA ligase-mediated RACE
RL-PCR	Reverse ligation-mediated PCR
RNAse	Ribonuclease
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT	Reverse transcription
RT-PCR	Reverse transcription PCR
SERA	Serine repeat antigen
SERP	Serine repeat protein
SDS	Sodium Dodecyl Sulphate
ss	Single stranded
SSP 2	Sporozoite surface antigen
STARP	Sporozoite surface threonine and aspartic acid rich protein
T	Tymidine
T_m	Melting temperature
$t_{1/2}$	Half-life
Taq	*Thermus aquaticus*
TBS	Tris buffered saline
TdT	Terminal deoxynucleotidyltransferase
TE	Tris-EDTA buffer
TIM	Triosephosphate isomerase
TMAC	Tetramethylammonium chloride
TNF α	Tumor necrosis factor α
TRAP	Thrombospondin related anonymous protein
TRIS	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
TS	Thymidyate synthetase
TVM	Tubovesicular membrane network
U	Units

UTR	Untranslated region
UV	Ultraviolet
V	Volts
VCAM	Vascular cell adhesion molecule
W	A or T nucleotides
X-gal	5-bromo-4-chloro-indolyl-β-D-galactoside
X-phosphate	5-bromo-4-chloro-3-indolyl-phosphate
Y	C or T nucleotides

CHAPTER 1

Literature review

Since the beginning of the written word (6000-5500 BC), deadly fevers have been recorded and were superstitiously believed to be due to bad air ('miasma' or 'mal-aria') emanating from swamps (Desowitz, 1991). It was only in 1880 that A. Laveran detected organisms in the blood of patients with acute fever and thereby identified the malaria parasite. In 1897, Ronald Ross established the mode of transmission of the infection and made the connection between parasite and mosquito, and wrote: (King and King, 1992; Hagan and Chauhan, 1997)

"I find thy cunning seeds
O, million murdering death."

Despite significant successes in the prevention of the disease, malaria still presents an immense problem in tropical and subtropical areas of the globe. It is a public health problem in more than 92 countries, placing a predicted 2400 million people at risk. An estimated 300-500 million people are infected annually and the mortality is estimated to be in the range of 1.5 to 2.7 million deaths each year. This makes malaria the most important insect-transmitted, infectious tropical disease. Africa in particular is severely affected by malaria: 90% of the world's cases occur in sub-Saharan Africa. Unfortunately, this is also one of the poorest continents which makes the containment of the malaria epidemic virtually impossible (Collins and Paskewitz, 1995; WHO, 1996). Factors contributing to the increase in notifications include favourable climatic conditions, migration of infected people to parasite-free regions, poor public health control programmes and resistance to existing insecticides and drugs.

This literature review will first introduce the causative agents of the disease and their biochemical properties. The pathological manifestation of the disease and the efficacy of different control strategies are then examined. In order to understand the biochemistry of the malaria parasite, some metabolic processes are considered which could be interfered with in order to contain the disease. One such target, polyamine biosynthesis and its rate controlling enzyme ornithine decarboxylase, is then examined in greater detail.

1.1 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 as an apicomplexan in the phylum sporozoa, order haemosporidida (parasitic in the blood of vertebrates). The genus *Plasmodium* consists of four species of morphologically distinct human malaria parasites: *P. falciparum, P. vivax, P. ovale* and *P. malariae* (Cox, 1993). All four human species are universally distributed although they are more prevalent in the tropics and subtropics. *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).

Human malaria parasites all undergo the same bi-phasic life cycle. Fig. 1.1 illustrates the sexual (in the insect host) and asexual phases (in the human host) 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 2-10 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 to commence erythrocytic schizogony consisting of four stages: merozoites develop into the ring stage (immature trophozoites visible in peripheral blood of the infected patient), which in turn mature into trophozoites and multinucleate schizonts; schizonts erupt after 48-72 hours and release as many as 36 daughter merozoites per schizont which initiate the next round of invasion and multiplication in erythrocytes. The rupture of the schizonts is typically associated with bouts of fever. Alternatively, some of the merozoites are capable of developing into gametocytes which are in turn ingested by a mosquito to complete the cycle (Cox, 1993).

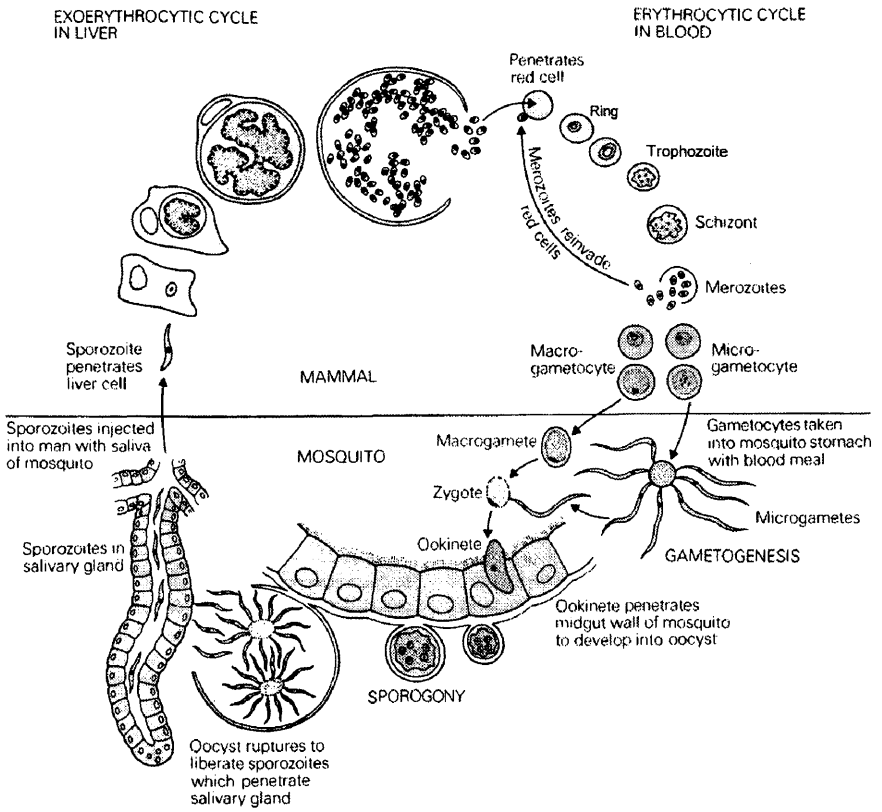

Figure 1.1: Bi-phasic life cycle of the *Plasmodium* parasite. Adapted from (Cox, 1993)

1.2 Malaria Pathogenesis.

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 and acidosis, severe liver failure, circulatory collapse and non-cardiac pulmonary oedema often resulting in death (Mendis and Carter, 1995; Ramasamy, 1998).

The above mentioned morbid conditions associated with the disease are ascribed to various host-parasite interactions. Toxins are released by the intra-erythrocytic parasite including malarial mitogens, toxic proteins and polar lipids, particularly glycophosphatidyl inositol (GPI; covalently bound to the merozoite surface antigens, MSA1 and MSA2). The main consequence of these bio-active molecules is to direct the systemic release of several pro-inflammatory cytokines, in particular tumor necrosis

factor α (TNFα), interferon γ (IFN γ), interleukin 1 (IL1) and IL6 from host macrophages/monocytes and γδ T-cells (Miller *et al.*, 1994; Hommel, 1997; Ramasamy, 1998).

One of the most distinctive pathophysiological characteristics which evolved for the survival of *P. falciparum,* is the 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) and to platelets, monocytes and lymphocytes (Berendt *et al.*, 1994). 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 (Ramasamy, 1998). To this end, electron-dense protrusions (70-90 nm) termed knobs are inserted on the surface of the erythrocyte membrane and include parasite-induced components such as the knob-associated histidine rich protein (HRP1) and a large (315 kDa) *P. falciparum*-infected erythrocyte membrane protein 3 (PfEMP 3) (Oh *et al.*, 1997). Other modifications include Pfalhesin (dimerised Band 3) and sequestrin (Berendt *et al.*, 1994; Hommel, 1997). Polymorphic 22-28 kDa proteins termed rosettins may mediate rosetting. The most potent modification is the expression of a family of *P. falciparum*-infected erythrocyte membrane proteins (200-300 kDa) called PfEMP1 that have been implicated as the main mediators of adhesion in both sequestration and possibly also in rosetting (Ramasamy, 1998).

The expression of host receptors for adherence by infected erythrocytes is to a large extent modulated by the parasite-induced local accumulation of different cytokines. These receptors include intracellular adhesion molecule 1 (ICAM1), vascular cell adhesion molecule 1 (VCAM1), E-selectin, P-selectin, CD 36, thrombospondin, glycoaminoglycan, chondroitin sulphate A and thrombomodulin (Hommel, 1997; Oh *et al.*, 1997).

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; Ramasamy, 1998).

1.3 Control 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 based on the following assumptions: 1) Human malaria could be eliminated by treatment with antimalarials and 2) Transmission could be eliminated by insecticides such as dichloro-diphenyl-trichloro ethane (DDT) (Krogstad, 1996). Malaria was subsequently eliminated from Europe, most of the Asian regions of Russia, the United States of America and most of the Caribbean (Krogstad, 1996). The eradication programme has however been abandoned since it became obvious that malaria cannot be dealt with as a single and uniform world-wide problem susceptible to one global control strategy (Collins and Paskewitz, 1995). Attention is now focused on creating alternative strategies consisting of chemotherapy/prophylaxis, vector control and vaccines.

1.3.1 Chemotherapy and –Prophylaxis.

1.3.1.1 Quinoline-containing Antimalarials.

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 4-amino-quinolines chloroquine, amodiaquine, mecaprime and sontaquine and derivatives such as

mefloquine and halofantrine. Of these, chloroquine is the most widely used antimalarial for both chemoprophylaxis and –therapy (Wellems, 1991; Cowman, 1995).

The quinoline-type antimalarials have their effect on the mature stages of the asexual intra-erythrocytic parasite. Chloroquine is concentrated in the acidic food vacuole of the parasite and causes the swelling, vesiculation and accumulation of undigested haemoglobin which suggests that the major action of these drugs is to block the function of the food vacuole. Normally, the haemoglobin is digested into usable short peptides and the released haeme 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. An enzyme activity called haeme polymerase appears to be involved (Foote and Cowman, 1994). Hypotheses for the mechanism of chloroquine action include increasing the vacuolar pH, the inhibition of enzymatic functions of the organelle, and the possible uncoupling of FP sequestration into haemozoin resulting in accumulation of toxic FP (Cowman, 1995).

In the early 1960's 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). Chloroquine resistant parasites expel the drug 40-50 times more rapidly compared to sensitive isolates. In 1997, Wellems' group identified a 36 kb segment of chromosome 7 related to chloroquine resistance. This segment contains a gene, called *cg2* that encodes a unique ~330 kDa protein with complex polymorphisms expected of the chloroquine resistance phenotype. The CG2 protein is found at the parasite periphery and also in association with haemozoin in the food vacuole (Su *et al.*, 1997).

Since the mechanism of expulsion is similar to the multi-drug resistant (MDR) phenotypes of mammalian tumor cells, the possible presence of P-glycoprotein and *mdr* genes in *P. falciparum* was investigated. Two targets were discovered, Pgh-1 and Pgh-2 encoded by *pfmdr-1* and *pfmdr-2* genes, respectively, but could not be linked to chloroquine resistance (Wellems *et al.*, 1990). *Pfmdr-1* gene seems to be amplified and overexpressed in all mefloquine-resistant field isolates which tend to be resistant to lower

levels of chloroquine, thus indicating an inverse *in vivo* relationship between chloroquine and mefloquine resistance. The same field isolates were also resistant to halofantrine and quinine, indicating cross-resistance between the structurally similar compounds (Foote and Cowman, 1994; Borst and Ouellette, 1995).

1.3.1.2 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, disrupts DNA synthesis, decreases methionine synthesis and also reduces the rate of Ser to Gly conversion. The two major enzymatic targets in the folate pathway are dihydrofolate reductase (DHFR) and dihydropteroate synthetase (DHPS).

The two most well-known DHFR inhibitors are pyrimethamine (usually used in chemotherapy) and proguanil (used for chemoprophylaxis). Proguanil is a pro-drug and is converted to its active metabolite cycloguanil by the mixed-function oxidase system of the liver. The DHFR inhibitors are competitive with the dihydrofolate substrate. In all protozoa, DHFR is a bi-functional enzyme with thymidylate synthase (TS) activity (Borst and Ouellette, 1995; Cowman, 1995).

In contrast to chloroquine resistance that has spread from two foci, resistance to the DHFR inhibitors developed independently in many different regions. At this stage, antifolate resistance is common in South America, Southeast Asia and also in Africa (Krogstad, 1996). A recent study by our group indicated that resistance has even spread as far south as the north-eastern parts of South Africa (Birkholtz *et al.*, 1998a; Birkholtz *et al.*, 1998c). Genetic analysis of DHFR genes from all the pyrimethamine resistant *P. falciparum* isolates tested so far indicated a single nucleotide point mutation (AGC-AAC) at codon position 108, changing Ser to Asn in the DHFR active site (Cowman *et al.*, 1988; Wellems, 1991). This suggests that the major mechanism of resistance could be a reduced binding affinity for the competitive inhibitor. Two additional mutations, Asn51Ile and Cys59Arg, in conjunction with Ser108Asn further increase the level of pyrimethamine resistance (Basco *et al.*, 1995). Recently, two additional mutations were also reported namely Val140Leu and Cys50Arg (Wang *et al.*, 1997).

Proguanil/cycloguanil resistance also mapped to a mutation at codon 108 except that Ser is changed to Thr (Foote and Cowman, 1994; Reeder *et al.*, 1996). A further synergistic mutation at codon 16 (Ala16Val) has been observed in one field isolate only. These mutations resulted in a parallel and moderate reduction in pyrimethamine resistance. Cross-resistant parasites to both pyrimethamine and proguanil have been observed where a codon 108 mutation occurs together with mutations at codons 51 and 59 and to another mutation at codon 164 (Ile164Leu) (Wellems, 1991; Cowman, 1995).

Pyrimethamine is usually used in combination with one of the sulpha drugs, sulphadoxine, which completely prevents folate synthesis by simultaneously inhibiting DHFR-TS and DHPS. DHPS is a bi-functional enzyme with 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (PPPK) which catalyses the preceding step in the pathway. The sulphone and sulphonamide drugs probably act as para-amino benzoic acid (a precursor of folate) analogues which inhibit DHPS, but could also give rise to toxic products (Cowman, 1995). Amino acid differences in DHPS from resistant and sensitive *P. falciparum* have been identified but not yet directly linked to the mechanism of resistance, although sulphadoxine metabolism is reduced in resistant parasites (Reeder *et al.*, 1996). These mutations include Ser436Phe or Ala, Ala437Gly, Lys540Glu, Ala581Gly or Tyr and Ala613Ser, with the substitutions at positions 437 and 581 being the most common (Wang *et al.*, 1997; Hyde, 1998).

1.3.1.3 Antimalarials under Development.

A new development in antimalarial therapy comes from a 2000-year-old remedy against malaria used in China. Artemisinin (Qinghaosu) was isolated in 1972 from the aerial part of *Artemisia annua*. Although artemisinin and some hydrophilic derivatives have been registered for use in China and several countries in Southeast Asia, it has not yet been approved in other countries (Olliaro and Trigg, 1995; Kirby, 1997).

1.3.2 Strategies for Vector Control.

The failure of the malaria eradication program and emergence of insecticide-resistant mosquitoes have shifted the focus to other strategies for the control of human-vector contact. 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). Vector control by larvidicing has been used in some instances but it is generally accepted to be inefficient in reducing transmission caused by the adult population. As a consequence biological control agents such as larvivorous fish and bacterial endospore toxins from *Bacillus sphaericus* have been investigated. Although successful, it is too expensive for common application. The most promising, although still speculative development, has been the possible introduction of genetically modified vectors (that are unable to host the malaria parasite) into natural populations (Collins, 1994; Collins and Paskewitz, 1995).

1.3.3 Malaria Vaccines.

The feasibility of malaria vaccines was first indicated in the early 1970's 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 (Hommel, 1997). 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 20 years since the first efforts, a 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. Three main strategies in malaria vaccine development have been investigated: 1) blocking infection at the pre-erythrocytic stage, 2) inhibition of the blood stage (asexual erythrocytic stage) where both parasite growth and pathogenesis is blocked and 3) a transmission blocking stage (Facer and Tanner, 1997). 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.

The most successful asexual blood-stage vaccine to date has been the multicomponent Spf66 developed by Patarroyo and colleagues. They identified potential antigens that gave protection to *Aotus* monkeys challenged by *P. falciparum,* and constructed a chimeric protein particle consisting of four epitopes. 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 (Amador and Patarroyo, 1996; Facer and Tanner, 1997).

An effective vaccine is envisaged as a cocktail of a multistage, -valent, -immune response vaccine (Doolan and Hoffman, 1997). The most promising candidates are recombinant DNA vaccines. It is possible to combine 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 just 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: (Nussenzweig and Long, 1994; Doolan and Hoffman, 1997; Facer and Tanner, 1997; Kwiatkowski and Marsh, 1997).

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	Development in hepatocyte
	Pfs 16	N/A	?
	Sporozoite surface threonine and aspartic acid rich protein	STARP	?
Asexual blood-stage	Merozoite surface antigen	MSA-1/2/3 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
*** 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	N/A	?
*** 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
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 dissertation is based on the assumption 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 an intervention strategy. Once a target protein has been identified and characterised it can be evaluated in terms of validity and considered in the design of novel antimalarials. The following

sections deal with the metabolic processes of *P. falciparum*, and primarily focus on polyamine biosynthesis.

1.4 Nutrient Requirements and Transport Pathways of the *P. falciparum* Parasite.

Although the erythrocyte provides protection to the intracellular parasite against the immune system of the host, this cell is unusual in the context of the average cell architecture and function. The mature erythrocyte has limited metabolic activity and protein synthesis is absent. The metabolically active intra-erythrocytic parasite requires glucose, purine nucleosides, L-glutamine and other essential amino acids, choline and *myo*-inositol for growth, as well as a mechanism to export toxic by-products of glycolysis e.g. lactate (Ginsburg and Stein, 1987; Elford *et al.*, 1995; Deitsch and Wellems, 1996).

To import nutrients, the parasite either increases the permeability (for trafficking of small solutes) or induces transport pathways in the host membrane (Ginsburg and Stein, 1987; Deitsch and Wellems, 1996). In order to reach the intracellular parasite, solutes need to cross three membranes: 1) the host erythrocyte membrane (EM), 2) the parasite vacuole membrane (PVM) and 3) the parasite plasmamembrane (PPM) itself. Alternatively, there is the direct access model in which the parasite has access to the extracellular environment through a 'metabolic window' or membranous duct (parasitophorous duct) involving membranous structures known as tubovesicular membrane networks (TVM) (Gero and Kirk, 1994; Deitsch and Wellems, 1996). Transport could be facilitated by parasite-induced transport proteins located at any of the three membrane structures, to support either the conventional or direct access models (Louw, 1998). Recently, our group has managed to identify a putative glucose transporter gene in *P. falciparum,* although the location and mechanism of action of the protein in the parasite is still unclear (K. Nel, personal communication).

1.5 Metabolic Processes and Enzymes as Potential Drug Targets in *P. falciparum.*

Studies on the metabolic pathways of *Plasmodia* have led to the identification of many characteristics unique to the parasite. Table 1.2 summarises the identified potential

targets for intervening with *P. falciparum* metabolism and indicates whether the respective genes have been cloned and sequenced.

Table 1.2: Summary of the major metabolic target proteins in *P. falciparum*. This table was compiled from references: (McKerrow *et al.*, 1993; Craig and Eakin, 1997; Subbayya *et al.*, 1997; Olliaro and Yuthavong, 1998).

Target protein	Function/Metabolic pathway	Gene cloned
Hexokinase (HK)	First enzyme in glycolysis	Yes
Aldolase	Glycolysis	Yes
Lactate dehydrogenase (LDH)	Last enzyme of glycolysis	Yes
Triosephosphate isomerase (TIM/TPI)	Glycolysis	Yes
Hypoxanthine-guanosine phosphoribosyl transferase (HGPRT)	Key enzyme in purine salvage pathway.	Yes
Adenylosuccinate lyase	Formation of AMP from IMP	No
Adenylosuccinate synthase (ASS)	Formation of AMP from IMP	No
Inosine monophosphate dehydrogenase (IMPDH)	Formation of GMP from IMP	No
Guanosine monophosphate synthase	Formation of GMP from IMP	No
Dihydroorotase	*De novo* pyrimidine synthesis	No
Dihydroorotase dehydrogenase (DHODH)	*De novo* pyrimidine synthesis	Yes
Carbamoyl phosphate synthetase	*De novo* pyrimidine synthesis	Yes
Orotoate phosphoribosyl transferase	*De novo* pyrimidine synthesis	?
Orotidine-5′-monophosphate decarboxylase	*De novo* pyrimidine synthesis	?
Dihydrofolate reductase-thymidilate synthase (DHFR-TS)	*De novo* pyrimidine synthesis/ Folate synthesis	Yes
Dihydropteroate synthetase-pyrophosphokinase (DHPS-PPPK)	Folate synthesis	Yes
Haeme polymerase ?	Degradation of haemoglobin	Yes (HRP II/ III/ IV)
Phosphatidyl choline cytidyl transferase (CTP)	Phospholipid synthesis	Yes
Sphingomyelin synthase	Tubovesicular membrane (TVM) synthesis	No
Falcipain (Cysteine protease)	Degradation of haemoglobin	Yes
Plasmepsin I (Aspartic protease)	Degradation of haemoglobin	Yes
Plasmepsin II (Aspartic protease)	Degradation of haemoglobin	Yes
Serine protease	Involved in MSA-1 processing	No
Ribonucleotide reductase	Cell Cycle regulated enzyme	Yes
DNA topoisomerase I and II	DNA replication	Yes
RNA polymerase (plastid encoded)	Transcription	Yes
DNA polymerase α	DNA replication	Yes
LSU rRNA (plastid encoded)	Protein synthesis	Yes
Tubulin (α, β subunits)	Cytoskeleton	Yes
Ornithine decarboxylase (ODC)	Polyamine biosynthesis	No
S-adenosyl methionine decarboxylase	Polyamine biosynthesis	No

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. The parasite has one mitochondrion but the contribution of electron transport to cellular ATP synthesis is relatively small (Opperdoes, 1995), although inhibitors of electron transport and mitochondrial proteins do have antimalarial effects. *P. falciparum* has a complete set of glycolytic enzymes but only four of the enzymes, namely hexokinase, aldolase, LDH and TIM (or TPI) (listed in Table 1.2) have either some sequence heterogeneity or a different tertiary structure in comparison to the host enzymes that could be exploited for antimalarial drugs (Subbayya *et al.*, 1997) (Sherman, 1979).

Virtually all parasites lack the enzymes for *de novo* synthesis of purines and they scavenge these nucleosides from their hosts. The enzymes needed for the salvage pathway are therefore potential targets for interruption of parasite growth and development (Table 1.2). The major purine source for the parasite is hypoxanthine (the base of inosine nucleosides) which is normally generated from ATP in the erythrocyte (Sherman, 1979; Fairlamb, 1989). Once in the parasite, hypoxanthine is converted to both adenosine and guanosine via IMP and the key enzyme, hypoxanthine-guanosine phosphoribosyltransferase (HGPRT), for which no unique structural features have been identified as yet (Berens *et al.*, 1995; Craig and Eakin, 1997). Furthermore, the conversion of IMP to AMP involves the enzymes adenylosuccinate lyase and –synthase (Table 1.2) 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. Orotoate phosphoribosyl transferase and orotidine-5′-monophosphate decarboxylase are separate enzymes, in contrast to the host which has a bi-functional enzyme; together with dihydroorotase, they hold promise for intervention strategies. Dihydroorotate dehydrogenase (DHODH) bridges pyrimidine synthesis with mitochondrial electron transport and is the proposed target for a new antimalarial, atovaquone (Berens *et al.*, 1995; Olliaro and Yuthavong, 1998).

De novo synthesised folate is a co-factor in one-carbon substitution reactions and is a precursor for thymidylate synthesis. The two bi-functional enzymes in this pathway, DHPS-PPPK and DHFR-TS have been thoroughly studied and are the targets of the antifolates as was discussed in section 1.3.1.2.

The primary source for amino acids 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. Although the mechanism is not yet defined it could provide the basis of an intervention strategy (Francis *et al.*, 1997; Louw, 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. Table 1.2 indicates two serine proteases, three cysteine, five aspartic and two aminopeptidase proteases which have been described for *P. falciparum* (McKerrow *et al.*, 1993).

Selective differences have been identified between enzymes involved in nucleic acid metabolism. *P. falciparum* DNA polymerase α shows only 14-17% homology with other eukaryotes, and ribonucleotide reductase genes (involved in DNA replication and repair) differ at the 3′-end from their human counterparts. DNA topoisomerase inhibitors (antitumor and antibiotic agents) have proved moderately effective against *P. falciparum.* In addition to the normal nuclear DNA, apicomplexan parasites (*Plasmodia, Toxoplasma, Babesia* etc.) have two maternally inherited extra-chromosomal DNA-containing organelles: a) A 6 kb, linear mitochondrial DNA and b) A plastid-like, 35 kb circular molecule. Both these DNA molecules are possible targets for selective inhibition (Olliaro and Yuthavong, 1998).

1.6 Polyamines.

The polyamines, spermine (N,N^1-Bis(3-aminopropyl)-1,4-butanediamine), spermidine, (N-(3-aminopropyl)-1,4-diaminobutane) and putrescine (1,4-butanediamine) are highly cationic, polybasic molecules found ubiquitously in all living cells, both pro- and eukaryotic, at relatively high concentrations. Their polycationic character gives them a high affinity for acidic cell constituents including nucleic acids, acidic proteins and phospholipids (Algranati and Goldemberg, 1989). Rapidly growing cells contain higher levels of polyamines and their biosynthetic enzymes than quiescent cells. New polyamine biosynthesis has been associated with the progress of cells through the cell cycle and cellular differentiation (Heby, 1989; Oka and Borellini, 1989).

In *P. falciparum,* polyamine biosynthesis reaches peak levels during the early trophozoite stage, whereas nucleic acid and protein synthesis peak in mature trophozoites (Gutteridge and Trigg, 1972). Inhibition of polyamine biosynthesis does not interfere with merozoite invasion or ring development, but prevents the transformation of trophozoites into schizonts and thus the *in vitro* proliferation and development of the parasites (Assaraf *et al.*, 1984; Bitoni *et al.*, 1987). These results indicate the importance of the polyamines in the schizogony of *Plasmodia*. *P. falciparum* contains 9 pmol putrescine, 33 pmol spermidine and 8 pmol spermine per 10^6 parasitised erythrocytes (Tabor and Tabor, 1984).

There is no polyamine biosynthetic machinery in erythrocytes and only traces of spermidine and spermine. The polyamines are therefore formed from ornithine as by-products of the urea cycle in *P. falciparum,* as is typical for other organisms. 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. 1.2). 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. The interconversion of the different polyamines occurs via the transfer of aminopropyl groups and oxidation (Pegg and McCann, 1982; Tabor and Tabor, 1984). 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). A putrescine transporter has been discovered in *P. knowlesi* (Singh *et al.*, 1997).

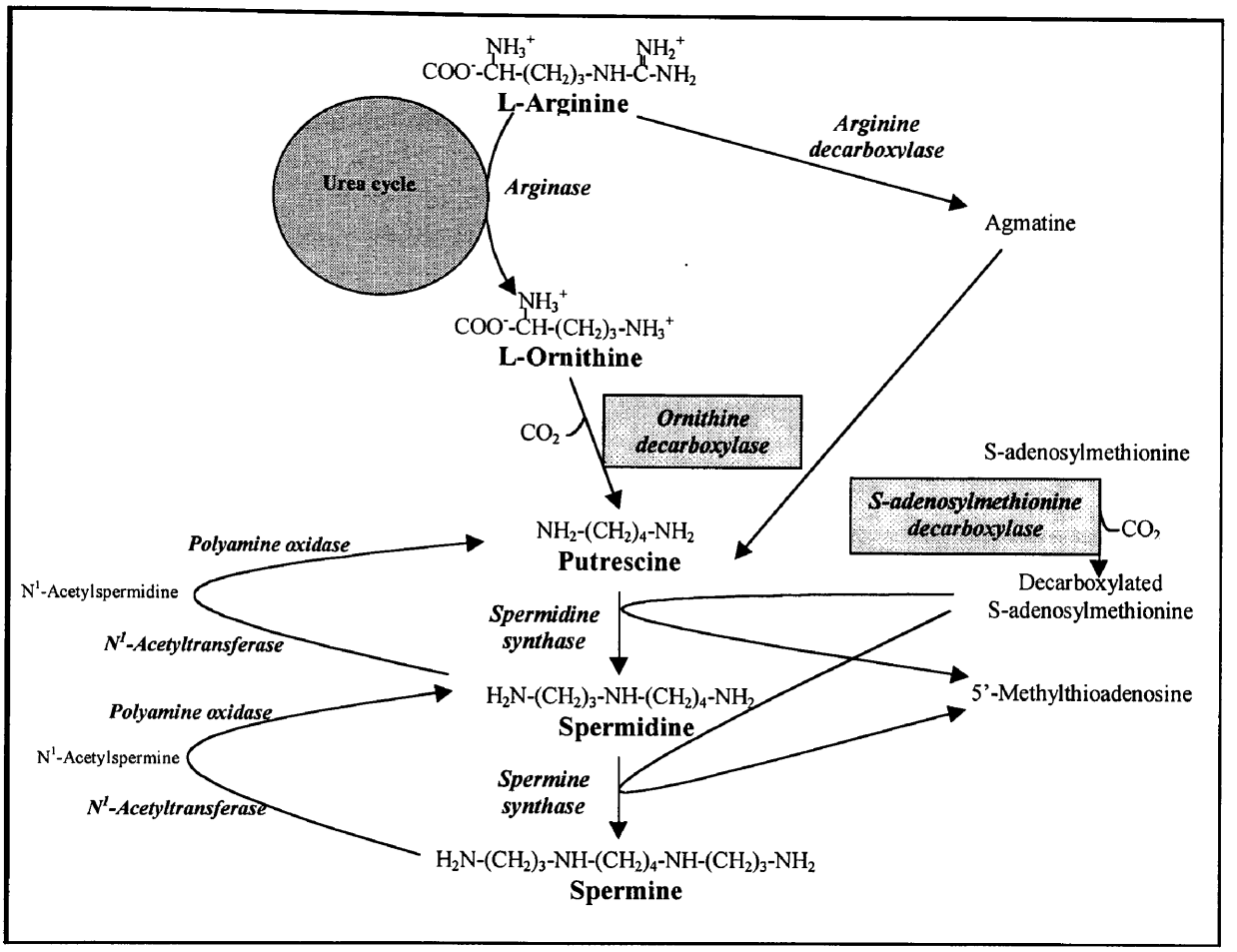

Figure 1.2: General pathway for the biosynthesis of the polyamines in pro- and eukaryotes. Enzymes are italicised and the two regulatory enzymes indicated in shadowed boxes. Prokaryotes can synthesise putrescine from agmatine.

Polyamines can be cytotoxic if they are present in excessive amounts in cells. The fact 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 anti-parasitic drugs. In the case of the polyamine synthesis pathway, the two decarboxylase enzymes are of importance since polyamine synthesis is totally dependent on their activities. Ornithine decarboxylase (ODC) is of particular importance, because it is the rate-limiting enzyme in the pathway, which makes it an ideal target to block polyamine biosynthesis totally (Heby, 1985).

1.7 Ornithine Decarboxylase (ODC: E.C.4.1.1.17).

1.7.1 Structure and Properties.

Mammalian ODC is an inducible enzyme that constitutes a minor fraction of the cellular protein content: 0.00014-0.05% in stimulated cells as opposed to 0.000002% in unstimulated cells (Heby, 1985; Heby and Persson, 1990). The enzyme is classified as a growth related enzyme located predominantly or exclusively in the cytoplasm. ODC activity was however also demonstrated in the nucleus, suggesting a cytoplasmic store of catalytically active ODC (Heby, 1985; Canellakis and Hayashi, 1989). The low cellular level and rapid loss of activity during purification made it difficult to isolate homogeneous ODC. The inclusion of sulphydryl reducing agents and non-ionic detergents in purification protocols stabilised ODC and led to the isolation of a 7600-fold purified enzyme activity in *P. falciparum* (Assaraf *et al.*, 1988).

ODC is a pyridoxal-5′-phosphate (PLP) dependent enzyme belonging to the Group IV decarboxylase family of proteins (including ornithine, arginine and diaminopimelic acid (DAP) decarboxylases) and acts physiologically on L-ornithine to form putrescine (Pegg *et al.*, 1994). *P. falciparum* ODC is an acidic protein with an optimal pH 8.0, correlating with other organisms (Kaye, 1984; Heby, 1985; Pegg, 1989). The reported Km values for L-ornithine are in the μM range (Km ~0.025 mM for *P. falciparum* (Assaraf *et al.*, 1988)).

Denaturing polyacrylamide gel electrophoresis of *P. falciparum* ODC indicated a MM of 51 kDa (Assaraf *et al.*, 1988). The mammalian and prokaryotic ODC is proposed to be a homodimeric enzyme of two identical monomers. ODC has a MM of about 100 000 with protomer MM ranging between 50-54 kDa, depending on the species (Heby, 1985). The dimer contains two active sites located at the interface between the monomers and is made up of parts of each subunit (Tobias and Kahana, 1993a). The active site is formed between Cys-360 on one subunit and Lys-169 and His-197 on the other (residue numbers are for the murine ODC), comprising of hydrophilic residues as shown in Fig. 1.3. The

amino acid composition of the dimer interface is apparently highly conserved between species.

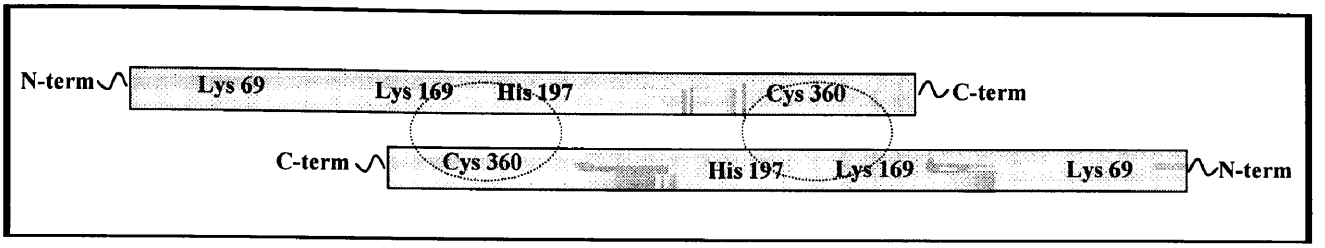

Figure 1.3: Diagrammatic representation of the murine ODC homodimer. The location of the active site at the interface between the subunits are indicated in ellipses with residues Cys-360, His-197, Lys-169. The PLP-binding site is located at Lys-69 (Pegg, 1989).

1.7.2 Turnover Rate of Mammalian ODC.

The rapid turnover rate (rate at which individual proteins are degraded and replaced by synthesis of new protein) of mammalian ODC was first reported by Russell and Snyder (Hayashi and Murakami, 1995). Since then numerous reports have indicated a $t_{1/2}$ (time required for 50% degradation of the protein) ranging from a few minutes to one hour (~5-35 minutes on average), making it the shortest $t_{1/2}$ yet reported for all known enzymes. This turnover rate is essential for rapid and dramatic changes in the enzyme level in response to various stimuli and thus regulation of ODC (Heby, 1985; Hayashi and Murakami, 1995).

Due to this distinctive characteristic, the following sections will focus on the multiple mechanisms involved in regulating both the activity and cellular concentration of ODC. Very little is known about the regulation of *P. falciparum* ODC and this necessitates extending the discussion to ODCs from other well-characterised organisms.

1.7.3 Regulation of ODC Concentration and Activity.

The induction of ODC activity by growth factors and hormones has been shown to be accompanied by a roughly parallel increase in the amount of ODC mRNA (Feinstein *et al.*, 1985). The accumulation of mRNA in stimulated tissues appears to be due to the stabilisation of the message leading to increased levels of ODC protein being translated (Heby and Persson, 1990; Hanson *et al.*, 1992). ODC mRNA contains 5′-untranslated

regions (5′-UTR) ranging from 245-310 nucleotides, which are extremely GC-rich and predicted to have extensive and thermodynamically stable secondary structures. These features could reduce the translational efficiency of the mRNA (Coffino, 1989).

Polyamines regulate ODC activity in a negative feedback mechanism by eliciting a rapid decay of pre-induced ODC activity (Seely and Pegg, 1983). Spermidine is the most effective and putrescine the least (Hayashi and Murakami, 1995). The polyamine-induced decay is time-dependent and is bigger than enzyme decay in the presence of cycloheximide (a potent protein synthesis inhibitor), indicating that it cannot be attributed solely to an inhibition of ODC synthesis (Hayashi *et al.* 1989a). Polyamine-induced ODC destabilisation is inhibited by cycloheximide but not by actinomycin D (an antibiotic and RNA synthesis inhibitor) (Kanamoto *et al.*, 1986), indicating that the effects of polyamines need *de novo* protein synthesis but not RNA synthesis, so that a protein that is induced at a post-translational level plays an essential role in destabilising ODC. Such a mechanism of enzyme inhibition via *de novo* synthesis of another inhibitory protein was a novel type of feedback inhibition and the inhibitor was therefore called antizyme (Heller *et al.*, 1976). Antizyme binds to ODC in a stoichiometric, reversible and non-competitive manner and is recycled after dissociation at high ionic strengths (Hayashi *et al.* 1989b)(Hayashi and Murakami, 1995). It is a protein of 20-27 kDa and has a $t_{1/2}$ comparable to that of ODC (18-66 min on average) (Hayashi *et al.*, 1996). Antizyme appears to be ubiquitous in all cells and tissues where ODC turns over rapidly and is repressed by polyamines (Hayashi *et al.* 1989b).

Antizyme is encoded in two open reading frames (ORF). The initiating frame (ORF1) is too short to encode a functional antizyme. The second and larger main frame (ORF2) overlaps ORF1 and is in the +1 frame to ORF1 but lacks an initiator AUG codon. This discontinuity implies an unusual mechanism involving translational frameshifting. This is a process in which the protein-synthesising machinery reads four bases as a codon for one amino acid and then continues reading triplets, and is elicited by the addition of polyamines (Rom and Kahana, 1994). The mechanism of stimulation of frameshifting by the polyamines remains to be clarified (Matsufuji *et al.*, 1995). Antizyme frameshifting is the first example of programmed frameshifting in the decoding of an eukaryotic gene and

the mechanism used is distinct from that of frameshifting used in the decoding of virus genes.

ODC may be degraded by proteasomes in an energy-dependent manner although Ub is not involved (Ciechanover *et al.*, 1984; Glass and Gerner, 1987). Proteasomes are known to exist as two types in the cytosol, namely 20S proteasomes (700 kDa, consisting of multiple subunits of 21-31 kDa) and 26S proteasomes (2000 kDa, which is made up by the 20S proteasome core and two additional regulatory subunits of 35-110 kDa) (Hayashi and Murakami, 1995). The 26S proteasome normally catalyses ATP-dependent degradation of poly-ubiquitinated proteins (Hayashi and Murakami, 1995). ODC degradation by purified 26S proteasome, which does not contain Ub or Ub-conjugating enzymes, required ATP for both the association of the proteasome subunits and also as an energy source and was antizyme dependent (Murakami *et al.*, 1992; Hayashi and Murakami, 1995). This is the first case of a non-ubiquitinated protein being degraded by the 26S proteasome and the first example of the *in vitro* degradation of a cellular protein by a specific ATP-dependent protease (Hayashi and Murakami, 1995).

ODC subunits rapidly dissociate and reassociate to allow antizyme binding at ODCs' N-terminus, fixing the enzyme in its inactive form. The binding of antizyme elicits a conformational change in ODC to expose its protease sensitive C-terminus containing the PEST-rich region (Mamroud-Kidron *et al.*, 1994). PEST-rich regions (regions rich in Pro, Glu, Ser, Thr and to a lesser extent Asp) are characteristic of proteins with $t_{1/2}$ less than two hours (Rogers *et al.*, 1986). The 26S proteasome then degrades ODC into different peptides of 5-11 amino acids, as indicated in Fig. 1.4 (Tokunaga *et al.*, 1994). These results suggested for the first time that the 26S proteasome acts as a multifunctional endoprotease complex (Hayashi and Murakami, 1995; Hayashi *et al.*, 1996).

In addition to its effect in mediating polyamine induced ODC degradation, antizyme seems to also effectively suppress cellular polyamine levels by inactivating the polyamine uptake system. Polyamine transport is known to be under negative feedback control by the polyamines and antizyme reversibly inactivates the polyamine transporter (Mitchell *et al.*, 1991). Thus, antizyme is a bifunctional regulatory protein that plays a pivotal role in the negative feedback system to prevent excess accumulation of cellular

polyamine (Fig. 1.4), although the exact molecular mechanisms have not been described (Hayashi and Murakami, 1995).

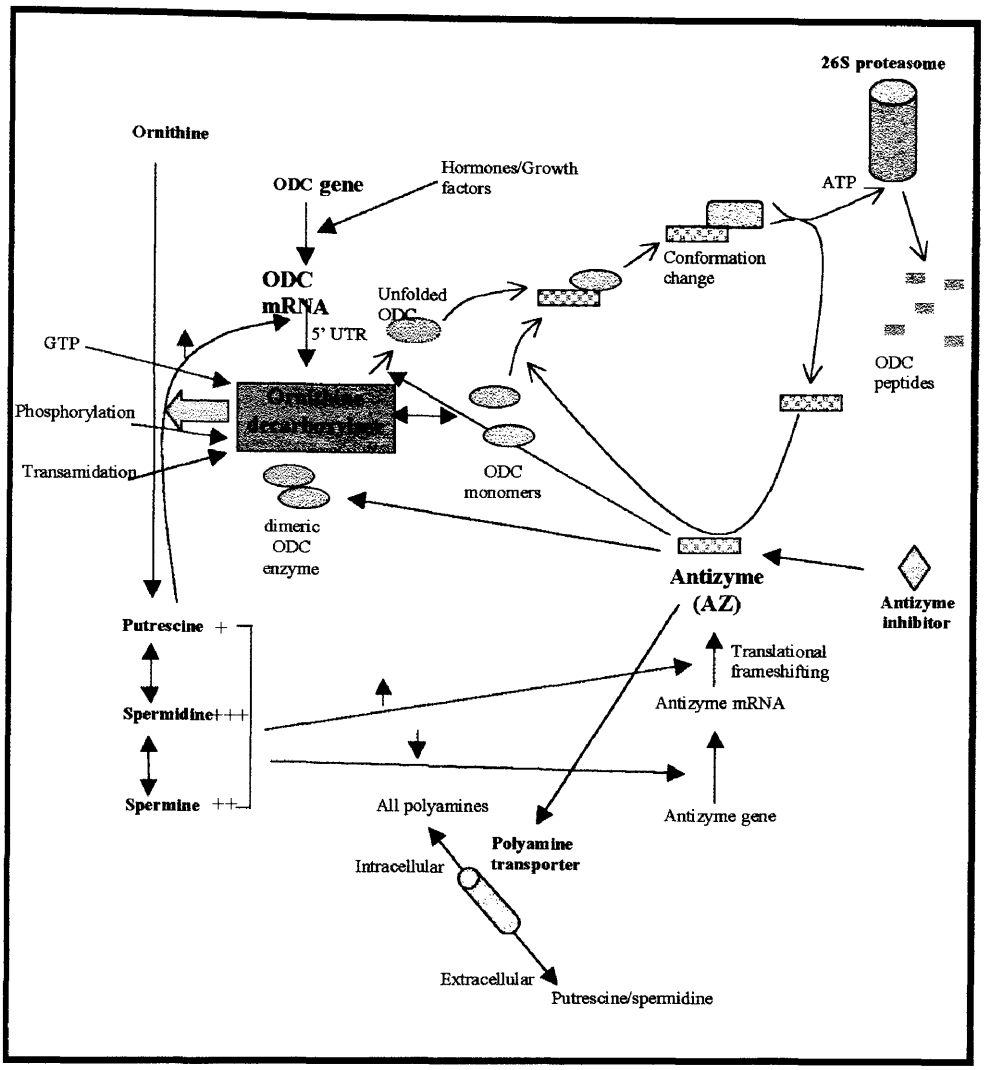

Figure 1.4: Schematic representation of the regulatory pathways involved in ODC regulation. Red arrows indicate inhibition and green stimulation or induction. Blue arrows indicate high or low levels of polyamines.

In 1982, Fujita and co-workers identified a specific macro-molecular inhibitor of antizyme in rat liver and named it the antizyme inhibitor (Fujita *et al.*, 1982). This protein is a dimer similar in size to ODC with monomer sizes of around 48-50 kDa (Hayashi and Murakami, 1995). Antizyme inhibitor (also called anti-antizyme) binds to antizyme with a higher affinity than ODC and reverses antizyme's inhibition of ODC activity in a time-dependent and stoichiometric manner (Fujita *et al.*, 1982). It has no ODC activity or any direct effect on ODC itself. Taken together, these results indicate that ODC activity depends not only on the amount of enzyme but also on the amount of antizyme as well as on the amount of antizyme inhibitor (Fig. 1.4).

1.7.4 ODC Regulation in the Parasitic Protozoa.

In the parasitic protozoa so far examined (*Trypanosoma brucei, P. falciparum, Leishmania donovani,* and *Crithidia fasciculata*), ODC has a relatively long half life as judged by the decay rate of enzyme activity in the presence of cycloheximide.

In African trypanosomes (*T. b. brucei*) a 35-amino acid C-terminal region is missing and 20 amino acids are added to the N-terminal region of the protein compared to the murine enzyme (Phillips *et al.*, 1987). The shortened C-terminal region does not contain one of the PEST-rich regions needed for degradation thereby stabilising the parasitic enzyme ($t_{1/2}$ > 6 hr) (Ghoda *et al.*, 1990). Recombining the *T. brucei* ODC with the C-terminus of mammalian ODC confers a short $t_{1/2}$ to the fused protein. The finding that trypanosome ODC behaves similarly in trypanosomes and mammalian cells with regard to stability suggests that similar rules for intracellular degradation apply in both protozoa and mammals (Ghoda *et al.*, 1990).

L. donovani ODC has ~200 amino acids added to the N-terminal and has a truncated C-terminus (28 amino acids) when compared to the murine ODC. The specific activity of this enzyme was not influenced by inhibition of protein synthesis for up to 20 hours, indicating stability in the absence of the C-terminal PEST-rich region. However, a second PEST-rich region at residues 591-602 was discovered but the enzyme activity was not altered in the presence of polyamines (Hanson *et al.*, 1992; Bacchi and Yarlett, 1995).

Inhibition of protein synthesis with cycloheximide indicated a decrease in ODC activity after 75 min in *P. falciparum.* ODC also seems to be translationally regulated in *P. falciparum* since addition of putrescine to *P. falciparum*-culture medium resulted in a rapid loss of ODC activity indicating a negative feedback mechanism which regulates the putrescine concentration (Assaraf *et al.*, 1987b; Assaraf *et al.*, 1988).

In contrast with the above, a recent publication described a metabolically unstable protozoan ODC lacking the C-terminal degradation domain. ODC from *Crithidia fasciculata* turns over rapidly with a $t_{1/2}$ of about 30 minutes and was not repressed by

polyamines. The deduced ODC amino acid sequence had 70% homology with *L. donovani* ODC and is extended by 265 N-terminal amino acids but lacks the crucial PEST-rich sequence at the C-terminal. It does contain two additional internal PEST-rich regions but no homology is apparent between these areas and mammalian ODCs (Svensson *et al.*, 1997). These findings suggest that other signals may mediate the fast turnover of this enzyme.

Antizyme activity could not be demonstrated in any of the protozoa studied so far. *T. brucei* has no antizyme activity and trypanosome ODC does not bind to murine antizyme (Hua *et al.*, 1995). *Tetrahymena pyriformis* also exhibits similar phenomena (Koguchi *et al.*, 1996). There is very little homology between the antizyme binding region of mouse ODC when compared to *T. brucei* and *C. fasciculata* (Svensson *et al.*, 1997). To date, there has not been any conclusive evidence for the existence of the 26 S proteasome complex required for antizyme-mediated ODC degradation in the protozoa.

1.7.5 ODC as an Inhibitory Target.

Due to its role as rate-limiting enzyme in polyamine biosynthesis, ODC serves as the obvious target for chemical intervention aimed at depletion of the polyamine pool. Early inhibitors of ODC included non-specific PLP antagonists, and a number of substrate and product analogues that proved both potent and specific. Another group of compounds are the unphysiological diamines which do not directly inhibit ODC but seem to repress the translation of the enzyme (Janne and Alhonen-Hongisto, 1989). A landmark in ODC inhibition was 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 (Heby, 1985; Janne and Alhonen-Hongisto, 1989).

The antiproliferative effects of polyamine depletion due to ODC inhibition 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, 1989). In 1990, DFMO was approved as the first drug in forty years for the treatment of

African sleeping sickness (*T. brucei*) (McCann and Pegg, 1992). *T. brucei* ODC is 60 times more sensitive to DFMO than the mammalian ODC (Janne and Alhonen-Hongisto, 1989). The seemingly inefficient inhibition of mammalian ODC might be due to its relatively short $t_{1/2}$ in comparison to the $t_{1/2}$ of the trypanosome ODC. This enzyme has been fully validated as a drug target in *T. brucei* with double-gene knockouts for ODC being totally deficient in ODC activity. The activity could be complemented by transfection with a genomic library clone containing the full-length *T. brucei* ODC gene (Wang, 1997).

DFMO also inhibits *P. falciparum* ODC *in vitro* and reduces the parasitaemia in patients by 42-70% (McCann and Pegg, 1992). 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). These results indicated a cytostatic rather than cytocidal effect of polyamine depletion. 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. 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).

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.

1.8 Strategies for the Isolation of *P. falciparum* ODC.

Efforts in isolating ODC from *P. falciparum* have so far only resulted in the isolation of a partially purified protein (Assaraf *et al.*, 1988). Possible reasons for the difficulty experienced could be the low abundance of the protein, especially in the parasite where it is active in only one of the four intra-erythrocytic stages. Therefore, alternative strategies must be applied to isolate sufficient amounts of the protein to allow extensive biochemical characterisations. Success has been achieved in various fields by isolating the corresponding gene followed by its expression to obtain the recombinant protein.

If the function of the encoded protein of a specific gene is to be studied, it is preferable to obtain the cDNA rather than the genomic DNA, which could contain introns and make expression of the active recombinant protein impossible. If a gene is therefore isolated from genomic DNA, the cDNA must subsequently also be isolated. Although this methodology is more labour-intensive and time-consuming, it may sometimes be the only way to identify the gene.

There are numerous approaches for isolating a specific gene. If a partial amino acid sequence is available, either from the N-terminal or from internal regions, (which is not the case for *P. falciparum* ODC), the homologous nucleotide sequence can be used as either a primer (for PCR) or as a probe. If no protein sequence is available, either primers or probes must be derived by identification of a conserved area in the amino acid sequences of the characterised protein from other organisms.

Labelled probes can be used to screen genomic or cDNA libraries of the organism to isolate a specific clone containing either the whole gene/cDNA or a partial sequence. The probe could also be used to perform *in situ* hybridisation of the organisms' chromosomes, thereby identifying a specific genetic locus spanning the gene. Chromosome walking and jumping (processes by which a specific area of the chromosome is sequentially sequenced until the desired gene is located) can then be applied to isolate the genomic DNA of the gene.

The development of the PCR in 1983 by Kary B. Mullis has revolutionised molecular genetics by providing a new approach for studying and analysing genes and their corresponding mRNA (Mullis, 1986). In conjunction with the reverse transcription reaction to create complementary DNA (cDNA) as a template, RT-PCR can be used to study mRNA almost at the level of a single cell.

Another interesting approach is to identify the gene based on the identification of the expressed protein. Expression cloning entails the identification of a particular protein expressed from a cDNA library (in an expression vector system) either through a functional analysis of the protein (through its ability to catalyse a specific reaction) or by immunological methods (with a specific antibody against the protein).

1.9 Research Objectives.

The aim of this study was the molecular characterisation of the cDNA of *P. falciparum* ODC. Once the cDNA is available, the recombinantly expressed protein can be characterised. The hypothesis is that this enzyme will, like some other malarial proteins e.g. DHFR, have insertions specific to the *P. falciparum* enzyme which could effect its activity and/or stability. The molecular characterisation of the cDNA of ODC will also enable the identification of possible polymorphisms unique to the parasitic enzyme.

The molecular identification and characterisation of the ODC cDNA entail the following:

- The design of suitable primers for the polymerase chain reaction amplification of the cDNA according to the rapid amplification of cDNA ends (RACE) protocol
- Cloning of the parasite ODC cDNA and characterisation by nucleotide sequence determination
- Identification of unique features of the deduced amino acid sequence of the *P. falciparum* ODC by comparison to other ODC protein sequences.

Chapter 2 addresses the identification of conserved amino acid sequences in ODCs of other related organisms used in the design of ODC-specific primers and their application in a 3′-RACE protocol. Chapter 3 presents the results of 5′-RACE protocols developed

to obtain the open reading frame of the cDNA. Chapter 4 deals with the molecular characterisation of the ODC cDNA of *P. falciparum* and is followed by the concluding discussion in Chapter 5.

Some of the results of this study have been presented at the Biochemistry in Africa congress (15th SASBMB/2nd FASBMB) (Birkholtz *et al.*, 1998b).

CHAPTER 2

Identification, cloning and nucleotide sequence of the 3′-fragment of the *P. falciparum* ODC cDNA obtained by 3′-RACE protocols.

2.1 INTRODUCTION.

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 of a cDNA library, which is a laborious and time-consuming task and usually yield 5′-truncated cDNA clones. Furthermore, the desired cDNA may be very rare and could be underrepresented in the library (Old and Primrose, 1994). In addition, cDNA library screening requires probes of sufficient length for high specificity hybridisations.

RT-PCR is extremely sensitive and rapid and is therefore an attractive approach for obtaining a specific cDNA sequence. Thermostable DNA polymerase (e.g. *Taq* from *Thermus aquaticus*) facilitates several million-fold amplification of specific templates. The combination with other polymerases (e.g. *Pfu* from *Pyrococcus furiosis*) with 3′-5′ exonuclease proofreading activity, further increases the specificity and fidelity of the reaction. Each component of a PCR, either physical (e.g. thermocycler, reaction tubes, cycle number and reaction temperatures) or chemical (e.g. concentrations of primers, dNTP, $MgCl_2$, DNA polymerase and template) needs to be optimised to obtain the best combination for reaction specificity and yield (Karcher, 1995; Kidd and Ruano, 1995). In this regard, an optimal reaction requires a low $MgCl_2$ concentration (excess $MgCl_2$ can reduce enzyme fidelity and may increase the level of non-specific amplification), a primer:template molar ratio of ~10^7, the least possible numbers of cycles (to prevent non-specific amplification occurring in the plateau phase) and an optimum DNA polymerase concentration (Kidd and Ruano, 1995).

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. Fig. 2.1 shows a 3′-RACE protocol where the reverse transcription reaction is driven by a differential display poly-T primer (DD-Poly-T). This primer contains a 3′-clamp to position the primer at the edge of the poly-A tail of the mRNA and an anchor sequence linked to the 5′-end to facilitate the RACE (Frohman, 1993; Karcher, 1995; Scheafer, 1995).

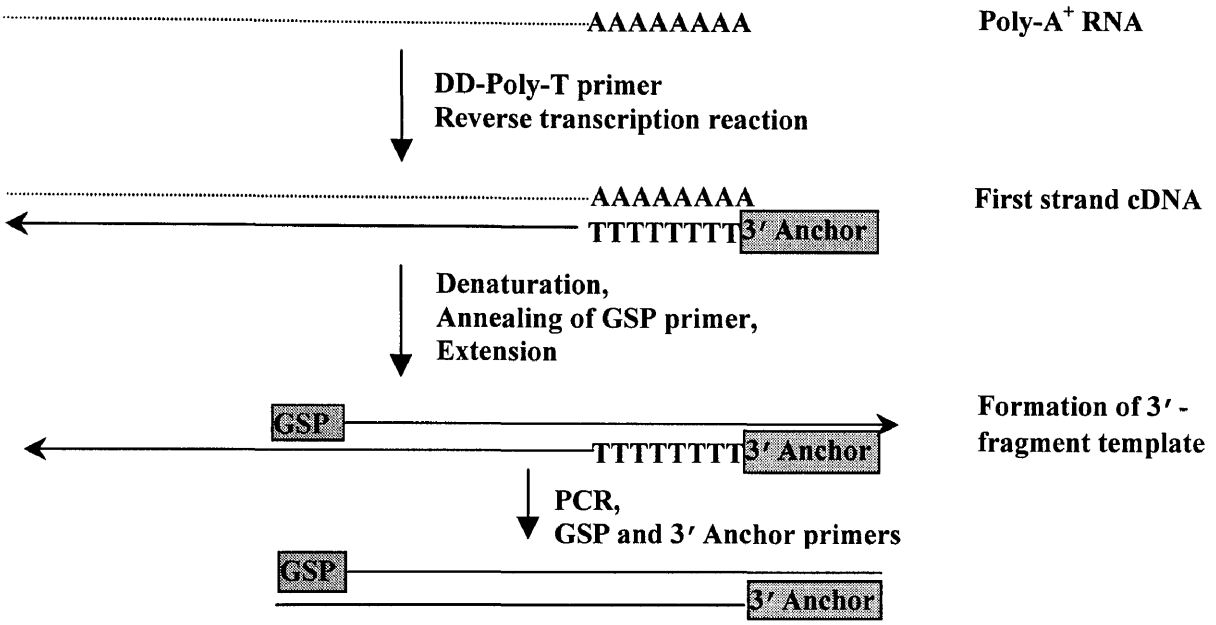

Figure 2.1: Identification of the full-length 3′-fragment of a specific cDNA by 3′-RACE protocols.

One of the most important factors affecting the specificity and quality of the PCR is the choice of primers. The design of primers for the amplification of genes for which no amino acid or nucleotide sequence data is available, requires sequence information from related organisms (Hyde *et al.*, 1989). These primers are without exception degenerate and the sequences are based on the codon preference of the specific organism. The requirements for a successful primer include: minimum self complementarity; high specificity; ΔG of between –5 and –10 kcal/mol at the 3′-termini; a AT/CG ratio complementary to the template; limited secondary structures (no intramolecular loops with negative ΔG values); primer length between 18 to 25 bp; a degeneracy below 128 and a small difference between the Tm of the template and less stable primer (Hyde and Holloway, 1993; Preston, 1993; Rychlik, 1993).

One of the most useful methods for determining the presence of a particular DNA/RNA sequence is the homologous hybridisation with a labelled probe. Non-radioactive labelling is 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. A phosphate group is removed from the colourless chromogenic substrate, 5-bromo-4-chloro-3-indolyl phosphate (BCIP/X-phosphate) by the alkaline phosphatase (Fig. 2.2). The indolyl product is then oxidised by nitroblue tetrazolium chloride (NBT) to the insoluble dimeric indigoid dye. This colour is amplified by the reduction of NBT to an insoluble, intense blue dye, NBT-formazan, to aid visualisation of hybrid molecules. This system is as sensitive as radioactive detection but is much safer to use and less time-consuming.

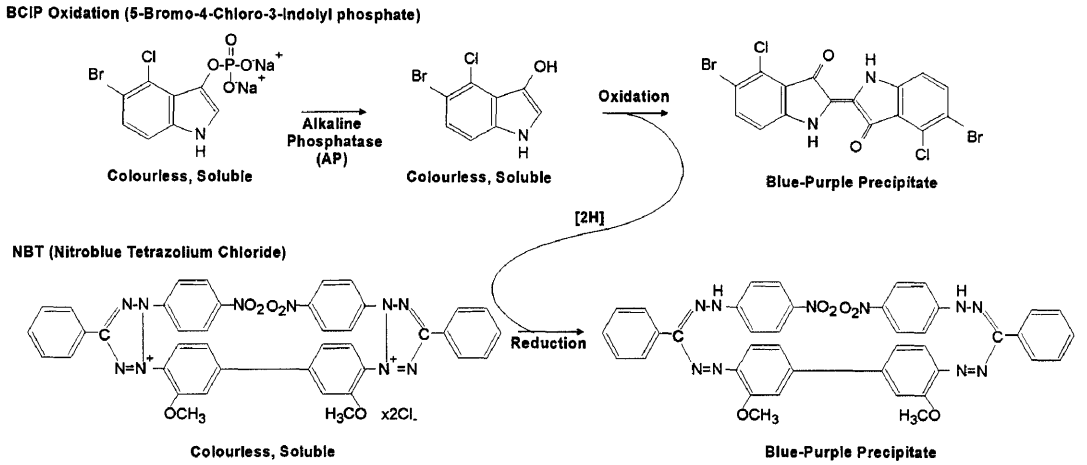

Figure 2.2: The reaction scheme of alkaline phosphatase with its chromogenic substrate. Adapted from (Karcher, 1995).

Cloning of PCR products involves blunt- or sticky-end cloning strategies or a combination of both. Another strategy makes use of the tendency of DNA polymerases (e.g. *Taq*) to add a non-template dependent adenosine to the 3′-end of each amplified strand. Cloning of A-tailed products consist of ligation to vectors containing T-residues as 3′-extensions (Zhou and Sanchez, 1997). This method does not require enzymatic manipulation of the PCR product prior to cloning as in sticky-end cloning, and is more efficient than blunt-end cloning. Selection of insert-containing plasmids on the basis of insertional inactivation of the *lac*Z gene entails disruption of the β-galactosidase α-

peptide sequence by the insert. This prevents α-complementation occurring between the host DNA and the vector, which normally produces a functional protein that can be assayed with a chromogenic blue substrate, X-gal (5-bromo-4-chloro-indolyl-β-D-galactoside).

The decisive step in identifying a specific gene is the determination of its nucleotide sequence. Techniques for this purpose include the Maxam and Gilbert technique and the more popular Sanger dideoxy-NTP (ddNTP) method (Sambrook *et al.*, 1989). The latter method was modified for automation by applying PCR in a cycle-sequencing protocol for incorporation of the ddNTPs. Detection is achieved by fluorophore tagging of either the sequencing primer or the ddNTPs. Each tag is discriminated in a single lane after electrophoresis, based on its fluorescence emission properties. Thermostable AmpliTaq DNA Polymerase FS (source: *Thermus aquaticus*) lacks 5′-3′ nuclease activity that reduces the discrimination against fluorescent labelled ddNTPs and permit a longer length of read. Four different fluorescent dyes identify the A, T, G, and C extension reactions. Novel high-sensitivity dyes contain a fluorescein donor dye (e.g. 6-carboxyfluorescein) linked to a dichlororhodamine (dRhodamine) acceptor dye. Automated sequencing has several advantages over the manual process: less template is needed, the shelf-life of the fluorescent dyes is much longer than for radio-isotopes, higher sensitivity and longer sequences are obtained and the process is not time-consuming.

When all of the above is taken into account, it should be possible to identify and characterise the cDNA of a putative anti-malarial target enzyme without prior knowledge of its amino acid or nucleotide sequence, as we describe in this chapter for the *P. falciparum* ODC cDNA.

While this study was in progress, the cDNA sequence of the *P. falciparum* ODC was deposited in Genbank (accession number: AF012551). No information was supplied on how the sequence was obtained. Correspondence with the research group (Dr. R. D. Walter, Biochemistry, Bernhard Nocht Institute for Tropical Medicine, Germany) with regard to the methods used and potential collaboration proved unsuccessful. At the time

of writing this dissertation (November 1998), the sequence and other details have still not been published. The methodology and results of the study presented in this dissertation were therefore developed and obtained independently.

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 and Jensen, 1976; Trager, 1994). 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. The culture was established in a 75 cm^3 culture flask (Corning, New York, USA) with 10 ml culture medium [1.04% (w/v) RPMI-1640 (Highveld Biologicals, South Africa); 24.96 mM HEPES (BDH); 22.22 mM D-glucose (Merck, Germany); 323.27 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 21.43 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 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 Quantification.

2.2.2.1 Spectrophotometrical Quantification

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).

2.2.2.2 Fluorometrical Quantification.

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.

2.2.2.3 Ethidium Bromide (EtBr) Dot Quantitation Method.

Ethidium bromide (2µg/ml) was added in equal volumes to a series of DNA standards (ranging from 2.5 µg/ml to 100 µg/ml) and spotted on a sheet of transparent plastic wrap stretched over a Spectroline TC-312 A UV transilluminator (312 nm wavelength). The concentrations of the DNA samples (especially of purified PCR fragments) were determined by comparison to the intensities of the standard series (Robyt and White, 1987; Sambrook *et al.*, 1989).

2.2.3 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 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 (136.9 mM NaCl, 1.45 mM KH_2PO_4, 2.68 mM KCl, 10.14 mM Na_2HPO_4, pH 7). The pellet was washed four times in total with 1×PBS and 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.

2.2.4 Analysis of RNA by Denaturing Agarose Electrophoresis.

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-H_2O (1 ml diethyl pyrocarbonate in 1L Milli Q H_2O) to a 30 ml agarose gel solution (Promega). The gel was run in a 1×MOPS buffer. Samples were denatured in 2 µl 10×MOPS, 3 µl formaldehyde and 10 µl deionised 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) (Sambrook *et al.*, 1989).

2.2.5 Complementary DNA (cDNA) Synthesis.

For the reverse transcription reaction, a DD-Poly-T primer (Genosys, Cambridge, England) was designed with the following sequence:

5′ GCT ACT ATT ACC ACA ACA CTC $(T)_{18}$ VN 3′

where V: A/G/C and N: A/T/G/C.

For all cDNA synthesis reactions, a recombinant reverse transcriptase (RT), Superscript II RT (Gibco BRL, Gaithersburg, USA), was used. It contains the *pol* gene of Moloney Murine Leukaemia virus (MMLV) and lacks RNAse H activity. This enzyme is capable of higher yields of full-length cDNA products.

2.2.5.1 cDNA Synthesis from Total RNA (DNA free).

Total RNA was pelleted from FORMAzol to avoid potential inhibitory effects of the formamide during RT-PCR. To 43.5 µg RNA in FORMAzol was added one-tenth the volume of 3 M sodium acetate and 2.5 volumes of cold absolute ethanol followed by incubation at -20°C for 16 hours. The RNA was sedimented by centrifugation at 13000×g for 30 min (4°C), the pellet was washed with 70% cold ethanol (in DEPC-H_2O) and airdried. The pellet was dissolved in 50 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.4) to which 20 U RNAsin was added (Promega, Wisconsin, USA). The concentration of total RNA was determined spectrophotometrically.

Contaminating genomic DNA was removed with RNAse free DNAse I (Boehringer Mannheim, Mannheim, Germany) in a reaction containing 10 mM $MgCl_2$, 1 mM DTT, 40 U RNAsin and 3.25 U DNAse I in a final volume of 100 µl with TE buffer. Incubation at 37°C for 30 min was followed by the addition of 25 µl of a DNAse Stop Mix (50 mM EDTA, 1.5 M sodium acetate, 1% SDS). The proteins were removed by phenol:chloroform:isoamyl alcohol extraction (25:24:1) followed by an extra chloroform:isoamyl alcohol (24:1) extraction. The DNA-free RNA was precipitated with

2.5 volumes of cold absolute ethanol for 5 hours at -20°C, collected by centrifugation at 13000×g for 15 min (4°C) and washed with 70% ethanol. The pellet was airdried, resuspended in DEPC-H_2O and the concentration determined spectrophotometrically.

5 µg of the treated total RNA was added to either 2.5 or 25 pmol DD-Poly-T primer (0.5 and 5 pmol primer/µg RNA respectively), denatured at 70°C for 10 min and snapcooled on ice. To this was added 4 µl 5×first strand buffer (250 mM Tris-HCl pH 8.3; 375 mM KCl; 15 mM $MgCl_2$), 2 µl 0.1 M DTT, 2 µl dNTPs (10 mM G/C; 15 mM A/T) and 200 U Superscript II RT (Gibco BRL, Gaithersburg, USA) in a final volume of 20 µl DEPC-H_2O. The reaction was incubated for 1 hour at 45°C and the enzyme denatured at 95°C for 3 min. The cDNA was stored at -70°C in silanised tubes.

2.2.5.2 cDNA Synthesis from Total RNA.

One µg of untreated total RNA in FORMAzol was used to which either 5, 10 or 20 pmol of DD-Poly-T primer was added. Following incubation at 70°C for 10 min, the solution was snapcooled on ice. The reverse transcription was initiated by the addition of 4 µl 5×first strand buffer (250 mM Tris-HCl pH 8.3; 375 mM KCl; 15 mM $MgCl_2$, Gibco BRL, Gaithersburg, USA), 2 µl 0.1 M DTT, 4 µl dNTPs (10 mM G/C; 15 mM A/T) in a final volume of 20 µl DEPC-H_2O. The reaction was equilibrated at 45°C for 2 min (hot-start) before addition of 200 U Superscript II RT (Gibco BRL) and incubation for 1 hour at 45°C. The enzyme was inactivated at 95°C for 3 min and the cDNA was stored at -70°C in silanised tubes.

2.2.5.3 Definition of the cDNA Concentrations.

The precise concentration of cDNA synthesised from total RNA is unknown and is defined as RNA equivalents. cDNA derived from DNA-free total RNA is referred to as processed total RNA-cDNA equivalents and cDNA from DNA-contaminated total RNA as unprocessed total RNA-cDNA equivalents. The concentration of the cDNA derived from processed total RNA was calculated from spectrophotometrical determinations of the total RNA concentration after DNAse I treatment. 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).

2.2.6 cDNA Quality and Quantity Assessment.

A control PCR was conducted on the known, highly abundant 18S small-subunit rRNA gene with the Bio1:Dig2 primer pair to estimate the representivity of the cDNA sample (Birkholtz *et al.*, 1998a). This primer pair should amplify a product of 408 bp. The quality, i.e. the number of truncations present, was monitored with the Bio1:3′ Anchor primer pair. The full-length product should be ~1800 bp. The 50 µl reactions contained 10 pmol each of primers Bio1-forward (ATC AGC TTT TGA TGT TAG GGT ATT G, Tm ~60°C) and Dig2-reverse (TAA TAA TAG AGG AAG CGT ATT, Tm ~50°C) or 3′ Anchor (see section 2.2.9), 0.2 mM GC/0.3 mM AT, 2 mM $MgCl_2$ and 250 ng processed total RNA-cDNA equivalents or 50 ng unprocessed total RNA-cDNA equivalents. The primers were obtained from Genosys, Cambridge, England. 2.5 U Takara ExTaq (Takara, Shuzo, Japan) was added in a hot-start protocol at 80°C (after a 3 min denaturation period at 94°C), followed by 30 cycles at 94°C for 30 sec, 48°C or 50°C for 30 sec and 72°C for 2 min.

2.2.7 Primer Design for RT-PCR.

The National Centre for Biotechnology Information (NCBI) SwissProt database was searched for known ODC amino acid sequences. Multiple-homologous alignments of ODC amino acid sequences of 13 related organisms with the CLUSTALW Program were used for the identification of conserved areas (Thompson *et al.*, 1994). 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 of ODC as well as to obtain a position specific sorting matrix (PSSM) of the BLOCKS alignment for graphical presentation (Pietrokovski, 1996). Further analysis using LAMA (Local Alignments of Multiple Alignments) was used to obtain alignments with other closely related proteins or protein families. The identified conserved areas were verified using the Basic Local Alignment Search Tool (BLAST) Program (Altschul *et al.*, 1990) for homology testing against the SwissProt database.

The amino acid sequence of the human ODC cDNA sequence was used as template 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 and GSP2, each with 64-fold degeneracy.

2.2.8 Dot-Blot of *P. falciparum* Genomic DNA with DIG-labelled GSP1.

To determine if primer GSP1 recognises a homologous segment in the *P. falciparum* genome, a dot-blot was performed with DIG-labelled GSP1. 100 pmol GSP1 was 3′-end labelled in a 20 µl reaction containing 1×tailing buffer, 5 mM $CaCl_2$, 0.05 mM DIG-ddUTP and 2.5 U terminal transferase (Boehringer Mannheim, Mannheim, Germany). 100 pmol human β-actin antisense primer was included for detection of human DNA contamination. The reactions were incubated for 15 min at 37°C, stopped by addition of 200 µM EDTA and probes were precipitated with one-tenth the volume of 4 M LiCl and three volumes of cold absolute ethanol at -70°C for 30 min. The probes were pelleted at 12000×g for 20 min (4°C) and washed once with 70% ethanol. Pellets were dried *in vacuo* and resuspended in Milli Q H_2O to a concentration of 10 pmol/µl. The DIG-labelled probes were stored at -20°C in silanised tubes.

Genomic DNA was isolated according to the TRI Reagent prescribed protocol (see section 2.2.3). One in ten dilutions (ranging from 2 µg/µl to 100 pg/µl) were made and the DNA was denatured at 95°C for 10 min and snapcooled on ice. 1 µl dot-blots of each dilution was spotted onto positively charged nylon membranes (Boehringer Mannheim). 1.25 and 2.5 pmol DIG-labelled GSP1 and β-actin probes were also spotted as controls of the labelling reaction. The DNA was crosslinked to the membrane for 3 min on an UV transilluminator (312 nm wavelength). 450 pmol GSP1 and 50 pmol β-actin probes were added to separate membranes in hybridisation buffer (5×SSC, 1% blocking agent, 0.1% N-lauroylsarkosine, 0.02% SDS). After incubation for 16 hours at 37°C, the membranes were washed three times with 5×SSC (5 min/wash) and incubated twice for 30 min each

in 5×SSC at 4°C. The membranes were subsequently washed twice at room temperature, (5 min/wash) in tetramethylammonium chloride (TMAC) wash buffer (3 M TMAC, 50 mM Tris-HCl, 2 mM EDTA, 0.1% SDS) to equalise the stabilities of AT and GC base pairs (DiLella and Woo, 1987). Two further incubations for 20 min at 55°C in TMAC were used to increase the stringency. Blocking for 30 min in 1×blocking solution (1% (v/v) blocking reagent in maleic acid buffer) while shaking, was followed by incubation in a 5000-fold dilution of anti-DIG alkaline phosphatase at room temperature for 30 min while shaking and washing twice with wash buffer (3% (v/v) TWEEN 20 in maleic acid buffer). The membranes were equilibrated in detection buffer (100 mM Tris-HCl, pH 9.5; 1 mM NaCl) for 2 min after which 45 µl NBT, 35 µl BCIP and 10 ml detection buffer were added to the membranes. Rinsing of the membrane with H_2O and air drying followed incubation for 8 hours in the dark. Images were captured on a CCD (charged coulped device) camera connected to a computer system.

2.2.9 Polymerase Chain Reaction.

Due to the sensitivity of the method, the PCR conditions had to be optimised empirically for highly specific amplifications. All the amplifications 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. This primer, 3′ Anchor (Genosys Products, Cambridge, England), had the following sequence:

5′ GCT ATC ATT ACC ACA ACA CTC 3′

Primers GSP1 and GSP2 (MWG Biotech, Germany) were tested on cDNA derived from various RNA sources under optimised reaction conditions as summarised in Table 2.1.

Table 2.1: Optimal concentrations of reaction components for GSP1 amplifications (50 µl reaction volume).

Reaction Component	Optimal Concentration
Takara ExTaq Reaction buffer	1×
dNTP	0.2 mM G/C; 0.3 mM A/T
$MgCl_2$	2 mM
3′ Anchor primer	10 pmol
GSP1 primer	40 pmol

The amount of total RNA-cDNA equivalents used in the 3′ RACE with GSP1 ranged from 6.25 ng to 250 ng. The specific amounts used for individual experiments are indicated in the results section.

2.5 U Takara ExTaq (a proprietary mixture of Takara Shuzo, Japan; Taq DNA polymerase-*Thermus aquaticus* and Pfu DNA polymerase-*Pyrococcus furiosis*) was added in a 50 µl reaction during a hot-start protocol at 80°C after a 3 min denaturation period at 94°C (Kidd and Ruano, 1995). A touchdown protocol from 60°C to 46°C was used in an attempt to increase the specificity of the PCR for GSP1 (Kidd and Ruano, 1995). The touchdown cycling profile consisted of 30 cycles at 94°C for 20 sec, 60°C for 15 sec (decreasing in increments of 0.5°C per cycle to 46°C) and 72°C for 2 min. Subsequently, 10 cycles at 94°C for 20 sec, 50°C for 15 sec and 72°C for 2 min were performed.

A normal cycling profile was also employed using a theoretically calculated annealing temperature of the GSP1 primer (Oligo program) and reaction components as indicated in Table 2.1. The normal cycling profile consisted of denaturing at 94°C for 3 min, a 2 min pause at 80°C for the hot-start addition of 1.25 U Takara ExTaq DNA Polymerase followed by 30 cycles at 94°C for 30 sec, 50°C for 30 sec and 72°C for 2 min.

In order to establish whether the PCR products were driven by both primers or by single primers, bands separated by agarose electrophoresis were excised and DNA eluted at 80°C in ~100 µl of 10 mM Tris-HCl (pH 8.5). Approximately 5 µl (2-20 ng) was used as template with either both or single primers for 25 cycles and reaction conditions as for the normal cycling profile described above. Reactions were hot-started with 1.25 U Takara ExTaq DNA Polymerase. The same reamplification conditions were also used to obtain sufficient PCR product for cloning purposes.

PCR was conducted in 0.6 ml thin-walled tubes (Quality Scientific Plastics, California, USA) in a DNA Engine (MJ PTC 200, MJ Research Inc., Massachusetts, USA) or in 0.2 ml thin-walled tubes in a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, California, USA) at ramp speeds of the 9600 system.

2.2.10 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.

2.2.11 Purification of Agarose-electrophoresed DNA Fragments.

2.2.11.1 Crystal Violet Visualisation of DNA Bands during Electrophoresis.

Direct visualisation of PCR products with crystal violet under white light was used for cloning purposes to prevent exposure of the AT-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 (Rand, 1996).

2.2.11.2 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. TAE/agarose gel electrophoresis of the DNA products must be used for this protocol as the borate in the TBE buffer inhibits the dissolution of agarose in NaI. 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 (10-40 µl) of Milli Q H_2O. The average recovery was ~50-100 ng/µl.

2.2.12 Cloning Protocols.

2.2.12.1 Preparation of Competent Cells.

Competent SURE (Stratagene, La Jolla, CA, USA) and DH5α *E. coli* (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×g 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.

2.2.12.2 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.

2.2.12.3 High-Pure Plasmid Isolation Kit (Boehringer Mannheim, 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.

2.2.13 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 (See Appendix I for vector map). Purified PCR products were tailed with adenosines at the 3′-ends in a reaction containing 1×Taq DNA polymerase buffer, 2.5 mM $MgCl_2$, 0.2 mM dATP and 5 U *Taq* DNA polymerase (Promega) and incubated at 70°C for 30 min. Later studies showed that a sufficient number of PCR products were A-tailed during amplification with ExTaq DNA polymerase, and this optional step was omitted. 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 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^- (see Appendix I for vector map) 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 (20 mg/ml in DMF) and 0.4 mM IPTG (for SURE cells only) was plated on LB-Agar plates containing 100 µg/ml ampicillin, and left at room temperature to allow evaporation 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 2.2.12.2). 500 ng plasmid DNA was digested with 10 U of *EcoR*I (Promega, Wisconsin, USA) in the appropriate 1×buffer

for four hours or overnight at 37°C. The digested plasmids were analysed on a 1% (w/v) agarose (Promega)/TAE gel in 1×TAE at 7.8 V/cm and visualised by staining with EtBr. Positive clones were stored in 15% (v/v) glycerol/LB at –70°C.

2.2.14 Restriction Mapping of Recombinant Clones.

The identities of different clones were established by restriction mapping once sequence data for some of the clones became available. Unique restriction sites were identified with the DNA Strider Program Version 1.0 (C. Marck, Commissariat a l'Enegrie Atomique, France). 500 ng plasmid DNA was digested with 10 U of *EcoR*V in the appropriate 1×buffer for 3 hours at 37°C. Digestion products were analysed by electrophoresis on 1% (w/v) agarose (Promega)/TAE gel at 7.8 V/cm and visualised with EtBr staining.

2.2.15 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, primers complementary to the T7 and SP6 promoters in the pGEM-T Easy vector (Promega, See Appendix I). 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 in a final reaction volume of 5 µl. The cycle-sequencing was performed in a DNA Engine (MJ PTC 200, MJ Research Inc., Massachusetts, USA) or Perkin Elmer GeneAmp PCR system 9700 (duration times in brackets) with 25 cycles of 96°C for 30 sec (10 sec), 50°C for 15 sec (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 User's 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.

2.2.16 Primer Walking Strategy for Nucleotide Sequencing of the 3′-fragment of the ODC cDNA.

The sequence of the 3'-fragment was used for the design of primer ODCR2 (GTA ATA ACA TAA ATA GGA CAT C, Tm ~49°C) to sequence the internal ~ 650 bp upstream from the poly-A tail as shown in Fig. 2.3. The T7 and SP6 directional sequencing was repeated twice for three positive clones (clones 6, 12, 13) and the internal sequence once for all three clones. The sequences were aligned using the CLUSTALW Program to obtain a consensus sequence.

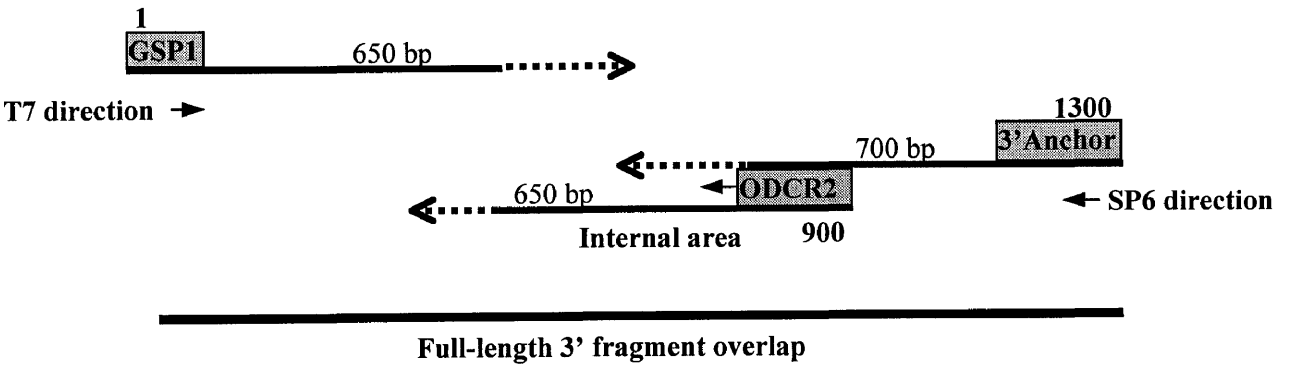

Figure 2.3: Schematic representation of the primer walking strategy used for nucleotide sequencing of the 3′-fragment of the ODC cDNA. The dotted lines indicate areas of overlaps. The relative fragment sizes are indicated in bp and the primer annealing sites in bold.

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 amino acid sequences from 13 different organisms as depicted in Fig. 2.4. The boxed areas indicate the consensus

regions used for primer design. The 13 eukaryotic organisms were chosen from 19 partially or fully characterised sequences available in the SwissProt databank based on the relationship between *P. falciparum* and the higher eukaryotes (Hyde and Holloway, 1993).

Figure 2.4: Multiple-alignment of ODC amino acid sequences from 13 different organisms. Relevant selections of the total sequence (~ 460 amino acids) indicate in **(A)** the region around residues 110 and **(B)** the region around residue 200 (of the human sequence). Red residues indicate amino acids with 100 % homology, blue indicates residues conserved in terms of properties. The boxed residues indicate those used in the primer design, the area in (A) for GSP2 and the one in (B) for GSP1. The organisms are *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*.

Both of the identified conserved areas grouped into BLOCKS for proteins of the group IV decarboxylases i.e. the family of PLP dependent decarboxylases including the ornithine, arginine and diaminopimelic acid decarboxylases according to the LAMA program. The BLOCKS results are graphically represented in Fig. 2.5

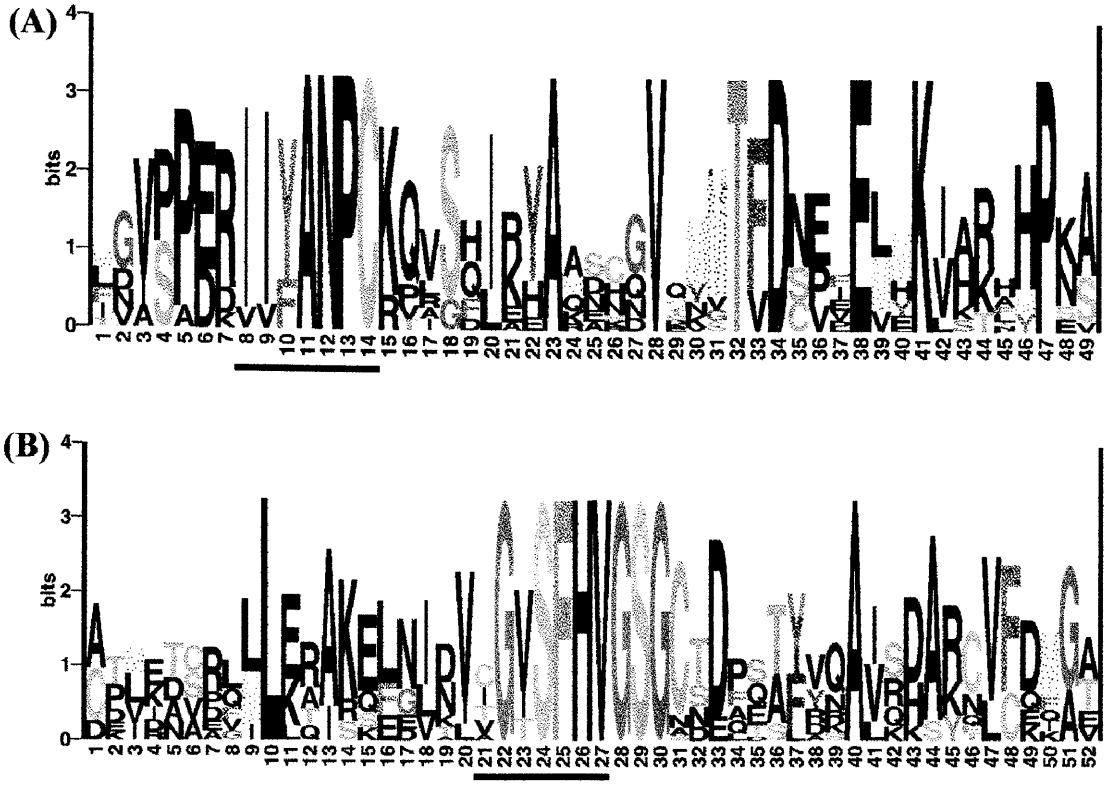

Figure 2.5: Signature motifs of the identified conserved areas in ODC. The conserved areas were identified in the alignment of the amino acid sequences of 13 different ODCs. The X-axis indicates the residue number and the Y-axis the relative abundance of the different amino acids present at a specific position. The horizontal black lines indicate the conserved areas used in the design of the GSP2 (A) and GSP1 primers (B).

The characteristics of the degenerate primers and the other primers used in the 3′-RACE are summarised in Table 2.2. Degenerate nucleotides were included in areas where more than one amino acid is encoded, and an inosine was included in GSP1 to decrease the degeneracy. Primer GSP1 has a real degeneracy of 128 but an effective degeneracy of only 64 due to the degenerate wobble introduced at the 3′-end with a W (A/T). The primers comply with the standards of an optimal primer set by Rychlik i.e. degeneracy of < 128, and a 3′-end ΔG of -6 kcal/mol (Rychlik and Rhoades, 1989). None of the primers had internal secondary structures or self-complementary regions with negative ΔG values, predicting that little primer-dimerisation or hairpin loop formation should occur.

Table 2.2: Summary of the characteristics of the different primers used in 3′-RACE.

Primer	GSP1	GSP2	3′ Anchor	DD-Poly-T [c]
Primer sequence (5′-3′)	TTW GGW GTW WCI TTT CAY GTW	RTT RTW TWT GCW AAT CCW TGT	GCT ATC ATT ACC ACA ACA CTC	GCT ATC ATT ACC ACA ACA CTC $(T)_{18}$VN
Length	21	21	21	41
% A/T	38.1	47.6	57.1	80.5
Degeneracy	64	64	1	12
Tm Oligo [a] (°C) a) or b)	a) 61.1 b) 54	a) 59.1 b) 52	a) 66.9 b) 60	a) 47.8 b) 40
Tm Equation (°C) [b]	max: 52 min: 50	max: 52 min: 48	56	max: 41 min: 37
-ΔG 3′-end (kcal/mol)	-6.6	-6.7	-6.4	Not determined

- [a] Tm Oligo: (Tm values at 1 M NaCl)
  a) % G/C method
  b) ($2^{\circ} \times [A+T] + 4^{\circ} \times [G+C]$)
- [b] Tm Equation: 69.3 + 0.41 (%G/C) - 650/length (Rychlik *et al.*, 1990). Min and max are the calculations for the minimum %GC and maximum %GC respectively.
- [c] 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. The Tm for the 3′ Anchor primer is calculated separately for use in PCR.
- W: A or T; R: A or G and Y: C or T according to the IUBMB degenerate nucleotide nomenclature.

Initial analysis with primer GSP1 as a DIG-labelled probe was performed on genomic DNA to determine if the homologous segment is present in the parasite genome. Faintly positive dots (at 100 and 10 ng genomic DNA) was obtained which suggested recognition of the complementary strand of the genomic DNA (Fig. 2.6). The β-actin probe that was included as negative control for human DNA contamination indicated that human DNA was absent from the *P. falciparum* genomic DNA preparation.

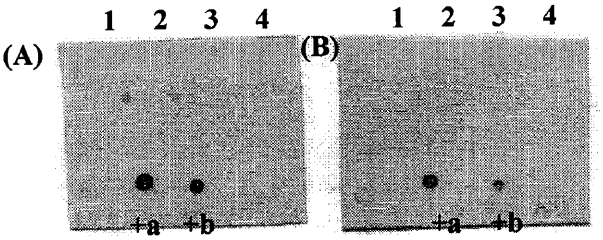

Figure 2.6: Dot-blot of Dig-labelled GSP1 (A) and β-actin probes (B) on *P. falciparum* genomic DNA. 1: 100 ng/µl; 2: 10 ng/µl; 3: 1 ng/µl and 4: 0.1 pg/µl genomic DNA spotted on the membrane and hybridised with the respective probes. +a: 2.5 and +b: 1.25 pmol of the respective labelled probes blotted as positive controls for the labelling reaction in both (A) and (B).

2.3.2 Quality of Isolated Total RNA.

Total RNA was isolated from representative fractions of all four stages of unsynchronised *P. falciparum* cultures. The isolated total RNA appeared intact as indicated by the distinct bands in the denaturing gel electropherogram (Fig. 2.7). The upper band indicates 28 S rRNA, the middle band 18 S rRNA and the lower band 5 S rRNA. The quality of the RNA is characterised by the higher intensity of the upper two bands compared to that of the lower band.

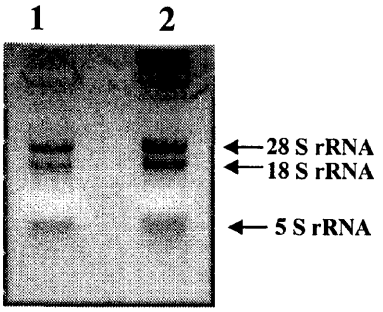

Figure 2.7: Denaturing gel electrophoresis of total RNA. RNA was isolated from *P. falciparum* PfUP1 *in vitro* cultures. Lane 1: 3 µg and lane 2: 5 µg total RNA.

2.3.3 Analyses of cDNA Synthesised from Total RNA.

cDNA synthesised from total RNA could contain contaminating genomic DNA which may give rise to bands after PCR that can be confused with bands originating from cDNA. The total RNA preparation was therefore treated with RNAse-free DNAse I to remove genomic DNA. The quantity of the genomic DNA before and after treatment of the total RNA was monitored by PCR of the highly abundant 18S small-subunit rRNA gene with the gene-specific primers Bio1:Dig2 (Fig. 2.8). The expected band of 408 bp was observed before but also after DNAse treatment, albeit at a low intensity. This indicated that DNAse I treatment did not remove all of the genomic DNA.

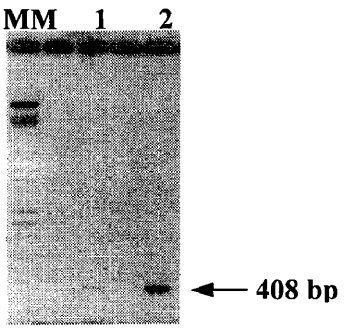

Figure 2.8: Bio1:Dig2 PCR of the 18S rRNA gene to monitor the presence of genomic DNA in *P. falciparum* total RNA. Products of 408 bp are observed (arrow) after (lane 1) and before (lane 2) removal of the genomic DNA. **MM**: *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The quality of cDNA synthesised from the processed total RNA was analysed in a control 3′-RACE using the highly abundant 18S small-subunit rRNA gene as template with the Bio1 and 3′ Anchor primers (Fig. 2.9 A). Although the full-length product of ~ 1800 bp was observed, other bands of smaller sizes suggested possible truncated versions of the gene. This was most likely as a consequence of the annealing of the DD-Poly-T primer to internal poly-A stretches in genes of the AT-rich *P. falciparum* genome. A hot-start procedure was subsequently introduced during reverse transcription of total RNA in order to promote transcription from poly-A stretches of at least 18 nucleotides, resulting in improved cDNA synthesis. This entailed the equilibration of the primer-annealed RNA for 2 min at 45°C before addition of the reverse transcriptase.

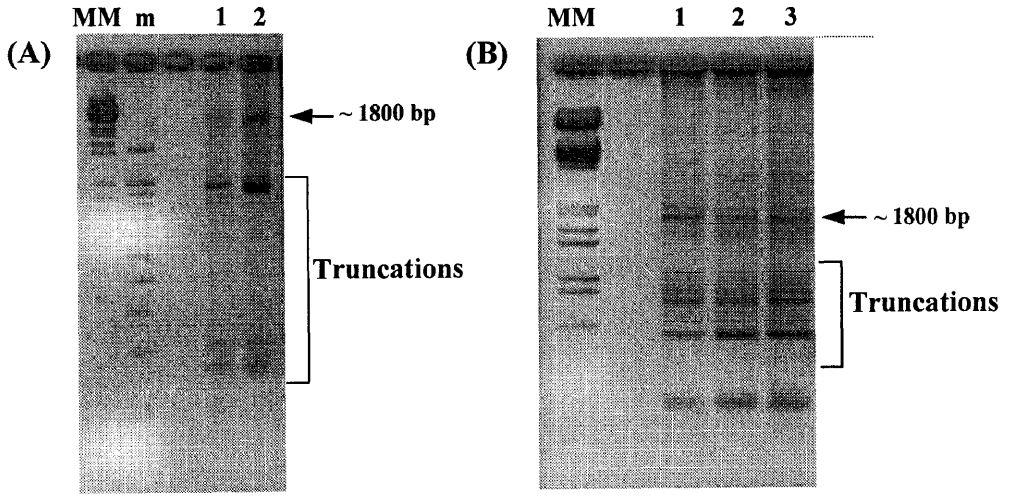

Figure 2.9: Comparison of cDNA synthesised with different ratios of DD-Poly-T primer to total RNA. The Bio1:3′ Anchor primer pair was used on the 18S rRNA gene as template. **(A)** 250 ng processed total RNA-cDNA equivalents synthesised with a 0.5:1 ratio (lane 1) and a 5:1 ratio (lane 2). **(B)** 50 ng unprocessed total RNA-cDNA equivalents synthesised with a 5:1 ratio (lane 1), a 10:1 ratio (lane 2) and a 20:1 ratio (lane 3). The full-length 18S rRNA product is indicated (~1800 bp) in each gel. **MM**: *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker. **m:** 100 bp DNA ladder as molecular mass marker.

The potential loss of gene copies during DNAse I treatment and precipitation steps was determined by comparing the results obtained with cDNA synthesised from unprocessed and processed total RNA with the Bio1:3′ Anchor primer pair. From Fig. 2.9B, it appears that the ~1800 bp bands produced with 50 ng unprocessed total RNA-cDNA equivalents had similar or higher intensities than those bands produced by 250 ng processed total RNA-cDNA equivalents (determined with spectrophotometry) from Fig. 2.9A. These results suggest that template copies were lost during the processing of the total RNA.

There was however an increase in the number of bands (possibly truncated versions) when unprocessed total RNA was used.

The ratio of DD-Poly-T primer to µg total RNA to be used during reverse transcription was also investigated. cDNA synthesis from processed total RNA was performed with 0.5 pmol DD-Poly-T to 1 µg total RNA and with a 5:1 ratio and the quality was evaluated by PCR with the Bio1:3′ Anchor primer pair of the 18S rRNA gene (Fig. 2.9 A). The effect of the increased ratio of the primer to total RNA is evident as the products in lane 2 (5:1) have a higher intensity compared to those in lane 1 (0.5:1), when the same concentration of 250 ng processed total RNA-cDNA equivalents were used (concentration determined spectrophotometrically). cDNA was also synthesised from unprocessed total RNA with ratios of 5, 10 and 20 pmol DD-Poly-T primer to 1 µg total RNA and analysed with the Bio1:3′ Anchor control 18S rRNA 3′-RACE (Fig. 2.9 B). It appears that in this case, a plateau phase was reached since higher ratios of the DD-Poly-T primer (10:1 and 20:1) did not further improve the yield of cDNA as evident by the similar band intensities (Fig. 2.9 B).

The results on the highly abundant 18S rRNA gene indicated that the correct ratio of DD-Poly-T primer to total RNA is imperative for obtaining the maximum yield of the specific cDNA species. Unprocessed total RNA also appeared to produce higher yields of cDNA compared to processed total RNA. Consequently, unprocessed total RNA and a 5:1 DD-Poly-T to RNA ratio for reverse transcription was used in subsequent studies. These principles were applied to the 3′ –RACE with GSP1 as indicated below.

2.3.4 3′-RACE on Total RNA with GSP1.

It should be noted at this stage that various experiments were conducted with GSP2 but only non-specifically amplified products were obtained (see discussion). These included different *P. falciparum* genes that have already been characterised by others (results not shown). The following sections therefore describe only the results obtained with GSP1.

Experiments with GSP1 were conducted on total RNA-cDNA equivalents (with and without DNAse I treatment) to establish if any loss of template copies for GSP1 had

occurred during processing of RNA with DNAse I (Fig. 2.10). A 3′-RACE with a touchdown protocol (65°C to 46°C over 30 cycles, and 10 additional cycles at 50°C) on cDNA derived from 250 ng processed total RNA-cDNA equivalents did not yield detectable amplification products. This is in contrast with the products observed at ~1500 bp, 800-900 bp and 400 bp (with a background smear throughout) with 250 ng cDNA derived from unprocessed total RNA. These results confirmed that some template copies were lost during DNAse treatment. In subsequent experiments only unprocessed total RNA was used. Since 3′-RACE is performed, i.e. amplification with an anchor primer attached via the poly-A tail, cDNA or genomic DNA without the complementary sequence for the anchor primer would not be expected to be amplified.

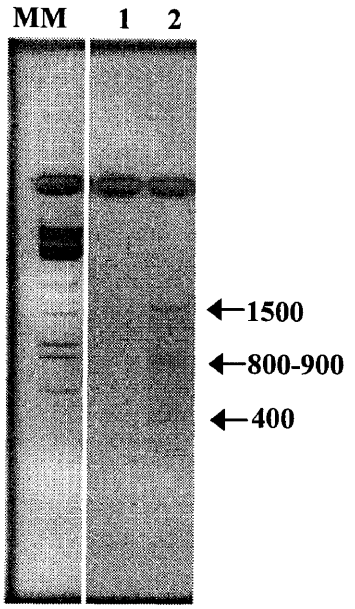

Figure 2.10: 3′-RACE products with GSP1 primer used in a touchdown protocol. The touchdown consisted of 65°C to 46°C over 30 cycles, and 10 additional cycles at 50°C. Lane 1: 250 ng cDNA derived from processed total RNA (5:1 ratio) and lane 2: 250 ng cDNA from unprocessed total RNA (5:1 ratio). Arrows indicate bands at 1500 bp, 800-900 bp and 400 bp. **MM**: *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The 3′-RACE was repeated with GSP1 on unprocessed total RNA at a fixed annealing temperature of 50°C for comparison to the results in Fig. 2.10, lane 2. This temperature was based on the predicted Tm given in Table 2.2. cDNA obtained with DD-Poly-T primer to total RNA ratios of 5, 10 and 20:1 were analysed to establish an optimal ratio to use for reverse transcription of the GSP1 template.

Various concentrations (6.25; 12.5 and 25 ng) of the processed total RNA-cDNA equivalents were used as starting material. An intense band of ~ 1300 bp was obtained

(Fig. 2.11), which is 500 bp more than the expected size of about 800 bp. Comparing the intensity of the 1300 bp band at 10:1 and 20:1 to the 5:1 ratio suggests that a plateau is already reached at the lower ratio. A faint smear can be observed around 800-900 bp in the 5:1 ratio samples (lanes 1-3). At higher ratios of the DD-Poly-T primer in the reverse transcription reaction, the intensity of the 1300 bp band decreases whereas the intensities of the smaller bands are more pronounced (Fig. 2.11, lanes 4-7).

The banding pattern of Fig. 2.11 is similar to that obtained in Fig. 2.10, although the sizes of the biggest bands are 1300 bp and 1500 bp, respectively. This may be an electrophoresis artefact since the sizes of the two other bands are identical (i.e. 800-900 bp smear and 400 bp band).

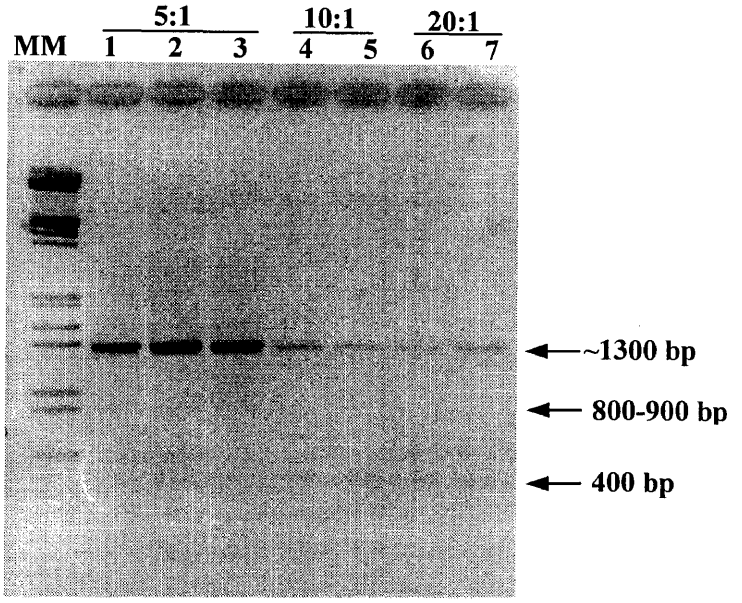

Figure 2.11: 3′-RACE with GSP1 with a fixed annealing temperature (50°C) on unprocessed total RNA. 5:1, 10:1 and 20:1 ratios of DD-Poly-T primer to total RNA during reverse transcriptase was monitored. Total RNA-cDNA equivalents of 6.25 ng (lanes 1,4 and 6), 12.5 ng (lanes 2, 5 and 7) and 25 ng (lane 3) were used for each sample indicated. The bands are indicated. **MM**: *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The yields of products at the fixed annealing temperature, with 6.25 ng of unprocessed total RNA-cDNA equivalents (5:1 ratio was used, Fig. 2.11, lanes 1-3), are higher when compared to the products obtained with 250 ng of unprocessed total RNA-cDNA equivalents used in a touchdown protocol (Fig. 2.10). At suboptimal ratios of primer to cDNA (i.e. constant concentration of primer but higher cDNA concentrations, usually ~10^7 molar excess primer is optimal), less amplification of the corresponding template

would be expected. However, this possibility was not further investigated with reduced cDNA concentration in the touchdown PCR.

The larger bands obtained with GSP1 (1300 bp band and 800-900 bp smear, Fig. 2.11) were used as templates for subsequent reamplification reactions with primer pairs or single primers to establish the origin of the bands. The results indicated that both the 1300 bp band and 800-900 bp smear were derived by amplification with the GSP1:3′ Anchor primer pair and not by single primer amplifications (Fig. 2.12).

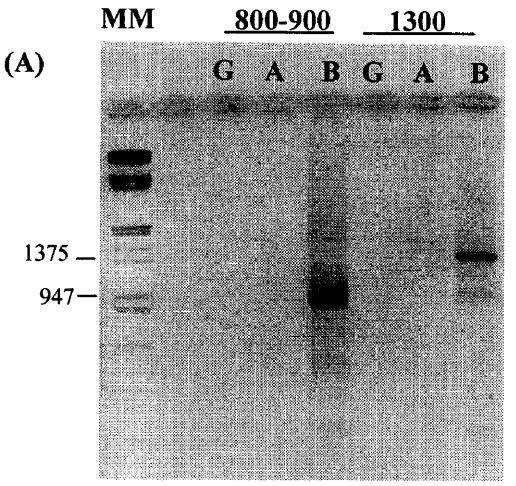

Figure 2.12: Reamplification of 3′-RACE products. PCRs were conducted with the GSP1:3′ Anchor primer pair or either primer separately. The excised bands from Fig. 2.11 were used as templates. G: GSP1 primer alone, A: 3′ Anchor primer alone and B: both present in the reaction. **MM**: *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The reamplified 1300 bp band and 800-900 bp smear were purified with the silica-based method and subsequently cloned into the pGEM-T Easy vector and propagated in DH5α *E. coli* cells. The clones of the 800-900 bp smear contained three different sized bands with nucleotide sequences corresponding to already characterised *P. falciparum* genes (e.g. *P. falciparum* RhopH3 gene and *P. falciparum* carbamoyl phosphate synthetase gene). The nucleotide sequences of three clones of the 1300 bp band showed high homology (93% identity) with the Genbank-deposited *P. falciparum* ODC cDNA sequence (Genbank accession number AF012551), from the T7 side. The SP6 side, surprisingly, showed a high homology (85% identity) with *E. coli* transposon Tn10. However, since the band could be amplified with both primers (Fig. 2.12), these results

suggested that gene-segment rearrangement occurred during propagation of the cloned fragments in DH5α *E. coli*.

To substantiate this postulate, the 1300 bp band was recloned into the pGEM-T Easy vector but propagated in SURE *E. coli,* as these cells are unable to support homologous recombination. Clones were screened for the presence of recombinants by digestion with *EcoR*I but, unexpectantly, slightly different sized inserts of the cloned 1300 bp PCR band were observed (Fig. 2.13). This indicates that PCR with degenerate primers does not mean that a single band visualised in electrophoresis necessarily should consist of a single product.

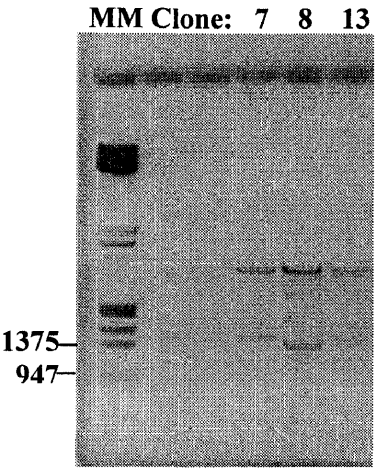

Figure 2.13: *EcoR*I restriction digestion analyses of the 1300 bp/pGEM-T Easy clones. Products were propagated in SURE *E. coli*. The three different clone numbers are indicated. **MM**: *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The identity of the cloned inserts was determined by restriction mapping rather than nucleotide sequencing (Fig. 2.14). The restriction map was based on the sequence of the T7 side of the *P. falciparum* ODC nucleotide sequence of the 1300 bp band, propagated in DH5α cells. An unique *EcoR*V site was identified in this sequence and was used to screen for clones containing the *P. falciparum* ODC sequence. Using this approach, only three of the ten clones were digested with *EcoR*V. The 1300 bp band thus contained more than one species of DNA, confirming the results presented in Fig. 2.13.

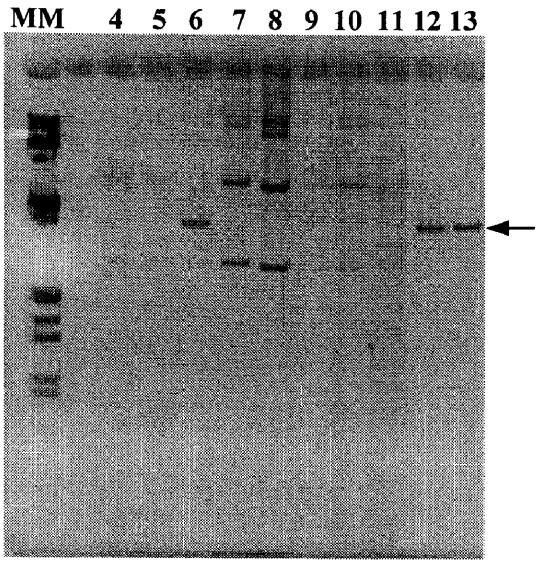

Figure 2.14: *EcoR*V restriction mapping of cloned 1300 bp GSP1 3′-RACE products. Products were recloned and propagated in SURE *E. coli*. The clone numbers are indicated. Clones 6,12,13 were linearised (arrow) with *EcoR*V digestion while the rest of the clones were undigested. **MM**: *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The nucleotide sequences of the three digested clones (clones 6, 12 and 13) as well as three undigested clones (clones 7, 8 and 10, Fig. 2.14) were determined from both the T7 and SP6 sides. Clones 7, 8 and 10 were identified as non-specific amplified *P. falciparum* genes already characterised (*P. falciparum* carbamoyl phosphate synthetase II gene and 5.8S rRNA and LSU rRNA genes). The nucleotide sequences of clones 6, 12 and 13 indicated high homology (92% identity) with the *P. falciparum* ODC cDNA (Genbank accession number AF012551) from both the T7 and SP6 sides. Gene-segment rearrangement was apparently prevented by recloning the 1300 bp band into SURE *E. coli* rather than DH5α cells.

The nucleotide sequence of the 1300 bp fragment was completed by a primer walking strategy designed to obtain the sequence of the internal ~650 bp stretch (Fig. 2.3). Three different plasmids were sequenced twice from the T7 and SP6 sides and the internal area once with the ODCR2 primer. The sequences were subsequently overlap-aligned and a consensus sequence derived.

Alignment of the consensus nucleotide sequence of the complete 1300 bp 3′-RACE product with the *P. falciparum* ODC sequence deposited in Genbank, is shown in Fig. 2.15. The GSP1 primer site was identified, although it differed in 6 places with the

Genbank sequence. The consensus nucleotide sequence is almost 100% homologues to the Genbank sequence although point mutations are evident (Fig. 2.15). Fourteen of the 17 point mutations included 9 transversions (i.e. pyrimidine to purine or vice versa) and 5 transitions (substitutions within a base type). Three A or T nucleotide insertions are evident around 576 bp, in a stretch of AAT-repeats but do not cause a frameshift (see chapter 4 for deduced amino acid sequence). The reliability of the *P. falciparum* ODC sequence in Genbank is uncertain since the information has not yet been published. However, our sequence is the result of multiple sequencing experiments on three different clones and the differences observed are therefore unlikely to be due to PCR, cloning and/or sequencing errors.

The stop codon (TAA) and poly-adenylation signals (AATAAA) could be identified. The sequence downstream of the AATAAA has a characteristic GT-rich nature (TAGGG) that further specifies the poly-A site (Adams *et al.*, 1993). Interestingly, the poly-A signal is only 5 nucleotides removed from the start of the poly-A tail, whereas this site is typically found 15-25 residues upstream of the poly-A tail in other mRNAs. 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. The whole 3′-fragment size is ~ 1321 bp (1339 bp including the 18-mer poly-A tail of the DD-Poly-T primer). The sequence of the *P. falciparum* ODC cDNA is further examined in Chapter 4 where it is compared to sequences from other organisms.

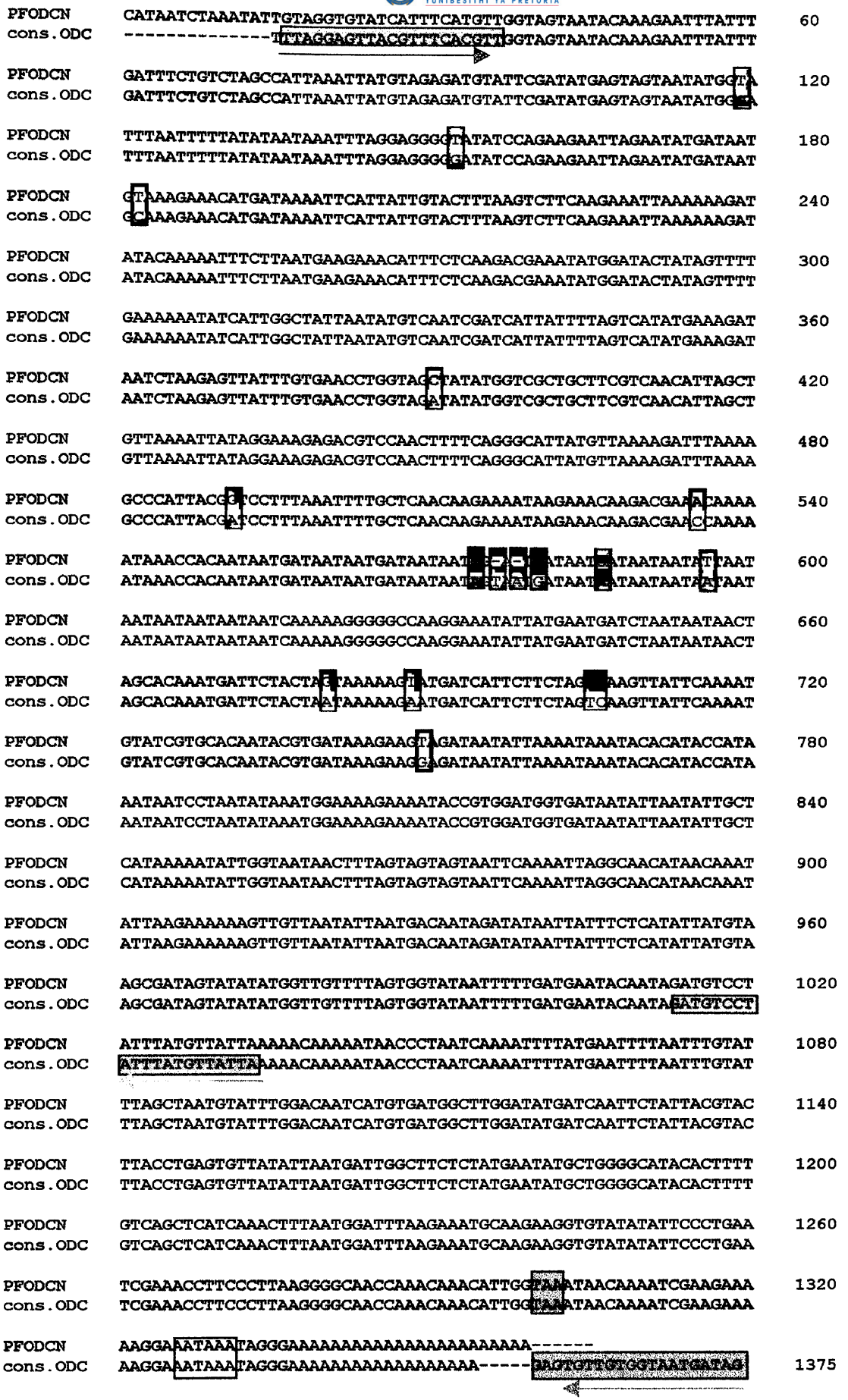

Figure 2.15: Alignment of *P. falciparum* ODC nucleotide sequences for the 1300 bp 3'-fragment. The PFODCN sequence was taken from Genbank (accession number AF012551). The sequence denoted cons.ODC is the sequence derived from the 1300 bp GSP1 3′-RACE product. The residues boxed in red indicate those that differ between the two sequences. Blue=GSP1 primer; purple=3′ Anchor primer; green=ODCR2 primer; yellow=poly-adenylation signal; orange=stop codon. Arrows indicate the directions of the primers.

2.4 DISCUSSION.

The construction and design of the ODC-specific primers used in the 3′-RACE were based on sequence data for the ODC protein characterised in related organisms. *P. falciparum* seems to be more closely related to the higher rather than the lower eukaryotes (Hyde *et al.*, 1989). The parasite genome is further distinguished by an AT content of ~81%, which is the most AT-rich eukaryotic genome yet characterised. Codon preferences more closely resemble those of the vertebrates than the invertebrates and lower organisms (Hyde *et al.*, 1989). Based on these facts, the number of organisms chosen in the multiple alignments for the identification of the conserved areas, were limited to only 13.

The primer sequences were based on two identified conserved areas (Fig. 2.4). 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. The primers further complied with the characteristics specified by Rychlik (Table 2.2) (Rychlik, 1993).

3′-RACE was initially performed with cDNA prepared from isolated mRNA with the GSP1 and GSP2 primers. In the case of GSP1, no amplification products were observed even after extensive manipulation of the RACE reaction conditions. For GSP2, a 350 bp product was obtained but nucleotide sequence determination indicated that this band was homologous to a *Dictyostelium discoidium* α-mannosidase gene (results not shown).

Other studies in our laboratory indicated that products of the expected sizes only emerged when ~300 ng of mRNA-cDNA equivalents were used (K. Clark, personal communication). This was in contrast to the 1 ng recommended by Frohman (Frohman, 1993). It was therefore speculated that some gene copies may have been be lost during mRNA isolation possibly due to saturation of the oligo(dT) beads by annealing to stretches of internal 'A' in either RNA or genomic DNA. This probably also happens in other organisms but is a more profound problem in malaria because of the high AT-content of its genome. This led to an investigation of the representativeness of synthesised cDNA. At least a 10-fold increase was observed in the yield of cDNA

prepared from total RNA compared to cDNA synthesised from a similar amount of mRNA (K. Clark, personal communication). Only cDNA synthesised directly from total RNA was therefore used in subsequent RACE protocols.

The isolated RNA was of the expected quality, which is essential for the subsequent RACE procedures (Fig. 2.7). Since total RNA was used, traces of contaminating genomic DNA had to be removed with RNAse-free DNAse I. However, this treatment was only partially successful (Fig. 2.8). Bauer indicated that DNAse I treatment in the presence of Mg^{2+} could result in overlapping double-stranded DNA fragments with the potential to recombine during PCR and still form a product. This can be prevented by using Mn^{2+} as cofactor but was not further investigated in this study (Bauer *et al.*, 1997).

The quality of the cDNA synthesised from DNAse I treated and untreated total RNA was monitored with 3′-RACE of the highly abundant 18 S rRNA gene as template (mRNA for this gene constitutes about 80% of the total number of transcripts (Lodish *et al.*, 1995)). The DNAse I treatment protocol resulted in the loss of gene copies, as 250 ng processed total RNA-cDNA equivalents was needed to obtain the predicted band of ~1800 bp (Fig. 2.9 A) compared to only 50 ng unprocessed total RNA-cDNA equivalents (Fig. 2.9 B). Successful PCR is dependent on at least 10^4 copies of the template to be amplified. Since RACE results are produced with anchor and gene-specific primers, traces of contaminating genomic DNA in cDNA prepared from unprocessed total RNA, is not expected to be problematic. It was subsequently decided to use only unprocessed total RNA for cDNA synthesis.

Possible truncations of the 18 S rRNA transcript are most likely due to the annealing of the DD-Poly-T primer to poly-A stretches internally in the transcript during the reverse transcription reaction (Fig. 2.9, K. Clark, personal communication). A hot-start procedure was subsequently introduced during reverse transcription to prevent annealing of the DD-Poly-T primer to nucleotide stretches of less than 18 A's. A higher number of smaller bands were observed in the unprocessed total RNA sample (Fig. 2.9 B) compared to the equivalent processed RNA sample. This might be due to the higher number of transcript copies present in the unprocessed RNA sample (only 50 ng total RNA-cDNA equivalents

were needed to obtain the ~1800 bp band). Even though more amplification products were observed with unprocessed total RNA, the benefits of retaining transcript copies outweigh the negative implications of increased numbers of truncations. This should be particularly true for medium to low abundance transcripts that could easily be lost during the processing of the total RNA, as was indeed indicated for the ODC transcript (Fig. 2.10). No products were obtained on processed total RNA (lane 1) which was not the case for unprocessed total RNA where three bands were observed (lane 2).

The cDNA synthesis was further optimised by increasing the ratio of DD-Poly-T primer relative to the concentration of total RNA in the reverse transcription reaction. A 5:1 ratio of DD-Poly-T primer to processed total RNA (Fig. 2.9 A, lane 2) resulted in a increase in the products observed compared to a 0.5:1 ratio. This indicates that the highly abundant 18 S rRNA message probably needed more DD-Poly-T primer to optimise reverse transcription of all the available messages. The amount of cDNA transcribed did not increase further at higher ratios (10:1 and 20:1) indicating that a plateau was probably reached (Fig. 2.9 B).

When this rationale was applied to the medium to low abundant ODC transcript on processed total RNA, optimum results were achieved with a 5:1 ratio (Fig. 2.11, lanes 1-3). Higher amounts of the DD-Poly-T primer during reverse transcription resulted in a decrease of the reverse transcribed ODC message but an increase of smaller products (lanes 4-7). In this instance, the higher amounts of the primer seemed to result in the annealing of the DD-Poly-T primer to internal poly-A stretches of either the ODC message or other unknown messages. This indicates that there is a specific relationship between the DD-Poly-T primer and the relative abundance of the message that should be optimised for each target. At too low concentrations of the DD-Poly-T primer, suboptimal reverse transcription occurs, as the relative concentration of the message is high. The primer concentration should then be increased until an optimum ratio of primer to template is reached that must not be breached, since this will result in excess primer annealing to internal sites. This principle is much the same as is required for the PCR itself, where an optimal molar ratio of ~10^7 primer:template is required.

To obtain specific results with PCR-based techniques requires optimisation of various reaction components. A touchdown protocol was initially used for GSP1 with unprocessed total RNA because of the uncertainty of the predicted annealing temperatures for degenerate primers (Kidd and Ruano, 1995). However, this resulted in only faint products that were not easily discriminated from a background smear (Fig. 2.10, lane 2). Comparing this to the distinct band obtained when a fixed annealing temperature was used (Fig. 2.11) indicates that there was not sufficient amplification of the template during the touchdown protocol. The fixed annealing temperature also allowed detection of products with as little as 6.25 ng unprocessed total RNA-cDNA equivalents synthesised with a 5:1 ratio (Fig. 2.11) compared to the faint products with 250 ng unprocessed total RNA-cDNA equivalents (5:1 ratio) after the touchdown protocol (Fig. 2.10, lane 2). The 6.25 ng total RNA-cDNA equivalents amounts to ~0.2 ng mRNA-cDNA equivalents (based on the assumption that mRNA constitutes 3% of the total RNA), which correlates with the 1 ng mRNA recommended for RACE by Frohman (Frohman, 1993). The 250 ng unprocessed total RNA-cDNA equivalents therefore contained ~7.5 ng mRNA-cDNA equivalents (Fig. 2.10, lane 2). This excess template could also have led to decreased amplification of the specific products due to a too low primer:template ratio.

The distinct band of 1300 bp obtained with GSP1 and unprocessed total RNA (5:1 ratio, Fig. 2.11) was ~500 bp larger than the expected size predicted by the multiple alignments. According to the amino acid sequence data available for other organisms, the ODC protein contains ~450 residues, which accounts for a cDNA of ~1350 bp. Therefore, the 3′-fragment found here is almost as large as the complete cDNA in other eukaryotes. An experiment to determine the nature of the band indicated that it was indeed due to exponential amplification with the GSP1:3′ Anchor primer pair (Fig. 2.12).

The sequence of the cloned 1300 bp band yielded an ODC sequence from the T7 but unexpectedly, a transposon Tn10 sequence from the SP6 sequencing sides of the vector. Presumably, gene-segment rearrangements took place during proliferation of clones in DH5α *E. coli* cells. These cells contain only the recA mutation, and although deficient in the major recombination gene involved in general recombination and radiation repair,

they are still able to perform homologous recombination. Transposon Tn10 is a prokaryotic transposon in the composite class I that contains elements for encoding the transposase enzyme needed for the transposition reaction to occur, flanked by inverted repeats (Adams *et al.*, 1993). The fact that the 1300 bp clones contained transposon Tn10 sequences might indicate either that a transposition reaction occurred (although the full transposon was not present) or that homologous recombination between the recombinant plasmid and the host genome (coincidentally with a transposon element) took place. The possibility cannot be excluded that the ODC nucleotide sequence might contain some localised sequence homology with the Tn10 that could have facilitated the homologous recombination. The rearrangement was prevented by using SURE cells that additionally contain the recB and recJ mutations, impairing the recB (DNA exonuclease V) and recJ proteins (involved in DNA repair). This reduces the recA-independent rearrangements and the cells are therefore deficient in homologous recombination (Adams *et al.*, 1993). Propagation in these cells resulted in clones of the 1300 bp fragment with the nucleotide sequence of the full-length 3′-fragment of *P. falciparum* ODC.

Further modifications introduced in the cloning strategy was to propagate the SURE *E. coli* at 30°C, a temperature that limits rearrangement and improves the competency of the cells (Hanahan *et al.*, 1991). Other studies in our laboratory have shown that high copy number plasmids (like the pBluescript-based vectors) could still facilitate gene rearrangements, even in a recombinantly negative host like SURE *E. coli* (K. Clark, personal communication). However, rearrangements could be prevented or reduced by limiting propagation at 30°C only to the exponential phase. Propagation to the saturation phase apparently favours gene shuffling due to the limited availability of nutrients.

The fact that a distinct band on an agarose gel could contain multiple products (not resolved by size in agarose electrophoresis) was demonstrated when the 1300 bp GSP1 band was recloned into SURE cells. Three types of clones with different *P. falciparum* nucleotide sequences (including ODC) were identified. Restriction digestion was used to discriminate between the different clones, which indicated that only 3 of the 10 clones contained the ODC 1300 bp fragment (Fig. 2.14). The rest contained *P. falciparum*

cDNA fragments already characterised, again indicating the possibility of mispriming with degenerate primers.

The 3′-fragment of the *P. falciparum* ODC gene is ~1339 bp in length (Fig. 2.15). The nucleotide sequence did not show significant homologies to any ODC genes from any other organism except *P. falciparum* (Genbank accession number AF012551). This is not surprising since the malaria genome is AT-rich whereas those of the other organisms tend to be GC-rich. The larger than expected size of the 3′-fragment might indicate the presence of insertions in the *P. falciparum* cDNA. The nucleotide sequence has a near 100% homology to the sequence of the *P. falciparum* ODC sequence in Genbank. There were however 17 sites where differences occurred between the two sequences. Our sequencing results were based on duplicate determinations from both sides (T7 and SP6) for all three positive clones and once for the internal ~650 bp sequence of each clone with the ODCR2 primer. The consistency of the point mutations in all the clones does not support PCR, cycle-sequencing or cloning artefacts. Until details of the Genbank sequence have been published, the significance of these differences cannot be assessed. Fourteen of the differences were single nucleotide substitutions, including 9 transversions and 5 transition mutations. Most of the nucleotide insertions were located internally (around 567 bp, Fig. 2.15) in an area that contains an AAT repeat only found in the *P. falciparum* ODC nucleotide sequence. The three nucleotide insertions should not affect the reading frame of the translated protein sequence as this will only encode an extra amino acid. This is further indicated in chapter 4 where the translated amino acid sequence of the full-length protein is analysed.

The GSP1 site was identified in the 1300 bp sequence (Fig. 2.15). It does not fully correspond to the Genbank sequence as six nucleotide differences are apparent. None of these differences occurred within the last three nucleotides at the 3′-end so that the necessary specificity was achieved in this part of the primer even though the 5′-end of the primer did misprime at these positions. In chapter 3 the sequence of the 5′-RACE fragment overlapped with the GSP1 site, showing that there are no differences between our sequence and the Genbank sequence (see later, Fig. 3.10).

The conserved poly-adenylation signal, AATAAA, was identified 24 nucleotides downstream of the TAA stop codon. The site was further characterised by the GT-rich nucleotides, TAGGG, prior to the start of the poly-A tail. The hexanucleotide AAUAAA is conserved in poly-A RNA and located typically 15-25 residues upstream of the poly-A site (Adams *et al.*, 1993). This site was, however, located 5 residues upstream of the poly-A tail in our sequence. The DD-Poly-T primer therefore annealed in this instance at the poly-A tail and not internally in the cDNA, which indicates that the ratio of DD-Poly-T primer to total RNA used in the 3′-RACE, was optimal for the ODC template (Fig. 2.15).

Several PCR products were obtained with the GSP2 primer but consisted of already characterised *P. falciparum* genes (results not shown). There could be several reasons for these misprimings, including low stringency amplification conditions or too low copy number of the specific template cDNA. In the case of GSP2, the mispriming was shown to be due to the absence of the 3′-end of the priming site (See chapter 4, Fig. 4.2). The consensus amino acid area used for the synthesis of GSP2 was (I/V)(I/V)(Y/F)AN**PC** but the *P. falciparum* ODC sequence has the residues IIFAN**TI**. This indicates the risk involved when a single consensus sequence of other species is used for the design of RACE protocols.

The nucleotide sequence obtained for the 3′-fragment of the *P. falciparum* ODC cDNA enables the design of ODC gene-specific primers without any degeneracy. This strategy is employed in the following chapter where the primers are used to obtain the upstream 5′-fragment of the cDNA with 5′-RACE.

CHAPTER 3

The application of 5′-RACE protocols in obtaining the full-length *P. falciparum* ODC nucleotide sequence.

3.1 INTRODUCTION.

Obtaining a full-length cDNA is one of the most important and often the most challenging task in characterising a gene. This is particularly true with respect to obtaining the 5′-ends of cDNA. In the RACE protocols of Frohman, cDNA is synthesised from RNA using either a DD-Poly-T primer or a downstream gene-specific primer. The 3′-end of the first-strand cDNA is then modified by the addition of a homopolymer tail (e.g. poly-A) using terminal deoxynucleotidyltransferase (TdT). The second-strand is synthesised using the DD-Poly-T primer again, attached to an anchor sequence in a protocol similar to 3′-RACE (Frohman, 1993).

Various modifications and improvements of the original RACE procedure have since been developed. These include anchor-PCR (where the homopolymer tail added is poly-G) or one-sided PCR, ligation-anchored PCR (LA-PCR) (Edwards *et al.*, 1991), RNA ligase-mediated RACE (RLM-RACE, where an anchor sequence is ligated to the RNA prior to cDNA synthesis) (Liu and Gorovsky, 1993) or reverse ligation-mediated PCR (RL-PCR) (Scheafer, 1995).

However, these methods do have some drawbacks. There are usually at least three enzymatic steps (RT, TdT-tailing and PCR) which are subject to failure. A high background of non-specific PCR products often necessitates performing a nested-PCR, with primers annealing internally to the first set, to achieve greater specificity. The homopolymer tailing reaction can be inefficient; the length of the tail is unknown and also introduces more nucleotides to the already unknown length of the 3′-end of the cDNA. Ligation of anchor nucleotides either to the single-stranded cDNA or directly to the RNA can be particularly inefficient and may lead to low amplification efficacy

especially for cDNAs with strong secondary structures at the 3′-end (Chenchik *et al.*, 1996).

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. 3.1.

Second-strand cDNA synthesis normally depends on priming by either RNAse H generated RNA fragments or hairpin loops at the 3′-end of the first-strand cDNA. In the double-stranded adaptor ligation method described here, the established Gubler and Hoffman method for second-strand synthesis was used (Gubler, 1987). The conditions were adjusted to promote synthesis by RNA fragments (created by RNAse H), rather than the hairpin loop. The resulting ds cDNA is then blunt-ended with T4 DNA polymerase to assure effective ligation to the synthetic ds adaptors. As indicated in Fig. 3.1, an adaptor primer homologous to the long strand of the adaptor, is used for either 3′- or 5′-RACE (Chenchik *et al.*, 1996).

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/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 as described above. 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. 3.1) (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. 3.1). 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, was used to obtain the 5′-end of the *P. falciparum* ODC cDNA. The nucleotide sequence of the 3′-fragment of the *P. falciparum* ODC cDNA obtained in Chapter 2 facilitated the design of primer ODCR1. The reverse primer is located in the most 5′-end of the known sequence of the 3′-fragment and downstream of the GSP1 forward primer to create an overlap between the 3′- and 5′-fragments. The full-length *P. falciparum* ODC sequence in Genbank (Accession number AF012551) was used in the design of the rest of the gene-specific primers used in the 5′-RACE.

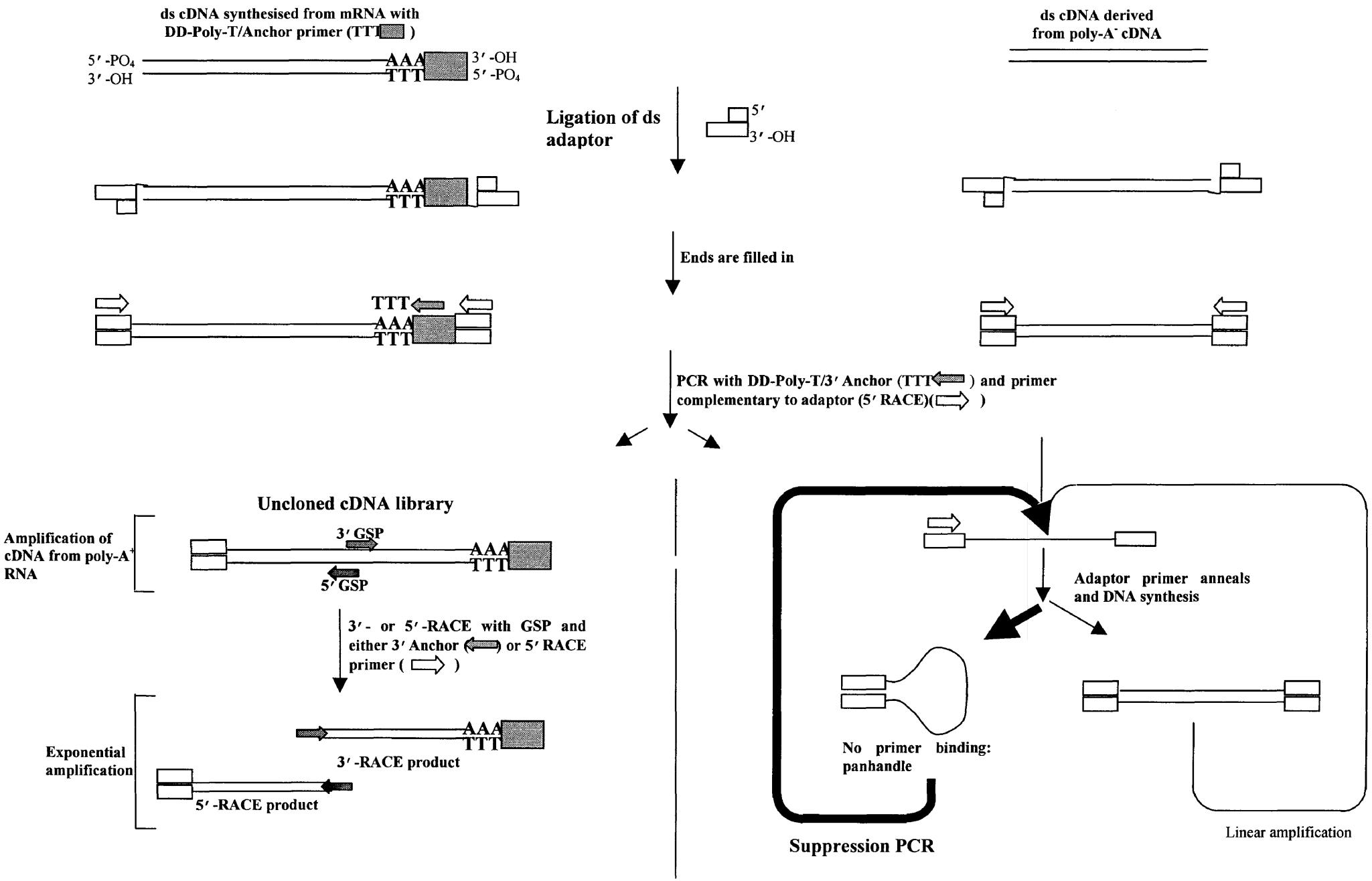

Figure 3.1: RACE protocols with double strand adaptor-ligated cDNA and suppression PCR. The adaptor is indicated with the open boxes. The DD-Poly-T and attached Anchor region are indicated with the solid boxes. The ligation to only the long strand of the adaptor is indicated with the small loop. The suppression PCR effect is shown on the right hand side of the schema and the formation of the uncloned cDNA library to the left.

3.2 MATERIALS AND METHODS.

3.2.1 Synthesis of an Uncloned *P. falciparum* cDNA Library with an In-house Designed Kit.

The ds adaptor was designed in-house with a short strand complementary to the 3′-end of the longer one and was not phosphorylated so that only the longer strand is ligated by its 3′-OH end to the 5′-PO_4 end of the ds cDNA (Fig. 3.1) (Lukyanov *et al.*, 1997). The characteristics (i.e. Tm and ΔG profile) of the adaptor were determined with the Oligo Version 4.0 Program (Rychlik and Rhoades, 1989). The sequences are as follows (MWG Biotech, Germany):

5′ TTA CAG GAC CAC ATC AAC TAT CGG GAG CGG CGG GGG CAG CA-OH 3′
3′ HO-CCC GTC GT 5′

The adaptor Tm is 78.4°C (according to the equation, see table 3.1) with a ΔG of -89.3 kcal/mol.

The protocol described by Lukyanov was followed as is, except that the adaptor was designed for use with a previously designed 5′RACE primer (see section 3.2.1.4) (Lukyanov *et al.*, 1997).

3.2.1.1 First-strand cDNA Synthesis.

1µg of *P. falciparum* total RNA was annealed to the DD-Poly-T primer (10 pmol) as described before (section 2.2.5 in Chapter 2). The cDNA synthesis was performed in a 10 µl reaction volume containing 1×first-strand buffer (50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$), 3 mM $MgCl_2$, 2 mM DTT and dNTPs (1 mM GC/1.5 mM AT). The reaction was equilibrated at 42°C for 2 min (hot-start) after which 200 U MMLV Superscript II RT (Gibco BRL, Gaithersburg, USA) was added followed by incubation at 42°C for 1 hr.

3.2.1.2 Second-strand cDNA Synthesis (Sambrook *et al.*, 1989).

The 10 µl first-strand cDNA solution (100 ng/µl total RNA-cDNA equivalents) was used to which was added 1×second-strand buffer (100 mM KCl, 10 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 20 mM Tris-HCl pH 7.5, 10 mM DTT, 0.05 mg/ml BSA, 0.15 mM β-NAD), 0.2

mM GC/0.3 mM AT, 1 U *E. coli* RNAse H (Gibco BRL), 4.8 U *E. coli* DNA ligase (Gibco BRL), 24 U *E. coli* DNA polymerase I (Promega, Wisconsin, USA) in a final volume of 80 µl. The reaction was incubated at 16°C for 90 min. Blunt-ends were created by adding 10 U T4 DNA polymerase (Boehringer Mannheim, Mannheim, Germany) and incubation for 45 min at 16°C. The reaction was stopped with 0.8 M EDTA/8 mg/ml glycogen. Proteins were extracted once with 100 µl phenol:chloroform:isoamyl alchohol (25:24:1) and once with 100 µl chloroform:isoamyl alchohol (24:1). The ds cDNA was precipitated with half the volume of 4 M ammonium acetate and 2.5 volumes of 95% ethanol at room temperature. After centrifugation at 13000×g for 20 min (room temperature), the pellet was washed with 70% ethanol, dried *in vacuo* and dissolved in 10 µl Milli Q H_2O and stored at –20°C. The concentration of the ds cDNA, assuming 100% recovery of the cDNA after precipitation, is 100 ng/µl total RNA-cDNA equivalents.

3.2.1.3 Ligation of the Double-stranded Adaptor to ds cDNA.

The long and short single-stranded oligonucleotides were annealed in equimolar amounts in Milli Q H_2O by heating at 60°C for 2 min, followed by cooling to room temperature over a 30 min period.

5 µl of the ds cDNA solution (i.e. 500 ng total RNA-cDNA equivalents) was ligated to 20 pmol of the adaptor in 1×DNA ligase buffer (66 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP) with 1 U T4 DNA ligase (Boehringer Mannheim, Mannheim, Germany) and 5% (v/v) PEG in a 10 µl reaction for 16 hours at 22°C. The enzyme was heat-inactivated at 70°C for 5 min. The adaptor-ligated ds cDNA concentration is theoretically 50 ng/µl total RNA-cDNA equivalents.

3.2.1.4 Amplification of the Adaptor-ligated ds cDNA.

The adaptor-ligated ds cDNA was amplified with a primer complementary to the long strand of the adaptor (5′RACE primer: 5′ TTA CAG GAC CAC ATC AAC TAT CGG G 3′, see table 3.1) and the DD-Poly-T primer (Chapter 2, section 2.2.5). A 50 µl reaction contained 1 µl of a five-fold dilution of the adaptor-ligated ds cDNA (~10 ng), 1× Takara ExTaq reaction buffer, 2 mM $MgCl_2$, 0.2 mM GC/0.3 mM AT and 10 pmol each of 5′RACE (MWG Biotech, Germany) and DD-Poly-T primers. 2.5 U Takara ExTaq

(Takara Shuzo, Japan) was added at 80°C in a hot-start protocol. PCR was conducted for 17 cycles (131 000-fold amplification) at 94°C for 30 sec and 68°C for 8 min in a Perkin Elmer GeneAmp system 9700 (PE Applied Biosystems, California, USA) (Lukyanov *et al.*, 1997). The amplified cDNA library was analysed on a 1.5% (w/v) agarose (Promega)/TAE gel at 78 V (5.2 V/cm) and visualised with EtBr staining. The amplified uncloned cDNA library represents only the mRNA fraction of the total RNA. If mRNA is taken as constituting 3% of the total RNA, the concentration of the uncloned cDNA library is 786 ng/µl mRNA-cDNA equivalents.

3.2.2 5′-RACE of the *P. falciparum* ODC cDNA.

3.2.2.1 3′-RACE Evaluation of the Adaptor-ligated, Uncloned ds cDNA.

3′-RACE was performed on the in-house amplified uncloned cDNA library using the GSP1:3′ Anchor primer pair to get an indication of the success of cDNA synthesis and adaptor ligation reactions. The 50 µl reactions contained 1×reaction buffer, 2 mM $MgCl_2$, 0.2 mM GC/0.3 mM AT, 40 pmol GSP1 and 10 pmol 3′ Anchor, 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 and 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.

3.2.2.2 Primer Design for 5′-RACE.

Several ODC-specific primers were designed for the 5′-RACE. The 3′-end sequence of the 1300 bp *P. falciparum* ODC (Chapter 2, section 2.3.3) was used for the design of primer ODCR1. The sequence of *P. falciparum* ODC deposited in Genbank (Accession number AF012551) was used to design the upstream primer ODCF1 and the nested primers ODCR3 and ODCF2 which was also used for sequencing purposes. The respective positions of the primers are indicated in Fig. 3.2. The primers were obtained from Gibco BRL, Gaithersburg, USA.

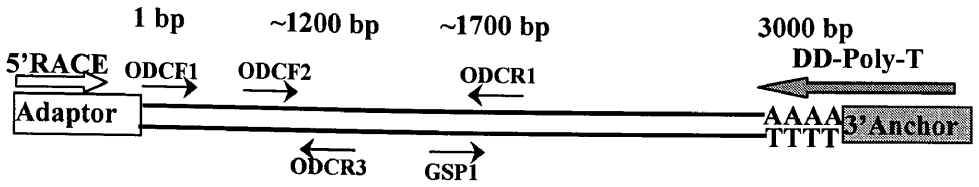

Figure 3.2: Respective positions of the adaptor and ODC-specific primers used in 5′-RACE protocols. Arrows indicate the direction of the primers. The predicted sizes of the RACE products in bp are given at the top of the figure.

3.2.2.3 5′-RACE on the Amplified Uncloned cDNA Library.

The in-house amplified uncloned cDNA library was used as template for the 5′-RACE with the ODCR1:5′ RACE primer pair. 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 GC/ 0.3 mM AT. 2.5 U Takara ExTaq (Takara Shuzo, Japan) was added in the 25 µl reaction 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 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. Fig. 3.3 summarises the different PCRs used to obtain and identify the ODC-specific band.

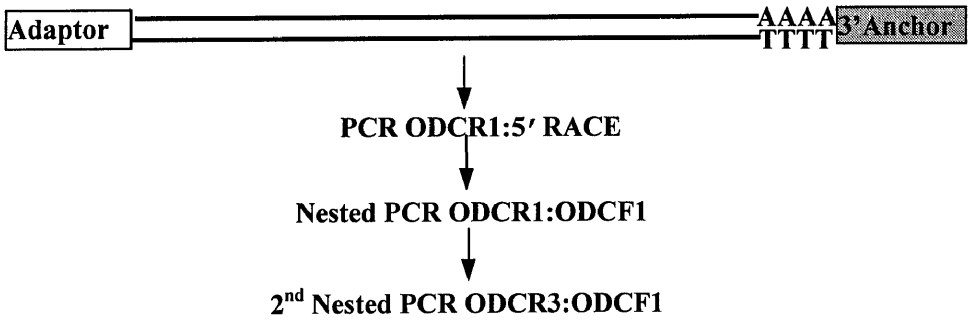

Figure 3.3: Summary of the nested PCR approach to obtain and identify the 5′-RACE ODC-specific band.

All the PCR's were performed on a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, USA). Products were analysed on a 1.5% (w/v) agarose (Promega)/TAE gel, which was electrophoresed at 78V (5.2 V/cm) and visualised by EtBr staining.

3.2.3 Direct PCR Product Nucleotide Sequence Determination.

The nucleotide sequence of the 1700 bp 5′-RACE fragment of the in-house amplified uncloned cDNA library was determined with direct automated fluorescent cycle sequencing (Section 2.2.15, PE Applied Biosystems, California, USA). The RACE product was purified from agarose with a silica-based method (Chapter 2, section 2.2.11.2) and 62.5 ng was used in a reaction containing 3.2 pmol of the respective primers (see Fig. 3.4) 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 in section 2.2.15 (Chapter 2). The ODC-specific primers ODCF1 and ODCR1 were used to obtain the sequences of the 5′- and 3′-end fragments. The internal ~700 bp sequence was obtained with a primer walking strategy employing the internal primers ODCR3 and ODCF1 (Fig. 3.4). The 1700 bp fragment was sequenced twice with each of all four primers. The alignments were done with the CLUSTALW Program (Thompson *et al.*, 1994) and the consensus sequences deduced with the Genetic Data Environment program (Smith *et al.*, 1994).

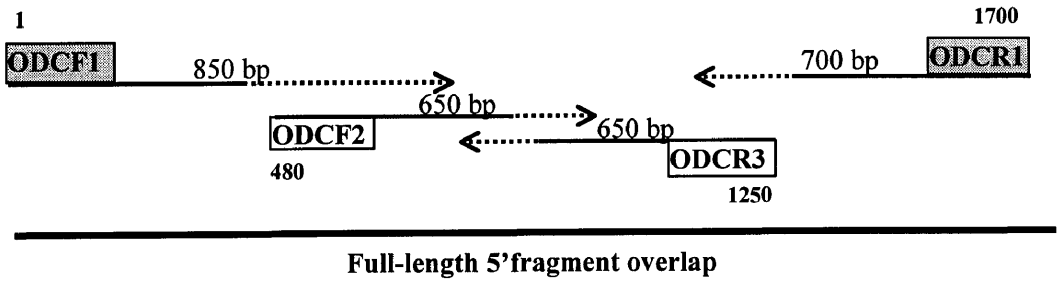

Figure 3.4: Schematic representation of the primer walking strategy for nucleotide sequencing of the 5′-fragment of the ODC cDNA. The shaded boxes indicate the external primers and the open boxes the nested primers. The bold numbers indicate the annealing positions of the sequencing primers. The fragment sizes are indicated in bp.

3.3 RESULTS.

Primers were designed for 5′-RACE based on the *P. falciparum* ODC nucleotide sequence deposited in Genbank (Accession number AF012551, primers ODCF1, ODCF2 and OCDR3) and the 3′-fragment of the cDNA obtained in Chapter 2 (primer ODCR1). The relative positions of the primers are indicated in Fig. 3.2. The characteristics and sequences of the primers are summarised in table 3.1. The primers were all non-degenerate. ODCF2 was used only for sequencing purposes.

Table 3.1: Summary of the characteristics of the primers used in 5′-RACE. R in the primer names denotes reverse and F, forward.

Primer	Primer sequence (5′-3′)	Length	T_m (°C) [a]	3′-end ΔG (kcal/mol)
ODCR1	**GCT** ACT CAT ATC GAA TAC ATC TCT AC	26	60	-5.7
ODCR3	GAA TTT ATA CAA ACT ACT GAT G	22	51	-6.9
ODCF1	GAA TTT TTA TAA TGG AAA GTA TAT G	25	52	-5.9
ODCF2	GTT GAT GAT ATG TAT GAG TAT G	22	53	-5.7
5′ RACE	TTA CAG GAC CAC ATC AAC TAT CGG G	25	63	-11

- [a] T_m equation=69.3+0.41(%GC)-650/length (Rychlik *et al.*, 1990)
- The nucleotides in bold were included for cloning strategies.

Ten ng of adaptor ligated ds cDNA (total RNA-cDNA equivalents) was amplified 131 000-fold to obtain the uncloned cDNA library. This library consists of only mRNA, which normally constitutes 1-5% of total RNA. The theoretical mRNA concentration in the amplified library is calculated as 3% of the total RNA, i.e. 10 ng total RNA-cDNA equivalents contained 0.3 ng mRNA. This was amplified 131 000-fold to a theoretical yield of 786 ng/µl mRNA-cDNA equivalents. The amplified cDNA (3 µg mRNA-cDNA equivalents) is indicated by a diffuse smear ranging from < 500 bp to over 5000 bp, which is in the expected size range (Fig. 3.5) (Lukyanov *et al.*, 1997).

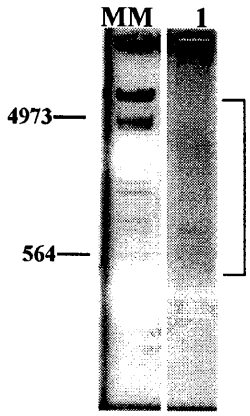

Figure 3.5: Amplified double-stranded adaptor-ligated cDNA. cDNA was derived from mRNA and amplified with the DD-Poly-T:5′RACE primer pair. Lane 1 indicates the characteristic smear obtained. **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

3′-RACE with GSP1:3′Anchor was used to monitor the various steps during the construction of the uncloned cDNA library. The expected 1300 bp band was observed for both the first-strand (50 ng total RNA-cDNA equivalents) and second-strand cDNA (100 ng total RNA-cDNA equivalents) although it was overamplified due to the use of excess template (the efficiencies of the various steps and recovery of precipitated products were uncertain), compared to earlier results where the band was distinct with as little as 6.25 ng unprocessed total RNA-cDNA equivalents (Chapter 2, Fig. 2.11). Comparison of the intensities of the 1300 bp band in lanes 1 and 2 (Fig. 3.6) indicates a lower than estimated recovery (<50%) of the ds cDNA since 50 ng first-strand cDNA (total RNA-cDNA equivalents) yielded a more intense band than 100 ng ds cDNA (total RNA-cDNA equivalents). No distinct band was observed after the adaptor-ligation reaction as was expected since only panhandle structures will be formed in the absence of the template that must be generated with the DD-Poly-T primer (lane 3). After amplification of the uncloned cDNA library, the 1300 bp band was sharper but high molecular mass bands were observed when 7.86 ng mRNA-cDNA equivalents were used.

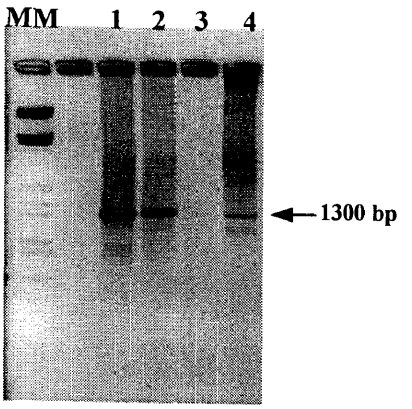

Figure 3.6: Control 3′-RACE PCR with the GSP1:3′ Anchor primer pair on various cDNA samples. 50 ng total RNA-cDNA equivalents of first-strand cDNA (lane 1), 100 ng total RNA-cDNA equivalents of double-strand cDNA (lane 2), 100 ng total RNA-cDNA equivalents of adaptor-ligated ds cDNA (lane 3) and 7.86 ng of the 17×amplified cDNA (lane 4) of the in-house amplified uncloned cDNA library. **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

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 when 15.72 ng mRNA-cDNA equivalents were used (lane 2) but not when 7.86 ng mRNA-cDNA equivalents were used (lane 1), as indicated in Fig. 3.7. 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.

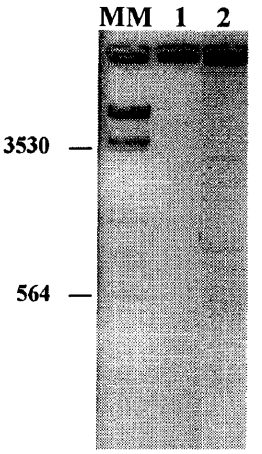

Figure 3.7: 5′-RACE on the amplified, uncloned cDNA library. Several bands were observed with the ODCR1:5′RACE primer pair only with 15.72 ng mRNA-cDNA equivalents (lane 2) but not with 7.86 ng (lane 1). **MM:** *EcoR*I-*Hind*III digested λ-phage DNA used as high molecular mass marker.

A nested PCR was performed with the ODCR1:ODCF1 primer pair on the 5′-RACE products from Fig. 3.7. The expected 1700 bp band was amplified (Fig. 3.8 A) 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. 3.8 B).

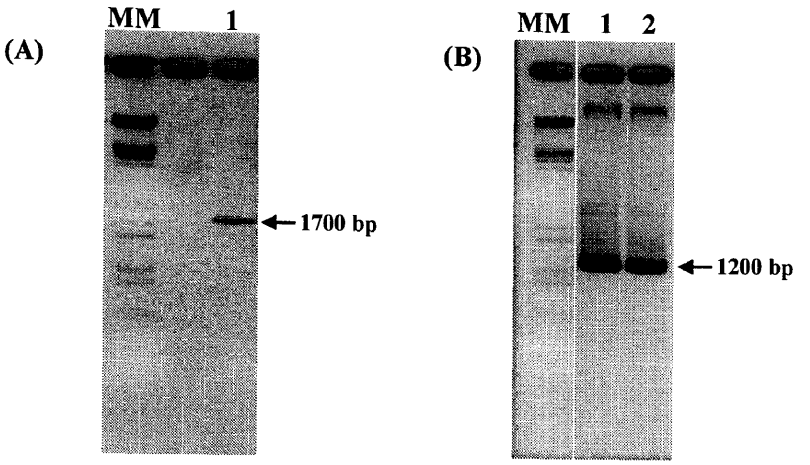

Figure 3.8: Nested PCR strategy employed to confirm specific 5′-RACE amplifications on the uncloned cDNA library. (A) The nested PCR was done with the ODCR1:ODCF1 primer pair on the products from Fig. 3.7. **(B)** The second nested PCR with ODCR3:ODCF1 on the products obtained in (A) with 50-fold (lane 1) and 100-fold dilutions (lane 2). **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The 1700 bp band from Fig. 3.8A was purified with the silica-based method (Fig. 3.9) and used directly for determination of its nucleotide sequence.

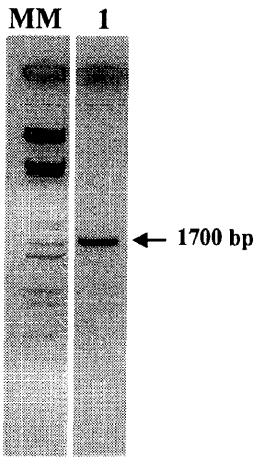

Figure 3.9: Silica purified 1700 bp full-length 5′-RACE product of *P. falciparum* ODC. The band in Fig. 3.8A was purified for direct nucleotide sequencing. **MM:** *Eco*RI-*Hind*III digested λ-phage DNA used as high molecular mass marker.

The nucleotide sequence of the full-length 1700 bp fragment 5′-RACE (Fig. 3.9) of the in-house amplified uncloned cDNA library was determined with the primer walking strategy and showed high homology with the Genbank *P. falciparum* ODC cDNA

sequence (Fig. 3.10). Virtually no difference was observed between our sequence obtained and the sequence deposited in Genbank. Due to the direct PCR product sequencing protocol used, 22 bp after the initiation codon and 16 bp after the annealing site for ODCR1 could not be identified (purple dashes in Fig. 3.10). The fragment is ~1678 bp in length including the ODCF1 and ODCR1 priming sites (25 and 26 bp respectively). An overlap of ~96 bp (including the GSP1 and ODCR1 primer sites) with the 3′-fragment sequence of Chapter 2 (Fig. 2.15) allowed the identification of the GSP1 primer site.

```
PFODC   GTGATCGATT TTTAGACTTT ATTCACAAGC AGTTGAATTT TTATAATGGA
5cons   ---------- ---------- ---------- ---------- ----------      50

PFODC   AAGTATATGT TCATGATAAA TTATGTATTC TGTGAGGAGA GTAACAACAT
5cons   ---------- ---------- ---------- -GTGAGGAGA GTAACAACAT     100

PFODC   GTCTAAAATG GTACCTGATG ATGATAATAA TAATTATAGT AGTGGTAAAA
5cons   GTCTAAAATG GTACCTGATG ATGATAATAA TAATTATAGT AGTGGTAAAA     150

PFODC   GTTGCGTTTA TTATCAAGAT TTAAATAAGA AAGAAAAAGA AGAATATTAT
5cons   GTTGCGTTTA TTATCAAGAT TTAAATAAGA AAGAAAAAGA AGAATATTAT     200

PFODC   CGCTTGAACA AAAAATTAAG AAACGACTTA TTTATTAATT CGAAACAATT
5cons   CGCTTGAACA AAAAATTAAG AAACGACTTA TTTATTAATT CGAAACAATT     250

PFODC   TTATGAATTA CATACATTTA CCGAACGAAC GGTTGGATTT ATGAGAGTGC
5cons   TTATGAATTA CATACATTTA CCGAACGAAC GGTTGGATTT ATGAGAGTGC     300

PFODC   AATATTTTGT TTATAAATTA AGAGATGTTG TTAAATGTGT AGAAAAAGAA
5cons   AATATTTTGT TTATAAATTA AGAGATGTTG TTAAATGTGT AGAAAAAGAA     350

PFODC   ACTTTGCTAG CTAGGAGTTC GTCTTGTTTA TTTATGTTTA ATAATATCAA
5cons   ACTTTGCTAG CTAGGAGTTC GTCTTGTTTA TTTATGTTTA ATAATATCAA     400

PFODC   ACGAAATGAC GTACATGATG ATTATGTAAC TAAGTCGTCA AATGGTGGTG
5cons   ACGAAATGAC GTACATGATG ATTATGTAAC TAAGTCGTCA AATGGTGGTG     450

PFODC   TAATAAAACA ATTAACGGAA AGAGATGTTG ATGATATGTA TGAGTATGCT
5cons   TAATAAAACA ATTAACGGAA AGAGATGTTG ATGATATGTA TGAGTATGCT     500

PFODC   TTAAATTTTT GTAAACAAAA TAAAATAGTT GTTGTAGATA CTAATACCTT
5cons   TTAAATTTTT GTAAACAAAA TAAAATAGTT GTTGTAGATA CTAATACCTT     550

PFODC   TTTTTTTGAT GCATCTAAAA GAAAGGAGAA CTTAATAAAA CTTGAAAAGG
5cons   TTTTTTTGAT GCATCTAAAA GAAAGGAGAA CTTAATAAAA CTTGAAAAGG     600

PFODC   TACAAACAAA TGAGAAAGAT GAATATGAAG AAAAAGATGA AGTGTATCGA
5cons   TACAAACAAA TGAGAAAGAT GAATATGAAG AAAAAGATGA AGTGTATCGA     650

PFODC   AGGGGTAATA ATGAATTGAG TTCGTTGGAT CATTTAGATA GTAAGAATAA
5cons   AGGGGTAATA ATGAATTGAG TTCGTTGGAT CATTTAGATA GTAAGAATAA     700

PFODC   TTTGATTCAT ATGTATTATG AAAAGAACAA ATGTGATATC ATAAATAAGG
5cons   TTTGATTCAT ATGTATTATG AAAAGAACAA ATGTGATATC ATAAATAAGG     750

PFODC   ATGATGAGAA TTCAACGATA GCGACGAATA ATAATGATAA TAATAATGAT
5cons   ATGATGAGAA TTCAACGATA GCGACGAATA ATAATGATAA TAATAATGAT     800

PFODC   AGTAGTTCTT ATGACAAAAG TATAACGATC AGCAGAAGCA GTAGCTGTAA
5cons   AGTAGTTCTT ATGACAAAAG TATAACGATC AGCAGAAGCA GTAGCTGTAA     850

PFODC   TAATAGCCAT TTGAGTTATA GTAGTTTTGA TAATAATCAT GGAAATGAAA
5cons   TAATAGCCAT TTGAGTTATA GTAGTTTTGA TAATAATCAT GGAAATGAAA     900

PFODC   AAATGAAAGA TTATATAAGT GTTGATGAAA ATAATAATAA TAATAATAAT
5cons   AAATGAAAGA TTATATAAGT GTTGATGAAA ATAATAATAA TAATAATAAT     950
```

```
PFODC   AATAAAAATA ATAATGTATT GTTAACTTTA CAAAGGAACA GTGATGATGA
5cons   AATAAAAATA ATAATGTATT GTTAACTTTA CAAAGGAACA GTGATGATGA   1000

PFODC   AAATGGTAAA GATAAAGATA ATGAAAAAAA TGACGTAAGT TTAGAAAACA
5cons   AAATGGTAAA GATAAAGATA ATGAAAAAAA TGACGTAAGT TTAGAAAACA   1050

PFODC   ATATGGAAAA GAATTATAAA GAAGAAATAT GGAATTATTA TACAAAAAAT
5cons   ATATGGAAAA GAATTATAAA GAAGAAATAT GGAATTATTA TACAAAAAAT   1100

PFODC   AAAGTGGAAG TAAAAACATT AGAAAAAGTA TTAAATGAAA ATATAGATAC
5cons   AAAGTGGAAG TAAAAACATT AGAAAAAGTA TTAAATGAAA ATATAGATAC   1150

PFODC   ATCAGTAGTT TGTATAAATT TACAGAAAAT ATTAGCTCAG TATGTTAGAT
5cons   ATCAGTAGTT TGTATAAATT TACAGAAAAT ATTAGCTCAG TATGTTAGAT   1200

PFODC   TTAAAAAGAA TCTTCCACAT GTTACTCCAT TCTATTCTGT AAAAAGTAAT
5cons   TTAAAAAGAA TCTTCCACAT GTTACTCCAT TCTATTCTGT AAAAAGTAAT   1250

PFODC   AATGATGAAG TTGTAATCAA ATTTTTATAT GGATTGAATT GTAATTTTGA
5cons   AATGATGAAG TTGTAATCAA ATTTTTATAT GGATTGAATT GTAATTTTGA   1300

PFODC   TTGCGCTTCG ATAGGTGAAA TAAGTAAAGT AATAAAATTA TTACCAAATT
5cons   TTGCGCTTCG ATAGGTGAAA TAAGTAAAGT AATAAAATTA TTACCAAATT   1350

PFODC   TATCAAGAGA TAGAATAATT TTTGCGAATA CAATTAAAAG TATTAATTCT
5cons   TATCAAGAGA TAGAATAATT TTTGCGAATA CAATTAAAAG TATTAATTCT   1400

PFODC   TTAATATATG CAAGAAAGGA AAATATTAAT TTATGTACTT TTGATAATTT
5cons   TTAATATATG CAAGAAAGGA AAATATTAAT TTATGTACTT TTGATAATTT   1450

PFODC   AGATGAATTA AAAAAAATAT ATAAATATCA TCCGAAATGT TCTTTAATAT
5cons   AGATGAATTA AAAAAAATAT ATAAATATCA TCCGAAATGT TCTTTAATAT   1500

PFODC   TACGTATTAA TGTAGATTTT AAAAATTACA AATCTTATAT GTCTTCAAAA
5cons   TACGTATTAA TGTAGATTTT AAAAATTACA AATCTTATAT GTCTTCAAAA   1550

PFODC   TATGGAGCTA ATGAATATGA ATGGGAAGAA ATGTTATTGT ATGCAAAAAA
5cons   TATGGAGCTA ATGAATATGA ATGGGAAGAA ATGTTATTGT ATGCAAAAAA   1600

PFODC   ACATAATCTA AATATTGTAG GTGTATCATT TCATGTTGG TAGTAATACA
5cons   ACATAATCTA AATATTGTAG GTGTATCATT TCATGTTGG TAGTAATACA   1650

PFODC   AAGAATTTAT TTGATTTCTG TCTAGCCATT AAATTATGTA GAGATGTATT
5cons   AAGAATTTAT TTGATTTCTG T--------- -------              1700

PFODC   CGATATGAGT AGTAATATGG
```

Figure 3.10: Nucleotide sequence of the 1700 bp 5′-fragment of the *P. falciparum* ODC cDNA. The 'pfodc' sequence indicates the Genbank sequence (accession number AF012551) and the '5cons' sequence indicates the 1700 bp sequence obtained with 5′-RACE. The area in pink indicates the position of the ODCF1 primer, the green area ODCF2, the yellow area ODCR3, the blue ODCR1 and the red area the position of GSP1 used in 3′-RACE. The arrows indicate the directions of the primers. The square indicates the initiation ATG codon. The purple dashes indicate the missing nucleotides due to the direct PCR sequencing protocol.

3.4 DISCUSSION.

5′-RACE was based on the principle of ligating an adaptor sequence to ds cDNA, to be used in creating an in-house uncloned cDNA library. The characteristics of the designed adaptor were such that suppression PCR technology could be applied. The high Tm, GC content and negative ΔG value of the adaptor ensured intramolecular annealing of cDNAs (derived from RNA other than mRNA) with only the full-length adaptor sequence at both ends due to the formation of a GC-clamp that is not easily denatured. This is an elegant method for enriching the mRNA-derived cDNA species. The method was also designed to amplify copies of cDNAs from low abundance mRNAs and to compensate for losses of template copies during the double-strand cDNA preparation protocol (see also chapter 2).

The uncloned cDNA library was successfully amplified 131 000-fold (estimated on the assumption of 2^n amplifications, where n=number of cycles of PCR) as indicated by the 500-5000 bp smear (Fig. 3.5). The spread of the smear is optimal as it was shown that starting with less total RNA usually decreases the average smear size and requires more cycles of PCR. This is undesirable due to preferential amplification of one sequence over another, thereby again disrupting the relative abundance of the transcripts (Lukyanov *et al.*, 1997).

The adaptor-ligated ds cDNA was synthesised successfully as monitored with the 3′-RACE. The 1300 bp band could be shown in both the first strand and second strand cDNA preparations (Fig. 3.6). In Chapter 2 (Fig. 2.11, lane 4), the 1300 bp band was observed with as little as 6.25 ng total RNA-cDNA equivalents (synthesised with a 10:1 ratio). This is ~10-times less template compared to the 50 ng total RNA-cDNA equivalents used in the 3′-RACE of the uncloned cDNA after first strand synthesis (Fig. 3.6, lane 1) and explains the slight overamplification observed. A theoretical concentration of 100 ng total RNA-cDNA equivalents (based on the assumption of 100% yield of ds cDNA after precipitation) after synthesis of ds cDNA (Fig. 3.6, lane 2) did not result in a higher intensity of the 1300 bp band as compared to the results with 50 ng total RNA-cDNA equivalents after first strand synthesis (lane 1). The intensity of the band

instead decreased, which might indicate that there was > 50% loss of template copies during the precipitation steps. This corresponds to the results in chapter 2 where template copies were also lost during treatment of RNA with DNAse I and the subsequent precipitation step. If ligation efficiency of the adaptor to the ds cDNA is also estimated at ~50%, the theoretical concentration of the amplified uncloned cDNA library used in 3′-RACE comes to 1.97 ng mRNA-cDNA equivalents (Fig. 3.6, lane 4). This indicates retrospectively that an estimated 10-times more template was used in Fig. 3.6, explaining the slight overamplification compared to the results in Fig. 2.11, lane 4.

The initial 5′-RACE on the amplified uncloned cDNA library did not result in the production of distinct bands with the ODCR1 primer, even when 15.72 ng mRNA-cDNA equivalents (corresponding to 3.93 ng mRNA-cDNA equivalents based on the above assumptions) of the library were used (Fig. 3.7). This is in contrast with the 3′-RACE (Fig. 3.6, lane 4) where the expected product was observed with 7.86 ng mRNA-cDNA equivalents. The ligation of the adaptor to full-length as well as incomplete cDNAs (shortened at the 3′-end of first-strand cDNA) effectively lowers the template concentration of similar length cDNAs. It is known that ODC mRNAs from other organisms contain a long GC-rich 5′-UTR with thermodymically stable secondary structures (Coffino, 1989). The efficiency of the reverse transcriptase in transcribing the UTR could also have been decreased by the secondary structures. This could have further decreased the efficiency of ligation to full-length cDNA and consequently, the relative abundance of the transcript.

The full-length ODC-specific fragment was obtained with the ODCF1:ODCR1 nested PCR and had the predicted size of ~1700 bp (Fig. 3.8 A). The authentic nature of the ODC-specific fragment was confirmed with PCR with the nested ODCR3:ODCF1 primer pair on the 1700 bp product (Fig. 3.8 B).

The 5′-RACE extensions were conducted at 68°C instead of the normal 72°C PCR extensions in order to favour the panhandle formation for the suppression PCR effect. This has an added benefit of amplifying sequences that are normally refractory to PCR. These include AT-rich areas that tend to dissociate at higher temperatures. Extension

temperatures of 60°C have been used successfully on the AT-rich *P. falciparum* genome without negatively affecting the performance of the DNA polymerase (Su *et al.*, 1996).

The nucleotide sequence of the 1700 bp fragment was obtained by direct sequencing of the PCR product, thereby eliminating problem-prone cloning strategies. The direct PCR sequencing did not have any background noise in the electropherogram (results not shown) which indicates that the silica-purified 1700 bp band (Fig. 3.9) was indeed extremely pure.

The consensus nucleotide sequence after overlap-alignment of the various segments sequenced in the primer walking strategy (Fig. 3.4) had 100% homology with the *P. falciparum* sequence deposited in Genbank (Fig. 3.10). Due to the direct PCR product sequencing, 22 bp after the initiation codon (annealing site for ODCF1 primer) and 16 bp upstream of the ODCR1 primer annealing site could not be determined. Therefore, the sequence of the upstream bases should be confirmed either with another sequencing primer (in the upstream direction) with direct PCR product sequencing or by cloning of the PCR product. Fortunately, the 16 bp missing at the 3′-end was already obtained from the overlapping sequence of the 3′-fragment of the cDNA (Chapter 2). The 5′-fragment of the cDNA spanned ~1678 bp including the 38 missing nucleotides and the sites for the two primers, ODCF1 (25 bp) and ODCR1 (26 bp). An overlap of ~96 bp (including the GSP1 and ODCR1 primer sites) with the 3′-fragment sequence from Chapter 2 (Fig. 2.15) will facilitate the assembly of the full-length nucleotide sequence as is further described in chapter 4. The GSP1 site was elucidated and indicated that there are no differences at this site between our sequence and the Genbank sequence (Fig. 2.15).

By obtaining the nucleotide sequence of the *P. falciparum* ODC cDNA, the amplified uncloned cDNA library was validated. Its use must, however, still be vindicated for other transcripts with a lower abundance than the medium abundance mRNA of ODC.

The following chapter will describe the assembly of the full-length nucleotide sequence of the *P. falciparum* ODC cDNA obtained with RACE procedures. The amino acid sequence of the protein is deduced and compared to sequences from other organisms. Molecular characterisation of the deduced amino acid sequence will also be described.

CHAPTER 4

Molecular characterisation of the deduced amino acid sequence of the full-length *P. falciparum* ODC cDNA.

4.1 INTRODUCTION.

Amino acid sequence data are very helpful in the molecular characterisation of a protein as far as predicting specific secondary structures and structural and functional motifs are concerned. This can be accomplished by the use of computer algorithms and comparisons to the protein sequence information in the available databanks on the World Wide Web.

The crystal structures of a mammalian and prokaryotic ODC have been determined (Momany *et al.*, 1995; Kern *et al.*, 1996). The sequence similarity between different species indicated that the enzymes may share some structural features (McCann and Pegg, 1992). The mammalian and prokaryotic ODC is proposed to be a homodimeric enzyme of two identical monomers. ODC has a MM of about 100 000 with protomer MM ranging between 50-54 kDa, depending on the species (Heby, 1985). Each sequence contains 12 cysteine residues which are apparently not all involved in disulphide bridges (Kaye, 1984; Heby and Persson, 1990). The lysine at position 69 forms a Schiff base with pyridoxal-5′-phosphate (PLP) (Pegg *et al.*, 1994). The active site was located by active site-directed mutational analyses (Coleman *et al.*, 1993)(Lu *et al.* 1991) and was confirmed with active site-directed irreversible inhibitors like DFMO (Poulin *et al.*, 1992). The dimer contains two active sites located at the interface between the monomers and is made up of parts of each subunit (Tobias and Kahana, 1993a). The active site is formed between Cys-360 on one subunit and Lys-169 and His-197 on the other (residue numbers are for the murine ODC), comprising hydrophilic residues as shown in Fig. 1.3 (Chapter 1). The subunits dissociate and reassociate extremely rapidly with randomisation occurring within five minutes (Coleman *et al.*, 1994; Hayashi and Murakami, 1995). Gly387 of the murine enzyme has been implicated in the formation of stable homodimers (Tobias *et al.*, 1993b).

Phosphorylation of ODC by a polyamine-dependent protein kinase increases the enzyme activity (Brown *et al.*, 1994; Reddy *et al.*, 1996). Conversely, post-translational cross-linking between Gln residues and four putrescine molecules by transglutaminase, leads to a complete loss of ODC activity. The *in vivo* significance of the latter modification remains uncertain (Bachrach, 1984)(Russell *et al.* 1983).

The rapid intracellular turnover rate of ODC may be explained by the presence of PEST-rich regions (regions rich in Pro, Glu, Ser, Thr and to a lesser extent Asp) in its amino acid sequence (refer to Chapter 1). These sequences are characteristic of proteins with $t_{1/2}$ less than two hours, and are flanked by areas of positively charged amino acids as stated in the PEST-hypothesis of Rechsteiner and co-workers (Rogers *et al.*, 1986). The mouse ODC contains two PEST-rich regions in the C-terminus, the one most C-terminal accounting for the intracellular instability.

In this chapter, the full-length *P. falciparum* ODC cDNA has been assembled from the 3′- and 5′-RACE results and the amino acid sequence deduced. This has then been analysed for the presence of specific characteristics (general ODC characteristics and possible qualities unique to the *P. falciparum* protein), in comparison to the ODC amino acid sequences of other organisms.

4.2 MATERIALS AND METHODS.

4.2.1 Assembly of the Full-length Nucleotide Sequence of the *P. falciparum* ODC cDNA obtained with 5′- and 3′-RACE.

The 3′- and 5′-consensus nucleotide sequences of the cDNA as described in chapters 2 and 3, respectively, were used to construct the full-length nucleotide sequence with the Genetic Data Environment Program (Smith *et al.*, 1994). The deduced amino acid sequence of the consensus nucleotide sequence was obtained in all three possible reading frames using the same computer program. Multiple alignments of the amino acid sequences of the *P. falciparum* ODC and other organisms were performed with the CLUSTALW Program (Thompson *et al.*, 1994).

4.2.2 Computer Analyses of the Deduced Amino Acid Sequence of the *P. falciparum* ODC cDNA.

Molecular weight and amino acid composition determinations were performed with the ANTHEPROT Program (Geourjon *et al.*, 1991). This package was also used to predict the secondary structure, hydropathy and antigenicity plots. The sequence was also analysed with the PROSITE database to identify structural motifs (Bairoch *et al.*, 1995). The phylogenetic relationship between the *P. falciparum* ODC sequence and the sequences from the 13 organisms initially used for primer design (chapter 2, section 2.2.7 and Fig. 2.4) was obtained with the PHYLIP package, Version 3.2, employing the Protpars algorithm (Felsenstein, 1989)

4.3 RESULTS.

The full-length *P. falciparum* ODC sequence was obtained by overlap-alignment of the 3′- and 5′-consensus nucleotide sequences (Chapters 2 and 3). Fig. 4.1 indicates the overlap-alignment strategy employed. A consensus sequence was obtained for the overlap (Appendix II). The full-length consensus sequence comprised 2922 bp including the 18 bp poly-A tail and ODCF1 primer. The open reading frame contained ~2847 bp which includes the start and stop codons (including the 22 bp missing in our sequence downstream to the start codon as indicated in chapter 3, Fig. 3.10).

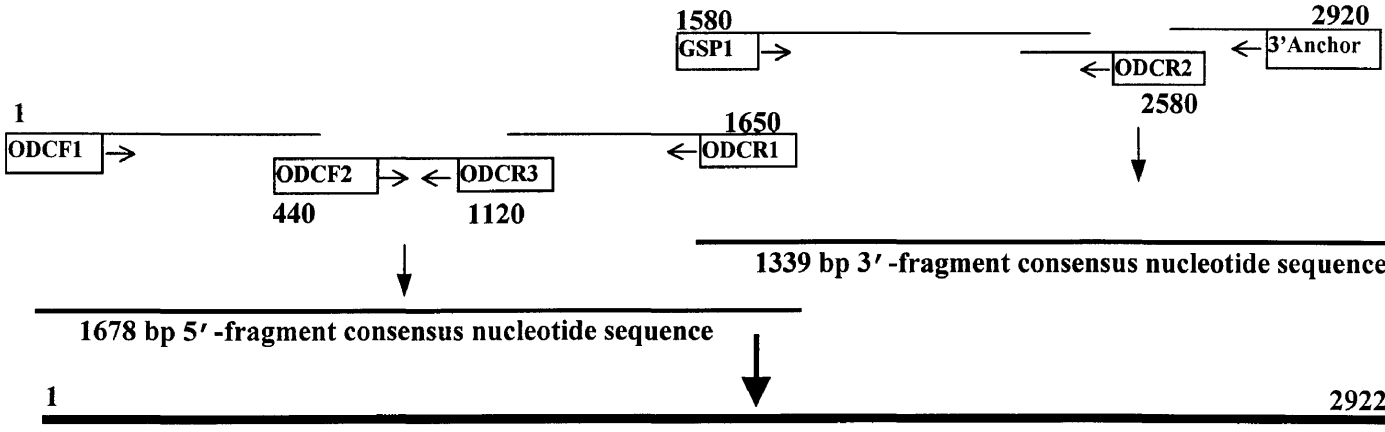

Figure. 4.1: Overlap-alignment strategy for obtaining the full-length consensus nucleotide sequence of the *P. falciparum* ODC cDNA. The primers used in the RACE and primer walking strategies are indicated. The open arrows indicate the directions of the primers. The relative positions of the primer annealing sites are indicated.

The consensus nucleotide sequence was translated to the corresponding amino acid sequence in the correct reading frame and compared to the *P. falciparum* ODC amino acid sequence in Genbank (Fig. 4.2). The deduced amino acid sequence contained ~ 939 residues, giving a predicted ODC monomer size of ~109.202 kDa as indicated with the ANTHEPROT program. The first 9 amino acids were not identified due to the lack of the first 22 nucleotides after the start codon (25 nucleotides in total including the start codon, see Fig. 3.10). The amino acid sequences between residues 10 and 545 were identical. The point mutations in the nucleotide sequence of the 3′-fragment (Chapter 2, Fig. 2.15) led to substitutions of some amino acids in the C-terminal of the protein as indicated by the blue residues in Fig. 4.2, but not a reading frame shift. Some of these substitutions were relatively conservative e.g. Val was changed to Ala. Others included drastic changes e.g. Arg635Ser; Thr684Pro, Asp697Ser, Gly669Asp and Tyr735Asn. An extra Asp was inserted at position 699. The overall identity was 98%. The conserved areas used in the design of the 3′-RACE primers GSP1 and GSP2 were identified. The consensus site used in the design of GSP2 was not conserved in the *P. falciparum* ODC sequence (I/V,I/V,Y/F AN**PC** changed to IIFAN**TI**, Fig. 4.2 and chapter 2, Fig. 2.4).

```
pfodc     MFMINYVFCEESNNMSKMVPDDDNNNYSSGKSCVYYQDLNKKEKEEYYRLNKKLRNDLFI   51
consams   ---------EESNNMSKMVPDDDNNNYSSGKSCVYYQDLNKKEKEEYYRLNKKLRNDLFI

pfodc     NSKQFYELHTFTERTVGFMRVQYFVYKLRDVVKCVEKETLLARSSSCLFMFNNIKRNDVH   111
consams   NSKQFYELHTFTERTVGFMRVQYFVYKLRDVVKCVEKETLLARSSSCLFMFNNIKRNDVH

pfodc     DDYVTKSSNGGVIKQLTERDVDDMYEYALNFCKQNKIVVVDTNTFFFDASKRKENLIKLE   171
consams   DDYVTKSSNGGVIKQLTERDVDDMYEYALNFCKQNKIVVVDTNTFFFDASKRKENLIKLE

pfodc     KVQTNEKDEYEEKDEVYRRGNNELSSLDHLDSKNNLIHMYYEKNKCDIINKDDENSTIAT   231
consams   KVQTNEKDEYEEKDEVYRRGNNELSSLDHLDSKNNLIHMYYEKNKCDIINKDDENSTIAT

pfodc     NNNDNNNDSSSYDKSITISRSSSCNNSHLSYSSFDNNHGNEKMKDYISVDENNNNNNNNK   291
consams   NNNDNNNDSSSYDKSITISRSSSCNNSHLSYSSFDNNHGNEKMKDYISVDENNNNNNNNK

pfodc     NNNVLLTLQRNSDDENGKDKDNEKNDVSLENNMEKNYKEEIWNYYTKNKVEVKTLEKVLN   351
consams   NNNVLLTLQRNSDDENGKDKDNEKNDVSLENNMEKNYKEEIWNYYTKNKVEVKTLEKVLN

pfodc     ENIDTSVVCINLQKILAQYVRFKKNLPHVTPFYSVKSNNDEVVIKFLYGLNCNFDCASIG   411
consams   ENIDTSVVCINLQKILAQYVRFKKNLPHVTPFYSVKSNNDEVVIKFLYGLNCNFDCASIG

pfodc     EISKVIKLLPNLSRDHIIFANTTKSINSLIYARKENINLCTFDNLDELKKIYKYHPKCSL   471
consams   EISKVIKLLPNLSRDHIIFANTTKSINSLIYARKENINLCTFDNLDELKKIYKYHPKCSL

pfodc     ILRINVDFKNYKSYMSSKYGANEYEWEEMLLYAKKHNLNIVGVSFHVGSNTKNLFDFCLA   531
consams   ILRINVDFKNYKSYMSSKYGANEYEWEEMLLYAKKHNLNIVGVSFHVGSNTKNLFDFCLA

pfodc     IKLCRDVFDMSSNMFNFYIINLGGGYPEELEYDNVKKHDKIHYCTLSLQEIKKDIQKFL   591
consams   IKLCRDVFDMSSNMFNFYIINLGGGYPEELEYDNVKKHDKIHYCTLSLQEIKKDIQKFL

pfodc     NEETFLKTKYGYYSFEKISLAINMSIDHYFSHMKDNLRVICEPGSYMVAASSTLAVKIIG   651
consams   NEETFLKTKYGYYSFEKISLAINMSIDHYFSHMKDNLRVICEPGRYMVAASSTLAVKIIG

pfodc     KRRPTFQGIMLKDLKAHYGPLNFAQQENKKQDETKINHNNDNNDNNDNNDNNDNNNNNN   711
consams   KRRPTFQGIMLKDLKAHYGPLNFAQQENKKQDEPKINHNNDNNDNNSNNDNNDNNNNNN

pfodc     QKGGQGNIMNDLIITSTNDSTSKKYDHSSRRVIQNVSCTIRDKRVDNIKINTHTINNPNI   771
consams   QKGGQGNIMNDLIITSTNDSTNKKNDHSSSQVIQNVSCTIRDKRGDNIKINTHTINNPNI

pfodc     NGKENTVDGDNINIAHKNIGNNFSSSNSKLGNITNIKKKVVNINDNRYNYFSYYVSDSIY   831
consams   NGKENTVDGDNINIAHKNIGNNFSSSNSKLGNITNIKKKVVNINDNRYNYFSYYVSDSIY

pfodc     GCFSGIIFDEYNRCPIYVIKNKNNPNQNFMNFNLYLANVFGQSCDGLDMINSITYLPECY   891
consams   GCFSGIIFDEYNRCPIYVIKNKNNPNQNFMNFNLYLANVFGQSCDGLDMINSITYLPECY

pfodc     INDWLLYEYAGAYTFVSSSNFNGFKKCKKVYIFPESKPSLKGQPNKHW               939
consams   INDWLLYEYAGAYTFVSSSNFNGFKKCKKVYIFPESKPSLKGQPNKHW
```

Figure 4.2: The deduced amino acid sequence for the full-length *P. falciparum* ODC cDNA. The 'pfodc' sequence indicates the Genbank sequence while the 'consams' sequence depicts the derived amino acid sequence from the consensus nucleotide sequence of the full-length ODC cDNA obtained with RACE. Blue residues indicate positions that are different between the two sequences. The red indicates the consensus area used in the design of GSP1 and the green the area used for GSP2 (chapter 2).

The *P. falciparum* ODC amino acid sequence obtained was also aligned with the corresponding sequences from four related organisms in order to identify species-specific differences (Fig. 4.3). The *P. falciparum* ODC contains an extra N-terminal ~273 amino

acids and is truncated by ~ 41 residues at the C-terminal end compared to the mammalian ODCs. Furthermore, five regions of >10 amino acids each were inserted in the *P. falciparum* sequence (residues 338-360, 566-603, 664-674, 686-709 and 714-821). One insertion contained ~108 amino acids (residues 714-821). The inserts were also Asn/Asp rich, especially regions 686-709 and 714-821. The residues that are proposed to be involved in the active site of eukaryotic ODCs (i.e. Lys, His and Cys) were conserved in the *P. falciparum* sequence as indicated by the red blocks in the alignment (residues Lys 489, His 517 and Cys 875). The binding site for PLP was also retained in *P. falciparum* ODC at Lys 435 (blue box in Fig. 4.3). Position Gly 914 was conserved as the proposed residue involved in dimerisation.

Amino acid analysis of the *P. falciparum* ODC protein composition with the ANTHEPROT Program indicated that the most abundant residue was Asn (15% of the total amino acids) with Lys second at 10%. Leu, Ile, Val, Ser, Asp and Glu were present at ~5-7% each while the rest of the amino acids was present at low frequencies (<5%). The protein contained 21 Cys residues. The average occurrence of amino acids in over 200 diverse proteins analysed revealed that the most abundant amino acids are normally Gly, Ala, Leu, Ser and Lys (Klapper, 1977). The prevalence of Asn in the *P. falciparum* ODC protein is thus atypical, as this amino acid usually constitutes only 4% of the total amino acids in other proteins (Klapper, 1977).

PROSITE identified eleven possible N-glycosylation sites (Asn-X-Ser/Thr) shown in Fig. 4.3 as orange boxes. The protein does contain 8% Ser and 3% Tyr residues and therefore O-glycosylation can not be excluded. Several possible phosphorylation sites, including Casein kinase II (e.g. TI/ERD or SSL/YD), Protein kinase C (e.g. TER), Tyrosine kinase (e.g. KKEKEEYY) and cAMP-dependent protein kinase (RRPT) sites, were identified. The protein also contains an amidation site (position 676, residues IGKR) indicated by a blue ellipse in Fig. 4.3. A helix-loop-helix dimerisation domain signature was indicated for residues 244-252 (DKSITISRS) (Fig. 4.3, red ellipse).

```
Human       ------------------------------------------------------------
Mouse       ------------------------------------------------------------
Trypanosoma ------------------------------------------------------------
Leishmania  ---------MGDHD----------VALCHVSRYNHANYW--------------AFVPLP
PLASMODIUM  EESNNMSKMVPDDDNNNYSSGKSCVYYQDLNKKEKEEYYRLNKKLRNDLFINSKQFYELH   60

Human       ------------------------------------------------------------
Mouse       ------------------------------------------------------------
Trypanosoma ------------------------------------------------------------
Leishmania  TVSDDTG-------------------CDSLHHDSASERIRMAPPAS-------ASKAGA
PLASMODIUM  TFTERTVGFMRVQYFVYKLRDVVKCVEKETLLARSSSCLFMFNNIKRNDVHDDYVTKSSN  120

Human       ------------------------------------------------------------
Mouse       ------------------------------------------------------------
Trypanosoma ------------------------------------------------------------
Leishmania  AEERLHPYERRLLDQYQIHLQPANRNPLSRADSAAGREETAQTPAQVQMVPVVAVADSTS
PLASMODIUM  GGVIKQLTERDVDDMYEYALNFCKQNKIVVVDTNTFFFDASKRKENLIKLEKVQTNEKDE  180

Human       ------------------------------------------------------------
Mouse       ------------------------------------------------------------
Trypanosoma ------------------------------------------------------------
Leishmania  DQHASVASSQDLVDLFFLEGSQAVDGLCFSPYPIYGWRTAEERRAAVCEVFKTYNVVTRL
PLASMODIUM  YEEKDEVYRRGNNELSSLDHLDSKNNLIHMYYEKN--KCDIINKDDENSTIATNNNDNNN  238

Human       ----------------------------------MNNFG---------NEEFDCHFLD-   15
Mouse       ----------------------------------MSSFT---------KDEFDCHILD-
Trypanosoma ------------MTTKSTPSSLS-VNCLVAQTEKSMDIVV---------NDDLSCRFL--
Leishmania  PASPAALAAAQRRYSRHRHSAIAPINKSAIETREQYWRRL---------SNLYTQKGVK-
PLASMODIUM  DSSSYDKSITISRSSSCNNSHLSYSSFDNNHGNEKMKDYISVDENNNNNNNNKNNNVLLT  298

Human       -------EGFTAKDILDQKINE-----VSSSDDKDAFYV----------------------   42
Mouse       -------EGFTAKDILDQKINE-----VSSSDDKDAFYV----------------------
Trypanosoma -------EGFNTRDALCKKIS------MNTCDEGDPFFV----------------------
Leishmania  -------DAASAADAAATTATNGAVPAAPAYEPEDPFYI----------------------
PLASMODIUM  LQRNSDDENGKDKDNEKNDVSLENNMEKNYKEEIWNYYTKNKVEVKTLEKVLNENIDTSV  358

Human       --ADLGDILKKHLRWLKALPRVTPFYAVKCNDSKAIVKTLAATGTGFDCASKTEIQLVQS   94
Mouse       --ADLGDILKKHLRWLKALPRVTPFYAVKCNDSRAIVSTLAAIGTGFDCASKTEIQLVQG
Trypanosoma --ADLGDIVRKHETWKKCLPRVTPFYAVKCNDDWRVLGTLAALGTGFDCASNTEIQRVRG
Leishmania  --IDLGRVVEQMARWRHELPMVRPYFAVKSNPQPAVLEVLSALGAGFDCASKEEIHMVLG
PLASMODIUM  VCINLQKILAQYVRFKKNLPHVTPFYSVKSNNDEVVIKFLYGLNCNFDCASIGEISKVIK  418

Human       --LGVPPERIIYANPCKQVSQIKYAANNGVQMMTFDSEVELMKVARAHPKAKLVLRIATD  146
Mouse       --LGVPAERVIYANPCKQVSQIKYAASNGVQMMTFDSEIELMKVARAHPKAKLVLRIATD
Trypanosoma --IGVPPEKIIYANPCKQISHIRYARDSGVDVMTFDCVDELEKVAKTHPKAKMVLRISTD
Leishmania  RQLVASPDDIIFANPCKQLGDLREAQACGVTYVTVDNPLEMEKISRLMPSAHAIIRIKTN
PLASMODIUM  LLENLSRDRIIFANTIKSINSLIYARKENINLCTFDNLDELKKIYKYHPKCSLILRINVD  438

Human       DSKAVCRLSVKFGATLRTSRLLLERAKELNIDVVGVSFHVGSGCTDPETFVQAISDARCV  206
Mouse       DSKAVCRLSVKFGATLKTSRLLLERAKELNIDVIGVSFHVGSGCTDPDTFVQAVSDARCV
Trypanosoma DSLARCRLSVKFGAKVEDCRFILEQAKKLNIDVTGVSFHVGSGSTDASTFAQAISDSRFV
Leishmania  DSKAQCSFSTKFGAPLEDVEGLLEAARQFNVTVCGVSFHVGSGNDDQSAYVSAVRDAYQV
PLASMODIUM  FKNYKSYMSSKYGANEYEWEEMLLYAKKHNLNIVGVSFHVGSNTKNLFDFCLAIKLCRDV  538

Human       FDMGAEVGFSMYLLDIGGGFPGSEDVK---------------------------------  233
Mouse       FDMATEVGFSMHLLDIGGGFPGSEDTK---------------------------------
Trypanosoma FDMGTELGFNMHILDIGGGFPGTRDAP---------------------------------
Leishmania  FQQAVQYGFKCTILDIGGGFPGTEVVEGS--------------------------------
PLASMODIUM  FDMSSNMGFNFYIINLGGGYPEELEYDNAKKHDKIHYCTLSLQEIKKDIQKFLNEETFLK  598

Human       -----LKFEEITGVINPALDKYFPSDS-GVRIIAEPGRYYVASAFTLAVNIIAKK-----  283
Mouse       -----LKFEEITSVINPALDKYFPSDS-GVRIIAEPGRYYVASAFTLAVNIIAKK-----
Trypanosoma -----LKFEEIAGVINNALEKHFPPDL-KLTIVAEPGRYYVASAFTLAVNVIAKK-----
Leishmania  ---GNTSFEAIARTIRPVLAELFGGG--DVTIISEPGRYFTAASHALLMNVFASR-----
PLASMODIUM  TKYGYYSFEKISLAINMSIDHYFSHMKDNLRVICEPGRYMVAASSTLAVKKIGKRRPTFQ  658
```

Figure 4.3: Multiple alignment of the ODC amino acid sequences. *P. falciparum* sequence (this study) was aligned with *Homo sapiens, Mus musculus, Trypanosoma brucei* and *Leishmania donovani* sequences. The residues in red indicate 100% homology while those in green indicate residues conserved in terms of properties. The shadowed, red boxes indicate the proposed active site residues, while the blue box indicates the PLP-cofactor binding site. The yellow box indicates the Gly possibly involved with dimerisation. The green boxes in the mammalian sequences indicate the positions of the identified PEST-rich regions. N-glycosylation sites are indicated in the *P. falciparum* sequence by orange boxes. An amidation site is identified by the blue ellipse and the helix-loop-helix dimerisation domain by a red ellipse.

The insertions present in the amino acid sequence prevented the modelling of the three-dimensional structure of the *P. falciparum* ODC on the known structures of other proteins. To obtain a prediction of the structural characteristics of the protein, the ANTHEPROT program was used. Fig. 4.4 predicts that the protein contains more α-helixes than β-sheets. Only two regions (between residues 233 and 291 as well as 697 and 871) favoured the dominance of β-sheet conformations.

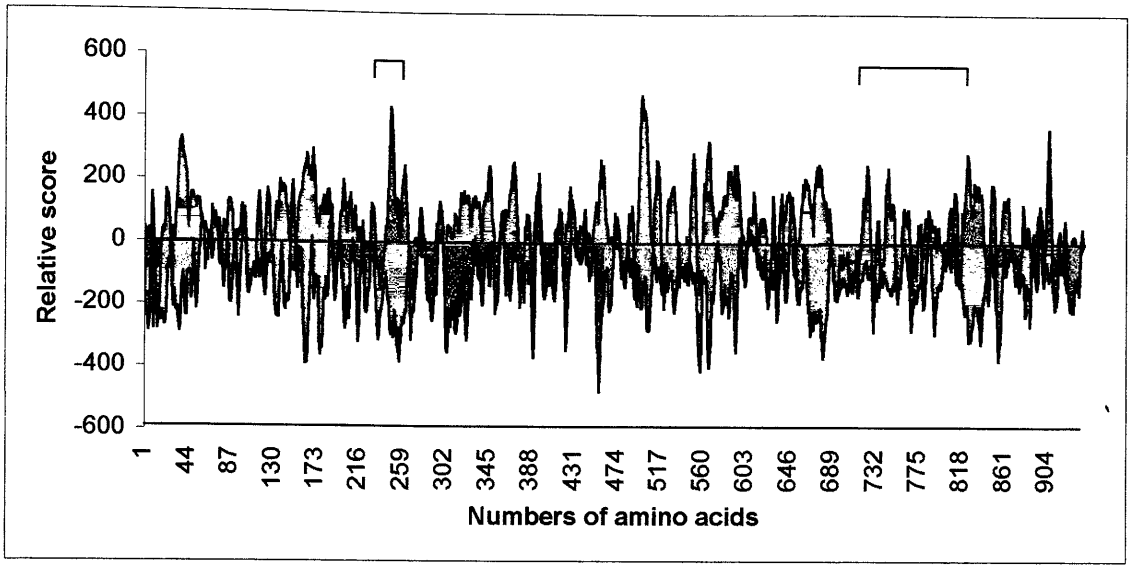

Figure 4.4: Secondary structure prediction of the deduced *P. falciparum* ODC amino acid sequence. The blue indicates areas where α-helixes are formed and the red where β-sheets are formed. The brackets indicate the two areas where β-sheet conformation is favoured.

The hydropathy plot prediction indicates a higher abundance of hydrophilic residues compared to hydrophobic amino acids (Fig. 4.5). The region including residues ~800 to 900 is the most apolar, and corresponds to the area with the highest β-sheet tendency (Fig. 4.4). No transmembrane anchoring or spanning regions (i.e. more than 20 hydrophobic residues) were observed.

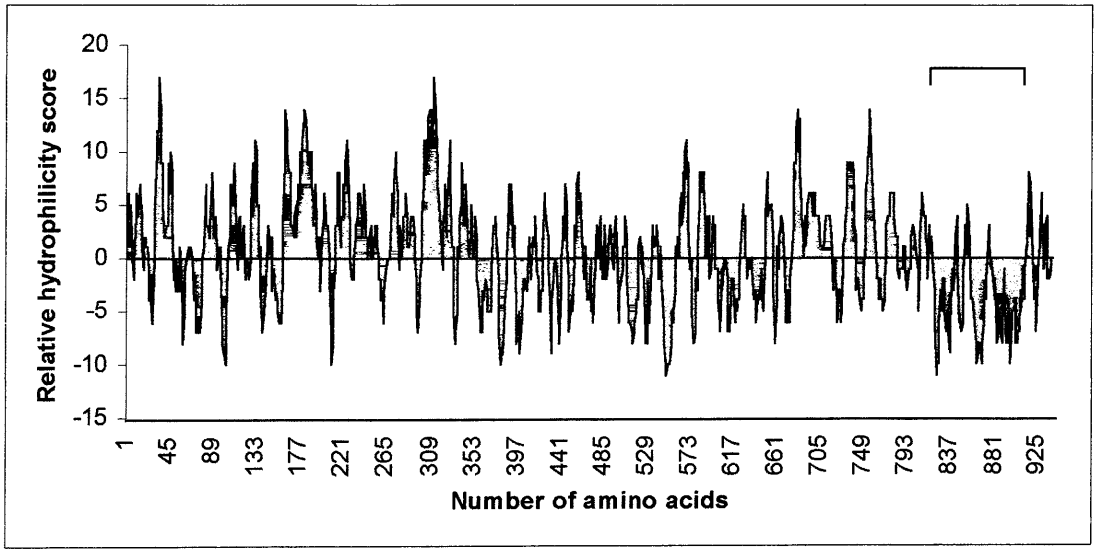

Figure 4.5: Hydrophilicity plot of the deduced *P. falciparum* ODC amino acid sequence. The brackets indicate the most hydrophobic residues.

The properties of the *P. falciparum* ODC amino acid sequence in terms of immunogenicity were investigated to obtain an idea of the possible function of the insertions in the sequence. The antigenicity profile in Fig. 4.6 showed negative antigenic scores throughout the protein. Only two areas, residues ~196-202 (SSLDHLD) and 664-673 (DLKAHVDPLN), showed antigenic properties based on antigenicity scores of ~3. Antigenic segments generally correspond roughly to the highest peaks in hydrophilicity plots, but the most hydrophilic region around residue 309 in *P. falciparum* ODC (Fig. 4.5) is not identified as a possible antigenic epitope in figure 4.6. These results suggest that the *P. falciparum* ODC protein does not contain immunogenic epitopes.

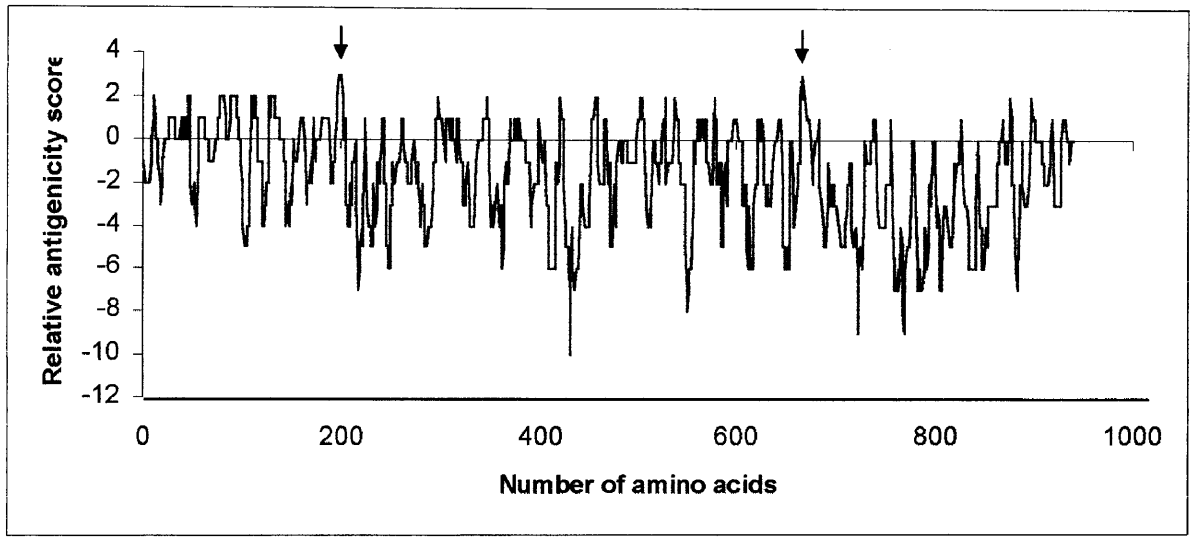

Figure 4.6: Antigenicity scores for the deduced *P. falciparum* ODC amino acid sequence. The arrows indicate the two areas with the highest antigenicity scores.

Because of the size of the *P. falciparum* ODC protein, it was of interest to determine its phylogenetic relationship to ODCs from the 13 other organisms used originally in 3′-RACE primer design (Fig. 2.4). The phylogenetic tree obtained with the PHYLIP package indicated that the *P. falciparum* protein grouped with the other parasitic protozoa (*Leishmania* and *Trypanosoma,* Fig. 4.7). The mammalian proteins were grouped separately.

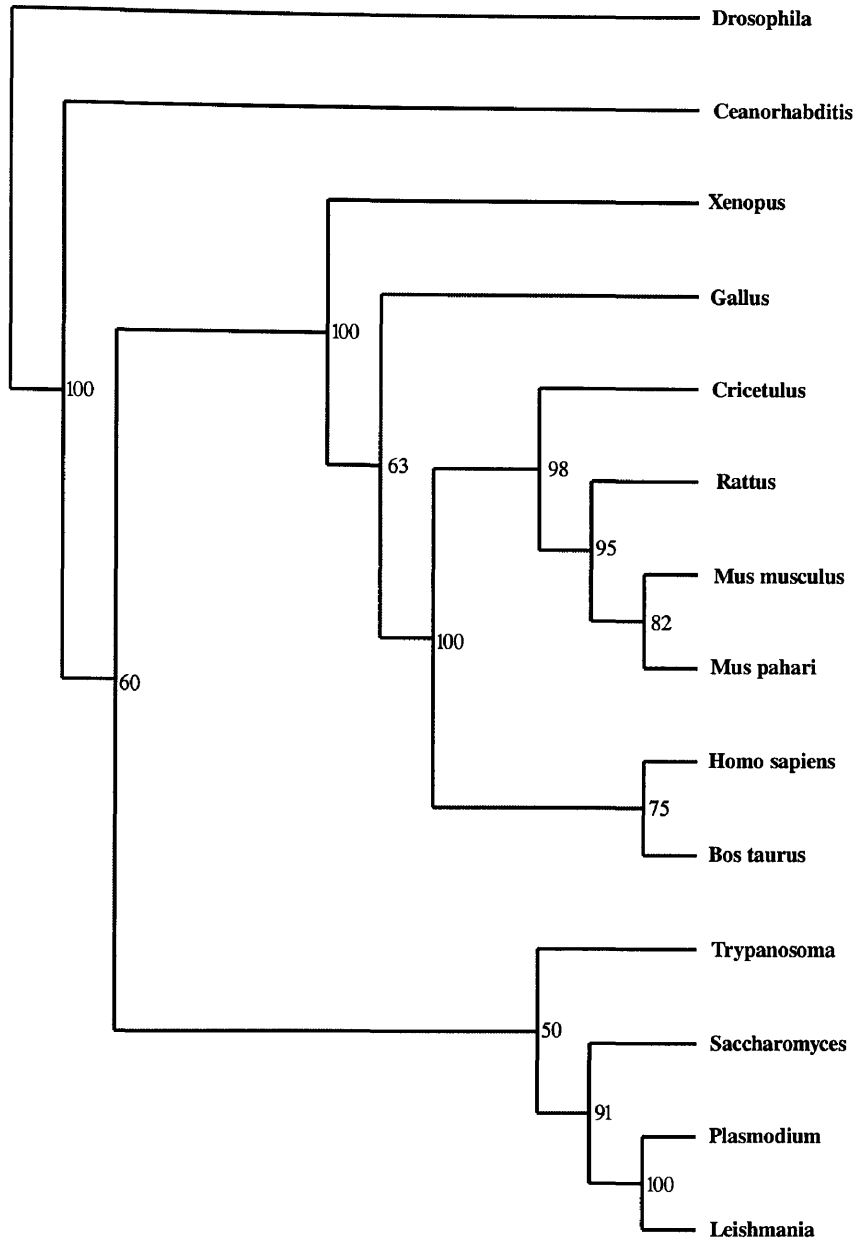

Figure 4.7: Unrooted phylogenetic tree of the ODC proteins for 14 different organisms including *P. falciparum*. The organisms are as in Fig. 2.4: *Drosophila melanogaster, Bos taurus, Ceanorhabditis elegans, Gallus gallus, Homo sapiens, Leishmania donovani, Mus musculus, Mus pahari, Rattus norvegicus, Trypanosoma brucei, Xenopus laevis, Saccharomyces cerevisiae* and *Cricetulus griseus* and *P. falciparum.* One hundred bootstrapped data sets were used to construct a consensus tree. The bootstrap values (%) are indicated on the tree as the confidence in the branch.

4.4 DISCUSSION.

The overlap-alignment of the consensus nucleotide sequences for the 5′- and 3′-fragment of the *P. falciparum* ODC cDNA indicated an open reading frame of ~ 2847 bp (Appendix II). This is approximately 1500 bp larger than the mammalian ODC genes

(1350 bp on average) but compares well with other protozoa like the *L. donovani* ODC gene (~2732 bp) and the *T. brucei* ODC gene (~1641 bp) (Phillips *et al.*, 1987; Hanson *et al.*, 1992).

The deduced ODC amino acid sequence was derived in the correct reading frame as indicated by the absence of multiple stop codons and inappropriate codons. Comparison to the *P. falciparum* sequence data deposited in Genbank indicated several nucleotide differences that led to amino acid substitutions, some of which were of a conservative nature e.g. Val-Ala. By contrast, drastic changes from uncharged to charged, polar amino acids (e.g. Ser-Arg) and a Thr-Pro was also observed. The significance of these differences cannot be assessed until further proof of reliability of the sequence in Genbank.

The deduced amino acid sequence of the *P. falciparum* ODC cDNA predicts a protein of ~939 residues and a MM of ~109.202 kDa. The amino acid sequence has areas of high homology with ODC protein sequences from other organisms, implying that the large cDNA encodes only the monomer of the homodimeric enzyme. This makes the *P. falciparum* ODC monomer approximately twice as large as mammalian ODC monomers (~50 kDa), due to the presence of the inserted amino acids (Heby, 1985). This is in stark contrast to the reported MM of 51 kDa for the *P. falciparum* ODC monomer as determined by SDS-PAGE (Assaraf *et al.*, 1988). Assaraf reported that the protein migrated as a major band of 51 kDa and a minor band of 49 kDa under non-reducing conditions. The authors did not give any explanation for the presence of the second band. It is possible that degradation or fragmentation of the protein occurred as the combined MM of the two bands is 100 kDa. This discrepancy in MM was also observed for the ODC of *S. cerevisiae*. The native protein had a MM of 86 kDa, but SDS-PAGE indicated a monomer size of 50 kDa (Fonzi, 1989).

It is conceivable that the amino acid inserts present in the *P. falciparum* ODC protein can be removed by post-translational proteolysis. Analysis of the residues at the beginning or ends of the inserts indicates the presence of several Lys and Arg residues that could be involved in cleavage by proteases like trypsin (see for example Lys at position 278, Fig. 4.3). The removal of all of the inserts surrounded by Lys or Arg (regions 566-603, 686-

709 and 714-821) and the N-terminal 273 residues results in a much smaller protein of ~499 amino acids with a MM of ~ 55 kDa. This could explain the ~51 kDa band obtained by Assaraf during SDS-PAGE (Assaraf *et al.*, 1988), and might indicate that the size of the *P. falciparum* ODC corresponds to proteins from other organisms. The possibility exists that the processed protein could be stabilised as a multi-chain molecule with disulphide bridges after proteolysis to form the active enzyme, much the same as e.g. chymotrypsin and insulin are formed from large precursors (Lodish *et al.*, 1995).

The *P. falciparum* ODC protein is predicted by PHYLIP to be closely related to those from the other parasitic protozoa, *L. donovani* and *T. brucei* (Fig. 4.7). Usually, proteins from this organism would more closely resemble those from human, mouse, *Xenopus*, *Trypanosoma* and chicken than the proteins from *Leishmania, Drosophila* and yeast (Hyde *et al.*, 1989). The grouping of ODC from *P. falciparum* could be due to the insertions present in the protein sequence.

The *P. falciparum* ODC protein is extended by ~273 residues at the N-terminus compared to the mammalian ODCs. This compares to the *L. donovani* protein that is extended by ~200 residues and the ~20 residues added to the *T. brucei* enzyme. In contrast, the *P. falciparum* protein lacks the last 41 residues at the C-terminus as compared to the mammalian protein. The C-terminus has been implicated in the short $t_{1/2}$ of the enzyme in mammals, as it contains the most important PEST-rich region (residues 423-449, mammalian numbering, Fig. 4.3). Organisms that lack this region have been shown to have stable ODC proteins, including *T. brucei* (lacking 35 residues) and *L. donovani* (28 residues shorter) (Ghoda *et al.*, 1990; Hanson *et al.*, 1992). The loss of the C-terminal residues supports the $t_{1/2}$ of >75 min reported for the *P. falciparum* protein (Assaraf *et al.*, 1988). The murine enzyme also contains a second PEST-rich region (residues 298-333) with a lower PEST score. This region is relatively homologous between the organisms and corresponds to residues 822-841 of *P. falciparum* ODC (Fig. 4.3). The possibility of this and other PEST-rich regions in *P. falciparum* ODC needs to be investigated.

The *P. falciparum* ODC protein sequence contained several features common to ODCs of other organisms. The proposed active site residues (residues Lys 489, His 517 and Cys 875, Fig. 4.3) were conserved in the *P. falciparum* ODC protein, indicating similar activities to ODCs already characterised (Lu *et al.*, 1991; Coleman *et al.*, 1993). The PLP binding site (Lys 435) was also conserved in the *P. falciparum* ODC sequence, confirming the dependency of enzyme activity on this cofactor and its classification as a group IV PLP-dependent decarboxylase (Pegg *et al.*, 1994). The mammalian ODC sequences contain 12 Cys residues, presumably not all of which are involved in disulphide bridges as enzyme activity is dependent on sulphydryl reducing agents during purification (Heby and Persson, 1990). This may suggest that a few sulphydryl groups are required for enzymatic activity and the rest may be involved in disulphide bridges for stabilisation of the enzyme. The *P. falciparum* ODC sequence contained 21 Cys residues but their function remains to be elucidated.

The *P. falciparum* ODC protein contained putative sites for phosphorylation by various protein kinases, which may play a role in the regulation of protein activity. Phosphorylation has been shown to increase the stability of mammalian enzymes to intracellular proteolysis compared with unphosphorylated ODC. The catalytic activity was also ~50% higher in the phosphorylated enzyme (Brown *et al.*, 1994). A possible transamidation site (IGKR, position 676) was identified but the significance is not clear at this stage. Transamidation with Lys as the amino-group donor to form a isopeptide bond with a carboxyl group of Glu occurs in the stabilisation of fibrin with Factor XIIIa (Lodish *et al.*, 1995). It has been shown that a Gly at position 387 of the murine enzyme is essential for the formation of stable homodimers (Tobias *et al.*, 1993b). This residue is also conserved in the *P. falciparum* ODC (residue 914, Fig. 4.3) and might therefore also be implicated in the dimerisation of the parasite enzyme.

Analyses of the amino acid sequence of *P. falciparum* ODC indicated the presence of features unique to the parasitic enzyme. An interesting aspect is the presence of $(NND)_x$- and $(N)_x$-repeat insertions in the *P. falciparum* ODC sequence that are not observed in any of the other organisms. *P. falciparum* does have various additional Asn-rich proteins in particular STARP (sporozoite threonine and asparigine rich protein) (Facer and Tanner, 1997) as well as the clustered asparigine rich protein (CARP) (Wahlgren *et al.*,

1991), and also includes proteins with Asn-repeats like the circumsporozoite protein (PNANP repeat) (Kwiatkowski and Marsh, 1997). These repeats are present in immunodominant domains normally associated with antigenic proteins on the surface of the parasite (see chapter 1, section 1.3.3). To determine if the $(NND)_x$-repeats found in the *P. falciparum* ODC could have any immunogenic significance, an antigenic profile of the amino acid sequence was obtained. The profile indicated only two possible regions with slight antigenic properties (Fig. 4.6). Both occur in regions containing inserted sequences in the parasite protein, although they did not appear to be located at the $(NND)_x$-rich areas and were not in the most hydrophilic segments of the protein. This indicates that *P. falciparum* ODC probably does not contain marked antigenic properties. Expression studies and mutational analyses are needed to ascertain the function of these regions in the *P. falciparum* ODC protein, whether immunological or related to structure-function aspects.

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 contribute to the structure-function relationship of the proteins.

The *P. falciparum* ODC is not unusual in containing species-specific inserts, as other well-characterised *P. falciparum* proteins like DHFR and P-type cation ATPases also have major insertions (Krishna *et al.*, 1993). In the case of the cationic ATPases, there were four regions of inserts (of ~120 residues) that varied in length between different parasite isolates. DHFR contains two areas where ~20 residues were inserted compared to the human enzyme (Vasques *et al.*, 1996)(F. Joubert, personal communication). 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).

Several possible N-glycosylation sites were identified. It has only recently been suggested that *P. falciparum* proteins can contain this form of glycosylation (Kimura *et*

al., 1996). In terms of formation of the active homodimer, a helix-loop-helix dimerisation domain was identified in the N-terminal extended region of the *P. falciparum* ODC. These domains are usually a characteristic of the DNA binding proteins (transcription factors) and the overall structure is similar to a leucine zipper, where hydrophobic residues in the C-terminal helix allows dimerisation. This motif has not been identified in any other ODC.

The secondary structure prediction indicated the abundance of α-helix forming residues and the protein has an overall hydrophilic nature, typical of globular proteins. Unfortunately, the major insertions in the amino acid sequence of the *P. falciparum* ODC protein are not conducive to the modelling of this protein on the known three-dimensional structures of other proteins. However, all of the molecular characterisations mentioned in this chapter are only predictions until they are confirmed empirically with the determination of the structure of the protein or analyses of the expressed protein activity.

The results presented in this chapter represent the last of the experimental 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 *P. falciparum* ODC cDNA.

CHAPTER 5

Concluding Discussion

Malaria remains one of the most important tropical infectious diseases of man. The emergence of multi-drug resistant parasites and the lack of a viable vaccine have stimulated the research 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, 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 attractive 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. By applying enhanced evolution techniques, the development of possible mechanisms of resistance could be predicted and considered in the design of second-line drugs (Stemmer, 1994).

Several metabolic proteins of the malaria parasite, *P. falciparum,* have been shown to be fundamental in parasite survival (see chapter 1, table 1.1). One enzyme, ornithine decarboxylase, seems indispensable for the growth and maturation of the asexual parasite. Its inhibition with a competitive inhibitor arrests parasite development in schizogony to the point of decreasing the total parasite burden. The protein from the parasite was partially purified (Assaraf *et al.*, 1988), but further characterisation requires large amounts of the pure protein. This could be achieved by the application of molecular biology techniques to isolate and express the gene encoding the protein.

This dissertation describes the successful isolation and identification of the ODC cDNA from *P. falciparum.* Conventional methods to obtain a gene include screening of a cDNA library constructed for the particular organism. However, more efficient methods have been devised in the last ten years. RT-PCR and RACE protocols were applied as it selects

and amplifies the specific mRNA/cDNA within a few steps. 3′-RACE was conducted on the *P. falciparum* cDNA with degenerate gene-specific primers for ODC derived from a consensus amino acid sequence after multiple alignment of the sequences of other organisms. The successful identification of the *P. falciparum* ODC cDNA is to our knowledge the first example where a *P. falciparum* cDNA was isolated by using degenerate primers in RACE protocols.

Several reasons for the difficulty experienced with RACE on the malaria genome were identified in chapter 2. The importance of the degenerate gene-specific primer used must not be underestimated, as this primer is crucial to the specificity of the reaction. Even after extensive manipulation of the amplification reaction components, no specific product could be obtained for primer GSP2. It was later shown that the conserved site for this primer is modified in *P. falciparum* ODC (Fig. 4.2). Therefore, two or more degenerate primers should be considered if enough consensus areas can be identified.

Success was achieved with primer GSP1 after careful manipulation of the cDNA synthesis and RACE procedure as described in chapter 2. The most important variable was found to be the template concentration, as too low template copies produced poor results. Loss of messages occurred when cDNA was prepared from isolated mRNA, possibly due to saturation of the oligo(dT) beads used for mRNA isolation. This could happen if internal poly-A stretches of mRNA, rRNA or genomic DNA anneals since the malaria genome is ~ 81% AT-rich (K. Clark, personal communication). This led to the use of total RNA for cDNA synthesis. Processing the total RNA to remove contaminating genomic DNA also resulted in the decrease of template copies, as ~50% of the RNA was lost in precipitation steps. The best results were achieved with unprocessed total RNA, shown to contain a high amount of template copies (Chapter 2, Fig. 2.11).

Truncated cDNAs were observed possibly as a result of excess DD-Poly-T primer annealing at poly-A stretches internally in the genes. This led to the introduction of a hot-start protocol during reverse transcription in an attempt to allow more specific annealing to at least 18 A's. It was also shown that the ratio of DD-Poly-T primer to total RNA was important in obtaining optimal reverse transcription of the messages for a particular gene.

This ratio should be arbitrarily determined for each transcript depending on its relative abundance. In the case of the ODC transcript, a 5:1 ratio seemed optimal (Fig. 2.11).

The novel application of suppression PCR technology in differential amplification of cDNA derived from mRNA, was accomplished for the *P. falciparum* ODC cDNA in Chapter 3 (Lukyanov *et al.*, 1997). This system could prove to be invaluable in obtaining rare messages and can be used to overcome many of the problems experienced in the 3′-RACE. The amplification of the mRNA fraction will eliminate poly-A$^-$ RNA and genomic DNA contamination (these will form panhandle structures), and will increase the relative abundance of the template copies in the amplification of the uncloned cDNA library. Exploiting this method allowed the application of 5′-RACE to obtain the full-length *P. falciparum* ODC cDNA, the first report of its kind for the malaria organism. Usually, the obtained 3′-RACE fragment of the specific cDNA is used as a labelled probe to screen a cDNA library and thereby isolate the full-length cDNA sequence. These were the strategies used for isolation of the ODC cDNA from other protozoa, e.g. *L. donovani* and *T. brucei.*

Because PCR was involved in obtaining the *P. falciparum* ODC sequence, there is always the possibility, albeit minor, of misincorporation of nucleotides during the amplification, so that the sequence should be verified. As indicated, there were strain-specific differences in the sequence obtained from the PfUP1 strain in this study and the sequence from the strain deposited in Genbank. Our sequence was verified by sequencing more than one clone at least twice with each sequencing primer. The 5′-fragment of the cDNA sequenced without cloning should also be cloned to obtain the most upstream nucleotides. However, there are several pitfalls in cloning that should be carefully considered. Rearrangements can occur if the transformed cells are grown to the saturation phase due to the stress induced by nutrient deficiencies. Recombination can be eliminated by logarithmic growth at 30°C in recombinantly negative *E. coli* cells such as SURE and the use of plasmids with moderate copy numbers (Hanahan *et al.*, 1991).

The identity of the cDNA should be confirmed by functional studies on the recombinantly expressed protein. For this purpose, the full-length cDNA must be

amplified. The DNA polymerase system used in this study can support the long and accurate amplification (LA-PCR) of the whole ~3000 bp cDNA. This can be accomplished with the ODCF1 primer and either a newly designed reverse primer that anneals around the stop codon, or the existing 3′ Anchor primer. Further studies on the gene should also include characterisation of the 5′-UTR of the mRNA, taking into consideration the proposed secondary structures present and the possible involvement of these in the translational regulation of the gene (Coffino, 1989).

The full-length *P. falciparum* ODC cDNA sequence will allow the direct expression of the protein as the cDNA contains only the already spliced exons of the gene. Expression systems that have been shown to be able to decode the AT-rich *P. falciparum* messages include the eukaryotic *Xenopus laevis* oocyte system (although only transient expression is possible) and stable expression in another AT-rich organism e.g. *Dictyostelium discoidium*. Prokaryotic expression systems have been found in our laboratory to be particularly troublesome in expressing *P. falciparum* proteins (F. Joubert, personal communication).

The expressed, purified protein will allow numerous investigations to be conducted. Most importantly, the kinetic properties of the enzyme should be determined and compared to those reported for the partially purified protein (Assaraf *et al.*, 1988). The active site residues could be investigated with PCR-mediated site-directed point mutagenesis. Another interesting aspect that needs to be resolved is the disagreement between the reported molecular mass for the ODC monomer, and the predicted molecular mass based on the cDNA sequence data (Assaraf *et al.*, 1988). Possible techniques for accurate determination of the molecular mass include mass-spectrometry and analytical centrifugation of the recombinant protein as migration of proteins in SDS-PAGE can give misleading molecular mass predictions. It will also be interesting to determine if *P. falciparum* ODC is post-translationally modified to form the active enzyme that could be a multi-chain molecule. The involvement of Cys residues in disulphide bridges to stabilise the processed protein should be investigated.

SUMMARY

Malaria is 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 chemotherapy and –prophylaxis drugs. The need for novel therapeutic targets and drugs is therefore enormous and the understanding of the biochemistry of the parasite is imperative. The aim of this study was the identification and molecular characterisation of the cDNA of one such metabolic target protein, ornithine decarboxylase (ODC), in the human malaria parasite *P. falciparum.*

The *P. falciparum* ODC cDNA was isolated by means of a modified RT-PCR technique, RACE. No sequence data were available and the primers used were based on consensus areas identified in the protein sequences from other related organisms. The isolation and identification of the cDNA with degenerate primers was successful in 3′-RACE, but necessitated the optimisation of the cDNA synthesis protocol and the use of total RNA as starting material. The sequence obtained facilitated the application of 5′-RACE with ODC-specific primers based on the 3′-RACE sequence data. The full-length ODC cDNA sequence was obtained by overlap-alignment of various segments. A novel suppression PCR technology was applied during the 5′-RACE in order to create an uncloned cDNA library of amplified cDNAs representing only the mRNA population.

The *P. falciparum* ODC cDNA contains an open reading frame of ~2847 bp and translates to a large 939 amino acid protein. The protein contained large internal insertions and was extended by ~273 N-terminal residues compared to ODCs from other organisms. Several possible signature motifs were identified for phosphorylation, glycosylation and transamidation. The *P. falciparum* ODC protein seems to contain more hydrophilic and α-helix forming residues. These characteristics should be further investigated after expression of the recombinant protein.

The isolation of the *P. falciparum* ODC cDNA facilitates the validation of this protein as an antimalarial target.

OPSOMMING

Malaria is `n tropiese infeksie-siekte wat elke jaar miljoene mense aantas. Die siekte raak al hoe ernstiger as gevolg van parasiete wat weerstandig raak teen die bestaande voorkomende en terapeutiese middels. Dit is dus uiters noodsaaklik om nuwe terapeutiese teikens te identifiseer waarteen nuwe medisynes ontwerp kan word. Hierdie studie is onderneem om `n metaboliese teikenproteïen naamlik ornitien dekarboksilase (ODC), van die menslike malaria parasiet, *P. falciparum,* se cDNA te identifiseer en te karakteriseer.

Die cDNA van ODC was geïsoleer deur `n gemodifiseerde vorm van die RT-PKR tegniek te gebruik. Hierdie metode, genaamd RACE, word gebruik as daar nie inligting beskikbaar is oor die volgorde van die betrokke geen nie. Die voorvoerders wat gebruik is in die 3′-RACE metode was gebasseer op `n konsensus-volgorde wat geïdentifiseer was in die aminosuur-volgordes van die proteïen van ander organismes. Totale RNA is gebruik vir die sintese van cDNA, `n proses wat verskeie optimiseringstappe vereis het vir die beste resultate. Die geenvolgorde wat verkry is met 3′-RACE is gebruik om nuwe spesifieke voorvoerders te ontwerp om die volledige geen te isoleer met 5′-RACE. Die vollengte geenvolgorde is bepaal deur verskeie fragmente aanmekaar te las. Tydens 5′-RACE is `n geamplifiseerde, ongekloneerde cDNA biblioteek gevorm deur 'n unieke onderdrukkings PKR metode te gebruik sodat slegs cDNA vanaf mRNA geamplifiseer word.

Die *P. falciparum* ODC cDNA bestaan uit `n oop leesraam van ~2847 nukleotiede wat vir `n groot proteïen van 939 aminosure kodeer. Die proteïen het verskeie interne invoegings en besit `n verlengde N-terminale ent van ~273 aminosure. Daar is ook verskeie moontlike motiewe geïndentifiseer in die aminosuur volgorde insluitend fosforilering, glikosilering en transamidering modifikasie plekke. Die proteïen is hidrofilies met heelwat α-heliks-vormende aminosure. Hierdie karaktereienskappe sal verder bestudeer moet word na verkryging van `n rekombinante proteïen. Die beskikbaarheid van die *P. falciparum* ODC cDNA sal ondersoek na die geldigheid van diè proteïen as `n anti-malaria teiken, bevorder.

REFERENCES

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

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*. U. Bachrach. Boca Raton, CRC Press. **I:** 143-155.

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.

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., Messer, G. and Bachrach, U. (1987b) Cytostatic effect of DL-α–difluoromethylornithine against *Plasmodium falciparum* and its reversal by diamines and spermidine. *Parasitology Res.* **73**, 313-318.

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., 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.

Bacchi, C.J. and Yarlett, N. (1995). Polyamine metabolism. In: *Biochemistry and molecular biology of parasites*. J. J. Marr and M. Muller. London, Academic Press Ltd.**:** 119-132.

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

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

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

Bauer, P., Rolfs, A., Regitz-Zagrosek, V., Hildebrandt, A. and Fleck, E. (1997) Use of manganese in RT-PCR eliminates PCR artifacts resulting from DNAse I digestion. *Biotechniques* **22**, 1128-1132.

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) Purine and pyrimidine metabolism. In: *Biochemistry and Molecular Biology of Parasites*. J. J. Marr and M. Muller. London, Academic Press Ltd.:89-118

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., Visser, L., Neitz, A.W.H. and Louw, A.I. (1998b) The nucleotide sequence of a *Plasmodium falciparum* ornithine decarboxylase gene. Biochemistry in Africa, Potchefstroom, South Africa.

Birkholtz, L., Visser, L., Louw, A.I., van der Linde, R. and Brink, A.J. (1998c) The prevalence of mixed-species and anti-folate resistant malaria infections in Mpumalanga. *S. A. Med. J.* **88**, 58-60.

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.

Borst, P. and Ouellette, M. (1995) New mechanisms of drug resistance in parasitic protozoa. *Ann. Rev. Microbiol.* **49**, 427-460.

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

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.

Canellakis, E.S. and Hayashi, S. (1989). The role of antizyme in the regulation of ornithine decarboxylase activity in eukaryotic cells. In: *The physiology of polyamines*. U. Bachrach. Boca Raton, CRC Press. **I:** 315-330.

Chenchik, A., Diachenko, L., Moqadam, F., Tarabykin, V., Lukyanov, S. and Siebert, P.D. (1996) Full-length cDNA cloning and determination of mRNA 5'- and 3'- ends by amplification of adaptor-ligated cDNA. *Biotechniques* **21**, 526-535.

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

Ciechanover, A., Finley, D. and Varshavsky, A. (1984) Ubiquitin dependence of selective protein degradation demonstration in the mammalian cell cycle mutant ts85 cells. *Cell* **37**, 57-66.

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.

Coffino, P. (1989). Molecular biology of eukaryotic ornithine decarboxylase. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 135-147.

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. (1994) Prospects for malaria control through the genetic manipulation of its vectors. *Parasitology Today* **10**, 370-371.

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

Cowman, A.F. (1995) Mechanisms of drug resistance in malaria. *Aust. NZ J. Med.* **25**, 837-844.

Cowman, A.F., Morry, M.J., Biggs, B.A., Cross, G.A.M. and Foote, S.M. (1988) Amino acid changes linked to pyrimethamine resistance in dihydrofolate reductase-thymidylate synthase genes of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **85**, 9109-9113.

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

Craig, S.P. and Eakin, A.E. (1997) Purine salvage enzymes of parasites as targets for structure-based inhibitor design. *Parasitology Today* **13**, 238-241.

Deitsch, K.W. and Wellems, T.E. (1996) Membrane modification in erythrocytes parasitised by *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **76**, 1-10.

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

DiLella, A.G. and Woo, S.L. (1987) Hybridisation of genomic DNA to oligonucleotide probes in the presence of tetramethylammonium chloride. *Methods in Enz.* **152**, 447-451.

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

Edwards, J.B.D.M., Delort, J. and Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNA's: a new tool for cloning 5' ends of mRNA's and for constructing cDNA libraries by *in vitro* amplification. *Nucl. Acids Res.* **19**, 5227-5232.

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

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

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. *Parasitology* **S99**, S93-S112.

Feinstein, S.C., Dana, S.L., McConlogue, L., Shooter, E.M. and Coffino, P. (1985) Nerve growth factor rapidly induces ornithine decarboxylase mRNA in PC12 rat pheochromocytoma cells. *Proc. Natl. Acad. Sci. USA* **82**, 5761-5766.

Felsenstein, J. (1989) PHYLIP- Phylogeny inference package, Version 3.2. *Cladistics* **5**, 164-166.

Fonzi, W.A. (1989) Biochemical and genetic charaterisation of the structure of yeast ornithine decarboxylase. *Biochem. Biophys. Res. Commun.* **162**, 1409-1416.

Foote, S.J. and Cowman, A.F. (1994) The mode of action and the mechanism of resistance to antimalarial drugs. *Acta Tropica* **56**, 157-171.

Francis, S.E., Sullivan, D.J. and Goldberg, D.E. (1997) Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annu. Rev. Microbiol.* **51**, 97-123.

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.

Fujita, K., Murakami, Y. and Hayashi, S. (1982) A macromolecular inhibitor of the antizyme to ornithine decarboxylase. *Biochem. J.* **204**, 647-652.

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.

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

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.

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

Glass, J.R. and Gerner, E.W. (1987) Spermidine mediates degradation of ornithine decarboxylase by a non-lysosomal, ubiquitin-independant mechanism. *J. Cell. Physiol.* **130**, 133-141.

Gubler, U. (1987) Second-strand cDNA synthesis: mRNA fragments as primers. *Methods Enzymol.* **152**, 330-335.

Gutteridge, W.E. and Trigg, P.I. (1972) Periodicity of nuclear DNA synthesis in the intraerythrocytic stage of *Plasmodium knowlesi*. *J. Protozool.* **19**, 378-381.

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

Hanahan, D., Jessee, J. and Bloom, F.R. (1991) Plasmid transformation of *Escherichia coli* and other bacteria. *Methods 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. (1989a) Multiple mechanisms for the regulation of mammalian ornithine decarboxylase In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 35-45

Hayashi, S. and Cannelakis, E. S. (1989b) Ornithine decarboxylase antizymes In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 47-57

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*. U. Bachrach. Boca Raton, CRC Press. **I:** 83-93.

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.

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

Hua, S.B., Li, X., Coffino, P. and Wang, C.C. (1995) Rat antizyme inhibits the activity but does not promote the degradation of mouse ornithine decarboxylase in *Trypanosoma brucei*. *J. Biol. Chem.* **28**, 10264-10271.

Hyde, J.E. (1996) Transformation of malaria parasites: the barriers come down. *Trends Microbiol.* **4**, 43-45.

Hyde, J.E. (1998) Antifolate drug resistance in the human malaria parasite *Plasmodium falciparum*. Biochemistry in Africa, Potchestroom, South Africa.

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.

Janne, J. and Alhonen-Hongisto, L. (1989). Inhibitors of ornithine decarboxylase: Biochemistry and applications. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 59-85.

Kanamoto, R., Utsunomiya, K., Kameji, T. and Hayashi, S. (1986) Effects of putrescine on synthesis and degradation of ornithine decarboxylase in primary cultured hepatocytes. *Eur. J. Biochem.* **154**, 539-544.

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

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

Kern, A., Oliveira, M.A., Chang, N., Ernst, S.R., Carroll, D.W., Momany, C., Minard, K., Coffino, P. and Hackert, M.L. (1996) Crystallisation of a mammalian ornithine decarboxylase. *Prot. Struc. Funct. Genet.* **24**, 266-268.

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

Kimura, E.A., Couto, A.S., Peres, V.J., Casal, O.L. and Katzin, A.M. (1996) N-linked glycoproteins are related to schizogony of the intraerythrocytic stage in *Plasmodium falciparum*. *J. Biol. Chem.* **271**, 14452-14461.

King, M. and King, E. (1992) In: *The story of medicine and disease in Malawi.* Blantyre, Malawi, Montford Press.

Kirby, G.C. (1997) Plants as a source of antimalarial drugs. *Trop. Doctor* **S1**, 7-11.

Klapper, M.H. (1977) The independent distribution of amino acid near neighbor pairs into polypeptides. *Biochem. Biophys. Res. Commun.* **78**, 1018-1024.

Koguchi, K., Murakami, Y. and Hayashi, S. (1996) Control of ornithine decarboxylase activity by polyamines and absence of antizyme in *Tetrahymena*. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **113**, 157-162.

Krishna, S., Cowan, G., Meade, J.C., Wells, R.A., Stringer, J.R. and Robson, K.J. (1993) A family of cation ATPase-like molecules from *Plasmodium falciparum*. *J. Cell Biol.* **120**, 385-398.

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

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

Liu, X. and Gorovsky, M.A. (1993) Mapping the 5' and 3' ends of *Tetrahymena thermophila* mRNA ends using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucl. Acids Res.* **21**, 4954-4960.

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

Louw, A.I. (1998) Molecular characterisation of facilitated transport proteins of the human malaria parasite. *S.A.J. Sci.* **94**, 257-277.

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.

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.

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.

Matsufuji, S., Mastufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J.F., Gesteland, R.F. and Hayashi, S. (1995) Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase. *Cell* **80**, 51-60.

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

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.

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

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

Mitchell, J.L.A., Hoff, J.A. and Bareyal-Layser, A. (1991) Stable ornithine decarboxylase in a rat hepatoma cell line selected for resistance to α-Difluoromethylornithine. *Arch. Biochem. Biophys.* **290**, 143-152.

Momany, C., Ernst, S., Ghosh, R., Chang, N. and Hachert, M.L. (1995) Crystallographic structure of a PLP-dependent ornithine decarboxylase from *Lactobacillus* 30a to 3.0A resolution. *J. Mol. Biol.* **252**, 643-655.

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

Murakami, Y., Masufuji, S., Kameji, T., Hayashi, S., Igarashi, K., Tamura, T., Tanaka, T. and Ichihara, A. (1992) Ornithine decarboxylase is degraded by the 26S proteasome without ubiquitination. *Nature* **360**, 597-599.

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

Oh, S.S., Christi, A.H., Palek, J. and Liu, S. (1997) Erythrocyte membrane alterations in *Plasmodium falciparum* malaria sequestration. *Curr. Opin. Hematol.* **4**, 148-154.

Oka, T. and Borellini, F. (1989). The role of ornithine decarboxylase and polyamine biosynthesis in cellular development and differentiation. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 7-19.

Old, R.W. and Primrose, S.B. (1994) Principles of gene manipulation. In: *Studies in Microbiology*. Oxford, Blackwell Science. Ltd.

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. and Trigg, P.I. (1995) Status of antimalarial drugs under development. *Bull. WHO* **73**, 565-571.

Opperdoes, F.R. (1995) Carbohydrate and Energy metabolism in aerobic protozoa. In: *Biochemistry and Molecular Biology of Parasites*. J. J. Marr and M. Muller. London, Academic Press Ltd.:19-32

Pegg, A.E. (1989). Characterisation of ornithine decarboxylase from various sources. In: *Ornithine Decarboxylase: Biology, Enzymology and Molecular Genetics*. S. Hayashi. Oxford, England, Pergamon Press, Inc. **Section 129:** 21-28.

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.

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.

Pietrokovski, S. (1996) Searching databases of conserved sequence regions by aligning protein multiple-alignments. *Nucl. Acids Res.* **24**, 3836-3845.

Poulin, R., Lu, L., Ackermann, B., Bey, P. and Pegg, A.E. (1992) Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by α-Difluoromethylornithine. *J. Biol. Chem.* **267**, 150-185.

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

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**.

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., Rieckmann, K.H., Genton, B., Lorry, K., Wines, B. and Cowman, A.F. (1996) Point mutations in the dihydrofolate reductase and dihydropteroate synthetase genes and *in vitro* susceptibility to pyrimethamine and cycloguanil of *Plasmodium falciparum* isolates from Papua New Guinea. *Am. J. Trop. Med. Hyg.* **55**, 209-213.

Robyt, J.F. and White, J.B. (1987) In: *Biochemical techniques.* Belmont, California, Brooks/Cole Publishing company.

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

Rom, E. and Kahana, C. (1994) Polyamines regulate the expression of ornithine decarboxylase antizyme *in vitro* by inducing ribosomal frame-shifting. *Proc. Natl. Acad. Sci.USA* **91**, 3959-3963.

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

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

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.

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

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

Seely, J.E. and Pegg, A.E. (1983) Effect of 1,3-diaminopropane on ornithine decarboxylase enzyme protein in thioacetamide-treated mice. *Biochem. J.* **216**, 701-707.

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

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.

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.

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.

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.

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

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, T. (1984) Polyamines. *Annu. Rev. Biochem.* **53**, 749-790.

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

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.

Tobias, K.E., Mamroud-Kidron, E. and Kahana, C. (1993b) Gly387 of murine ornithine decarboxylase is essential for the formation of stable homodimers. *Eur. J. Biochem.* **218**, 254-250.

Tokunaga, F., Goto, T., Koide, T., Murakami, Y., Hayashi, S., Tamura, T., Tanaka, K. and Ichihara, A. (1994) ATP- and Antizyme-dependant endoproteolysis of ornithine decarboxylase to oligopeptides by the 26S proteasome. *J. Biol. Chem.* **269**, 17382-17385.

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.

Vasques, J.R., Gooze, L., Kim, K., Gut, J., Petersen, C. and Nelson, R.G. (1996) Potential antifolate resistance determinants and genotypic variation in the bifunctional dihydrofolate reductase-thymidilate synthase gene from human and bovine isolates of *Cryptosporidium parvum*. *Mol. Biochem. Parasitol.* **79**, 153-165.

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.

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**.

Wang, P., Lee, C., Bayoumi, R., Djimde, A., Doumbo, O., Swedberg, G., Dao, L., Mshinda, H., Tanner, M., Watkins, W.M., Sims, P.F.G. and Hyde, J.E. (1997) Resistance to antifolates in *Plasmodium falciparum* monitored by sequence analysis of dihydropteroate synthetase and dihydrofolate reductase alleles in a large number of field samples of diverse origins. *Mol. Biochem. Parasitol.* **89**, 161-177.

Wellems, T.E. (1991) Molecular mechanisms of drug resistance in *Plasmodium falciparum* malaria. *Parasitology Today* **7**, 110-112.

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.

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

Zhou, M. and Sanchez, C.E. (1997). Universal TA cloning. In: *Gene cloning and analysis: current innovations*. B. Scheafer. Norfolk, U.K., Horizon Scientific Press: 165-178.

Appendix I. Cloning vectors.

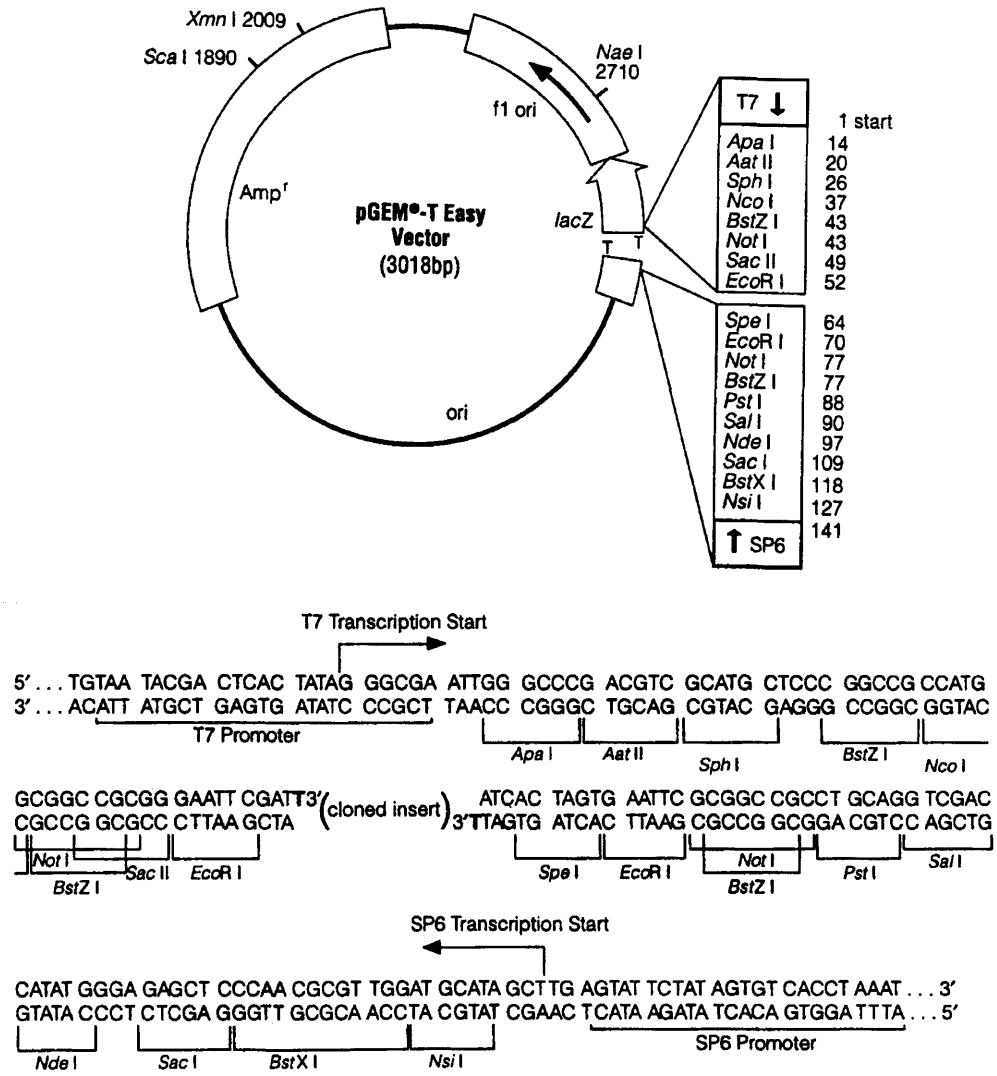

Figure Ia: Circular map of the pGEM-T Easy vector system used in A/T cloning strategies. The ampicillin resistance and *lacZ* genes are indicated as well as the multiple cloning site (MCS) flanked by the T7 and SP6 promotors. The positions of the T-overhangs are indicated in the MCS.

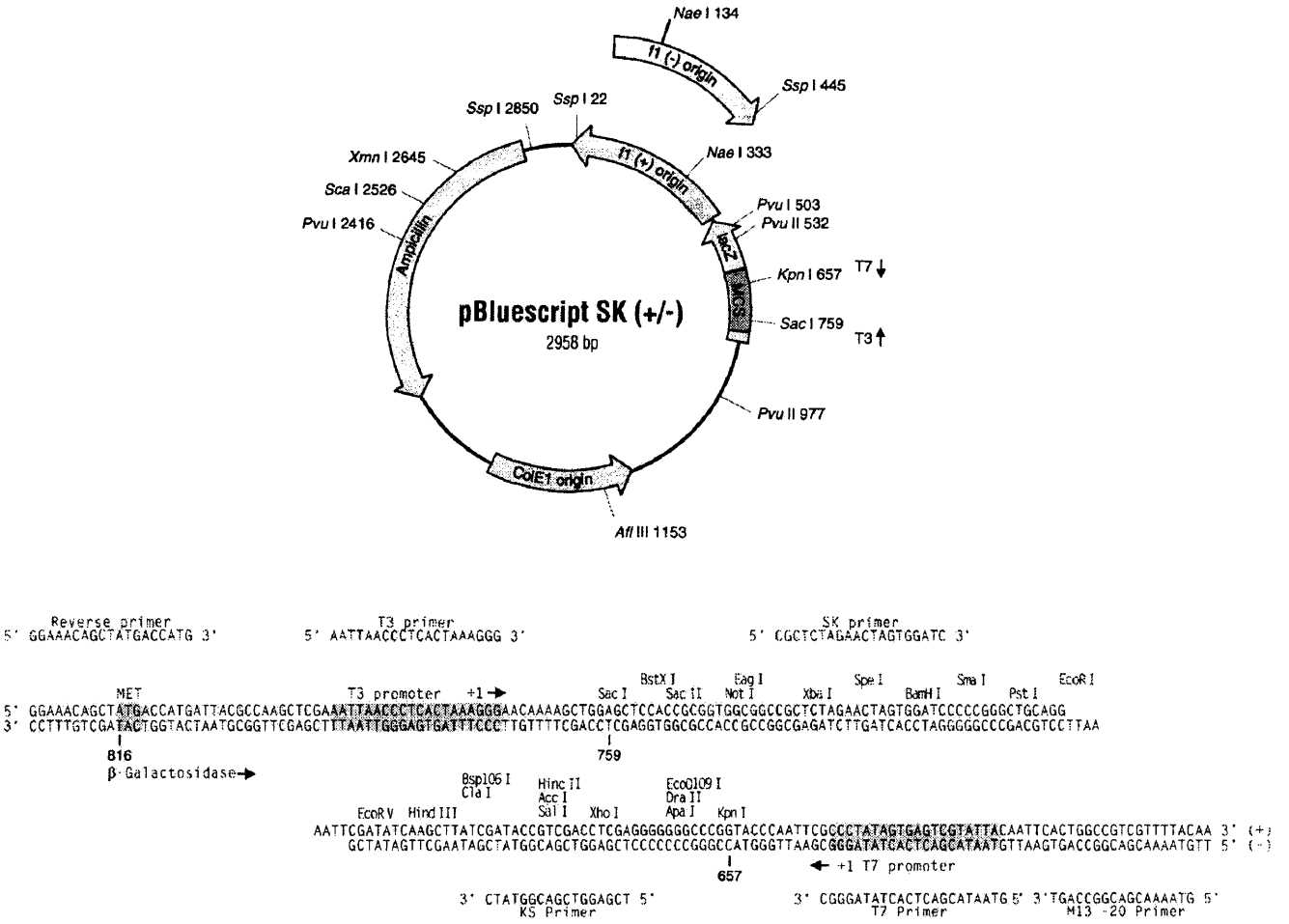

Figure Ib: Circular map of the pBluescript SK⁻ cloning vector. The multiple cloning site is indicated.

Appendix II. Overlap alignment of the full-length *P. falciparum* ODC nucleotide sequence.

```
          1          11         21         31         41         50
PFODC     GTGATCGATT TTTAGACTTT ATTCACAAGC AGTTGAATTT TTATAATGGA
5cons     ---------- ---------- ---------- ---------- ----------
ODCnucl   ---------- ---------- ---------- ---------- ----------

          51         61         71         81         91        100
PFODC     AAGTATATGT TCATGATAAA TTATGTATTC TGTGAGGAGA GTAACAACAT
5cons     ---------- ---------- ---------- -GTGAGGAGA GTAACAACAT
ODCnucl   ---------- ---------- ---------- -GTGAGGAGA GTAACAACAT

          101        111        121        131        141       150
PFODC     GTCTAAAATG GTACCTGATG ATGATAATAA TAATTATAGT AGTGGTAAAA
5cons     GTCTAAAATG GTACCTGATG ATGATAATAA TAATTATAGT AGTGGTAAAA
ODCnucl   GTCTAAAATG GTACCTGATG ATGATAATAA TAATTATAGT AGTGGTAAAA

          151        161        171        181        191       200
PFODC     GTTGCGTTTA TTATCAAGAT TTAAATAAGA AAGAAAAAGA AGAATATTAT
5cons     GTTGCGTTTA TTATCAAGAT TTAAATAAGA AAGAAAAAGA AGAATATTAT
ODCnucl   GTTGCGTTTA TTATCAAGAT TTAAATAAGA AAGAAAAAGA AGAATATTAT

          201        211        221        231        241       250
PFODC     CGCTTGAACA AAAAATTAAG AAACGACTTA TTTATTAATT CGAAACAATT
5cons     CGCTTGAACA AAAAATTAAG AAACGACTTA TTTATTAATT CGAAACAATT
ODCnucl   CGCTTGAACA AAAAATTAAG AAACGACTTA TTTATTAATT CGAAACAATT

          251        261        271        281        291       300
PFODC     TTATGAATTA CATACATTTA CCGAACGAAC GGTTGGATTT ATGAGAGTGC
5cons     TTATGAATTA CATACATTTA CCGAACGAAC GGTTGGATTT ATGAGAGTGC
ODCnucl   TTATGAATTA CATACATTTA CCGAACGAAC GGTTGGATTT ATGAGAGTGC

          301        311        321        331        341       350
PFODC     AATATTTTGT TTATAAATTA AGAGATGTTG TTAAATGTGT AGAAAAAGAA
5cons     AATATTTTGT TTATAAATTA AGAGATGTTG TTAAATGTGT AGAAAAAGAA
ODCnucl   AATATTTTGT TTATAAATTA AGAGATGTTG TTAAATGTGT AGAAAAAGAA

          351        361        371        381        391       400
PFODC     ACTTTGCTAG CTAGGAGTTC GTCTTGTTTA TTTATGTTTA ATAATATCAA
5cons     ACTTTGCTAG CTAGGAGTTC GTCTTGTTTA TTTATGTTTA ATAATATCAA
ODCnucl   ACTTTGCTAG CTAGGAGTTC GTCTTGTTTA TTTATGTTTA ATAATATCAA

          401        411        421        431        441       450
PFODC     ACGAAATGAC GTACATGATG ATTATGTAAC TAAGTCGTCA AATGGTGGTG
5cons     ACGAAATGAC GTACATGATG ATTATGTAAC TAAGTCGTCA AATGGTGGTG
ODCnucl   ACGAAATGAC GTACATGATG ATTATGTAAC TAAGTCGTCA AATGGTGGTG

          451        461        471        481        491       500
PFODC     TAATAAAACA ATTAACGGAA AGAGATGTTG ATGATATGTA TGAGTATGCT
5cons     TAATAAAACA ATTAACGGAA AGAGATGTTG ATGATATGTA TGAGTATGCT
ODCnucl   TAATAAAACA ATTAACGGAA AGAGATGTTG ATGATATGTA TGAGTATGCT

          501        511        521        531        541       550
PFODC     TTAAATTTTT GTAAACAAAA TAAAATAGTT GTTGTAGATA CTAATACCTT
5cons     TTAAATTTTT GTAAACAAAA TAAAATAGTT GTTGTAGATA CTAATACCTT
```

```
ODCnucl   TTAAATTTTT GTAAACAAAA TAAAATAGTT GTTGTAGATA CTAATACCTT

          551         561         571         581         591         600
  PFODC   TTTTTTTGAT GCATCTAAAA GAAAGGAGAA CTTAATAAAA CTTGAAAAGG
  5cons   TTTTTTTGAT GCATCTAAAA GAAAGGAGAA CTTAATAAAA CTTGAAAAGG
ODCnucl   TTTTTTTGAT GCATCTAAAA GAAAGGAGAA CTTAATAAAA CTTGAAAAGG

          601         611         621         631         641         650
  PFODC   TACAAACAAA TGAGAAAGAT GAATATGAAG AAAAAGATGA AGTGTATCGA
  5cons   TACAAACAAA TGAGAAAGAT GAATATGAAG AAAAAGATGA AGTGTATCGA
ODCnucl   TACAAACAAA TGAGAAAGAT GAATATGAAG AAAAAGATGA AGTGTATCGA

          651         661         671         681         691         700
  PFODC   AGGGGTAATA ATGAATTGAG TTCGTTGGAT CATTTAGATA GTAAGAATAA
  5cons   AGGGGTAATA ATGAATTGAG TTCGTTGGAT CATTTAGATA GTAAGAATAA
ODCnucl   AGGGGTAATA ATGAATTGAG TTCGTTGGAT CATTTAGATA GTAAGAATAA

          701         711         721         731         741         750
  PFODC   TTTGATTCAT ATGTATTATG AAAAGAACAA ATGTGATATC ATAAATAAGG
  5cons   TTTGATTCAT ATGTATTATG AAAAGAACAA ATGTGATATC ATAAATAAGG
ODCnucl   TTTGATTCAT ATGTATTATG AAAAGAACAA ATGTGATATC ATAAATAAGG

          751         761         771         781         791         800
  PFODC   ATGATGAGAA TTCAACGATA GCGACGAATA ATAATGATAA TAATAATGAT
  5cons   ATGATGAGAA TTCAACGATA GCGACGAATA ATAATGATAA TAATAATGAT
ODCnucl   ATGATGAGAA TTCAACGATA GCGACGAATA ATAATGATAA TAATAATGAT

          801         811         821         831         841         850
  PFODC   AGTAGTTCTT ATGACAAAAG TATAACGATC AGCAGAAGCA GTAGCTGTAA
  5cons   AGTAGTTCTT ATGACAAAAG TATAACGATC AGCAGAAGCA GTAGCTGTAA
ODCnucl   AGTAGTTCTT ATGACAAAAG TATAACGATC AGCAGAAGCA GTAGCTGTAA

          851         861         871         881         891         900
  PFODC   TAATAGCCAT TTGAGTTATA GTAGTTTTGA TAATAATCAT GGAAATGAAA
  5cons   TAATAGCCAT TTGAGTTATA GTAGTTTTGA TAATAATCAT GGAAATGAAA
ODCnucl   TAATAGCCAT TTGAGTTATA GTAGTTTTGA TAATAATCAT GGAAATGAAA

          901         911         921         931         941         950
  PFODC   AAATGAAAGA TTATATAAGT GTTGATGAAA ATAATAATAA TAATAATAAT
  5cons   AAATGAAAGA TTATATAAGT GTTGATGAAA ATAATAATAA TAATAATAAT
ODCnucl   AAATGAAAGA TTATATAAGT GTTGATGAAA ATAATAATAA TAATAATAAT

          951         961         971         981         991        1000
  PFODC   AATAAAAATA ATAATGTATT GTTAACTTTA CAAAGGAACA GTGATGATGA
  5cons   AATAAAAATA ATAATGTATT GTTAACTTTA CAAAGGAACA GTGATGATGA
ODCnucl   AATAAAAATA ATAATGTATT GTTAACTTTA CAAAGGAACA GTGATGATGA

          1001        1011        1021        1031        1041        1050
  PFODC   AAATGGTAAA GATAAAGATA ATGAAAAAAA TGACGTAAGT TTAGAAAACA
  5cons   AAATGGTAAA GATAAAGATA ATGAAAAAAA TGACGTAAGT TTAGAAAACA
ODCnucl   AAATGGTAAA GATAAAGATA ATGAAAAAAA TGACGTAAGT TTAGAAAACA

          1051        1061        1071        1081        1091        1100
  PFODC   ATATGGAAAA GAATTATAAA GAAGAAATAT GGAATTATTA TACAAAAAAT
  5cons   ATATGGAAAA GAATTATAAA GAAGAAATAT GGAATTATTA TACAAAAAAT
ODCnucl   ATATGGAAAA GAATTATAAA GAAGAAATAT GGAATTATTA TACAAAAAAT
```

```
          1101        1111        1121        1131        1141        1150
  PFODC   AAAGTGGAAG TAAAAACATT AGAAAAAGTA TTAAATGAAA ATATAGATAC
  5cons   AAAGTGGAAG TAAAAACATT AGAAAAAGTA TTAAATGAAA ATATAGATAC
ODCnucl   AAAGTGGAAG TAAAAACATT AGAAAAAGTA TTAAATGAAA ATATAGATAC

          1151        1161        1171        1181        1191        1200
  PFODC   ATCAGTAGTT TGTATAAATT TACAGAAAAT ATTAGCTCAG TATGTTAGAT
  5cons   ATCAGTAGTT TGTATAAATT TACAGAAAAT ATTAGCTCAG TATGTTAGAT
ODCnucl   ATCAGTAGTT TGTATAAATT TACAGAAAAT ATTAGCTCAG TATGTTAGAT

          1201        1211        1221        1231        1241        1250
  PFODC   TTAAAAAGAA TCTTCCACAT GTTACTCCAT TCTATTCTGT AAAAAGTAAT
  5cons   TTAAAAAGAA TCTTCCACAT GTTACTCCAT TCTATTCTGT AAAAAGTAAT
ODCnucl   TTAAAAAGAA TCTTCCACAT GTTACTCCAT TCTATTCTGT AAAAAGTAAT

          1251        1261        1271        1281        1291        1300
  PFODC   AATGATGAAG TTGTAATCAA ATTTTTATAT GGATTGAATT GTAATTTTGA
  5cons   AATGATGAAG TTGTAATCAA ATTTTTATAT GGATTGAATT GTAATTTTGA
ODCnucl   AATGATGAAG TTGTAATCAA ATTTTTATAT GGATTGAATT GTAATTTTGA

          1301        1311        1321        1331        1341        1350
  PFODC   TTGCGCTTCG ATAGGTGAAA TAAGTAAAGT AATAAAATTA TTACCAAATT
  5cons   TTGCGCTTCG ATAGGTGAAA TAAGTAAAGT AATAAAATTA TTACCAAATT
ODCnucl   TTGCGCTTCG ATAGGTGAAA TAAGTAAAGT AATAAAATTA TTACCAAATT

          1351        1361        1371        1381        1391        1400
  PFODC   TATCAAGAGA TAGAATAATT TTTGCGAATA CAATTAAAAG TATTAATTCT
  5cons   TATCAAGAGA TAGAATAATT TTTGCGAATA CAATTAAAAG TATTAATTCT
ODCnucl   TATCAAGAGA TAGAATAATT TTTGCGAATA CAATTAAAAG TATTAATTCT

          1401        1411        1421        1431        1441        1450
  PFODC   TTAATATATG CAAGAAAGGA AAATATTAAT TTATGTACTT TTGATAATTT
  5cons   TTAATATATG CAAGAAAGGA AAATATTAAT TTATGTACTT TTGATAATTT
ODCnucl   TTAATATATG CAAGAAAGGA AAATATTAAT TTATGTACTT TTGATAATTT

          1451        1461        1471        1481        1491        1500
  PFODC   AGATGAATTA AAAAAAATAT ATAAATATCA TCCGAAATGT TCTTTAATAT
  5cons   AGATGAATTA AAAAAAATAT ATAAATATCA TCCGAAATGT TCTTTAATAT
ODCnucl   AGATGAATTA AAAAAAATAT ATAAATATCA TCCGAAATGT TCTTTAATAT

          1501        1511        1521        1531        1541        1550
  PFODC   TACGTATTAA TGTAGATTTT AAAAATTACA AATCTTATAT GTCTTCAAAA
  5cons   TACGTATTAA TGTAGATTTT AAAAATTACA AATCTTATAT GTCTTCAAAA
ODCnucl   TACGTATTAA TGTAGATTTT AAAAATTACA AATCTTATAT GTCTTCAAAA

          1551        1561        1571        1581        1591        1600
  PFODC   TATGGAGCTA ATGAATATGA ATGGGAAGAA ATGTTATTGT ATGCAAAAAA
  5cons   TATGGAGCTA ATGAATATGA ATGGGAAGAA ATGTTATTGT ATGCAAAAAA
ODCnucl   TATGGAGCTA ATGAATATGA ATGGGAAGAA ATGTTATTGT ATGCAAAAAA

          1601        1611        1621        1631        1641        1650
  PFODC   ACATAATCTA AATATTGTAG GTGTATCATT TCATGTTGG TAGTAATACA
  3cons   ---------- -----TTTAG GAGTTACGTT TCACGTTGG TAGTAATACA
  5cons   ACATAATCTA AATATTGTAG GTGTATCATT TCATGTTGG TAGTAATACA
ODCnucl   ACATAATCTA AATATTGTAG GTGTATCATT TCATGTTGG TAGTAATACA
```

```
          1651        1661        1671        1681        1691        1700
  PFODC   AAGAATTTAT TTGATTTCTG TCTAGCCATT AAATTATGTA GAGATGTATT
  3cons   AAGAATTTAT TTGATTTCTG TCTAGCCATT AAATTATGTA GAGATGTATT
  5cons   AAGAATTTAT TTGATTTCTG T
ODCnucl   AAGAATTTAT TTGATTTCTG TCTAGCCATT AAATTATGTA GAGATGTATT

          1701        1711        1721        1731        1741        1750
  PFODC   CGATATGAGT AGTAATATGG TATTTAATTT TTATATAATA AATTTAGGAG
  3cons   CGATATGAGT AGTAATATGG GATTTAATTT TTATATAATA AATTTAGGAG
ODCnucl   CGATATGAGT AGTAATATGG GATTTAATTT TTATATAATA AATTTAGGAG

          1751        1761        1771        1781        1791        1800
  PFODC   GGGTATATCC AGAAGAATTA GAATATGATA ATGTAAAGAA ACATGATAAA
  3cons   GGGGATATCC AGAAGAATTA GAATATGATA ATGCAAAGAA ACATGATAAA
ODCnucl   GGGGATATCC AGAAGAATTA GAATATGATA ATGCAAAGAA ACATGATAAA

          1801        1811        1821        1831        1841        1850
  PFODC   ATTCATTATT GTACTTTAAG TCTTCAAGAA ATTAAAAAAG ATATACAAAA
  3cons   ATTCATTATT GTACTTTAAG TCTTCAAGAA ATTAAAAAAG ATATACAAAA
ODCnucl   ATTCATTATT GTACTTTAAG TCTTCAAGAA ATTAAAAAAG ATATACAAAA

          1851        1861        1871        1881        1891        1900
  PFODC   ATTTCTTAAT GAAGAAACAT TTCTCAAGAC GAAATATGGA TACTATAGTT
  3cons   ATTTCTTAAT GAAGAAACAT TTCTCAAGAC GAAATATGGA TACTATAGTT
ODCnucl   ATTTCTTAAT GAAGAAACAT TTCTCAAGAC GAAATATGGA TACTATAGTT

          1901        1911        1921        1931        1941        1950
  PFODC   TTGAAAAAAT ATCATTGGCT ATTAATATGT CAATCGATCA TTATTTTAGT
  3cons   TTGAAAAAAT ATCATTGGCT ATTAATATGT CAATCGATCA TTATTTTAGT
ODCnucl   TTGAAAAAAT ATCATTGGCT ATTAATATGT CAATCGATCA TTATTTTAGT

          1951        1961        1971        1981        1991        2000
  PFODC   CATATGAAAG ATAATCTAAG AGTTATTTGT GAACCTGGTA GCTATATGGT
  3cons   CATATGAAAG ATAATCTAAG AGTTATTTGT GAACCTGGTA GATATATGGT
ODCnucl   CATATGAAAG ATAATCTAAG AGTTATTTGT GAACCTGGTA GATATATGGT

          2001        2011        2021        2031        2041        2050
  PFODC   CGCTGCTTCG TCAACATTAG CTGTTAAAAT TATAGGAAAG AGACGTCCAA
  3cons   CGCTGCTTCG TCAACATTAG CTGTTAAAAT TATAGGAAAG AGACGTCCAA
ODCnucl   CGCTGCTTCG TCAACATTAG CTGTTAAAAT TATAGGAAAG AGACGTCCAA

          2051        2061        2071        2081        2091        2100
  PFODC   CTTTTCAGGG CATTATGTTA AAAGATTTAA AAGCCCATTA CGGTCCTTTA
  3cons   CTTTTCAGGG CATTATGTTA AAAGATTTAA AAGCCCATTA CGATCCTTTA
ODCnucl   CTTTTCAGGG CATTATGTTA AAAGATTTAA AAGCCCATTA CGATCCTTTA

          2101        2111        2121        2131        2141        2150
  PFODC   AATTTTGCTC AACAAGAAAA TAAGAAACAA GACGAAACAA AAATAAACCA
  3cons   AATTTTGCTC AACAAGAAAA TAAGAAACAA GACGAACCAA AAATAAACCA
ODCnucl   AATTTTGCTC AACAAGAAAA TAAGAAACAA GACGAACCAA AAATAAACCA

          2151        2161        2171        2181        2191        2200
  PFODC   CAATAATGAT AATAATGATA ATAAT-G-A- TAATAATGAT AATAATATTA
  3cons   CAATAATGAT AATAATGATA ATAATAGTAA TGATAATAAT AATAATAATA
ODCnucl   CAATAATGAT AATAATGATA ATAATAGTAA TGATAATAAT AATAATAATA
```

```
          2201        2211        2221        2231        2241        2250
  PFODC   ATAATAATAA TAATAATCAA AAAGGGGGCC AAGGAAATAT TATGAATGAT
  3cons   ATAATAATAA TAATAATCAA AAAGGGGGCC AAGGAAATAT TATGAATGAT
ODCnucl   ATAATAATAA TAATAATCAA AAAGGGGGCC AAGGAAATAT TATGAATGAT

          2251        2261        2271        2281        2291        2300
  PFODC   CTAATAATAA CTAGCACAAA TGATTCTACT AGTAAAAAGT ATGATCATTC
  3cons   CTAATAATAA CTAGCACAAA TGATTCTACT AATAAAAAGA ATGATCATTC
ODCnucl   CTAATAATAA CTAGCACAAA TGATTCTACT AATAAAAAGA ATGATCATTC

          2301        2311        2321        2331        2341        2350
  PFODC   TTCTAGAAAA GTTATTCAAA ATGTATCGTG CACAATACGT GATAAAGAAG
  3cons   TTCTAGTCAA GTTATTCAAA ATGTATCGTG CACAATACGT GATAAAGAAG
ODCnucl   TTCTAGTCAA GTTATTCAAA ATGTATCGTG CACAATACGT GATAAAGAAG

          2351        2361        2371        2381        2391        2400
  PFODC   TAGATAATAT TAAAATAAAT ACACATACCA TAAATAATCC TAATATAAAT
  3cons   GAGATAATAT TAAAATAAAT ACACATACCA TAAATAATCC TAATATAAAT
ODCnucl   GAGATAATAT TAAAATAAAT ACACATACCA TAAATAATCC TAATATAAAT

          2401        2411        2421        2431        2441        2450
  PFODC   GGAAAAGAAA ATACCGTGGA TGGTGATAAT ATTAATATTG CTCATAAAAA
  3cons   GGAAAAGAAA ATACCGTGGA TGGTGATAAT ATTAATATTG CTCATAAAAA
ODCnucl   GGAAAAGAAA ATACCGTGGA TGGTGATAAT ATTAATATTG CTCATAAAAA

          2451        2461        2471        2481        2491        2500
  PFODC   TATTGGTAAT AACTTTAGTA GTAGTAATTC AAAATTAGGC AACATAACAA
  3cons   TATTGGTAAT AACTTTAGTA GTAGTAATTC AAAATTAGGC AACATAACAA
ODCnucl   TATTGGTAAT AACTTTAGTA GTAGTAATTC AAAATTAGGC AACATAACAA

          2501        2511        2521        2531        2541        2550
  PFODC   ATATTAAGAA AAAAGTTGTT AATATTAATG ACAATAGATA TAATTATTTC
  3cons   ATATTAAGAA AAAAGTTGTT AATATTAATG ACAATAGATA TAATTATTTC
ODCnucl   ATATTAAGAA AAAAGTTGTT AATATTAATG ACAATAGATA TAATTATTTC

          2551        2561        2571        2581        2591        2600
  PFODC   TCATATTATG TAAGCGATAG TATATATGGT TGTTTTAGTG GTATAATTTT
  3cons   TCATATTATG TAAGCGATAG TATATATGGT TGTTTTAGTG GTATAATTTT
ODCnucl   TCATATTATG TAAGCGATAG TATATATGGT TGTTTTAGTG GTATAATTTT

          2601        2611        2621        2631        2641        2650
  PFODC   TGATGAATAC AATAGATGTC CTATTTATGT TATTAAAAAC AAAAATAACC
  3cons   TGATGAATAC AATAGATGTC CTATTTATGT TATTAAAAAC AAAAATAACC
ODCnucl   TGATGAATAC AATAGATGTC CTATTTATGT TATTAAAAAC AAAAATAACC

          2651        2661        2671        2681        2691        2700
  PFODC   CTAATCAAAA TTTTATGAAT TTTAATTTGT ATTTAGCTAA TGTATTTGGA
  3cons   CTAATCAAAA TTTTATGAAT TTTAATTTGT ATTTAGCTAA TGTATTTGGA
ODCnucl   CTAATCAAAA TTTTATGAAT TTTAATTTGT ATTTAGCTAA TGTATTTGGA

          2701        2711        2721        2731        2741        2750
  PFODC   CAATCATGTG ATGGCTTGGA TATGATCAAT TCTATTACGT ACTTACCTGA
  3cons   CAATCATGTG ATGGCTTGGA TATGATCAAT TCTATTACGT ACTTACCTGA
ODCnucl   CAATCATGTG ATGGCTTGGA TATGATCAAT TCTATTACGT ACTTACCTGA
```

```
          2751        2761        2771        2781        2791        2800
  PFODC   GTGTTATATT AATGATTGGC TTCTCTATGA ATATGCTGGG GCATACACTT
  3cons   GTGTTATATT AATGATTGGC TTCTCTATGA ATATGCTGGG GCATACACTT
ODCnucl   GTGTTATATT AATGATTGGC TTCTCTATGA ATATGCTGGG GCATACACTT

          2801        2811        2821        2831        2841        2850
  PFODC   TTGTCAGCTC ATCAAACTTT AATGGATTTA AGAAATGCAA GAAGGTGTAT
  3cons   TTGTCAGCTC ATCAAACTTT AATGGATTTA AGAAATGCAA GAAGGTGTAT
ODCnucl   TTGTCAGCTC ATCAAACTTT AATGGATTTA AGAAATGCAA GAAGGTGTAT

          2851        2861        2871        2881        2891        2900
  PFODC   ATATTCCCTG AATCGAAACC TTCCCTTAAG GGGCAACCAA ACAAACATTG
  3cons   ATATTCCCTG AATCGAAACC TTCCCTTAAG GGGCAACCAA ACAAACATTG
ODCnucl   ATATTCCCTG AATCGAAACC TTCCCTTAAG GGGCAACCAA ACAAACATTG

          2901        2911        2921        2931        2941        2950
  PFODC   GTAAATAACA AAATCGAAGA AAAAGGAAAT AAATAGGGAA AAAAAAAAAA
  3cons   GTAAATAACA AAATCGAAGA AAAAGGAAAT AAATAGGGAA AAAAAAAAAA
ODCnucl   GTAAATAACA AAATCGAAGA AAAAGGAAAT AAATAGGGAA AAAAAAAAAA

          2951        2961
  PFODC   AAAAAAAAAAA
  3cons   AAAAAA-----GAGT GTTGTGGTAA TGATAGA
ODCnucl   AAAAAA
```

Figure II: Nucleotide alignments of the full-length *P. falciparum* ODC cDNA. The consensus sequence is indicated by 'ODCnucl' and the Genbank sequence by the 'PFODC'. '5cons' indicates the 5′-fragment and '3cons' the 3′-fragment. The colour code for the primers are as follows: OCRF1=purple, ODCF2=pink, ODCR3=light green, GSP1=red, ODCR1=light blue, ODCR2=dark blue and 3'Anchor=burgundy. The start (ATG) and stop (TAA) codons are boxed. The polyadenylation signal is indicated in the shadowed box.

The relevance of the probable phosphorylation sites should be investigated, especially the effect of phosphorylation/dephosphorylation on the enzyme activity (Brown *et al.*, 1994). The lack of the important PEST-rich region in the sequence indicate that the *P. falciparum* enzyme might be more stable than the mammalian proteins (Assaraf *et al.*, 1988). Since the mammalian counterpart of this enzyme is one of the most interestingly regulated enzymes known, it would be exciting to determine the extent of regulation of this enzyme in the malaria parasite. Furthermore, the presence and function of the extended N-terminal of the *P. falciparum* ODC protein must be investigated. To this end, creation of truncated deletion mutants of the gene (with specific restriction enzymes or exonucleases) and expression of the smaller proteins would provide information on the functional role of the extra ~273 residues. This same principle can be applied to investigate the role of the internal insertions present in the *P. falciparum* ODC protein. It is possible that these malaria-specific insertions could be targeted to inhibit the enzyme if they are essential to catalytic activity. Alternatively, the possibility also exists that these areas are involved in the dimerisation of the protein as the identified helix-loop-helix dimerisation domain (in the N-terminal extension) may play an important role in the stabilisation of the dimeric protein. If dimerisation can be prevented, it could result in the accumulation of inactive monomers and a decrease in the number of catalytically active homodimers. It is therefore conceivable that these structural differences in the parasite enzyme could be exploited in the design and development of new selective drugs.

After biochemical characterisation of the *P. falciparum* ODC protein in terms of structure-function relationships, the ultimate focus will be the validation of the protein as a possible antimalarial target. In this context, knockout mutants can be created and the effect on parasite survival determined (Wang, 1997). The expected decreased survival should be reversible by gene complementation experiments. The technology of transformation of the malaria parasite has been described in the last 5 years (Hyde, 1996). The application of antisense oligonucleotides for the purpose of inhibition of gene transcription or mRNA translation in *P. falciparum* is gaining popularity and is under investigation in our laboratory (D. Schulze, personal communication). These techniques could facilitate the specific inhibition of ODC protein expression and will allow the further *in vivo* investigation of expression of the ODC gene and the importance of the translated protein to the parasite.

Malaria will remain a problem to most of the tropical areas of the world unless the biochemistry and molecular biology of the parasite are better understood. The isolation of the *P. falciparum* ODC cDNA promises a more comprehensive understanding of polyamine metabolism of the parasite and, ultimately, the validation of ODC as an antimalarial target. The methods established in this dissertation could hopefully simplify the search for other putative antimalarial targets.