Effects of co-expression of chaperone combinations on production of soluble plasmodial protein PfAdoMetDC in *E. coli*

A thesis submitted in fulfillment of the requirement for the degree

Of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY AND MICROBIOLOGY

UNIVERSITY OF ZULULAND

By

MAKHOBA XOLANI HENRY

20051930

CHAPTER ONE

1 LITERATURE REVIEW

1.1 Malaria as a health burden

Malaria is a major cause of death, especially in developing countries. According to the World Health Organization (WHO) report released in March 2013, 3.3 billion people are at risk of contracting malaria. Most deaths occur among children under the age of 5 years and in Africa, a child dies every minute from malaria (WHO, 2013). It has been shown that approximately 80% of malaria deaths occur in 14 countries and more than 40% of the total global annual deaths occur in the Democratic Republic of Congo and Nigeria (WHO, 2010).

Malaria is caused by *Plasmodium* parasites. The *Anopheles* mosquitos are the known malaria vectors responsible for the transmission of these parasites through their bites (Figure 1). There are five parasite species that cause malaria in humans: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi* (White, 2008; Hector, 2014). Of these five species, *P. falciparum* is the most deadly (Snow *et al*., 2005).

When certain forms of blood stage parasites are ingested by a female *Anopheles* mosquito during a blood meal from a human, it undergoes cycles of growth and multiplication in the mosquito. After 10–18 days, the parasites are found in the mosquito's salivary glands (Sinnis and Sim, 1997; Fujioka and Aikawa, 2002) (Figure 1). When the *Anopheles* mosquito takes a blood meal from a human, the parasites are injected with the mosquito's saliva to start an infection in the host's liver cells. The mosquito transmits the infection from one human to another, acting as a vector (Fujioka and Aikawa, 2002). Unlike the infected human, the mosquito vector does not suffer from the presence of the parasites as it lacks red blood cells (Snow *et al*., 2005). Other

than transmission by mosquitoes, malaria can also be transmitted via blood transfusions or through the sharing of needles. Mother to child transmission during pregnancy has also been documented (Fujioka and Aikawa, 2002). In the greater scheme of things, modes of transmission other than via the mosquito are believed to be very rare and unimportant (Snow *et al*., 2005).

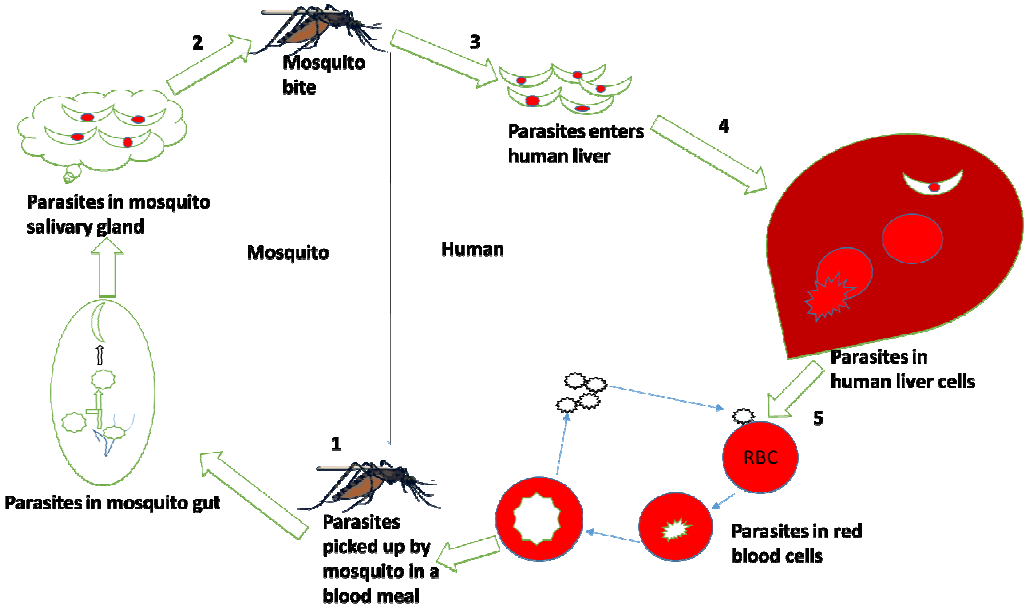

Figure 1: **The malaria life cycle in humans**. **(1)** After parasites have been sucked up, Oocysts develop in the gut wall of the mosquito. Sporozoites then develop in the oocyte and the Sporozoites migrate to the salivary glands. **(2)** When the mosquito bites, Sporozoites are injected into the human, who then becomes the second host to the parasites. **(3** and **4)** The Sporozoites enter the liver cells where they multiply for about 7 to 14 days, producing between 10,000 and 30,000 daughter cells called merozoites. These daughter cells then invade the red blood cells. **(5)** In the red blood cells, further multiplication occurs by asexual reproduction. Between 8 and 16 merozoites are produced every 48 or 72 hours, depending on the species of *Plasmodium.* Merozoites are then released through the bursting of red blood cells. This release of toxic substances causes the fibrile attacks of the disease. **(6)** After a number of such cycles, male and female gametocytes are produced (the sexual stage) and taken up by a feeding mosquito. The *Plasmodium* life cycle is completed by sexual reproduction, resulting in new sporozoites (Sinnis and Sim, 1997; Fujioka and Aikawa, 2002).

1.2 Production of recombinant malarial proteins

The heterologous production of malarial proteins in *E.coli* is often hindered by low yields of pure and soluble protein (Chowdhury *et al*., 2009). This could be due to the incompatibility of these proteins with the prokaryotic folding machinery which tends to delay folding relative to translation (Gupta *et al*., 2010). The low yield of pure and soluble protein makes isolating recombinant proteins in large quantities for structural drug discovery studies problematic.

Mehlin and colleagues (2006) analyzed 1000 genes from *P. falciparum* heterologously produced in *E. coli* and reported that only 337 were successfully expressed. Of these, as few as 63 were reported as soluble proteins. The high A/T content of *P. falciparum* genes has been implicated in the low expression of proteins (Flick *et al*., 2004). However, high molecular weight (> 56 kDa), basic isoelectric point (pI > 6) and lack of homology with *E. coli* proteins were also observed as major factors affecting production of malarial proteins (Mehlin *et al* ., 2006). A universal system that could produce pure, soluble and active recombinant proteins in *E. coli* remains a core objective of the industry of biotechnology.

1.3 The challenges of protein folding in the cell

Escherichia coli is frequently used as a host for the genetic manipulation and production of recombinant proteins, for both commercial and therapeutic purposes. Its known advantages include low cost, fast growth, and the ability to grow labeled proteins (Niraj *et al*., 2006). However, the production of foreign proteins in *E. coli* is occasionally problematic. Some of the challenges include the strength and translation ability of messenger RNA (mRNA). The breakdown of recombinant protein by proteases of the host system poses an additional problem,

affecting the cellular system function as the products become toxic (Makrides, 1996). Transcription and translation are tightly regulated such that ribosomes release a protein chain every 35 seconds resulting in over-crowding of the cytosol (Baneyx and Mujacic, 2004). Proteins become misfolded and form inclusion bodies (Hartl and Hayer-Hartl, 2002). Approaches such as codon harmonization, co-expression of molecular chaperones and chaperone combinations have been proposed as possible solutions to protein folding, as these approaches improve the production of target proteins (de Marco *et al*., 2007; Birkholtz *et al*., 2008; Stephenes *et al*., 2011).

1.3.1 Codon harmonization approach

Codon-harmonization is a technique that refers to the process of recognizing the low usage frequency codons within the native host and changing the code of the gene as expressed by codons which are recognized as low usage frequency codons by a non-natural host (Birkholtz *et al*., 2008). This change ensures that the positional codon occurs with a low to intermediate frequency by allowing the translational machinery to move and pause at the correct sites. It thus allows the folding of the secondary and tertiary structures to take place in a way similar to that which it occurs in the natural host. This approach has been successful in the heterologous production of some *Plasmodium falciparum* proteins, such as liver stage antigen (LSA-1) and *Plasmodium falciparum* sexual stage antigen (Pfs25) (Hillier *et al*., 2004; Kumar *et al*., 2014).

1.3.2 Molecular chaperones

Molecular chaperones are proteins that bind to exposed patches of newly synthesized proteins to protect them from misfolding and aggregation, without becoming part of the final product. They are classified into three categories based on their functional activities: folding, holding and

disaggregation chaperones (Baneyx, 2004). The major constituents that make up the molecular machinery of *E. coli* are summarized in Table 1.1

Table 1.1: Major groups and functions of molecular chaperones

Molecular chaperone	Target proteins	Size (kDa)	Function	References
Trigger factor (TF)	8 amino acid residue motif enriched with aromatic residues	48	Folding chaperone	Patzelt *et al*., 2001
DnaK	Segment of 4 to 5 hydrophobic amino acids	70	Folding chaperone	Hartl and Hartl, 2002
GroEL	Folds enriched in hydrophobic residues	60	Folding chaperone	Baneyx, 2004
ClpB	Segment enriched with aromatic residues	100	Disaggregation	Mayer and Bukau, 2005

Heat shock proteins (HSP) are a group of proteins induced by heat shock. The most prominent members of this group are a group of functionally related proteins involved in the folding and unfolding of other proteins. Most Hsps, but not all, can be induced by stress (De Mario, 1999). The up-regulation of heat shock proteins is a key part of the heat shock response and is induced primarily by the heat shock factor (HSF) (Wu, 1995). Hsps are found in almost all living organisms, from bacteria to humans.

Shonhai and colleagues (2007), reported that there are six heat shock proteins 70s (Hsp70s) from *Plasmodium falciparum*. Of these *P. falciparum*, PfHsp70 has been extensively studied, due to the molecular chaperone activity that it possesses. A previously conducted study shows that PfHsp70 improves the heterologous production of Plasmodium *falciparum* GTP cyclohydrolase I

(PfGCHI) in *E. coli* (Stephens *et al*., 2011). However, co-overproduction of individual molecular chaperones in *E. coli* has not always been successful in assisting target proteins to fold properly; for this reason this approach is not reliable (Agashe *et al*., 2004; Chang *et al*., 2005). Stephens and colleagues (2011) have suggested that the target protein is not completely soluble upon co-expression with PfHsp70 as an individual molecular chaperone. This would mean that certain proteins require more than one molecular chaperone to become soluble, and fold properly. Certain individual molecular chaperones such as ClpB and DnaK have been found to solubilize aggregated protein and facilitate the folding of newly synthesized polypeptides (Hartl and Hayer-Hartl, 2002; Mayer and Bukau, 2005). De Marco and colleagues (2007) reported that 26 of 50 target proteins investigated showed an increase in final protein yields and solubility upon co-expression with chaperones from the *E. coli* system (de Marco *et al*., 2007). Using molecular chaperones with different functional activities could, therefore, possibly improve the production, solubility, purity and activity of recombinant proteins.

The focus of the present study is on an investigation into the effects of combined molecular chaperones from *E. coli* and *P. falciparum,* by cloning them into the same vector and co-expressing with a malarial protein as the target gene in *E. coli*. Previous studies have focused on the use of molecular chaperones from one source such as *E. coli* or *P. falciparum* (de Marco *et al*., 2007; Stephens *et al*., 2011). This is thus the first report on the use of combined chaperones from different sources as possible tools to improve malarial drug target proteins.

1.4 Quality of recombinant proteins

Poor yield in the biosynthesis of target proteins is a major cause for recombinant enzymes and pharmaceuticals being excluded from the market (Ferrer-Miralles *et al*., 2009). Protein

misfolding and of consequential cell response triggers occur frequently in recombinant protein production (Mosser *et al*., 2000). However, what protein quality means might be contested. The general consensus is that a soluble form of a given protein is the most desirable form in the protein production process, despite the potential occurrence of soluble aggregates and the presence of functional protein species in protein aggregation (de Marco *et al*., 2007). A reduced temperature, recombinant gene dosage, promoter strength and supplementation of external molecular chaperones are all proposed protein solubility enhancers (de Marco *et al*., 2007). Over-expression of recombinant proteins is often ideal for the production of large amounts of protein for commercial or structural studies. However, protein over-expression is difficult to achieve due to the formation of so-called inclusion bodies. Inclusion bodies are regarded as "dead" end subjects or wasted products, which are the result of poor folding machinery in the cell (Dobson *et al*., 2001). These inclusion bodies reduce the quality of the protein.

1.5 Molecular chaperones as folding catalysts

The traditional view was that many proteins fold spontaneously (Anfinsen, 1973). This view was revised when it was found that many proteins in living cells will not fold correctly without the assistance of molecular chaperones. Chaperones are now defined as a family of cellular proteins which mediate the correct folding of other polypeptide without becoming the final functional structures (Fink, 1999). This definition implies that the dominant function of molecular chaperones is to transiently interact with other proteins, thereby preventing the formation of illegitimate interactions that might otherwise lead to deleterious protein aggregation. Molecular chaperones bind to exposed hydrophobic surfaces of polypeptides that will ultimately be buried in the final folded state (de Marco *et al*., 2007).

1.6 Heat Shock Proteins as Molecular Chaperones

Heat shock proteins (Hsps) form a major part of molecular chaperones and represent an abundant and conserved class of proteins that fulfill the role of maintaining the viability of cells under different conditions of stress (Handrick and Hartl, 1994). These proteins were first identified by Ritosa (1962) when *Drosophila melanogaster* cells were exposed to increased temperature. Heat shock proteins are up-regulated in response to stress such as high (heat shock) or low temperature (cold shock), reduced oxygen, water or nutrients, and high concentrations of cytotoxic compounds including free radicals, antibiotics or heavy metals. The role of heat shock proteins is more pronounced in the presence of stress stimuli attributed to different client binding and co-chaperone regulation (Callahan *et al*., 2002; Tutar *et al*., 2005).

1.6.1 Protein folding, misfolding and aggregation

The folding process of a protein in a cell is complicated by the over-crowded cellular environment (Bukau *et al*., 2000). Proteins constitute a large portion of the cell, however, amounting to at least half of the cell's dry weight (Lodish *et al*., 2000). Unfavorable physiological conditions such as extreme temperatures, heat and other cellular stressors, contribute to the misfolding of the proteins and the subsequent development of diseases such as Alzheimer's disease, Parkinson's disease Huntington's disease and many other degenerative diseases (Dobson, 2004; Chaudhuri *et al*., 2006). In addition, the challenges involved in the folding of large multi-domain proteins result in misfolding of the proteins, which, in turn, ultimately results in protein aggregation (de Marco *et al*., 2007). The folding rate of these large

multi-domains can also be higher than its synthesis which may give rise to misfolded products (Barral *et al*., 2004).

1.7 *E. coli* chaperones and their role in protein folding

HSPs such as 40 kilodalton (kDa), (HSP40), 60-kDa (HSP60), 70-kDa (HSP70), 90-kDa (HSP90) and 100-kDa (HSP100) are named according to their molecular weight. Table 1.2, summarizes some of the functions of heat shock proteins during the folding of polypeptides. The most interesting characteristic of the molecular chaperone is that it possesses the ability to recognize and bind newly synthesized proteins in order to prevent misfolding and aggregation. However, the high concentration of misfolded and aggregated proteins could overwhelm the molecular chaperones in the cells (Hartl and Hayer-Hartl, 2002). Cells therefore need a rapid and precise system to eliminate aggregated and misfolded proteins. Some of the molecular chaperones are known to be ATP-independent; they fall under the so-called, ‘holding’ group of molecular chaperones. The function of ‘holding’ chaperones is to bind nascent peptides and pass them on to ATP-dependent counterparts that are capable of folding the peptides into functional forms.

Table 1.2: The role of major heat shock proteins

Heat shock proteins	Role	References
Small heat shock protein	Small heat shock proteins are known to exist in all organisms. Their molecular weight ranges between 16-30 kDa. They are upregulated in various diseases, such as neurodegenerative disease and damaged motor neuron cells.	Sun *et al.*, 2005; van den Ijssel., 1999
Heat shock protein 40 (Hsp40)	Heat shock protein 40 acts as a co-chaperone partner of Hsp70. It also recognizes and binds unfolded proteins, handing them over to Hsp70 for proper folding. Hsp40 stimulates the hydrolysis of ATP by ATPase.	Li *et al*., 2009
Heat shock protein 60 (Hsp60)	Heat shock protein 60 is in the mitochondrial matrix and is responsible for the transportation and refolding of proteins from cytoplasm to mitochondrial matrix. Hsp60 plays an important role as a chaperone in proper folding of the linear amino acid chains as they acquire a three-dimensional structure.	Cheng *et al.*, 1990
Heat shock protein 70 (Hsp70)	Heat shock protein 70 is involved in the shielding of newly synthesized proteins. It is also involved in protein translocation, the degradation of misfolded proteins and the disaggregation of protein units. Hsp70 protects cells from thermal stress.	Daugaard *et al*., 2007
Heat shock protein 90 (Hsp90)	Heat shock protein 90 is involved in transcriptional regulation, signal transduction, and cell cycle control.	Bagatell *et al.*, 2000
Heat shock protein 100 (Hsp100)	Heat shock protein 100 is involved in repairing denatured proteins through the resolubilisation of heat-denature and misfolded proteins.	Krobitsch *et al*., 1998

1.7.1 Heat shock protein 70 (Hsp70)

DnaK is the *E. coli* representative of the ubiquitous Hsp70 family, which specializes in the binding of exposed hydrophobic regions in unfolded polypeptides (Hartl and Hayer-Hartl, 2002). Hsp70s are composed of three domains: N-terminal (ATPase) domain (45 kDa), the substrate binding domain (SBD-15 kDa) and a C-terminal domain which is 10 kDa in size (Flaherty *et al.*, 1990). The DnaK substrate bind release cycle requires co-factors such as DnaJ and a nucleotide exchange factor (GrpE) (Hartl and Hayer-Hartl, 2002). DnaJ also acts as a chaperone by binding to the newly synthesized proteins, handing them over to DnaK for proper folding (Hartl and Hayer-Hartl, 2002). Furthermore, DnaJ binds to DnaK to stimulate its ATPase activity (Hartl and Hayer-Hartl, 2002).

The controlled release of a substrate protein from the chaperone, often driven by ATP hydrolysis, promotes folding into its native state. Repeated cycles of bind and release are necessary for productive folding. In the ATP bound state, DnaK binds and releases the substrate faster, whilst the process occurs at a slower rate in the ADP bound state (Hartl and Hayer-Hartl, 2002) (Figure 2).

GrpE (23 kDa) acts as a nucleotide exchange factor by binding and modulating the binding pocket of DnaK, causing the release of bound ADP (Hartl and Hayer-Hartl, 2002). The cycle continues through further binding of ATP, resulting in the release of the substrate which could fold or be handed over to other chaperones if the substrates need further assistance for folding (Hartl and Hayer-Hartl, 2002).

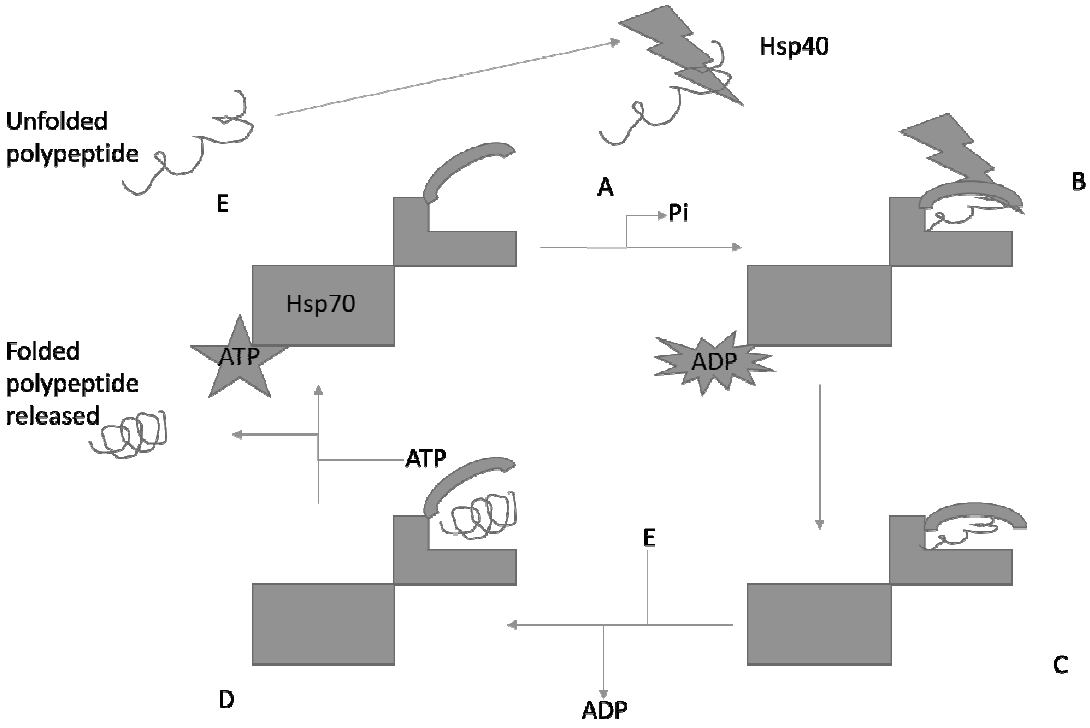

Figure 2: Reaction cycle of DnaK system The Hsp70 folding cycle of the substrate is controlled by the hydrolysis of ATP to ADP. **(A)** Hsp40 binds to a nascent or misfolded polypeptide handed to Hsp70-ATP. **(B)** The ATPase activity of Hsp70 is stimulated by Hsp40, resulting in the closed conformation of the Hsp70 nucleotide binding cleft. **(C)** The ADP-bound Hsp70 complex is stabilized **(D)** The nucleotide exchange factor is bound. **(E)** The polypeptide is released upon the uptake of the ATP. However, if the protein is not properly folded it can still be recruited by Hsp40 and enter the cycle again (Adapted from Edkins and Boshoff, 2014).

1.7.2 Heat Shock protein 40 (Hsp40)

Hsp40, also known as DnaJ in prokaryotes, play a significant role as co-chaperone partner of Hsp70 (Suh *et al*., 1999). The main role of Hsp40 is to bind to newly synthesized proteins, which then hand them over to Hsp70 to reach their final folding structure. *E. coli* DnaJ proteins have three domains: a highly conserved J-domain of approximately 70, a glycine and phenylalanine rich region (G/F domains) and a cysteine rich region (C domains) containing 4 motifs resembling a zinc-finger domain (Rajan and D'Silva, 2009).

The zinc-finger domain has been identified to have four CXXCXGXG repeats present in two separate clusters where each cluster co-ordinates with a zinc ion (Rajan and D'Silva, 2009) (Figure 3, Type I). This domain is important for sequestering the denatured substrate and assisting Hsp70 during the protein folding reaction (Walsh *et al*., 2004). The C-terminal region is less well conserved, however, and consists of a β-sandwich surrounded by a short C-terminal α-helix, followed by sequences essential for dimerization (Rajan and D'Silva, 2009). This region plays a major role in substrate binding and its sequestration into Hsp70 during the ATPase cycle and is thought to provide specificity for the Hsp70: DnaJ machine (Craig *et al*., 2006).

Almost all Hsp40s are characterized by the presence of a canonical J-domain and are also termed J-proteins or DNAJ in humans as most members have a molecular weight above 40 kDa. The functional activity of the J-domain is to interact with the ATPase domain of Hsp70 and stimulate the basal ATPase activity of Hsp70 (Wittung-Stefshede *et al*., 2003). Based on the domain organization, J-proteins are classified into 4 types namely: Type I, Type II, Type III and Type IV. Type I J-proteins show the presence of all domains found in DnaJ (Rajan and D'Silva, 2009) (Figure 3). They contain an N-terminal J-domain that is separated from the rest of the protein by

a long flexible linker ‘G/F region’ of 50-100 amino acids. The distal portion of the G/F region is the zinc-binding cysteine-rich sequence named the ‘zinc-finger domain’ which is the signature motif of type I proteins and distinguishes type I from other types of J-proteins (Rajan and D’Silva, 2009). The zinc-finger domain is followed by the C-terminal domain (Rajan and D'Silva, 2009).

Type II proteins possess all the domains except the zinc-finger domain. Type III J-proteins, on the other hand, contain a C-terminal J-domain and lack both G/F and zinc-finger domains. Type IV constitutes a group of recently identified proteins that are classified by the absence of the histidine, proline, aspartic acid motif (HPD) in their primary sequence (Botha *et al*., 2007; Rajan and D'Silva, 2009). Most but not all of them have a less well conserved DKE motif and is also termed as J-like proteins. The Type IV proteins share some similarity to their Type III counterparts, with their J-like domain towards the C-terminus of the protein (Rajan and D'Silva, 2009).

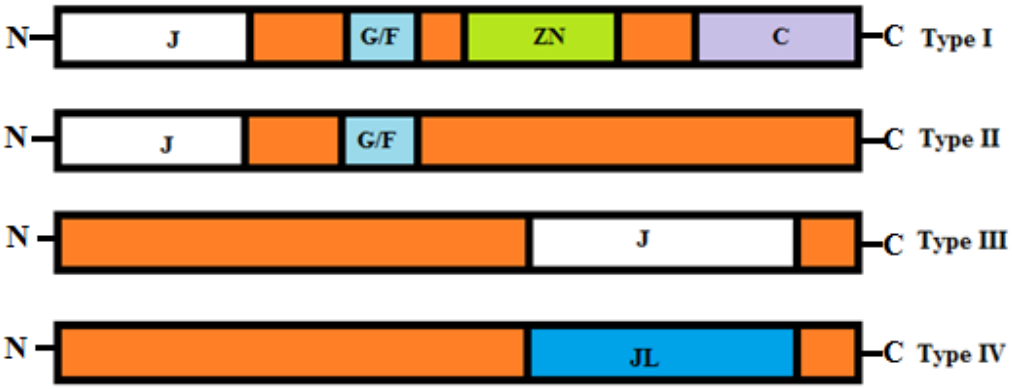

Figure 3: Four types of J-domains. J domain, glycine, zinc, lysine, cystein. (Adapted from Rajan and D'Silva, 2009).

1.7.3 Trigger factor

The trigger factor (TF) is a molecular chaperone that binds to bacterial ribosomes where it interacts with newly synthesized proteins promoting proper folding (Gupta *et al*., 2010). *E. coli* TF is 432 residues in length and made up of three domains: N-domain, C-domain, and the peptidyl-prolyl cis/trans isomerase (PPlase) domain which interconnects the *cis* and *trans* isomers with the proline amino acid (Shi *et al*., 2007) (Figure 4). The N-domain has an exposed loop which mediates the binding to the ribosomal protein tunnel exit terminal (Kramer *et al*., 2002; Marz *et al*., 2006). The N-terminal is connected by means of a long linker connected to the second PPlase domain which is located at the opposite end of the molecule and displays PPlase activity (Gupta *et al*., 2010) (Figure 4).

Figure 4: Linear structure of trigger factor. TF has got three domains: N-terminal which binds to the ribosomal subunit, followed by PPlase domain and C-terminal domain. However, the PPlase domain is connected to N-terminal via a long linker. The C-terminal is positioned between the molecules and provides binding to newly synthesized proteins (Adapted from Gupta *et al*., 2010).

The third domain comprises the C-terminal region, with its two protruding arms. The third domain forms the main part of TF with and is a crucial part of chaperone activity (Marz *et al*., 2006). It has been shown that the affinity of the TF ribosome-nascent complex for newly synthesized proteins increases with the length and hydrophobicity of these proteins (Rutkowska *et al*., 2008). The TF substrate binding resident time depends on the exposure of TF to the hydrophobic patches on the newly synthesized protein (Lakshmipathy *et al*., 2010). A previous study, in which two forms of *E. coli* proteomes with weak and strong interactions with TF were compared, has suggested that the high affinity of TF with nascent chains was due to the exposed linear hydrophobic regions during translation (Rutkowska *et al*., 2008; Lakshmipathy *et al*., 2010). TF interacts mainly with short nascent chains, whereas the bacterial Hsp70 homologue, DnaK, recognizes longer nascent chains adjacent to TF (Gupta *et al*., 2010). It has been suggested that the combined removal of *E. coli* TF and DnaK at 37ºC can be lethal (Deuerling *et al*., 1999; Teter *et al*., 1999). Furthermore, the functional co-operation of *E. coli* TF and DnaK has been reported to improve the folding yield of multi-domain proteins such as Beta-gal (five domains) and eukaryotic luciferase (two domains), but delays their folding during translation (Agashe *et al*., 2004). It has been suggested in this study that both chaperones recognize similar hydrophobic regions in newly synthesized polypeptides (Agashe *et al*., 2004).

1.7.4 GroEL

The chaperonin GroEL/Hsp60 is found in prokaryotes, mitochondria and chloroplasts of eukaryotes. It is a cylindrical shaped protein that belongs to the chaperonin family of molecular chaperones (Zeilsta-Ryalls *et al*., 1991). Chaperonins are subtype of molecular chaperones, but not all chaperones are chaperonins. GroEL chaperonin plays a major role in assisting with the folding of polypeptide, in cooperation with the lid-like co-chaperonin protein complex GroES (Hsp10) (Farr *et al*., 2000). Binding of the substrate and ATP causes a structural change which allows for the interaction of the binary complex with a separate lid structure, GroES (Farr *et al*., 2000; Brinker *et al*., 2001).

The process of protein folding within the cell involves encapsulation and release of the substrate protein. Substrate binds to hydrophobic regions on the interior rim of the open cavity of GroEL (Figure 5). The release of the substrate to the cytosol is due to the hydrolysis of ATP and binding of the new substrate to the opposite cavity which sends an allosteric signal to GroEL to release the substrate (Jim *et al*., 2006). At least 20-30% of proteins in the prokaryotic cytoplasm travel via the DnaK or GroEL chaperone machineries before adopting their final three dimensional structure (Jim *et al*., 2006). The role of GroEL and its partner GroES is to promote protein folding by sequestrating newly synthesized polypeptides in a cage-like structure. It has been suggested that substrates of up to 60 kDa in size can be encapsulated and their confinement in the GroEL/GroES cage may result in accelerated folding (Brinker *et al*., 2001) (Figure 5). GroES dissociates from GroEL every 10-15 seconds in a reaction dependent on the GroEL ATPase, thus allowing for the release of folded substrate and the capture of substrate that has not folded completely (Hartl and Hayer-Hartl, 2002).

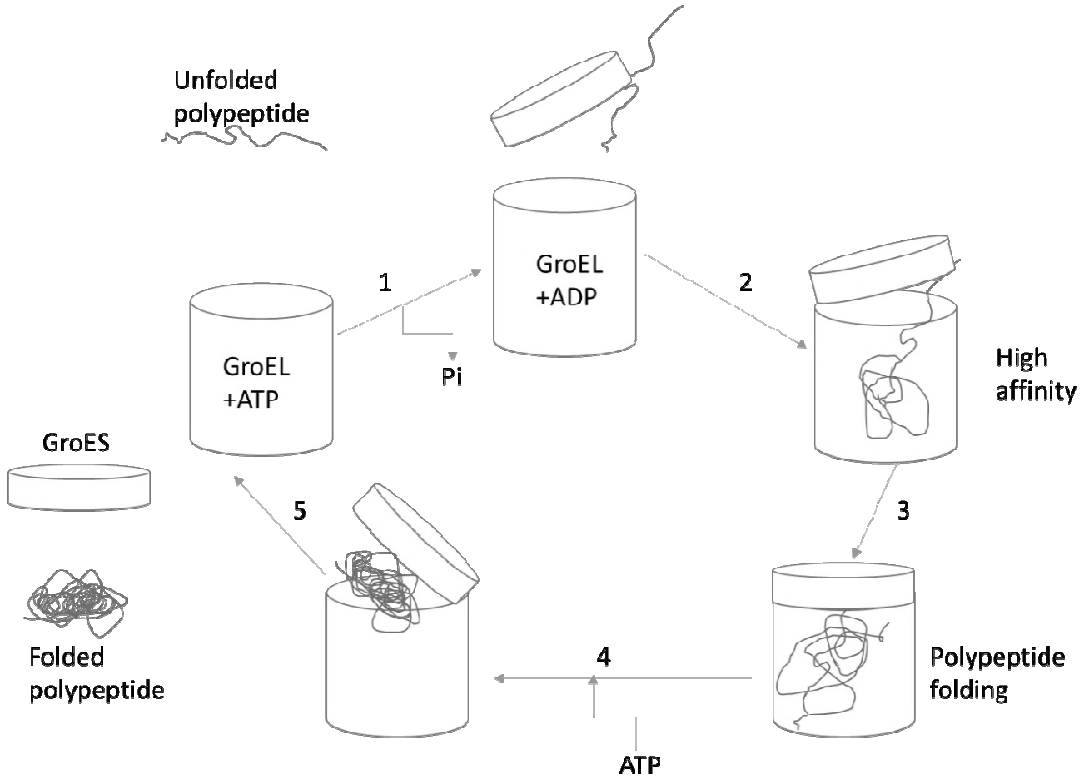

Figure 5: Reaction cycle of GroEL and GroES during the folding of newly synthesized protein (1) ATP hydrolysis results in high binding affinity and the substrate can bind; (2) In ADP form GroEL binds the substrate brought by the GroES co-chaperone; (3) Nascent polypeptide folds inside GroEL; (4) ADP converts to ATP; (5) The binding affinity of the substrate to GroEL becomes low and releases the folded polypeptide (Adapted from Zeilsta-Ryalls *et al*., 1991).

Most processes in living cells are performed in a coordinated manner between multiple proteins, often forming multi-protein complexes. The chaperone-assisted protein folding pathway is a very good example of this, as many chaperones have been identified as forming a network that interacts with substrates and assist with proper protein folding (Mogk *et al*., 1999). It has been suggested that DnaK and GroEL are cooperative in assisting the folding pathways of the newly synthesized proteins (Hartl and Hayer-Hartl, 2002). However, despite the fact that these chaperones are cooperative, they do not form a stable complex as compared to their eukaryotic counterparts during protein folding (Cueller *et al*., 2008).

1.7.5 ClpB

The ClpB (Hsp 100) family includes a wide range of proteins involved in different cellular functions. ClpB functions include refolding of aggregated proteins, degradation of misfolded proteins, transcription regulation and tolerance of high temperature in cells (Scirmer *et al*., 1996; Wawiznow *et al*., 1996). *E. coli* ClpB belongs to the Hsp100 protein family and plays an important role in protein degradation and disaggregation (Li and Sha , 2003). ClpB is made up of two nucleotide binding domains, NBD1 and NBD2 (Figure 6) (Jingzhi and Bingdong, 2002). Studies conducted show that ClpB co-operates with DnaK/DnaJ and GrpE to form a multi-chaperone system that solubilizes aggregated proteins and folds them into their active forms (Goloubinoff *et al*., 1999; Zolkiewski, 1999).

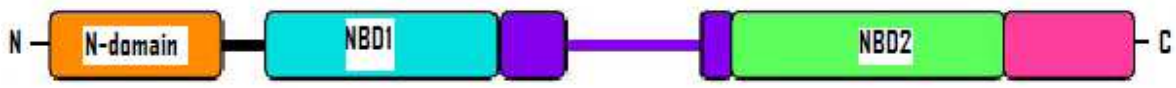

Figure 6: Linear structure of ClpB. The structure of ClpB is linear and consists of N-domain, nucleotide binding domain 1(NBD1), nucleotide binding domain 2 (NBD2) and C-terminal domain (Adapted from Jingzhi and Bingdong, 2002).

A previously conducted study suggests that, when firefly luciferase was denatured and subjected to ClpB on its own it was not reactivated. On the other hand, ClpB together with DnaK/DnaJ/GrpE greatly activated firefly luciferase by suppressing its aggregation. Furthermore, hydrophobic patches of polypeptide are exposed such that DnaK binds and further promotes refolding (Michal, 1999; Goloubinoff *et al*., 1999). Therefore, the coordination function of ClpB and DnaK/DnaJ/GrpE requires ATP hydrolysis.

1.8 Plasmodium falciparum heat shock protein 70

Heat shock protein 70 (Hsp70s) play an important role in folding of newly synthesized proteins. At least six Hsp70s have been characterized from *P. falciparum*: PfHsp70, PfHsp70-2, PfHsp70-3, PfHsp70-x, PfHsp70-y and PfHsp70-z (Shonhai *et al*., 2007). Of these, PfHsp70 has been identified as a cytosolic molecular chaperone that possesses ATPase activity and suppresses protein aggregation (Shonhai *et al*., 2008). Structurally, PfHsp70 is made of three domains namely: N-terminal domain, substrate binding domain and C-terminal domain (Figure 7).

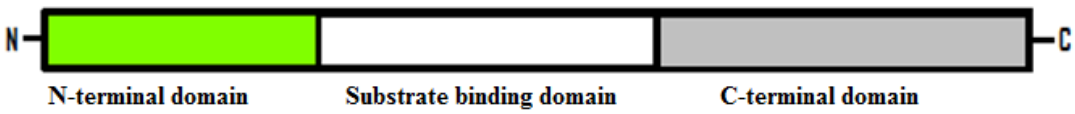

Figure 7: The structure of PfHsp70 showing: N-terminal domain, substrate binding domain and C-terminal domain (Adapted from Shonhai *et al*., 2005).

The molecular weight of this protein is approximately 74 kDa and it possesses an EEVD motif that enables it to interact with co-chaperones via the C-terminus (Ramya *et al.*, 2006). Stephens and colleagues (2011), have suggested that the PfHsp70 improves the expression, solubility and activity of PfGCHI where heterologous production in the *E. coli* was problematic. This suggests that molecular chaperones from the same target-species protein may help the protein to fold in *E.*

coli. The combination of malarial chaperones and *E. coli* chaperones co-expressed with a malaria drug target proteins could improve protein expression, purity and solubility.

1.8.1 PfHsp70 and its derivatives show chaperone activity in *E. coli* cells

The chimeric protein KPf, made up of the ATPase domain from *E. coli* DnaK and substrate binding domain from PfHsp70, was generated by Shonhai and colleagues (2005) (Figure 9). KPf was able to suppress thermosensitivity in the *E. coli dnaK756* strain (a DnaK mutant strain that is thermosensitive) (Retalede *et al*., 2009). Shonhai and co-workers (2005) have suggested that both PfHsp70 and its derivative KPf operated as chaperones in the *E. coli dnaK 756* cells. This indicates that they functionally replaced DnaK and demonstrates the potential functional equivalence of these Hsp70 homologs.

PfAdoMetDC protein has been identified as a possible drug target (Muller *et al*., 2001; Pegg, 2006). A previous study showed that PfAdoMetDC expression in *E. coli* is not a problem, but the major challenge rests upon its purification. It has been suggested that it co-purified with endogenous DnaK of the *E. coli* host system (Williams *et al*., 2011). This would have jeopardized the isolation of pure protein in large quantities for structural studies, thereby delaying the discovery of malarial drugs or compounds targeted to PfAdoMetDC protein. Furthermore, its co-purification with DnaK suggests that PfAdoMetDC is not produced as a fully folded protein.

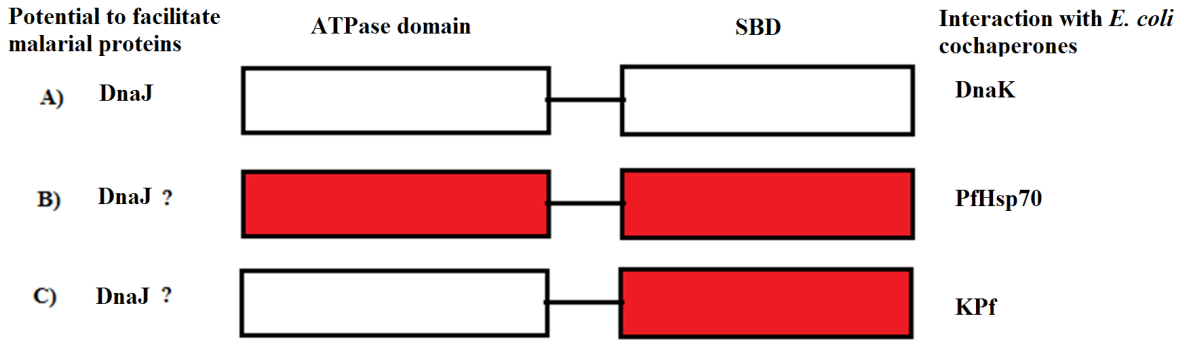

Figure 8: Possible cooperation between *E. coli* co-chaperones and malarial chaperones to facilitate malarial protein folding. (A) Known interaction between *E. coli* cochaperones DnaJ with DnaK **(B)** One focus of this study is to investigate whether *E. coli* cochaperones DnaJ functions cooperatively with PfHsp70 **(C)** This study also investigates the possible cooperative functioning of the *E. coli* cochaperone, DnaJ with KPf. The white boxes in figure 8 above represent the ATPase and substrate binding domain of DnaK and the red boxes represent ATPase and substrate binding domains of PfHsp70. The ATPase domain from DnaK and SBD from PfHsp70 were linked to form the KPf encoding segment (Adapted from Shonhai *et al*., 2005).

Hsp40s present substrates to Hsp70. They can be perceived as regulating the specificity of Hsp70 (Figure 8 A). In a prior study of PfHsp70, it could not be ascertained whether this molecular chaperone interacted with Hsp40 proteins in *E. coli*, or whether it bound to misfolded proteins independently, thereby preventing them from aggregating (Figure 8 B) (Shonhai *et al*., 2005). However, possible interaction between PfHsp70 and DnaJ could not be ruled out, as Hsp40s possesses a conserved J-domain that facilitates their interaction with the highly conserved N-terminal ATPase domain of Hsp70. It is possible that KPf cooperates with the DnaJ co-chaperone, as it has an ATPase domain from *E. coli* DnaK (Figure 8 C). The study therefore proposed that the *E. coli* system could better accommodate KPf because of the ATPase domain it possesses from *E.coli* DnaK. In addition to that, the substrate binding domain may recognize the PfAdoMetDC protein since it has SBD from PfHsp70, which improves the quality of the protein. It is this part of the hypothesis that the current study seeks to investigate.

1.9 Molecular chaperones as functional partners

Molecular chaperones have different functional characteristics, though they are co-operative in the folding processes of nonnative polypeptides. The trigger factor binds newly synthesized proteins upon their exit from ribosomes, passes them on to Hsp70s (homologue of *E. coli* DnaK) and then finally on to the cylindrical chaperonin system GroEL/GroES which receives a substrate after its interaction with DnaK (Hartl and Hayer-Hartl, 2002) (Figure 9).

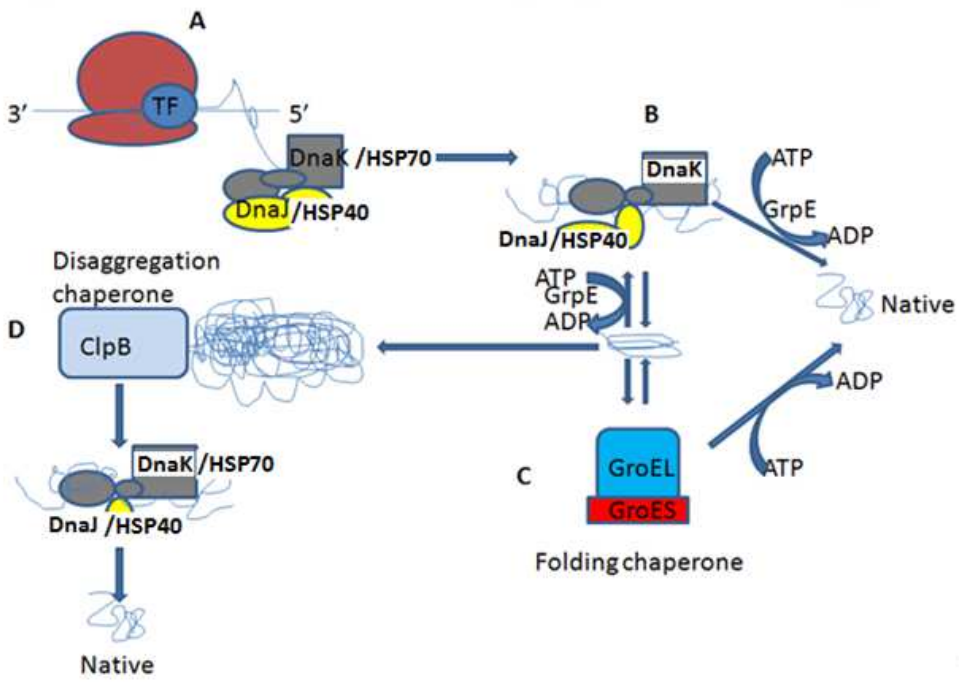

Figure 9: Folding pathways of the newly synthesized protein. (A) A newly synthesized protein exits from the ribosome and binds to the trigger factor (TF) as a first chaperone; **(B)** The protein is then passed on to DnaK/Hsp70 for proper folding; **(C)** DnaK hands it over to GroEL/GroES for further folding; **(D)** Next is ClpB, the disaggregation chaperone, and from here the proteins then move over to DnaK/Hsp70 for folding purposes. (Adapted from Hartl and Hayer-Hartl, 2002).

The most important features of the proteins that determine their biologically active form are their solubility and activity (de Marco *et al*., 2005). However, increased solubility does not mean a

corresponding or equal improvement in protein activity (de Marco *et al*., 2005). Some groups have reported poor or no activity despite producing highly soluble proteins (Butz *et al*., 2003). Apparently, combined chaperones have been reported as a promising approach to obtaining an over-production of protein and active target proteins (Alibolandi *et al*., 2010, Cui *et al*., 2011). It was concluded from the same study that the combinations of DnaK/DnaJ/GrpE could not facilitate the production of recombinant hbFGF protein. Other researchers have shown success by producing target proteins using combined chaperones of *E. coli* origin (de Marco *et al*., 2007). However, not all the sets of chaperones that were used succeeded in meeting the required objectives, which include a high yield, and a soluble, active and pure protein.

Combinational expression of different chaperone systems may improve solubility and protein activity (Nishihara *et al*., 1998; Schlieker *et al*., 2002). The question of its effectiveness has not been answered experimentally, due to the fact that most co-expression studies focus more on achieving high yields of the target protein instead of the biochemical properties of the co-expression system (de Marco *et al*., 2007). On the other hand, it is very important to understand whether molecular chaperones, individual or combined, were specific in promoting the expression of certain target recombinant proteins expressed in *E. coli*. The incorrect selection of chaperones could negatively affect host cell physiology and the production of heterologous proteins.

1.10 Research Hypothesis, Broad Objectives and Approach

A major function of molecular chaperones that interact with nascent polypeptides chains is to protect exposed hydrophobic amino acids in the highly crowded environment of the cytosol. However, while increased folding yields are observed in many specific cases, the over-expression of these chaperones, individually or in combination does not generally improve the folding of eukaryotic multi-domain proteins upon recombinant expression (Agash *et al*., 2004; Cheng *et al*., 2005; Haack *et al*., 2009). The fact that different chaperones prefer to bind specific proteins in the *E. coli* system requires that molecular chaperones be used in promoting recombinant proteins in *E. coli*. Hsp70s of *P.falciparum* origin PfHsp70 and its chimeric derivative KPf have exhibited chaperone activity in *E. coli* cells (Shonhai *et al*., 2005).

It is envisaged that the optimized partnerships of these chaperones, in combination with other *E. coli* chaperones, will improve the heterologous production of malarial protein in *E. coli*. This is because, as PfHsp70 and KPf are both of plasmodial origin, they are expected to recognize the malarial proteins co-expressed in *E. coli*. However, the approach of matching these chaperones with the rest of the *E. coli* proteins folding machinery will be important in achieving maximum folding efficiency of the recombinant malarial protein expressed in *E. coli*.

Based on this hypothesis, the following broad objectives of this study are:

1) Bioinformatics analysis of possible binding of PfAdoMetDC to molecular chaperones;

2) Expression of PfAdoMetDC in *E. coli* supplemented with heterologously expressed Hsp70 chaperones;

3) Co-expression of PfAdoMetDC in *E. coli* supplemented with heterologously expressed GroEL chaperone system;

4) Purification of PfAdoMetDC recombinant protein; and

5) *In vitro* biochemical characterization of PfAdoMetDC co-expressed with molecular chaperones *E. coli.*

CHAPTER TWO

Bioinformatics analysis of PfAdoMetDC protein as a potential substrate of Hsp70 and GroEL molecular chaperones

2.1 INTRODUCTION

This chapter focuses on the bioinformatics based assessment to establish the protein folding pathway that PfAdoMetDC is likely to go through in the cell. Bioinformatics is the general used approach in the management of biological information using computer technology. Therefore, computers are used as main tools to gather, store, analyze and integrate biological and genetic information which can then be used towards gene-based drug discovery and development (Kiefer *et al*., 2009). It has been established that molecular chaperones bind to the hydrophobic patches of the newly synthesized polypeptides, protecting them from misfolding (Lakshimipathy *et al*., 2010). *E. coli* expresses its own proteins endogenously, hence expression of foreign genes in E. coli could result in the overcrowding (de Marco *et al*., 2007). The overcrowding of *E. coli* as host has been suggested to cause the titration of the folding machinery; thus, most proteins become vulnerable to misfolding, losing their functional activity (de Marco *et al*., 2007). To overcome this challenge, supplementation with molecular chaperones has been suggested as a solution (Birkholtz *et al*., 2008). Previous studies have shown some success in the use of both individual and combined molecular chaperones (de Marco *et al*., 2007; Stephens *et al*., 2011).

The occurrence of hydrophobic regions in the linear sequence of nascent polypeptide chains results in high binding affinity of molecular chaperones such as Hsp70 and GroEL (Lakshmpathy *et al*., 2010). Kaiser and colleagues (2004) showed that certain proteins such as luciferase possess multiple hydrophobic regions that are easily recognized by molecular chaperones to bind to in order to prevent them from misfolding and aggregation. Thus, the presence of linear hydrophobic regions in the primary sequence of the newly synthesized protein could be used to predict whether the protein has weak or strong interactions with molecular

chaperones (Lakshimipathy *et al.*, 2010). The hydrophilic and hydrophobic profile of a protein along with its amino acid sequences can be evaluated using computer software as previously described by (Kyte and Doolittle 1982). The Kyte and Doolittle method can be used to predict or analyze the binding strength of the protein based on its hydrophobic profile. The information obtained from this chapter could be more helpful to gain insight as to whether PfAdoMetDC as a target protein could be a suitable substrate for Hsp70 and GroEL chaperones. Therefore, the hydrophobicity profile of PfAdoMetDC protein was predicted using a method developed by Kaiser and colleagues (2006). DnaK binding sites on PfAdoMetDC were also predicted as described (Rudiger *et al*., 1997).

The aims of this study were to:

1) Conduct comparative protein sequence analyses between *E. coli* AdoMetDC and PfAdoMetDC in order to identify binding motifs of DnaK and GroEL
2) Assess the physicochemical properties of *E. coli* AdoMetDC and PfAdoMetDC to identify them as potential substrate of Hsp70 and GroEL chaperones
3) To predict the hydrophobicity of *E. coli* AdoMetDC and PfAdoMetDC profile using Kyte and Doolittle predictive tools
4) To establish binding motifs of *E. coli* AdoMetDC and PfAdoMetDC by predicting binding sites of DnaK using a method developed by Rudiger and co-workers (1997)
5) Depict the 3D model highlighting the distribution of Hsp70 binding sites on *E. coli* AdoMetDC and PfAdoMetDC using SWISS-MODEL and PyMol

2.2 Experimental Procedures

2.2.1 Protein sequence alignment

2.2.2 Sequence alignment of PfAdoMetDC and *E. coli* AdoMetDC

The sequence of *P. falciparum* AdoMetDC (Accession number: E.C.4.1.1.50) was obtained from PlasmoDB (www.PlasmoDB.org). The *E. coli* AdoMetDC (Accession number: A7ZW69.1) protein sequence was obtained from the National Center for Biotechnology information (NCBI) website (http://www.ncbi.nlm.nih.gov). Using Bioedit_v7.0.5.3 software (Hall, 1999), sequence identity and similarity between PfAdoMetDC and *E. coli* AdoMetDC was determined. The sequence alignment was performed using Multiple ClustalW tool (Thompson *et al*., 1994; Hall, 1999).

2.2.2 Physicochemical properties

2.2.2.1 Comparing physicochemical properties of *E. coli* AdoMetDC and PfAdoMetDC proteins

AdoMetDC from *E. coli* and *P. falciparum* species were analyzed using online tools such as GenScript, Psort II (Li *et al*., 2014), PROPKA 3.1 (Sondergaard *et al*., 2011) and FoldIndex© (Kilambi and Gray, 2012) to determine if they share any similarities. This was done by comparing the optimum pH of substrate, molecular weight (kDa), isoelectric point (pI) and pKa value.

2.2.3 Hydrophobicity profiles and DnaK binding sites

2.2.3.1 Analysis of hydrophobicity of *E. coli* AdoMetDC and PfAdoMetDC proteins

The hydrophobicity profile of *E. coli* AdoMetDC and PfAdoMetDC was predicted using Kyte and Doolittle (1982) method. Hydropathy profiles can be used to examine the surface features of proteins in order to generate hypotheses that can be confirmed experimentally (Krystek *et al*., 2001). The peaks that are on the positive side represent the hydrophobic profile of the protein, while those on the negative side represent the hydrophilic profile of the protein. DnaK/Hsp70 binding sites on *E. coli* AdoMetDC and PfAdoMetDC was predicted using Rudiger and colleagues (1997) method. This method is based on differential scoring of the statistical energy contribution of each amino acid in a five residue core and four residue flanking regions, together constituting the proposed DnaK binding motif (Rudiger *et al*., 2001). The combined energy value obtained for a given sequence is taken as a measure of the likelihood that DnaK binds to this sequence. Therefore, peptides that registered scores less than -5 are predicted as DnaK binders, while peptides that registered scores greater than -4 are regarded as non-binders (Rudiger *et al*., 1997).

2.2.4 Homology Modelling

2.2.4.1 Homology Modelling of *E. coli* AdoMetDC and PfAdoMetDC

Models of *E. coli* AdoMetDC and PfAdoMetDC were generated using the software package Modeller version 8.2 (Fiser *et al*., 2000) and the protein modelling server SWISS-MODEL (Schewede *et al*., 2003). The models were subsequently visualized using PyMol version 0.99rc6 to highlight the predicted binding sites of Hsp70 (DeLano, 2002).

2.2.5 Analysis of suitable substrate for GroEL

2.2.5.1 The assessment of the structural features of *E. coli* AdoMetDC and PfAdoMetDC that determine its eligibility as a GroEL substrate

In *E. coli*, the ribosome-associated trigger factor together with DnaK/Hsp70 system assist the *de novo* folding of at least 340 cytosolic proteins within a broad size range between 16 - 167 kDa (Deuerling *et al*., 1999), whereas GroEL chaperone machinery helps to fold 250–300 newly synthesized proteins, highly preferring those with a size of 20 - 60 kDa (Houry *et al*., 1999). In addition to that the hydrophobicity profile of the substrates is a major factor that promotes the interaction between GroEL and the substrate (Coyle *et al*., 1997). Therefore, the analysis of both *E. coli* AdoMetDC and PfAdoMetDC features would be compared to the previously published data (Coyle *et al*., 1997; Hartl and Hayer-Hartl, 2002; Kerner *et al*., 2005). The physicochemical features of typical GroEL substrate have been extensively analyzed (Kerner *et al*., 2005). Based on the previous studies, the following features are crucial determinants peptides as GroEL substrates: isoelectric point (pI), pH, hydrophobicity profile and sizes. Using these guidelines, the physicochemical features of *E. coli* AdoMetDC and PfAdoMetDC were analyzed to confirm their eligibility as GroEL substrates.

2.3 Results

2.3.1 Comparison of PfAdoMetDC sequences with *E. coli* AdoMetDC to identify binding motifs of DnaK and GroEL

Firstly, a protein sequence alignment of *E. coli* AdoMetDC and PfAdoMetDC was conducted in order to identify conserved GroEL and DnaK binding motifs (Figure 2.1). *E. coli* AdoMetDC is supposed to be well recognized by *E. coli* DnaK and GroEL chaperones since they are from the same species (Hartl and Hayer-Hartl, 2002). There are conserved residues in PfAdoMetDC that could possible interact with DnaK and GroEL chaperones, these amino acids are highlighted by black boxes (Figure 2.1) (Coyle, *et al*., 1997). The sequence of *E. coli* AdoMetDC served as a as a reference point for the assessment of PfAdoMetDC as both GroEL and DnaK substrate. PfAdoMetDC shares a sequence identity of 33% and similarity of 50.09% with *E. coli* AdoMetDC which suggests that they are closely rated homologs (Figure 2.1). Altogether, this indicates that PfAdoMetDC though not from the same species with *E. coli* AdoMetDC could be recognized by Hsp70 and GroEL chaperones.

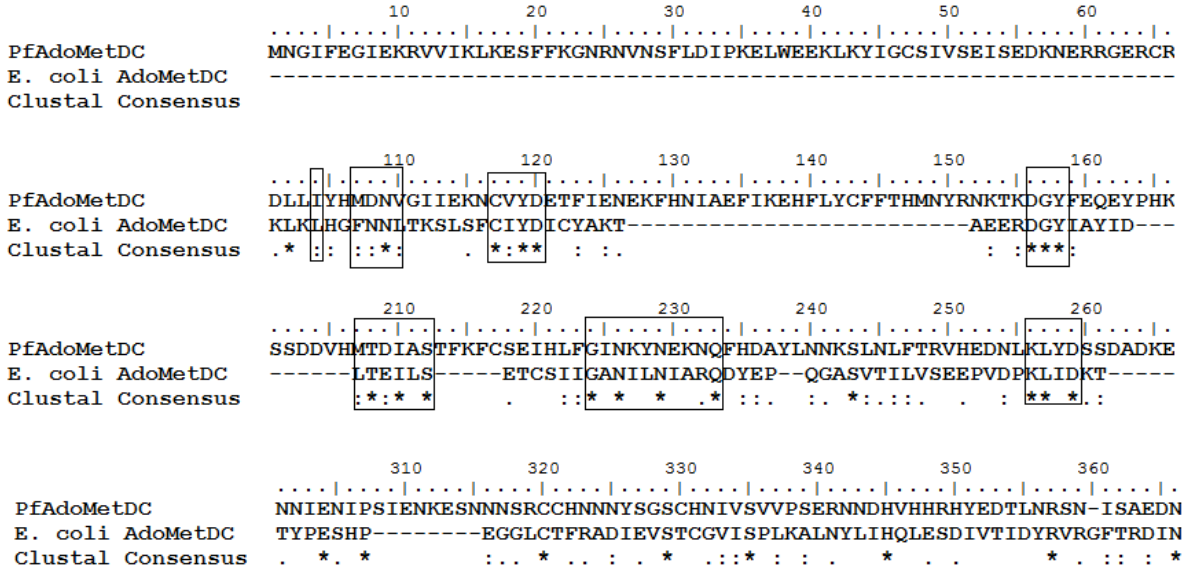

Figure 2.1: Sequence alignment of PfAdoMetDC and *E. coli* AdoMetDC. ClustalW alignments for PfAdoMetDC (accession number: E.C.4.1.1.50) and *E. coli* AdoMetDC (accession number: A7ZW69.1) were performed using Bioedit programme ClustalW alignment option (Thompson et al., 1994). Charged, polar and hydrophobic amino acids are highlighted by black box, that are thought to be play a critical role in protein folding). Symbols: "*": identical amino acids. ":" : conserved amino acids. ".": semi-conserved amino acids.

2.3.2 Predictive analysis of *E. coli* AdoMetDC and PfAdoMetDC structural features

Based on the physicochemical analysis of *E. coli* AdoMetDC and PfAdoMetDC the following features were established: pH optimum stability, attribute of substrate, molecular weight isoelectric point (pI) and hydrophobicity (Table 2.1). At physiological pH both proteins are acidic, *E. coli* AdoMetDC has a pI of 6.0 and PfAdoMetDC with pI of 6.35. Mehlin and colleagues (2006) demonstrated that proteins with pI values above 6.0 are difficult to express in *E. coli.* (Table 2.1). *E. coli* AdoMetDC is 30 kDa size while PfAdoMetDC is 60 kDa size (Table 2.1). Typically, substrates of GroEL are of similar size to GroEL itself (Kerner *et al*., 2005). Hence, PfAdoMetDC could interact better with GroEL than *E. coli* AdoMetDC; thus shield exposed hydrophobic amino acids residues that can give rise to aggregation in the highly

crowded environment of the cytosol. Hsp70 molecular chaperone binds to the substrate before it hands it over to GroEL chaperone provided if there is a need for further folding of the substrate (Hartl and Hayer-Hartl, 2002). The same mechanism of cooperation is expected to take place *in vivo* with PfAdoMetDC as substrate if further folding is required.

Table 2.1: Physicochemical properties of *E. coli* AdoMetDC and PfAdoMetDC as potential substrates of DnaK and GroEL chaperones

	***E. coli* AdoMetDC**			**PfAdoMetDC**			
Features	**Net status**	**DnaK**	**GroEL**	**Net status**	**DnaK**	**GroEL**	**References**
pH for optimum stability	3.6	√	√	3.6	√	√	Hui et al., 2005
pKa (kcal/mol)	33.6	√	√	33.6	√	√	Mats et al., 2011
Molecular weight (kDa)	30*	√	√	60*	√	√	Hartl and Hartl, 2002; Baneyx, 2004; Fujiwara et al., 2010
Isoelectric point (pI)	6.0 #	√	√	6.35 #	√	√	Mchlin et al., 2006
Hydrophobicity	0.465	√	√	0.417	√	√	Hartl and Hayer-Hartl, 2002

Symbols: "√" DnaK and GroEL potential substrate; "* " suitable substrate of DnaK and GroEL; "#" most suitable candidate of DnaK and GroEL.

2.3.3 Hydrophobicity profiles of *E. coli* AdoMetDC and PfAdoMetDC and predictive DnaK binding sites in *E. coli* AdoMetDC and PfAdoMetDC proteins

The hydrophobicity profile of both *E. coli* AdoMetDC and PfAdoMetDC was predicted to further investigate if PfAdoMetDC was a suitable substrate for GroEL and DnaK chaperones. It is important to note that *E. coli* AdoMetDC is smaller in size than PfAdoMetDC, thus the Y-axis (score) values are different from PfAdoMetDC values (Figure 2.2). Based on the analysis,

PfAdoMetDC protein has fewer hydrophobic residues than *E. coli* AdoMetDC (Table 2.1; Figure 2.2 A and B). However, there are residues that are observed to have high registered scores on PfAdoMetDC, at position 1-101 on the β-sheet (Figure 2.2 B). Their difference in hydrophobicity scores could influence the ability of either Hsp70 or GroEL binding (Table 2.1; Figure 2.2 A and B).

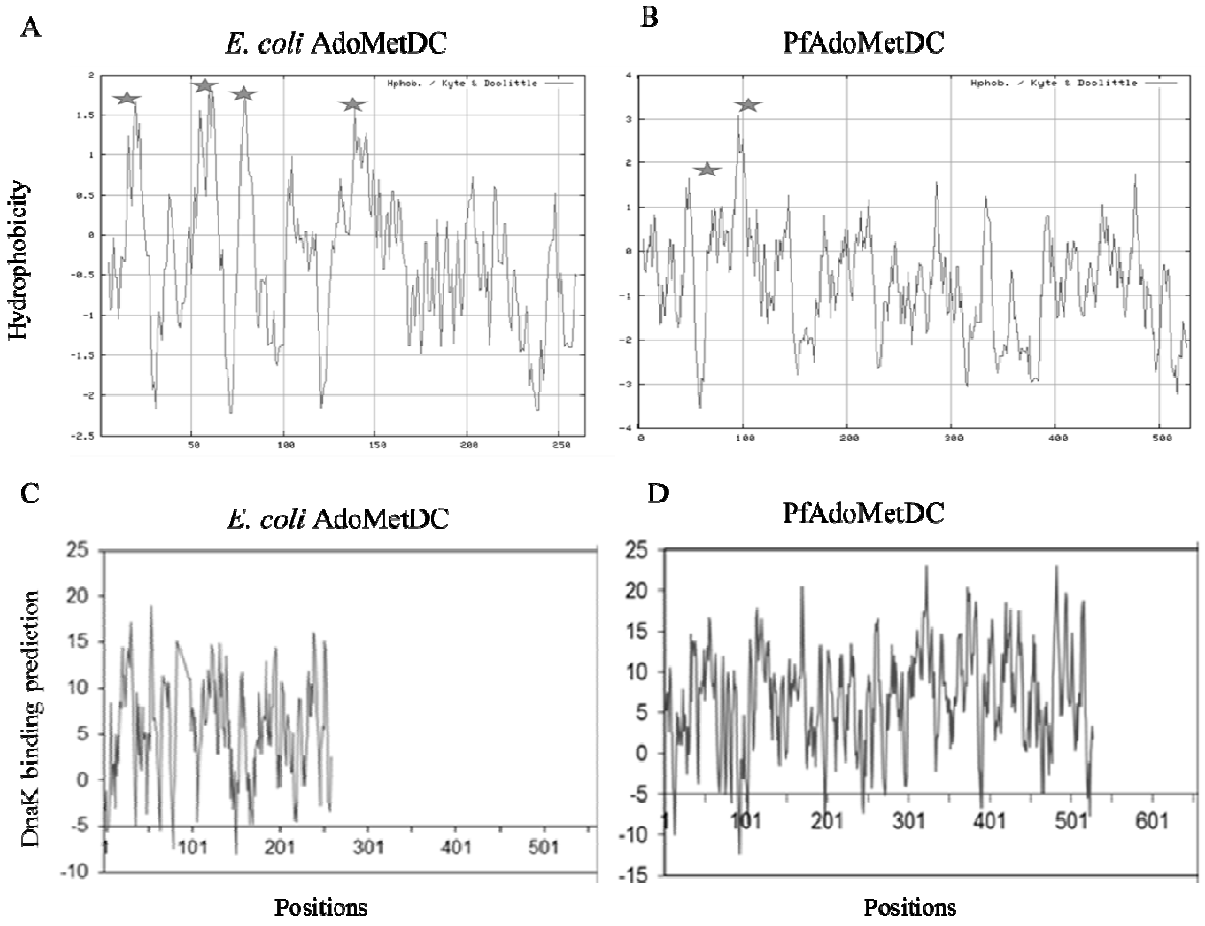

Figure 2.2: The prediction of hydrophobicity profile of *E. coli* AdoMetDC and PfAdoMetDC protein and prediction of DnaK binding on *E. coli* AdoMetDC and PfAdoMetDC substrates. (**A)** and (**B**) Kyte and Doolittle hydrophobicity profile analysis of PfAdoMetDC indicates the abundance of hydrophilic (negative scores) compared to hydrophobic profile (positive scores, denoted with stars) amino acids residues. (**C)** and (**D**) Prediction of potential DnaK binding sites on *E. coli* AdoMetDC and

PfAdoMetDC proteins. Segments with less than -5 indicate potential DnaK binding sites within PfAdoMetDC protein sequence.

Therefore, the ability of DnaK to bind to *E. coli* AdoMetDC and PfAdoMetDC was predicted as described (Rudiger *et al*., 1997). Figure 2.2 (C and D) shows the binding scores of DnaK on *E. coli* AdoMetDC and PfAdoMetDC protein sequences. Based on the analysis it appears that DnaK displays high binding affinity on PfAdoMetDC compared to *E. coli* AdoMetDC, the difference in size of these two proteins could be a contributing factor. Therefore, PfAdoMetDC profile suggests that it is a suitable candidate or substrate of DnaK (Figure 2.2 C). It is tempting to conclude that PfAdoMetDC possesses more hydrophobicity residues based on this finding. A previous study conducted suggests that smaller proteins do not need more assistance during their folding. However, larger proteins interact with molecular chaperones such as DnaK and GroEL in order to stabilize them during their folding process (Hartl and Hayer-Hartl, 2002).

2.3.4 Identification of predicted binding sites of Hsp70 in *E. coli* AdoMetDC and PfAdoMetDC proteins

To further demonstrate the possible binding sites of Hsp70 molecular chaperone on *E.coli* AdoMetDC and PfAdoMetDC 3D models were generated. It is important to note these were not full length proteins. *E.coli* AdoMetDC has α-helix at the N-terminal and β-sheet at the C-terminal (Figure 2.3 A). In contrast PfAdoMetDC has a β-sheet at the N-terminal and α-helix at C-terminal Figure 2.4 B). Based on the analysis of the 3D protein, hydrophobic residues are more pronounced at the α-helix compared to the β-sheet of the protein, thus proposed binding sites of Hsp70s are at the N-terminal residues 20-149 (Rudiger *et al*., 2001). The analysis shows that PfAdoMetDC structure has a long loop, which contributes to the flexibility of the protein (Figure 2.4) (Birkholtz, 2004). It is believed these loops are prone to be attacked by proteolysis if the

protein is not folded properly (Fontana *et al*., 2004). Based on the hydrophobicity profile of PfAdoMetDC protein, it has more hydrophobic regions in the β-sheet at residues 10-81 (Rudiger *et al*., 2001). These are proposed sites at which Hsp70s bind during its folding (Figure 2.1; Figure 2.4 B) (Hartl and Hayer-Hartl, 2002).

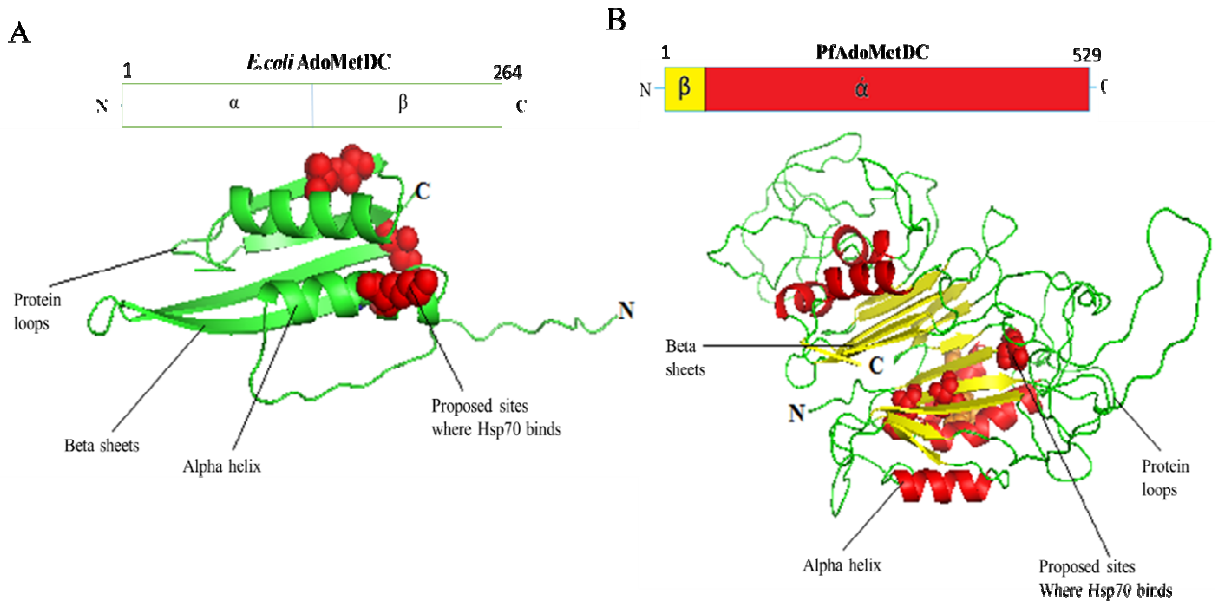

Figure 2.4: **Modeling of *E. coli* AdoMetDC and PfAdoMetDC showing the proposed Hsp70 binding sites.** Red color represent alpha helix, yellow, represent beta sheets, green color are loops and spheres in red and orange are hydrophobic site that Hsp70 prefers to bind.

2.4 Discussion

The role of molecular chaperones is to bind to hydrophobic patches of the substrate until it folds properly (Hartl and Hayer-Hartl, 2002). On the other hand, suitable client proteins for interaction with molecular chaperones are proteins with exposed hydrophobic regions. Even though *E. coli* AdoMetDC and PfAdoMetDC are from different species based on the sequence alignment results show that they share hydrophobic amino acids (such as leucine and glycine) that could be

recognized by DnaK and GroEL chaperones (Figure 2.1) (Lakshimipathy *et al*., 2010). Molecular chaperones such as Hsp70 and GroEL bind to proteins approximately 20-60 kDa in size, preferably with exposed hydrophobic patches (Hartl and Hartl, 2002; Baneyx, 2004). *E. coli* AdoMetDC was predicted to be 30 kDa and PfAdoMetDC 60 kDa (Table 2.1). Based on the analysis, *E. coli* AdoMetDC has pI value of 6.0 and PfAdoMetDC protein has pI value of 6.35 (Table 2.1). Mehlin and colleagues (2006) suggested that malarial proteins that were difficult to express in *E. coli* was due to the fact that have a pI above 6. The effects of the protein size on recombinant protein production in *E. coli* system has been discussed (Bikholtz *et al*., 2008). Their size difference suggests that both GroEL and Hsp70 prefers binding to larger proteins compared to the smaller proteins during their folding process (Hartl and Hayer-Hartl, 2002). This indicates that PfAdoMetDC might have more binding motifs that could be recognized by Hsp70 and GroEL chaperones, compared to *E. coli* AdoMetDC protein (30 kDa) (Table 2.1).

Kyte and Doolittle analysis indicated that *E. coli* AdoMetDC has a more pronounced hydrophobic profile compared to PfAdoMetDC suggesting that the protein is a good candidate for molecular chaperones (Figure 2.2). Williams and colleagues (2011) suggested that PfAdoMetDC protein expresses in *E. coli* BL21 Star^{TM} cells, however the major challenge was that the PfAdoMetDC co-purified with a contaminant protein of 70 kDa, *E. coli* DnaK. Hydrophobicity seems to be the driving force for Hsp70 to interact with newly synthesized proteins, the exposure of hydrophobic regions by PfAdoMetDC was shown to result in the increased recruitment of DnaK/Hsp70 (Figure 2.2).

DnaK binding sites on PfAdoMetDC were predicted to be more frequent compared to *E. coli* AdoMetDC based on the predictive tool that determines the possible binding motifs of DnaK on substrates (Figure 2.2) (Rudiger *et al*., 1997). The results show that DnaK binds on the β-sheet of the protein with high pronounced hydrophobicity residues. On the other hand, DnaK binds on the αhelix on *E. coli* AdoMetDC (Figure 2.3 A). DnaK is known to interact with both side chains and peptide backbone, while DnaJ only binds to side chains, to ensure that these chaperones interact with hydrophobic sequences without strong preference on the sequence (Rudiger *et al*., 2001).

It has been suggested that for chaperones to distinguish between folded and unfolded substrates, hydrophobic sequences are normally found on the interior of the folded substrates (Hartl and Hayer-Hartl, 2002). Together, these features appear to provide the necessary promiscuity that permits them to bind a wide range of substrates (Srinivasan *et al*., 2012). A proteomic study using a temperature sensitive GroEL mutant strain, demonstrated that DnaK and DnaJ bound all the GroEL substrates, indicating that the GroEL-dependent proteins also contain predicted DnaK-DnaJ binding sites (Srinivasan *et al*., 2012). This therefore shows that like DnaK-DnaJ system, GroEL could possibly interact with a relatively large number of protein substrates including PfAdoMetDC protein.

A study conducted showed that there are three classes of GroEL interacting substrate proteins based on their chaperone dependence (Kerner *et al*., 2005). Class I, include proteins that can fold spontaneously whose refolding yield could be optimized by chaperone interaction. However, their solubility is independent of GroEL (Coyle *et al*., 1997). GroEL has ability to bind exposed hydrophobic surfaces on substrates (Coyle *et al*., 1997). Hartl and Hayer-Hartl (2002),

demonstrated that trigger factor and DnaK/DnaJ/GrpE chaperones interact with nascent polypeptides upstream of the chaperonin and most of the proteins with class I behavior have the opportunity to fold before reaching GroEL. Class II proteins are unable to fold spontaneously, therefore they solely depend on chaperones assistance for them to fold to their native structures (Coyle *et al*., 1997). The solubility of the proteins of class II have been suggested to not get affected by overexpression of chaperonins, meaning that their solubility status does not change (Coyle *et al*., 1997). Class III proteins on the other hand constitute proteins that are absolutely dependent on the assistance of GroEL both *in vitro* and *in vivo* for folding to their native state (Coyle *et al*., 1997). Such proteins are highly aggregation prone and are unable to fold spontaneously. It is suggested that DnaK is able to interact with class III substrates to suppress their aggregation but cannot promote their folding (Coyle *et al*., 1997). Such proteins are transferred downstream for subsequent folding by GroEL (Coyle *et al.*, 1997). This indicates that molecular chaperones are cooperative and certain substrates require more than one set of chaperon system to reach their native structures. In addition to that the size and pronounced features of PfAdoMetDC strongly suggests that it is a suitable candidate for GroEL binding. Therefore, it is important to assess whether PfAdoMetDC could benefit from these molecular chaperones during its synthesis until it folds. In chapters subsequent to the current chapter, experimental work that was conducted to investigate the effect of co-expressing PfAdoMetDC in *E. coli* supplemented with exogenous plasmodial and *E. coli* chaperones will be described.

CHAPTER THREE

Co-expression of recombinant PfAdoMetDC with molecular chaperone combinations

3.1 INTRODUCTION

It has been suggested that the molecular chaperones of *E. coli* may occur at concentrations that are inadequate to meet the protein folding demands of the cell during the production of recombinant proteins (Carrio and Villaverde, 2002; de Marco and De Marco, 2004). Consequently, it has been proposed that expression of plasmodial recombinant proteins in *E. coli* in the presence of molecular chaperones of similar origin could solve this problem (Birkholtz *et al*., 2008; Stephens *et al*., 2011). *E. coli* serves as the host of choice for the production of recombinant proteins. However, one of the challenges of producing recombinant proteins in *E. coli* is that the products are occasionally released off the ribosomes as insoluble inclusion bodies. In addition, the use of strong promoters and high inducer concentration can generate yields of recombinant proteins exceeding 50% of the total cellular protein (Baneyx and Mujacic, 2004). Under such circumstances, the rate of protein production overwhelms the protein folding machinery, resulting in the generation of poor quality, misfolded recombinant proteins (Salvador, 2005).

In addition, the malarial genome is AT-rich which constitutes a further complication (Flick *et al.*, 2004). Mehlin and co-workers (2006) analyzed 1000 genes from *P. falciparum* that were heterologously expressed in *E. coli* and reported that only 337 were successfully expressed. Of these, only 63 were reported as soluble proteins. Apart from the AT-bias of the genome, proteins with the following features were also found to be difficult to express: high molecular weight (> 56 kDa), basic isoelectric point (pI > 6) and lack of homology with *E. coli* proteins (Mehlin *et al*., 2006). PfAdoMetDC is an essential protein involved in the biosynthesis of polyamines, making it a potential anti-malarial drug target (Muller *et al*., 2001; Wells *et al*., 2006). Obtaining

a pure and active product in fairly large quantities is important in facilitating the validation of drug targets through biochemical assays. Although the recombinant form of PfAdoMetDC has been expressed in *E. coli*, the protein co-purifies with *E. coli* proteins, and amongst them is DnaK (Williams *et al*., 2011). DnaK belongs to the Hsp70 family of chaperones whose main responsibility is to bind misfolded proteins to facilitate their fold (Bukau *et al*., 2006). It is therefore plausible that PfAdoMetDC is released off the ribosomes in a misfolded form, attracting DnaK. This suggests that the protein may not have been appropriately folded and this may have influenced the reported biophysical features of the protein. Therefore, the aim of this study was to apply the various combinations of molecular chaperones in the production of PfAoMetDC protein in *E. coli*. Briefly, PfHsp70 is a molecular chaperone of plasmodium origin, whereas KPf is a chimeric molecular chaperone of the ATPase domain of DnaK and the substrate binding domain of the PfHsp70 molecular chaperone (Shonhai *et al*., 2005).

The aims of this study were:

1) To co-express PfAdoMetDC with following sets of molecular chaperones:

Number	**Chaperone combinations**
i	DnaK+DnaJ
ii	KPf+DnaJ
iii	PfHsp70+DnaJ
iv	DnaK+DnaJ+GroEL
v	KPf+DnaJ+GroEL
vi	PfHsp70+DnaJ+GroEL

2) To analyse the solubility of PfAdoMetDC co-expressed with various combinations of molecular chaperones;

3) To purify PfAdoMetDC co-expressed with the above combinations of molecular chaperones; and

4) To clone DnaJ, KPf and GroEL in pQE30 vector to add N-terminal His-tag to facilitate their purification.

3.2 Materials and Methods

3.2.1 Materials

The following plasmids and reagents were used in this study: pBB535 encoding (*DnaK* and *DnaJ* genes) and pBB542 (*DnaK-DnaJ-GroEL*), and kindly donated from Dr. Bernd Bukau's lab (Germany; Deuerling *et al*., 1999; de Marco *et al*., 2007). The vector provides, with CoIE1 origin of replication, a T7 promoter (regulated by the product of lacl gene, induced by IPTG) and spectinomycin resistance. The construct pASKA-IBA/PfAdoMetDC, host to the codon harmonized *PFADOMETDC* gene which encodes for the α-subunit of the protein (approximately 60 kDa), was donated by Prof. Lyn-Marie Birkholtz (University of Pretoria, South Africa). The vector has the ampicillin resistance gene, the transcription for which is under the *tet*A operator, and is induced by AHT.

Construct pQE30/PfHsp70 has routinely been used to express PfHsp70 in *E. coli* (Matambo *et al*., 2004; Shonhai *et al*., 2005; Shonhai *et al*., 2008). The pQE60/*KPf* construct used in the present study was provided by Prof. A. Shonhai (University of Venda; Shonhai *et al*., 2005). The plasmid mini prep kit used in this study was purchased from Zymogen and the pQE30 vector was purchased from (Qiagen, Germany). The rest of the reagents used in this study are listed in Table 1 (Appendix D).

Table 3.1: *E. coli* strains and plasmids used for protein expression

Name of *E. coli* strain	Key features	Source/references
***E. coli* BB 1553 cells (MC4100 ΔdnaK52 Cm^R SidB1)**	• dnaK gene substituted with cat cassette • low DnaJ levels • grows at 30°c • does not grow at 40°c • chloramphenicol tretracycline	Bukau and Walker,1990; Ratelade *et al*.,2009
***E. coli* BL21 (DE3) Star™ cells**	• high efficiency protein expression • grows at 37°c • does not interfere with IPTG induction	Studier *et al*., 1990
***E. coli* XL1 Blue cells**	• high efficiency protein expression • grows at 37°c • does not interfere with IPTG induction	Bullock *et al*., 1987
Anti-strep tag II	• delectation of Strep-tag® II proteins	IBA, Germay
Anti-His tag	• for the detection of His-tag® proteins	Pierce, USA
Anti-DnaK	• probing of *E. coli* DnaK protein	Stressgen
Anti-PfHsp70	• delectation of KPf and PfHsp70 protein	Pesce *et al*., 2008
Anti-Hsp60	• detection of GroEL protein	Kindly donated by Prof. Alister Craig (Liverpool University, UK).

3.2.2 Methods

3.2.2.1 Cloning of DnaJ, KPf and GroEL in pQE30 vector to add N-terminal His-tag and facilitate their purification

DNA segments encoding DnaJ, KPf and GroEL were cloned into pQE30 plasmid vector in frame with an N-terminal His-tag to facilitate purification using nickel affinity chromatography. The aim was to purify these proteins to further investigate whether there is any possible interaction between them by conducting *in vitro* studies. Briefly these constructs were as follows: the DnaJ encoding segment was first amplified by polymerase chain reaction (PCR) from pBB535. The forward primer 5´-ATCACGGATCCATGGTCTAAGCAAGATATTATTACG-3´ and reverse primer 5´-TTGGCTGCAGTTAGCGGGTCAGGTCG-3´ were used with restriction sites ((*Bam*HI and *Pst*I respectively (underlined)). The amplified product and pQE30 was digested with *Bam*HI and *Pst*I. The *DnaJ* encoding segment was then ligated between *Bam*HI and *Pst*I restriction sites in the pQE30 plasmid vector. The integrity of the resulting pQE30/DnaJ construct was confirmed by restriction analysis as well as DNA sequencing.

The *KPf* encoding segment was first amplified by PCR from pQE60/*KPf* plasmid. A forward primer, 5´-ATCACGGATCCATGGTGAAACTCTGG-3´ and reverse primer, 5´-TAATTAAGCTTTTCCACTTGGCATTCC-3´ were used for PCR amplification. The forward and reverse primers contained *Bam*HI and *Hind*III restriction sites (underlined) respectively. The amplified product was digested with *Bam*HI, *Hind*III and pQE30 vector. The *KPf* encoding segment was then inserted between *Bam*HI and *Hind*III restriction sites in the pQE30 plasmid vector. The integrity of the resultant pQE30/*KPf* was confirmed by restriction analysis as well as DNA sequencing.

The *GroEL* encoding segment was amplified from pBB542 plasmid. A forward primer, 5´-TCCGCATGCATGGCACTAAAGAC-3´, and reverse primer, 5´-TAATTAAGCTTTTACATCATGCCGCCC-3´, were used for PCR amplification. The forward and reverse primers contained *Sph*I and *Hind*III restriction sites (underlined) respectively. The amplified product was digested with *Sph*I, *Hind*III and the pQE30 vector. The *GroEL* encoding segment was then inserted between the *Sph*I and *Hind*III restriction sites in the pQE30 plasmid vector. The integrity of the resultant pQE30/*GroEL* was confirmed by restriction analysis as well as DNA sequencing. Primers were synthesized by Inqaba biotechnical Industries (PTY) Ltd, South Africa. The PCR reaction is explained in detail in Appendix B.

3.2.2.2 The cloning strategies used to substitute DnaK with either KPf or PfHsp70

Site directed mutagenesis was carried out by designing primers that introduced *Bam*HI before the start codon and *Sma*I after the stop codon of the *DnaK* gene in order to replace the latter with *PfHsp70* or *KPf* by introducing restriction enzymes (Appendix B, table B.1). The introduction of *Bam*HI and *Sma*I allowed for the excision of *DnaK*. Figure 3.1, shows the cloning strategy that was used to replace *DnaK* with *PfHsp70* or *KPf* in pBB535 (7061 bp) and pBB542 (9177 bp). This cloning strategy resulted in the following constructs: PfHsp70+DnaJ, KPf+DnaJ, PfHsp70+DnaJ+GroEL and KPf+DnaJ+GroEL.

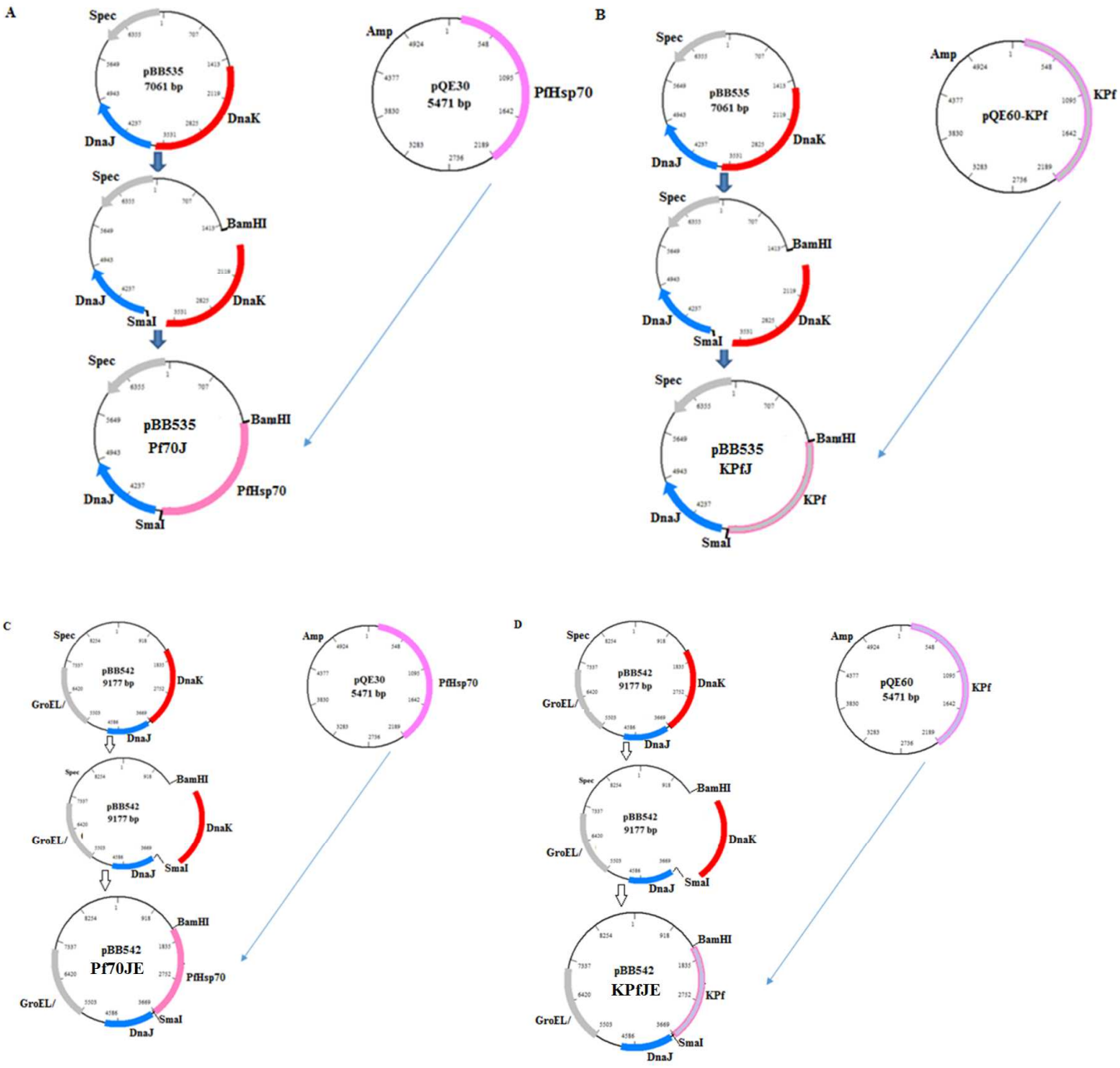

Figure 3.1: **The cloning strategy of PfHsp70 or KPf to substitute DnaK**. The *Bam*HI restriction site was introduced before the start codon of DnaK and the *Sma*I restriction site was introduced after the stop codon of DnaK with site directed mutagenesis. This resulted to the following constructs: **(A)** pBB535-Pf70J; **(B)** pBB535-KPfJ; **(C)** pBB542-Pf70JE; and **(D)** pBB542-KPfJE.

3.2.2.3 Protein expression and purification

The monofunctional PfAdoMetDC protein was expressed in *E. coli* BL21 Star™ (DE3) cells (Williams *et al*., 2011). This was then followed by co-expression of the chaperone sets; the co-expression of chaperones was initiated by the addition of 1 mM IPTG when cells reached an optical density (OD) of 0.2. The co-transformed cells were incubated further, until the OD reached 0.7. The expression of the target gene was then initiated by the addition of 2 ng/ml anhydrotetracycline (AHT).

To confirm whether the protein was expressed in a soluble form, either on its own or in the presence of suplementary chaperones in *E. coli* BL21 Star™ (DE3) cells, the cell suspension was centrifuged at 5000 g for 20 minutes at 4ºC and resuspended in 5 ml lysis buffer (0,01 mM Tris, pH 7.5; 10 mM Imidazole, containing 1 mM (aminoethyl benzenesulforyl floride hydrochloride (AEBSF)) and 1 mg/ml of lysozyme)) (Gitau *et al*., 2012). The cells were frozen overnight at -80ºC then thawed rapidly and sonicated mildly. The cell lysate was centrifuged at 12000 g at 4ºC for 20 minutes (Stephens *et al*., 2011). The supernantant was collected (soluble fraction) and the pellet (insoluble fraction) was resuspended in 3 ml of phosphate buffered saline solution (PBS, pH 7.5) . This was followed by an analysis of the samples using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

The strep-tagged PfAdoMetDC purification was conducted under native conditions using the strep-tactin column purification system (Williams *et al*., 2011). After packing the column with 1 ml strep-tactin, the column was washed twice with 10 ml of Buffer W (150 mMNacl, EDTA, 100 mMTris, pH 8). The supernatant on the column was allowed to flow three times through, followed by three washes with Buffer W. A protein was then eluted with 3 ml of Buffer E (100

mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, and 2.5 mM D-Desthiobiotin). The column was regenerated with 10 ml of Buffer R (1 mM 4-hydroxyl azobenzene-2-carboxylic acid (HABA) in Buffer W), which was then washed out with buffer W until the strep-tactin column became white or clear. The column was stored at 4^{o}C in Buffer W.

3.2.2.4 Purification of His-tagged DnaJ, GroEL and Hsp70s proteins

The His-tagged proteins were expressed in *E. coli* XL1 Blue cells as previously described (Matambo *et al*., 2004; Gitau *et al*., 2012). A colony was picked and inoculated in 5 ml 2x YT liquid broth supplemented with 100 mg/ml of ampicillin, then grown overnight at 37^{o}C on a shaking platform for 16 hours. After transferring into 225 ml of liquid broth, this was allowed to grow up to 0.5 $OD_{600.}$ The induction of the protein was initiated by the addition of 1 mM IPTG, and the cultures were grown for 5 hours before the cells were harvested following a previously employed protocol (Stephens *et al*., 2011). The HisPur Ni-NTA Resin purification system (Pierce, USA) was used following the supplier's standard protocol, with minor changes.

3.2.2.5 Protein quantification and Western analysis

To determine the concentration of the protein after its purification, the Bradford method was used for protein quantification (Bradford, 1976). The expression of PfAdoMetDC and molecular chaperones was confirmed by a Western blot analysis. This was done by the transfer of proteins, onto a nitrocellulose membrane for 1 hour at 100 V after an SDS-PAGE. The membrane was washed twice with TBS (50 mM Tris, 150 mM NaCl) and blocked with a 5% blocking buffer of 5 g non-fat milk powder in 100 ml of TBS for 1 hour. Following that was the incubation of the membrane with primary antibody (1:1000). The monofunctional PfAdoMetDC was tagged with

Strep-tag II; monoclonal anti-strep-tag II antibodies (IBA, Germany) were used for its detection. KPf and PfHsp70 were detected using anti-Hsp70 antibodies (Pesce *et al*., 2008). For DnaK detection, anti-DnaK antibodies were used (Stressgen). GroEL was detected using anti-Hsp60 antibodies. After overnight incubation with the primary antibody, the membrane was washed twice with TBST (50mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20, pH7.4) for 20 minutes, followed by incubation with the secondary antibody in 5% blocking buffer for 1 hour on ice with a shaking platform. The membrane was washed four times with TBST for 15 minutes each. This was then followed by the addition of ECL with hydrogen peroxide to the membrane in order to view the reaction of antibodies with specific proteins in the dark room.

3.3 Results

3.3.1 Confirmation of pBB535-Pf70J constructs

PfHsp70 was PCR amplified from pQE30/*PfHsp70* and the product size of 2054 base pairs was analyzed on agarose gel (Figure 3.2 B). After digestion with restriction enzymes, the product was cloned into pBB535 expression vector in *Bam*HI and *Sma*I restriction sites in place of *DnaK* (7191 bp) (Figure 3.2 A). The size of the fragment that corresponds with the PCR product was recovered when digested with both enzymes (Figure 3.2 C, lane 4). The integrity of the construct was further confirmed through sequencing.

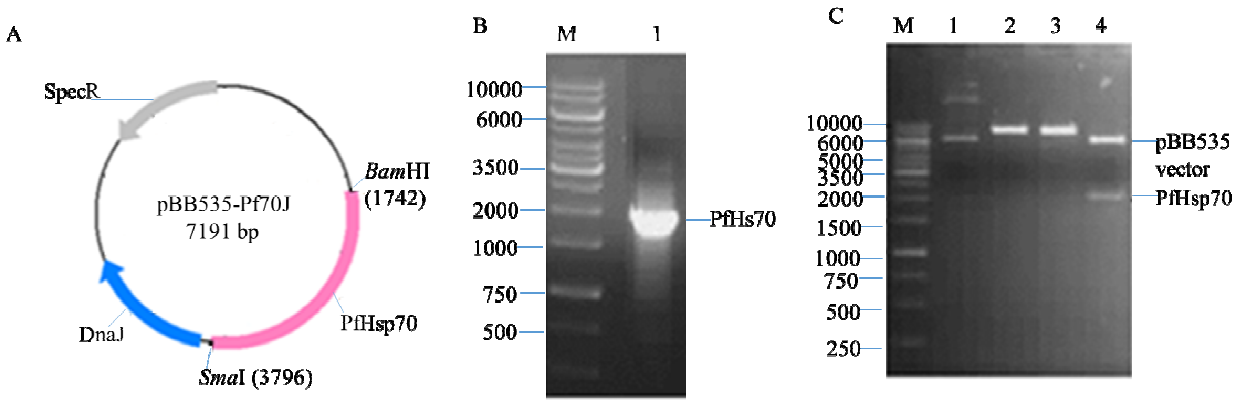

Figure 3.2 Diagnostic analysis of pBB535-Pf70J. (A) Plasmid map of pBB535-Pf70J construct. **(B)** Ethidium bromide stained agarose gel of PfHsp70 PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pBB535-Pf70J: lane M, DNA molecular markers; lane 1, undigested pBB535-Pf70J; lane 2 plasmid digested with *Bam*HI; lane 3 plasmid digested with *Sma*I; and lane 4 plasmid digested with both *Bam*HI and *Sma*I.

3.3.2 Confirmation of pBB535-KPfJ

KPf gene was PCR amplified from pQE60/*KPf* plasmid (2009 bp) (Figure 3.3 B) and cloned into pBB535 expression vector in *Bam*HI and *Sma*I restriction sites in place of *DnaK* (7146 bp). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered when digested with both enzymes (Figure 3.3 B, C, lane 4). The integrity of the construct was further confirmed through sequencing.

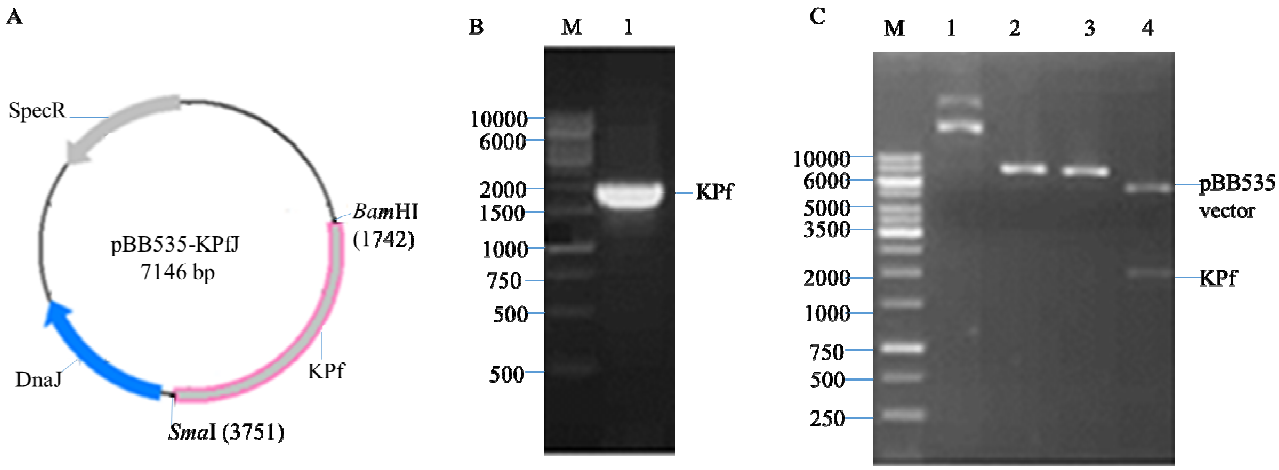

Figure 3.3 Diagnostic analysis of pBB535-KPfJ. (A) plasmid map of pBB535-KPfJ construct. **(B)** Ethidium bromide stained agarose gel of KPf PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pBB535-KPfJ. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Bam*HI; lane 3, plasmid digested with *Sma*I; and lane 4, plasmid digested with both *Bam*HI and *Sma*I.

3.3.3 Confirmation of pBB542-Pf70JE

The *PfHsp70* gene was PCR amplified from pQE30/*KPf70* plasmid (2054 bp) (Figure 3.4 B) and cloned into pBB542 expression vector in *Bam*HI and *Sma*I restriction sites in place of *DnaK* (9314 bp) (Figure 3.4 A). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered (Figure 3.4 C, lane 4). The integrity of the construct was further confirmed through sequencing.

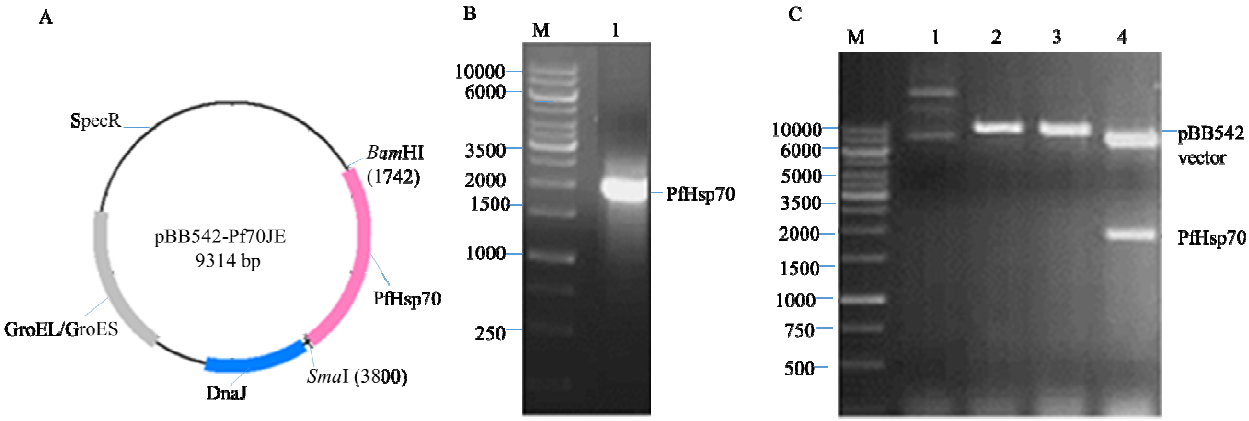

Figure 3.4 Diagnostic analysis of pBB542-Pf70JE. (A) Plasmid map of pBB542-Pf70JE construct. **(B)** Ethidium bromide stained agarose gel of KPf PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pBB542-Pf70JE. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Bam*HI; lane 3, plasmid digested with *Sma*I; and lane 4, plasmid digested with both *Bam*HI and *Sma*I.

3.3.4 Confirmation of pBB5542-KPfJE

KPf gene was PCR amplified from pQE60/*KPf* plasmid (2009 bp) (Figure 3.5 B) and cloned into pBB542 expression vector in *Bam*HI and *Sma*I restriction sites in place of *DnaK* (9268 bp) (Figure 3.5 A). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered when digested with both enzymes (Figure 3.5 C, lane 4). The integrity of the construct was further confirmed through sequencing.

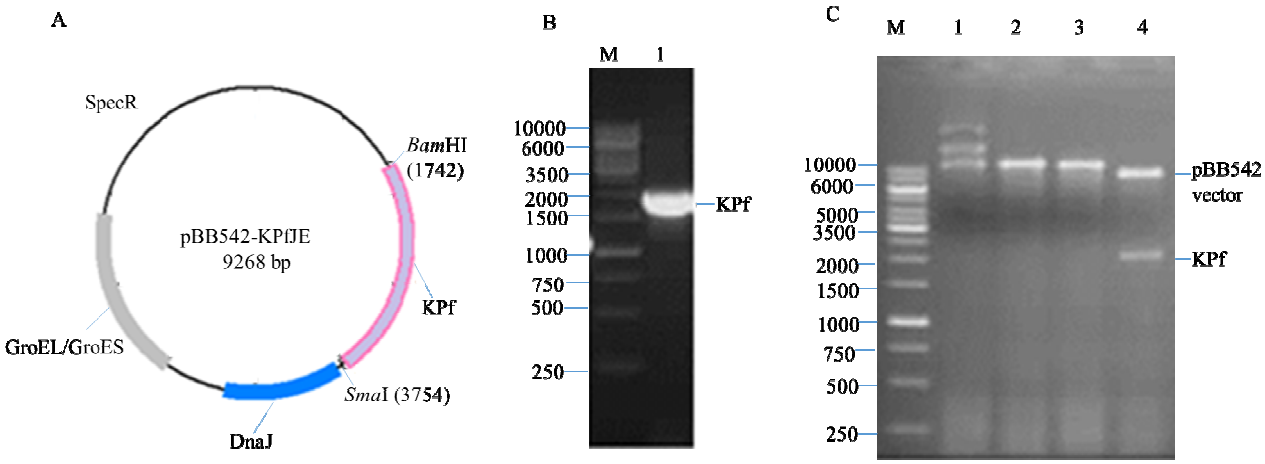

Figure 3.5 Diagnostic analysis of pBB542-KPfJE. (A) Plasmid map of pBB542-KPfJE construct. **(B)** Ethidium bromide stained agarose gel of KPf PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pBB542-KPfJE. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Bam*HI; lane 3 plasmid digested with *Sma*I; and lane 4, plasmid digested with both *Bam*HI and *Sma*I.

3.3.5 Confirmation of pQE30/DnaJ construct

The *DnaJ* gene was PCR amplified from pBB535(*DnaK-DnaJ*) plasmid (1286 bp) (Figure 3.6 B) and cloned into pQE30 expression vector in *Bam*HI and *Pst*I restriction sites (4592 bp) (Figure 3.6 A). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered (Figure 3.6 C, lane 4). The integrity of the construct was further confirmed through sequencing.

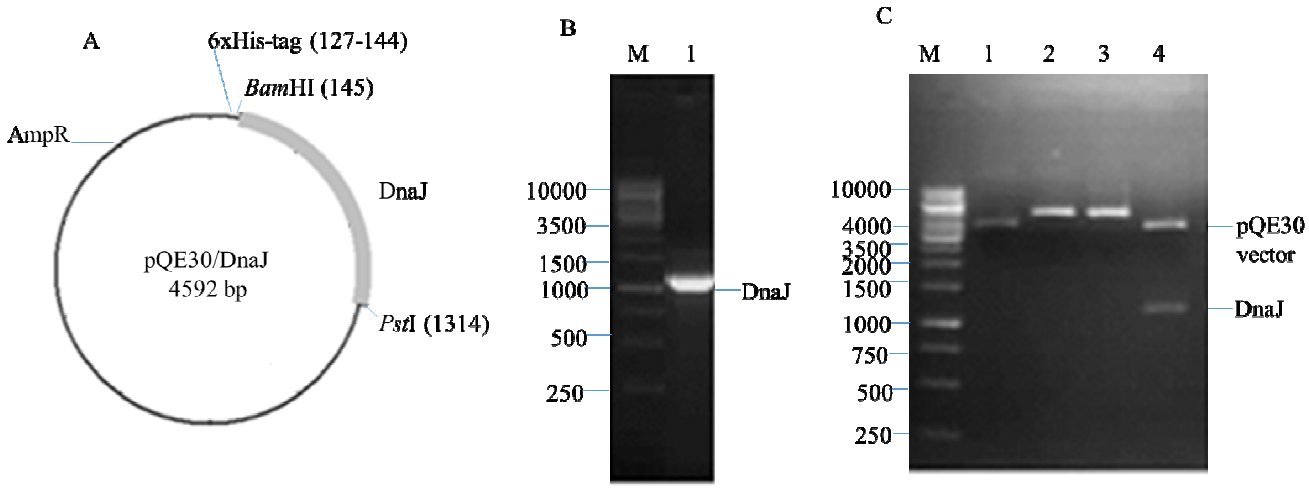

Figure 3.6: Diagnostic analysis of pQE30/*DanJ*. (A) Plasmid map of pQE30/*DnaJ*. **(B)** Ethidium bromide stained agarose gel of DnaJ PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pQE30/DnaJ. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Bam*HI; lane 3, plasmid digested with *Pst*I; and lane 4, plasmid digested with both *Bam*HI and *Pst*I.

3.3.6 Confirmation of pQE30/*GroEL* construct

The *GroEL* gene was PCR amplified from pBB542 (*DnaK-DnaJ-GroEL-GroES*) plasmid (1804 bp) (Figure 3.7 B) and cloned into pQE30 expression vector in *Hind*III and *Sph*I restriction sites (Figure 3.7 A). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered (Figure 3.7 C, lane 4). The integrity of the construct was further confirmed through sequencing.

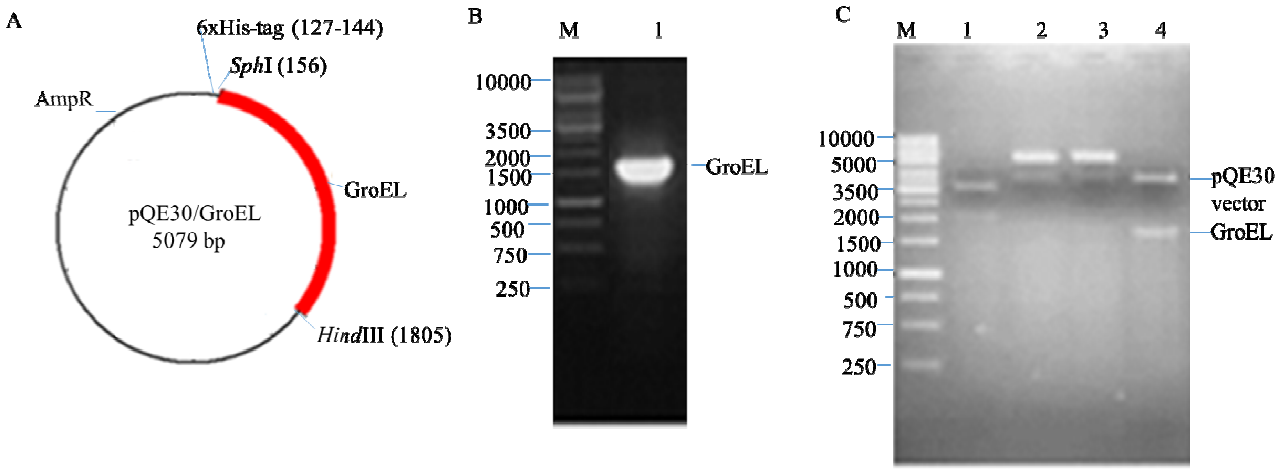

Figure 3.7: Diagnostic analysis of pQE30/*GroEL*. (A) Plasmid map of PQ30/*GroEL*. **(B)** Ethidium bromide stained agarose gel of *GroEL* PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pQE30/GroEL. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Sph* I; lane 3, plasmid digested with *Hind* III; and lane 4, plasmid digested with both *Sph*I and *Hind*III.

3.3.7 Confirmation of pQE30/*KPf* construct

The *KPf* gene was PCR amplified from pQE60/*KPf* plasmid (2009 bp) (Figure 3.8 B) and cloned into pQE30 expression vector in *Bam*HI and *Hind*III restriction sites (5195 bp) (Figure 3.8 A). The same enzymes were then used for a restriction analysis of the construct and the size of the fragment that corresponds with the PCR product was recovered (Figure 3.8 C, lane 4). The integrity of the construct was further confirmed through sequencing.

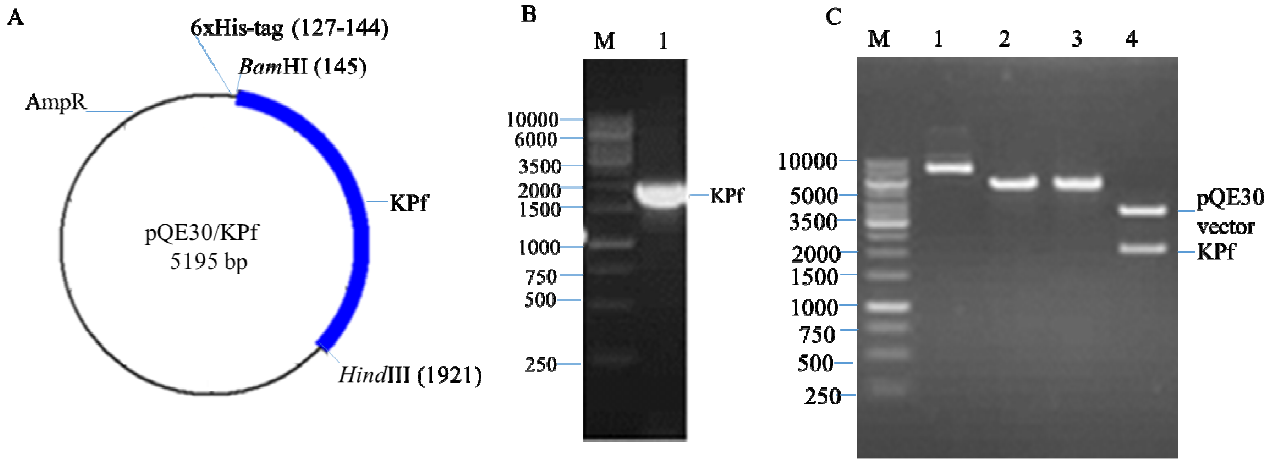

Figure 3.7: Diagnostic analysis of pQE30/*KPf*. (A) Plasmid map of pQE30/*KPf*. **(B)** Ethidium bromide stained agarose gel of KPf PCR product: lane M, DNA molecular markers; lane 1, PCR product. **(C)** Endonuclease restriction digest of pQE30/*KPf*. DNA was loaded onto the gel in the following order: lane M, DNA molecular markers; lane 1, undigested plasmid; lane 2, plasmid digested with *Bam*HI; lane 3, plasmid digested with *Hind*III; and lane 4, plasmid digested with both *Bam*HI and *Hind*III.

3.4 Expression and purification of his-tagged recombinant molecular chaperones

3.4.1 Expression of recombinant DnaJ protein

Recombinant DnaJ protein was expressed in *E. coli* XL1 blue cells. Figure 3.8 shows the SDS-PAGE analysis of the samples that were collected pre-induction and post-induction. The expected size of the protein was around 40 kDa species. The visibility of the protein pre-induction was also noticeable (Figure 3.8 lane 0). However, the protein showed an increase in expression in samples collected hourly, post-induction (Figure 3.8, lane 1-5). The expression of the protein was confirmed by means of a Western blot analysis (lower panel).

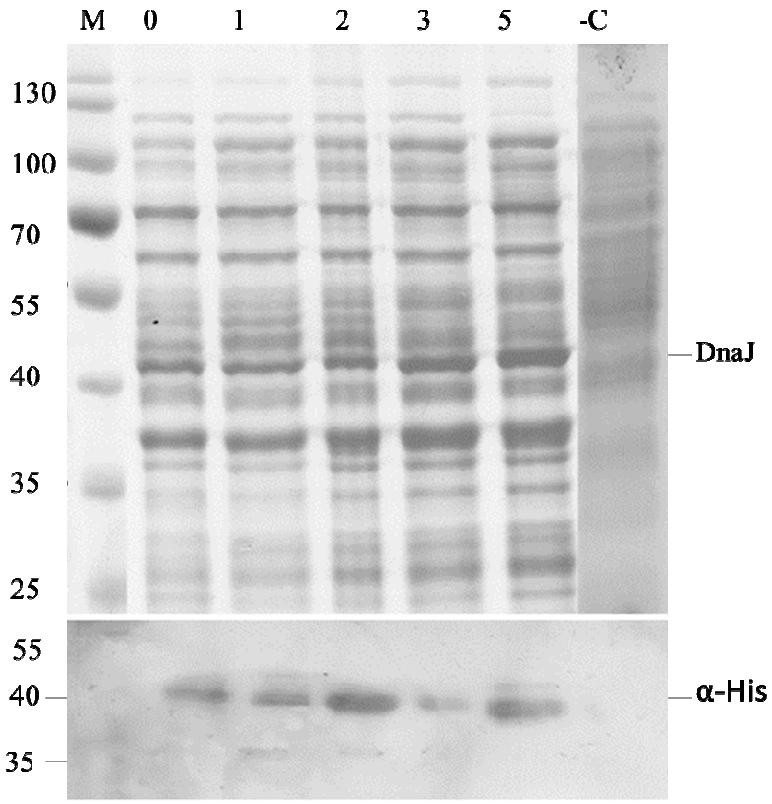

Figure 3.8: Analysis of the expression of pQE30/DnaJ in XL1 blue cells. SDS-PAGE (upper panel) and Western analysis (lower panel) of the expression of *E. coli* Hsp40 (DnaJ) in XL1 Blue cells. Lane M, page ruler pre-stained protein ladder (sizes indicated on the left hand side); lane 0, pre-induction sample; lanes 1-3, hourly samples taken post-induction (total protein extracts); lane 5, sample taken overnight (total protein extracts); lane –C, negative control. Western blot analysis conducted using a monoclonal anti-polyHistidine-peroxidase antibody.

3.4.2 Expression of recombinant GroEL protein

Recombinant his-tagged GroEL protein was expressed in *E. coli* XL1 blue cells. The expected size of the protein was around 60 kDa species (Figure 3.9). The protein showed an increase in expression in hourly collected samples post-induction (Figure 3.9, lane 1-3). The expression of the protein was confirmed by means of a Western blot analysis using anti-His antibodies (lower panel).

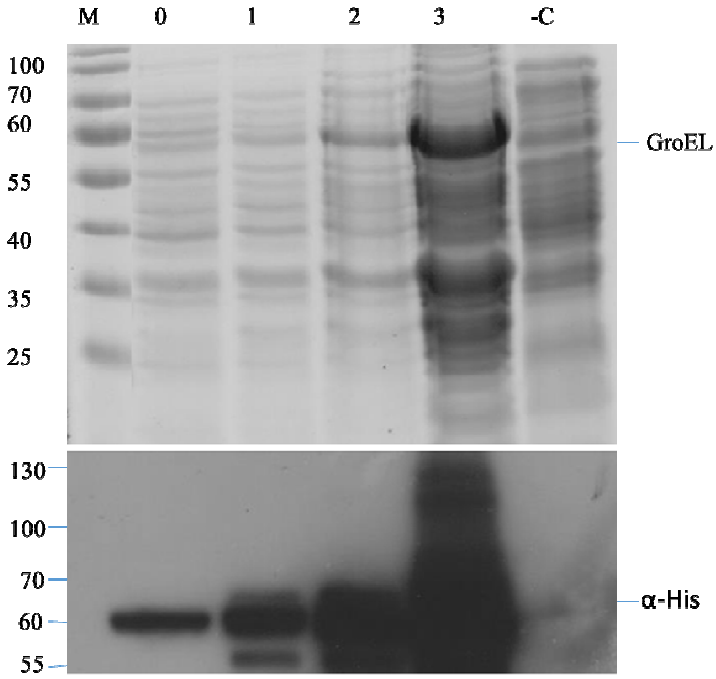

Figure 3.9: Analysis of the expression of pQE30/GroL in XL1 blue cells. SDS-PAGE (upper panel) and Western analysis (lower panel) of the expression of *E. coli* GroEL in XL1 Blue cells. Lane M, page ruler pre-stained protein ladder (sizes indicated on the left hand side); lane 0, pre-induction sample (total protein extract); lane 1-3, hourly samples taken post-induction (total protein extracts); lane -C, negative control cells transformed with pQE30 plasmid after induction with IPTG. Western blot analysis conducted using monoclonal anti-polyHistidine-peroxidase antibody.

3.4.3 Expression of recombinant KPf protein

Recombinant his-tagged KPf protein was expressed in *E. coli* XL1 blue cells. Figure 3.10, shows the SDS-PAGE analysis of the protein samples. The expected size of the protein is around 70 kDa species. The visibility of the protein pre-induction was also noticeable (Figure 3.10 lane 0). The protein showed an increase in expression in samples collected hourly post-induction (Figure 3.10, lane 1-5). The expression of the protein was confirmed by means of a Western blot analysis (lower panel).

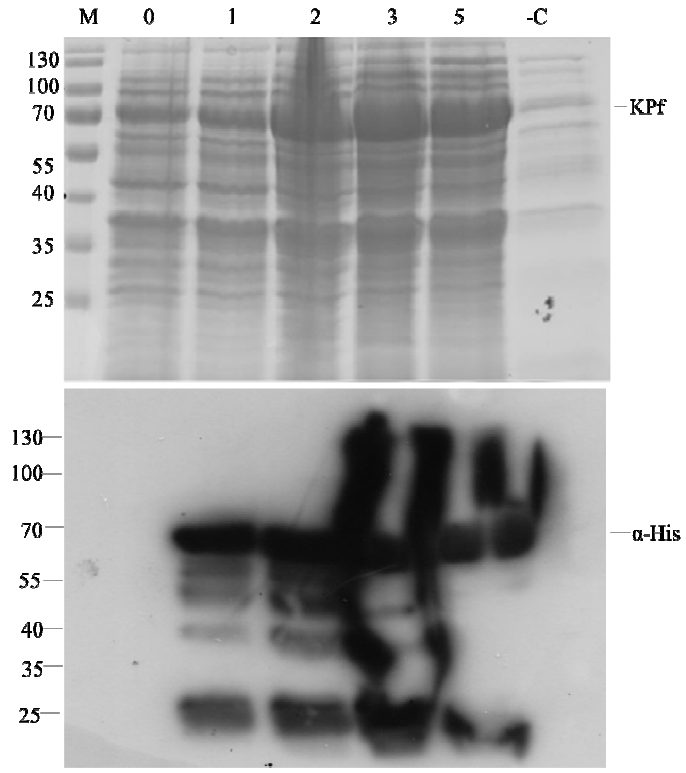

Figure 3.10: Analysis of the expression of pQE30/KPf in XL1 blue cells. SDS-PAGE (upper panel) and Western analysis (lower panel) of the expression of KPf in XL1 Blue Cells. Lane M, page ruler, pre-stained protein ladder (sizes indicated on the left hand side); lane 0, pre-induction sample (total protein extract); lanes 1-5, hourly samples taken post-induction (total protein extracts); lane -C, negative control cells transformed with pQE30 plasmid after induction with IPTG. Western blot analysis conducted using monoclonal anti-polyHistidine-peroxidase antibody.

3.5 Expression of PfAdoMetDC from *E.coli* BL21 (DE3) Star™ cells

3.5.1 Expression of recombinant PfAdoMetDC protein

The expression of recombinant PfAdoMetDC on its own was conducted as previously described (William *et al*., 2011). Figure 3.11, shows the SDS-PAGE analysis of the expressed protein, which was confirmed by a Western blot analysis using anti-strep tag antibodies. Based on the SDS-PAGE, the expression of PfAdoMetDC protein increased over time as post-induction. This was confirmed by a Western blot analysis.

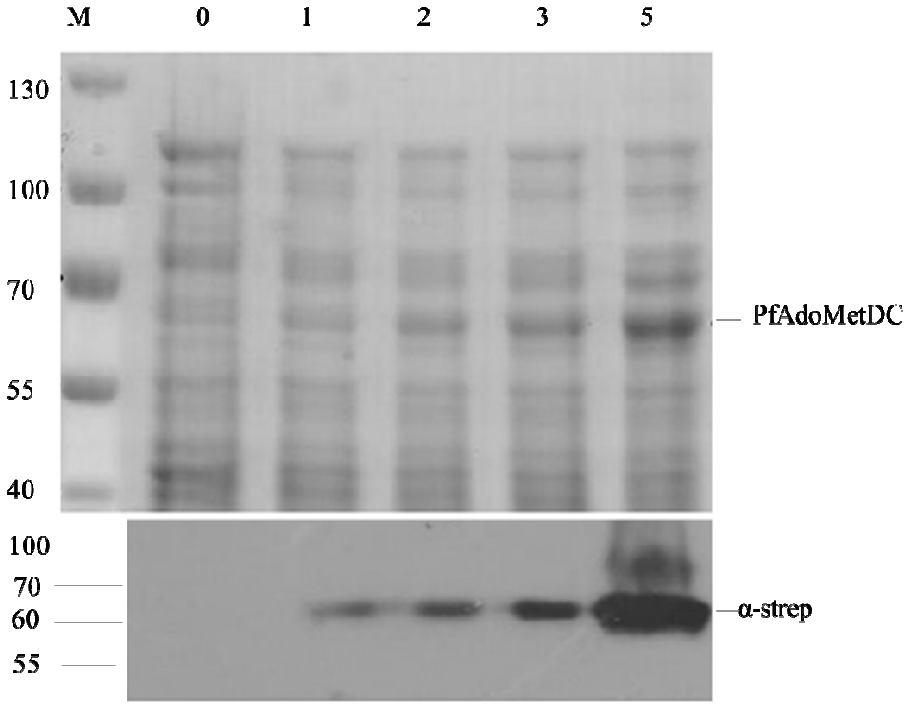

Figure 3.11: Analysis of the expression of pASK-IBA/PfAdoMetDC in *E. coli* BL21 star™ cells. This shows the SDS-PAGE (upper panel) and Western analysis (lower panel) of the expression of PfAdoMetDC in BL21 cells. Lane M, page ruler pre-stained protein ladder (sizes indicated on the left hand side); lane 0, pre-induction sample (total protein extract); lanes 1-5, hourly samples taken post-induction (total protein extracts). Western blot analysis conducted using monoclonal anti-strep tag antibody.

3.5.2 Co-expression of PfAdoMetDC with molecular chaperones

The following chaperone combinantions were co-expressed with PfAdoMetDC in *E. coli*: DnaK+DnaJ; KPf+DnaJ and PfHsp70+DnaJ. The chaperones were expressed as species of 70 kDa (representing DnaK, KPf and PfHsp70), 60 kDa (representing GroEL) and 40 kDa (representing DnaJ) (Figures 3.12 A, B and C respectively). The expression of the chaperones was initiated by using IPTG followed by the addition of AHT for the induction of PfAdoMetDC as a co-expression partner. Molecular chaperones and PfAdoMetDC were successfully expressed in *E. coli* BL21 (DE) star^{TM} cells as a host system (Figure 3.12). However, expression of recombinant DnaJ did not appear to significantly increase above the endogenous level (Figure 3.12). Since significantly expressed DnaJ did not rise above the endogenous levels of the protein, it was assumed that only the supplementary Hsp70 chaperones had an influence on the production of PfAdoMetDC. The expression of *E. coli* DnaK chaperone was confirmed by a Western blot analysis using anti-DnaK antibodies (Figure 3.12 D). In addition, the expression of PfHsp70 and KPf chaperones was confirmed using anti-PfHsp70 antibodies (Figure 3.12 E and F). The expression PfAdoMetDC was confirmed by a Western blot analysis using anti-strep tag antibodies (Figure 3.12 G, H and I).

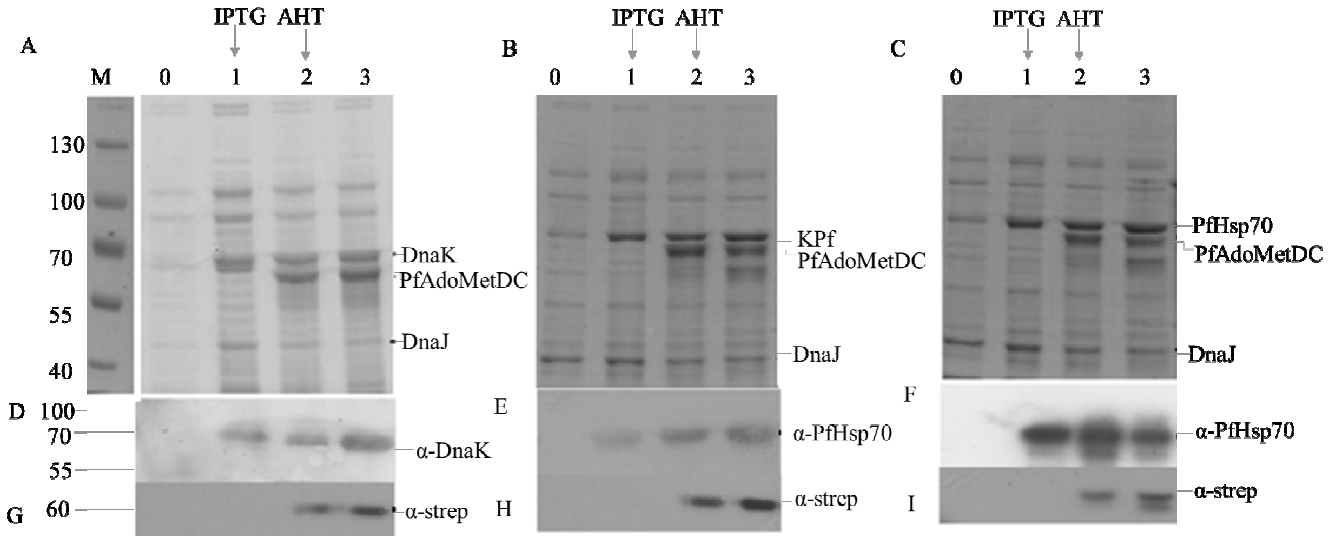

Figure 3.12 Analysis of co-expression of combined chaperones from plasmodial and *E. coli* origin with PfAdoMetDC in BL21 (DE3) Star™cells. (A) SDS-PGE and Western blot analysis of the co-expression of DnaK+DnaJ with PfAdoMetDC. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane 0, sample taken before induction; Lane 1, sample taken after induction with 1 mM IPTG; Lanes 2-3, sample taken after induction with 2 ng/ml AHT. **(B)** Co-expression of KPf+DnaJ with PfAdoMetDC. Lane 0, sample taken before induction; Lane 1, sample taken after induction with 1 mM IPTG; Lanes 2-3, sample taken after induction with 2 ng/ml AHT. **(C)** Co-expression of PfHsp70+DnaJ with PfAdoMetDC; Lane 0, sample taken before induction; Lane 1, sample taken after induction with 1 mM IPTG; Lanes 2-3, sample taken after induction with 2 ng/ml AHT. **(D)** Western blot analysis using anti-DnaK. **(E** and **F)** Anti-PfHsp70 antibodies. **(G**, **H** and **I)** Use of anti-strep tag antibodies**.**

3.5.3 Expression of Hsp70-GroEL chaperone combinations

The pBB542 plasmid encoding DnaK+DnaJ+GroEL was expressed in *E. coli* BL21 (DE3) star™ cells as previously described (de Marco *et al.*, 2007). Figure 3.13, shows the SDS-PAGE analysis of the proteins and their respective sizes (70, 60, and 40). Molecular chaperones from *E. coli* were successfully expressed (Figure 3.13 A) and a Western blot analysis using anti-DnaK antibodies and anti-GroEL was performed to confirm their expression (Figure 3.13 D and G) and The expression of KPf+DnaJ+GroEL and PfHsp70+DnaJ+GroEL combined chaperones in *E.*

coli BL21 (DE3) star™ cells. The expression of these chaperone sets were confirmed by SDS-PAGE and a Western blot analysis (Figure 3.13 B and C and Figure 3.13 E and F). The expression of Hsp60 was confirmed by a Western blot analysis using anti-GroEL antibodies (Figure 3.13 H and I).

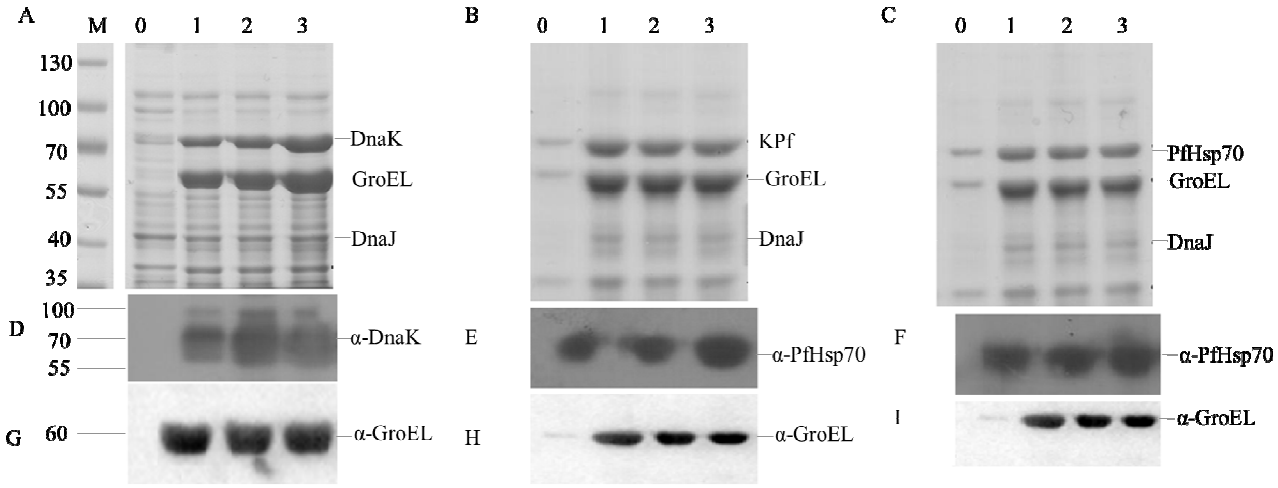

Figure 3.13: Analysis of expression of GroEL and various Hsp70-DnaJ combinantions in *E. coli* BL21 (DE3) star™ cells. (A) SDS-PGE and Western blot analysis expression of DnaK+DnaJ+GroEL. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane 0, sample taken before induction; Lanes 1-3 sample taken after induction with 1 mM IPTG. **(B)** Expression of KPf+DnaJ+GroEL; Lane 0, sample taken before induction; Lane 1-3, sample taken after induction with 1 mM IPTG. **(C)** Expression of PfHsp70+DnaJ+GroEL. Lane 0, sample taken before induction, Lanes 1-3, sample taken after induction with 1mM IPTG. **(D)** Western blot analysis using anti-DnaK. **(E** and **F)** Anti-PfHsp70 antibodies. (**G**, **H** and **I**) Anti-GroEL antibodies were used to probe for GroEL.

3.5.4 Co-expression of PfAdoMetDC with Hsp70-DnaJ-GroEL chaperone combinations

The following chaperone combinantions were co-expressed with PfAdoMetDC in *E. coli*: DnaK+DnaJ+GroEL, KPf+DnaJ+GroEL, PfHsp70+DnaJ+GroEL (Figure 3.14 A, B and C). The proteins were subjected to SDS-PAGE analysis and further confirmed by a Western blot analysis (Figure 3.14 A, B and C). The expression of PfAdoMetDC in the presence of DnaK+DnaJ+GroEL chaperone system seemed to have separated well based on the SDS-PAGE (Figure 3.14 B and C) compared to the chaperone combination system (Figure 3.14 A). However, a Western blot analysis using anti-strep tag antibodies shows that PfAdoMetDC was expressed, even though it is not clearly observed on the SDS-PAGE (Figure 3.14 G, H and I). It is also worth noting that the Western blot analysis revealed that PfAdoMetDC was produced after the addition of AHT, suggesting that the PfAdoMetDC protein may have been hidden by GroEL since they are a similar size (Figure 3.14, G, H and I on lane 3 on the Westerns). The GroEL expression was also confirmed by a Western blot analysis using anti-GroEL antibodies (Figure 3.14 J, K and L).

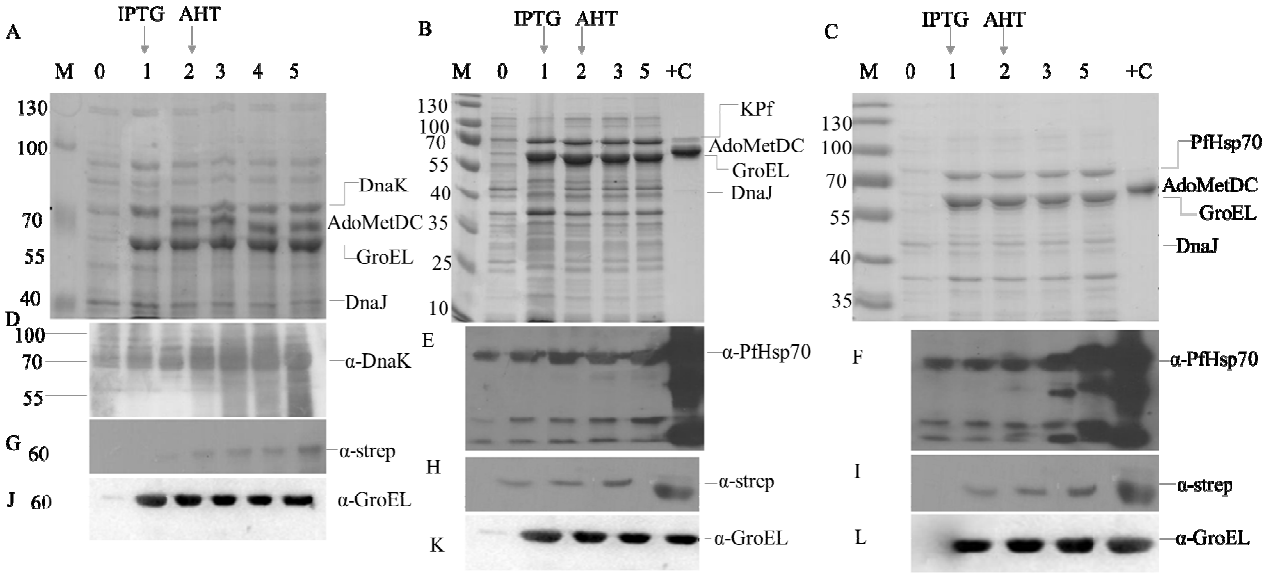

Figure 3.14: Co-expression of PfAdoMetDC with Hsp70-DnaJ-GroEL chaperone combinantions. (A) SDS-PGE and Western blot analysis co-expression of DnaK+DnaJ+GroEL with PfAdoMetDC. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane 0, sample taken before induction, Lane 1, sample taken after induction with 1 mM IPTG; Lanes 2-5, sample taken after induction with 2 ng/ml AHT. **(B)** Co-expression of KPf+DnaJ+GroEL with PfAdoMetDC. Lane 0, sample taken before induction; Lane 1, sample taken after induction with 1 mM IPTG; Lane 2-5, sample taken after induction with 2 ng/ml AHT. **(C)** Co-expression of PfHsp70+DnaJ+GroEL with PfAdoMetDC. Lane 0, sample taken before induction; Lane 1, sample taken after induction with 1 mM IPTG; Lane 2-5, sample taken after induction with 2 ng/ml AHT. **(D)** Western blot analysis using anti-DnaK. (**E** and **F)** Both probed by anti-PfHsp70 antibodies. **(G**, **H** and **I)** Confirmed by using anti-strep tag antibodies. **(J**, **K** and **L)** Probed using anti-GroEL antibodies. Lane +C, represents a positive control (purified PfAdoMetDC)..

Based on a Western blot analysis using anti-PfHsp70 to probe for either KPf or PfHsp70, the antibodies also detected what could be degradation products or incomplete translation of either KPf or PfHsp70 (Figure 3.14, lane E and lane F). This could have compromised their compatability to facilitate PfAdoMetDC folding. However, full length species of KPf and PfHsp70 were produced and these may have provided chaperone function during the co-expression of PfAdoMetDC.

3.5.5 Solubility study of PfAdoMetDC protein

Solubility studies were performed to investigate the effects of supplementary chaperones on the solubility of PfAdoMetDC produced in *E. Coli* BL21(DE3) star^{TM} cells. The protein was recovered in the supernanant fraction indicating that the protein was soluble and confirmed by Western blot analysis using anti-strep tag antibodies (Figure 3.15 A). Molecular chaperones were also observed, both in pellet and supernatant, based on the SDS-PAGE analysis. It is possible that molecular chaperones associated with aggregated proteins in an attempt to resolubilize the proteins.

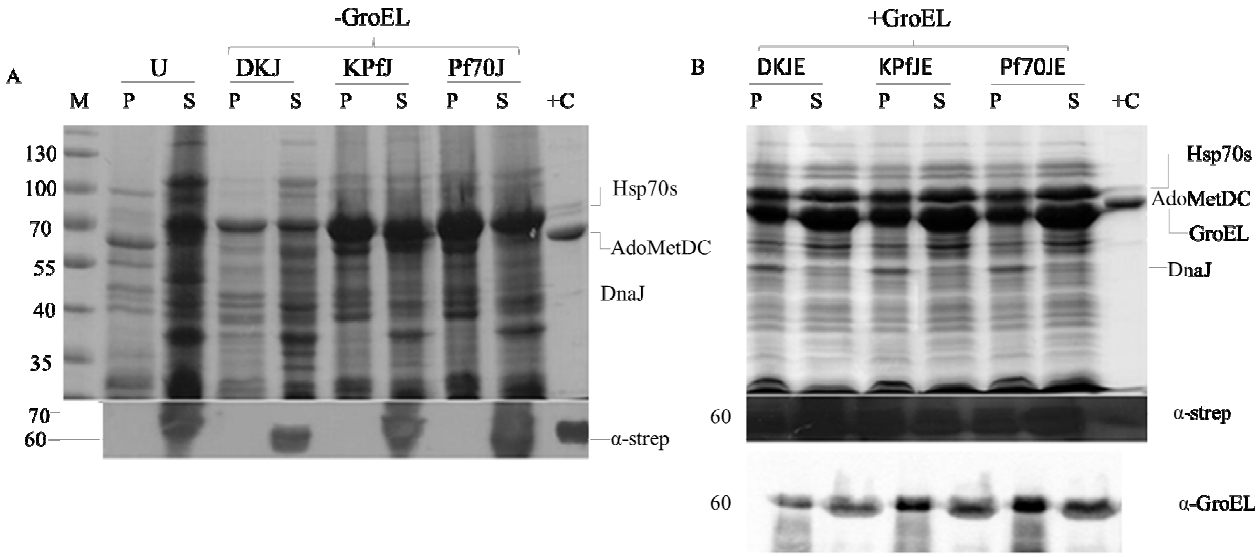

Figure 3.15: Solubility of PfAdoMetDC co-expressed with Hsp70-DnaJ-GroEL combinations. (A) Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane S, soluble fraction; Lane P, insoluble protein of AdoMetDC protein expressed on its own. (U, without the supplementation of molecular chaperones). PfAdoMetDC co-expressed with DnaK+DnaJ; PfAdoMetDC co-expressed with KPf+DnaJ; PfAdoMetDC co-expressed with PfHsp70+DnaJ. Western blot analysis using anti-strep tag antibodies. **(B)** Lane S, soluble fraction; Lane P, insoluble proteins; PfAdoMetDC co-expressed with KPf+DnaJ+GroEL; PfAdoMetDC co-expressed with PfHsp70+DnaJ+GroEL was compared to DnaK+DnaJ+GroEL; Lane +C, represents a positive control (purified PfAdoMetDC). Western blot analyses were conducted using anti-strep tag antibodies and anti-Hsp60 antibodies.

To further confirm the solubility status of the co-expressed PfAdoMetDC protein with the various Hsp70-GroEL chaperone combinations, the soluble and pellet fractions were subjected to SDS-PAGE analysis. Figure 3.15 B, shows that more PfAdoMetDC occurred in the soluble fraction compared to the pellet fraction, suggesting that the supplementation of Hsp70-GroEL chaperones was effective in solubiliz PfAdoMetDC soluble. This therefore allowed the protein to be purified under native conditions. Since both GroEL and PfAdoMetDC are 60 kDa species, a Western blot analysis was performed using antibodies directed to Strep tag II to detect the strep-tagged PfAdoMetDC protein (Figure 3.15 B, low panel). GroEL was detected using anti-Hsp60 antibodies (Figure 3.17 B, low panel). The molecular chaperones appeared in lighter levels in the supernatant than in the pellet fraction (Figure 3.15, 70 and 60 kDa).

3.5.6 Purification of PfAdoMetDC protein co-expressed with chaperone combinations

The protein was purified using the strep-tactin column system as previously described (Birkholtz, 2004; Williams *et al*., 2011). The PfAdoMetDC preparations were analysed on SDS-PAGE to compare their purity (Figure 3.16). It was observed that the PfAdoMetDC co-expressed with molecular chaperones of plasmodial origin displayed a better quality of the target protein compared to the protein that was expressed in the absence of supplementary molecular chaperones (Figure 3.16, lane 1, lane 3 and 4). However, the protein that was co-expressed with DnaK+DnaJ supplementation co-purified with DnaK, similar to the protein that was produced in cells that were not supplemented with the chaperones (Figure 3.16 lane 1 and 2). On the Western blot analysis using anti-DnaK antibodies, the results confirmed the presence of DnaK which co-purified with PfAdoMetDC (Figure 3.16, low panel), suggesting that PfAdoMetDC was not folded properly. On the other hand, PfAdoMetDC co-expressed in cells supplemented with the

PfHsp70+DnaJ and KPf+DnaJ system did not co-purify with DnaK, since anti-DnaK antibodies did not detect DnaK (Figure 3.16 A, lane 3 and 4). This suggests that both KPf and PfHsp70 might have bound to PfAdoMetDC protein until it folded properly, thereby inhibiting endogenous DnaK binding.

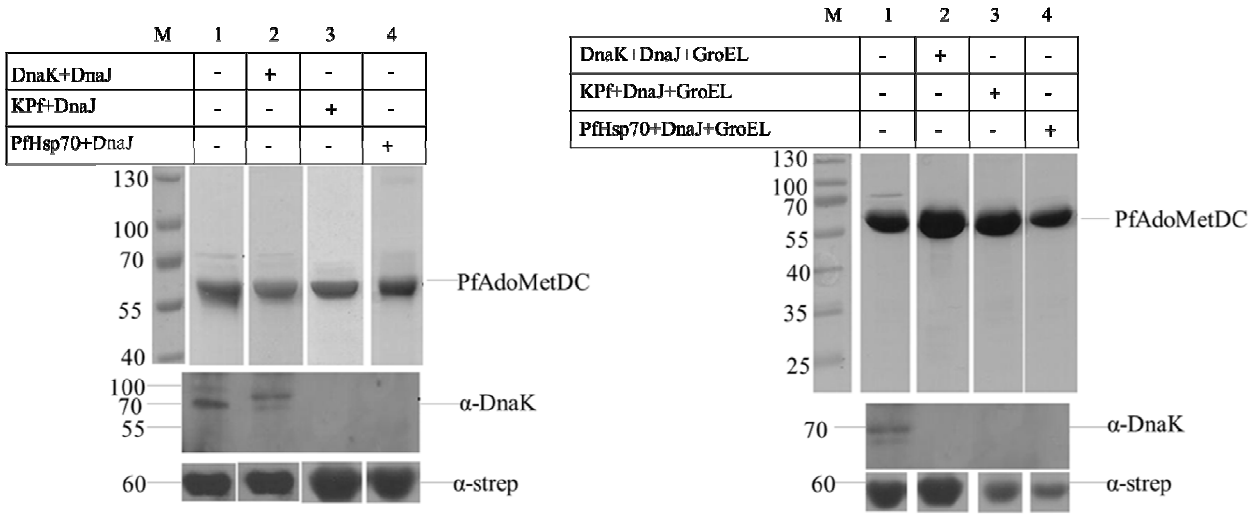

Figure 3.16: Purification of PfAdoMetDC protein co-expressed with Hsp70-Hsp40 and Hsp70-Hsp60 in *E.coli* BL21 (DE3) Star™ cells. (A) SDS-PAGE and Western blot analysis of co-expressed PfAdoMetDC with Hsp70-Hsp40 chaperones. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane 1, eluted PfAdoMetDC protein; Lane 2, PfAdoMetDC co-expressed with DnaK+DnaJ; Lane 3, eluted PfAdoMetDC protein co-expressed with KPf+DnaJ; Lane 4, eluted PfAdoMetDC protein co-expressed with PfHsp70+DnaJ. **(B)** SDS-PAGE and Western blot analysis of co-expressed PfAdoMetDC with Hsp70-Hsp60 chaperones. Western blot analysis conducted using anti-DnaK and anti-strep tag antibodies respectively. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane 1, eluted protein expressed on its own in BL21 cells; Lane 2, protein co-expressed with DnaK+DnaJ+GroEL; Lane 3, eluted protein co-expressed with pBB 542 KPf+DnaJ+GroEL; Lane 4, eluted protein co-expressed with PfHsp70+DnaJ+GroEL.

The following combinations of molecular chaperones were co-expressed with PfAdoMetDC in *E. coli:* DnaK+DnaJ+GroEL; KPf+DnaJ+GroEL; PfHsp70+DnaJ+GroEL, which reduced the contaminant proteins that co-purified with PfAdoMetDC (Figure 3.16 lane 1 and lanes 2, 3, 4). Anti-DnaK antibodies were utilized to confirm whether DnaK co-purified with PfAdoMetDC co-expressed with each chaperone combination system (Figure 3.16, low panel). The results show that anti-DnaK antibodies only detected DnaK associated with PfAdoMetDC that expressed cells, in this instance, not supplemented with chaperones. The antibodies must have detected endogenous DnaK. Furthermore, anti-strep tag antibodies were then used to confirm PfAdoMetDC co-expressed with DnaK+DnaJ+GroEL; KPf+DnaJ+GroEL; and PfHsp70+DnaJ+GroEL system. This shows that the system managed to assist the protein to complete its folding, and endogenous DnaK could not bind to the protein. PfAdoMetDC protein was expressed on *E. coli* BB1553 cells *(*Δ*dnaK* strain). The latter *E. coli* BB1553 cells lack endogenous DnaK to the PfAdoMetDC protein expression from *E. coli* BL21 (DE3) star^{TM} cells which overproduce endogenous DnaK (Figure 3.17). Due to the endogenous expression of DnaK that co-purified with PfAdoMetDC, it was speculated that expression of the protein in the absence of DnaK may address the issue. *E. coli* BB1553 was able to produce recombinant PfAdoMetDC protein (Figure 3.17, lane "Δ*dnaK"*). This finding suggests that DnaK was not required for the production of PfAdoMetDC in *E. coli* BB1553 cells. However, it was necessary to determine whether the PfAdoMetDC produced in *E. coli* BB1553 cells was functional (see section 4.2.2; Figure 4.4).

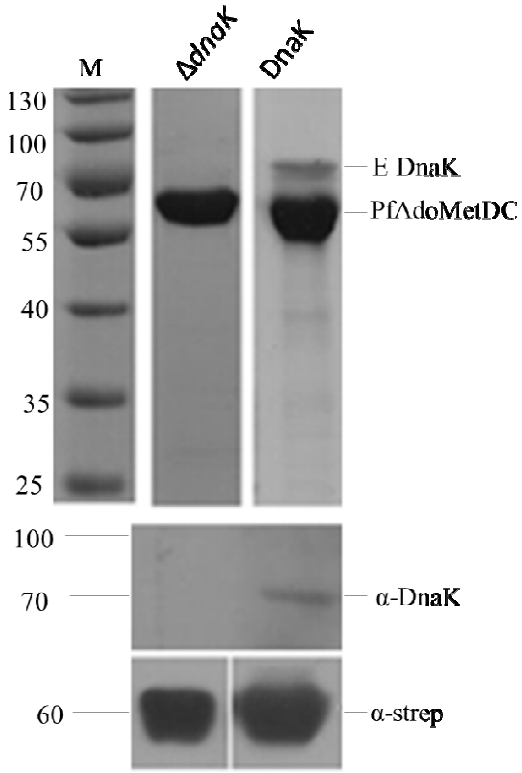

Figure 3.17: SDS-PAGE and Western blot analysis for the purification of PfAdoMetDC expressed in *E. coli* Δ*dnaK* strain and *E. coli* BL21 (DE3) star™ cells. Western blot analysis of PfAdoMetDC using anti-strep tag antibodies. Lane M, page ruler pre-stained ladder (sizes indicated on the left hand side); Lane Δ*dnaK*, PfAdoMetDC expressed in strain that does not produce endogenous DnaK; Lane DnaK, PfAdoMetDC expressed in *E. coli* BL21 star™ cells that produce endogenous DnaK.

3.5.7 Purification of His-tagged Hsp70, GroEL and DnaJ

Recombinant DnaK, PfHsp70, KPf, DnaJ and GroEL proteins were expressed in XL1 Blue cells as His-tagged species of 70 kDa (DnaK, PfHsp70,and KPf), 60kDa (GroEL) and 40 kDa (DnaJ) (Figure 3.18). This was done in preparation to investigate whether molecular chaperones like KPf and PfHsp70 could co-operate with DnaJ in suppressing the heat-induced MDH aggregation *in vitro*. All five proteins were successfully purified by nickel affinity chromatography under the native system, as confirmed in SDS-PAGE analysis. Since all these proteins were His-tagged, Western blot analyses were conducted using anti-His antibodies (Figure 3.18).

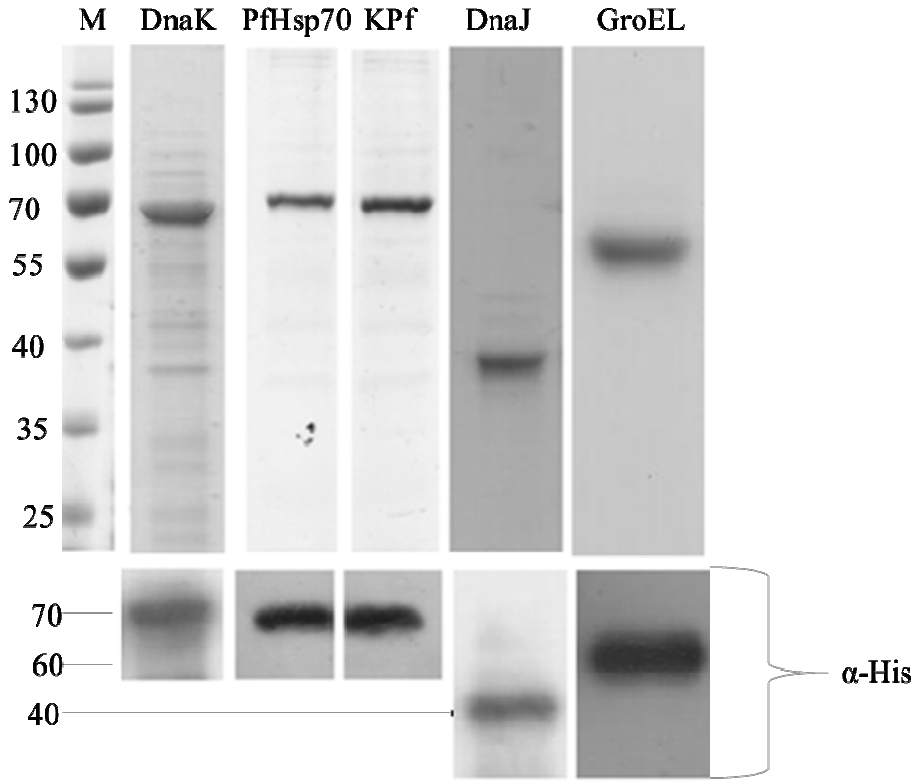

Figure 3.18: Purification of His-tagged Hsp70s, DnaJ and GroEL by nickel affinity chromatography. SDS-PAGE analysis confirming the successful native purification of DnaK, PfHsp70, KPf, DnaJ and GroEL overproduced in *E. coli* XL1 Blue cells. Anti-His antibodies used in Western blot analysis to detect Hsp70s, DnaJ and GroEL proteins respectively. Lane M, page ruler pre-stained ladder is shown in the extreme left hand side lane, with sizes indicated on the left hand side.

3.6 DISCUSSION

The co-operation of chaperones of *E. coli* origin in partnership with newly synthesized proteins has been demonstrated (de Marco *et al*., 2007; Alibolandi *et al*., 2010). In the current study, the effect of combined *E. coli* chaperones (DnaJ+GroEL) with *P. falciparum* chaperones (KPf and PfHsp70) in improving the quality of PfAdoMetDC in *E. coli* was investigated. The quality of PfAdMetDC final product is generally accompanied by contaminating proteins of 70 kDa (Williams *et al*., 2011). Structural protein studies require pure protein, therefore combined molecular chaperones were used as co-expression partners to improve the quality of the PfAdoMetDC protein expressed in *E. coli*.

PfHsp70 has previously been shown to protect *E. coli* and yeast cells against stress and may aid folding of other proteins during adverse cellular conditions (Shonhai *et al*., 2005; Bell *et al*., 2011). In addition, PfHsp70 has been used to improve the solubility and yield of recombinant malarial protein PfGCHI, produced in *E. coli* cells (Stephens *et al*., 2011). This current study demonstrated that PfHsp70 improved the quality of recombinant PfAdoMetDC produced in *E. coli* BL21 (DE3) star^{TM} cells, providing further evidence that this chaperone could be used to facilitate the production of a broad spectrum of malarial proteins in *E. coli*. As expected, KPf also improved the quality of PfAdoMetDC protein. This was due to the fact that KPf has a substrate binding domain from PfHsp70, which may have had an immense contribution towards recognition of the substrate, and thus allowing it to bind to it for folding. In addition, KPf could be recognized by the *E. coli* BL21 system since it has the ATPase domain from *E. coli* DnaK. As mentioned in the literature, DnaJ interacts with the ATPase domain of Hsp70 and stimulates its

basal ATPase activity. It is therefore possible that DnaJ may cooperate with KPf because of its ATPase domain (Wittung-Stefshede *et al*., 2003).

It was previously reported that purified recombinant PfAdoMetDC associates with *E. coli* DnaK (Williams *et al*., 2011), suggesting that the former thereby exposes hydrophobic patches to which DnaK is bound. In addition, the rate of translation in bacteria is much higher (approximately 20 amino acids per second) than in eukaryotes (approximately 4 amino acids per second). The slower translation rates in eukaryotes are consistent with the production of multi-domain proteins, which require more time for folding (Vabulas *et al*., 2010). The rate of PfAdoMetDC synthesis in *E. coli* may have taken place too fast, giving the protein too little time to fold. This will have led to the generation of a product that was not fully folded and bound by DnaK for an extended residence time. The residence time for DnaK on peptides varies from 30 s to 25 minutes; only proteins that interact only with DnaK upon synthesis are released fast (Sekhar *et al*., 2012). On the other hand, proteins that exhibit low cellular abundance, a higher number of DnaK-binding sites, and those that tend to assume dynamic structural intermediates (slow folding proteins which do not easily bury their hydrophobic patches) exhibit higher DnaK residence time (Calloni *et al*., 2012). Typically, such proteins require DnaK for their maintenance (Calloni *et al*., 2012). Consequently, the extended binding of DnaK to peptides may slow their folding, resulting in detrimental consequences. For this reason, it has been proposed that the expression of recombinant proteins in DnaK minus strains of *E. coli* be used to circumvent DnaK contamination (Ratelade *et al*., 2009).

In the current study, the expression of PfAdoMetDC in the presence of Hsp70s of plasmodial origin prevented the association of the former with DnaK, suggesting that the protein's folded

conformation was improved. This was confirmed by the comparatively higher activities registered by PfAdoMetDC, which were produced in the presence of PfHsp70 and KPf, compared to the protein that was co-expressed with supplementary heterologously expressed DnaK. This is in line with a previous study which demonstrated that co-expression of PfGCHI with PfHsp70 improved the solubility and activity of the former (Stephens *et al*., 2011). The fact that KPf possesses the DnaK ATPase domain may facilitate its recognition by *E. coli* co-chaperones. At the same time, the *P. falciparum* substrate binding domain is likely to recognize PfAdoMetDC as a substrate. Both PfHsp70-1 and KPf are *plasmodial* in origin, and are thus expected to recognize the malarial proteins co-expressed in *E. coli*. The role of PfHsp70 and KPf in protein folding could be similar to DnaK. It was reported that both proteins showed chaperone activity in *E. coli* cells possessing a functionally compromised DnaK. Chaperone activity was shown by protecting proteins from misfolding, thus maintaining cell viability (Shonhai *et al*., 2005). De Marco and coworkers (2007) have demonstrated that DnaK+DnaJ+GroEL chaperones are cooperative in assisting newly synthesized protein to fold. Therefore, both KPf and PfHsp70 chaperones could cooperate with GroEL in some way with DnaK, thereby assisting nonnative polypeptides to complete their folding.

The solubility of PfAdoMetDC was affected by supplemented chaperones (Figure 3.15). However, the protein was not clear when analyzed on SDS-PAGE (Figure 3.14), as compared to the protein co-expressed in the presence of *E. coli* chaperone combination. Part of the reason the protein was not properly separated in the gel, or as observed, could be due to the fact that both GroEL and PfAdoMetDC share the same sizes (60 kDa). In addition, the issue of the degradation products of Hsp70s may have affected the protein expression. The Western blot analysis using anti-strep tag antibodies provided enough evidence that one of the 60 kDa expressions was

PfAdoMetDC. Anti-GroEL antibodies detected the presence of GroEL protein as well (Figure 3.17). However, the PfAdoMetDC expressed in cells supplemented with Hsp70-GroEL, was easy to free of contaminants suggesting that it had better fold, and DnaK therefore did not co-purify with it (Figure 3.16). The finding of this study shows that KPf+DnaJ, PfHsp70+DnaJ, KPf+DnaJ+GroEL and PfHsp70+DnaJ+GroEL chaperone sets played an important role in assisting PfAdoMetDC to fold up to the point where the endogenous Hsp70 chaperone that seemed to bind longer and co-purify with the final product was reduced. The role of Hsp40 is to recognizes a substrate and hand it over to Hsp70 for binding. On the other hand, the GrpE/nucleotide exchange factor mediates the binding and release of the substrate by binding to Hsp70, thereby affecting nucleotide exchange. This suggests that this cooperation displayed by these molecular chaperones could have taken place in order to assist PfAdoMetDC to complete its folding. PfAdoMetDC purified from *E. coli* BB1553 cells (deficient in DnaK function) displayed high quality protein (Figure 3.17). However, the protein did not fold properly due to fact that PfAdoMetDC, which was not supplemented with molecular chaperones, was completely digested by the proteinase K. This suggests that molecular chaperones play a key role in assisting non-native proteins to complete their folding.

CHAPTER FOUR

In vitro biochemical characterization of PfAdoMetDC co-expressed with molecular chaperones

4.1 Introduction

Different techniques have been developed to probe the conformation of protein in solution. These include circular dichroism (CD), which is the most commonly used technique for monitoring the structure of protein in solution (Sharon *et al*., 2005). This technique allows for evaluation of the overall structural features of the secondary structure of protein and also to quantify relative proportions of alpha helix and beta sheets and random coils. NMR spectroscopy is also a useful and informative technique for structural elucidation of small protein in solution (O'Cannell *et al*., 2009). However, this technique is plagued by instrumentation requirements and needs a millimolar concentration of a non-aggregating protein solution. Moreover, limited success has been had in providing the information on the partly folded and fluctuating states of protein due to resonance broadening and lack of sufficient chemical shift dispersion. X-ray crystallography is one of the techniques that can be used for protein structure analysis, provided that suitable protein crystals are available, although it is of no use with a dynamic protein system (Elspeth *et al*., 2014). It appears that none of these techniques is superior to the others; each technique has its advantages and drawbacks.

Limited proteolysis is a simple technique used to probe the molecular features of polypeptides in their native or partly folded states (Figure 4.1) (Fontana *et al*., 2004). It is preferable to use this biochemical technique at those loops which display inherent conformational flexibility, as the protein core remains quite rigid and thus resistant to proteolysis. A survey of cleavage sites in a variety of proteins of known three dimensional structure revealed that they never occur at the level of alpha helices, but occur largely on the loops (Fontana *et al*., 2004). Therefore, limited proteolysis experiments can be used to identify the sites of enhanced flexibility or local

unfolding of a polypeptide chain. There is evidence that suggests that this technique can produce results that are in agreement with the NMR technique and other computational approaches by detecting the unfolding proteins (Fontana *et al*., 2004).

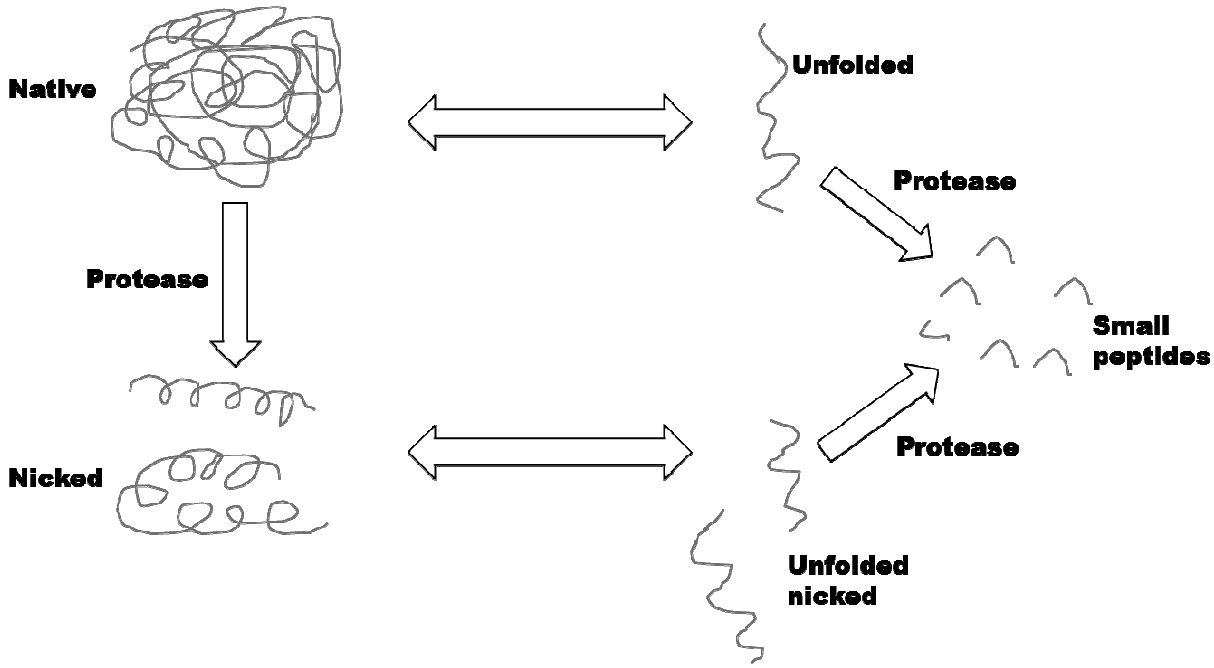

Figure 4.1: A schematic mechanism of proteolysis of a globular protein. A dual representation mechanism of protein degradation is shown; the one involving a substrate for proteolysis and the other, an unfolded and native form of the protein. In the latter case proteolysis is limited and occurs at flexible site(s), leading to a nicked protein species that can unfold and then be degraded down into small polypeptides (Adapted from Fontana *et al*., 2004).

The production of protein in a non-native host has provided the means for studying the protein from prokaryotes and eukaryotes, and for exploiting it commercially in the production of medicines and industrial biocatalysts (Terpe *et al*., 2006; Littlechild *et al*., 2007). Cells synthesize less protein than they produce. However, in order for the cells to make use of the protein, a larger quantity of protein is needed than the cells produce. This problem is solved by

overexpression of the protein of interest in a suitable host (Widersten, 1998; Villaverde and Carrio, 2003).

Overproduction of recombinant protein in *E. coli* is often accompanied by misfolding and segregation into insoluble aggregates known as inclusion bodies (Carrio and Villaverde, 2001). This problem is magnified when overexpression of the recombinant protein is carried out (de Marco *et al*., 2007). Although the formation of inclusion bodies could be one of the methods used for protein purification, however, it is not guaranteed that the large yield of protein would be refolded in a biologically active form (de Marco *et al*., 2005). Protein recovered using this method are often obtained in low concentrations compared to time-intensive procedures, and protein activity is not guaranteed (de Marco *et al*., 2005).

Fermentation engineering, which involves reducing cultivation temperature, can be used to reduce protein aggregation (Georgiou and Valax, 1996). *In vivo* studies have shown that protein folding is assisted by molecular chaperones. Chaperones promote the proper isomerization and cellular targeting of other polypeptides by transiently interacting with folding intermediates and foldases (Baneyx, 1999). This accelerates the rate-limiting pathways involved in folding steps and provides a tool to combat the problem of inclusion body formation (de Marco and de Marco, 2004). In addition, some proteins co-purifiy with contaminant proteins, which compromises the quality of the final product (Williams *et al*., 2012).

Overexpression of native *E. coli* chaperones has been shown to improve the protein expression of several heterologous proteins, though with with varying success; production of ScFv antibody fragment with DnaK/DnaJ/GrpE increased 100 fold; Interleukin-1 solubility was increased only with increased overexpression of DnaK/DnaJ (Kolaj *et al*., 2009). The most studied molecular

chaperones in the cytoplasm of *E. coli* are ATP-dependent DnaK-DnaJ-GrpE and GroEL-GroES (de Marco *et al*., 2007), indicating that they cooperate in assisting newly synthesized protein for folding. Some recombinant proteins, such as human thromboxane, human lysosome or *Agrobacterium tumefaciens* β-glucosidase, are often found in inclusion bodies. When these proteins were co-expressed with combinatory chaperones, GroEL/GroES, they were recovered as soluble proteins (Nishihara *et al*., 2000). The solubility of aconitase protein was improved by 40% upon co-expression with GroEL/GroES, whilst its activity was only increased by 1.5 fold (Kolaj *et al*., 2009). Despite widespread success, this system does not always provide a solution to protein insolubility; some proteins fail to be refolded by GroEL/GroES and some have poor enzyme activity, which contributes to an overall decrease in cell viability (Kolaj *et al*., 2009).

There is evidence that co-overproduction of DnaK+DnaJ+GroEL could prevent the formation of aggregation, and improve production of target proteins (de Marco *et al*., 2007). To date, no study has been conducted showing that co-overproduction of combined molecular chaperones from different sources such as *E. coli* and *P.falciparum*:PfHsp70+DnaJ-GroEL and KPf+DnaJ+GroEL could improve protein quality, solubility and activity. Therefore, the aim of the study at hand was to investigate the effects of these chaperone tools on co-expressed PfAdoMetDC protein in *E. coli*.

The aims of this study were:

1) To conduct limited proteolysis on PfAdoMetDC protein co-expressed with various chaperone combinations;
2) To analyze the activity of PfAdoMetDC protein co-expressed with various Hsp70s and combined with DnaJ chaperone ;

3) To explore the *in vitro* chaperone activity of Hsp70s, GroEL and DnaJ, in isolation and in combination.

4.2 Experimental Procedures

4.2.1 Partial Proteolysis

Limited proteolysis was employed to assess the conformation/structure of PfAdoMetDC after its expression. The principle of this method is that the protein is incubated with a relatively low amount of protease which cuts at exposed regions, such as loops and other flexible regions of the protein. Therefore 4.2 mg/ml of protein was digested with 1 mg/ml proteinase K in a buffer of 10 mM Tris-HCl, pH 7.7; 10 mM magnesium acetate; 60 mM NH_4Cl and 30 mM β-mercaptoethanol at 37^{o}C for 45 minutes. Aliquotes were withdrawn from the mixture at specific time intervals for electrophoretic analysis. The samples were analysed on SDS-PAGE and Western blot to identify the digested products. The appearance of different bands from untreated protein denotes the digestion of the protein, treated protein was visualised on commasie stain.

4.2.2 PfAdoMetDC activity assays

The monofunctional PfAdoMetDC was assayed for enzyme activity directly after purification in order to check the effects of the chaperone sets on protein quality. The assays contained 5 ug enzyme, 100 uM S-adenosy-L-methionine chloride (Sigma-Aldrich) and 50 nCi S-[Carboxyl-14 C] adenosyl-L-methionine (55 mCi/mmol, Amersham Biosciences) in a total volume of 250 µl assay buffer (50 mM KH2PO4 pH 7.5, 1 mM EDTA, 1 mM DTT), and were perfomed as previously described (Bikhroltz, 2004; Williams *et al*., 2011).The activity assays were perfomed in duplicate for three assays and specific enzyme activities were expressed as the amount of CO_2 produced in nmol/min/mg.

4.2.3 Malate Dehydrogenase Aggregation Suppression Assays

Aggregation suppression assays were used to assess the ability of the chaperone proteins (recombinant PfHsp70, KPf, DnaJ and DnaK) and GroEL partnerships to suppress the heat-induced aggregation of L-Malate Dehydrogenase (MDH) at 48ºC over time. The MDH (0,72 µM) and chaperone protein(s) of interest (0,36 µM) were suspended in assay buffer (100 mM NaCl, 50 mM Tris, pH 7.4). The accumulation of protein aggregation was monitered by changes in absorbance at 360 nm in a thermo-controlled spectrophotometer. As a reference, MDH aggregation was monitored in the absence of chaperone proteins at 48ºC over time. Similarly, in the absence of MDH, the thermal stability of the chaperone proteins under investigation was assesed to ensure accurate monitoring of the MDH aggregation in the experiment. Furthermore, BSA (0,72 µM) was used in place of chaperone proteins to ensure that the observed MDH aggregation suppression was specific to chaperone activity, and to further demonstrate the thermal instability of MDH in the absence of chapeorone intervention.

4.3 Results

4.3.1 Limited proteolysis of PfAdoMetDC using proteinase K

Limited proteolysis was used in order to assess the conformation of PfAdoMetDC co-expressed with the following combinantion of chaperones: DnaK+DnaJ, KPf+DnaJ and PfHsp70+DnaJ. PfAdoMetDC produced in *E. coli* BL21 cells without supplementary chaperones, but with express endogenous DnaK, was completely degraded by Proteinase K within 5 minutes. This generated small fragments that could not be detected by either the SDS-PAGE or Western blot analyses (Figure 4.1). However, PfAdoMetDC expressed in the presence of supplementary DnaK and KPf chaperones was fairly stable and withstood the action of proteinase K for 30 minutes (Figure 4.1). Although the co-expression of PfAdoMetDC with PfHsp70 improved the stability of the former, digestion fragments were generated at a higher rate than was observed for supplementation with DnaK and KPf (Figure 4.1). A close analysis of the products of Proteinase K digestion shows that, after 5 minutes, PfAdoMetDC supplemented with heterologous DnaK was the most stable. Interestingly, the proteolysis products of PfAdoMetDC co-expressed with different chaperone sets exhibited varied profiles (Figure 4.1). This indicates that the protein produced in each case had unique conformations. In addition, different sized products of digestion were detected. PfAdoMetDC contains a Strep-tag at its C-terminus; the α-Strep antibody can recognize the products of digestion that have the Strep-tag (Birkholtz,2004). Therefore, those products that were not detected by antibodies may have had the Strep-tag cleaved off.

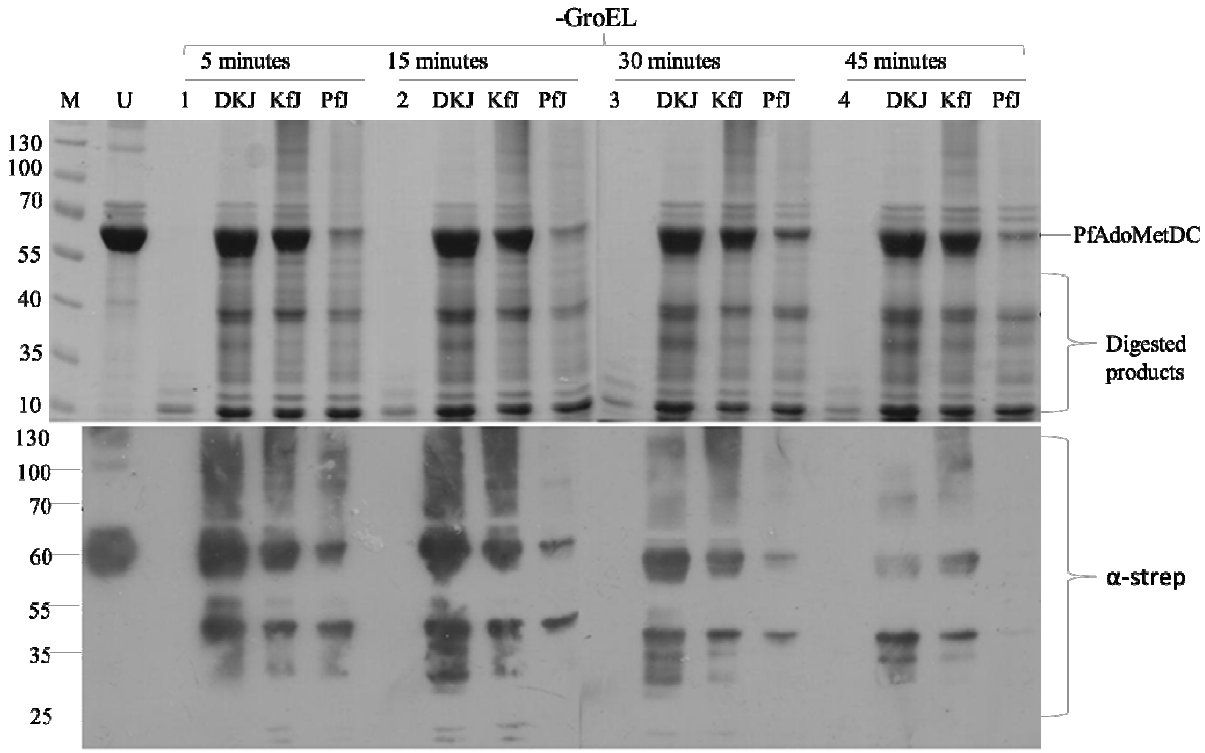

Figure 4.2: PfAdoMetDC attains a unique conformation depending on the Hsp70 co-expression partner. The limited proteolysis by the PK enzyme to substrate ratio; E: S=1:500. Lane M, page ruler pre-stained protein ladder (sizes indicated on the left hand side); Lane U, untreated protein; Lanes 1,2,3, and 4, digested protein that was expressed on its own; Lane DKJ, digested protein co-expressed with DanK+DnaJ; Lane KfJ, digested protein co-expressed with KPf+DnaJ; Lane PfJ, digested protein co-expressed with PfHsp70+DnaJ. Western blot analysis conducted using anti-strep tag antibodies.

PfAdoMetDC was further co-expressed with the GroEL chaperone system: DnaKa+DnaJ+GroEL; KPf+DnaJ+GroEL and PfHsp70+DnaJ+GroEL. Figure 4.2, shows the SDS-PAGE analysis of the protein that was exposed to proteinase K over time (minutes). The results show that the protein was properly folded, as it showed stability upon exposure to the proteinase K (Figure 4.2). Unfolded proteins are suitable candidates for the protease to bind to, as they expose the target sites (flexible loops) (Fontana *et al*., 2004).

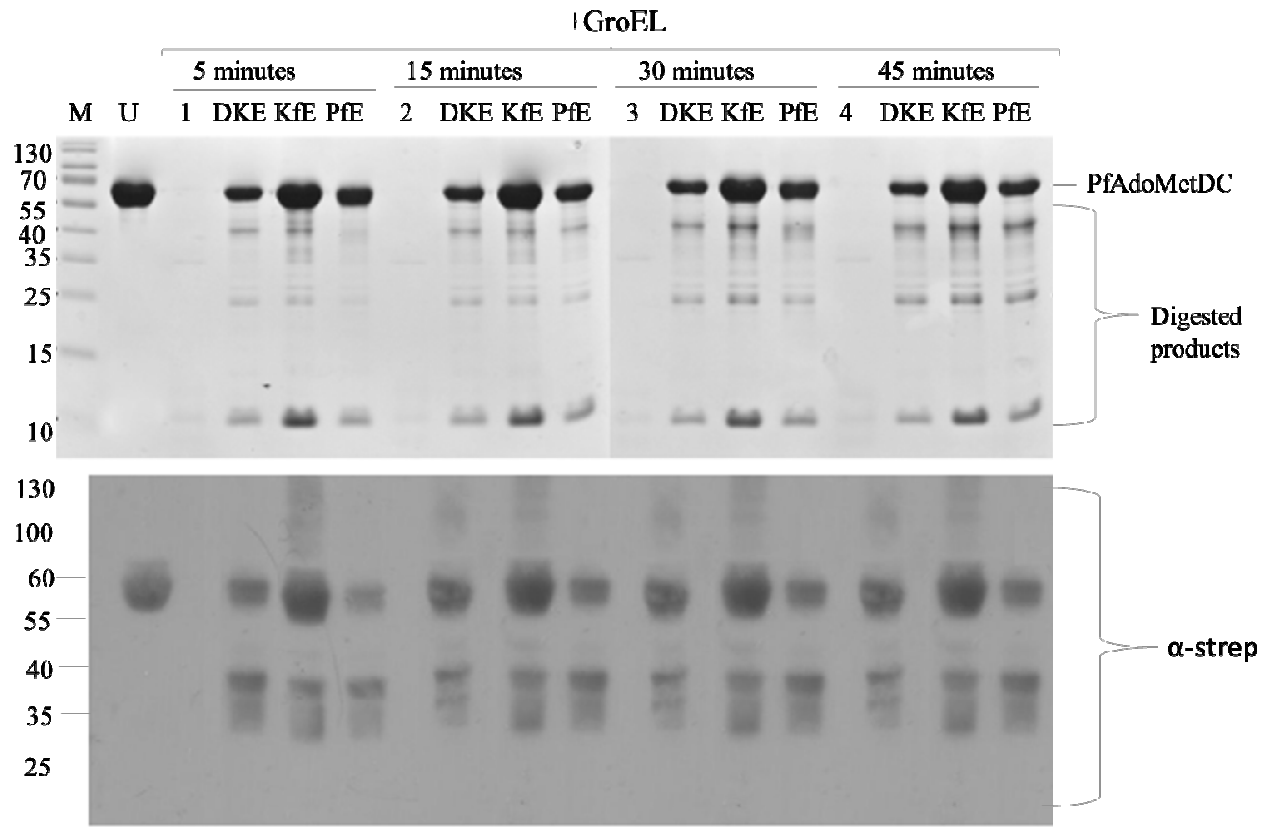

Figure 4.3: Limited proteolysis of co-expressed PfAdoMetDC with supplementation of Hsp70+GroEL chaperone tools. Time course of co-expressed PfAdoMetDC proteolysis monitored by SDS-PAGE and confirmed by the Western blot analysis**.** Limited proteolysis by the PK enzyme to substrate ratio; E: S=1:500. Lane M, page ruler pre-stained protein ladder (sizes indicated on the left hand side); Lane U, untreated protein; Lanes 1,2,3 and 4, digested protein that was expressed on its own; Lane DKE, PfAdoMetDC co-expressed with DanK+DnaJ+GroEL; Lane KfE, PfAdoMetDC co-expressed KPf+DnaJ+GroEL; Lane PfE, PfAdoMetDC co-expressed with PfHsp70+DnaJ+GroEL. Lower pane; Western blot analysis conducted using anti-strep tag antibodies.

PfAdoMetDC expressed in *E.coli* Δ*dnaK* strain was exposed to proteinase K to probe its conformation. Even though the protein was purified without any contaminant proteins, particularly DnaK, upon its exposure to protease it was completely digested compared to proteins that were co-expressed with the supplementation of molecular chaperones. Figure 4.3 shows the SDS-PAGE analysis of the digested protein and Western blot using anti-strep tag antibodies. The anti-bodies could not detect the breakdown proteins, suggesting that the C-terminal strep tag was degraded by proteinase.

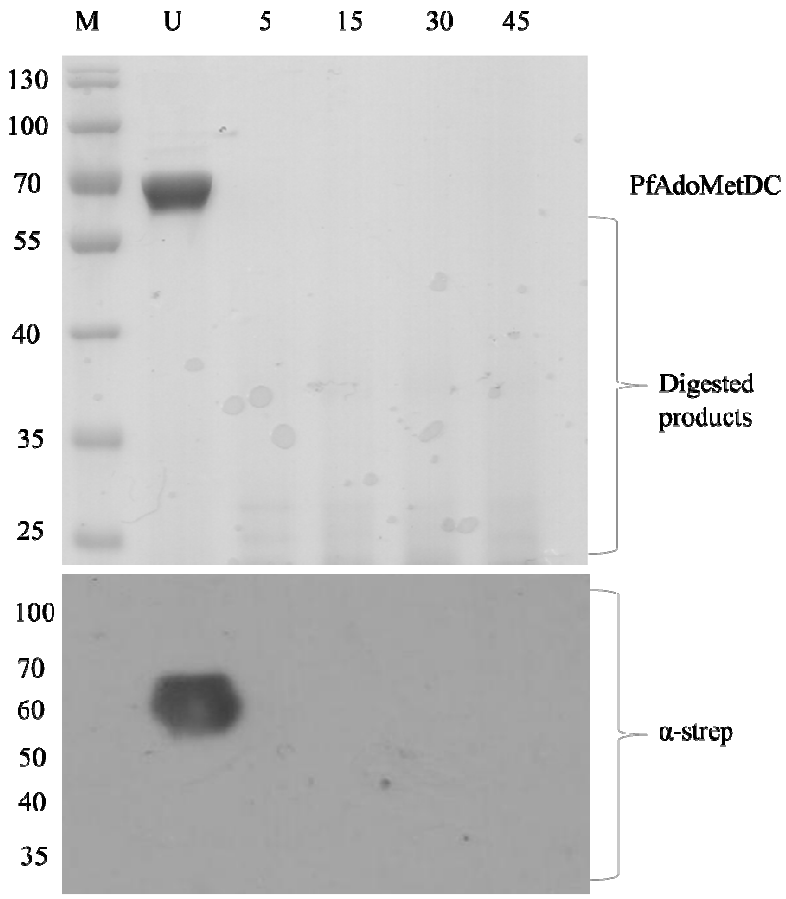

Figure 4.4: Time course limited proteolysis of PfAdoMetDC protein expressed in Δ*dnaK* strain cells. The limited proteolysis by the PK enzyme to substrate ratio; E: S=1:500. Lane M, page ruler pre-stained protein ladder; Lane U, undigested PfAdoMetDC; Lane 5-45, sample withdrawn from the reaction over time (minutes). Digested products were detected using anti-strep tag antibodies.

4.3.2 Different folding conditions influence the activity of recombinant PfAdoMetDC

The activity of PfAdoMetDC purified from *E. coli* BL21 (DE3) star^{TM} cells and un-supplemented with chaperones was presented as 100% (Figure 4.4). The PfAdoMetDC activity co-expressed with DnaK+DnaJ was less active, with PfAdoMetDC activity of a 1.14 fold increase, compared to KPf+DnaJ (with a 3.24 fold increase) and PfHsp70+DnaJ (with a 2.77 fold increase) (Figure 4.4). Unexpectedly, the activity of PfAdoMetDC produced the *E. coli* dnaK minus strain and exhibited the highest activity, with a 3.85 fold increase.

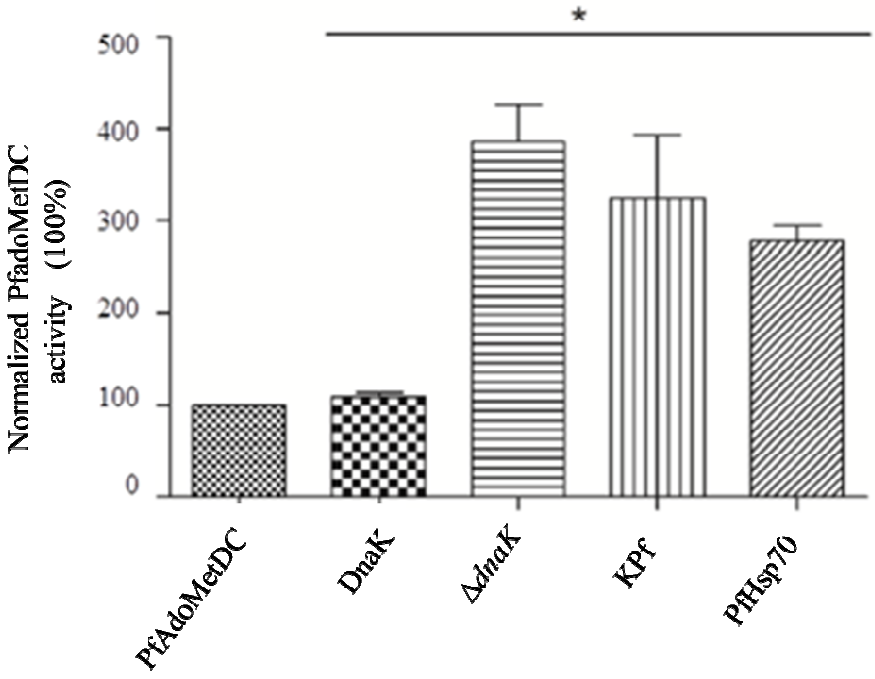

Figure 4.5: Co-expression with malarial Hsp70 derivative (KPf) and PfHsp70 improves the activity of PfAdoMetDC. The protein activity of PfAdoMetDC was compared to the co-expressed protein with molecular chaperones DnaK, KPf, PfHsp70, and PfAdoMetDC protein, expressed in cells that do not produce endogenous DnaK (Δ*dnaK*) and given as a percentage.

4.3.3 Protein aggregation suppression assays

Protein aggregation is achieved by the holdase and refoldase activity of molecular chaperones. The current study explored the potential of *E. coli* co-chaperones (DnaJ, GroEL) to enhance the aggregation suppression activity of different Hsp70s. *E. coli* DnaK and KPf share the same ATPase domain. The ATPase domain of Hsp70 facilitates its interaction with Hsp40 co-chaperones; KPf may thus interact with DnaJ thereby facilitating its chaperone function in the folding of PfAdoMetDC. Recombinant DnaK, PfHsp70, KPf, DnaJ and GroEL were successfully purified by nickel chromatography, separated by SDS-PAGE, and identified by Western blot analysis as 70 kDa, 60 kDa and 40 kDa proteins of high purity using α-His

antibody (see section 3.5.7; Figure 3.18). DnaK (*E. coli* Hsp70) and PfHsp70 are capable of suppressing heat-induced protein aggregation *in vitro* (Shonhai *et al*., 2008). In the current study, it was investigated whether KPf could suppress aggregation of MDH. The assessment of the Hsp70 chaperones actvity was further investigated in the presence of, and in the absence of, the co-chaperones *E. coli* DnaJ and GroEL. The maximum absorbance representing the aggregation of MDH was presented as an arbitrary value of 100% (Figure 4.5). In the presence of BSA, the aggregation of MDH was only slightly reversed (data not shown). All the molecular chaperones (DnaJ, DnaK, KPf, PfHsp70 and GroEL) suppressed the aggregation of MDH (Figure 4.5). In the presence of DnaJ, both PfHsp70 and KPf exhibited improved activity, comparable to that of the DnaK-DnaJ combination (Figure 4.5). The addition of GroEL to the Hsp70+Hsp40 combinations did not significantly influence suppression of aggregation activity (Figure 4.5).

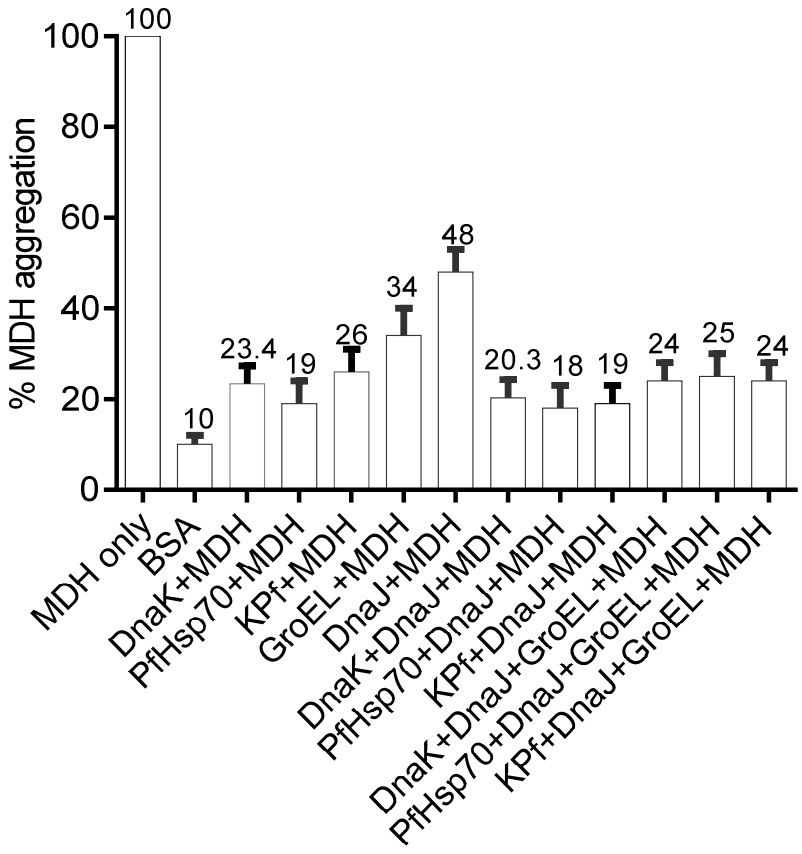

Figure 4.6s: The percentage of the MDH aggregation in the presence of the recombinant purified chaperone proteins. The heat-induced aggregation of 0.72 µM MDH was monitored at 48^{O}C at an absorbance of 360 nm. The degree of aggregation was monitored in the presence of, and in the absence of, the molecular chaperones indicated. The aggregation of 0.72 µM MDH in the absence of chaperone proteins was 100%. The aggregation of 0.72 µM BSA was also monitored, on its own and in the presence of MDH as control is 88% (data not shown). The aggregation of MDH in the presence of DnaJ, DnaK, GroEL, PfHsp70, KPf alone (0.36 µM), and the Hsp70s+Hsp40; Hsp70s+GroEL+DnaJ combinations were assessed and the aggregation was determined relative to the MDH control and BSA as a control. This experiment was repeated at least three times, with the standard deviations shown as *error bars*.

4.4 Discussion

It has been proposed that the supplementary expression of molecular chaperones could improve the yield and quality of recombinant malarial proteins produced in *E. coli* (Birkholtz *et al*., 2008; Stephens *et al*., 2011). Findings from this study show that PfAdoMetDC produced in *E. coli* BL21 star^{TM} cells in the presence of PfHsp70 and its derivative, KPf, was easier to purify, compared to the protein that was produced in the presence of only heterologous *E. coli* DnaK. In addition, findings from this study suggest that supplementation of heterologous *E. coli* DnaK to complement endogenous DnaK did not result in the improved purity of recombinant PfAdoMetDC.

PfAdoMetDC expressed in the presence of various chaperone conditions in *E. coli* and subsequently subjected to limited proteolysis exhibited unique proteolytic fragments, suggesting that the protein had various conformational states. The protein that was produced in the presence of endogenous DnaK (without supplementary exogenous Hsp70) was most susceptible to proteolysis compared to the protein that was co-expressed with heterologously expressed Hsp70s (DnaK, KPf and PfHsp70) (Figure 4.1). However, results show that PfAdoMetDC supplemented with DnaK+DnaJ and KPf+DnaJ combinations indicated that the protein was more stable when exposed on Proteinase K. This could be due to the fact that, *E. coli* DnaK and KPf share the same ATPase domain. The ATPase domain of Hsp70 facilitates its interaction with Hsp40 co-chaperones. It is possible that DnaK and KPf may have been able to interact with *E. coli* DnaJ, thereby improving their chaperone function to produce more stable PfAdoMetDC protein. This proves that the presence of the Hsp70 chaperones influenced the folded status of the protein in unique ways.

Kerner and colleagues (2005), demonstrated that proteins that were supplemented with molecular chaperones, such as GroEL, were protected from the external proteinase K. However, those that were not supplemented with molecular chaperones were degraded due to the exposure of the flexible loops. This is in line with findings from the current study, because the PfAdoMetDC protein that was co-expressed with either Hsp70-DnaJ or Hsp70-GroEL molecular chaperone sets was not completely degraded when exposed to the proteinase K (Figure 4.1 and Figure 4.2). On the other hand, PfAdoMetDC protein that was expressed without the supplementation of molecular chaperones was degraded completely (Figure 4.3). These results suggest the importance of obtaining a properly folded protein, as it can tolerate harsh conditions. This also indicates the influence of molecular chaperones as folding mediators of newly synthesized proteins in preventing the protein from misfolding and maintaining its stability.

Supplementing the *E. coli* host cell with exogenous DnaK did not improve the activity and folding status of PfAdoMetDC. However, supplementing the *E. coli* cell with plasmodial chaperones (PfHsp70 and its derivative chimera KPf) improved the purification of PfAdoMetDC and, based on activity assays, the protein was of better quality than that produced in the absence of, and in the presence of, supplementary DnaK (Figure 4.4).

Surprisingly, the expression of PfAdoMetDC in *E. coli* DnaK deficient cells resulted in the production of a highly active protein (Figure 4.4). *E. coli* BB1553 cells that lack DnaK were capable of expressing recombinant PfAdoMetDC. Not only did this result in the reduction of contaminants associated with the purified PfAdoMetDC, but the PfAdoMetDC produced in these cells demonstrated higher activity than the protein produced in the presence of Hsp70 chaperones

(Figure 4.4). Rateleda and colleagues (2009) show that using an *E. coli* system that does not produce endogenous DnaK could improve the quality of recombinant proteins expressed in *E. coli.* However, despite the purity of the protein produced on *E. coli* BB1553 cells, the folding behavior of this protein was in sharp contrast to its purity and activity. This is in line with the question asked earlier about what the quality of the recombinant protein could be, or what it could mean (see section 1.4; Chapter one). PfAdoMetDC produced on the DnaK minus strain was very pure and active compared to PfAdoMetDC protein produced under supplementation of molecular chaperones. Its conformation, though, showed that the protein was not properly folded when exposed to proteolytic conditions (Figure 4.3), while proteins produced in the presence of supplementation of molecular chaperones demonstrated their stability, possibly suggesting these were of better fold (Figure 4.1 and Figure 4.2). Taken together, these results show that molecular chaperones have more influence on the folding of proteins in most cases. *E. coli* DnaK and KPf share the same ATPase domain. The ATPase domain of Hsp70 facilitates its interaction with Hsp40 co-chaperones. It is possible that DnaK and KPf may have been able to interact with *E. coli* DnaJ, thus improving its chaperone functional activity in producing a stable PfAdoMetDC protein. This led to the need for further confirmation of these findings by conducting *in vitro* biochemical assays to assess possible cooperation of the chaperones.

It is known that Hsp70 proteins promiscuously bind to hydrophobic patches of misfolded proteins (Walter and Buchner, 2002). However, it is also known that Hsp40 (*E. coli* DnaJ) serves as a substrate scanner for Hsp70 by binding to misfolded proteins first, and then passing them to Hsp70 to facilitate their refolding (Bukau *et al.*, 2006). Therefore, Hsp40 does not only stimulate the ATPase activity of Hsp70, but it also regulates functional specialization of Hsp70. It was proposed that KPf may interact with DnaJ through its ATPase domain. The possible interaction

between DnaJ and KPf may have facilitated the folding of the substrate in *E. coli*. However, Hsp70 also protects cells by binding to misfolded proteins, stabilizing them until they refold (Buchner and Walter, 2002). This holdase function of Hsp70 does not require co-operation with Hsp40 (Bukau *et al*., 2006). In the presence of DnaJ, GroEL, PfHsp70, KPf, and indeed DnaK, all demonstrated an improved capacity to suppress the thermally-induced aggregation of MDH *in vitro* (Figure 4.5). Furthermore, DnaJ itself was also capable of suppressing MDH aggregation. It is not clear if the activity of the Hsp70s in the presence of DnaJ was through co-operative chaperone-co-chaperone based function, or through independent functions between the particular Hsp70 and the DnaJ. However, this shows the crucial role played by the individual molecular chaperone, and in combination, during the folding stages of the newly synthesized polypeptide. It is also worth noting that GroEL showed the highest MDH suppression aggregation compared to DnaJ on its own (Figure 4.5). This is line with the study that showed that the overproduction of GroEL, but not of other chaperones, restored the growth of Δ*tig*Δ*dna*K52 cells at 30^{O}C and suppressed protein aggregation, including protein ≥ 60kDa which normally requires TF and DnaK for folding (Vorderwulbecke *et al*., 2004). The trigger factor, as mentioned (see section 1.7.3; Chapter one), is the first chaperone to bind to newly synthesized polypeptides exiting from ribosomes then hand them over to DnaK (Hartl and Hayer-Hartl, 2002). When a protein requires further folding, DnaK hands them over to GroEL, which encapsulate them until they complete the folding (Hartl and Hayer-Hartl, 2002). It has been suggested that GroEL could interact with substrate proteins that are regarded as DnaK/TF chaperones (Vorderwulbecke *et al*., 2004). This shows the crucial role played by molecular chaperones in protein folding, thus maintaining the cells viability. This also suggests that molecular chaperones are cooperative during the folding of newly synthesized polypeptides.

It has been shown that DnaK cooperates with DnaJ and GroEL during the folding of newly synthesized proteins (Alibolandi *et al*., 2010). On the other hand, a study conducted shows that KPf and PfHsp70 were able to suppress the thermosensitivity of the minus DnaK strain (Shonhai *et al*., 2005), suggesting that they operated in the same way that DnaK would have functioned if it were present in the system. This therefore demonstrates the possibility that both KPf and PfHsp70 may have cooperated with GroEL in the same way DnaK cooperates with DnaJ and GroEL during protein folding. In addition to that, the fact that KPf shares the same ATPase domain with DnaK might have added the advantage of KPf to cooperate with DnaJ and GroEL (Shonhai *et al*., 2005; Alibolandi *et al*., 2010). Therefore these biotechnological tools could be very useful for various proteins not only recombinant proteins from *plasmodium falciparum.*

CHAPTER FIVE

5.1 Conclusions and Future Perspectives

This study demonstrated the role played by molecular chaperones on the folding of newly synthesized proteins. This study developed combined molecular chaperones from different species that were very effective in removing contaminant proteins that copurify with the PfAdoMetDC protein. The most effective combination in this regard was, KPf+DnaJ+GroEL and PfHsp70+DnaJ+GroEL, which further demonstrated that PfAdoMetDC was folded properly, such that proteinase K could not digest the protein completely. It further provided the evidence that combined molecular chaperones could be a solution to the occasional folding challenges of recombinant protein (de Marco *et al*., 2007).

Molecular chaperones of the same species as the target protein could improve its production and quality of protein (Birkholtz *et al*., 2008). Based on a bioinformatics analysis, PfAdoMetDC has less hydrophobicity, but in spite of this the DnaK binding predictive tool showed that DnaK could bind to this protein. As such, the possible sites were predicted to be on the β-sheets of the protein, between 1-101 positions, suggesting that the protein could interact with molecular chaperones, depending on the exposed hydrophobic region in the protein. The question regarding the origin of contaminant proteins, particularly Hsp70, was investigated. This was done by expressing PfAdoMetDC using DnaK minus strain. The final product of the protein was pure and active. Therefore, this study is in line with a previous study that demonstrated that the production of recombinant protein on the *E. coli* Δ*dnaK* strain reduces DnaK background (Retelade *et al*., 2009; Williams *et al*., 2011).

The production of PfAdoMetDC protein on this system provided the evidence of the source of contaminating proteins, showing that it was endogenous DnaK. Interestingly, despite the fact that

PfAdoMetDC protein was produced on this system as quality protein, and showed the highest activity, its folding status showed that it required the supplementation of molecular chaperones for it to fold properly. The protein was completely degraded by proteinase K, indicating the crucial role played by molecular chaperones during protein folding. This is proven by the fact that PfAdoMetDC supplemented with molecular chaperones was not completely digested, suggesting that the protein was protected until it completed its folding. In addition, the protein activity was improved, compared to the protein that was supplemented with DnaK and the one that was not supplemented with molecular chaperones. Furthermore, molecular chaperones were able to suppress the MDH aggregation as individual and combined chaperones, indicating that molecular chaperones are cooperative during the folding of newly synthesized polypeptide.

5.2 Limitations to the experimental approach and how they could be overcome in future projects

Despite the successful purification of PfAdoMetDC protein, the activity assays for the protein that was co-expressed with Hsp70-GroEL was not evaluated for technical reasons. Also to further investigate the cooperation of these molecular chaperones by conducting ATPase activity assays, it would be interesting to clone and purify GroES protein to assess its contribution in protein folding. In addition, determination of the residence time of each chaperone on PfAdoMetDC would clarify their contribution to its folding process. Sub-cloning of the regions of PfAdoMetDC that were observed to be most hydrophobic based on predictive tools could also shed light on the effects of its subdomains on its folding fate.

References

Agashe, V. R., Guha, S., Chang, H. C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F. U., and Barral, J. M. (2004). Function of trigger factor and DnaK in multi-domain protein folding: increase in yield at the expense of folding speed. *Cell*, **117**: 199–209.

Alibolandi, M., Mirzahoseini, H., Nehi, F.M., Tabatabaian, G., Amini, H., and Sardari, S. (2010). Improving recombinant protein solubility in *Escherichia coli*: Identification of best chaperone combination which assists folding of human basic fibroblast growth factor. *Afri. J. Biotechn*, **9**:8100-8109.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*.**181**: 223-230.

Baneyx, F. and Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol*. **22**:1399–1408.

Bell, L. S., Chiang, A. N., and Brodsky, J. L. (2011). Expression of a Malarial Hsp70 improves Defects in Chaperone-Dependent Activities in ssa1 Mutant Yeast. *PLos ONE,* **6** (5): e 20047.doi.10.1371/journal.pone.0020047.

Birkholtz, L. M., Wrenger, C., Joubert, F., Wells, G. A., Walter, R. D, Louw, I.A. (2004). Parasite specific inserts in the bifunctional S-Adenosylmethionine decarboxylase/ornithine decarboxylase in *Plasmodium falciparum* modulate catalytic activities and domain interactions. *Biochem. J*.**377:** 439-448.

Birkholtz, L. M., Blatch, G. L., Coetzer, T. L., Hoppe, H.C., Human, E., Morris, E.J., Ngcete, Z., Oldfield, L., Roth, R., Shonhai, A., Stephens, L., and Louw, A.I. (2008). Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malaria Journal* 2008, **7**:197 doi: 10.1186/1475-2875-7-197.

Botha, M., Pesce, E. R., and Blatch, G. L. (2007). The Hsp40 proteins of *Plasmodium falciparum* and other apicomplexa: Regulating chaperone power in the parasite and the host. *Int J Biochem Cell Biol*, **39** (10):1781-1803.

Bukau, B. and Walter, G.C. (1990). Mutations altering heat shock specific subunit of RNA polymerase suppress major cellular defects of E. coli mutants lacking the DnaK chaperone. *EMBO J*. 9: 4027-4036

Bukau, B., Jonathan, W., and Arthur, H. (2006). Molecular chaperones and Protein Quality control. *Cell*, **04**: 014.

Butz, J.A., Niebauer, R.T., and Robinson, A.S. (2003). Co-expression of molecular chaperones does not improve the heterologous expression of mammalian G-protein coupled receptor expression in yeast. *Biotechnol. Bioeng*. **84**: 292-304.

Bradford, M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of protein utilizing the Principle of Protein-Dye Binding. *Anal. Biochem*, **72**: 248-354.

Buchner, J. and Walter, S. (2002). Molecular chaperones-Cellular Machines for Protein Folding. *Angew. Chem. Int. Ed*, **41**:1098-1113.

Carrio, M.M and Villaverde, A. (2002). Construction and deconstruction of bacterial inclusion bodies. *J. Biotechnol*, **96**: 3-12.

Cheng, M. Y., Hartl, F. U., and Horwich, A. L. (1990). The mitochondrial chaperone Hsp 60 is required for its own assembly. *Nature*, **348**: 455-458.

Cheng, H.C., Kaiser, C.M., Hartl, F.U. and Barral, J.M. (2005). De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J. Mol.Biol*, **353**:397–409.

Calloni, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G. G., Hayer-Hartl, M., and Hartl, F. U. (2012). DnaK Functions as a central Hub in *E. coli* chaperone network. *Cell Reports*, **1**: 251-264.

Cui, S., Lin, X., and Shen, J. (2011). Effects of co-expression of molecular chaperones on heterologous soluble expression of the cold-active lipase Lip-948. *Protein. Expr. Purif*, **77**: 166-172.

Daugaard, M., Rohde, M., and Ja¨a¨ttela, M. (2007). The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Lett*, **581**: 3702-3710.

Deuerling, E., Schulze-Specking, A., Tomoyasau, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature*, **400**: 693-696.

de Marco, A and De Marco, V. (2004). Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expressing tuning. *J. Biotechnol*, **109**: 45-52.

de Marco, A., Vigh, L., Diamant, S., and Goloubinoff, P. (2005). Native folding of aggregationprone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell Stress Chaperon* **10**: 329-39.

de Marco, A., Deuerling, E., Mogk, A., Tmoyasu, T., and Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli. BMC Biotechnol,* **7**:32.

Dobson, C. M. (2004). Principles of protein folding, misfolding and aggregation. *Semin Cell, Dev Biol*, **15**: 3-16.

Durme, V. J., Maurer-Stroh, S., Gallardo, R., Wilkinson, H., Rousseau, F., and Schymkowitz, J. (2009). Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. *PLos Comp Biol*, 8: e1000475.doi10.1371.

Edkins, A. L and Boshoff, A. (2014). General structural and functional features of heat shock proteins. Book Chapter in: Heat shock proteins of malaria. Editors: Addmore Shonhai and Greg Blatch. Springer, ISBN 978-94-007-7437-7; pp 5 -45.

Ellis, R. J. (2006). Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem. Sci*. **31**, 395–401.

Fink, A. L. (1999). Chaperone-mediated protein folding. *Physiol Rev* **79**: 425-49.

Flaherty, K. M., DeLuca-Flaherty, C., McKay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70 K heat-shock cognate protein. Nature, 346:623–628.

Flick, K., Ahuja, S., Chene, A., Bejarano, M. T., Chen, Q. (2004). Optimized expression of Plasmodium falciparum erythrocyte membrane protein 1 domains in Escherichia coli. *Malar J,* **3**: 50.

Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., Picotti, P., and Zambonin, M. (2004). Probing protein Structure by Limited Proteolysis. *Act. Biochem. Pol,* **51**: 299-321.

Gitau, G.W., Mandal, P., Blatch, G.L., Przyborski, J., and Shonhai, A. (2012). Characterisation of the *Plasmodium falciparum* Hsp70-Hsp90 organising protein (PfHop). *Cell Stress Chaperone*, 17: 191-202.

Georgio, G and Valax, P. (1996). Expression of correctly folded proteins in *Escherichia coli Curr. Issue. Biotech.* **2**. 190-196.

Gupta, R.,Lakshmipathy, S. K., Chang, H.C., Etchells, S. A., and F.U. Hartl. (2010). Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett*., **584:** 3620–3624.

Hartl, F. U and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **259**: 1852-1858.

Heby, O., Persson and Rentala, M. (2003).Targeting the polyamine biosynthetic enzymes: a promising approach to therapy of Afrcan sleeping sickeness, chagas' disease, and leishmaniasis. *Amino acids*, 2: 359-366.

Kramer, G., Rutkowska, R.D., Wergryzn, H., Patzelt, T.A., Kurz, F., Marz, T., Rauch, S. V.Deuerling, E., and Bukau, B. (2002). Functional Dissection of Escherichia coli Trigger factor: Unraveling the Function of Individual Domains. *J. Bacteriol*, **12**: 3777-3784.

Kumar, R., Dhalib, S., Srikanthb, R., Ghosha, S. K., and Srivastavaa, S. (2014). Comparative proteomics of mitosis and meiosis in Saccharomyces cerevisiae. *J. Proteomics*, 109: 1-15.

Kyte. J., and Doolittle. R. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol*. **157**: 105-132.

Kolaj, O., Spada, S., Robin, S., and Wall, J.G. (2009). Use of folding modulators to improve heterologous protein production in Escherichia coli. *Microb Cell Fact* **8**: 9-26.

Lakshmipathy, S. K., Gupta, R, Pinkert, S., Etchells, S. A and Hartl, F. U. (2010). Versatility of trigger factor interactions with ribosome-nascent chain complexes. *J. Biol Chem*, **285**: 27911-23.

Li, J and Sha, B. (2003). Crystal Structure of the E. coli Hsp100 ClpB N-Terminal Domain. *Structure*, **3**: 323-328.

Li, J., Qian, X., and Sha, B. (2009). Heat shock protein 40: structural studies and their functional implications. *Protein Pept Lett*, **16**: 606-12.

Mayer, M. P and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular machenism. *Cell. Mol. Life Sci*, **62**: 670-684.

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

Marz, F., Hoffmann, A., Rutkowska, A., Zachmann-Brand, B., Bukau, B., and Deuerling, E. (2006). The c-terminal domain of *Escherichia coli* trigger factor represent the central module of its chaperone activity. *J Biol Chem*, **281**(42): 31963-71.

Matambo, T., Odunuga, O.O., Boshoff, A., Blatch, G. (2004). Overexpression, purification and characterisation of the plasmodium falciparum heat shock protein 70. *Prot Expr Purif*, **33**: 214-222.

Mehlin, C., Boni, E., Buckner, F. S., Engel, L., Feist, T., Gelb, M. H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P. J., Reddy, J. T., Sampson, J. N., Subramanian, E.,Van Voorhis, W. C., Worthey, E., Zucker, F., and Hol, W. G. J. (2006). Heterologous expression of proteins from *Plasmodium falciparum*: Results from 1000 genes. *Mol Biochem Parasitol*, **148**:144-160.

Mosser, D.D., Caron, A.W., Bourget, L., Merii, A.B., Sherman, M.Y., Morimoto, R.I., and Massie, B. (2000). The chaperon function of hsp70 is required for protection against stressed-induced apoptosis. *Mol Cell Biol*, **20**: 7146-7159.

Muller, S., Coombs, S. H., and Walter, R.D. (2001). Targeting polyamines of parasitic protozoain chemotherapy. *Trends in parasitology*, **17**: 242-249.

Nishihara, K., Kanemori, M., Yanagi, H., and Yura, T. (2000). Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Applied and Evironmental Microbiology*, **3**: 884-889.

Patzelt, H. S., Rudiger, D., Brehmer, G., Kramer, S., Vorderwulbecke, E., Schaffitzel, A., Waitz, T., Hesterkamp, L., Dong, J., Schneider-Mergener, B., Bukau, E. Deuerling. (2001). Binding specificity Escherichia coli trigger factor. *Proc.Natl.Acad*.Sci.USA **98**:1424.

Pesce, E.R., Acharya, P., Tatu, U., Nicoll, W.S., Shonhai, A., Hoppe, H. C., and Blatch, G.L. (2008). The *plasmodium falciparum* heat shock protein 40, Pfj4, associates with heat shock protein 70 and shows similar heat induction and localisation patterns. *J Biochem Cell Biol*, **40**: 2914-2926.

Pegg, E. A. (2006). Regulation of ornithine decarboxylase. *J. Biol. Chem*, 281: 14529-14532.

Rajan, V. B and D'Silva, P. (2009). Arabidopsis thaliana J-class heat shock proteins: cellular stress sensors. *Funct Integr Genomics*, **9**: 433–446.

Ramya, T.N., Surolia, N., and Surolia. A. (2006). 15-Deoxyspergualin modulates *Plasmodium falciparum* heat shock protein function. *Biochem. Biophys. Res*. **348**: 585–592.

Ratelade, J., Miot, M., Johnson, E., Betton, J., Mazodier, P., and Benaroudj, N. (2009). Production of Recombinant Proteins in the Ion-Deficient BL21 (DE3) strain *Escherichia coli* in the absence of the DnaK chaperone. *Appl and Env Microbiol*, **75**: 3803-3807.

Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in Drosophila. *Experientia*. **18**: 571-573.

Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *The EMBO Journal*, 16: 1501-1507.

Schroder, H., Langer, T., Hartl, F. U., and Bukau, B. (1993). DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBOJ*. 12: 4137-44.

Shi, Y. Y., Hong, X. G., and Wang, C. C. (2005). The C-terminal (331–376) sequence of Escherichia coli DnaJ is essential for dimerization and chaperone activity: a small angle X-ray scattering study in solution. *J. Biol. Chem*. **280**: 22761–22768.

Shi, Y., Fan, D., Li, S. X., Zhang, H. J., Perrett, S., Zhou, J. M. (2007). *Protein Sci*. **16**: 1165-1175.

Shonhai, A., Botha, M., De Beer, T.A.P., Boshoff, A., and Blatch, G.L. (2008). Structure-function study of a *Plasmodium falciparum* Hsp70 using three dimensional modelling and *in vitro* analyses. *Protein and Peptide Letters*, **15**:1117–1125.

Shonhai, A., Boshoff. A., and Blatch, G. L. (2005) *Plasmodium falciparum* heat shock protein70 is able to suppress the thermosensitivity of an *Escherichia coli* DnaK mutant strain. *Biol Gen Genomics*, **274**:70-78.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H.Y., and Hay, S. I. (2005). "The global distribution of clinical episodes of Plasmodium falciparum malaria". Nature, **434** : 214–7.

Stephens. L. L., Shonhai, A., and Blatch, G. L. (2011). Co-expression of the *Plasmodium falciparum* molecular chaperon, PfHsp70, improves the heterologous production of the antimalarial drug target GTP cyclohydrolase I, PfGCHI. *Protein. Expr. Purif*, Doi:10.1016/j.pep.

Srinivasan, S. R., Gilles, A., Chang, L., Thompson A. D., and Gestwick, J. E. (2012). Molecular chaperones DnaK and DnaJ share predicted binding sites on the most proteins in the E. coli proteome. *Mol Biosyst*, **9**:2323-2333.

Teter, S. A., Houry, W. A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F. U. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell*, **97**: 755–765.

Thiel, L. (2005) "Gathering of data on the funding mechanisms for malaria research and the status of malaria research in South Africa". Malaria Initiative for Africa (SAMI) Report.

Wang, C.; Delcros, J.G., Cannon, L., Konate, F., Carias, H., Biggerstaff, J., Gardner, R., and Phanstiel, O. (2003). Defining the molecular requirements for the selective delivery of polyamine-conjugates into cells containing active polyamine transporters. *J. Med. Chem*. **46**: 5129-5138.

Walsh, P., Bursac, D., Law, Y. C., Cry, D., and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep*. **5**: 567-571.

Wells, G.A., Birkholtz, L.M, Joubert, F., Walter. R.D. and Louw. A.I. (2006). Novel properties of malarial S-adenosylmethionine decarboxylase as revealed by structural modeling. *JMGM*:, **24**: 307–318.

Williams, M., Sprenger, J., Human, E., Al-Karadaghi, S., Persson, L., Louw, A.I., and Birkholtz, L. M. (2011). Biochemical characterisation and novel classification of monofunctional S-Adenosylmethionine decarboxylase of *Plasmodium falciparum. Mol. Biochem. Parasitol*, **180**: 17-26.

Wu, C. (1995). Heat shock transcription factors: structure and regulation. *Annu. Rev. Cell Dev. Biol*, **11**:441–469.

World Health Organization. World Malaria Report 2013. Geneva: WHO (2013). Available from: http://www.who.int/malaria/publications/world_malaria_report_2013/en.

Zeilsta-Ryalls, J., Fayet, O., and Georgopoulos, C. (1991). The universally conserved GroE (Hsp60) chaperonins. *Annu Rev Microbiol*, **45:** 301-25.

Appendix A

General experimental procedures

A. 1 Preparation of competent *E. coli* BL21 (DE3) Star cells™ and *E.coli* XL1Blue cells

Colony of E.coli XL1Blue was inoculated in 5ml 2x YT broth (1.6 g tryptone, 1.0 g yeast, 1.5 g agar, 0.5 g NaCl per 100 ml preparation in deionized water) broth and grown overnight with shaking at 37 ºC. The overnight culture was diluted 1:200 into 50ml YT broth and thereafter grown with shaking to early log phase of absorbance 0.3-0.6 measured at 600nm. The cells were harvested by centrifuging at 5,000g for 10 minutes at 4 ºC. The cells were kept on ice from this point onwards. The cells were resuspended in 10ml 0.1 M $MgCl_2$ and left on ice for 30 minutes. The suspension was centrifuged for 10 minutes at 4000g at 4 ºC. The cells were pelleted as before and gently resuspended in 10 ml ice cold 0.1 M $CaCl_2$ and then followed by incubation on ice for 4 hours. Centrifugation was carried out at 4000g at 4 ºC for 10 minutes. The competent cells were aliquoted by adding 1 volume of sterile 30% glycerol mixing and stored at -80 ºC until use.

A. 2 Transformation of competent cells

Volume of 2 µl of plasmid DNA was added into an aliquote of 100 µl of competent cell.

The cells were then incubated on ice for 30 minutes followed by heat shocking at 42 ºC for 45 seconds and immediately placed on ice for 10 minutes. Volume of 900 µl of 2x YT broth was added and then incubated at 37 ºC for 1 hour. The cells were transferred on 2x YT plates containing the desired antibiotics followed by incubation at 37 ºC overnight.

A. 3 Small scale extraction of plasmid DNA

Plasmid DNA was extracted using Zyppy^{TM} Plasmid Miniprep Kit according to supplier's protocol.

A. 4 Large scale extraction of plasmid DNA

Plasmid DNA was extracted using Qiagen Kit according to supplier's protocol.

A. 5 Restriction digest of Plasmid DNA using enzymes

Plasmid DNA was digested using the desired diagnostic restriction enzymes as described below. The reagents were set up as follows:

Sterile deionised water	16 µl
10x restriction buffer	2 µl
DNA (100-200 ng)	2 µl

The reaction was initiated by addition of 2 units of restriction enzymes. The restriction was allowed to proceed for 2-3 hours at 37 ºC. The reaction was stopped by addition of 4 µl of 10x DNA loading buffer (0.25% bromophenol blue and 30% glycerol). The product was then analyzed by agarose gel electrophoresis as described in (Appendix A. 6).

A. 6 Agarose gel electrophoresis

To prepare (w/v) 0.8% agarose gel, the required amount of agarose was completely dissolved in 1x TAE buffer (40 mm, 20 mM acetic acid and 1 mM EDTA) by heating with frequent agitation. The agarose was then cooled to 55 ºC prior to addition of ethidium bromide (0.5 µg/ml). The

agarose gel was allowed to polymerise for 15-30 minutes at room temperature. The gel was placed in the electrophoresis chamber and covered with 1x TAE buffer. Volume of 4 µl of 10x DNA loading buffer (0.25% bromophenol blue + 30% glycerol) was added to 20 µl of sample followed by loading of the samples into the wells. Electrophoresis was allowed to proceed at 100 volts for 1 hour. The gel was then visualized using UV light (Gene Genius Bioimaging System Syngene, USA).

A. 7 Extraction of DNA from agarose gel

Resolved DNA was extracted from the agarose gel using the Zymoclean™ Gel DNA Recovery Kit according to the manufacturer's protocol. Briefly, after identification of the DNA fragment using UV light, the DNA was excised from the gel using a sterile scalpel and transferred to a 1.5 ml micro-centrifuge tube and 3 volumes of ADB buffer added to each volume of agarose excised from the gel. The agarose was then incubated at 55 ºC for 10 minutes until the gel slice was completely dissolved. The melted agarose solution was transferred to a Zymo–spin column in a collection tube followed by centrifugation at ≥10 000 x *g* for 30–60 seconds and the flow–through was discarded. The DNA was washed twice with wash buffer and centrifuged at ≥ 10 000 x *g* for 30 seconds. The flow–through was discarded and the DNA was eluted using ≥ 6 µl sterile distilled water and centrifugation at ≥10 000 x *g* for 30 – 60 seconds to elute DNA.

A. 8 Ligation reaction

Ligation reaction using (QIAGEN PCR Cloning Kit*)* was set up using 100 ng of required plasmid and 500 ng of DNA fragment used as the insert. Included in the reaction mix, was 5 µl of 2x rapid ligation buffer, 1 unit of T4 DNA ligase and deionized sterile water to make to a total

volume of 10 µl The ligation was left to proceed overnight at 4 ºC and thereafter the ligation product was transformed into JM109 super competent cells.

A. 9 DNA Sequencing

Following restriction digest of the respective plasmid using desired restriction enzymes, sequencing was performed to further confirm the integrity of the plasmids. DNA samples were prepared using Zyppy^{TM} Plasmid Miniprep Kit according to supplier's protocol. The DNA was eluted in sterile ultra-pure water to exclude interference of sequencing with salt contamination.

The sequencing reaction mixture was set up as follows: 250-500 ng of the DNA, 3.2 pmol forward or reverse primers respectively. Volume of 2 µl of 5X Big DyeR Terminator Sequencing Buffer and 4 µl of Big Dye Terminator (Applied Biosystems) was added to the mixture. Sterile triple distilled water was then added to a final volume of 10 µl. The following thermal clycling parameters were performed using a GeneAmp PCR System 9700 (version 3.05; Applied Biosystems): initial 1 cycle denaturation at 96 ºC for 2 minutes followed by 30 cycles of denaturation, annealing and elongation, 96 ºC for 30 seconds, desired annealing temperature for 30 seconds and 72 ºC for 30 seconds respectively and final elongation, 1 cycle at 72 ºC for 7 minutes. The product was purified and concentrated using Zymo–Spin I TM columns (Zymo Research, California, USA) according to the manufacturer's instructions followed by elution in 15 µl of water and vacuum dried. Following resuspension of the DNA in Hi-Di buffer, sequencing was performed using ABI 3100 Genetic analyzer (Applied Biosystems, USA) with a capillary electrophoresis. Analyzing of the sequence was performed using BioEdit Sequence Alignment Editor (version 7.0.4.1).

A. 10 Preparation of Chelating Sepharose Fast Flow beads

Chelating sepharose beads obtained from (GE Healthcare, Life Sciences) was prepared as per the manufacturer's instruction. In brief, the 50% slurry of Chelating Sepharose Fast Flow was charged with nickel as described below. The beads were washed by gently shaking the container of Chelating Sepharose Fast Flow to re-suspend the gel. Enough slurry to bind the histidine tagged protein was transferred into a 50 ml centrifuge tube. The storage buffer was removed by centrifugation at 1500 x g for 2-5 minutes and the supernatant discarded. The gel was then resuspended in five gel volumes of distilled water followed by gentle agitation until the gel is fully re-suspended. The gel was re-sedimented by centrifugation at 1500 x g for 2-5 minutes and the supernatant discarded.

The beads were then charged by addition of 1 volume of 0.1 M NiSO4 solution followed by gentle agitation for 15 min until the gel is fully re-suspended. The unbound nickel was removed by centrifugation at 1500 x g for 2-5 minutes and the supernatant discarded. The beads were washed using five gel volumes of distilled water with gentle agitation for 5 minutes followed by centrifugation at 1500 x g for 2-5 minutes and the supernatant discarded. The washing steps were repeated two more times after which the beads were re-suspend in one gel volume of Chelating Sepharose Fast Flow Start Buffer (20 mM Na2HPO4, 0.5 M NaCl, 10 mM imidazole pH 7.4).

After use, the regeneration of the beads was done by washing with 10 column volumes of

0.05 M EDTA. Residual EDTA was removed by washing with 2-3 column volumes of 0.5 M NaCl.

A. 11 Determination of protein concentration using Bradford assay

Protein concentration was determined by Bradford's method (Bradford, 1976). Bovine serum albumen (BSA) standards were prepared using concentration ranging from 0 to 2.5 mg/ml in 0.15 M Nacl. Bradfords reagent 200 µl (Sigma Aldrich, USA) was added to 10 µl of protein and the reaction incubated in the dark at room temperature for 5 minutes. Absorbance was read at 595 nm using (Das plate reader, Italy).The recombinant protein was similarly treated and the protein concentration determined by extrapolation from the standard curve as indicated in (Appendix C; Figure C.3). The readings were prepared in triplicate and the average obtained.

A. 12 Sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were treated by boiling in SDS sample buffer (0.25% Coomasie Brilliant blue (R250); 2% SDS; 10% glycerol (v/v); 100 mM tris; 1% mercaptoethanol) in a ratio of 4:1 for 5 minutes at 95 °C and resolved using 12% acrylamide resolving gel prepared as shown in (Table A.1). The gel were then transferred into the electrophoresis tank and electrophoresis buffer (25 mM Tris, pH 8.3 250 mM glycine and 0.1 % (w/v) SDS) was added. The boiled samples were loaded in respective wells and prestained protein molecular weight markers (ThermoFisher Scientific, USA) were also loaded. The electrophoresis was performed at 150 volts for hour using the BioRad Mini protein 3 electrophoresis system (Biorad, U.S.A).

Table A.1 Preparation of SDS-PAGE

Reagent (ml)	12% Separating gel	5% Stacking gel
30% Polyacrylamide	2.08	0.235
10% Ammonium persulphate	0.025	0.00875
10% SDS	0.05	0.0175
1.5 M Tris (pH 8.8)	1.25	
1.0 M Tris (pH 6.8)		0.437
Distilled water	1.58	1.05
TEMED	0.0025	0.003

A. 13 Western Blot

Proteins were resolved in 12% acrylamide gel as described above (**Appendix A.12**). Removal of SDS-PAGE gel from the glass plates after completion of electrophoresis process and cutting off of the stacking gel was done. The what-man filter papers, gel, two scotchbrite fibre pads and nitrocellulose were immersed in western transfer paper and left to equilibrate at 8 ºC for 30 minutes. Preparation of the gel for transfer was done as follows: filter paper was placed on a scotch brite pad; Gel was placed on the filter paper ensuring no air bubbles are trapped; nitrocellulose was placed over the gel; another filter paper was laid on top of the nitrocellulose, followed by another scotch brite pad. The transfer of the protein on the nitrocellulose membrane was performed by running at 100 volts for 1 hour. The membrane was removed from the sandwich and rinsed using transfer buffer as well as removal of adhering gel on the nitrocellulose membrane using a cotton swab. The blot was stained with ponceau stain to determine the success of the transfer followed by visualizing the band using chemiluminescence. The membrane was blocked in 10 ml of (5% nonfat milk in TBS) for 1 hour on a rotary shaker set at 1000 g. The membrane was washed three times in TBS-Tween for 10 minutes followed by incubation of the membrane with primary antibody for 1 hour. Unbound primary antibody were removed by washing of the membrane three times using TBS-Tween for 10 minutes each wash. The membrane was incubated with secondary antibody for 1 hour followed by washing of the membrane 3 times using TBS-Tween. Images were acquired using X-ray film (Pierce CL-Xposure™ Film, Thermo Scientific, USA).

Appendix B

Primers used for site-directed mutagenesis

Table B.1 Generation of modified plasmids used to replace DnaK with either PfHsp70 or KPf segments

Modified plasmid	Codon changes	Forward and reverse primers	Reasons for mutation	Diagnositc restriction enzymes
pBB535	acgtaa tcgccc	gactctcttccg**ggtacc**atgccataccgcg caaagacaaaaaataa**cccggg**ataaacgggta	Introduction of *Bam*HI and *Sma*I	*Bam*HI and *Sma*I
pBB542	acgtaa tcgccc	gactctcttccg**ggtacc**atgccataccgcg caaagacaaaaaataa**cccggg**ataaacgggta	Introduction of *Bam*HI and *Sma*I	*Bam*HI and *Sma*I

Codon changes introduced are highlighted in bold

Table B. 2: Pipetting instructions for the site-directed mutagenesis reaction

Component	50 µL reaction	Final concentration
H2O	Add to 50 µL	
10x reaction buffer	5 µL	1x
10 mM dNTPs	1 µL	200 µM each
Forward primers	X µL	125 ng
Reverse primers	X µL	125 ng
Template DNA	X µL	10 ng
***Pfu* ultra HF DNA polymerase (2 U/µL)**	1 µL	0.02 U/µL

Table B.2: Cycling instruction for the site-directed mutagenesis reaction.

Cycle step	Temperature	Time	Number of cycles
Initial Denaturation	98º C	30 s	1
Denaturation	98º C	5-10 s	
Extension	72º C	10-30 s/kb	25
Final extension	72º C	5-10 min	1
	4º C		

Table B. 3 Reaction set up for PCR

Dream taq Green PCR master mix (2x)	**25 µl**
Forward primer	**0.5 µM**
Reverse primer	**0.5 µM**
Template DNA	**10 pg- 1 µg**
Water, nuclease free	**Up to 50 µl**

Table B. 4 Thermal cycling conditions.

Step	Tm °C	Time	Number of cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 s	
Annealing	Tm-5	30 s	25
Automated flouresecent	72	1 min/kb	
Extension			
Final extensioon	72	12	1

Appendix C

C. 1 Bradford standard curve

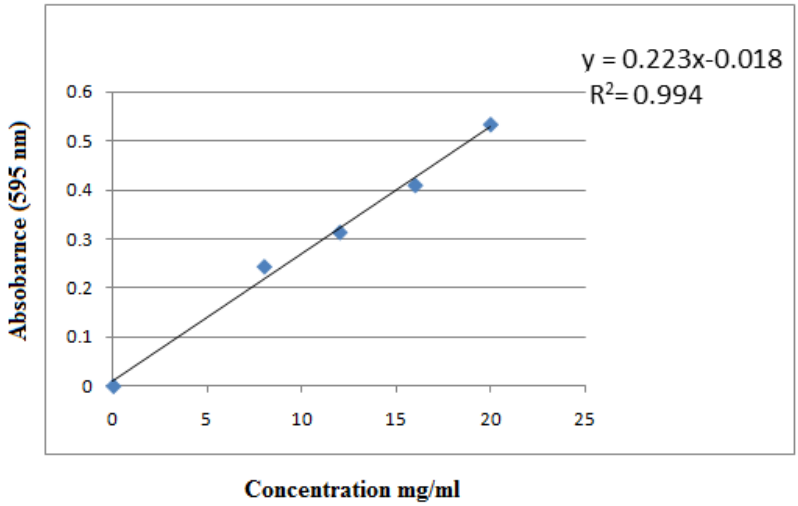

Figure C. 1 Bradford standard curve for protein concentration determination. Bovine serum albumen (BSA) standards of concentration ranging from 0 to 25 mg/ml were prepared and absorbance was read at 595 nm using (Biorad, U.S.A). The linear equation: y = 0.223x-0.018; R^2= 0.994 was used to calculate the protein concentration.

Appendix D (Preparation of solutions)

PBS buffer (pH 7.4)

137mM sodium chloride, 27mM potassium chloride, 4.3mM sodium hydrophosphate 1 and 1.4mM potassium hydrophosphate were dissolved in ddH_20 and autoclaved.

TE buffer (pH 8.0)

0.1211 g Tris and 0.0370 g EDTA 100 ml were mixed together with distilled water and autoclaved.

Sodium hydroxide

115 g of sodium hydroxide was dissolved in 250ml of distilled water.

SDS (10%)

1 g of SDS was dissolved in 10 ml of distilled water

DNA extraction solutions

Solution 1 (pH 8.0)

0.9008 g glucose, 0.628 g Tris chloride and 0.372 g EDTA were mixed together and the volume was made up to 100 ml

Solution 2

2ml of 10M sodium hydroxide and 10 ml of SDS were mixed together and the volume was made up to 100 ml

Solution 3 (pH 5.0)

29.442 g of potassium chloride was dissolved in distilled water and the volume was made up to 100 ml.

30% glycerol

30 ml of glycerol was mixed together with 70 ml of distilled water to make 100 ml of 30% glycerol.

IPTG

1 g of IPTG was dissolved in 4.2 ml of distilled water and filtered through filter into 1 ml eppendorf tube then stored at -20 ^{0}C

10X TAE buffer (pH 7.8)

48.4 g of Tris, 20 ml of 0.5M EDTA and 11.42 ml of glacial acetic acid were mixed together with distilled water and made up to 1000 ml.

Ampicillin

1 g of ampicillin was dissolved in 10 ml of distilled water and filtered through a filter unit into eppendorf tube100 µg/ml

Ampicillin agar plates

1.6 g of pancreatic tryptone, 1.5 g nutrient agar, 0.5 g of sodium chloride and 1 g of yeast were mixed together with distilled water and the volume was made up to 100 ml in a flask and autoclaved. The mixture was allowed to cool and 10 µl of ampicillin was added and poured into Petri dishes and allowed to solidify.

Liquid broth

1.6 g of pancreatic tryptone, 0.5 g of sodium chloride and 1 g of yeast were mixed together with distilled water and the volume was made up to 100 ml in a flask and autoclaved.

Ammonium persulphate (10%)

1 g of Ammonium persulphate was dissolved in 10 ml of distilled water.

Acrylamide stock solution

300 ml of distilled water was placed in a beaker and 150 g of acrylamide was added to distilled water while stirring and 4.0 g of bisacrylamide was added while stirring and allowed to dissolve. The volume of the solution was made up to 500 ml with distilled water. The acrylamide/Bisacrylamide solution was filtered with a cloth and placed in a dark container and stored in the fridge.

Ethanol 70%

70 ml of alcohol was mixed with 30 ml of distilled water and the volume was made up to 100 ml

0.1M Calcium chloride

1.47 g of calcium chloride was dissolved in 100 ml of distilled water

0.1M Magnesium chloride

2.0 g of magnesium chloride was dissolved in 100 ml of distilled water

Ethidium bromide (10 mg/ml)

1 g of ethidium bromide was dissolved in 100 ml of distilled water and stored in a bottle covered with aluminum foil

Tris

1.5M Tris (pH 8.8)

36.3 g of Tris was dissolved in 200 ml of distilled water

0.5M Tris (pH 6.8)

12.11 g of Tris was dissolved in 200 ml of distilled water

Destaining solution

800 ml of distilled water, 7.5 ml glacial acetic acid and 100 ml of ethanol were mixed together and the volume was made up to 1000 ml

Short protocol of the Strep-Tactin chromatography cycle

1. After the protein extract has entered the *Strep*-Tactin matrix, wash column 5 times with 1 CV (column volume) Buffer W.

2. Elute recombinant protein by the addition of 6 times 0.5 CV Buffer E.

3. Regenerate the column by the addition of 3 times 5 CV Buffer R.

4. Equilibrate the column by the addition of 2 times 4 CV Buffer W prior to the next purification run.

5. Store the column at 4° C overlaid with 2 ml Buffer W or R.

Appendix E

Special reagents and chemicals

Table D.1 List of reagent and Special reagents

Name of reagents	Vendor/Supplier
Agarose	Merck
Ammonium persulphate	Sigma-Aldrich
Ampicillin	Calbiochem
Chemiluminescence Western blotting kit	Amarsham
2-mercaptoethanol	Merck
Bovine serum albumin	Melford
Bromophenol blue	Merck
Calcium chloride	Merck
Coomasie brilliant blue R250	Merck
Ethidium bromide	Merck
Glacial acetic acid	Merck
Glycerol	Merck
Glycine	Merck
Imidazole	Merck
Lambda DNA	Fermantas Life sciences
Lysozyme	Merck
Methanol	Merck
Magnesium chloride	Merck
Monoclonal mouse anti-His antibody	Merck

Phenylmethylsufonyl fluoride	Merck
Polyacrylamide	Merck
Ponceau S	Sigma-Aldrich
Restriction enzymes	Merck
Sodium chloride	Merck
Sodium dodecyl sulphate	Merck
Sodium hydroxide	Merck
TEMED	Sigma-Aldrich
Tris	Merck
Tryptone	Oxoid, England
Tween 20	Radchem
Yeast	Merck
Broth	Merck
Nutrient agar	Merck
Strep-tagged Antibody	Merck
Glucose	Merck
L+Arabinose	Merck
Anhydrotetracycline	Merck
Kanamycine	Merck
Strep-tactin	IBA Germany
Buffer W	IBA Germany
Buffer E	IBA Germany
Buffer R	IBA Germany
Silver staining kit	Fermantas life sciences

Protein ladder	Fermantas life sciences
Anti-strep tag Antibodies	Merck

Abstract

Plasmodium falciparum S-adenosylmethionine decarboxylase (PfAdoMetDC) plays an important role during the synthesis of polyamines, such as putrescine, spermidine and spermine. Polyamines play a major role in the survival of the malaria parasite. They are critical components of cell growth and division, particularly in rapidly proliferating cells that include cancerous cells and numerous parasites. Hence, PfAdoMetDC protein is regarded as an ideal drug target for malaria. Even though *E. coli* is the most established expression system for recombinant protein expression, the production of this malarial drug target protein in *E. coli* still remains an obstacle towards development of compounds or drugs to prevent its function. In this study, an approach that improves the quality of PfAdoMetDC protein produced in *E. coli* was developed based on the use of molecular chaperones as co-expression partners. For structural studies aimed at designing drugs or compounds obtaining a pure protein is crucial. Therefore, PfAdoMetDC was co-expressed with six different combinations of molecular chaperones. *E. coli* BL21 (DE3) star™ cells were transformed with plasmids in six combinations: combination 1 (pBB535) encoding DnaK with DnaJ; combination 2 (pBB535-Pf70J) encoding PfHsp70 with DnaJ; combination 3 (pBB535-KPfJ) encoding KPf with DnaJ; combination 4 (pBB542) encoding DnaK with DnaJ and GroEL; combination 5 (pBB542-Pf70JE) encoding PfHsp70 with DnaJ and GroEL; and combination 6 (pBB542-KPfJE) encoding KPf with DnaJ and GroEL. The effects of these combinations on the conformation of purified PfAdoMetDC protein was assessed by limited proteolysis and also by conducting activity assays to evaluate the activity of the protein. The PfAdoMetDC protein expressed with supplementation of PfHsp70+DnaJ and KPf+DnaJ combinations was not completely degraded by proteinase K and its activity was higher compared

to the protein that was not supplemented with molecular chaperones. Almost similar results were observed on PfAdoMetDC co-expressed with DnaJ+GroEL, combined with the respective Hsp70s. The protein was not completely degraded, suggesting that PfAdoMetDC protein was protected by these chaperones during its folding. PfAdoMetDC protein, produced using *E. coli* Δ*dnaK* strains, showed the highest activity compared to co-expressed protein with molecular chaperones. It was, however, completely degraded by proteinase K. This suggests that Hsp70 molecular chaperones are important for the fold and stability of proteins. To further confirm any possible cooperation amongst these molecular chaperones, MDH aggregation suppression assays were conducted to mimic what might be taking place inside the cell. The recombinant proteins DnaK, KPf, PfHsp70, GroEL and DnaJ were evaluated for their ability to suppress MDH aggregation, both as individual chaperones and as combined chaperones. There were indications that these molecular chaperones may be cooperative during protein folding. This study also revealed that the biotechnological tools developed in this study may be useful to other recombinant proteins, and not only those from *plasmodium falciparum*.

DECLARATION

I, Xolani Henry Makhoba, declare that this is my own unaided work hereby submitted for the degree of Doctor of Philosophy of University of Zululand in the Faculty of Science and Agriculture. It has not been submitted for any degree for examination at any other university.

_____

Xolani Henry Makhoba

DATED THIS _____ DAY OF _____ _____AT

_____

DEDICATION

This thesis is dedicated to my late Mother N Makhoba, my young Sister Nokwanda Makhoba and my nieces Emihle and Asemahle for selflessly supporting me in every way throughout my studies.

PERSONAL ACKNOWLEDGEMENTS

Mr S.S. Cele and Mrs Cele

Thank you for the immense support, love and encouragement.

ACKNOWLEDGEMENTS

I would like to acknowledge the following individuals for their contributions to this work:

My supervisor, Professor Addmore Shonhai – who endeavored to share his passion and immense enthusiasm for Science with me during the course of this study. Professor Shonhai inspired me to remain motivated and encouraged me to think creatively and independently as a Scientist in my own right.

Professor Lyn-Marie Birkholtz – who provided me with pASK-IBA/PfAdoMetDC plasmid and valuable technical support.

Professor Bernd Bukau – who kindly donated pBB535 and pBB542 plasmid and the predictive tool for DnaK binding on substrates.

Dr. Adelle Burger – who took her valuable time to proofread my thesis, shared her experience with me, encouraged me and kept me grounded.

The University of Zululand, Department of Biochemistry and Microbiology – where I conducted this study.

My lab mates and friends – who supported me.

And, last but not least, I wish to give all the praise to the almighty Jehovah God who is the pillar of my strength in good times and bad times.

ACKNOWLEDGEMENT OF FUNDING

I wish to thank University of Zululand Research Committee for funding this project and the National Research Foundation (South Africa) Scarce Skills Development Fund Programme for the scholarship. I would also like to thank the Carl Storm International Diversity Fellowship Program in the United States for granting me a travelling grant to attend the Gordon Research Conference.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Symbols

Abbreviations of units	Symbol Interpretation
%	percent
µl	microlitre
A360	absorbance at 360 nanometres
A600	absorbance at 600 nanometres
bp	base pair
kDa	kilodalton
µM	micromolar
nmol	nanomoles
nm	nanometres
ºC	degree Celsius
ml	millilitre
l	litres
w/v	weight per volume
v/v	volume per volume
µg	microgram
ng	nanogram
g	gram
β	beta
α	Alpha

IUPAC amino acid single letter codes (SLC) and corresponding DNA codons

SLC	Amino	Abbreviation	DNA codons
A	Alanine	Ala	GCT, GCC, GCA, GCG
C	Cysteine	Cys	TGT, TGC
D	Aspartic Acid	Asp	GAT, GAC
E	Glutamic Acid	Glu	GAA, GAG
F	Phenylalanine	Phe	TTT, TTC
G	Glycine	Gly	GGT, GGC, GGA, GGG
H	Histidine	His	CAT, CAC
I	Isoleucine	Ile	ATT, ATC, ATA
K	Lysine	Lys	AAA, AAG
L	Leucine	Leu	CTT, CTC, CTA, CTG, TTA, TTG
M	Methionine	Met	ATG
N	Asparagine	Asn	AAT, AAC
P	Proline	Pro	CCT, CCC, CCA, CCG
Q	Glutamine	Gln	CAA, CAG
R	Arginine	Arg	CGT, CGC, CGA, CGG, AGA,AGG
S	Serine	Ser	TCT, TCC, TCA, TCG, AGT, AGC
T	Threonine	Thr	ACT, ACC, ACA, ACG
V	Valine	Val	GTT, GTC, GTA, GTG
W	Tryptophan	Trp	TGG
Y	Tyrosine	Tyr	TAT, TAC

LIST OF RESEARCH OUTPUTS

Makhoba, X.H., Coertzen, D., Birkholtz, L.M., and Shonhai, A. Rehosting of molecular chaperones in *E. coli* improves the quality of recombinantly expressed *Plasmodium falciparum* AdoMetDC (Under review).

CONFERENCE PROCEEDINGS

(1) Presented at international conference: Gordon Research Conference, Polyamines in Biology and Disease, Waterville Valley Resort, NH **United States**, June 15-22, 2013. **Xolani H. Makhoba[1]**, Lyn-Marie Birkholtz[2], Addmore Shonhai. Characterisation of the structure-function status of *Plasmodium falciparum* AdoMetDC recombinant protein co-expressed with variable molecular chaperone combinations in *E. coli*

(2) Oral poster presentation: VI International Congress on Stress Proteins in Biology and Medicine, from 18th to 22nd August 2013, held in Sheffield, **United Kingdom**. **Xolani H. Makhoba[1]**, Dina le Rou[2] Lyn-Marie Birkholtz[2], Addmore Shonhai[1]. Exploring the effect of various chaperone combinations towards facilitating the production of recombinant *Plasmodium falciparum* AdometDC protein in *E. coli*

(3) Poster presentation: 6th MIM PAN-AFRICAN MALARIA CONFERENCE, from 6th to 11th October 2013, held in Durban South Africa. **Xolani H. Makhoba[1]**, Dina le Rou[2] Lyn-Marie Birkholtz[2], Addmore Shonhai[1]. Use of *Plasmodium falciparum* Hsp70 and its derivative to improve the production and quality of recombinant PfAdoMetDC.