CHARACTERISATION OF HUMAN Hsj1a: AN HSP40 MOLECULAR CHAPERONE SIMILAR TO MALARIAL Pfj4

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

MASTER OF SCIENCE

of

RHODES UNIVERSITY

by

CARYN McNAMARA

May 2006

DECLARATION

I, Caryn McNamara, declare that this is my own unaided work hereby submitted for the Master of Science degree at Rhodes University. It has not been submitted for any degree for examination in any other university.

_____

Caryn McNamara

May 2006

Dated this _____ day of _____ 2006, at _____.

ABSTRACT

Protein folding, translocation, oligomeric rearrangement and degradation are vital functions to obtain correctly folded proteins in any cell. The constitutive or stress-induced members of each of the heat shock protein (Hsp) families, namely Hsp70 and Hsp40, make up the Hsp70/Hsp40 chaperone system. The Hsp40 J-domain is important for the Hsp70-Hsp40 interaction and hence function. The type-II Hsp40 proteins, *Homo sapiens* DnaJ 1a (Hsj1a) and *Plasmodium falciparum* DnaJ 4 (Pfj4), are structurally similar suggesting possible similar roles during malarial infection. This thesis has focussed on identifying whether Hsj1a and Pfj4 are functionally similar in their interaction with potential partner Hsp70 chaperones. Analysis *in silico* also showed Pfj4 to have a potential chaperone domain, a region resembling a ubiquitin-interacting motif (UIM) corresponding to UIM1 of Hsj1a, and another highly conserved region was noted between residues 232-241. The highly conserved regions within the Hsp40 J-domains, and those amino acids therein, are suggested to be responsible for mediating this Hsp70-Hsp40 partner interaction. The thermosensitive *dnaJ cbpA Escherichia coli* OD259 mutant strain producing type-I *Agrobacterium tumefaciens* DnaJ (*Agt*DnaJ) was used as a model heterologous expression system in this study. *Agt*DnaJ was able to replace the lack of two *E. coli* Hsp40s *in vivo*, DnaJ and CbpA, whereas *Agt*DnaJ(H33Q) was unable to. *Agt*DnaJ-based chimeras containing the swapped J-domains of similar type-II Hsp40 proteins, namely Hsj1*Agt* and Pfj4*Agt*, were also able to replace these in *E. coli* OD259. Conserved J-domain amino acids were identified and were substituted in these chimeras. Of these mutant proteins, Hsj1*Agt*(L8A), Hsj1*Agt*(R24A), Hsj1*Agt*(H31Q), Pfj4*Agt*(L11A) and Pfj4*Agt*(H34Q) were not able to replace the *E. coli* Hsp40s, whilst Pfj4*Agt*(Y8A) and Pfj4*Agt*(R27A) were only able to partially replace them. This shows the leucine of helix I and the histidine of the loop region are key in the *in vivo* function of both proteins and that the arginine of helix II is key for Hsj1a. The histidine-tagged Hsj1a protein was also successfully purified from the heterologous system. The *in vitro* stimulated ATPase activity of human Hsp70 by Hsj1a was found to be approximately 14 nmol P_i/min/mg, and yet not stimulated by Pfj4, suggesting a possible species-specific interaction is occurring.

DEDICATION

In loving memory of my father

Douglas Ralph McNamara

13.08.1944 - 10.11.2003

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor Gregory Blatch for the opportunity to work on this project. Whilst at first chaperones seemed very daunting subject, they are truly fascinating and a source of constant amazement.

Thank you to Dr William Nicoll, who coined the original idea that lead to the commencement of this MSc project, and co-supervised the first year of the project.

To those who inspired me with a passion for science through their teaching in various aspects and fields: Mrs Alison Page, Dr Brett Pletschke, and Profs Rosemary Dorrington and Gregory Blatch. Your enthusiasm for science is truly contagious.

To the Department of Biochemistry, Microbiology and Biotechnology, thank you for all the learning and inspiration that has happened within your walls over the last six years.

To the entire Chaperone Research Group.

To my friends over the last two years, whether within the lab or outside of it, thanks for all the support. It takes a lot to keep someone sane whilst they are confined to a lab.

To Jason Economou, for giving me so much happiness and motivation.

Also, last but certainly not least, to my family. I could never have done any of this without your unfailing support throughout my time spent so far away from home. Thank you so much. Words cannot describe... And to my mother, you truly are *Simply The Best.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

α	Alpha
β	Beta
λ	Lambda
° C	Degrees Celsius
μg	Microgram(s)
μℓ	Microlitre(s)
μM	Micromolar
μmol	Micromole(s)
%	Percentage
A	Absorbance
A. thaliana	*Arabidopsis thaliana*
A. tumefaciens	*Agrobacterium tumefaciens*
ADP	Adenosine diphosphate
*Agt*DnaJ	*A. tumefaciens* DnaJ
Amp^R	Ampicillin resistance
APS	Ammonium persulphate
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
Bag-1	Bcl2-associated athanogene 1 protein
BiP	Binding protein
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BSA	Bovine serum albumin
C. elegans	*Caenorhabditis elegans*
CbpA	Curved DNA binding protein A
cDNA	Complementary DNA
CHIP	C-terminus of heat shock cognate 70 stress protein-interacting protein
D. bucksii	*Drosophila bucksii*

D. melanogaster	*Drosophila melanogaster*
D. rerio	*Danio rerio*
Da	Daltons
DNA	Deoxyribonucleic acid
DnaJ	*E. coli* Hsp40
DnaK	*E. coli* Hsp70
dsDNA	Double-stranded deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra-acetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
g	Gram(s)
g	Gravitational force
G. zeae	*Gibberella zeae*
GF-rich region	Glycine-phenylalanine rich region
H. marismortui	*Haloarcula marismortui*
H. sapiens	*Homo sapiens*
HPD motif	Histidine, proline, aspartic acid motif
Hip	Hsp70-interacting protein
His-tagged	6 x Histidine tag
Hop	Hsp90-Hsp70 organising protein
Hsc	Heat shock cognate
Hsc70	Hsc of 70 kDa
HSF	Heat shock factor
Hsj1	*H. sapiens* DnaJ protein 1
Hsj1a	Hsj1 (Isoform a)
Hsj1b	Hsj1 (Isoform b)
HRP	Histidine-rich protein
HSE	Heat shock element
HSP	Heat shock protein family
HSP40	HSP family of 40 kDa

HSP70	HSP family of 70 kDa
Hsp	Heat shock protein
Hsp40	Hsp of 40 kDa
Hsp70	Hsp of 70 kDa
ICAM-1	Intercellular adhesion molecule 1
IPP	Isopentanyl phosphate
IPTG	Isopropyl-β-D-thiogalactopyranoside
kBp	Kilobase pairs
kDa	Kilodaltons
l	Litre
M	Molar
M. musculus	*Mus musculus*
mg	Milligram(s)
ml	Millilitre(s)
mM	Millimolar
mmol	Millimole(s)
mRNA	Messenger RNA
NAC	Nascent-polypeptide associated complex
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NCBI	National Centre for Biotechnology Information
NEF	Nucleotide exchange factor
ng	Nanogram
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
P. berghei	*Plasmodium berghei*
P. chabaudi	*Plasmodium chabaudi*
P. falciparum	*Plasmodium falciparum*
P. vivax	*Plasmodium vivax*
P. yoelii yoelii	*Plasmodium yoelii yoelii*
Pfj1	*P. falciparum* DnaJ protein 1

Pfj4	*P. falciparum* DnaJ protein 4
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PEXEL	*Plasmodium* export element
PfEMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
PfEMP3	*Plasmodium falciparum* erythrocyte membrane protein 3
PfHRP	*Plasmodium falciparum* histidine rich protein
pH	- log $[H^+]$
Pi	Inorganic phosphate
PMSF	Phenylmethylsulphonylfluoride
PV	Parasitophorous vacuole
RAC	Ribosome-associated complex
RESA	Ring-infected erythrocyte surface antigen
RNA	Ribonucleic acid
rpm	Revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SAP	Shrimp alkaline phosphatase
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sHSP	Small heat shock protein
sp.	Species
SPR	Surface plasmon resonance
TBE buffer	Tris-Borate-EDTA buffer
TBS	Tris buffered saline
TBST	Tris buffered saline containing Tween 20
TE buffer	Tris-EDTA buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
TF	Trigger factor
Tris	Tris-2-amino-2-(hydroxynethyl)-1,3-propandiol

U	Unit
UIM	Ubiquitin-interacting motif
UV	Ultraviolet
V	Volt
v/v	Volume/volume ratio
w/v	Weight/volume ratio
X. axonopodis	*Xanthomonas axonopodis*
X. campestris	*Xanthomonas campestris*
X. oryzae	*Xanthomonas oryzae*
YT agar	Yeast tryptone agar
YT broth	Yeast-tryptone broth

Amino Acids	Abbreviation	Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Iso	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Nucleic Acids	Code
Adenosine	A
Cytosine	C
Guanine	G
Thymidine	T

The nomenclature used for the twenty amino acids and the four nucleic acids are listed above as outlined by the International Union of Pure and Applied Chemists (IUPAC) and the International Union of Biochemistry and Molecular Biology (IUBMB).

Chapter 1

Introduction

Chapter 1
INTRODUCTION

1.1. CELLS AND THEIR STRESS RESPONSES

Within the cellular environment at any given time, there are many tasks that a cell must undertake to ensure the correct functioning of all its components so that it may continue to function correctly within its natural environment. The exposure of a cell to cellular stress leads to various stress-responses at the cellular level (Morimoto *et al.*, 1997). Stress responsive genes undergo various levels of regulation directly affecting proteins that may undergo varied production from the norm. This would be in order to protect the cell from fluctuations in homeostasis that could potentially have a negative effect. These cellular stresses are a result of factors including, but not limited to, temperature, pH, exposure to ultraviolet light, toxic chemicals, mutagens and other pharmacologically active molecules (Morimoto *et al.*, 1997).

1.1.1. The protein folding problem

Without the presence of cellular stresses, protein folding is already a challenging process. Cells must ensure that their nascent polypeptides fold into the correct three-dimensional conformations, as this is essential to protein functioning. Each organism has evolved a specific manner to cope with the production of its proteins, whilst the average protein size varies greatly across different organisms: from bacteria (35 kDa in *Escherichia coli*) to eukaryotes (52 kDa in *Homo sapiens*) (Cagney *et al.*, 2003; Mayer and Bukau, 2005).

Nascent polypeptides within a eukaryotic cell are synthesized in a vectorial manner with the N-terminal regions being translated before the subsequent C-termini by the ribosomes (Zimmerman and Monton, 1993; Ellis and Hartl, 1999). Subsequent to synthesis, the protein chain still needs to be folded into its correct three-dimensional conformation and this relies heavily on the manner of protein assembly.

Scenarios available for protein folding are that: (i) all the domains of a protein could be folded separately and then assembled, or (ii) proteins could fold vectorially residue by residue or domain by domain as the protein is being produced. The rate of protein synthesis is slower than the rate of protein folding. If the N-terminal regions of individual domains are folded before whole domains are fully synthesized (Feldman and Frydman, 2000; Pfund *et al.*, 2001) this could result in the incorrect three-dimensional conformation or in the degradation of the protein due to protease susceptibility (Ellis and Hartl, 1999) if the protein is normally folded in another manner.

Proteins have the ability to correctly fold in a few seconds or less but the manner in which nascent polypeptides achieve their tertiary and quaternary conformations in such a limited time frame occurs by an elusive mechanism. It was considered that the three-dimensional folding of a protein should potentially take a long time to fold into the correct conformation if the nascent polypeptide chain amino acid residues had to first adopt all possible conformations at random before settling on the most thermodynamically stable confirmation (Zwanzig *et al.*, 1992). In trying to establish how proteins attain their correctly folded state in such a short time frame from such a vast number of possibilities, Levinthal performed a study of mathematical analysis in the 1960's (Zwanzig *et al.*, 1992). The results of this study were first published in 1968 (Levinthal, 1968). Using a simple model, Levinthal showed that in assigning a small energy bias to locally unfavourable amino acid residue configurations in the three-dimensional situation drastically reduced the randomness of protein folding. This lower number of folding possibilities yielded a three-dimensional folding time that was possible in the *in vivo* situation by excluding those conformations that were assigned an energy bias for the disallowed bond configurations (Zwanzig *et al.*, 1992).

Since the initial studies by Levinthal, it was found that all the steric information needed to synthesize a new protein chain and allow it to fold into its correct three-dimensional conformation within a cell, is contained solely in the primary amino acid residue sequence of the translated nascent polypeptide (Ellis *et al.*, 1998; Ellis and Hartl, 1999; Pfund *et al.*, 2001; Deuerling and Bukau, 2004).

The addition of stresses to cells poses further problems such as above average protein misfolding leading to aggregation and degradation. Protein aggregation within the cell is likely and is sometimes favoured due to the high concentration of proteins and ribonucleic acids (RNA) in the cellular milieu being ordinarily in the region of 300 mg/ml (Zimmerman and Monton, 1993). This phenomenon is termed macromolecular crowding, or also as the excluded volume effect (Young *et al.*, 2004), and occurs due to the increased likelihood of the hydrophobic regions of proteins being transiently exposed on the protein intermediate surface and interacting causing potentially negative interactions. These protein interactions can potentially cause a negative impact on cellular functioning if essential proteins unfold or are misfolded (Morimoto *et al.*, 1997; Ellis, 2001).

1.1.2. Changes at the genetic level

The effects resulting from an applied stress on a cell involve a highly ordered pattern, series or cascade of responses that are genetically defined (Morimoto *et al.*, 1997). Upon stress-induced activation, heat shock factors (HSFs) begin a cascading series of events within the cell. HSFs are highly conserved in structure and consist of a helix-turn-helix DNA-binding domain that is approximately 100 amino acid residues in length (Bonner *et al.*, 1994). The HSFs bind as homotrimers in a sequence-specific manner to the heat shock elements (HSEs), which are conserved promoter DNA recognition sites with at least two or more contiguous inverted repeats of nGAAn (Bonner *et al.*, 1994). The binding of HSFs to HSEs within the promoters can directly cause elevated levels of transcription, which in turn lead to a rapid increase in the production of those inducible gene products (Liu *et al.*, 1997; Morimoto *et al.*, 1997).

The initial eukaryotic experiments involving the study of heat shock stresses were performed on the common fruit fly, *Drosophila bucksii*. Specific polytene chromosome puffing patterns were observed in the larvae due to the chromosomal replication in the somatic stages of the salivary gland cells where the normal house keeping genes of the cell were expressed. However, when larvae were subjected to heat shock temperatures of 30 °C instead of 25 °C, the puffing pattern was observed to have changed (Ritossa, 1962; Tissieres *et al.*, 1974; Ritossa, 1996) due to the shutdown of the house keeping genes and the up-regulation and transcription of another set of mRNAs (Pauli *et al.*, 1992). It was shown that these proteins were also produced in higher quantities when subjected to other cellular stresses as well (Feige and Mollenhauer, 1992) but due to the initial work of heat shock in *Drosophila* these inducible gene products were termed heat shock proteins (HSPs) (Tissieres *et al.*, 1974; Georgopoulos and Welsh, 1993).

Later it was found that the higher HSP production was actually in response to the high amount of unfolded or misfolded proteins that were being produced under stress conditions (Georgopoulos and Welsh, 1993) and the rapid HSP production was important in attempting to allowing the return of the cellular machinery to normal house keeping functions after the removal of a stress. This is important because those stresses that are applied over a prolonged time period, have been known to result in deoxyribonucleic acids (DNA) and RNA production arrest (Morimoto *et al.*, 1997; Liu *et al.*, 1997).

1.2. MOLECULAR CHAPERONES

The term molecular chaperones was originally coined by Laskey *et al.* in 1978 whilst working on nucleoplasmin and testing the principle of protein self-assembly. They found that an excess of this protein aided the formation of nucleosomes and proposed that the protein was able to chaperone nucleosomal formation (Laskey *et al.*, 1978). Ellis later expanded this term to encompass "*molecular chaperones to describe a class of cellular proteins whose function is to ensure that the folding of certain other polypeptide chains and their assembly into oligomeric structures occur correctly*" (Ellis, 1987). Since then, molecular chaperones have also been implicated in protein degradation when proteins cannot be successfully refolded (Chapple *et al.*, 2004; Westhoff *et al.*, 2005; Höhfeld *et al.*, 2006).

1.2.1. Chaperones

Within the cell, chaperones act as regulators of cellular processes ranging from and involving signalling, transcription, translation, translocation, cytoskeletal and organelle networks (Soti *et al.*, 2005). This process involves successive cycles of binding and release followed by molecular chaperone dissociation so as not remain part of the final folded protein product (Ellis, 1987; Beissinger and Buchner, 1998; Young *et al.*, 2004). Chaperones have been found in a variety of systems including prokaryotic, eukaryotic and viral systems (Cheetham and Caplan, 1998; Beissinger and Buchner, 1998).

Chaperones interact and bind to exposed hydrophobic patches, surfaces on proteins, or to protein substrates of similar hydrophobic nature (Richarme and Kohiyama, 1993). The recognised patches were originally thought to consist of small hydrophobic regions (Flynn *et al.*, 1991; Ellis and Hartl, 1999). In the case of the 70 kDa heat shock protein (Hsp70) from *E. coli*, DnaK, this has since been found to consist of a hydrophobic core region of 4 to 5 amino acid residues that is flanked by 2 basic residues (Naylor *et al.*, 1999). Negative residues were excluded from the core of the motif and were not favourable in the flanking regions in experiments using a peptide scanning approach on the *E. coli* Hsp70, DnaK (Zhu and Zhao *et al.*, 1996; Rudiger *et al.*, 1997). The basic flanking residues were thought to form electrostatic interactions with the negatively charged surroundings of the substrate-binding cavity in DnaK (Naylor *et al.*, 1999).

1.2.2. Co-chaperones

Co-chaperones are the partner proteins to chaperones and participate in chaperone function by assisting the chaperones with the refolding of other proteins (Caplan, 2003). This gives them the ability to prevent polypeptide aggregation by binding to the substrate protein and the chaperone concurrently thus mediating specificity of the protein refolding process by the chaperones (Caplan, 2003). There are numerous families of co-chaperones, but the main two are: (i) the 40 kDa heat shock protein (HSP40) family that interact with HSP70 family and (ii) the tetratricopeptide repeat (TPR) motif proteins that interact with HSP70 or HSP90 (Caplan, 2003; Smith, 2004).

Table 1: The main functions of Hsp70 co-chaperones

Co-chaperone	Motif	Function[1]
Hsp40	J	Catalyses the ATP hydrolysis by Hsp70
Hip	TPR	Stabilises the Hsp70-ADP form in the ATPase cycle
Hop	TPR	Acts as an adaptor molecule to connect HSP70-HSP90
CHIP	TPR	U-box domain with ubiquitin ligase activity for ubiquitin-mediated degradation
Tom70	TPR	Mitochondrial pre-protein import
Nucleotide exchange factor		Function[1]
GrpE		Nucleotide exchange factor in prokaryotes and mitochondria of eukaryotes and chloroplasts
Bag1		Nucleotide exchange factor protein family for Hsp70 and proteasome binding in eukaryotes
$Fes1p_{cytosol}$ / $Sls1p_{ER}$		Nucleotide exchange factor in yeast
BAP		Nucleotide exchange factor
Ssa1p		Nucleotide exchange factors in yeast

[1] Adapted and expanded from Kabani *et al.*, 2002; Kabani *et al.*, 2003; Caplan, 2003; Smith, 2004; Höhfeld *et al.*, 2006.

For the sake of clarity, it is important to note that the heat shock protein families are denoted by capitalised HSPs, whilst those individual members of these families, the heat shock proteins themselves are referred to as Hsps.

1.2.3. Heat shock proteins as molecular chaperones

HSPs are found ubiquitously in all organisms (Georgopoulos and Welsh, 1993; Szabo *et al.*, 1994). HSPs and heat shock cognate proteins (HSCs) are the two representative types of families of molecular chaperones (Ingolia and Craig, 1982; Ellis and van der Vies, 1991; Rudiger *et al.*, 1997). The difference between these two types of protein families is that the HSPs are induced and up-regulated with an increase in cellular stress phenomena whilst the HSCs are constitutively expressed regardless of whether stress is present (Ingolia and Craig, 1982; Ellis and van der Vies, 1991). However, since both of these families are highly similar in structure and function, the term HSP is often used collectively to encompass both of these types (Becker and Craig, 1994).

There are two major groups of molecular chaperones that are known to interact with nascent polypeptides: the large group consists of those molecular chaperones that are greater than 200 kDa in size, and the small group of molecular chaperones which consists of those less than 200 kDa in size (Ellis and Hartl, 1999). The small group protein families are named according to their molecular weights of the proteins in kilo Daltons (Table 2): HSP100, HSP90, HSP70, HSP60, HSP40 and the small HSPs (sHSPs).

Hsp100s prevent the aggregation of proteins in solution and also aid in general thermotolerance (Pflanz and Hoch, 2000). Hsp90s have been found to be abundant yet specialised in their function and act with regulatory signal transduction pathway components such as steroid hormone receptors and proto-oncogenic kinases (Agashe and Hartl, 2000; Barral *et al.*, 2004). Hsp70 chaperones are the most highly conserved group important in protein folding, protein translocation and protein assembly into correct oligomeric structures (Gething and Sambrook, 1992; Georgopoulos and Welsh, 1993; Cheetham *et al.*, 1996; Beissinger and Buchner, 1998). Hsp70s function via an adenosine triphosphate (ATP) regulated cycle of substrate binding and release, utilising Hsp40 co-chaperones and nucleotide-exchange factors in this process (Barral *et al.*, 2004). The chaperonin system is made up of the 57 kDa GroEL double-ring complex chaperonin and the 10 kDa GroES lid-like subunit in *E. coli* acting together to capture and fold partially unfolded polypeptides in an Anfinsen cage (Fenton and Horwich, 1997; Beissinger and Buchner, 1998; Barral *et al.*, 2004).

Hsp40s are important as they act as co-chaperones interacting with their Hsp70 chaperone partner proteins to regulate the Hsp70 folding activity in nascent polypeptide chain folding (Georgopoulos and Welsh, 1993; Cheetham *et al.*, 1996). Some Hsp40s have also been implicated in degradation through a ubiquitin-mediated process (Bailly *et al.*, 1997; Young *et al.*, 1998; Höhfeld *et al.*, 2006). The degradation of incorrectly folded proteins is aided by the ubiquitous, ATP-independent sHSPs that often occur as large oligomeric complexes consisting of 12 – 42 subunits that range from approximately 15 – 42 kDa in size (Smith *et al.*, 1998; Barral *et al.*, 2004; Haslbeck, 2006).

Table 2: The groups of HSPs and their main functions

Chaperone family	Approximate size (kDa)	Main functions
HSP100	100	Aggregation prevention and thermotolerance
HSP90	90	Signal transduction
HSP70[1]	70	Precursor stabilisation, folding, translocation, oligomeric structure assembly and degradation
HSP60	57 and 10	Capture within the Anfinsen cage and folds
HSP40[1]	30-45	Folding and interaction with Hsp70s to stimulate ATPase activity, ubiquitin-mediated degradation
sHSP	8-25	Degradation and aggregation prevention

[1] This study focuses mainly on the HSP40 co-chaperone family and to a lesser extent also on the HSP70 chaperone family with which they interact.

Since the main function for molecular chaperones through evolution has been to prevent excess protein aggregation, it is expected that denatured proteins would not be allowed to accumulate freely within the cytosol. It is assumed that different molecular chaperones act on various stages of folding during protein synthesis and they may transfer proteins between themselves using various mechanisms: either (i) along a defined sequential pathway with the evidence for this being seen from studies on rhodanese (Langer *et al.*, 1994) or (ii) in a lateral manner according to experiments on firefly luciferase that have showed transfer back and forth suggesting a network type interaction between proteins and molecular chaperones (Buchberger *et al.*, 1996; Ellis and Hartl, 1999). These networks can be extensive for example the network of 627 yeast proteins and 160 binding partners that are able to interact with yeast Hsp90 in either a physical or genetic fashion (Nardai *et al.*, 2006).

1.3. HSP70S

The HSP70 chaperone family of stress-inducible and constitutive proteins help with tasks ranging from the stabilisation of nascent polypeptides prior to multimeric complex assembly, the translocation into various organelles such as the endoplasmic reticulum (ER) and mitochondria, the oligomeric rearrangement of proteins, the resolution of protein aggregates (Becker and Craig, 1994; Laufen *et al.*, 1999) and the implication in protein degradation through the ubiquitination pathways (Chapple *et al.*, 2004; Westhoff *et al.*, 2005; Höhfeld *et al.*, 2006) or the lysomal pathways (Agarraberes and Dice, 2001).

Table 3: Hsp70s from various systems

Origins	Hsp70 Name	Location in cell	Size (kDa)
Prokaryotes	DnaK	Cytosol	70
Plasmodium falciparum	PfHsp70	Cytosol	70
Yeasts	Ssa1-4p	Cytosol	70
	Ssb1/2p	Cytosol	70
	Ssc1p	Mitochondria	70
	Kar2p	Endoplasmic reticulum	70
Plants	ctHsp70	Chloroplasts	70
Mammals	Hsc72 (Hsp70)	Cytosol/Nucleus	72
	Hsp73 (Hsc70)	Cytosol/Nucleus	73
	Hsp73 (Prp73)	Cytosol	73
	Mtp70	Mitochondria	70
	Grp75	Mitochondria	75
	Grp78 (BiP)	Endoplasmic reticulum	78

Adapted and expanded from Cheetham *et al.*, 1992; Georgopoulos and Welsh, 1993; Becker and Craig, 1994; Cyr *et al.*, 1994; Pfund *et al.*, 2001; Banumathy *et al.*, 2002; Shonhai *et al.*, 2005.

This diverse range of functions for proteins located throughout various organelles in the cells of various evolutionary systems (Table 3) has lead to the association of Hsp70s in a promiscuous fashion with misfolded substrates and to associate more selectively with other folded substrates, with substrate selection being thought to be the role of the co-chaperones (Laufen *et al.*, 1999; Hennessy *et al.*, 2005b).

The Hsp70 expression is highly induced under various types of stress conditions because the proteome differs at any given point in time as the demand for HSPs needed by a cell changes (Cyr *et al.*, 1994). The function is mediated by binding to exposed hydrophobic surfaces of nascent polypeptides preventing disadvantageous interactions between a polypeptides own domains or with other misfolded polypeptides, which could potentially cause detrimental cell development, protein aggregation and misfolding (McCarty *et al.*, 1995). A well-characterised member of the HSP70 chaperone family is the *E. coli* Hsp70, known as DnaK (Bardwell and Craig, 1984). The co-chaperone to this protein is the *E. coli* Hsp40, known as DnaJ (Bardwell *et al.*, 1986).

1.3.1. Hsp70 structure

All Hsp70 proteins consist of two major domains or regions: the N-terminal ATPase domain of approximately 45 kDa and the C-terminal substrate-binding domain of approximately 25 kDa (containing two sub-domains).

Two groups working independently, found that the removal of the N-terminal region or of the C-terminal region in mutants of DnaK lead to the inability to bind to their respective co-chaperone partner, DnaJ, proving that both of these structural domains are required for correct Hsp70 function (Gassler *et al.*, 1998; Suh *et al.*, 1999). The three crucial elements for substrate-binding are the hydrophobic pocket and the arch making up the substrate-binding cavity, along with the helical lid of the adjacent domain (Mayer *et al.*, 2000).

1.3.1.1. The N-terminal ATPase domain

Crystal structures for this domain have been determined for bovine Hsc70 (Flaherty *et al.*, 1990), DnaK with the nucleotide-exchange factor GrpE (Harrison *et al.*, 1997) and Hsc70 with nucleotide-exchange factor Bag-1 (Sondermann *et al.*, 2001), and human Hsp70 (Osipiuk *et al.*, 1999). Recently, the crystal structure of the partial full-length bovine Hsc70 was elucidated (Figure 1; Jiang *et al.*, 2005). This domain has been found to be a highly conserved (Georgopoulos and Welsh, 1993) and has a nucleotide-binding site, as can be seen in the *E. coli* Hsp70 example, DnaK (Flaherty *et al.*, 1990).

The domain has two-fold symmetry with four α/β sub-domains known as IA, IB, IIA and IIB (Flaherty *et al.*, 1990; Beissinger and Buchner, 1998). This symmetry is important in the binding and release of the nascent client-proteins in the ATPase-regulated cycle between the Hsp70s and Hsp40s (Cheetham *et al.*, 1994; Buchberger *et al.*, 1995; Suh *et al.*, 1999).

1.3.1.2. The C-terminal substrate-binding domain

The C-terminal substrate-binding domain has two sub-domains within it. Crystal structures for this domain have been determined for rat Hsc70 (Morshauser *et al.*, 1999), *E. coli* DnaK without substrates (Pellechia *et al.*, 2000) and with substrates (Zhu and Zhao *et al.*, 1996), and also from bovine Hsc70 (Jiang *et al.*, 2005).

The substrate-binding pocket sub-domain consists of β-sheets (Mayer *et al.*, 2000). It is important in the regulation of heat shock responses, binding to substrate and for protein translocation through intracellular membranes and ranges between 15 and 20 kDa in size (Beaucamp *et al.*, 1998).

The α-helical lid sub-domain is approximately 10 kDa in size and is less highly conserved and located C-terminally (Mayer *et al.*, 2000). It was thought to potentially act as a lid structure closing over the associated polypeptide substrate and holding it in place until the folding or refolding of the protein could occur successfully within the cell (Suh *et al.*, 1999). Mayer and co-workers found the C-terminal domain to exist from residues 509-607 in *E. coli* DnaK and proposed that residues 536-538 made up a helical lid hinge that remains open in the ATP state and closed in the ADP state (Mayer *et al.*, 2000). The conserved EEVD motif (King *et al.*, 2001) at the end of the C-terminus has been found to be important in inter-domain regulation between a Hsp70 and a human Hsp40 homologue (Suh *et al.*, 1999). Recent studies, however, by Aron and co-workers in a yeast system suggest another role for this domain. Earlier work on yeast using a modified enzyme-linked immunosorbent assay (ELISA) showed a second site of interaction between the Hsp40, SisI, and the Hsp70, SsaI (Qian *et al.*, 2002) leading to the proposal by Aron and co-workers that these proteins undergo a bipartite interaction with the Hsp70, SsaI (Aron *et al.*, 2005).

The interaction between more than one site (the ATPase domain of Hsp70 with the J-domain of Hsp40, and the C-terminal domain of SsaI with the C-terminus of Hsp40) is suggested to potentially be involved in orientation and alignment of the two proteins to interact with each other assisting the transfer of substrates from the co-chaperone to the chaperone. Structurally, the domain undergoes a conformational change after the client-protein binds forming an α-helical lid over the client-protein binding cleft, although the lid mechanism is not understood (Beissinger and Buchner, 1998; Slepenkov *et al.*, 2003).

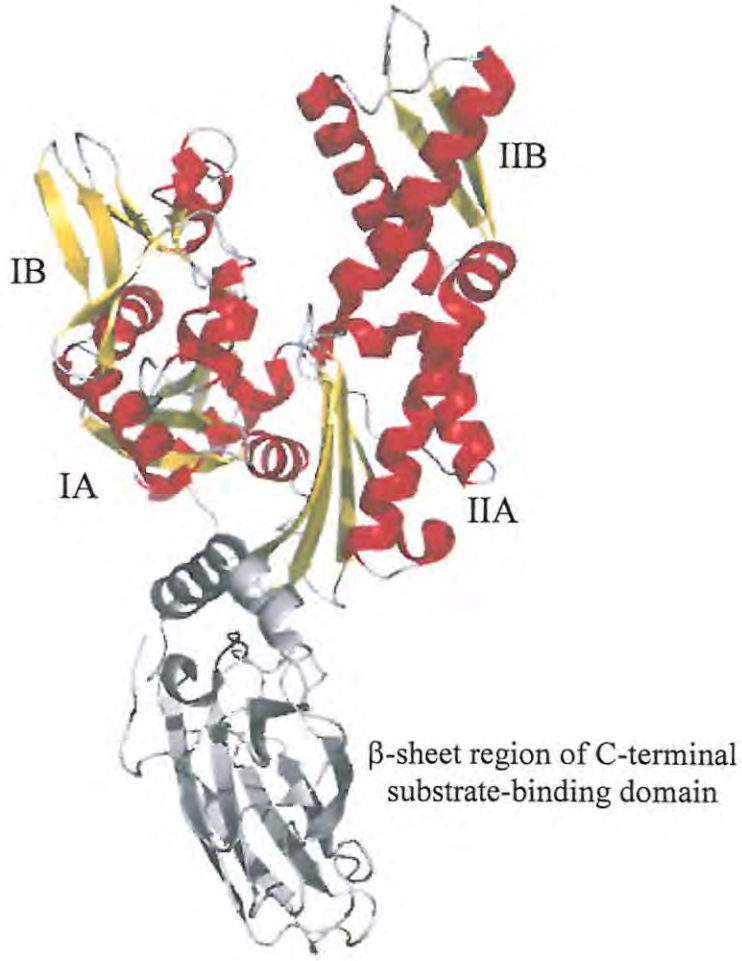


Figure 1: Ribbon representation of the structure of bovine Hsc70

Bovine Hsc70 is depicted with the N-terminal ATPase domain shown in colour and the C-terminal substrate-binding domain of the protein shown in grey. The α-helices of the ATPase domain are shown in red and the β-sheets are shown in yellow. (PDB accession no: 1YUW; Jiang *et al.*, 2005). The figure was visualised in *PyMOL* (DeLano, 2005).

1.3.2. ATPase activity

There are two main aspects involved in ATPase activity of Hsp70s: (i) the hydrolysis activity of the ATP substrate, which is facilitated by the Hsp40 co-chaperones where ATP is hydrolysed to adenosine diphosphate (ADP) and inorganic phosphate (P_i), and (ii) the nucleotide exchange activity, which is facilitated by external nucleotide exchange factors (NEFs) (Russell *et al.*, 1999). Studies on the stress-inducible Hsp70s and on the constitutive Hsc70s have been performed in order to study the ATPase activity of these chaperones (Benaroudj *et al.*, 1996; Minami *et al.*, 1996; Kanazawa *et al.*, 1997; Chamberlain and Burgoyne, 1997a; Edkins *et al.*, 2004).

The release of bound substrate protein from Hsp70 (potentially refolded, misfolded or unfolded) during the refolding, assembly and transport processes is only able to occur due to the presence of ATP (Benaroudj *et al.*, 1996). In the presence of ATP, there is a low affinity for misfolded and unfolded substrate, whilst conversely, in the presence of ADP, there is a high affinity for substrate.

ATP is also needed for the complete dissociation of Hsp70 oligomers to stabilise it in the monomeric form (Benaroudj *et al.*, 1996). ADP binding is able to shift the equilibrium toward the oligomeric species (Benaroudj *et al.*, 1996). A stabilised oligomeric form of Hsp73 was suggested when a study on the interaction of mammalian cytosolic Hsp73 was shown to form stable complexes (Palleros *et al.*, 1991). The equilibrium between these various Hsp70 conformations was found to be regulated by peptides, nucleotides, substrates and co-chaperones (McCarty *et al.*, 1995; Benaroudj *et al.*, 1996; Benaroudj *et al.*, 1997). Hsp70s were found to bind to themselves via the substrate-binding site present within the C-terminal domain as this mimics Hsc70-substrate binding (Benaroudj *et al.*, 1996; Benaroudj *et al.*, 1997).

1.4. HSP40S

The HSP40 co-chaperone family has members that occur ubiquitously throughout cells of various origins (Borges *et al.*, 2005), with at least 44 Hsp40 genes present within the human genome, 34 in the *Drosophila melanogaster* genome, 33 in the *Caenorhabditis elegans* genome, 20 in *Saccharomyces cerevisiae* and 93 in *Arabidopsis thaliana* (Venter *et al.*, 2001). The family is identified as those Hsps of approximately 40 kDa size that are identified by the presence of the J-domain (Bardwell *et al.*, 1986; Ohki *et al.*, 1986). The J-domain is a domain of approximately 70 amino acid residues that bear structural similarity to the first 73 amino acid residues of the *E. coli* Hsp40, known as DnaJ (Bardwell *et al.*, 1986; Ohki *et al.*, 1986).

Hsp40s interact with their respective partner Hsp70s to stimulate the ATPase activity that is required for nascent client-proteins to be cycled to aid in substrate folding (Cyr *et al.*, 1994). They are also involved in protein degradation (Chapple *et al.*, 2004; Westhoff *et al.*, 2005), clathrin uncoating (Fliss *et al.*, 1999; Jiang *et al.*, 2000; Young *et al.*, 2003) and viral infection (Campbell *et al.*, 1997).

The initial characterisation of the function and chaperone activity of Hsp40 co-chaperones was done by analysing DnaJ showing that it was involved in the ATPase cycle with DnaK, as it had the ability to stimulate ATP hydrolysis (Lieberek *et al.*, 1991; Caplan, 2003). DnaJ can also function as an independent chaperone in the absence of the partner Hsp70, DnaK in the *E. coli* system (Langer *et al.*, 1992; Szabo *et al.*, 1994; Borges *et al.*, 2005).

1.4.1. Hsp40 structure

Hsp40s are usually found as dimers, as in the cases of the human Hsp40s, DjA1 and DjB4 (Borges *et al.*, 2005) and yeast Hsp40, Sis1 (Aron *et al.*, 2005). In *E. coli* DnaJ, there are four domains (Cheetham and Caplan, 1998): the J-domain, the glycine/phenylalanine (G/F)-rich region, the cysteine-rich region, and the less conserved and less characterised C-terminal domain (Kelley, 1998).

1.4.1.1. The J-domain

The J-domain is a region of approximately 70 amino acid residues making up four α-helices (helices I – IV). The four α-helices are stabilised by the hydrophobic residues present within the first three helices (Kelley, 1998). The J-domain of the Hsp40 co-chaperone interacts with the partner Hsp70 chaperone (Cheetham and Caplan, 1998) thereby causing the ATPase stimulation of Hsp70 (McCarty *et al.*, 1995; Hennessy *et al.*, 2000).

High-resolution solution NMR structures are known for various DnaJ proteins: the J-domains of a number of Hsp40s such as human DnaJ homologue 1 (PDB accession no: 1HDJ, Qian *et al.*, 1996) and *E. coli* DnaJ (1XBL, Pellechia *et al.*, 1996; 1BQO, 1BQZ, Huang *et al.*, 1999). The J-domain structures show an overall 54 % sequence identity at the amino acid level (Cheetham *et al.*, 1994; Qian *et al.*, 1996; Kelley, 1998). Most J-domains share the very characteristic shape shown below (Figure 2).

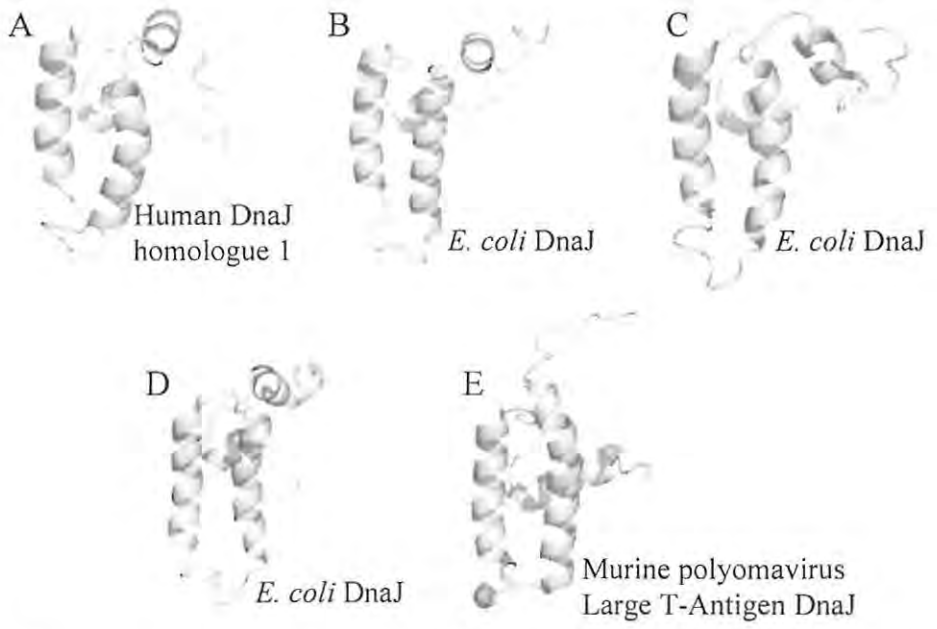


Figure 2: Ribbon representations of the J-domain structures of Hsp40s of various origins

Hsp40s are defined by the presence of the J-domain. (A) Human DnaJ homologue 1 (PDB accession no: 1HDJ, Qian *et al.*, 1996); (B) *E. coli* DnaJ (PDB accession no: 1XBL, Pellechia *et al.*, 1996); (C) *E. coli* DnaJ (PDB accession no: 1BQZ, Huang *et al.*, 1999); (D) *E. coli* DnaJ (PDB accession no: 1BQ0, Huang *et al.*, 1999); (E) Murine polyomavirus Large T Antigen (PDB accession no: 1FAF, Berjanskii *et al.*, 2000). The figures were visualised in *PyMOL* (DeLano, 2005).

A highly conserved HPD motif occurs in almost all J-domains (Wall *et al.*, 1994; Mayer *et al.*, 1999; Suh *et al.*, 1999; Genevaux *et al.*, 2002; Landry, 2003; Hennessy *et al.*, 2005a), with one of the few known exceptions being the substituted YPY motifs in ring-infected erythrocyte surface antigen (RESA) proteins (Bork *et al.*, 1991). The HPD motif is sometimes also referred to as the J-box (Ohtsuka and Hata, 2000) and the residues of this motif are important in Hsp70-Hsp40 functional interaction (Wall *et al.*, 1994; Mayer *et al.*, 1999; Suh *et al.*, 1999). The motif occurs in the solvent exposed loop region between helices II and III of the J-domain (Figure 3). Mutations of these amino acid residues have been shown to cause a loss of functional interaction between the Hsp70 chaperone and the partner Hsp40 co-chaperone (Hennessy *et al.*, 2005a; Hennessy *et al.*, 2005b).

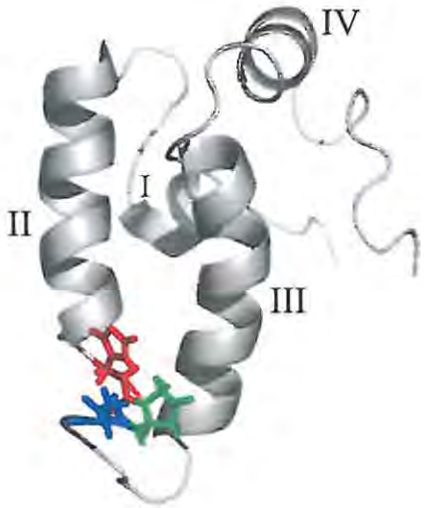


Figure 3: Ribbon representation of the J-domain of Hdj1 showing the HPD motif

The J-domain interacts with partner Hsp70 chaperones. The HPD motif (or J-box) is found within the loop region between helices II and III. The J-domain is shown in grey with the HPD motif residues shown in colour: the histidine residue is shown in red, the proline residue in green and the aspartic acid residue in blue. The four helices are shown as ribbons and are numbered in Roman numerals (PDB accession no: 1HDJ; Qian *et al.*, 1996). The figure was visualised in *PyMOL* (DeLano, 2005).

Other amino acid residues have also been found to be conserved within the J-domains of Hsp40s, as well as the HPD motifs. Some work from previous studies of J-domain residue mutagenesis is outlined (Table 4).

Table 4: Effect of substitutions within the J-domain

Protein	Substitutions	Helix	Effect of substitution
T Ag	L13V	I	Viral DNA synthesis reduction
*Agt*DnaJ	Y7A; L10A	I	Destabilises structure
DnaJ	KQD(3-5)AAA; E8A;	I	No effect
Hdj1-DnaJ chimera	Y24A; RR(25,26)AA	II	Loss of function
T Ag-Ydj1 chimera	W24R; L29P; M30T; Y34N; A37V; C38R	II	Non-functional
*Agt*DnaJ	E20A; K22A; K27A	II	No effect
*Agt*DnaJ	R26A; RK(26-27)AA	II	J-domain activity abolished
DnaJ	RE(19,20)AA; RK(22,23)AA; R27A; L28A; A29G; MK(30,31)AA; Y32A	II	No effect
DnaJ	RK(22,23)AA; KR(26,27)AA; Y25A; K26A; KR(26,27)AA; K26E	II	J-domain activity abolished
DnaJ	EE(17-18)AA	II	Destabilises structure
DnaJ	YKR(25-27)AAA	II	Toxic
*Agt*DnaJ	H33Q	Loop	No ATPase stimulation
*Agt*DnaJ	HD(33,35)YY	Loop	Failure to complement
Hdj1-DnaJ chimera	K35G	Loop	Function affected
Hdj1-DnaJ chimera	N36G	Loop	No effect
DnaJ	RNQ(36-38)GGG	Loop	Failure to complement
DnaJ	H33Q; P34F	Loop	J-domain activity abolished
T Ag-Ydj1 chimera	K51E; K53R; M55T; N56D; Y59N	III	Non-functional
DnaJ	KE(41-42)AA; E44A; K46A; KE(48-49)AA; KE(51-52)AA; Y54A; E55A	III	No effect
DnaJ	F47A	III	J-domain activity abolished
DnaJ	SQ(60,61)AA; KR(62,63)AA; DQ(66,67)AA	IV	No effect
*Agt*DnaJ	D59A	IV	Failure to complement
*Agt*DnaJ	D59N; D67A; D67N	IV	No effect

Adapted from Li *et al.*, 2001; Fewell *et al.*, 2002; Genevaux *et al.*, 2002; Hennessy *et al.*, 2005b.

In addition to individual residues being mutated within given domains, domain swapping experiments have also been done in order to determine how specific interactions are and how protein interactions differ when whole domain units are swapped (Deloche *et al.*, 1997; Campbell *et al.*, 1997; Kelley and Georgopoulos, 1997; Fewell *et al.*, 2002; Genevaux *et al.*, 2002; Hennessy *et al.*, 2005b; Genevaux *et al.*, 2003).

1.4.1.2. GF-rich region

This region is approximately 30 to 40 amino acid residues in length and named due to the high glycine (G) and phenylalanine (F) content. It lacks secondary structure and has been thought to function as a linker region between the J-domain and the rest of the DnaJ molecule, and hypothesized that this region is potentially important in substrate binding to Hsp70 (Szabo *et al.*, 1996; Johnson and Craig, 2001). The J-domain alone was found to be incapable of stimulation of the ATPase hydrolysis activity of DnaK without the presence of the GF-rich region downstream (Karzai and McMacken, 1996). Aron and co-workers have recently suggested two possible alternatives for the function of this region: (i) that the region may have a direct effect on the spatial conformation of the amino acid residues of the J-domain, with differing GF-rich regions causing changes in the overall J-domain conformation, and (ii) that the region could interact directly with the Hsp70s and thus affect the Hsp40 structure (Aron *et al.*, 2005). These hypotheses both correspond to results obtained by Fan and co-workers linked to Hsp70 specificity, namely when the two *S. cerevisiae* Hsp40s (Ydj1 and Sis1) with their central domains swapped between them, they showed a change in their substrate-binding specificities (Fan *et al.*, 2004). However, recent studies by another group suggest that the region is involved in substrate transfer from Hsp40 to Hsp70 (Borges *et al.*, 2005).

1.4.1.3. Cysteine-rich domain

This domain contains four zinc-finger-like motifs (CXXCXGXG, where X can represent any charged or polar amino acid) (Bardwell *et al.*, 1986; Ohki *et al.*, 1986). Various binding and cross-linking experiments (Szabo *et al.*, 1996) show zinc-finger-like motifs are needed to recognise denatured state proteins (Banecki *et al.*, 1996; Lu and Cyr, 1998a). These four motifs form two highly conserved zinc centres (Linke *et al.*, 2003). In the case of DnaJ, there are two zinc ions that interact in a tetrahedral coordination with the sulphur atoms present in the cysteine residues of the motifs (Szabo *et al.*, 1996). DnaJ is stabilised in its correct three-dimensional conformation by the following two structures: (i) zinc centre I is a high-affinity binding site for unfolded proteins and the autonomous DnaK-independent chaperone activity of DnaJ is largely reduced when this is absent and (ii) zinc centre II is essential in the DnaJ functioning mediating interactions between the DnaJ and DnaK where the absence of this causes substrate not to lock onto DnaJ (Lu and Cyr, 1998b; Linke *et al.*, 2003).

Studies with *S. cerevisiae* Hsp40s, namely Ydj1 and Sis1, showed that the structural dissimilarities of proteins yielded two different mechanisms of chaperone interaction (Lu and Cyr, 1998b). SisI which lacks this domain will however assist refolding *in vitro*, although with less efficiency than type-I Ydj1 (Linke *et al.*, 2003). Studies on chimeras of these proteins with domain-swapped cysteine-rich regions suggest that these regions cooperate with the GF-rich regions in functioning (Fan *et al.*, 2005).

1.4.2. Nomenclature and classification considerations

Various nomenclature exists for the HSP40 co-chaperone family. The naming of the *E. coli* Hsp40 as DnaJ was due to the reference to DNA in the initial work. Due to the presence of the J-domain, Hsp40s are also known as the DnaJ-like or DnaJ-related proteins (Cheetham and Caplan, 1998; Borges *et al.*, 2005).

One type of classification attempted to divide Hsp40 proteins into two groups based on their presence or absence of the GF-rich region (Kelley, 1998; Hennessy et al., 2005b). However, the similarity in the domain structure of Hsp40s, has lead to a classification system that is based on the *E. coli* Hsp40, DnaJ (Cheetham and Caplan, 1998). There are three Hsp40 classes based on the presence of the structural domains (Figure 4). Type-I Hsp40s are the most similar proteins to *E. coli* DnaJ, only differing in the C-terminal region, and have autonomous chaperone activity and may function in an Hsp70-dependent or Hsp70-independent fashion (Borges *et al.*, 2005). Type-II Hsp40s only have a J-domain and a glycine/pheylalanine-rich region but differ in their C-terminal region, but have no autonomous chaperone activity of their own and are Hsp70-dependent for activity (Borges *et al.*, 2005). The type-III Hsp40s only have the characteristic J-domain and differ beyond this (Cheetham and Caplan, 1998), and thus do not necessarily bind non-native polypeptides or act as true co-chaperones (Sha *et al.*, 2000).

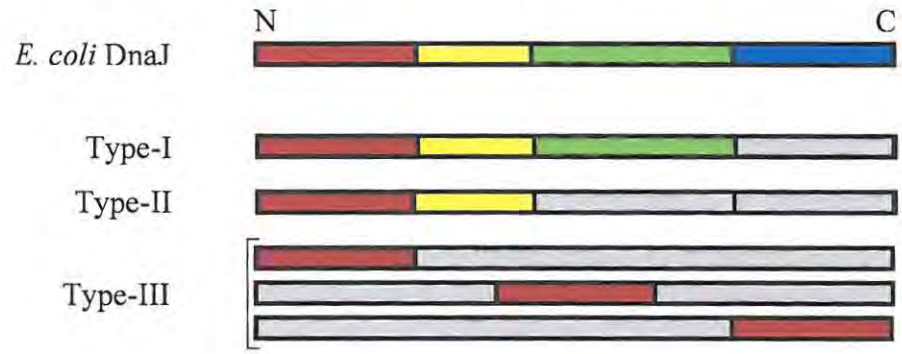


Figure 4: Schematic representation of the three classes of Hsp40s

The classes are shown as likened to *E. coli* DnaJ. The J-domains are shown in red, the GF-rich regions are shown in yellow, the cysteine-rich regions are shown in green and the *E. coli* C-terminal end is shown in blue. The non-domain regions are shown in grey. N and C represent the N- and C-termini of the proteins respectively. This figure is not to scale (adapted from Cheetham and Caplan, 1998; Hennessy *et al.*, 2000; Borges *et al.*, 2005).

In the type-I and type-II Hsp40 groups, the J-domain is N-terminally located, whereas with the type-III Hsp40s, the J-domain can be located at any position throughout the protein. These three Hsp40 classes have also been termed Subfamily A, B and C respectively in some cases (Ohtsuka and Hata, 2000; Borges *et al.*, 2005).

1.5. HSP70 AND HSP40 INTERACTION (ATPASE CYCLE)

Most of the understanding from the mechanism through which Hsp40s interact with Hsp70s is through extensive studies on the ATP hydrolysis activity and nucleotide exchange activity using the *E. coli* system, including the chaperone DnaK, the co-chaperone DnaJ and the nucleotide exchange factor GrpE (McCarty *et al.*, 1995; Cheetham and Caplan, 1998; Laufen *et al.*, 1999). However, regulation of the ATPase cycle is very complex in higher eukaryotes (Caplan, 2003).

In the *E. coli* system, the interaction between the substrate with DnaK in its low affinity, ATP-bound state (McCarty *et al.*, 1995) is initiated when DnaJ interacts with DnaK and transfers the substrate to DnaK (Laufen *et al.*, 1999). ATP is hydrolysis is stimulated by DnaJ in the presence of the substrate (McCarty *et al.*, 1995). The ATP hydrolysis activity locks the substrate into the substrate-binding cavity of DnaK (Liberek *et al.*, 1991; Cheetham *et al.*, 1994) by a closing of the cavity (Laufen *et al.*, 1999). Various cofactors (two K^+ ions and a Mg^{2+} ion) function by interacting with the nucleotides within a conserved region of the N-terminal ATPase domain on DnaK (Wilbanks and McKay, 1995; Naylor *et al.*, 1999). DnaK has a high affinity for substrates in the ADP-bound state with a slow exchange rate (Flynn *et al.*, 1989; Laufen *et al.*, 1999).

Experiments using domain shuffling showed that substrate-binding cavities from different Hsp70s could be interchanged without causing any significant effects on chaperone function suggesting that the Hsp70 domains have no role in determining specificity of chaperone function (James *et al.*, 1997; Naylor *et al.*, 1999). Although conversely, the domain swapping of the central regions of Hsp40s in yeast, have been shown to affect substrate specificities (Fan *et al.*, 2004). These results suggest that specificity is determined by the co-chaperones.

The nucleotide exchange activity involves the dissociation of Hsp70 from the substrate when the nucleotide exchange factors (NEFs) bind to it and accelerate ADP release (McCarty *et al.*, 1995; Russell *et al.*, 1999; Caplan, 2003). In the case of the *E. coli* system, this would involve DnaK releasing ADP when GrpE was bound (McCarty *et al.*, 1995), and in the case of eukaryotes, the binding of Bag-1 would cause the dissociation of ADP from Hsp70 (Takayama *et al.*, 1997). As the Hsp40 dissociates, the substrate is released, with the simultaneous exchange of the bound-ADP for ATP (Laufen *et al.*, 1999).

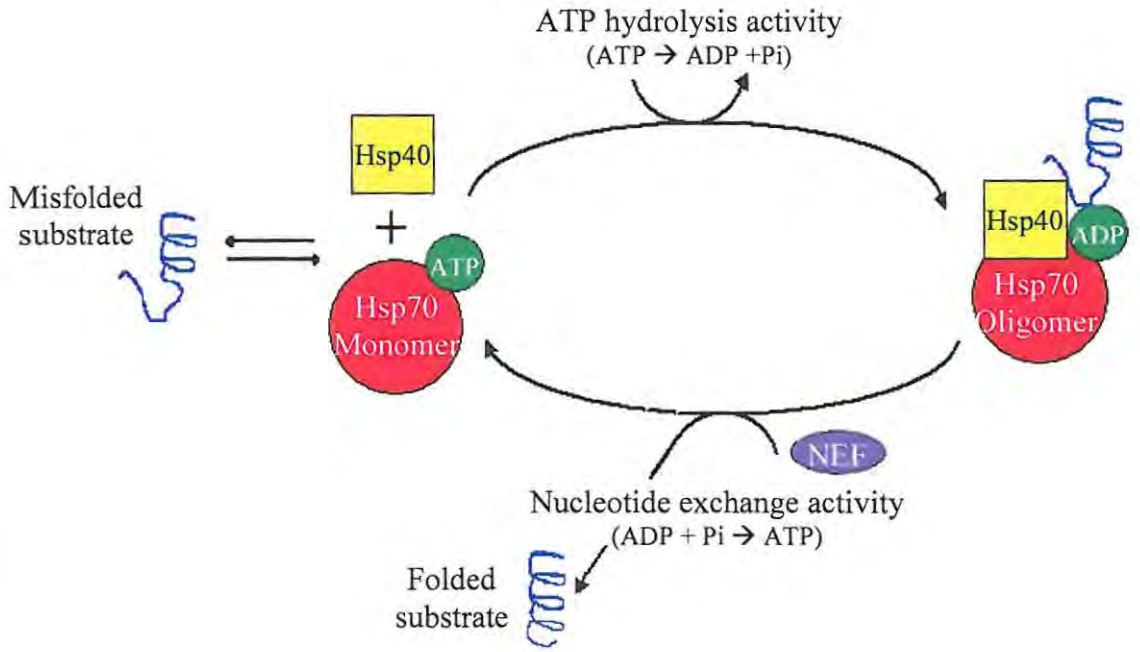


Figure 5: Schematic representation of the ATP cycle of Hsp70, with Hsp40

This shows the mechanism whereby nascent polypeptides and misfolded or unfolded substrates are aided in folding. Hsp70 and Hsp40 are shown in red and yellow respectively. ATP and ADP are shown in green when bound to the Hsp70-Hsp40 complex. The NEF is shown in purple (adapted from McCarty *et al.*, 1995; Gassler *et al.*, 1998; Suh *et al.*, 1999; Feldman and Frydman, 2000).

At a molecular level, the DnaJ concentration is an order of magnitude lower than the DnaK concentration (Laufen *et al.*, 1999). The stoichiometry of DnaJ in relation to DnaK with the addition of a heat stress to the *E. coli* system remains the same, however, there is an up-regulation in the levels showing a 2 to 3 fold increase with stress (Laufen *et al.*, 1999).

1.6. *HOMO SAPIENS* DNAJ 1 (HSJ1)

In 1992, five distinct protein members of the HSP70 chaperone family were known in human cells, however at this point in time, no human homologue of the HSP40 co-chaperone family had been described (Cheetham *et al.*, 1992). Studies by Cheetham and co-workers then isolated two alternatively spliced transcripts of human cDNA homologues of DnaJ from a cDNA library obtained from the human brain-frontal-cortex (Cheetham *et al.*, 1992).

1.6.1. Hsj1 structure

The two transcripts were found to be similar to *E. coli* DnaJ. A J-domain region was detected, as was a GF-rich region in the position corresponding to the GF-rich region of DnaJ. There was no cysteine-rich region present. The C-terminal regions did not bear high sequence similarity to other known DnaJ sequences that were known (Cheetham *et al.*, 1992). These were termed *H. sapiens* DnaJ 1s (Hsj1a and Hsj1b) from the *Hsj1* gene (GenBank accession no. X63368) (Cheetham *et al.*, 1992) and were designated Type-II Hsp40s (Cheetham and Caplan, 1998). More recently, residues in both proteins have been found to represent a putative novel protein interaction chaperone domain (Chapple *et al.*, 2004) and two ubiquitin-interacting motifs (UIMs) have also been identified (Westhoff *et al.*, 2005; Höhfeld *et al.*, 2006).

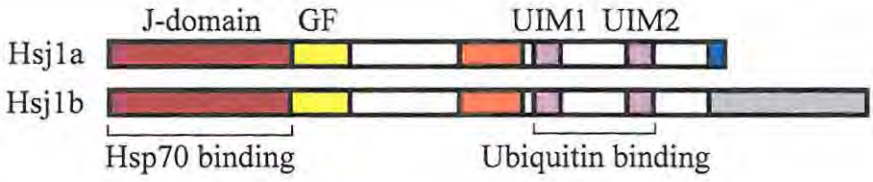


Figure 6: Schematic representation of the domain structure of Hsj1a and Hsj1b

Hsj1a is the shorter isoform consisting of 277 amino acid residues. It contains a J-domain (1-70) shown in red, a GF-rich region (70-91) in yellow, a putative chaperone binding domain (160-207) in orange and two UIMs (210-223: LALGLELSRRE; 249-266: LQLAMAYSLSE) in lilac. The Hsj1a-specific C-terminus is shown in blue and the Hsj1b C-terminus is shown in grey. This figure is not to scale (adapted from Cheetham *et al.*, 1992; Chapple *et al.*, 2004; Höhfeld *et al.*, 2006).

Hsj1a is the smaller of the two spliced isoforms and exists as a protein consisting of 277 amino acid residues. It is cytoplasmic and nuclear in localisation due to the lack of a prenylation motif (Chapple and Cheetham, 2003). Hsj1b is the larger of the two spliced isoforms and exists as a 324 amino acid residue protein. It is targeted to the cytoplasmic face of the endoplasmic reticulum (ER) for correct functioning (Casey and Seabra, 1996) due to the geranyl geranyl farnesylation moiety. The moiety is added at the point of post-translational processing to the cysteine residue of the CaaX box prenylation motif (CLIL) at the C-terminal end of the protein via a thioester linkage, which causes the cleavage of the aaX residues (Zhu *et al.*, 1993; Fahr *et al.*, 1995; Nambara and McCourt, 1999; Chapple and Cheetham, 2003).

1.6.2. Hsj1 characterisation

Distribution studies of the Hsj1s showed protein expression to be restricted almost exclusively to the human brain (mostly to the frontal-cortex and the hippocampus and to a lesser degree to the cerebellum) (Cheetham *et al.*, 1992). Both proteins are also found within photoreceptors (Chapple *et al.*, 2004).

Western blotting studies expected apparent molecular masses of 32 and 38 kDa for the Hsj1a and Hsj1b proteins respectively, yet two bands were detected at 36 and 42 kDa respectively (Cheetham *et al.*, 1992). Initially this phenomenon was assumed to be due to differences attributed to post-translational modification, however, *in vitro* transcription and *in vivo* translation of Hsj1a produced a protein of 36 kDa suggesting that no extensive post-translational modification occurred. This was thought to be attributed to an anomalous electrophoretic ability (Cheetham *et al.*, 1992).

Studies on the physiological function of Hsc70-catalysed clathrin uncoating of vesicles have shown that Hsj1 proteins inhibit the uncoating of clathrin-coated vesicles by interfering with the normal interaction between endogenous vesicle-associated Hsp40s and Hsc70, but not directly with the vesicle suggesting they activate Hsc70 into an altered substrate binding form (Cheetham *et al.*, 1996). Hsj1 has also been shown to effect processing of the archetypal G-protein-coupled-receptor rhodopsin (Chapple *et al.*, 2004).

1.6.3. Hsj1 ATPase activity

Hsj1a and Hsj1b were found to have no ATPase activity of their own, however, they do have intrinsic chaperone activity allowing them to deliver substrate to Hsp70 (Westhoff *et al.*, 2005). The addition of Hsj1a or Hsj1b to porcine Hsc70 ATPase reactions showed the ATPase activity of Hsc70 was enhanced approximately 5 to 6 fold (Cheetham *et al.*, 1994), whereas the DnaJ with DnaK stimulation is approximately two fold (Liberek *et al.*, 1991; Cheetham *et al.*, 1994) and that of yeast Type-I Hsp40, Ydj1, with yeast Hsp70, Ssa1, is approximately ten fold (Cyr *et al.*, 1992; Cheetham *et al.*, 1994). Titrations of Hsj1a and Hsj1b into the Hsc70 ATPase reactions showed maximal stimulation in the equimolar ranges (Cheetham *et al.*, 1994), although Hsj1b was slightly more effective at stimulating Hsc70 ATPase activity than Hsj1a (Cheetham *et al.*, 1994). However, there is no information available on the ATPase stimulation of Hsj1a or Hsj1b of the inducible Hsp70. Similar to the ATPase cycle in *E. coli* (Figure 5), Hsj1a and Hsj1b were found to increase the rate of ATP hydrolysis (Cheetham *et al.*, 1994). They were also found: (i) not to have synergistic effects on ATPase stimulation of Hsc70, (ii) to have a halved Hsc70 ATPase activity in the absence of K^+ and (iii) not to affect the rate of nucleotide release (Cheetham *et al.*, 1994).

1.6.4. Hsj1 links to the degradation pathway

Since chaperones have the ability to bind misfolded and potentially damaged proteins, this suggests that they may also be well suited to be involved in the degradation of those misfolded proteins (Höhfeld *et al.*, 2006). The ubiquitin-proteasome system is the main degradation system involved in the degradation of misfolded or unfolded proteins within the cytoplasm of eukaryotes (Höhfeld *et al.*, 2006). However, proteins other than chaperones have also been found to use this mechanism in degradation, such as the human 26 S Protease Subunit 5a (Young *et al.*, 1998) and the yeast proteins, Rad6 and Rad18 (Bailly *et al.*, 1997).

The Hsp70 chaperone machinery is functionally connected to the ubiquitin-proteasome system by various co-chaperones, for example, through the promotion of polyubiquitination by the U-box motif of the C-terminus of heat shock cognate 70 stress protein-interacting protein (CHIP) (Höhfeld *et al.*, 2001; Chapple *et al.*, 2004). CHIP interacts through a series of tetratricopeptide repeats, negatively regulating Hsp70 ATPase chaperone activity (Chapple *et al.*, 2004; Höhfeld *et al.*, 2006). It is suggested that similar to CHIP, Hsj1s may target proteins for degradation (Chapple *et al.*, 2004; Figure 7).

Numerous proteins containing ubiquitin-interacting motifs (UIMs) have been found to be important in proteasomal sorting (Westhoff *et al.*, 2005). UIMs have been identified in three broad classes of proteins: (i) proteins involved in ubiquitination or ubiquitin metabolism (ii) proteins involved in receptor endocytosis and (iii) proteins where the UIM motif has not been identified.

The UIMs of Hsj1 were identified by bioinformatics and have the conserved sequence FXXAXXXSXXAc (where F is hydrophobic and Ac is acidic) (Hofmann and Falquet, 2001; Chapple *et al.*, 2004). These UIM motifs bind a 76 amino acid protein known as ubiquitin (Vijay-Kumar *et al.*, 1987; Figure 7) to form mono- and polyubiquitylated proteins (Chapple *et al.*, 2004), which are then targeted for degradation by the proteasome (Höhfeld *et al.*, 2006). Hsj1s are thought to fulfil an escort function during the sorting of substrate to the proteasome in the cytoplasm and at the ER membrane (Höhfeld *et al.*, 2006).

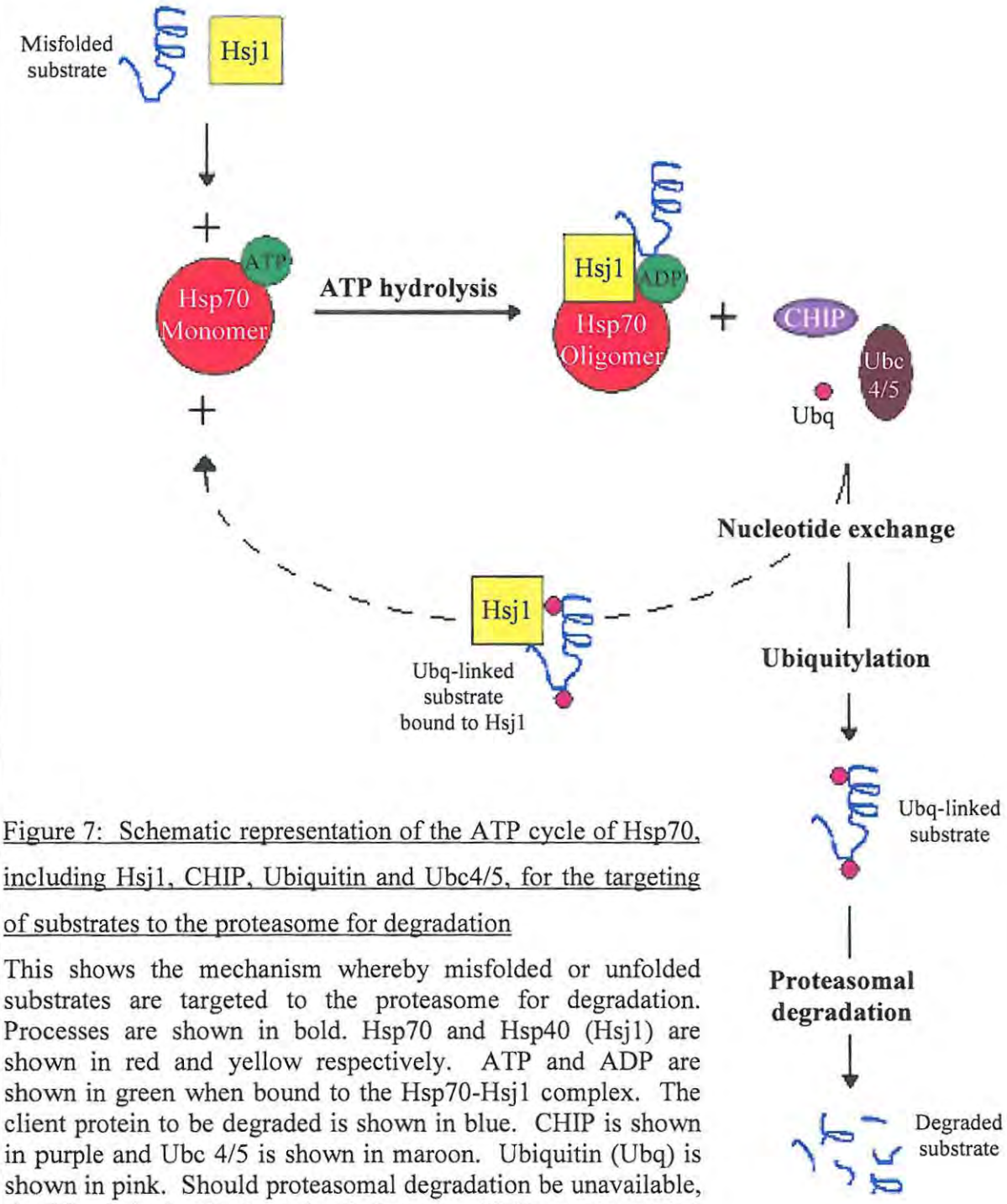


Figure 7: Schematic representation of the ATP cycle of Hsp70, including Hsj1, CHIP, Ubiquitin and Ubc4/5, for the targeting of substrates to the proteasome for degradation

This shows the mechanism whereby misfolded or unfolded substrates are targeted to the proteasome for degradation. Processes are shown in bold. Hsp70 and Hsp40 (Hsj1) are shown in red and yellow respectively. ATP and ADP are shown in green when bound to the Hsp70-Hsj1 complex. The client protein to be degraded is shown in blue. CHIP is shown in purple and Ubc 4/5 is shown in maroon. Ubiquitin (Ubq) is shown in pink. Should proteasomal degradation be unavailable, the Ubq-linked substrates that is bound to Hsj1, can be cycled again. For simplicity, Hsp70 and Hsj1 ubiquitylation is omitted (adapted from Höhfeld *et al.*, 2001; Höhfeld *et al.*, 2006).

The proteasome is a large oligomeric complex with a central proteolytic cavity where the ubiquitin-linked substrates can be degraded separate from the remainder of the cellular contents (Höhfeld *et al.*, 2006). The substrates are sorted and cleaved via a multiple enzyme cascade involving a series of activations and conjugations that are diverse enough to be able to recognise thousands of different substrates in a different manner, although the mechanism in the regulation and balance is unknown (Höhfeld *et al.*, 2006).

1.7. MALARIA AND *PLASMODIUM FALCIPARUM* DNAJ 4 (PFJ4)

Malaria is a disease that is caused by the *Plasmodium* species of parasite infecting the human host (Martinko, 2000; www.who.org). This disease is involved with the transfer of a parasite from a poikilothermic insect vector (with an ambient body temperature corresponding to the range of 25 to 30 °C) to a homeothermic human host (with a body temperature of 37 °C) (Martinko, 2000). The most dangerous form of malaria is caused by *P. falciparum* and causes cerebral malaria that can be fatal (Martinko, 2000). It is likely that the parasitic chaperones, as well as potentially host chaperones, are able to aid the parasite in surviving this heat-shock change and also aid the parasite in its establishment within the various cells that the parasite infects throughout the various stages of its life cycle within the host (Banumathy *et al.*, 2002).

1.7.1. Malarial pathogenesis

P. falciparum infection often results in erythrocyte cytoadherence to the vascular endothelium within the host brain mainly via the remodelling of the host erythrocyte membrane. This process is central to cerebral malarial pathogenesis (Banumathy *et al.*, 2002). The parasite encodes, synthesizes and traffics several proteins to the host erythrocyte membrane to be assembled into knob structure complexes on the host erythrocyte membrane surface to act as sites of attachment to the vascular endothelium (Banumathy *et al.*, 2002). *P. falciparum* erythrocyte membrane proteins 1 (PfEMP1) and 3 (PfEMP3) are found within infected erythrocytes, is known to be an adherence receptor involved in specifically binding to CD36 and ICAM-1 (Baruch *et al.*, 1997; Banumathy *et al.*, 2002). *P. falciparum* histidine-rich protein (PfHRP) is also known to be present in knob structures at the erythrocyte membrane (Banumathy *et al.*, 2002).

Chaperones present within the erythrocyte cytoplasm are involved in the transport and assembly of malarial remodelling proteins destined for the erythrocyte membranes (Sargeant and Marti *et al.*, 2006[1]). Proteins cannot be transcribed within host erythrocytes, yet various chaperones are found within infected erythrocyte cytosols and have been found to be remnants from erythrocytic precursors (Banumathy *et al.*, 2002).

[1] In cases where the authorship of a journal article has been specified as to having two first co-authors, the article has been referenced in text as Co-author1 and Co-author2 *et al.* to give due credit to both first authors.

Banumathy and co-workers used systematic localisation studies to study a suite of chaperone systems that were known to be available to the parasite within the infected host erythrocyte and found no evidence of those parasite-encoded chaperones being exported into the erythrocyte cytoplasm (Banumathy *et al.*, 2002). Co-fractionation studies of detergent-resistant complexes suggested that human Hsp70 was possibly involved in the facilitation of knob-structure assembly (Banumathy *et al.*, 2002).

However, more recently, various laboratories have shown that a short signal sequence exists for the export of certain parasitic proteins via the parasitophorous vacuole (PV) into the host erythrocyte (Sargeant and Marti *et al.*, 2006). This 5 amino acid signal sequence is known as the *Plasmodium* export element (or PEXEL) and has the sequence (R/K)XLX(E/Q) (Hiller *et al.*, 2004; Marti *et al.*, 2004). This signal sequence has allowed for bioinformatic analysis using the algorithm ExportPred to predict the "exportome" of three lineages of *Plasmodium* by the identification of this motif across these lineages (Sargeant and Marti *et al.*, 2006). Analyses identified 43 *P. falciparum* Hsp40-like proteins containing J-domain regions, of which 19 of these proteins were predicted to be exported to the erythrocyte (Sargeant and Marti *et al.*, 2006). GFP-fusion experiments confirmed that a type-II Hsp40 protein (known as PFE0055c) is exported to the erthrocyte cytosol (Sargeant and Marti *et al.*, 2006).

1.7.2. Malarial chaperones

Chaperones and co-chaperones have been characterised in the *P. falciparum* system. At least six Hsp70 chaperones have been identified in this species (Sargeant and Marti *et al.*, 2006), but of these, only PfHsp70 has been characterised and shown to have typical Hsp70 molecular chaperone features (Matambo *et al.*, 2004; Shonhai *et al.*, 2005). For the co-chaperones, initially 4 *P. falciparum* Hsp40s were characterised (also known as PfHsp40s of Pfjs) (Watanabe, 1997). These four proteins were denoted *P. falciparum* DnaJ 1 through 4: Pfj1, Pfj2, Pfj3 and Pfj4 (Watanabe, 1997). More of these Hsp40-like proteins have been identified since the release of the *P. falciparum* genome project (Gardner *et al.*, 2002; Bahl *et al.*, 2003) and in total the 43 Hsp40-like proteins are divided as 2 type-I Hsp40s, 8 type-II Hsp40s and 33 type-III Hsp40s (Sargeant and Marti *et al.*, 2006).

Table 5: Host- and parasite-encoded chaperone localisation in infected erythrocytes

Cells	Chaperone	Site of localisation[1]
Human frontal cortex cells	Hsj1a	Cytoplasmic and nuclear
	Hsj1b	Cytoplasmic face of ER
Infected human erythrocytes	Hsp70[2]	Membrane-associated and cytosolic
	Hsp90[2]	Membrane-associated and cytosolic
	Hip48	Cytosolic
	Hop60	Cytosolic
	Pfj 1-4	Unknown
	PFE0055c	EXPORTED: Cytosolic
Malarial parasites	PfBiP	ER
	PfHsp60	Mitochondrial
	PfHsp70	Cytosolic
	PfHsp90	Cytosolic
	P58 (Hip)	Cytosolic
	Pfj 1-4	Unknown

[1] Adapted and expanded from Banumathy *et al.*, 2002; Chapple and Cheetham, 2003; Sargeant and Marti *et al.*, 2006. [2] A distinctive distribution of chaperones is seen with regards to these chaperone groups in uninfected human erythrocytes (soluble and cytoplasmic) versus infected human erythrocytes (mainly membrane-associated and some cytoplasmic).

1.7.3. Pfj4

Pfj4 is a 244 amino acid protein that falls into the type-II Hsp40 class (Watanabe, 1997). It is a protein encoded by a gene located on chromosome 12 in *P. falciparum* and it is the second most abundant of the Pfj 1-4 proteins (Watanabe, 1997). Pfj4 has similar domain and motif structures to the human type-II Hsp40, Hsj1a, consisting of a highly conserved J-domain, a GF-rich region and similar C-terminal regions.

Results by Banumathy and co-workers had showed that those previously investigated host and parasite Hsps had a non-overlapping distribution between the parasite and the infected host erythrocyte (Banumathy *et al.*, 2002). However, studies on additional Hsps have now shown that malarial co-chaperones such as certain type-II Hsp40s (PFE0055c) are exported into the host erythrocyte cytosol (Table 5; Sargeant and Marti *et al.*, 2006). This suggests that an inter-species chaperone pair could potentially exist that bridges the host/parasite barrier involving a host chaperone and a parasitic co-chaperone, or *vice versa.* These pairs could potentially be key in understanding how the parasite mechanism functions for *P. falciparum.* However, since the localisation and other cell biology of the Hsp70s and Hsp40s (such as Pfj4) is not fully understood, this information must also be elucidated in conjunction with finding functional chaperone pairs.

1.8. PROBLEM STATEMENT

Malaria is a disease causing over 3 million deaths per annum with over 40% of the world's population at risk. Despite this, little is known about the parasitic mechanism of the disease at the molecular level.

Chaperones and co-chaperones have been implicated in cytoprotection when cells are under stress, including disease stresses. The Hsp40 co-chaperones are present in relatively higher amounts in cells under stress. Hsj1 potentially confers cytoprotection to neuronal cells, whilst Pfj4 is found in the cerebral malaria parasite (*Plasmodium falciparum*) within the brain-sequestered infected erythrocytes and in no other *Plasmodium* species. Pfj4 could potentially be important in conferring cytoprotection to the parasite by (i) aiding the establishment of the parasite within the erythrocytes by protecting the parasite from host defences, or by (ii) aiding the establishment of the infected erythrocytes. However, only gene and sequence data are available for Pfj4 and no other information is known about the structure and cell biology of this protein.

Hsj1a and Pfj4 are structurally similar. It is known that both contain the J-domain and glycine/phenylalanine functional domains as well as a less characterised C-terminus. Hsj1a has two UIM motifs, as well as having a putative chaperone-binding motif corresponding to the putative C-terminal motif identified in Pfj4. The similarity between Hsj1 and Pfj4 suggests that these proteins could have similar roles either within host neurons and/or the parasite. Also, the cross-interaction of these proteins across the host and parasite is unknown.

Ultimately downstream, yet outside the scope of this project, if Hsj1 and Pfj4 were to function with their partner Hsp70 proteins by sufficiently different mechanisms of interaction to each other, one or both of these proteins may have potential to be studied further from a therapeutic perspective as potential anti-malarial targets.

1.9. HYPOTHESIS AND BROAD OBJECTIVES

There exists a high level of similarity between the amino acid residues of, and the J-domains themselves, of Hsp40s suggesting some functional similarity. The hypothesis for this thesis suggests that two similar Type-II Hsp40 co-chaperones, namely Hsj1a and Pfj4, have J-domains and selected equivalent residues within these domains, that are functionally similar in their interaction with potential partner Hsp70 chaperones.

The broad objectives of this study were to find a suitable system where the two protein J-domains could be tested and compared in their functional interaction with potential partner Hsp70s, and subsequently, to test specific amino acid residues selected in order to investigate their individual roles in this interaction.

1.10. APPROACH

1.10.1. Bioinformatic analysis

Multiple amino acid alignments were performed in order to select amino acid residues for substitution through studies of the levels of conservation of amino acid residues within the J-domains of various Hsp40 proteins. The presence of other regions consisting of high amino acid residue conservation was also investigated as this could suggest new, putative domains.

1.10.2. Replacement of the J-domain of *Agt*DnaJ with alternate J-domains

The coding regions for the Hsj1a and Pfj4 coding regions were each domain swapped with the coding region of *Agt*DnaJ J-domain and were tested for the ability to reverse the thermosensitivity of the thermosensitive *E. coli* strain, OD259. This was done in order to test whether these J-domains could successfully replace the *Agt*DnaJ J-domain so as to provide a foundation for further mutagenesis studies.

1.10.3. Rational site-directed mutagenesis of selected J-domain amino acid residues

Equivalent highly conserved amino acid residues from both the Hsj1a and Pfj4 J-domains were selected from multiple alignments and reference to literature for substitution by site-directed mutagenesis to investigate and compare the effect of these residues on J-domain function.

1.10.4. Biochemical analysis of Histidine-tagged Hsj1a

The pHsj1a plasmid construct was created in order to over-produce the Histidine-tagged Hsj1a protein. The protein was purified and the oligomeric state of the protein was investigated using Fast Performance Liquid Chromatography (FPLC). The stimulation of the ATPase activity of human Hsp70 was also investigated using a modified form of an ammonium molybdate and ascorbic acid ATPase assay.

Chapter 2

Materials and Methods

Chapter 2
MATERIALS AND METHODS

2.1. MATERIALS

2.1.1. Media, Chemicals and Strains

Growth media were made up using chemicals from the following companies: tryptone and yeast extract (Oxoid, England), sodium chloride (Saarchem, South Africa), agar bacteriological (Biolab Diagnostics, South Africa) and MacConkey agar (Biolab Diagnostics, South Africa). All chemical compounds other than those used in growth media are listed (Table A.1; Appendix 1). Restriction endonuclease companies are given in the text. Antibiotics used were ampicillin and kanamycin sulphate (Roche, Germany) and chloramphenicol (Boehringer-Mannheim, Germany). Various *E. coli* strains were used throughout the study (Table 6).

Table 6: Strains of *E. coli* used in this study

Name	Genotype
E. coli XL1-blue[1]	*rec*A1 *end*A1 *gyr*A96 *thi hsd*R17 *sup*E44 *rel*A1 *lac*$^-$ [F` *pro*AB$^+$ *lac*I^q *lac*Z ΔM15 Tn*10* (*tet*^{r})]
E. coli OD259[2]	MC4100 *ara*D139 Δ*ara*714 Δ*cbp*A ::*kan dnaJ*::Tn10-42
E. coli JM109[3]	*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17 (r_K-,m_K+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F`, *tra*D36, *pro*AB, *lac*I^{q}ZΔM15]

[1] *E.coli* XL1-blue (Bullock *et al.*, 1987; Sambrook and Russell, 2001) was obtained from Stratagene, USA. [2] *E. coli* OD259 (Kelley and Georgopoulos, 1997; Deloche *et al.*, 1997) was a generous gift from Dr Olivier Deloche (Departement de Biochemie Medicale, Centre Medicale Universitaire, Faculte de Medecine, Universite de Geneva, Switzerland). [3] *E. coli* JM109 (Messing *et al.*, 1981; Yanisch-Perron *et al.*, 1985; Sambrook and Russell, 2001) was purchased from Promega, USA.

2.1.2. Vectors and Plasmid Constructs

The pQE30 vector system (Qiagen, Germany) was used throughout this study (Figure A6.2; Appendix 6). This vector system is advantageous as it has a T5 promoter region. Those vector systems with T7 promoters (e.g. pCMV-Tag3a) are problematic as they can only be used in strains such as *E. coli* BL21 (DE3). Vector systems with T5 promoters can be used in strains ranging from *E. coli* XL1-blue, *E. coli* OD259 and *E. coli* JM109. Various pQE30-based plasmid constructs were used for this study (Table 7).

Table 7: Plasmid constructs used in this study

Plasmid name[1,2]	Protein name	Protein description[3]	Size[4]
pRJ30[5]	*Agt*DnaJ	*A. tumefaciens* DnaJ	42.5
pRJ30(H33Q)[5]	*Agt*DnaJ(H33Q)	*A. tumefaciens* DnaJ (H33Q)	42.5
pCMV-3a(Hsj1a)[6]	Hsj1a	*H. sapiens* DnaJ 1a	36
pHsj1*Agt*[7]	Hsj1*Agt*	*A. tumefaciens* DnaJ [Hsj1 J-domain]	42.6
pPfj4*Agt*[8]	Pfj4*Agt*	*A. tumefaciens* DnaJ [Pfj4 J-domain]	42.1
pPfj4*Agt*(Y7A)[8]	Pfj4*Agt*(Y7A)	*A. tumefaciens* DnaJ [Pfj4 J-domain (Y7A)]	42.1
pPfj4*Agt*(L10A)[8]	Pfj4*Agt*(L10A)	*A. tumefaciens* DnaJ [Pfj4 J-domain (L10A)]	42.1
pPfj4*Agt*(R26A)[8]	Pfj4*Agt*(R26A)	*A. tumefaciens* DnaJ [Pfj4 J-domain (R26A)]	42.1
pPfj4*Agt*(H33Q)[8]	Pfj4*Agt*(H33Q)	*A. tumefaciens* DnaJ [Pfj4 J-domain (H33Q)]	42.1
pPfj4*Agt*(D59A)[8]	Pfj4*Agt*(D59A)	*A. tumefaciens* DnaJ [Pfj4 J-domain (D59A)]	42.1

[1] All plasmid constructs listed above are within the pQE30 vector system encoding 6 x Histidine-tagged proteins (except for pCMV-3a(Hsj1a) where the pCMV mammalian expression vector system is used encoding a Tag3a-tagged Hsj1a protein). [2] All plasmid constructs listed above have the gene for ampicillin resistance except the pCMV mammalian expression vector system that has the neomycin resistance gene. [3] Square brackets indicate those swapped domains. [4] Sizes given represent the protein sizes in kiloDaltons (kDa). [5] pRJ30 and pRJ30(H33Q) were created by Dr F. Hennessy (Hennessy *et al.*, 2005a). [6] pCMV-Tag3a(Hsj1a) was a generous gift from Dr M. E. Cheetham (Institute of Opthalmology, University College London, UK) (Chapple and Cheetham, 2003). [7] pHsj1*Agt* was created by Dr G. L. Blatch (Department of Biochemistry, Microbiology and Biotechnology, Rhodes University). [8] pPfj4*Agt* and mutant derivatives were created by Dr W. S. Nicoll (Department of Biochemistry, Microbiology and Biotechnology, Rhodes University).

2.2. METHODS

2.2.1. Bioinformatic analysis of the J-domains of various Hsp40s

Various protein sequence alignments of Hsp40s were performed in order to determine the identities and similarities between the proteins of interest and other similar proteins. Alignments were performed using ClustalW (http://www.ch.embnet.org/software/ClustalW.html; Higgins *et al.*, 1994) and the alignments were shaded using BOXSHADE (version 3.21; http://www.ch.embnet.org/software/BOX_form.html). This multiple alignment was then used to create a phylogenetic tree showing the relatedness of these Hsp40s. Phylogenetic trees were viewed in TreeView (version 1.6.6; Page, 1996).

The primary amino acid sequences for two Type-II Hsp40s were obtained from the National Centre for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov) for Hsj1a (Genbank accession no. CAA44969; Cheetham *et al.*, 1992) and for Pfj4 (Genbank accession no. BAB17689; Watanabe, 1997). A pairwise protein sequence alignment was performed in order to attempt to identify regions that may be previously undiscovered or uncharacterised putative motifs or domains.

The multiple sequence alignment of the J-domain amino acid residues of a small selection of Hsp40s were investigated alongside the Hsj1a and Pfj4 J-domains. Since both Hsj1a and Pfj4 are Type II Hsp40s, these and four other commonly studied Type II proteins were selected. These were included alongside the Type I Hsp40 that was probably the most characterised Hsp40, *E. coli* DnaJ, and the Type I *Agt*DnaJ protein was also included as it was used in this study both as the control protein and as the protein into which the J-domains of interest were domain-swapped into. The primary amino acid sequences for these six Hsp40s were obtained from the NCBI website (http://www.ncbi.nlm.nih.gov) for the following proteins: *Agt*DnaJ (Genbank accession no. AAR84666; Hennessy *et al.*, 2005), DnaJ (Genbank accession no. AAA00009; Ohki *et al.*, 1986), Hdj1 (Genbank accession no. P25685; Raabe and Manley, 1991) and Sis1 (Genbank accession no. P25294; Luke *et al.*, 1991).

Homology models were generated using the online version of SWISS-MODEL (http://swissmodel.expasy.org; Peitsch, 1995; Guex and Peitsch, 1997; Schwede *et al.*, 2003), using the following known NMR structures from the Protein Data Bank (PDB) (http://www.rscb.org) as templates: *Homo sapiens* DnaJ 1 (1HDJ; Qian *et al.*, 1996) and *E. coli* DnaJ structures (1XBL, Pellechia *et al.*, 1992; 1BQZ and 1BQO, Huang *et al.*, 1999). The NMR structure of an *E. coli* DnaJ (1XBL; Pellechia *et al.*, 1992) and the homology models were visualised using PyMOL (http://www.pymol.org; DeLano, 2005).

2.2.2. Competent *E. coli* cell production

The following competent *E. coli* cell production uses a modified calcium chloride production method (Dagert and Erhlich, 1979). An overnight culture (25 mℓ) of relevant cells was grown up in 2 x YT broth (16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl) to early log phase (A_{600nm} = 0.3 - 0.6) with shaking at 200 rpm, at 30 °C for *E. coli* OD259 or at 37 °C for *E. coli* XL1-blue. The cells were collected by centrifugation at 5000 g for 15 minutes at 4 °C. The pellet was resuspended in 1 culture volume ice-cold 0.1 M $MgCl_2$ and incubated on ice for 20 minutes and cells were collected by centrifugation at 5000 g for 15 minutes at 4 °C. Cells were resuspended in ½ culture volume of 0.1 M $CaCl_2$ and incubated on ice for 2 hours. Cells were collected by centrifugation at 5000 g for 15 minutes at 4 °C and were resuspended in $^1/_{10}$ culture volume of 0.1 M $CaCl_2$ and incubated on ice for 30 minutes. An equal volume of ice-cold 30 % sterile glycerol was added to the cells, which were aliquotted and stored at – 80 °C until use.

2.2.3. Transformation of plasmid constructs into competent *E. coli* cells

The following transformation procedure uses a method modified from various protocols (Cohen *et al.*, 1973; Inoue *et al.*, 1990; Tu *et al.*, 2005). The respective plasmid (2 µl) was incubated on ice with 50 µl of the relevant competent cells for 30 minutes. The cells were heat-pulsed for 45 seconds, at 37 °C for *E. coli* OD259 or for 42 °C for *E. coli* XL1-blue or JM109. The cells were incubated on ice for 2 minutes followed by the addition of 900 µl 2 x YT broth and were incubated, at 30 °C for 90 minutes for *E. coli* OD259 and at 37 °C for 1 hour for *E. coli* XL1-blue or JM109.

Samples were centrifuged at 16000 g for 1 minute and 800 µl of the supernatant was removed. The pellet was resuspended in the remaining supernatant volume and this was spread onto 2 x YT agar plates (16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl, 15 g/l agar). Agar plates contained varying antibiotics (100 µg/ml ampicillin for *E. coli* XL1-blue or *E. coli* JM109 transformed cells and 100 µg/ml ampicillin and 35 µg/ml kanamycin for *E. coli* OD259). Plates were incubated overnight, at 30 °C for *E. coli* OD259 or at 37 °C for *E. coli* XL1-blue or JM109.

2.2.4. Plasmid DNA isolation for screening of mutants

Alkaline lysis is a relatively cheap method of obtaining plasmid DNA and was used for the screening of large numbers of colonies (Birnboim and Doly, 1979; Sambrook and Russell, 2001). *E. coli* transformants containing the potential DNA of interest were inoculated into 5 ml 2 x YT broth (with relevant antibiotic; Section 2.2.2) and grown overnight, at 30 °C for *E. coli* OD259 and at 37 °C for *E. coli* XL1-blue, with shaking at 200 rpm. The overnight culture (3 ml) was centrifuged at 16000 g for 30 seconds and the supernatant was discarded. The pellet was resuspended in Solution 1 (100 µl; 50 mM Glucose, 25 mM Tris, pH 8.0, 10 mM EDTA) and allowed to stand for 5 minutes at room temperature. Solution II (200 µl; 0.2 M NaOH, 1 % (w/v) sodium dodecyl sulphate [SDS]) was added and mixed by inversion. Solution III (150 µl; 1.5 M potassium acetate, 12 % glacial acetic acid) was added. Samples were centrifuged at 16000 g for 10 minutes. The supernatant (350 µl) was extracted with 100 µl chloroform and was mixed vigourously. Samples were centrifuged at 16000 g for 10 minutes and 300 µl of the upper phase was mixed with 300 µl isopropanol. This was incubated together for 10 minutes and was centrifuged at 16000 g for 10 minutes. The supernatant was discarded. The pellets were air-dryed and resuspended in 30 µl TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA).

2.2.5. Plasmid DNA isolation for DNA sequencing

DNA was prepared as per the manufacturer's instructions for the QIAprep Spin Miniprep Kit (Catalogue no. 27106; Qiagen, Germany). The constituents of the buffers within the kit are not revealed. Selected colonies were picked and inoculated into 5 ml 2 x YT broth (containing relevant antibiotic). These were grown overnight with shaking at 200 rpm, at either 37 °C for *E. coli* XL1-blue competent cells or at 30 °C for *E. coli* OD259 competent cells. The overnight culture (3 ml) was centrifuged at 16000 g for 1 minute. The supernatant was removed and the pellet resuspended in 250 µl Suspension Buffer P1. Lysis Buffer P2 (250 µl) was then added followed by 350 µl Neutralisation Buffer N3. The samples were centrifuged at 16000 g for 10 minutes. The supernatant was added to the columns and was centrifuged at 16000 g for 1 minute. The flow-through was discarded and 500 µl Buffer PB was added and centrifuged as before. The flow-through being discarded and 750 µl Wash Buffer PE was added and centrifuged as before. The flow-through was discarded and an additional wash step was repeated to remove residual buffer. The DNA from the column was eluted with 2 x 20 µl Elution Buffer (10mM Tris, pH 8.5) for restriction endonuclease analysis samples by centrifugation at 16000 g for 1 minute, whereas samples for DNA sequencing were eluted with 2 x 20 µl sterile distilled, deionised H_2O.

2.2.6. Diagnostic restriction endonuclease analysis

Diagnostic restriction endonuclease analysis uses type II restriction endonucleases that cut specific known sequences. Restriction enzyme digests were made up using 16 µl distilled, deionised H_2O, 2 µl 10 x restriction enzyme buffer, 2 µl DNA (usually of 100 – 300 ng/µl) and 10 U restriction enzyme. Restriction digests were incubated at the relevant temperature for approximately 2 hours. Samples were prepared for agarose gel electrophoresis with 5 x DNA loading buffer (0.25 % bromophenol blue, 30 % glycerol) and were resolved on a 0.8 – 1 % agarose gel made up in TBE buffer (0.045 M Tris, pH 8.3, 0045 M borate, 0.001 M EDTA) at 100 V.

The molecular mass marker (Lambda DNA restricted with *Pst* I; Figure A3.1; Appendix 3) was resolved alongside the samples. Agarose gels were then stained in a TBE buffer with 2 μg/ml ethidium bromide and visualised using a ChemiDoc Imager on a UV setting with the Quantity One software package (BioRad, USA).

2.2.7. DNA sequencing

Double-stranded DNA (200 – 500 ng) was added to 2 μl 5 x Big Dye Terminator (version 3.1) Sequencing Buffer (constituents not revealed), 3.2 pmol pQE30 primer (either forward or reverse; Table A4.2; Appendix 4), 4 μl Big Dye Terminator (version 3.1) and was made up to a 20 μl final reaction volume with sterile distilled, deionised H_2O. The samples were thermal cycled in a GeneAmp PCR System 9700 version 3.05 (Applied Biosystems, USA). An initial denaturation step of 96 °C for 10 seconds was performed before 25 thermal cycles (denaturation at 96 °C for 10 seconds, annealing at 52 °C for 5 seconds and extension at 60 °C for 4 minutes) followed by 37 °C for 2 minutes.

Alcohol precipitation was used to remove excess Big Dye from the sample. EDTA (5 μl of 125 mM) and 60 μl 99.6 % denatured ethanol was added to the samples and these were incubated at room temperature for 15 minutes and centrifuged at 16000 g for 30 minutes at room temperature. The supernatant was removed and 100 μl ice cold 70 % ethanol was added. The samples were centrifuged at 16000 g for 15 minutes at 4 °C. The ethanol was air-dried off at 95 °C for 5 minutes. Dried samples were resuspended in 10 μl Hi-Di Formamide and denatured at 95 °C for 5 minutes. Samples were injected into an ABI-PRISM 3100 Genetic Analyser (Applied Biosystems, USA) for capillary electrophoresis on an 80 cm capillary. Data Collection Software (version 1.0.1; Applied Biosystems, USA) was used. Results were analysed using Chromas (version 2.3; Technelysium, Australia).

2.2.8. Site-directed mutagenesis

Amino acid residues were selected for mutagenesis of the pHsj1*Agt* construct (Table 8) based on the rationale described (Table 10). Site-directed mutagenesis was performed using primers that were engineered to include specific changes at the DNA level to introduce the substituted amino acid residue, as well as specific changed restriction sites included using silent mutations for ease of screening (Table 8).

Table 8: Engineered changes in restriction sites of pHsj1*Agt* with mutagenic substitutions of selected amino acid residues

Substitution[1]	Helix	Forward primer[2]	Reverse primer[2]	Changed restriction sites	Position[3]
Y5A	I	Hsj1(Y5A)F	Hsj1(Y5A)R	Introduction of *Tfi* I site	169
L8A	I	Hsj1(L8A)F	Hsj1(L8A)R	Removal of *Bfa* I site	173
R24A	II	Hsj1(R24A)F	Hsj1(R24A)R	Introduction of *Bsi WI* site	216
H31Q	Loop	Hsj1(H31Q)F	Hsj1(H31Q)R	Removal of *Ban* I site	240
D59A	IV	Hsj1(D58A)F	Hsj1(D58A)R	Introduction of *Sca* I site	315

[1] Numbering used is for that of the Hsj1 J-domain. [2] Further description of the primers can be found in Table A4.3; Appendix 4. [3] Position numbering corresponds to the base pair numbering that is given on the plasmid maps for each of the mutants (Figures 17 – 21).

Site-directed mutagenesis was performed using the QuikChange Site-directed Mutagenesis Kit (Stratagene, USA). Mutagenesis reactions were setup using 5 µl of 10 x reaction buffer (100 mM KCl, 100 mM $(NH_4)_2SO_4$, 200 mM Tris, pH 8.8, 20 mM $MgSO_4$, 1 % Triton X-100, 1 mg/ml nuclease-free bovine serum albumin [BSA]), 5 – 50 ng dsDNA template (pHsj1*Agt*), 125 ng forward primer (Table 8), 125 ng reverse primer (Table 8) and 1 µl dNTP mix (concentration not revealed). This was made up to 50 µl with sterile distilled water. *Pfu* Turbo (2.5 U) was added. Thermal cycling was performed as follows: 95 °C for 30 seconds, 16 subsequent thermal cycles (denaturation at 95 °C for 30 seconds, annealing at 55 °C for 1 minute and extension at 68 °C for 4.5 minutes), followed by a final 37 °C hold for 2 minutes. A sample of 10 µl of this was removed and was treated with 5 x DNA loading buffer for agarose gel electrophoresis analysis on a 0.8 % agarose gel. A 1 µl volume of the *Dpn* I restriction enzyme (10 U/µl; Stratagene, USA) was added to the remainder of the reaction mix which was incubated for 2 hours at 37 °C. Of this post-*Dpn* I digestion sample, 10 µl was removed for agarose gel electrophoresis analysis as above to determine whether the *Dpn* I digestion of the parental DNA was successful. A portion of the post-*Dpn*I sample (2 µl) was then transformed as per the transformation protocol (Section 2.2.3) into *E. coli* XL1 blue competent cells.

2.2.9. *In vivo* complementation assays in the *dnaJ cbpA* strain, *E. coli* OD259

Whilst *E. coli* DnaJ is probably the most well-characterised Hsp40 system (Bardwell *et al.*, 1986; Ohki *et al.*, 1986), this study used an alternative system previously characterised in our laboratory, where *Agt*DnaJ was heterologously produced using an *E. coli* expression system (Boshoff *et al.*, 2004; Hennessy *et al.*, 2005a). This system has been used successfully to study *Agt*DnaJ at both the *in vitro* and *in vivo* levels. Furthermore, *Agt*DnaJ was preferred to DnaJ for certain experiments as *Agt*DnaJ was easier to over-produce in large quantities than DnaJ, and because our laboratory had more available plasmids for comparisons of *Agt*DnaJ than DnaJ, making experimental set up using *Agt*DnaJ less laborious. For these reasons, *Agt*DnaJ was used for J-domain swapping experiments.

Freshly transformed colonies were used for complementation assays. The positive and negative controls (pRJ30 and pRJ(H33Q) respectively), the chimeric constructs (pHsj1*Agt* and pPfj4*Agt*) and mutant plasmids were transformed into *E. coli* OD259 (Section 2.2.3). Single transformant colonies were inoculated into 5 ml 2 x YT broth (100 µg/ml ampicillin and 35 µg/ml kanamycin) for overnight growth with shaking at 200 rpm at 30°C. The overnight culture (1 ml) was diluted 1:100 into 100 ml 2 x YT broth and was monitored until A_{600nm} = 1.8 – 2.0. Samples were diluted until A_{600nm} = 0.2 and tenfold dilutions were made ($10^0 - 10^{-8}$). An aliquot of 2 µl of each dilution was spotted onto 2 x YT agar plates (with 100 µg/ml ampicillin, 50 µg/ml kanamycin and 50 µM isopropyl-β-D-thiogalactopyranoside [IPTG]). Plates were generated in quadruplicate and were sealed with Parafilm M (Pechiney Plastic Packaging, USA) to prevent drying out. The quadruplicate plates were separated and placed for incubation at the following temperatures: 30 °C, 37 °C, 40 °C and 42 °C. Plates were left at these temperatures overnight for > 16 hours, with the 30 °C plate being left for > 20 hours due to slower growth rates at this lower temperature. Bacterial cell growth was recorded using photography with a Canon PowerShot A85 digital camera (Canon, Japan). Each complementation assay experiment was performed three times independently.

2.2.10. Production of the pCMcHsj1a construct

The Hsj1a coding region was excised from pCMV-3a(Hsj1a) by restriction endonuclease digestion with *Bam* HI (Roche, Germany) at 37 °C. The fragment was resolved on a 0.8 % agarose gel in 0.5 x TBE buffer and the fragment of interest was purified using the GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences, England) as per the manufacturer's instructions. The band was excised, weighed and 10 µl of capture buffer was added for every 10 mg of excised gel. The tube was incubated at 60 °C for 15 minutes.

The sample was incubated in the column for 1 minute at room temperature and centrifuged in a bench-top micro-centrifuge for 30 seconds at 16000 g. The column was washed with 500 µl Wash Buffer (Amersham Biosciences, England) and centrifuged for 30 seconds at 16000 g. Elution Buffer (50 µl; Amersham Biosciences, England) was added and incubated at room temperature for 1 minute before centrifuging for 1 minute at 16000 g to elute the purified DNA.

The gel purified Hsj1a coding region was amplified using PCR amplification by a reaction of 200 ng purified Hsj1a coding region, 2.5µl 20 µM Hsj1aF forward primer and 2.5 µl 20 µM Hsj1aR reverse primer (Table A4.1; Appendix 4), 5 µl 10 x buffer (100 mM Tris, pH 8.3, 500 mM KCl,), 1 µl dNTPs (10 mM) and 0.75 µl *Taq* enzyme (1 U/µl), to a final volume of 50 µl. This was cycled as follows: 95 °C for 5 minutes, then 28 cycles of 3 temperatures (95 °C for 30 seconds, 53 °C for 30 seconds, 72 °C for 1 minute), followed by a 72 °C hold for 5 minutes. The Hsj1a coding region with flanking *Bam* HI and *Hind* III sites was ligated into pGEM-T Easy using the pGEM-T Easy Vector System (Promega, USA; Figure A6.1; Appendix 6). The ligation reaction was setup as follows: 5 µl 2 x Rapid Ligation Reaction Buffer (60 mM Tris, pH 7.8, 20 mM MgCl2, 20 mM DTT, 2 mM ATP, 10 % PEG MW8000), 1 µl pGEM-T Easy vector (50 ng), *x* µl Hsj1a mix (3 insert : 1 vector molar ratio) and 1 µl T4 DNA ligase (3 U/µl) to a final volume of 10 µl. The reactions were incubated at room temperature for 1 hour.

The ligation mix was transformed into super-competent *E. coli* JM109 cells (Promega, USA) (Section 2.2.3) but the transformed cells were plated out onto MacConkey-lactose plates (100 μg/ml ampicillin) instead for red/white selection. The *lacZ* gene encodes β-galactosidase, which is an enzyme that converts lactose in MacConey media into glucose and galactose monomers. The gene is disrupted by an insert thus those colonies with inserts will not be able to metabolise lactose media and will have no access to glucose. Since no metabolism can occur, no lactic acid is generated by these colonies to turn the colonies red from the media. Single white colonies were selected for screening and inoculated into 5 ml 2 x YT broth (with 100 μg/ml ampicillin) and grown overnight at 37 °C with shaking at 200 rpm. Plasmid DNA was isolated using the small-scale alkaline lysis method (Section 2.2.4) and the DNA was screened and analysed using by restriction endonuclease analysis. The putative pGEM-T-Hsj1a colonies were then confirmed by DNA sequencing (data not shown).

Subsequently, the Hsj1a coding region was purified from pGEM-T-Hsj1a by excision from an agarose gel using *Bam* HI and *Hind* III diagnostic restriction enzymes. The fragment was ligated into *Bam* HI and *Hind* III cut pQE30 as before for pGEM-T-Hsj1a, to yield pCMcHsj1a.

2.2.11. Over-expression studies of 6 x His-tagged proteins

An overnight 25 mℓ culture was grown in 2 x YT broth (with relevant antibiotic) in *E. coli* XL1 blue at 37 °C or in *E. coli* OD259 at 30 °C with shaking at 200 rpm. This culture was then diluted 1:10 in 250 mℓ 2 x YT fresh broth (with relevant antibiotic) and was grown to log phase (A_{600nm} = 0.6 – 1.0.). Two 1 mℓ samples were taken for analysis at each time point interval. Samples were taken pre-induction. The culture was induced with 1 mM IPTG and post-induction samples were taken hourly, from 1 – 6 hours, and overnight. A 1 ml sample was used to measure absorbance at A_{600nm}, whilst the second 1 ml sample was centrifuged at 16000 g for 1 minute and resuspended in 150 μl PBS (137 mM NaCl, 27 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4) per each 0.5 A_{600nm} absorbance unit. Samples were treated with 5 x Loading Buffer (0.5 M Tris, pH 6.8, 2 % SDS, 0.05 % β-mercaptoethanol, 1 % bromophenol blue, 10 % glycerol) and were boiled for 5 minutes prior to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Section 2.2.12).

2.2.12. Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Proteins were studied using discontinuous SDS-PAGE to separate protein monomers based on their molecular size. Using a modified SDS-PAGE method (Laemmli, 1970), a resolving gel (0.375 M Tris, pH 8.8, 0.1 % SDS, 12 % acrylamide/bis-acrylamide mix (30 : 0.8), 0.1 % ammonium persulphate, 0.02 % TEMED) was set and a stacking gel (0.125 M Tris, pH 6.8, 0.1 % SDS, 4 % acrylamide/bis-acrylamide mix (30 : 0.8), 0.1 % ammonium persulphate, 0.02 % TEMED) was layered on top of the resolving gel. The protein samples and SDS-PAGE Broad Range Markers (Biorad, USA; Table A3.1; Appendix 3) were applied to the sample wells. The gels were placed into BioRad Mini-Protean 3 Apparatus (BioRad, USA) in SDS-PAGE Running Buffer (25 mM Tris, pH 8.0, 250 mM glycine, 0.1 % SDS) and resolved at 150 V for approximately 1 hour. Gels were subsequently placed in Coomassie Stain Solution (0.25 % Coomassie Brilliant Blue G250, 40 % methanol, 7 % glacial acetic acid) from 2 to 5 hours and subsequently in Destain Solution (40 % methanol, 7 % glacial acetic acid) until bands were clearly visible. Both the gels and the cellophane (Sigma-Aldrich, USA) were equilibrated in a Soaking Solution (10 % glycerol, 20 % ethanol) before drying each gel between cellophane.

2.2.13. Western analysis

Western analysis was performed using the ECL Advanced Kit (Amersham Biosciences, England). SDS-PAGE gels to be subjected to Western analysis were resolved in duplicate as described (Section 2.2.12). One gel was stained in Coomassie stain and protein in the second gel was electrophoretically transferred onto Hybond-C Extra Nitrocellulose (Amersham Biosciences, England) at 100 V for 1 hour in a BioRad Western apparatus (BioRad, USA) with Western Transfer Buffer (25 mM Tris, pH 8.3, 192 mM glycine, 20 % methanol) cooled to 4 °C. The post-transfer SDS-PAGE gel was then stained with Coomassie stain to visualise a decrease in the proteins to ensure that protein transfer was successful.

The nitrocellulose membrane was covered in Ponceau S stain (0.005 % Ponceau S, 1 % glacial acetic acid) and all proteins present on the membrane could then be visualised and photographed. Excess Ponceau S Stain was washed off with water and the membrane was placed into the ECL Advanced 5 % Blocking Solution in TBS-Tween (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1 % Tween 20) overnight on ice. The membrane was washed with 4 x 15 minute washes in TBS-Tween at room termperature.

Primary mouse anti-$(His)_6$ antibody (Amersham Biosciences, England) was diluted 1:5000 in ECL Advanced 5 % Blocking Solution and incubated with the nitrocellulose membrane for 1 hour. 4 x 15 minute washes were repeated as above. Secondary anti-mouse IgG antibody conjugated to horse-radish peroxidase (Amersham Biosciences, England) was diluted 1:5000 in ECL Advanced 5 % Blocking Solution and incubated with the nitrocellulose membrane for 1 hour. Washes were repeated again as above. All the antibody incubations and washes were performed at room temperature.

Lumigen TMA-6 Solution A and Lumigen TMA-6 Solution B were mixed in a 1:1 ratio (500 µl total volume) and was applied to the nitrocellulose membrane. The nitrocellulose membrane was immediately imaged using a ChemiDoc Imager with Quantity One software (BioRad, USA) for Western analysis (Filter Setting: I, Image Mode: Chemi) at room temperature and photographed with an 8-bit camera (BioRad, USA).

2.2.14. Purification of His-tagged Hsj1a

2.2.14.1. Nickel sepharose bead regeneration

The Histidine-tagged full-length Hsj1a protein was purified from the lysates of *E. coli* XL1-blue [pCMcHsj1a] construct using nickel chelate affinity chromatography, using a native batch procedure.

A 5 ml 50 % slurry of Chelating Sepharose Fast Flow (Amersham Biosciences, Sweden) was incubated in 2 mℓ Cleaning Buffer I (2 M NaCl) for 5 minutes with rocking at room temperature. The beads were collected at 3000 g for 3 minutes and washed with 1 mℓ Regeneration Buffer (20 mM Tris, pH 8.0, 300 mM NaCl, 10 mM Imidazole) for 5 minutes with rocking at room temperature. The beads were incubated in 2.5 mℓ cleaning buffer II (1 M NaOH) for 5 minutes with rocking at room temperature and were collected as above. The beads were incubated with 1 mℓ Native Lysis Buffer (40 mM Tris, pH 8.0, 100 mM NaCl, 10 mM Imidazole) for 5 minutes with rocking at room temperature and were collected as above. The beads were incubate with 2.5 ml 70 % ethanol for 5 minutes with rocking at room temperature and were collected as above. The beads were incubated with 2.5 mℓ 0.1 M $NiSO_4$ for 5 minutes at room temperature and were collected as before. Three washes with 5 ml water were done collecting beads between each wash as above. The beads were stored as 50 % slurry in Native Lysis Buffer at 4 °C.

2.2.14.2. Lysate preparation for nickel-chelate affinity chromatography

E. coli XL1-blue [pCMcHsj1a] was inoculated into a 25 ml 2 x YT broth (100 µg/ml ampicillin) and was grown at 37 °C at 200 rpm overnight. The overnight culture was diluted 10-fold into 250 mℓ 2 x YT broth (100 µg/ml ampicillin) and was grown to mid-log growth phase (A_{600nm} = 0.6 – 1). The cells were induced for protein production by the addition of 1 mM IPTG. Cells were grown at 37 °C for 5 hours at 200 rpm. The culture was then harvested by centrifugation at 3000 g for 30 minutes and the pellet resuspended in 10 ml Native Lysis Buffer. The cells were aliquotted into 1 ml fractions with 1 µl PMSF (10 mM) added and were frozen overnight at –20 °C. The cells were thawed and cell lysis was achieved by lysozyme treatment (100 µg/ml) for 2 hours on ice and sonication (30 second bursts for 1 minute). Cell debris was removed by centrifugation at 16000 g for 5 minutes at 4 °C.

2.2.14.3. Nickel-chelate affinity chromatography purification

The supernatant is now known as the Cleared Lysate and was added to a 50 % slurry of 1 ml of nickel-charged Chelating Sepharose Fast Flow (Section 2.2.14.2). The mixture was gently rocked overnight at 4 °C. The charged resin was collected by centrifugation at 3000 g for 2 minutes. The resin was washed once with 10 ml Native Wash Buffer (40 mM Tris, pH 8.0, 100 mM NaCl, 100 mM Imidazole). The proteins that were bound to the charged resin were eluted by the addition of two 1 ml volumes and 1 x 0.5 ml volume Native Elution Buffer (40 mM Tris, pH 8.0, 100 mM NaCl, 1 M Imidazole). The beads were also kept for analysis after all the elution steps to determine how much protein was still bound after elutions. All fractions were collected and analysed by discontinuous SDS-PAGE using a 4 % polyacrylamide stacking gel and a 12 % polyacrylamide resolving gel. The proteins were visualised using staining with Coomassie Stain Solution (Section 2.2.12; Laemmli, 1970). No further dialysis and concentration, or G25 chromatography, was needed to remove imidazole from the sample, as it did not appear to affect subsequent ATPase assays.

2.2.15. Protein quantification

Protein quantification was determined using the Bradford's assay (Bradford, 1976). For 5 minutes, 10 µl of protein sample and 200 µl Bradford's Reagent (Sigma-Aldrich, USA) were incubated together. Absorbance was read at 595 nm and the protein concentration was calculated from a BSA (Fraktion V; Roche, Germany) standard curve that was prepared (0 – 1000 ng/µl; Figure A5.1; Appendix 5).

2.2.16. Fast Performance Liquid Chromatography analysis of Histidine-tagged Hsj1a

Molecular mass markers were used for comparative size analysis. Blue dextran (2000 kDa), catalase (232 kDa), aldolase (158 kDa) and human transferrin (77 kDa) were generous donations from Dr A. Boshoff (Department of Biochemistry, Microbiology and Biotechnology, Rhodes University). BSA Fraktion V (66 kDa; Roche, Germany) and chicken egg white albumin grade V (45 kDa; Sigma-Aldrich, USA) were purchased.

The molecular mass markers and the purified Hsj1a protein were analysed on an AKTA basic 10 FPLC (UV-900 monitor; P-903 pump; INV-907 valve; M-925 mixer; 10 mm flow cell; CU-900 PCI) (Amersham Pharmacia Biotech, England) and injected onto a Superdex 200 10/30 HR column (Amersham Pharmacia Biotech, England). The proteins were eluted through the column using FPLC buffer (50 mM Tris, pH 7.5, 0.15 M NaCl, 0.1 % Sodium azide) at a flow rate of 0.5 ml/min collecting 1 ml fractions in a Frac-900 fraction collector (Amersham Pharmacia Biotech, England). The absorbance at 280 nm was used to detect the absorbance due to aromatic residues within the proteins. Results were analysed using Unicorn software (version 4.11 AA; build 211) (Amersham Pharmacia Biotech, England).

2.2.17. *In vitro* ATPase activity assays

Purified protein was analysed by SDS-PAGE to investigate the extent of purity. ATP hydrolysis by human Hsp70 (a kind donation of Dr D. Toft, Mayo Clinic, Minnesota, USA) was measured using a modified method of the ascorbate/molybdate colourimetric assay (Lanzetta *et al.*, 1979; Chifflet *et al.*, 1988; Chamberlain and Burgoyne, 1997a; Chamberlain and Burgoyne, 1997b; Edkins *et al.*, 2004). The assay detects a colorimetric change with the release of inorganic phosphate during the reaction between Hsp70 and a partner Hsp40.

Hsp70 (0.4 µM) was equilibrated at 37 °C for 5 minutes in ATPase Reaction Buffer (10 mM Hepes, pH 7.4, 10 mM $MgCl_2$, 20 mM KCl, 0.5 mM DTT) prior to starting the reaction with 600 µM ATP. Samples of 50 µl were removed at regular time intervals (0 minutes; 30 minutes; and 1 – 5 hours) and were added to 50 µl 10 % SDS in a 96-well microtitre plate. The effect of the Hsp40s on the basal human Hsp70 activity were similarly determined but with the addition of equimolar concentrations (0.4 µM) of Hsj1a, or Pfj4 (a kind donation of Mr A. Shonhai, Rhodes University). Control reactions were setup to monitor spontaneous ATP degradation and comprised a reaction with water replacing the Hsp40 and/or Hsp70. A potassium hydrogen phosphate standard curve was prepared (0 – 10 nmol phosphate; Figure A5.2; Appendix 5) using K_2HPO_4 and KH_2PO_4. For colorimetric detection of phosphate, 50 µl 1 % ammonium molybdate (in 1 N HCl) and 50 µl 6 % ascorbic acid was then added to the stopped samples and incubated for 45 minutes at room temperature. To this, 125 µl 2 % sodium citrate/2 % acetic acid was added to develop at room temperature for 30 minutes. The A_{850nm} was measured.

Chapter 3

Results

Chapter 3
RESULTS

3.1. BIOINFORMATIC ANALYSIS

3.1.1. Preliminary bioinformatic analysis of Hsp40s

Pfj4 was detected through protein BLAST (BLASTp) as being similar to the human Hsj1a. This similarity to a human neuronal protein, and the fact the little is known of the cell biology of Pfj4, is interesting as the possibility exists for the involvement of these two proteins in the malarial scenario. Results of a BLASTp of Pfj4 showed no other similar proteins from any other *Plasmodium* species and hence this protein is only found within *Plasmodium falciparum* that is known to cause the sometimes fatal, cerebral form of malaria.

A multiple alignment of four type-II Hsp40s from various evolutionary origins (yeast, human and parasitic) and two type-I Hsp40s from prokaryotic origins was performed (data not shown). A phylogram was generated from this multiple alignment in order to determine the potential evolutionary relationships between these selected Hsp40s (Figure 8). As was expected, the type-I proteins of prokaryotic origin were found to be the most distant from the type-II Hsp40s of eukaryotic origin. Also interestingly, as expected, from BLASTp results, the Hsj1a and Pfj4 proteins were the most closely related of the analysed type-II Hsp40s.

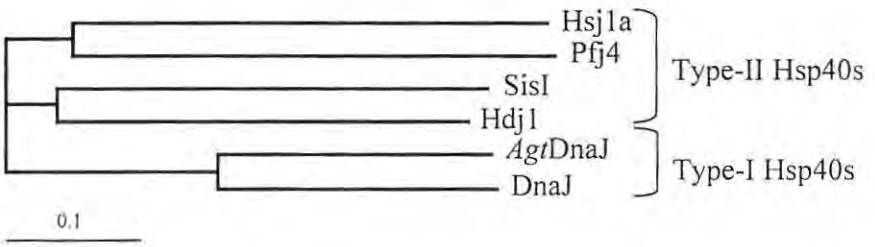


Figure 8: Phylogram showing the relation of selected Hsp40s of various origins

H. sapiens DnaJ 1a (Hsj1a; GenBank accession no. CAA44969), *P. falciparum* DnaJ 4 (Pfj4; GenBank accession no. BAB17689), *S. cerevisiae* DnaJ (Sis1; GenBank accession no. P25294), *H. sapiens* DnaJ 1 (Hdj1; GenBank accession no. P25685), *A. tumefaciens* DnaJ (*Agt*DnaJ; GenBank accession no. AAR84666), *E. coli* DnaJ (DnaJ; GenBank accession no. AAA00009) are shown. The phylogram was generated of the six selected Hsp40s using ClustalW and was viewed using TreeView (v. 1.6.6; Page, 1996).

All of the above Hsp40 co-chaperones are of similar function, aiding the functioning of their partner Hsp70 chaperones (except for Pfj4 whose cell biology and function is not yet known). A comparison of the percentage identities of the amino acid residues comprising these proteins, at both the full-length and J-domain levels, show the similarity of the proteins and their J-domains, and hence a possible relation to protein function.

Table 9: Percentage identities[1] of full-length Hsp40s[2] and their J-domains[3]

	*Agt*DnaJ	DnaJ	Sis1	Hdj1	Hsj1a	Pfj4
*Agt*DnaJ		58	28	27	22	21
DnaJ	65		26	26	22	22
Sis1	52	52		36	22	20
Hdj1	48	50	50		25	25
Hsj1a	48	54	47	58		27
Pfj4	47	48	44	44	51	

[1] A multiple alignment of the six Hsp40s was performed and pair-wise percentage identities were obtained for each pair-wise protein combination using ClustalW. [2] Those numerical values represented above the diagonal line represent the amino acid identities of the full-length Hsp40s. [3] Those numerical values represented below the diagonal line represent the amino acid identities of the J-domain regions of the Hsp40s only.

It can be seen that the level of identity between these full-length Hsp40s ranges from 20 % to as high as 58 % (Upper half; Table 9). The highest percentage identity is shared by the two type-I prokaryotic proteins, DnaJ and *Agt*DnaJ, as was expected from the knowledge of the domain structure of the type-I proteins (Cheetham and Caplan, 1998) and from the phylogram results (Figure 8). The type-II proteins share significantly lower percentage amino acid identities across their full-lengths than the type-I proteins due to a greater difference at the level of domain structure, ranging from 20 – 36 % (Table 9). The two type-II Hsp40s of interest, Hsj1a and Pfj4, share an amino acid identity of 27 % at their full-length amino acid level (shaded grey; Table 9).

The J-domains of Hsp40s are more highly conserved than the C-terminal remainders of the proteins (Cheetham and Caplan, 1998; Hennessy *et al.*, 2000) so the amino acid identities of these domains are significantly higher than the percentage identities of the full-length proteins. The identities between the J-domains range between 44 % and 65 % (Lower half; Table 9). The prokaryotic type-I proteins share a 65 % pair-wise amino acid identity across their J-domains. The human Hsj1a and malarial Pfj4 share a high pair-wise J-domain amino acid identity of 51 % (shaded grey; Table 9).

Interestingly, the two human type-II proteins, Hsj1a and Hdj1, share a 58 % pair-wise J-domain amino acid identity, which is higher than that of Hsj1a and Pfj4, even though these proteins are not as closely related as Hsj1a and Pfj4 according to the phylogram (Figure 8). This could suggest that a high level of conservation across the J-domains could be important in the species specificity of Hsp40s, although this hypothesis needs to be tested.

3.1.2. Conserved domains and motifs

In general, a high level of conservation exists across proteins of a family that are functionally related, for example, those members of the HSP40 co-chaperone family. This conservation suggests a possible importance in functional interaction. Due to the potential evolutionary relatedness and conservation of Hsj1a and Pfj4 to each other than to other Hsp40s, this could suggest a similarly of function and/or the mechanism of interaction for these two proteins in their respective host and parasite systems. For this reason, the conservation of these two Hsp40s was investigated further to identify potentially important domains or motifs that could be fundamental in their function.

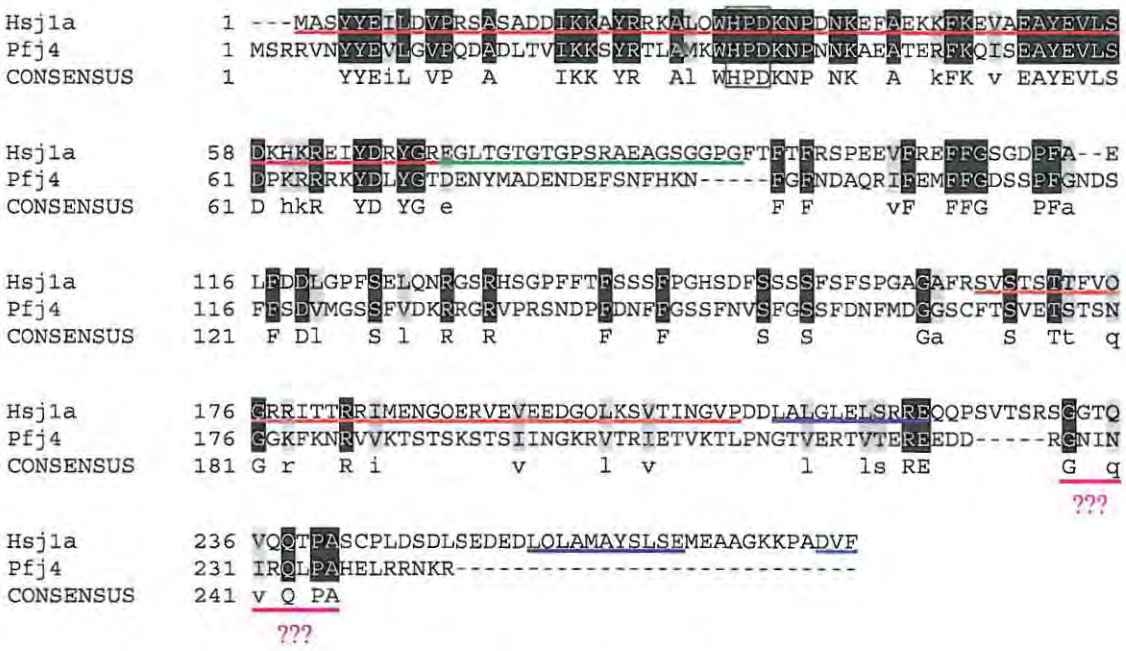


Figure 9: Pair-wise alignment of two type-II Hsp40 proteins, Hsj1a and Pfj4

The amino acid residues of *H. sapiens* DnaJ 1a (Hsj1a; GenBank accession no. CAA44969) and *P. falciparum* DnaJ 4 (Pfj4; GenBank accession no. BAB17689) are pair-wise aligned. Identical amino acid residues are shown in black. Similar amino acid residues are shown in grey. As defined for Hsj1a: The red line shows the J-domain, whilst the bright green line shows the previously proposed GF-rich region. The dark green box shows the HPD motif. The orange line shows the putative chaperone-binding domain and the purple lines show the two Ubiquitin Interacting Motifs (UIMs). The blue line shows the C-terminus that is specific to Hsj1a. Dashes indicate gaps due to the alignment program. Aligned using ClustalW, where the consensus is given as letters with the identical amino acid residues shown as capitals, whilst the similar amino acid residues are shown as lower case highlighting the Hsj1a residue. For both Hsj1a and Pfj4: The pink line highlights a conserved region not previously identified but which may represent an interesting motif of unknown function (question marks). Shaded using BOXSHADE (v. 3.21).

The pair-wise alignment of Hsj1a and Pfj4, that are potentially functionally similar, is highlighted with the known regions of Hsj1a in order to investigate the presence of similar regions in Pfj4 (Figure 9). A region of high identity was found across the N-terminus at the J-domain (red) as was expected from previous J-domains studies showing levels of high conservation (Hennessy *et al.*, 2000). The C-terminal remainder of the protein showed lower similarity, being fairly evenly interspersed with similar residues, except for those three regions corresponding to: (i) the initial region of the putative chaperone-binding domain corresponding to residues 166-209 in Hsj1a (orange), (ii) the UIM1 domain corresponding to residues 212-222 in Hsj1a (purple), and (iii) a region of clustering of conserved amino acid residues seen corresponding to residues 232-241 in Hsj1a and residues 227-236 in Pfj4 (pink).

A GF-rich region for Hsj1a was previously identified and described by Cheetham and co-workers (Chapple *et al.*, 2004). The lower level of conservation across the GF-rich region of Hsj1a (bright green; Figure 9) is due to the fact that no glycine (G) and phenylalanine (F) residues are found within the region corresponding to the first half of this GF-rich region in Pfj4. Only 2 F residues in Pfj4 occur within this corresponding Hsj1a GF-rich region, with the first F residue only appearing at residue position 85 in Pfj4. This F residue actually appears to be the start of an extensive F and serine (S) region extending from residues 85 to 165 of Pfj4. This corresponds to almost the entire region of Hsj1a between the J-domain and putative chaperone domain (residues 94 to 169) containing a high level of F and S residues. This suggests that this region should be termed a FS-rich region in Hsj1a and Hsj1b.

Based on this knowledge, a few scenarios exist for Pfj4: either (i) the GF-rich region exists as predicted leaving Pfj4 as a type-II protein and it needs re-naming in this case, (ii) the boundaries of the GF-rich region of Pfj4 need to be re-evaluated to incorporate those F residues and not correspond to Hsj1a as per Cheetham and Caplan's designation, or (iii) Pfj4 should be considered as having a *pseudo* GF-region. However, in order to keep with that Hsp40 nomenclature and domain structure as defined by Cheetham and Caplan in 1998, it is suggested that Pfj4 be known to have a *pseudo* GF-rich region (i.e. a lengthy region [residues 85 to 165] that occurs after the J-domain but is actually a GFS-rich region).

The region of highest conservation within the initial portion of the putative chaperone-binding domain of Hsj1a from residues 166-185, could suggest that the equivalent region of Pfj4 (FTSV...NRVV) may also be a putative chaperone-binding domain within Pfj4, or alternatively the equivalent residues of the entire putative chaperone-binding domain within Pfj4 (FTSV...VKTL) could be involved.

The region of Pfj4 that corresponds to the eleven amino acid residues comprising the Hsj1a UIM1, has two identical and three similar amino acid residues to Hsj1a. This could also be an equivalent UIM-like motif (Hofmann and Falquet, 2001) potentially targeting ubiquitin-linked malarial proteins for degradation within the parasite in a similar fashion to Hsj1a in human cells (Banumathy *et al.*, 2002; Westhoff *et al.*, 2005). It has been suggested that such a short sequence motif is probably likely to form a short α-helix that can be embedded into different protein folds (Hofmann and Falquet, 2001).

Through this study, another region of conservation was also detected by the pair-wise alignment of Hsj1a and Pfj4. The region corresponds to amino acid residues 232-241 in both Hsj1a and Pfj4 and has a pairwise consensus of GXX(Q/N)(V/I)XQXPA, where X represents any amino acid residue. A BLASTp of the potential motif from Pfj4 (GNINIRQLPA) using the "nearly exact short sequences" option, showed that other proteins from various evolutionary backgrounds contained regions of relatively high similarity to this potential motif: *P. yoelii yoelii* (GenBank accession no. XP_728590), *P. chabaudi chabaudi* (GenBank accession no. XP_742451), *P. berghei* (GenBank accession no. 676696), *Xanthomonas campestris* (GenBank accession no. AAY48210), *X. axonopodis* (GenBank accession no. AAM37987), *X. oryzae* (GenBank accession no. BAE68306), *Gibberella zeae* (GenBank accession no. XP_389207), *Danio rerio* (GenBank accession no. CAI21342), *H. sapiens* (Hsj1a; GenBank accession no. CAA44969), and others. Similarly, for Hsj1a, proteins were detected from the following evolutionary backgrounds: *Haloarcula marismortui* chaperone DnaJ (GenBank accession no. YP_137730), *Mus musculus* T cell receptor (GenBank accession no. AAA64783), *P. vivax* meroziote surface proteins (GenBank accession no. AAV71114), and others. Whilst this region was identified, no further work was continued on this due to the time constraints of the project. Hence, it is yet to be described as a true type of functional motif although the conservation across a wide range of evolutionarily diverse proteins is interesting as it may play a more general function than a chaperone-specific one.

3.1.3. J-domain analysis

In continuing the study of the J-domain, we will to try to establish the importance of those individual amino acid residues within the J-domain for Hsj1a. Since Hsj1a is of known function and cell biology, by comparison to the same mutations to the equivalent residues of Pfj4, we propose to potentially extrapolate residue function for Pfj4.

Hennessy and co-workers showed that a significant overall amino acid residue conservation exists between evolutionary diverse organisms using multiple alignments of numerous J-domains from various organisms (Hennessy *et al.*, 2000). When focussing on the first 70 amino acid residues forming the J-domain regions of the selected Hsp40s (Figure 10), as expected, it can be seen that many amino acid residues are conserved within this region.

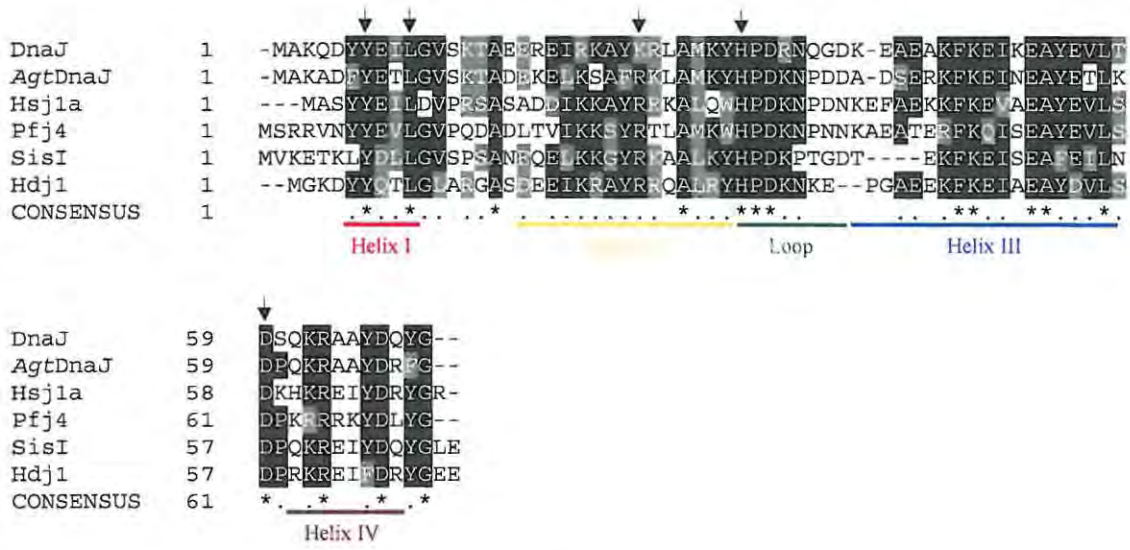


Figure 10: Multiple alignment of the J-domains of various Hsp40s

The amino acid residues of the J-domains of *E. coli* DnaJ (DnaJ; GenBank accession no. AAA00009), *A. tumefaciens* DnaJ (*Agt*DnaJ; GenBank accession no. AAR84666), *H. sapiens* DnaJ 1a (Hsj1a; GenBank accession no. CAA44969) *P. falciparum* DnaJ 4 (Pfj4; GenBank accession no. BAB17689), *S. cerevisiae* DnaJ (Sis1; GenBank accession no. P25294), *H. sapiens* DnaJ 1 (Hdj1; GenBank accession no. P25685) are aligned. Identical amino acid residues are shown in black. Similar amino acid residues are shown in grey. Helix I is underlined in red, helix II in orange, the loop region in green, helix III in blue and helix IV in purple as previously defined (Genevaux *et al.*, 2002). Those amino acid residues selected for mutagenesis are marked above with black arrows. Dashes indicate gaps due to the alignment program. Aligned using ClustalW. Shaded using BOXSHADE (v. 3. 21).

It can also be seen that a higher level of conservation exists within the four α-helical secondary structures that comprise the J-domain, whilst relatively fewer conserved residues are found within the non-helical regions. Of the helical residues, 76 % are conserved, whilst of those non-helical residues (excluding the HPD motif) show only 47 % conservation. For reasons of conservation, and the fact that secondary structures are important in the overall conformation of the protein affecting the orientation of residues within the J-domain, the amino acid residues selected to be investigated in this study by mutagenic substitution were selected from these α-helical regions.

The only exception to the α-helical selection was the histidine (H) residue of the HPD motif that is found within the loop region of the J-domain and has been found to be important in functional interaction with partner Hsp70s in all previously characterised Hsp40s known to contain this motif. The rationale for these specific substitutions are given below (Table 10).

Table 10: Rationale for amino acid substitutions to be performed in the Hsj1 J-domain to investigate the function of these potential key amino acid residues

Number[1]	Amino acid	Nature	Rationale	Substitution[2]
5	Tyrosine (Y)	Aromatic	Potential structural residue	Alanine (A)
8	Leucine (L)	Hydrophobic	Potential structural residue	Alanine (A)
24	Arginine (R)	Positive	Potential functional residue	Alanine (A)
31	Histidine (H)	Positive	Negative control	Glutamine (Q)
58	Aspartic acid (D)	Negative	Potential functional residue	Alanine (A)

[1] Numbering used is for that of the Hsj1 J-domain. [2] Alanine is often used for mutagenic substitutions due to (i) this amino acid being unlikely to introduced new charged interactions that may potentially negatively impact the structure of the protein, and (ii) this amino acid having a side chain of significant size whereas amino acids with bulky side chains could cause further structural hindrances.

Previous structural data have shown that helix II and III of most J-domains are relatively well-ordered α-helices anti-parallel to one another, flanking disordered solvent-exposed loop regions that each contain an HPD motif (Pellechia *et al.*, 1996; Qian *et al.*, 1996; Huang *et al.*, 1999; Berjanskii *et al.*, 2000). This similar overall conservation in folding has suggested an evolutionary conserved mechanism for the interaction of the J-domain with Hsp70 (Genevaux *et al.*, 2002), and the importance in inter-molecular and intra-molecular force formation, and thus potentially in function (Hennessy *et al.*, 2000; Genevaux *et al.*, 2002).

3.1.4. Homology modelling

Whilst homology modelling is used as a tool as a rough guide to investigate the potential structures of proteins and potential orientations of amino acid residues, as a general rule, crystal or NMR structures should be used for this type of investigation. However, with structures not being available for all proteins, homology modelling is the closest approximation available for this task and since the J-domains in this study share a high level of conservation (above 44 %; Table 9), these should serve as good approximations of the potential protein structures and residue orientations. Thus, in order to investigate these properties before mutagenic substitutions of Hsj1a and Pfj4 (for which structures are not known), homology modelling of the J-domains of *Agt*DnaJ, Hsj1a and Pfj4 was performed in comparison with the NMR structure of DnaJ (Pellechia *et al.*, 1996).

The homology modelled J-domains of *Agt*DnaJ, Hsj1a and Pfj4 all show a folded structure similar to that of the structure of DnaJ (Figure 11A), as was expected. Each J-domain consists of the four α-helices (Figure 11) forming this folded shape of helices II and III flanking a loop region that has come to be associated with most J-domains.

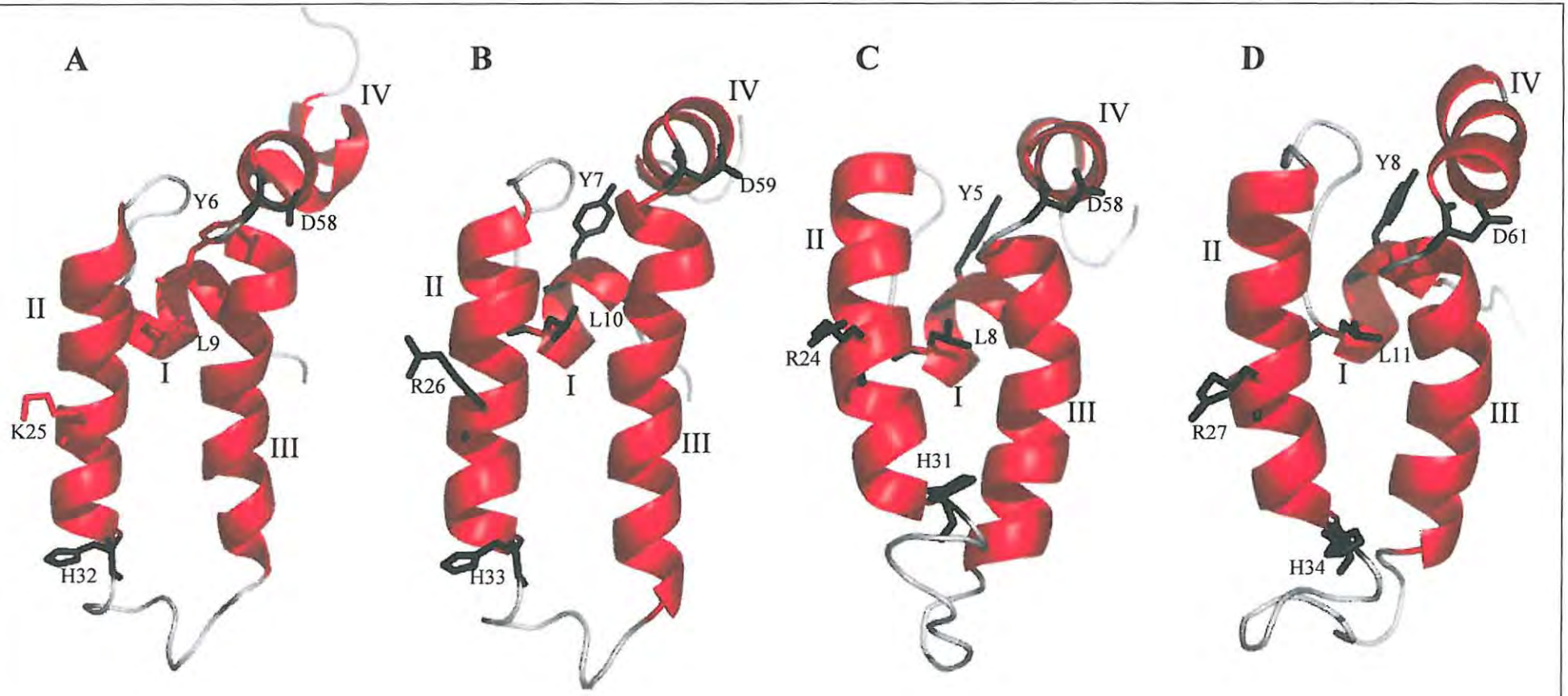


Figure 11: Ribbon representations of the crystal structure of the J-domain of DnaJ compared to the homology models of the J-domains of *Agt*DnaJ, Hsj1a and Pfj4

(A) The DnaJ J-domain crystal structure of *E. coli* DnaJ (DnaJ; PDB accession number: 1XBL). (B) The *Agt*DnaJ J-domain was homology modelled onto *H. sapiens* DnaJ (PDB accession number: 1HDJ), *M. musculus* DnaJ (PDB accession number: 1WJZ) and three sequences from *E. coli* DnaJ (PDB accession numbers: 1XBL, 1BQO, 1BQZ). The key residues identified as potentially important to be studied (Y7, L10, R26, H33 and D59) are shown in sticks and coloured in black. (C) The Hsj1 J-domain was homology modelled onto *H. sapiens* DnaJ 1 (PDB accession number: 1HDJ) and three sequences from *E. coli* DnaJ (PDB accession numbers: 1XBL, 1BQO, 1BQZ). The key residues identified as potentially important to be studied (Y5, L8, R24, H31 and D58) are shown in sticks and coloured in black. (D) The Pfj4 J-domain was homology modelled onto *H. sapiens* DnaJ (PDB accession number: 1HDJ) and three sequences from *E. coli* DnaJ (PDB accession numbers: 1XBL, 1BQO, 1BQZ). The key residues identified as potentially important to be studied (Y8, L11, R27, H34 and D61) are shown in sticks and coloured in black. The four α-helices (I – IV) are coloured red. The figures were modelled in *Swiss-Model* (Peitsch, 1995). The figures were visualised in *PyMOL* (DeLano, 2005).

3.2. HETEROLOGOUS EXPRESSION

3.2.1. The *Agt*DnaJ heterologous expression system

*Agt*DnaJ was previously heterologously produced for studies using an *E. coli* expression system (Boshoff *et al.*, 2004; Hennessy *et al.*, 2005a). *Agt*DnaJ was shown to be 58 % identical to DnaJ (Table 9) and capable of replacing *E. coli* DnaJ and CbpA in a mutant strain lacking those Hsp40s (Hennessy *et al.*, 2005a) and was used for J-domain swapping experiments in this study.

3.2.2. pHsj1*Agt* and pPfj4*Agt* parental plasmids

Domain-swapping has been used successfully in aiding the determination of the function and cross-functionality of J-domains across various families and species (reviewed in Hennessy *et al.*, 2005b). In this study, the less-characterised Hsj1a and Pfj4 J-domains were domain-swapped into the characterised *Agt*DnaJ in order to investigate whether these domains, and key residues therein, are important in J-domain functioning.

The Hsj1*Agt* chimera and the Pfj4*Agt* chimera were made (Appendix A2) by Prof G. L. Blatch and Dr W. S. Nicoll respectively (both of Department of Biochemistry, Microbiology and Biotechnology, Rhodes University) for work in previous studies (unpublished data).

The 4576 bp pHsj1*Agt* plasmid encoded the Hsj1 J-domain amino acids fused to the downstream regions of *Agt*DnaJ (Figure A2.1; Appendix 2). The identity of the pHsj1*Agt* plasmid (Figure 12A) was confirmed using analysis of DNA restricted with various type-II restriction enzymes (Figure 12B). Separate restriction experiments with *Bam* HI (Lane 2; Figure 12) and with *Hind* III (Lane 3; Figure 12) were seen to linearise the plasmid yielding DNA fragments of 4576 bp in length, whilst a digest with *Eco* RI yielded fragments of 3391 bp and 1185 bp (Lane 4; Figure 12). Plasmid identity was also confirmed by DNA sequencing (data not shown).

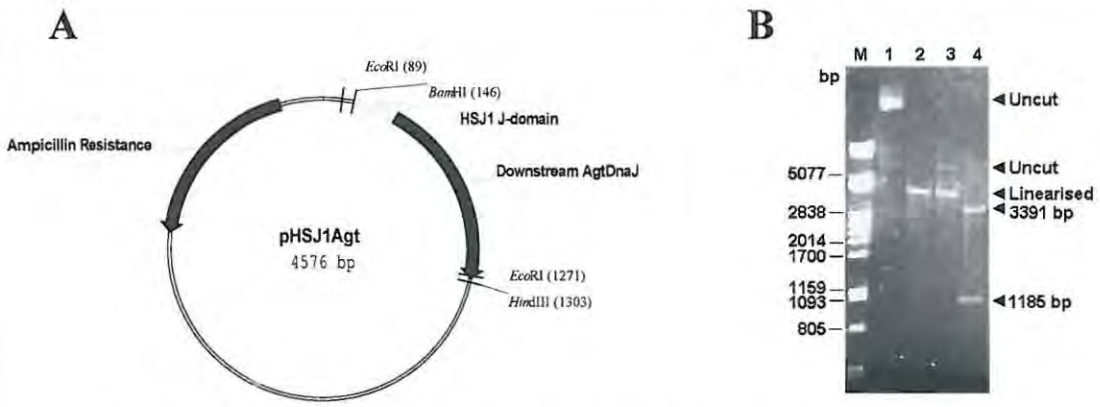


Figure 12: Confirmation of the identity of the pHsj1*Agt* plasmid

(A) Plasmid map of pHsj1*Agt* showing the Hsj1 J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt* was restricted with various restriction enzymes (Lanes 1 - 4) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Bam* HI (4576 bp); Lane 3 – *Hind* III (4576 bp); Lane 4 – *Eco* RI (3391 bp, 1185 bp).

The heterologous production of the Hsj1*Agt* protein in *E. coli* XL1-blue [pHsj1*Agt*] and *E. coli* OD259 [pHsj1*Agt*] was performed to begin the basic characterisation of this protein.

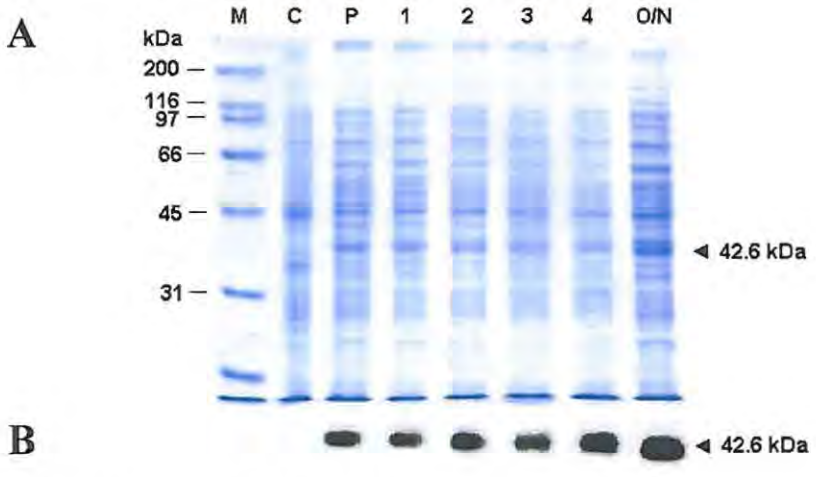


Figure 13: Heterologous production of Hsj1*Agt*

(A) Histidine-tagged Hsj1*Agt* was over-produced by *E. coli* XL1-blue [pHsj1*Agt*] using IPTG induction (lanes P – O/N). The total protein samples were resolved by SDS-PAGE. Lane M – Molecular mass marker (BioRad Broad Range SDS Marker); Lane C – *E. coli* XL1-blue [pQE30] control sample; Lane P – Pre-induction sample; Lane 1 – Post-induction 1 hour sample; Lane 2 – Post-induction 2 hour sample; Lane 3 – Post-induction 3 hour sample; Lane 4 – Post-induction 4 hour sample; Lane O/N – Post-induction overnight sample. (B) Histidine-tagged Hsj1*Agt* over-production was confirmed by Western analysis of the above samples. Western analysis of the 42.6 kDa Hsj1*Agt* protein is matched to the same region of the SDS-PAGE gel (black marker arrows).

After heterologous expression using *E. coli* XL1-blue [pHsj1*Agt*], the Hsj1*Agt* protein was not clearly identified within the total protein extract (Figure 13A). Faint bands of similar intensities were noted due to the leaky expression of the pQE30-based construct and no clearly over-produced band was seen on the SDS-PAGE gel (Lanes P – O/N; Figure 13A). Western analysis, however, did show a clear band present at the expected region of 42.6 kDa (Figure 13B). The control lane (Lane C; Figure 13B) showed no detection of a histidine-tagged protein in the 42.6 kDa region as expected, while in the same region for the test lanes the presence of the Hsj1*Agt* protein was detected (Lanes P – O/N; Figure 13B). The amount of protein produced increased with the length of induction. Significant increases in the accumulation of the protein started to occur from 4 hours onwards with the 1 – 3 hour samples (Lanes 1 – 3; Figure 13B) being approximately the same intensity as that of the pre-induction sample (Lane P; Figure 13B). The protein levels increased further with overnight induction. Because such a large protein band was detected in the overnight sample (Lane O/N; Figure 13B), large-scale degradation of the protein did not appear to have occured. If this were the case, the protein would not be detected at 42.6 kDa. Similar results were obtained with *E. coli* OD259 [pHsj1*Agt*] (data not shown).

Similar initial studies were performed on the plasmid, pPfj4*Agt* (Figure A2.7; Appendix 2). The heterologous production of the protein by *E. coli* XL1-blue [pPfj4*Agt*] was investigated.

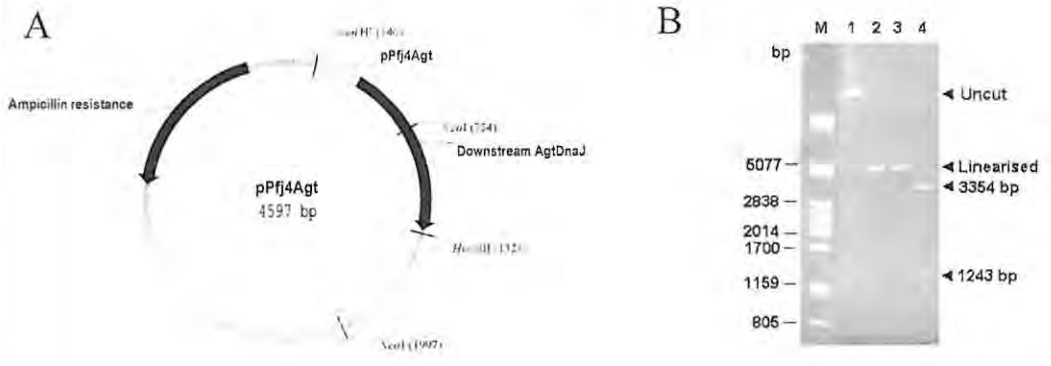


Figure 14: Confirmation of the identity of the pPfj4*Agt* plasmid

(A) Plasmid map of pPfj4*Agt* showing the Pfj4 J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pPfj4*Agt* was restricted with various restriction enzymes (Lanes 1 - 4) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Bam* HI (4597 bp); Lane 3 – *Hind* III (4597 bp); Lane 4 – *Nco* I (3354 and 1243 bp).

The pPfj4*Agt* plasmid (Figure 14A) encoded the Hsj1 J-domain amino acids fused to the downstream regions of *Agt*DnaJ. This was confirmed using restriction enzyme analysis (Figure 14B) and DNA sequencing (data not shown). Uncut DNA could be seen (Lane 1; Figure 14B) along with a *Bam* HI linearised (Lane 2; Figure 14B) or *Hind* III linearised (Lane 3; Figure 14B) fragment of 4597 bp. The pPfj4*Agt* DNA was restricted with the *Nco* I restriction enzyme and the two *Nco* I restriction sites at positions 754 and 1997 lead to two fragments. The two fragments from the restriction event with the *Nco* I restriction enzyme were 3354 bp and 1243 bp (Lane 4; Figure 14B).

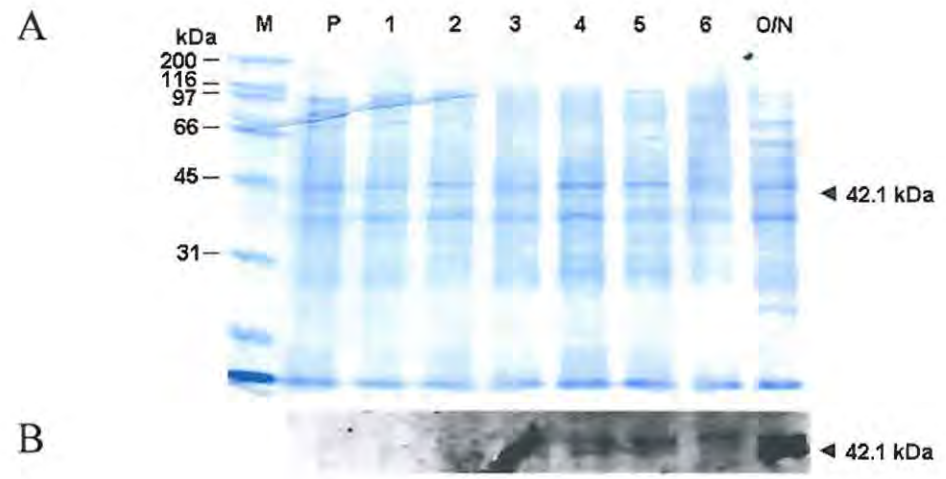


Figure 15: Heterologous production of Pfj4*Agt*

(A) Pfj4*Agt* was over-produced by *E. coli* XL1-blue [pPfj4*Agt*] using IPTG induction (lanes P – O/N). The total protein samples were resolved by SDS-PAGE. Lane M – Molecular mass marker (BioRad Broad Range SDS Marker); Lane P – Pre-induction sample; Lane 1 – Post-induction 1 hour sample; Lane 2 – Post-induction 2 hour sample; Lane 3 – Post-induction 3 hour sample; Lane 4 – Post-induction 4 hour sample; Lane 5 – Post-induction 5 hour sample; Lane 6 – Post-induction 6 hour sample; Lane O/N – Post-induction overnight sample. (B) Pfj4*Agt* over-production was confirmed by Western analysis of the above samples. Western analysis of the 42.1 kDa histidine-tagged Pfj4*Agt* protein is matched to the same region of the SDS-PAGE gel (black marker arrows).

After heterologous expression using *E. coli* XL1-blue [pPfj4*Agt*], the Pfj4*Agt* protein could not be clearly identified within the total protein extract as an over-expressed band (Figure 15A). Only faint bands of similar intensities were noted and no clearly over-produced band was seen.

Western analysis, however, did show a clear band pattern present at the expected region of 42.1 kDa (Figure 15B). The 42.1 kDa region showed the detection of histidine-tagged Pfj4*Agt* protein as present. Significant increases in the accumulation of the protein started to occur from 4 hours onwards with the 1 – 3 hour samples (Lanes 1 – 3; Figure 15B) being approximately the same intensity as that of the pre-induction sample (Lane P; Figure 15B). The protein levels increased further with overnight induction. Because such a large protein band was detected in the overnight sample (Lane O/N; Figure 15B), large-scale degradation of the protein did not appear to have occurred. If this were the case, the protein would not have been be detected at 42.1 kDa. Similar results were obtained with *E. coli* OD259 [pPfj4*Agt*] (data not shown).

The presence of a doublet was noted within the Western analysis samples (Figure 15B). This was probably due to cleavage of the Histidine-tag, which is approximately equivalent to 1 kDa of protein size. This remaining protein corresponded to the slightly smaller band pattern that was seen at approximately 41.1 kDa (Figure 15B), under the 42.1 kDa protein band.

3.2.3. Mutagenesis of pHsj1*Agt*

The coding region of the Hsj1*Agt* protein, produced successfully by *E. coli* XL1-blue [pHsj1*Agt*] and *E. coli* OD259 [pHsj1*Agt*], was used as the template plasmid for site-directed mutagenesis substitutions of those selected amino acid residues within the Hsj1 J-domain (Table 10).

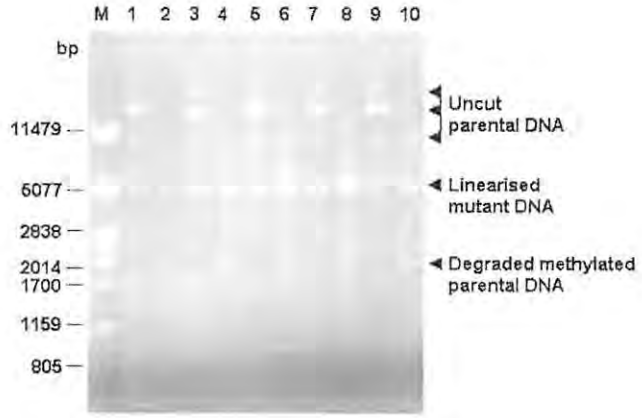


Figure 16: Analysis of mutagenesis amplification products of pHsj1*Agt*

Each of the five pHsj1*Agt* mutant plasmids were synthesized by whole plasmid amplification (Lanes 1, 3, 5, 7, 9). Two complimentary oligonucleotides containing the desired mutation and flanked by unmodified sequence were used for the synthesis of each of the five pHsj1*Agt* mutants. The DNA was then restricted with *Dpn* I cleaving the methylated recognition sites within the parental DNA template (Lanes 2, 4, 6, 8, 10). The Pre-*Dpn* I sample and Post-*Dpn* I sample for each mutant was resolved (odd and even numbered lanes respectively) by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lanes 1 – 2: pHsj1*Agt*(Y5A); Lanes 3 – 4: pHsj1*Agt*(L8A); Lanes 5 – 6; pHsj1*Agt*(R24A); Lanes 7 – 8: pHsj1*Agt*(H31Q); Lanes 9 – 10: pHsj1*Agt*(D58A).

Each of the pHsj1*Agt* mutagenesis amplification products resolved as linearised fragments at 4576 bp (Lanes 1 – 10; Figure 16). This was expected, as these fragments should be equivalent to the linearised pHsj1*Agt* plasmid. Various forms of uncut parental DNA (covalently-closed circular and nicked and open) were seen in all the pre-*Dpn* I samples (Odd lanes; Figure 16) and yet not within the post-*Dpn* I samples (Even lanes; Figure 16) as this methylated parental DNA was degraded by restriction with *Dpn* I. The degradation product fragments from the restriction event of the parental DNA with *Dpn* I could be seen in the post-*Dpn* I samples at approximately the 2000 bp region, due to the frequency of occurrence on the *Dpn* I restriction sites. The presence of the desired mutagenesis amplification products after successful degradation of parental DNA suggested that the desired mutagenesis was successful.

A small portion of each of the mutagenesis reactions was transformed into the relevant strains and the plasmid DNA was extracted from cells grown to mid-log phase for restriction analysis screening of the mutant plasmids.

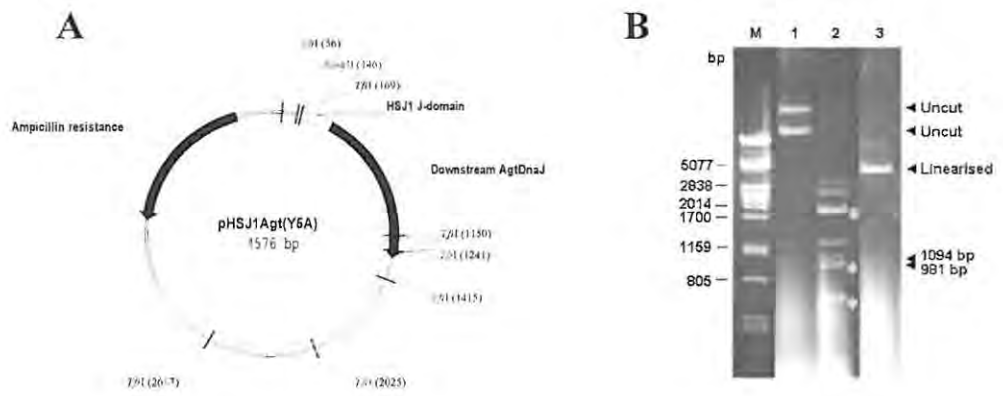


Figure 17: Confirmation of the identity of the pHsj1*Agt*(Y5A) mutated plasmid

(A) Plasmid map of pHsj1*Agt*(Y5A) showing the Hsj1(Y5A) J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt*(Y5A) was restricted with various restriction enzymes (Lanes 1 - 4) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Tfi* I (1965bp, 981 bp, 642 bp and 610 bp doublet, and 174 bp, 113 bp and 91bp fragments that are not visible); Lane 3 – *Bam* HI (4576 bp). The larger of the three DNA fragments that are fully restricted are indicated with stars, whereas other bands indicate partial digests.

The pHsj1*Agt*(Y5A) plasmid (Figure 17A), encoding the Y5A-mutated Hsj1a J-domain with downstream *Agt*DnaJ, was confirmed using restriction enzyme analysis (Figure 17B) and DNA sequencing (Figure A2.1; Appendix 2). Two forms of uncut DNA could be seen (Lane 1; Figure 17B) along with a linearised fragment of 4576 bp (Lane 3; Figure 17B). Diagnostic *Tfi* I restriction enzyme digestion (Lane 2; Figure 17) indicated that partial restriction had occurred. The pHSJ1*Agt* parental DNA had a profile of fragments of the following predicted sizes when restricted with *Tfi* I: 1965 bp, 1094 bp, 642 bp, 610 bp, 174 bp and 91 bp. In the case of the Y5A mutant plasmid, the 1094 bp fragment of the parental pHsj1*Agt* was restricted further by the diagnostic *Tfi* I restriction enzyme and a 981 bp diagnostic fragment was released (lane 2). This 981 bp fragment was formed by the introduction of an additional *Tfi* I restriction site as engineered into the mutagenesis primers. The 1094 bp fragment was still visible as the restriction was partial.

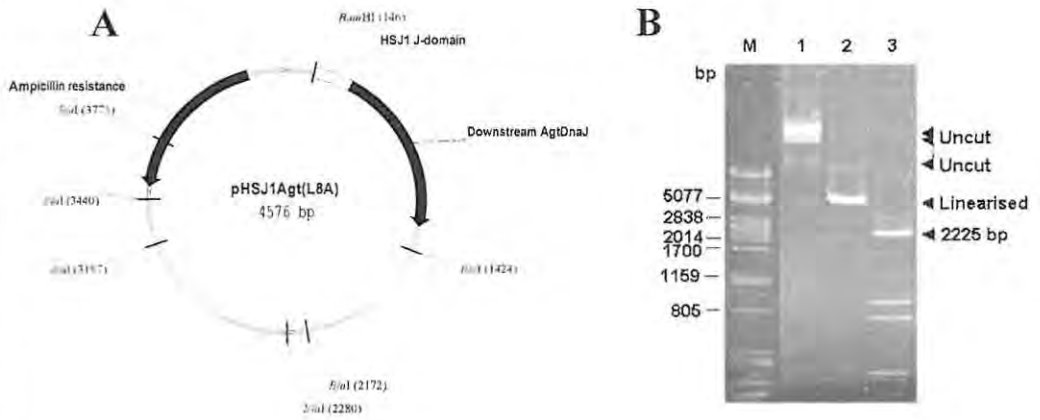


Figure 18: Confirmation of the identity of the pHsj1*Agt*(L8A) mutated plasmid

(A) Plasmid map of pHsj1*Agt*(L8A) showing the Hsj1(L8A) J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt*(L8A) was restricted with various restriction enzymes (Lanes 1 - 3) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Bam* HI (4576 bp); Lane 3 – *Bfa* I (2225 bp, 907 bp, 748 bp, 335 bp, and the 253 bp and 108 bp fragments that are not visible).

The pHsj1*Agt*(L8A) plasmid (Figure 18A), encoding the L8A-mutated Hsj1a J-domain with downstream *Agt*DnaJ, was confirmed using restriction enzyme analysis (Figure 18B) and DNA sequencing (Figure A2.2; Appendix 2). Three forms of uncut DNA could be seen (Lane 1; Figure 18B) along with a *Bam* HI-digested linearised fragment of 4576 bp (Lane 2; Figure 18B). The diagnostic *Bfa* I restriction enzyme digestion (Lane 3; Figure 18B) gave the expected DNA fragments. The pHsj1*Agt* parental DNA had a profile of fragments of the following predicted sizes when restricted with *Bfa* I: 1251 bp, 974 bp, 907 bp, 748 bp, 335 bp, 253 bp and 108 bp. In the case of the L8A mutant plasmid, the 1251 bp and 974 bp fragments of the parental pHsj1*Agt* were seen as one diagnostic fragment of 2225 bp due to the removal of a *Bfa* I restriction site. This was due to the engineering of the mutagenesis primers (Table A4.3; Appendix 4) leading to a larger fragment from the addition of the 1251 and 974 bp fragments.

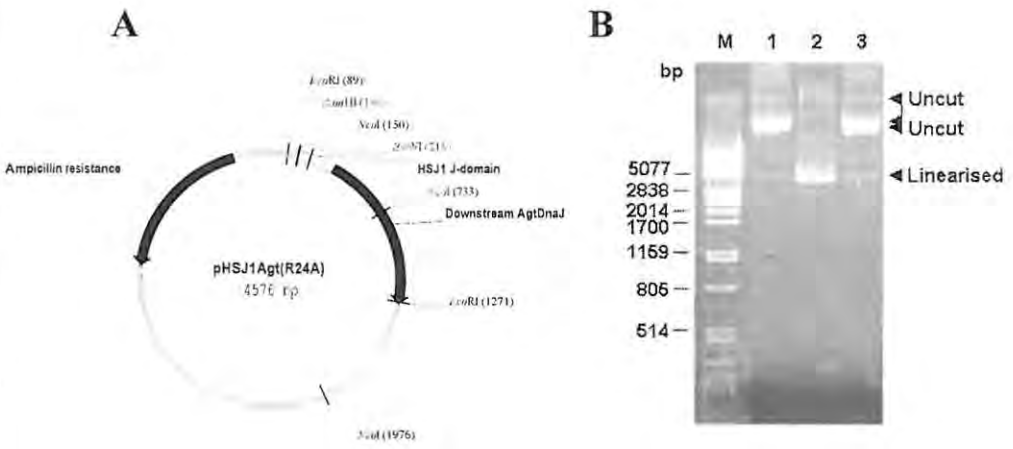


Figure 19: Confirmation of the identity of the pHsj1*Agt*(R24A) mutated plasmid

(A) Plasmid map of pHsj1*Agt*(R24A) showing the Hsj1(R24A) J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt*(R24A) was restricted with various restriction enzymes (Lanes 1 - 3) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Bam* HI (4576 bp); Lane 3 – *Bsi* WI (4576 bp).

The pHsj1*Agt*(R24A) plasmid (Figure 19A), encoding the R24A-mutated Hsj1a J-domain with downstream *Agt*DnaJ, was confirmed using restriction enzyme analysis (Figure 19B) and DNA sequencing (Figure A2.3; Appendix 2). Three forms of uncut DNA could be seen (Lanes 1 and 3; Figure 19B) and a fragment linearised by *Bam* HI of 4576 bp (Lane 2; Figure 19B). The diagnostic *Bsi* WI restriction enzyme digestion (Lane 3; Figure 19B) gave the expected linearised fragment. The uncut DNA could still be seen as the restriction event was only partial thus the linearised fragment was faint in comparison to the uncut fragments. The pHsj1*Agt* parental DNA did not contain a *Bsi* WI restriction site, and one *Bsi* WI restriction site was added by being engineered into the mutagenesis primers. The pHsj1*Agt*(R24A) mutant plasmid thus only showed one linearised fragment (Lane 3; Figure 19B).

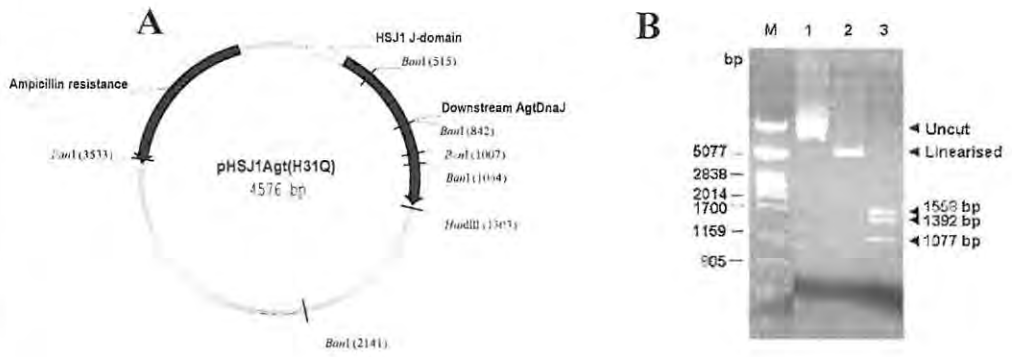


Figure 20: Confirmation of the identity of the pHsj1*Agt*(H31Q) mutated plasmid

(A) Plasmid map of pHsj1*Agt*(H31Q) showing the HSJ1(H31Q) J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt*(H31Q) was restricted with various restriction enzymes (Lanes 1 - 3) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Hind* III (4576 bp); Lane 3 – *Ban* I (1558 bp, 1392 bp, 1077 bp, and smaller fragments that are not visible: 576 bp, 327 bp, 165 bp).

The pHsj1*Agt*(H31Q) plasmid (Figure 20A), encoding the H31Q-mutated Hsj1a J-domain with downstream *Agt*DnaJ, was confirmed using restriction enzyme analysis (Figure 20B) and DNA sequencing (Figure A2.4; Appendix 2). Uncut DNA could be seen (Lane 1; Figure 20B) along with a linearised fragment of 4576 bp (Lane 2; Figure 20B). The diagnostic *Ban* I restriction enzyme digestion (Lane 3; Figure 20B) gave the expected DNA fragment profiles. The pHsj1*Agt* parental DNA had a profile of fragments of the following predicted sizes when restricted with *Ban* I: 1392 bp, 1283 bp, 1077 bp, with the smaller fragments not visible: 327 bp, 275 bp, 165bp, 57 bp. In the case of the pHsj1*Agt*(H31Q) mutant plasmid, the removal of a *Ban* I site lead to the formation of a 1558 bp diagnostic fragment by the addition of the 1283 bp and 275 bp parental fragments (Lane 3; Figure 20B).

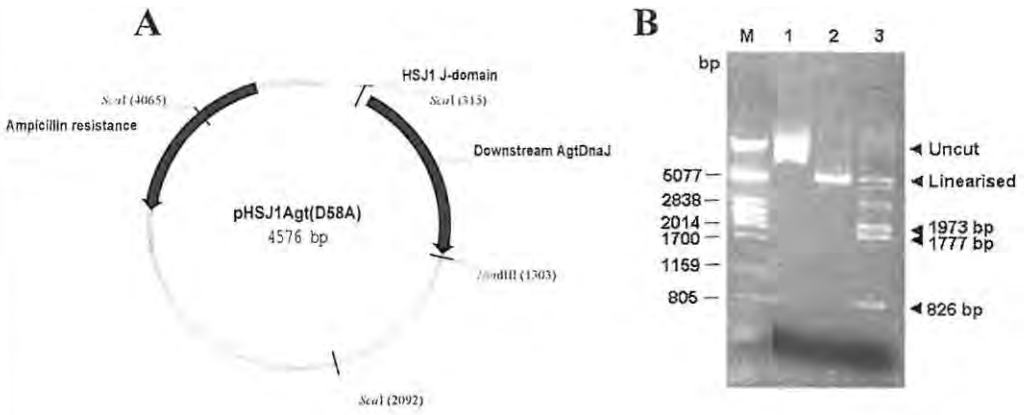


Figure 21: Confirmation of the identity of the pHsj1*Agt*(D58A) mutated plasmid

(A) Plasmid map of pHsj1*Agt*(D58A) showing the Hsj1(D58A) J-domain (grey block) and the downstream *Agt*DnaJ coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pHsj1*Agt*(D58A) was restricted with various restriction enzymes (Lanes 1 - 3) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Hind* III (4576 bp); Lane 3 – *Sca* I (1973 bp, 1777 bp, 826 bp).

The pHsj1*Agt*(D58A) plasmid (Figure 21A), encoding the D58A-mutated Hsj1a J-domain with downstream *Agt*DnaJ, was confirmed using restriction enzyme analysis (Figure 21B) and DNA sequencing (Figure A2.5; Appendix 2). Three forms of uncut DNA could be seen (Lane 1; Figure 21B) along with a linearised fragment of 4576 bp (Lane 2; Figure 21B). The diagnostic *Sca* I restriction enzyme digestion (Lane 3; Figure 21B) gave the expected DNA fragment profile. The pHsj1*Agt* parental plasmid had two Sca I restriction sites, giving two fragments of 2603 bp and 1973 bp when restricted with *Sca* I. In the case of the pHsj1*Agt*(D58A) mutant plasmid, the introduction of a *Sca* I site lead to the formation of two fragments 1777 bp and 826 bp from the further restriction of the 2603 bp fragment. Those bands found between the linearised and the 1973 bp band are due to partial restriction.

The equivalent substitutions of the pPfj4*Agt* plasmid were engineered by Dr William S. Nicoll (Dept. of Biochemistry, Microbiology and Biotechnology, Rhodes University) for a previous study (unpublished data). These mutations were confirmed by restriction enzyme digestion and DNA sequencing (data not shown).

3.2.4. *In vivo* complementation assays

E. coli OD259 is a thermosensitive strain lacking DnaJ and CbpA, which are Hsp40 proteins needed to aid in survival of the strain under heat shock temperatures (i.e. above 37 °C and below 16 °C). Transforming this strain with plasmids that produce substitutes for these proteins will result in the reversal of strain thermosensitivity and growth will be seen at the non-permissive temperatures.

Heterologously-expressed *Agt*DnaJ (*E. coli* OD259 [p*Agt*DnaJ]) is able to reverse the thermosensitivity of the thermosensitive strain, *E. coli* OD259, whilst *Agt*DnaJ(H33Q) is unable to perform this function (Hennessy *et al.*, 2005a). These properties allow for the investigation into the *in vivo* function of protein chimeras in this system, with the intention of attempting to determine the function of heterologously-expressed chimeric domains.

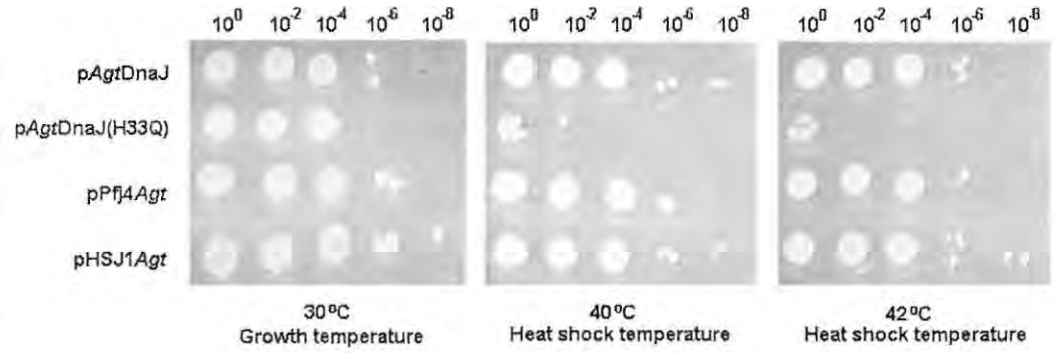


Figure 22: Hsj1*Agt* and Pfj4*Agt* were able to reverse the thermosensitivity of *E. coli* OD259

E. coli OD259 cells were transformed with the relevant plasmids (left). Cells were grown by spotting a dilution series of cells ($10^0 - 10^{-8}$) onto agar plates containing IPTG for induction to test reversal of thermosensitivity at the various temperatures: 30, 40 and 42 °C. *Agt*DnaJ and *Agt*DnaJ(H33Q) were positive and negative controls respectively. Pfj4*Agt* and Hsj1*Agt* were test experiments.

In interpreting the results for these *in vivo* complementation assays, the importance of growth at the level of individual colonies needs emphasis. Growth of a large numbers of colonies clumped together could be due to a mass colony shielding effect, however, it is growth of colonies at the single colony level that are important as no shielding can occur by other bacterial cells and hence growth is due to viable cells.

The production of *Agt*DnaJ from the pRJ30 plasmid was able to reverse the thermosensitivity of the *E. coli* OD259 strain (Top row; Figure 22). At the partial heat shock temperature of 40°C (10^{-8}) and at heat shock temperature of 42 °C (10^{-6}), *Agt*DnaJ was able to assist the transformed strain as growth was seen. Conversely, the production of *Agt*DnaJ(H33Q) from the pRJ30(H33Q) plasmid was not able to reverse the thermosensitivity of the transformed strain as growth was only detected at the 10^{-2} dilution at 40 °C and in the undiluted sample at 42 °C (Second row; Figure 22). Both the Hsj1*Agt* and Pfj4*Agt*, producing *E. coli* OD259 transformants were able to grow as well at 30 °C as they were able to grow under heat shock conditions at both 40 °C and 42 °C. This shows that the J-domains of Hsj1 and Pfj4 were able to successfully substitute for the *Agt*DnaJ J-domain in *Agt*DnaJ because growth could be seen in dilutions as high as 10^{-8} in cells producing Hsj1*Agt* and as high as 10^{-6} for cells producing Pfj4*Agt* at the 42 °C heat shock temperature.

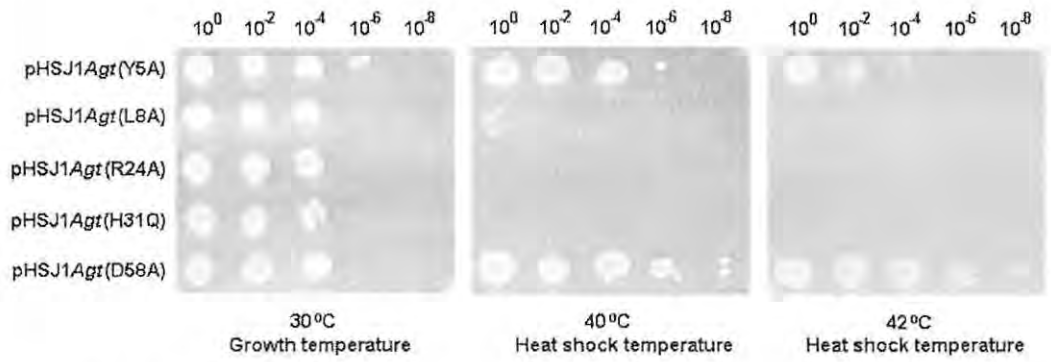


Figure 23: Hsj1*Agt*(Y5A) and Hsj1*Agt*(D58A) were able to reverse the thermosensitivity of *E. coli* OD259

E. coli OD259 cells were transformed with the relevant plasmids (left). Cells were grown by spotting a dilution series of cells ($10^{0} - 10^{-8}$) onto agar plates containing IPTG for induction to test reversal of thermosensitivity at the various temperatures: 30, 40 and 42 °C. *Agt*DnaJ and *Agt*DnaJ(H33Q) were repeated as positive and negative controls respectively (data not shown). Hsj1*Agt* mutants were tested for their ability to reverse the thermosensitivity of *E. coli* OD259.

Since previous results (Figure 22) showed that the Hsj1*Agt* protein was able to reverse the thermosensitivity of *E. coli* OD259, similar studies were performed on the mutant versions of this protein in order to investigate whether highly conserved amino acid residues were important to the *in vivo* function of this Hsp40. The production of each of the Hsj1*Agt* mutant proteins in *E. coli* OD259 did not all yield the same results with regards to thermosensitivity reversal (Figure 23).

Whilst all the *E. coli* OD259 transformants producing mutant proteins were able to grow at the non-permissive 30 °C temperature up to at least a 10^{-4} dilution, the Hsj1*Agt*(Y5A) mutant protein (Top row; Figure 23) was able to partially reverse the thermosensitivity of *E. coli* OD259 and yield growth in the 10^{-4} dilution range at 42 °C. The Hsj1*Agt*(D58A) mutant protein (Bottom row; Figure 23) was able to fully reverse the thermosensitivity of *E. coli* OD259 and yield growth up to the 10^{-8} dilution range at 42 °C. The Hsj1*Agt*(L8A), Hsj1*Agt*(R24A) and Hsj1*Agt*(H31Q) mutant proteins were all unable to reverse the thermosensitivity of *E. coli* OD259 showing no growth at 42 °C.

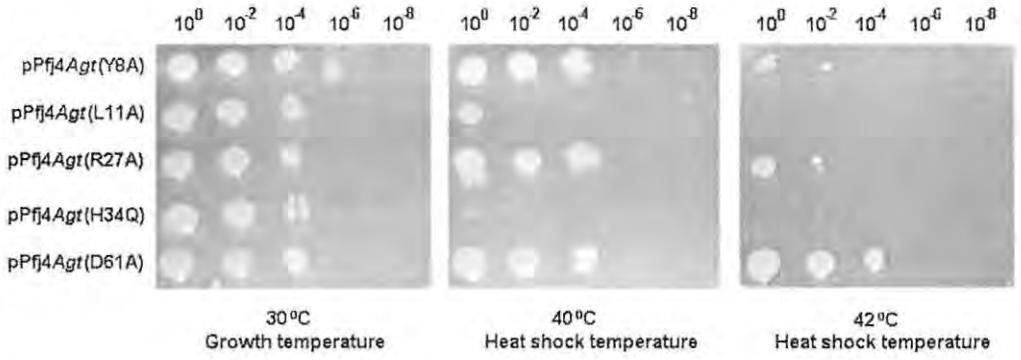


Figure 24: Pfj4*Agt*(D61A) was able to reverse the thermosensitivity of *E. coli* OD259, whilst Pfj4*Agt*(Y8A) and Pfj4*Agt*(R27A) can only partially reverse thermosensitivity

E. coli OD259 cells were transformed with the relevant plasmids (left). Cells were grown by spotting a dilution series of cells ($10^{0} - 10^{-8}$) onto agar plates containing IPTG for induction to test reversal of thermosensitivity at the various temperatures: 30, 40 and 42 °C. *Agt*DnaJ and *Agt*DnaJ(H33Q) were repeated as positive and negative controls respectively (data not shown). Pfj4*Agt* mutants were tested for their ability to reverse the thermosensitivity of *E. coli* OD259.

Similarly, the Pfj4*Agt* proteins with mutations equivalent to the Hsj1 J-domain mutations were investigated for their ability to reverse the thermosensitivity of *E. coli* OD259 and to compare this ability with those results obtained for the equivalent mutations in Hsj1*Agt*. The production of the Pfj4*Agt* mutant proteins (Figure 24) do not all yield the same results with regards to reversal of thermosensitivity of *E. coli* OD259 when compared to the Hsj1*Agt* mutants (Figure 23).

Whilst all the *E. coli* OD259 transformants producing mutant proteins were able to grow at the non-permissive 30 °C temperature up to at least a 10^{-4} dilution in much the same way as the cells producing the Hsj*Agt* mutant proteins, the Pfj4*Agt*(Y8A) (Top row; Figure 24) and the Pfj4*Agt*(R27A) (Third row; Figure 24) mutants were only partially able to reverse the thermosensitivity of the thermosensitive strain and yield growth in the 10^{-4} dilution range at 40 °C and in the 10^{-2} dilution range at 42 °C.

Both the Pfj4*Agt*(L11A) (Second row; Figure 24) and Pfj4*Agt*(H34Q) (Fourth row; Figure 24) mutant proteins were unable to reverse the thermosensitivity of *E. coli* OD259. No growth was detected for these mutants at the heat shock temperature of 42 °C. The Pfj4*Agt*(D61A) mutant protein was able to effectively reverse the thermosensitivity of *E. coli* OD259. The non-permissive 30 °C temperature showed growth up to the 10^{-4} dilution range. The results were duplicated for both the heat shock temperatures showing that this mutant protein is able to function in reversing thermosensitivity of *E. coli* OD259 at heat shock temperatures.

In order to show that those results where no growth was noted, was due to the Hsj1*Agt* protein not being able to substitute for DnaJ or CbpA in *E. coli* OD259, and not due to a lack of protein production, the heterologous protein production for each of the mutated proteins was investigated (Figure 25). Hsj1*Agt* and its mutant derivatives were all detected by Western analysis of *E. coli* OD259 transformant cell extracts (Figure 25B). A similar result was also found for Pfj4*Agt* and its mutant derivatives (Dr W. S. Nicoll, unpublished data).

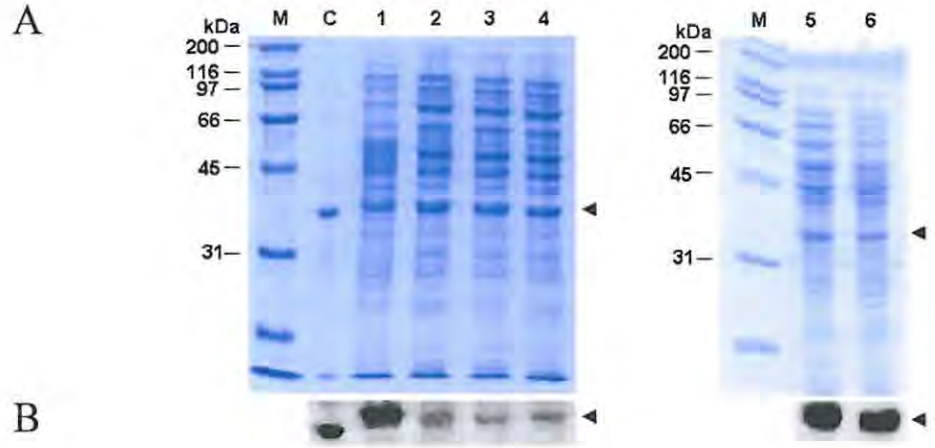


Figure 25: Analysis of the production of Hsj1*Agt* and the Hsj1*Agt* mutant proteins

(A) *E. coli* OD259 cells were transformed and the pHsj1*Agt* parental and mutant plasmids were induced for production with IPTG (lanes 1 - 6). The total protein samples were resolved by SDS-PAGE. Lane M – Molecular mass marker (BioRad Broad Range SDS Marker); Lane C – Purified histidine-tagged Hsj1a (positive control for Western analysis); Lane 1 – [pHsj1*Agt*] control sample; Lane 2 – [pHsj1*Agt*(Y5A)] sample; Lane 3 – [pHsj1*Agt*(L8A)] sample; Lane 4 – [pHsj1*Agt*(R24A)] sample; Lane 5 – [pHsj1*Agt*(H31Q)] sample; Lane 6 – [pHsj1*Agt*(D58A)] sample. (B) Hsj1*Agt* chimera and mutant over-production was confirmed by Western analysis of the above samples. Western analysis of the 42.6 kDa Hsj1*Agt* protein is matched to the same region of the SDS-PAGE gel (black marker arrows).

The histidine-tagged Hsj1a was used as positive control for Western analysis (Lane C; Figure 25A). The 42.6 kDa Hsj1*Agt* protein and its mutant versions could be seen in neighbouring lanes (Lanes 1 – 6; Figure 25A) within the total protein extract as resolved slightly above the 36 kDa Hsj1a. The Hsj1*Agt*, Hsj1*Agt*(H31Q) and the Hsj1*Agt*(D58A) samples showed intense bands with Western analysis showing protein production. The Hsj1*Agt*(Y5A), Hsj1*Agt*(L8A) and Hsj1*Agt*(R24A) total protein samples (Lanes 2 – 4; Figure 25B) showed light bands with Western analysis suggesting lower amounts of these proteins being produced in the strain relative to the Hsj1*Agt* and the Hsj1*Agt*(H31Q) and Hsj1*Agt*(D58A) mutant proteins. The variable protein levels across the lanes could be due to the production of each of the various mutant proteins affecting the growth rate of the cell cultures differently, with this then being reflected in the single time-point sampling method.

3.2.5. Hsj1a characterisation

3.2.5.1. Hsj1a purification

The region encoding histidine-tagged Hsj1a was excised from another plasmid construct and ligated into the pQE30 vector system as described previously (Section 2.2.10; Chapter 2). The reason for this was two-fold, because: (i) the currently available Hsj1a expression system constructs for *E. coli* (pET-based constructs) would be limited to certain *E. coli* strains for heterologous expression (*E. coli* BL21[DE3]), and (ii) the other plasmids used in this study were pQE30-based and for effective comparison of protein production similar vector systems were preferable.

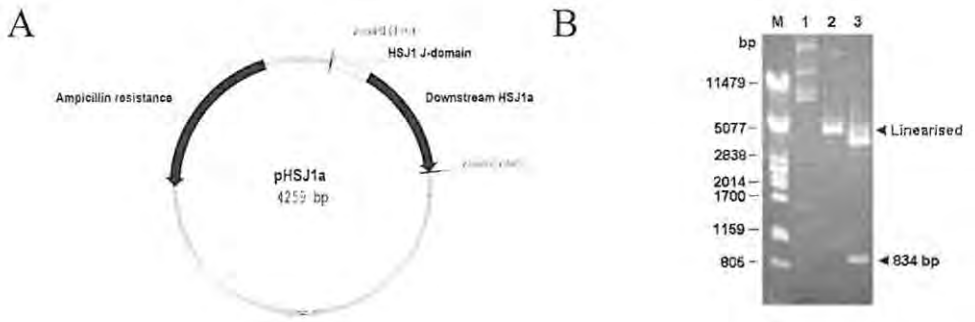


Figure 26: Confirmation of the identity of the pCMcHsj1a plasmid

(A) Plasmid map of pCMcHsj1a showing the Hsj1 J-domain (grey block) and the remainder of the Hsj1a coding regions (downstream black arrow). The β-lactamase coding sequence for ampicillin resistance is also highlighted (black arrow). The figure was generated using Vector NTI (version 9.0.0.) (InforMax, USA). (B) pCMcHsj1a was restricted with various restriction enzymes (Lanes 1 - 3) and the DNA fragments were resolved by agarose gel electrophoresis. Lane M – Molecular mass marker (Lambda DNA restricted with *Pst* I); Lane 1 - Uncut; Lane 2 – *Bam* HI (4259 bp); Lane 3 – *Bam* HI and *Hind* III (3425 and 834 bp).

The sequence for the pCMcHsj1a plasmid, encoding histidine-tagged full-length Hsj1a (Figure 26A), was confirmed using restriction enzyme analysis (Figure 26B) and DNA sequencing (Figure A2.8b; Appendix 2). Uncut DNA (Lane 1; Figure 26B) and the expected *Bam* HI linearised fragment of 4259 bp (Lane 2; Figure 26B) were detected. The pCMcHsj1a DNA was restricted with the *Bam* HI and *Hind* III restriction enzymes and restriction sites at

positions 146 and 986 lead to the two expected fragments of 3425 bp and 834 bp (Lane 3; Figure 26B).

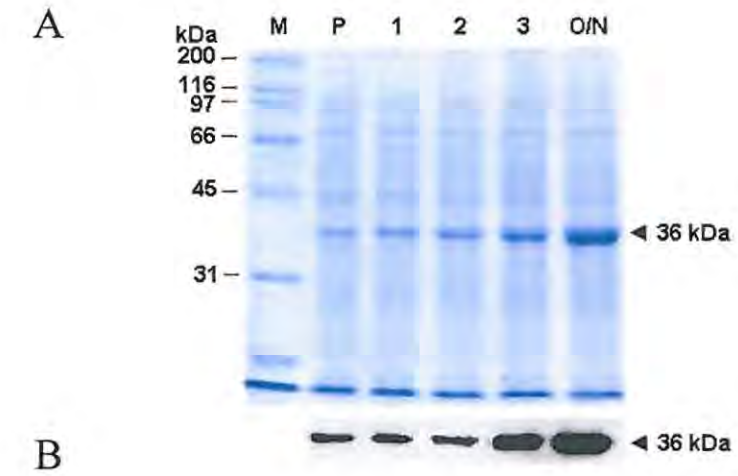


Figure 27: Histidine-tagged Hsj1a was successfully heterologously over-produced in *E. coli*

(A) Histidine-tagged Hsj1 was over-produced in *E. coli* XL1-blue [pCMcHsj1a] using IPTG induction (lanes P – O/N). The total protein samples were resolved by SDS-PAGE. Lane M – Molecular mass marker (BioRad Broad Range SDS Marker); Lane P – Pre-induction sample; Lane 1 – Post-induction 1 hour sample; Lane 2 – Post-induction 2 hour sample; Lane 3 – Post-induction 3 hour sample; Lane O/N – Post-induction overnight sample. (B) Histidine-tagged Hsj1 over-production was confirmed by Western analysis of the above samples. Western analysis of the 36 kDa Hsj1 protein is matched to the same region of the SDS-PAGE gel (black marker arrows).

The successful over-production of human histidine-tagged Hsj1a was achieved in *E. coli* XL1-blue (Figure 27). Both the SDS-PAGE analysis (Figure 27A) and Western analysis (Figure 27B) showed clear protein bands representing the over-production of the protein. The protein was over-produced after 3 hours. Initially over the first two hours of IPTG induction there was no difference in the amount of protein produced.

The histidine-tagged protein was then purified using a native batch purification method (Figure 28). The sonicated sample was centrifuged and the soluble fraction was kept as the cleared lysate (Lane CL; Figure 28A). This showed that there was a significant amount of Hsj1a protein present in the supernatant portion of the sample in the soluble form suggesting that it could potentially be purified successfully using native methods incorporating non-denaturing washes and elutions.

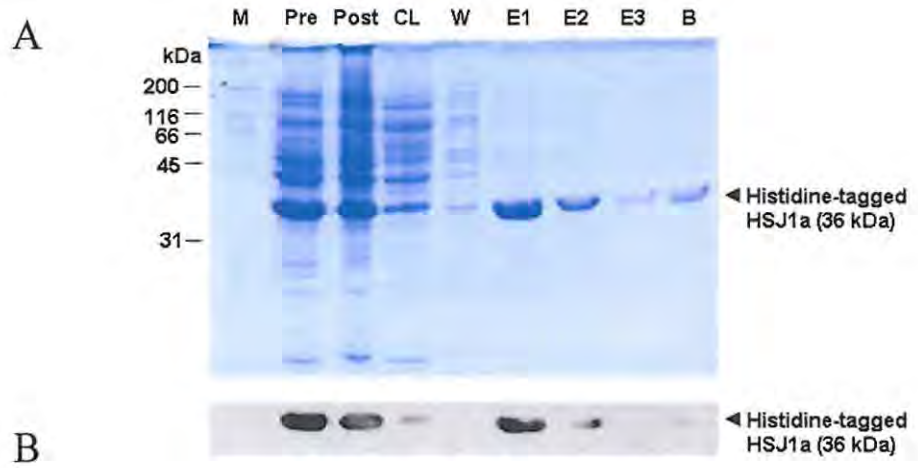


Figure 28: Histidine-tagged Hsj1a was successfully purified by a native batch method

Histidine-tagged Hsj1 was over-produced in *E. coli* XL1-blue [pCMcHsj1a] using IPTG induction and samples from various stages of the purification protocol were resolved by SDS-PAGE. Lane Pre – Total protein sample of *E. coli* XL1-blue [pCMcHsj1a] pre-sonication; Lane Post – Total protein sample of *E. coli* XL1-blue [pCMcHsj1a] post-sonication; Lane CL – Cleared lysate sample (soluble fraction of post sonication sample); Lane W –Wash sample (cleared lysate sample post non-denaturing wash); Lane E1-E3 – Elution samples (wash sample post non-denaturing elutions); Lane B – Bead sample.

The non-denaturing wash sample (Lane W; Figure 28A) revealed that the removal of most non-specific binders to the nickel beads had occurred and this could be clearly seen when comparing the many protein bands in the wash sample with those from the non-denaturing elution samples (Lane E1 – 3; Figure 28A) containing only the specifically bound Hsj1a protein. Each of the successive elutions contained relatively less protein and it was also seen that a large portion of the Hsj1a protein still remained bound to the charged nickel beads with analysis of the bead sample (Lane B; Figure 28A). Approximately 7.5 mg of protein was purified per litre of broth culture using this method from the cleared lysate sample. Purification was almost to homogeneity, however, there were very faint protein bands present in the regions of 50, 70 and 200 kDa. The protein bands present at approximately 36 kDa were confirmed as histidine-tagged Hsj1a protein by Western analysis clearly showing the presence of a histidine-tagged protein (Figure 28B). These results were seen to correspond to those obtained previously by Cheetham and co-workers (Cheetham *et al.*, 1992; Cheetham *et al.*, 1994).

3.2.5.2. Hsj1a oligomeric analysis

A preliminary investigation into the oligomeric state of Hsj1a was conducted using the histidine-tagged Hsj1a purified from the *E. coli* heterologous expression system (Lane E1; Figure 28). This was performed using the comparison of proteins of known size by size exclusion FPLC. The Hsj1a sample was eluted under the same conditions (Section 2.2.16; Chapter 2) and the fractions obtained were analysed by Western analysis to determine which fractions obtained the histidine-tagged Hsj1a protein.

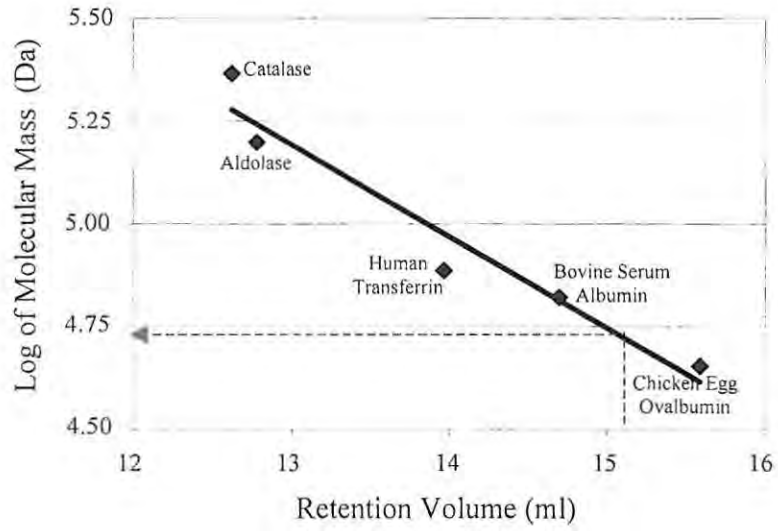


Figure 29: Standard curve of the fplc size exclusion chromatography of histidine tagged Hsj1a

A superdex 200 FPLC column was used. The standard curve was comprised of catalase (232 kDa), aldolase (158 kDa), human transferrin (77 kDa), bovine serum albumin (66.2 kDa) and chicken egg ovalbumin (45 kDa). The standard curve data points represent averages of replicates of the various standards. The histidine-tagged Hsj1a molecular mass (46 kDa) was calculated based on the retention volume obtained (15.18 ml), as shown by the dashed arrow.

The histidine-tagged Hsj1a protein was confirmed as being present by Western analysis and was found to have an elution volume of 15.18 ml (Figure 29) eluting before the 45 kDa chicken egg ovalbumin, and hence suggesting the histidine-tagged Hsj1a has a larger molecular mass.

From comparison to those proteins of known size shown on the standard curve, Hsj1a was found to have a size of approximately 46 kDa (red dashed line; Figure 29). This was unexpected since oligomers of multiples of 36 kDa, depending on whether the protein was found to be monomeric (36 kDa) or dimeric (72 kDa) etc, were not found. The 46 kDa size suggests neither a monomer or a dimer. This preliminary result could suggest that histidine-tagged Hsj1a exists in solution as either (i) a monomer having a large extended conformation, (ii) a dimer having a tightly compacted conformation, or (iii) a higher oligomer which is degraded before or during the FPLC analysis and only a portion of the protein remains attached to the histidine-tag which is then detected as a 46 kDa fragment.

3.2.6. *In vitro* ATPase assays

In vitro studies of the ATPase activity of the co-chaperones, histidine-tagged Hsj1a and histidine-tagged Pfj4, were assessed. This was performed using a modified ascorbate/molybdate colourimetric ATPase assay (Section 2.2.17; Chapter 2; Chifflet *et al.*, 1988; Edkins *et al.*, 2004).

The steady state basal ATPase activity of human Hsp70 was monitored *in vitro* by assessing ATP hydrolysis as measured by release of inorganic phosphate over time. The initial velocities of the reactions were calculated over the linear regions of the reaction curves and the specific activities (nmol Pi/min/mg human Hsp70) were calculated (Figure 30).

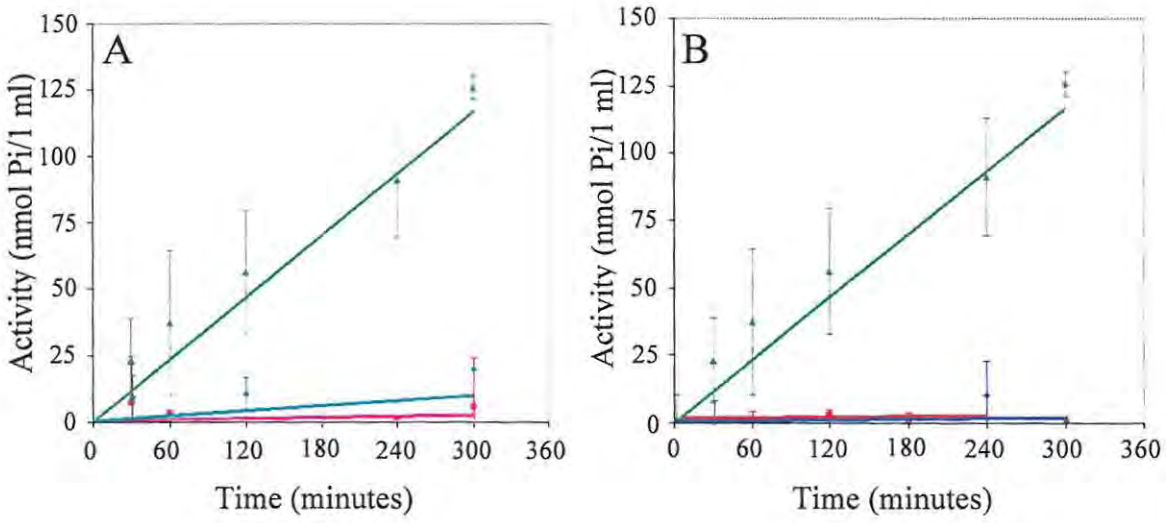


Figure 30: Histidine-tagged Hsj1a was able to stimulate the ATPase activity of human Hsp70, whilst histidine-tagged Pfj4 could not

Activities were given for the ATPase assay reactions. Curves are defined as follows by reference to those constituents included in the ATPase assay reaction (from 0.4 μM Hsp40, 0.4 μM human Hsp70 and/or 600 μM ATP, excluding the ATPase buffer constituents). (A) The basal ATPase activity (Hsp70 and ATP) is shown in turquoise, and is deemed non-detectable due to the low activity values obtained. The Hsj1a and ATP reaction is shown in pink. The Hsj1a, ATP and Hsp70 reaction is shown in green. (B) The Pfj4 and ATP reaction is shown in dark blue. The Pfj4, ATP and Hsp70 reaction is shown in red. Similarly, the Hsj1a, ATP and Hsp70 reaction is again shown in green for comparative purposes.

There was no inorganic phosphate released from the control reaction samples containing human Hsp70 alone or ATP alone, indicating no contamination of P_i in either of these constituents used in the experiment (data not shown). The lack of inorganic phosphate in the ATP alone control also indicated no spontaneous breakdown of the ATP. The specific activity findings indicated that human Hsp70 in the presence on ATP had such a low basal specific ATPase activity (Turquoise; Figure 30A), that it was deemed non-determinable (ND).

The histidine-tagged Hsj1a alone reaction (data not shown) and the histidine-tagged Hsj1a and ATP reaction (Pink; Figure 30A) both showed no significant ATPase activity as expected since a type-II co-chaperone is known to have no autonomous ATPase activity and no Hsp70 protein was present to cause ATP hydrolysis. Histidine-tagged Hsj1a in the presence of ATP and human Hsp70 was able to stimulate the ATPase activity of human Hsp70 and a specific activity of approximately 14 nmol/P_i/min/mg was detected (Green; Figure 30A).

Freshly purified, soluble Pfj4 was a generous gift of Mr Addmore Shonhai, Dept. of Biochemistry, Microbiology and Biotechnology, Rhodes University. The histidine-tagged Pfj4 alone reaction (data not shown) and the histidine-tagged Pfj4 reaction in the presence of ATP (Dark blue; Figure 30B) both showed no ATPase activity as expected since a type-II co-chaperone is known to have no autonomous ATPase chaperone activity and no Hsp70 protein was present to cause ATP hydrolysis. Interestingly, histidine-tagged Pfj4 with ATP and human Hsp70 was unable to stimulate the ATPase activity of human Hsp70 (Red; Figure 30B). This means that whilst human Hsj1a and human Hsp70 are able to function as an effective chaperone – co-chaperone pair, Pfj4 cannot function with human Hsp70 in this way under the conditions used in this experiment.

Chapter 4

Discussion

Chapter 4
DISCUSSION

4.1. BIOINFORMATIC ANALYSIS

Multiple sequence alignments were used to investigate the relatedness and domain structure of Hsj1a and other selected Hsp40s, and to identify conserved regions, and amino acid residues for mutagenic substitutions that are potentially crucial within the J-domain. The results showed that of those selected proteins, the proteins were more closely related to each other within their relevant type classes as expected (Figure 8; Hennessy *et al.*, 2000). In addition, Pfj4, of unknown cell biology and function, was the most closely related protein to Hsj1a, due to the potential conservation of certain domains and motifs across their lengths (Figure 9). This study has identified a region of Pfj4 that corresponded significantly to the known UIM1 of Hsj1a (Figure 9). This finding suggested that Pfj4 may in fact function in a manner similar to that of Hsj1a, utilising ubiquitin or similar compounds by linking them to target proteins for a cellular purpose (such as that of degradation for the Hsj1a-proteasomal system). Also, the first half of the putative chaperone domain of Hsj1a is significantly similar in Pfj4 (Pfj4 residues: 166-184; Figure 9) suggesting the Pfj4 may in fact have a chaperone domain of its own, but the Pfj4 domain may perhaps be shorter in length or may differ in the second half of the domain for reasons of Hsp70 interaction specificity. This study also identified an additional region within the C-terminus of Pfj4 that showed as conserved with respect to Hsj1a, but this region has not been investigated further (Pfj4 residues: 227-236; Figure 9). This potential motif conservation suggests a similar mechanism of Hsp70-Hsp40 interaction, either (i) for the two proteins within the malarial or human systems respectively, or (ii) for the two proteins in a cross-reacting manner between these two systems.

Homology models of the J-domains of Hsj1a and Pfj4 were created alongside that of *Agt*DnaJ to examine the possibility of similar spatial arrangements of these J-domains for which no structures are available. The J-domains showed similar structural composition to the known structure of *E. coli* DnaJ, showing two α-helices flanking each side of a loop region.

Studies of J-domains swaps have shown that there is a level of specialisation required to instruct correct Hsp70-Hsp40 partner interactions at a domain level. Furthermore, at an amino acid level, studies have shown that the variation in amino acid residue arrangement potentially affects specificity or promiscuity of these J-domain-based Hsp70-Hsp40 interactions (Hennessy *et al.*, 2000; Hennessy *et al.*, 2005b). Type-I and type-II proteins have a greater degree of J-domain conservation overall than the type-III proteins, whilst type-III proteins are thought to be involved in more specialised functioning than the more general type-I and type-II Hsp40s (Hennessy *et al.*, 2000). Since type-I and type-II Hsp40s are more highly conserved overall across their J-domain regions, this would suggest that domain-swapping of J-domains could be more easily done between these types as the level of interactions could be similar, when compared to type-III Hsp40s (Hennessy *et al.*, 2005b). This suggested that domain-swapping of J-domains from various Hsp40s into *Agt*DnaJ would allow for the investigation into the functioning of those J-domains.

4.2. HETEROLOGOUS EXPRESSION

In this study, we confirmed the use of a heterologous expression system using *E. coli* to produce *Agt*DnaJ. Our results support those of Hennessy and co-workers (Hennessy *et al.*, 2005a) which showed that the transformation of the pRJ30 plasmid, encoding the *Agt*DnaJ protein, in *E. coli* OD259 reversed the thermosensitivity of the strain, and that conversely the pRJ30(H33Q) plasmid, encoding *Agt*DnaJ(H33Q), was unable to do so. This system was subsequently used for testing whether diverse type-II J-domains, could effectively replace the type-I J-domain of *Agt*DnaJ in this system yielding functional chimeric proteins as a means of testing J-domain function of Hsp40s. Similarly to the *Agt*DnaJ studies, the protein chimeras containing each of the J-domains from Hsj1a and Pfj4 swapped into *Agt*DnaJ, namely Hsj1*Agt* and Pfj4*Agt*, were able to aid effective growth of *E. coli* OD259 under heat shock conditions in the *in vivo* assays conducted. This showed that two diverse type-II J-domains, could each effectively replace the type-I J-domain of *Agt*DnaJ in this system yielding functional chimeric proteins. This proposed a system whereby mutagenesis of those previously identified key J-domain amino acid residues of Hsj1a and Pfj4 could be performed to test the function of individual conserved J-domain residues.

4.3. AMINO ACIDS IMPORTANT IN THE J-DOMAIN

Several amino acid residues were identified for mutagenesis investigation, equivalent to those studied in previous work (Genevaux *et al.*, 2002; Hennessy *et al.*, 2005a). Interestingly, the critical functional residues appear to be located within helices II and the loop region, which project to the solvent-exposed face of the J-domain (Genevaux *et al.*, 2002), whilst those critical structural residues appear to be located within helix I. This could add significance to the suggested model of Hsp70-Hsp40 interaction as recently suggested by Nicoll and co-workers (Nicoll *et al.*, 2005).

In this study, Hsj1*Agt*(Y5A) and Pfj4*Agt*(Y8A) were only partially able to reverse the thermosensitivity of *E. coli* OD259. The equivalent *Agt*DnaJ mutant protein, *Agt*DnaJ(Y7A) was shown to be unable to reverse the thermosensitivity of *E. coli* OD259, possibly due to the conserved tyrosine residue being implicated in structural integrity of the J-domain, along with other residues such as L10 and L57 in *Agt*DnaJ (Hennessy *et al.*, 2005a). This suggested that the reduction of the tyrosine side-chain to an alanine side-chain of relatively reduced size, could be hindering the formation of necessary intra-molecular forces that are key in the maintenance of J-domain structural integrity as previously suggested.

Hsj1*Agt*(L8A) and Pfj4*Agt*(L11A) also did not reverse thermosensitivity, suggesting that this conserved leucine residue was also important in maintenance of J-domain structure. The equivalent mutant protein, *Agt*DnaJ(L10A), also showed similar non-functionality (Hennessy *et al.*, 2005a).

Analysis of the *in vivo* function of proteins Hsj1*Agt*(R24A) and Pfj4*Agt*(R27A) suggested that the R24 residue of Hsj1a was more important than the R27 residue of Pfj4 since Pfj4*Agt*(R27A) was still partially able to reverse the thermosensitivity of *E. coli* OD259. However, in both J-domains this residue appeared to be necessary for optimal protein function. Results from Hennessy and co-workers suggested that the positively charged helix II residues were important in the potential interaction with the ATPase domain of Hsp70s (Hennessy *et al.*, 2005a; Hennessy *et al.*, 2005b). Thus substitution of these residues may render the Hsp40s unable to interact with the endogenous DnaK from *E. coli* OD259 yielding reduced or no growth under heat shock. These results were consistent with those obtained for the equivalent K26A substitution in *E. coli* DnaJ (Genevaux *et al.*, 2002).

The Hsj1*Agt*(H31Q) and Pfj4*Agt*(H34Q) mutant proteins, could not reverse the thermosensitivity of *E. coli* OD259, suggesting that this histidine residue was also conserved for function in these two type-II Hsp40s. *E. coli* DnaJ(H33Q) was also unable to compensate for the lack of endogenous Hsp40s in *E. coli* OD259 (Genevaux *et al.*, 2002). Similarly, *Agt*DnaJ(H33Q) could not substitute for missing Hsp40s in *E. coli* OD259, and was found to be unable to stimulate the ATPase activity of *Agt*DnaK (Hennessy *et al.*, 2005a). These findings were not unexpected since this histidine is part of the conserved HPD tripeptide motif known to be essential for Hsp40 function (Genevaux *et al.*, 2002; Landry, 2003; Hennessy *et al.*, 2005a).

In helix IV of *Agt*DnaJ, more than 50 % of the helix consists of charged residues (Hennessy *et al.*, 2005a). The conservative D59N substitution in *Agt*DnaJ did not disrupt the *in vivo* function of *Agt*DnaJ, whereas a non-conservative D59A substitution caused a loss in ability to reverse *E. coli* OD259 thermosensitivity (Hennessy *et al.*, 2005a). This suggested that there was a structural and/or functional role for a negatively charged residue at this position in *Agt*DnaJ. Whilst this might be the situation for *Agt*DnaJ, non-conservative changes in Hsj1*Agt* and Pfj4*Agt* of the equivalent D59A mutations, suggested that this residue was not critical in the functioning of these J-domains. However, should the aspartic acid residue be involved in structural maintenance of the helix IV region of the J-domain, or in the interaction with the remainder of the protein beyond the J-domain, the presence of the *Agt*DnaJ downstream regions may in fact have a negative affect on the results obtained for this study. For this reason, mutagenesis studies on the full-length Hsj1a and Pfj4 proteins may be required to supplement this information further.

These results could imply that of these conserved residues, the helix I residues could to be important in structural integrity, whilst the helix II and loop region substitutions performed could be critical in function, from a comparison to other known systems. However, we must not totally exclude the possibility that these residues could be important in either a functional or structural manner respectively, which can only be clarified once the structures of the Hsj1a and Pfj4 proteins are determined. The purpose and importance of the negatively charged aspartic acid residue corresponding to D59 in *Agt*DnaJ, is still unknown.

OLIGOMERIC ANALYSIS

Histidine-tagged Hsj1a was found to have a size of approximately 46 kDa (Figure 29), which suggested neither a monomeric (36 kDa) nor a dimeric (72 kDa) form of the protein. Histidine-tagged Hsj1a thus either exists in solution as: a large extended monomer, a tightly compacted dimer, or a higher order oligomer that was degraded before or during the FPLC analysis where only a portion of the protein remained intact and detectable. Other studies have shown that Hsp40s, such as yeast SisI, occur as dimers (Langer *et al.*, 1992; Sha *et al.*, 2000; Shi *et al.*, 2005; Wu *et al.*, 2005) and hence it would be suspected that the second two scenarios were more likely than the first scenario given above. However, this study is inconclusive in providing a true reflection of the oligomeric structure of Histidine-tagged Hsj1a.

4.4. IDENTIFICATION OF CHAPERONE PAIRS

It is an established fact that Hsp70-Hsp40 interactions can be promiscuous (Table 11). Furthermore, certain *P. falciparum* Hsp40s have been shown to be exported into the host erythrocyte cytosol (Sargeant and Marti *et al.*, 2006) suggesting they may interact with the human Hsp70. Therefore the ability of both Hsj1a and Pfj4 to interact with human Hsp70 was investigated.

This study has shown that histidine-tagged Hsj1a was also able to interact with the inducible human Hsp70. Human inducible Hsp70 had a stimulated ATPase activity of approximately 14 nmol P_i/min/mg in the presence of histidine-tagged Hsj1a. Hsj1a was previously found to interact with constitutive bovine Hsc70 (Cheetham *et al.*, 1994) with an approximate 5 to 6 fold stimulation (Table 11).

Conversely to the histidine-tagged Hsj1a scenario, histidine-tagged Pfj4 was unable to interact with inducible human Hsp70. Now whilst some type-II Hsp40s move into the host erythrocyte cytosol in the malarial situation, the cell biology of type-II Pfj4 is unknown. Thus Pfj4: (i) could be exported and interact with the human Hsp70, such as inducible human Hsp70, (ii) could interact with another human Hsp70 isoform, or (iii) could interact with a parasitic Hsp70, such as cytosolic PfHsp70 or ER PfBiP for example.

Whilst the results of this study suggested that Pfj4 did not functionally interact with inducible human Hsp70, it has not confirmed the occurrence of a corresponding Hsp70 chaperone partner to Pfj4, as this was outside the scope of the project and more Hsp70s need to be tested for this.

Table 11: Chronological comparison of basal and stimulated ATPase activities of some Hsps

Hsp	Source	Basal ATPase activity (nmol/min/mg)	Stimulated ATPase activity (nmol/min/mg)	Approximate Fold Stimulation	Stimulated By
DnaK		NR	NR	2	DnaJ[1]
Ssa1	*S. cerevisiae*	NR	NR	10	Ydj1[2]
Bip	Bovine liver	0.35	2.1	6	Synthetic peptides[3]
Hsc70	Bovine brain	NR	NR	5 - 6	Hsj1a[4]
Hsc70	Bovine brain	NR	NR	5 - 6	Hsj1b[4]
Hsc70	Bovine brain	0.2	1.5	8	DnaJ and GrpE[5]
Hsc70	Bovine brain	0.2	1.25	6	Hsp40[5]
Hsc70	*R. norvegicus*	1	7	7	Hdj1 and Bag1[6]
Hsc70	Bovine brain	1.08	13.98	13	Csp1[7]
Hsp70	Recombinant *H. sapiens*	0.5	4.66	9	Csp1[7]
Ssa1	Recombinant *S. cerevisiae*	5	28	6	Ydj1[8]
Ssa1	Recombinant *S. cerevisiae*	5	24	5	Sis1[8]
Ssa1	*S. cerevisiae*	1.25	9	7	Ydj1[9]
Hsp70	Recombinant His-tagged *T. cruzi*	40	60	1.5	Tcj2 and RCMLA[10]
Hsc70	Bovine brain	1.2	16.5	14	Tcj2[10]
Hsp70	Recombinant His-tagged *H. sapiens*	ND	ND		Pfj4[11]
Hsp70	Recombinant His-tagged *H. sapiens*	ND	14		Hsj1a[11]

Adapted and expanded from Edkins *et al.*, 2004. [1] Liberek *et al.*, 1991; [2] Cyr *et al.*, 1992; [3] Blond-Elquindi *et al.*, 1993; [4] Cheetham *et al.*, 1994; [5] Minami *et al.*, 1996; [6] Kanazawa *et al.*, 1997; [7] Chamberlain and Burgoyne, 1997b; [8] Lu and Cyr, 1998; [9] Fewell *et al.*, 2001; [10] Edkins *et al.*, 2004; [11] This study. NR = not reported. ND = non-detectable.

Hence, Hsj1a was found to interact with both constitutive (Cheetham *et al.*, 1994) and inducible forms (this study) of Hsp70. And Pfj4 was found to not interact with an inducible human Hsp70.

Chapter 5

Conclusions and Future Work

Chapter 5
CONCLUSIONS AND FUTURE WORK

5.1. CONCLUSIONS

5.1.1. Bioinformatic analysis

Hsj1a links misfolded or unfolded proteins to the proteasomal system for degradation. It was interesting that Pfj4 was found to exhibit similar domain structure to Hsj1a, having a potential chaperone domain, UIM-like motif, and additional region of significant conservation of unknown function. The domain similarity lead to the proposal that Pfj4 may have been involved in a similar mechanism of interaction within the malarial parasite, to that of Hsj1a.

5.1.2. Heterologous expression

The use of *E. coli* to successfully produce Hsp40 proteins that are able to replace the lacking endogenous Hsp40s within *E. coli* OD259 (namely *Agt*DnaJ, and the J-domain chimeras and mutants thereof) was confirmed. Proteins were produced for *in vivo* and *in vitro* testing of protein structure and function at the whole protein, domain and amino acid levels.

5.1.3. Amino acids important in the J-domain

Equivalent amino acid substitutions were made on Hsj1a and Pfj4 J-domains based on rationale of key residues previously identified within other studied Hsp40s, such as *E. coli* DnaJ and *Agt*DnaJ. These residues were shown to be important in J-domain structure and Hsp70-Hsp40 interaction. Results from this study showed that of these conserved residues, the helix I residues equivalent to Y7 and L10 of *Agt*DnaJ were potentially important in structural integrity, whilst the helix II and loop region substitutions performed equivalent to R26 and H33 of *Agt*DnaJ respectively, seemed to suggest these residues were critical in function. The purpose and importance of the negatively charged aspartic acid residue corresponding to D59 of *Agt*DnaJ, however, was still unknown.

5.1.4. Oligomeric analysis

The preliminary investigation into determining the oligomeric state of Hsj1a from FPLC analysis was inconclusive.

5.1.5. Identification of chaperone pairs

Hsj1a was found to be able to act as a co-chaperone to inducible human Hsp70, stimulating its ATPase activity, whilst Pfj4 was unable to stimulate the ATPase activity of human Hsp70. This suggested that Pfj4 functions with another Hsp70 chaperone partner, potentially a malarial Hsp70 within the malarial parasite, instead of interacting across species with a host Hsp70.

5.2. FUTURE WORK

Many experiments could still be carried out on these two proteins of interest. Some of these experiments are mentioned here.

5.2.1. Improved purification of Hsj1a and Pfj4 for structural studies

Whilst no DnaK contamination was detected in this study, purification of histidine-tagged Hsj1a and histidine-tagged Pfj4 from a DnaK$^-$ *E. coli* strain is suggested. This would ensure that no endogenous DnaK contamination results which may be detrimental to later downstream experiments, such as ATPase assays or an oligomeric analysis. Alternatively, the FPLC could be repeated under different conditions, as the determination of the structure of this protein would be advantageous and could be performed by solution NMR or X-ray diffraction of crystals. In addition to FPLC, the less costly method of ultracentrifugation could be used to determine the Stokes' radius of the protein to determine the oligomeric state. A bioinformatic analysis could also be used to determine whether Hsj1a has a dimerization motif. If such a motif does exist, mutagenic substitutions, or deletions of the motif could be used to study the effects of the oligomerization state of Hsj1a.

5.2.2. Studies on Hsj1a and Pfj4 J-domains chimeric proteins and their mutants

The examination of the Hsj1a and Pfj4 chimeric proteins and their mutant derivatives that were created in this study could be done. This would involve an investigation of the ATPase activities of these chimeric proteins and mutant derivatives, using ATPase assays as performed in this study on the full-length Hsj1a and Pfj4 proteins. These chimeric proteins and mutant derivatives could also be purified from the improved purification method to ensure no Hsp70 contamination (Section 5.2.1).

Binding studies on the purified Hsj1a chimeric proteins and each of the mutant proteins could also be performed to determine the effect that specific residues have on the binding of the J-domain to identified Hsp70 chaperone partners, such as inducible human Hsp70. This could be done using the sophisticated techniques of surface plasmon resonance (SPR), spectroscopy, or equivalent technologies.

5.2.3. Studies on full-length Histidine-tagged Pfj4

5.2.3.1. UIM-motif analysis for Pfj4

Pfj4 potentially has a region that corresponds to UIM1 of Hsj1a. Further studies into whether this is a true UIM-like motif are needed and could be carried out using domain-swapping of these regions to investigate potential function by means of ubiquitin degradation assays. Similarly, substitutions of highly conserved residues within the potential UIM-motif could also be conducted in order to determine which residues within the motif are of the highest functional importance.

5.2.3.2. Localisation studies for Pfj4

The localisation of Pfj4 is still unknown. Pfj4 could be studied using bioinformatic analysis to determine whether it has an export signal to translocate it into the host erythrocyte cytosol in the infected scenario. Localisation studies could be performed in both the parasite and in the infected erythrocyte. This could be done along with immunoprecipitation in order to find the localisation of Pfj4-containing complexes, and any interacting proteins could be identified in this manner.

5.2.3.3. Identification of an Hsp70-Hsp40 chaperone partner for Pfj4

Whilst this study suggests that Pfj4 could not interact with inducible human Hsp70, a chaperone partner should be identified for this type-II malarial protein. Additional ATPase assays could be performed with Pfj4 and various *P. falciparum* Hsp70s, such as PfHsp70 and PfBiP. Ideally, all recombinant *E. coli* proteins used in these studies should be purified from a DnaK$^-$ strain of *E. coli*. Alternative sources can also be used to purify Hsp40s from, such as using yeast expression systems, eukaryotic insect or mammalian expression systems. Natural protein sources, such as Pfj4 from *P. falciparum* parasites and Hsj1a from human tissue, could also advantageous as post-translational modifications would be ensured.

Chapter 6

References

REFERENCES

Agarraberes, F. A., Dice, J. F. (2001) A molecular chaperone complex at the lysosomal membrane is required for protein translocation, *Journal of Cell Science*, **114**, 2491-2499.

Agashe, V. R., Hartl, F. U. (2000) Roles of molecular chaperones in cytoplasmic folding, *Seminars in Cell and Developmental Biology*, **11**, 15-25.

Aron, R., Lopez, N., Walter, W., Craig, E. A., Johnson, J. (2005) *In vivo* bipartite interaction between the Hsp40 Sis1 and Hsp70 in *Saccharomyces cerevisiae*, *Genetics*, **169**, 1873-1882.

Bahl, A., Brunk, B., Crabtree, J., Fraunholz, M. J., Gajria, B., Grant, G. R., Ginsburg, H., Gupta, D., Kissinger, J. C., Labo, P., Li, L., Mailman, M. D., Milgtam, A. J., Pearson, D. S., Roos, D. S., Schug, J., Stoeckert, C. J., Whetzel, P. (2003) PlasmoDB: the *Plasmodium* genome resource, A database integrating experimental and computational data, *Nucleic Acids Research*, **31**, 1, 212-215.

Bailly, V., Lauder, S., Prakash, S., Prakash, L. (1997) Yeast DNA repair proteins Rad6 and Rad18 form a heterodimer that has ubiquitin conjugating, DNA binding, and ATP hydrolytic activities, *Journal of Biological Chemistry*, **272**, 37, 23360-23365.

Banecki, B., Liberek, K., Wall, D., Wawrzynow, A., Georgopoulos, C., Bertoli, E., Tanfani, F., Zylicz, M. (1996) Structure-function analysis of the zinc finger region of the DnaJ molecular chaperone, *Journal of Biological Chemistry*, **271**, 25, 14840-14848.

Banumathy, G., Singh, V., Tatu, U. (2002) Host chaperones are recruited in membrane-bound complexes by *Plasmodium falciparum*, *Journal of Biological Chemistry*, **277**, 6, 3902-3912.

Bardwell, J. C. A., Tilly, K., Craig, E., King, J., Zylicz, M., Georgopolous, C. (1986) The nucleotide sequence of the *Escherichia coli* K12 *dnaJ*$^{+}$ gene, *Journal of Biological Chemistry*, **261**, 4, 1782-1785.

Barral, J. M., Broadley, S. A., Schaffar, G., Hartl, F. U. (2004) Roles of the molecular chaperones in protein misfolding diseases, *Seminars in Cell and Developmental Biology*, **15**, 17-29.

Baruch, D. I., Ma, X. C., Singh, H. B., Bi, X., Pasloske, B. L., Howard, R. J. (1997) Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequences, *Blood*, **90**, 9, 3766-3775.

Beaucamp, N., Harding, T. C., Geddes, B. J., Williams, J., Uney, J. B. (1998) Overexpression of Hsp70i facilitates reactivation of intracellular proteins in neurones and and protects them from denaturing stress, *Federation of European Biochemical Socities Letters*, **441**, 215-219.

Becker, J., Craig, E. A. (1994) Heat shock proteins as molecular chaperones, *European Journal of Biochemistry*, **219**, 1, 11-23.

Benaroudj, N., Fouchaq, B., Ladjimi, M. M. (1997) The COOH-terminal peptide binding domain is essential for self-association of the molecular chaperone HSC70, *Journal of Biological Chemistry*, **272**, 13, 8744-8751.

Benaroudj, N., Triniolles, F., Ladjimi, M. M. (1996) Effect of nucleotides, peptides, and unfolded proteins on the self-association of the molecular chaperone HSC70, *Journal of Biological Chemistry*, **271**, 31, 18471-18476.

Beissinger, M., Buchner, J. (1998) How chaperones fold proteins, Biological Chemistry, **379**, 245-249.

Berjanskii, M. V., Riley, M. I., Xie, A., Semenchenko, V., Folk, W. R., Van Doren, S. R. (2000) NMR structure of the N-terminal J domain of murine polyomavirus T antigens: Implications for DnaJ-like domains and for mutations of T antigens, *Journal of Biological Chemistry*, **275**, 36094-36103.

Birnboim, H., Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucleic Acids Research*, **7**, 1513-1523.

Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., Gething, M. (1993) Peptide dependent stimulation of the ATPase activity of the molecular chaperone BiP is the result of conversion of oligomers to active monomer. *Journal of Biological Chemistry*, **268**, 12730-12735.

Bonner, J. J., Ballou, C., Fackenthal, D. L. (1994) Interactions between DNA-bound trimers of the yeast heat shock factor, *Molecular and Celullar Biology*, **14**, 1, 501-508.

Borges, J. C., Fischer, H., Craievich, A. F., Ramos, C. H. I. (2005) Low resolution structural study of two human HSP40 chaperones in solution, *Journal of Biological Chemistry*, **280**, 14, 13671-13681.

Bork, P., Sander, C., Valencia, A., Bukau, B. (1991) A module of the DnaJ heat shock proteins found in malaria parasites. *Trends in Biochemical Sciences*, **17**, 129.

Boshoff, A., Hennessy, F., Blatch, G. L. (2004) The *in vivo* and *in vitro* characterisation of DnaK from *Agrobacterium tumefaciens* RUOR, *Protein Expression and Purification*, **38**, 161-169.

Bradford, M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein-dye binding, *Analytical Chemistry*, **72**, 248-254.

Buchberger, A., Theyssen, H., Schröder, H., McCarty, J. S., Virgallita, G., Milkereit, P., Reinstein, J., Bukau, B. (1995) Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication, *Journal of Biological Chemistry*, **270**, 16903-16910.

Buchberger, A., Schröder, H., Hesterkamp, T., Schonfeld, H. J., Bukau, B. (1996) Substrate shuttling between the DnaK and GroEL systems indicate a chaperone network promoting protein folding, *Journal of Molecular Biology*, **261**, 3, 328-333.

Bullock, W. O., Fernandez, J. M., Short, J. M. (1987) XL1-blue: A High Efficiency Plasmid Transforming recA *Escherichia coli* strain with β-galactosidase selection, *Biotechniques*, **5**, 4, 376-379.

Cagney, G., Amiri, S., Premawaradena, T., Lindo, M., Emili, A. (2003) *In silico* proteome analysis to facilitate proteomics experiments using mass spectrometry, *Proteome Science*, **1**, 5.

Campbell, K. S., Mullane, K. P., Aksoy, I. A., Stubdal, H., Zalvide, J., Pipas, J. M., Silver, P. A., Roberts, T. M., Schaffhausen, B. S., DeCaprio, J. A. (1997) DnaJ/hsp40 chaperone domain of SV40 large T antigen promotes efficient viral DNA replication, *Genes and Development*, **11**, 1098-1110.

Caplan, A. J. (2003). What is a co-chaperone?, *Cell Stress and Chaperones*, **8**, 2, 105-107.

Casey, P. J., Seabra, M. C. (1996) Protein prenyltransferases, *Journal of Biological Chemistry*, **271**, 10, 8744-8751.

Chamberlain, L. H., Burgoyne, R. D. (1997a) Activation of the ATPase activity of heat-shock proteins Hsc70/Hsp70 by cysteine-string protein, *Biochemistry*, **322**, 853-858.

Chamberlain, L. H., Burgoyne, R. D. (1997b) The molecular chaperone function of the secretory vesicle cysteine string proteins, *Journal of Biological Chemistry*, **272**, 50, 31420-31426.

Chapple, J. P., Cheetham, M. E. (2003) The Chaperone Environment at the Cytoplasmic Face of the Endoplasmic Reticulum Can Modulate Rhodopsin Processing and Inclusion Formation, *Journal of Biological Chemistry*, **278**, 21, 19087-19094.

Chapple, J. P., van der Spuy, J., Poopalasundaram, S., Cheetham, M. E. (2004) Neuronal DnaJ proteins HSJ1a and HSJ1b: a role in linking the Hsp70 chaperone machine to the ubiquitin-proteasome system?, *Biochemical Society Transactions*, **32**, 4, 640-642.

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

Cheetham, M. E., Brion, J., P., Anderton, B. H. (1992) Human homologues of the bacterial heat-shock protein DnaJ are preferentially expressed in neurons, *Biochemistry Journal*, **284**, 469-476.

Cheetham, M. E., Jackson, A. P., Anderton, B. H. (1994) Regulation of 70-kDa heat-shock-protein ATPase activity and substrate binding by human DnaJ-like proteins, HSJ1a and HSJ1b, *European Journal of Biochemistry*, **226**, 99-107.

Cheetham, M. E., Anderton, B. H., Jackson, A. P. (1996) Inhibition of hsc70-catalysed clathrin uncoating by HSJ1 proteins, *Biochemistry Journal*, **319**, 103-108.

Chifflet, S., Torriglia, A., Chiesa, R., Tolosa, S. (1988) A method for the determination of inorganic phosphate in the presence of labile organic phosphate and high concentrations of protein: application to lens ATPases, *Analytical Biochemistry*, **168**, 1-4.

Cohen, S. M., Chang, A. C. Y., Boyer, H. W., Helling, R. B. (1973) Construction of Biologically Functional Bacterial Plasmids *in vitro*, *Proceedings of the National Academy of Science USA*, **70**, 11, 3240-3244.

Cyr, D. M., Lu, X., and Douglas, M. G. (1992) Regulation of Hsp70 function by a eukaryotic DnaJ homolog, *Journal of Biological Chemistry*, **267**, 20927-20931.

Cyr, D. M., Langer, T., Douglas, M. G. (1994) DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70, *Trends in Biochemical Sciences*, **19**, 176-181.

Dagert, M., Ehrlich, S. D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells, *Gene*, **6**, 1, 23-38.

DeLano, W.L. (2005) The PyMOL Molecular Graphics System, DeLano Scientific LLC, California.

Deloche, O., Kelley, W. L., Gerogopolous, C. (1997) Structure-function Analyses of the Ssc1p, Mdj1p and Mge1p *Saccharomyces cerevisiae* Mitochondril Proteins in *Escherichia coli*, *Journal of Bacteriology*, **179**, 19, 6066-6075.

Deuerling, E., Bukau, B. (2004) Chaperone-assisted folding of newly synthesized proteins in the cytosol, *Critical Reviews in Biochemistry and Molecular Biology*, **39**, 261-277.

Edkins, A. L., Ludewig, M. H., Blatch, G. L. (2004) A *Trypanosoma cruzi* heat shock protein 40 is able to stimulate the adenosine triphosphate hydrolysis activity of heat shock protein 70 and can substitute for a yeast heat shock protein 40, *International Journal of Biochemistry and Cell Biology*, **36**, 1585-1598.

Ellis, R. J. (1987) Proteins as molecular chaperones, *Nature*, **328**, 378-379.

Ellis, R. J. (2001) Macromolecular crowding: obvious but underappreciated, *Trends in Biochemical Science*, **26**, 10, 597-604.

Ellis, R. J., Dobson, C., Hartl, U. (1998) Sequence does specify protein conformation, *Trends in Biochemical Sciences*, **23**, 12, 458.

Ellis, R. J., Hartl, F. U. (1999) Principles of protein folding in the cellular environment, *Current Opinions in Biology*, **9**, 102-110.

Ellis, R. J., van der Vies, S. M. (1991) Molecular chaperones, *Annual Review of Biochemistry*, **60**, 321-327.

Fahr, L., Mitchell, D. A., Deschenes, R. J. (1995) Farnesylation and proteolysis are sequential, but distinct steps in the CaaX box modification pathway, *Archives of Biochemistry and Biophysics*, **318**, 1, 113-121.

Fan, C. Y., Lee, P., Ren, H. Y., Cyr, D. M. (2004) Exchangeable chaperone modules contribute to specification of Type I and Type II Hsp40 cellular function, *Molecular Biology of the Cell*, **15**, 761-773.

Fan, C. Y., Ren, H. Y., Lee, P., Caplan, A. J., Cyr, D. M. (2005) The type I Hsp40 zinc finger-like region is required for Hsp70 to capture non-native polypeptides from Ydj1, *Journal of Biological Chemistry*, **280**, 1, 695-702.

Feige, U., Mollenhauer, J. (1992) Heat shock proteins, *Experientia*, **48**, 621-622.

Feldman, D. E., Frydman, J. (2000) Protein folding *in vivo*: the importance of molecular chaperones, *Current Opinion in Structural Biology*, **10**, 26-33.

Fenton, W. A., Horwich, A. L. (1997) GroEL-mediated protein folding, *Protein science*, **6**, 743-760.

Fewell, S. W., Pipas, J. M., Brodsky, J. L. (2002) Mutagenesis of a functional chimeric gene in yeast identifies mutations in the simian virus40 large T antigen J domain, *Proceedings of the National Academy of Science USA*, **99**, 2002-2007.

Fewell, S. W., Travers, K. J., Weissman, J. S., Brodsky, J. L. (2001) The action of molecular chaperones in the early secretory pathway, *Annual Review of Genetics*, **35**, 149-191.

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

Fliss, A. E., Rao, J., Melville, M. W., Cheetham, M. E., Caplan, A. J. (1999) Domain requirements of DnaJ-like (Hsp40) molecular chaperones in the activation of a steroid hormone receptor, *Journal of Biological Chemistry*, **274**, 48, 34045-34052.

Flynn, G. C., Chappell, T. G., Rothman, J. E. (1989) Peptide binding and release by proteins implicated as catalysts of protein assembly, *Science*, **245**, 385-390.

Flynn, G. C., Pohl, J., Flocco, M. T., Rothman, J. E. (1991) Peptide-binding specificity of the molecular chaperone BiP, *Nature*, **353**, 726-730.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*, *Nature*, **419**, 498-511.

Gassler, C. S., Buchberger, A., Laufen, T., Mayer, M. P., Schröder, H., Valencia, A., Bukau, B. (1998) Mutations in the DnaK chaperone affecting the interaction with the DnaJ cochaperone, *Proceedings of the National Academy of Science USA*, **95**, 15229-15234.

Genevaux, P., Lang, F., Schwager, F., Vartikar, J. V., Rundell, K., Pipas, J. M., Georgopoulos, C., Kelley, W. L. (2003) Simian virus 40 T antigen and J domains: analysis of Hsp40 cochaperone functions in *Escherichia coli*, *Journal of Virology*, **77**, 19, 10706-10713.

Genevaux, P., Schwager, F., Georgopoulos, C., Kelley, W. L. (2002) Scanning mutagenesis identifies amino acid residues essential for the *in vivo* activity of the *Escherichia coli* DnaJ (Hsp40) J-domain, *Genetics*, **162**, 1045-1053.

Georgopoulos, C., Welsh, W. J. (1993) Role of major heat shock proteins as molecular chaperones, *Annual Review of Cell Biology*, **9**, 601-634.

Gething, M. J., Sambrook, J. (1992) Protein folding in the cell, *Nature*, **355**, 32-45.

Gething, M. J. (1997) Protein folding: the difference with prokaryotes, *Nature*, **388**, 329-331.

Guex, N., Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling, *Electrophoresis*, **18**, 2714-2723.

Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F. U., Kuriyan, J. (1997) Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK, *Science*, **276**, 431-435.

Hartl, F. U. (1996) Molecular Chaperones in Cellular Protein Folding, *Nature*, **381**, 571-580.

Haslbeck, M. (2006) Recombinant expression and in vitro refolding of the yeast small heat shock protein Hsp42, *International Journal of Biological Macromolecules*, Article in press.

Hennessy, F., Boshoff, A., Blatch, G. L. (2005a) Rational mutagenesis of a 40 kDa heat shock protein from *Agrobacterium tumefaciens* identifies amino acid residues critical to its *in vivo* function, *International Journal of Biochemistry and Cell Biology*, **37**, 177-191.

Hennessy, F., Cheetham, M. E., Dirr, H. W., Blatch, G. L. (2000) Analysis of the levels of conservation of the J domain among the various types of DnaJ-like proteins, *Cell Stress and Chaperones*, **5**, 347-358.

Hennessy, F., Nicoll, W. S., Zimmermann, R., Cheetham, M. E., Blatch, G. L. (2005b) Not all J domains are created equal: implications for the specificity of Hsp40-Hsp70 interactions, *Protein Science*, **14**, 1697-1709.

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

Hiller, N. L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., Haldar, K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infections, *Science*, **306**, 1934-1937.

Hofmann, K., Falquet, L. (2001) A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems, *Trends in Biochemical Sciences*, **26**, 6, 347-350.

Höhfeld, J., Cyr, D. M., Patterson, C. (2001) From the cradle to the grave: molecular chaperones that may choose between folding and degradation, *European Molecular Biology Organisation Journal*, **2**, 885-890.

Höhfeld, J., Bohse, K., Genau, M., Westhoff, B. (2006) in **Networking of Chaperones by Co-chaperones**, From creator to terminator – co-chaperones that link molecular chaperones to the ubiquitin/proteasome system, (Blatch, G. L., ed.), *Published online at www.eureka.com.*

Huang, K., Flanagan, J. M., Prestegard, J. H. (1999) The influence of C-terminal extension on the structure of the "J-domain" in *E. coli* DnaJ, *Protein Science, 8, 203-214.*

Ingolia, T., Craig, E. A. (1982) *Drosophila* gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock, *Proceedings of the National Academy of Sciences USA*, **79**, 525-529.

Inoue, H., Nojima, H., Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids, *Gene*, **96**, 1, 23-28.

James, P., Pfund, C., Craig, E. A. (1997) Functional specificity among Hsp70 molecular chaperones, *Science*, **275**, 5298, 387-389.

Jiang, R., Gao, B., Prasad, K., Greene, L. E., Eisenberg, E. (2000) Hsc70 chaperones clathrin and primes it to interact with vesicle membranes, *Journal of Biological Chemistry*, **275**, 12, 8439-8447.

Jiang, J., Prasad, K., Lafer, E. M., Sousa, R. (2005) Structural basis of interdomain communication in the Hsc70 chaperone, *Molecular Cell*, **20**, 513-524.

Johnson, J .L., Craig, E. A. (2001) An essential role for the substrate-binding region of Hsp40s in *Saccharomyces cerevisiae*, *Journal of Cell Biology*, **19**, 851-856.

Kabani, M., Beckerich, J. M., Brodsky, J. L. (2003) The yeast Sls1p and Fes1p proteins define a new family of Hsp70 nucleotide exchange factors, *Current Genomics*, **4**, 273.

Kabani, M., Beckerich, J. M., Brodsky, J. L. (2002) Nucleotide exchange factor for the yeast Hsp70 molecular chaperone Ssa1p, *Molecular and Cellular Biology*, **22**, 4677-4689.

Kanazawa, M., Terada, K., Kato, S., Mori, M. (1997) HSDJ, a human homolog of DnaJ, is farnesylated and is involved in protein import into mitochondria, *Journal of Biochemistry*, **121**, 890-895.

Karzai, A. W., McMacken, R. (1996) A bipartite signalling mechanism involved in DnaJ-mediated activation of the *Escherichia coli* DnaK protein, *Journal of Biological Chemistry*, **271**, 11236-11246.

Kelley, W. L. (1998) The J-domain family and the recruitment of chaperone power, *Trends in Biochemical Sciences*, **23**, 222-227.

Kelley, W. L., Georgopoulos, C. (1997) The T/t common exon of simian virus 40, JC, and BK polyomavirus T antigens can functionally replace the J-domain of the *Escherichia coli* DnaJ molecular chaperone, *Proceedings of the National Academy of Science USA*, **94**, 3679-3684.

King, F. W., Wawryzynow, A., Höhfeld, J., Zylicz, M. (2001) Co-chaperones Bag-1, Hop and Hsp40 regulate Hsc70 and Hsp90 interactions with wild-type or mutant p53, *European Molecular Biology Organisation Journal*, **20**, 22, 6297-6305.

Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, *Nature*, **227**, 680-685.

Landry, S. J. (2003) Structure and energetics of an allele-specific genetic interaction between dnaJ and dnaK: correlation of nuclear magnetic resonance chemical perturbations in the J-domain of Hsp40/DnaJ with binding affinity for the ATPase domain of Hsp70/DnaK, *Biochemistry*, **42**, 17, 4926-4936.

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K., Hartl, F. U. (1992) Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding, *Nature*, **356**, 683-689.

Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., Candia, O. A. (1979) An improved assay for nanomole amounts of inorganic phosphate, *Analytical Biochemistry*, **100**, 95-97.

Laskey, R. A., Honda, B. M., Mills, A. D., Finch, T. T. (1978) Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA, *Nature*, **275**, 416-420.

Laufen, T., Mayer, M. P., Beistle, C., Klostermeier, D., Mogk, A., Reinstein, J., Bukau, B. (1999) Mechanism of regulation of Hsp70 chaperones by DnaJ cochaperones, *Proceedings of the National Academy of Science USA*, **96**, 5452-5457.

Levinthal, C. (1968) Are there pathways for protein folding, *Journal de Chimie Physique*, **65**, 44-45.

Li, H., Söderbärg, K., Houshmand, H., You, Z. Y., Magnusson, G. (2001) Effect on polyomavirus T-antigen function of mutants in a conserved leucine-rich segment of the DnaJ domain, *Journal of Virology*, **75**, 2253-2261.

Lieberek, K., Marszalek, J., Ang, D., Georgopoulos, C., Zylicz, M. (1991) *Escherichia coli* DnaJ and GrpE heat proteins jointly stimulate ATPase activity of DnaK, *Proceedings of the National Academy of Science USA*, **88**, 2874-2878.

Linke, K., Wolfram, T., Bussamer, J., Jakob, U. (2003) The roles of the two zinc binding sites in DnaJ, *Journal of Biological Chemistry*, **278**, 44457-44466.

Liu, C. P., Perrett, S., Zhou, J. M. (2005) Dimeric trigger factor stably binds folding-competent intermediates and cooperates with the DnaK-DnaJ-GrpE chaperone system to allow refolding, *Journal of Biological Chemistry*, **280**, 14, 13315-13320.

Liu, X. D., Liu, P. C. C., Santoro, N., Thiele, D. J. (1997) Conservation of a stress response: human heat shock transcription factors functionally substitute for yeast HSF, *European Molecular Biology Organisation Journal*, **16**, 21, 6466-6477.

Lu, Z., Cyr, D. M. (1998a) The Conserved Carboxyl Terminus and Zinc Finger-like Domain of the Co-chaperone Ydj1 Assist Hsp70 in Protein Folding, *Journal of Biological Chemistry*, **273**, 5970-5978.

Lu, Z., Cyr, D. M. (1998b) Protein folding activity of Hsp70 is modified differentially by the Hsp40 co-chaperones Sis1 and Ydj1, *Journal of Biological Chemistry*, **273**, 43, 27824-27830.

Luke, M. M., Sutton, A., Arndt, K. T. (1991) Characterization of SIS1, a *Saccharomyces cerevisiae* homologue of bacterial dnaJ proteins, *Journal of Cell Biology*, **114**, 623-638.

Marti, M., Good, R. T., Rug, M., Kneupfer, E., Cowman, A. F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte, *Science*, **306**, 1930 – 1933.

Martinko, J. M. (2000) in **Brock: Biology of Microorganisms**, 9th edition, Chapter 24: Animal-transmitted, vectorborne, and common source microbial diseases, (Corey, P. F., ed.), pp. 966-969, Prentice-Hall, Inc., New Jersey.

Matambo, T., Odunuga, O. O., Boshoff, A., Blatch, G. L. (2004) Over-production, purification, and characterization of the *Plasmodium falciparum* heat shock protein 70, *Protein Expression and Purification*, **33**, 214-222.

Mayer, M. P., Laufen, T., Paal, K., McCarty, J.S., Bukau, B. (1999) Investigation of the interaction between DnaK and DnaJ by surface plasmon resonance spectroscopy, *Journal of Molecular Biology*, **289**, 1131-1144.

Mayer, M. P., Bukau, B. (2005) Hsp70 chaperones: cellular functions and molecular mechanisms, *Cellular and Molecular Life Sciences*, **62**, 670-684.

Mayer, M. P., Schröder, H., Rudiger, S. Paal, K., Laufen, T., Bukau, B. (2000) Multistep mechanism of substrate binding determines chaperone activity of Hsp70, *Nature Structural Biology*, 7, 7, 586-593.

McCarty, J. S., Buchberger, A., Reinstein, J., Bukau, B. (1995) The role of ATP in the functional cycle of the DnaK chaperone system, *Journal of Molecular Biology*, **249**, 126-137.

Messing J., Crea, R., Seeburg, P. H. (1981) A system for shotgun DNA sequencing, *Nucleic Acids Residues*, **9**, 309-321.

Minami, Y., Höhfeld, J., Ohtsuka, K., Hartl, F. U. (1996) Regulation of the heat-shock protein 70 reaction cycle by the mammalian DnaJ homologue Hsp40, *Journal of Biological Chemistry*, **271**, 19617-19624.

Morimoto, R. I., Kline, M. P., Bimston, D. N., Cotto, J. J. (1997) The heat-shock response: regulation and function of heat-shock proteins and molecular chaperones, *Essays in Biochemistry*, **32**, 17-29.

Morshauer, R. C., Hu, W., Wang, H., Pang, Y., Flynn, G. C., Zuiderweg, E. R. (1999) High-resolution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70, *Journal of Molecular Biology*, **289**, 1387-1403.

Nambara, E., McCourt, P. (1999) Protein farnesylation in plants: a greasy tale, *Current Opinion in Plant Biology*, **2**, 388-392.

Nardai, G., Vegh, E. M., Prohaszka, Z., Csermely, P. (2006) Chaperone-related immune dysfunction: an emergent property of distorted chaperone networks, *Trends in Immunology*, Article in press.

Naylor, D. J., Hoogenraad, N. J., Hoj, P. B. (1999) Characterisation of several Hsp70 interacting proteins from mammalian organelles, *Biochimica and Biophysica Acta*, **1431**, 443-450.

Nicoll, W. S., Boshoff, A., Ludewig, M. H., Hennessy, F., Jung, M., Blatch, G. L. (2005) Approaches to the isolation and characterization of molecular chaperones, *Protein Expression and Purification*, **46**, 1, 1-15.

Ohki, M., Tamura, F., Nishimura, S., Uchida, H. (1986) Nucleotide Sequence of the *Escherichia coli dnaJ* Gene and Purification of the Gene Product, *Journal of Biological Chemistry*, **261**, 4, 1778-1781.

Ohtsuka, K., Hata, M. (2000) Mammalina HSP40/DNAJ homologs: cloning of novel cDNAs and a proposal for their classification and nomenclature, *Cell Stress and Chaperones*, **5**, 2, 98-112.

Osipiuk, J., Walsh, M. A., Freeman, B.C., Morimoto, R. I., Joachimiak, A. (1999) Structure of a new crystal form of human Hsp70 ATPase domain, *Acta Crystallographica*, **D55**, 1105-1107.

Page, R. D. M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers, *Computer Applications in the Biosciences*, **12**, 357-358.

Palleros, D. R., Welch, W. J., Finch, A. L. (1991) Interaction of hsp70 with unfolded proteins: effects of temperature and nucleotides on the kinetics of binding, *Proceedings of the National Academy of Science USA*, **88**, 5719-5723.

Pauli, D., Arrigo, A. P., Tissieres, A. (1992) Heat shock response in *Drosophila*, *Experientia*, **48**, 623-629.

Peitsch, M. C. (1995) Protein modeling by E-mail, *Bio/Technology*, **13**, 658-660.

Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C., Wuthrich, K. (1996) NMR structure of the J-domain and the Gly/Phe-rich region of the *Escherichia coli* DnaJ chaperone, *Journal of Molecular Biology*, **260**, *236-250.*

Pellechia, M., Montgomery, D. L., Stevens, S. Y., Vander Kooi, C. W., Feng, H. P., Gierasch, L. M., Zuiderweg, E. R. P. (2000) Structural insights into substrate binding by the molecular chaperone DnaK, *Nature Structural Biology*, **7**, 298-303.

Pflanz, R., Hoch, M. (2000) Dtrap-1 encodes a novel member of the heat shock protein super family of proteins and is expressed in derivatives of all three germ layers during *Drosophila* embryogenesis, *Mechanisms of development*, **96**, 219-222.

Pfund, C., Yan, W., Craig, E. A. (2001) in **Molecular chaperones in the cell**, The roles of the major cytoplasmic chaperones in normal eukaryotic growth: Hsc70 and its co-factors, pp. 119 – 141, Oxford university press, Oxford.

Qian, X., Hou, W., Zhengang, L., Sha, B. (2002) Direct interactions between molecular chaperones heat-shock protein (Hsp) 70 and Hsp40: yeast Hsp70 SsaI binds the extreme C-terminal region of yeast Hsp40 SisI, *Biochemistry Journal*, **361**, 27-34.

Qian, Y. Q., Patel, D., Hartl, F. U., McColl, D. J. (1996) Nuclear magnetic resonance solution structure of the human Hsp40 (HDJ-1) J-domain, *Journal of Molecular Biology*, **260**, *224-235.*

Raabe, T., Manley, J. L. (1991) A human homologue of the *Escherichia coli* DnaJ heat-shock protein, *Nucleic Acids Residues*, **19**, 23, 6645.

Richarme, G., Kohiyama, M. (1993) Specificity of the *Escherichia coli* chaperone DnaK (70-kDa heat shock protein) for hydrophobic amino acids, *Journal of Biological Chemistry*, **268**, 32, 24074-24077.

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

Ritossa, F. (1996) Discovery of the heat shock response, *Cell Stress and Chaperones*, **1**, 97-98.

Rudiger, S., Germeroth, L., Scheinder-Mergener, J., Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries, *European Molecular Biology Organisation Journal*, **16**, 1501-1507.

Russell, R., Karzai, A. W., Mehl, A. F., McMacken, R. (1999) DnaJ dramatically stimulates ATP hydrolysis by DnaK: insight into targeting of Hsp70 proteinsto polypeptide substrates, *Biochemistry*, **38**, 16412-16418.

Sambrook, J., Russell, D. W. (2001) in **Molecular Cloning: A Laboratory Manual**, 3rd edition, Cold Spring Harbor Laboratory Press, New York.

Sargeant, T. J., Marti, M., Caler, E., Carlton, J. M., Simpson, K., Speed, T. P., Cowman, A. F. (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites, *Genome Biology*, 7, 2, R12.

Schwede, T., Kopp, J., Guex, N., Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server, *Nucleic Acids Research*, **31**, 3381-3385.

Sha, B., Lee, S., Cyr, D. M. (2000) The crystal structure of the peptide-binding fragment from the yeast Hsp40 protein Sis1, *Structure*, **8**, 8, 799-807.

Shi, Y. Y., Hong, X. G., Wang, C. C. (2005) The C-terminal (331-376) sequence of *Escherichia coli* DnaJ is essential for dimerization and chaperone activity, *Journal of Biological Chemistry*, **280**, 24, 22761-22768.

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

Slepenkov, S. V., Patchen, B., Peterson, K. M., Witt, S. N. (2003) Importance of the D and E helices of the molecular chaperones DnaK for ATP binding and substrate release, *Biochemistry*, **42**, 5867-5876.

Sondermann, H., Scheufler, C., Schneider, C. Höhfeld, J., Hartl, F. U., Moarefi, I. (2001) Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsc70 nucleotide exchange factors, *Science*, **291**, 1553-1557.

Soti, C., Pai, C., Papp, B., Csermely, P. (2005) Molecular chaperones as regulatory elements of cellular networks, *Current Opinion in Cell Biology*, **17**, 210-215.

Smith, D. F., (2004) Tetratricopeptide repeat cochaperones in steroid receptor complexes, *Cell Stress and Chaperones*, **9**, 2, 109-121.

Smith, D. F., Whitesell, L., Katsanis, E. (1998) Molecular chaperones: biology and prospects for pharmacological intervention, *Biochemistry*, **42**, 5867-5876.

Suh, W. C., Lu, C. Z., Gross, C. A. (1999) Structural features required for the interaction of the Hsp70 molecular chaperone of DnaK with its cochaperone DnaJ, *Journal of Biological Chemistry*, **274**, 30534-30539.

Szabo, A., Langer, T., Schröder, H., Flanagan, J., Bukau, B., Hartl, F. U. (1994) The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system-DnaK, DnaJ and GrpE, *Proceedings of the National Academy of Science USA*, **91**, 10345-10349.

Szabo, A., Forszun, R., Hartl, F. U., Flanagan, J. (1996) A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrate, *European Molecular Biology Organisation Journal*, **15**, 408-417.

Takayama, S., Bimston, D. N., Matsuzawa, S., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R., Reed, J.C. (1997) BAG-1 modulates the chaperone activity of Hsp70/Hsc70, *European Molecular Biology Organisation Journal*, **16**, 4887-4896.

Tissieres, A., Mitchell, H. K., Tracey, U. (1974) Protein synthesis in the salivary glands of Drosophila melanogaster: relation to chromosome puffs, *Journal of Molecular Biology*, **84**, 389-398.

Tu, Z., He, G., Li, K. X., Chen, M. J., Chang, J., Chen, L., Yao, Q., Liu, D. P., Ye, H., Shi, J., Wu, X. (2005) An improved system for competent cell preparation and high efficiency plasmid transformation using different *Escherichia coli* strains, *Electronic Journal of Biotechnology*, **8**, 1, 114-120.

Venter, J.C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.* (2001) The sequence of the human genome, *Science*, **291**, 1304-1351.

Vijay-Kumar, S., Bugg, C. E., Cook, W. J. (1987) Structure of ubiquitin refined at 1.8 A resolution, *Journal of Molecular Biology*, **194**, 3, 531-544.

Wall, D., Zylicz, M., Georgopoulos, C. (1994) The NH_2-terminal 108 amino acids of the *Escherichia coli* DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication, *Journal of Biological Chemistry*, **269**, 5446-5451.

Wall, D. Zylicz, M., Georgopoulos, C. (1995) The conserved G/F motif of the DnaJ cochaperone is necessary for the activation of the substrate binding properties of the DnaK cochaperones, *Journal of Biological Chemistry*, **270**, 2139-2144.

Watanabe, J. (1997) Cloning and characterization of heat shock protein DnaJ homologues from *Plasmodium falciparum* and comparison with ring infected erythrocyte surface antigen, *Molecular and Biochemical Parasitology*, **88**, 253-258.

Westhoff, B., Chapple, J. P., van der Spuy, J., Höhfeld, J., Cheetham, M. E. (2005) HSJ1 is a neuronal shuttling factor for the sorting of chaperone clients to the proteasome, *Current Biology*, **15**, 1058-1064.

Wilbanks, S. M., McKay, D. B. (1995) How potassium affects the activity of the molecular chaperone Hsc70, *Journal of Biological Chemistry*, **270**, 2251-2257.

Wu, Y., Li, J., Jin, Z., Fu, Z., Sha, B. (2005) The crystal structure of the C-terminal fragment of yeast Hsp40 Hdj1 reveals novel dimerization motif for Hsp40, *Journal of Molecular Biology*, **346**, 1005-1011.

Yanisch-Perron, C., Vieira, J., Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene*, **33**, 1, 103-199.

Young, J. C., Agashe, V. R., Siegers, K., Hartl., F. U. (2004) Pathways of chaperone-mediated protein folding in the cytosol, *Nature Reviews: Molecular Cell Biology*, **5**, 781-791.

Young, J. C., Barral, J. M., Hartl, F. U. (2003) More than folding: localized functions of cytosolic chaperones, *Trends in Biochemical Sciences*, **28**, 10, 541-547.

Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M., Rechsteiner, M. (1998) Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a, *Journal of Biological Chemistry*, **273**, 10, 5461-5467.

Zhu, J. K., Bressan, R. A., Hasegawa, P. M. (1993) Isoprenylation of the plant molecular chaperone ANJ1 facilitates membrane association and function at high temperature, *Proceedings of the National Academy of Science USA*, **90**, 8557-8561.

Zhu, X., Zhao, X., Burkholder, W., Gragerov, A., Ogata, C. M., Gottesman, M. E., Hendrickson, W. A. (1996) Structural analysis of substrate binding by the molecular chaperones DnaK, *Science*, **272**, 1606-1614.

Zimmerman, S. B., Monton, A. P. (1993) Macromolecular Crowding: Biochemical, Biophysical, and Physiological Consequences, *Annual Reviews of Biophysics and Biomolecular Structures*, **22**, 27-65.

Zwanzig, R., Szabo, A., Bagchi, B. (1992) Levinthal's paradox, *Proceedings of the National Academy of Science USA*, **80**, 20-22.

NOTE:
In cases where the authorship of a journal article has been specified as to having two first co-authors, the article has been referenced in text as Co-author1 and Co-author2 et al. *to give due credit to both first authors.*

CUT REFS

Brodsky, J. L. (1996) Post-translational protein translocation: not all hsc70s are created equal, *Trends in Biochemical Sciences*, **21**, 122-126.

Brown, H., Hien, T. T., Day, N., Mai, N. T. H., Chuong, L. V., Chau, T. T. H., Loc, P. P., Phu, N. H., Bethell, D., Farrar, J., Gatter, K., White, N., Turner, G. (1999) Evidence of blood-brain barrier dysfunction in human cerebral malaria, *Neuropathology and Applied Neurobiology*, **25**, 331-340.

Caplan, A. J., Cyr, D. M., Douglas, M. G. (1993) Eukaryotic homologues of *E. coli* DnaJ: A diverse protein family that functions with HSP70 stress proteins, *Molecular Biology of the Cell*, **4**, 555-563.

Fourie, A. M., Sambrook, J. F., Gething, M. J. H. (1994) Common and divergent peptide binding specificities of hsp70 molecular chaperones, *Journal of Biological Chemistry*, **269**, 30470-30478.

Frydman, J. (2001) Folding of newly translated proteins *in vivo*: the role of molecular chaperones, *Annual Review of Biochemistry*, **70**, 603-647.

Georgopoulos, C. (1992) The emergence of chaperone machines, *Trends in Biochemical Sciences*, **17**, 295-299.

Haldar, K., Hiller, N. L., van Ooij, C., Bhattacharjee, S. (2005) *Plasmodium* parasite proteins and the infected erythrocyte, *Trends in Parasitology*, **21**, 9, 402-403.

Morimoto, R. I. (1993) Cells in stress: transcriptional activation of heat shock genes, *Science*, **259**, 1409-1410.

Zhong, M., Orosz, A., Wu, C. (1998) Direct sensing of heat and oxidation by *Drosophila* heat shock transcription factor, *Molecular Cell*, **2**, 101-108.

APPENDIX 1: MATERIALS

Table A1.1: Compounds and suppliers used in this study

Compound name	Supplier	Country of Origin
Acrylamide	Sigma-Aldrich	Germany
Agarose	Hispanagar	Spain
Ammonium molybdate	Merck	Germany
Ammonium persulphate	Saarchem	South Africa
Ammonium sulphate	Saarchem	South Africa
Ascorbic acid	Merck	Germany
Bis-acrylamide	Sigma-Aldrich	Germany
Borate	Saarchem	South Africa
Bromophenol Blue	Sigma-Aldrich	Germany
Bovine Serum Albumin	Roche	Germany
Calcium chloride	Saarchem	South Africa
Chloroform	BDH	England
Coomassie Brilliant Blue G250	Sigma-Aldrich	Germany
Disodium hydrogen phosphate	Saarchem	South Africa
Dithiothreitol	USB	USA
Ethylene diamine tetra-acetic acid	Saarchem	South Africa
Ethanol	BDH	England
Ethidium bromide	Sigma-Aldrich	Germany
Glacial acetic acid	Saarchem	South Africa
Glucose	Saarchem	South Africa
Glycerol	Saarchem	South Africa
Glycine	Sigma-Aldrich	Germany
Hepes	Sigma-Aldrich	Germany
Hi-Di Formamide	Applied Biosystems	USA
Imidazole	Sigma-Aldrich	Germany
Isopropyl-β-D-thiogalactopyranoside	Roche	Germany
Isopropanol	BDH	England
Lambda DNA	Promega	USA
Lysozyme	Roche	Germany
Magnesium chloride	Saarchem	South Africa
Magnesium sulphate	Sigma-Aldrich	Germany
Mercaptoethanol	Merck	Germany
Methanol	Saarchem	South Africa
Nickel sulphate	Sigma-Aldrich	Germany
Phenylmethylsulphonylfluoride	Sigma-Aldrich	Germany
Ponceau S	Sigma-Aldrich	Germany
Potassium acetate	Saarchem	South Africa
Potassium chloride	Saarchem	South Africa
Potassium dihydrogen phosphate	Pro Analysi	South Africa

Sepharose Fast Flow, Chelating	Amersham Biosciences	Sweden
Sodium azide	Sigma-Aldrich	Germany
Sodium chloride	Saarchem	South Africa
Sodium citrate	Saarchem	South Africa
Sodium dodecyl sulphate	BDH	England
Sodium hydroxide	Saarchem	South Africa
N,N,N',N'-tetramethylethylenediamine	Sigma-Aldrich	Germany
Tris	Sigma-Aldrich	Germany
Triton X-100	Sigma-Aldrich	Germany
Tween 20	Saarchem	South Africa

APPENDIX 2: DNA SEQUENCING

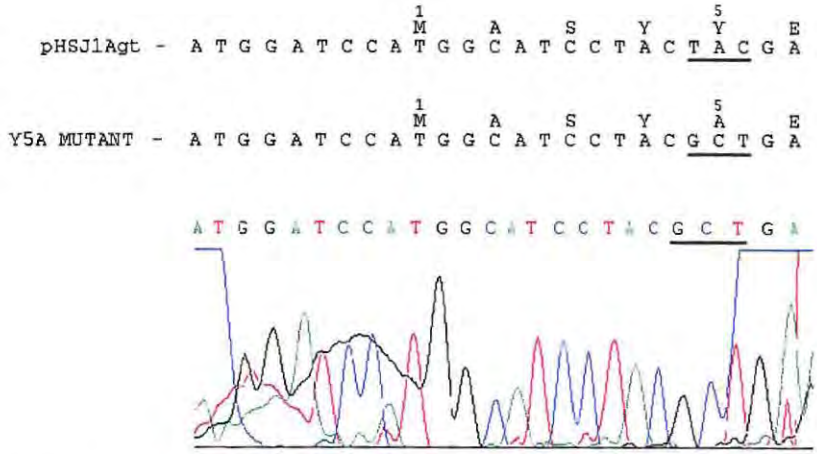


Figure A2.1: pHsj1a*Agt*(Y5A) DNA sequencing results

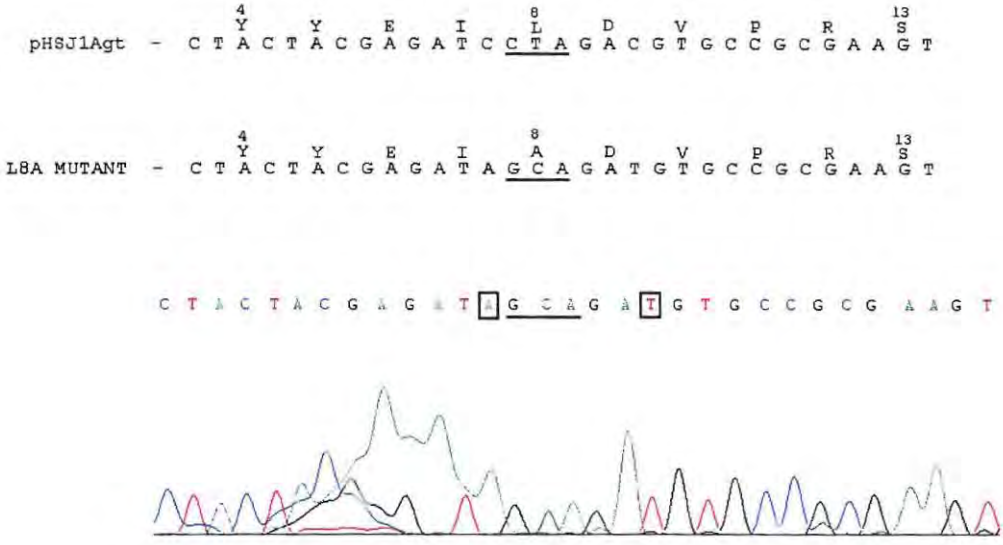


Figure A2.2: pHsj1a*Agt*(L8A) DNA sequencing results

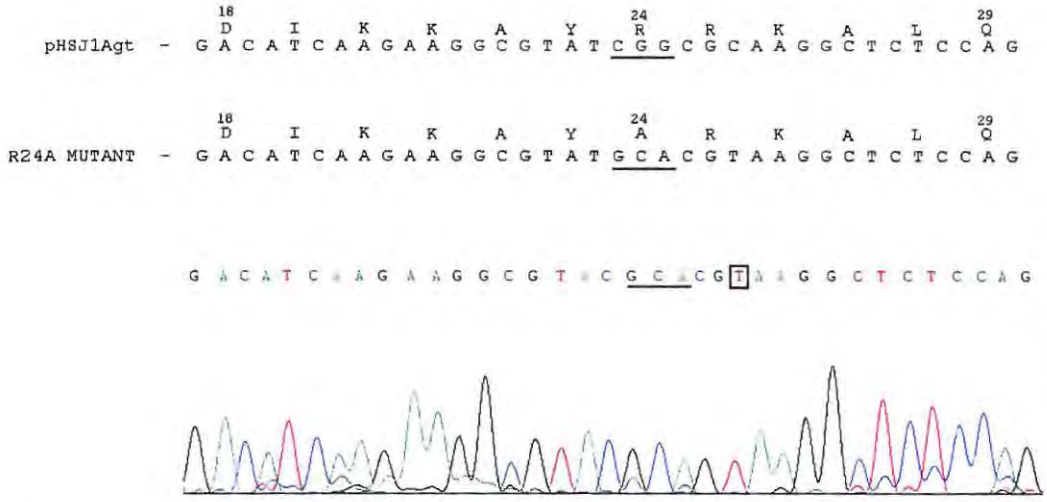


Figure A2.3: pHsj1a*Agt*(R24A) DNA sequencing results

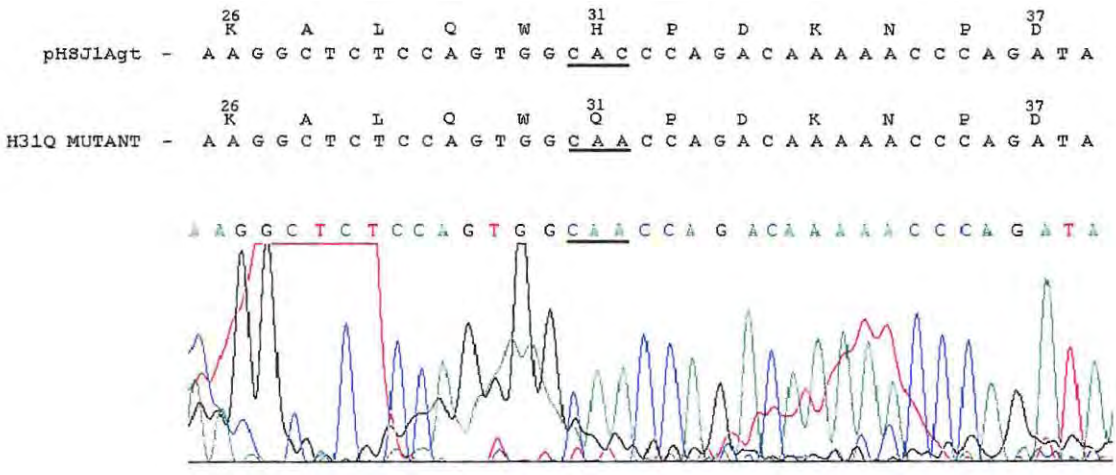


Figure A2.4: pHsj1a*Agt*(H31Q) DNA sequencing results

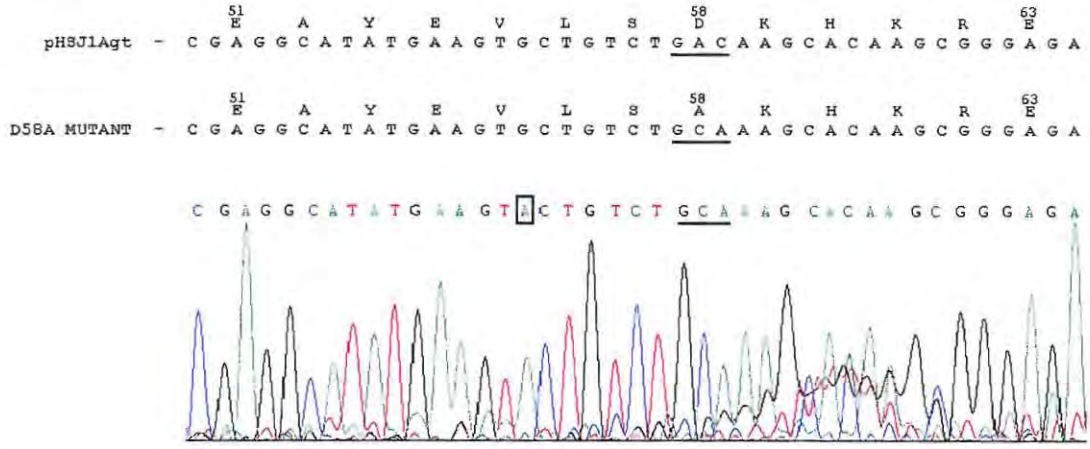


Figure A2.5: pHsj1a*Agt*(D58A) DNA sequencing results

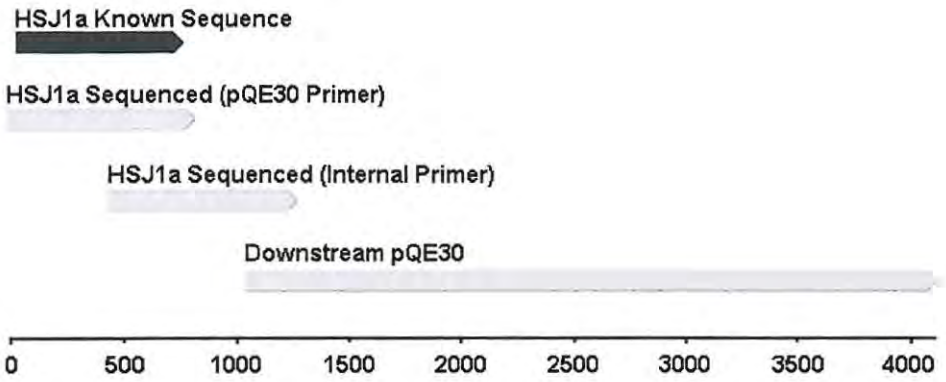


Figure A2.6: pHsj1a plasmid contig overlays for DNA sequencing

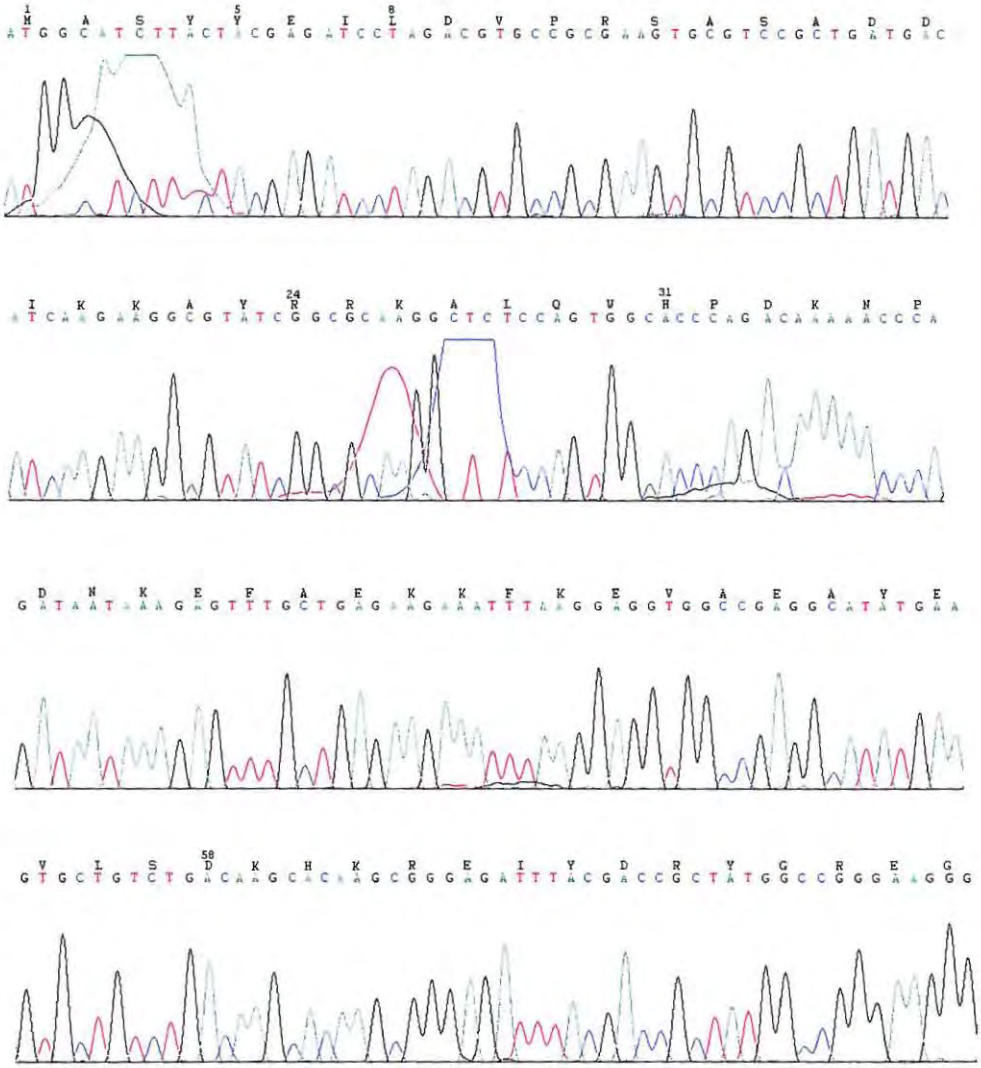
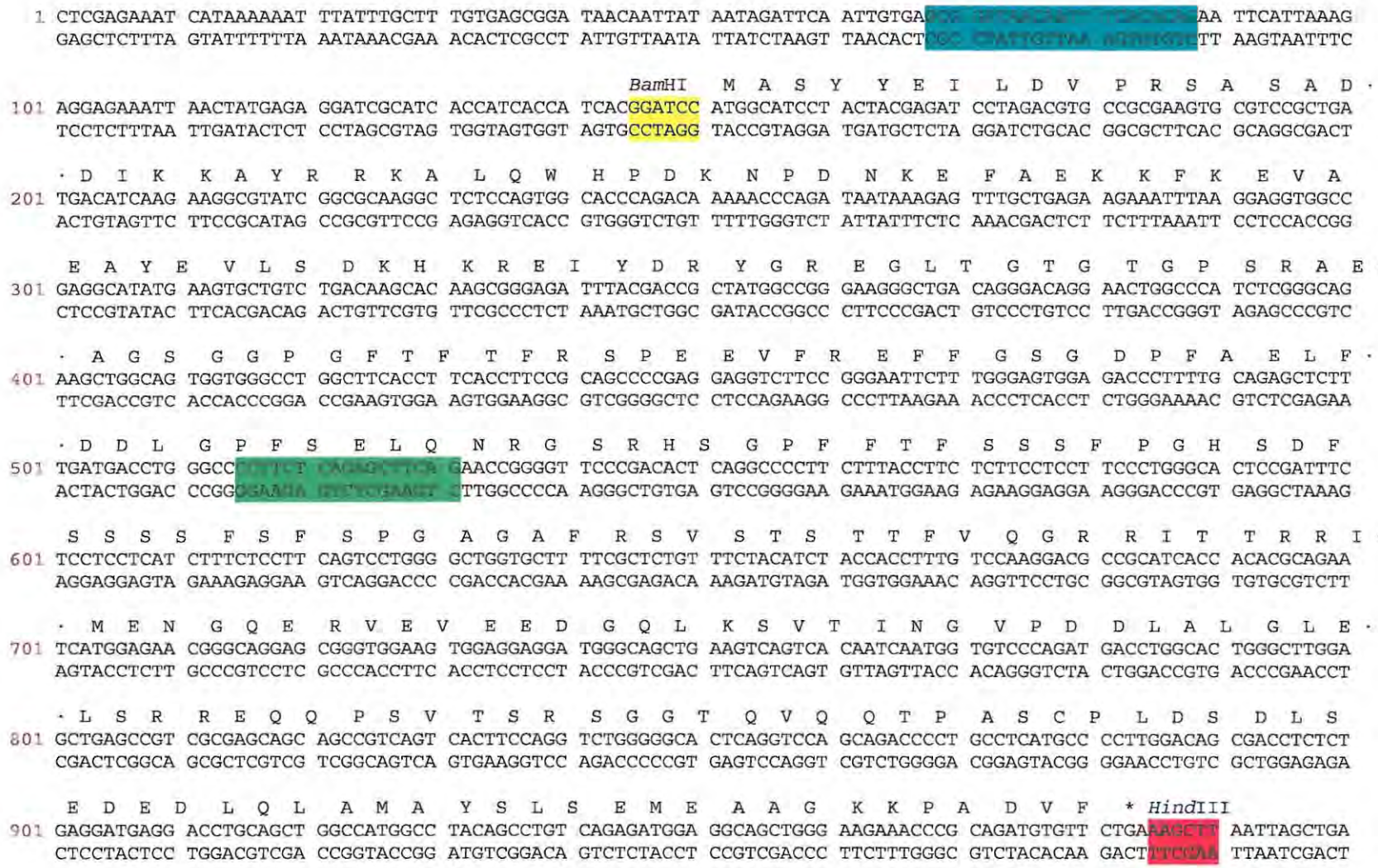


Figure A2.7: DNA sequencing chromatogram results for the J-domain region of pHsj1a


1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA ATTGTGA AA TTCATTAAAG
GAGCTCTTTA GTATTTTTTA AATAAACGAA ACACTCGCCT ATTGTTAATA TTATCTAAGT TAACACT TT AAGTAATTTC
BamHI M A S Y Y E I L D V P R S A S A D
101 AGGAGAAATT AACTATGAGA GGATCGCATC ACCATCACCA TCACGGATCC ATGGCATCCT ACTACGAGAT CCTAGACGTG CCGCGAAGTG CGTCCGCTGA
TCCTCTTTAA TTGATACTCT CCTAGCGTAG TGGTAGTGGT AGTGCCTAGG TACCGTAGGA TGATGCTCTA GGATCTGCAC GGCGCTTCAC GCAGGCGACT
D I K K A Y R R K A L Q W H P D K N P D N K E F A E K K F K E V A
201 TGACATCAAG AAGGCGTATC GGCGCAAGGC TCTCCAGTGG CACCCAGACA AAAACCCAGA TAATAAAGAG TTTGCTGAGA AGAAATTTAA GGAGGTGGCC
ACTGTAGTTC TTCCGCATAG CCGCGTTCCG AGAGGTCACC GTGGGTCTGT TTTTGGGTCT ATTATTTCTC AAACGACTCT TCTTTAAATT CCTCCACCGG
E A Y E V L S D K H K R E I Y D R Y G R E G L T G T G T G P S R A E
301 GAGGCATATG AAGTGCTGTC TGACAAGCAC AAGCGGGAGA TTTACGACCG CTATGGCCGG GAAGGGCTGA CAGGGACAGG AACTGGCCCA TCTCGGGCAG
CTCCGTATAC TTCACGACAG ACTGTTCGTG TTCGCCCTCT AAATGCTGGC GATACCGGCC CTTCCCGACT GTCCCTGTCC TTGACCGGGT AGAGCCCGTC
A G S G G P G F T F T F R S P E E V F R E F F G S G D P F A E L F
401 AAGCTGGCAG TGGTGGGCCT GGCTTCACCT TCACCTTCCG CAGCCCCGAG GAGGTCTTCC GGGAATTCTT TGGGAGTGGA GACCCTTTTG CAGAGCTCTT
TTCGACCGTC ACCACCCGGA CCGAAGTGGA AGTGGAAGGC GTCGGGGCTC CTCCAGAAGG CCCTTAAGAA ACCCTCACCT CTGGGAAAAC GTCTCGAGAA
D D L G P F S E L Q N R G S R H S G P F F T F S S S F P G H S D F
501 TGATGACCTG GGCC TTCT CAGAGCTTCA GAACCGGGGT TCCCGACACT CAGGCCCCTT CTTTACCTTC TCTTCCTCCT TCCCTGGGCA CTCCGATTTC
ACTACTGGAC CCGG GAAGA TTCTCGAAGT CTTGGCCCCA AGGGCTGTGA GTCCGGGGAA GAAATGGAAG AGAAGGAGGA AGGGACCCGT GAGGCTAAAG
S S S S F S F S P G A G A F R S V S T S T T F V Q G R R I T T R R I
601 TCCTCCTCAT CTTTCTCCTT CAGTCCTGGG GCTGGTGCTT TTCGCTCTGT TTCTACATCT ACCACCTTTG TCCAAGGACG CCGCATCACC ACACGCAGAA
AGGAGGAGTA GAAAGAGGAA GTCAGGACCC CGACCACGAA AAGCGAGACA AAGATGTAGA TGGTGGAAAC AGGTTCCTGC GGCGTAGTGG TGTGCGTCTT
M E N G Q E R V E V E E D G Q L K S V T I N G V P D D L A L G L E
701 TCATGGAGAA CGGGCAGGAG CGGGTGGAAG TGGAGGAGGA TGGGCAGCTG AAGTCAGTCA CAATCAATGG TGTCCCAGAT GACCTGGCAC TGGGCTTGGA
AGTACCTCTT GCCCGTCCTC GCCCACCTTC ACCTCCTCCT ACCCGTCGAC TTCAGTCAGT GTTAGTTACC ACAGGGTCTA CTGGACCGTG ACCCGAACCT
L S R R E Q Q P S V T S R S G G T Q V Q Q T P A S C P L D S D L S
801 GCTGAGCCGT CGCGAGCAGC AGCCGTCAGT CACTTCCAGG TCTGGGGGCA CTCAGGTCCA GCAGACCCCT GCCTCATGCC CCTTGGACAG CGACCTCTCT
CGACTCGGCA GCGCTCGTCG TCGGCAGTCA GTGAAGGTCC AGACCCCCGT GAGTCCAGGT CGTCTGGGGA CGGAGTACGG GGAACCTGTC GCTGGAGAGA
E D E D L Q L A M A Y S L S E M E A A G K K P A D V F * HindIII
901 GAGGATGAGG ACCTGCAGCT GGCCATGGCC TACAGCCTGT CAGAGATGGA GGCAGCTGGG AAGAAACCCG CAGATGTGTT CTGAAAGCTT AATTAGCTGA
CTCCTACTCC TGGACGTCGA CCGGTACCGG ATGTCGGACA GTCTCTACCT CCGTCGACCC TTCTTTGGGC GTCTACACAA GACTTTCGAA TTAATCGACT

```
1001 GCTTGGACTC CTGTTGATAG ATCCAGTAAT GACCTCAGAA CTCCATCTGG ATTTGTTCAG AACGCTCGGT TGCCGCCGGG CGTTTTTTAT TGGTGAGAAT
     CGAACCTGAG GACAACTATC TAGGTCATTA CTGGAGTCTT GAGGTAGACC TAAACAAGTC TTGCGAGCCA ACGGCGGCCC GCAAAAAATA ACCACTCTTA

1101 CCAAGCTAGC TTGGCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAAATCAC TGGATATACC ACCGTTGATA TATCCCAATG GCATCGTAAA
     GGTTCGATCG AACCGCTCTA AAAGTCCTCG ATTCCTTCGA TTTTACCTCT TTTTTTAGTG ACCTATATGG TGGCAACTAT ATAGGGTTAC CGTAGCATTT

1201 GAACATTTTG AGGCATTTCA GTCAGTTGCT CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG CCTTTTTAAA GACCGTAAAG AAAAATAAGC
     CTTGTAAAAC TCCGTAAAGT CAGTCAACGA GTTACATGGA TATTGGTCTG GCAAGTCGAC CTATAATGCC GGAAAAATTT CTGGCATTTC TTTTTATTCG

1301 ACAAGTTTTA TCCGGCCTTT ATTCACATTC TTGCCCGCCT GATGAATGCT CATCCGGAAT TTCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA
     TGTTCAAAAT AGGCCGGAAA TAAGTGTAAG AACGGGCGGA CTACTTACGA GTAGGCCTTA AAGCATACCG TTACTTTCTG CCACTCGACC ACTATACCCT

1401 TAGTGTTCAC CCTTGTTACA CCGTTTTCCA TGAGCAAACT GAAACGTTTT CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA
     ATCACAAGTG GGAACAATGT GGCAAAAGGT ACTCGTTTGA CTTTGCAAAA GTAGCGAGAC CTCACTTATG GTGCTGCTAA AGGCCGTCAA AGATGTGTAT

1501 TATTCGCAAG ATGTGGCGTG TTACGGTGAA AACCTGGCCT ATTTCCCTAA AGGGTTTATT GAGAATATGT TTTTCGTCTC AGCCAATCCC TGGGTGAGTT
     ATAAGCGTTC TACACCGCAC AATGCCACTT TTGGACCGGA TAAAGGGATT TCCCAAATAA CTCTTATACA AAAAGCAGAG TCGGTTAGGG ACCCACTCAA

1601 TCACCAGTTT TGATTTAAAC GTGGCCAATA TGGACAACTT CTTCGCCCCC GTTTTCACCA TGGGCAAATA TTATACGCAA GGCGACAAGG TGCTGATGCC
     AGTGGTCAAA ACTAAATTTG CACCGGTTAT ACCTGTTGAA GAAGCGGGGG CAAAAGTGGT ACCCGTTTAT AATATGCGTT CCGCTGTTCC ACGACTACGG

1701 GCTGGCGATT CAGGTTCATC ATGCCGTTTG TGATGGCTTC CATGTCGGCA GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGGCG
     CGACCGCTAA GTCCAAGTAG TACGGCAAAC ACTACCGAAG GTACAGCCGT CTTACGAATT ACTTAATGTT GTCATGACGC TACTCACCGT CCCGCCCCGC

1801 TAATTTTTTT AAGGCAGTTA TTGGTGCCCT TAAACGCCTG GGGTAATGAC TCTCTAGCTT GAGGCATCAA ATAAAACGAA AGGCTCAGTC GAAAGACTGG
     ATTAAAAAAA TTCCGTCAAT AACCACGGGA ATTTGCGGAC CCCATTACTG AGAGATCGAA CTCCGTAGTT TATTTTGCTT TCCGAGTCAG CTTTCTGACC

1901 GCCTTTCGTT TTATCTGTTG TTTGTCGGTG AACGCTCTCC TGAGTAGGAC AAATCCGCCC TCTAGAGCTG CCTCGCGCGT TTCGGTGATG ACGGTGAAAA
     CGGAAAGCAA AATAGACAAC AAACAGCCAC TTGCGAGAGG ACTCATCCTG TTTAGGCGGG AGATCTCGAC GGAGCGCGCA AAGCCACTAC TGCCACTTTT

2001 CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG
     GGAGACTGTG TACGTCGAGG GCCTCTGCCA GTGTCGAACA GACATTCGCC TACGGCCCTC GTCTGTTCGG GCAGTCCCGC GCAGTCGCCC ACAACCGCCC

2101 TGTCGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACT ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT
     ACAGCCCCGC GTCGGTACTG GGTCAGTGCA TCGCTATCGC CTCACATATG ACCGAATTGA TACGCCGTAG TCTCGTCTAA CATGACTCTC ACGTGGTATA

2201 GCGGTGTGAA ATACCGCACA GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG
     CGCCACACTT TATGGCGTGT CTACGCATTC CTCTTTTATG GCGTAGTCCG CGAGAAGGCG AAGGAGCGAG TGACTGAGCG ACGCGAGCCA GCAAGCCGAC
```

```
2301 CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
     GCCGCTCGCC ATAGTCGAGT GAGTTTCCGC CATTATGCCA ATAGGTGTCT TAGTCCCCTA TTGCGTCCTT TCTTGTACAC TCGTTTTCCG GTCGTTTTCC

2401 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
     GGTCCTTGGC ATTTTTCCGG CGCAACGACC GCAAAAAGGT ATCCGAGGCG GGGGGACTGC TCGTAGTGTT TTTAGCTGCG AGTTCAGTCT CCACCGCTTT

2501 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT
     GGGCTGTCCT GATATTTCTA TGGTCCGCAA AGGGGGACCT TCGAGGGAGC ACGCGAGAGG ACAAGGCTGG GACGGCGAAT GGCCTATGGA CAGGCGGAAA

2601 CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC
     GAGGGAAGCC CTTCGCACCG CGAAAGAGTA TCGAGTGCGA CATCCATAGA GTCAAGCCAC ATCCAGCAAG CGAGGTTCGA CCCGACACAC GTGCTTGGGG

2701 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
     GGCAAGTCGG GCTGGCGACG CGGAATAGGC CATTGATAGC AGAACTCAGG TTGGGCCATT CTGTGCTGAA TAGCGGTGAC CGTCGTCGGT GACCATTGTC

2801 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG
     CTAATCGTCT CGCTCCATAC ATCCGCCACG ATGTCTCAAG AACTTCACCA CCGGATTGAT GCCGATGTGA TCTTCCTGTC ATAAACCATA GACGCGAGAC

2901 CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA
     GACTTCGGTC AATGGAAGCC TTTTTCTCAA CCATCGAGAA CTAGGCCGTT TGTTTGGTGG CGACCATCGC CACCAAAAAA ACAAACGTTC GTCGTCTAAT

3001 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG
     GCGCGTCTTT TTTTCCTAGA GTTCTTCTAG GAAACTAGAA AAGATGCCCC AGACTGCGAG TCACCTTGCT TTTGAGTGCA ATTCCCTAAA ACCAGTACTC

3101 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA
     TAATAGTTTT TCCTAGAAGT GGATCTAGGA AAATTTAATT TTTACTTCAA AATTTAGTTA GATTTCATAT ATACTCATTT GAACCAGACT GTCAATGGTT

3201 TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
     ACGAATTAGT CACTCCGTGG ATAGAGTCGC TAGACAGATA AAGCAAGTAG GTATCAACGG ACTGAGGGGC AGCACATCTA TTGATGCTAT GCCCTCCCGA

3301 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
     ATGGTAGACC GGGGTCACGA CGTTACTATG GCGCTCTGGG TGCGAGTGGC CGAGGTCTAA ATAGTCGTTA TTTGGTCGGT CGGCCTTCCC GGCTCGCGTC

3401 AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT
     TTCACCAGGA CGTTGAAATA GGCGGAGGTA GGTCAGATAA TTAACAACGG CCCTTCGATC TCATTCATCA AGCGGTCAAT TATCAAACGC GTTGCAACAA

3501 GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
     CGGTAACGAT GTCCGTAGCA CCACAGTGCG AGCAGCAAAC CATACCGAAG TAAGTCGAGG CCAAGGGTTG CTAGTTCCGC TCAATGTACT AGGGGGTACA
```

```
3601 TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC
     ACACGTTTTT TCGCCAATCG AGGAAGCCAG GAGGCTAGCA ACAGTCTTCA TTCAACCGGC GTCACAATAG TGAGTACCAA TACCGTCGTG ACGTATTAAG

3701 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC
     AGAATGACAG TACGGTAGGC ATTCTACGAA AAGACACTGA CCACTCATGA GTTGGTTCAG TAAGACTCTT ATCACATACG CCGCTGGCTC AACGAGAACG

3801 CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC
     GGCCGCAGTT ATGCCCTATT ATGGCGCGGT GTATCGTCTT GAAATTTTCA CGAGTAGTAA CCTTTTGCAA GAAGCCCCGC TTTTGAGAGT TCCTAGAATG

3901 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
     GCGACAACTC TAGGTCAAGC TACATTGGGT GAGCACGTGG GTTGACTAGA AGTCGTAGAA AATGAAAGTG GTCGCAAAGA CCCACTCGTT TTTGTCCTTC

4001 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT
     CGTTTTACGG CGTTTTTTCC CTTATTCCCG CTGTGCCTTT ACAACTTATG AGTATGAGAA GGAAAAAGTT ATAATAACTT CGTAAATAGT CCCAATAACA

4101 CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA
     GAGTACTCGC CTATGTATAA ACTTACATAA ATCTTTTTAT TTGTTTATCC CCAAGGCGCG TGTAAAGGGG CTTTTCACGG TGGACTGCAG ATTCTTTGGT

4201 TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT CGTCTTCAC
     AATAATAGTA CTGTAATTGG ATATTTTTAT CCGCATAGTG CTCCGGGAAA GCAGAAGTG
```

Figure A2.8: DNA sequence of pHsj1a plasmid, showing the amino acid sequence of Hsj1a overlayed

The DNA sequencing primers are highlighted with pQE30F shown in turquoise and pQE30R shown in purple. The Hsj1a internal coding region DNA sequencing primer is shown in green. The *Bam*HI restriction site is shown in yellow and the *Hind*III restriction site is shown in red.

APPENDIX 3: MOLECULAR MASS MARKERS

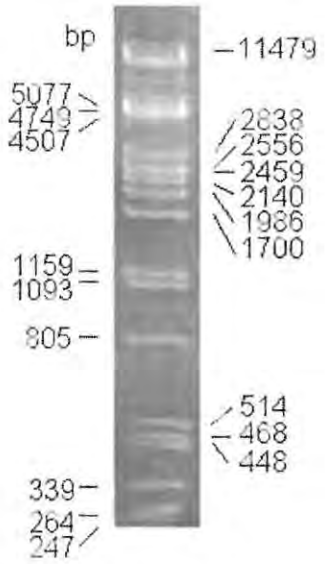


Figure A3.1: DNA Fragment Sizes obtained for Lamda DNA restricted with *Pst* I

Table A3.1: Mass of proteins making up the BioRad SDS Broad Range Marker[1]

Protein name	Size (kDa)
Rabbit skeletal muscle myosin	200
E. coli β-Galactosidase	116.25
Rabbit muscle Phosphorylase b	97.4
Bovine serum albumin (BSA)	66.2
Hen egg white ovalbumin	45
Bovine carbonic anhydrase	31
Soybean trypsin inhibitor	21.5
Hen egg white lysozyme	14.4
Bovine pancreatic trypsin inhibitor (Aprotinin)	6.5

[1] BioRad Catalogue No. 161-0317

APPENDIX 4: PRIMERS

Table A4.1: PCR amplification primers[1]

Name	Sequence	Tm[2]	% GC[3]
Hsj1aF	5′-*GGATCC*ATGGCATCTTACTAC-3′	60.6	47.6
Hsj1aR	3′-*AAGCTT*TCAGAACACATCTGC-5′	61.3	42.9

[1] Restriction sites shown in italics, silent mutations are underlined. [2] Tm and [3] %GC values were all calculated using GeneRunner.

Table A4.2: DNA sequencing primers

Name	Sequence	Tm[1]	% GC[2]
M13F	5′-GGTTTTCCCAGTCACGAC-3′	58.4	55.6
M13R	5′-GGAAACAGCTATGACCATG-3′	56.7	47.4
pQE30F	5′-GCGGATAACAATTTCACACAG-3′	62.6	42.9
pQE30R	5′-AGTTCTGAGGTCATTACTGG-3′	55.0	45.0
Hsj1IntF	5′-CCTTCTCAGAGCTTCAG-3′	49.5	52.9

[1] Tm and [2] %GC values were all calculated using GeneRunner.

Table A4.3: Site-directed mutagenesis primers[1]

Name	Sequence[2]	Tm[3]	% GC[4]
Hsj1(Y5A)F	5′-GGCATCCTAC**GCT**GA*GATTC*TAGACGTGC-3′	80.3	55.2
Hsj1(Y5A)R	5′-GCACGTCTA*GAATC*TC**AGC**GTAGGATGCC-3′	80.3	55.2
Hsj1(L8A)F	5′-CTACTACGAGATA***GCA***GATGTGCCGCGAAG-3′	80.5	53.3
Hsj1(L8A)R	5′-CTTCGCGGCACATC***TGC***TATCTCGTAGTAG-3′	80.5	53.3
Hsj1(R24A)F	5′-AAGAAGG*CGTAC***GCA**CGTAAGGCTCTCC-3′	81.6	57.1
Hsj1(R24A)R	5′-GGAGAGCCTTACG**TGC***GTACG*CCTTCTT-3′	81.6	57.1
Hsj1(H31Q)F	5′-GCTCTCCAGT*GG***CAA**CCAGACAAAAACCC-3′	83.3	55.2
Hsj1(H31Q)R	5′-GGGTTTTTGTCTGG***TTG***CCACTGGAGAGC-3′	83.3	55.2
Hsj1(D58A)F	5′-CATATGA*AGTACT*GTCT**GCA**AAGCACAAGC-3′	76.1	43.3
Hsj1(D58A)R	5′-GCTTGTGCTT**TGC**AGAC*AGTACT*TCATATG-3′	76.1	43.3

[1] Introduction or removal of restriction endonuclease recognition site are shown in grey and italics, silent mutations are shown underlined and mutated residue codons are shown in bold. [2] Restriction endonucleases are *Tfi* I, *Bfa* I, *Bsi* WI, *Ban* I, and *Sca* I respectively for each of the mutants shown as listed from top to bottom. [3] Tm and [4] %GC values were all calculated using GeneRunner.

APPENDIX 5: STANDARD CURVES

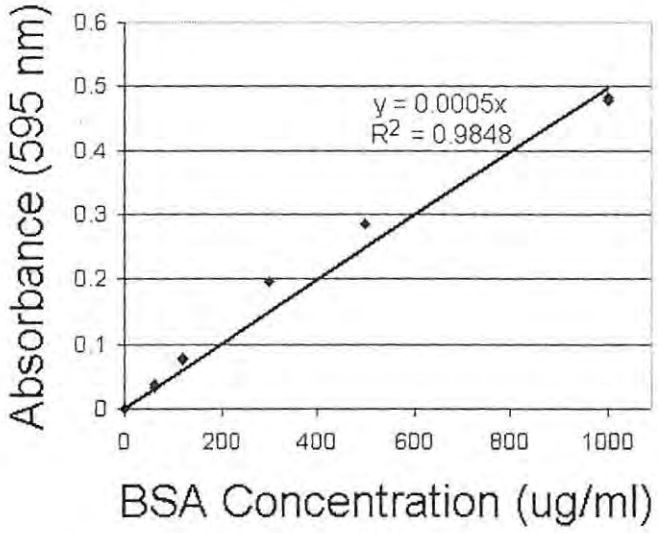


Figure A5.1: Bovine Serum Albumin (BSA) standard curve

The BSA standard curve was used to calculate the protein concentration of various protein samples for set up of ATPase activity assays of human Hsp70.

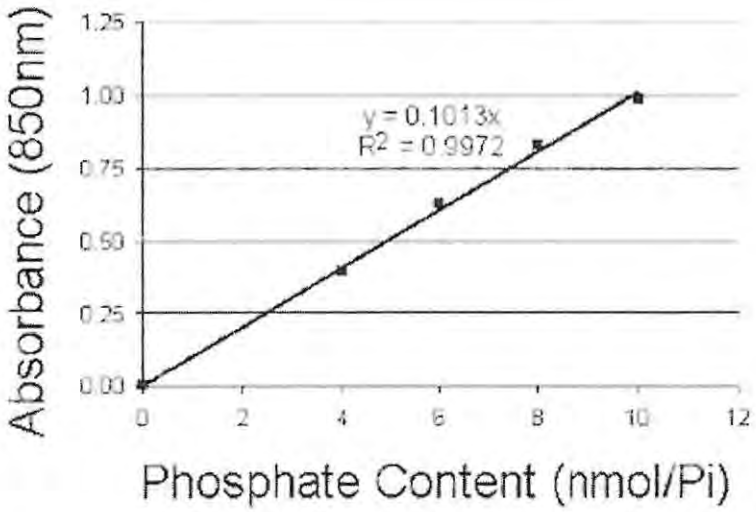


Figure A5.2: Phosphate standard curve

The phosphate standard curve was used to calculate the inorganic phosphate content for ATPase activity assays of human Hsp70.

APPENDIX 6: VECTORS

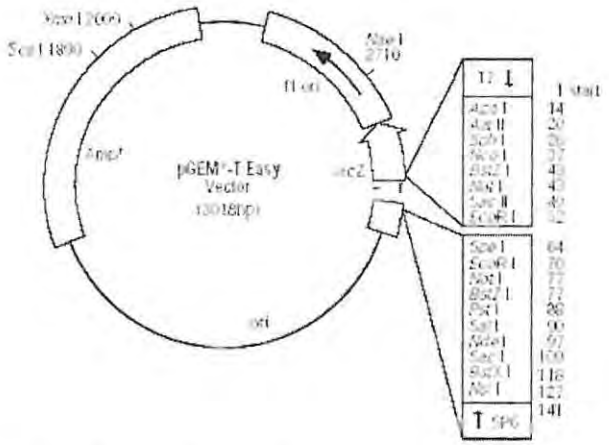


Figure A6.1: pGEM-T Easy vector system

Available from Promega
http://www.tcd.ie/Genetics/staff/Noel.Murphy/recombinant%20dna%20ge4021/pgem.pdf
Accessed 18/01/2006.

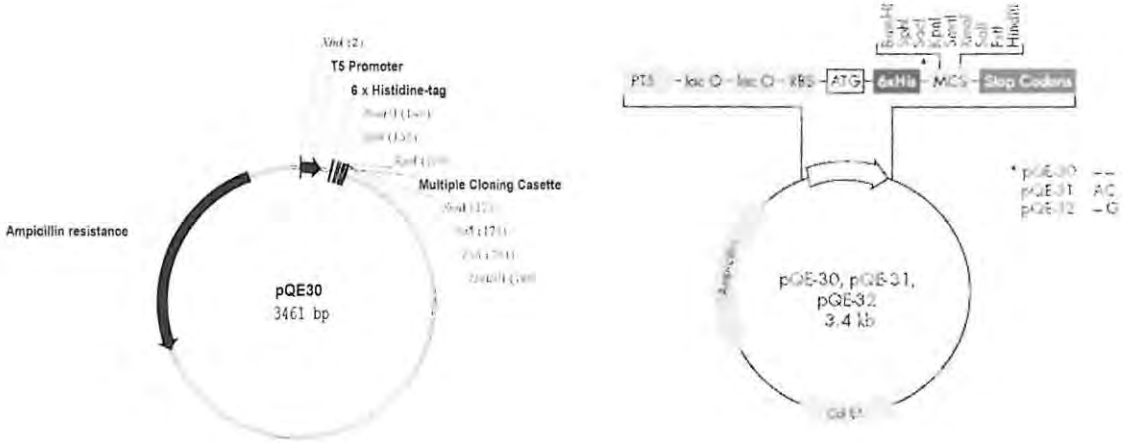


Figure A6.2: pQE30 vector system

Available from Qiagen; http://www1.qiagen.com/literature/pqesequences/pqe3x.pdf;
Accessed 18/01/2005.

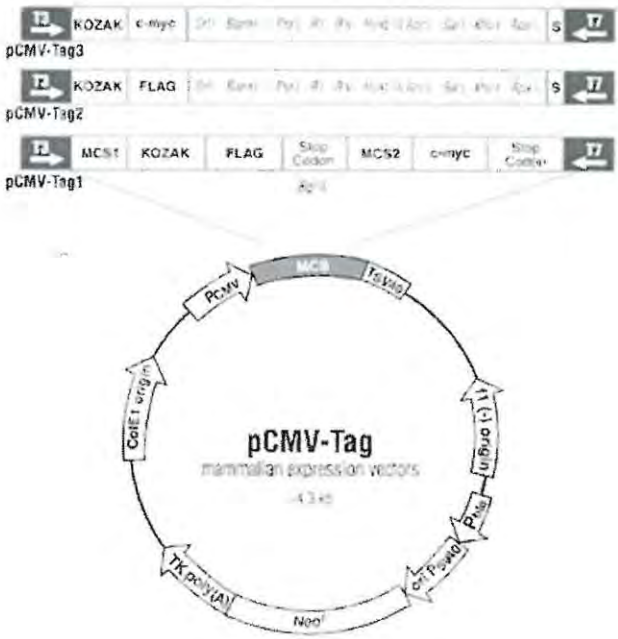


Figure A6.3: pCMV-Tag vector system, showing Tag-3a

Available from Stratagene; http://www.stratagene.com/products/displayProduct.aspx?pid=242; Accessed 18/01/2006.