Chapter 3: Delineation of the roles of structural features in the PfAdoMetDC parasite-specific inserts on bifunctional activity

3.1 Bifunctional enzymes in *P. falciparum*

Besides PfAdoMetDC/ODC, several other bifunctional genes have been described in *P. falciparum*. Bifunctional enzymes are particularly abundant in the folate biosynthesis pathway of the malarial parasite, i.e. PPPK/DHPS, dihydrofolate synthase/folylpolyglutamate synthase (DHFS/FPGS) and DHFR/TS (Nirmalan *et al.*, 2002).

The functional advantages of the bifunctional arrangement of Plasmodial proteins remain uncertain. It has been proposed that some of these proteins originated from gene-fusion [e.g. glucose-6-phosphate dehydrogenase/6-phosphoglucolactonase, *PfGlc6PD/6PGL* of *P. berghei* (Clarke *et al.*, 2001)] or gene-duplication events [e.g. guanylyl cyclase α and β (Carucci *et al.*, 2000)] resulting in acquired functionalities such as substrate channeling, coordinated regulation of both the abundance and activities of the proteins and intramolecular interactions. These evolutionary acquired characteristics may be important for the parasites during immune evasion. For example, the pentose phosphate pathway bifunctional enzyme Glc6PD/6PGL, is involved in protecting the parasites against oxidative stress during erythrocyte infection, while at the same time the efficiency of the pathway is improved by the bifunctional nature of the two enzymes as the product of the Glc6PD reaction is immediately available for hydrolysis by 6PGL (Clarke *et al.*, 2001). Another example where substrate channelling takes place is the direct transfer of dihydrofolate from TS to DHFR. Positive residues along the surface of the two enzymes' active sites form an electrostatic highway for the channelling of the negatively charged substrate (Knighton *et al.*, 1994). The advantages of a direct transfer include shorter transit times for pathway intermediates, no loss of intermediates by diffusion and the protection of labile solvents from the surrounding environment (Ovadi, 1991).

Substrate channelling, however, is not a functional advantage of PfAdoMetDC/ODC, as the reaction product synthesised by PfODC does not serve as the substrate for PfAdoMetDC and *vice versa*. Studies done by Wrenger *et al*. also showed that mutations in the active site of one domain or an inhibitor against it does not effect the activity of the neighbouring domain (Wrenger *et al.*, 2001). It has also been shown that specific regions within the two domains are involved in intramolecular interactions resulting in the correct folding and stabilisation of the bifunctional protein (Birkholtz *et al.*, 2004). A biological advantage of the bifunctional

arrangement of PfAdoMetDC/ODC in *P. falciparum* might thus be that the abundance and activity of the two proteins is regulated in a combined manner for the maintenance of optimum polyamine levels.

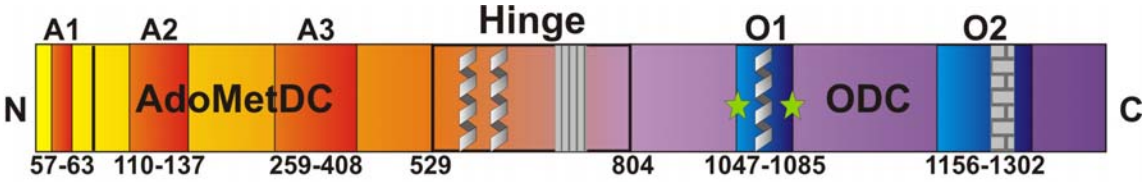
3.2 Parasite-specific inserts in PfAdoMetDC/ODC

Plasmodium proteins are unusually longer in size compared to their yeast orthologues, which is mostly due to the insertion or expansion of LCRs (Gardner *et al.*, 2002) present in nearly all of *P. falciparum* proteins (Pizzi and Frontali, 2001). LCRs in proteins characteristically consist of a limited set of amino acids, often contain repeats of one or more residues and do not adopt definite structures (Wootton, 1994). Similarly, parasite-specific inserts also exist as non-globular regions and often contain LCRs, consisting mainly of Lys and Asn residues. These are species-specific, rapidly diverging areas within proteins that form flexible domains extending from the protein core (Pizzi and Frontali, 2000; Pizzi and Frontali, 2001).

A number of theories exist regarding the evolution and function of LCRs in proteins. The prevalence of LCRs at antigenic loci has linked these areas to antigenic diversity for reduced effectiveness of the human immune response (Hughes, 2004). The parasite-specific inserts in PfAdoMetDC/ODC, however, do not show significant antigenicity suggesting a different role of these areas (Birkholtz, 2002). LCRs have alternatively been implied in mRNA stability (Xue and Forsdyke, 2003) and protein-protein interactions (Karlin *et al.*, 2002). A systematic study of LCRs by DePristo *et al.* led them to propose that the abundance and the amino acid composition of LCRs in *P. falciparum* is merely as a result of the A+T rich genome of the parasites and that these areas have not been evolutionary maintained for their functional and adaptive roles (DePristo *et al.*, 2006). Areas rich in positively charged Asn residues have been described as prion-domains, which mediate protein-protein interactions and contribute to either homo- or hetero-aggregation of polypeptides (Singh *et al.*, 2004). Polar residues (Asp and Glu) have also been predicted to mediate protein-protein interactions via modular polar zipper domains because of the capacity of their side chains to form hydrogen bonded networks (Perutz *et al.*, 1994; Michelitsch and Weissman, 2000). A highly probable explanation for the presence of parasite-specific inserts in PfAdoMetDC/ODC and the high abundance of charged residues therein might thus be the formation of intramolecular protein-protein interactions for the stabilisation of the heterotetrameric bifunctional complex.

The bifunctional PfAdoMetDC/ODC enzyme contains six Plasmodial conserved parasite-specific inserts, designated A1, A2, A3 (in the PfAdoMetDC domain), the hinge region, O1 and O2 (in the PfODC domain) (Müller *et al.*, 2000; Birkholtz *et al.*, 2004; Wells *et al.*, 2006). The roles that these insertions play in the activities of the proteins have previously been

investigated (Birkholtz *et al.*, 2004). Secondary structure prediction analysis of the parasite-specific inserts revealed several Plasmodial conserved structures within the inserts. The secondary structures in the hinge region and the PfODC inserts were investigated in a previous study (Roux, 2006). Figure 3.1 summarises the effects of various point mutations and deletions on domain activities within PfAdoMetDC/ODC.



Mutation	% AdoMetDC activity	% ODC activity	Reference
Monofx AdoMetDC + hinge (1-660)	119	-	Wrenger *et al.*, 2001
Central hinge deletion (573-752)	67	50	Birkholtz *et al.*, 2004
1st helix breaker in hinge	87	66	Roux, 2006
2nd helix breaker in hinge	36	33	Roux, 2006
Beta sheet deletion in hinge (573-752)	89	127	Roux, 2006
Monofx ODC (805-1419)	-	3	Krause *et al.*, 2000
Monofx ODC + hinge (660-1419)	-	6	Krause *et al.*, 2000
Insert O1 deletion (1047-1085)	23	6	Birkholtz *et al.*, 2004
1st Gly to Ala displacement in O1	17	0	Roux, 2006
2nd Gly to Ala displacement in O1	33	3	Roux, 2006
Helix breaker in O1	6	0	Roux, 2006
Insert O2 deletion (1156-1302)	48	3	Birkholtz *et al.*, 2004
Conserved repeat deletion in O2	72	54	Roux, 2006
Deletion of inserts O1 and O2	30	3	Birkholtz *et al.*, 2004

Figure 3.1: The bifunctional *P. falciparum* AdoMetDC/ODC protein together with results of previous structure-function relationship studies.
Only the heterodimeric form of the bifunctional protein is shown. The N-and C-termini as well as the positions of the parasite-specific inserts are indicated (A1, A2, A3, hinge, O1, O2) (Birkholtz *et al.*, 2003; Birkholtz *et al.*, 2004; Wells *et al.*, 2006). Secondary structures within the parasite-specific inserts are also shown. The β-sheet is represented by a striped area while the repeat (NND)x region in insert O2 is shown as a bricked area. The activities are normalised as a percentage of the wild type PfAdoMetDC/ODC activity. Helices were disrupted with the insertion of Pro point mutations.

The functions of the unique parasite-specific inserts and the conserved structures within these are important in studies aimed at targeting these areas in PfAdoMetDC/ODC with possible chemo-preventative inhibitors.

3.3 Identification of secondary structures

Secondary structures are responsible for the local conformations of proteins and are also involved in the initiation of protein folding and its conformational stabilisation (Richardson, 1981). The formation of secondary structures are favoured by some amino acids and discouraged by others, and it is these structural propensities of residues that form the basis of a number of algorithms designed to predict secondary structures. More than three

decades ago, Chou and Fasman documented that amino acids have different propensities to initiate the folding of regular secondary structures such as α-helices and β-sheets (Chou and Fasman, 1974). They created a prediction model that assigns values to each residue and then a simple algorithm is applied to determine the most probable secondary structure that will form. The conformational parameters of each amino acid were calculated by the frequency of the given amino acid within a protein, its frequency within a given type of secondary structure, and the fraction of amino acids occurring in that particular type of structure. These parameters ultimately give an indication of the probability that an amino acid will be found in a helix, sheet or coil. Nucleation sites for the folding of a particular structure can thus be located and extended until a stretch of amino acids are encountered that are not suitably found in that type of structure or until a stretch is encountered that has a greater disposition for another type of secondary structure (Chou and Fasman, 1974). Since then, several models have been developed to predict secondary structure folding based on the primary sequence of amino acids in a polypeptide.

The Garnier method is very similar to the Chou-Fasman method but it additionally takes into account the immediate environment of the respective residue when predicting a probable structure (Sternberg, 1996).

3.3.1 Hydrophobic cluster analysis

In the “Hydrophobic core collapse” model, protein-folding cooperativity is due to the assembly of non-polar residues within a core. Subsequent development of hydrophobic clusters and cores results in the formation of secondary structures. This model is consistent with data showing a concurrent collapse of protein and some secondary structure formation in the early stages of protein folding (Chan *et al.*, 1995).

The established connection between hydrophobic interactions and consequent protein folding has led to the development of hydrophobic cluster analysis (HCA) as an efficient method to represent an amino acid sequence, 2-dimensionally, in the absence of a 3-dimensional structure. These hydrophobic clusters tend to fold into secondary structures and are thus useful to identify functional areas within a protein (Gaboriaud *et al.*, 1987). Indeed, hydrophobic amino acids dominate the internal core of a protein, whereas hydrophilic amino acids mostly occur on the protein surface to protect the core from the solvent. As a result, the hydrophobic amino acids tend to cluster into a stable and compact structure that is typical of a particular fold (Callebaut *et al.*, 1997).

HCA allows one to compare a protein sequence with unknown structure, such as that of the parasite-specific inserts, with a known structure from, for example, a related species, where the main chain folding of the latter protein provides a starting point for the structural modelling of the former protein (Gaboriaud *et al.*, 1987). This is because a marked difference in amino acid identity does not necessarily imply major differences in the overall 3-dimensional folding of the proteins (Argos, 1987).

In this study HCA, in addition to secondary structure prediction programmes, was thus applied to determine the overall structural content of the seemingly functionless parasite-specific inserts in the PfAdoMetDC domain of *P. falciparum* AdoMetDC/ODC. Structural delineation of the parasite-specific inserts provides important insights towards the rational design and development of specific inhibitory compounds targeting these areas.

3.3.2 The roles of secondary structures in protein-protein interactions and motif formation

Any specified residue within a protein may participate in various interactions, including those within a secondary structural element as well as those in long-range interactions a considerable distance away in the amino acid sequence (Kwok *et al.*, 2002). The identification of secondary structures represents the first step in describing the tertiary and super-secondary topologies of proteins such as barrels and bundles.

3.3.2.1 α-Helices and protein-protein interactions

Periodical secondary structures such as helices and sheets are the main components of homodimeric interfaces where 40% of the interface residues are helical ones (Guharoy and Chakrabarti, 2007). α-Helices also serve as important recognition regions for the binding and interaction of other proteins, DNA and RNA (Davis *et al.*, 2007). The majority of integral membrane proteins adopt α-helical transmembrane segments that may form well-packed bundles. These helical structures are stabilised within the membrane as the backbone amide-carbonyl hydrogen bonds are satisfied within the helix and are isolated from the non-polar environment of the surrounding lipid bilayer (White and Wimley, 1999).

Glycophorin A (GpA) is arguably the best-studied example showing the role that helices play in intramolecular protein-protein interactions and the dimerisation of protein monomers. GpA is an integral membrane sialoglycoprotein in the human erythrocyte that associates to form a dimer via side-by-side interactions between transmembrane α-helices. It was discovered early that the dimeric form of the protein cannot be separated into its monomeric forms with

sodium dodecylsulphate (SDS), which demonstrates the strength of this interaction mediated solely by the transmembrane region (Bormann *et al.*, 1989; Lemmon *et al.*, 1992).

Leucine zippers or coiled-coils are perhaps the simplest protein units that can be classified as small dimeric proteins (Wendt *et al.*, 1997). These consist of two peptides in a roughly α-helical conformation, wound around each other into a bundle of right-handed helices, coiled about one another to form a left-handed superhelix quaternary structure (Gromiha and Parry, 2004). The fundamental structure consists of a repeating seven-residue motif (*abcdefg*) where the *a* and *d* positions are characteristically occupied by non-polar residues comprising the hydrophobic core of the helix-helix interface, while residues at the peripheral positions *e* and *g* are mostly charged residues forming salt-bridges with each other and hydrophobic contacts with the core residues. The side chains of one helix protrude into the cavities formed by the side chains of the adjacent helix in a regular manner termed knobs-into-holes or leucine zipper packing. Leucine zippers in the basic-leucine-zipper (bZIP) family of transcription factors mediate the dimerisation of two basic protein domains to form a DNA-binding site, showing the involvement of helices in intra- and intermolecular protein-protein interactions, respectively. Leucine zippers can be homodimeric as in the yeast GCN4 transcription factor or heterodimeric as in the jun-fos type transcription factors (Wendt *et al.*, 1997).

Another example of helix involvement in intermolecular protein-protein interactions is the transmembrane domains of SNARE proteins. The mechanism of fusion between a vesicle to a target membrane protein (e.g. Ca^{2+}-mediated neurotransmitter release) is regulated by interactions via a set of proteins termed the soluble *N*-ethylmaleimide-sensitive factor (NSF) attachment protein receptors (SNARE) complex. This complex includes SNAP25 (synaptosome-associated protein) and transmembrane proteins that are located on the target membrane (t-SNARE) and the vesicle membrane (v-SNARE) (Ungar and Hughson, 2003). The fusion capability of this complex involves sequence-specific interactions between the v- and t-SNAREs via extremely stable coiled-coils holding the vesicle close to the target membrane for subsequent fusion. These interactions can only be broken by NSF-catalysed hydrolysis of ATP dissociating the complex or the introduction of an interference protein (Kroch and Fleming, 2006).

Proteins mostly perform their functions through their interactions with other molecules or with itself, and the identification of these inter- and intramolecular binding sites are important for functional annotation studies, the understanding of molecular recognition processes and rational drug target development and design strategies. It will thus prove useful to identify interacting areas within the parasite-specific inserts of PfAdoMetDC/ODC that can be

targeted with compounds to prevent these interactions from taking place and in such a manner influence protein activity.

3.3.2.2 β-Sheets, barrels and turns

β-Sheets are not common as binding sites in protein-protein interactions, which are mostly fulfilled by the helical secondary structures. Chakraborty *et al*., however, reported that a four-stranded β-sheet excised from the HIV gp120 protein, forms the binding site of the neutralising CD4-binding antibody namely 17b (Chakraborty *et al.*, 2005). Initially, it was also thought that α-helical transmembrane bundles formed all integral membrane proteins including transporters and ion channels. Now it is evident that transmembrane β-barrels are also involved in the formation of transbilayer porins, K^+ and related voltage-gated ion channels and epithelial Na^+ channels (Sansom and Kerr, 1995).

β-Barrel membrane proteins are found in the outer membranes of mitochondria, chloroplasts and Gram-negative bacteria and span the membranes as β-sheets. Alternating hydrophilic and hydrophobic residues facing the membrane environment and hydrophilic interior of the barrel, respectively, forms the entire transmembrane protein (Schneider *et al.*, 2007). β-Turns are also important motifs involved in the biological recognition of various proteins and peptides. They form regular recognition motifs for antibody-peptide complexes and peptide hormones, such as angiotensin II, gonadotrophin releasing hormone, somatostatin and many more. They are often evolutionary conserved and are believed in some cases to initiate protein folding (Tran *et al.*, 2005).

Following the identification of conserved secondary structures within PfAdoMetDC/ODC, disrupting an α-helix or deleting a β-sheet can give an indication of their possible involvement in protein-protein interactions and activity. The cyclic nature of Pro makes it an unique amino acid as its amide group lacks the hydrogen atom necessary for hydrogen bond stabilisation in regular secondary structures (Orzáez *et al.*, 2004). The replacement of a residue that has a high frequency in α-helices with a Pro residue can thus be employed to disrupt the folding event from taking place. For example, the replacement of an Ile residue, a strong helix-former, with a Pro residue, a helix breaker, within the 42-amino acid amyloid β-peptide, which is responsible for increased levels of oxidative stress in Alzheimer's disease, resulted in the disruption of a helix that caused a decrease in amyloid peptide aggregation, oxidative stress and neurotoxicity (Kanski *et al.*, 2002).

An important aspect to consider with regards to Pro residues is *cis-trans* isomerisation. The Pro hypothesis states that protein molecules, especially slow-folding ones, possess non-native isomers of peptide bonds between Pro and the next residue in a polypeptide chain

and that key steps in the refolding event are limited by the rate of slow reisomerisation of the incorrect Pro-peptide bonds. The enzyme responsible for the *cis*-in-equilibrium-*trans* isomerisation of Pro-peptide bonds is prolyl isomerase and its activity is dependent on the accessibility of the particular Pro bond in the refolding protein chain (Lang *et al.*, 1987). The introduction of a Pro residue may therefore result in general conformational changes instead of a desired conformational disruption in the region of the helix. This enzyme is present in *P. falciparum* (PDB entry: Pf08_0121).

Prior research has focussed on delineated areas of the parasite-specific inserts in the PfODC domain as well as the hinge region (Table 3.1). These studies showed that the bifunctional protein depends on the presence of these areas for their proper functioning and/or complex formation. The present study therefore focussed on the parasite-specific inserts within the PfAdoMetDC domain to determine their roles in bifunctional activity. The presence of *Plasmodia* conserved secondary structures within these was also identified, which could possibly be responsible for the activities and/or stabilisation of the bifunctional PfAdoMetDC/ODC complex via protein-protein interactions as in the examples given above.

3.4 Methods

3.4.1 Delineation of structural features within the PfAdoMetDC parasite-specific inserts

It was previously thought that low-complexity-containing parasite-specific inserts do not contain any significant structures and mostly form non-globular, surface exposed loops (Pizzi and Frontali, 2001). The application of secondary structure prediction algorithms in past and present studies on the parasite-specific inserts in PfAdoMetDC/ODC have, however, predicted the presence of several conserved α-helices and β-sheets in both the domains as well as the hinge region (Roux, 2006). Similarly, the DHFR/TS crystal structure was also shown to contain parasite-specific inserts with definite structures (Yuvaniyama *et al.*, 2003).

The structural alignment of the PfAdoMetDC domain was previously performed (Wells *et al.*, 2006). The AdoMetDC domains of *Plasmodium* spp. are highly divergent from other eukaryotic enzymes mostly due to the presence of *Plasmodium*-specific inserts, which is why a structure-based approach was followed in order to optimise the alignment. Initial studies suggested the presence of one large PfAdoMetDC insert (Müller *et al.*, 2000; Birkholtz *et al.*, 2004), but subsequent multiple alignments, however, identified three inserts within this domain i.e. A1, A2 and A3 (Wells *et al.*, 2006).

Manual adjustments were made during homology modeling of the PfAdoMetDC domain to prevent the disruption of secondary structures that involve areas of the insert and the main protein chain (Wells *et al.*, 2006). It is, however, still possible that the inserts themselves adopt defined secondary structures that may confer structural functionalities to these areas. The sequences of the parasite-specific inserts were thus investigated for the presence of such structural features. The A2 and A3 parasite-specific inserts sequences of three *Plasmodia* species (*P. falciparum*, *P. berghei* and *P. yoelii*) were subjected to numerous secondary structure prediction algorithms from different web servers listed in Figure 3.3 (G.A. Wells, personal communication). During the homology modeling of PfAdoMetDC, insert A1 was moved two residues downstream to occupy residues Lys57 to Glu63. This was done in order to prevent its involvement in the second β-strand of PfAdoMetDC and as such was thus not investigated for the presence of any secondary structures (Wells *et al.*, 2006).

In addition, hydrophobic cluster analysis diagrams of the A2 and A3 inserts were constructed by subjecting their insert amino acid sequences to the DrawHCA programme (Gaboriaud *et al.*, 1987). During HCA, the amino acid sequence is written on a classical α-helix with 3.6 amino acids per turn and smoothed on a cylinder (Figure 3.2 a). In such a representation, residues *i* and *i*+18 (after five helical turns) will have similar positions parallel to the axis of the cylinder (e.g. Figure 3.2 a, G246 and E264). An easier representation of the 3-dimensional helical structure is subsequently obtained when the cylinder is cut parallel to its axis and unrolled (Figure 3.2 b). The widely separated distribution of some amino acids upon unrolling (e.g. Figure 3.2 b, F253 and L254) are then corrected by duplicating the representation, which gives a better indication of each amino acid's immediate environment (Figure 3.2 c).

Sets of adjacent hydrophobic amino acids (Ile, Leu, Phe, Trp, Met, Tyr, Val) are finally encircled to create hydrophobic clusters. Particular amino acids are highlighted: a star for Pro, which confers the greatest constraint to the polypeptide chain; a diamond for Gly, which confers the largest freedom to the chain; and squares with a dot or not for Ser and Thr, respectively, as these two polar amino acids can mask their polarity, particularly within helices (Figure 3.2 d).

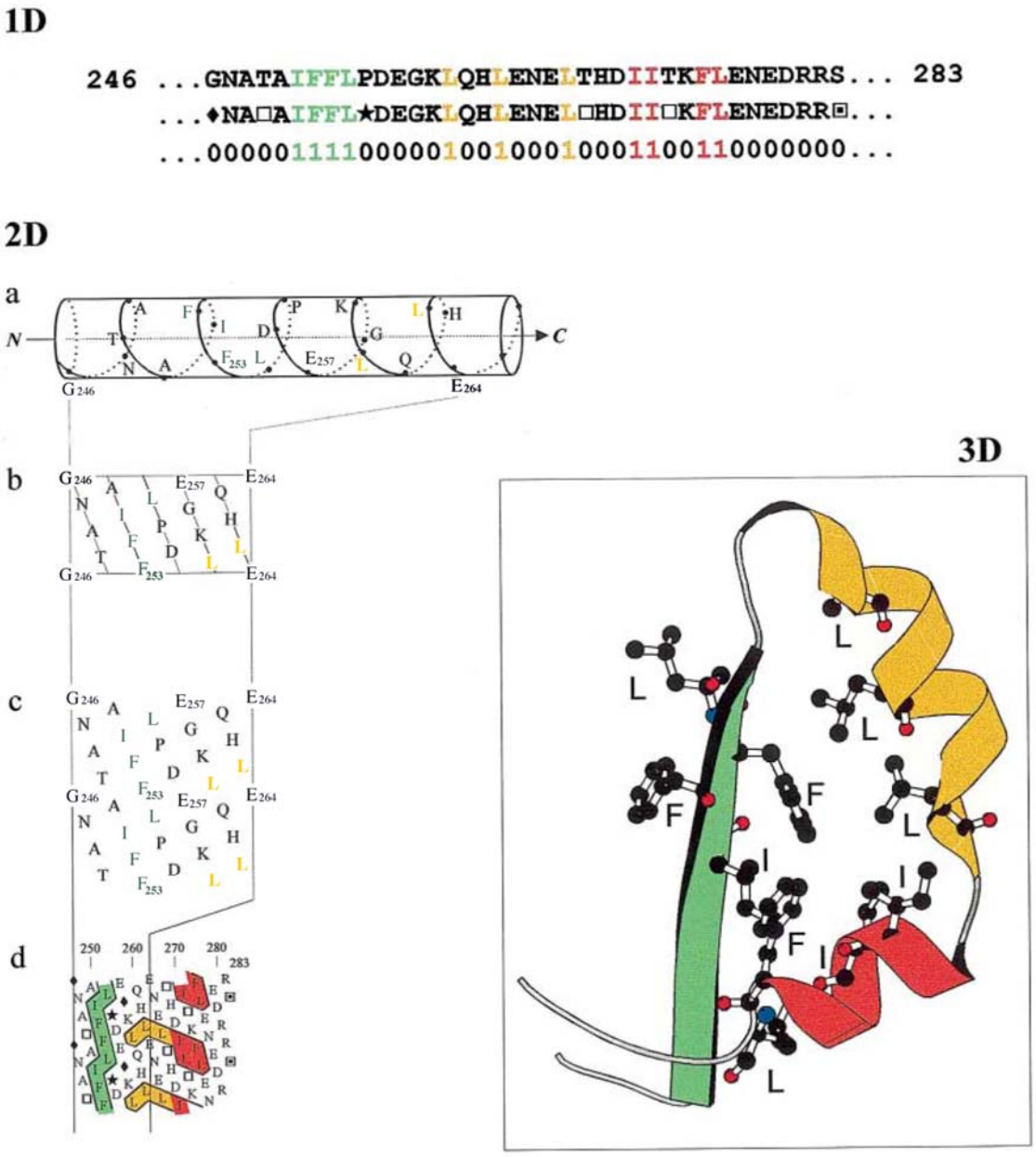


Figure 3.2: Hydrophobic cluster analysis: between sequence and structure.
The linear 1-dimensional sequence (1D) of a segment of human α_1-antitrypsin is shown with coloured hydrophobic amino acids (first line). The same sequence is represented using the HCA code for the Gly, Pro, Thr and Ser amino acids (second line). The direct translation into a numbered code where 1=hydrophobic and 0=hydrophilic is also represented (third line). The sequence is displayed on a cylinder (a), which is cut parallel to its axis and unrolled into a bi-dimensional diagram (b). In (c), this diagram is compacted and duplicated. Hydrophobic amino acids form clusters (d), which are illustrated by the corresponding 3-dimensional structure (3D). The vertical cluster (green) is associated with a β-strand whereas the horizontal one (yellow and orange) corresponds to α-helices. Sequence stretches separating clusters correspond to loops or hinges between domains. A hydrophobic cluster always starts and ends with a hydrophobic amino acid coded with a 1.

3.4.2 Nucleotide deletion of the PfAdoMetDC parasite-specific inserts and mutagenesis of delineated areas

3.4.2.1 Cloning of *PfAdoMetDC/ODC* into pASK-IBA3

The *P. falciparum AdoMetDC/ODC* gene (~4300 bp) was cloned into a pASK-IBA3 vector [~3100 bp, (Institut für Bioanalytik, Germany)] and used as template during the mutagenesis reactions (total template size ~7400 bp and called pA/Owt henceforth) (Müller *et al.*, 2000).

3.4.2.2 Mutagenesis PCR

The size of the A3 parasite-specific insert prompted an investigation into available deletion mutagenesis methods and led to the development of the RE-mediated inverse PCR method (Williams *et al.*, 2007). The applicability of the method was tested on and proven to be highly efficient when the largest A3 insert in the PfAdoMetDC domain was removed (Chapter 2). This section will therefore only include the deletion mutagenesis of the smaller A1 and A2 parasite-specific inserts.

The partial overlapping primer method by Zheng *et al*. was applied in the design of the primers for the deletion of the A1 (7 amino acids) and A2 (28 amino acids) inserts (Zheng *et al.*, 2004). The principles of this method are described in section 2.2.1.4. The primers for the deletion of the A1 (A1F and R) and A2 (A2F and R) are listed in Table 3.1 and mutagenesis PCR was performed as stipulated in section 2.2.1.4 to create the A/OΔA1 and A/OΔA2 mutations, respectively.

Table 3.1: The mutagenic primers and their properties for the deletion of the smaller parasite-specific inserts in PfAdoMetDC

Primer	Length (nt)	Tm* (°C)	Primer Sequence (5' to 3')
A1F	46	73	cggaaataagtgaggac^cgatgtcgtgtgtatttattgtcagagag
A1R	51	73	taaatacacacgacatcg^gtcctcacttatttccgatacaatactacaac
A2F	53	68	catatggataat^gaacatttcctttattgttttttacacatatgaattaccg
A2R	48	65	caataaaggaaatgttc^attatccatatgatatattaataaatcaacc

* The Tm's were calculated according to the Rychlik *et al*. formula:
69.3 + 0.41(%GC) – (650/N) where N is the number of nucleotides (Rychlik *et al.*, 1990). Blue letters indicate where the 5'-ends of the forward and reverse primers overlap. ^ Indicates where the deletions of the inserts are located.

For the disruption of the α-helix within insert A2 identified above, a specific residue conserved between the different secondary structure predictions and located in the centre of the specific area was disrupted through point mutation. The strong Ile helix-former at position 132 was replaced with a conformationally rigid Pro residue using the overlapping primer design method (Zheng *et al.*, 2004). In the case of the first β-sheet in insert A3, the conserved eight amino acid structure predicted (Val284 to Tyr291) by the secondary

structure algorithms was simply deleted with the RE-mediated inverse PCR method described in Chapter 2 (Williams *et al.*, 2007).

The mutagenesis primers for the disruption of the α-helix in A2 and the deletion of the β-sheet in A3 are given in Table 3.2. The A2aF and R primers were used to introduce a Pro as helix breaker at position 132 while the A3bF and R primers were used for the deletion of the β-sheet in insert A3. The A3b primers contain 5'-overhanging CGC ends followed by unique *Bam*HI RE sites.

Table 3.2: The mutagenic primers and their properties for the disruption of the secondary structures within the PfAdoMetDC parasite-specific inserts

Primer	Length (nt)	Tm* (°C)	Primer Sequence (5' to 3')
A2aF	25	56	cat aat **cca** gct gaa ttc ata aaa
A2aR	22	56	c agc **tgg** att atg gaa ttt ctc
A3bF	27	63	CGC ggatcc aag aat gaa agt aca ttg
A3bR	29	67	CGC ggatcc cat tcc tgt atc ttc ata tg

* The Tm's were calculated according to the Rychlik *et al*. formula:
69.3 + 0.41(%GC) – (650/N) where N is the number of nucleotides (Rychlik *et al.*, 1990). Blue letters indicate where the 5'-ends of the forward and reverse primers overlap. The nucleotides written in bold specify the Ile132Pro point mutation while the capital letters represent the 5'-overhanging ends followed by the *Bam*HI restriction enzyme sites (underlined) for the RE-mediated inverse PCR primers.

The pA/Owt plasmid was used as template during PCR mutagenesis for the disruption of the α-helix and deletion of the β-sheet to produce constructs A/OpA2a and A/OΔA3b, respectively, where "a" represents an α-helix and "b" a β-sheet. The mutagenesis reactions for the partial overlapping and RE-mediated inverse PCR methods were followed as described in sections 2.2.1.4 and 2.2.1.6, respectively.

3.4.2.3 Post-PCR manipulations

The PCR products were visualised with agarose electrophoresis gels as described in section 2.2.1.7 to ensure that PCR products of approximately 7400 bp (gene size ~4300 bp, vector size ~3100 bp) were obtained.

All the PCR products were subsequently treated with 10 U of *Dpn*I (Fermentas, Canada) for 3 hrs at 37°C. The same quantity of *Bam*HI (Fermentas, Canada) was additionally added to the *Dpn*I digestion reaction of the RE-mediated inverse PCR A/OΔA3b product in a dual compatibility buffer Tango™ (33 mM Tris-OAc, 10 mM Mg-OAc, 66 mM K-OAc, 0.1 mg/ml BSA, pH 7.9). The insert-deleted products were purified with the High-Pure PCR product purification kit (Roche Diagnostics, Germany) while the secondary structure AOpA2a and A/OΔA3b mutants were cleaned with the NucleoSpin® Extract II PCR cleanup kit (Macherey-

Nagel, Germany). The pure DNA samples were finally circularised for 16 hrs with 3 U of T4 DNA Ligase (Promega, USA) at 22°C (A/OΔA1, A/OΔA2 and A/OpA2a) and 4°C (A/OΔA3b) for blunt- and sticky-ended ligation, respectively. The mutant, circular DNA samples (pA/OΔA1, pA/OΔA2, pA/OpA2a and pA/OΔA3b) were subsequently electroporated into DH5α cells as described in section 2.2.1.8.

3.4.2.4 Plasmid isolation and nucleotide sequencing

The plasmids from five clones for each mutation were inoculated in 5 ml LB-amp (LB-medium with 50 µg/ml ampicillin) and allowed to grow overnight at 37°C and 250 rpm. Plasmids were subsequently isolated from the overnight cultures with the use of the High Pure Plasmid Isolation kit from Roche (section 2.2.1.9) and the peqGOLD Plasmid Miniprep Kit I (Biotechnologie, Germany), which is based on the rapid alkaline-SDS lysis method.

Verification of the insert deletions was performed with nucleotide sequencing as described in section 2.2.1.11. The SamRSf1 (5'-aatgaacggaatttttgaag-3') primer was used for the detection of the A/OΔA1 mutant and K215A2 (5'-ccgaacagaatccaaacgtagaag-3') was used to detect the A/OΔA2 insert deletion. The K215A2 primer was also used to detect the A/OpA2a mutation. A helix-disrupted mutant will contain Pro instead of Ile at position 132. The K215A1 (5'-gcttctacgtttgcattctgttcgg-3') sequencing primer was used for the verification of the β-sheet deletion within insert A3.

3.4.3 Recombinant protein expression and isolation of the mutated PfAdoMetDC/ODC proteins

The previously transformed DH5α cells containing the pASK-IBA3 vectors (Institut für Bioanalytik, Germany) with the wild type *PfAdoMetDC/ODC* gene (pA/Owt) and those lacking the three individual PfAdoMetDC parasite-specific inserts (pA/OΔA1, pA/OΔA2 and pA/OΔA3, created in Chapter 2) were grown overnight in 5 ml LB-amp (LB-medium with 50 µg/ml ampicillin) at 37°C with agitation at 250 rpm. The plasmids were isolated as described in section 2.2.1.9. The isolated wild type (pA/Owt) and mutant (pA/OΔA1, pA/OΔA2 and pA/OΔA3) plasmids were subsequently recombinantly expressed.

The plasmids were heat shock transformed into the AdoMetDC and ODC deficient EWH331 *E. coli* protein expression cells (Hafner *et al.*, 1979).

3.4.3.1 Heat shock transformation

Cells were grown overnight in 5 ml LB-medium at 30°C with moderate shaking. The culture was diluted 1:10 ml in 50 ml LB-glucose (20 mM glucose in LB-medium) and grown at 30°C

with moderate shaking until an OD550 of 0.3 was reached. The culture was then transferred into 50 ml centrifuge tubes and incubated on ice for 10 min after which the cells were collected by centrifugation at 1500 x g and 4°C for 15 min. The pellets were dissolved in 8.35 ml cold CCMB (80 mM $CaCl_2$, 20 mM $MnCl_2$, 10 mM $MgCl_2$, 10 mM KOAc, 10% v/v glycerol, pH 6.4) and incubated on ice for 20 min. These were centrifuged for 10 min as above. The pellets were finally resuspended in 2.1 ml CCMB, aliquoted into ice-cold Eppendorfs (100 µl) and stored at -70°C (Hanahan *et al.*, 1991).

The prepared heat shock competent EWH331 cells were thawed on ice to which 10 ng of the different plasmids were separately added, followed by incubation on ice for 30 min. The cells were heat shocked for 90 sec at 42°C and immediately transferred to ice for 2 min. A volume of 800 µl preheated LB-glucose (20 mM glucose in LB-medium) was added to each and incubated at 37°C for 30 min with agitation (Sambrook *et al.*, 1989). The transformed cells were subsequently plated on LB-amp agar plates (1% w/v agar in LB-medium with 100 µg/ml ampicillin) by two general routes to ensure that single colonies would be obtained. Volumes of 100 µl of each of the inoculations were plated. This would result in the growth of individual colonies if the transformation reactions were highly efficient while the growth of colonies from such a small volume would be scarce if only a few cells were successfully transformed. In the latter case, to ensure that transformed cells would be plated, the cells were firstly concentrated by a short centrifugation step. The pelleted cells were resuspended in the remaining LB-glucose and plated on the agar plates.

3.4.3.2 Protein expression induction and expression

Single colonies were picked from each plate and inoculated overnight in 10 ml LB-amp (LB-medium with 50 µg/ml ampicillin) in a shaking incubator at 37°C and 250 rpm. The overnight cultures were diluted 1:100 in 500 ml LB-amp and grown at 37°C with agitation until an optical density of 0.5 at 600 nm was reached. Protein expression of the constructs containing the wild type and insert-deleted mutants were induced with 200 ng/ml anhydrotetracycline [AHT, (Institut für Bioanalytik, Germany)] for the controlled expression from the *tet* promoter. The constitutive expression of the *tet* repressor keeps the promoter in a repressed state until the repression is chemically relieved by the addition of AHT. The system is thus highly regulated and expression leakage is kept to a minimum. The cultures were subsequently grown overnight at 22°C with shaking at 200 rpm to allow recombinant expression to take place. An additional feature of the pASK-IBA3 vectors is the presence of a C-terminal *Strep*-tag (NH_2-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-OH) that can be used for the isolation of the proteins on a *Strep*-Tactin affinity column (Institut für Bioanalytik).

3.4.3.3 Isolation of the *Strep*-tagged recombinant proteins

The cells were harvested by centrifugation at 1500 x g for 30 min at 4°C in a Beckman J-6 centrifuge (Beckman, USA). The pellets were dissolved in 1 ml ice-cold Wash Buffer (150 mM NaCl, 1mM EDTA, 100 mM Tris, pH 8). Approximately 0.1 mg/ml lysozyme (Roche Diagnostics, Germany) was added to each to degrade the cell membranes followed by 0.1 mM Phenylmethylsulfonyl fluoride [PMSF, (Roche Diagnostics)] as protease inhibitor. The suspensions were incubated on ice for 30 min after which PMSF was once again added. The cells were disrupted with the use of a Sonifier Cell Disruptor B-30 instrument (Instrulab, South Africa) with an attached flat microtip by pulsing the suspension for 10 cycles of 45 sec each with 1 min rest periods in between at an output control of 3. PMSF was added to each suspension to prevent the exposed proteins from being degraded by serine proteases. The suspensions were transferred into ice-cold 10 ml ultracentrifuge tubes, balanced to the second decimal, and ultracentrifuged at 4°C for 1 hr at 100 000 x g in a Beckman Avanti J-25 centrifuge with a fixed angle rotor (Beckman). The supernatants were transferred into ice-cold sterile 10 ml tubes.

The recombinant proteins were purified from the total soluble protein extracts with *Strep*-Tactin affinity chromatography (Institut für Bioanalytik, Germany). The *Strep*-tag fused to the recombinant proteins has a high affinity for the streptavidin in the matrix. Each protein extract was loaded into a Chromabond® 15 ml PP column (Macherey-Nagel, Germany) containing a 1 cm^3 bed volume of *Strep*-Tactin beads at 4°C and was allowed to run through three times. The beads were subsequently washed three times with 10 ml Wash Buffer each (150 mM NaCl, 1mM EDTA, 100 mM Tris, pH 8) and the bound protein was finally eluted five times from the streptavidin with the same 5 ml of Elution Buffer (150 mM, NaCl, 1mM EDTA, 2.5 mM desthiobiotin, 100 mM Tris, pH 8), and collected in four fractions on ice. The desthiobiotin in the Elution Buffer reversibly competes with binding to the streptavidin and thus releases the *Strep*-tag fused proteins. The beads were finally regenerated for future use with Regeneration Buffer [150 mM NaCl, 1mM EDTA, 1mM 4-hydroxy azobenzene-2-carboxylic acid [HABA, (Sigma-Aldrich, UK)], 100 mM Tris, pH 8]. HABA, in turn, displaces the desthiobiotin from the affinity beads such that it can be used for the next protein isolation experiment. The protein samples were kept at 4°C until further use.

3.4.3.4 Protein concentration determination

The concentrations of the isolated proteins were determined by means of the Bradford assay (Bradford, 1976). Binding of the dye to the basic and aromatic residues in the protein leads to an absorbency shift from 465 to 595 nm, which can be detected by a spectrophotometer. Bovine serum albumin [BSA, (Promega, USA)] was used as the standard protein to set up a standard curve from which the unknown protein's concentration could be determined. A 1

mg/ml stock BSA solution was used to prepare a standard dilution series consisting of the following concentrations in µg/ml: 200, 100, 50, 25, 12.5 and 6.25. A 96 well ELISA plate was used to which 150 µl of the Bradford dye (Pierce, USA) was added per 50 µl of each standard dilution and the isolated protein fractions in duplicate. The A595 was read with a Multiskan Ascent scanner (Thermo Labsystems, USA) and the trend line resulting from the standard curve of absorbency at 595 *versus* standard dilution was deemed reliable if the R^2 value was >95%.

3.4.3.5 SDS-PAGE analysis

The isolated proteins were separated with SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970). The gel was composed of a 7.5% separating (7.5% w/v acrylamide, 0.375 M Tris-HCl, pH 8.8, 0.1% w/v SDS) and 4% stacking gel (4% w/v acrylamide, 0.13 M Tris-HCl, pH 6.8, 0.1% w/v SDS). The samples were diluted 4:1 in sample buffer (0.06 M Tris-HCl, pH 6.8, 0.1% v/v glycerol, 2% w/v SDS, 0.05% v/v β-mercaptoethanol, 0.025% v/v bromophenol blue) and boiled for 5 min. Electrophoresis was carried out with SDS electrophoresis running buffer (0.25 M Tris-HCl, pH 8.3, 0.1% w/v SDS, 192 mM glycine) in a Biometra electrophoresis system (Biometra, Germany). A 10-200 kDa protein ladder (Fermentas, Canada) was loaded as a molecular size marker together with each prepared protein sample. The electrophoresis was run at 60 V until the front reached the separating gel after which the voltage was increased to 100 V.

The protein bands were visualised with silver staining, a highly sensitive method for the detection of low protein quantities on a gel. The gels were firstly fixed for 30 min with 30% v/v ethanol and 10% v/v acetic acid followed by gel sensitisation for 30 min with 30% v/v ethanol, 0.5 M NaOAc, 0.5% v/v glutaraldehyde and 0.2% w/v $Na_2S_2O_3$. The gels were subsequently washed three times for 10 min each with $dddH_2O$. Protein bands were silver stained for 30 min with 0.1% w/v $AgNO_3$ and 0.25% v/v formaldehyde and washed briefly with $dddH_2O$. Staining development was carried out in 2.5% w/v Na_2CO_3 and 0.01% v/v formaldehyde and terminated with 0.05 M EDTA [adapted from (Merril *et al.*, 1981)]. Densitometric analyses to determine the relative quantities of each band on the gels (Appendix A and B, Figures 1A and 1B) were obtained with a VersaDoc Imaging System Model 3000 (BioRad, USA) and the Quantity One™ Software programme.

3.4.3.6 MS analysis of the mutant PfAdoMetDC/ODC proteins

The pA/Owt, pA/OΔA2 and pA/OΔA3 vectors were recombinantly expressed and purified as described in section 3.4.3 and separated on a 7.5% SDS-PAGE (section 3.4.3.6). The proteins were visualised with colloidal Coomassie staining. This method of staining is based on the colloidal properties of Coomassie Brilliant Blue G-250 whereby a shift in the dye

occurs from the colloidal form to the molecular dispersed form upon the addition of methanol that enables the dye to disperse into the gel. The proteins are stained with very high sensitivity while background staining is kept at a minimum (Neuhoff *et al.*, 1988). A 0.1% w/v Coomassie Brilliant Blue G-250 solution (Sigma-Aldrich, UK) containing 10% w/v ammonium sulphate and 2% v/v phosphoric acid was prepared and subsequently diluted 4:1 with methanol. The diluted solution was used to stain the gel for at least 24 hrs after which it was briefly washed in 25% v/v methanol and 10% v/v acetic acid. The gel was destained with 25% v/v methanol for approximately 30 min. The stained gel was dried and the desired protein bands were cut from the gel and sent to Dr P.F.G. Sims (Manchester Interdisciplinary Biocentre, UK) for mass spectrometry (MS) analysis.

3.4.3.7 AdoMetDC and ODC activity assays

The various protein samples were tested for their PfAdoMetDC and PfODC activities in the bifunctional protein and compared to those of the wild type activities. These activities were determined by measuring the release of the ^{14}C-labelled reaction product, $^{14}CO_2$. PfAdoMetDC uses the substrate S-adenosyl-[carboxy-^{14}C]methionine [56.2 mCi/mmol, (Amersham Biosciences, UK)] while PfODC uses L-[1-^{14}C]ornithine [47.7 mCi/mmol, (Amersham Biosciences)] (Krause *et al.*, 2000; Müller *et al.*, 2000; Birkholtz *et al.*, 2004). The PfAdoMetDC and PfODC protein activity determination reaction setup is shown in Table 3.3.

Table 3.3: The AdoMetDC and ODC radioactivity assay reactions

Reagent	Final concentration	
	AdoMetDC	ODC
AdoMet	99 µM	-
Ornithine	-	90 µM
[^{14}C]AdoMet (10x dilution)	12.5 nCi (1 µM)	-
[1-^{14}C]ornithine	-	125 nCi (10 µM)
DTT	1 mM	1 mM
EDTA	1 mM	1 mM
PLP		40 µM
KH_2PO_4 (pH 7.5)	40 mM	-
Tris-HCl (pH 7.5)	-	40 mM
Enzyme	~5 µg	~5 µg
Total volume	**250 µl**	

Decarboxylase enzyme activity determination reactions were performed in 50 ml glass tubes in duplicate for each insert-deleted mutant protein together with its corresponding wild type PfAdoMetDC/ODC protein. The blank reactions contained H_2O. Whatman no.2 filter papers (Merck, Germany) were folded lengthwise and inserted into 2 ml open-ended inner tubes to which 40 µl hydroxide of hyamine (PE Applied Biosystems, USA) was added to trap the released $^{14}CO_2$ in the form of hyamine carbonate. The inner tubes were then placed inside the glass tubes and sealed with rubber stoppers. The reactions were allowed to take place at

37°C for 30 min in a shaking water bath and terminated by the precipitation of the proteins with the injection of 500 µl of 30% w/v trichloro-acetic acid. The acidic mixtures were neutralised with ~40 mM $NaHCO_3$ while the free $^{14}CO_2$ was diluted at the same time with the release of excess unlabeled CO_2. The tubes were once again incubated for 30 min at 37°C. The filter papers were separately transferred to 4 ml Pony-Vial H/I tubes (PE Applied Biosystems) to which 4 ml of Ultima Gold XR scintillation fluid (PE Applied Biosystems) was added. The radioactivity was counted with a Tri-Carb series 2800 TR liquid scintillation counter (PE Applied Biosystems). The results were analysed with the QuantaSmart™ Software programme and were calculated as the mean of three independent experiments carried out in duplicate and expressed as specific activity (SA) in nmol/min/mg. The results of the mutant specific activities are also expressed as a percentage of the wild type SA.

The specific activity is calculated as follows (the relative quantity of each protein was taken into account to calculate the amount of protein in the assay sample, Appendix A and B):

$$SA = \frac{\text{CPM x nmol substrate}}{\text{mg protein x min x total CPM}}$$

In order to account for the smaller size of the mutant proteins, the normalised activities to the full-length wild type proteins were calculated with the following formula:

$$\% \text{ Activity} = SA \times \frac{\text{MW mutant}}{\text{MW AdoMetDC/ODC}} \times 100$$

All of the above experiments were also repeated with the plasmids containing the disrupted secondary structures within the A2 and A3 parasite-specific inserts.

3.4.3.8 Statistical analysis

The significance of the results was calculated with a non-directional two-tailed t-Test assuming unequal variances. Results with *p* values smaller than 0.05 meant that there was a significant difference between the control and sample population.

3.5 Results and Discussion

3.5.1 The identification of structural features within the PfAdoMetDC inserts

Based on the results from secondary structure prediction programmes, a conserved nine amino acid α-helix was identified within insert A2, which contains six amino acids that show preference for helix formation (Figure 3.3, A2). Three possible secondary structures could be seen when insert A3 was investigated. The first β-sheet (residues 268-272) consists of five

residues of which four favour β-sheet formation. This structure was, however, not investigated due to its small size. The second, larger eight-residue sheet (residues 284-291) of which five amino acids have high propensities to form β-sheets was considered as a structure that may have an important function within the A3 insert. While the last sheet consisting of residues 387-392, of which only three show weak preference for β-sheet formation, was also not investigated (Figure 3.3, A3).

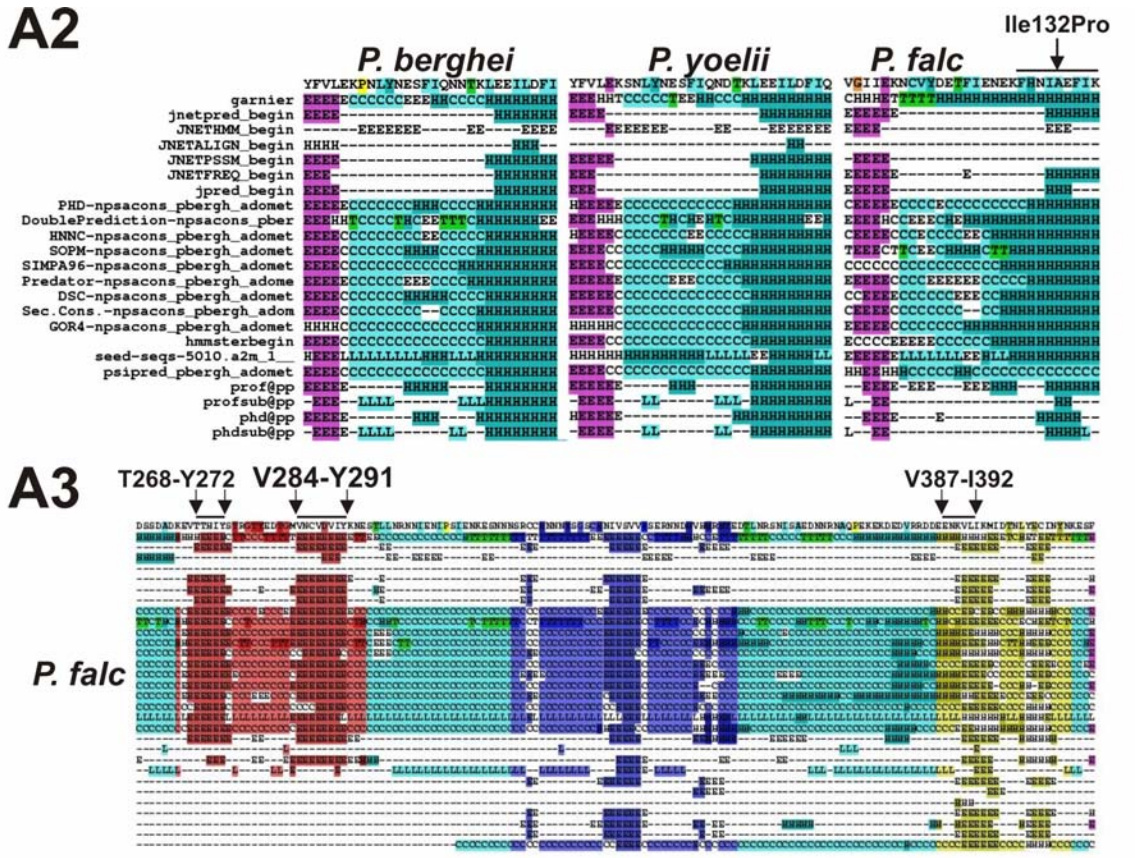


Figure 3.3: Secondary structure prediction results of the A2 and A3 insert sequences.
The legends on the left list the secondary structure prediction programmes that were used. The α-helix within A2 is indicated with a black bar. The arrow shows the position where a helix breaker (Ile132Pro) can be introduced during site-directed mutagenesis. β-sheets within insert A3 are indicated with black bars in the bottom panel. The arrows indicate the start and end positions of the areas. For simplicity, only the *P. falciparum* sequence is shown. Dashed lines are areas where secondary structures are absent and correspond to loops. Abbreviations: H, helix; C/L, coils; E, β-sheet; T, turns.

HCA diagrams of the A2 and A3 parasite-specific inserts confirmed the results obtained with the secondary structure prediction programmes above (Figure 3.3). Figure 3.4 gives the amino acid sequences of the inserts together with the two-state codes of these. The A2 insert contains one particular structure encoded by 10010011 (Figure 3.4, A2). This coding sequence represents a frequently encountered α-helix cluster consisting of eight amino acids, four of which are hydrophobic (Callebaut *et al.*, 1997). The cluster also starts and ends with hydrophilic residues and is separated from the previous cluster with four non-

hydrophobic residues. For insert A3, two specific clusters could be identified (Figure 3.4, A3). These structures most likely correspond to β-sheets as they are arranged in vertical-type clusters; they are also once again flanked with hydrophobic residues and are distinct as they are separated from adjacent structures with at least four hydrophilic residues (Callebaut *et al.*, 1997).

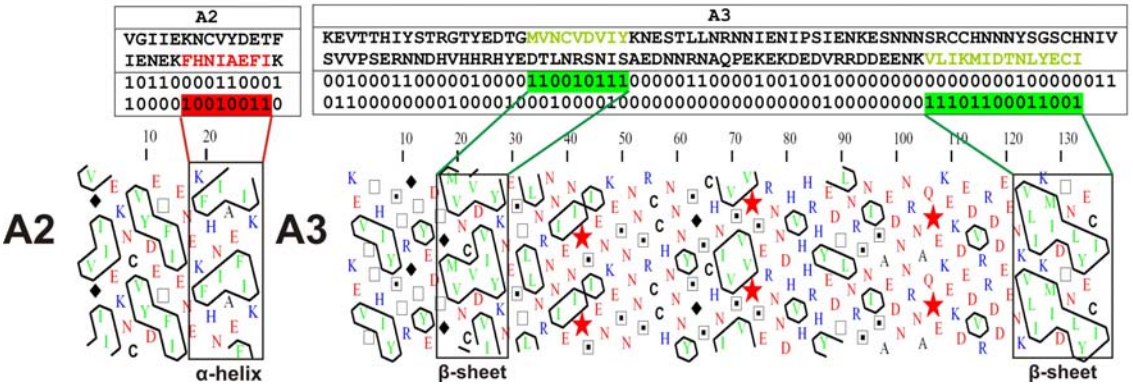


Figure 3.4: HCA plots for the A2 and A3 parasite-specific inserts in *P. falciparum* AdoMetDC/ODC.
The HCA plot of the A2 parasite-specific insert is given on the left while the plot for A3 is given on the right. The amino acid sequences and HCA codes of the two respective inserts are given above the plots. The strong and moderate hydrophobic amino acids (Val, Ile, Leu, Phe, Met, Trp, Tyr) are grouped together in clusters and shown in green, the basic amino acids (Lys, His, Arg) are shown in blue, while the acidic amino acids (Asp, Glu) are shown in red. Particular amino acids are highlighted: a star for Pro, a diamond for Gly and squares with a dot or not for Ser and Thr, respectively. The boxes indicate the secondary structures in the two plots: an α-helix in A2 (red) and two β-sheets in A3 (green) (Callebaut *et al.*, 1997).

Based on these results obtained with both secondary structure prediction programmes (Figure 3.3) and HCA (Figure 3.4), two structures were further investigated: the α-helix in A2 and the second β-sheet in insert A3. The third β-sheet in A3 is positioned close to the edge of the parasite-specific insert and its folding might therefore involve residues outside of the 137 residue insert. In addition, this β-sheet in insert A3 predicted with HCA was also identified with secondary structure prediction programmes above but these showed that the cluster is interspersed with several α-helix preference residues and was thus not further investigated.

3.5.2 Deletion of the parasite-specific inserts in PfAdoMetDC

The ~3100 bp pASK-IBA vector together with the ~4300 bp wild type *PfAdoMetDC/ODC* gene gives a total linear construct size of ~7400 bp (A/Owt). Deletion of the A1 and A2 inserts (A/OΔA1 and A/OΔA2) results in PCR products of roughly the same size as the wild type (Figure 3.5).

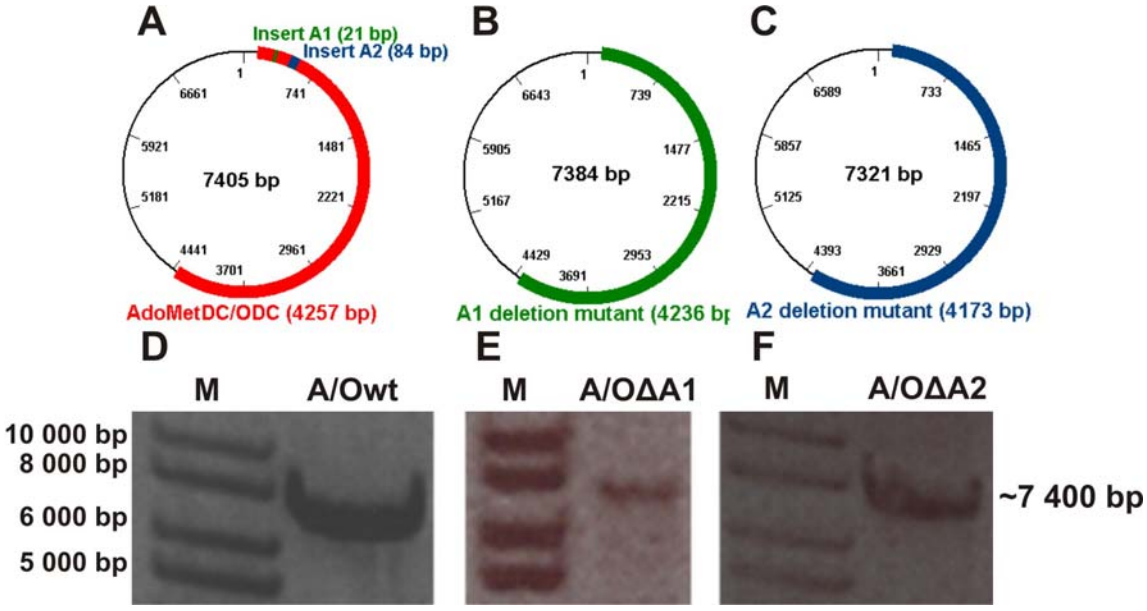


Figure 3.5: Agarose gel electrophoresis of the A/OΔA1 and A/OΔA2 PCR products.
Schematic vector diagrams of the wild type (A), the 21 bp deletion mutant A/OΔA1 (B) and the 84 bp deletion mutant A/OΔA2 (C) *PfAdoMetDC/ODC* genes inserted into pASK-IBA3 vectors (vector size ~3100 bp). The agarose electrophoresis gel of the deletion mutagenesis PCR products is given in the bottom panel showing the A/Owt (D), A/OΔA1 (E) and A/OΔA2 (F) PCR products of ~7400 bp. (M) 1 kb DNA marker.

The deletion of the A1 and A2 parasite-specific inserts removed 21 and 84 nt, respectively, and thus resulted in PCR products of approximately 7400 bp (Figure 3.5 E and F), which corresponds to the size of wild type template (Figure 3.5 D). The independent deletion mutagenesis results of each of the A1 and A2 inserts were verified with nucleotide sequencing (Figure 3.6). The deletion of the largest A3 parasite-specific was performed in Chapter 2 (Figure 2.3).

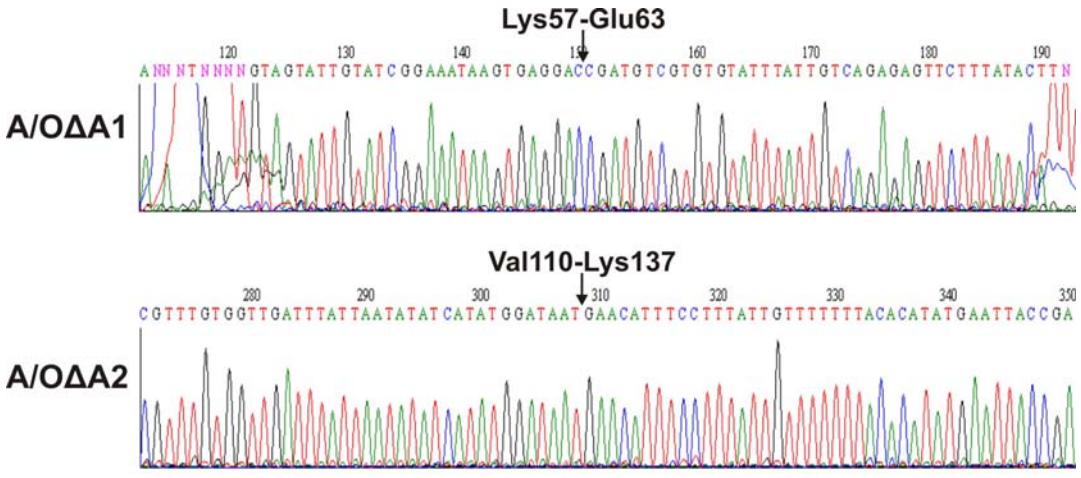


Figure 3.6: Nucleotide sequencing chromatograms of the PfAdoMetDC/ODC insert-deleted constructs.
The arrows indicate the positions where the A1 (top panel, A/OΔA1) and A2 (bottom panel, A/OΔA2) parasite-specific inserts were deleted.

Automated nucleotide sequencing of the A/OΔA1 clones resulted in two out of five mutant clones (40% mutagenesis efficiency) while the deletion of the A2 insert with the overlapping primer method was only 20% successful. All three of the parasite-specific inserts were thus successfully removed (Figure 2.3 for A/OΔA3 and Figure 3.6 for A/OΔA1 and A/OΔA2), which can be used to determine the individual effect of each insert on the activity of the corresponding as well as the neighbouring decarboxylase domain.

3.5.3 Mutagenesis of delineated areas within the PfAdoMetDC parasite-specific inserts

The PCR products obtained with the site-directed mutagenesis reactions were visualised with agarose gel electrophoresis. Once again a DNA band of ~7400 bp was expected for each of the mutagenesis products (Figure 3.5 A), which can be seen for both the A/OpA2a and A/OΔA3b products below (Figure 3.7 B and C, respectively).

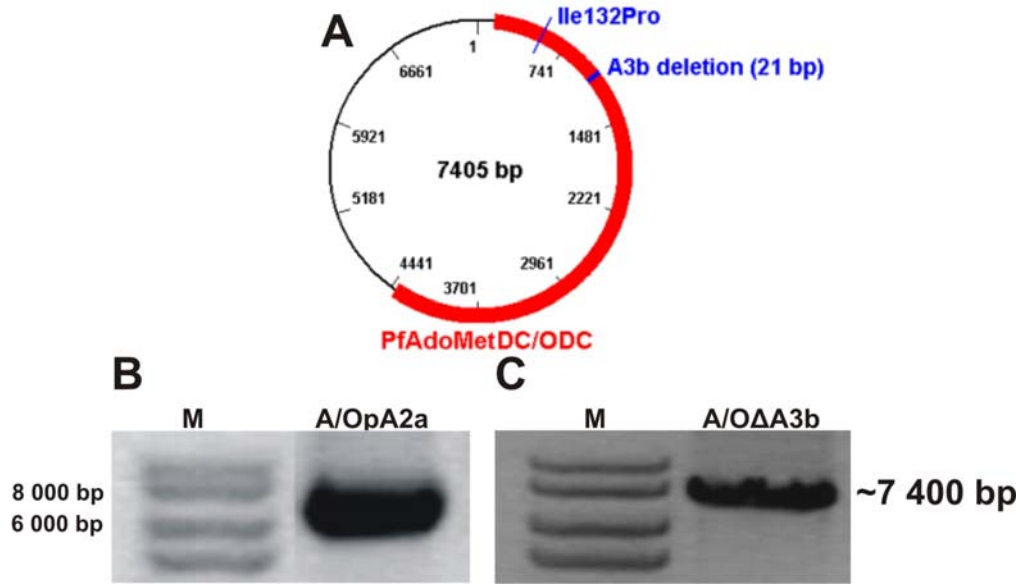


Figure 3.7: Agarose gel electrophoresis of the A/OpA2a and A/OΔA3b PCR products.
(A) A schematic diagram of *PfAdoMetDC/ODC* (~4300 bp) cloned into pASK-IBA3 (~3100 bp). The location of the Ile132Pro point mutation and A3b deletion is indicated in blue. (B) 1% Agarose gels were used to visualise the A/OpA2a (B) and A/OΔA3b (C) PCR products with the expected sizes of ~7400 bp. (M) 1 kb DNA ladder.

Nucleotide sequencing of the mutant clones was subsequently performed to determine whether the point mutation at amino acid position 132 was successfully incorporated and if the β-sheet in insert A3 was deleted. The K215A2 reverse primer was used to detect the presence of the Ile132Pro point mutations in the five clones. The reverse complement of the nucleotide sequences shows the CCA mutation as TGG, which was seen in three out of the five clones analysed giving a 60% mutagenesis efficiency with the overlapping primer

method (Figure 3.8). The K215A1 sequencing primer was used for the verification of the β-sheet deletion within insert A3. The sequencing results obtained indicated that the sheet was once again absent in three out of the five clones analysed also resulting in a 60% mutagenesis efficiency (Figure 3.8).

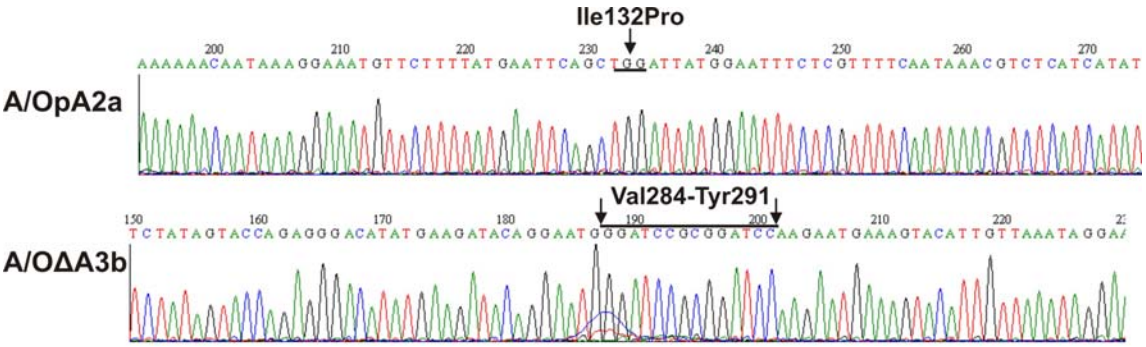


Figure 3.8: Nucleotide sequencing chromatograms of the A/OpA2a and A/OΔA3b constructs. The black bar indicates the position of the Ile132Pro point mutation within pA/OpA2a (top panel). The longer black bar in the bottom panel indicates the position of the *Bam*HI restriction enzyme sites that were inserted in place of the A3 parasite-specific insert in pA/OΔA3b with the RE-mediated inverse PCR method.

3.5.4 Influence of the PfAdoMetDC inserts on bifunctional activity

3.5.4.1 Isolation of the mutant PfAdoMetDC/ODC proteins

The wild type and insert-deleted *PfAdoMetDC/ODC* sequences cloned into the pASK-IBA3 vectors were transformed into heat shock competent *E. coli* EWH cells. The recombinant proteins were subsequently isolated via cell disruption, ultracentrifugation and *Strep*-tag affinity chromatography. There was no significant difference between the amount of isolated protein of the wild type and the mutant samples ($p > 0.05$) indicating that the inserts may not have had an influence on the overall expression levels of the recombinant proteins in the soluble fractions (Figure 3.9). The average total amount of protein obtained from 1L culture per sample was 200 µg.

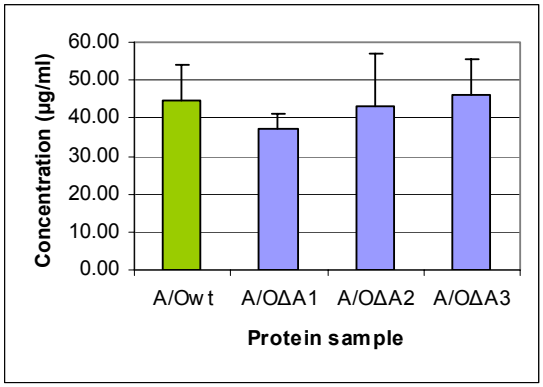


Figure 3.9: Protein expression levels of the wild type PfAdoMetDC/ODC and the PfAdoMetDC insert-deleted proteins.
The wild type protein (A/Owt) is shown in green and the mutant ones in blue (A/OΔA1, A/OΔA2 and A/OΔA3). The concentrations were determined from three independent experiments. The standard deviations of the mean are indicated as error bars on each graph.

The recombinantly expressed proteins were analysed with SDS-PAGE. Protein bands of ~160 kDa (heterodimer: ~150 kDa consisting of covalently linked PfAdoMetDC/ODC and the 9 kDa β-subunit of PfAdoMetDC) were obtained after recombinant protein expression and isolation of the wild type PfAdoMetDC/ODC as well as the PfAdoMetDC insert-deleted proteins (Figure 3.10). The amount of each protein loaded onto the SDS polyacrylamide gel averaged at 2.5 µg per lane, but this did not reflect the amount of the desired ~160 kDa protein. Inspection of the PAGE result showed that the amount of ~160 kDa protein decreased from the A/Owt to the A/OΔA3 protein sample while the intensities of contaminating proteins were the highest for the insert-deleted mutated protein samples (Appendix A, Table 1A). Lower amounts of *Strep*-tagged purified ~160 kDa protein in the mutant samples could not be attributed to the presence of misfolded proteins in inclusion bodies as this would have resulted in the absence of these proteins in the soluble protein fractions and a significant decrease in the total amount of protein isolated. In a previous study the crude *E. coli* lysates were also visualised with SDS-PAGE and the results showed that these proteins are absent in the *Strep*-tag purified proteins samples [results not shown, (Niemand, 2007)].

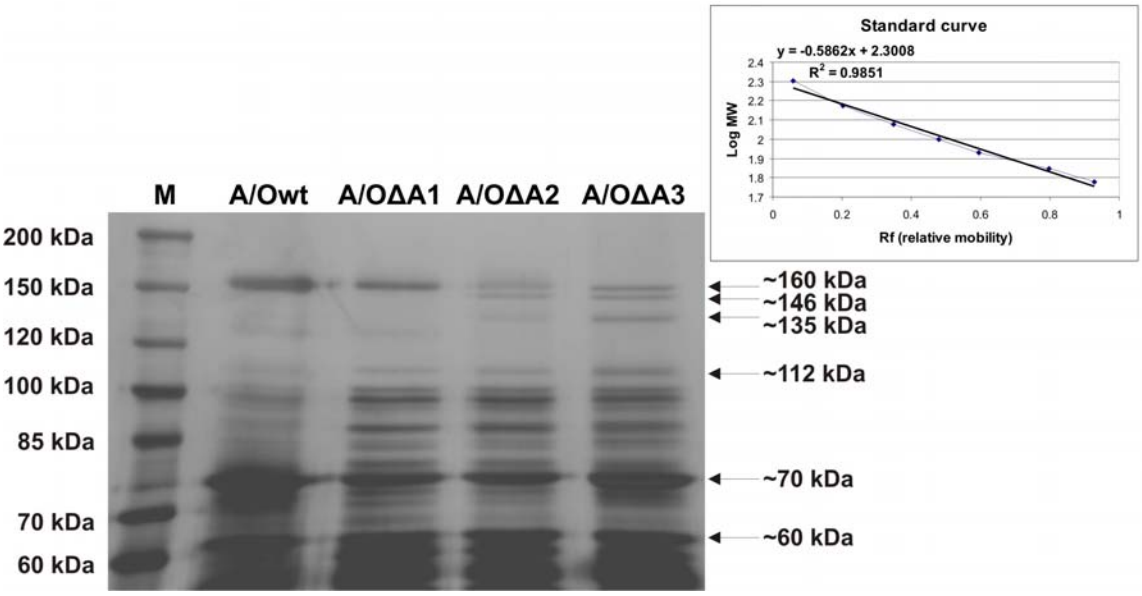


Figure 3.10: SDS-PAGE analysis of the expressed wild type and the bifunctional PfAdoMetDC insert-deleted recombinant proteins.
The protein size-determining marker is shown in lane M. The denatured bifunctional heterodimeric protein is seen as a ~160 kDa protein band. Calibration curve: distance moved against logMW for each of the standard proteins in lane M.

The expression and purification of *P. falciparum* proteins is a common obstacle for many researchers in the field and may be attributed to various factors such as the high A+T content of the malarial genome, the increased sizes of the proteins due to the high occurrence of LCRs and parasite-specific inserts and extensive RNA secondary structures (Pizzi and Frontali, 2001; Gardner *et al.*, 2002; Birkholtz *et al.*, 2003; Mehlin *et al.*, 2006). It is thus not surprising that these proteins are difficult to express in adequate and pure quantities in heterologous protein expression systems. Mehlin *et al*. recently expressed 1000 *P. falciparum* proteins in *E. coli* but managed to obtain soluble quantities of only 63 proteins that were sufficient for subsequent purification. The size, degree of disorder and pI of these proteins negatively influenced their soluble expression in the heterologous system (Mehlin *et al.*, 2006). Furthermore, the stretches of A and T bases in the genes often result in frame shift mutations and false internal start codons leading to truncated protein products of smaller sizes.

In a previous study, apart from the correctly sized ~160 kDa PfAdoMetDC/ODC protein band, three major proteins sized ~112 kDa, ~70 kDa and ~60 kDa were obtained with denaturing PAGE, which can also be seen in the figure above. Various experiments were subsequently performed in this laboratory in order to identity these contaminating proteins (Niemand, 2007). Size-fractionation studies suggested that contaminating proteins associate with the full-length PfAdoMetDC/ODC protein to form high molecular weight complexes of ~600 and ~400 kDa, which is in-line with previous studies that showed that the PfODC domain

mediates the dimerisation of the bifunctional protein via an aromatic amino acid zipper, several hydrophobic interactions and a salt bridge (Birkholtz *et al.*, 2003; Birkholtz *et al.*, 2004). Any additional proteins containing PfODC would thus form tight interactions with PfODC in the main protein and result in the co-elution of these proteins during *Strep*-tag affinity chromatography. The ~60 kDa protein was not detected with *Strep*-tag Western immunodetection and was thus regarded as either a C-terminal truncated protein or a host-derived protein.

Explanations for the presence of the ~112 and ~70 kDa *Strep*-tagged proteins include post-translational degradation, in frame ribosomal slippage due to the high abundance of mRNA secondary structures and false initiation codons. Ribosomal slippage on mRNA secondary structures as the cause of the differently expressed proteins was excluded as these expressed contaminating proteins was found to only possess C-terminal *Strep*-tags and no N-terminal tags. Subsequent analysis of the *PfAdoMetDC/ODC* gene revealed the presence of two possible prokaryotic Shine-Dalgarno-like sequences downstream of start codons, which may lead to the translation of the ~112 and ~70 kDa proteins from internal mRNA initiation sites to produce the N-terminally truncated proteins. These internal sequences have no effect on the *in vivo* translation of the proteins within the eukaryotic host but it can cause the expression of these truncated proteins from these sites when using the prokaryotic *E. coli* expression system. MS analysis of the major contaminating bands revealed some interesting results: the ~112 kDa protein is an N-terminal truncation of the full-length protein and the ~70 kDa and 60 kDa bands were identified as *E. coli* 70 kDa (with some traces of N-terminally truncated PfAdoMetDC/ODC) and 60 kDa heat shock proteins (HSP), respectively. The purification of these two proteins with the full-length protein implies that these chaperones are tightly bound to PfAdoMetDC/ODC, possibly in an effort to cope with the expression and folding of the recombinant protein. Overall, these studies revealed that the presence of internal translation initiation sites within the gene sequence results in the production of a multi-component protein complex and thus prevents the homogenous purification of a recombinant protein (Niemand, 2007).

A possible explanation for the higher intensity of contaminating proteins in the soluble fraction of the mutated protein samples in Figure 3.10 above, may be that the deletion of regions that are possibly necessary for their correct folding can place further stress on the expression of the recombinant proteins leading to the co-elution of chaperone proteins and possible degradation products. The presence of the contaminating proteins in the purified protein samples may interfere with downstream experiments and needs to be taken into consideration especially in the determination of the effects that the inserts have on the activities of the full-length bifunctional protein. However, the activities obtained after the

deletion of each of the parasite-specific inserts in the PfAdoMetDC domain will be normalised to the activity of the wild type PfAdoMetDC/ODC protein sample, which will even out the effects of the contaminating proteins in all the samples. At present, the only route around this problem of heterogeneous protein isolates is to re-synthesise the *PfAdoMetDC/ODC* gene to remove internal Shine-Dalgarno-like sequences together with codon harmonisation to ease the burden on the heterologous expression system thus achieving better protein expression and isolation.

The electrophoresis gel also shows the presence of additional smaller sized protein bands that appeared upon the removal of the larger parasite-specific inserts (Figure 3.10, lanes A/OΔA2 and A/OΔA3). The approximate MW of these unknown proteins in lanes A/OΔA2 and A/OΔA3 were determined with the use of a calibration curve drawn up from the Rf (relative mobility) values *versus* the logMW of the standard proteins (Figure 3.10, inserted graph). Substitution of these unknown proteins' Rf values into the linear regression equation determined these to be ~146 kDa and ~135 kDa in size. The determination of protein sizes with SDS-PAGE is, however, not a reliable method as the conformation of the proteins can interfere with their mobility through the gel matrix. The formation of these smaller proteins with correspondingly smaller sizes were subsequently analysed with MS.

3.5.4.2 MS analysis of the PfAdoMetDC insert-deleted proteins

The recombinant A/OΔA2 and A/OΔA3 proteins were separated with SDS-PAGE and visualised with an MS compatible colloidal Coomassie staining protocol. The gel was dried and the ~146 kDa shadow bands directly beneath the ~160 kDa bands in the last two lanes in Figure 3.10 were cut out and analysed with MS after trypsin digestion. Table 3.4 lists the different peptides identified with MS.

Table 3.4: Peptides identified with MS

2. Plasmodium_falciparum_3D7|MAL10|PF10_0322|Pf

Mass: 168063 **Score:** 48 **Queries matched:** 7

Annotation|Plasmodium_falciparum_TIGR|(protein coding)

S-adenosylmethionine decarboxylase-ornithine decarboxylase

Query	Observed	Mr(expt)	Mr(calc)	Delta	Score	Expect	Rank	Peptide
498	593.74	1185.47	1183.67	1.80	8	18	1	K.NNNVLLTLQR.N
616	629.24	1256.47	1255.65	0.82	3	57	10	R.YMVAASSTLAVK.I
629	635.16	1268.30	1267.58	0.72	16	4.3	1	K.NIGNNFSSSNSK.L
929	779.31	1556.60	1554.79	1.81	48	0.0012	1	K.IVVVDTNTFFDASK.R
382	545.26	1632.76	1631.87	0.89	15	5.9	1	K.IHYCTLSLQEIKK.D
828	734.93	2201.77	2202.11	-0.34	13	8	1	K.VLIKMIDTNLYECINY NK.E

All the peptides that were identified were complementary to the PfAdoMetDC/ODC sequence (Figure 3.11), indicating that the smaller protein bands seen in Figure 3.10 (lanes A/OΔA2 and A/OΔA3, sized ~146 kDa) isolated together with the A2- and A3-deleted proteins are PfAdoMetDC/ODC albeit with a very low Mascot score of only 48. The Mascot score gives the probability that the observed protein match between experimental data and a protein sequence is a random event and is regarded insignificant when it scores below 74. In this case, the low score may be attributed to the low intensity of the protein band that was tested, which was thus insufficient for effective MS analysis.

```
   1 MNGIFEGIEK RVVIKLKESF FKGNRNVNSF LDIPKELWEE KLKYIGCSIV SEISEDKNER
  61 RGERCRVYLL SESSLYIFDD SLFIKTCGKT RVLFFIPFVV DLLIYHMDNV GIIEKNCVYD
 121 ETFIENEKFH NIAEFIKEHF LYCFFTHMNY RNKTKDGYFE QEYPHKSLED EKKFFEFFFK
 181 NVQMYNTHLP MEKMHYIFFY SSDDVHMTDI ASTFKFCSEI HLFGINKYNE KNQFHDAYLN
 241 NKSLNLFTRV HEDNLKLYDS SDADKEVTTH IYSTRGTYED TGMVNCVDVI YKNESTLLNR
 301 NNIENIPSIE NKESNNNSRC CHNNNYSGSC HNIVSVVPSE RNNDHVHHRH YEDTLNRSNI
 361 SAEDNNRNAQ PEKEKDEDVR RDDEENKVLI KMIDTNLYEC INYNKESFLY NEFYFTPCGY
 421 SCNVSEKNNY FCVHYSPEDS VSYVSVEVSS NLSCDRFLDF IHKQLNFYNG KYMFMINYVF
 481 CEESNNMSKM VPDDDNNNYS SGKSCVYYQD LNKKEKEEYY RLNKKLRNDL FINSKQFYEL
 541 HTFTERTVGF MRVQYFVYKL RDVVKCVEKE TLLARSSSCL FMFNNIKRND VHDDYVTKSS
 601 NGGVIKQLTE RDVDDMYEYA LNFCKQNKIV VVDTNTFFFD ASKRKENLIK LEKVQTNEKD
 661 EYEEKDEVYR RGNNELSSLD HLDSKNNLIH MYYEKNKCDI INKDDENSTI ATNNNDNNND
 721 SSSYDKSITI SRSSSCNNSH LSYSSFDNNH GNEKMKDYIS VDENNNNNNN NKNNNVLLTL
 781 QRNSDDENGK DKDNEKNDVS LENNMEKNYK EEIWNYYTKN KVEVKTLEKV LNENIDTSVV
 841 CINLQKILAQ YVRFKKNLPH VTPFYSVKSN NDEVVIKFLY GLNCNFDCAS IGEISKVIKL
 901 LPNLSRDRII FANTIKSINS LIYARKENIN LCTFDNLDEL KKIYKYHPKC SLILRINVDF
 961 KNYKSYMSSK YGANEYEWEE MLLYAKKHNL NIVGVSFHVG SNTKNLFDFC LAIKLCRDVF
1021 DMSSNMGFNF YIINLGGGYP EELEYDNAKK HDKIHYCTLS LQEIKKDIQK FLNEETFLKT
1081 KYGYYSFEKI SLAINMSIDH YFSHMKDNLR VICEPGRYMV AASSTLAVKI IGKRRPTFQG
1141 IMLKELKDHY DPLNFAQQEN KKQDETKINH NNDNNDNNDN NDNNINNNNN NQKGGQGNIM
1201 NDLIITSTND STSKKNDHSS SQVIQNVSCT IRDKEGDNIK INTHTINNPN INGKENTVDG
1261 DNINIAHKNI GNNFSSSNSK LGNITNIKKK VVNINDNRYN YFSYYVSDSI YGCFSGIIFD
1321 EYNRCPIYVI KNKNNPNQNF MNFNLYLANV FGQSCDGLDM INSITYLPEC YINDWLLYEY
1381 AGAYTFVSSS NFNGFKKCKK VYIFPESKPS LKGQPNKHW
```

Figure 3.11: Peptides identified with MS aligned with the PfAdoMetDC/ODC sequence.
The positions of the three PfAdoMetDC parasite-specific inserts are shown in green while the peptides that were identified with MS analysis are shown in red. The methionine residues at the beginning of the protein sequences as possible translation initiation sites are highlighted in red.

The smaller contaminating proteins could be formed as a result of post-translational degradation of the expressed protein or the presence of false translation initiation sites downstream of the start codon that occur in-frame to the particular open reading frame and are exposed as a result of conformational changes upon the deletion of the inserts (Niemand, 2007). Unfortunately, these bands could not be detected with anti-*Strep* antibodies due to their close proximity to the mutated PfAdoMetDC/ODC protein (results not shown). Western analysis would indicate whether the deletions resulted in N-terminal truncations of the full-length protein, as these would still contain their C-terminal *Strep*-tags. It is also possible that the deletions result in conformational changes leading to proteolysis sites becoming accessible to proteolytic enzymes or start codons becoming accessible to the protein translation initiation complex, which are usually buried, resulting in the formation of smaller protein products. The latter is a possibility due to the high abundance of secondary structures present on mRNA strands. If Shine-Dalgarno-like sequences are situated

downstream and in close proximity to exposed Met start codons, the formation of the ~146 kDa protein could be as a result of translation initiation taking place from Met107 situated between inserts A2 and A3 while the ~135 kDa protein could be formed if translation takes place from Met184, 191 or Met194. The exact identities of these protein bands and the manner in which they are produced thus require further investigation.

3.5.4.3 Activity assays of the PfAdoMetDC insert-deleted proteins

The isolated recombinant proteins were subjected to radioactivity assays during which ^{14}C-labelled substrates were co-incubated with the isolated enzymes to generate $^{14}CO_2$. The labelled product was captured and counted with a liquid scintillation counter. The results obtained were normalised against the specific activity (nmol/min/mg) of the wild type protein and expressed as a percentage of its activity (Figure 3.12). The altered sizes of the mutated proteins (A/Owt = 156 090 Da, A/OΔA1 = 155 430 Da, A/OΔA2 = 153 120 Da and A/OΔA3 = 141 020 Da) were taken into consideration. In addition, densitometric analyses of the band intensities were performed and the percentage contribution of the bifunctional protein was used to determine the amount compared to the total protein concentration (Appendix A, Table 1A). The percentage contributions of the truncated versions of the bifunctional proteins to the total protein concentrations were not taken into consideration for the activity determinations since monofunctional PfODC has negligible activity (Krause *et al.*, 2000).

The small A1 insert occurs in all *Plasmodium* species of known sequences and has a high similarity with the other *Plasmodium* sequences (*P. yoelii*, *P. berghei*, *P. chaubaudi* and *P. knowlesi*). The second insert in the PfAdoMetDC domain is also highly conserved while insert A3 is the most divergent one (Wells *et al.*, 2006). Previous mutagenesis studies on a region encompassing insert A3 severely effected PfAdoMetDC activity but had no impact on the formation of the bifunctional complex (Birkholtz *et al.*, 2004), which is in agreement with the PfAdoMetDC homology model where this insert is positioned on the edge opposite to the dimerisation boundary of the domain. This insert, however, was identified before the PfAdoMetDC model was solved, which means that the deletion removed part of the core structure (residues 214 to 410) of the protein possibly resulting in the observed activity loss (Wells *et al.*, 2006). Multiple alignments and the homology model of the PfAdoMetDC domain resulted in the identification of three parasite-specific inserts, which forms the focus of this study.

Deletion of insert A1 within the PfAdoMetDC domain of the bifunctional protein significantly decreased this domain's activity by 47%. The results also show 98% inhibition of PfAdoMetDC activity upon the deletion of the A2 and A3 inserts (Figure 3.12).

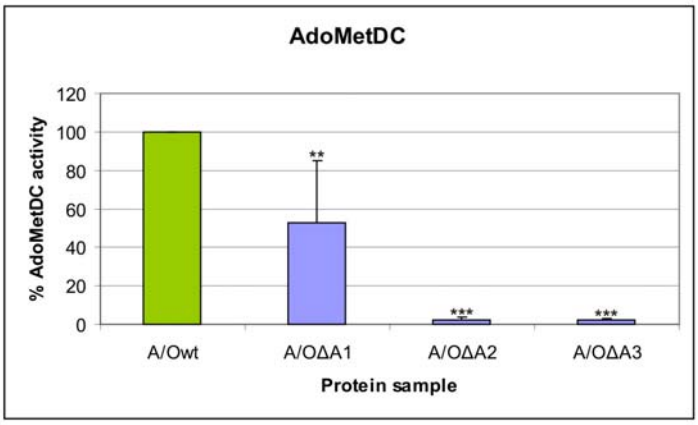


Figure 3.12: AdoMetDC activity assays of the wild type PfAdoMetDC/ODC and bifunctional PfAdoMetDC insert-deleted proteins.
Wild type A/Owt activity is shown in green at 100% while the various mutant proteins activities normalised to the A/Owt activity are shown in blue. The values were determined from three independent experiments carried out in duplicate. The standard deviations of the mean are indicated as error bars on each graph. Significant differences at a confidence level of 95% are represented as follows: ** for $p < 0.01$and *** for $p < 0.001$.

The effects of the deletion of these inserts in the bifunctional protein on the neighbouring PfODC domain were also investigated. Figure 3.13 shows the normalised PfODC activities upon the deletion of the PfAdoMetDC inserts in the PfAdoMetDC/ODC protein. From this figure, it is clear that the deletion of these inserts also resulted in significant losses in activity of the neighbouring PfODC domain where each deletion decreased PfODC activity by 95, 82 and 90%, respectively.

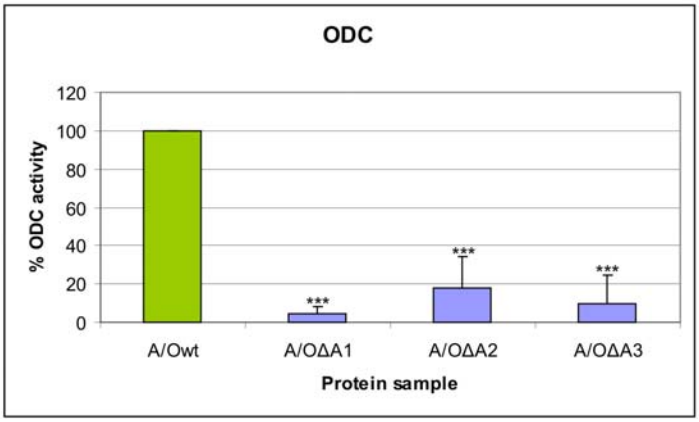


Figure 3.13: ODC activity assays of the wild type PfAdoMetDC/ODC and bifunctional PfAdoMetDC insert-deleted proteins.
Wild type A/Owt activity is shown in green at 100% while the various mutant proteins activities normalised to the A/Owt activity are shown in blue. The values were determined from three independent experiments carried out in duplicate. The standard deviations of the mean are indicated as error bars on each graph. Significant differences at a confidence level of 95% are represented as follows: *** for $p < 0.001$.

While the deletion of the seven amino acid A1 insert halved PfAdoMetDC activity, it had a greater effect on the adjacent PfODC domain with a 95% decrease in activity. This supports previous studies where the PfODC domain was shown to be effected to a greater extent than that of the PfAdoMetDC domain (Birkholtz *et al.*, 2004; Roux, 2006). The A1 insert is similar

to O1 in the sense that it is short and highly conserved amongst the *Plasmodium* species. Previous studies on the O1 insert showed that this highly conserved area, positioned close in linear amino acid sequence to the neighbouring domain and hinge region, plays essential roles in the activities of both domains in the bifunctional protein and also mediates the association of the monofunctional PfAdoMetDC and PfODC proteins into an active bifunctional heterotetrameric protein (Birkholtz *et al.*, 2004). It is thus interesting to note that, despite the linear location of the conserved A1 insert close to the N-terminus of the bifunctional protein (residues 57-63), its deletion resulted in only 47% active PfAdoMetDC protein and 5% active PfODC protein, which may be due to similar effects as seen for O1. The reason for such a severe loss in activity might be due to conformational changes communicated to the PfODC domain with direct consequences on the function of the active site or the dimerisation of this domain via protein-protein interactions. Studies by Wrenger *et al*. showed that the active sites of the two domains function independently (Wrenger *et al.*, 2001). However, here the removal of a *Plasmodia* conserved parasite-specific insert in the PfAdoMetDC domain had, in terms of catalytic activity, a detrimental effect on the PfODC domain corroborating reports by Birkholtz *et al*. that interdomain interactions are essential for both activities in the bifunctional complex (Birkholtz *et al.*, 2004).

The two smallest, non-LCR containing inserts namely A1 and O1 are similar in their size, in their conservation across the *Plasmodium* species and their apparent functional roles in the domain activities of PfAdoMetDC/ODC. The content of positively charged Lys and Asn residues in inserts A1 and O1, which are often involved in protein-protein interactions via hydrogen bond formation, are also considerably high at 30 and 25%, respectively (Singh *et al.*, 2004). These regions therefore need to be investigated further as sites for possible inhibitor targeting, which could result in decreased levels of the polyamines within the parasites.

The deletion of the larger A2 and A3 inserts mostly depleted both PfAdoMetDC (98% activity inhibition) and PfODC (82% and 90% inhibition, respectively) activities in the bifunctional protein. Previous deletion mutagenesis studies on the A1 (as identified before the homology model of PfAdoMetDC was solved), O1 and O2 inserts in the bifunctional protein effected not only the activity of the domain within which the insert resides but also the activity and/or conformation of the neighbouring domain. The surface localisations of these inserts imply that they may participate in intra- as well as interdomain interactions, which are brought about by long-range interactions propagated throughout the full-length protein between the surface localised inserts and the respective active sites (Myers *et al.*, 2001; Yuvaniyama *et al.*, 2003). A 3-dimensional structure of the PfAdoMetDC domain together with the parasite-specific inserts where the 3-dimensional location of the inserts are shown will shed more light

on the manner in which these unique areas interact with the remainder of the protein. These inserts might also be involved in the dimerisation of the PfODC monomers, which would explain the severe depletion of this domain's activity. The two active sites of the PfODC domain are formed as a result of its dimerisation and an interference with this conformation would thus lead to an inactive enzyme (Müller *et al.*, 2000).

The results therefore suggest that the parasite-specific inserts in the PfAdoMetDC domain play important roles in the catalytic activities of both domains, which is in agreement with studies performed on the hinge and the PfODC domain inserts where the latter domain was shown to be more dependent on interdomain interactions than PfAdoMetDC (refer to Figure 3.1) (Birkholtz *et al.*, 2004; Roux, 2006). Similar observations have been made in another *P. falciparum* bifunctional protein; monofunctional TS, which is located at the C-terminal of DHFR/TS, is only active within the hybrid complex with the N-terminal DHFR domain (Shallom *et al.*, 1999). It is therefore highly plausible that the bifunctional arrangement of PfAdoMetDC/ODC and DHFR/TS is to maintain optimal catalytic activities and/or stabilities of their C-terminal domains via interdomain interactions across the entire protein, including the inserts. This is supported by experimental evidence that the N-terminal domains of both these bifunctional proteins are functionally active in the absence of their respective C-terminal domains but not *vice versa* (Shallom *et al.*, 1999; Wrenger *et al.*, 2001).

The presence of a number of contaminating proteins following affinity chromatography may also be interfering with the bifunctional protein by forming complexes with it thereby reducing the amount of the desired active protein. These unwanted proteins might be interfering with the effectiveness of the radioactivity assays and need to be taken into consideration when analysing the results. The normalised activities of the mutant proteins are, however, a means of standardising the influence of contaminating proteins.

A possible explanation for the presence of parasite-specific inserts in PfAdoMetDC/ODC and the high abundance of charged residues therein might thus be their involvement in intramolecular protein-protein interactions and the stabilisation of the active, heterotetrameric bifunctional complex. Information on the involvement of the inserts in the PfAdoMetDC domain on complex formation are, however, still outstanding. The possibility exists that the extreme decreases in protein activities after large deletions of entire parasite-specific inserts are due to general conformational changes and cannot be excluded. The deletions could result in unstable proteins or multiple conformations thereof and it was thus hypothesised that the disruption of definite structures within the parasite-specific inserts would be a more sensible approach to study the roles of these areas.

Since a previous study focused on the predicted secondary structures in the hinge and PfODC parasite-specific inserts (Roux, 2006), the specific roles of the inserts in the PfAdoMetDC domain in the activity of both proteins, whether through multiple conformations interfering with discrete interdomain interactions or the removal of such large areas from the protein, needed to be investigated further. The following section will look at delineated areas within the parasite-specific inserts in this domain as possible sites for protein-protein interactions.

3.5.5 The roles of secondary structures within the PfAdoMetDC inserts on bifunctional activity

3.5.5.1 Recombinant protein expression of the newly created constructs

Due to unforeseen experimental complications, protein purification and subsequent activity analysis on the A/OΔA3b mutant protein is still in progress. The proteins encoded by the wild type pA/Owt and α-helix disrupted pA/OpA2a plasmids were recombinantly expressed and isolated with *Strep*-tag affinity chromatography. As can be seen in Figure 3.14, there was once again no significant difference between the expression levels of the wild type (A/Owt) and mutant (A/OpA2a) proteins ($p > 0.05$), which averaged at 70 µg/ml. The total amount of protein obtained from 1L culture was approximately 350 µg.

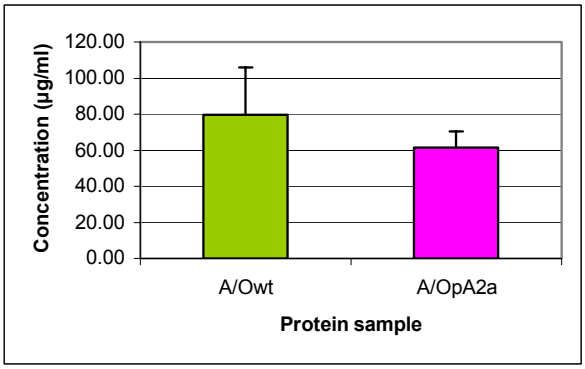


Figure 3.14: Protein expression levels of the wild type PfAdoMetDC/ODC and A/OpA2a proteins.
The wild type protein (A/Owt) is shown in green and the mutant (A/OpA2a) in pink. The concentrations were determined from three independent experiments. The standard deviations of the mean are indicated as error bars on each graph.

The correct size protein bands of ~160 kDa were obtained after recombinant protein expression and isolation of the wild type PfAdoMetDC/ODC as well as the helix-disrupted A/OpA2a protein. The amount of protein loaded onto the SDS-polyacrylamide gel averaged at 2.1 µg per lane (Figure 3.15). As can be seen from the figure below, even though similar

amounts of total protein were eluted and loaded onto the SDS-PAGE, the amount of the mutant ~160 kDa A/OpA2a protein was much lower than the A/Owt full-length protein. The intensity of the co-eluted contaminating proteins (especially at ~70 kDa) during A/OpA2a isolation is again much higher possibly as a result of recombinant protein degradation and the binding of host chaperone proteins to assist with the folding of the protein (discussed in section 3.5.3.1). The intensities of each band on the SDS-PAGE gel are given in Table 1B (Appendix B) as a percentage of the total protein amount loaded per gel lane.

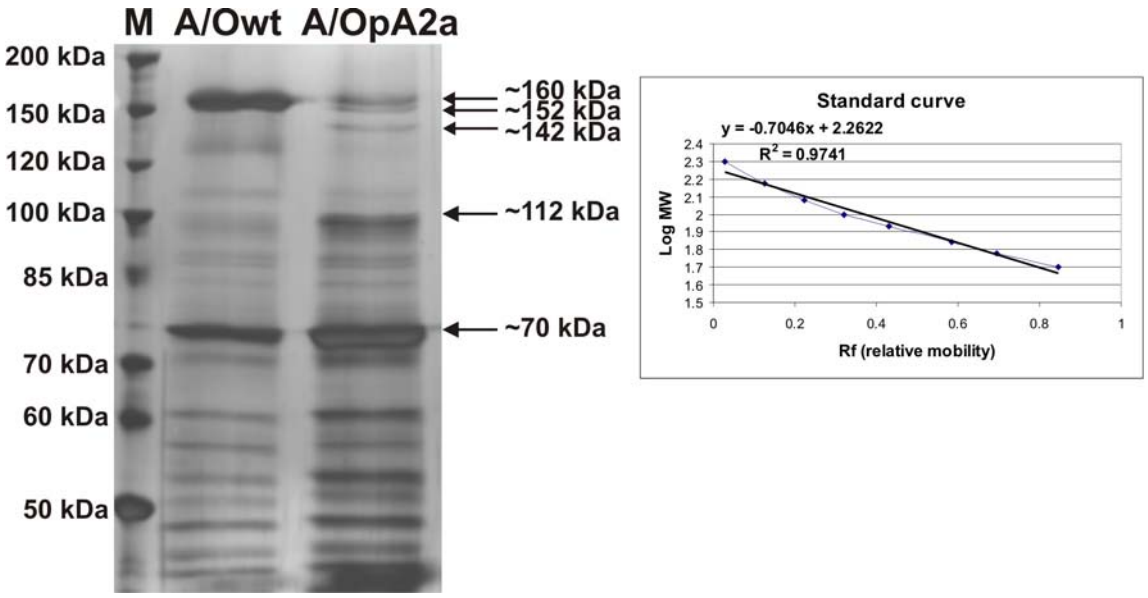


Figure 3.15: SDS-PAGE analysis of the wild type and A/OpA2a recombinant proteins.
(M) 10 to 200 kDa protein ladder. Calibration curve: distance moved against logMW for each of the standard proteins in lane M.

The disruption of the helix in insert A2 resulted in definitive increases in especially the amount of ~112 kDa truncated PfAdoMetDC/ODC and ~70 kDa HSP70/truncated PfODC proteins (Figure 3.15). It seems that, in comparison to the SDS-PAGE profile of the wild type protein, the mutated protein is placing severe stress on the heterologous system for its expression. The helix may be involved in protein-protein interactions throughout the full-length protein, which is abolished upon its disruption resulting in correlative increases in the amount of *E. coli* chaperone protein in an attempt to aid the correct folding of the recombinant protein. Once again unknown proteins of smaller sizes can be seen directly underneath the mutant protein (Figure 3.15 lane A/OpA2a sized ~152 and ~142 kDa), which slightly differs in size from those calculated in Figure 3.10, possibly due to the differences in resolution between the two gels and the mere estimation of sizes with a Rf calibration curve. These ~150 kDa PfAdoMetDC/ODC protein bands (identified with MS in section 3.5.3.2) could thus be produced as a result of any changes that are made within the region of the A2 and A3 inserts.

As mentioned previously, the *cis-trans* isomerisation of the Pro residue inserted to disrupt a helix could lead to a conformational change of the entire protein and not the exclusive disruption of the helix. It is thus possible that such a conformational change could result in the formation of misfolded versions of the protein, which would lead to the aggregation of insoluble proteins in inclusion bodies. The observation that there was no significant difference between the amount of isolated wild type and mutant protein possibly indicates that the introduction of the Pro residue did not enhance the formation of inclusion bodies as this would lead to large increases in the amount of protein in the insoluble fraction and decreases in the amount of protein in the soluble fraction as compared to that of the wild type sample. However, the possibility still remains that the Pro point mutation could have caused a general conformational change that led to the appearance of additional truncated versions of the protein in the mutant sample without resulting in insoluble misfolded proteins. Regarding the non-truncated, mutated bifunctional protein ~161 kDa in size, a general conformational change upon the introduction of the Pro residue should lead to an overall decrease in PfODC activity as this domain would be in the incorrect conformation for proper functioning. In other words, if the disruption of the helix resulted in a general conformational change then the protein downstream of the mutation would be more inactive compared to the when the entire insert was removed as in the previous study.

The isolated proteins were subsequently subjected to activity assays to determine whether the delineated area in insert A2, i.e. the conserved α-helix, contributes to the activities of the proteins in the bifunctional PfAdoMetDC/ODC complex.

3.5.5.2 Activity assays of the secondary structure-mutated proteins

Previous studies in which the conserved α-helix in the O1 insert was disrupted, resulted in PfODC activity depletion while the PfAdoMetDC enzyme retained a mere 6% activity. These effects were subsequently ascribed to the involvement of this delineated area in protein-protein interactions for the formation of the heterotetrameric complex (Birkholtz *et al.*, 2004; Roux, 2006). Similarly, the severe consequences on PfAdoMetDC and PfODC activities observed upon the deletion of the entire A2 insert above, led to the identification of a conserved α-helix within this region, which possibly serves as a protein interaction site for the subsequent formation of the active bifunctional PfAdoMetDC/ODC complex.

The isolated A/Owt and A/OpA2a recombinant proteins were subjected to radioactivity assays. Once again, the amount of ~161 kDa bifunctional protein in each of the total protein samples were taken into account during the determination of the specific activities of the proteins (Appendix B, Table 1B) and the percentage contributions of the truncated versions

of the bifunctional proteins were not taken into consideration since monofunctional PfODC has negligible activity (Krause *et al.*, 2000). Figure 3.16 shows the normalised activity of PfAdoMetDC in the mutant recombinant protein. The results show very little residual PfAdoMetDC domain activity of 15% when only the secondary structure within the A2 insert is disrupted compared to the 98% activity loss when the entire insert was removed.

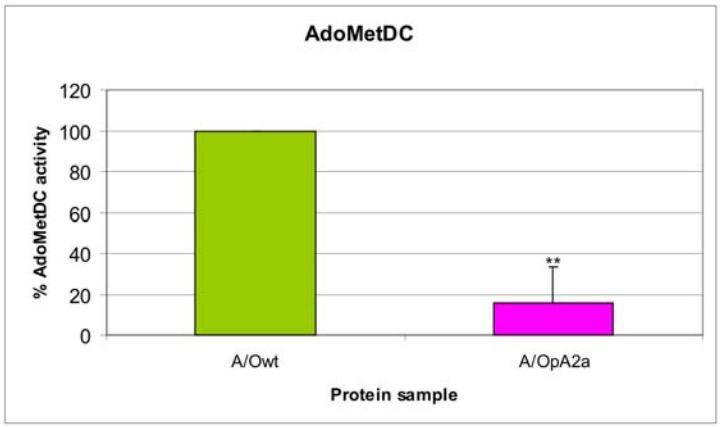


Figure 3.16: AdoMetDC activity assays of the of the wild type PfAdoMetDC/ODC and bifunctional A/OpA2a mutant proteins.
Wild type A/Owt activity is shown in green at 100% while the mutant proteins activity normalised to the A/Owt activity is shown in pink. The values were determined from three independent experiments carried out in duplicate. The standard deviations of the mean are indicated as error bars on each graph. Significant differences at a confidence level of 95% are represented as follows: ** for $p < 0.01$.

Figure 3.17 shows the normalised PfODC activity of the insert A2 helix-disrupted PfAdoMetDC/ODC protein against wild type protein activity. The disruption of only the α-helix in the A2 parasite-specific insert of PfAdoMetDC/ODC resulted in a significant decrease to 45% of wild type PfODC activity within the bifunctional PfAdoMetDC/ODC protein whereas the deletion of the entire A2 insert resulted in a decrease to only 18% of wild type PfODC activity (Figure 3.13).

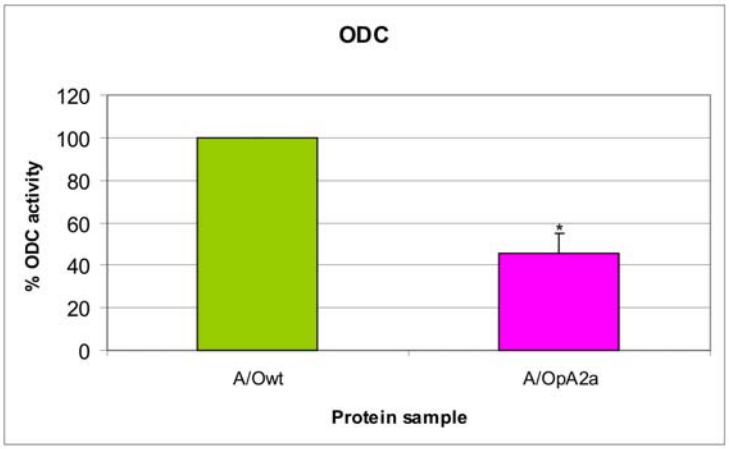


Figure 3.17: ODC activity assays of the wild type PfAdoMetDC/ODC and bifunctional A/OpA2a mutant proteins.

Wild type A/Owt activity is shown in green at 100% while the mutant proteins activity normalised to the A/Owt activity is shown in pink. The values were determined from three independent experiments carried out in duplicate. The standard deviations of the mean are indicated as error bars on each graph. Significant differences at a confidence level of 95% are represented as follows: * for $p < 0.05$.

The increase in PfODC activity observed here, indicates that the Pro hypothesis, which states that the use of a Pro residue in order to disrupt an α-helix could lead to undesired conformational changes within the protein, is not valid here as such a conformationally changed protein would not be active at all. These results thus indicate that the secondary structure within the A2 insert is indeed essential for the correct functioning of the PfAdoMetDC domain and that it also interacts with the adjacent domain via long-range interactions.

From the combined results of this study, Birkholtz *et al*. and Roux it thus seems that PfODC is more vulnerable to changes in the bifunctional protein, i.e. it is more dependent on the hinge and PfAdoMetDC domain for stabilisation and activity, and it is at the same time also the domain where initialisation of heterotetrameric complex formation was thought to take place (Birkholtz *et al.*, 2004; Roux, 2006). The deleterious influences of the insert mutations on the activities of both domains indicate possible allosteric-type effects that the two domains exert on one another for the maintenance of optimum intracellular polyamine levels. The reaction product of the PfODC enzyme does not stimulate the activity of PfAdoMetDC (Wrenger *et al.*, 2001) while PfODC is strongly feedback regulated by its own product putrescine (Krause *et al.*, 2000). PfODC is at the same time also dependent on the presence of the hinge region as well as PfAdoMetDC, which further controls the levels of putrescine and spermidine since the product of PfAdoMetDC is utilised by PfSpdSyn to synthesise spermidine from putrescine and dcAdoMet. Inactive PfAdoMetDC thus creates a bottleneck for subsequent reactions in the polyamine metabolism of *P. falciparum*, which is mediated by the presence of unique parasite-specific inserts. Targeting these inserts in the PfAdoMetDC domain with a specific compound would thus have severe effects on the activity of PfODC leading to a temporary halt in polyamine synthesis until these products are obtained extracellularly.

The analysis of the role of the β-sheet in the A3 insert on the activities of PfAdoMetDC/ODC is being investigated

In conclusion, this chapter discussed studies performed on the secondary structures in the PfAdoMetDC parasite-specific inserts of PfAdoMetDC/ODC to delineate possible interacting areas within the bifunctional heterotetrameric complex. From the results obtained it is clear that these secondary structures, particularly an α-helix in insert A2, are important for the activity of the PfAdoMetDC/ODC protein. The results obtained here are important for the

design and synthesis of possible non-active site based antimalarial drugs targeting these unique areas within the bifunctional PfAdoMetDC/ODC protein. The next chapter will focus on the conserved α-helix in the O1 parasite-specific insert to prove the involvement of secondary structures within the parasite-specific inserts in mediating interdomain protein-protein interactions.