In cellulo and biophysical exploration of the role of GGMP residues of Hsp70

A dissertation submitted in fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN BIOCHEMISTRY,

FACULTY OF SCIENCE, ENGINEERING, AND AGRICULTURE

UNIVERSITY OF VENDA

By

Dongola Tendamudzimu Harmfree

11633824

Supervisor: Prof A. Shonhai

Co-supervisor: Dr T. Zininga

Abstract

Hsp70 is a prominent molecular chaperone. Structurally, Hsp70 is composed of two domains, C-terminus substrate binding domain (SBD) and nucleotide binding domain (NBD). The NBD of Hsp70 is responsible for its ATPase activity. Some Hsp70s of parasites particularly those of apicomplexa are marked by GGMP residues. For example, *Plasmodium falciparum* Hsp70-1 (PfHsp70-1) which occurs in the cytosol and nucleus harbors seven GGMP repeats on the C-terminus upstream of the EEVD motif. The function of GGMP residues of Hsp70 is largely unclear but they were recently reported to be involved in substrate and co-chaperone binding. Therefore, the main aim of this study was to investigate the role of the GGMP residues of Hsp70s using *E. coli* Hsp70 (DnaK) and chimeric protein, KPf as models. Chimeric protein KPf is made up of NBD of DnaK and SBD of PfHsp70-1. *E. coli* DnaK lacks the GGMP residues that are present in PfHsp70-1. DnaK-G (DnaK modified to include GGMP residues) was created to elucidate the function of these residues. Exogenously expressed KPf and DnaK are known to reverse the thermosensitivity of *E. coli dnaK756* cells. The native DnaK of *E. coli dnaK756* cells is functionally compromised making this strain heat sensitive. $KPf_{617\text{-}647}$ and $KPf_{\Delta G}$ mutants were previously created by conservative substitution and deleting the GGMP residues of KPf, respectively. Circular dichroism and tryptophan fluorescence highlighted that the insertion of GGMP residues did not drastically change the secondary and tertiary structure of DnaK. Cytoprotection of *E. coli dnaK756* cells could not be recovered when cells were heterologous expressing DnaK-G. DnaK recovered the cytoprotection of the same cells as expected. Furthermore, mutations on GGMP residues adversely impacted the chaperone function of KPf with respect to cytoprotection of *E. coli dnaK756* cells. On the other hand, the absence of the GGMP residues results in KPf losing its chaperone function. This suggests that the SBD of PfHsp70-1 requires the GGMP residues to function, whereas *E. coli* DnaK does not require these residues. It was further noted that insertion of the GGMP residues into DnaK led to toxicity in *E. coli dnaK103* cells under permissive growth temperature. A proteomic study conducted to establish the functional deficiencies of DnaK-G established that the presence of the GGMP residues compromised the capability of DnaK-G to bind some substrates. Altogether the findings suggest that GGMP repeats account for the specialized function of PfHsp70-1 in distinction to that of *E. coli* DnaK.

Keywords: Heat shock proteins, molecular chaperones, GGMP residues DnaK, complementation assay

Declarations

I, **Tendamudzimu Harmfree Dongola**, declare that the work and contents of this dissertation my own and have not been submitted for any other degree at any university or institution. This thesis does not contain any other person's writing unless specifically acknowledged and referenced accordingly.

Signature (Student):  Date: 07/09/2021

Table of contents

Lists of outputs

1. Makumire, S., **Dongola, T.H**., Chakafana, G., Tshikonwane, L., Chauke, C.T., Maharaj, T., Zininga, T., Shonhai, A. (2021). Mutation of GGMP Repeat Segments of *Plasmodium falciparum* Hsp70-1 Compromises Chaperone Function and Hop Co-Chaperone Binding. *International Journal of Molecular Sciences*. **22**(4). DOI: 10.3390/ijms22042226. PMID: 33672387; PMCID: PMC7926355

Dedication

I dedicate this dissertation to my daughter Thendo Zwonaka Dongola, my uncle Ntswikiseni Reckson Dongola and to my uncle Solomon Mukwevho who gave me the necessary support throughout my studies. Lastly, I dedicate this dissertation to my late parents who should be proud of the things I have achieved so far.

Acknowledgments

I would like to thank God above everything who gave me wisdom, strength, and his spirit during this study. Special thanks to my supervisors who had hopes in me, encouraged me to do my best and supported me financially during the study, Prof A. Shonhai, and Dr. T. Zininga. Furthermore, I would also like to pass my gratitude to Dr. N.E Madala for his support, he made things looked easy for me during the study and helped me to grow as a person and in my career. I would also like to thank Stellenbosch University central for analytical facility (CAF) for their help in identifying protein fragments with their LCMS. Thanks to Dr Stanley Makumire for proof-reading my dissertation. Lastly, I would like to thank Protein Biochemistry and Malaria (ProBioM) group for their support.

List of symbols

Units abbreviations	Interpretation of symbols
Bp	base pair
kDa	kilodalton
°C	degree Celsius
µl	microlitre
µg	microgram
g	gram
α	alpha
g/f	glycine-phenylalanine
β	beta
%	percent
µl	microlitre
A600	absorbance at 600 nanometres
ml	millilitre
l	litres
w/v	weight per volume
v/v	volume per volume
mM	millimolar

1.0 Literature review

1. Introduction

1.1 Molecular chaperones

Molecular chaperones are involved in protein quality control and protein homeostasis. Molecular chaperones are proteins that interact with, stabilize or assist other proteins in acquiring their functionally active conformation. Moreover, molecular chaperones do not form part of the final conformations of the folded protein. When proteins fold *in vivo*, some of the proteins fail to reach their native state (Hoffmann *et al*., 2010). This is because the *in vivo* environment is too crowded which results in protein aggregation. Consequently, molecular chaperones are required to minimize protein aggregation and maintain substrates in competent folding form (Muñoz *et al*., 2015). There are various distinct classes of molecular chaperones that facilitate specialized functions (Hartl *et al*., 2011; Kim *et al*., 2013).

1.2 Heat shock proteins

Heat shock proteins (Hsps) are prominent molecular chaperones that are important in protein folding. Hsps are highly conserved molecules that occur in almost all life forms and are localized in different subcellular compartments in the cell. Hsps are grouped based on their functions as “holdases” that keep substrates in a folding competent form, thus reducing the chances of protein aggregation. “Foldases” are those chaperones that assist misfolded or partially folded proteins to fold to their native states (Hoffmann *et al*., 2010). Most Hsps are expressed when the cell is under stressful conditions. When these Hsps are expressed, they play a crucial role in cell survival and development (Przyborski *et al*., 2015). Hsps are classified according to their molecular weight average sizes, examples of these are; Hsp110 (110 kDa), Hsp100 (100 kDa), Hsp90 (90 kDa), Hsp70 (70 kDa), Hsp60 (60 kDa), Hsp40 (40 kDa), and small Hsp families (15-29 kDa) (Datta *et al*., 2017).

1.3 Heat shock protein 40

The causative agent of malaria, *Plasmodium falciparum* expresses several Hsp40 members of Hsp40 numbering up to fifty-one (Botha *et al*., 2007; Njunge *et al*., 2007). Structurally, Hsp40 is composed of four functional domains, J-domain, glycine-phenylalanine rich domain (G/F rich), zinc finger region, and substrate binding domain on the C-terminus domain. The J-domain consists of four α-helices (helix I, II, III, and IV) and histidine-proline-aspartic acid

(HPD) motif located at the loop region between Helix II and III (Figure 1). The highly conserved HPD motif promotes Hsp40 and Hsp70 interaction (Hennessy *et al*., 2005; Pesce *et al*., 2008). Hsp40s are grouped based on their domain organization (type I, II, III, and IV). Type I Hsp40s possess all domains, type II Hsp40s possesses all the domains except the zinc finger region. Type III Hsp40s possess only the conserved J-domain with HPD but at variable positions along their sequence and other Hsp40 domains (Botha *et al*., 2007). Type IV Hsp40s possess a J-domain that lacks the HPD motif and the J domain position is variable (Figure 1.1). Type I and Type II Hsp40s have the capabilities to recruit substrates to Hsp70 thus stimulating ATP hydrolysis (Botha *et al*., 2011).

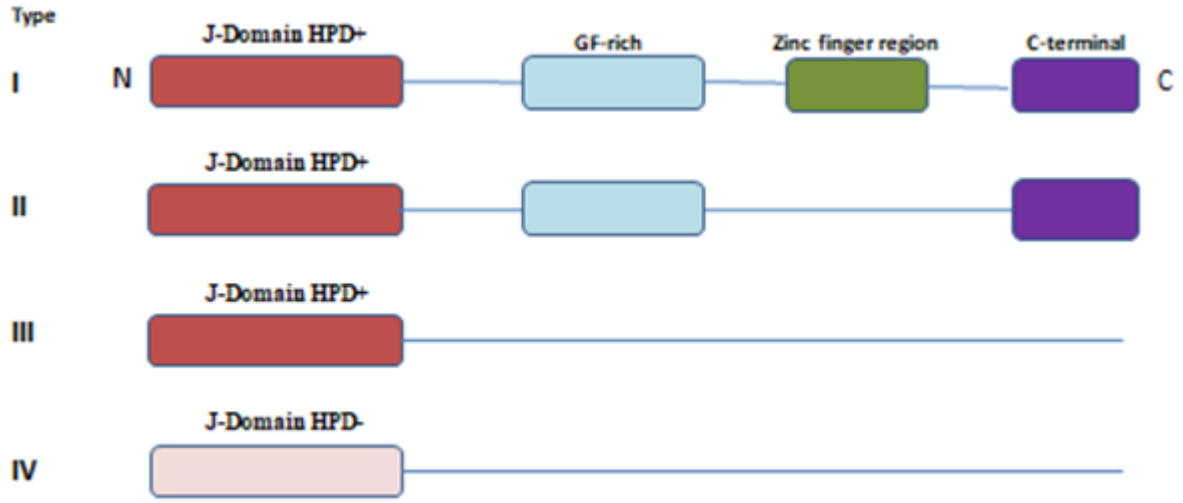


Figure 1.1: Schematic representation of Hsp40 subtypes

PfHsp40 type I possesses full J domain conservation thus J-domain with HPD motif, GF rich region, and zinc finger region. Type II DnaJ proteins have J domain and gly/phe rich region (G/F rich) motif. Type III DnaJ proteins have conserved J domain only and type IV possesses a J-domain with HXD motif (Rug and Maier 2011).

1.4 Heat shock protein 70

Heat shock protein 70 (DnaK in prokaryotes) acts as molecular chaperones by protecting cells under heat stress (Bukau and Walker, 1989: Flaherty *et al*., 1990: Shonhai *et al*., 2005). Hsp70s are generally stress-induced proteins. However, some Hsp70s are constitutively expressed and are named heat cognate proteins 70 (Hsc70) (Shonhai *et al*., 2007: Fakhari *et al*., 2013). Prokaryotic and mammalian Hsp70s are conserved (Yu *et al*., 2015: Chakafana *et al*., 2019). *Plasmodium falciparum* expresses 6 Hsp70s while humans express 17 Hsp70s (Chakafana *et al*., 2019). Hsp70 isoforms are localized in different subcellular compartments. Hsp70s are custodians of protein quality control and hence regulate parasite development (Zininga and

Shonhai, 2019). It is thought that the minor structural variations present in Hsp70s account for their functional difference (Chakafana *et al*., 2019). Variations in Hsp70 counterparts may account for substrates and peptides binding preference as previously reported (Zininga *et al*., 2016; Mabate *et al*., 2018).

The NBD and SBD are connected by a highly conserved linker domain that is important for their allosteric communication (Figure 1.2) (Mayer *et al*., 2000: Chakafana *et al*., 2019). The SBD is composed of two subdomains, the β-sandwich sub-domain, and the α-helical subdomain lid segment (Flaherty *et al*., 1990: Figure 1.2).

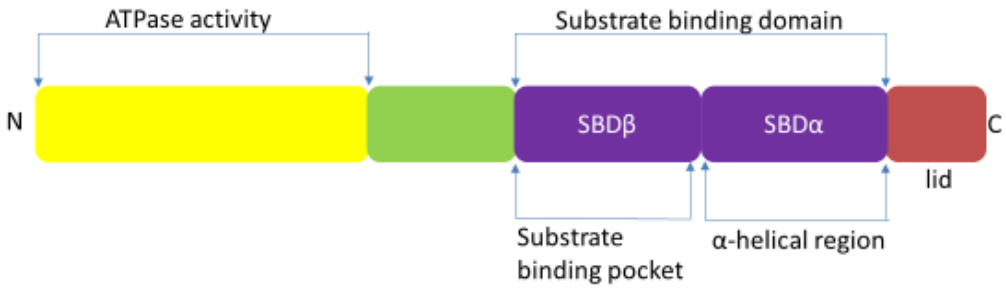


Figure 1.2: Diagram of conserved Hsp70 domains

The conserved structure of Hsp70 possesses the N-terminus binding domain (NBD) (ATP-binding domain) (in yellow), substrate binding domain (SBD) composed of the β- and α-region (in blue) and the C-terminus lid (in maroon). The SBD and NBD are joined by a conserved linker (in green) responsible for interdomain communication (Shaner and Morano, 2007).

The bound peptide is recognized by the C-terminus β-subdomain substrate of Hsp70. A substrate bound to Hsp70 is released through an allosteric signal from ATP-bound NBD to the SBD influencing the opening of the lid. The SBD β- and α- subdomains are responsive to ATP-driven conformational changes, thus regulating the affinity for peptides (Kityk *et al*., 2012). Hsp70 has a high affinity for substrates when bound to ADP and a low affinity for substrate when bound to ATP (Nguyen *et al*., 2017).

Hsp70s are versatile molecules, as their co-operation with other Hsps such as Hsp40, Hsp60, Hsp90, Hsp100, Hsp110, and small Hsps enhance their function (Mogk *et al*., 2015). Hsp70s are classified into two sub-family groups, namely canonical Hsp70s (resemble *E. coli* Hsp70 known as DnaK) and non-canonical Hsp70s (Hsp110/glucose-regulated protein Grp170) (Chakafana *et al*., 2019; Chakafana and Shonhai, 2021). Hsp110 and Grp170 are closely related chaperones (Chakafana and Shonhai, 2021). Notable differences in nucleotide binding affinity and ATP hydrolysis between canonical Hsp70s and non-canonical Hsp70s have been reported

(Zininga *et al*., 2016). The chaperone function of canonical Hsp70s is regulated by nucleotides while the function of Hsp110/Grp170 is nucleotide independent (Shonhai *et al*., 2007; Xu *et al*., 2012; Chakafana and Shonhai, 2021). Hsp110/Grp170 suppresses the aggregation of proteins subsequently handing them to canonical Hsp70s for further folding (Mattoo *et al*., 2014; Mogk *et al*., 2015; Chakafana *et al*., 2021).

Hsp110 members are specialized Hsp70-like proteins that exhibit distinct features from canonical Hsp70s (Chakafana *et al*., 2019; Chakafana and Shonhai, 2021). Hsp110 is localized in the cytosol and is thought to function as nucleotide exchange factors (NEF) of canonical Hsp70s (Dragovic *et al*., 2006: Gauley *et al*., 2008). NEF mediates the exchange of ATP to ADP, thus supporting substrate release in Hsp70 (Dragovic *et al*., 2006). Hsp110 suppresses protein aggregation through its role of stabilizing misfolded proteins thus preventing protein misfolding (Andreasson *et al*., 2008; Zininga *et al*., 2015). Structurally, Hsp110s possess the SBD-β and the SBD-α subunits and harbour an extended acidic insertion and a distinct linker segment (Zininga *et al*., 2014; Chakafana *et al*., 2019). The acidic insertion of Hsp110 and Hsp70 is thought to regulate substrate binding (Chakafana and Shonhai, 2021). Hsp110 has been previously reported to possess 7 β strands which account for substrates preferences (Xu *et al*., 2012; Chakafana and Shonhai, 2021). Hsp110 substrates are generally rich in aliphatic residues (Chakafana and Shonhai, 2021).

Grp170 is localized in the endoplasmic reticulum (ER) and is not only induced during glucose starvation but also under broad cell stress conditions such as hypoxia and proteasome inhibitors (Zhang *et al*., 2017; Chakafana and Shonhai, 2021). Grp170 has been reported to be effective in blocking the aggregation of proteins and assisting in the translocation of polypeptides into the ER (Kudyba *et al*., 2019).

The NBD of canonical Hsp70s and non-canonical Hsp70s (Hsp110/Grp170) is conserved. On the other hand, their SBDs show great structural variation (Chakafana *et al*., 2019). Canonical and non-canonical Hsp70s exhibit distinct substrate preferences. Canonical Hsp70s have their substate binding cleft characterized by $L_{1,2}$ and $L_{3,4}$ loops located in the SBDβ which could account for variation in substrate preference as compared to non-canonical Hsp70s (Chakafana *et al*., 2019).

1.5 Hsp70 functional cycle

Hsp70 forms a functional cycle with Hsp40 co-chaperone (type I and type II) localized in the cytosol (Hennesy *et al*., 2005: Botha *et al*., 2011). Hsp70 functional cycle begins when Hsp40 recruits an unfolded substrate to Hsp70 (Kityk *et al*., 2012). Subsequently, ATP hydrolysis into ADP results in the closing of the lid and stabilization of peptide binding (Kityk *et al*., 2012; Figure 1.3). NEF swap ADP for ATP binding by Hsp70 and this process facilitates the release of bound substrates as Hsp70 shifts into a low affinity state when bound to ATP.

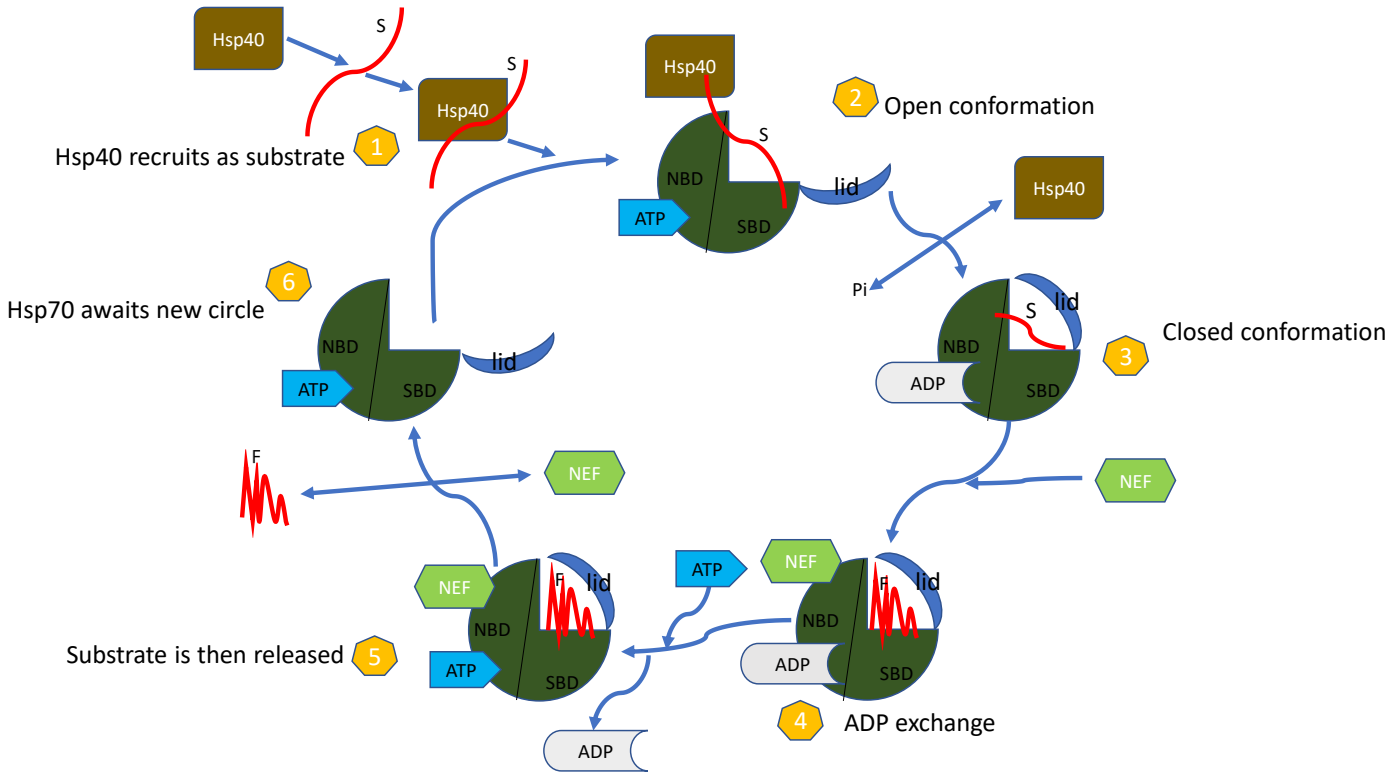


Figure 1.3: The Hsp70 functional cycle

The Hsp70 folding cycle is mediated by ATP hydrolysis and nucleotide exchange factors. Hsp70-Hsp40 chaperone complex starts when Hsp40 recruits misfolded substrate to Hsp70 in its low substrate's affinity (ATP-bound state). ATP is hydrolyzed and Hsp70 assumes a high affinity substrate binding status associated with lid closure in ADP. NEF catalyzes the exchange of ADP for ATP resulting in folded substrate release when the α-helical lid opens due to ATP binding.

1.6 *E. coli* DnaK

DnaK is needed by *E. coli* cells for protein homeostasis (Bukau and Walker 1989; Haslbeck *et al*., 2019). DnaJ is the co-chaperone of DnaK. DnaK co-operates with a co-chaperone Grop-like gene (GrpE), which serves as its NEF (Richter *et al*., 2010). On the other hand, σ^{32} is a transcriptional factor for *E. coli* (Liberek *et al*., 1995). σ^{32} binds to the bacterial enzyme RNA polymerase to activate the transcription machinery (Roncarati and Scarlato, 2017). DnaK also interacts with Hsp100 (ClpB) a molecular chaperone that depends on ATP (Mogk *et al*., 2015). IbpA and IbpB are members of small heat shock proteins which are 48% identical (Chuang *et al*., 1993). Both IbpA and IbpB co-operate with Hsp100 and Hsp70 in mediating refolding of aggregating proteins (Baneyx and Mujacic, 2004). DnaK interacts with bacterial Hsp90 to refold misfolded proteins (Thomas and Baneyx, 1998; Kravats *et al*., 2017). DnaJ enhances the ATPase activity of DnaK (Wild *et al*., 1996). Altogether, DnaJ and GrpE help DnaK to perform its function under heat stress (Ahn and Im, 2020). When an unfolded substrate is recruited to DnaK by DnaJ, ATP is hydrolyzed promoting substrate binding to DnaK (Kampinga *et al*., 2010). One of the approaches used to study the role of Hsp70 in *E. coli* is complementation assay (Shonhai *et al*., 2005). To study the role of DnaK/Hsp70 by complementation involves the use of an *E. coli* strain whose DnaK gene is either deleted or mutated (Shonhai *et al*., 2005). Replacement of the DnaK protein by a non-functional Hsp70 leads to normal growth of the *E. coli* cells. Otherwise, *E. coli* cells whose DnaK gene is non-functional are known to be unable to grow at elevated temperatures (Shonhai *et al*., 2005; Makumire *et al*., 2021). Hsp70 Various mutated *E. coli* strains have been used to study the function of Hsp70 (Section 2;2.10). Low demand maintenance of *E. coli* cells offers a good model for *in vivo* studies to study the function of Hsp70s.

1.7 Hsp70 functional specificity

DnaK's role in the folding of proteins, disaggregation of aggregated proteins, and disassembly of protein complexes has been well established (Clerico *et al*., 2015). *E. coli* DnaK is localized in the cytosol and has been previously reported to protect cells under heat stress conditions (Shonhai *et al*., 2005: Makhoba *et al*., 2016: Makumire *et al*., 2021). DnaK has a high affinity for peptides that are enriched with hydrophobic residues (Rüdiger *et al*., 2001). The preference of DnaK for hydrophobic enriched substrates could define the difference in functional features of plasmodial Hsp70 and DnaK as *P. falciparum* Hsp70 prefer to bind asparagine enriched peptides (Lebepe *et al*., 2020). *P. falciparum* (malaria causative agent) Hsp70-1 prefers binding

to asparagine rich peptides as compared to its counterparts human Hsp70 and DnaK (Makumire *et al*., 2021). DnaK lacks the GGMP residues present in PfHsp70-1 (Makumire *et al*., 2021). Moreover, the chimera KPf made up of NBD of DnaK and SBD of PfHsp70-1 is known to protect *E. coli* cells with compromised DnaK function (Shonhai *et al*., 2005). At least 30% of the proteome of *P. falciparum* is glutamate/asparagine rich and hence is deemed to be prone to aggregation (Singh *et al*., 2004; Pallarès *et al*., 2018). The presence of the GGMP residues in the PfHsp70-1 SBD was recently reported to be crucial for substrate binding of the chaperone (Makumire *et al*., 2021). Furthermore, PfHsp70-1 is reported to have high ATPase activity and its chaperone function as compared to its human counterpart and DnaK (Lebepe *et al*., 2020; Matambo *et al*., 2004; Anas *et al*., 2020; Makumire *et al*., 2021). The presence of the GGMP residues in PfHsp70-1 is thus thought to account for PfHsp70-1 functional efficiency and specificity.

PfHsp70-1 and DnaK are conserved structurally and they both are canonical Hp70s. However, the C-terminus is marked with seven GGMP residues (Makumire *et al*., 2021). Thus, the chimeric protein, KPf harbors the SBD of PfHsp70-1 which is then characterized by the GGMP repeats (Shonhai *et al*., 2005; Makhoba *et al*., 2016; Figure 1.4).

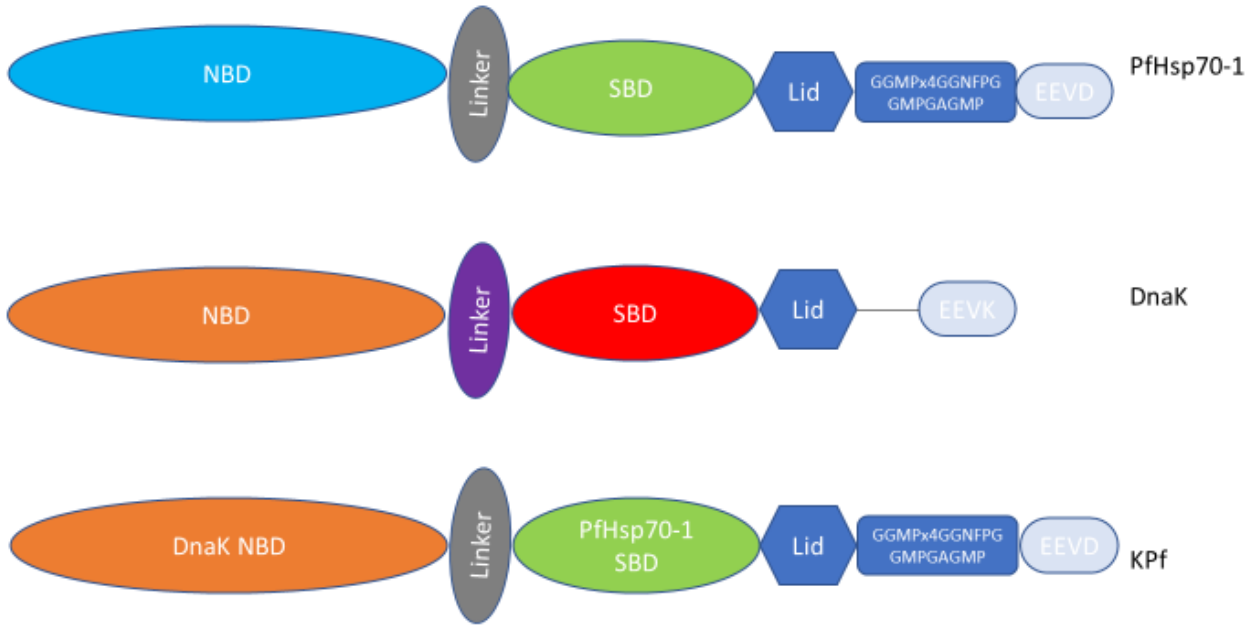


Figure 1.4: Structural organization of DnaK, PfHsp70-1, KPf

Both PfHsp70-1 and DnaK possess the NBD (light blue for PfHsp70-1 and orange for DnaK) and SBD (green for PfHsp70-1 and red for DnaK) interconnected by a flexible linker (grey for PfHsp70-1 and purple for DnaK). At the C-terminus, PfHsp70-1 possesses the GGMP residues (blue) upstream to the EEVD motif (light purple). DnaK

lacks the GGMP residues in the C-terminus. A chimeric protein, KPf possesses the NDB of DnaK and SBD of PfHsp70-1. NBD: nucleotide binding domain, SBD: substrate binding domain.

1.8 GGMP residues of Hsp70

In some eukaryotic organisms, Hsp70 possesses GGMP residues which are tetra-peptide repeats located in the C-terminus (Wang *et al*., 2015). Most of the Hsp70s that possess the GGMP residues are localized in the cytosol. In addition, their presence in this exclusive organelle suggests a distinct role of the residues (Makumire *et al*., 2021). PfHsp70-1 possesses seven GGMP tetra-peptide repeat residues in the C-terminus upstream of the EEVD motif (Matambo *et al*., 2004; Zininga *et al*., 2015; Makumire *et al*., 2021). The EEVD motif is responsible for PfHsp70-1 interaction with PfHop (*P. falciparum* Hsp70/Hsp90 organizing protein) (Gitau *et al*., 2012; Chakafana *et al*., 2019; Makumire *et al*., 2020). GGMP residues are not only present in parasitic Hsp70s but also Hsp70s from various organisms especially members of apicomplexa (Makumire *et al*., 2021). In addition, GGMP residues are absent in *E. coli* DnaK (Makumire *et al*., 2021). Conserved repeats of GGMP residues were reported to be partially responsible for the unique functional features of PfHsp70-1 such as high ATPase activity of PfHsp70-1 (Lebepe *et al*., 2020; Makumire *et al*., 2021). The dominant presence of GGMP repeats in Hsp70s of parasitic organisms may reflect their role in regulating parasite virulence (Makumire, 2019). In place of GGMP residues, yeast Hsp70 possesses GGAP residues that were reported to regulate substrate binding (Gong *et al*., 2018). The stability of PfHsp70-1 is regulated by GGMP residues (Makumire *et al*., 2021). There is a need to further investigate the role of this important functional motif.

1.9 Problem statement

Hsp70s are molecular chaperones that play a central role in protein quality control to maintain protein homeostasis. PfHsp70-1 possesses GGMP repeat residues on the C-terminus adjacent to the EEVD motif. A previous study suggested that PfHsp70-1 GGMP residues play a role in the host immune response (Matambo *et al*., 2004). However, an independent study contradicted these findings as PfHps70-1 has been shown to lack immunomodulatory activity (Pooe *et al*., 2017). Recent evidence showed that the GGMP repeat residues are important for substrate binding and the conformational stability of PfHsp70-1 (Makumire *et al*., 2019). PfHsp70-1 upregulation under heat stress implies that the chaperone is crucial in protecting the parasite under stressful conditions (Zininga *et al*., 2015). Furthermore, PfHsp70-1 expression was reported throughout all the blood stages of the malaria parasite life cycle (Pallavi *et al*., 2010).

PfHsp70-1 has high ATPase activity as compared to its human counterpart (Makumire *et al*., 2019). Moreover, the GGMP residues are implicated in upregulating the ATPase activity of PfHsp70-1 (Makumire *et al*., 2021). Hsp70s are potential drug targets due to their role as protein folding machines in parasitic and diseased cells (Shonhai *et al*, 2010). It is well documented that Hsps are prospective drug targets and are implicated in malaria drug resistance (Zininga and Shonhai, 2019; Daniyan *et al*., 2019). Therefore, this study sought to investigate the function of the GGMP repeat residues of Hsp70. The lack of GGMP residues in DnaK and the presence of GGMP residues in chimeric protein KPf offers a model for studying the role of the GGMP residues *in cellulo*. Furthermore, it is important to explore if deletion or substitution of GGMP repeat residues of the chimeric protein KPf, would affect its role in cytoprotection of *E. coli* cells.

1.10 Hypothesis

The GGMP residues promote the chaperone function of parasitic Hsp70s.

1.11 Aim

To investigate the role of GGMP residues on the chaperone activity of PfHsp70-1

1.12 Objectives

- To establish the distribution of GGMP residues in Hsp70s from various organisms
- To determine the secondary and tertiary structure confirmations of DnaK and DnaK-G
- To investigate the role of GGMP motif on Hsp70 using *E. coli* complementation assay
- To investigate the role of GGMP residues of Hsp70s using a proteomics approach

2.0 Methodology

2.1 Materials

Plasmids used in this study are listed (Table 2.1) and include pQE30/DnaK, pQE60/DnaK, mutants pQE60/KPf, pQE60/KPf$_{\Delta G}$ (whose GGMP repeat residues is deleted), pQE60/KPf$_{617-647}$ (whose GGMP repeat residues were substituted conservatively), KPfV$_{436F}$ (has substituted valine at position 436 with phenylalanine) and pQE30/DnaK-G (DnaK with GGMP repeat residues inserted). Anti-His antibody (Rabbit raised, Thermo Scientific, USA) was used to confirm recombinant proteins (Table 2.1). Strains used in this study are also listed (Table 2.2).

Table 2.1: Constructs description

Plasmids	Protein encoded for	Antibiotic resistance	His-tag	References
1. pQE30/DnaK	DnaK	Amp^R	✓	(Makumire *et al*., 2021)
2. pQE30/DnaK-G	DnaK-G	Amp^R	✓	(Makumire *et al*., 2021)
3. pQE60/DnaK	DnaK	Amp^R	✓	(Burkholder *et al*., 1994)
4. pQE60/KPf	KPf	Amp^R	x	(Shonhai *et al*., 2005; Makhoba *et al*., 2016)
5. pQE60/KPf$_{\Delta G}$	KPf$_{\Delta G}$	Amp^R	x	(Makumire *et al*., 2021)
6. pQE60/KPf6$_{17-647}$	KPf$_{617-647}$	Amp^R	x	(Makumire *et al*., 2021)
7. pQE60/KPf$_{V436F}$	KPfV436F	Amp^R	x	(Shonhai *et al*., 2005)

Amp^R: ampicillin resistant, ✓: symbolizes the presence of his-tag, x: symbolizes the absence of his-tag

Table 2.2: Description of *E. coli* strains

Strains	Characteristics	PT	NPT	Antibiotics resistance	References
***E. coli dnaK756* (BB2362)**	DnaK expressed by this strain has three substitutions at residue 32, 455, and 468 resulting in the strain having low GrpE affinity.	37°C	>40°C	Kanamycin, Tetracycline	(Buchberger *et al*., 1999 Shonhai *et al*,. 2005)
***E. coli dnaK103* (BB2392)**	Expresses truncated DnaK and high levels of DnaJ	30°C	40°C	Chloramphenicol, Tetracycline	(Spence *et al*., 1990; Mayer *et al*., 2000; Shonahi *et al*., 2005)
***E. coli* Δ*dnaK52* (BB1553)**	DnaK gene is truncated and expresses a low amount of DnaJ	30°C	40°C	Chloramphenicol Tetracycline	(Paek and Walker, 1987; Shonhai *et al*., 2005)

PT: permissive growth temperature, NPT: non-permissive growth temperature

2.2 Bioinformatics studies

Multiple sequence alignment was conducted to establish the occurrence of GGMP residues across Hsp70 homologs. Amino acids were first retrieved from (https://www.ncbi.nlm.nih.gov/protein/) and Uniprot (https://www.uniprot.org). Multiple sequence alignment was then conducted using Bioedit sequence alignment. The conservation percentage of residues was further calculated with the Bioedit tool.

2.3 Designing of KPf mutant constructs

Conservative amino acids substitution (methionine to leucine; glycine/proline to alanine; and asparagine to aspartic acid) and deletion of the GGMP repeat residues from residues 601-657 of KPf molecule to construct $KPf_{617\text{-}647}$ and $KPf_{\Delta G}$ respectively. Amino acids were substituted based on biochemical property similarities (Betts, 2003: Makumire *et al*., 2021). The focus of mutation was on GGMP repeat residues (Section 3. Figure 3.1.2).

2.4 Confirmation of plasmid constructs

E. coli JM109 DE3 competent cells (Appendix A.1) were transformed with the following constructs; pQE30/DnaK, pQE60/DnaK, mutants pQE60/KPf, pQE60/$KPf_{\Delta G}$, pQE60/$KPf_{617\text{-}647}$, and pQE30/DnaK-G. Plasmids were purified using A Zippy™ mini prep kit following the manufacturer's instruction (Appendix A.3). Purified DNA was analyzed with restriction digest using *Bam*HI and *Hind*III enzymes (Thermo Scientific, USA) to digest and confirm their integrity. Digested DNA was analyzed by 0.8% agarose gel electrophoresis

2.5 Comparative analysis of three-dimensional models of KPf, DnaK, and their GGMP mutants

An *in silico* study was conducted to predict the possible implications of introducing mutations on the structure of KPf and DnaK. Phyre^2 (http://www.sbg.bio.ic.ac.uk/phyre2) was used to generate the three-dimensional models of KPf and DnaK relative to their GGMP mutants (Kelley *et al*., 2015). KPf and its GGMP variants were used in the *in cellulo* study, hence it was important to predict structural variation between KPf variants. PDB files were then retrieved from Phyre^2 and visualized using Chimera version 1.9 (Pettersen *et al*., 2004). Superpose (http://wishart.biology.ualberta.ca/Superpose/) was subsequently used to super-position the three-dimensional structures. The root mean square deviation (RMSD) was used to determine comparative structural similarities (Pettersen *et al*., 2004). Hydrogen bonding was measured using Chimera (www.cgl.ucsf.edu/chimera).

2.6 Recombinant protein expression

Chemically competent *E. coli* JM109 cells were transformed with pQE30/DnaK, pQE60/DnaK, pQE60/KPf, pQE60/$KPf_{\Delta G}$, pQE60/$KPf_{617\text{-}647}$, and pQE30/DnaK-G constructs. The transformed cells were used to express the his-tagged recombinant proteins following a previously described protocol (Zininga *et al*., 2016). A single colony of each transformant was picked and inoculated into fresh 50ml 2 YT broth containing 100 µg/ml ampicillin.

Transformants were cultured overnight at 37°C with shaking incubator at 150 rpm. After overnight incubation, culture was diluted at a 1:10 ratio and then incubated until at OD A_{600} = 0.6. The culture was then induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) to induce protein expression. Pre induction sample was collected, and post-induction hourly samples were collected until the 6th hour of post-induction for SDS-PAGE analysis. Cells were harvested by centrifuging at 5000 xg for 20 minutes at 4 °C and pellet fractions were suspended in lysis buffer (100 Mm Tris-HCL, pH 7.4, 300 Mm NaCl, 10 mM imidazole, 1 mM EDTA, 1 mM PMSF, and 1 mg/mL lysozyme). To allow cell lysis, *E. coli* cells were incubated at room temperature for 1 hour with gentle shaking and then stored at -80°C. Expression of protein was analyzed by SDS-PAGE analysis (Appendix A.4) and confirmed with Western blot (Appendix A.5).

2.7 Protein purification

Recombinant proteins were purified using sepharose nickel affinity chromatography as previously described (Zininga *et al*., 2015). Cell lysates were centrifuged for 20 minutes at 4°C at 5000 xg (Shonhai *et al*., 2008). The supernatant was added to HisPur^{TM} Nickel-charged nitrilotriacetic acid (Ni-NTA) (Thermo Scientific, USA) immobilized affinity chromatography column and incubated for 1 hour to enhance binding of recombinant protein. Unbound proteins were washed with 2-bed volume of wash buffer I (10 ml Tris-HCl pH 7.5; 300 mM NaCl; 25 mM imidazole; 1 mM PMSF) and wash buffer II (10 ml Tris-HCl pH 7.5; 300 mM NaCl; 80 mM imidazole; 1 mM PMSF). Proteins that were bound to the beads were eluted with elution buffer I (10 ml Tris-HCl pH 7.5; 300 mM NaCl; 250 mM imidazole; 1 mM PMSF) and elution buffer II (10 ml Tris-HCl pH 7.5; 300 mM NaCl; 500 mM imidazole; 1 mM PMSF). Collected wash and elution samples were prepared for analysis with 12% SDS-PAGE (Appendix A.4) and western blot (Appendix A.5).

2.8 Determination of the secondary structure of DnaK and DnaK-G

Far-UV circular dichroism (CD) spectroscopy was used to determine the secondary structure of DnaK and its mutant under different temperatures (from 20 ºC to 90 ºC) as previously described (Zininga *et al*., 2016). Chirascan Plus CD spectrometer (Applied Photophysics, Surrey, UK) instrument that is equipped with a Peltier that controls temperature was used for analysis. A quartz cuvette (Hellma, Singapore) was used during Spectral scans for each protein (2 µM) recorded from 250 and 190 nm. Derived spectra were then deconvoluted using Dichroweb server (http://dichroweb.cryst.bbk.ac.uk) to determine α-helical, β-sheet, β-turn,

and unordered regions. Secondary structure at 22 nm was monitored to determine the thermal stability of each protein. Each protein's folded fraction was expressed as a molar residue ellipticity (deg.cm^2.dmol^{-1}). Equation 1 was used to determine the folded state of each protein.

$$((\theta)t - (\theta)h) / ((\theta)l - (\theta)h) \quad (1)$$

where (θ)t represents the molar ellipticity at a given temperature, (θ)h is the highest temperature, and (θ)l is the lowest temperature (Zininga *et al*., 2015).

2.9 Tryptophan fluorescence spectroscopic analysis of DnaK and DnaK-G

The organization of the tertiary structure of DnaK and DnaK-G was determined by tryptophan fluorescence spectroscopy. The assay was conducted as previously described but with minor modifications (Zininga *et al*., 2016). DnaK and DnaK-G (2 µM) recombinant proteins were incubated in the presence of varying concentrations (0-6 M) of denaturant Guanidine hydrochloride (GHCl) in Tris buffer (10 mM Tris-HCl pH 7.5) at 20 ºC for 1 hour. The assay was repeated in the presence of nucleotides, ATP, and ADP. Tertiary structure changes of DnaK and DnaK-G was analyzed in the presence of varying concentration (0-500 nM) Hsp70 substrate peptides, ANNNMYRR, and ALLLMYRR (Mabate *et al*., 2018; Lebepe *et al*., 2020; Makumire *et al*., 2021). Peptides ALLLMYRR have been reported to be derived from a chicken mitochondrial precursor, aspartate aminotransferase (Goeckeler *et al*., 2008). Leucine was substituted with asparagine in ALLLMYRR peptides to form ANNNMYRR peptides (Mabate *et al*., 2018). Fluorescence emission spectra were measured between 300 nm and 450 nm at 295 nm excitation using Jasco FP-6300 spectrofluorometer (JASCO, Tokyo, Japan). Data was recorded in triplicates.

2.10 Complementation assay

A complementation assay has been used in the past to study the function of Hsp70 (Makhoba *et al*., 2016). A complementation assay was conducted to investigate the role of PfHsp70-1 GGMP repeat residues *in cellulo* as previously reported (Shonhai *et al*., 2005; Makumire *et al*., 2021). Three different types of strains were used, *E. coli dnaK756* (BB2362, *E. coli dnaK756* recA::Tc^R pDMI,1) which expresses DnaK that has low GrpE affinity and its ATPase basal activity (Buchberger *et al*., 1999). *E. coli dnaK756* is resistant to bacteriophage lambda (Georgopoulos, 1977). *E. coli dnaK103* (BB2393/C600 *dnaK103* thr::Tn10::Tet^R) express truncated DnaK which is non-functional due to the amber mutation on the *dnaK* gene.

Moreover, *E. coli dnaK103* expresses enough DnaJ levels to meet physiological requirements (Spence *et al*., 1990; Suppini *et al*., 2004). *E. coli* Δ*dnaK52* (BB1553, MC4100 *ΔdnaK52*::Cm^R sidB1) expresses a dnaK gene that is interrupted by *cat* cassette resulting in *E. coli ΔdnaK52* cells exhibiting thermosensitivity (Paek and Walker 1987).

Using a similar approach, *E. coli dnaK756* cells were transformed with the following constructs pQE30/DnaK, pQE60/DnaK and mutants pQE60/KPf, pQE60/$KPf_{\Delta G}$, pQE60/$KPf_{617\text{-}647}$ and pQE30/DnaK-G (Shonhai *et al*., 2005). A single colony was inoculated into 5 ml 2x YT broth and then incubated at 37°C with shaking overnight. The overnight culture was transferred to fresh 45ml 2x YT broth and then incubated under the same experimental conditions until OD=0.6 was reached. Double strength yeast tryptone (2x YT) broth was supplemented with 10 µg/ml tetracycline, 50 µg/ml kanamycin, and 100 µg/ml ampicillin. At mid-log phase OD=0.6, cells were induced with 1mM IPTG and allowed to grow until OD=2. Cell cultures were standardized to 0.2 and then serial dilutions were done to each culture before spotting on an agar plate that had 50 µM IPTG, 10 µg/ml tetracycline, 10 µg/ml kanamycin, and 100 µg/ml ampicillin. Plates were incubated at permissive temperature 37°C and non-permissive temperature 40°C overnight. The same approach was followed for *E. coli dnaK103* cells except that cells were selected using 34 µg/ml chloramphenicol, 10 µg/ml tetracycline, and 100 µg/ml ampicillin. Furthermore, cells were incubated at 30 °C.

2.11 Co-affinity chromatography

The effect of GGMP residues of Hsp70 on its interactome in an *E. coli* system was investigated using Pierce™ Co-affinity chromatography kit (Thermo Scientific, USA) as per the manufacture's instruction with minor modification (Figure 2.2). HisPur cobalt resin beads (50 µl) were added into a Pierce spin column. Beads were equilibrated using 400 µl wash buffer (1:1 TBS: Pierce lysis buffer and 10 mM imidazole) and centrifuged at 1250 xg for 1 minute, wash step was repeated 5 times. Purified recombinant proteins (Figure S3) (bait) (pQE30/DnaK, and pQE30/DnaK-G) were added into the Pierce spin column containing beads and incubated for 4 hours with a gentle rocking motion. Pierce spin columns were centrifuged at 1250 xg for 1 minute and then followed by 5 times washes as previous wash step. *E. coli dnaK756* cells were used as prey, in terms of preparation, *E. coli dnaK756* cells were expressed and harvested according to standard protocol. In terms of preparation, cells were lysed with lysis buffer for 10 minutes (Zininga *et al*., 2015). Following the lysis step, cytosol was obtained by centrifugation at 5000 xg for 20 minutes. Tris-buffered saline (TBS) (200 µl) (50 Mm Tris, 150 mM NaCl, pH 7.5) and 200 µl of the pierce lysis buffer were added to the supernatant

fraction (prey proteins) to standardize and then incubated for 1 hour on ice with gentle rocking. After 30 minutes of incubation, lysis buffer containing 2 µg/ml of prey protein was added to the Pierce spin column containing beads and bait proteins followed by incubation overnight at 4 °C. After overnight incubation, 4% formaldehyde was added to make weak bonds between interacting partners stronger which is then followed by 7 minutes incubation before quenching the reaction by adding 10mM glycine for 3 minutes. Pierce spin columns were centrifuged at 1250 xg for 1 minute and then followed by 5 washes as per the previous wash step. Flow-through and washes were collected for SDS-PAGE analysis. Prey-bait proteins were eluted with 200 µl elution buffer (pierce wash buffer + 290 mM imidazole). Elutions were collected for SDS-PAGE analysis. As a negative control, non-treated beads were incubated with prey proteins and eluted with the same elution. Gels were stained using Pierce silver stain and then unique protein bands were cut and sent for Liquid chromatography-mass spectrometry (LCMS) analysis. Illustrated below is the protocol on how the assay was conducted (Figure 2.3).

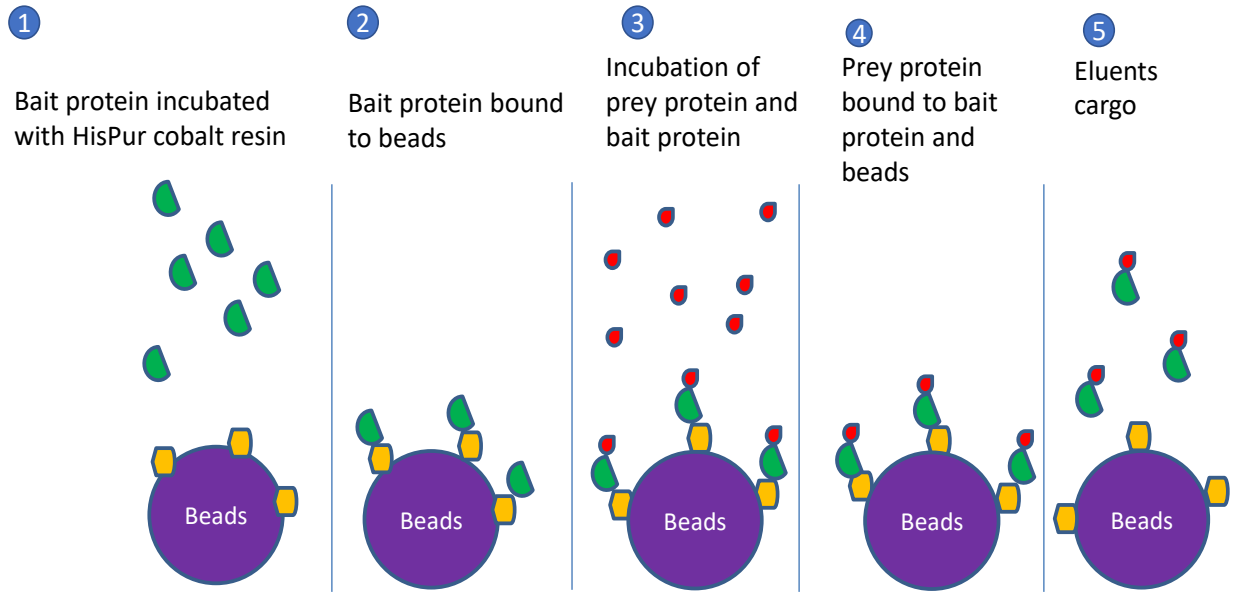


Figure 2.3: Co-affinity chromatography schematic diagram

HisPur cobalt resin beads were incubated with bait protein to allow maximum binding, when bait protein and HisPur cobalt resin beads are bound together, prey protein was added and incubated to allow binding of prey and bait protein. Bound prey-bait proteins cargo is then eluted for analysis.

2.12 Proteomics determination of DnaK interactors using liquid chromatography mass spectrometer (LCMS)

Unique bands that were cut from co-affinity chromatography SDS-PAGE were analyzed using LCMS. Gel pieces were destained with 50/50 200 mM NH_4HCO_3:Acetonitrile (Sigma, USA) until they were clear. Samples were dehydrated before reduction at room temperature for 15 minutes with 2 mM triscarboxyethyl phosphine (TCEP; Fluka) in 25 mM NH_4HCO_3. 20 mM S-Methyl methanethiosulfonate (Sigma, USA) was used to thiomethylate cysteine residuals for 30 minutes at room temperature. Gel pieces were then dehydrated and washed with 25 mM NH_4HCO_3. After dehydrating gel pieces, proteins were then digested at 37 ºC with trypsin (Pierce) solution (20 ng/µL). Extraction of peptides was performed using 50 µL water and 50% acetonitrile. Samples were resuspended immediately after drying them down in 30 µL 2% acetonitrile: water; 0.1% FA. An in-house manufactured C_{18} stage tip (Empore Octadecyl C_{18} extraction discs; Supelco) was used to remove residual digest reagents. The C_{18} membrane was activated with 30 µL methanol (Sigma, USA) and equilibrated with 30 µL 2% acetonitrile: water; 0.05% TFA before sample loading to the stage tip. The bound sample was washed with 30 µL 2% acetonitrile:water; 0.1% TFA before elution with 30 µL 50% acetonitrile:water 0.05% TFA. The eluate was dried by evaporation. Peptides were prepared for LC-MS analysis by dissolving them into 2% acetonitrile: water; 0.1% FA.

Trap column 200mm x 100µm C_{18} (Thermo Scientific) and CSH 25 cm x 75 µm 1.7µm particle size C_{18} column (waters) analytical column equipped on Thermo scientific ultimate 3000 RSLC was used to perform liquid chromatography. 2% acetonitrile: water; 0.1% FA; Solvent A: 2% acetonitrile: water; 0.1% FA and Solvent B: 100% acetonitrile: water was loaded by the solvent system employed. Loading solvent was used to load samples on the trap column at 2 µL/min rates from a temperature-controlled autosampler at 7ºC. Samples were loaded for 5 minutes before elution onto the analytical column, 2%-30% from 5 minutes to 1 hour 5 minutes, and then increased to 30-50% from 1 hour 5 minutes to 1 hour 20 minutes. mass spectrometer received the outflow performed for 45°C at the chromatography.

Thermo Scientific Fusion mass spectrometer equipped with a Nanospray Flex ionization source was used to perform mass spectrometry. The stainless steel was used to inject a sample. Spray voltage set at 1.8kV and ion transfer capillary set at 280°C were used to collect data in positive mode. Polysiloxane was used to calibrate spectra internally by m/z = 445.12003 and 371.10024. Orbitrap detector set at 120 000 resolution was used to perform MS scans with AGC target at

E5 and maximum injection time of 40 MS over the scan range of 350-1650. Data were obtained in centroid mode.

Proteome discoverer v1.4 (Thermo Scientific) was used to analyze raw data generated by the spectrometer Sequest algorithm. Interrogation of the database was performed against a concatenated database created using the UniProt *E. coli* with the cRap contaminant database. Precursor and fragment mass tolerance were set to 10ppm and 0.02 Da, respectively. Oxidation (M), Demamidation (NQ), and acetylation of N-terminus were allowed as dynamic modifications and thiomethyl of C as static modification. To perform peptide validation, a Target-decoy PSM validator node was used. Data was viewed and analyzed using Scaffold (Proteome Software, Inc., Portland, OR 97219, Oregon, USA) derived from MS/MS sequencing results. Scaffold verifies peptide identifications assigned by Proteome discoverer using the X! Tandem database searching program (Craig, 2003; Searle, 2008). Scaffold then probabilistically validates these peptide identifications using PeptideProphet (Keller, 2002) and derives corresponding protein probabilities using ProteinProphet (Nesvizhskii, 2003, Searle, 2010).

3.0 Results

3.1 Cytosolic Hsp70s from Apicomplexa species predominantly possess GGMP residues

To compare the occurrence of the GGMP residues, various Hsp70s were aligned using Bioedit. GGMP residues occur mostly in cytosolic Hsp70s of parasites that of Apicomplexa and Kinetoplastids origins. PfHsp70-1 possesses 7 GGMP residues while *T. gondii*, *Cryptosporidum parvum*, *T. cruzi* Hsp70s have six, eleven, and ten respectively (Figure 3.1). Notably, amongst kinetoplastids, *Leishmania donovani* Hsp70s possess only two GGMP repeats. Moreover, *Cyclospora cayanensis* Hsp70 from kinetoplastids origin organism contains few copies number of GGMP residues with only 3 copies. Other cytosolic localized Hsp70s of parasitic organisms such as *Entamoeba histolica*, and Coccodian, possess 2 GGMP residues. Only two non-parasitic organisms cytosolic localized Hsp70s namely, *Homo sapiens* and *Felis catus* possess 2 GGMP residues in their Hsp70s. Makumire and co-authors (2021) recently reported that the GGMP residues occur mostly in apicomplexa organisms. This study supports and confirms the same findings.

Figure 3.1.1: Multiple sequence alignment of the substrate binding domain of various Hsp70s

Multiple sequence alignment of *Leishmania donovani* (UniprotKB: P17804); *Trypanosoma cruzi* (Uniprot: Q4DTM8); *Trypanosoma brucei* HSP74 (UniprotKB: P11145); *Toxoplasma gondii* Hsp70 (UniprotKB: A0A125YXI9); *Plasmodium falciparum* 3D7 Hsp70 (PfHsp70-1, NCBI accession number: XP_001349336.1);

Plasmodium ovale curtisi (GenBank: SBS81157.1); *Plasmodium vivax* (NCBI: XP_001614972.1); *Plasmodium knowlesi* strain H (NCBI: XP_002258136.1); *Plasmodium malariae* (Genbank: XP_028860418.1), *Cryptosporidium parvum Iowa II* (NCBI: XP_625373.1); *Cyclospora cayetanensis* (NCBI: XP_022588287.1); *Giardia lamblia* ATCC 50803 (GenBank: EDO80296.1); *Entamoeba histolytica* (GenBank: AAA29102.1); *E. coli* Hsp70 (DnaK, UniProtKB: A1A766.1); *Schistosoma haematobium* (Uniprot: AOA6A5DLH8); *Candinda albicans* (Uniprot: Q24896) *M. tuberculosis* Hsp70 (UniProtKB: P9WMJ9.1); *Saccharomyces cerevisiae* HSP70 (Ssa1p, GenBank: AAC04952.1); *E. coli* Hsp70 (DnaK, UniProtKB: A1A766.1); Trichonomas vaginalis (Uniprot: A2EPF1); Homo sapiens heat shock cognate 71 kDa (Hsc70, NCBI: NP_006588.1); HspA9 (NCBI, NP_004125.3); Homo sapiens heat shock protein 70s (HspA14; NCBI: NP_057383.2); HspA12B (NP_001317093.2); HspA2 (NCBI: NP_068814.2); HspA12A (NCBI: NP_001317093.1); HspA6 (NCBI: NP_002146.2); HspA4 (NCBI: NP_002145.3); HspA13 (NCBI: NP_008879.3); HSP70 (Ssa1p, GenBank: AAC04952.1); *Felis catus* (Uniprot: A0A2I2UZMU); *Arabiodopsis thaliana* (Uniprot: Q9CHA8); *Rattus norvegicus* (Uniprot: P0DMW1) and *Saccharomyces cerevisiae* (Uniprot: P40150). The red rectangle represents the region where GGMP residues are located. Numbering of residues is based on HspA12A.

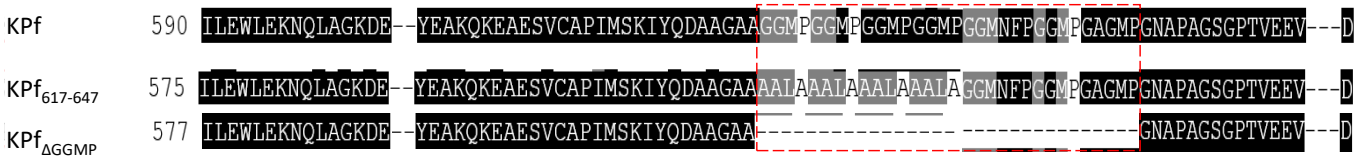


Figure 3.1.2: C-terminus sequence alignment of KPf derivatives

Mutations in KPf were generated by conservative amino acids substitution of GGMP residues to create $KPf_{617-647}$ and deletion of the whole residue to create $KPf_{\Delta GGMP}$.

The conservation of Hsp70s across aligned cytosolic Hsp70s from different species that possess the GGMP residues was compared. Sequence percentage identity matrix was analyzed after alignment using ClustalW. There is high conservation across cytosolic Hsp70s that possess the GGMP residues. For example, Human cytosolic Hsp70 and *P. falciparum* Hsp70 have identity scores of 73.02%. The identity percentage score for all cytosolic Hsp70s aligned in this study is greater than 57%. This implies that cytosolic Hsp70s that possess the GGMP residues are generally closely related and may perform similar functions. Furthermore, the distinct function of these proteins may also be accounted to the number of copies of GGMP residues they possess.

Table 3.2. Percentage similarity comparison of selected Hsp70s that possess the GGMP residues

	G. Lamblia	*E. Histolyca*	*F. catas*	*Leshmania*	*P. falciparum*	*Cyclospora*	*Coccodian*	*H. sapien*
G. Lamblia	-	60.96	65.26	61.65	59.15	60.96	57.48	65.78
E. Histolyca	60.96	-	72.32	71.27	66.62	70.28	71.92	73.64
F. catas	65.26	72.32	-	73.01	73.49	72.88	71.48	73.20
L. denovani	61.65	71.27	73.01	-	71.01	71.88	70.35	72.02
P. falciparum	59.15	66.62	73.49	71.01	-	77.01	77.35	72.83
C. cayetansis	60.96	70.28	72.88	71.88	77.01	-	90.87	73.33
C. albicans	57.48	68.92	71.48	70.51	77.35	90.87	-	99.85
H. sapien	65.78	73.64	73.20	72.83	73.02	73.33	99.85	-

G. Lamblia; Giardia lamblia, *E. Histolyca*; *Entamoeba histolyca, F. catas*; *Felis catas, L. denovani*; *Leshmania denovani, P. falciparum*; *Plasmodium falciparum, C. albicans*; *Coccodian albicans, Homo sapiens.*

3.2 Confirmation of pQE60/KPf and pQE60/DnaK-G plasmid variants

Restriction digest and DNA sequencing were used to verify pQE60/KPf, pQE60/KPf$_{\Delta G,}$ and pQE60/KPf$_{617\text{-}647}$ plasmids (Figure 3.2A-E). Plasmid pQE60/KPf expressing chimeric KPf protein was digested with *Bam*HI and *Hind*III. Plasmid pQE60/KPf$_{\Delta G}$ expressing a mutant version of KPf with a complete deletion of GGMP residues was confirmed by restriction digest using *Hind*III and *Bam*HI. A plasmid (pQE60/KPf$_{617\text{-}647}$) expressing KPf mutant with GGMP residues substituted conservatively was confirmed using *Nco*I and *Hind*III enzymes. A single digest with either of the enzymes resulted in a linearized plasmid of 5312 bp and 5393 bp for pQE60/KPf$_{\Delta G}$ and pQE60/KPf$_{617\text{-}647}$, respectively. Double digest of pQE60/KPf$_{\Delta G}$ with both enzymes resulted in two separate bands at around 3431 bp and 1906, which corresponds to

pQE60 and KPf$_{\Delta G}$ insert. Restriction of pQE60/KPf$_{617-647}$ with both *Nco*I and *Hind*III enzymes resulted in two bands, 3431 bp and 1995 bp corresponding to pQE60 and KPf$_{617-647}$. DnaK and KPf were also confirmed using the *Hind*III enzyme resulting in 5323 bp and 5395 bp respectively (Figure S1). Plasmids integrity of pQE30 and pQE60 vector were confirmed using *Hind*III. Digestion of each vector using *Hind*III resulted in a linearized plasmid of 3431 bp and 3461 bp respectively.

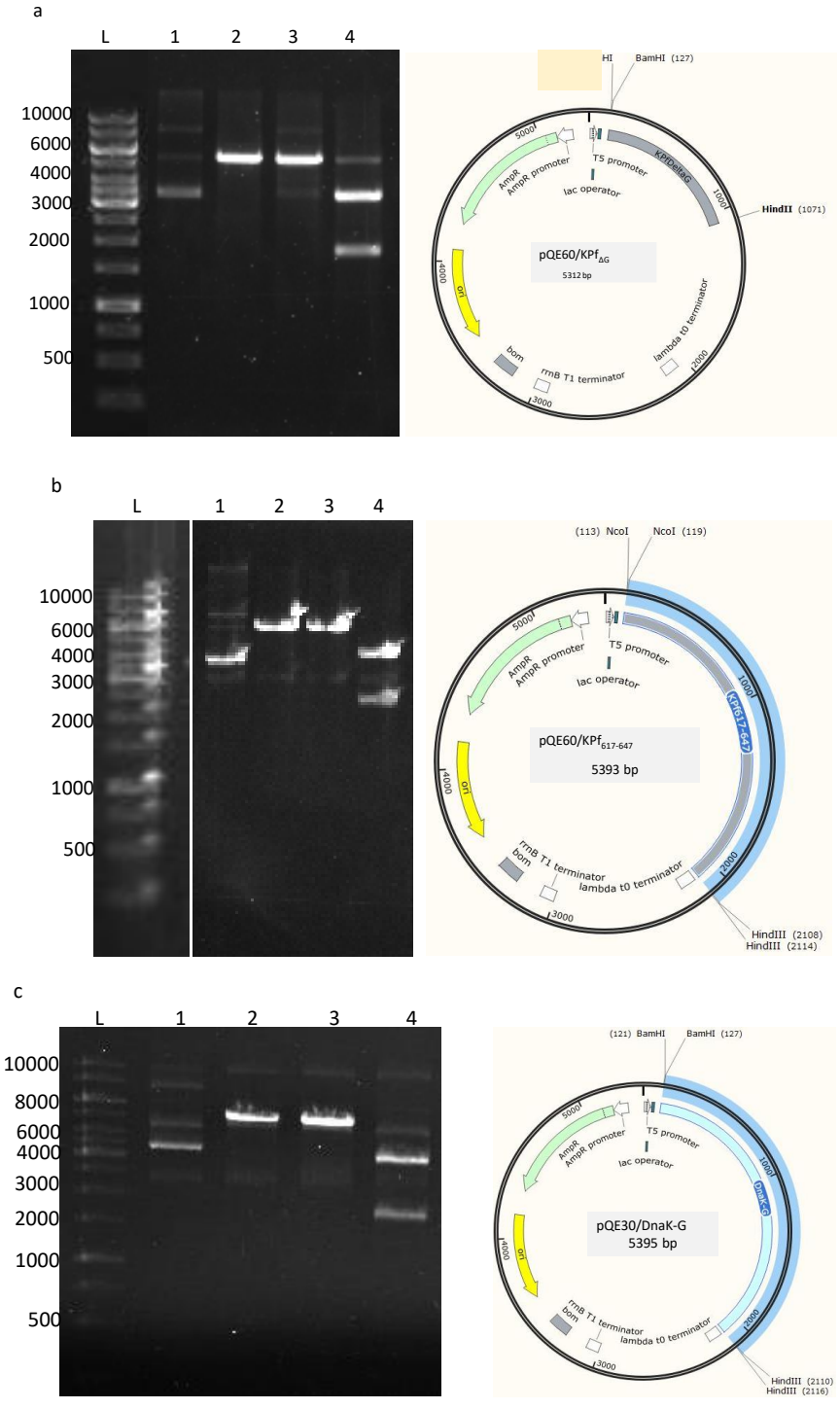


Figure 3.2: Agarose gel analysis of pQE60/KPf$_{\Delta G}$, pQE60/KP$_{617-647}$, pQE30/DnaK-G plasmids, and their plasmid maps

Analysis of KPf variants DNA by restriction digest. (a) pQE60/*KPf*$_{\Delta G}$ and (b) pQE60/*KPf*$_{617-647}$ agarose gel and plasmid map showing restriction sites. pQE60/KPf$_{617-647}$ restriction sites cut with the following enzymes *Nco*l and *Hind*III while pQE60/*KPf*$_{\Delta G}$ restriction sites cut with *Bam*HI and *Hind*III enzymes. Lane L; represents DNA ladder, lane 1; uncut DNA, lane 2; single digest with one of the respective enzymes, lane 3; single digest with the second enzyme of respective DNA and 4; double digest with both enzymes. Full agarose gels for cropped gels are also provided (Appendix A1.1).

3.3 Expression and purification of DnaK and DnaK-G

SDS-PAGE analysis reveals that recombinant DnaK and DnaK-G proteins were expressed (Figure 3.3A and B) and purified successfully (Figure 3.3C and D). The presence of expressed and purified proteins was confirmed by Western blot analysis using an anti-his antibody (bottom panel). Both DnaK (Figure 3.3A) and DnaK-G (Figure 3.3B) proteins migrated at 71 kDa and 72 kDa respectively as observed in SDS-PAGE. Both proteins were expressed in bulk upon IPTG induction (1-6 hours). Both proteins were not expressed before induction (Figure 3.4A and B, lane 0). pQE30 plasmid was used as a control (Figure S4).

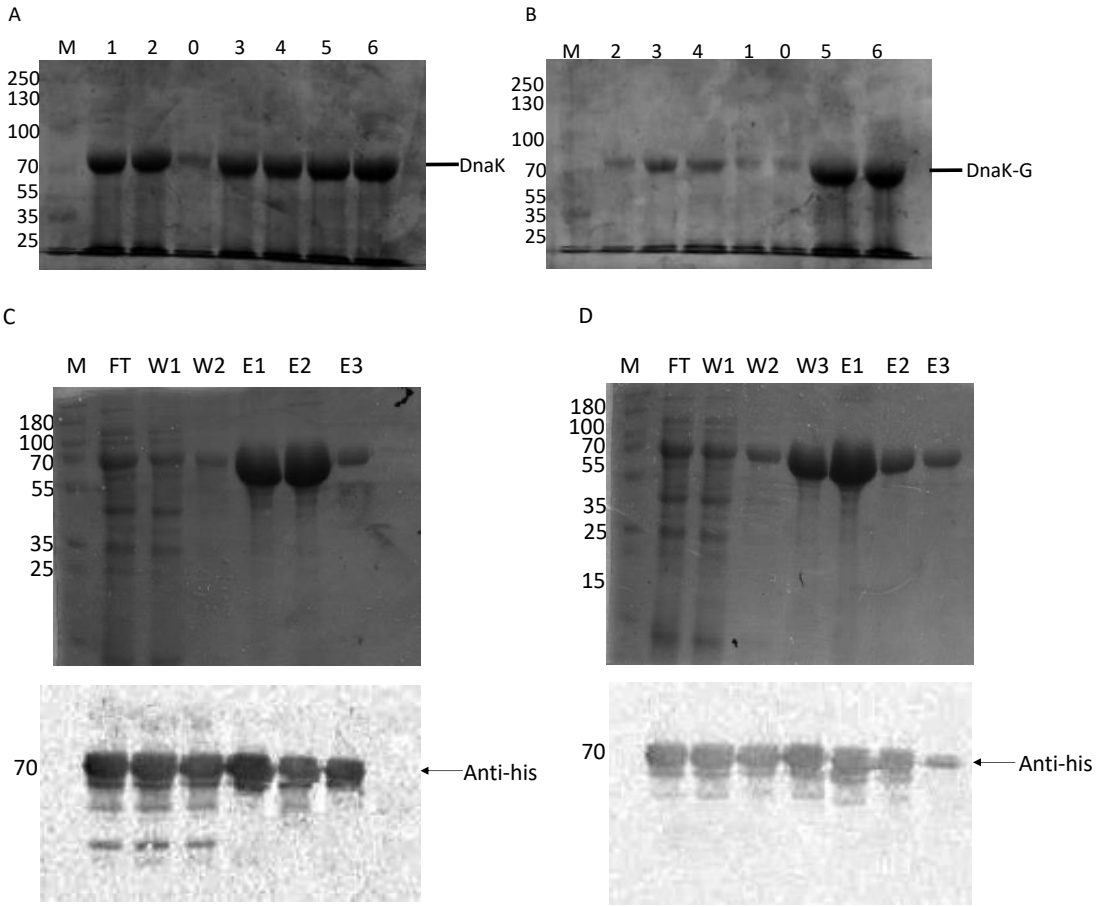


Figure 3.3: DnaK and DnaK-G expression and purification

Expression of DnaK (a) and DnaK (b) in *E. coli JM109 cells* analyzed by SDS-PAGE. Proteins were further purified, DnaK (c) and DnaK-G (d). The following lanes represent, Lane M; protein marker (kDa), Lane 0; pre-induction sample, Lane 1-6; hourly samples after induction with IPTG. Both proteins were further confirmed by Western blot using anti-his antibody (bottom panel). The following lanes represent, Lane M; protein marker (kDa), Lane FT; flow-through of non-binding proteins; Lane W1; samples washed with lysis buffer, Lane W2-W3 samples washed with wash buffer, Lane E1-E3 proteins eluted with elution buffer, and anti-his: represents antibody used to detect DnaK and DnaK-G proteins.

3.4 GGMP residues appear important for the structural integrity of Hsp70

Three-dimensional models of KPf and DnaK relative to those of their derivatives were superimposed and analyzed using Chimera version 1.15. KPf and $KPf_{617\text{-}647}$ exhibit no apparent variation (Figure 3.4A). Furthermore, the conservative substitution of the GGMP residues did not result in the reorientation of $KPf_{617\text{-}647}$. On the other hand, KPf and $KPf_{\Delta G}$ exhibited notable predicted structural variation (Figure 3.4B). KPf possesses an extended loop (highlighted in cyan) in the SBDα as compared to $KPf_{\Delta G}$. $KPf_{\Delta G}$ exhibits a coiled region (highlighted in green) stabilized by 52 hydrogen bonds as a result of the loss of the extended C-terminus loop (Figure 3.1B). This coiled region led to the reorientation of the EEVD motif in the C-terminus. A previous study reported variations within the loops of the SBDβ of DnaK-G as a result of the insertion of GGMP residues (Makumire *et al*., 2021). Moreover, introduced residues did not alter the hydrogen bonding of DnaK. Furthermore, DnaK-G lost a section of its α-helical loop from ^{608}QHAQ-AKDD^{645}, which is part of the SBDα (Figure 3.1C). The insertion of GGMP residues led to DnaK having an extended loop in the C-terminus as compared to the wild-type protein. Findings from this study support previously reported findings that the insertion of GGMP affects the conformation of DnaK-G SBDβ (Makumire *et al*., 2021). Furthermore, the conformational integrity of the DnaK lid is predictably changed by the insertion of the GGMP residues.

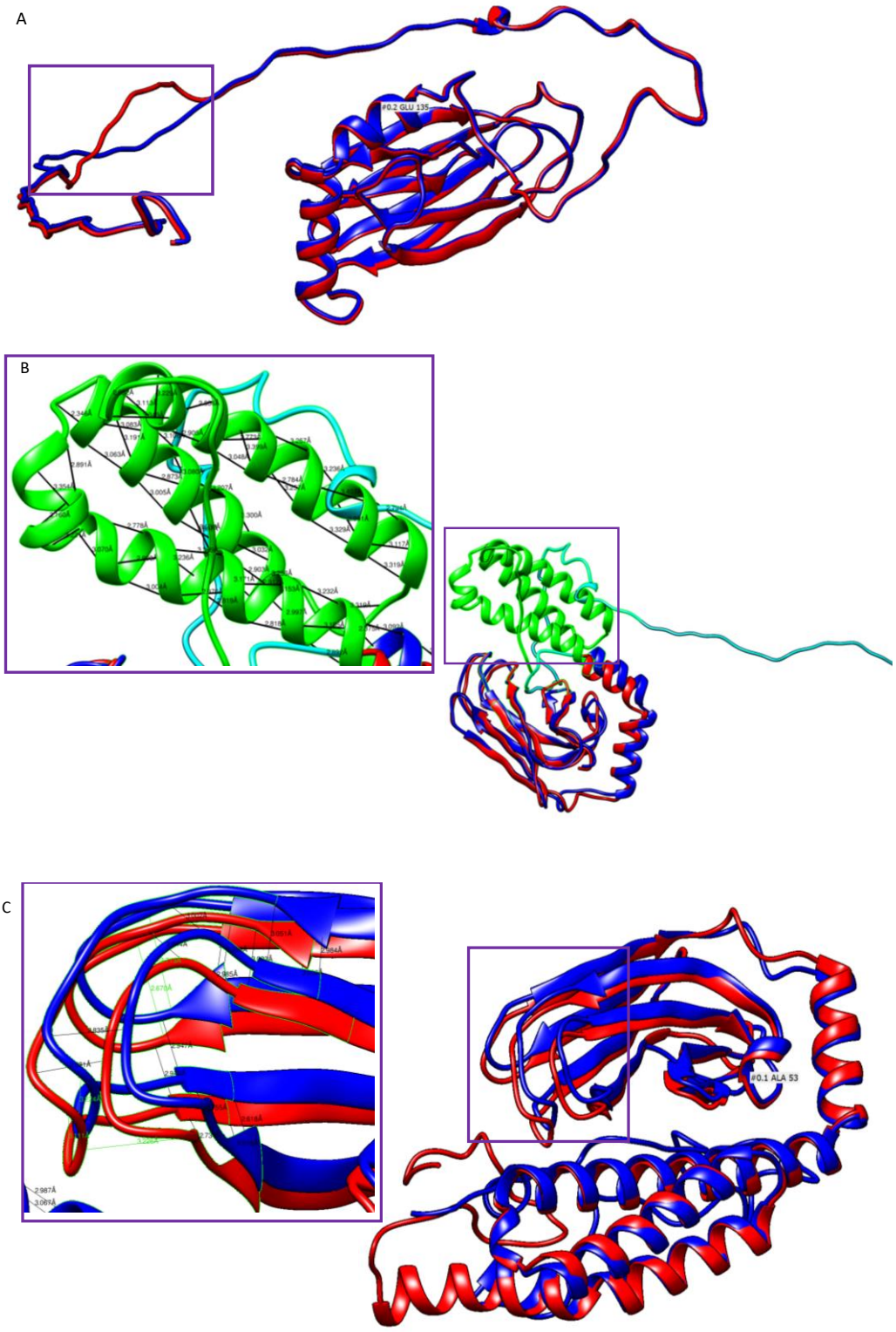


Figure 3.4: Superimposed 3D structure of DnaK, KPf, and their GGMP mutants

Superimposed 3-dimensional structure models of DnaK, KPf, and their GGMP mutants substrate binding domain. a) KPf (red) versus $KPf_{617-647}$ (blue), b) KPf (red) versus $KPf_{\Delta G}$ (blue) and c) DnaK (red) and DnaK-G (blue).

3.5 GGMP residues do not augment DnaK stability

DnaK and DnaK-G secondary structures were determined using CD at wavelengths between 260 and 190 nm. Positive peaks were exhibited at around 194 nm and negative peaks at around 208 nm and 222 nm respectively (Figure 3.5A). The results obtained by CD spectroscopy suggest that DnaK is predominantly an α-helical protein as previously reported (Lebepe *et al*., 2020). The observed α-helical content of DnaK is similar to that previously reported (Lebepe *et al*., 2020: Makumire *et al*., 2021). On the other hand, the DnaK-G CD spectrum had shallow peaks at 208 and 222 nm. A previous study reported that the insertion of GGMP into DnaK led to a shallower CD spectrum (Makumire *et al*., 2021). It was further reported that the shallower CD spectrum suggests less α-helical content compared to DnaK (Makumire *et al*., 2021).

The effect of the inserted GGMP residues in regulating the thermal stability of DnaK was determined using tryptophan fluorescence spectroscopy. Both DnaK and DnaK-G were subjected to heat stress, there were no distinct conformational changes between DnaK and DnaK-G as both proteins were able to maintain their secondary structure at around 65°C (Figure 3.5B). Deconvoluted data sets confirmed that both DnaK and DnaK-G possess more than 50% α-helical content. Furthermore, both DnaK and DnaK-G lost their α-helical structure at around 80°C (Table 3.2). However, at 20°C to 75°C, both proteins maintained their α-helical content. At 80°C, both proteins lost their α-helical content from 80% to around 35%. The RMSD shows that the data is reliable as it is less than 1. The RMSD scores for DnaK and DnaK-G at 25°C are the same which symbolizes that there is no structural variation between the two proteins (Kufareva and Abagyan, 2012). This may suggest that the observed shallower curve in DnaK-G did not adverse its heat stability.

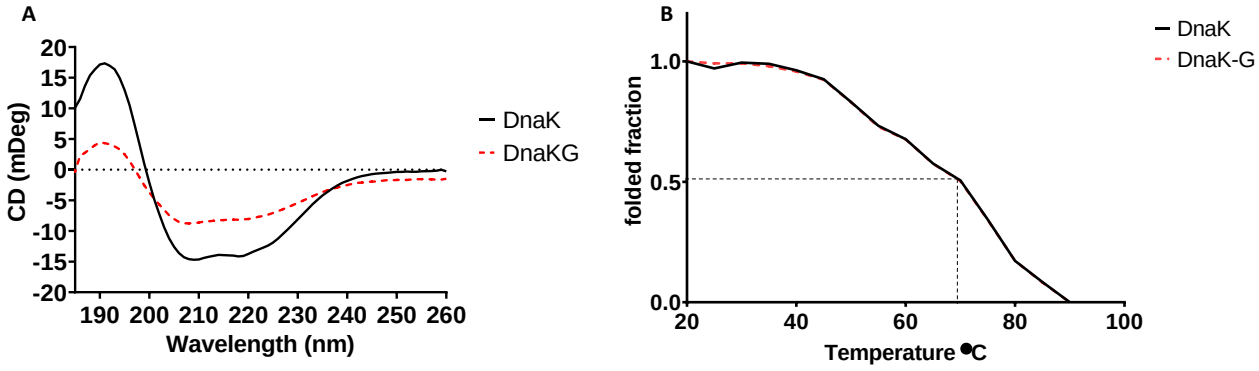


Figure 3.5: The GGMP residues did not modulate the thermal stability of DnaK

a) CD spectra of DnaK and DnaK-G; (B) Thermal Stability of DnaK and DnaK-G as determined by tryptophan fluorescence spectroscopy.

Table 3.2 Secondary structure comparison of the composition of DnaK and DnaK-G at 25℃

	α-helix	β-sheet	β-turns	Unordered	RMSD
DnaK	80.6	1.0	6.4	15	0.647
DnaK-G	80.6	1.0	6.4	15	0.647

RMSD: root mean square deviation. Represented data sets are obtained from the Dichroweb server with CDSSTR reference set 4. Two-way ANOVA *($p<0.001$)* was used for statistical analysis.

The role of GGMP repeat residues on the conformation of DnaK and DnaK-G was further investigated using tryptophan fluorescence as previously described (Zininga *et al*., 2016). DnaK has been reported to possess one (1) tryptophan residue in the NBD (W102). GHCL is a known protein denaturant (Makumire, 2020). Incubation of DnaK with ADP resulted in no change in the emission spectra. On the other hand, the presence of ATP decreased the emission spectra to 332 nm from 336 nm in the absence of nucleotides (Figure 3.5A). DnaK-G was conformationally unresponsive to either of the nucleotides hence the emission spectra were unchanged (Figure 3.5B). As a control, DnaK and DnaK-G were both denatured by GHCL in a concentration dependent manner, as both proteins exhibited red spectral shifts as the concentration of GHCL was increased (Figure 3.5C-D). Emission maxima in the absence of GHCL were recorded at around 333 nm and 338 nm for DnaK-G and DnaK respectively. Emission maxima increased as the concentration of GHCL was increasing as observed in 6M GHCL where the emission maxima red shift to 347 nm and 345 nm for DnaK-G and DnaK respectively. Observed findings suggest that the insertion of GGMP residues results in DnaK being sensitive to GHCL. The experiment was further performed in the presence of peptide substrates, ALLLMYRR and ANNNMYRR respectively. A blue shift was observed for DnaK in the presence of ALLLMYRR peptides substrates. The emission spectra shifted by 2 nm in the absence of peptide substrates to 125 mM peptide substrate concentration (Figure 3.5E). The emission spectra then leveled off in the presence of 125 mM to 1000 mM of the peptide. The insertion of GGMP residues in DnaK did not drastically change the binding affinity of DnaK-G to ALLLMYRR as there was a blue shift in the emission spectra as observed in wild-type DnaK (Figure 3.5F). Observed findings may suggest that the GGMP residues do not prefer binding to substrates that are enriched with hydrophobic residues. A blue shift in the emission spectra was also observed for DnaK in the presence of ANNNMYRR. DnaK emission spectra

shifted in concentration dependent manner. The insertion of GGMP residues in DnaK led to the reduced affinity of DnaK to bind to ANNNMYRR peptides as no change was observed in the emission spectra even as the concentration of peptide substrates was increased (Figure 3.5G-H). The drop in affinity of DnaK-G to bind to ANNNMYRR may suggest that DnaK does not need the GGMP residues in binding peptides.

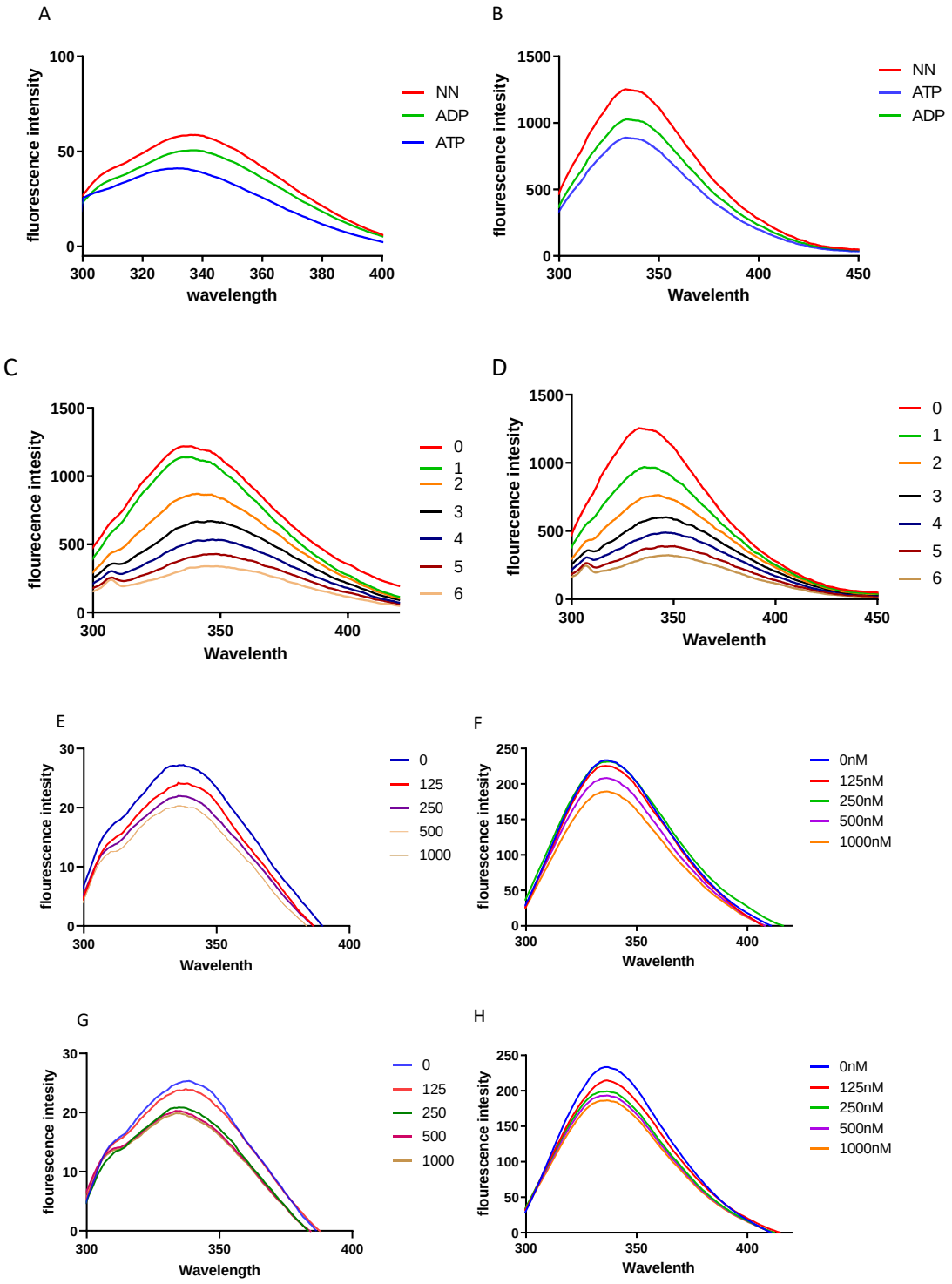


Figure 3.5: The insertion of the GGMP residues did not improve the stability of DnaK

Tryptophan fluorescence emission spectra analysis of (a) DnaK and (b) DnaK-G emission spectra in the presence of nucleotides, ATP, and ADP. (c) DnaK and (d) DnaK-G spectra emission changes in the presence of varying concentrations of GHCL. Conformational shifts in DnaK (e) and DnaK-G (f) in the presence of substrate peptides ALLLMYRR in varying concentrations. (g) DnaK and (h) DnaK-G conformational change in the presence of substrate peptides ANNNMYRR in varying concentrations. Fluorescence intensity was measured in a.u unit.

3.7 Deletion and substitution of the GGMP repeat residues abrogated the *in cellulo* function of KPf

To elucidate the role of GGMP residues *in cellulo* using DnaK, KPf, and KPf derivatives, complementation assay was conducted as previously described (Shonhai *et al*., 2005; Makhoba *et al*., 2016). KPf is a chimeric protein with DnaK NBD and PfHsp70-1 SBD (Shonhai *et al*., 2005). This chimeric protein has been shown to reverse *E. coli dnaK756* strain thermosensitivity. The current study, KPf with GGMP mutations was used to study the role of GGMP residues in Hsp70 *in cellulo* function. Various plasmids were transformed in *E. coli dnaK756* cells which were then subjected to a permissive temperature of 37°C (Figure 3.6A). *E. coli dnaK756* cells were able to grow at permissive temperature irrespective of the transformed plasmid. Cells that were transformed with positive controls (pQE60/DnaK and pQE60/KPf) were able to grow at a non-permissive temperature of 43.5°C. Cells that were transformed with the control plasmid (pQE60/KPf$_{V436F}$), did not grow under non-permissive temperature as expected (Makhoba *et al*., 2016). Cells transformed with pQE30/DnaK-G, pQE60/KPf$_{617\text{-}647,}$ and pQE60/KPf$_{\Delta G}$ did not grow at 43.5°C (Figure 3.6B). The data set suggest that the GGMP residues are important for KPf function in a cellular setting. On the other hand, GGMP residues are detrimental to the chaperone function of *E. coli* DnaK.

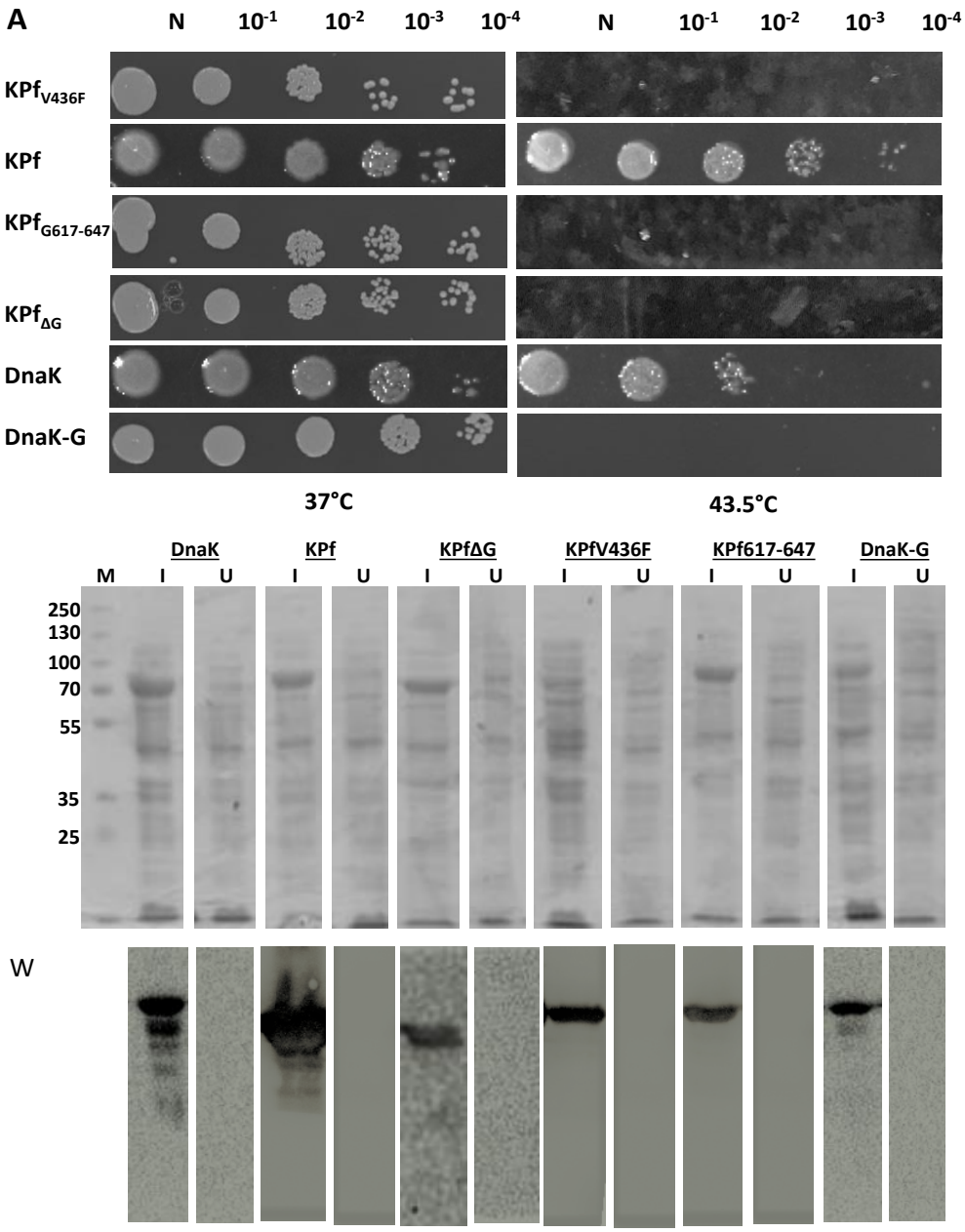
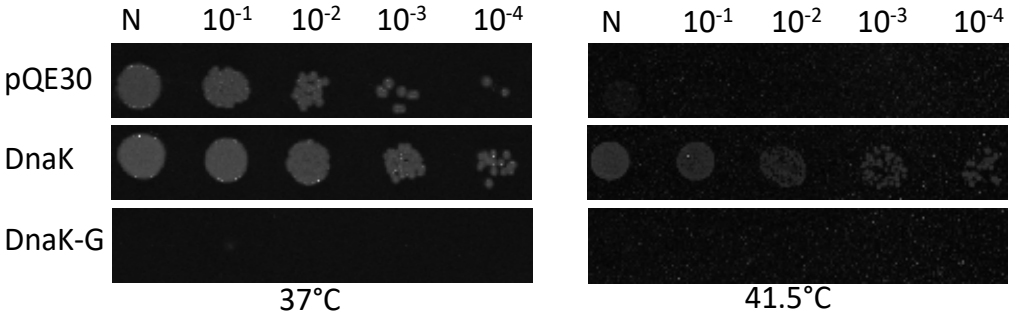


Figure 3.6: Deletion and substitution of the whole GGMP section abrogate the function of KPf *in cellulo*

a) E. coli dnaK756 cells expressing DnaK, KPf, and GGMP mutants were incubated at 37 °C and 43.5 °C. Cells transformed with pQE30/KPf$_{V436F}$ were used as a negative control. (B) Expression of KPf, DnaK, and GGMP mutants was analyzed using SDS-PAGE. Samples taken before induction are represented by "U" while samples that were taken after induction are represented by "I". Molecular marker in kDa is represented by "M". "W" represents Western blot analysis using an anti-his antibody. "N"; undiluted cells, 10^{-1}; 1:10 dilution, 10^{-2}; 1:100 dilution, 1;1000 dilution and 1;10000 dilution.

3.8 Insertion of GGMP repeat residues into DnaK resulted in a protein that was toxic to *E. coli dnaK103* cells

E. coli dnaK103 cells can grow at a permissive temperature of 30°C. Furthermore, because this strain expresses a non-functional DnaK, this strain cannot grow at a temperature above 40°C (Spence *et al*., 1990; Mayer *et al*., 2000; Shonhai *et al*., 2005). Cells transformed with pQE30/DnaK were used as a positive control. DnaK can reverse the thermosensitivity of these cells (Shonhai, 2007). As expected, cells transformed with pQE30/DnaK were able to grow at permissive temperature (37°C) and non-permissive temperature (43.5°C). Cells transformed with pQE30 were used as a negative control. At the permissive temperature, cells transformed with pQE30 managed to grow hence at the non-permissive temperature they did not grow. Although SDS-PAGE shows DnaK-G expression, insertion of the GGMP residues in DnaK resulted in toxicity to the cells at both permissive and non-permissive temperatures (Figure 3.7). The toxicity could suggest that DnaK-G could be failing to fold properly due to the inserted GGMP residues resulting in toxicity. A previous study reported that the absence of DnaK in *E. coli* cells can result in the lack of a folding pathway (Shonhai, 2007).



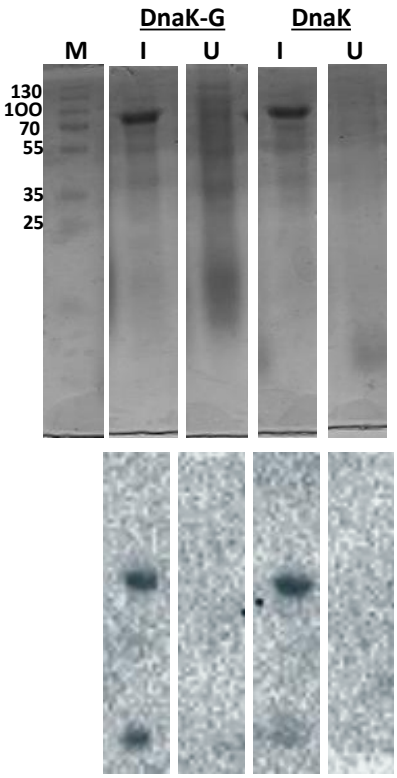


Figure 3.7: GGMP repeat residues insertion to DnaK became toxic to *E. coli dnaK103* cells

(a) Plasmids that express DnaK and DnaK-G were transformed into *E. coli dnaK103* cells and incubated at 37°C and 41.5°C respectively. Cells that were transformed with pQE30 were used as a negative control. (B) Expression of exogenous DnaK and DnaK-G was monitored using SDS-PAGE. Samples taken before induction are represented by "U" while samples taken after induction are represented by I. Molecular marker in kDa is represented by "M". "W" represents Western blot analysis using anti-his antibody. "N" represents samples that were not diluted, 10^{-1}, 10^{-2},10^{-3},10^{-4} is a serial dilution.

3.9 Insertion of GGMP residues in DnaK led to the loss of select DnaK interactors

DnaK was reported to interact with other co-chaperones to properly fold proteins that are aggregating (Shannon *et al*., 2014). This study was conducted as a follow-up to the *in cellulo* study. As a negative control, expressed *E. coli dnaK756* cells and *E. coli* Δ*dnaK52* cells (prey proteins) were passed through magnetic beads. As expected, beads did not passively interact with prey proteins (Appendix B: Figure S2). Purified DnaK protein was used as a positive control and was immobilized onto the beads. Upon DnaK incubation with prey proteins, DnaK was eluted with a distinct set of proteins. Insertion of the GGMP residues resulted in DnaK losing its ability to bind to proteins that the wild type DnaK was eluted with (Figure 3.8). This supports the *in cellulo* study for DnaK-G failing to reverse thermosensitivity of *E. coli dnaK756* cells. The inability of DnaK-G to reverse thermosensitivity of *E. coli dnaK756* cells could be because the protein lost the ability to bind to some substrates and select co-chaperones.

Eluted proteins from co-affinity chromatography were further identified using LCMS and confirmed as DnaK interactors using the String database (http://string-db.org/). In total 418 proteins were identified. Out of 418 proteins, 393 proteins were identified to have interacted with DnaK, and of the latter, 81 proteins were unique (Table 3.2). Moreover, it is unclear if identified proteins interact with DnaK directly or via co-chaperones. Out of these 81 DnaK unique proteins, some were predicted to interact with DnaK using *in silico* approaches (http://string-db.org/:Figure 3.9A, Table 3.2). Among proteins identified to interact with DnaK are its previously known co-chaperones such as GroEL, Clp family proteins, and GrpE and DnaJ. Furthermore, DnaK interacted with ribosomal proteins such as ribosomal RNA small subunit methyltransferase. On the other hand, 25 proteins were identified to be unique DnaK-G interactors as they were not associated with DnaK (Table 3.2). The insertion of the GGMP residues in DnaK led to the loss of some of the ribosomal proteins that interacted with a DnaK. Furthermore, DnaK interacted with ATP hydrolyzing/binding proteins, such as clpP, intermembrane phospholipid transport system, ABC transporter system. On the other hand, DnaK-G did not bind proteins that recognize ATP. Surprisingly, the insertion of GGMP residues into DnaK resulted in DnaK attracting ATP-dependant phosphofructokinase (Table 3.2). Moreover, there were different proteins from various organelles that have different functions that were identified to interact with DnaK such as phosphopentomutase, trysine recombinase, glutamate-cysteine. The insertion of the GGMP residues led to the attraction of proteins that are mostly associated with transcription such as transcription terminator protein, transcription regulator proteins (Table 3.2). This could suggest that the GGMP is involved in attracting proteins that oversee transcription. Some proteins identified to interact with DnaK were also identified in proteins that were bound to DnaK. Moreover, none of the proteins that interacted with DnaK-G are amongst proteins predicted to interact with DnaK. There was a total of 312 proteins identified in both DnaK and DnaK-G. Some of the identified proteins are DnaK co-chaperones such as GrpE, DnaJ, clpB, and malate dehydrogenase (MDH). Consequently, the insertion of the GGMP residues led to the loss of typical predicted DnaK interactors. Furthermore, the insertion of GGMP resides led to the loss of clpP which is known to interact with DnaK in the SBD (Fernández *et al*., 2018). Considering the site of mutation, it is possible that lid repositioning led to the loss of clpP binding to DnaK. The false discovery rate for identified proteins was set to 1% which suggests that 99% of data is correctly identified (Zhang *et al*., 2007).

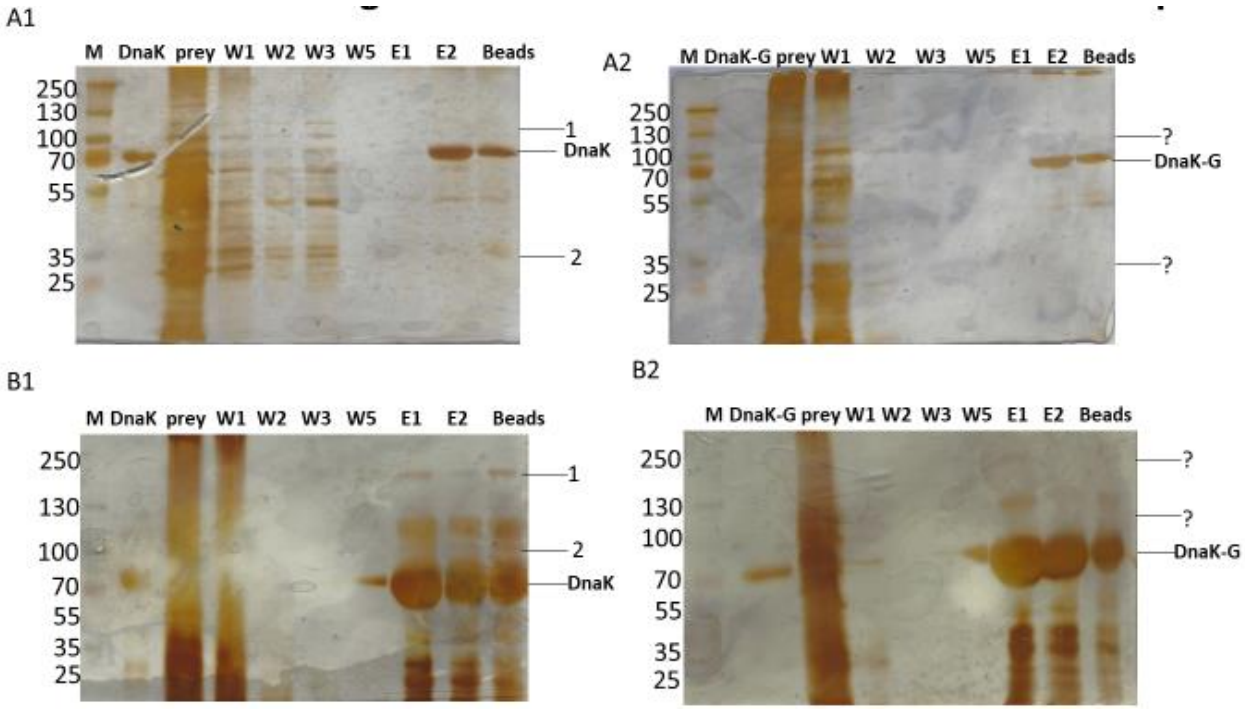


Figure 3.8: DnaK's ability to form a functional network is abrogated by the insertion of the GGMP residues

Purified DnaK and DnaK-G were immobilized on beads and incubated with prey whole lysates of *E. coli dnaK756* (A1-A2) and *E. coli* Δ*dnaK52* cells (B1-B2), respectively. W1-W5 are washes after prey-bait incubation, E1-E2 are elutions and beads after elutions. Arrows and numbers represent unique bands that are present on DnaK while a question mark (?) represents the same position in DnaK-G where the unique bands were observed in DnaK.

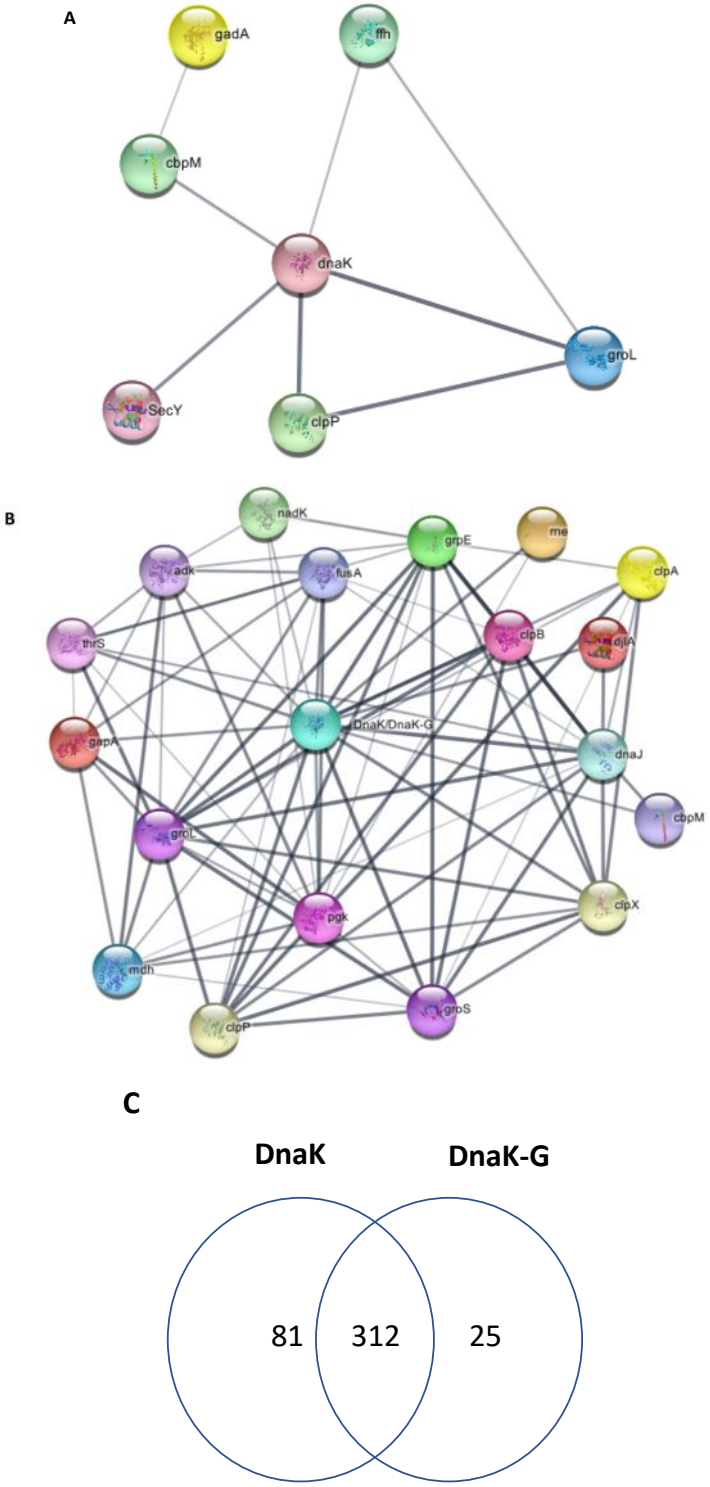


Figure 3.9: DnaK functional network as predicted by string database

a) Representation of predicted DnaK unique interactors and b) representation of DnaK/DnaK-G predicted unique proteins. Data were retrieved from string database, http://string-db.org/). (C) represents a Venn diagram of identified protein interactors of DnaK DnaK-G.

Table 3.2: DnaK and DnaK-G unique proteins identified by LCMS

DnaK unique interactors	DnaK-G unique interactors
Phosphopentomutase deoB	GTP cyclohydrolase 1folE
Pyridoxine/pyridoxal/pyridoxamine kinase pdxK	Putative uncharacterized protein ygaQ2
Uncharacterized HTH-type transcriptional regulator ynej	Ribosomal RNA small subunit methyltransferase E rsmE
3',5'-cyclic adenosine monophosphate phosphodiesterase cpdA	Virulence transcriptional regulatory protein phop
Glutamate decarboxylase alpha* gadA	Ribonucleotide monophosphatase nagD
Adenylosuccinate lyase purB	Transcription termination/antitermination protein nusG
DNA-cytosine methyltransferase dcm	Aldehyde-alcohol dehydrogenase adhE
Pyridoxine 4-dehydrogenase pdxl	Spermidine N (1)-acetyltransferase speG
HTH-type transcriptional regulator rpiR	Outer membrane protein C ompC
ECF RNA polymerase sigma-E factor rpoE	Bifunctional protein glmU
HTH-type transcriptional repressor purr	Histidine-binding periplasmic protein hisJ
ATP-dependent Clp protease proteolytic subunit clpP*	*E. coli* Protein YdcF
Acyl-CoA thioesterase 2 tesB	50S ribosomal protein L7/L12 rplL
Pyridoxal phosphate phosphatase ybhA	Elongation factor P-like protein pal
Ribosomal RNA small subunit methyltransferase rmsG	Peptidoglycan-associated lipoprotein efpL
Intermembrane phospholipid transport system ATP-binding protein MlaF	Alkyl hydroperoxide reductase ahpC
HTH-type transcriptional regulator pgrR	Glutamate--cysteine ligase gshA
ATP-dependent Clp protease proteolytic subunit clpP*	3-oxoacyl-[acyl-carrier-protein] synthase 3 fabH
Pyruvate kinase I pykF	Transcription termination factor rhO
Tryptophanase tnaA	60 kDa chaperonin geotm
Ribosomal RNA small subunit methyltransferase A rsmA	Elongation factor Tu 2 tu2
30S ribosomal protein rpcB	2-dehydro-3-deoxyphosphooctonate aldolase kdsA
transcriptional regulator TreR	Tyrosine recombinase xerC
50S ribosomal protein L3 yerP	ATP-dependent 6-phosphofructokinase pfkA
22 kDa peptidyl-prolyl cis-trans isomerase fkbP	
HTH-type transcriptional regulator cynR	
L-serine dehydratase 2 sdhM	
tRNA dimethylallyl transferase miaA	

Glucosamine-6-phosphate deaminase nagB
Phosphopentomutase deoB
3,4-dihydroxy-2-butanone 4-phosphate synthase ribB
Arabinose operon regulatory protein araC
Stringent starvation protein B sspB
Thymidylate synthase thyA
Dihydroxy-acid dehydratase 3 ilvD
Probable GTP-binding protein EngB
Protein translocase subunit SecY*
Colicin-E1 immunity protein imm
Histidine biosynthesis bifunctional protein hisB
Transcriptional regulator KdgR
HTH-type transcriptional repressor NemR
30S ribosomal protein S18
Signal recognition particle protein ffh*
Uncharacterized ABC transporter ATP-binding protein YbbA
Trehalose-6-phosphate hydrolase treC
N6-adenosine threonylcarbamoyltransferase tRNA
Anaerobic dimethyl sulfoxide reductase chain B dmsB
Delta-aminolevulinic acid dehydratase hemB
D-cysteine desulfhydrase dcyD
ATP-dependent dethiobiotin synthetase BioD
Flavodoxin 1 fldA
Protein HflK
LPS-assembly lipoprotein LptE
30S ribosomal protein S13 rpsM
Cyclopropane-fatty-acyl-phospholipid synthase cfa
3-bisphosphoglycerate-dependent phosphoglycerate mutase gpmA
Phosphoenolpyruvate synthase
Galactose-proton symporter galp

50S ribosomal protein L3 OS=Escherichia coli rlpC
Guanine nucleotide-binding protein G(i) subunit alpha gnaI
Toluene 1,2-dioxygenase system ferredoxin subunit todB
CH60_BUCAI
2-dehydro-3-deoxyphosphooctonate aldolase kdsa
ATP synthase subunit beta
6-phosphogluconate dehydrogenase, decarboxylating 6pgd
6-phosphogluconate dehydrogenase, decarboxylating 6pgd
3-oxoacyl-[acyl-carrier-protein] synthase fabH
Protein-L-isoaspartate O-methyltransferase pcm
NAD-dependent dihydropyrimidine dehydrogenase subunit PreT
50S ribosomal protein L3 rplc
Glycerol-3-phosphate dehydrogenase gpsA
Thymidine kinase tdk
Protein MtfA
60 kDa chaperonin clpA*

***** Represent proteins predicted to interact with DnaK by string database

4.0 Discussion and conclusion

PfHsp70-1 C-terminus domain is marked by the GGMP residues (Matambo *et al*., 2004). Parasitic Hsp70s possess the GGMP residues but these residues are enhanced in cytosolic Hsp70s of apicomplexa organisms in general (Makumire *et al*., 2021). Furthermore, *in vitro* studies have shown that the GGMP residues are important for PfHsp70-1 ATPase activity and substrate binding (Makumire *et al*., 2021). However, these residues have not been extensively characterized *in cellulo*. Observations made in complementation assay and co-affinity chromatography in this study show that the GGMP residues are important for the chaperone function of parasitic Hsp70. Furthermore, LCMS revealed that the insertion of GGMP residues in DnaK compromises its chaperone function and ability to interact with other substrates. Moreover, GGMP residues are involved in substrate synthesis and folding. Observations made in this study suggest that the GGMP residues are specifically important for Hsp70s of parasitic origin. This study also confirms that GGMP residues have enhanced numbers in Hsp70 of apicomplexa organisms. This could suggest that the GGMP residues may account for specialized features of Hsp70 from apicomplexa (Lebepe *et al*., 2020; Makumire *et al*., 2021).

In silico study revealed that conservative substitution of the GGMP residues did not lead to any structural defects of $KPf_{617\text{-}647}$ from the chimeric protein KPf. The loss of chaperone function of $KPf_{617\text{-}647}$ may be due to the absence of the GGMP residues as the protein retained its tertiary structure upon conservative substitution. On the other hand, deletion of the GGMP residues led to structural defects in the SBD domain of $KPf_{\Delta G}$. Observed changes could signify the importance of GGMP residues in the structural orientation of Hsp70s. A previous study reported that the SBD confers the stability of KPf under heat stress (Lebepe *et al*., 2020). This suggests that defects in the SBD led to reduced stability of KPf under stress conditions. Furthermore, the SBD of Hsp70 interacts directly with intermediates that are folding and some parts of folded proteins to stabilize them (Mashaghi *et al*., 2016). Defects in the SBD may also lead to loss of KPf interaction with intermediates. Lid repositioning was previously reported to account for functional defects and low affinity of Hsp70 with co-chaperones (Makumire *et al*., 2021). The insertion of the GGMP residues led to some changes in the SBD of DnaK-G in comparison to the wild-type DnaK. The observed unique structural orientation in $KPf_{\Delta G}$ and DnaK-G may account for the loss of chaperone function and consequently, the binding of substrates. Findings obtained in this study may suggest that the GGMP residues are important for parasitic Hsp70's unique structural orientation.

A previous study reported that the GGMP residues are predominant Hsp70s of parasites (Makumire *et al*., 2021). The dominance of the GGMP residues symbolizes the importance of these residues in parasitic organisms' development and pathogenesis. The presence of GGMP residues in cytosol localized Hsp70 could suggest that these residues play a vital role in Hsp70 function mediating quality control in the cell. Furthermore, PfHsp70-1 is reported to have a special function that could be implicated in the frequency of GGMP residues in apicomplexa organisms (Lebepe *et al*., 2020; Makumire *et al*., 2021). *Plasmodium* species Hsp70s possess most copies of GGMP residues as compared to other organisms within their domain. Taking into consideration the elevation of temperature when the human host is infected with malaria and the role of Hsp70s in response to heat, GGMP residues may be important for the specialized functions of PfHsp70-1. Furthermore, it has been previously reported that both KPf and PfHsp70-1 preferentially binds to asparagine rich peptides as compared to DnaK (Lebepe *et al*., 2020). The proteome of the malaria parasite composes of at least 30% of asparagine and glutamate repeat segments (Pallarès *et al*. 2018). This strongly suggests that the presence of the GGMP residues in PfHsp70-1 could be playing a vital role in substrate selectivity.

Structural analysis of the recombinant form of DnaK and its mutant form using CD spectroscopy further revealed that the insertion of GGMP residues in DnaK did not abrogate nor enhance the heat stability of DnaK. An independent study revealed that DnaK and PfHsp70-1 are both stable at temperatures up to 60ºC (Lebepe *et al*., 2020). The binding of ATP to Hsp70 is important in the Hsp70 functional cycle to release bound substrates. It has been previously demonstrated that DnaK exhibits a blue shift when bound to ATP. A blue spectral shift has also been reported to be a result of hydrophobicity increase around the tryptophan residue micro-environment (Qi *et al*., 2013; Kityk *et al*., 2012; Makumire, 2019). Tryptophan fluorescence conducted in this study generated a similar phenomenon, which strongly suggests the hydrophobicity around the tryptophan residue. Previously, it was reported that binding of substrate peptide ALLLMYRR and ANNNMYRR results in quenching of fluorescence and marginally blue shift (Makumire, 2019). This study reveals that DnaK-G did not respond to asparagine rich peptides. The observed findings could be accounted for by the altered DnaK-G SBD. The insertion of GGMP residues in DnaK may have blocked substrates from binding to the chaperone. The binding of substrate peptides in the SDB was reported to cause structural changes that did not modulate the hydrophobicity around the tryptophan residue (Makumire, 2019). The same phenomenon could account for the observed findings

when taking into consideration the position of the tryptophan residue in DnaK located in the NBD at position W102. This suggests that DnaK does not need the GGMP residues to perform its function. Furthermore, the GGMP residues may be specific for parasitic Hsp70s function.

A complementation assay revealed that the absence of the GGMP residues in $KPf_{617-647}$ and $KPf_{\Delta G}$ abrogated the chaperone function of these mutants. This suggests that the absence of the GGMP residues results in KPf losing its chaperone function *in cellulo*. Furthermore, the absence of the GGMP residues may have resulted in KPf mutants failing to recognize their substrates under heat stress. A recent study suggests that substitution of the flanks of the GGMP residues on KPf does not abrogate the chaperone function of KPf (Makumire *et al*., 2021). This study reveals that the GGMP residues are important for parasitic Hsp70 chaperone function in a cellular context. A recent study suggested that PfHsp70-1 folds substrates more efficiently than human Hsp70 (Anas *et al*., 2021). Most notable, KPf cytoprotected *E. coli dnaK756* cells more efficiently than DnaK (Figure 3.6). GGMP residues may account for efficient cytoprotection of *E. coli dnaK756* cells by KPf. This suggests that the presence of GGMP repeats may account for the efficient chaperone function of KPf. GGMP residues were previously speculated to account for the efficiency of PfHsp70-1 to fold proteins (Makumire *et al*., 2021). PfHsp70-1 was previously reported to protect cells under heat stress (Shonhai *et al*., 2005). Insertion of GGMP residues in DnaK resulted in DnaK losing its chaperone function.

DnaK forms a functional network with other co-chaperones to help *E. coli* cells survive under heat stress (Chung *et al*., 2006). The toxicity and failure of DnaK-G to reverse thermosensitive of *E. coli dnaK103* and *dnaK756* cells may suggest that DnaK does not need these residues to perform its function. The toxicity of DnaK-G may be due to *E. coli dnaK103* cells expressing a non-functional native DnaK which may not have helped DnaK-G to fold properly. On the other hand, a native DnaK with a low GrpE affinity expressed by *E. coli dnaK756* cells may have helped DnaK-G to fold resulting in the strain growing at the permissive temperature. Furthermore, the failure of DnaK-G to cytoprotect cells may suggest that GGMP may be specific for parasitic Hsp70 functions and their insertion is detrimental to DnaK. Observed findings are also supported by three-dimensional structure prediction and the tryptophan fluorescence assay. DnaK-G did not respond to nucleotides binding which is important in the Hsp70-Hsp40 functional cycle. The failure of DnaK-G to respond to nucleotide could account for the failure of DnaK-G to reverse thermosensitivity of *E. coli dnaK756* cell. The failure of

DnaK-G to respond to nucleotides suggests that DnaK-G is neither opening the lid nor closing the lid for substrate binding and release. These findings are supported by a previous study that suggested that the insertion of the GGMP residues in DnaK may disrupt the integration of the lid (Makumire *et al*., 2021).

Co-affinity chromatography revealed that the GGMP residues may enhance the function of PfHsp70-1 and preferential binding to other substrates. DnaK-G associated with proteins that are known DnaK chaperones and substrates. The insertion of GGMP residues may have been detrimental to DnaK-G resulting in irreversible interaction between DnaK chaperones and substrates. The irreversible interaction could be accounted for in the toxicity observed in *E. coli dnaK103* cells heterologous expressing DnaK-G (Figure 3.7). Furthermore, the insertion of GGMP residues in DnaK resulted in DnaK-G interacting with 25 unique proteins that did not interact with wild-type DnaK. Moreover, various proteins from different functional pathways were identified amongst proteins that interacted with DnaK-G. Most of the proteins associated with DnaK-G only could be attributed to the GGMP insertion belong to the transcription and translation factors including elongation factors such as ribosomal RNA small subunit methyltransferase, transcription termination, elongation factor. This group of proteins is reported to drive the translation elongation of growing peptides (Parker, 2001). A variety of groups of identified proteins in DnaK-G implicates the GGMP motif involvement in gene transcription and translation (He *et al*., 2000). This study supports the findings of a previous study that suggest that the GGMP residues are essential for chaperone function (Makumire *et al*., 2021). Furthermore, the interaction of DnaK-G with ATP-dependent 6-phosphofructokinase, GroEl like chaperone protein, glutamate-cysteine ligase further supports the functional versatility of GGMP residues. Hsp70s are reported to have a versatile function (Mayer and Bukau, 1998). These observations may reveal that the PfHsp70-1 function may be modulated by the presence of these GGMP residues. This study further confirms DnaK co-chaperone partners *in cellulo*. DnaK is known to interact with co-chaperones such as GroEL and DnaJ to fold misfolded proteins (Calloni *et al*., 2012). The interactome of DnaK has been reported to overlap with GroEL substrates (Calloni *et al*., 2012). Furthermore, an ATP-dependent direct association has been reported between σ^{32}, DnaK, DnaJ, and GrpE (Arsène *et al*., 2000). Although most of the co-chaperones are identified, the abrogation of some DnaK functions when inserted GGMP residues supports that GGMP residues are specific for parasitic PfHsp70-1. A high basal ATPase activity has been reported in PfHsp70-1 as compared to DnaK

(Lebepe *et al*., 2020; Makumire *et al*., 2021). Furthermore, PfHsp70-1 has also been previously reported to be an efficient chaperone as compared to human Hsp70 (Anas *et al*., 2020). Taking all that together, the number of identified proteins between DnaK and DnaK-G proves that, GGMP residues are involved in substrates synthesis and folding in parasites. The insertion of GGMP residues in DnaK resulted in DnaK failing to bind its chaperone, Hsp100. The role of the GGMP residues in substrate synthesis and proper folding could be accounted for the better cytoprotection that KPf provided *E. coli dnak756* cells as compared to DnaK. Hsp100-DnaK complex was previously reported to require the remodeling of ATPase-dependent (Fernández *et al*., 2018). DnaK-G's failure to be conformationally modulated by nucleotides could account for its substrate capability which was compromised. Furthermore, lid repositioning in DnaK as a result of inserted GGMP residues cannot be ruled out as responsible for the compromised function. Notably, DnaK-G also lost substrates that bound to the NBD of DnaK. In conclusion, this study reveals that GGMP residues are important for PfHsp70-1 chaperone function. However, this motif is detrimental to DnaK function. A follow-up study is needed to further investigate the role of these residues in *Plasmodium falciparum* 3D cells. Another study is needed to design and identify possible compounds that can inhibit the role of the GGMP residues of Hsp70 to combat the spread and infections of parasitic diseases.

References

Andreasson, C., Fiaux, J., Rampelt, H., Mayer, M.P., Bukau, B. (2008). Hsp110 is a nucleotide-activated exchange factor for Hsp70. *J Biol Chem*. **283**: 8877–8884

Anas, M., Shukla, A., Tripathi, A., Kumari, V., Prakash, C., Nag, P., Kumar, L.S., Sharma, S.K., Ramachandran, R., Kumar, N. (2020). Structural–functional diversity of malaria parasite's PfHsp70-1 and PfHsp40 chaperone pair gives an edge over human orthologs in chaperone-assisted protein folding. *Biochem J*. **477**:3625–3643

Ahn, Y.J., Im, E. (2020). Heterologous expression of heat shock proteins confers stress tolerance in *Escherichia coli*, an industrial cell factory: A short review. *Biocatal Agric Biotechnol*. **29**

Assimon, A.V., Southworth, D.R., Gestwick, J.E. (2015). Specific binding of tetratricopeptide repeat (TPR) proteins to heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90) is regulated by affinity and phosphorylation. *J Biochem*. **54**: 7120–7131

Arsène, F., Tomoyasu, T., Bukau, B. (2000). The heat shock response of *Escherichia coli*. *Int J Food Microbiol*. **55**:3-9

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

Betts, M.J., Russell, R.B. (2003). Amino acid properties and consequences of substitutions. Chapter 4. J. *Bioinform Comput.* Biol. Edited by Michael R. Barnes and Ian C. Gray Copyright John Wiley & Sons, Ltd. ISBN: 0-470-84393-4 (HB); 0-470-84394-2 (PB)

Biebl, M.M., Buchner, J. (2019). Structure, function, and regulation of the Hsp90 machinery. *Cold Spring Harb Perspect Biol*. a03401

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., Louw AI. (2008). Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malar J*. **7**:197

Botha, M., Pesce, E.R., 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**:1781-1803

Bukau, B., G. C. Walker. (1989). Delta *dnaK* mutants of *Escherichia coli* have defects in chromosome segregation and plasmid maintenance at normal growth temperatures. *J Bacteriol*. **171**: 6030-6038

Bukau, B., Weissman, J., Horwich, A. (2016). Molecular chaperones and protein quality control. *Cell Stress chaperones.* **125**: 443–451

Buchberger, A., Gässler, C.S., Büttner, M., McMacken, R., Bukau B. (1999). Functional defects of the *dnaK756* mutant chaperone of *Escherichia coli* indicate distinct roles for amino- and carboxyl-terminal residues in substrate and co-chaperone interaction and interdomain communication. *Chem Biol*. **274**:38017-26

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

Chakafana, G., Zininga, T., Shonhai, A. (2019a). Comparative structure-function features of Hsp70s of *Plasmodium falciparum* and human origins. *Biophys Rev*. **11**:1–12

Chakafana, G., Zininga, T., Shonhai, A. (2019b). The link that binds: The linker of Hsp70 as a helm of the protein's function. *Biomolecules*. 9:543

Chakafana, G.; Shonhai, A. (2021). The role of non-canonical Hsp70s (Hsp110/Grp170) in Cancer. *Cell J*. **10**:254

Cheetham, M. E., Caplan, A. J. (1998). Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. *Cell Stress Chaperones*. **3**: 28-36

Chuang S.E., Burland V., Plunkett G., Daniels D.L., Blattner F.R. (1993). Sequence analysis of four new heat-shock genes constituting the hslTS/ibpAB and hslVU operons *in Escherichia coli*. *Gene*. **134**: 1-6

Chung, H., Bang, W., Drake, M. (2006). Stress response of *Escherichia coli*. *Compr Rev Food Sci F*. **5**: 52-64

Clerico, E.M., Tilitsky, J.M., Meng, W., Gierasch, L.M. (2015). How hsp70 molecular machines interact with their substrates to mediate diverse physiological functions. *J Mol Biol*. **427**:1575-88

Datta, K., Rahalka, K., Dinesh, D.K. (2017). Classification and its involvement in health and disease. *J Pharm Health Care Sci*. **4**:175

Demand, J., Luders, J., Hohfeld. (1998). The carboxy-terminal domain of Hsc70 provides binding sites for a distinct set of chaperone cofactors. *Mol Cell Biol*. **18**: 2023-2028

Dragovic, Z., Broadley, S.A., Shomura, Y., Bracher. A., Hartl, F.U. (2006). Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. *EMBO J*. **25**:2519–2528

Fernández-Higuero, J.A., Aguado, A., Perales-Calvo, J. et al. (2018). Activation of the DnaK-ClpB Complex is Regulated by the Properties of the Bound Substrate. *Sci Rep*. **8**:5796

Fakhari, D.E., Saidi, L.G., Wahlster, L. (2013). Molecular chaperones and protein folding as therapeutic targets in Parkinson's disease and other synucleinopathies. *Acta Neuropathol Commun*. **1**:79

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

Gauley, J., Young, J.T.F., Heikkila, J.J. (2008). Intracellular localization of the heat shock protein, HSP110, in Xenopus laevis A6 kidney epithelial cells. *Comp Biochem*. **151**:133-138

Gitau, G.W., Mandal, P., Blatch, G.L., Przyborski, J., Shonhai, A. (2012). Characterization of the *Plasmodium falciparum* Hsp70-Hsp90 organizing protein (*Pf*Hop). *Cell Stress Chaperones*. **17**: 191-202

Goeckeler, J.L., Petruso, A.P., Aguirre, J., Clement, C.C., Chiosis, G., Brodsky, J.L. (2008). The yeast Hsp110, Sse1p, exhibits high affinity peptide binding. *FEBS Lett*. **582**: 2393-2396

Georgopoulos, C.P. (1977). A new bacterial gene (groPC) which affects lambda DNA replication. *Mol Genet*. **151**: 35-39

Hartl, F.U., Bracher, A., & Hartl, M.H. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*. **475**:324-325

He, H., Chen, C., Xie, Y., Asea, A., Calderwood, S. K. (2000). HSP70 and heat shock factor 1 cooperate to repress Ras-induced transcriptional activation of the c-fos gene. *Cell stress chaperones*. **5**:406–411

Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. **10**:845–858

Kityk, R., Kopp, J., Sinning, I., Mayer, M.P. (2012). Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Mol Cell*. **48**: 863-874

Kudyba, H.M., Cobb, D.W., Fierro, M.A., Florentin, A., Ljolje, D., Singh, B., Lucchi, N.W., Muralidharan, V. (2019). The endoplasmic reticulum chaperone PfGRP170 is essential for asexual development and is linked to stress response in malaria parasites. *Cell Microbiol*. 21:e13042

Harrison, C. (2010). GrpE, a nucleotide exchange factor for DnaK. *Cell Stress Chaperones*. **8**:218-24

Haslbeck, M. , Weinkauf , S. , Buchner , J. , (2019). Small heat shock proteins: simplicity meets complexity. *J Biol Chem*. **294**:2121 – 2132

Heiserman, J.P., Chen, L., Kim, B.S., Kim, S.C., Tran, A.L., Siebenborn, N., Knowlton, A.A. (2015). TLR4 mutation and Hsp60-induced cell death in adult mouse cardiac myocytes. *Cell Stress Chaperones*. **20**:527–535

Hennessy, F., Nicoll, W.S., Zimmermann, R., Cheethan, M.E., Blatch, G.L. (2007). Not all J domains are created equal implications for the specificity of Hsp40-Hsp70 interactions. *Protein Sci*. **14**:1697-1709

Heiny, S.R., Spork, S., Przyborski, J.M. (2012). The apicoplast of the human malaria parasite *P. falciparum*. *Endocytobiosis Cell Res*. **23**:91–95

Hoffmann, A., Bukau, B., Kramer, G. (2010). Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta Mol Cell Res*. **1803**:650-661

Jha, P., Laskar, S., Dubey, S., Bhattacharyya, M.K., Bhattacharyya, S. (2017). *Plasmodium* Hsp40 and human Hsp70: A potential cochaperone-chaperone complex. *Mol Biochem Parasitol Mol Biochem Par*. **21:**10-13

Kampinga, H.H., Andreasson, C., Barducci, A., Cheetham, M.E., Cyr, D., Emanuelsson, C., Genevaux, P., Gestwicki, J.E., Goloubinoff, P., Huerta-Cepas, J., Kirstein, J., Liberek, K., Mayer, M.P., Nagata, K., Nillegoda, N.B., Pulido, P., Ramos, C., De Los Rios, P., Rospert, S., Rosenzweig, R., Marszalek, J. (2019). Function, evolution, and structure of J-domain proteins. *Cell stress chaperones*. **24:**7–15

Kampinga, H.H., Craig, E.A. (2010). The Hsp70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol*. **11**:579-92

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., & Hartl, F.U. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annu Rev Biochem*. **83**:153-161

Kravats, A.N., Doyle, S.M., Hoskins, J.R., Genest, O., Doody, E. Wickner, S. (2017). Interaction of *E. coli* Hsp90 with DnaK involves the DnaJ binding region of DnaK. *J Mol Biol*. *429*:858 – 872

Kudyba, H.M., Cobb, D.W., Fierro, M.A., Florentin, A., Ljolje, D., Singh, B., Lucchi, N.W., Muralidharan, V. (2019). The endoplasmic reticulum chaperone PfGRP170 is essential for asexual development and is linked to stress response in malaria parasites. *Cell Microbiol*. **21**:e13042

Kufareva, I., Abagyan, R. (2012). Methods of protein structure comparison. *Methods Mol Biol*. **857**: 231-257

Lebepe, C.M., Matambanadzo, P.R., Makhoba, X.H., Achilonu, I., Zininga, T., Shonhai, A. (2020). Comparative characterization of *Plasmodium falciparum* Hsp70-1 relative to *E. coli* DnaK reveals the functional specificity of the parasite chaperone. *Biomolecules*. **10**:856

Liberek, K., Wall, D., Georgopoulos, C. (1995). The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. *Proc Natl Acad Sci*. **14**:6224-6228

Mabate, B., Zininga, T., Ramatsui, L., Makumire, S., Achilonu, I., Dirr, H.W., Shonhai, A. (2018). Structural and biochemical characterization of *Plasmodium falciparum* Hsp70-x reveals functional versatility of its C-terminal EEVN motif. *Proteins*. **86**:10

Makhoba, X.H., Burger, A., Coertzen, D., Zininga, T., Birkholtz, L.M., Shonhai, A. (2016) Use of a chimeric Hsp70 to enhance the quality of recombinant *Plasmodium falciparum* S-adenosylmethionine decarboxylase protein produced in *Escherichia coli*. *PLoS One.* **11:** e0152626

Makhoba, X.H., Pooe, O.J., Mthembu, M.S. (2015). Molecular chaperone assisted expression systems: obtaining pure soluble and active recombinant proteins for structural and therapeutic purposes. *Genomics Proteomics Bioinformatics*. **8**:212-216

Makumire, S. (2019). Investigation of the role of the GGMP residues of *Plasmodium falciparum* Hsp70-1 on the chaperone function of the protein and its interaction with a co-chaperone, PfHop. Ph.D., Thesis. University of Venda

Makumire, S., Dongola, T.H., Chakafana, G., Tshikonwane, L., Chauke, C.T., Maharaj, T., Zininga, T., Shonhai, A. (2021). Mutation of GGMP repeat segments of *Plasmodium falciparum* Hsp70-1 compromises chaperone function and Hop co-chaperone Binding. *Int J Mol Sci*. **22**: 2226

Makumire, S., Zininga, T., Vahokoski, J., Kursula, I., Shonhai, A. (2020). Biophysical analysis of *Plasmodium falciparum* Hsp70-Hsp90 organizing protein (PfHop) reveals a monomer that is characterized by folded segments connected by flexible linkers. *PLoS One*. **28**:e0226657

Mashaghi, A., Bezrukavnikov, S., Minde, D. P., Wentink, A. S., Kityk, R., Zachmann-Brand, B., Mayer, M.P., Kramer, G., Bukau B., Tans, S. J. (2016). Alternative modes of client binding enable functional plasticity of Hsp70. *Nature*. **539**. 448–451

Matambo, T., Odunuga, O.O., Boshoff, A., Blatch, G. (2004). Overproduction, purification and characterization of the *Plasmodium falciparum* heat shock protein 70. *Protein Expr Purif*. **33**: 214-222

Mattoo, R.U., Farina Henriquez Cuendet, A., Subanna, S., Finka, A., Priya, S., Sharma, S.K., Goloubinoff, P. (2012). Synergism between a foldase and an unfoldase: Reciprocal dependence between the thioredoxin-like activity of DnaJ and the polypeptide-unfolding activity of DnaK. *Front Mol Biosci*. **1**:7

Mayer, M.P., Bukau, B. (1998). Hsp70 chaperone systems: diversity of cellular functions and mechanism of action. *Chem Biol*. **379**:261-8

Mayer, M.P., Schröder, H., Rüdiger, S., Paal, K., Laufen, T., Bukau, B. (2000). Multistep mechanism of substrate binding determines chaperone activity of Hsp70. *Nat Struct Mol Biol*. **7**: 586-583

Miller, C.M.D., Smith, N.C., Johnson, A.M. (1999). Cytokines, nitric oxide, heat shock proteins and virulence in Toxoplasma. *J Parasitol*. **10**: 418-422

Misra, G., Ramachandran, R. (2009). Hsp70-1 from *Plasmodium falciparum*: protein stability, domain analysis and chaperone activity. *Biophysic Chem.* **142**:55-6

Mogk, A., Schlieker, C., Friedrich, K.L., Schönfeld, H., Vierling, E., Bukau, B. (2003). Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/ DnaK. *Biophys Chem*. **278**:31033 –31042

Mogk, A., Kummer, E., Bukau, B. (2015). Cooperation of Hsp70 and Hsp100 chaperone machines in protein disaggregation. *Biophysics J*. **2**:22

Muñoz, P.L.A., Minchaca, Z.A., Mares, R.E., Ramos, M.A. (2015). Protein folding and molecular chaperones of protozoa. Nova Science Publishers, Inc, ISBN: 978-1-63483-543-5

Muralidharan, V., Oksman, A., Pal, P., Lindquist, S., Goldberg, D.E. (2012). *Plasmodium falciparum* heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nature*. **3**:1310

Nguyen, B., Hartich, D., Seifert, U., Rios, P. (2017). Thermodynamic bounds on the ultra and infra-affinity of Hsp70 for its substrates. *Biophys J*. **113**:362–370

Njunge, J.M., Mandal, P., Przyborski, J.M., Boshoff, A., Pesce, E.R., Blatch, GL. (2015). PFB0595w is a *Plasmodium falciparum* J protein that co-localizes with PfHsp70-1 and can stimulate its in vitro ATP hydrolysis activity. *Int J Biochem Cell Biol.* **62**: 47–53

Njunge, J.M., Ludewig, M.H, A., Boshoff, A., Pesce, E.R., Blatch, G.L. (2013). Hsp70s and J proteins of *Plasmodium parasites* infecting rodents and primates: Structure, function, clinical relevance, and drug targets. *Curr Pharm Des*. **19**: 387–403

Nyakundi, D.O., Vuko, L.A., Bentley. S.J., Hoppe, H., Blatch, G.L., Boshoff A. (2016). *Plasmodium falciparum* Hep1 is required to prevent the self aggregation of PfHsp70-3. *PLoS One*. **11**:e0156446

Oh, H.J., Easton, D., Murawski, M., Kaneko, Y., Subjeck, J.R. (1999). The chaperoning activity of hsp110. Identification of functional domains by use of targeted deletions. *J Biol Chem*. **274**:15712–15718

Parker, J. (2001). "Elongation Factors; Translation". *Encyclopedia of Genetics*. 610–611

Paek, K.H., Walker, G.C. (1987). *Escherichia coli* DnaK null mutants are inviable at high temperature. *J Bacteriol*. **69**: 283–290

Pallavi, R., Archarya, P., Chandran, S., Daily, J.P., Tatu, U. (2010). Chaperone expression profiles correlate with distinct physiological states of *Plasmodium falciparum* in malaria patients. *Malar J*. **9**: 236

Pesce, E.R., Acharya, P., Tatu, U., Nicoll, W., Shonhai, A., Hoppec, C.H., Blatch, G.H. (2008). The *Plasmodium falciparum* heat shock protein 40, Pfj4, associates with heat shock protein 70 and shows similar heat induction and localization patterns. *Int J Biochem Cell Biol*. **40**:2914–2926

Pooe, O.J., Kollisch, G., Heine, H., Shonhai, A. (2017). *Plasmodium falciparum* heat shock protein 70 lacks immune modulatory activity. *Protein Pept Lett*. **24**:503–510

Prapapanich, V., Chen, S., Toran, E.J., Rimerman, R.A., Smith, D.F. (1996). Mutational analysis of the hsp70-interacting protein Hip. *Mol Cell Biol*. 6200-6207

Przyborski, J.M., Diehl, M., Blatch, G.L. (2015). Plasmodial Hsp70s are functionally adapted to the malaria parasite life cycle. *Frontiers in Molecular Biosciences*. **2**: 34

Qi, R., Sarbeng, E.B., Liu, Q., Le, K.Q., Xu, X., Xu, H., Yang, J., Wong, J.L., Vorvis, C., Hendrickson, W.A., Zhou, L., Liu, Q. (203). Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP. *Nat Struct Mol Biol*. **20**: 900-907

Requena, J.M, Montalvo, A.M., Fraga J. (2015). Molecular chaperones of Leishmania: central players in many stress related and unrelated physiological processes. *Biomed Res Int*. **2015**:25

Richter, K., Haslbeck, M., Buchner, J. (2010). The heat shock response: life on the verge of death. *Mol Cell*. **40**:253 – 266

Roncarati, D., Scarlato, V. (2017). Regulation of heat - shock genes in bacteria: from signal sensing to gene expression output. *FEMS Microbiol. Rev*. **41**:549 – 574

Shaner, L., Morano, A. (2007). All in the family: atypical Hsp70 chaperones are conserved modulators of Hsp70 activity. *Cell Stress Chaperones*. **12**:1-8

Shonhai, A. (2007). Molecular characterization of the chaperone properties of *Plasmodium falciparum* heat shock protein 70. Ph.D., Thesis. Rhodes University

Shonhai, A., Boshoff, A., Blatch, G.L. (2005). Molecular characterization of the chaperone properties of *Plasmodium falciparum* heat shock protein 70. *Mol Gen Genet*. **274**: 70–78

Shonhai, A. (2010). *Plasmodial* heat shock proteins: targets for chemotherapy. *FEMS Immunol Med Microbiol.* **58**: 61-74

Shonhai, A. (2014). Role of Hsp70s in development and pathogenicity of *Plasmodium* species:

In: Shonhai A, Blatch G, Editors. *Heat Shock Proteins of Malaria*, Springer New York;47-70

Shonhai, A., Boshoff, A., Blatch, G.L. (2007). The structural and functional diversity of Hsp70 proteins from *Plasmodium falciparum*. *Protein Science*. **16**: 1803-1818

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

Shonhai, A., Maier, A.G., Przyborski, J., Blatch, G.L. (2011). Intracellular protozoan parasites of humans: The role of molecular chaperones in development and pathogenesis. *Protein Peptides*. **18**:143–157

Shonhai, A., Boshoff, A., 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

Singh, G.P., Chandra, B.R., Bhattacharya, A., Akhouri, R.R., Singh, S.K., Sharma, A. (2004). Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in Plasmodium falciparum. *Mol. Biochem. Parasitol*. **137**:307–319

Pallarès I, de Groot NS, Iglesias V, Sant'Anna R, Biosca A, Fernàndez-Busquets X, Ventura S. (2018). Discovering putative prion-like proteins in *Plasmodium falciparum*: a computational and experimental analysis. *Front Microbiol*. **9**:1737

Spence, J., Cegielska, A., Georgopoulos, C. (1990). Role of *Escherichia coli* heat shock proteins DnaK and HtpG (C62.5) in response to nutritional deprivation. *J Bacteriol*. **172**: 7157-7166

Pesce, E., Blatch, G. (2014). Plasmodial Hsp40 and Hsp70 chaperones: Current and future perspectives. *Parasitol J*. **141**: 1167-1176

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. (2004). UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem*. **25**. 1605–1612

Suppini, J-P., Amor, M., Alix, J-H., Ladjimi, M. M. (2004). Complementation of an *Escherichia coli* DnaK defect by Hsc70-DnaK chimeric proteins. *J Bacteriol*. **186**: 6248–6253

Wang, L., Yang, S., Zhao, K., Han, L. (2015). Expression profile of heat shock protein 70 gene in response to heat stress in *Agrotis c-nigrum* (Lepidoptera: Noctuidae). *J Insect Sci*. **15**:564-569

Wall, D., Zylicz, M., Georgopoulos, C. (1999). The NH2-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. *J Biol Chem*. **269**: 5446-5451

Wild, J., Rossmeissl, P., Walter, W. A., & Gross, C. A. (1996). Involvement of the DnaK-DnaJ-GrpE chaperone team in protein secretion in *Escherichia coli*. *J Bacteriol*. **178**:3608–3613

Wu, J., Newton, A. (1996). Isolation, Identification, and Transcriptional Specificity of the Heat Shock Sigma Factor s32 from Caulobacter crescentus. *J Bacteriol*. **178**: 2094–2101

Xu, X., Sarbeng, E.B., Vorvis, C., Kumar, D.P., Zhou, L., Liu, Q. (2012). Unique peptide substrate binding properties of 110-kDa heat-shock protein (Hsp110) determine its distinct chaperone activity. *J Biol Chem*. **287**:5661–5672

Yeo, S.J., Liu, D.X., Park, H., Korean. (2015). Potential Interaction of *Plasmodium falciparum* Hsp60 and Calpain. *J Parasitol*. **53**:665-73

Zhang, B., Chambers, M.C., Tabb, D.L. (2007). Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J Proteome Res*. **6**:3549–3557

Zhang, P., Leu, J.I.J., Murphy, M.E., George, D.L., Marmorstein, R. (2014). Crystal structure of the stress-inducible human Heat Shock Protein 70 Substrate-binding domain in complex with Peptide Substrate. *Plos One.* **9**: e103518

Zhang, Q., Ma, C., Oberli, A., Zinz, A., Engels, S., Przyborski, J.M. (2017). Proteomic analysis of exported chaperone/co-chaperone complexes of *P. falciparum* reveals an array of complex protein-protein interactions. *Sci Rep*. **7**:42188

Zhang, T., Ploetz, E.A., Nagy, M., Doyle, S.M., Wickner, S., Smith, P.E., Zolkiewski, M. (2012). Flexible connection of the N-terminal domain in ClpB modulates substrate binding and the aggregate reactivation efficiency. *Proteins*. **80**:2758–2768

Zininga, T., Achilonu, I., Hoppe, H., Prinsloo, E., Dirr, H., Shonhai, A. (2016). *Plasmodium falciparum* Hsp70-z, an Hsp110 homolog, exhibits independent chaperone activity and interacts with Hsp70–1 in a nucleotide-dependent fashion. *Cell Stress Chaperones*. **21**:499-513

Zininga, T., Ramatsui, L., Makhado, P.B., Makumire, S., Achilinou, I., Hoppe, H., Dirr, H., Shonhai, A. (2017). Epigallocatechin-3-Gallate Inhibits the chaperone activity of *Plasmodium falciparum* Hsp70 chaperones and abrogates their association with functional partners. *Molecules*. **22**: 2139

Zininga, T., Achilonu, I., Hoppe, H., Prinsloo, E., Dirr, H.W., Shonhai, A. (2015). Overexpression, purification and characterization of the *Plasmodium falciparum* Hsp70-z (PfHsp70-z) Protein, *PLoS ONE* **10**: e0129445

Zininga, T., Shonhai, A. (2014). Are heat shock proteins druggable candidates? *Biosci Biotechnol Biochem*. **10**, 209–210

Zininga, T., Shonhai, A. (2019). Small molecule inhibitors targeting the heat shock protein system of human obligate protozoan parasites. *Int j mol sci*. **20**: 5930

Appendix A: Supplementary methodology

A.1 Competent cells preparation

A colony of BB2393/BB2362 cells was picked from a 2YT (2x yeast-tryptone) agar plate (1.6g tryptone, 0.5g yeast, 1g yeast, and 1.5g agar per 100 ml distilled water) and inoculated into 5ml 2x YT broth containing 1.6g tryptone, 0.5 yeast, 1g yeast. The culture was incubated overnight at 37°C and shaking at 160 rpm. The overnight culture was transferred to 45ml 2 YT broth and allowed to grow to OD 0.3-0.6 at 600 nm. Cells were then centrifuged at 5000 xg for 20 minutes at 4°C. Pellet was resuspended with ice-cold 10 ml 0.1M $MgCl_2$ and allowed to stand in ice for 30 mins. Cells were then centrifuged for 10 minutes at 5000 xg, 4°C and then resuspended in 5 ml 01M $CaCl_2$ and incubated in ice for 4 hours with a gently shaking. Cells were then centrifuged for 10 minutes at 5000 xg, 4°C and resuspended in 5ml $CaCl_2$ and ice-cold 30% glycerol. Competent cells were then aliquoted into 2 ml tubes then stored at -80°C fridge for future use.

A.2 Transformation

100 µl of competent cells were added into microcentrifuge tubes in an icebox. 2 µl of DNA was added into competent cells aliquot and gently mixed. Cells were incubated for 30 minutes on ice. Cells were then heat shocked at 42°C for 60 seconds and then placed on ice for 10 minutes. 900 µl of 2 YT broth was added into the cells and incubated for 1 hour at 37°C with shaking. Cells were then streaked on 2 YT agar plates containing appropriate antibiotics.

A.3 DNA extraction and restriction digestion

A colony from successfully transformed cells was picked from each transformation plate and inoculated into 2 YT broth supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C shaking. A Zippy^{TM} Mini-Prep kit (Zymo Research, USA) procedure provided by the supplier was followed to extract. In brief, overnight culture was centrifuged in 2 ml Eppendorf tubes for 2 minutes at 12000 xg. Resuspension buffer was then added to resuspend the pellet after discarding the supernatant until the pellet was completely resuspended. Cells are then lysed with lysis buffer for few minutes but less than 5 minutes, which are then neutralized with neutralization solution. Cells are then centrifuged for 5 minutes at 12000 xg, then the supernatant was carefully transferred into a Zymo^{TM} spin column without disturbing the pellet. The spin column was then centrifuged for 1 minute and then the flow-through was discarded. 500 ml was added into the spin column and the spin column was then centrifuged for 1 minute

at 12000 xg. The washing step was repeated under the same conditions and procedure. The flow-through was discarded and the tube was centrifuged again at 12000 xg. The spin column was transferred into a sterile 1.5 ml eppendorf tube and then 50 µl of elution buffer was added and incubated for 2 minutes at room temperature followed by centrifugation for 2 minutes at 12000 xg. Extracted DNA was confirmed with necessary enzymes. The table below shows how DNA samples for agarose gel were prepared.

Table A.1: Restriction digest reaction mixture

	Tube 1 (volumes in µL)	**Tube 2(*Hind*III) (volumes in µL)**	**Tube 3 (*Bam*HI) (volumes in µL)**	**Tube 4 (volumes in µL)**
H_2O	16	15	14	13
Buffer	2	2	2	2
DNA	2	2	2	2
Enzyme	-	1	2	3
Total	20	20	20	20

A.4 Analysis of proteins using SDS-PAGE

SDS-PAGE was used to separate proteins according to their sizes while the role of SDS was to impart negative charge on proteins so that they can have the same charge, furthermore, proteins migrate from negative to the positive pole in the gel. Protein samples were heated at 95℃ in the presence of SDS sample buffer (0.25% Coomassie Brilliant blue (R250); 2 % SDS; 10% glycerol (v/v); 100 Mm Tris; 1 % β-mercaptoethanol) in a ratio of 4:1 for 10 minutes at 95℃ and resolved in 12% acrylamide resolving gel prepared as shown in the table.

Table A.2: x2 preparations for running gel of SDS-PAGE

Regents	**x2**
Bis-acrylamide	30 % (w/v)
Tris pH 8.8	1.5 M
SDS	10 % (w/v)
H_20	3.16 mL
APS	10 % (w/v)

TEMED	20 µL

Table A.3: x2 preparations for stacking gel of SDS-PAGE

Reagents	**x2**
Bis	30 % (w/v)
Tris pH 6.8	0.5 M
SDS	10% (w/v)
H_20	2.1 mL
APS	10% (w/v)
TEMED	20 µL

A.5 Silver staining

The SDS-PAGE gel was used for silver staining. Pierce silver staining kit was used to stain gels (Thermo scientific, USA) as described per supplier's protocol as described below:

1. The gel was washed for 5 minutes in ultrapure water.
2. The wash step was repeated.
3. Gel was fixed in 30% ethanol:10% acetic acid solution (i.e., 6:3:1 water:ethanol:acetic acid) for 15 minutes.
4. The gel was washed in 10% ethanol solution for 5 minutes.
5. The wash step was repeated.
6. The gel was washed in ultrapure water for 5 minutes.
7. The wash step was repeated.
8. Sensitizer working solution was prepared by mixing 50 µL Silver stain sensitizer with 25 mL water.
9. The gel was incubated in a sensitizer working solution for exactly 1 minute, then wash with two changes of ultrapure water for 1 minute each.
10. Stain working solution was prepared by mixing 0.5 mL of silver stain enhancer with 25 mL stain.
11. The gel was incubated in a stain working solution for 30 minutes.
12. The developer working solution was prepared by mixing 0.5 mL of silver stain enhancer with 25 mL developer.
13. 5% acetic acid solution was prepared as a stop solution.
14. Quickly gel was washed with two changes of ultrapure water for 20 seconds each.

15. Immediately developer working solution was added and incubated until protein bands appear.
16. When the desired band intensity was reached, the developer working solution was stopped with a prepared Stop Solution (5% acetic acid).

A.6 Immunoblotting (Western blot)

Expressed proteins were confirmed by western blot. Proteins that were on SDS-PAGE gel were transferred to a Nitrocellulose membrane. Prior to transferring proteins, the nitrocellulose membrane was activated using 70% ethanol for 5 seconds followed by washing the membrane with deionized water then equilibrating the nitrocellulose membrane on ice-cold western transfer buffer (25 mM Tris; 192 mM glycine, 20% methanol) for 5 minutes. Proteins were transferred for 1 hour at 100 volts. The membrane was blocked by incubating the membrane for 1 hour on 5% (w/v) skimmed milk prepared in TBS (50 mM Tris, 150 mM NaCl, pH 7.5) on ice. The membrane was washed using TBS-Tween (50 mM Tris, 150 mM NaCl, 0.1% (w/v) Tween 20) 3 times for 5 minutes each, then followed by incubating the membrane with anti-His antibody in 5% skimmed milk at 4℃ on a shaker for 1 hour. The membrane was washed with TBS-Tween 20 to remove unbound antibodies 3 times for 5 minutes per wash. For visualization of bands chemiluminescence (ECL) developing reagents was used and Chemidoc™ MP Imaging System was used (BioRad, USA) to visualize the image.

Appendix B: Supplementary data

A.1 DnaK and KPf agarose gel

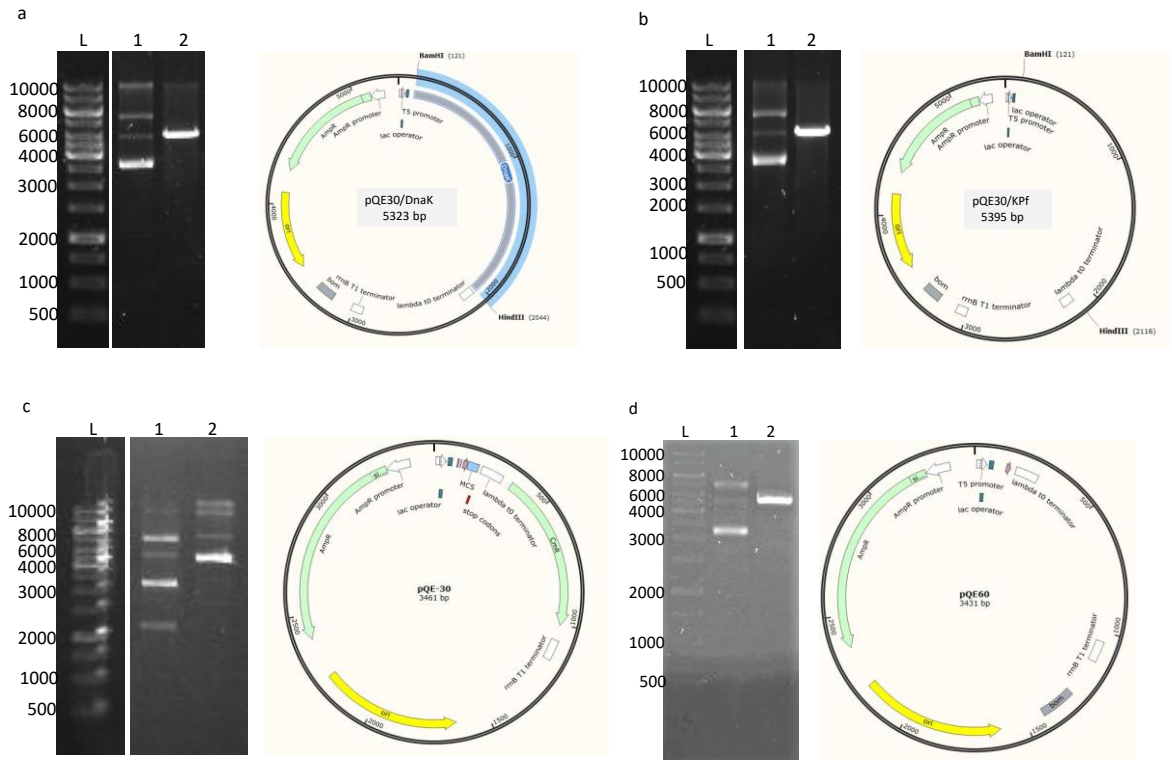


Figure S1: pQE30/DnaK, pQE60/KPf, pQE60 and pQE30 restriction digest analysis.

Agarose gel for a) DnaK, b) KPf, c) pQE30 and d) pQE30. Lane 1; Uncut DNA, Lane 2; single digest with *Hind*III enzyme

A1.1. Full agarose gel

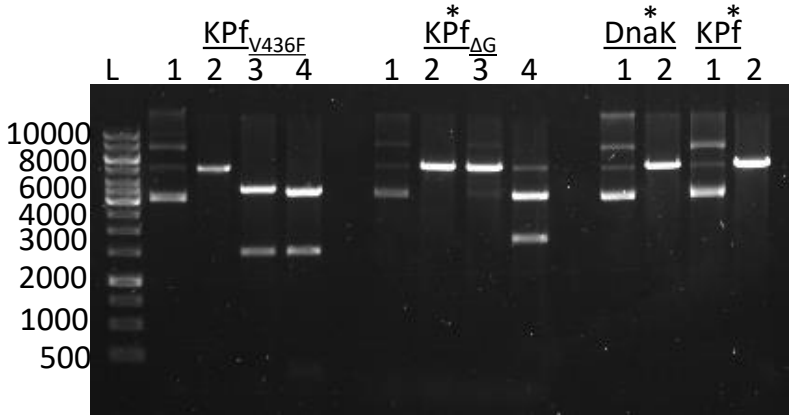


Figure S1.1: Restriction analysis of pQE60/KPf$_{V436F}$, pQE60/KPf, pQE60/KPf$_{\Delta G}$, pQE60/DnaK

Lane 1; Uncut DNA, Lane 2; single digest with one of the respective enzymes, Lane 3; single digest with the second respective enzyme, Lane 4; double digest with respective enzymes. Stars represent gels that were sliced.

A2. Silver staining gel for *E. coli dnaK756* cells and *E. coli* Δ*dnaK52* incubated with magnetic beads.

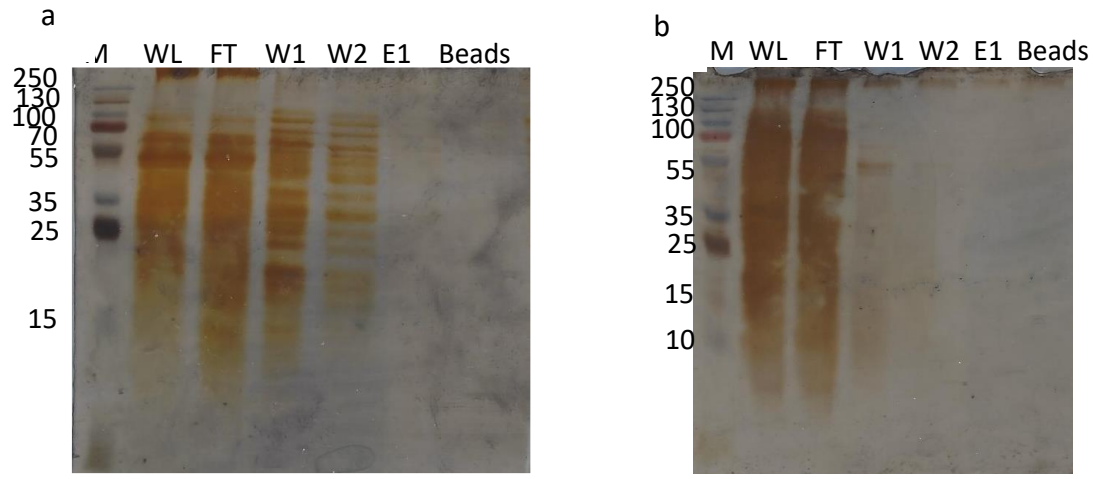


Figure S2: Investigation of direct interaction between *E. coli dnaK756* cells, *E. coli* Δ*dnaK52* cells, and magnetic beads using co-affinity chromatography

Co-affinity chromatography was conducted to investigate the direct interaction between *E. coli* strains *dnaK756* cells, Δ*dnaK52,* and magnetic beads. Eluents were further analyzed using silver stain gel, *E. coli dnaK756* cells (a), and *E. coli* Δ*dnaK52* (b). Lane M; protein marker, Lane WL; *E. coli* strain whole lysate, Lane FT; non-binding proteins samples, Lane W1-W2 samples collected after washes, E1; eluted proteins and Beads; magnetic beads used.

A.3 SDS-PAGE analysis of DnaK and DnaK-G purified proteins

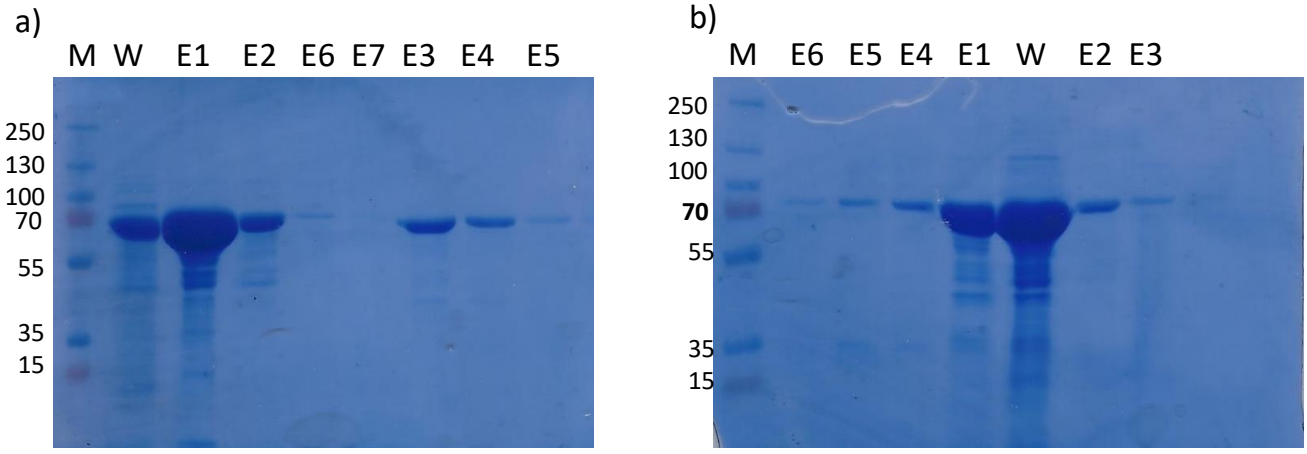


Figure S3: Purification of DnaK and DnaK-G

Wild type protein purification for (a) DnaK and (b) DnaK-G analyzed by SDS-PAGE. Lane M: molecular weight marker (kDa); Lane W; washes and Lane E1-E5.

A4. pQE30 expression gel and Western blot

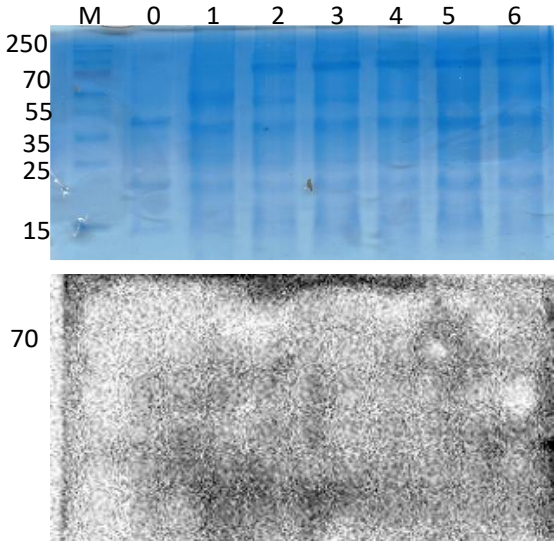
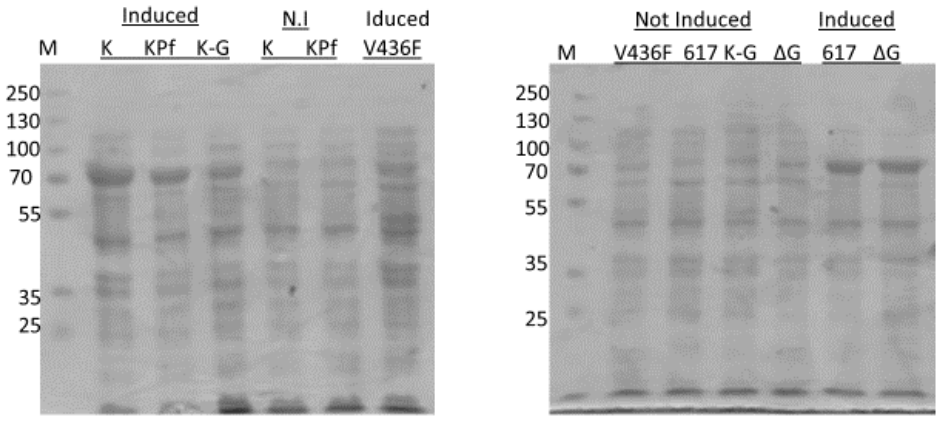


Figure S4: SDS-PAGE analysis of pQE30 plasmid

pQE30 plasmid expression was analyzed using SDS-PAGE**.** The following lanes represent, Lane M: protein marker (kDa), Lane 0; pre-induction sample, lane 1-6; hourly samples after induction.

A.4. Confirmation of DnaK and KPf constructs



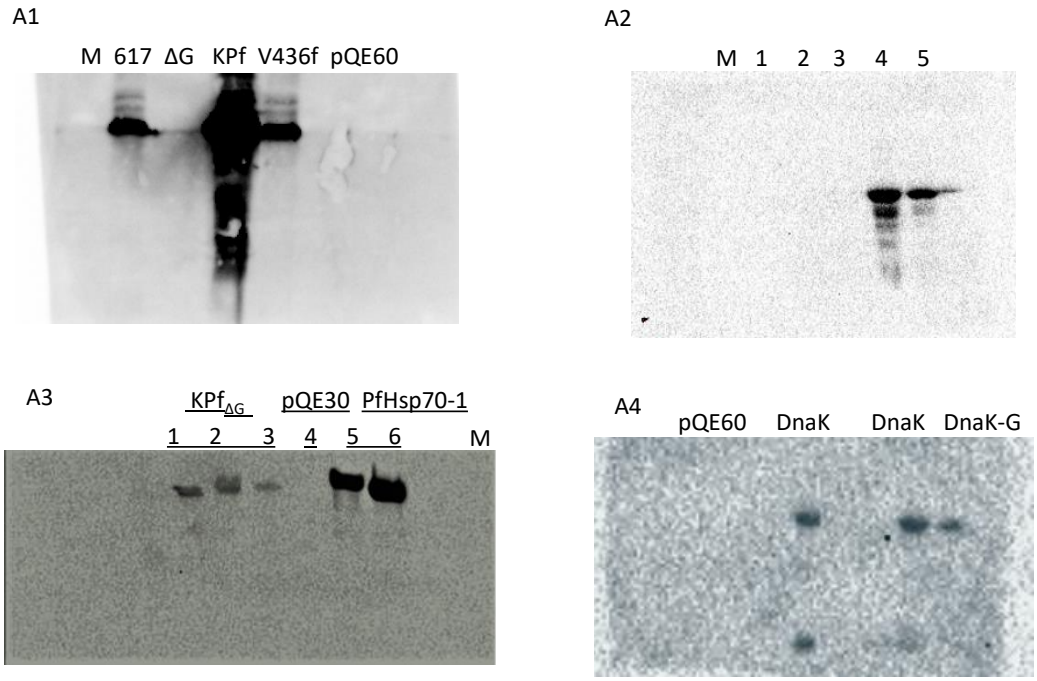


Figure S5: DnaK and KPf variants expression analysis and confirmation

SDS-PAGE (upper panel) analysis of DnaK and KPf variants expression and western blot (lower panel) to confirm. A1; Western blot analysis of KPf constructs. A2; Western blot analysis of DnaK constructs, 1; pQE30, 2; uninduced DnaK, 3; uninduced DnaK-G 4; induced DnaK and 5; induced DnaK-G.

A5. Identification of DnaK and DnaK-G samples

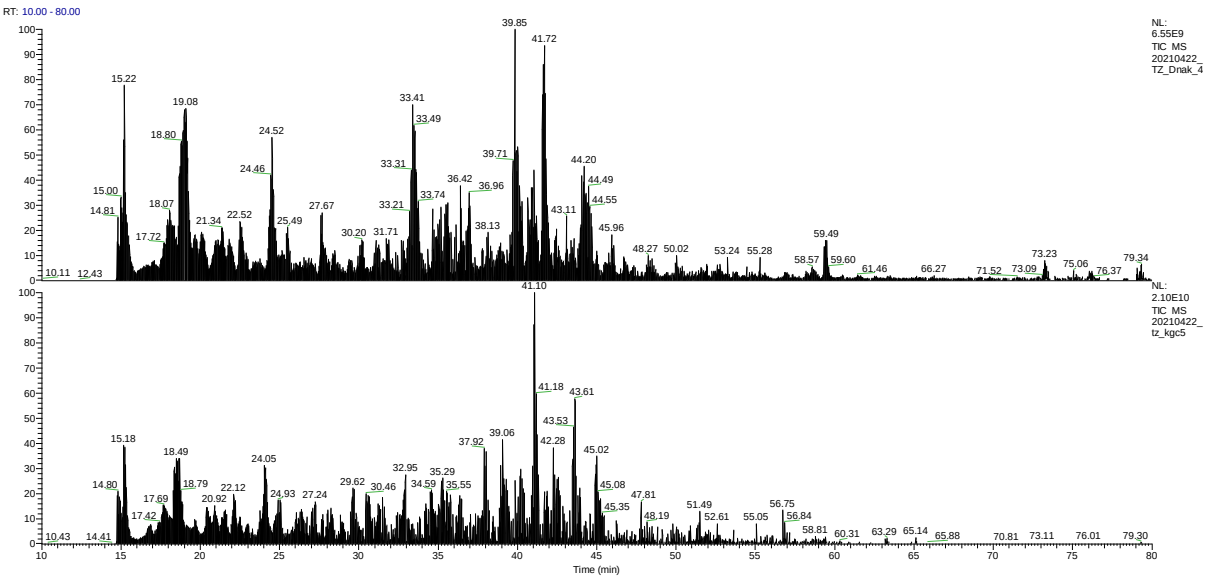


Figure S6: LCMS total ion chromatograms

Total ion chromatograms of DnaK (top), followed by DnaK-G (bottom)

Appendix C: List of Reagents

REAGENT	SUPPLIER
Acetic acid	Merck, Germany
Adenosine triphosphate	Sigma, U.S.A
Agarose	Whitehead scientific, South Africa
Ammonium molybdate	Merck, Germany
Ammonium persulphate	Merck, Germany
Ampicillin	Sigma, U.S.A
Bovine serum albumin	Sigma, U.S.A
Bromophenol blue	Sigma, U.S.A
Calcium chloride	Merck, Germany
Chloramphenicol	Sigma, U.S.A
Coomasie brilliant blue R250	Merck, Germany
Diethiothreitol	Sigma, U.S.A
Ethidium bromide	Sigma, U.S.A
Glacial acetic acid	Merck, Germany
Glycerol	Merck,
Glycine	Merck, Germany
Imidazole	Sigma, U.S.A
Isopropyl-1-thio-D-galacopyranoside	Sigma, U.S.A
Lysozyme	Merck, Germany
Magnesium chloride	Merck, Germany
Methanol	Merck, Germany
Monoclonal anti-His6-HRP antibodies	Sigma, U.S.A

Ni-NTA resin	**Thermo Scientific, U.S.A**
Nitrocellulose membrane	**Pierce, U.S.A**
PagerRuler Prestained Protein Ladder	**Thermo Scientific, U.S.A**
Peptone	**Merck, Germany**
Phenylmethylsufonyl fluoride	**Sigma, U.S.A**
Polyacrylamide	**Merck, Germany**
Polyethylene glycol 2000	**Sigma, U.S.A**
Polyethylenimine	**Sigma, U.S.A**
Ponceau S	**Sigma, U.S.A**
Potassium chloride	**Merck, Germany**
Potassium dihydrogen phosphate	**Merck, Germany**
Proteinase-K	**Sigma, U.S.A**
Restriction enzymes	**Thermo Scientific, U.S.A**
Snakeskin™ pleated dialysis tubing	**Pierce, U.S.A**
Sodium chloride	**Merck, Germany**
Sodium dodecyl sulphate	**Merck, Germany**
Sodium hydroxide	**Merck, Germany**
TEMED	**Sigma, U.S.A**
Tris	**Merck, Germany**
Tryptone	**Merck, Germany**
Tween 20	**Merck, Germany**
Urea	**Melford, UK**
Yeast extract powder	**Merck, Germany**
β-mercaptoethanol	**Sigma, U.S.A**