CHAPTER 3

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

3.1) INTRODUCTION.

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

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

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

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

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

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

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

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

Multiwavelength stopped-flow spectroscopy results led to the proposal of the reaction mechanism of ornithine decarboxylation by *T. brucei* ODC (Fig. 3.1) (Brooks and Phillips, 1997; Osterman, *et al.*, 1999). The first step in catalysis is the linking of PLP to Lys_{69} via a Schiff-base and the formation of the external aldimine (second Schiff-base) with ornithine. Decarboxylation then occurs via a quinoid intermediate followed by protonation at the Cα to form putrescine and transimination to generate free ODC and product (Brooks and Phillips, 1997; Osterman, *et al.*, 1999).

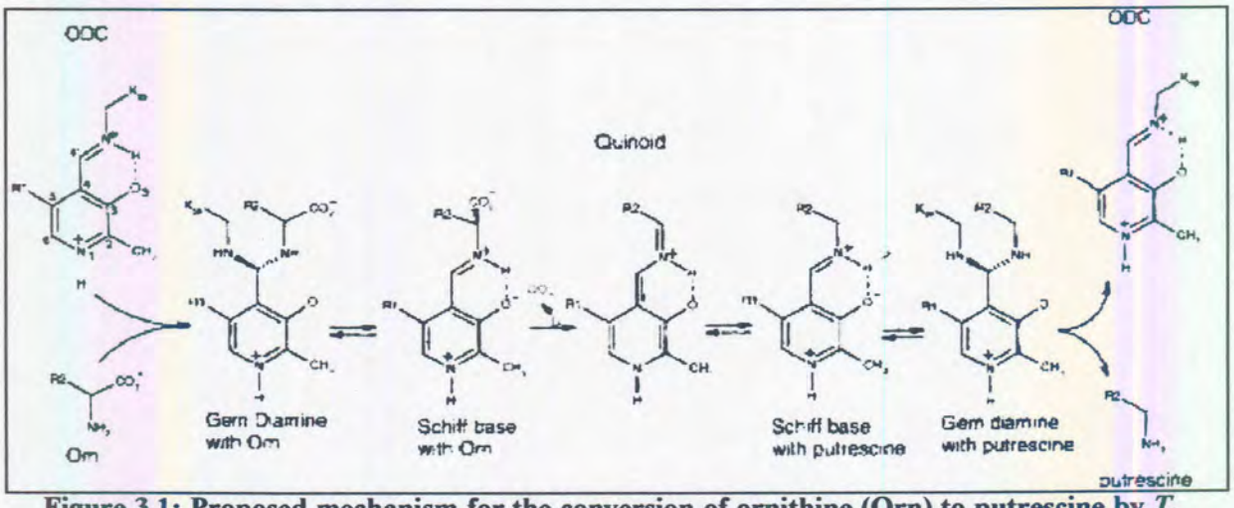

Figure 3.1: Proposed mechanism for the conversion of ornithine (Orn) to putrescine by *T. brucei* ODC. R1 is $CH_2OPO_3^{2-}$ and R2 is $(CH_2)_3NH_3^+$. Adapted from (Brooks and Phillips, 1997).

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

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

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

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

The decarboxylation reaction starts with Schiff-base formation between the pyruvoyl-cofactor and the α-amino group of the substrate, AdoMet (Fig. 3.2 B). The resulting electron sink facilitates removal of the α-carboxylate group from the substrate, which is replaced with a proton. Finally, the Schiff-base is hydrolysed to yield the decarboxylated product (Ekstrom, *et al.*, 1999).

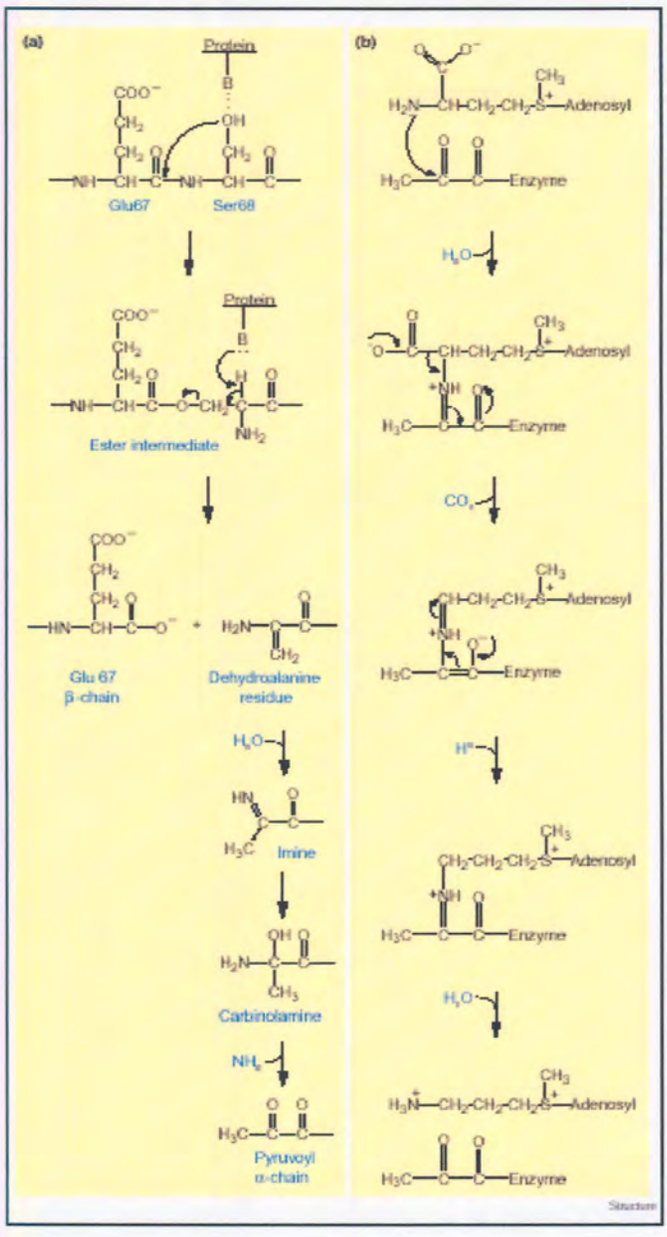

Figure 3.2: Proposed reaction mechanism for the autocatalytic intramolecular activation of AdoMetDC (A) and the decarboxylation of AdoMet (B). Adapted from (Ekstrom, *et al.*, 1999; Ekstrom, *et al.*, 2001).

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

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

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

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

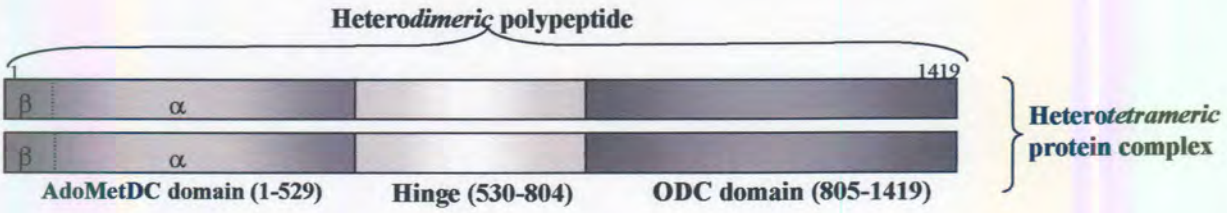

Figure 3.3: Schematic organisation of the bifunctional AdoMetDC/ODC from *P. falciparum*. AdoMetDC chains are indicated with α and β. Adapted from (Wrenger, *et al.*, 2001)

3.1.4) Recombinant protein expression and analyses.

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

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

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

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

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

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

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

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

3.2) MATERIALS AND METHODS.

3.2.1) Recombinant expression of His-Tag fusion proteins.

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

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

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

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

plasmid containing the full-length PfODC of 2918 bp as template (250 ng/µl, Chapter 2), 5 pmol of each primer, 1x Takara ExTaq reaction buffer (Takara Shuzo, Japan), 0.2 mM of each dNTP, 2 mM $MgCl_2$ and 1.25 U Takara ExTaq DNA Polymerase (Takara Shuzo, Japan) in 0.2 ml thin-walled tubes in a Perkin Elmer GeneAmp PCR system 9700 (PE Applied Biosystems, California, USA). Cycling parameters consisted of denaturing at 94°C for 3 min, a 2 min pause at 80°C for the hot-start addition of polymerase followed by 30 cycles at 94°C for 30 sec, 57°C for 30 sec and 60°C for 3 min. The PCR products were analysed on a agarose gel, purified after electrophoresis, cloned into pGem-T Easy vector and the nucleotide sequence determined as described in Chapter 2, section 2.2.7-11.

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

The AdoMetDC domain was amplified with the Sampetf1 and Sampetr1 primer pair (Table 3.1) based on the published *PfAdometdc/Odc* sequence (Genbank accession number AF0934833). The 50 µl reactions contained 1×reaction buffer, 2 mM $MgCl_2$, 0.2 mM dGTP,dCTP/0.3 mM dATP,dTTP, 10 pmol of each primer and 2.5 U Takara ExTaq DNA polymerase (Takara Shuzo, Japan) added in a hot-start protocol using 25 ng total RNA:cDNA equivalents of single-stranded cDNA as template. PCR was performed for 30 cycles at 94°C for 30 sec, and 65°C for 2.5 min. The PCR product was purified from the agarose gel and cloned into pGEM-T Easy as described in sections 2.2.7-2.2.11. The AdoMetDC insert was cut from the cloning vector with *Bam*HI and

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

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

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

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

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

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

3.2.2) Recombinant expression of Strep-Tag fusion proteins.

The *P. falciparum* monofunctional AdoMetDC and ODC domains as well as the bifunctional PfAdoMetDC/ODC protein were expressed as Strep-Tag fusion proteins as described in (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). The engineered Strep-Tag (NH_2-WSHPQFEK-COOH) with highly selective binding properties for derivatised streptavidin coupled to Sepharose called Strep-Tactin can be fused to the N- or C-terminal ends of proteins (Institut für Bioanalytik, Göttingen, Germany). Proteins expression from any pASK-IBA vector occurs under tight control of the tetA promotor/operator. The tet repressor is encoded on the expression vector and is constitutively expressed under the β-lactamase promotor. Protein expression is induced with low levels of anhydrotetracycline. The tetA promotor is not leaky and not functionally coupled to any cellular regulation mechanisms or genetic background. Recombinantly expressed Strep-Tag fusion proteins are then isolated with affinity chromatography on a Strep-Tactin derivatised Sepharose column under physiological conditions using any physiological buffer. Protein elution occurs in the presence of desthiobiotin (Sigma), a competitive biotin analogue with a high affinity to Strep-Tactin. Reversibly bound desthiobiotin is removed with 4-hydroxy azobenzene-2-carboxylic acid (HABA, Sigma) to regenerate the column. The Strep-Tag does not need to be removed from the proteins since there is no interference with folding or activity as it has no ion-exchange properties and does not induce protein aggregation.

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

Monofunctional ODC expressed as a Strep-tag fusion protein was further purified with SE-HPLC. The proteins from the Strep-Tactin affinity column in buffer W (100 mM Tris-HCl, pH 8.0, 1 mM EDTA) was filtered through a 0.22 µm filter (Millex GV4, Millipore Corporation, USA) and applied to a size-exclusion column ($G2000W_{XL}$, 7.8 mm × 30 cm, 100 kDa exclusion size, TosoHaas, USA). Buffers were filtered through a 0,22 µm filter and degassed for 30 min prior to application to the column. The column was equilibrated using isocratic conditions in buffer W at a flow speed of 1 ml/min and

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

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

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

3.2.5) Quantitation of proteins (Bradford, 1976).

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

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

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

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

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

3.2.6.1) Staining of SDS-PAGE gels.

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

3.2.7) AdoMetDC and ODC enzyme activity assays.

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

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

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

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

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

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

3.3) RESULTS.

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

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

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

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

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

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

Fig. 3.4 indicates the different constructs for the various expression systems used. The 1419 residue bifunctional PfAdoMetDC/ODC was expressed in pASK-IBA3 vector with a C-terminal Strep-tag, monofunctional AdoMetDC (660 residues) and ODC (758 residues) in pASK-IBA7 with N-terminal Strep-tags as described in (Krause, *et al.*, 2000; Müller, *et al.*, 2000; Wrenger, *et al.*, 2001). His-tag based pET-15b vectors were used for expression of the 596 residue AdoMetDC and 614 residue ODC monofunctional proteins with N-terminal His-tags.

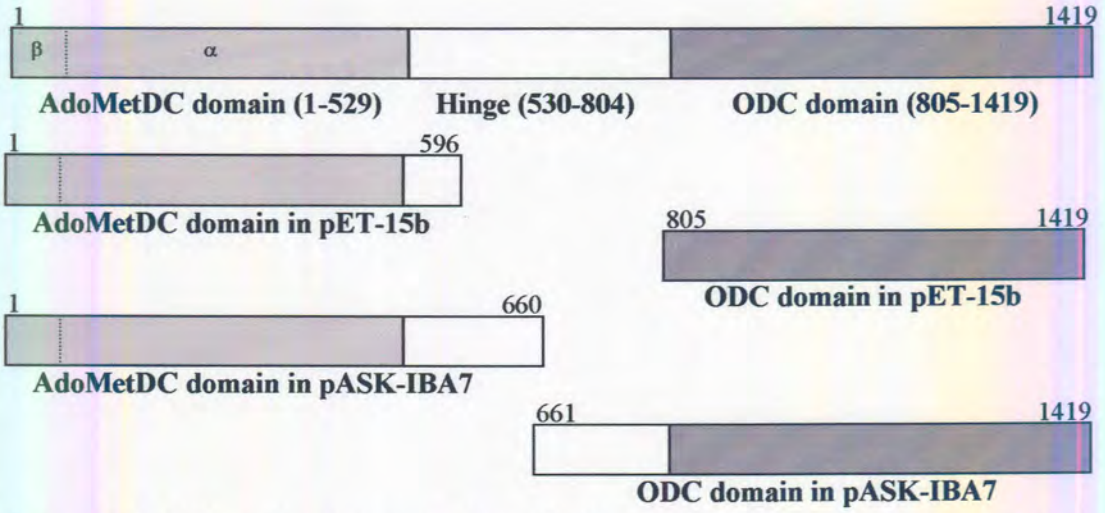

Figure 3.4: Schematic representation of the cloning strategy for expression of monofunctional AdoMetDC and ODC or bifunctional PfAdoMetDC/ODC. The top panel indicates the bifunctional PfAdoMetDC/ODC with the sizes of the domains given in number of amino acids. Monofunctional proteins were expressed as His-tag fusion proteins (pET-15b), or as Strep-tagged proteins (pASK-IBA7) with the residue sizes given. The α- and β-chains of AdoMetDC are separated with a dashed line. Parts of the hinge region (empty box) were included in some of the constructs.

3.3.3) Recombinant expression of monofunctional AdoMetDC and ODC domains.

3.3.3.1) His-tag expression system.

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

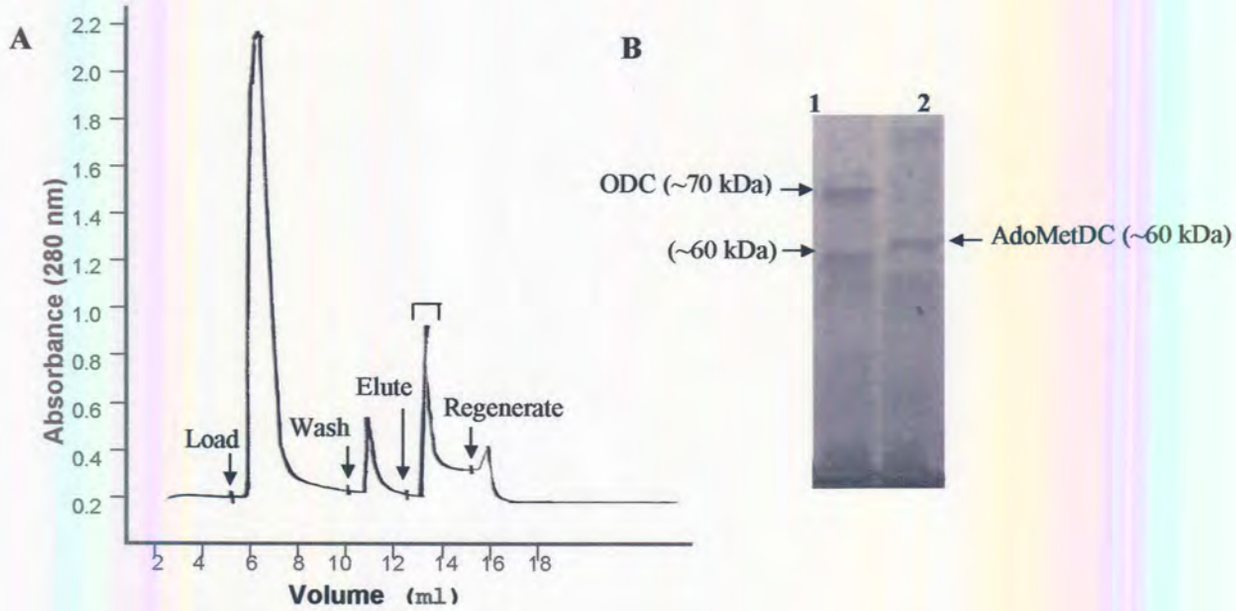

Figure 3.5: His-tag fusion protein expression of monofunctional AdoMetDC or ODC. (A) Typical elution profile of a His-tagged based purification of recombinantly expressed AdoMetDC or ODC. The elution of the recombinant proteins is indicated with the horizontal bracket. **(B)** SDS-PAGE of recombinantly expressed monofunctional AdoMetDC and ODC. Lane 1: ODC protein at ~70 kDa with a contaminating band at ~60 kDa and lane 2: AdoMetDC protein at ~60 kDa. Recombinant protein was visualised with silver staining.

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

3.3.3.2) Strep-tag expression system.

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

al., 2001). Proteins were purified with Strep-Tactin affinity chromatography and eluted in ~ 3 ml in a single peak as indicated in Fig. 3.6. The maximal yield was ~ 250 µg per litre culture for ODC and ~ 150 µg per litre culture for AdoMetDC.

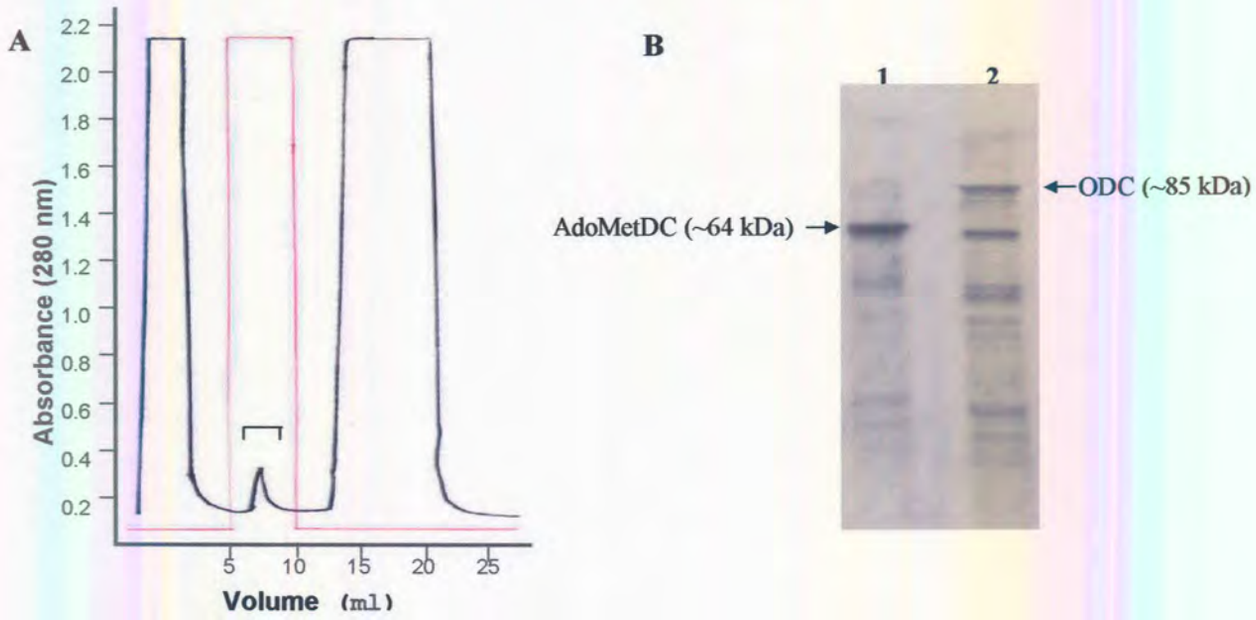

Figure 3.6: Expression of monofunctional AdoMetDC and ODC as Strep-tag proteins. (A) Typical elution profile of the Strep-tag proteins with Strep-Tactin affinity chromatography. The peak containing the eluted protein is indicated with a horizontal bracket. The red line indicates 100% elution buffer. **(B)** SDS-PAGE of the monofunctional proteins. Lane 1: AdoMetDC protein (64 kDa) and lane 2: ODC protein (~85 kDa). Proteins were visualised with silver staining.

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

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

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

Monofunctional ODC was purified with a calibrated size-exclusion HPLC (r^2=0.967 after linear regression, Fig. 3.7 A). ODC eluted in a single peak corresponding in size to the homodimeric form of the protein of ~170 kDa (Fig. 3.7 B). Activity analyses indicated the presence of active ODC in this peak. A large peak at ~10 min (~30 kDa) was present in the ODC separations, possibly indicating the presence of the contaminating proteins from the affinity chromatography step. SDS-PAGE analyses of the ODC-containing single peak showed the presence of the ~85 kDa band and the smaller ~60 kDa band (Fig. 3.7 C). The smaller contaminating bands present after the affinity chromatography was removed by the SE-HPLC. No monomeric ODC was visible in any of the fractions corresponding in size to 85 kDa.

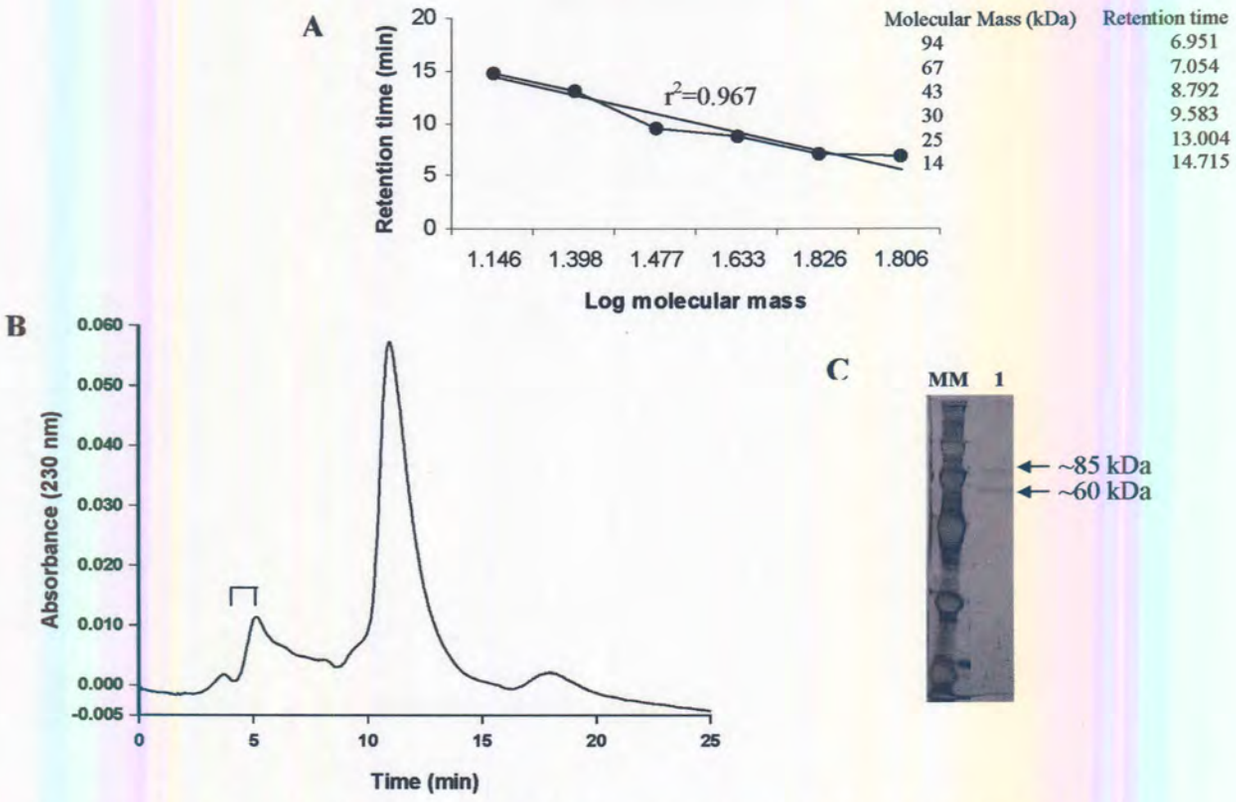

Figure 3.7: Size-exclusion HPLC of the monofunctional ODC purified with affinity chromatography. (A) Calibration curve for the size-exclusion HPLC. **(B)** Elution profile of homodimeric ODC (170 kDa). The peak containing ODC activity is indicated with a horizontal bracket. **(C)** SDS-PAGE of fractions corresponding to the ODC-specific peak indicating only two bands at ~85 and 60 kDa. MM: Low molecular weight protein markers.

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

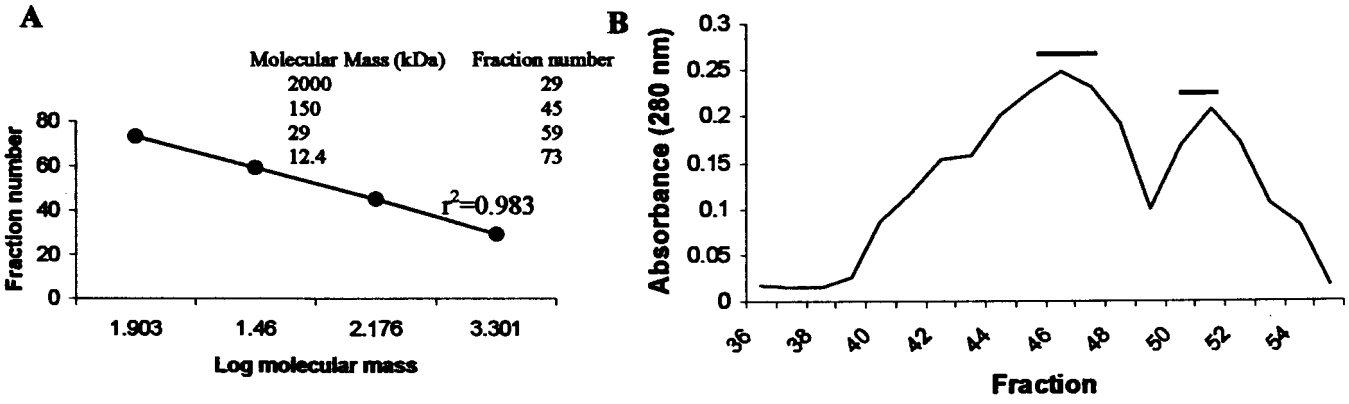

Figure 3.8: SE-FPLC curve for separation of the monofunctional AdoMetDC. (A) Calibration curve for the SE-FPLC. **(B)** Elution profile of AdoMetDC. The horizontal bars indicate AdoMetDC activity.

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

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

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

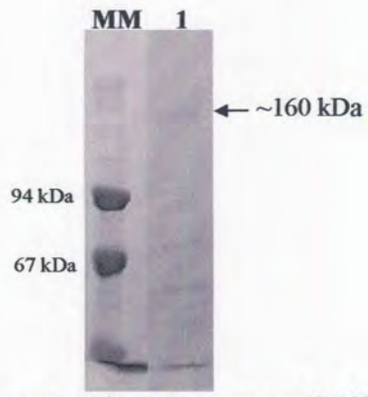

Figure 3.9: SDS-PAGE of the recombinantly expressed bifunctional PfAdoMetDC/ODC. Lane 1 indicates the heterodimeric form of the protein at ~160 kDa. MM: Low molecular weight protein molecular mass marker. Proteins were visualised with silver staining.

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

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

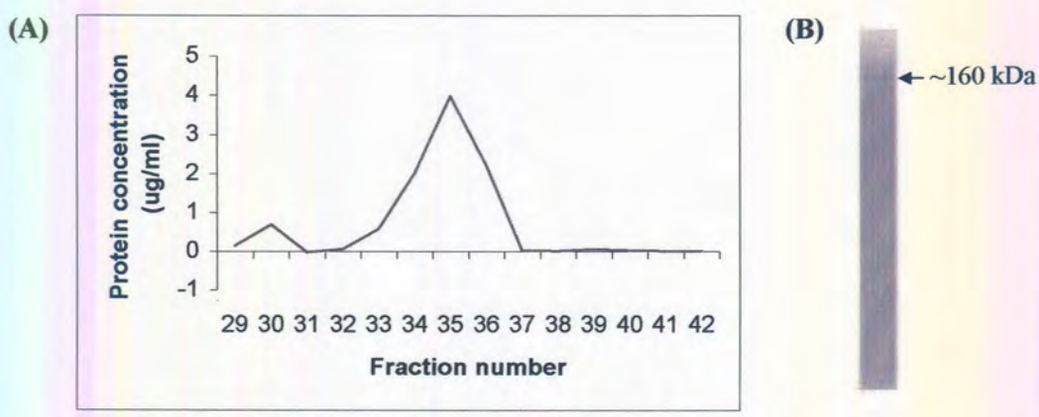

Figure 3.10: SE-FPLC purification of the bifunctional PfAdoMetDC/ODC. (A) Indicates a typical elution profile for PfAdoMetDC/ODC with the size of the protein in fraction 35 corresponding to ~330 kDa. **(B)** SDS-PAGE of fraction 35 indicating a single ~160 kDa band for denatured PfAdoMetDC/ODC.

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

ODC and AdoMetDC activities of the Strep-Tag expressed monofunctional and bifunctional proteins were assayed. Active PfODC is an obligate homodimer of ~170 kDa and is expressed containing part of the hinge region (see also Fig. 3.4). PfAdoMetDC is able to form a heterotetramer of 145 kDa (two 64 kDa and two 9 kDa fragments) but this is not essential for activity, however, the recombinantly expressed protein also contains part of the hinge region. The active form of PfAdoMetDC/ODC is a heterotetrameric structure of two 147 kDa subunits (two polypeptides containing both AdoMetDC and ODC activities) and two 9 kDa subunits (processed β-subunits of AdoMetDC) forming a ~330 kDa complex. Fig. 3.11 schematically represents the active forms of the protein.

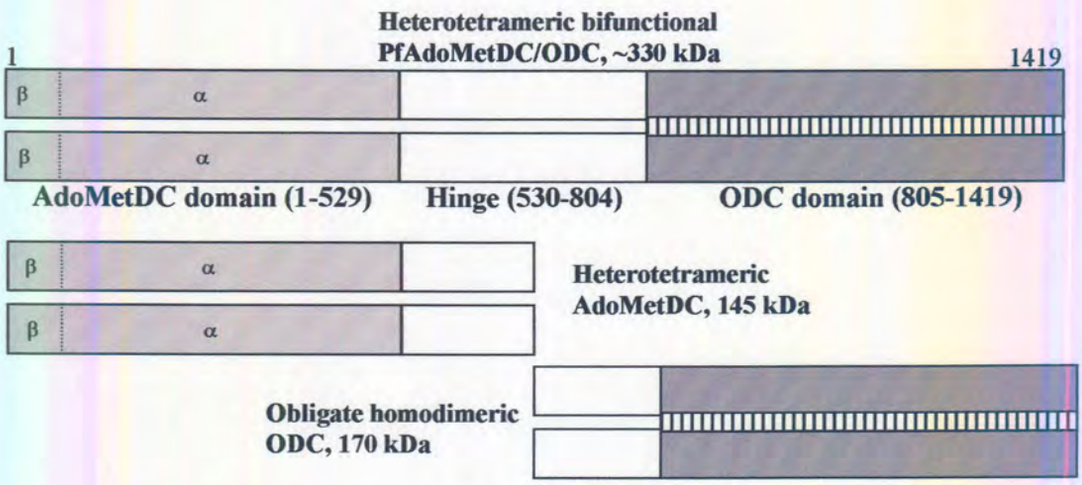

Figure 3.11: Schematic representation of the active forms of the monofunctional AdoMetDC and ODC or the bifunctional PfAdoMetDC/ODC. The dimerisation of the ODC domain is indicated with vertical lines between the monomers, and the α- and β-chains of AdoMetDC are separated with a dashed line.

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

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

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

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

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

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

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

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

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

After removal of the parasite-specific regions and hinge region, the separate domains show moderate identity with the respective proteins from other organisms. The *P. falciparum* AdoMetDC domain has a 18.1% sequence identity with the human homologue, and sequence identities of 17.8%, 15% and 14.2% with the murine, *T. brucei* and *L. donovani* enzymes, respectively. The ODC domain shows a more pronounced sequence identity of 27.6%, 26.6%, 30.7% and 18.6% with the human, mouse, *T. brucei* and *L. donovani* monofunctional proteins, respectively.

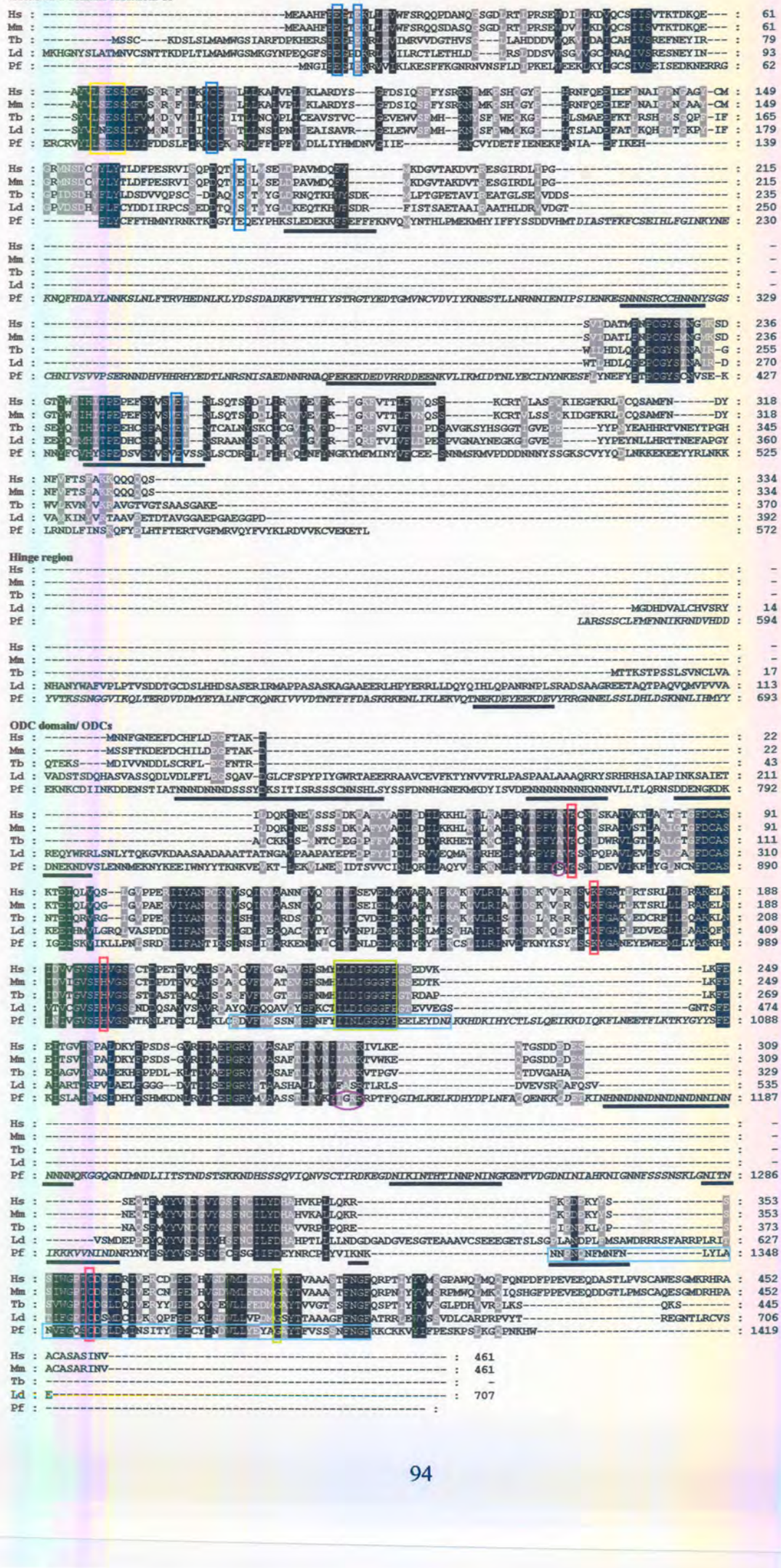

Figure 3.12: Multiple alignment of the bifunctional PfAdoMetDC/ODC amino acid sequence with homologues of the monofunctional AdoMetDC and ODC from other organisms. The deduced amino acid sequence of the PfAdoMetDC/ODC was aligned to the monofunctional sequences from *H. sapiens* (Hs), *M. musculus* (Mm), *T. brucei* (TB) and *L. donovani* (Ld). The AdoMetDC and ODC domains are indicated as well as the hinge region as defined by Müller *et al.* (Müller, *et al.*, 2000). Residues with >80% conserved properties are indicated in black boxes, and >60% in grey. The processing site of AdoMetDCs is indicated in a yellow box, and residues important for activity of AdoMetDCs in blue boxes. Active site residues of ODCs are indicated with red boxes, whereas the proposed amidation site is in a purple circle. The Gly-rich area and Gly involved in dimerisation of ODC are in green boxes. Possible PEST-rich regions are in light blue boxes. The parasite-specific inserts are indicated in italics. Low-complexity areas are indicated by horizontal bars.

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

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

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

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

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

Secondary structure analyses with the GOR IV algorithm indicate a preference for α-helices in the AdoMetDC domain (Fig. 3.13). The ODC domain exhibits a α-helix-β-sheet domain close to the hinge region with a strong β-sheet domain in the C-terminal of the protein (ODC domain). The hinge region and parasite-specific inserts contain unstructured areas or coils.

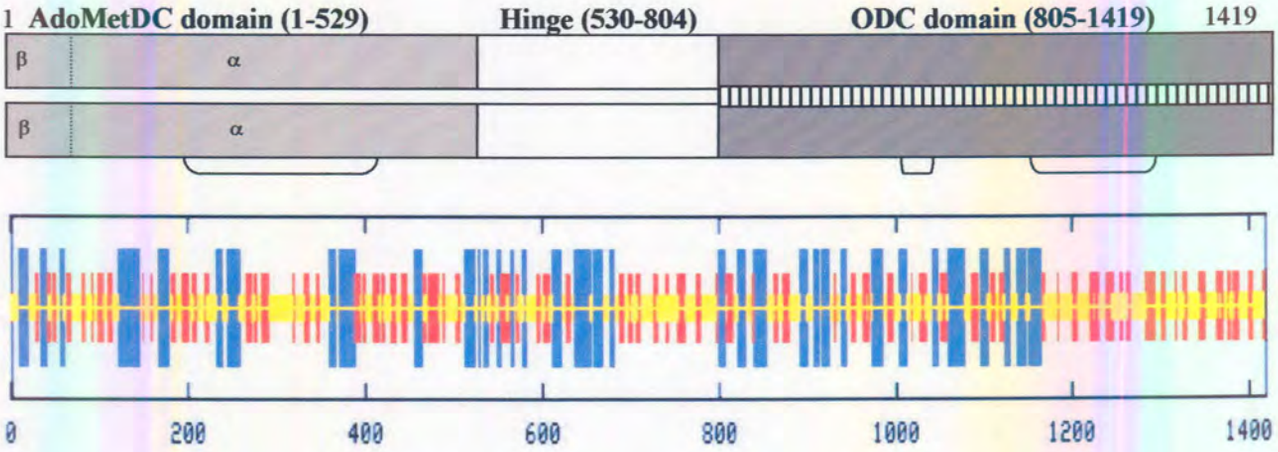

Figure 3.13: Secondary structure prediction of the PfAdoMetDC/ODC amino acid sequence. Top panel indicates the schematic representation of the bifunctional PfAdoMetDC/ODC for reference. The bottom panel indicates α-helices in blue and β-sheets in red. Unstructured areas or coils are indicated in yellow. The horizontal brackets indicate the relative positions of the parasite-specific inserts.

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

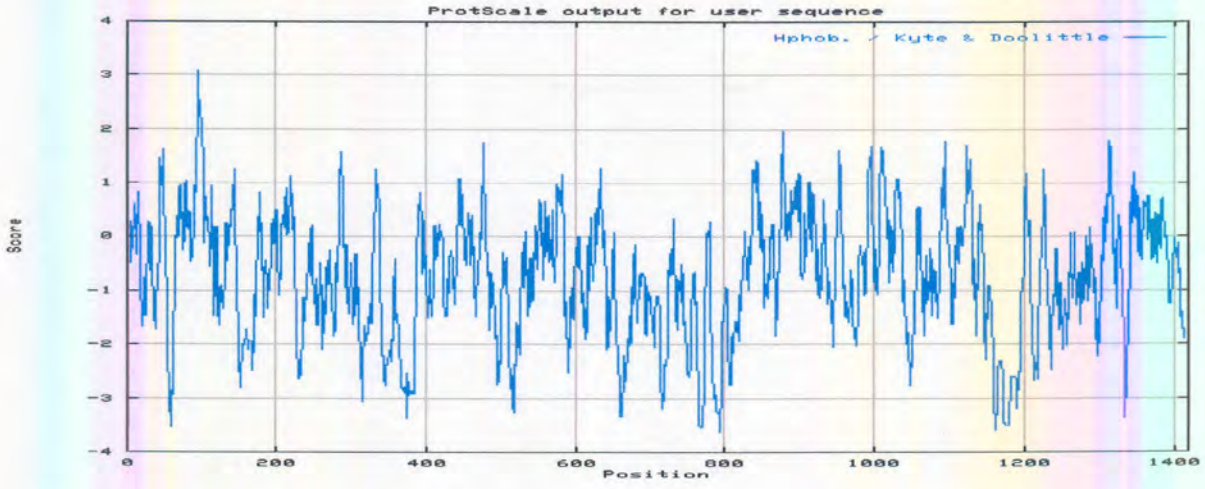

Figure 3.14: Hydrophobicity plot of the deduced PfAdoMetDC/ODC amino acid sequence. Hydrophilic residues are scored negatively.

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

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

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

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

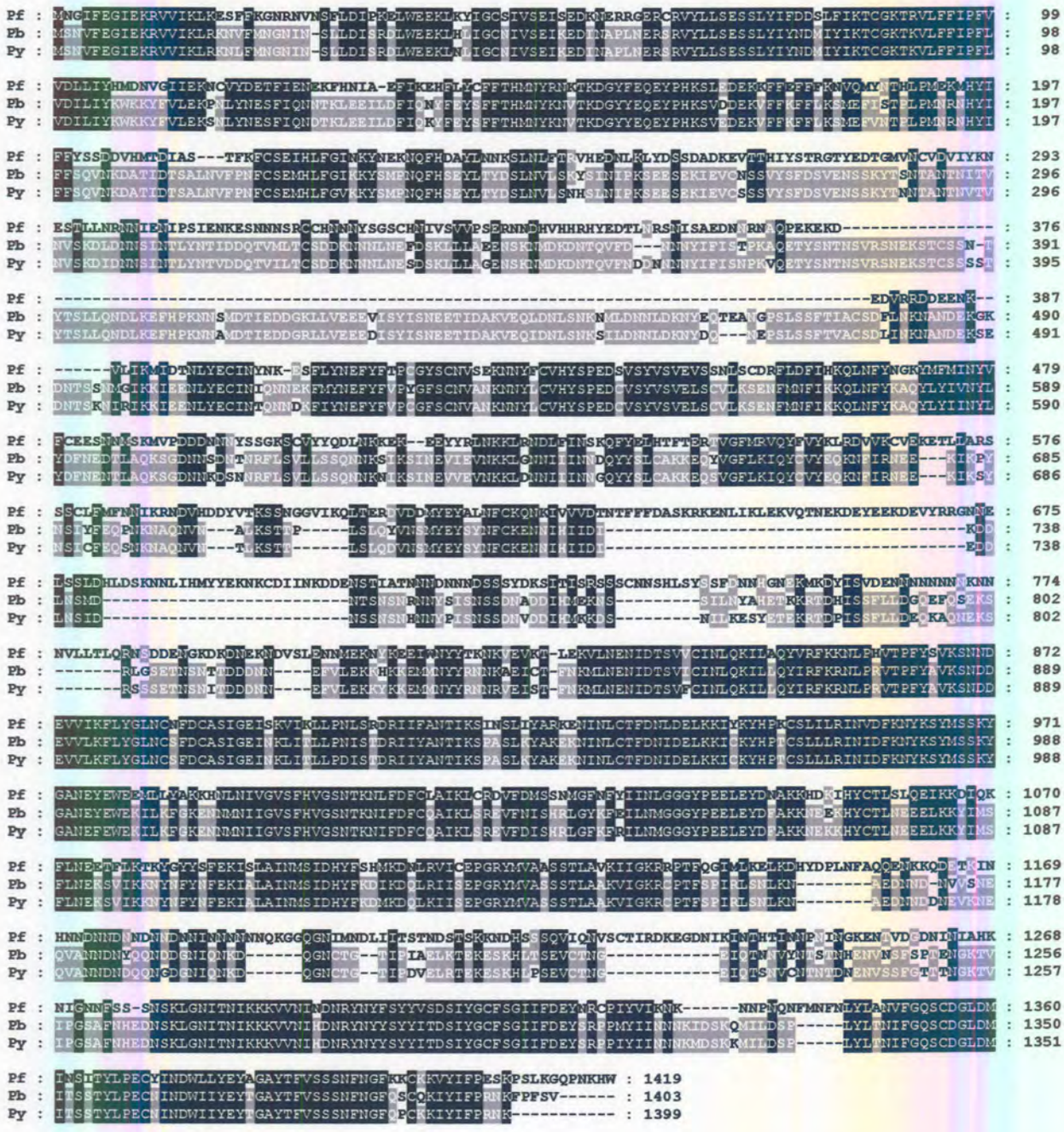

Figure 3.15: Multiple sequence alignment of the deduced amino acid sequence of the bifunctional AdoMetDC/ODC from three *Plasmodium* species. Pf: *P. falciparum*, Pb: *P. berghei* and Py: *P. yoelii*. Residues >80% conserved are boxed in black and >60% conserved in grey.

3.4) DISCUSSION.

3.4.1) Heterologous expression of the decarboxylase proteins.

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

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

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

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

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

3.4.2) Multimeric states of the monofunctional and bifunctional proteins.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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