OVER-EXPRESSION, PURIFICATION AND BIOCHEMICAL CHARACTERISATION OF TRYPANOSOMAL HEAT SHOCK PROTEINS

Submitted in Fulfilment of the Requirements for the Degree Master of Science in Biochemistry by Research

by

Adrienne Lesley Edkins
Rhodes University, Grahamstown
August 2002

ABSTRACT

The molecular chaperone process of assisted protein folding, characteristic of members of the Heat Shock Protein 70 kDa (Hsp70) and Heat Shock Protein 40kDa (Hsp40) families, is essential for cytoprotection in stressful cellular conditions. Examples of such conditions are heat shock or invasion by pathogens. The Hsp70/Hsp40 process of assisted protein folding is dependent on ATP (governed by the intrinsic ATPase activity of Hsp70) and the ability of molecular chaperones to recognise and bind non-native protein conformations. Here, we analyse and attempt to characterise the molecular chaperone activity of an inducible, cytoplasmic Hsp70 (TcHsp70) from *Trypanosoma cruzi* and its interactions with its potential partner Hsp40s, Tcj1, Tcj2, Tcj3 and Tcj4. A bioinformatic analyses of the primary sequences of the trypanosomal proteins revealed that they all contained the canonical domains that define other members of the Hsp70 and Hsp40 family. Tcj2 and Tcj4 showed deviations from the consensus sequence in their substrate binding regions, which may have implications for their substrate binding specificities. TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 were over-expressed recombinantly as 6xHis-tag fusion proteins in *Escherichia coli*. His-TcHsp70, Tcj1-His and His-Tcj2 were successfully purified by Nickel-affinity chromatography for functional analyses to assess the molecular chaperone activity of His-TcHsp70 in terms of its ATPase activity and substrate binding ability. The basal ATPase activity of His-TcHsp70 was determined as 40 nmol Pi/min/mg, significantly higher than that reported for other Hsp70s. This basal ATPase activity was stimulated to a maximal level of 60 nmol Pi/min/mg in the presence of His-Tcj2 and a model non-native substrate, reduced carboxymethylated α-lactalbumin (RCMLA). Using native polyacrylamide gel electrophoresis and Western analysis, His-TcHsp70 was shown to form discrete complexes when in the presence of Tcj1-His, His-Tcj2 and/or RCMLA. These complexes potentially represent His-TcHsp70 – RCMLA or His-TcHsp70 – Tcj interactions, that may be indicative of chaperone activity. *In vivo* complementation assays showed that Tcj2, but not Tcj3, was able to overcome the temperature sensitivity of the *ydj1* mutant *Saccharomyces cerevisiae* strain JJ160, suggesting that Tcj2 may be functionally equivalent to the yeast Hsp40 Ydj1.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the following people for their support and assistance during this study:

- My supervisor, Professor Greg Blatch, for his unfailing dedication, support and advice.
- Rhodes University and the National Research Foundation for their generous funding.
- Mr R. Cross and the members of the Electron Microscopy Unit at Rhodes University
- Ms P. Gentz, Ms F. Hennessy and Mr A. Venter for their scientific contributions to this work
- Mr G.M. Jevons, Ms P. Gentz, Ms F. Hennessy and Dr. G. Jones for their support, friendship and encouragement.
- All the members of the Chaperone Research Group at Rhodes University for their continued support and friendship

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

µg	Microgram(s)
µM	Micromolar (µmoles/litre)
µL	Microlitre(s)
µmol	Micromole(s)
A	Absorbance
ATP	Adenosine Triphosphate
ADP	Adenosine Diphosphate
ATPase	Adenosine triphosphatase
B. taurus	*Bos taurus* (bovine / cow)
BSA	Bovine Serum Albumin
E. coli	*Escherichia coli*
H.sapiens	*Homo sapiens* (human)
BLAST	Basic Local Alignment Search Tool
°C	Degrees Centigrade
cDNA	Complimentary DNA
DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
Csp	Cysteine String Protein
bp	Base Pairs
kDa	Kilodaltons
EDTA	Ethylene Diaminetetra-acetic Acid
g	Gravitational Force
M	Molar (moles / Litre)
IPTG	Isopropyl β-D Thiogalactoside
Hsp	Heat Shock Protein
Hsp70	Heat Shock Protein 70kDa
Hsp40	Heat Shock Protein 40kDa
Hsp90	Heat Shock Protein 90kDa
Hop	Hsp70/Hsp90 Organising Protein (generic term)
mSTI1	*M.musculus* Stress Inducible Protein 1 (mouse equivalent to Hop)
TcSTI1	*T.cruzi* Stress Inducible Protein 1 (mouse equivalent to Hop)
M.musculus	*Mus musculus* (mouse)
L.braziliensis	*Leishmania braziliensis* (parasite)
DnaJ	*E.coli* Hsp40
DnaK	*E.coli* Hsp70
Tcj	*T.cruzi* Hsp40
TcHsp70	*T.cruzi* Hsp70
T.cruzi	*Trypanosoma cruzi*
T.brucei	*Trypanosoma brucei*
6xHis-tag	Hexahistidine purification tag

mg	Milligram
G	Gram
mL	Millilitre
His-TcHsp70	*T.cruzi* Hsp70 with an N terminal hexahistidine tag
Tcj1-His	*T.cruzi* Tcj1 with an C terminal hexahistidine tag
His-Tcj2	*T.cruzi* Tcj2 with an N terminal hexahistidine tag
Tcj3-His	*T.cruzi* Tcj3 with an C terminal hexahistidine tag
His-Tcj4	*T.cruzi* Tcj4 with an N terminal hexahistidine tag
IgG	Immunoglobulin G
IgG-POD	Peroxidase-labelled IgG
JJ160	*S.cerevisiae ydj1* mutant strain
HPD motif	Histidine – Proline – Aspartic Acid tripeptide
G/F region	Glycine /Phenylalanine Rich Region
N-terminus	N- terminal region of a protein
C-terminus	C- terminal region of a protein
SDS	Sodium Dodecylsulphate
SDS-PAGE	Sodium Dodecylsulphate Polacrylamide Gel Electrophoresis
PAGE	Polacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
Tris	Tris-2-amino-2-(hydroxymethyl)-1,3,propandiol
TBS	Tris Buffered Saline
TBST	Tris Buffer Saline containing 0.1% Tween-20
RPM	Revolutions per minute
S.cerevisiae	*Saccharomyces cerevisiae* (baker's yeast)
2xYT broth	Yeast-tryptone broth
2xYT agar	Yeast-tryptone agar
V	Volts
V/V	Volume to Volume Ratio
W/V	Weight to Volume Ratio
UV	Ultraviolet
WY26	*S.cerevisiae sis1* mutant strain
YMM/A glucose	Yeast Minimal Medium/Agar, carbon source is glucose
YMM/A galactose	Yeast Minimal Medium/ Agar, carbon source is galactose
YMM/A (HIS^-)	Yeast Minimal Medium/Agar lacking Histidine
YMM/A (LEU^-)	Yeast Minimal Medium/ Agar lacking Leucine
YMM/A (HIS^- URA^-)	Yeast Minimal Medium/Agar lacking Histidine and Uracil
YMM/A (LEU^-TRP^-)	Yeast Minimal Medium/ Agar lacking Leucine and Tryptophan

CHAPTER 1

LITERATURE REVIEW

1.1 INTRODUCTION

1.1.1 Protein Folding and the Cellular Environment

The biological functioning of proteins require that they assume a defined, biophysical structure. The information defining this structure is encoded in the primary amino acid sequence. Anfinsen (1973) demonstrated this in his renowned paper, where he reported the *in vitro* refolding of denatured ribonuclease. Ribonuclease is a small protein and it took 20 minutes to assume its correct structure, far longer than would be feasible *in vivo*. One can imagine then, how long it would take for larger, oligomeric proteins to attain their correct tertiary structure under similar conditions. In addition, the crowded cellular environment increases the chance of unproductive side reactions that could lead to erroneous conformations. How then do proteins attain their correct conformations in a time frame that ensures the process will be viable in the cell? The answer may lie with a group of highly conserved proteins themselves, whose specific function is to aid the folding and assembly of other proteins in the cell.

1.1.2 Heat Shock Proteins (Hsps) and the Molecular Chaperone Concept

Heat Shock Protein is the collective name given to a group of ubiquitous and highly conserved proteins, having essential roles in physiological and stressful cellular environments (Whitley et al., 1999). As their name suggests, the expression of these proteins was first observed in *Drosophila* salivary glands that had been exposed to a heat shock of 37°C (Ritossa, 1962). However, subsequent research has revealed that the expression of these Hsps is induced by a number of stressful conditions, including heat, amino acid analogues, free radicals and heavy metals (Hendrick and Hartl, 1993). The Hsps therefore could be more correctly termed 'stress proteins' due to their fundamental role within the stressed cell. In addition, constitutively expressed homologues of Hsps (termed Hscs), with roles in basic physiological and house-keeping processes in the cell have been identified. While the constitutive Hscs are expressed continuously, the expression of Hsps is controlled by the binding of heat shock transcription factors (HSFs) to specific heat shock elements (HSEs) in the gene. Stress induces the phosphorylation of the HSFs, leading to the formation of trimeric HSF species that bind the HSEs and stimulate expression of the Hsps (Kiang and Tsokos, 1998).

Hsps have been identified in almost all organisms, from bacteria to lower and higher eukaryotes. They are classified and named on the basis of their molecular size, for example Hsp70 will have a molecular weight of approximately 70kDa. To date a number of Hsp families have been described. These are the Hsp100, Hsp90, Hsp70, Hsp60, Hsp40 and the small Hsp (sHsps; ~18 – 43 kDa) Families (Frydman, 2001; Fink, 1999). The major Hsp families and their roles in the cell are summarised in Table 1.1.

Table 1.1: Overview of the members of the major Hsp families and their predominant role in the functioning of the cell.

Family	Member	Location	Size (kDa)	Functions
Hsp100				
Prokaryote	Clp	Cytosol	100	Disaggregation of proteins and thermotolerance
Higher Eukaryote	Hsp104	Cytosol	104	
Hsp90				
Prokaryote	HtpG	Cytosol	90	Protein kinase and transcription factor control; prevention of aggregation
Higher Eukaryote	Hsp90	Cytosol	90	
	Grp94/Grp96	ER	94/96	
Hsp70				
Prokaryote	DnaK	Cytosol	70	Chaperone mediated protein folding of nascent/ misfolded proteins; targeting of proteins for degradation; antigen presentation; thermotolerance
Yeast	Ssa	Cytosol	70	
	Ssb	Cytosol/ Mitochondria	70	
	Ssc	Cytosolic	70	
	Kar2	ER		
Higher Eukaryote	Hsc70	Cytosol/Nucleus	70	
	Hsp70	Cytosol/Nucleus	70	
	Grp78 (BiP)	ER	70	
	p75	Mitochondria	78	
			75	
Hsp60				
Prokaryote	GroEL	Cytosol	60	Protein secretion; mitochondrial protein import; stabilization of prefolded proteins towards folding; export of precursors to intermembrane space.
Yeast	Mif4	Mitochondria	60	
Higher Eukaryote	Hsp60	Mitochondria	60	
Hsp40				
Prokaryote	DnaJ	Cytosol	40	Co-chaperones for DnaK/Hsp70; binding of misfolded/nascent proteins/polypeptides.
Higher Eukaryote	Hsp40	Cytosol	40	

Data from Fink (1999); Hendrick and Hartl (1993)

Each of these Hsps has a defined structure that is conserved throughout members of that particular family from different organisms. It is this structure that defines the individual functions of the different families of Hsps. In addition to the different families of Hsps, subsets of Hsps within the families have been detected in different compartments of the cell, such as the nucleus, endoplasmic reticulum, mitochondria and cytosol. Therefore, although Hsps may be confined to a family group, the different members of that family will be specifically tailored to the roles they are required to perform within their particular organelle (Frydman, 2001; Fewell et al., 2001).

The ability of some, but not all, Hsps to act as molecular chaperones is integral to their protein folding and protective roles in the cell. A molecular chaperone is a protein that is capable of interacting with and stabilizing non-native protein structures or nascent polypeptides, to prevent aggregation and promote the formation of correct, functional conformations (Hendrick and Hartl, 1993; Fink, 1999). They are essential for the refolding of stress-denatured proteins, assembly of oligomeric proteins and *de novo* folding of nascent polypeptides. Molecular chaperones are catalysts in the folding process, in the sense that they do not actively fold proteins, but rather assist folding by preventing non-productive or aggregation reactions and do not constitute part of the final protein structure. In addition, they control the translocation of proteins to different subcellular compartments as well as the degradation of protein aggregates that would otherwise disrupt the cellular environment. The ability of Hsps to function as molecular chaperones is central to their cytoprotective role in stressful cellular conditions, where protein aggregation is often the result (Whitley et al., 1999).

Hsps rarely function singularly as molecular chaperones. Interactions between different Hsp classes occur to modulate distinct chaperone functions. Thus, the molecular chaperone activity of a particular Hsp is controlled, and able to fulfil a number of tasks, through different interactions. Here, the role of Hsp70s as molecular chaperones and their interactions with the Hsp40 and Hsp90 families in that regard will be focused on.

1.1.3 Heat Shock Protein 70 (Hsp70)

The Hsp70 molecular chaperone family is composed of members that are essential for correct protein folding, transport of proteins to different sub-cellular compartments and degradation of unstable proteins in the cell (Kiang and Tsokos, 1998). Hsp70 has been shown to discriminate with respect to substrate, binding only denatured or non-native proteins (Fink, 1999). Hsp70s have been identified as having both constitutive (Hsc70) and heat-inducible members (Hsp70), and have been found in the cytosol (Hsp70/DnaK/Ssa), ER (BiP/Grp78) and mitochondria (p75; Ssc) (Table 1.1). The constitutive form of Hsp70 (Hsc70) is essential in physiological functions involved in metabolism, while the heat inducible form is required for assisted protein folding in times of stress (Mayer and Bukau, 1998).

Hsp70, and its constitutively expressed homologue Hsc70 (heat shock cognate protein) possess a quaternary structure composed of two defined domains of different function that work in collaboration. Hsp70 is composed of a conserved N-terminal ATPase domain of ~44kDa, an ~18 kDa substrate binding region and an ~10kDa variable C-terminal region of largely undefined function (James et al., 1997). Recent studies have targeted the C terminal regions of Hsp70 as being involved in specific interactions with other proteins, such as Hop (Hsp70/Hsp90 Organising Protein). The basic structure of Hsp70 is conserved throughout Hsp70s from different organisms and is specifically suited to chaperone activity (Bork et al., 1992).

The ATPase domain of Hsp70 (Figure 1.1) forms two distinct, equally sized structural lobes, separated from each other by a deep cleft, which is purported to be the binding site of ATP (Flaherty et al., 1990). These two domains display a similar folding pattern to each other, both consisting of two subdomains and possessing phosphate binding sites. The overall structure of the ATPase domain of Hsp70s displays a striking similarity to that of actin, despite the disparate functions of these two proteins (Bork et al., 1992). ATP hydrolysis by the ATPase domain of Hsp70s is central to the role of the protein as a chaperone during assisted protein folding (Wawrzynow et al., 1995). Work by various groups has indicated that certain highly conserved residues within the ATPase domain, namely lysine 71, threonine 199 and aspartic acid 10, are essential for ATP hydrolysis (Huang et al., 1993; McCarty and Walker, 1991; O'Brien et al., 1996).

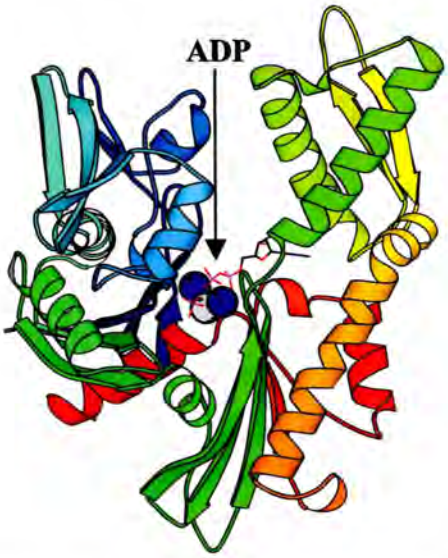


Figure 1.1: Ribbon representation of the structure of the ATPase domain of Bovine Hsp70 (3HSC) visualised using Molscript (Kraulis, 1991).

The two similar subdomains of the ATPase domain of bovine Hsp70, form the cleft that is the binding site of ATP/ADP (Flaherty et al., 1990).

The substrate-binding region of Hsp70 (Figure 1.2) is composed of a compact ß-sandwich at the N-terminal end, formed by two ß–sheets, each comprising 4 antiparallel ß- strands, followed by an extended region of 5 α-helices (Zhu et al., 1996). The ß-sandwich is described as the substrate-binding region, with the α-helical structure involved in stabilization of the complex. The substrate is bound to Hsp70 in an extended conformation, via a hydrophobic channel defined by the ß-sandwich region (Rudiger et al., 1997). The α-helical domain is capable of rotation supporting the suggestion that substrate binding by Hsp70 is facilitated by conformational changes in the chaperones (Buchberger et al., 1995; Zhu et al., 1996). The current suggestion for the recognition of the substrate by the chaperone is due to exposed hydrophobic regions, in non-native forms of the protein substrate. This is consistent with the role of Hsps in the refolding of misfolded proteins, as these hydrophobic residues, normally in the core of the protein, would become exposed in the event of incorrect folding (Richarme and Kohiyama, 1993).

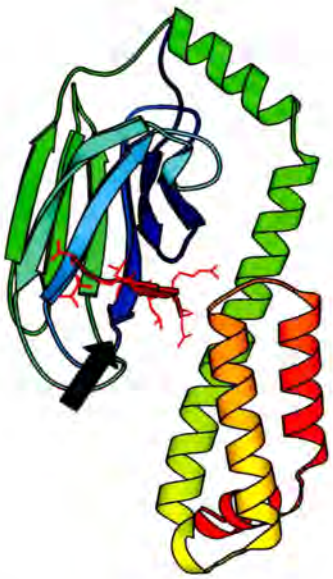


Figure 1.2: Ribbon representation of the Substrate binding region of DnaK (1DKX) visualised using Molscript (Kraulis,1991).

The compact β-sandwich at the N-terminal region of the substrate binding domain is composed of four antiparallel β-sheets proposed to be involved in binding of the substrate, followed by the extended region of five alpha helices thought to act to stabilize the complex. The peptide substrate in the complex is shown as sticks in red and indicated by the arrow (Zhu et al., 1996).

Although the two functional domains of Hsp70 are distinct, they are known to cooperate during assisted protein folding. Besides being essential for the chaperone activity of Hsp70, the ATPase domain and substrate binding regions of Hsp70 have been implicated in oligomerisation of Hsp70 (Blond-Elguindi et al., 1993). Independent studies have suggested that Hsp70 is capable of self-association, potentially through its substrate binding region. The presence of a competing substrate results in the formation of active monomers that are capable of assisting the refolding of these non-native substrates (Benaroudji et al., 1996). This oligomerisation may be a way of bringing the substrate into close proximity with a high concentration of Hsp70s, allowing the rapid formation of Hsp70-substrate complexes facilitating chaperone-assisted protein folding (King et al., 1999).

1.1.4 Heat Shock Protein 40 (Hsp40)

One of the major partners of the Hsp70 molecular chaperones comes from the Hsp40 family of proteins. Certain Hsp40s act as co-chaperones in the Hsp70-assisted protein folding process, via regulation of the ATPase activity of Hsp70. The Hsp40 family of chaperones is commonly referred to as the J proteins, as they all possess a functional motif known as the J domain, in conjunction with other domains that define their function. Hsp40 proteins can be classified as Type I, II or III, depending on the degree of conservation of the functional motifs they contain, with respect to *Escherischia coli* (*E.coli*) DnaJ (Figure 1.3) (Cheetham and Caplan, 1998). The type and number of motifs contained in an Hsp40 will govern its function, with most of the Hsp70 chaperone interactions being modulated by Type I Hsp40s (Figure 1.3 [A]). More specific functions are likely to be mediated by Type III Hsp40s (Henessey et al., 2000)

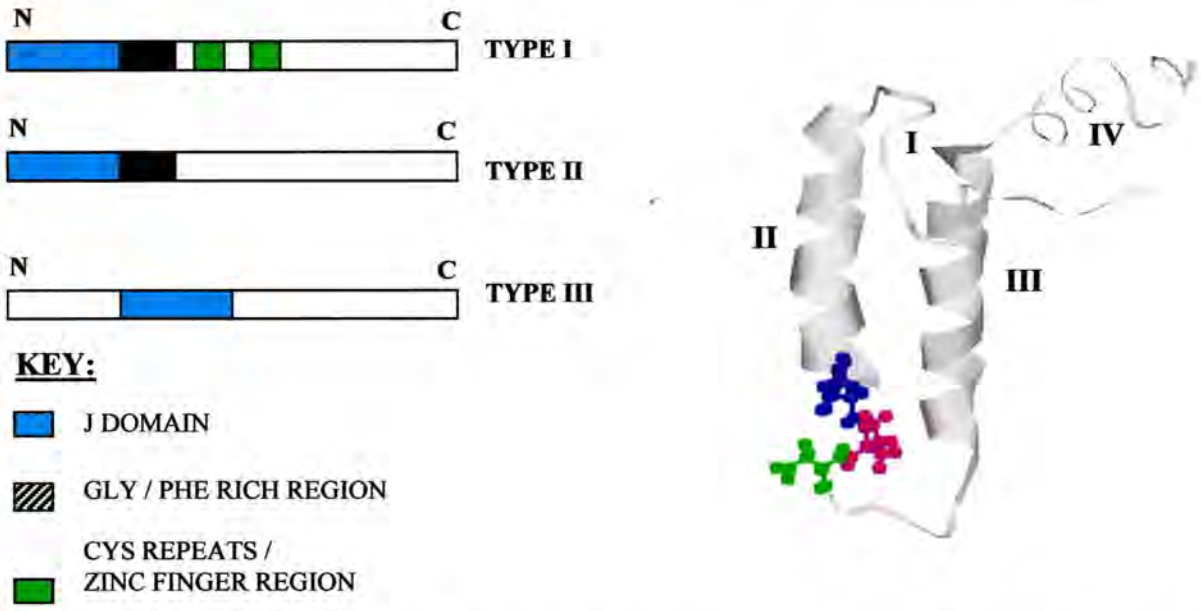


Figure 1.3: Classification of Hsp40s in terms of their structural motifs and structure and of the J-domain.

(A) Classification of DnaJ/Hsp40 family of molecular chaperones according to conserved domains. Type I J proteins possess an N-teriminal J domain, a G/F rich region and two cysteine repeats, forming a zinc finger region. C-terminal region is less well conserved. Type II J proteins possess the N terminal J domain and the G/F rich region only, while the Type III J proteins possess only the J domain, anywhere in the protein. The function of the particular J protein will be defined by the number and types of motifs that they possess. (Cheetham and Caplan, 1998)
(B) Ribbon representation of the structure of the N- terminal J-domain of the *E. coli* DnaJ chaperone protein (1XBL) (Pellechia et al., 1996) , visualised using Rasmol (Sayle and Milner-White, 1995). The three residues constituting the HPD motif, thought to be involved in the interaction with the ATPase domain of Hsp70, are highlighted and displayed as ball-and-stick. (His^{33} is shown in blue, Pro^{34} in pink and Asp^{35} in green.)

The J domain is a stretch of ~70 amino acids with a high degree of similarity to *E. coli* DnaJ, contained in all Hsp40s (Figure 1.3[B]). Structurally, the J domain consists of four α-helices with a loop region containing a highly conserved HPD (histidine-proline-aspartic acid) tripeptide motif (Figure 1.3[B]). It is this HPD motif of the J domain that is proposed to be involved in the stimulation of the ATPase domain of Hsp70, which facilitates chaperone activity (Kelley, 1998). A bioinformatic analysis of the J domains of 223 proteins revealed that there is a general similarity between the J domains of even functionally distinct proteins (Hennessy et al., 2000). In some cases, the J domains of certain Hsp40s, such as the J domain of the Simian virus 40 (SV40) large T-antigen, may be able to functionally replace the J domain of an unrelated Hsp40 (Kelley and Georgopoulos, 1997; Sullivan et al., 2000; Fewell et al., 2002). This suggests that although the J domain is required for the physical interaction with Hsp70, other regions of the Hsp40 must govern the specificity of the interaction.

Following the J domain in Type I and II Hsp40s is the largely unstructured G/F region, thought to be important as a 'hinge' region. It has been proposed that this linker region is important for determining the specificity of the interaction between Hsp70 and Hsp40 (Caplan et al., 1993). C-terminal to the G/F rich region in Type I Hsp40s is a region of cysteine repeats. The cysteine repeats are in the form of CysXXCysXGlyXGly, where X is any amino acid, usually a charged or polar residue (Caplan et al., 1993). This cysteine rich region coordinates two zinc atoms in a tetrahedral arrangement to form a zinc finger-like region, that is involved in the binding of denatured protein substrates and can itself prevent protein aggregation (Banecki et al., 1996; Szabo et al., 1996; Lu and Cyr, 1998; Martinez-Yamout et al., 2000). DnaJ has also been reported as having the ability to bind native proteins and peptides, potentially via this zinc finger-like region (Feifel et al., 1998). Modifications of sequence in this cysteine rich region would therefore affect the ability or specificity of Hsp40s to bind substrate proteins. These substrate binding regions of Hsp40s are often functionally distinct and tailored for the recognition of precise substrates (Johnson and Craig, 2001). In addition to these domains, certain eukaryotic Hsp40s contain a C-terminal CaaX motif (where C is cysteine, 'a' is an aliphatic amino acid and X is any amino acid), which is the site of potential protein prenylation (Glomset et al., 1990). This modification increases the hydrophobicity of the protein, promotes association with intracellular membranes and influences interaction with other proteins (Marshall, 1993; Kanazawa et al., 1997).

1.1.5 Protein Folding By The Hsp70/Hsp40 Complex

The interaction between Hsp70 and Hsp40 occurs to fulfill a variety of functions within the cell, not all of which are involved with molecular chaperoning. The Type of Hsp40, and therefore the number of the canonical domains it possesses, will influence the type of reaction. Table 1.2 summarises the interactions between Hsp70s and their potential partner Hsp40s and the functions these partnerships fulfill in the cell.

Table 1.2: Hsp70/Hsp40 Partnerships and Major Role fulfilled by the Interaction.

Hsp70 Molecular Chaperone	Hsp40 Co-Chaperone	Organism	Major Function/Role	Reference
Hsp70	Auxilin	*B. taurus*	Clathrin uncoating	Kelley (1998)
Kar2p/BiP	Sec63p	*S. cerevisiae*	Protein translocation	Kelley (1998)
Ssa	Ydj1	*S. cerevisiae*	Molecular chaperoning	Lu and Cyr (1998)
Ssa	Sis1	*S. cerevisiae*	Translation initiation	Horton et al. (2001)
BiP	ERdj4	*M. musculus*	Stimulation of BiP ATPase Activity	Shen et al. (2002)
Hsp70	Hdj1	*H. sapiens*	Associated with nucleus	Cyr et al. (1994)
Hsc70	Hsj1	*H. sapiens*	Neuron specific function	Kelley (1998)

The classical, and most widely studied, Hsp70/Hsp40 partnership is that which governs chaperone-assisted protein folding in the cell. As stated previously, the chaperones do not physically fold proteins, but rather facilitate the folding process by preventing non-productive side reactions such as aggregation. Chaperone-mediated folding by the Hsp70/Hsp40 (DnaK/DnaJ/GrpE in *E. coli*) is cyclical in nature and controlled by ATP (Figure 1.4) (Mayer et al., 2000). The affinity of Hsp70 for substrate is governed by the binding and hydrolysis of ATP, that is, when ATP is bound to Hsp70 the chaperone has a lower affinity for protein substrate, than when ADP is bound (Hiromura et al., 1998). There are two models that describe the mechanism of Hsp70/Hsp40 protein folding, in both cases the rate-limiting step of the system being the cleavage of the γ-phosphate of ATP (McCarty et al., 1995). The difference between the two models is with respect to the binding of the substrate to the Hsp70. One model proposes that Hsp70 binds the substrate and then Hsp40 binds, while the other model proposes that Hsp40 binds the substrate and delivers it to the substrate binding domain of Hsp70 (Mayer et al., 2000; Fink, 1999). In both cases, Hsp40 binds to the ATPase domain of Hsp70 and stimulates the hydrolysis of ATP to ADP. The ADP bound Hsp70 undergoes a conformational change that results in a higher affinity for the protein substrate and tight binding of the protein to the substrate domain of Hsp70. The regeneration of the ATP bound form of Hsp70 by nucleotide exchange results in a drop in affinity for the substrate and hence release of the substrate protein to fold correctly.

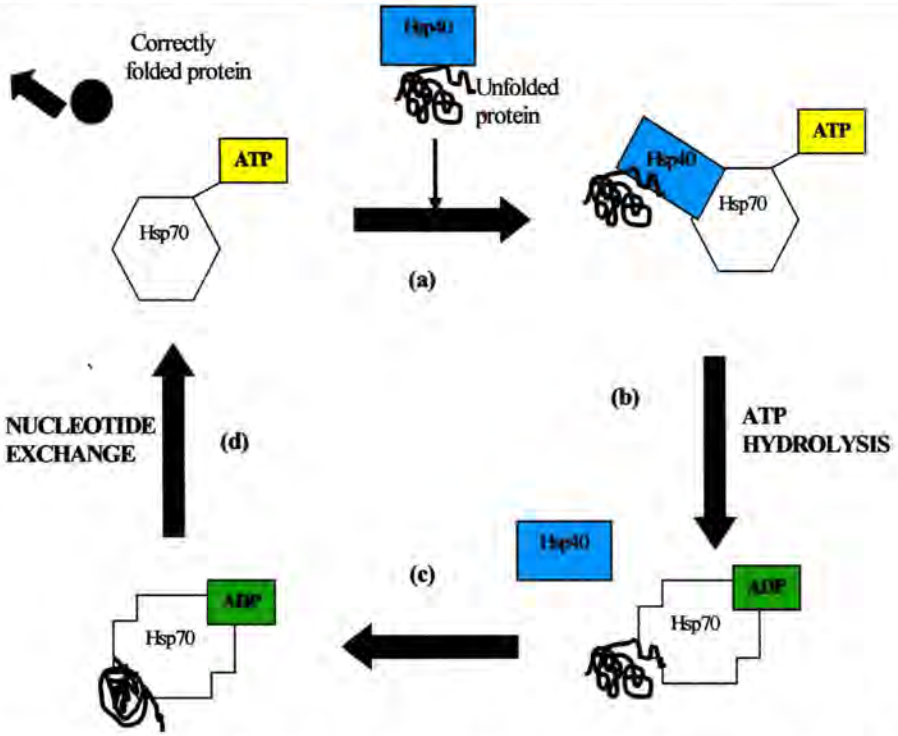


Figure 1.4: Proposed model of chaperone-assisted protein folding by the Hsp70/Hsp40 (DnaK/DnaJ/GrpE) complex.

(a) Hsp40/DnaJ forms a transient association with the unfolded protein (substrate) via exposed hydrophobic residues. **(b)** Hsp40/DnaJ binds to the ATP-bound form of Hsp70/DnaK , and transfers the misfolded protein to the substrate binding cleft of Hsp70/DnaK. Hsp40/DnaJ stimulates hydrolysis of ATP to ADP by the ATPase domain of Hsp70/DnaK. ADP bound form of Hsp70/DnaK has a higher affinity for the substrate, than the ATP bound form, and thus binds the misfolded protein tightly. Hsp40/DnaJ leaves the complex. **(c)** Misfolded protein is prevented from aggregation or non-productive folding pathways. **(d)** ADP bound form of Hsp70/DnaK is converted to the ATP bound form via nucleotide exchange. In prokaryotes this exchange is catalysed by the co-chaperone GrpE. A cytosolic GrpE homologue is yet to be discovered in eukaryotes. ATP bound Hsp70/DnaK has a lower affinity for the substrate and so the protein is released, to fold to its normal conformation or re-associate with another chaperone. The chaperones and co-chaperones are now free to continue the chaperone-assisted protein folding cycle. (Fink, 1999; Hendrik and Hartl, 1993).

In prokaryotes, nucleotide exchange is catalyzed by the co-chaperone GrpE, for which a cytosolic eukaryote homologue has not been fully identified. Recent work on the interaction of Hsp70s with the anti-apoptotic protein Bag-1 has shown that Bag-1 could be functioning as a eukaryotic version of GrpE to destabilize the ADP-bound form of Hsp70 and facilitate nucleotide exchange (Gassler et al., 2001; Brehmer et al., 2001). Brehmer and coworkers (2001) hypothesize that the eukaryotic Hsp70, due to a particular structural difference between the eukaryotic and prokaryotic Hsp70 ATPase domains, interacts differently with nucleotide exchange factors. Brehmer et al. (2001) identified residues on the opposite lobes of the ATPase domain of Hsp70 that they propose form salt bridges and thus hinder the nucleotide exchange process. Prokaryotic Hsp70 contains two potential salt bridges (necessitating the use of GrpE during nucleotide exchange), whereas the eukaryotic homologues possess only one potential salt bridge, possibly eliminating the requirement for a factor such as GrpE (Brehmer et al., 2001).

1.1.6 Interactions of Hsp70 and Hsp40 with other Heat Shock Proteins

Besides its role as a molecular chaperone in the refolding of denatured proteins, Hsp70 and Hsp40 interact with another molecular chaperone, Hsp90, during the assembly of specific receptors (Rajapandi et al., 2000). What is interesting about this interaction is that Hsp70 and Hsp90 do not bind each other, but rather their interaction is modulated by a protein known as Hop (Hsp70/Hsp90 Organising Protein) / STI1 (Stress Inducible Protein 1) (Johnson et al., 1998; Chen and Smith, 1998). Hop / STI1 contains three tetratricopeptide repeat (TPR) domains, with the Hsp70 binding site contained in TPR1 and the Hsp90 binding site in TPR2. The interaction between Hop / STI1 and Hsp70 and Hsp90 is specific and occurs through the C terminal EEVD motifs of both Hsp70 and Hsp90. Recently, the sites of interaction have been mapped to distinct amino acid residues, with the C terminal 'GPTIEEVD' peptide of Hsp70 essential for recognition and binding between the two proteins (Brinker et al., 2002). This specific interaction not only illustrates the diversity of the cellular function of Hsp70s, but also demonstrates the importance of the functionally conserved regions of the protein.

1.1.7 Heat Shock Proteins and Molecular Chaperones In Disease

The critical role of Hsps as molecular chaperones has made them the focus of interest in the stress conditions associated with pathology of the cell. Current research has implicated incorrect functioning of molecular chaperones in diseases such as Alzheimer's, Parkinson's and Creutzfeld-Jacob's diseases, where protein aggregates are hallmarks of the disorders (Slavotinek and Biesecker, 2001; Muchowski et al., 2000; Dobson, 1999). Molecular chaperones have also been targeted as potential therapeutic agents in certain disorders, via induction of specific Hsps (Smith et al., 1998). However, most of the research conducted is aimed at stimulating the host (human) Hsp expression in order to rectify protein-folding problems associated with certain disorders. Some work has reported on the role of the Hsps of the actual pathogen in the course of the disease (Smith et al., 1998). The Hsps of certain pathogens have been implicated in protecting the invading organism from host immunity, contributing to the pathogen's virulence and playing an important role in differentiation that occurs during the pathogen's lifecycle (Multhoff et al., 1998; Sharma, 1992). Hsps have also been identified as having potential roles in inflammation, antigen processing via the MHC Class I and Class II systems, receptor-mediated endocytosis and processing of host proteins (Krautz et al., 1998; Polla, 1988).

1.1.8 *Trypanosoma cruzi* And Chaga's Disease

Trypanosoma cruzi is the causative agent of the parasitic Chaga's disease (American trypanosomiasis). The African version of this disease is 'Sleeping Sickness' (African trypanosomiasis), caused by a parasite of the same genus, *Trypanosoma brucei*. Chaga's disease is characterized by an acute parasitemic phase, lasting 2 months, followed by a chronic phase that prevails until death (Krautz et al., 1998).

Some research has suggested that Hsps are related to the ability of *T. cruzi* to cause Chaga's disease. This parasite undergoes a lifecycle that is split between a cold-blooded vector and a warm-blooded host. The very course of the disease involves a heat shock and so it appears logical that the parasitic Hsps will be involved to some degree. In addition the fever associated with Chaga's disease will subject the parasite to elevated temperatures (Tafuri, 1999), which will induce the synthesis of Hsps (Olson et al., 1994). Parasitic Hsp70 proteins are capable of inducing a strong humoral and cellular immune response (Krautz et al., 1998).

1.1.9 *T. cruzi* Hsp70/Hsp40 Molecular Chaperones

Trypanosomal Hsp70 proteins have been identified in the cytoplasm, mitochondria and endoplasmic reticulum (Krautz et al., 1998). We are focusing on the characterization of the cytoplasmic, inducible Hsp70 (henceforth referred to as TcHsp70). Five trypanosomal Hsp40 co-chaperones (Tcj1, Tcj2, Tcj3, Tcj4 and Tcj6) have also been described (Tibbetts et al., 1998; Salmon et al., 2001).

The genes encoding the Hsp70 molecular chaperones are the best characterized of the Hsp gene families in *T.cruzi*. The *HSP70* gene is arranged as a tandem repeat of ten coding sequences and is transcribed as a polycistronic pre – mRNA that is post-transcriptionally modified, to yield the mRNA encoding each of the specific Hsp70 proteins (Olson et al., 1994). The expression of these genes is increased fourfold when the growth temperature of the parasite is increased from 26°C to 37°C, analogous to heat shock (Olson et al., 1994). In addition, the expression of Hsp70 proteins increases approximately threefold when the parasite is exposed to heat shock from 26°C to 37°C (Olson et al., 1994). Olson et al. (1994) reported that Hsp70 of *T.cruzi* has an extremely high ATPase activity, approximately 100 times greater than that of human Hsp70 (Olson et al., 1994). Recent structure-function

analyses on a novel *E. coli* Hsp70, HscC, has shown that it too has a high basal ATPase activity and in addition is independent of nucleotide exchange (Kluck et al., 2002). Kluck et al., (2002) propose that HscC is a member of a novel subfamily of Hsp70s, which have adapted these characterisitics to fulfill a specialized function.

TcHsp70 has been identified as being inducible on heat shock of the parasite, such as when it moves from its insect vector (26°C) to its mammalian host (37°C) (Krautz et al., 1998). TcHsp70 has been identified as one of the immunodominant antigens in patients infected with *T. cruzi* (Ahmed et al., 1999; Krautz et al., 1998; Skeiky et al., 1995). This antigenicity may be due to an extended GGMP repeat, found to be highly antigenic, contained in the C-terminus of TcHsp70 (Kumar and Zheng, 1998). TcHsp70 has been proposed to have a cytoprotective role in *T. cruzi*, when the parasite is in its host cell, by overcoming stress conditions such as oxidative stress, elevated temperature and even virulence (Polla, 1991). This action would require that TcHsp70 function as a molecular chaperone in the parasite. TcHsp70 was shown to stimulate cell proliferation in mice cells, resulting in apoptosis and cell death (Maranon et al., 2000). In addition, certain Hsp70s have been shown to associate with membranes (Arispe and De Maio, 2000) that may suggest a role for them in the processing and translocation of parasitic and host proteins (Polla, 1991).

In addition to TcHsp70, four Hsp40 homologues (Tcj1, Tcj2, Tcj3 and Tcj4), which could potentially act as co-chaperones for TcHsp70, have been identified (Tibbetts et al., 1998). All of the Hsp40 proteins contain the J domain that allows them to interact with Hsp70 and stimulate ATPase activity. All of the proteins, save Tcj1, are classified as Type I J proteins, as they all contain the J domain, glycine/phenylalanine rich region and the cysteine repeats. Tcj1 is a Type III J protein, possessing only a J domain. This is potentially the smallest Hsp40 described to date and shows little similarity (at most 25%) to Hsp40s from other organisms. Of the four identified, Tcj2 and Tcj3 are the most similar to each other, with Tcj3 being most similar in sequence to *E. coli* DnaJ. The *T. cruzi* Hsp40 proteins do not seem to be highly susceptible to heat shock, with Tcj2 showing the only marked increase (39%) in expression after heat shock. This suggests that Tcj2 may be important in stress conditions and thus be the partner protein for the stress-induced TcHsp70, while the non-inducible Tcj3 may interact with the cognate form of TcHsp70. The other trypanosomal Hsp40 isolated is Tcj6. This is a Type II Hsp40 and has been shown to be functionally similar to the yeast Hsp40, Sis1, and involved in initiation of translation in *T.cruzi* (Salmon et al., 2001)

Heat shock proteins have also been identified in the African strain, *Trypanosoma brucei,* which causes African trypanosomiasis. Interestingly, there is significant sequence identity between the Hsp70 proteins from the American (*T. cruzi*) and African (*T. brucei*) parasites. Sequence alignments indicate that the proteins show 84% identity and over 90% similarity at the primary amino acid sequence level. However, the lack of crystal structures for these proteins make it impossible to compare their tertiary structures, although molecular modeling programs allow model structures to be generated and used to make functional predictions. It could be suggested that the differences in sequence may account for the differences in function and activity between the parasitic and human Hsp70 proteins especially, enabling the parasite to colonize the human system and survive the stress conditions encountered therein.

Most of the research conducted on the *T.cruzi* Hsp70 and Hsp40 proteins to date deals with the individual characteristics and cellular biology of the two families. There is little reported on the possible interaction between the Hsp70 and Hsp40 proteins and the possible role of this in Chaga's disease. Characterization of the Hsp70 and Hsp40 chaperones from *T.cruzi* will not only yield results that could potentially be informative in elucidating the role of these proteins in the disease, but will contribute to the general body of scientific knowledge in this discipline.

1.2 HYPOTHESIS AND MOTIVATION

The majority of research conducted on the roles of Hsps in pathological states focuses on the use of molecular chaperones to prevent or treat disease, however little has been done on any potential role that Hsps may play in causing disease. *T. cruzi*, the causative agent of the incurable Chaga's disease (South American trypanosomiasis), experiences a heat shock during its infection of the human host. It seems possible that the trypanosomal Hsps may be involved in the protection of the parasitic cell from protein aggregation, thus acting as cytoprotective agents. Indeed, trypanosomal Hsps are found to be antigenic, suggesting that they must at some point in the lifecycle come into contact with the immune system of the host. All these factors suggest that the Hsps of *T. cruzi* may be integrally involved in protection of the invading parasite and may play a role in the pathogenicity. We hypothesise that TcHsp70 functions as a molecular chaperone and that it interacts with one or all of the Tcj proteins in chaperone-assisted protein folding. We propose that through this mechanism of chaperone activity, TcHsp70 may play a cytoprotective role in the parasite.

1.3 AIMS AND OBJECTIVES

In order to test the hypothesis that TcHsp70 functions as a classical Hsp70 chaperone and to study its interactions with the Tcj proteins, we have laid out the following objectives of this study:

1. To over-express the TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 proteins in a suitable heterologous system(s).
2. To purify the proteins to homogeneity, such that they may be used in functional biochemical studies.
3. To determine the basal ATPase activity of TcHsp70 *in vitro*, as previous work has indicated it to be significantly higher than reported for most Hsp70 proteins, and to investigate the effect of the presence of potential co-chaperones (Tcjs) on the ATPase activity of TcHsp70.
4. To determine the *in vitro* molecular chaperone activity of TcHsp70 in the absence and presence of potential co-chaperones (Tcjs), using model protein substrates.
5. To determine the molecular chaperone ability and cytoprotective capacity of Tcj proteins using *in vivo* complementation systems.
6. To perform bioinformatic analysis in conjunction with the laboratory work, to study primary sequence and tertiary structural differences and similarities between the trypanosomal Hsps and other Hsps. In particular, a comparison between the parasitic and mammalian Hsps will be performed in order to attempt to account for functional differences.

CHAPTER 2

BIOINFORMATIC ANALYSIS OF TRYPANOSOMAL HEAT SHOCK PROTEINS: PREDICTION OF FUNCTION AND DEVELOPMENT OF BIOLOGICAL REAGENTS

2.1 INTRODUCTION

Bioinformatics has a central role in science, especially with the recent explosions in computer technology and the enormous amount of information generated by the advent of gene and protein sequencing. Bioinformatics is the study of genes and proteins, using computers, in order to make rational predictions about their possible structure and function. If the information for the structure of a protein is contained within its primary amino acid sequence, then analysis of that sequence may provide a significant amount of information, allowing for educated predictions to be made and subsequently tested experimentally.

The only bioinformatic analysis on the trypanosomal Hsps to date was performed by Tibbetts et al. (1998), who performed a preliminary characterization of the Tcj proteins. Bioinformatic analyses were used to examine the protein sequences and to propose rational predictions with respect to their potential functions or interactions due to differences or similarity in the sequences, and used to support experimental findings. Alignments were performed to examine the similarity between the TcHsp70 and other well-characterised Hsp70s to define canonical domains essential to their function. Alignments of the four Tcjs were performed to identify those most similar to each other and to classify them to determine which could be potential partners for TcHsp70. Alignments of the Tcjs and Hsp40s of eukaryotic and prokaryotic origin were performed to determine the similarity of the Tcjs to other well described Hsp40s. Homology modeling was used to examine the potential similarity or differences between the predicted structure of certain trypanosomal Hsp domains to the equivalent domains of known structure in other Hsps. The composite analysis of the trypansomal proteins was used to identify peptide sequences that are potentially antigenic and may be used later for peptide-directed antibody synthesis to specifically detect each of the trypanosomal proteins.

2.1.1 Composite Protein Analysis and Peptide-directed Antibody Synthesis

The specific antibody-antigen reaction has been widely used in science for detection purposes. The advantage of peptide-directed antibody synthesis is that it does not require large concentrations of pure protein. Algorithms have been developed to assess the antigenic nature of peptide sequences in terms of criteria required for a region to be considered antigenic. Secondary structure predictions and antigenic analysis using software such as Generunner™ assess the following characteristics of peptide sequences – charge density, hydrophobicity,

hydropathy, flexibility and surface probability. The hydrophobicity along a sequence is analysed using the Hopp-Woods method to assign values to each amino acid derived from the transfer free energies of the side chains between ethanol and water (Hopp and Woods, 1981). The Kyte-Doolittle analysis combines the water-vapour free energies of the amino acids with their preference for an internal or external environment to generate a hydropathy scale (Kyte and Doolittle, 1982). The flexibility of the different regions of the protein are assessed by the Karplus-Schultz method, with individual amino acids being assigned as flexible or rigid (Karplus and Schultz, 1988) The Emini Surface Probability indicates which regions of sequence within the protein are most likely surface exposed, based on values assigned to specific amino acids according to the position they most often occupy in proteins of known structure. (Emini et al., 1985; Janin et al., 1978). These analyses are combined with secondary structure predictions to output an overall antigenic index for a peptide sequence within the final protein (Garnier et al., 1986; Chou and Fasman, 1978; Engelman et al., 1986). This peptide sequence can then be used to generate antibodies in a suitable system.

2.2 EXPERIMENTAL PROCEDURES

2.2.1 Analysis of Primary Amino Acid Sequence of TcHsp70

The primary amino acid sequences for TcHsp70 and Hsp70s from other organisms were downloaded from the NCBI Website (**http://www.ncbi.nlm.nih.gov/Entrez**) using the Entrez Browser. Alignments were performed using the web-based alignment package ClustalW (Corpet, 1988)

2.2.2 Modelling of TcHsp70 ATPase Domain

Low stringency homology modelling of the ATPase domain of TcHsp70 on the ATPase domain of *Bos taurus* Hsp70 (3HSC; Flaherty et al., 1990) was performed using SWISS – MODEL at **http://www.expasy.ch/SWISS – MODEL/swiss-model/** (Peitsch, 1996). Visualisation of the final models, and comparison of this model to known structures, was performed using Rasmol (Sayle and Milner-White, 1995).

2.2.3 Analysis of Primary Amino Acid Sequence of Tcj Proteins

Primary amino acid sequences for Tcj1, Tcj2, Tcj3 and Tcj4 were obtained from the NCBI Website at **http://www.ncbi.nlm.nih.gov/Entrez**. All alignments were performed using ClustalW (Corpet, 1998)

2.2.4 Modelling of Tcj J Domains and Substrate Binding Regions

The J domains and cysteine repeat regions of the Tcj proteins were modeled (described Section 2.2.2) on *E.coli* DnaJ J domain (1XBL; Pellichia et al., 1996) and substrate binding region (1EXK; Martinez-Yomout et al., 2000). These models were visualized using Rasmol (Sayle and Milner-White, 1995) and Molscript (Kraulis, 1991).

2.2.5 Composite Analysis of Trypansomal Heat Shock Proteins to determine Potential Antigenic Regions for Antibody Design

TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 amino acid sequences were analysed in order to determine which of the regions along the sequence could represent potential epitopic regions and therefore could be used for the synthesis of antibodies to the specific proteins. The algorithms of Karplus-Schultz, Hopp-Woods, Kyte-Doolittle, Emini and Jameson-Wolf were used to analyse all of the trypanosomal sequences. Potential amino acid sequences for peptide-directed antibody synthesis were selected and a BLAST search (Altschul et al., 1997) performed to ensure they were unique and specific to the target protein. The Tcj2 peptide was selected for peptide-directed antibody synthesis (Section 3.2.2.4).

2.3 RESULTS AND DISCUSSION

2.3.1 Bioinformatic Analysis of Primary Sequence of TcHsp70

Hsp70s comprise three domains that define their structure, a 44 kDa N-terminal ATPase domain, an 18 kDa substrate binding domain and a 10 kDa C-terminal region that modulates interactions with other proteins. This structure is observed for all Hsp70s and is essential for the particular function of Hsp70. The alignment of the primary amino acid sequences of TcHsp70, *Bos taurus* Hsp70 and *E. coli* DnaK is shown (Figure 2.1).

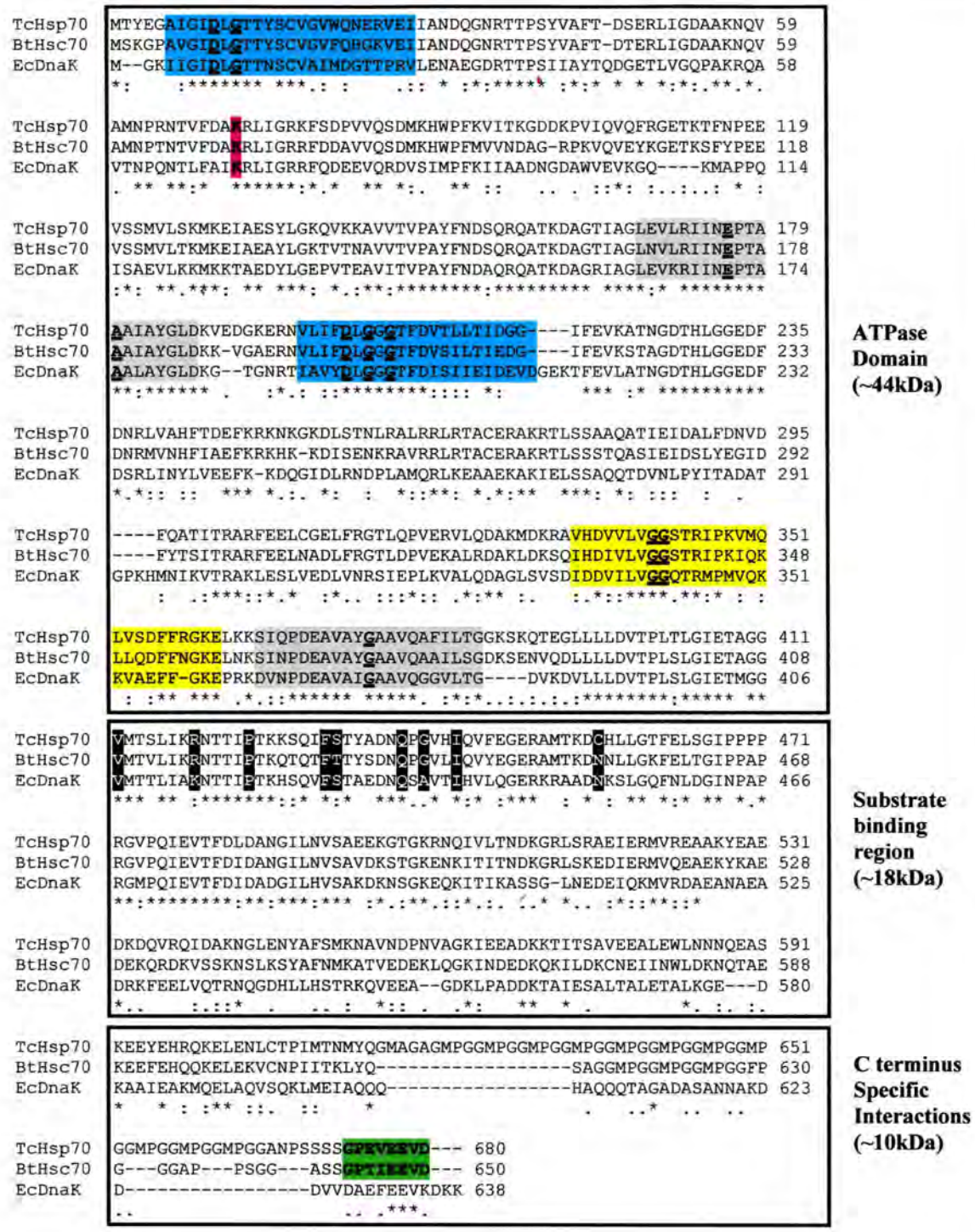


Figure 2.1: Defining the Canonical Hsp70 Domains in TcHsp70

Alignment of primary amino acid sequences of TcHsp70, Bovine Hsc70 and *E.coli* DnaK to define canonical, structural domains of Hsp70 proteins. Identical residues are indicated by * and similar residues by : or .. **ATPase domain structural elements** – blue shading indicates the phosphate binding sites; yellow shading indicates the region of adenosine coordination and grey shading indicates the linker regions (connect 1 and 2). Essential residues in each of the shaded areas are underlined (Bork et al.,1992). In addition, Lysine 71, essential for ATPase activity is shaded pink (O'Brien et al., 1996). **Substrate binding** – amino acids involved in substrate binding are shaded black (Montgomery et al., 1999; Zhu et al., 1996). **C terminal structural elements** – Green shading indicates residues involved in interaction of Hsp70 with Hop/STI1, that mediates interaction of Hsp70 with other proteins, such as Hsp90 (Freeman et al., 1995; Brinker et al., 2002).

The ATPase domain, substrate binding and C-terminal regions are indicated (Figure 2.1). The ATPase domain contains specific phosphate and adenosine binding regions, which are conserved. There is conservation of the functional residues within these domains that enable the ATPase domain to bind and hydrolyse ATP. In addition, the conserved linker regions are detected, as are the functional residues within the linker region (Bork et al., 1992). Lysine 71, essential for ATP hydrolysis, is also conserved in all of the sequences (O'Brien et al., 1996).

In addition to conservation of sequence within the ATPase domain, TcHsp70 contains the conserved amino acids in the substrate binding, purported to be involved in polypeptide binding (Montgomery et al., 1999; Zhu et al., 1996). The EEVD motif typically found at the C-terminus of eukaryotic Hsp70s is also highly conserved. This motif allows interaction of Hsp70 with STI1 to form the Hsp70-STI1-Hsp90 heterocomplex within the eukaryotic cell that fulfills chaperone functions such as the assembly of hormone receptors. The alignment of the primary sequence of TcHsp70 with the well-characterised bovine Hsc70 and *E.coli* DnaK shows that TcHsp70 contains the canonical domains required for classification of the protein as an Hsp70.

Having defined the Hsp70 canonical domains within TcHsp70, a primary sequence alignment with Hsp70s from potential host and reservoir organisms of *T.cruzi* was performed (Figure 2.2). The alignment of TcHsp70 with Hsp70s from potential hosts or reservoirs reveals that there is the expected similarity between the sequences governing the overall structure of the ATPase domain. The phosphate binding regions, adenosine coordinating region and linker region sequences are all conserved. The only exception is in the N-terminal phosphate binding sequence, where TcHsp70 has the sequence –**NER**– in contrast with the –**HGK**– of the host organisms (Figure 2.2, boxed).

In addition, there is an Ala to Phe change in the second phosphate binding region. The changes in TcHsp70 may influence the structure of the phosphate binding region of TcHsp70 ATPase domain. This may be significant, as the reported basal ATPase activity of TcHsp70 is appreciably higher than that of other Hsp70s (Olsen et al., 1994). Alterations in the phosphate binding regions may affect the rate of ATP hydrolysis by TcHsp70, and hence the basal ATPase activity.

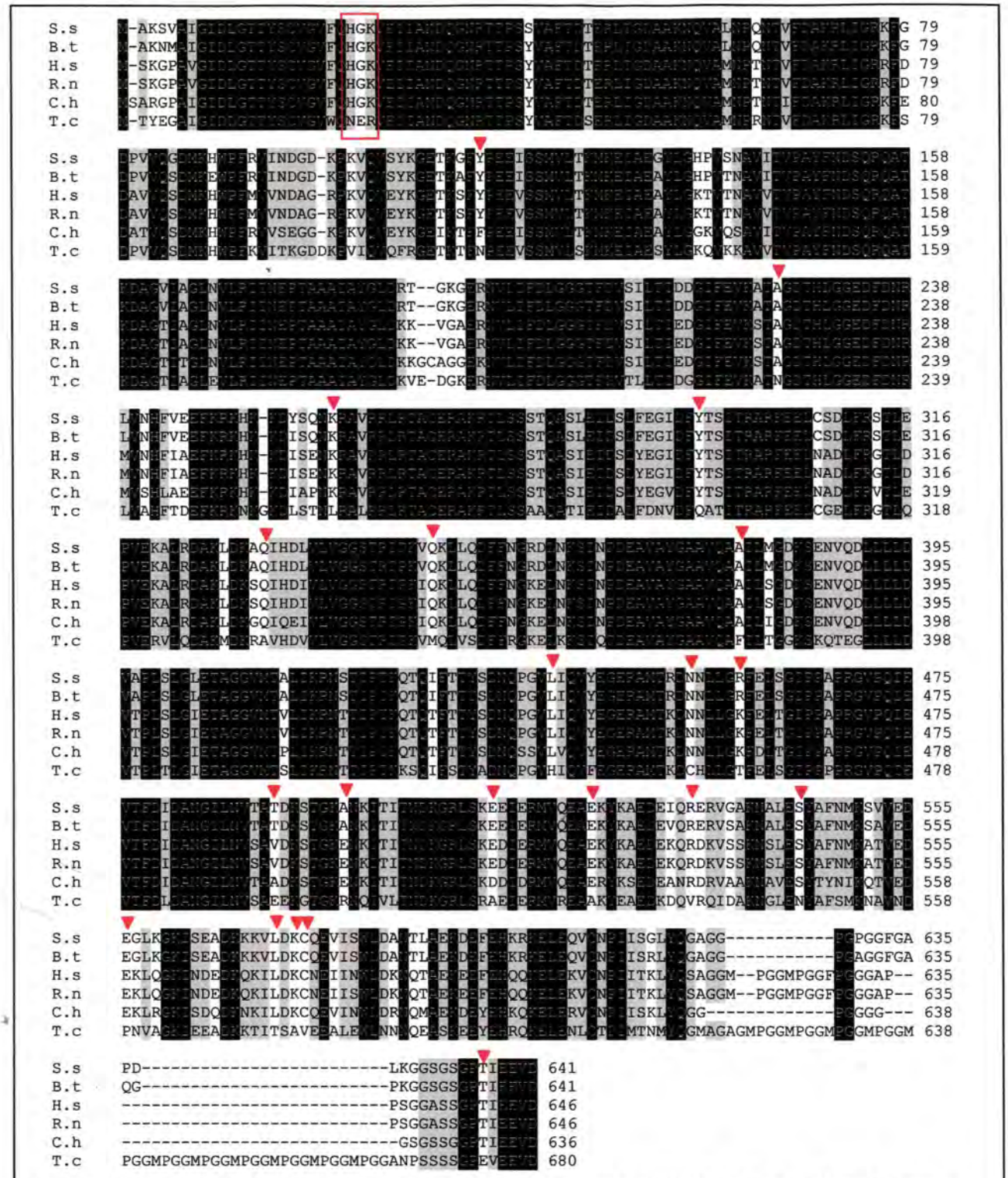


Figure 2.2: Alignment of Primary Amino Acid Sequences of TcHsp70 and Hsp70s from Host organisms

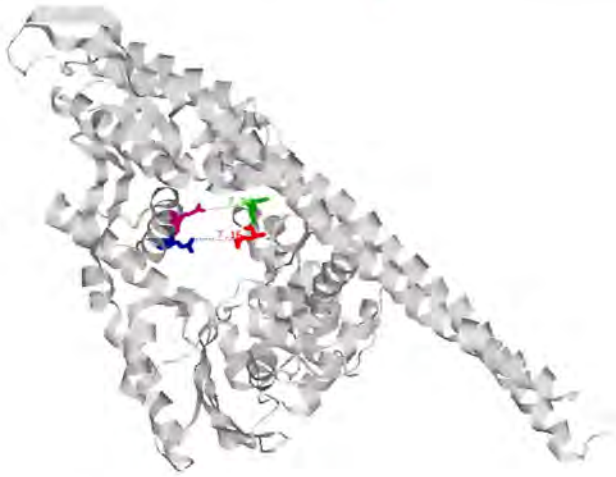
Multiple sequence alignment of TcHsp70 and Hsp70s from potential organisms, generated using ClustalW, internet based alignment package (Corpet, 1988). Residues highlighted in black indicate identical residues and residues shaded grey indicate those that are similar. Highly conserved sequences will most likely constitute amino acids essential for structural integrity. This is expected, as all Hsp70s assume a similar structure. Amino acids that vary significantly at certain positions will most likely affect the functional properties of the specific Hsp70. This alignment reveals the presence of distinct differences in sequence between TcHsp70 and the Hsp70s of potential host origin. In many cases there is a difference of charge or hydrophobicity in the TcHsp70 (indicated by red arrows). The difference in the N terminal phosphate binding region is boxed (red). The majority of the differences occur within the C terminal region of TcHsp70, which may affect its interaction with other proteins, such as STI1. T.c -*Trypanosoma cruzi*; C.h -*Capra hircus* (goat); H.s -*Homo sapiens* (human); R.n -*Rattus norvegicus* (rat); S.s -*Sus scrofa* (pig); B.t- *Bos taurus*(bovine/cow).

2.3.2 Homology Modelling of the ATPase Domain of TcHsp70

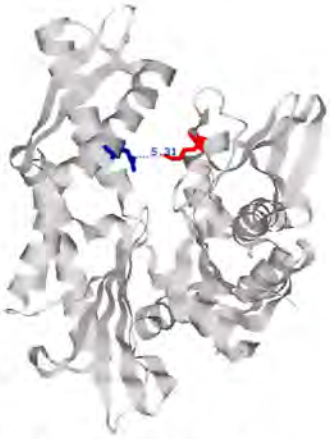
The potential structural effect of small changes in sequence in the ATPase domain of TcHsp70 was examined by modelling this domain on the known structure of the bovine Hsc70 ATPase domain. In addition to observing the overall structure, the potential salt bridge formation between the two lobes of the ATPase domain of TcHsp70 was examined. Scientists have long been searching for a cytosolic eukaryotic equivalent for the nucleotide exchange factor GrpE, and have recently identified the protein Bag-1 as a potential candidate. Brehmer and co-workers (2001) propose that nucleotide exchange in eukaryotic Hsp70s is governed by differences in the number of salt bridges formed as compared with prokaryotic Hsp70s.

Figure 2.3 shows the potential salt bridge formation between the two structural lobes of the ATPase domains of DnaK in complex with GrpE, *B. taurus* Hsp70 and *H. sapiens* Hsp70, compared with the modelled structure of the TcHsp70 ATPase domain. The figure details the potential formation of two salt bridges in DnaK/GrpE complex and the corresponding one in the eukaryotic Hsp70s. The Dnak/GrpE complex indicates the Hsp70 in the conformation that would favour nucleotide exchange, with distances of greater than 7Å between the four residues involved in the salt bridge formation. The distance between the two salt bridge residues in the corresponding human Hsp70 is 8.06 Å, suggesting that this may facilitate nucleotide exchange without a cofactor such as GrpE. However, the corresponding distance in the bovine and trypanosomal Hsp70s is 5.31 and 5.71Å respectively. This is smaller than the distance observed in the nucleotide-exchange conformation, which suggests a possible requirement for a nucleotide exchange factor.

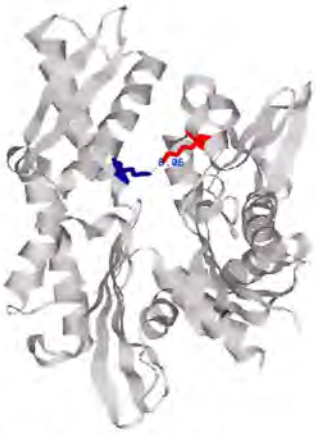
Recently, the protein Bag-1 has emerged as a potential eukaryotic GrpE functional homologue. It has been suggested that the number and characterisitics of the salt bridges affect the different interactions of prokaryotic DnaK with GrpE and the eukaryotic Hsp70s and Bag-1 (Brehmer, 2001). However, the homology model of the TcHsp70 ATPase domain was based on the structure of the ATPase domain of bovine Hsc70 and thus the actual structure of the ATPase domain of TcHsp70 and salt bridge distances may be different.



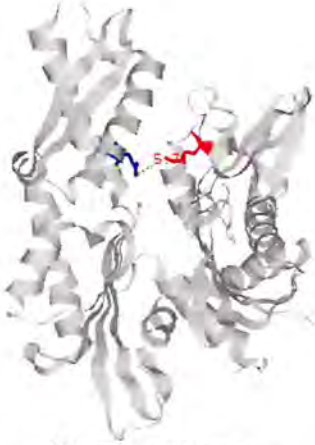
(a) ***E. coli*** **ATPase domain of DnaK in complex with GrpE (1DKG)**



(b) ***B. taurus*** **Hsp70 ATPase domain (3HSC)**



(c) ***H. sapiens*** **Hsp70 ATPase domain (1HJO)**



(d) ***T. cruzi*** **Hsp70 ATPase domain**

Figure 2.3: Ribbon representation of ATPase domain of Hsp70 proteins, to show potential salt bridge formation.

(a) *E. coli* DnaK in complex with GrpE, potential to form two salt bridges between Lys55 (red) - Glu267 (blue) (7.16 Å) and Arg56 (green)-Glu264 (pink) (7.90 Å); (b) *Bos taurus* Hsp70, potential formation of salt bridge between Lys55 (red) and Glu268 (blue) residues (distance 5.31 Å); (c) *Homo sapiens* Hsp70, potential formation of salt bridge between Lys56 (red) and Glu231 (blue) residues (distance 8.06 Å); (d) *T.cruzi* ATPase domain, modelled on *B.taurus* ATPase domain (3HSC). The formation of potential salt bridges between Lys56 (red) and glu (271) (5.71Å) is indicated. As GrpE is involved in nucleotide exchange in the DnaK chaperone cycle, it is proposed that the distances shown between the residues is representative of DnaK in an open conformation, during regeneration of the ATP from ADP. The fact that eukaryotic Hsp70s only have the potential to form a single salt bridge and the fact that the distance between the two residues (Lys56 and Glu231) in the *H.sapiens* Hsp70 is greater than 7.90 Å has lead to the proposal that this explains the absence of a cytosolic GrpE homologue in eukaryotes. However, the distance between the topologically equivalent residues in the TcHsp70 model is smaller (5.71Å) does not fit with this hypothesis. However, the TcHsp70 model is unrefined and based on the bovine structure. Thus the actual structure of the TcHsp70 ATPase domain may differ significantly. Models visualised using Rasmol (Sayle and Milner-White,1995).

2.3.3 Bioinformatic Analysis of the C-terminal Region of TcHsp70

The C-terminal regions of TcHsp70 also show a number of amino acid differences (shaded black) when compared to those of host origin. These changes in amino acid often result in a change in charge or hydrophobicity at that position. This could indicate a different substrate range for TcHsp70 or may result in an overall change in surface charge, as has been shown for parasitic Hsp70s of malarial origin. Whether these changes will give TcHsp70 some advantage over its host counterparts would have to be examined experimentally, for example using rational protein engineering. A closer observation of the C-terminal regions of TcHsp70 compared to its hosts' Hsp70s shows that there are specific changes at certain positions in the TcHsp70 sequence (Figure 2.4)

Figure 2.4: Comparison of the C-terminal region of TcHsp70 with Hsp70s of Host Origin

A multiple sequence alignment of the C termini of TcHsp70 and Hsp70s from organisms known to act as reservoirs or hosts for *T. cruzi*. The alignment reveals distinct amino acid differences in the sequence of TcHsp70 C terminus when compared to the other Hsp70s. The changes in amino acids include charge differences (such as positive as opposed to negative charge; or charged as opposed to not-charged/polar) and changes in sizes. These differences are shaded in black. In addition, TcHsp70 contains an extended GGMP repeat (boxed). The most significant difference can be observed at the very end of TcHsp70, in the region of the Hsp70s known to interact with STI1. TcHsp70 contains a change (blue) in the GPTIEEVD motif (yellow). Tcru -*Trypanosoma cruzi*; Chir -*Capra hircus* (goat); Hsap -*Homo sapiens* (human); Rnov -*Rattus norvegius* (rat); Sscr -*Sus scrofa* (pig).

The changes noted in the TcHsp70 sequence are not conservative ones, as they may result in a change of charge or hydrophobicity (Figure 2.4). This may potentially have an effect on the substrate binding regions of TcHsp70. It could suggest that the substrates bound by TcHsp70 are distinctly different to those bound by its hosts' Hsp70s. TcHsp70 also contains an extended GGMP repeat, that is distinctly different from those of its potential hosts (Figure 2.4, boxed). This sequence has been found to be highly antigenic and may be responsible for the generation of immune responses to TcHsp70 during infection. In addition, bioinformatic prediction package PSORT, has indicated that this region as a potential transmembrane-like region and may be involved in the association of TcHsp70 with cellular membranes.

TcHsp70 shows a deviation from the consensus sequence in its C-terminal GPTIEEVD motif. Instead of the consensus sequence (GPTIEEVD), TcHsp70 contains the sequence GPEVEEVD. This motif is involved in binding to STI1, and may have an effect on the formation of the Hsp70/STI1/Hsp90 heterocomplex in trypanosomes. STI1 is composed of three TPR domains (TPR1, TPR2A and TPR2B), with distinct binding sites on the different TPR domains for Hsp70 and Hsp90. EEVD is the general recognition sequence for binding by STI1, however the region of binding is governed by the Thr and Ile preceding that sequence. STI1 binds Hsp70 via its TPR1 domain, and the Ile determines the specificity of this binding in the sequence TIEEVD. By contrast, TPR2A of STI1, binds Hsp90 (Brinker et al., 2002). To observe whether this deviation from the consensus sequence is distinctive of parasitic Hsp70s an alignment of the C-terminal regions of different parasitic Hsp70s was performed (Figure 2.5)

The alignment shows that all the Hsp70s of parasitic origin show a deviation from the consensus sequence GPTIEEVD in their C-terminal motif. In all of the sample sequences, there is a substitution of the Ile with a Val. Although this is a conservative change it may affect the specificity of TcHsp70 binding to the *T. cruzi* STI1 (TcSTI1). The TPR1 domain of STI1 shows a specificity of binding for hydrophobic amino acids, and therefore it is likely that Val will still direct/promote this interaction. However, the change of Thr to a charged Lys in some of the parasitic Hsp70s may suggest a different specificity of binding. TcHsp70 is distinct from the other parasitic Hsp70s as it is the only one to contain a Thr to Glu change. This is a significant change and may suggest certain differences in the interaction of TcHsp70 with TPR 1 domain of TcSTI1 (Brinker et al., 2001).

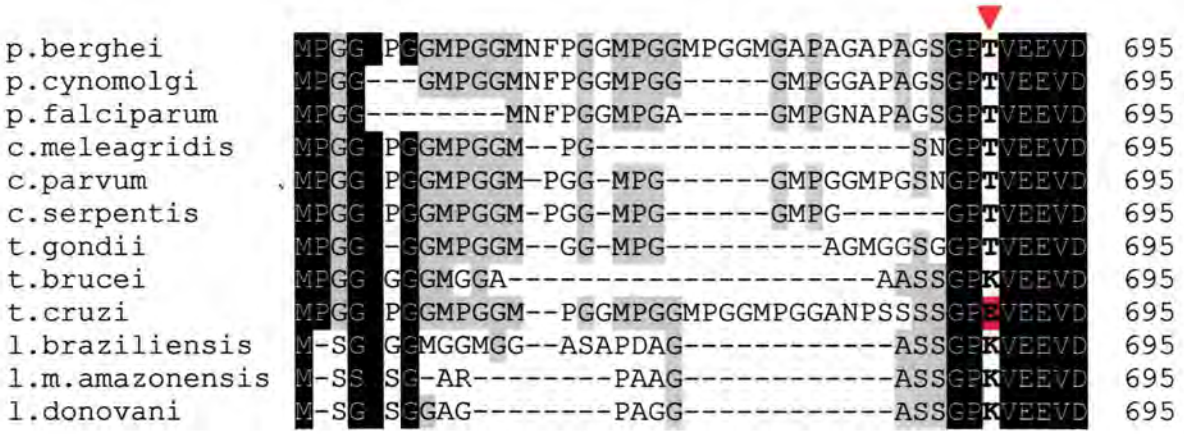


Figure 2.5: Alignment of C terminal regions of TcHsp70 with Hsp70s of Parasitic Origin

Residues in black indicate those of high conservation (100% of sequences), while the residues in grey indicate those of low conservation (50% of sequences). All of the Hsp70s of parasitic origin display a deviation from the consensus sequence GPTIEEVD (either GPTVEEVD or GPKEEVD, red arrow). The corresponding C terminal region of TcHsp70 (shaded red) is unique among the parasite Hsp70s (GPEVEEVD).

p. berghei- *Plasmodium berghei*; p. cynomogi- *Plasmodium cynomologi*; p. falciparum- *Plasmodium falciparum*; c. meleagridis- *Cryptosporidium meleagridis*; c. parvum- *Cryptosporidium parvum*; c. serpentis- *Cryptosporidium serpentis*; t. gondii- *Toxoplasma gondii*; t. brucei- *Trypanosoma brucei*; t. cruzi- *Trypanosoma cruzi*; l. braziliensis- *Leishmania braziliensis*; l. m. amazonensis- *Leishmania mexicana amazonensis*; l. donovani- *Leishmania donovani.*

To investigate the potential implications of a difference in the STI1 interaction motif of the parasitic Hsp70s, a sequence comparison of parasitic and human STI1 proteins was performed. Figure 2.6 defines the TPR 1 (Hsp70 binding) and TPR2A (Hsp90 binding) domains of the *T. cruzi* (TcSTI1), *Leishmania major* (LmSTI1) and *H.sapiens* (HsSTI1) STI1 proteins. The residues involved in general binding of Hsp70 and Hsp90 are indicated (Figure 2.6). These residues are conserved in all the sequences. The residues (Ala^{49} and Lys^{50}) thought to be involved in determining the specificity of Hsp70 binding to the TPR1 domain are highlighted (Figure 2.6, green). The parasitic STI1 proteins show a change in their sequences, with an Ala instead of a Lys at position 50. These differences may be indicative of a different interaction of the parasitic Hsp70s with their corresponding STI1 proteins.

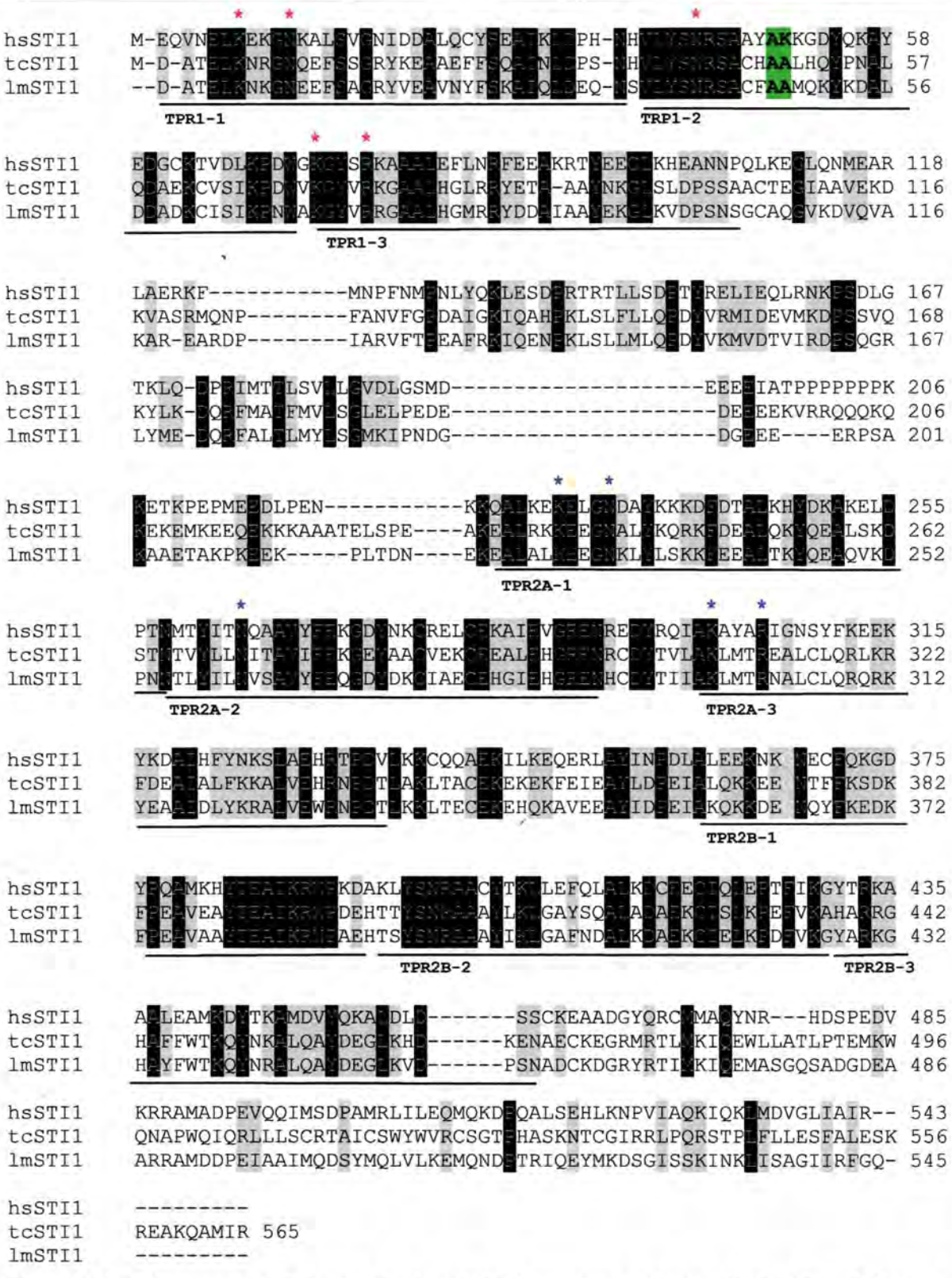


Figure 2.6: Alignment of Primary Amino Acid Sequences of Parasitic STI1 and Human STI1

Aligment of the primary amino acid sequences of the *H. sapiens* (hsSTI1), *T .cruzi* (tcSTI1) and *Leishmania major* (lmSTI1) STI1 proteins. These residues are conserved in all the sequences. Black shading indicates identical residues, while grey shading indicates similar residues. The residues comprising the TPR domains of the proteins are underlined. For HsSTI1, those residues (K8, N43, K73 and R77) in the TPR1 domain known for association with Hsc70 are indicated (*), as are those residues (K229, N233, N264, K301 and R305) in TPR2A known for association with Hsp90 (*). The residues highlighted in green indicate those implicated in the specificity of the Hsp70 interaction with the TPR1 domain of STI1

2.3.4 Definition of Canonical Domains and Classification of Tcj Proteins

The classification of Hsp40s into Type I, II or III depends on the degree of conservation of functional motifs that they share with *E. coli* DnaJ. This classification of Hsp40s is important to determining the potential role of the proteins, as typically Type I, II and III Hsp40s have distinct functions. Gènerally, Hsp40s involved in the Hsp70/Hsp40 protein-assisted folding process are Type I proteins. The preliminary characterization of the Tcj proteins was previously reported by Tibbetts et al. (1998). They classified Tcj2, Tcj3 and Tcj4 as Type I Hsp40s and Tcj1 as a Type III Hsp40. Tcj6 has been classified as a Type II Hsp40 (Salmon,2001).

The alignment of the Tcjs and the residues that comprise the J domain, G/F rich region and cysteine repeat region are indicated (Figure 2.7) All of the Tcjs contain an N-terminal J domain containing the highly conserved HPD motif, which is the motif known to interact with the ATPase domain of Hsp70s. All of the J domains show significant similarity, however, the J domains of Tcj2, Tcj3 and Tcj4 are most alike. The Tcj2, Tcj3 and Tcj4 sequences contain the G/F rich linker region, which is proposed to mediate the specificity of the Hsp70/Hsp40 partnership. The G/F rich region is extended in Tcj4, with the G/F rich regions of Tcj2 and Tcj3 being most similar to each other. Tcj2, Tcj3 and Tcj4 also contain the cysteine repeat region. This is a region of sequence containing four repeats of the form CysXXCysXXGlyXGly, where X is any amino acid (usually charged or polar) which is proposed to coordinate two zinc atoms forming a zinc finger like structure. It is this region that is supposedly involved in substrate binding by Hsp40s. This analysis confirms that Tcj1 is a Type III Hsp40. Tcj2, Tcj3 and Tcj4 are Type I Hsp40s, as described by Tibbetts et al. (1998), but only Tcj3 contains the consensus sequence in their final cysteine repeat region. Tcj2 and Tcj4 have deviations from the consensus sequence in their final cysteine repeat, Tcj2 contains the sequence CysXXCysXGlyX**Gln**, while Tcj4 contains the sequence CysXXCysX**MetXGln** as opposed to the consensus CysXXCysXGlyXGly. This deviation from the consensus sequence may suggest that these proteins bind substrates different to those Hsp40s containing the consensus sequence in their final cysteine repeat. Tcj2 and Tcj4 also contain the C-terminal CaaX motif, where 'a' is any amino acid and X is usually a charged or polar amino acid, such as glutamine. This is a protein prenylation modification site, involved in the interaction of Hsp40s with lipid membranes.

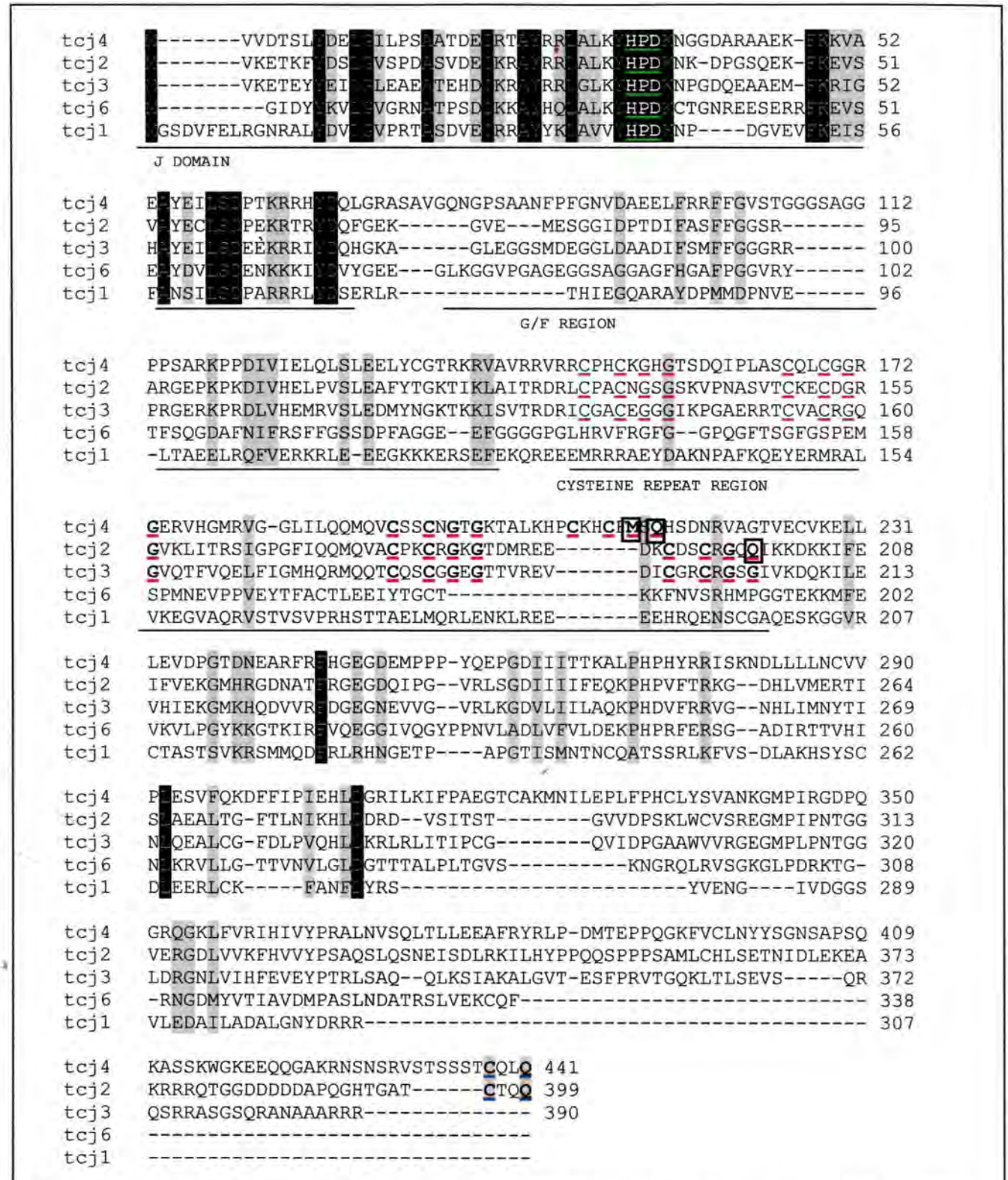


Figure 2.7: Multiple Sequence Alignment of Primary Amino Acid Sequences of Tcj Proteins

Alignment of the amino acid sequences of the Tcj proteins was performed using Clustal W (Corpet, 1988). Residues highlighted in black are identical and those residues highlighted grey are similar. The J domain, G/F rich region and cysteine repeat regions are indicated by underlining. In addition, the highly conserved HPD motif is underlined in green and the residues constituting the cysteine repeats are underlined in pink. The C-terminal prenylation site (CaaX motif) in Tcj2 and Tcj4 is underlined in blue. Tcj2 and Tcj4 also contain a deviation from the consensus sequence (CxxCxGxG) in their final cysteine repeat (boxed)

2.3.5 Homology Modelling of J domain and Cysteine Repeat Regions of Tcj Proteins

Primary sequence analysis has been used to classify and analyse the Tcj protein sequences. The potential structural significance of similarities and differences in the primary sequence was analysed by modelling the J domain and cysteine repeat region on those with a similar sequence and known structure. The N termini of all the Tcj proteins show the essential conserved structural and functional residues of the J domain. To confirm this, the N-terminal regions of Tcj1, Tcj2, Tcj3 and Tcj4 were homology modeled (Figure 2.8) on the known structure of the *E. coli* DnaJ J domain (1XBL; Figure 1.3).

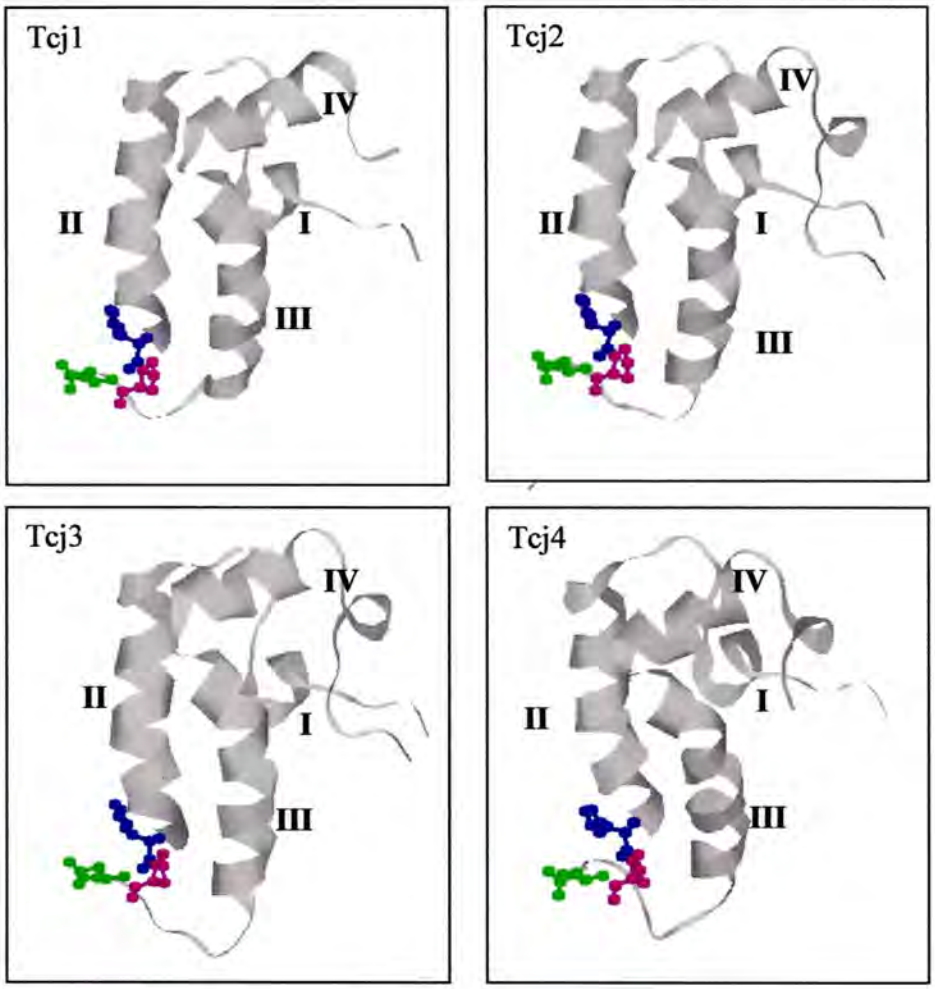


Figure 2.8: Ribbon Representation of Homology Modelled Structures of the J domains of the Tcj Proteins

The N-terminal regions of Tcj1, Tcj2, Tcj3 and Tcj4 containing the conserved J domain sequence was modelled using SWISS-MODEL on the known J domain structure of DnaJ (1XBL). The unrefined models were visualised using Rasmol (Sayle and Milner-White, 1995). The highly conserved tripeptide, His (blue), Pro (pink) and Asp (green), motif is highlighted. This HPD motif is the point of interaction of the Hsp40 with the ATPase domain of its partner Hsp70. The four Tcj J domains display a similar overall structure of four alpha helices (marked I –IV) with differences in the size of the loop region between helices II and III. This difference may result in orientation changes in the HPD motif, thereby potentially affecting the interaction with Hsp70.

The overall predicted structure of the J domain of the four Tcj proteins is similar, as expected because the J domain is a defined motif with a defined structure that is governed by the conserved structural residues. These structural residues are conserved throughout J domains of different function, therefore all of the J domains are composed of the four alpha helices, in the classical J domain structure. The difference in the potential structures arises in the loop region containing the HPD motif between α-helices II and III. There are predicted differences in the size and fold of this loop in the predicted models for the Tcjs that may suggest a change in the spatial orientation of the HPD motif, which may influence the interaction of these proteins with TcHsp70. On observation of the unrefined models, the predicted J domains of Tcj1 and Tcj2 appear most similar in structure, as do those of Tcj3 and Tcj4. However, as stated, the models are unrefined and therefore may be inaccurate and since all of the models are based on the same template (*E. coli* DnaJ 1XBL) some distinct similarity between the four is expected.

In contrast to the J domains, the substrate binding regions of Tcj2 and Tcj4 show sequence deviations from the consensus sequence in the sequence of their final cysteine repeat. Instead of the classical CysXXCysXGlyXGly motif in DnaJ (and Tcj3), the final cysteine repeat in Tcj2 has the sequence CysXXCysXGlyX**Gln** and Tcj4 the sequence CysXXCysX**Met**X**Gln** (Figure 2.9 [A]). The potential impact of these deviations from the consensus sequence on the predicted structure of the substrate binding domains of Tcj2 and Tcj4 was investigated by modelling their potential structures on that of DnaJ (1EXK) (Figure 2.9).

The four cysteine repeat regions of DnaJ form a zinc finger like structure, which coordinates two zinc atoms (Figure 2.9 [B]). The amino acids involved in the coordination (cysteine repeats) are shown in stick configuration (Figure 2.9 [A] and [B]). The structure is composed of three anti-parallel beta sheets linked by random coils that form the zinc finger like projection. The loop region in the base of the zinc finger is small. This structure will probably be suited to the type of substrate bound by DnaJ. The variation from the consensus sequence in the final cysteine repeats of Tcj2 and Tcj4, suggest an alteration to the structure of this region in Tcj2 and Tcj4. The predicted Tcj2 structure appears to have a more 'relaxed' conformation, with an extra pair of anti-parallel beta sheets and a larger loop region. The single amino acid deviation (Gly to Gln) suggests a potential abrogation of the coordination of the zinc atom at that position, potentially resulting in the predicted structure, which differs from the known structure of DnaJ.

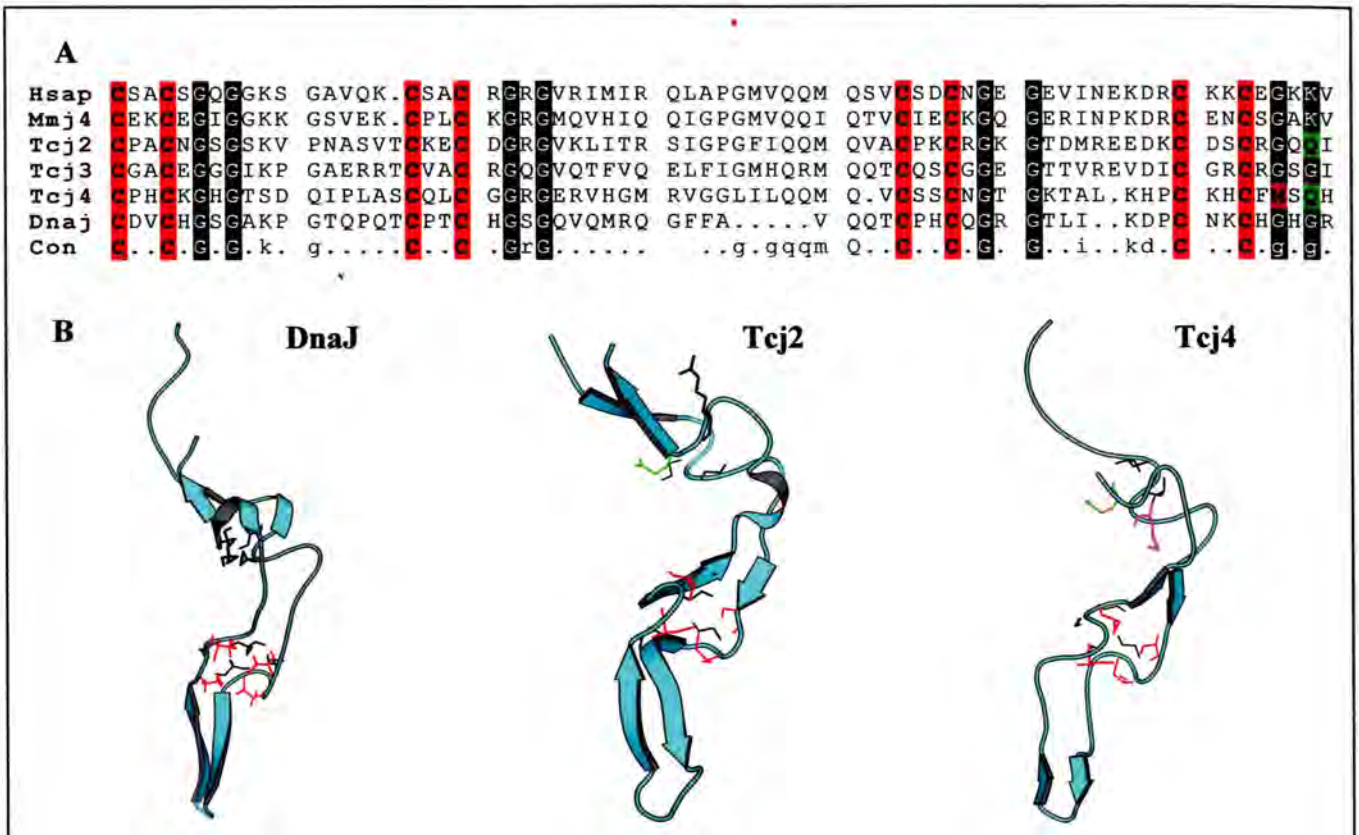


Figure 2.9: Analysis of the Proposed Structure of Substrate Binding Motif in Tcj2 and Tcj4

(A) Alignment of the cysteine repeat region of Tcj2 and Tcj4 with the equivalent regions of other Type I Hsp40s. Hsap- *H.sapiens* Hsp40; Mmj4- *Mus musculus* Hsp40; DnaJ- *E. coli* Hsp40; Tcj2-Tcj4- *T. cruzi* Hsp40s. The alignment reveals the difference in the final cysteine repeat of Tcj2 and Tcj4. Amino acids comprising the cysteine repeats are highlighted: Cys (red), Gly (black), Gln (green) and Met (pink).
(B) Ribbon representation of the homology models of the cysteine repeat substrate binding region of Tcj2 and Tcj4. The known structure of the equivalent region of DnaJ was used as the template. The residues highlighted in the alignment are represented as sticks and coloured accordingly. DnaJ contains the consensus sequence of four cysteine repeats (CysXXCysXXGlyXGly), that forms a zinc finger like domain known to coordinate two zinc atoms. The deviation from the consensus sequence in final cysteine repeat of Tcj2 (CysXXCysXGlyXGln) and Tcj4 (CysXXCysXMetXGln) would suggest abrogation of the zinc binding capacity and may alter the structure of the substrate binding motif. Models generated by SWISS-MODEL (Peitsch, 1995) and visualised with Molscript (Kraulis, 1991).

These potential structural changes in the substrate binding domain may be indicative of the type of substrate with which Tcj2 would normally associate. The greater variation in the Tcj4 final cysteine repeat appears to have an even more drastic effect on the predicted zinc finger like structure. The predicted Tcj4 structure is composed of only two anti-parallel beta sheets and a larger proportion of random coil. This suggests that potentially the real structure may differ significantly and that *E. coli* DnaJ was not a good model. This predicted structure suggests the complete abrogation of zinc binding capacity at that position. The predicted structure of Tcj4 substrate binding regions is distinctly different from the known DnaJ

structure. Both Tcj2 and Tcj4 are classified as Type I Hsp40s. Their J domains are very similar and yet, the sequence deviation in their substrate binding regions may alter their substrate binding domain structure, potentially discriminating between the types of substrates they may bind and therefore suggesting a possible difference in their physiological function. This is also true for Tcj3, which contains the consensus sequence in its substrate binding domain. As Tcj3 contains the consensus sequence, the structure of its substrate binding domain may be more similar to that of DnaJ, possibly meaning that the two proteins may have similar substrate binding specificities.

2.3.6 Sequence similarity between Tcjs and Hsp40s of Prokaryotic and Eukaryotic Origin

The comparison of the Type I Tcj protein sequences to sequences of Hsp40s of known function allow for the prediction as to which Hsp40 of known function they are most similar. Tcj1 was not included as it is classified as a Type III protein, and thus there is little expected similarity beyond the J domain, between it and the Type I proteins. Inclusion of the Tcj1 sequence would therefore adversely influence the alignments and consensus sequences defined. Sequence data is available for many well-characterised eukaryotic and prokaryotic Hsp40s. Yeast Hsp40s were selected as the eukaryotic Hsp40 representatives as both *T. cruzi* and *S. cerevisiae* are unicellular eukaryotes and DnaJ from *E. coli* chosen as the prokaryotic representative, as this is one of the best characterized Hsp40s.

Figure 2.10 presents the primary sequence alignment of Tcj2, Tcj3, Tcj4 and the prokaryotic Hsp40 DnaJ. There is significant sequence similarity between the J domain regions of the four Hsp40s. This is expected, even though DnaJ is of prokaryotic origin and the Tcjs are of parasitic origin, as the J domain is a highly conserved structural motif. All of the J domains contain the Hsp70 interacting HPD motif within the J domain. There is also a distinct similarity in the G/F rich regions of the four proteins. Tcj3 contains the consensus sequence in common with DnaJ in its cysteine repeat region, while Tcj2 and Tcj4 have a deviation from the consensus sequence in their final cysteine repeat (discussed Section 2.3.4).

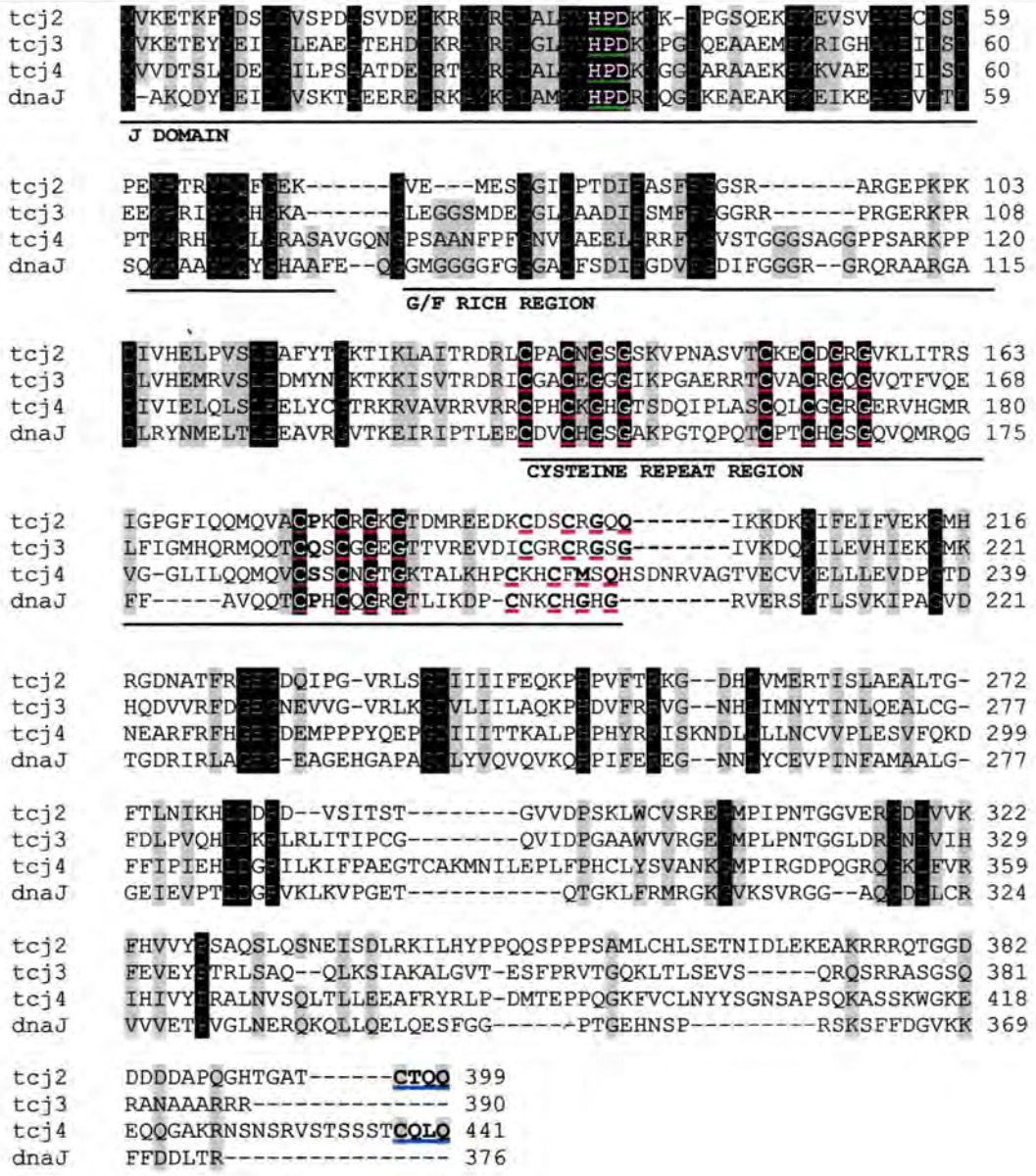


Figure 2.10: Comparison of the Primary Amino Acid Sequences of Tcj2, Tcj3, Tcj4 and *E. coli* DnaJ

Sequence alignment was performed using ClustalW (Corpet, 1988). Residues shaded in black are identical and residues in grey are similar. The J domain, G/F rich region and cysteine repeat region are underlined. The comparison of the primary amino acid sequences of Tcj proteins and DnaJ shows a high similarity in the N-terminal regions (J domain) with the similarity decreasing in the C terminal regions. The highly conserved HPD motif is underlined in green, the cysteine repeat regions of the proteins in pink and the CaaX prenylation motif in blue.

The deviation from the consensus sequence of the final cysteine repeat in the substrate binding regions suggests a possible difference between Tcj2, Tcj4 and DnaJ substrate specificities. In addition, DnaJ does not contain the CaaX motif found in Tcj2 and Tcj4. The fact that Tcj3 contains the consensus sequence in its cysteine repeat in common with that of DnaJ, suggests that these two proteins may have similar or overlapping substrate specificities. A similar alignment of the Tcj Type I proteins with the major yeast Hsp40s, Ydj1 and Sis1, was performed (Figure 2.11).

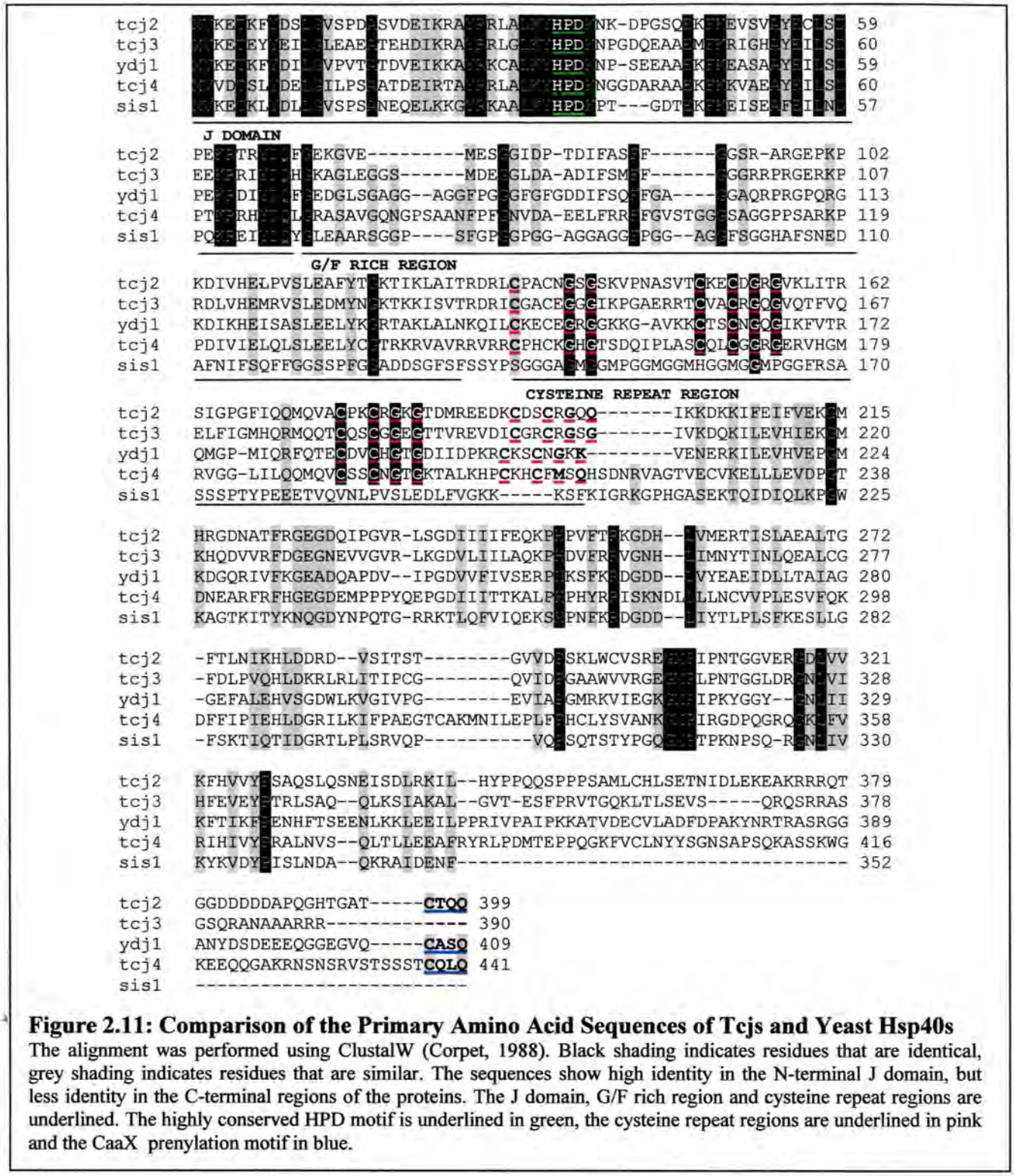


Figure 2.11: Comparison of the Primary Amino Acid Sequences of Tcjs and Yeast Hsp40s
The alignment was performed using ClustalW (Corpet, 1988). Black shading indicates residues that are identical, grey shading indicates residues that are similar. The sequences show high identity in the N-terminal J domain, but less identity in the C-terminal regions of the proteins. The J domain, G/F rich region and cysteine repeat regions are underlined. The highly conserved HPD motif is underlined in green, the cysteine repeat regions are underlined in pink and the CaaX prenylation motif in blue.

All of the Hsp40s show significant similarity in their N-terminal sequences. The essential HPD and other structural residues of the J domain are conserved between the Hsp40s (Figure 2.11). In addition there is conservation of the G/F rich region among all the Hsp40s used. The fourth cysteine repeat of Ydj1 contains a deviation from the consensus sequence (Gly to Lys). This suggests that Tcj2 or Tcj4, also containing deviations in the consensus sequence in this region, may potentially be similar in structure and therefore the proteins may have similar substrate specificities. Ydj1 also contains the CaaX prenylation motif found in Tcj2 and Tcj4.

2.3.7 Identification of Antigenic Regions for Peptide-directed Antibody Synthesis

The primary amino acid sequences of TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 were analysed to determine regions of potential antigenicity, for peptide-directed antibody synthesis. The composite analysis of a peptide sequence uses published algorithms to predict secondary structure and potential locations of regions of sequence (such as interior to the protein, surface exposed or transmembrane) with the final protein. In addition, the characteristics of each amino acid, such as hydrophobicity and flexibility, are reported. The results of the composite analysis can be studied in combination to determine regions of the protein containing peptide sequences that fulfill the properties of a potential antigen, and could thus be good targets for antibody synthesis.

For a peptide to be considered antigenic it must fulfill certain criteria. The antigenic region must be surface exposed, allowing interaction with an antibody in solution. Surface exposed regions are also most likely to be charged and/or hydrophilic, as opposed to hydrophobic regions that are more likely to be buried within the core of the protein. In addition to being surface exposed and hydrophilic/charged, an antigenic region must be flexible for the interaction with an antibody. This composite analysis was performed for all of the trypanosomal heat shock proteins, with a view to producing antibodies specific to each protein.

Using the criteria outlined above, amino acid sequences for peptide-directed antibody synthesis were chosen. Figure 2.12 shows the composite analysis for the amino acid sequence of TcHsp70 and the proposed antigenic peptide sequence. The region of TcHsp70 in which that peptide is found is boxed on the composite analysis figure (Figure 2.12).

The peptide sequence chosen fulfils all of the criteria of a potential epitope. The region is predicted to be flexible, hydrophilic, with a net negative charge and surface exposed. In addition, it shows a peak on the Jameson-Wolf Antigenic Index, identifying it as a good target for peptide-directed antibody synthesis. Observation of the composite analysis of TcHsp70 shows that there are a number of regions within the protein that fulfill these criteria. This is interesting, as TcHsp70 has been shown to be one of the immunodominant antigens in organisms infected with *T. cruzi* (Ahmed et al., 1999; Krautz et al., 1998; Skeiky et al., 1995). However, a peptide in the C-terminus of TcHsp70 was chosen as the target for antibody

synthesis, as the C-terminus shows the lowest degree of conservation across the range of Hsp70s. Antibodies raised to a peptide sequence in the N terminus, for example, may not be specific to TcHsp70 as this area of the protein contains the ATPase domain, a highly conserved motif found in all Hsp70s. The peptide sequence was BLAST searched and found to be specific to TcHsp70 alone. A high antigenic index is also recorded for a region at the extreme C-terminus of TcHsp70, corresponding to the extended GGMP repeat of TcHsp70, which has been previously determined as a highly antigenic motif (Kumar and Zheng, 1998).

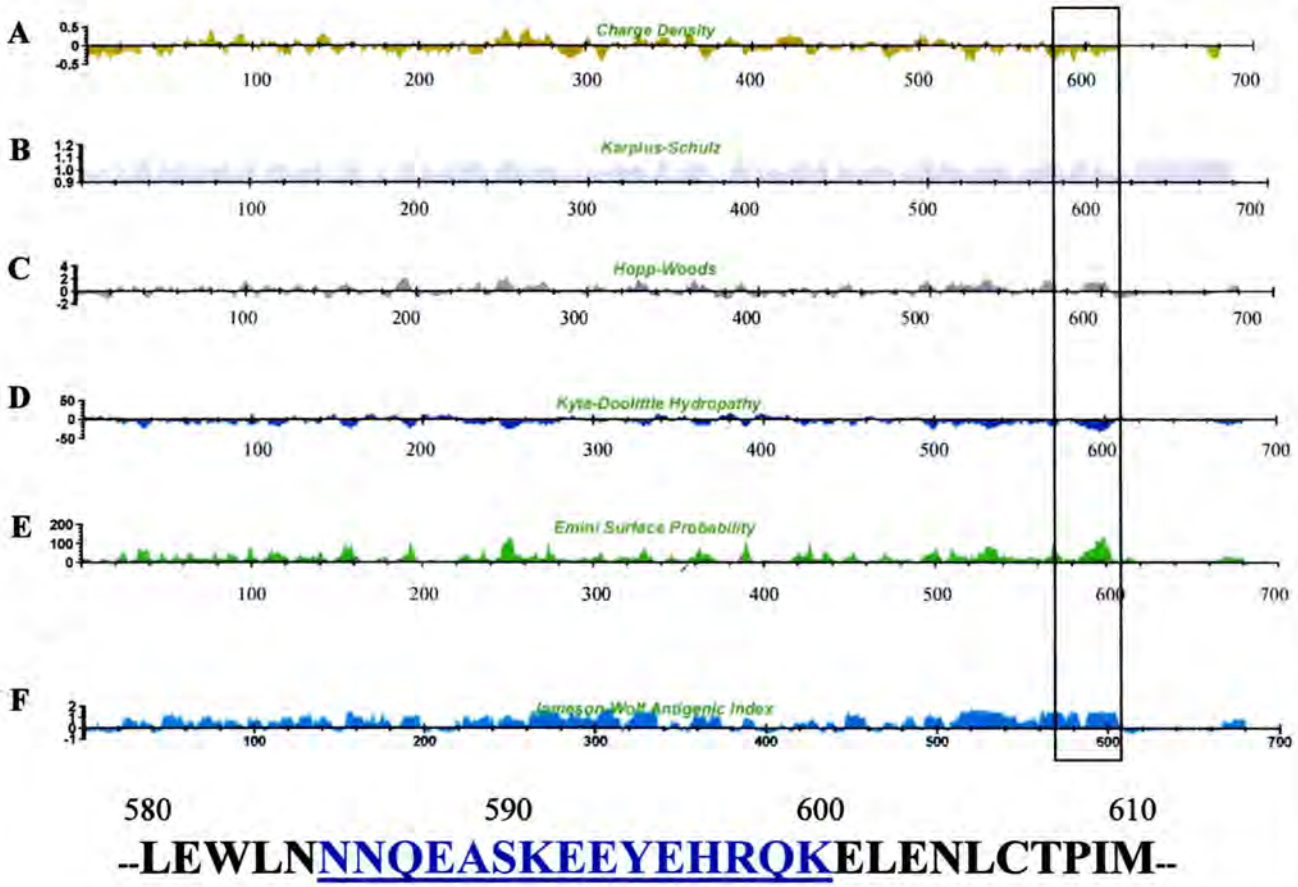


Figure 2.12: Composite Analysis of TcHsp70 Primary Sequence to determine potential Epitopic Regions

The peptide sequence identified as a potentially unique antigenic region (indicated below the graphical analyses) is shown in blue and the region of the protein in which it is found is boxed. Each graphical analysis represents the data for each amino acid along the length of the protein with the number on the x-axis indicative of the amino acid position from the N-terminus as 1. (A) Charge density analysis measures the average charge on each amino acid at physiological pH. Positive peaks indicate an overall positive charge and negative peaks indicate an overall negative charge. (B) The Karplus-Schultz analysis is a measure of the flexibility of each amino acid, with peaks indicating regions of potential high flexibility (Karplus and Schultz, 1988). (C) The Hopp-Woods analysis is a measure of hydrophobicity, with positive peaks indicative of hydrophilic regions and negative peaks indicative of hydrophobic regions (Hopp and Woods, 1981). (D) The Kyte-Doolittle analysis is a hydropathy scale, with positive peaks assigned to hydrophobic regions and negative peaks assigned to hydrophilic regions (Kyte and Doolittle, 1982). (E) The Emini Surface Probability analyses the preference of amino acids for surface or internal protein environments. Peaks indicate likely surface exposed regions (Emini et al., 1985; Janin et al., 1978). (F) The Jameson Antigenic Index combines the analyses above (A-E) and together with secondary prediction algorithms of Garnier-Robson, Chou-Fasman and Engelman-Steiz-Goldman, to generate an overall antigenicity analysis. Peaks are considered potential epitopic regions (Jameson and Wolf, 1988 ; Garnier et al., 1978; Chou and Fasman, 1978; Engelman et al., 1986).

Figure 2.13 shows the composite analysis of Tcj1 and shows the potential antigenic peptide. This region is indicated as potentially highly antigenic. A peptide outside of the J domain was chosen to ensure specificity of the antibody, as the J domain is a widely sequestered motif in a number of proteins. As the C-terminal regions of Hsp40s share the lowest level of similarity this is the preferred region to identify an epitope specific to Tcj1. As Tcj1 is a Type III Hsp40, there is also little similarity with other Hsp40s outside of the J domain.

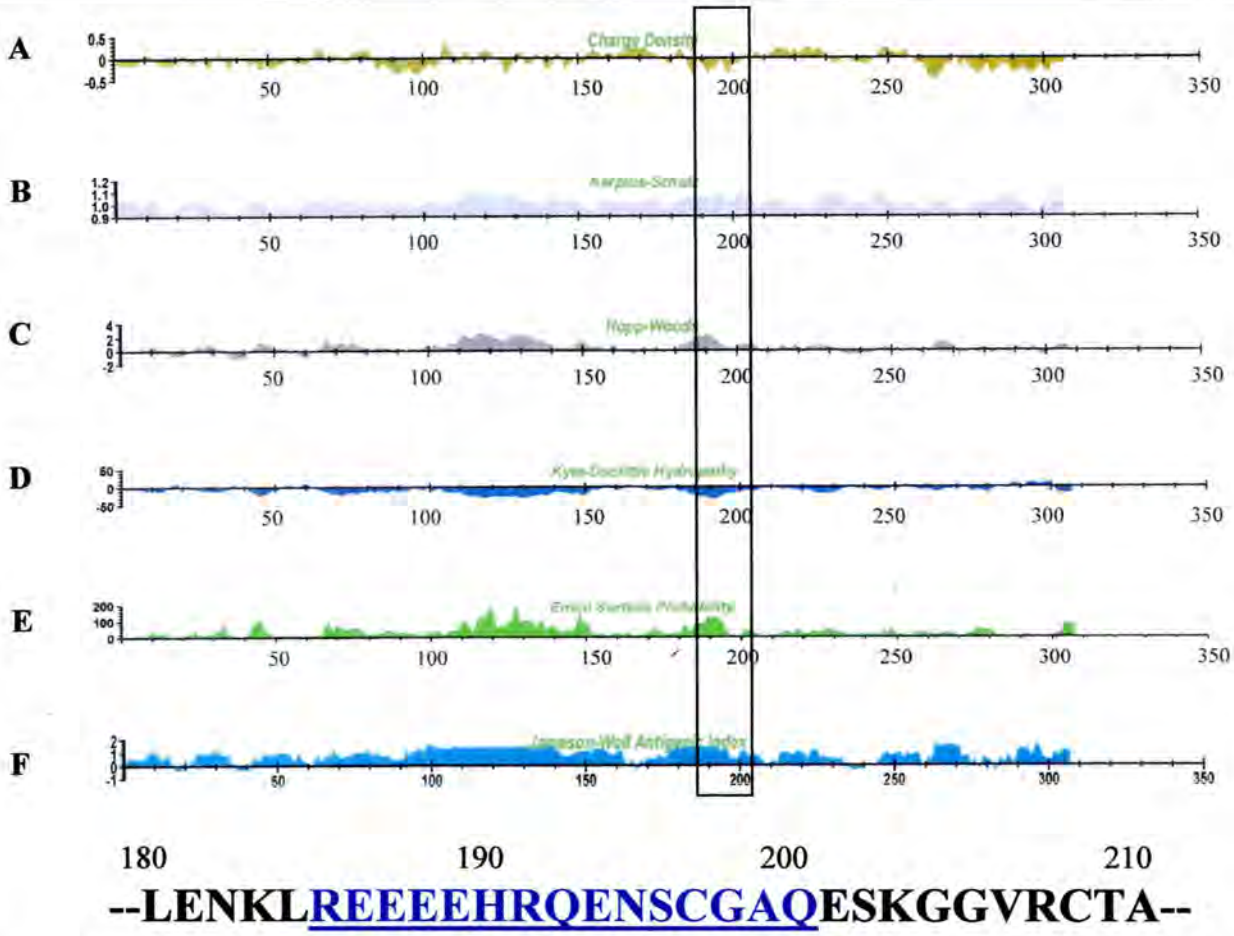


Figure 2.13: Composite Analysis of Tcj1 Primary Sequence to determine potential Epitopic Regions

The peptide sequence identified as a potentially unique antigenic region (indicated below the graphical analyses) is shown in blue and the region of the protein in which it is found is boxed. Each graphical analysis represents the data for each amino acid along the length of the protein with the number on the x-axis indicative of the amino acid position from the N-terminus as 1. (A) Charge density analysis measures the average charge on each amino acid at physiological pH. Positive peaks indicate an overall positive charge and negative peaks indicate an overall negative charge. (B) The Karplus-Schultz analysis is a measure of the flexibility of each amino acid, with peaks indicating regions of potential high flexibility (Karplus and Schultz, 1988). (C) The Hopp-Woods analysis is a measure of hydrophobicity, with positive peaks indicative of hydrophilic regions and negative peaks indicative of hydrophobic regions (Hopp and Woods, 1981). (D) The Kyte-Doolittle analysis is a hydropathy scale, with positive peaks assigned to hydrophobic regions and negative peaks assigned to hydrophilic regions (Kyte and Doolittle, 1982). (E) The Emini Surface Probability analyses the preference of amino acids for surface or internal protein environments. Peaks indicate likely surface exposed regions (Emini et al., 1985; Janin et al., 1978). (F) The Jameson Antigenic Index combines the analyses above (A-E) and together with secondary prediction algorithms of Garnier-Robson, Chou-Fasman and Engelman-Steiz-Goldman, to generate an overall antigenicity analysis. Peaks are considered potential epitopic regions (Jameson and Wolf, 1988; Garnier et al., 1978; Chou and Fasman, 1978; Engelman et al., 1986).

Figure 2.14 shows the composite analysis of the amino acid sequence of Tcj2. This region fulfils all the criteria of an antigenic region, being hydrophilic, flexible and surface exposed. A region in the C-terminus, outside all of the common Hsp40 motifs (J domain, G/F region, cysteine repeat region), was chosen to ensure specificity. Designing an antibody to the C-terminus of a protein will ensure that the full-length protein is detected, as opposed to an N-terminal antibody that would detect both truncated and full-length protein.

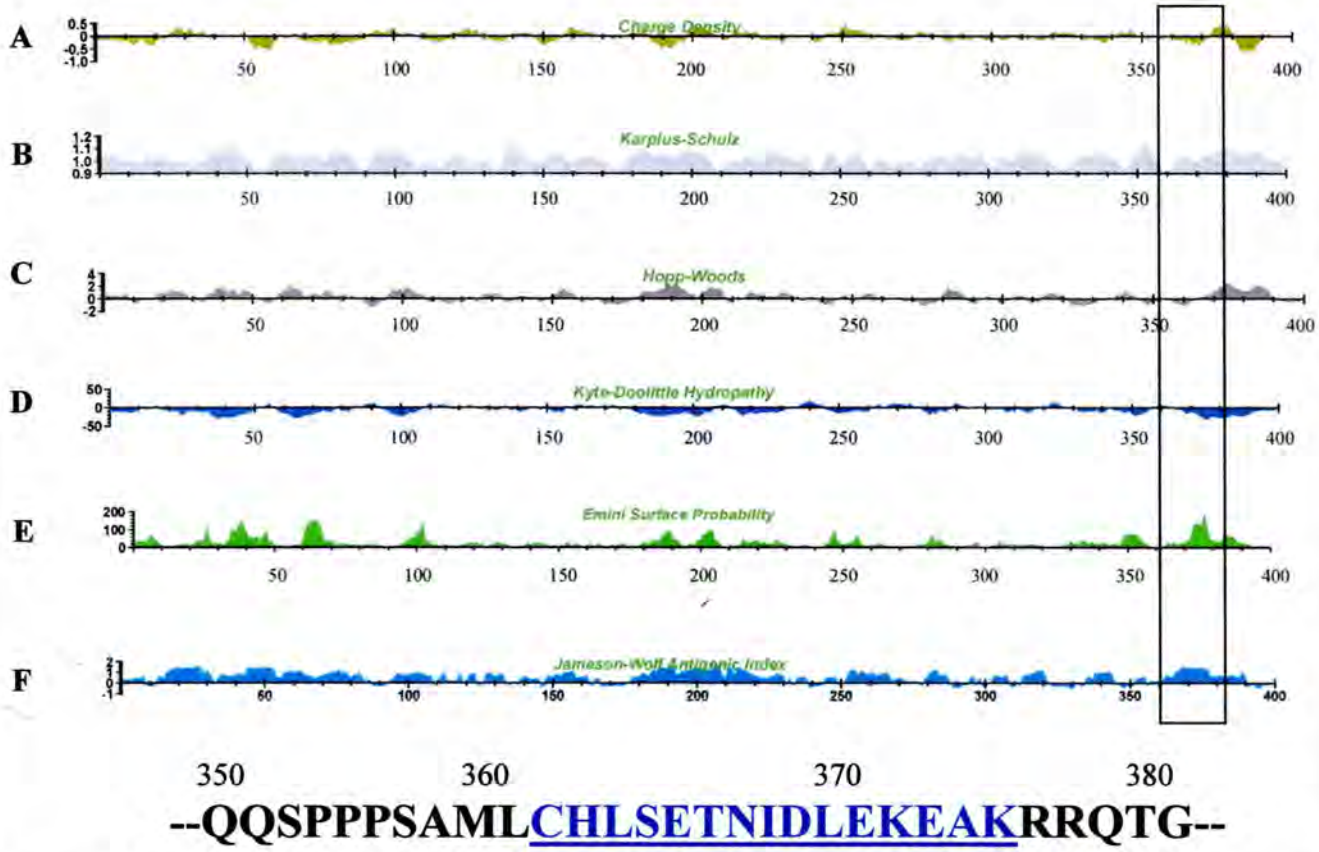


Figure 2.14: Composite Analysis of Tcj2 Primary Sequence to determine potential Epitopic Regions

The peptide sequence identified as a potentially unique antigenic region (indicated below the graphical analyses) is shown in blue and the region of the protein in which it is found is boxed. Each graphical analysis represents the data for each amino acid along the length of the protein with the number on the x-axis indicative of the amino acid position from the N-terminus as 1. (A) Charge density analysis measures the average charge on each amino acid at physiological pH. Positive peaks indicate an overall positive charge and negative peaks indicate an overall negative charge. (B) The Karplus-Schultz analysis is a measure of the flexibility of each amino acid, with peaks indicating regions of potential high flexibility (Karplus and Schultz, 1988). (C) The Hopp-Woods analysis is a measure of hydrophobicity, with positive peaks indicative of hydrophilic regions and negative peaks indicative of hydrophobic regions (Hopp and Woods, 1981). (D) The Kyte-Doolittle analysis is a hydropathy scale, with positive peaks assigned to hydrophobic regions and negative peaks assigned to hydrophilic regions (Kyte and Doolittle, 1982). (E) The Emini Surface Probability analyses the preference of amino acids for surface or internal protein environments. Peaks indicate likely surface exposed regions (Emini et al., 1985; Janin et al., 1978). (F) The Jameson Antigenic Index combines the analyses above (A-E) and together with secondary prediction algorithms of Garnier-Robson, Chou-Fasman and Engelman-Steiz-Goldman, to generate an overall antigenicity analysis. Peaks are considered potential epitopic regions (Jameson and Wolf, 1988; Garnier et al., 1978; Chou and Fasman, 1978; Engelman et al., 1986).

Figure 2.15 shows the composite analysis of the Tcj3 primary sequence. The region chosen is in the C-terminus of Tcj3, which is shown to be highly antigenic. It is very likely surface exposed, hydrophilic and flexible.

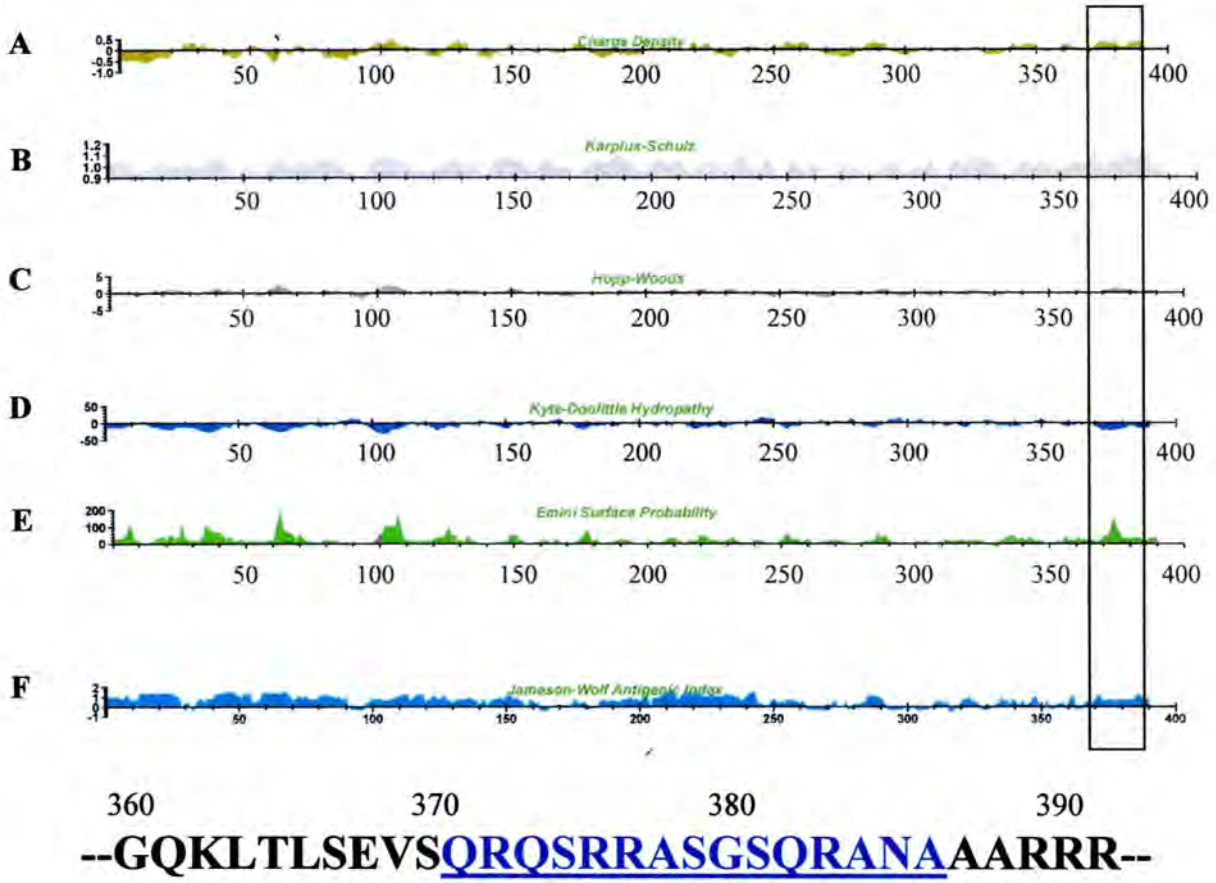


Figure 2.15: Composite Analysis of Tcj3 Primary Sequence to determine potential Epitopic Regions

The peptide sequence identified as a potentially unique antigenic region (indicated below the graphical analyses) is shown in blue and the region of the protein in which it is found is boxed. Each graphical analysis represents the data for each amino acid along the length of the protein with the number on the x-axis indicative of the amino acid position from the N-terminus as 1. (A) Charge density analysis measures the average charge on each amino acid at physiological pH. Positive peaks indicate an overall positive charge and negative peaks indicate an overall negative charge. (B) The Karplus-Schultz analysis is a measure of the flexibility of each amino acid, with peaks indicating regions of potential high flexibility (Karplus and Schultz, 1988). (C) The Hopp-Woods analysis is a measure of hydrophobicity, with positive peaks indicative of hydrophilic regions and negative peaks indicative of hydrophobic regions (Hopp and Woods, 1981). (D) The Kyte-Doolittle analysis is a hydropathy scale, with positive peaks assigned to hydrophobic regions and negative peaks assigned to hydrophilic regions (Kyte and Doolittle, 1982). (E) The Emini Surface Probability analyses the preference of amino acids for surface or internal protein environments. Peaks indicate likely surface exposed regions (Emini et al., 1985; Janin et al., 1978). (F) The Jameson Antigenic Index combines the analyses above (A-E) and together with secondary prediction algorithms of Garnier-Robson, Chou-Fasman and Engelman-Steiz-Goldman, to generate an overall antigenicity analysis. Peaks are considered potential epitopic regions (Jameson and Wolf, 1988; Garnier et al., 1978; Chou and Fasman, 1978; Engelman et al., 1986).

Figure 2.16 shows the composite analysis of the Tcj4 amino acid sequence. The peptide chosen is located in the C-terminus and has a high antigenic index. It is hydrophilic, flexible and most probably surface exposed.

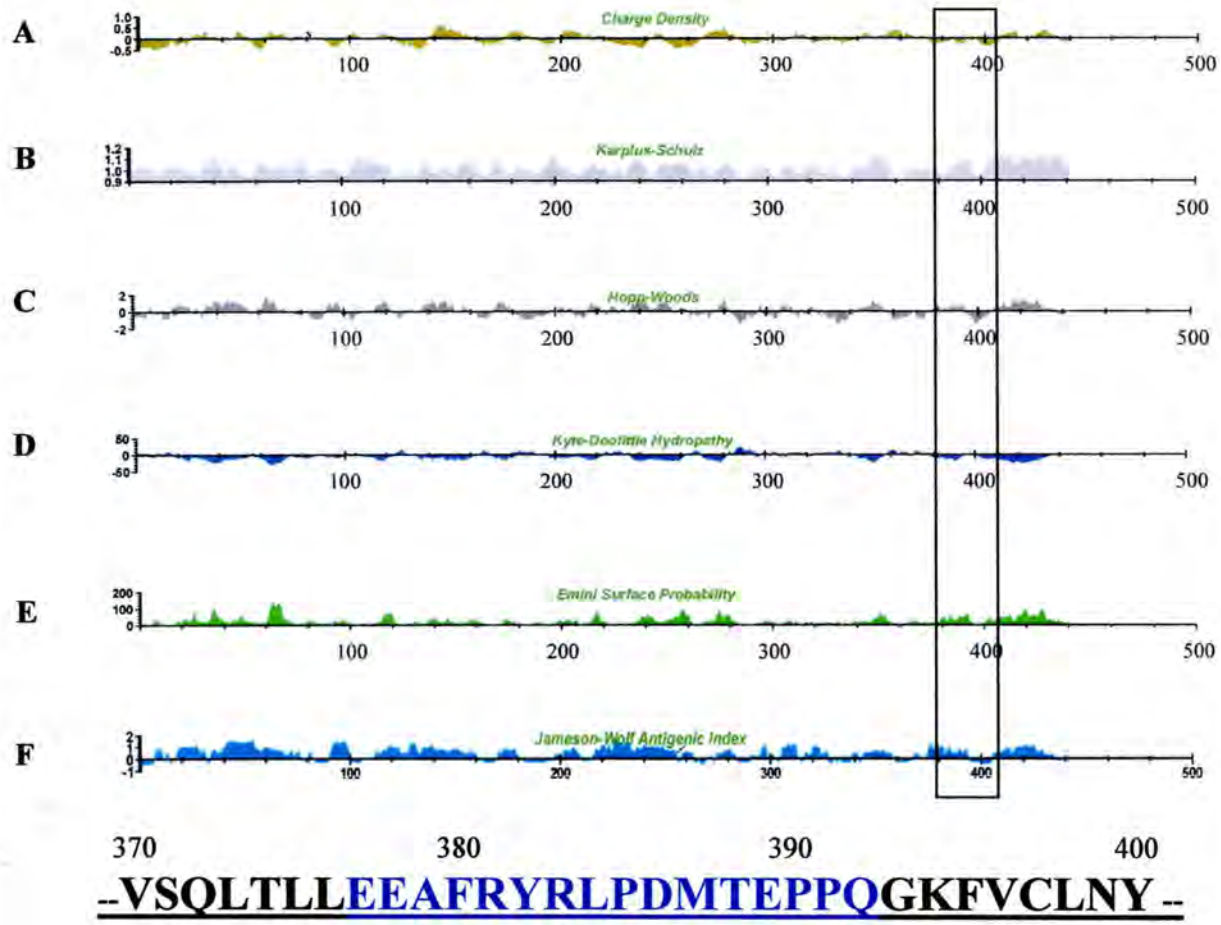


Figure 2.16: Composite Analysis of Tcj4 Primary Sequence to determine potential Epitopic Regions

The peptide sequence identified as a potentially unique antigenic region (indicated below the graphical analyses) is shown in blue and the region of the protein in which it is found is boxed. Each graphical analysis represents the data for each amino acid along the length of the protein with the number on the x-axis indicative of the amino acid position from the N-terminus as 1. (A) Charge density analysis measures the average charge on each amino acid at physiological pH. Positive peaks indicate an overall positive charge and negative peaks indicate an overall negative charge. (B) The Karplus-Schultz analysis is a measure of the flexibility of each amino acid, with peaks indicating regions of potential high flexibility (Karplus and Schultz, 1988). (C) The Hopp-Woods analysis is a measure of hydrophobicity, with positive peaks indicative of hydrophilic regions and negative peaks indicative of hydrophobic regions (Hopp and Woods, 1981). (D) The Kyte-Doolittle analysis is a hydropathy scale, with positive peaks assigned to hydrophobic regions and negative peaks assigned to hydrophilic regions (Kyte and Doolittle, 1982). (E) The Emini Surface Probability analyses the preference of amino acids for surface or internal protein environments. Peaks indicate likely surface exposed regions (Emini et al., 1985; Janin et al., 1978). (F) The Jameson Antigenic Index combines the analyses above (A-E) and together with secondary prediction algorithms of Garnier-Robson, Chou-Fasman and Engelman-Steiz-Goldman, to generate an overall antigenicity analysis. Peaks are considered potential epitopic regions (Jameson and Wolf, 1988; Garnier et al., 1978; Chou and Fasman, 1978; Engelman et al., 1986).

A detailed analysis of Figures 2.13 to 2.16 shows a similar composite analysis profile in the sequences comprising the J domains (Figures 2.13 to 2.16, residues 1-70) of all of the Tcj proteins and substrate binding regions (Figures 2.14 to 2.16, residues 150-200) of Tcj2, Tcj3 and Tcj4. All the analyses show a distinct peak at the peptide sequence 50 amino acids into the proteins (Figures 2.13 to 2.15). This region corresponds approximately to the HPD motif within all of the J domains, the region of Hsp40s known to interact with Hsp70s. A similar peak is observed with the substrate binding domains of Tcj2, Tcj3 and Tcj4 (Figures 2.14 to 2.16, residues 150-200). These regions score high on the antigenic index, as they are generally charged or hydrophilic, surface exposed and flexible. The characteristics making the potential interactions of the Tcj proteins with Hsp70s and substrates possible (that is charge/hydrophilic nature, flexibility and surface exposure) also make these regions potentially antigenic.

2.4 CONCLUSION

The bioinformatic analysis of TcHsp70 has revealed the presence of all the conserved Hsp70 domains required for chaperone function. Major differences in sequence were noted in the region of TcHsp70, which may influence its' binding to TcSTI1. Slight amino acid deviations within the phosphate binding domains of TcHsp70 may suggest a difference in the ATPase activity of this protein.

The bioinformatic analyses of the four trypanosomal Hsp40s has classified them into two groups – Tcj2, Tcj3 and Tcj4 as Type I Hsp40s, and Tcj1 as a Type III Hsp40. Tcj6 was shown to be a Type II Hsp40. The Type I Hsp40s are the most likely candidates for co-chaperoning TcHsp70, although the deviations from the consensus sequence in the cysteine repeat domains of Tcj2 and Tcj4 may influence their substrate binding capabilities. Tcj1 is unique in that it is the smallest Hsp40 recorded to date and thus may fulfill a novel partnership with TcHsp70. Composite bioinformatic analysis has identified sequences in the trypanosomal Hsps that may be potential candidates for peptide-directed antibody synthesis.

CHAPTER 3

OVER-PRODUCTION AND PURIFICATION OF TRYPANOSOMAL HEAT SHOCK PROTEINS

3.1 INTRODUCTION

3.1.1 Over-Production of Recombinant Proteins in Heterologous Systems

The laboratory analysis of a protein *in vitro* often requires high concentrations of pure product, which older 'conventional' methods of purification often cannot provide. Heterologous production systems offer a solution to this problem. Rather than purifying a protein from its source, which may potentially be rare, pathogenic or expensive, the protein can be produced specifically, relatively inexpensively and often in high concentrations by harnessing the protein production machinery of well-characterised organisms, such as *E. coli* or *Saccharomyces cerevisiae*. The gene encoding the target protein is introduced to the cell as part of an expression vector. The basic requirements for mRNA transcription and protein translation, specific to prokaryotic or eukaryotic expression systems, along with a selectable marker are all encoded on the plasmid (Balbas, 2001).

The most common problems arising are due to the expression of eukaryotic proteins in a prokaryotic host cell, such as problems with protein folding and protease degradation, especially if the target protein is produced at high concentrations. Certain target genes may contain rare codons, which could lead to truncated translation products or unstable proteins in *E. coli*. These problems can be reduced by manipulating growth conditions of the host cell to lower the levels of protein expression, introducing plasmids encoding tRNAs that recognise rare codons, modifying the gene of interest to the codon usage of the particular host cell or using protease deficient host cells.

3.1.2 Production of Target Proteins as Fusion Proteins for Efficient Purification

Although heterologous systems can be employed to efficiently over-express target sequences, the over-produced proteins often still need to be purified. Expressing proteins as fusions to an affinity tag simplifies purification, as during affinity purification the protein of interest is purified along with the affinity tag, which may later be removed using a protease cleavage site engineered between the protein and the tag. Commonly used fusion techniques are the GST (Glutathione-*S*-Transferase) (Smith and Johnson, 1988) and 6xHis (hexahistidine) tag fusion systems (Porath et al.,1975).

The GST-fusion system uses a GST protein as a fusion tag, engineered in frame with the target protein, for affinity purification of large quantities of pure proteins using glutathione covalently linked to a resin, such as sepharose. However, GST is a relatively large tag (26 kDa) that dimerises under normal conditions and if not cleaved may alter the structure or activity of the target protein. The 6xHis affinity tag is a stretch of six His residues, which can be engineered at either the N- or C- terminus of the target protein. This tag is useful as its small size (0.84kDa) does not generally interfere with the structure or function of the target protein and so it does not necessarily have to be cleaved. This tag is also useful, as it allows purification of fusion proteins under native or denaturing conditions, thus making it adjustable to the particular properties of the target protein

3.1.3 Production of 6xHis-Tag Fusion Proteins from pET Vector Constructs

The trypanosomal Hsps analysed in this report were all expressed from pET vector (Novagen, USA) constructs as either N- or C- terminal 6xHis fusion proteins. We examined the heterologous production of TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 as 6xHis fusion proteins and attempted to purify the proteins in their native states for use in functional and biochemical assays.

3.2 EXPERIMENTAL PROCEDURE

3.2.1 Reagents and Chemicals

All reagents used in the over-production and purification of the trypanosomal Hsps are detailed in Appendices (Section 6.1).

3.2.2 Over-Production of Trypanosomal Heat Shock Proteins

3.2.2.1 Constructs and Host Cells for Over-Production and Purification of Trypanosomal Hsps

Constructs containing coding regions of TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4 were a kind gift of Dr D.M. Engman from Northwestern University Medical School, Chicago. Trypanosomal heat shock coding sequences from an epimastigote cDNA library were provided as pET

vector (Novagen, USA) expression constructs, incorporating either a C-terminal or N-terminal 6xHis tag for nickel affinity purification. The TcHsp70 coding sequence was inserted in frame with an N-terminal 6xHis tag and an ampicillin resistance marker in the pET14b vector. The coding sequences for Tcj2 and Tcj4 were inserted into the pET28a vector, incorporating an N-terminal 6xHis tag and a kanamycin resistance marker. Tcj1 and Tcj3 coding sequences were inserted into pET23b vector, incorporating a C-terminal 6xHis tag and an ampicillin resistance marker. The host cells for the expression of His-TcHsp70, His-Tcj2 and Tcj3-His were *E. coli* BL21(DE3) [genotype *hsdS gal* (*λcIts* 857 *ind*1 *S*am7 *nin*5 *lac*UV5-*T7* gene *1*)] and the host cells for the expression of Tcj1-His and His-Tcj4 were *E. coli* BL21(DE3)(pLysS).

3.2.2.2 Transformation of Competent Host Cells with Expression Constructs

E. coli BL21(DE3) cells were treated to make them competent to take up plasmid DNA using a modification of the calcium chloride method (Dagert and Ehrlich, 1979). 50 mL of fresh 2xYT broth (1.6% tryptone, 1% yeast extract; 0.5% NaCl) was inoculated with 0.25 mL of *E. coli* BL21(DE3) from an overnight culture. The culture was grown at 37°C with shaking until an A_{600} of 0.3 to 0.6 was reached, whereupon the cells were harvested by centrifugation (5000x *g* for 5 minutes at 4°C) and the cell pellet resuspended in one culture volume (50 mL) of ice cold 0.1 M $MgCl_2$. After 1 minute incubation on ice, the cells were harvested by centrifugation as before and resuspended in half the culture volume (25 mL) of ice cold 0.1 M $CaCl_2$. Cells were held on ice for 2 hours, collected by centrifugation and the final pellet resuspended in one-tenth the culture volume (5 mL) of ice cold 0.1 M $CaCl_2$. Aliquots (150 µL) were mixed with equal volume (150 µL) of sterile 30% glycerol and stored at –70°C until required. The procedure was followed for the preparation of *E. coli* BL21(DE3)(pLysS) cells, with the exception that the 2xYT broth was supplemented with chloramphenicol (35 µg/ mL) to maintain the pLysS plasmid.

An aliquot of each dsDNA plasmid solution (~200 ng) was added to competent host cells (100 µL; specific to each plasmid, see above) and the mixture incubated on ice for 20 minutes. Cells were heat shocked at 42°C for 2 minutes, after which prewarmed (37°C) 2xYT broth (900 µL) was added. Cells were incubated at 37°C for 1 hour to allow expression of the antibiotic resistance gene (specific to each plasmid, see above). A portion of the cell suspension (100 µL) was removed and plated on to 2xYT agar containing the respective

antibiotic (ampicillin concentration –100 μg/mL; kanamycin concentration–50 μg/mL; chloramphenicol concentration–35 μg/mL). The remaining culture (900 μL) was harvested by centrifugation at 12000x *g*, a portion (800 μL) of the supernatant discarded and the pellet resuspended in the remaining volume (100 μL). This concentrated cell suspension was plated on to 2xYT agar plates containing antibiotics as stated previously. Plates were inverted and incubated overnight at 37°C. Transformants, namely those capable of growth in the presence of the antibiotic, were picked, streaked on to fresh 2xYT agar containing the correct antibiotic and analysed to confirm the identity of the plasmid. Control reaction containing sterile water instead of DNA was used to detect contamination and was treated using the exact conditions as the other transformation reactions.

3.2.2.3 Confirmation of Transformant Identity prior to Protein Production

Transformants were analysed to confirm the identity of the plasmids using a small-scale alkaline lysis plasmid extraction method and diagnostic restriction enzyme digests (Birnboim and Doly, 1979). A sample (2 mL) of overnight culture of the respective antibiotic-resistant transformant was centrifuged at 12000x *g* for 1 minute. The cell pellet was resuspended in Solution I (100 μL; 25 mM Tris-HCl, pH 8.0; 50 mM glucose; 10 mM EDTA) and incubated at room temperature for 5 minutes. Solution II was added (200 μL; 0.2 M NaOH; 1% SDS), the mixture incubated on ice for 2 minutes, and then Solution III (150 μL; 3 M potassium acetate, pH 5.0) was added. The mixture was mixed by inversion, incubated on ice for 5 minutes and centrifuged at 12000x *g* for 10 minutes. The supernatant was discarded and 1 volume (450 μL) of isopropanol added to the pellet. The mixture was incubated at room temperature for 5 minutes before the pellet was collected by centrifugation (12000x *g* for 30 min at 4°C) and air-dried. The pellet was resuspended in TE buffer (400 μL; 10 mM Tris-HCl, pH 7.5; 1 mM EDTA) and 1/10 volume of 3 M sodium acetate (40 μL) was added to the pellet, followed by 2 volumes (880 μL) of ice-cold absolute ethanol. The mixture was held at –20°C for 30 minutes and centrifuged at 12000x *g* for 30 minutes at 4°C. The resultant pellet was washed in ice cold 70% ethanol, centrifuged (12000x *g* for 1 min at 4°C), air dried and resuspended in sterile water (25 μL).

Each of the plasmid preparations was subjected to a series of diagnostic digests that were analysed by 0.8% agarose gel electrophoresis (AGE) in TBE buffer (45 mM Tris-Borate, pH 8.0; 1 mM EDTA) for 1 hour at 100 V and visualised under ultraviolet (UV) light.

3.2.2.4 Over-Production of 6xHis-tagged Trypanosomal Hsps in *E. coli* Host Cells

The production of all 6xHis-tagged trypanosomal heat shock proteins in *E. coli* was induced and monitored as follows. 10 mL of overnight culture of the respective transformant was used to inoculate 200 mL of sterile 2xYT broth containing the respective antibiotic (ampicillin 100 μg/mL; kanamycin 25 μg/mL; chloramphenicol 35 μg/mL). This culture was grown (37°C with shaking) to an A_{600} of 0.6-0.8 before protein production was induced by addition of IPTG (1 mM). Samples were collected prior to induction and at hourly intervals thereafter and analysed by SDS-PAGE (Laemmli, 1970). All samples were resuspended to equivalent A_{600}, treated with sample treatment buffer (95°C, 5 min; 50 mM Tris-HCl, pH 6.8; 2% SDS; 0.05% β-mercaptoethanol; 0.2% bromophenol blue) and analysed by discontinuous SDS-PAGE (12% acrylamide resolving gel; 4% acrylamide stacking gel) at 200 V for 45 min (Running buffer: 25 mM Tris-HCl; 250 mM Glycine, pH 8.0; 0.1% SDS). Proteins were visualised by staining with Coomassie Brilliant Blue (0.25% Coomassie G250 in 40% methanol, 7% glacial acetic acid).

Western analysis and chemiluminescence-based immunodetection was used to detect the presence of the 6xHis tag on the over-produced proteins. Discontinuous SDS-PAGE was conducted to resolve the proteins prior to electrophoretic transfer to nitrocellulose membrane using 100 V for 1 hour (Transfer Buffer: 25 mM Tris-HCl; 192 mM Glycine, pH8.3; 20% Methanol). Protein transfer on to the membrane was confirmed by Ponceau staining (0.5% Ponceau S in 1% glacial acetic acid). After destaining with distilled water, the membrane was blocked in 5% fat free milk powder in Tris Buffered Saline (TBS; 50 mM Tris-HCl, pH 7.5; 150 mM NaCl) for 1 hour at room temperature or overnight at 4°C. The membrane was incubated with mouse anti-6xHis antibody (primary antibody, Amersham Pharmacia Biotech, USA; 1 in 5 000 dilution in 5% fat free milk powder in TBS) at room temperature for 1 hour and then washed three times with TBS containing 0.1% Tween-20 (TBST; 50 mM Tris-HCl, pH 7.5; 150 mM NaCl, 0.1% Tween-20). The secondary antibody (Boehringer Mannheim, Germany; anti-mouse IgG-peroxidase labelled [POD]/anti-rabbit IgG-POD, 1 in 5 000 dilution in 5% fat free milk powder in TBS) was added to the membrane and incubated at room temperature for 45 min. The membrane was washed three times with TBST. The 6xHis tag signal was detected after incubation of the membrane with the chemiluminescent substrate (luminol). The film was placed in a HyperFilm Developing Cassette (Agfa, USA) wrapped in clingflim and high performance autoradiography film (Agfa, USA) exposed to the membrane

and hence to light emitted due to the oxidation of luminol by the peroxidase-conjugated secondary antibody. Exposure time of the film was varied from 1 min to 60 min depending on signal intensity. Subsequent to exposure to the membrane, the film was placed in developer (Agfa, USA) for between 1 and 3 minutes, the reaction stopped by passing the film through stop solution (2% glacial acetic acid) and fixed in fixer solution (Agfa, USA). After rinsing in water, the film was dried and the signal pattern recorded.

3.2.3 Purification of Trypanosomal Heat Shock Proteins

3.2.3.1 Preparation of Cleared Lysates for Ni-affinity Purification.

Sterile 2xYT broth (200 mL) containing the respective antibiotic (ampicillin 100 μg/mL, kanamycin 50 μg/mL or chloramphenicol 35 μg/mL) was inoculated with 10 mL overnight culture of *E. coli* BL21 (DE3) or *E. coli* BL21 (DE3) (pLysS) containing constructs encoding the 6xHis-tagged trypanosomal Hsps and grown (37°C with shaking) to A_{600} of 0.6-0.8. Protein production was induced by addition of IPTG to a final concentration of 1 mM. After 5 hours at 37°C with shaking, the culture was harvested by centrifugation (5000x *g* at 4°C for 15 min). For purification of His-TcHsp70, Tcj1-His, Tcj3-His and His-Tcj4, the cell pellet was resuspended in Denaturing Lysis Buffer (5 mL; 40 mM Tris, pH 8.0; 8 M urea; 100 mM NaCl; 10 mM Imidazole). For the purification of His-Tcj2, the cell pellet was resuspended in Native Lysis Buffer (5 mL; 40 mM Tris, pH8.0; 100 mM NaCl; 10 mM Imidazole) and cell lysis achieved by lysozyme (100 μg/mL) treatment for 1 hour and sonication (30 sec bursts for 5 min) at 4°C. In both cases insoluble cell debris was removed by centrifugation (10 000x *g* for 15 min at 4°C) and the supernatant (cleared lysate) collected for nickel affinity (Ni-affinity) purification of the 6xHis-tagged proteins.

3.2.3.2 Batch Ni-affinity Purification of His-TcHsp70, Tcj1-His, Tcj3-His and His-Tcj4.

Cleared lysate was added to Ni-charged chelating sepharose resin (1 mL). The mixture was gently agitated at room temperature for at least 1 hour, after which the resin was collected by centrifugation (1000x *g* for 1 min). The resin was washed once or twice with Denaturing Lysis Buffer (40 mM Tris-HCl, pH 8.0; 8 M urea; 100 mM NaCl; 10 mM Imidazole) and two or three times with Native Wash Buffer 1 (40 mM Tris-HCl, pH 8.0; 100 mM NaCl; 100 mM

Imidazole). Proteins bound to the resin were eluted three times by addition of Native Elution Buffer (1 mL; 40 mM Tris-HCl, pH 7.5; 100 mM NaCl; 1 M Imidazole). Each time the resin was collected by centrifugation (1000x *g* for 1 min at room temperature) and all fractions collected were analysed by discontinuous SDS-PAGE (described Section 3.2.2.4) and proteins visualised by Coomassie Brilliant Blue staining. Protein concentration of the elutions was determined by Bradfords assay (Bradford, 1976). In all cases, a protein sample (10 µL) was incubated with Bradford's Reagent (200 µL) at room temperature in a flat-bottomed 96 well microtitre plate for 5 min, before A_{595} was measured. A standard curve of protein concentration (40 µg/mL – 160 µg/mL) was constructed using bovine serum albumin (BSA) and used to determine the protein concentration in the samples. Aliquots of elutions (50 µL) containing sufficient pure proteins as determined by SDS-PAGE and Bradford's Assay were stored at –70°C for later use.

3.2.3.3 Column Ni-affinity Purification of His-Tcj2

Cleared lysate was added to Ni-charged chelating sepharose resin (1 mL). The mixture was gently agitated at room temperature for at least 1 hour, after which the resin was loaded in a 10 mL syringe, plugged with glass wool. The flow through was collected and the resin washed twice with Native Wash Buffer 1 (40 mM Tris-HCl, pH 8.0; 100 mM NaCl; 100 mM Imidazole) and twice with Native Wash Buffer 2 (40 mM Tris-HCl, pH 8.0; 200 mM NaCl; 200 mM Imidazole). Proteins bound to the resin were eluted with three column volumes of Native Elution Buffer (1 mL; 40 mM Tris-HCl, pH 7.5; 100 mM NaCl; 1 M Imidazole). All fractions were collected and analysed by discontinuous SDS-PAGE (described Section 3.2.2.4) and proteins visualised by Coomassie Brilliant Blue staining. Aliquots (50 µL) of elutions containing sufficient pure proteins as determined by SDS-PAGE and Bradford's Assay (described Section 3.2.2.4) were stored at –70°C for later use.

3.2.3.4 Testing of Peptide-directed Antibody to C-terminal Sequence in Tcj2

The synthesis of the antibody to the predicted continuous epitope ^{N}-CHLSETNIDKEKEAK-^{C} of Tcj2 (Section 2.2.5 and Section 2.3.7) was performed by Dr Nastainczyk, in the laboratory of Professor R. Zimmerman (Universität des Saarlandes, Germany). The peptide was synthesised and coupled via the N-terminal cysteine to haemocyanin by a bifunctional linker and a rabbit immunised with this peptide-hapten complex to generate antibodies. Serum was

collected prior to immunisation and at weekly intervals for 8 weeks after immunisation. Serum was used directly as primary antibody at a 1 in 10000 dilution for Western analysis and chemiluminescence-based immunodetection (Section 3.2.1.4) of Tcj2.

3.2.4 Spectrofluorometric Analysis of Purified 6xHis-Tagged Trypanosomal Proteins

Successfully purified trypanosomal Hsps (His-TcHsp70, Tcj1-His and His-Tcj2) were analysed for the presence of large and small soluble aggregates using spectrofluorimetry prior to biochemical assays. Sample elutions of His-TcHsp70, His-Tcj2 and Tcj1-His were diluted 1 in 100 in water to a final volume of 2 mL. The light scattering ability of these preparations was analysed using a fluorescence spectrophotometer. The samples were excited at 360 nm (detection of small, soluble aggregates) and 500 nm (detection of large aggregates) and their emission over the wavelength range of 300 nm to 600 nm recorded. A denatured / aggregated control was prepared by boiling BSA (10 μg/mL) for 15 min. A native BSA sample (10 μg/mL) was included as a negative control.

3.3 RESULTS AND DISCUSSION

3.3.1 Over-Production and Purification of His-TcHsp70

Previous work on the biochemical characterisation of TcHsp70 was reported by Olson et al. (1994). Of interest to us was to over-produce and purify TcHsp70 as a 6xHis-fusion from the pET14b/TcHsp70 construct for use in biochemical analysis. The pET14b/TcHsp70 plasmid was digested with a series of diagnostic digests to release the TcHsp70 insert and confirm the identity of the plasmid (Figure 3.1). The TcHsp70 coding sequence was inserted into the *Bam*HI site of pET14b, but digestion with this enzyme did not release the fragment (data not shown). This is most likely an artefact of the cloning procedure used to generate the plasmid. Alternative diagnostic digests were chosen after examination of the plasmid map. The pET14b/TcHsp70 plasmid was linearised by digests with *Eco*RI and *Nde*I (Figure 3.1 Lanes 4 and 5). Digestion with *Pst*I yielded two fragments of 2674 bp and 4361 bp (Figure 3.1 Lane 3), while double digest with *Eco*RI and *Nde*I produced DNA fragments of 2888 bp and 4147 bp (Figure 3.1 Lane 6).

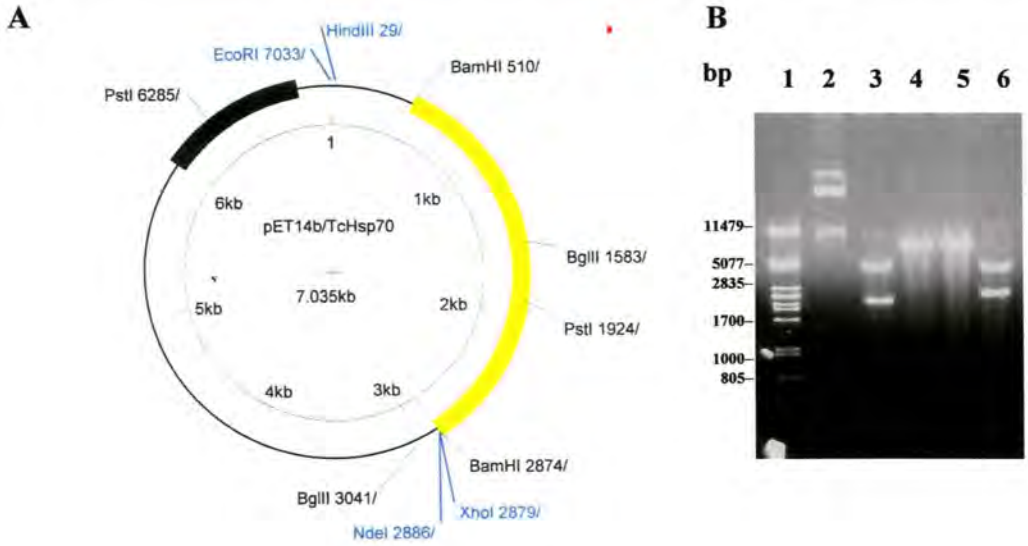


Figure 3.1: Confirmation of the Identity of the pET14b/TcHsp70 Plasmid

(A) Plasmid map of pET14b/TcHsp70 showing diagnostic restriction sites. The TcHsp70 coding region is shown in yellow and the ampicillin resistance coding region is shown in black.
(B) Confirmation of the identity of the pET14b/TcHsp70 plasmid by restriction digests. Lane 1 – *Pst*I digest of Lambda DNA (sizes shown in bp); Lane 2 – Uncut pET14b/TcHsp70; Lane 3 – *Pst*I digest yielding fragments of sizes 2674 and 4361bp; Lane 4 – pET14b/TcHsp70 linearised by digestion with *Eco*RI (7035bp); Lane 5 – pET14b/TcHsp70 linearised by digestion with *Nde*I (7035bp); Lane 6 – *Eco*RI / *Nde*I double digest of pET14b/TcHsp70 resulting in two bands of 2888bp and 4147bp.

The over-production of His-TcHsp70 in *E. coli* BL21(DE3) [pET14b/TcHsp70], and detection of the N-terminal 6xHis fusion tag is shown in Figure 3.2. Whole cell extracts taken at hourly intervals after the induction of the production of His-TcHsp70 were analysed by discontinuous SDS-PAGE and proteins visualised by Coomassie staining (Figure 4.2[A]). A gradual increase in the band at 71kDa corresponded to the subunit molecular mass of His-TcHsp70 and indicated the over-production of the protein. His-TcHsp70 was produced at high levels, with the maximum level of induction being achieved at 4 hours post addition of IPTG (1 mM) and maintained for 20 hours (Figure 3.2[A] Lanes 6 and 8). Low basal levels of production of His-TcHsp70 were detected in the uninduced *E. coli* BL21(DE3) [pET14b/TcHsp70] sample (Figure 3.2[A] Lane 2).

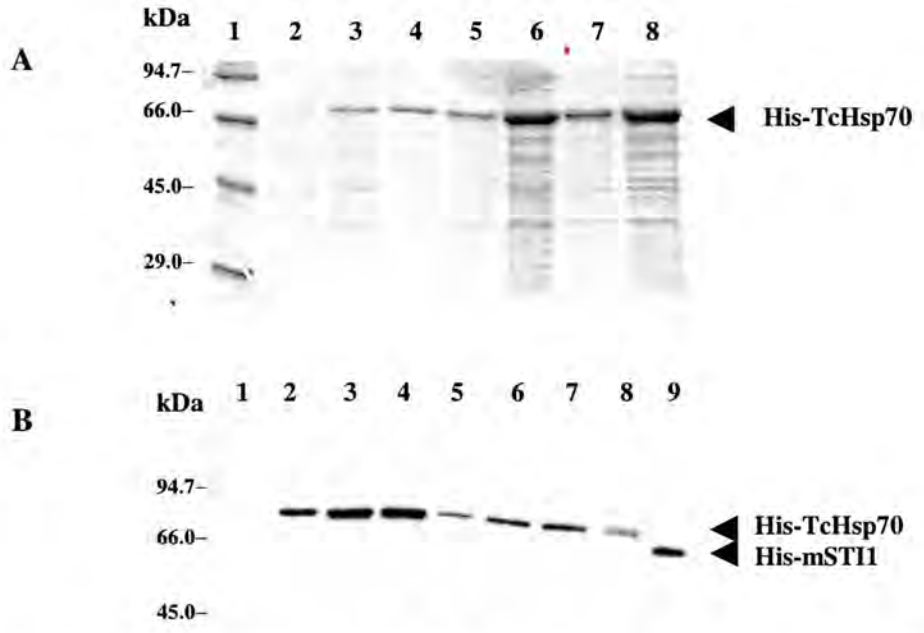


Figure 3.2: Over-production of His-TcHsp70 in *E.coli* BL21(DE3)[pET14b/TcHsp70]

(A) SDS-PAGE analysis of production of His-TcHsp70 in*E. coli* BL21(DE3)[pET14b/TcHsp70]. Lane 1 – broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E.coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract; Lanes 3 to 8 – *E. coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by the addition of 1mM IPTG. Over-production of His-TcHsp70 (71kDa) indicated (arrow).
(B) Detection of 6xHis tag on His-TcHsp70 by Western analysis and chemiluminesence-based immunodetection. Lane 1 – *E.coli* BL21(DE3) whole cell extract (negative control); Lane 2 – uninduced *E.coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract; Lanes 3 to 8 respectively – *E. coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by addition of 1mM IPTG; Lane 9 – His-mSTI1 (positive control; 63kDa).

Figure 3.2(B) illustrates the detection of the 6xHis tag on His-TcHsp70. While the levels of the protein were seen to increase with time (Figure 3.2[A]), there appeared to be a gradual reduction in the amount of actual 6xHis-tagged protein, possibly due to cleavage of the tag by proteases inherent to *E. coli* BL21(DE3) or shielding of the 6xHis-tag due to aggregation that is resistant to SDS-based denaturation.

His-TcHsp70 was successfully purified using a combination of denaturing/native Ni-affinity chromatography (Figure 3.3). The cleared lysate was prepared, bound to the resin and preliminary washes carried out in the presence of 8 M urea (denaturing conditions). The resin was subsequently washed with native buffers, to remove the urea, and allow the proteins to renaturate, before elution by competitive binding of high concentrations of imidazole.

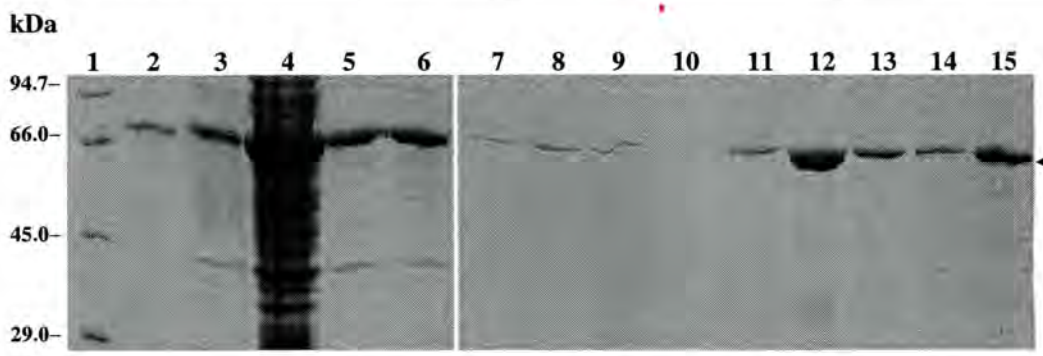


Figure 3.3: Purification of His-TcHsp70 by Nickel Affinity Chromatography

His-TcHsp70 was successfully purified from *E. coli* BL21(DE3)[pET14b/TcHsp70] cleared lysate by batch Ni-affinity chromatography using a combination of denaturing and native purification procedures.
Lane 1 – Broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E. coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract; Lane 3 – induced *E. coli* BL21(DE3)[pET14b/TcHsp70] whole cell extract; Lane 4 – induced *E. coli* BL21(DE3)[pET14b/TcHsp70] cleared lysate; Lane 5 – flow through; Lanes 6 and 7 – denaturing washes; Lanes 8 and 9 – native washes with Native Lysis Buffer; Lanes 10 and 11 – native washes with Native Wash Buffer 1; Lanes 12, 13 and 14 – elutions 1, 2 and 3 respectively; Lane 15 – Ni-affinity sepharose beads.

The resin was saturated with His-TcHsp70, as indicated by the levels of unbound His-TcHsp70 in the flow through fraction (Figure 3.3 Lane 5). A portion of His-TcHsp70 was still bound to the resin after competitive elution in the presence of 1 M imidazole, potentially due to non-specific associations arising due to misfolded protein (Figure 3.3 Lane 15). Sufficient concentrations of His-TcHsp70 for biochemical assays were purified by this method (typical concentration range of 300 μg/ mL to 500 μg/ mL).

Native Ni-affinity purification of His-TcHsp70 was attempted, but the procedure was not efficient for the purification of His-TcHsp70 (data not shown) as only low concentrations of His-TcHsp70 were purified (data not shown). The low yields were attributed to lack of binding to the Ni-charged resin, possibly due to the masking of the 6xHis tag during the folding of His-TcHsp70 to its tertiary structure. This would explain why the 6xHis-tag was detectable on a denaturing SDS-PAGE and Western analysis, but non-functional under the native purification conditions. This proposal was supported by the elution of some His-TcHsp70 during the native washes (Figure 3.3 Lanes 8 to 11). As the resin was washed with native buffer and the urea removed, His-TcHsp70 could potentially begin to refold to its correct conformation, potentially masking the 6xHis-tag, and releasing the protein from the resin.

3.3.2 Over-Production and Purification of Tcj1-His

Tcj1-His is a Type III J protein, according to the classification of Cheetham and Caplan (1998), containing only the J domain in common with *E. coli* DnaJ. It is the smallest Hsp40 (~34kDa) identified and may fulfil a role novel to the parasitic system. Tcj1-His was encoded on the plasmid pET23b/Tcj1. The confirmation of this construct by restriction enzyme analysis is shown in Figure 3.4.

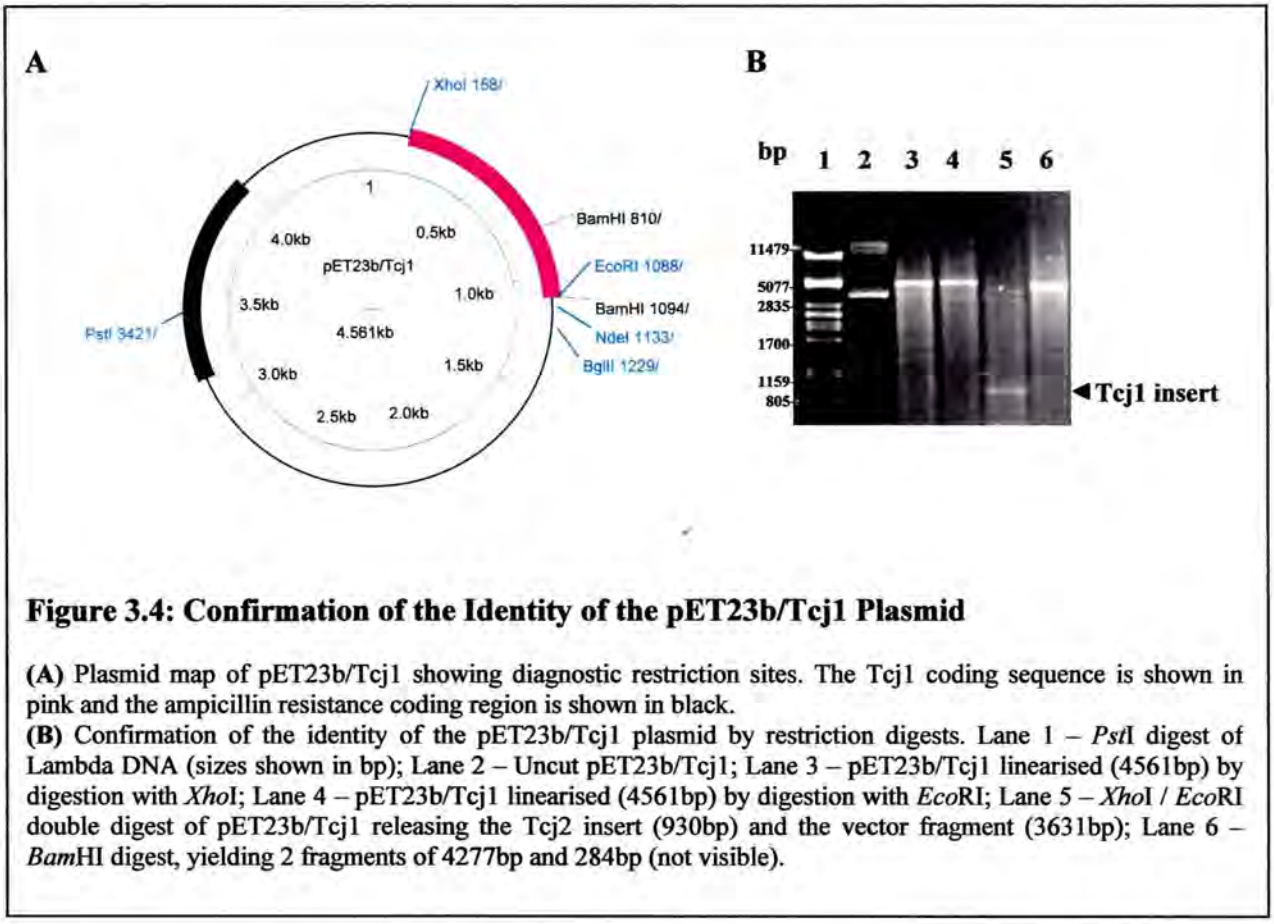


Figure 3.4: Confirmation of the Identity of the pET23b/Tcj1 Plasmid

(A) Plasmid map of pET23b/Tcj1 showing diagnostic restriction sites. The Tcj1 coding sequence is shown in pink and the ampicillin resistance coding region is shown in black.
(B) Confirmation of the identity of the pET23b/Tcj1 plasmid by restriction digests. Lane 1 – *Pst*I digest of Lambda DNA (sizes shown in bp); Lane 2 – Uncut pET23b/Tcj1; Lane 3 – pET23b/Tcj1 linearised (4561bp) by digestion with *Xho*I; Lane 4 – pET23b/Tcj1 linearised (4561bp) by digestion with *Eco*RI; Lane 5 – *Xho*I / *Eco*RI double digest of pET23b/Tcj1 releasing the Tcj2 insert (930bp) and the vector fragment (3631bp); Lane 6 – *Bam*HI digest, yielding 2 fragments of 4277bp and 284bp (not visible).

The Tcj1 coding sequence was inserted in frame with a C-terminal 6xHis-tag into the *Eco*RI and *Xho*I sites of pET23b. Restriction digests with each of these enzymes linearises pET23b/Tcj1 (4561bp) (Figure 3.4 Lanes 3 and 4), while a *Eco*RI / *Xho*I digest releases the Tcj1 coding region (930bp) and the pET23b vector fragment (3631bp) (Figure 3.4 Lane 5). The plasmid was digested with *Bam*HI to yield two fragments (284bp and 4277bp) (Figure 3.4 Lane 5).

The production of Tcj1-His in *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] was monitored over time, as described for His-TcHsp70. The whole cell extract samples were analysed by SDS-PAGE and Coomassie staining [Figure 3.5(A)] and the presence of the 6xHis-tag detected by Western analysis [Figure 3.5(B)].

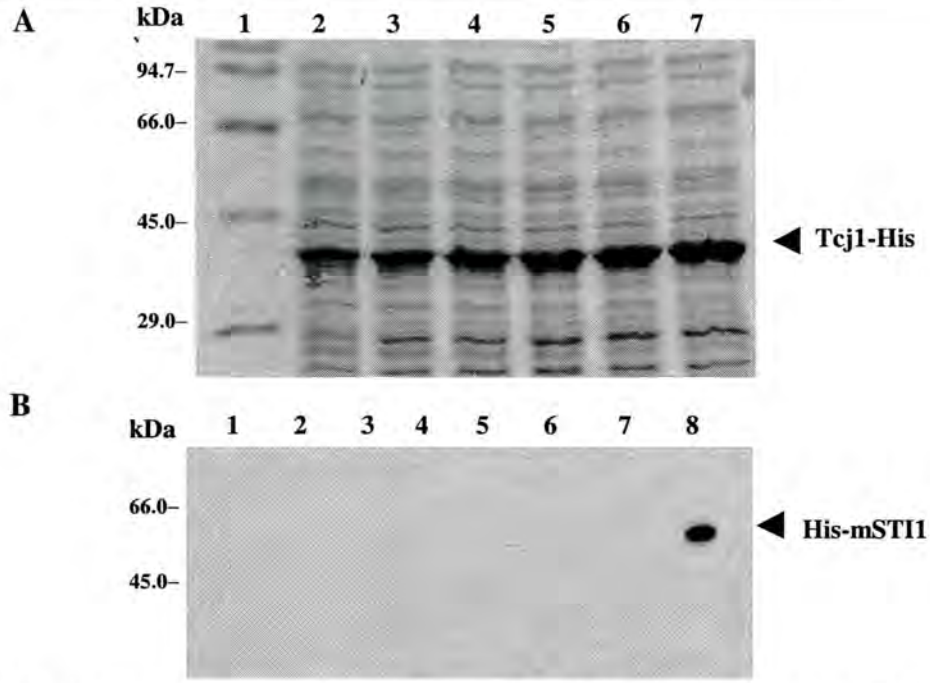


Figure 3.5: Over-production of Tcj1-His in *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj1]

(A) SDS-PAGE analysis of production of Tcj1-His in *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1]. Lane 1 – broad range SDS-PAGE molecular weight markers (sizes shown in kDa); Lane 2 – induced *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract; Lanes 3 to 7 respectively – *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract samples taken 1, 2, 3, 4, and 5 hours after induction by the addition of 1mM IPTG. Overproduction of Tcj1-His (~34kDa) is indicated (arrow).
(B) Detection of 6xHis tag on Tcj1-His by Western analysis and chemiluminesence-based immunodetection. Lane 1 – Uniduced *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract; Lanes 2 to 6 respectively – *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract samples taken 1, 2, 3, 4, and 5 hours after induction by addition of 1mM IPTG;Lane 7 – *E.coli* BL21(DE3) whole cell extract (negative control) Lane 8 – His-mSTI1 (positive control; 63kDa).

Tcj1-His (subunit molecular mass of ~34kDa) was produced at high levels, even in *E. coli* BL21(DE3)(pLysS), a host cell which traditionally lowers expression levels. The basal levels of expression are so high, that there is little difference between the levels of Tcj1-His in the induced and uninduced whole cell extracts (Figure 3.5 [A] Lanes 2-7). The 6xHis tag on Tcj1-His was not detectable using this particular anti-6xHis antibody (Figure 3.5[B]). A similar situation was observed with Tcj3-His (described Section 3.3.4), suggesting that this could be a characteristic of the C-terminal 6xHis-tags. It could be possible in this case that the C-terminal 6xHis-tags were forming an alternative conformation that was masked such that they were not recognised by the anti-His antibody. Lack of detection of the 6xHis-tag could also be due to cleavage or loss of the tag after protein production.

The *E. coli* BL21(DE3)(pLysS)[Tcj1-His] cleared lysate was prepared and bound to the Ni-charged resin in the presence of 8 M urea (denaturing conditions). The resin was washed with denaturing buffer and then with native buffers, prior to elution under native conditions. The Ni-affinity column was saturated (Figure 3.6 Lane 5) and a portion of Tcj1-His was still bound to the resin after elutions (Figure 3.6 Lane 15). Tcj1-His was efficiently purified by denaturing/native Ni-affinity chromatography and the successful purification of Tcj1-His by Ni-affinity chromatography indicated that the 6xHis-tag had not been cleaved. The concentrations of the elutions were typically 500μg/mL (Figure 3.6 Lanes 12-14).

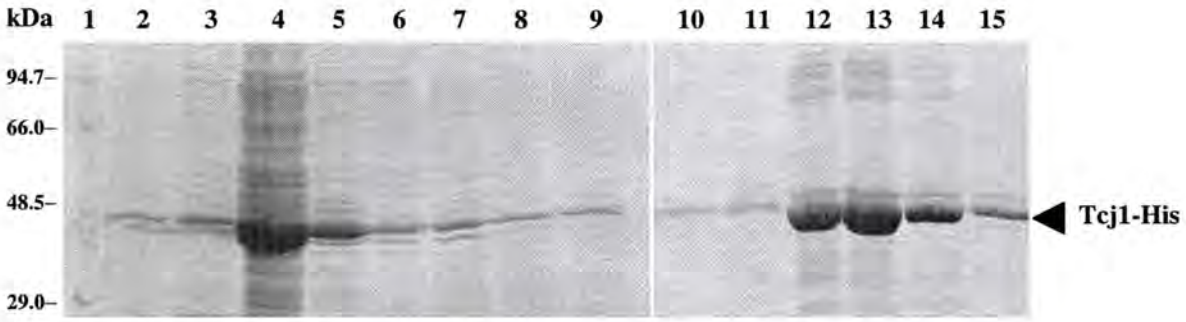


Figure 3.6: Purification of Tcj1-His by Nickel Affinity Chromatography

Tcj1-His was successfully purified from *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] cleared lysate by batch nickel affinity chromatography using a combination of denaturing and native purification procedures.
Lane 1 – Broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract; Lane 3 – induced *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract; Lane 4 – induced *E. coli* BL21(DE3)(pLysS)[pET23b/Tcj1] cleared lysate; Lane 5 – flow through; Lanes 6 and 7 – denaturing washes; Lanes 8 and 9 – native washes with Native Lysis Buffer; Lanes 10 and 11 – native washes with Native Wash Buffer 1; Lanes 12, 13 and 14 – elutions 1, 2 and 3 respectively; Lane 15 – Ni-affinity sepharose beads.

The denaturing/native combination Ni-affinity purification procedure was chosen for the purification of Tcj1-His as it was determined to be insoluble under native conditions (data not shown). This insolubility could have resulted from the extremely high concentrations of the protein produced. Native purification of Tcj1-His was not successful, as insoluble protein was detected even after the alteration of growth conditions, such as lower temperatures, or different induction parameters (data not shown).

3.3.3 Over-Production and Purification of His-Tcj2

Tcj2 is a Type I J protein, containing all three motifs in common with *E.coli* DnaJ, with a deviation from the consensus sequence in the final cysteine repeat region. Type I Hsp40s are generally the most likely co-chaperone for Hsp70s. In addition, Tcj2 is the only one of the four trypanosomal Hsp40s shown to be inducible by heat shock in *T. cruzi* (26°C to 37°C), which suggests it may be the partner protein for the inducible TcHsp70.

The pET28a/Tcj2 plasmid encodes the Tcj2 coding sequence fused to an N-terminal 6xHis-tag. The identity of the plasmid was confirmed by restriction analysis prior to protein studies (Figure 3.7). The pET28a/Tcj2 plasmid was linearised by digestion with *Pvu*I, *Xho*I and *Nde*I (6495bp) (Figure 3.7 Lanes 3, 4 and 5). The Tcj2 coding sequence (1207bp) was released by digestion of pET28a/Tcj2 with the enzymes used for construction of the plasmid, *Xho*I and *Nde*I (Figure 3.7 Lane 6).

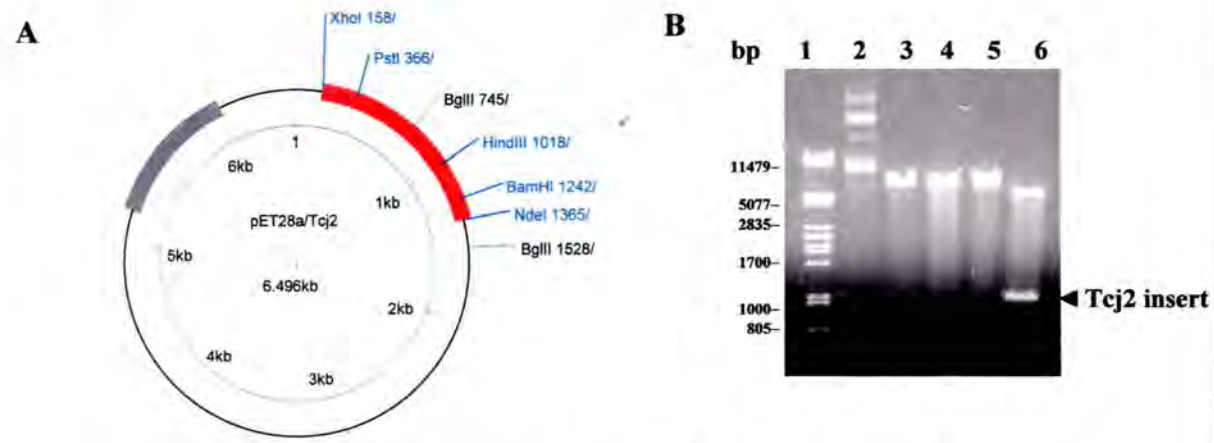


Figure 3.7: Confirmation of the Identity of the pET28a/Tcj2 Plasmid

(A) Plasmid map of pET28a/Tcj2 showing diagnostic restriction sites. The Tcj2 coding sequence is shown in red and the kanamycin resistance coding sequence is shown in dark grey.
(B) Confirmation of identity of pET28a/Tcj2 plasmid by restriction digests. Lane 1 – *Pst*I digest of Lambda DNA (sizes shown in bp); Lane 2 – Uncut pET28a/Tcj2; Lane 3 – *Pvu*I digest (6495bp); Lane 2 – pET28a/Tcj2 linearised by digestion with *Xho*I (6495bp); Lane 5 – pET28a/Tcj2 linearised by digestion with *Nde*I (6495bp); Lane 6 – *Xho*I/*Nde*I digest of pET28a/Tcj2 releasing the Tcj2 insert (1207bp) and vector fragment (5288bp).

The production of His-Tcj2 in *E. coli* BL21(DE3)[pET28b/Tcj2] at hourly intervals was monitored. Whole cell extract samples were prepared and analysed by SDS-PAGE and Coomassie staining. The presence of the 6xHis tag was confirmed by Western analysis (Figure 3.8 [A] and [B]).

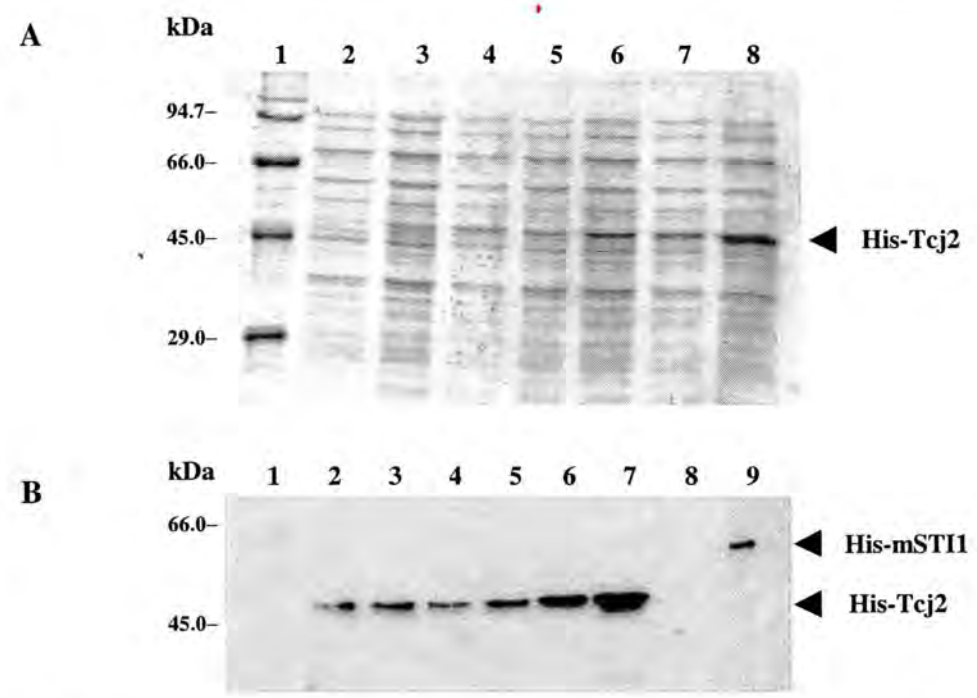


Figure 3.8: Over-production of His-Tcj2 in *E.coli* BL21(DE3)[pET28a/Tcj2]

(A) SDS-PAGE analysis of production of His-Tcj2 in *E.coli* BL21(DE3)[pET28a/Tcj2]. Lane 1 – broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lanes 3 to 8 respectively – *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by the addition of 1mM IPTG. Over-production of His-Tcj2 (~45kDa) is indicated (arrow).
(B) Detection of 6xHis tag on His-Tcj2 by Western analysis and chemiluminesence-based immunodetection. Lane 1 – Uniduced *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lanes 2 to 7 respectively – *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by addition of 1mM IPTG;Lane 8 – *E.coli* BL21(DE3) whole cell extract (negative control) Lane 9 – His-mSTI1 (positive control; 63kDa).

His-Tcj2 was produced as indicated by the increase in a protein band equivalent to the subunit molecular mass of ~45kDa [Figure 3.8 (A)]. The levels of production of His-Tcj2 were significantly lower than those of His-TcHsp70 and Tcj1-His. There was no detectable basal expression of His-Tcj2, as indicated by the SDS-PAGE and Western analysis (Figure 3.8 [A] Lane 2 and Figure 3.8 [B] Lane 1). The production of His-Tcj2 only occurs post induction and continues for 20 hours without any apparent degradation or decrease in the 6xHis-tagged species (Figure 3.8 [A] Lane 8 and Figure 3.8 [B] Lane 7). The levels of His-Tcj2 are greatest after induction for 20 hours, indicating that the protein was sufficiently stable to accumulate within the cells that had reached stationary phase.

The preparation of *E. coli* BL21(DE3)[pET28b/Tcj2] cleared lysate, binding to the Ni-charged resin, washing steps and elutions were all conducted under native conditions (Figure 3.9). The resin was saturated as indicated by the presence of free His-Tcj2 in the flow through (Figure 3.9 Lane 5). His-Tcj2 was successfully purified during native Ni-affinity purification. His-Tcj2 was found to be soluble and successfully purified by native Ni-affinity chromatography. The concentrations of the His-Tcj2 elutions were typically in the region of 200 μg/mL (Figure 3.9 Lanes 12-14).

Figure 3.9: Purification of His-Tcj2 by Nickel Affinity Chromatography

His-Tcj2 was successfully purified from *E.coli* BL21(DE3)[pET28a/Tcj2] using column native nickel affinity chromatography
Lane 1 – Broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lane 3 – induced *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lane 4 – *E.coli* BL21(DE3)[pET28a/Tcj2] cleared lysate; Lane 5 – flow through; Lanes 6 and 7 – native washes with Native Wash Buffer 1; Lanes 10 and 11 – native washes with Native Wash Buffer 2; Lanes 12, 13 and 14 – elutions 1, 2 and 3 respectively

4.3.4 Over-Production and Purification of Tcj3-His

Tcj3 is a Type I J protein, containing the J domain, G/F region and consensus cysteine repeat region in common with DnaJ (Cheetham and Caplan, 1998). The fact that it is a Type I protein, makes it a potential partner for TcHsp70, although it is not induced by heat shock in *T. cruzi*. Tcj3-His was expressed from the pET23b/Tcj3 plasmid as a C-terminal 6xHis-fusion protein. The pET23b/Tcj3 construct was confirmed by restriction digest analysis prior to protein studies (Figure 4.10). The pET23b/Tcj3 plasmid was linearised by single digests with *Xho*I and *Eco*RI (4810bp) (Figure 3.10 Lanes 4 and 5) and the coding region for Tcj3 released by a digest with both these enzymes (1179bp and vector fragment 3631bp) (Figure 3.10 Lane 6). The plasmid was digested by *Pst*I, yielding two fragments (1665bp and 3145bp) (Figure 3.10 Lane 3).

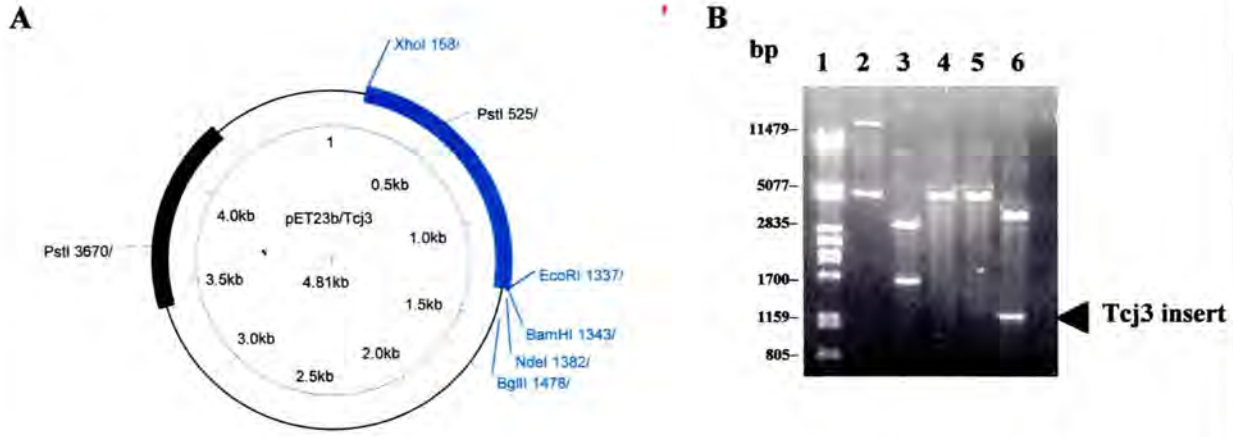


Figure 3.10: Confirmation of the Identity of the pET23b/Tcj3 Plasmid

(A) Plasmid map of pET23b/Tcj3 showing diagnostic restriction sites. The Tcj3 coding sequence is shown in blue and the ampicillin resistance coding sequence is shown in black.
(B) Confirmation of the identity of the pET23b/Tcj3 plasmid by restriction digests. Lane 1 – *Pst*I digest of Lambda DNA (sizes shown in bp); Lane 2 – Uncut pET23b/Tcj3; Lane 3 – *Pst*I digest yielding fragments of 3145bp and 1665bp; Lane 4 – pET23b/Tcj3 linearised by digestion with *Xho*I (4810bp); Lane 5 – pET23b/Tcj3 linearised by digestion with *Eco*RI (4810bp); Lane 6 – *Eco*RI/*Xho*I digest of pET23b/Tcj3 releasing the Tcj3 insert (1179bp) and the vector fragment (3631bp)

The over-production of Tcj3-His was monitored in *E. coli* BL21(DE3)[pET23b/Tcj3], with hourly samples after induction being analysed by SDS-PAGE and Coomassie staining. Tcj3-His was produced at levels similar to His-Tcj2 [Figure 3.11(A)]. A protein band corresponding to the subunit molecular weight of Tcj3-His (~45kDa) appeared to accumulate after 20 hours (Figure 3.11 [A] Lane 8). The 6xHis-tag on Tcj3-His was not detectable by Western analysis, using this particular anti-His antibody. This was interesting, as Tcj3-His carries a C-terminal 6xHis-tag, similar to Tcj1-His, which was also undetectable. The lack of detection of both of the C-terminal tag 6xHis-fusion proteins seems to support the proposal that these tags are not recognised by the antibody. It is possible that the 6xHis-tag in these cases are attaining alternate conformations or being hidden, which make them undetectable by the anti-His antibody. Tcj3-His could not be purified by either native, denaturing or denaturing/native combination Ni-affinity chromatography (data not shown). This suggests that the 6xHis-tag on Tcj3-His is either non-functional or being cleaved. Alternatively, there could be inefficient translation of the protein in *E.coli*, leading to a truncated protein and thus no 6xHis-tag. This would explain the lack of detection and purification of Tcj3-His.

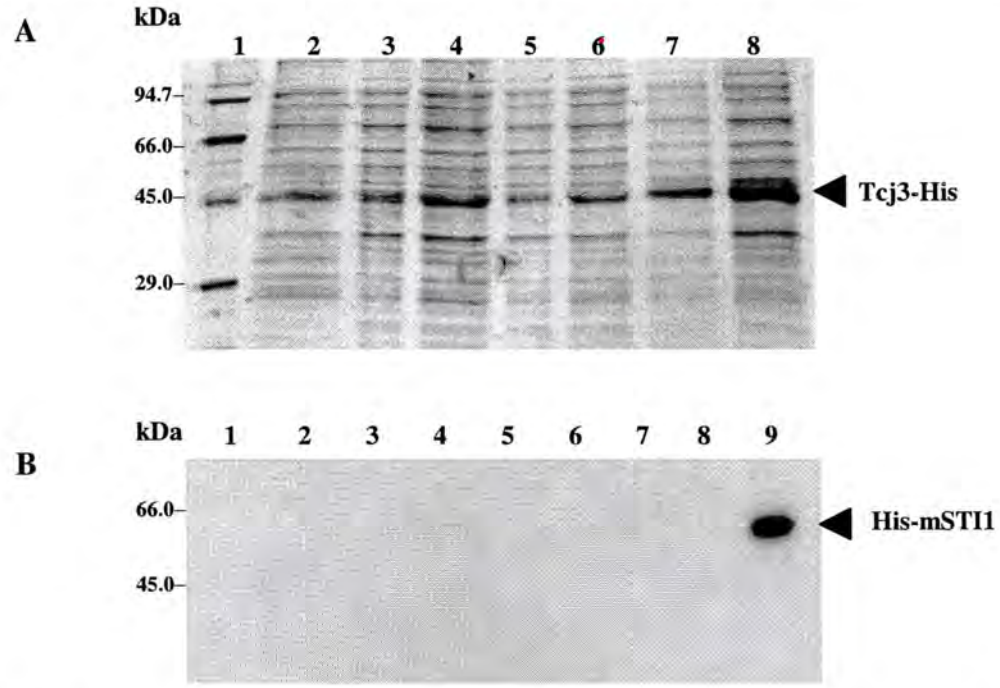


Figure 3.11: Over-production of Tcj3-His in *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3]

(A) SDS-PAGE analysis of production of Tcj3-His *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3]. Lane 1 – broad range SDS molecular mass markers (sizes shown in kDa); Lane 2 – uninduced *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3] whole cell extract; Lanes 3 to 8 respectively – *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by the addition of 1mM IPTG. Over-production of Tcj3-His (~45kDa) is indicated (arrow).
(B) Detection of 6xHis tag on His-Tcj2 by Western analysis and chemiluminesence-based immunodetection. Lane 1– uninduced *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3] whole cell extract; Lanes 2 to 7 respectively – *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj3] whole cell extract samples taken 1, 2, 3, 4, 5 and 20 hours after induction by addition of 1mM IPTG; Lane 8 – *E.coli* BL21(DE3) whole cell extract (negative control) Lane 9 – His-mSTI1 (positive control; 63kDa).

3.3.5 Over-Production and Purification of His-Tcj4

Tcj4, also a Type I J protein, is also a potential partner for TcHsp70. It has a deviation from the consensus sequence, different from that observed in Tcj2, in its final cysteine repeat suggesting a deviation in the structure of its substrate binding site from that of known structure. This does not mean that it may not interact with TcHsp70, but suggests that its substrates may be significantly different to those bound by Tcj2 and Tcj3. Tcj4 is not inducible by heat shock in *T. cruzi*.

His-Tcj4 was over-expressed from pET28a/Tcj4 as an N-terminal fusion protein in *E. coli* BL21(DE3)(pLysS). Restriction analysis was performed prior to the protein analysis, to confirm the identity of the pET28a/Tcj4 construct (Figure 3.12). pET28a/Tcj4 was linearised by single digests with *Xho*I, *Nde*I and *Bam*HI (Figure 3.12 Lanes 3, 4 and 6). A restriction digest with the restriction enzymes used for inserting the coding sequence for Tcj4, *Xho*I and *Nde*I, liberated the insert fragment (1344bp) from the vector fragment (5289bp) (Figure 3.12 Lane 5). In addition, the construct was digested with *Kpn*I that cuts within the Tcj4 coding region, producing two fragments of 433bp and 6200bp (Figure 3.12 Lane 7).

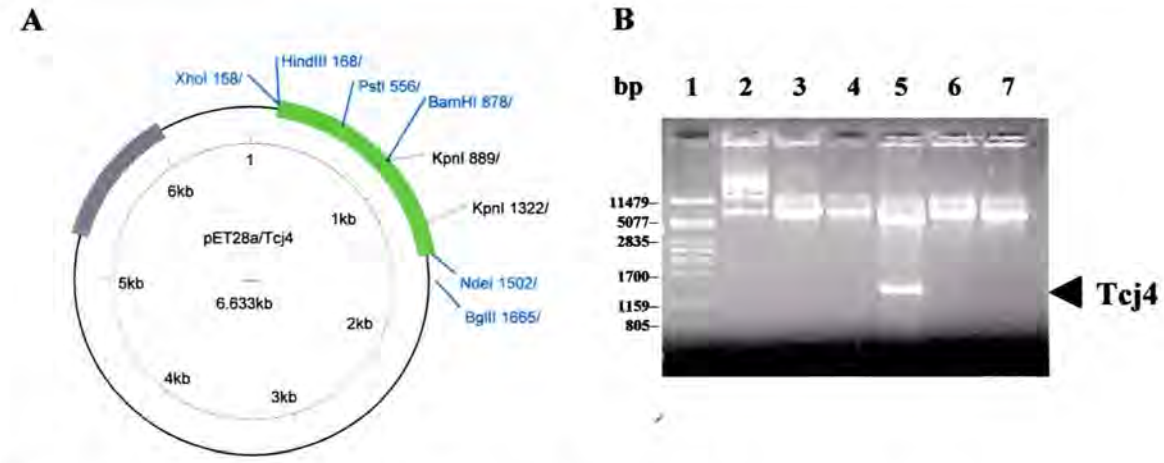


Figure 3.12: Confirmation of the Identity of the pET28a/Tcj4 Plasmid

(A) Plasmid map of pET28a/Tcj4 showing diagnostic restriction sites. The Tcj4 coding sequence is shown in green and the kanamycin resistance coding sequence is shown in dark grey.
(B) Confirmation of identity of pET28a/Tcj4 plasmid by restriction digests. Lane 1 – *Pst*I digest of Lambda DNA; Lane 2 – Uncut pET28a/Tcj4; Lane 3 – pET28a/Tcj4 linearised by digestion with *Xho*I (6633bp); Lane 4 – pET28a/Tcj4 linearised by digestion with *Nde*I (6633bp); Lane 5 – *Xho*I / *Nde*I double digest of pET28a/Tcj4 releasing the Tcj4 insert (1344bp) and vector fragment (5289bp); Lane 6 – *Bam*HI digest (6633bp); Lane 7 – *Kpn*I digest (6633bp).

The production of His-Tcj4 in *E. coli* BL21(DE3)(pLysS)[pET28a/Tcj4] was monitored at hourly intervals. The SDS-PAGE analysis of the timed induction profile of His-Tcj4 indicated no obvious increase in a band that would correspond to the subunit molecular mass of His-Tcj4 (~49kDa). Only a faint signal was detected by Western analysis, from approximately 3 hours after induction (data not shown). Denaturing/native combination Ni-affinity purification was chosen to attempt to purify His-Tcj4, as previous work had indicated it to be insoluble (data not shown).

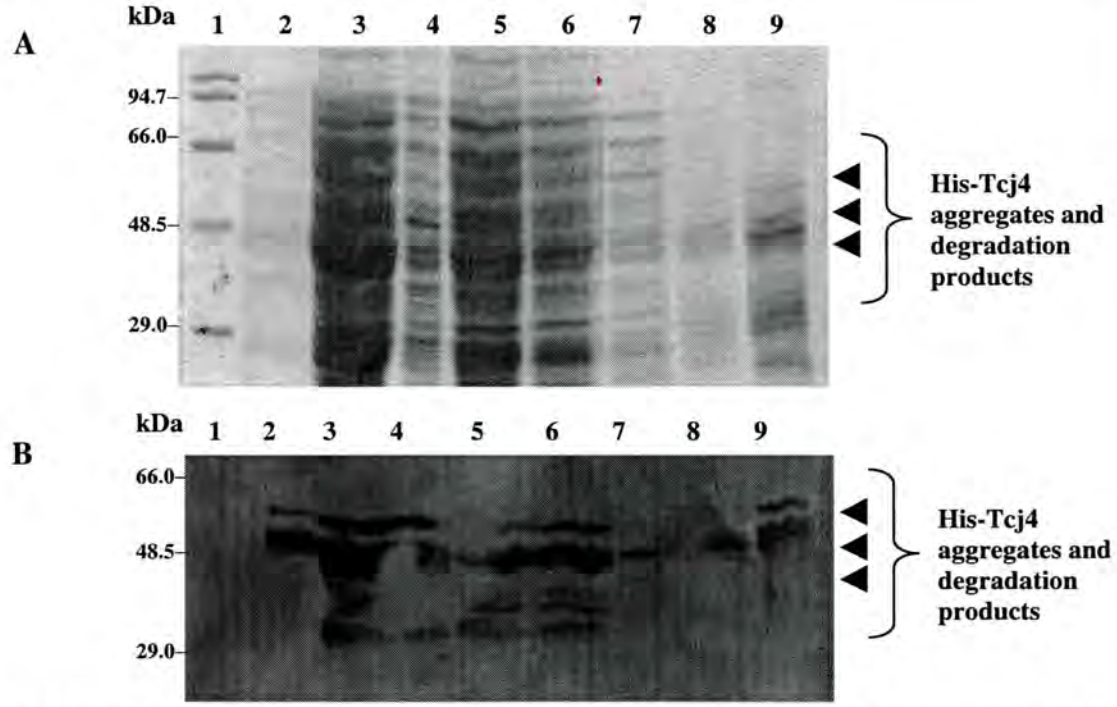


Figure 3.13: Unsuccessful Purification of His-Tcj4 by Nickel Affinity Chromatography
His-Tcj4 was not successfully purified from *E. coli* BL21(DE3)(pLysS)[pET28a/Tcj4] by a combination of native and denaturing batch Ni-affinity chromatography.**(A)** SDS-PAGE analysis of purification samples of His-Tcj4 and (B) corresponding Western analysis and chemiluminescence-based immunodetection. For both (A) and (B) Lane 1 – Broad range SDS-PAGE molecular mass markers (sizes shown in kDa); Lane 2 – induced *E. coli* BL21(DE3)(pLysS)[pET28a/Tcj4] whole cell extract; Lane 3 – *E. coli* BL21(DE3)(pLysS)[pET28a/Tcj4] cleared lysate; Lane 4 – flow through; Lane 5 – denaturing wash with Denaturing Lysis Buffer; Lanes 6 and 7 – native washes with Native Wash Buffer 1; Lane 8 and 9 – elutions 1 and 2 respectively. Arrows indicate the appearance of aggregates and/or degradation products containing His-Tcj4 observed on both SDS – PAGE and Western analysis.

The Western analysis [Figure 3.13(B)] revealed the presence of aggregation and/or degradation products of His-Tcj4 in all of the fractions collected during the purification. The initial induced *E. coli* BL21(DE3)(pLysS)[pET28a/Tcj4] culture displayed a small proportion of aggregates and/or potential degradation products (Figure 3.13 Lane 2). The actual purification appeared to promote aggregation and degradation, as the amount of aggregates and degradation products increased when cleared cell lysate was prepared (Figure 3.13 Lanes 2 and 3). A large proportion of His-Tcj4 did not bind the Ni-charged resin and can be observed in the flow through (Figure 3.13 Lane 4) and wash fractions (Figure 3.13 Lanes 5 to 7). The small amount of His-Tcj4 that did bind to the resin also appeared to be aggregated or degraded (Figure 3.13 Lanes 8 and 9). It was possible that the environment of *E.coli* was less than ideal for the over-production of His-Tcj4, which could have lead to aggregation and/or subsequent degradation within the cell. Purification would have therefore promoted a process that had already begun in the cell. The production of His-Tcj4 could have been too high to allow the protein to fold to a stable structure. A possible solution, would be to attempt the production and purification of His-Tcj4 in *E.coli* BL21(DE3) containing pLysE, a plasmid that is better at promoting the stable production of especially toxic proteins in *E.coli*.

3.3.6 Specificity of Interaction of Antibody to C-terminal Peptide Sequence of Tcj2

Of all the composite analyses performed, Tcj2 was chosen as the candidate for peptide-directed antibody synthesis. Serum extracted from the host rabbit after immunisation was used to confirm the ability of the antibody raised to that peptide to recognise Tcj2 (Figure 3.14).

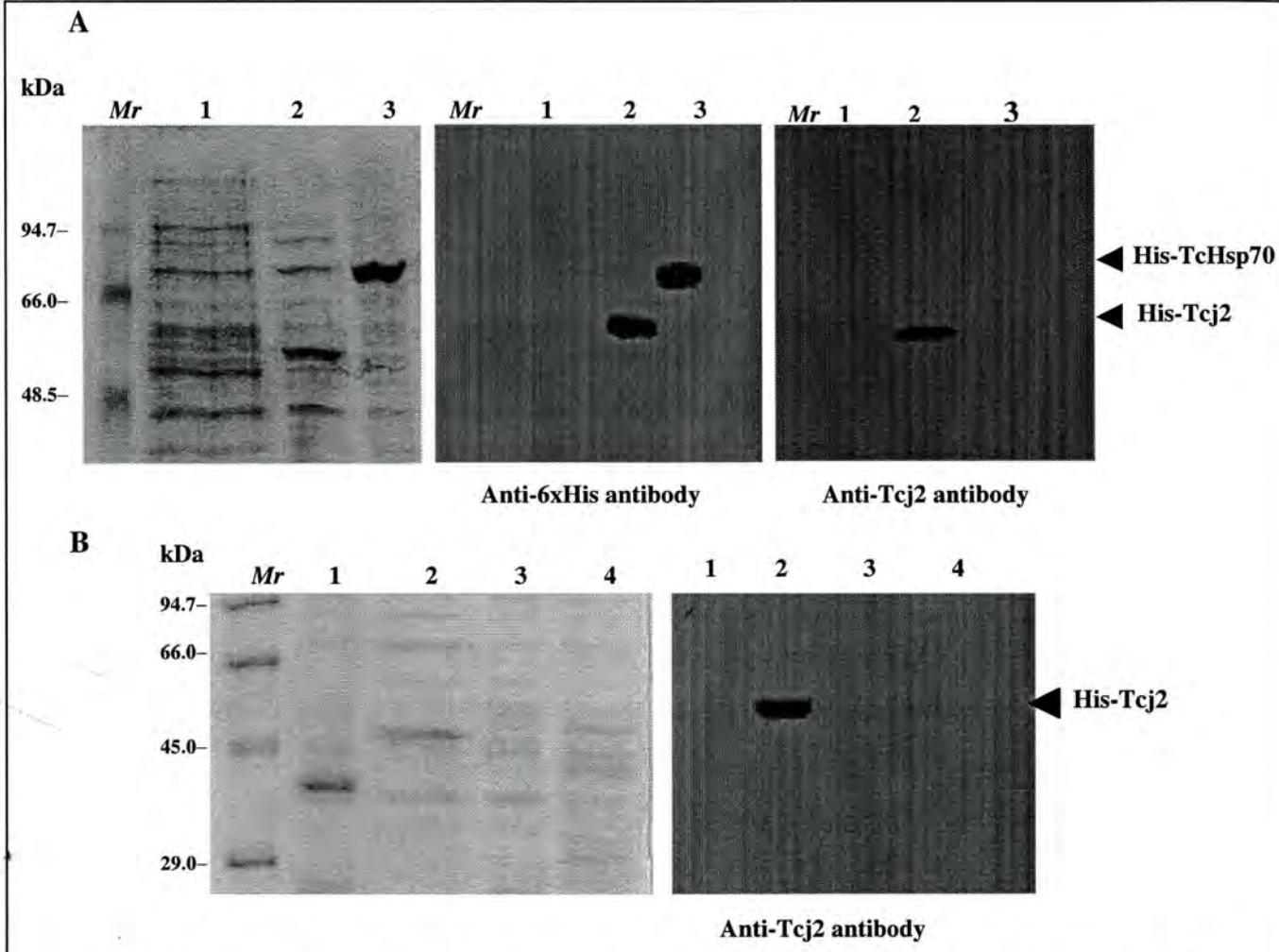


Figure 3.14: Specific Detection of Tcj2 using antibody designed to a C-terminal peptide in Tcj2

(A) Testing of the antibody designed to a C-terminal peptide of Tcj2 using Western analysis and chemiluminescence-based immunodetection. The left panel shows the SDS-PAGE. Western analysis and chemiluminescence-based immunodetection was performed using the anti-6xHis antibody (middle panel) and the antibody to the C-terminal peptide in Tcj2 (right panel) In all figures: *Mr* – molecular mass markers (sizes in kDa shown); Lane 1 – *E.coli* BL21(DE3) whole cell extract (negative control); Lane 2 – *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lane 3 –*E.coli* BL21(DE3)[pET14a/TcHsp70] whole cell extract (positive control for anti-6xHis antibody, negative control for anti-Tcj2 antibody).

(B) SDS-PAGE (left panel) and corresponding Western analysis (right panel), showing the detection of His-Tcj2 protein, and not the other Tcj proteins, by chemiluminescence-based immunodetection using the antibody to a C-terminal peptide within Tcj2. In all figures: *Mr* – molecular mass markers (sizes shown in kDa); Lane 1 – *E.coli* BL21(DE3)(pLysS)[pET23b/Tcj1] whole cell extract; Lane 2 – *E.coli* BL21(DE3)[pET28a/Tcj2] whole cell extract; Lane 3 – *E.coli* BL21(DE3)[pET23b/Tcj3] whole cell extract; Lane 4 – *E.coli* BL21(DE3)(pLysS)[pET28a/Tcj4] whole cell extract.

As observed in Figure 3.6, the antibody directed to the specific peptide in the C terminus of Tcj2 was capable of recognising His-Tcj2 in a Western analysis. Cell extract of induced transformed *E. coli* BL21(DE3) cultures producing Tcj1-His, His-Tcj2, Tcj3-His, His-Tcj4 and His-TcHsp70 were used in the analysis. The peptide-directed antibody was tested to determine its ability to recognise Tcj2 using Western analysis and chemiluminescence-based immunodetection (Figure 3.14 [A]). The presence of the 6xHis-tagged trypanosomal Hsps was simultaneously confirmed by western blotting with anti-6xHis antibody. The Tcj2 peptide-directed antibody is specifically able to recognise His-Tcj2 and not His-TcHsp70 (Figure 3.14[A]). In addition, the specificity of the peptide-directed antibody to Tcj2 was tested using whole cell extracts of cultures producing each of the Tcj proteins (Figure 3.14 [B]). The antibody specifically detected Tcj2 (Figure 3.14[B] Lane 2) and not the other trypanosomal Hsp40s (Figure 3.14 [B] Lanes 1, 3 and 4). This antibody was employed during the functional analysis of the chaperone abilities of TcHsp70 and the ability of Tcj2 to act as a co-chaperone (Chapter 4).

3.3.7 Spectrofluorimetric Analysis of Successfully Purified Trypanosomal Hsps

Successfully purified trypanosomal Hsps (His-TcHsp70, Tcj1-His and His-Tcj2) were analysed to assess the degree of aggregation, especially in the case of proteins purified by denaturing/native combination Ni-affinity purification. The proteins were analysed for the presence of small, soluble aggregates and large aggregates by spectrofluorimetry. The protein samples were excited at specific wavelengths (350 nm for small soluble aggregates and 500 nm for large aggregates) and their emission recorded over a wavelength range of 300 nm to 600 nm. The presence of small, soluble aggregates is indicated by a large emission at 350 nm, when the sample is excited at that wavelength. Similarly, for large aggregates, a large emission is noted at 500 nm after excitation of the sample at that wavelength. Figure 3.15 shows the spectrofluorimetric analysis of the purified proteins, including a control containing aggregated protein (heat denatured BSA) and a control relatively free of aggregated protein (native BSA).

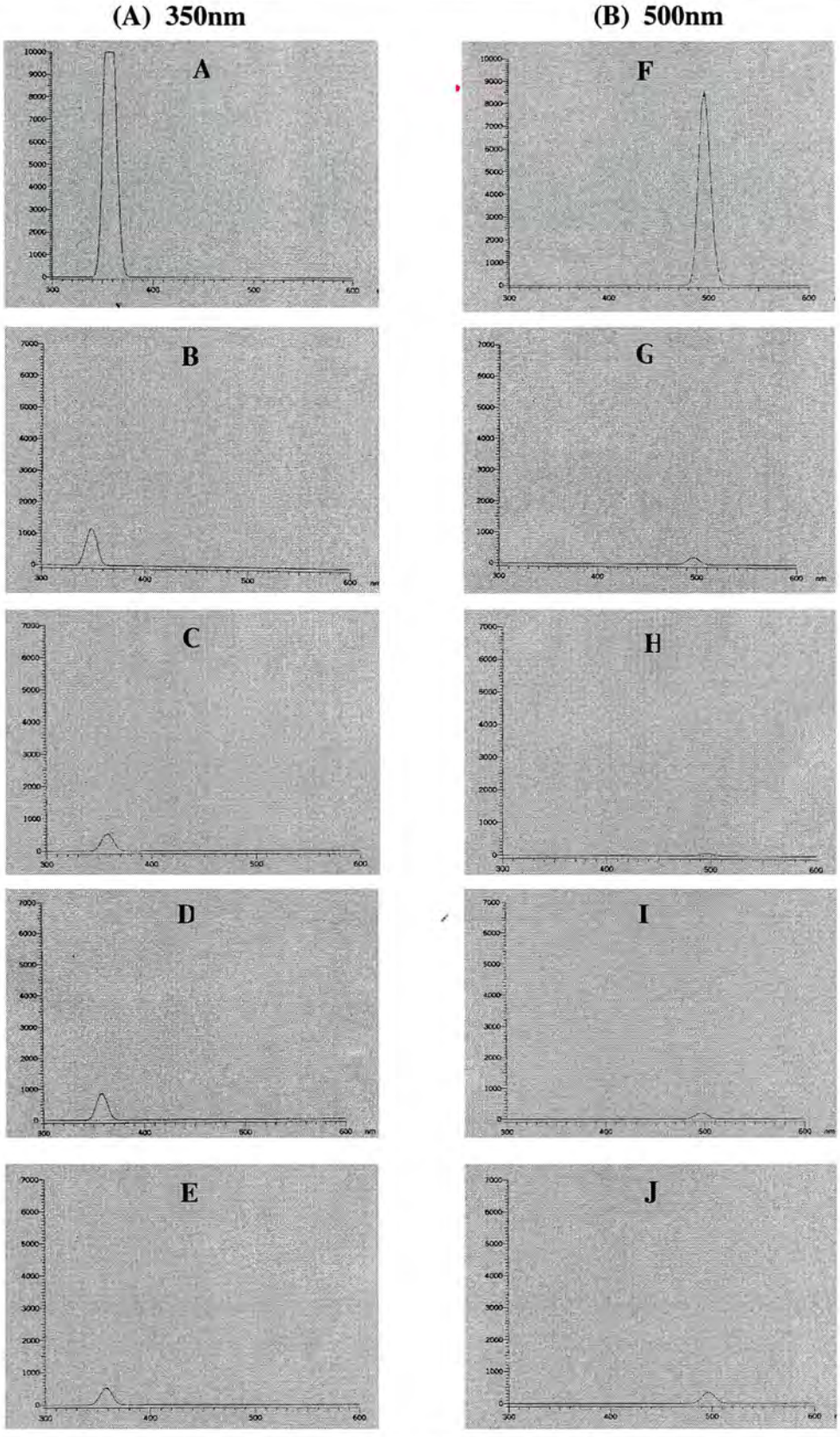


Figure 3.15: Spectrofluorometric Analysis of Purified His-TcHsp70, Tcj1-His and His-Tcj2

All purified protein preparations were diluted 1 in 100 and excited at 360nm (A) to detect soluble aggregates and at 500nm (B) to detect large aggregates, and their emission from 300nm to 600nm measured. Equivalent amounts of protein were used in all reactions. The y-axis indicates the emission in arbitrary units and the x-axis indicates the wavelength of emission (range 300-600nm) A and F – thermally denatured BSA (aggregation control); B and G – Native BSA (native control); C and H – His-TcHsp70; D and I – Tcj1-His; E and J – His-Tcj2.

The large peaks in Figure 3.15 (A) and (F) detail the absorbance of the denatured BSA when it was excited at 350 nm and 500 nm respectively. These peaks indicated the presence of both small soluble and large aggregates representative of a protein that had lost its native conformation through thermal denaturation. By contrast, the peaks at 360 nm and 500 nm observed for the native BSA sample [Figure 4.16 (B) and (G)] were distinctly smaller, by comparison with the denatured sample. The corresponding spectra after excitation at 360 nm and 500 nm for His-TcHsp70 [(C) and (H)], Tcj1-His [(D) and (I)] and His-Tcj2 [(E) and (J)] resemble that of the non-denatured control that should be relatively free of aggregates. This suggested that purified His-TcHsp70, Tcj1-His and His-Tcj2 were largely in their native conformations, not aggregated, and thus could be used in functional biochemical assays.

3.4 CONCLUSION

His-TcHsp70, Tcj1-His and His-Tcj2 were successfully purified using Ni-affinity chromatography. Attempts to purify Tcj3-His and His-Tcj4 were unsuccessful. The successfully purified 6xHis-tagged Hsps were determined to be predominantly in their native state and were used in the functional analyses.

CHAPTER 4

IN VITRO AND *IN VIVO* ANALYSIS OF CHAPERONE ACTIVITY OF TRYPANOSOMAL HEAT SHOCK PROTEINS

4.1 INTRODUCTION

4.1.1 *In Vitro* Characterisation of Molecular Chaperone Activity of Hsp70s

Molecular chaperone was a term coined to describe proteins capable of interacting with and stabilizing non-native, alternative or nascent protein structures and, through a controlled mechanism or activity, facilitating their release to attain a functional conformation (Hendrick and Hartl, 1993). Molecular chaperones rarely act alone, often aided by so-called co-chaperones, which regulate a particular characteristic of partner proteins to fulfil the requirements for chaperone-assisted protein folding. While not all Hsps are molecular chaperones, the Hsp70/Hsp40 cycle of assisted-protein folding has all these elements. The process of substrate binding and release is limited by the ATPase activity of the Hsp70, which is stimulated by the presence of its partner Hsp40. Thus, the molecular chaperone activity of a specific Hsp70 can be characterised *in vitro* by determining its ATPase activity, its ability to bind non-native protein substrates and its interaction with potential co-chaperones (Freeman et al., 1998).

The stable binding of a non-native protein to Hsp70 is dependent on ATP hydrolysis, while the release of the protein requires nucleotide exchange. In essence, the efficiency of substrate binding/release and, thus assisted – protein folding, is dependant on the rate at which the Hsp70 hydrolyses ATP and exchanges nucleotides. ATP hydrolysis and nucleotide exchange (ATPase activity) is therefore an intrinsic quality of Hsp70 molecular chaperones and can be used to characterise the chaperoning ability of a particular Hsp70. Typically, Hsp70 display a low basal ATPase activity, due to a low intrinsic ATP hydrolysis rate. Hsp40 co-chaperones stimulate the ATP hydrolysis rate of Hsp70 during times of assisted – protein folding (Table 4.1). The hydrolysis of ATP is often the rate-limiting step of the Hsp70/Hsp40 chaperone assisted protein folding cycle. In times of stress, the ATPase activity can be upregulated by the presence of Hsp40 such that Hsp70/Hsp40 assisted protein folding can proceed to counteract the stressful conditions. Therefore, when assessing the chaperone ability of a specific Hsp70, analysis of ATP hydrolysis by that protein is an essential element.

By definition, molecular chaperones need to interact with and bind non-native protein substrates. This essential function of molecular chaperones can be analysed by binding/complex formation studies or by protein unfolding/refolding assays. In the case of

unfolding/refolding assays, the renaturation of a denatured protein substrate, such as citrate synthase or luciferase, due to the presence of a molecular chaperone can be monitored using light scattering or activity assays (Minami et al., 1996). The ability of chaperones to recognise and bind to non-native substrates, such as reduced carboxymethylated α-lactalbumin (RCMLA) (Freeman et al., 1998), to form complexes can be analysed by gel filtration high performance liquid chromatography (HPLC) or native gel electrophoresis. While the ability of an Hsp70 to promote refolding of a chemically unfolded protein is a definite sign of molecular chaperone activity, the binding of a non-native substrate is also a fundamental quality of a chaperone.

4.1.2 *In Vivo* Characterisation of Molecular Chaperone Activity of Hsp70s

The ability to genetically modify certain organisms has provided an *in vivo* approach to assessing molecular chaperone activity. In this case, the protein of interest can be introduced into a bacterial or yeast strain that is deficient in a gene encoding certain chaperone (Deloche et al., 1997; Johnson and Craig, 2001). This chaperone-deficient strain will often display alternative phenotypes, such as temperature sensitivity, which can be used to monitor the ability of the chosen protein to functionally replace the chaperone deficiency. This approach to chaperone analysis gives an overall assessment, including all of the characterisitics assessed *in vitro*. However, the analysis is an overall one and does not allow for dissection of individual parameters. This approach is advantageous in the respect that it provides information into not only the ability of the chosen protein to act as a chaperone *in vivo*, but also indicates which well characterised chaperone it is most similar to and thus what types of other proteins it has potential to interact with (Deloche et al., 1997; Johnson and Craig,2001).

We aimed to attempt the characterisation of His-TcHsp70 as a potential molecular chaperone in terms of its *in vitro* basal ATPase activity and ability to recognise and bind non-native substrates. The ability of either Tcj1-His and His-Tcj2 to stimulate the ATPase activity of His-TcHsp70 and to influence substrate binding was investigated. The ability of Tcj2 and Tcj3 to act as functional equivalents of the yeast Hsp40s, Ydj1 and Sis1 was investigated using *in vivo* yeast complementation systems.

4.2 EXPERIMENTAL PROCEDURES

4.2.1 Reagents and Chemicals

All reagents used in functional analyses are detailed in Appendices (Section 6.1).

4.2.2 *In Vitro* Analysis of Chaperone Activity of Trypanosomal Hsps

4.2.2.2 Protein Preparations used in Biochemical Assays

All trypanosomal heat shock proteins used in biochemical assays (His-TcHsp70, Tcj1-His and His-Tcj2) were produced recombinantly in *E.coli* and purified by Ni-affinity chromatography. All protein preparations used were at least 80% pure, analysed for aggregation and found to be soluble (Chapter 3).

4.2.2.3 Determination of Basal ATPase Activity of His-TcHsp70

ATPase hydrolysis by His-TcHsp70 was measured using a colorimetric assay to detect the release of inorganic phosphate during the reaction. His-TcHsp70 (0.4 μM) was equilibrated at 37°C in ATPase buffer (10 mM Hepes, pH 7.4; 10 mM $MgCl_2$; 20 mM KCl; 0.5 mM DTT) before the addition of ATP (600 μM) to start the reaction. Samples (50 μL) were removed at regular time intervals and added to an equal volume (50 μL) of 10% SDS in a 96 well microtitre plate. A standard curve (Appendix 6.1; Figure 6.2) of phosphate concentrations ranging from 20 nmol to 200 nmol was prepared using potassium hydrogen phosphate (stock solution 1 mM). Colorimetric detection of phosphate in samples and standard solutions was achieved by addition of 1.25% ammonium molybdate in 6 M sulphuric acid (50 μL), followed by 9% ascorbic acid in triple distilled water (50 μL). The reactions were incubated at room temperature for 30 min with gentle agitation and the A_{660} measured. Using the data collected, initial velocities were determined and ATP hydrolysis calculated as Specific Activity (nanomoles of inorganic phosphate released per minute per milligram of His-TcHsp70).

4.2.2.4 Effect of RCMLA Substrate and Tcjs on Basal ATPase Activity of His-TcHsp70

The effect of the potential co-chaperones (Tcj1-His and His-Tcj2) and a model Hsp70 substrate (reduced carboxymethylated α-lactalbumin, RCMLA) on the basal ATPase activity of His-TcHsp70 was determined. ATPase assays were performed as described (Section 4.2.2.3), with the addition of either Tcj1-His or His-Tcj2 at submolar (0.2 μM), equimolar (0.4 μM) and molar excess (0.8 μM) concentrations, or RCMLA (2 μM), or a combination of RCMLA and each of the Tcj proteins.

4.2.2.5 Analysis of Substrate Binding and Complex Formation by Trypanosomal Hsps

The binding of RCMLA by His-TcHsp70 in the presence and absence of either Tcj1-His or His-Tcj2 was analysed by continuous, native/non-denaturing polyacrylamide gel electrophoresis (PAGE), adapted from Freeman et. al. (1998). His-TcHsp70 (0.25 μg/μL) was incubated with RCMLA (0.25 μg/μL) at 37°C for 30 min in buffer B (20 mM HEPES, pH 7.2; 5 mM $MgCl_2$; 100 mM NaCl). The same protocol was used to analyse the binding of RCMLA by His-TcHsp70 in the presence of Tcj1-His (0.125 μg/μL) or His-Tcj2 (0.125 μg/μL), and to detect the formation of complexes between His-TcHsp70, Tcj1-His and His-Tcj2. Control reactions containing only His-TcHsp70, RCMLA, Tcj1-His or His-Tcj2 were included. In all cases, native loading dye (5 μL) was added after the 30 min incubation and the whole reaction loaded on to a 10% native polyacrylamide gel (pH 8.8) and electrophoresised in native bath buffer (pH 9.0) at 200 V for 45 min.

Western analysis using chemiluminescence-based immunodetection was used to detect 6xHis-tagged (anti-6xHis antibody) and Tcj2 species (antibody to C-terminal peptide sequence of Tcj2). Western analysis using chemiluminescence-based immunodetection was performed as described (Section 3.2.2.4), with the modification that the pH of the transfer buffer was adjusted to 9.0.

4.2.3 *In Vivo* Analysis Of Chaperone Activity of Trypanosomal Hsps

4.2.3.1 Construction of Tcj2-Encoding Constructs for Expression in *Saccharomyces cerevisiae*

The Tcj2 coding region was inserted downstream from the *gal* promoter into the low copy number URA3 plasmid pKG6 and the low copy number TRP1 plasmid pKG4. The Tcj2 coding sequence was Polymerase Chain Reaction (PCR) amplified from pET28a/Tcj2 using the EXPAND™ High Fidelity PCR System (Boehringer Mannheim, Germany). Primers were designed to incorporate a yeast Kozak sequence for protein expression and unique restriction sites for *Eco*RI and *Xho*I (forward primer 5'-GAATTCAAGATGGTTAAGGAGACTAAGTTT-3' and reverse primer 5'-CTCGAGCTACTGTTGCGTACAAGT-3'). DNA template (pET28a/Tcj2, 100 ng), forward and reverse primers (500 nM each) and deoxynucleotides (dNTPs, 200 µM each) were combined with 2.6 Units EXPAND™ High Fidelity DNA-dependent DNA polymerase (*Pwo/Taq* enzyme mix) in EXPAND™ Buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 0.5% Tween 20 (v/v), 0.5% Nonidet P40 (v/v), 50% glycerol, 1.5 mM $MgCl_2$). The PCR cycling parameters were as follows: 94°C for 2 min; followed by 25 cycles of 94°C for 2 min, 55°C for 1 min, 72°C for 2.5 min; and a final incubation of 72°C for 10 min. An aliquot of the reaction was analysed by 0.8% agarose gel electrophoresis (AGE) and UV visualisation for the amplification of the Tcj2 coding sequence. The EXPAND™ High Fidelity PCR system produces PCR products with 3` single deoxyadenosine overhangs, for easy cloning using the pGEM®-T Easy Vector System I (Promega, USA) for amplification in *E. coli*.

The PCR product was purified from a 0.8% agarose gel using the Nucleospin® Extract 2 in 1 Gel Purification Kit (Macherey-Nagel, Germany) according to the manufacturers specifications, with the modification that the PCR product bound to the column was eluted in sterile triple distilled water (30 µL). This PCR product was checked by 0.8% AGE and its concentration determined spectrophotometrically (A_{260}) before inserting into the pGEM®-T Easy Vector System I (Promega, USA). Ligation reactions were set up as follows: pGEM®-T Easy vector (50 ng) and PCR product were combined in concentration ratios of 1:1, 1:3 and 1:6 with 2.5 µL of 2x rapid ligation buffer (60 mM Tris-Cl, pH 7.8; 20 mM $MgCl_2$; 20 mM DTT; 1 mM ATP, 10% PEG) and 1 U T4 DNA ligase (Promega, USA) in a total volume of 5 µL. Control reactions containing no vector, no PCR product and water alone were included.

Ligations were incubated at room temperature for 3 hours and then overnight at 4°C. After incubation, 2.5 µL of the ligation was transformed (Section 3.2.2.2) into *E. coli* DH5α [geneotype *supE*44 Δ*lac*U169 (ϕ80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1], plated on to 5% (w/v) MacKonkey agar (10 g/L lactose) containing ampicillin (100 µg/mL) and incubated overnight at 37°C. Single white/pink transformant colonies were selected, inoculated into 2xYT broth (5 mL) containing ampicillin (100µg/mL) and grown overnight at 37°C with shaking for small-scale plasmid isolation (Section 3.2.2.3). Correct clones were identified by digestion with *Eco*RI and *Xho*I, to release the Tcj2 coding sequence. A single correct pGEM®-T Easy/Tcj2 clone was selected and DNA for subcloning into pKG4 and pKG6 prepared using the High Pure Plasmid Isolation Kit (Roche, Germany), according to the manufacturer's specifications. The only modification was that plasmid DNA bound to the filter column was eluted in sterile triple distilled water (50 µL) instead of elution buffer. Restriction enzyme analysis and sequencing was used to verify the identity of the pGEM®-T Easy/Tcj2 construct and to confirm the integrity of the Tcj2 coding sequence.

A bulk double digest of pGEM®-T Easy/Tcj2 with *Eco*RI and *Xho*I was performed and the Tcj2 insert gel purified (described in this section). pKG4 and pKG6 vectors were digested with *Eco*RI and *Xho*I and treated with shrimp alkaline phosphatase (Roche, Germany) to remove 5' phosphate groups and prevent self-ligation. Ligation reactions for insertion of the Tcj2 coding sequence into both pKG4 and pKG6 were set up. The *Eco*RI/*Xho*I Tcj2 fragment was combined with *Eco*RI/*Xho*I linearised pKG4 or pKG6 in concentration ratios of 1:1, 1:3 and 1:6 with 2.5 µL of 10x T4 ligation buffer (330 mM Tris-HCl, pH 8.8; 660 mM potassium acetate; 100 mM magnesium acetate; 5 mM DTT, 0.1% BSA) and 1U of T4 ligase (Amersham Pharmacia Biotech, USA) in a total volume of 5µL. Ligation reactions were incubated at room temperature for 3 hours and then overnight at 4°C before transformation into *E.coli* DH5α (described Section 3.2.2.2). Transformants capable of growth on 2xYT agar containing ampicillin (100 µg/mL) were selected for small-scale plasmid isolation and screened for the Tcj2 insert by digestion with *Eco*RI and *Xho*I. Putative pKG4/Tcj2 and pKG6/Tcj2 constructs were isolated using the High Pure Plasmid Isolation Kit (described in this Section) and further restriction analysis performed to confirm their identity, after which the confirmed plasmid construct was stored at –20°C until transformation into *S. cerevisiae* strains.

4.2.3.2 Construction of Tcj3-Encoding Plasmids for Expression in *S. cerevisiae*

The Tcj3 coding sequence was inserted downstream of the *gal* promoter into the low copy number URA3 plasmid pKG6. The Tcj3 coding sequence was PCR-amplified from pET23b/Tcj3 using primers to introduce unique cloning sites *Bam*H1 and *Eco*R1 for insertion into the yeast expression vector pKG6 and a yeast Kozak sequence for protein expression (forward primer 5'-GGATCCGAGAAGATGGTAAAGGAAACAGAG-3' and reverse primer 5'-GAATCCAAGCTTTTATCTCCGCCGGGCCGC-3'). The PCR amplification of the Tcj3 coding sequence using the EXPAND™ High Fidelity PCR System was performed as described (Section4.2.3.1) for Tcj2, with the exception that the annealing temperature was reduced from 55°C to 50°C. PCR products were visualised under UV light after 0.8% AGE

PCR product from the amplification of the Tcj3 coding sequence was inserted into pGEM®-T Easy vector and subsequently excised and inserted into pKG6. The exact procedure employed to generate pKG4/Tcj2 and pKG6/Tcj2 (described Section 4.2.3.1) was followed to create pKG6/Tcj3, with the exception that *Bam*H1 and *Eco*R1 restriction digests were used to construct and confirm the plasmid identity. The integrity of the amplified Tcj3 fragment was confirmed by restriction enzyme analysis and sequencing of the pGEM-T Easy/Tcj3 plasmid construct.

4.2.3.3 Host Cells and Constructs for Complementation of Chaperone-Deficient *S.cerevisiae* strains

The ability of Tcj2 and Tcj3 to complement *in vivo* for the absence of the yeast Hsp40s, Ydj1 and Sis1, was investigated using *Saccharomyces cerevisiae* mutant strains JJ160, a *ydj1* strain [genotype ***mat a*** *trp1-1 ura3-1 leu2-3,112 his3-11,15 ade2-1 can1-100 GAL2+ met2-Δ1 lys2-Δ2 ydj1::HIS3*] and WY26, a *sis1* strain [genotype ***mat*** *α trp1-1 ura3-1 leu2-3,112 his3-11,15 ade2-1 can1-100 GAL2+ met2-Δ1 lys2-Δ2 sis1::LEU2*], respectively. The control plasmids for JJ160 and WY26 carry copies of the wild type genes disrupted in the knockout strains and distinct auxotrophic markers for selection. pRS317-*YDJ1*, the control plasmid for JJ160, carries the *YDJ1* coding region and the *LYS2* selectable marker. The control plasmid for WY26 is pYW98-*SIS1*, encoding wild type Sis1 and carrying the *TRP1* selectable marker. Both yeast strains and control plasmids were a kind donation of Dr Elizabeth Craig (University of Wisconsin Medical School, USA). pKG4/Tcj2, pKG6/Tcj2

and pKG6/Tcj3 contain the coding sequences of Tcj2 and Tcj3 under the control of the *gal* promoter, which is induced by growth in medium where the sole carbon source is galactose. The pKG6 constructs carry the *URA3* selectable marker and pKG4 constructs carry the *TRP1* selectable marker.

Competent cells of both strains JJ160 and WY26 were prepared using the Frozen – EZ Transformation II™ Kit (Zymo Research, USA). A yeast culture (10 mL) in selective medium (Yeast Minimal Medium [YMM] glucose [HIS$^-$] for *S. cerevisiae* JJ160; YMM glucose [LEU$^-$] for *S. cerevisiae* WY26) was grown at 23°C with shaking until an A_{600} of 0.8-1.0 was reached. The cells were harvested by centrifugation (500x *g* for 10 min) and washed with EZ 1 solution (10 mL). The mixture was centrifuged (500x *g* for 10 min), the supernatant discarded and the cells resuspended in EZ 2 solution (1 mL) and stored in aliquots (60 µL) at –70°C until use.

4.2.3.4 Complementation of *S.cerevisiae ydj1* strain JJ160

pKG6/Tcj2, pKG6/Tcj3, pKG6 (negative control) and pRS317-*YDJ1* (positive control) were transformed into competent *S. cerevisiae* strain JJ160 using the Frozen– EZ Transformation II™ Kit, according to the manufacturers instructions. Plasmid solution (100 ng) was added to competent JJ160 (15 µL) with EZ 3 solution (100 µL) and incubated at room temperature for 1 hour, vortexing every 15 min. After incubation, the whole transformation mixture was plated on to yeast minimal agar (YMA) selective for each plasmid 9YMA glucose [URA$^-$ HIS$^-$] for JJ160 containing pKG6 constructs; and YMA glucose [LYS$^-$ HIS$^-$] for pRS317-*YDJ1*) and incubated at 23°C for 5 days. A single yeast transformant for each of the plasmids was selected to inoculate the corresponding yeast minimal medium (YMM galactose [URA$^-$ HIS$^-$] for JJ160 containing pKG6 constructs; and YMM galactose [LYS$^-$ HIS$^-$] for pRS317-YDJ1) (5 mL) and grown at 23°C with shaking for 3 days. The density was determined and all of the cultures corrected to the same A_{600}. Tenfold serial dilutions (10^0 to 10^{-5}) were prepared for each of the cultures and an aliquot (10 µL) of each of the dilutions spotted on to the correct selective minimal agar for growth and protein expression (YMA galactose [URA$^-$ HIS$^-$] for JJ160 containing pKG6 constructs; and YMA galactose [LYS$^-$ HIS$^-$] for pRS317-*YDJ1*). Separate plates were prepared for incubation at 25°C, 30°C and 34°C for 5 days. Yeast growth at the different temperatures was noted and recorded using digital photography.

4.2.3.5 Complementation of *S.cerevisiae sis1* strain WY26

pKG4/Tcj2, pKG4 (negative control) and pYW98-SIS1 (positive control) were transformed into competent *S. cerevisiae* strain WY26 using the Frozen– EZ Transformation II™ Kit. Each plasmid solution (100 ng) was added to competent WY26 (15 µL) with EZ 3 solution (100 µL) and incubated at 30°C for 1 hour, vortexing every 15 min. The whole transformation mixture was plated on to yeast minimal agar (YMA glucose [TRP^- LEU^-] for all plasmids) and incubated at 30°C for 5 days.

A single yeast transformant for each of the plasmids was selected to inoculate yeast minimal medium (5 mL) (YMM galactose [TRP^- LEU^-]) and grown at 30°C with shaking for 3 days. The density was determined and all of the cultures corrected to the same A_{600}. Tenfold serial dilutions (10^0 to 10^{-5}) were prepared for each of the cultures and aliquots (10 µL) of each of the dilutions spotted on to selective minimal agar for growth and protein expression (YMA galactose [TRP^- LEU^-] with and without 0.5% 5-FOA). Plates were incubated at 30°C for 5 days and yeast growth recorded using digital photography.

4.2.2.6 Detection of Tcj2 in Yeast Cultures using Western Analysis with Chemiluminescence-based Immunodetection

Western analysis was used to confirm production of Tcj2 in transformed *S.cerevisiae* strains JJ160 and WY26. Yeast protein extracts were prepared for SDS-PAGE (described Section 3.2.2.4) as follows. Yeast cells from cultures (5 mL) of *S. cerevisiae* JJ160[pKG6/Tcj2] and *S. cerevisiae* WY26[pKG4/Tcj2] were harvested by centrifugation (12000x *g* for 5 min) and resuspended in ESB (30 µL; 80 mM Tris-HCl, pH 6.8; 2% SDS; 10% glycerol; 1.5% DTT; 0.1 mg/mL bromophenol blue) and boiled at 100°C for 3 min. Glass beads were added up to the level of the liquid and the samples vortexed vigorously for 2 mins. Additional ESB (30 µL) was added, the samples vortexed briefly and boiled for 1 min before analysis by SDS-PAGE (described Section 3.2.2.4). Western analysis and chemiluminescence-based immunodetection for the detection of Tcj2 using the antibody to a C-terminal peptide in Tcj2 was performed (described Section 3.2.2.4). Protein extracts of *S. cerevisiae* strains JJ160 and WY26 were prepared similarly and used as negative controls.

4.3 RESULTS AND DISCUSSION

4.3.1 Basal ATPase Activity of His-TcHsp70 and Effect of the presence of a Model Protein Substrate and Potential Co-Chaperones

The basal ATPase activity of His-TcHsp70 was monitored *in vitro* by assessing the release of phosphate, as a result of the hydrolysis of ATP to ADP. Initial velocities of the reactions were calculated and a specific activity (nmol Pi/min/mg His-TcHsp70) calculated. After the basal ATPase activity was determined, the effect of the presence of a model Hsp70 substrate, RCMLA was examined (Figure 4.1).

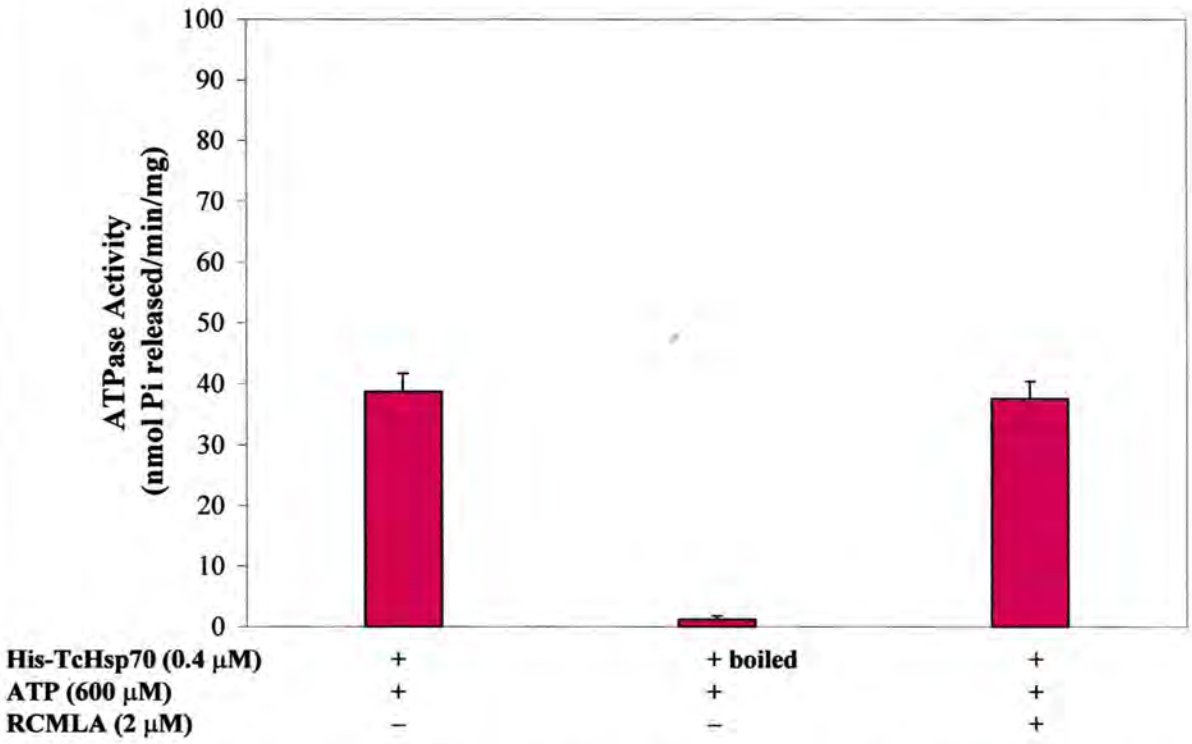


Figure 4.1: ATPase Activity of His-TcHsp70 in the presence and absence of RCMLA

The *in vitro* ATPase activity of His-TcHsp70 was monitored on the basis of ATP hydrolysis was determined in the absence and presence of a model protein substrate (RCMLA). ATP hydrolysis is represented as the nanomoles of inorganic phosphate produced per minute per mg of His-TcHsp70. Final concentrations of His-TcHsp70, ATP and RCMLA are indicated.

The ATPase activity was a direct result of the presence of active His-TcHsp70, as a sample of thermally denatured His-TcHsp70 did not display ATPase activity, in both the presence and absence of RCMLA. The basal ATPase activity of His-TcHsp70 was determined to be ~ 40nmol/min/mg His-TcHsp70. The basal ATPase activity of TcHsp70 was significantly higher than that reported for other Hsp70s of eukaryotic and prokaryotic origin (Table 4.1). The majority of Hsp70s have low basal ATPase activities, which are stimulated significantly by the presence of co-chaperones to promote Hsp70/Hsp40 mediated protein folding. This high basal ATPase activity of His-TcHsp70 has implications to the possible functioning of this protein as a chaperone. As ATP hydrolysis has been shown to be the limiting factor for the binding of substrates by Hsp70s during assisted-protein folding, our results suggest that unlike previous systems TcHsp70 may have the ability to bind and release proteins in the absence of stimulation by Hsp40 co-chaperones. This characteristic could be of an advantage to the parasite in its host. The elevated body temperature of a mammalian host (37°C) would ensure that the parasite was under constant heat shock.

The basal ATPase activity of His-TcHsp70 does not appear to be modified by the presence of the model substrate RCMLA alone. This was expected, as the stimulation of Hsp70 ATPase activity is specific and is modulated via interaction of the J domain with the N-terminal ATPase domain of Hsp70s, rather than by substrate binding alone. Maximal ATPase activity stimulation has been shown to occur in the presence of both substrates and a co-chaperone (Laufen et al., 1999), but not in the presence of substrate alone.

The effect of the Tcjs on the ATPase activity of His-TcHsp70 in the absence and presence of RCMLA was also determined. Figure 4.2 shows the effect of Tcj1-His on His-TcHsp70 ATP hydrolysis in the absence and presence of RCMLA.

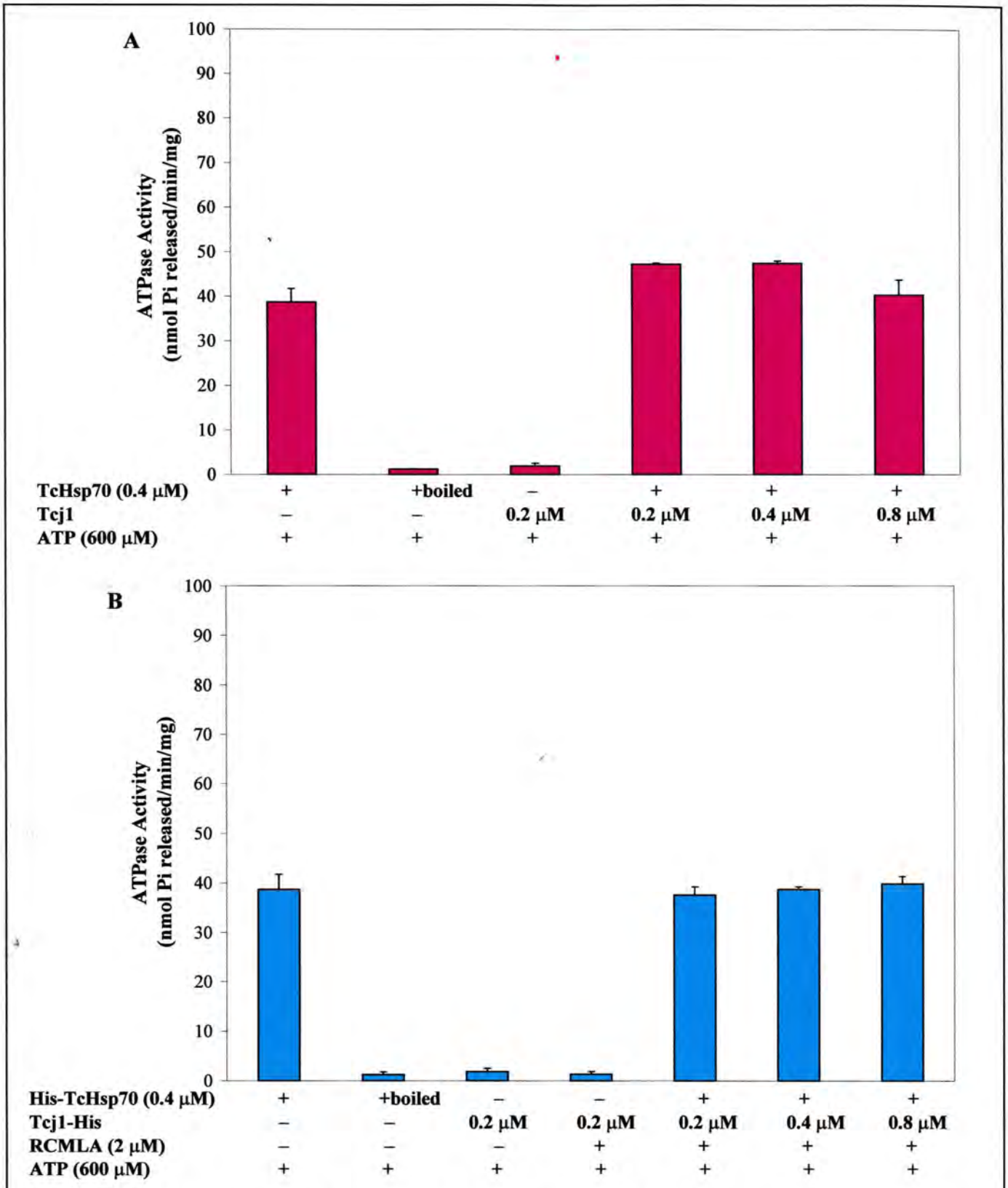


Figure 4.2: ATPase Activity of His-TcHsp70 in the presence of Tcj1-His

The effect of the presence of different concentrations of the potential co-chaperone Tcj1-His on the *in vitro* ATPase activity of His-TcHsp70 was monitored on the basis of ATP hydrolysis was determined in the absence (A) and presence (B) of a model protein substrate (RCMLA). ATP hydrolysis is represented as the nanomoles of inorganic phosphate produced per minute per mg of His-TcHsp70 (nmol Pi/min/mg TcHsp70). The final concentrations of His-TcHsp70, Tcj1-His, RCMLA and ATP are indicated.

Tcj-His alone, or in the presence of RCMLA, did not display ATPase activity. Tcj1-His did not appear to stimulate the ATPase activity of His-TcHsp70 to a significant level. A slight stimulation was observed at submolar and equimolar concentrations of Tcj1-His. However, there was no stimulation when His-TcHsp70 and Tcj1-His are combined in a 1:2 concentration ratio. Similarly, there is no stimulation of the ATPase activity of His-TcHsp70 by Tcj1-His at any concentration in the presence of RCMLA (Figure 4.2). The slight increase in stimulation of ATP hydrolysis in the presence of Tcj1-His alone could be due to Tcj1-His acting as a substrate for His-TcHsp70. Although the model substrate, RCMLA, did not stimulate the ATPase activity of His-TcHsp70, it could be that Tcj1-His, being of trypanosomal origin, was a more homologous substrate for His-TcHsp70. In addition, as Tcj1-His was purified under denaturing conditions, a portion of the protein may be unfolded even after renaturation. This unfolded portion may be capable of acting as a substrate for His-TcHsp70. This would explain the lack of stimulation by Tcj1-His in the presence of RCMLA, a competing substrate. This lack of stimulation of ATP hydrolysis by His-TcHsp70 in the presence of Tcj1-His was not unexpected. Tcj1-His is a Type III Hsp40, and by that token, not the "classical" Hsp70 – stimulating Hsp40. This did not mean that His-TcHsp70 and Tcj1-His do not interact. The two proteins may form a physiological partnership to fulfil a function within the parasite that it not dependent on ATP hydrolysis. The basal ATPase activity of His-TcHsp70 may be sufficient to accomplish this potential Tcj1-His mediated task within the parasite. Alternatively, Tcj1-His may act to stimulate the ATPase activity of His-TcHsp70 in the presence of a more specialised substrate, for example, the flagellar protein, which has been proposed to be a substrate for Tcj1 (Dr D. Engman, personal communication).

The ability of His-Tcj2 to stimulate the ATPase activity of His-TcHsp70 was investigated in the presence and absence of RCMLA (Figure 4.3). No significant stimulation of the basal ATPase activity of His-TcHsp70 was observed in the presence of His-Tcj2 when supplied at submolar concentrations, in the presence or absence of RCMLA (Figure 4.3). His-Tcj2 stimulated the basal ATPase activity of His-TcHsp70 when it was supplied at equimolar and molar excess concentrations. The stimulation of His-TcHsp70 basal ATP hydrolysis (~40 nmol Pi/min/mg) by His-Tcj2 was increased by the presence of RCMLA, to the maximal stimulation of ~60 nmol Pi/min/mg.

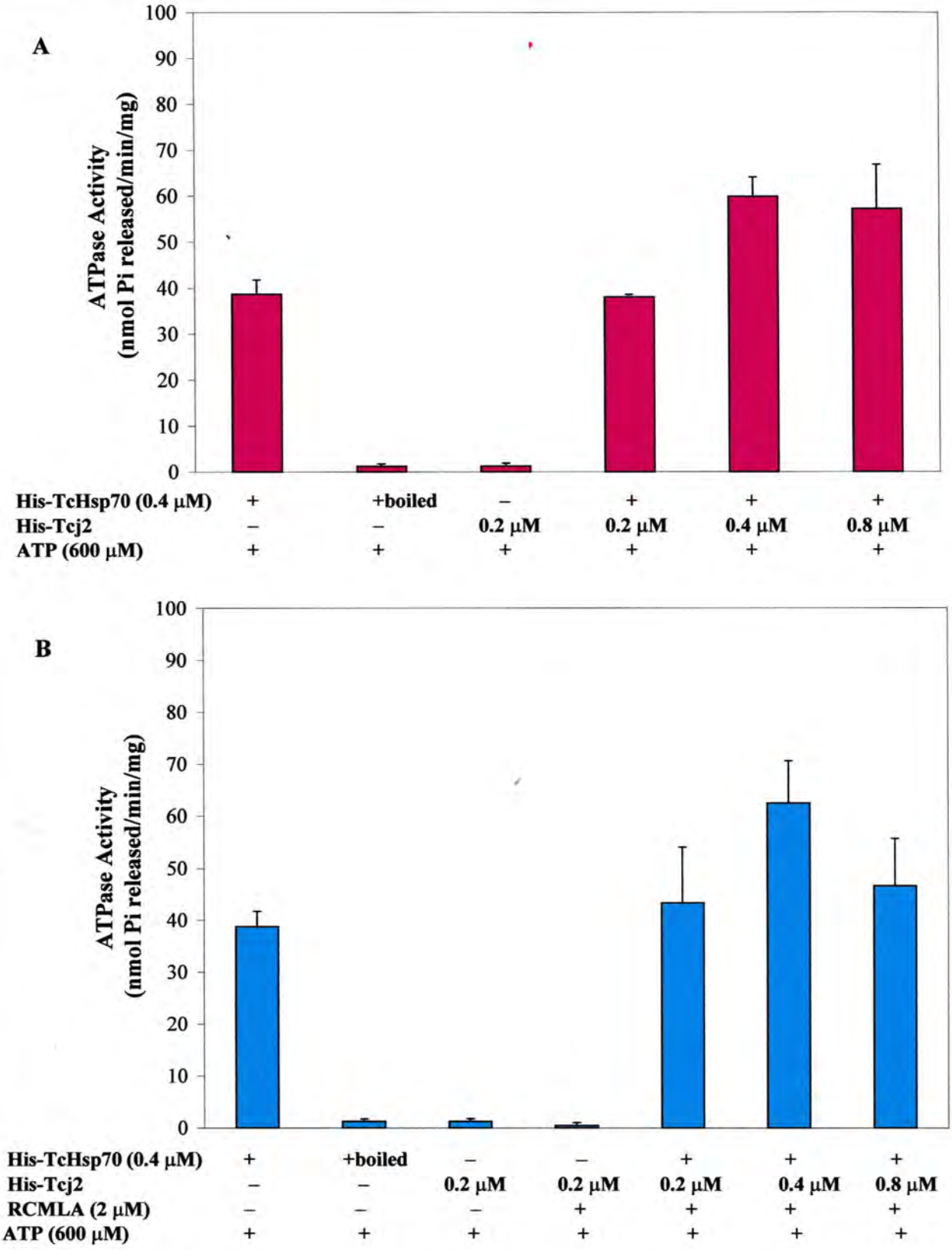


Figure 4.3: ATPase Activity of His-TcHsp70 in the presence of His-Tcj2

The effect of the presence of different concentrations of the potential co-chaperone His-Tcj2 on the *in vitro* ATPase activity of His-TcHsp70 monitored on the basis of ATP hydrolysis was determined in the absence (A) and presence (B) of a model protein substrate (RCMLA). ATP hydrolysis is represented as the nanomoles of inorganic phosphate produced per minute per mg of His-TcHsp70 (nmol Pi/min/mg TcHsp70). Final concentrations of His-TcHsp70, His-Tcj2, RCMLA and ATP are indicated

Although His-TcHsp70 was found to have a high basal ATPase activity, the fold stimulation (~1.5) in the presence of His-Tcj2 was low. This was different from most other Hsp70s, which exhibit a low basal ATPase activity that was stimulated significantly by the presence of a partner Hsp40 (Table 4.1). However, RCMLA may not be ideal and thus the experiment would have to be repeated using an alternative assay and a range of potential substrates to confirm the current results.

Table 4.1: Reported ATPase Activities for different Hsp70 Proteins

Heat Shock Protein	Source	Basal ATPase Activity	Stimulated ATPase Activity	Fold Stimulation (approx)	Stimulated by	Reference
Bip	Bovine liver	0.35 nmol/min/mg	2.1 nmol/min/mg	6	Synthetic peptides	Blond-Elguindi et. al. (1993)
Ssa1	Recombinant *S. cerevisiae*	5 nmol/min/mg	28 nmol/min/mg	6	Ydj1	Lu and Cyr (1998)
Ssa1	Recombinant *S. cerevisiae*	5 nmol/min/mg	24 nmol/min/mg	5	Sis1	Lu and Cyr, (1998)
Ssa1p	*S. cerevisiae*	1.25 nmol/min/mg	9 nmol/min/mg	7	Ydj1	Fewell et. al. (2001)
DnaK	*E.coli*	0.108 mol/min/mol	7.68 nmol/min/mg	71	DnaJ and GrpE	Montgomery et al. (1999)
DnaK	*E.coli*	0.108 mol/min/mol	0.5 nmol/min/mg	5	Synthetic peptides	Montgomery et al. (1999)
Hsc70	Bovine brain	0.2 nmol/min/mg	1.5 nmol/min/mg	8	DnaJ and GrpE	Minami et. al. (1996)
Hsc70	Bovine brain	0.2 nmol/min/mg	1.25 nmol/min/mg	6	Hsp40	Minami et. al. (1996)
Hsc70	Rat	1 nmol/min/mg	7 nmol/min/mg	7	Hdj1 and Bag1	Kanazawa et. Al. (1997)
Hsc70	Rat	1 nmol/min/mg	6.5 nmol/min/mg	7	Hdj2 and Bag1	Kanazawa et. Al. (1997)
Hsc70	Rat	1 nmol/min/mg	5 nmol/min/mg	5	Hdj3 and Bag1	Kanazawa et. Al. (1997)
Hsc70	Bovine brain	1.08 nmol/min/mg	13.98 nmol/min/mg	13	Csp1	Chamberlain and Burgoyne (1997)
Hsp70	Recombinant human	0.5 nmol/min/mg	4.66 nmol/min/mg	9	Csp1	Chamberlain and Burgoyne (1997)
Hsp70	Recombinant human (GST tagged)	5.2 nmol/min/mg	-	-	-	Olson et. Al. (1994)
Hsp70	Recombinant human (GST tagged)	520 nmol/min/mg	-	-	-	Olson et. Al. (1994)
Hsp70	**Recombinant *T.cruzi***	**40 nmol/min/mg**	**60 nmol/min/mg**	**1.5**	**Tcj2 and RCMLA**	**-**

The ATPase activities of recombinant TcHsp70 and human Hsp70, as GST fusion proteins, were previously reported by Olson et. al. (1994). The values reported for GST-TcHsp70 and GST-Hsp70 are approximately ten fold higher than those determined by this and previous studies (Table 4.1).

4.3.2 Substrate Binding by His-TcHsp70 and the Effect of the presence Potential Co-Chaperones

The ability of His-TcHsp70 to bind to and form complexes with RCMLA was analysed by 10% native/non-denaturing PAGE and Western analysis (Figure 4.4). A range of species of different mobilities are observed (Figure 4.4; labelled 1-5).

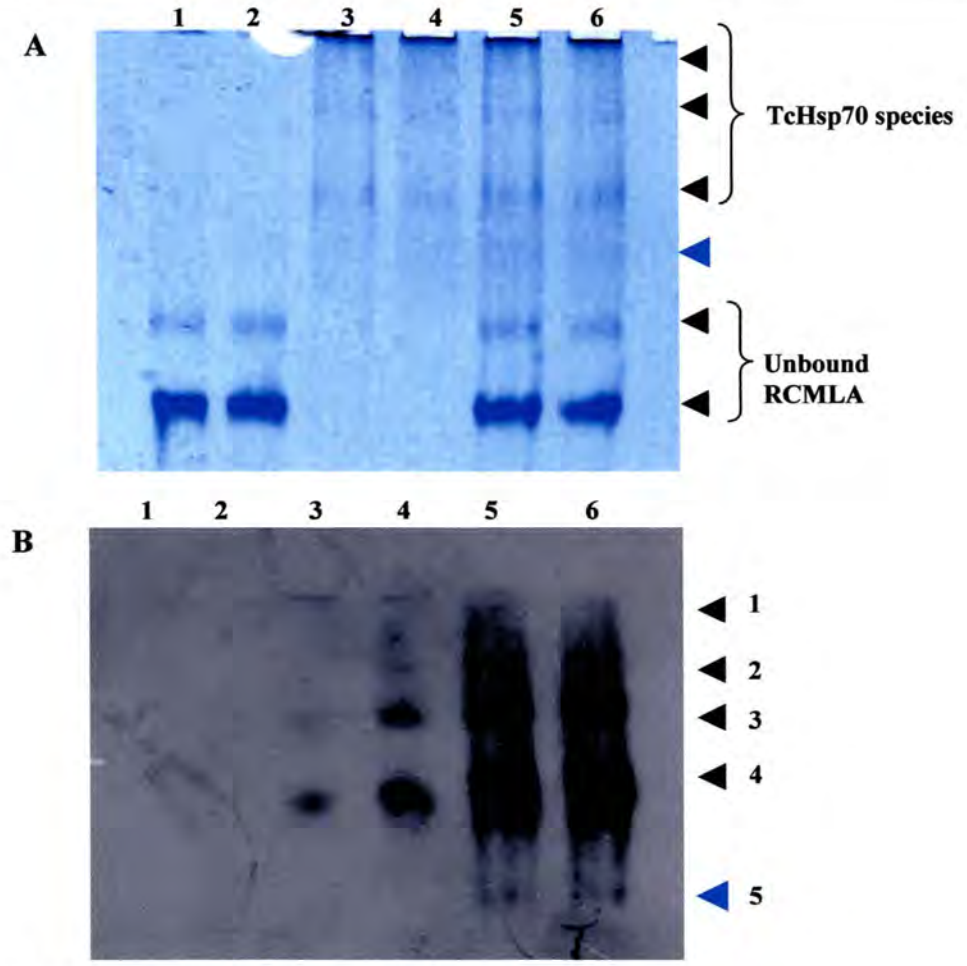


Figure 4.4: Analysis of RCMLA Binding by His-TcHsp70

The ability of His-TcHsp70 to bind a model protein substrate (RCMLA) was investigated by Native/Non-denaturing Polyacrylamide Gel Electrophoresis and Western analysis to detect the 6xHis tag on His-TcHsp70. The final concentrations of both of His-TcHsp70 and RCMLA was 0.25 μg/μL

(A) Coomassie stained Native PAGE of RCMLA binding by His-TcHsp70. Species corresponding to unbound RCMLA and His-TcHsp70 alone are indicated by black arrows and labelled accordingly. The blue arrow indicates a species of a different relative mobility arising due to the combination of His-TcHsp70 and RCMLA

(B) Western analysis of Native PAGE of RCMLA binding by His-TcHsp70. Species containing His-TcHsp70 were detected by chemiluminescence-based immunodetection with anti-6xHis antibody. Arrows (1-5) indicate the range of species containing His-TcHsp70, with the blue arrow indicating a species of different mobility (as observed on the PAGE).

For (A) and (B): Lane 1 – RCMLA alone; Lane 2 – RCMLA in the presence of 600μM ATP; Lane 3 – His-TcHsp70 alone; Lane 4 – TcHsp70 in the presence of 600 μM ATP; Lane 5 – TcHsp70 and RCMLA; Lane 6 – TcHsp70 and RCMLA in the presence of 600 μM ATP.

Species containing His-TcHsp70 were detected by chemiluminesence-based Western analysis with the anti-6xHis antibody. The lack of antibodies to RCMLA made the detection of species containing RCMLA not possible. The PAGE shows species of distinct mobilities, which correspond to unbound RCMLA (Figure 4.4 [A] Lanes 1, 2, 5 and 6) and His-TcHsp70 (Figure 4.4 [A] Lanes 3, 4, 5 and 6). The species containing His-TcHsp70 were also detected on the Western analysis using anti-6xHis antibodies. At least 4 species of different mobilities can be observed in the lanes containing His-TcHsp70 in the absence (Figure 4.4 Lane 3) and presence (Figure 4.4 Lane 4) of ATP on both the PAGE and Western analysis and an increase in species 3 and 4 in the presence of ATP (Figure 4.4 Lane 4). These species of different mobility could correspond to different oligomers (trimers, dimers etc) of His-TcHsp70. The formation of different oligomers could potentially be a mechanism to keep the Hsp70 inactive, until the presence of a model substrate or co-chaperone stimulates the formation of an active oligomeric state, eg monomer. This phenomenon of Hsp70 oligomers has been observed in similar studies (Blond-Elguindi et al., 1993; Palleros et al., 1991). In addition, polymerisation by Hsp70 has been proposed as a mechanism for bringing high concentrations of Hsp70 in contact with protein substrates to promote rapid formation of Hsp70-substrate complexes that facilitate protein refolding (King et al., 1999).

The formation of a new species of a different mobility (Figure 4.4 blue arrow, species 5, Lanes 5 and 6) was observed on both the PAGE and Western analysis in the lanes containing both His-TcHsp70 and RCMLA in the absence (Figure 4.4 Lane 5) and presence (Figure 4.4 Lane 6) of ATP. This species had a higher mobility than species 1 to 4, and is not observed in the reactions containing only His-TcHsp70. This suggests that this species is arising due to the presence of both His-TcHsp70 and RCMLA in the reaction. As the species was detected with anti-6xHis antibody, it was certain to contain His-TcHsp70. However, the lack of anti-RCMLA antibodies makes it not possible to say for certain whether the species contains RCMLA. It could have been a complex of His-TcHsp70 and RCMLA, due to binding of the non-native substrate by His-TcHsp70 acting as a molecular chaperone. It could also have been an alternative conformation, such as an active monomer of His-TcHsp70 that was formed due to the presence of a non-native substrate, which would also be consistent with potential chaperone activity of His-TcHsp70 (Blond-Elguindi et al., 1993). The formation of the higher mobility species 5 appeared to be independent of ATP, as it arises in both the absence (Figure 4.4 Lane 5) and presence (Figure 4.4 Lane 6) of 600 μM ATP.

In addition to the formation of species 5, there was an increase in the levels of species 1 – 4 in the reactions containing His-TcHsp70 and RCMLA. This suggested a change in the oligomeric state of the His-TcHsp70 population due to the presence of substrate RCMLA. It was also possible, that the increase in species 1 – 4 was due to binding of the different oligomers to RCMLA. However, to make a conclusive statement in this regard, Western analysis with an anti-RCMLA antibody would have to be performed.

4.3.3 Effect of Tcj1-His and His-Tcj2 on RCMLA Substrate Binding by His-TcHsp70

The effect of the presence of the potential co-chaperones, Tcj1-His and His-Tcj2, on the ability of His-TcHsp70 to bind RCMLA was tested by native PAGE. In addition, it has been proposed that Hsp40s may bind non-native substrates and deliver them to their respective Hsp70. The native PAGE and corresponding Western analysis shown in Figure 4.5 detail the effect of Tcj1-His on the binding of RCMLA by His-TcHsp70. Tcj1-His was not detectable with this particular anti-His antibody, and so only species containing His-TcHsp70 were detected. Five species of different mobility were observed (Figure 4.5 arrows 1-5). Low mobility species marked 1 – 4 (Figure 4.5 black arrows) were observed in all of the reactions containing His-TcHsp70 (Figure 4.5 Lanes 3 – 10), suggesting that these were the different oligomeric forms of His-TcHsp70, as observed in Figure 4.4. In the case of His-TcHsp70 and RCMLA (Figure 4.5 Lanes 7 and 8), the formation of these species appeared to be promoted by the presence of ATP. A decrease in the species with the lowest mobility (species 1) was observed in lanes 9 and 10, when His-TcHsp70 was combined with both Tcj1-His and RCMLA. This appeared to support the conclusion that His-TcHsp70 oligomerisies when in the absence of both substrate and co-chaperones. The drop in the level of this low mobility species would suggest the conversion of the largest oligomers into potentially smaller oligomeric or monomeric active species. Species 5 (Figure 4.5 blue arrow) is observed distinctly in Lanes 5 – 10. This was a potential complex between His-TcHsp70 and RCMLA observed previously (Figure 4.4, species 5). This species was also observed in the lanes containing His-TcHsp70 and Tcj1-His (Figure 4.5 Lanes 5 and 6), even though no RCMLA is present. This species may be a His-TcHsp70/Tcj1-His complex, in which Tcj1-His is being recognised as a co-chaperone or substrate and bound by His-TcHsp70 or may be a particular oligomeric state of His-TcHsp70 that was generated by the presence of substrate and/or Tcj1-His.

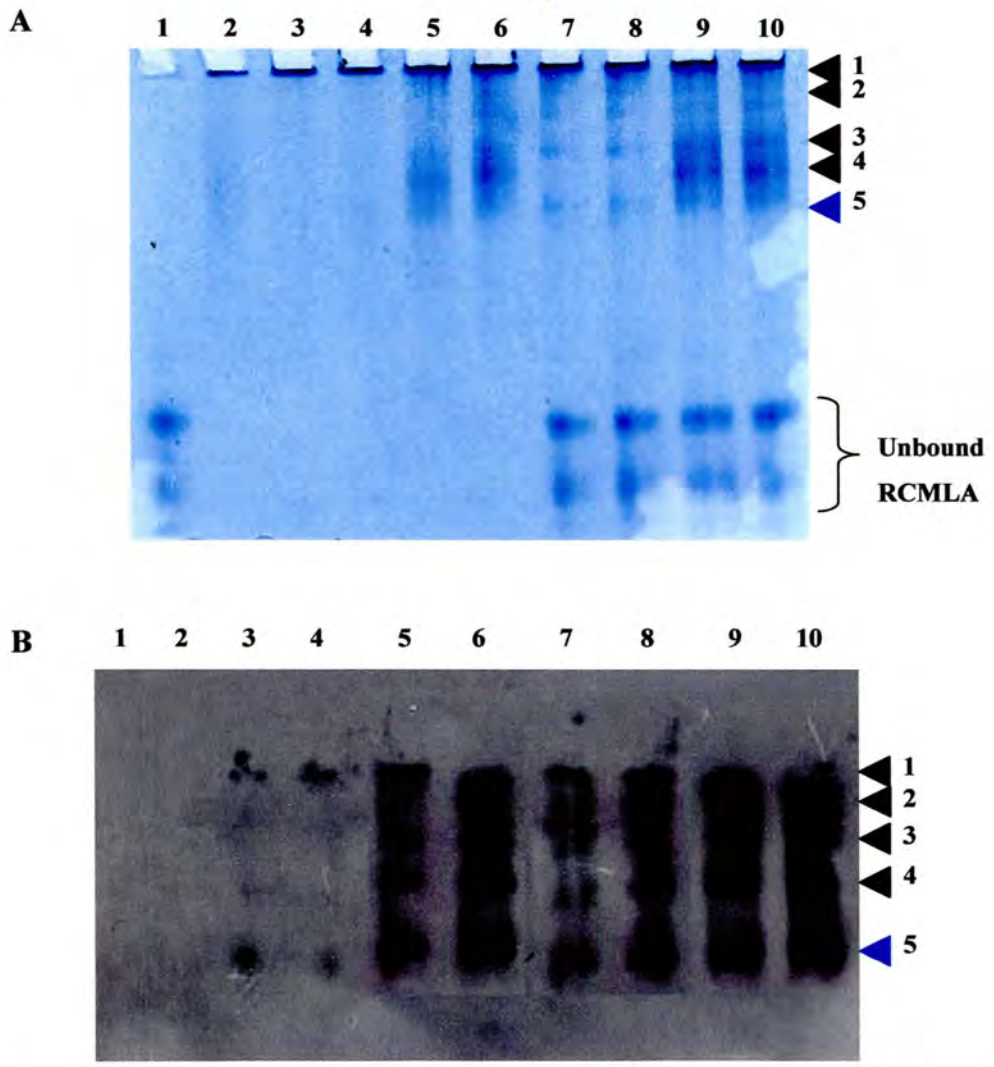


Figure 4.5: Analysis of RCMLA Binding by His-TcHsp70 in the presence of Tcj1-His

The ability of His-TcHsp70 to bind a model protein substrate (RCMLA) and the effect of the presence of Tcj1-His was investigated by Native/ Non-denaturing Polyacrylamide Gel Electrophoresis and Western blotting. Final concentration of His-TcHsp70 and RCMLA was 0.25 μg/μL and Tcj1-His was 0.125 μg/μL

(A) Coomassie stained Native PAGE of RCMLA binding by His-TcHsp70.Unbound RCMLA is also indicated.

(B) Western analysis of Native PAGE of RCMLA binding by His-TcHsp70. Species containing His-TcHsp70 were detected by chemiluminescence-based immunodetection using anti-6xHis antibody. The 6xHis tag on Tcj1-His is not detectable with this particular antibody. Black arrows indicate the change in levels of a number of species of different mobility containing His-TcHsp70. The blue arrow indicates the potential His-TcHsp70 species observed in Figure 4.4.

For (A) and (B): Lane 1 – RCMLA alone; Lane 2 – Tcj1-His alone; Lane 3 – His-TcHsp70 alone; Lane 4 – TcHsp70 in the presence of 600 μM ATP; Lane 5 – TcHsp70 and Tcj1-His; Lane 6 – TcHsp70 and Tcj1-His in the presence of 600 μM ATP; Lane 7 – His-TcHsp70 and RCMLA; Lane 8 – His-TcHsp70 and RCMLA in the presence of 600 μM ATP; Lane 9 – His-TcHsp70, Tcj1-His and RCMLA; Lane 10 – His-TcHsp70, Tcj1-His and RCMLA in the presence of 600 μM ATP.

Similar experiments were conducted to observe the effect of the presence of His-Tcj2 on the ability of His-TcHsp70 to potentially bind RCMLA (Figure 4.6). The ability to detect His-Tcj2 species specifically with the antibody to the C terminal peptide of Tcj2 made it possible to locate His-Tcj2 within the complexes. This gave a better indication of the interaction between His-Tcj2 and His-TcHsp70. Both His-TcHsp70 and His-Tcj2 were also detected by the anti-His antibody. Figure 4.6 shows the native PAGE and Western analysis for the analysis of RCMLA binding by His-TcHsp70 in the presence of His-Tcj2. The PAGE is not clear, but serves to indicate the presence of unbound RCMLA in the reactions (Figure 4.6 Lanes 7-10).

As observed in the previous experiments (Figures 4.4 and 4.5) a number of 6xHis-tagged species of different mobilities (Figure 4.6 species 1 – 6) were observed on the Western analysis using anti-6xHis antibody (Figure 4.6[B]). His-Tcj2 was detected (Figure 4.6 Lane 2) as a predominantly low mobility species. The species of His-TcHsp70 previously observed in Figure 4.4 were also observed (Figure 4.6 Lanes 3 and 4), although again in this case the formation of the high mobility species appears to be dependent on the presence of ATP (Figure 4.6 Lane 4). The formation of potential His-TcHsp70/RCMLA complexes (Figure 4.6 species 5) and the potential formation of the other species of His-TcHsp70 (Figure 4.6 species 1-3) observed in Figure 4.4 were visualised in lanes 7 and 8. What was of particular interest in this case, were the changes in the species observed in reactions containing a combination of His-TcHsp70 and His-Tcj2 (Figure 4.6 Lanes 5, 6, 9 and 10). Both His-TcHsp70 and His-Tcj2 were detectable with the anti-6xHis antibody, and so the species observed in Figure 4.6(B) resulted from the detection of either or both of the two 6xHis-tagged proteins. The corresponding species in which only His-Tcj2 was being detected are shown in the respective lanes of Figure 4.6(C). Observation of lanes 5 and 6 (His-TcHsp70 and His-Tcj2 reactions) in both figures show that the species of lower mobility contained His-Tcj2 species (and potentially His-TcHsp70 species) and that the higher mobility species contained only His-TcHsp70 species. The presence of His-Tcj2 appeared to promote the formation of an especially high mobility species (Figure 4.6 [B]species 6) containing only His-TcHsp70, and not His-Tcj2. This species disappeared in the presence of RCMLA (Figure 4.6 [B]Lanes 9 and 10). If species 6 was an active conformer of His-TcHsp70, it could potentially have been incorporated into a substrate/Hsp70 complex (potentially species 5) when RCMLA is present, explaining its disappearance (Figure 4.6 [B] Lanes 9 and 10).

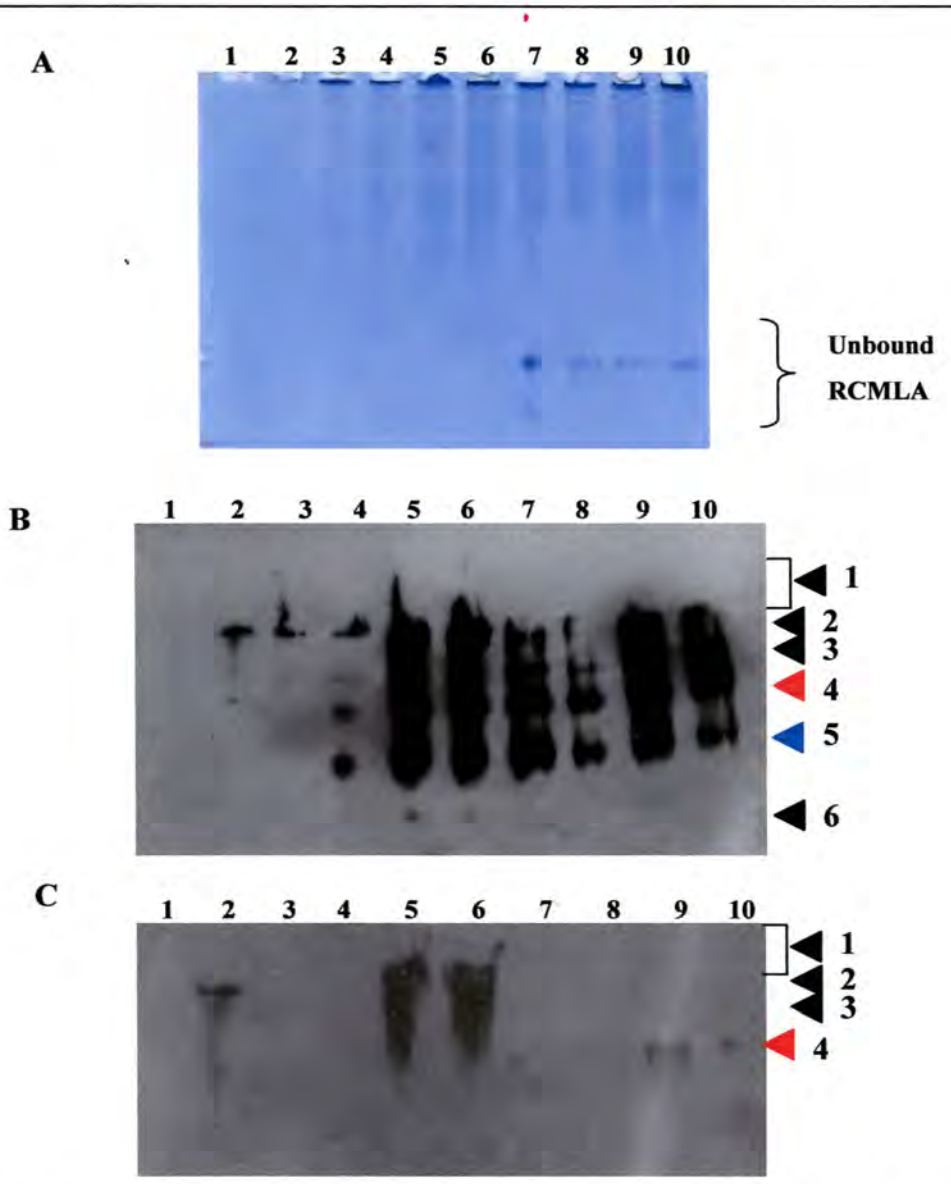


Figure 4.6: Analysis of RCMLA Binding by His-TcHsp70 in the presence of His-Tcj2

The ability of His-TcHsp70 to bind a model protein substrate (RCMLA) in the presence of His-Tcj2 was investigated by Native/ Non-denaturing Polyacrylamide Gel Electrophoresis and Western analysis. Final concentration of His-TcHsp70 and RCMLA was 0.25 μg/μL and His-Tcj2 was 0.125 μg/μL

(A) Coomassie stained Native PAGE of RCMLA binding by His-TcHsp70 in the presence of His-Tcj2. Unbound RCMLA is indicated in lanes 7-10.

(B) Western blot of Native PAGE of RCMLA binding by His-TcHsp70 in the presence of His-Tcj2. Species containing His-TcHsp70 and His-Tcj2 were detected by probing with anti-6xHis antibody.

(C) Western analysis showing species only containing His-Tcj2. The primary antibody used was designed specifically to detect a C-terminal amino acid sequence unique to Tcj2.

For (A), (B) and (C): Lane 1 – RCMLA alone; Lane 2 – His-Tcj2 alone; Lane 3 – His-TcHsp70 alone; Lane 4 – TcHsp70 in the presence of 600 μM ATP; Lane 5 – TcHsp70 and His-Tcj2; Lane 6 – TcHsp70 and His-Tcj2 in the presence of 600 μM ATP; Lane 7 – His-TcHsp70 and RCMLA; Lane 8 – His-TcHsp70 and RCMLA in the presence of 600 μM ATP; Lane 9 – His-TcHsp70, His-Tcj2 and RCMLA; Lane 10 – His-TcHsp70, His-Tcj2 and RCMLA in the presence of 600 μM ATP. Black arrows indicate the potential His-TcHsp70 oligomeric species previously observed in Figures 4.4 and 4.5. The blue arrow indicates a species, arising due to the presence of His-TcHsp70 and RCMLA, while the red arrow indicates a species containing His-Tcj2 arising in the presence of His-TcHsp70 and RCMLA.

A change in the mobility of the His-Tcj2 species is observed in the presence of both RCMLA and His-TcHsp70. The presence of His-TcHsp70 and a model non-native substrate lead to the formation of a higher mobility His-Tcj2 species (Figure 4.6 [C] species 4), also detected in Figure 4.6(B). This may suggest an interaction of His-Tcj2 with either His-TcHsp70 or RCMLA. It could be a complex between His-Tcj2 and RCMLA. However, a species of similar mobility was observed (Figure 4.6 [B] Lanes 5 and 6) where His-TcHsp70, but not RCMLA was present, suggesting that it could be a complex formed due to an interaction of His-TcHsp70 and His-Tcj2. This conclusion was supported by the finding that His-Tcj2 stimulated the ATPase Activity of His-TcHsp70 (discussed Section 4.3.1). There was also a possibility that this species 4 contained all three components, His-TcHsp70, His-Tcj2 and RCMLA, but without anti-RCMLA antibodies no conclusive statement in this regard could be made. In addition, similar binding studies with His-Tcj2, RCMLA and a native non-substrate such as BSA, in the absence of His-TcHsp70 would have to be conducted.

4.3.4 *In Vivo* Analysis of Chaperone Activity of Tcj2 and Tcj3 using *S. cerevisiae* strain JJ160

Temperature sensitive *S. cerevisiae ydj1* mutant strain JJ160 was used to investigate the ability of either Tcj2 or Tcj3 to functionally complement for Ydj1. The control plasmid pRS317-*YDJ1* encoding wild type Ydj1 and other plasmids encoding functional equivalents and homologues of Ydj1 rescue the temperature sensitivity of *S. cerevisiae* JJ160 and allow growth at temperatures (30 and 34°C) above the permissive temperature (23°C).

The coding sequences for Tcj2 and Tcj3 were inserted into the low copy number yeast expression plasmid pKG6. The identity of the pKG6/Tcj2 and pKG6/Tcj3 plasmid constructs was confirmed by restriction enzyme analysis prior to transformation into *S. cerevisiae* JJ160. The pKG6/Tcj2 plasmid was linearised by digestion with each of *Eco*RI and *Xho*I (6563bp). A digest with both of these enzymes released the Tcj2 coding region (1209bp) and the pKG6 vector fragment (5354bp) (Figure 4.7 [A]). The pKG6/Tcj3 plasmid was linearised by digestion with *Bam*HI and *Eco*RI (6904bp), while a digest with both of these enzymes released the Tcj3 coding region (1238bp) and the vector fragment (5666bp). In addition pKG6/Tcj3 was digested by *Sma*I to produce two fragments (374bp and 6530bp) and *Pst*I to produce three fragments (834bp, 2362bp and 3708bp) (Figure 4.7 [B]).

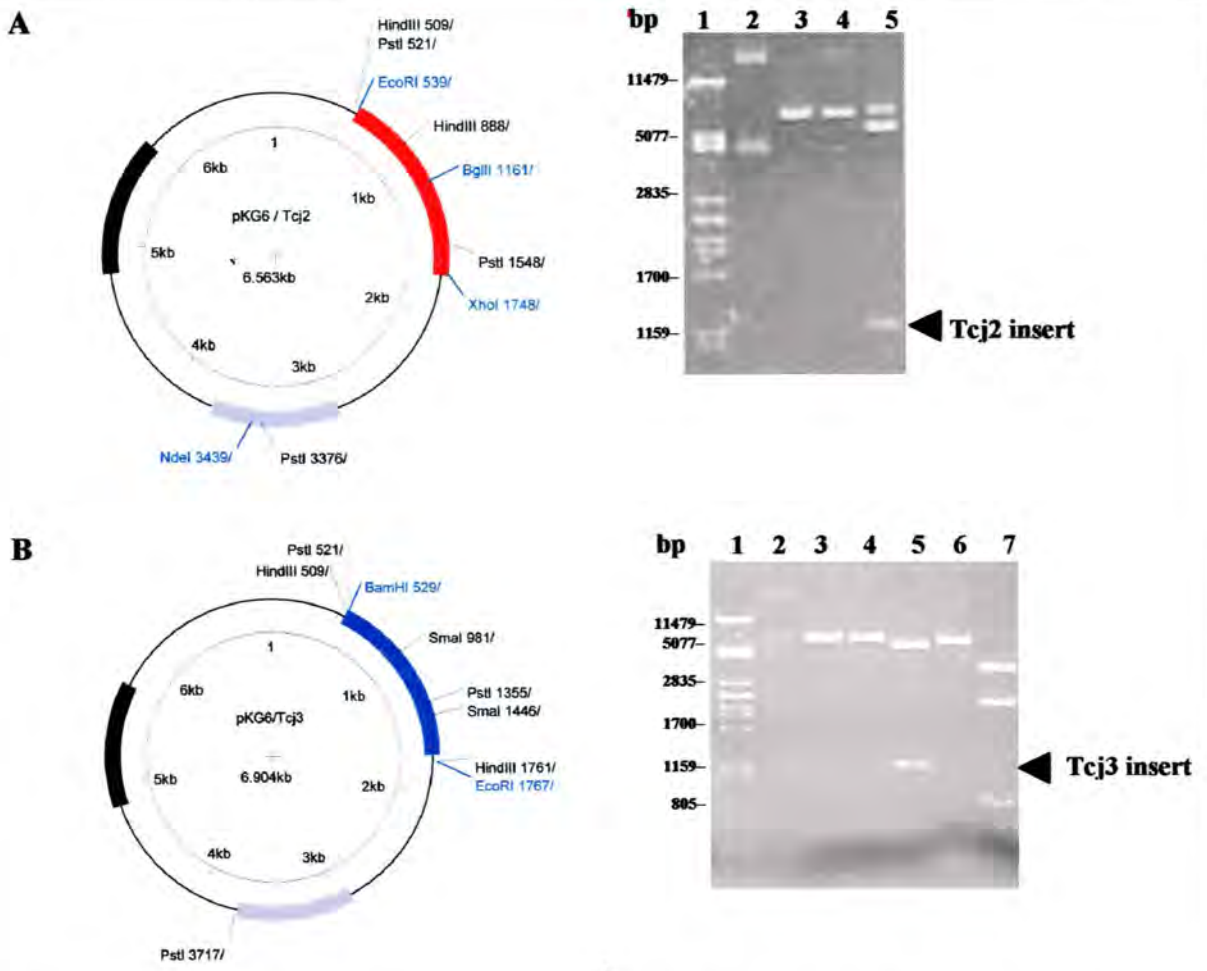


Figure 4.7: Confirmation of the Identity of the pKG6/Tcj2 and pKG6/Tcj3 Plasmids

(A) Plasmid map and 0.8% agarose gel confirming identity of pKG6/Tcj2 by restriction enzyme analysis. Tcj2 coding region (red) was inserted into *Eco*RI and *Xho*I sites of the plasmid pKG6. The vector carries an ampicillin resistance coding region (black) for selection in *E.coli* and a URA3 marker (grey) for selection in yeast. Lane 1 – Lambda *Pst*I digest marker (sizes shown in bp); Lane 2 – Uncut pKG6/Tcj2; Lane 3 –*Eco*RI digest; Lane 4 –*Xho*I digest; Lane 5 – *Eco*RI/*Xho*I digest releasing Tcj2 insert.
(B) Plasmid map and 0.8% agarose gel confirming identity of pKG6/Tcj3 by restriction enzyme analysis. Tcj3 coding region (blue) was cloned into *Bam*HI and *Eco*RI sites of the plasmid pKG6. The vector carries an ampicillin resistance coding region (black) for selection in *E.coli* and a URA3 marker (grey) for selection in yeast. Lane 1 – Lambda *Pst*I digest marker (sizes shown in bp); Lane 2 – Uncut pKG6/Tcj3; Lane 3 –*Bam*HI digest; Lane 4 –*Eco*RI digest; Lane 5 – *Bam*HI/*Eco*RI digest releasing Tcj3 insert; Lane 6 –*Sma*I digest; Lane 7 – *Pst*I digest.
In both cases, restriction enzymes used in the analysis were chosen to yield a pattern distinct to each of the constructs.

Serial dilutions of *S. cerevisiae* JJ160 transformed with the respective plasmids were plated on to selective media and incubated at the different temperatures to examine the ability of Tcj2 and/or Tcj3 to complement for the absence of Ydj1 (Figure 4.8).

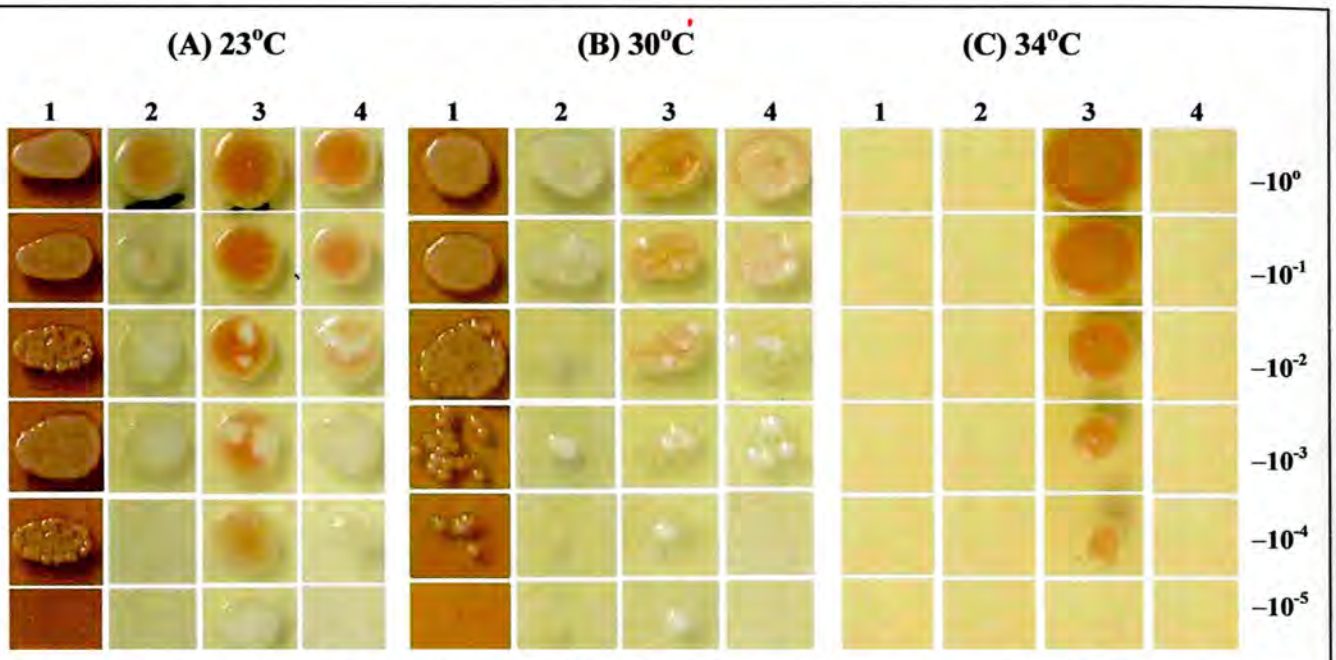


Figure 4.8: Analysis of the ability of Tcj2 and Tcj3 to functionally replace Ydj1 in the *ydj1* mutant strain *S. cerevisiae* JJ160

The ability of either Tcj2 or Tcj3 to complement for the absence of the yeast chaperone Ydj1 was analysed using the *ydj1* mutant *S. cerevisiae* strain JJ160. Serial dilutions (shown to the right) of overnight cultures were plated on to selective media (YMM galactose [HIS^{-} URA^{-}]) for pKG6 constructs and (YMM galactose [$LYS^{-}HIS^{-}$]) for pRS317-*YDJ1*] and incubated at (A) 23°C, (B) 30°C, (C) 34°C, to select for complementation of the absence of Ydj1.
For (A), (B) and (C): Lane 1 – *S. cerevisiae* JJ160[pRS317-*YDJ1*] (Ydj1 positive control); Lane 2 – *S. cerevisiae* JJ160[pKG6] (negative control); Lane 3 – *S. cerevisiae* JJ160[pKG6/Tcj2]; Lane 4 – *S. cerevisiae* JJ160[pKG6/Tcj3].

Figure 4.8(A) showed that growth of all the transformants (*S. cerevisiae* JJ160[pRS317-YDJ1] positive control, *S. cerevisiae* JJ160[pKG6] negative control, *S. cerevisiae* JJ160[pKG6/Tcj2] producing Tcj2 and *S. cerevisiae* JJ160[pKG6/Tcj3] producing Tcj3) occurred at the permissive temperature of 23°C (Figure 4.8 [A] Lanes 1 to 4 respectively). At this temperature *S. cerevisiae* JJ160 did not require Ydj1 to survive. However, raising the temperature (analogous to a heat shock) has been shown to result in the death of *S. cerevisiae* JJ160, unless Ydj1 or a functional equivalent is present. Growth of *S. cerevisiae* JJ160[pKG6], the vector transformed negative control strain, was impaired at 30°C and completely absent at 34°C (Figure 4.8 [B] and [C] Lane 2). *S. cerevisiae* JJ160[pRS317-*YDJ1*], the positive control strain producing Ydj1, was able to grow at 30°C but unable to grow at 34°C (Figure 4.8 [B] and [C] Lane 1). *S. cerevisiae* JJ160[pKG6/Tcj3] strain producing Tcj3 was not able to survive at 34°C and showed impaired growth at 30°C (Figure 4.8 [B] and [C] Lane 4). The growth pattern of *S. cerevisiae* JJ160[pKG6/Tcj3] resembled that of the negative control, *S. cerevisiae* JJ160[pKG6] (Figure 4.8 [B] and [C] Lanes 4 and 2 respectively). The lack of complementation by Tcj3 suggests that this protein was not

functioning as a Ydj1 homologue. However, this did not exclude the possibility that Tcj3 is a co-chaperone for TcHsp70. The lack of purified His-Tcj3 restricted any analysis of its ability to interact with His-TcHsp70 as a co-chaperone. As can be observed, *S. cerevisiae* JJ160[pKG6/Tcj2] producing Tcj2 (Figure 4.8 [B] and [C] Lane 3) was able to survive at 30°C and 34°C even at a dilution of 10^{-5}, supporting growth better than the positive control (Figure 4.8 [B] and [C] Lane 1). Figure 4.8 displays the ability of Tcj2 to complement for the absence of the yeast Hsp40, Ydj1, and rescue *S. cerevisiae* JJ160 from temperature sensitivity. The production of Tcj2 in *S. cerevisiae* JJ160[pKG6/Tcj2] was confirmed by Western analysis using the Tcj2 specific antibody (Figure 4.9)

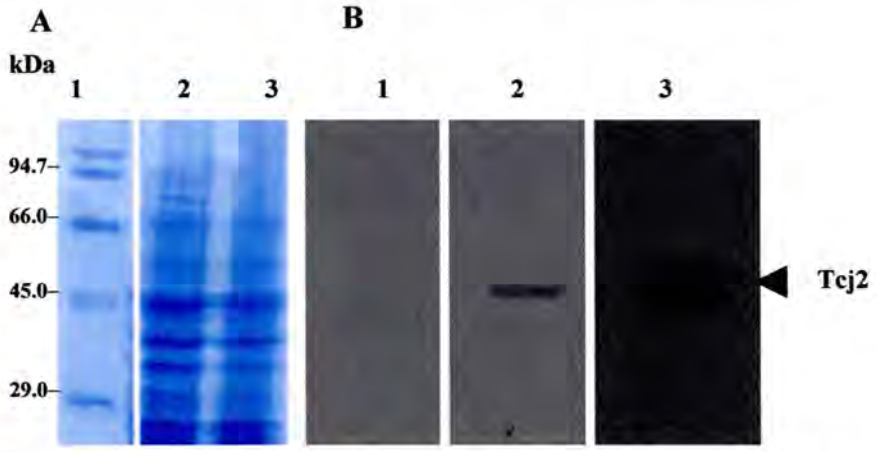


Figure 4.9: Detection of Tcj2 Production in *S.cerevisiae* JJ160[pKG6/Tcj2]

The production of Tcj2 in the *S.cerevisiae* JJ160[pKG6/Tcj2] complementation cultures was confirmed by Western analysis and chemiluminescence-based immunodetection with the antibody to the C-terminal peptide of Tcj2.
(A) Coomassie stained SDS-PAGE. Lane 1 – broad range molecular mass markers; Lane 2 – *S. cerevisiae* JJ160 whole cell extract; Lane 3 – *S. cerevisiae* JJ160[pKG6/Tcj2] whole cell extract.
(B) Western analysis; Lane 1 – *S. cerevisiae* JJ160 whole cell extract (negative control); Lane 2 – *S. cerevisiae* JJ160[pKG6/Tcj2] whole cell extract; Lane 3 – Purified His-Tcj2 (~3µg, positive control).

Purified His-Tcj2 (3µg) was used as a positive control and to make a preliminary examination of the levels of expression of Tcj2 in *S. cerevisiae* JJ160[pKG6/Tcj2]. Tcj2 was detected in the protein extract prepared from the complemented *S. cerevisiae* JJ160[pKG6/Tcj2] culture. The levels of protein detected are significantly lower than that in the purified sample (Figure 4.9 [B]). In addition, Tcj2 was not detected as a predominant protein band on SDS-PAGE analysis of the protein profile of *S. cerevisiae* JJ160[pKG6/Tcj2] (Figure 4.9 [A] Lane 3) compared to that of *S. cerevisiae* JJ160 (Figure 4.9 [A] Lane 2). This suggests that Tcj2 in itself and not over-production of a potential chaperone was responsible for the complementation of *S. cerevisiae* JJ160.

This result signifies that Tcj2 was able to functionally complement for the absence of Ydj1. This result provided evidence that Tcj2 was a chaperone/co-chaperone and suggested that Tcj2 may perform similar functions or interact with similar proteins as does Ydj1. The alignment of the primary sequences of Tcj2 and Ydj1 revealed that the proteins show a significant level of amino acid identity and similarity (Figure 4.10).

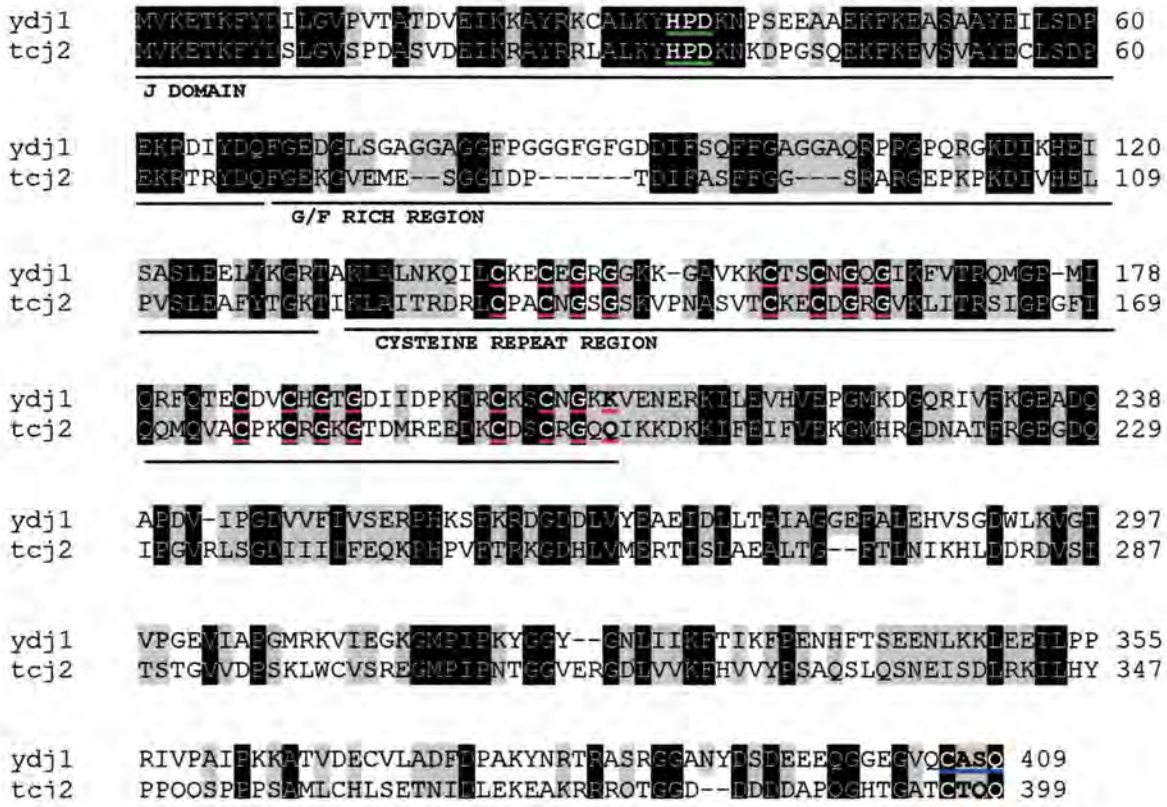


Figure 4.10: Alignment of the Primary Amino Acid Sequences of Tcj2 and Ydj1

Comparison of the amino acid sequence of Tcj2 and Ydj1 reveals a significant identity and similarity between the two primary amino acid sequences, in both the J domain (residues 1-70,) and C terminal regions. The positions of J domain, G/F rich region and cysteine repeat region are indicated (underlined). The highly conserved HPD motif in the J domain is underlined in green, the cysteine repeats are underlined in pink. Both Tcj2 and Ydj1 have a deviation from the consensus sequence in their final cysteine repeat region. Instead of a gly, Tcj2 has a gln and Ydj1 has a lys. The C terminal prenylation site (CaaX) is underlined in blue. This motif is the potential site of prenylation, allowing the proteins to anchor in membranes. Black shading indicates identical residues, while grey shading indicates similar residues.

The similarity between Tcj2 and Ydj1 was displayed throughout the sequences and not confined to certain motifs (Figure 4.10). The similarity between the N-terminal regions of the two proteins was expected, as this area constituted the J domain, which was a conserved motif found in all Hsp40s, not all of which have similar functions. The conservation of sequence suggests that the J domain of Tcj2 may be similar to the J domain of Ydj1 and may be able to interact with the partner yeast Hsp70 of Ydj1 (Ssa1).

The similarity between the C-terminal regions of Tcj2 and Ydj1 was also significant. The G/F rich regions of the two Hsp40s are similar. This linker region has been proposed to govern specificity of the interaction between Hsp40 and Hsp70 and therefore could direct the interaction of Tcj2 with Ydj1's partner Hsp70, Ssa1, allowing Tcj2 to function in place of the missing Ydj1. In addition, the similarity between the cysteine repeat regions of Tcj2 and Ydj1 suggests that Tcj2 may be able to bind substrates normally bound by Ydj1. Previous work by Johnson and Craig (2001) indicated that the C-terminal regions of yeast Hsp40 are essential for the functioning of the Hsp40s in *S. cerevisiae*. In addition, both of the Hsp40s contain the C-terminal CaaX motif. The sequence similarity between Tcj2 and Ydj1 was consistent with the ability of Tcj2 to complement the Ydj1-deficient *S. cerevisiae* JJ160.

4.3.5 *In Vivo* Analysis of Chaperone Activity of Tcj2 using *S. cerevisiae* strain WY26

The ability of Tcj2 to complement for the absence of the Type II yeast Hsp40 Sis1, was investigated using the *S. cerevisiae sis1* mutant strain WY26. Sis1 is an essential Hsp40 in yeast and thus the *S. cerevisiae* strain WY26 contains the plasmid pYW17 encoding the *SIS1* gene and a URA3 marker for viability. *S. cerevisiae* WY26 can be cured of pYW17 by growth in the presence of 5-fluoroorotic acid (5-FOA), a toxic uracil analogue, leading to death of the cells, in the absence of wild type Sis1 or a Sis1 functional equivalent. Lethality on 5-FOA is rescued by the presence of Sis1 functional equivalents or the control plasmid pYW98-*SIS1*, encoding wild type Sis1.

The identity of the pKG4 plasmid used in the complementation was confirmed by restriction enzyme analysis as before (Figure 4.11). The pKG4/Tcj2 plasmid was linearised by digestion with *Eco*RI and *Xho*I (6454bp) and a digestion with both the enzymes released the Tcj2 coding region (1209bp) and the vector fragment (5245bp). The pKG4/Tcj2 plasmid was digested by *Hind*III to produce three fragments (379bp, 1789 and 4286bp) and by *Sma*I to produce two fragments (509bp and 5945bp) (Figure 4.11).

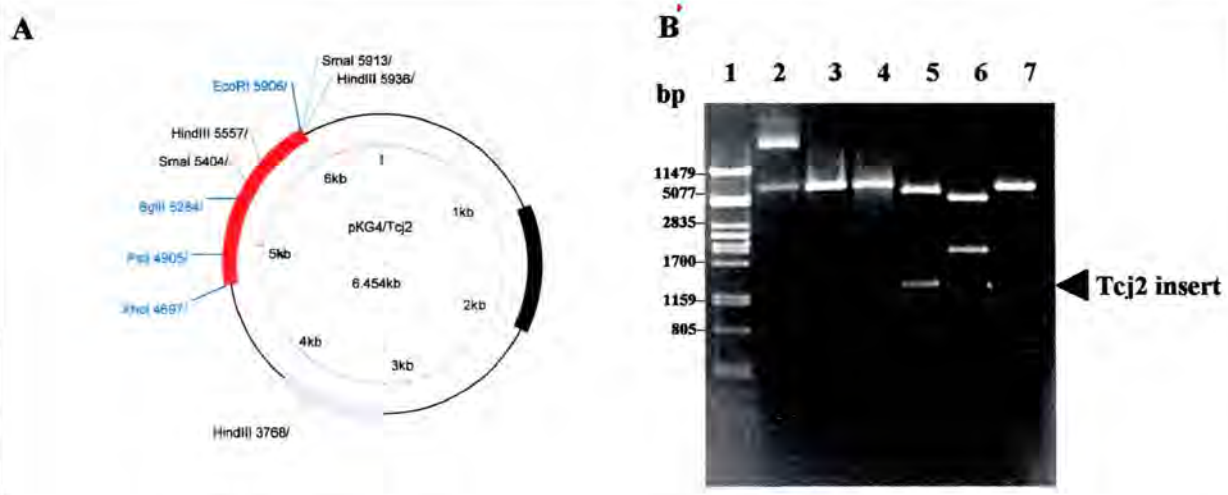


Figure 4.11: Confirmation of the Identity of the pKG4/Tcj2 Plasmid

(A) Plasmid map and (B) 0.8% agarose gel confirming the identity of the pKG4/Tcj2 plasmid by restriction digest analysis. The Tcj2 coding sequence (red) was inserted into *Eco*RI and *Xho*I sites. The plasmid carries an ampicillin resistance coding sequence (black) for selection in *E.coli* and a *TRP1* marker for selection in *S. cerevisiae*.
Restriction enzymes used in analysis were chosen to produce a pattern distinct to pKG4/Tcj2 construct. Lane 1 – *Pst*I digest of Lambda DNA used as markers (sizes shown in bp); Lane 2 – Uncut pKG4/Tcj2; Lane 3 – *Eco*RI digest; Lane 4 – *Xho*I digest; Lane 5 – *Eco*RI/*Xho*I digest releasing Tcj2 insert; Lane 6 – *Hind*III digest; Lane 7 – *Sma*I digest.

Serial dilutions of cultures of *S. cerevisiae* WY26[pKG4] (negative control), *S. cerevisiae* WY26[pYW98-*SIS1*] (positive control) and *S. cerevisiae* WY26[pKG4/Tcj2] were plated on to selective plates without 5FOA (Figure 4.12[A]) and with 5FOA (Figure 4.12[B]) and incubated at 30°C. In the absence of 5FOA all of the transformants, *S. cerevisiae* WY26[pKG4] negative control, *S. cerevisiae* WY26[pYW98-*SIS1*] positive control and *S. cerevisiae* WY26[pKG4/Tcj2], were capable of growth, which in this case was being supported by the presence of the pYW17 plasmid encoding wild type Sis1 (Figure 4.12[A]). In the presence of 5FOA (Figure 4.12[B]), *S. cerevisiae* WY26 was cured of the pWY17 plasmid, allowing for the presence of the test plasmids (pKG4 and pKG4/Tcj2) to be evaluated for their ability to support growth. The negative control vector transformed strain *S. cerevisiae* WY26[pKG4] did not grow in the presence of 5FOA (Figure 4.12[B] Lane 2). The positive control transformed strain *S. cerevisiae* WY26[pYW98-*SIS1*] is only capable of supporting growth at low dilutions in the presence of 5FOA (Figure 4.12[B] Lane 1). *S. cerevisiae* WY26[pKG4/Tcj2] is not capable of growth in the presence of 5FOA and resembles the negative control (Figure 4.12[B] Lane 3). This suggests that Tcj2 is not capable of functionally replacing Sis1, and therefore Tcj2 is not a functional equivalent of Sis1.

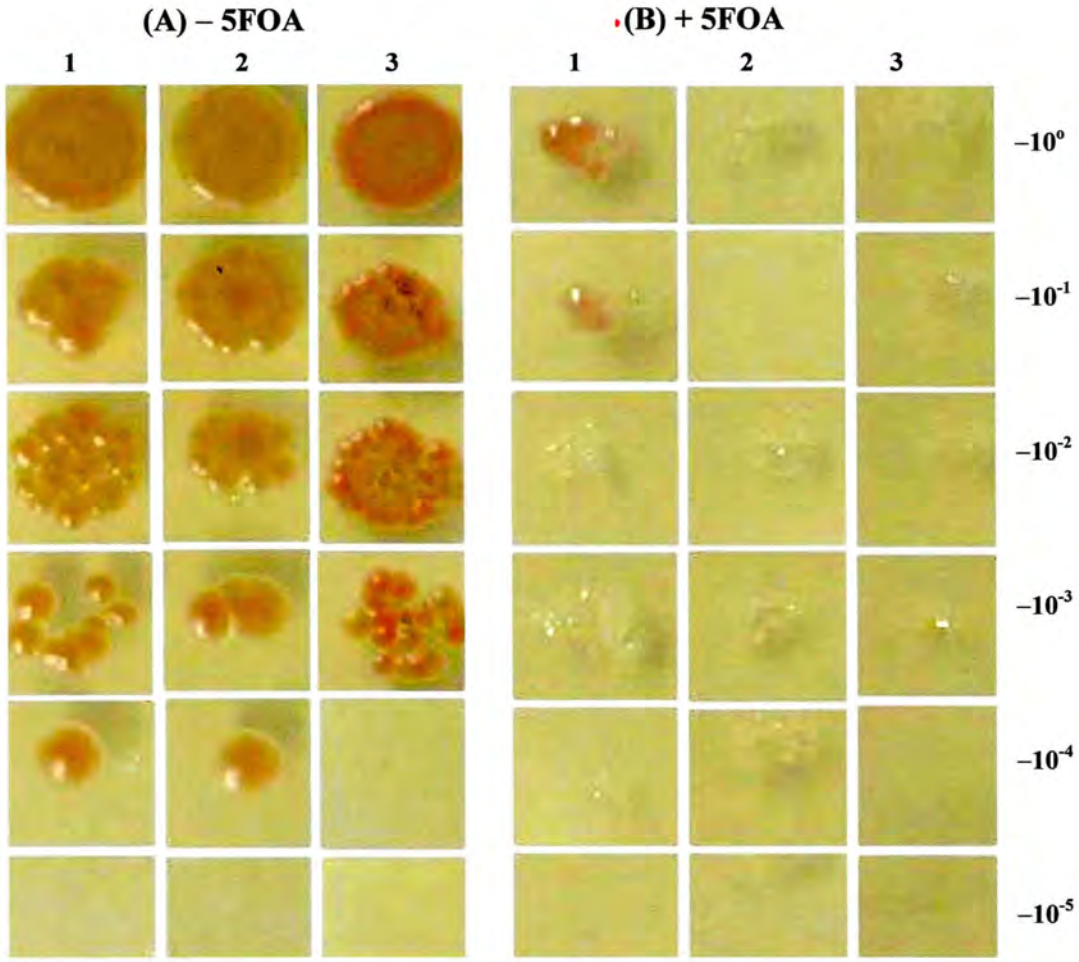


Figure 4.12: Analysis of the ability of Tcj2 to functionally replace Sis1 in the *sis1* mutant strain *S. cerevisiae* WY26

The ability of Tcj2 to complement for the absence of the yeast Hsp40, Sis1, was testing using the sis1 mutant *S. cerevisiae* strain WY26. Serial dilutions (shown to the right of the figure) were plated on to (A) yeast minimal medium without 5FOA (YMM galactose [$LEU^{-}TRP^{-}$]) to detect normal growth and (B) yeast minimal medium (YMM [$LEU^{-}TRP^{-}$]) supplemented with 0.1% 5-FOA, to select for plasmids capable of supporting the growth of *S. cerevisiae* WY26.
In both (A) and (B): Lane 1 – *S. cerevisiae* WY26[pYW98-SIS1] (positive control, expressing wild type Sis1); Lane 2 – *S. cerevisiae* WY26[pKG4] (negative control); Lane 3 – *S. cerevisiae* WY26[pKG4/Tcj2].

Western analysis with the antibody specific to Tcj2 was used to detect the levels of production of the protein within *S. cerevisiae* WY26[pKG4/Tcj2] (Figure 4.13). However, no signal on the Western analysis corresponding to Tcj2 was observed in the *S. cerevisiae* WY26[pKG4/Tcj2] protein extract (Figure 4.13 [B] Lane 1). In addition, a predominant band corresponding to Tcj2 is not observed on the SDS-PAGE analysis, which could be due to low production of Tcj2 within this particular strain.

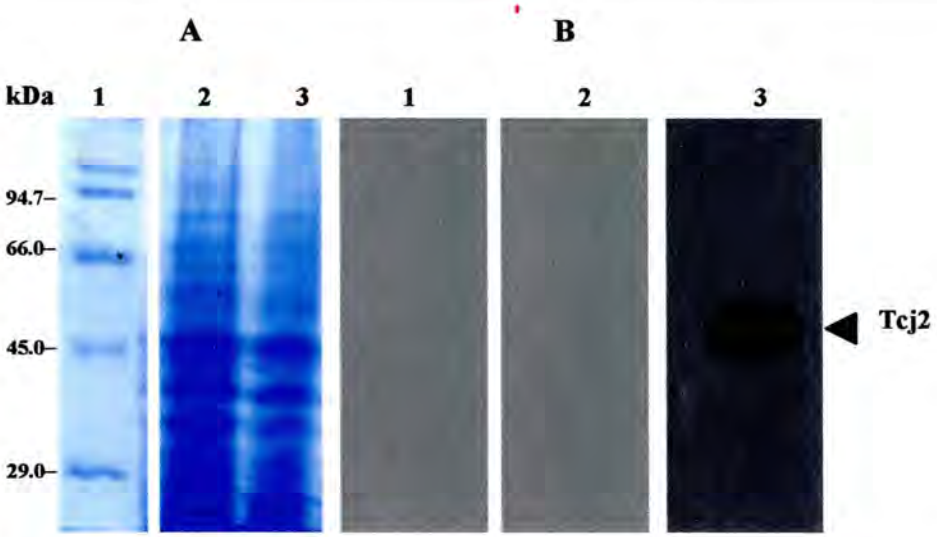


Figure 4.14: Detection of Tcj2 Production in *S. cerevisiae* WY26[pKG4/Tcj2]

The expression of Tcj2 in the *S. cerevisiae* WY26[pKG4/Tcj2] complementation cultures was detected by Western analysis and chemiluminescence-based immunodetection with the antibody to the C-terminal peptide sequence in Tcj2.
(A) Coomassie stained SDS-PAGE. Lane 1 – broad range molecular mass markers (sizes shown in kDa); Lane 2 – *S. cerevisiae* WY26 whole cell extract (negative control); Lane 3 – *S. cerevisiae* WY26[pKG4/Tcj2] whole cell extract.
(B) Corresponding Western analysis. Lane 1 – *S. cerevisiae* WY26 whole cell extract (negative control); Lane 2 – *S. cerevisiae* WY26[pKG4/Tcj2] whole cell extract; Lane 3 – Purified His-Tcj2 (~3μg, positive control).

Despite the lack of detection of Tcj2, it seemed unlikely that Tcj2 would complement for the absence of Sis1, even though Sis1 and Ydj1 have some functional overlap within the yeast cell. The presence of Sis1 was shown to overcome the temperature sensitivity of *S. cerevisiae* JJ160, but Ydj1 cannot complement for the absence of the essential Sis1 in *S. cerevisiae* WY26 (Johnson and Craig, 2001). If Tcj2 was a true Ydj1 homologue, as suggested by the *S. cerevisiae* JJ160 complementation, it would not be expected to complement for the absence of Sis1. If Tcj2 was a Sis1 homologue, it could potentially be able of complementation of both *S. cerevisiae* JJ160 and *S. cerevisiae* WY26. While Ydj1 and Sis1 share functional similarity, in that both of their J domains can interact with Ssa1, it is their substrate binding domains that govern their particular functions within the yeast cell (Johnson and Craig, 2001). The alignment of Tcj2 and Ydj1 showed that the two proteins shared similarity throughout their sequences. By contrast, Tcj2 and Sis1 share a high degree of similarity in the N-terminal region, but low similarity in the C-terminal region (Figure 4.14). The J domains of Tcj2 and Sis1 were similar, suggesting that Tcj2 would be capable of interaction with Ssa1, as Sis1 does. However, the difference in the substrate binding regions would potentially lead to the inability of Tcj2 to chaperone Sis1 substrates.

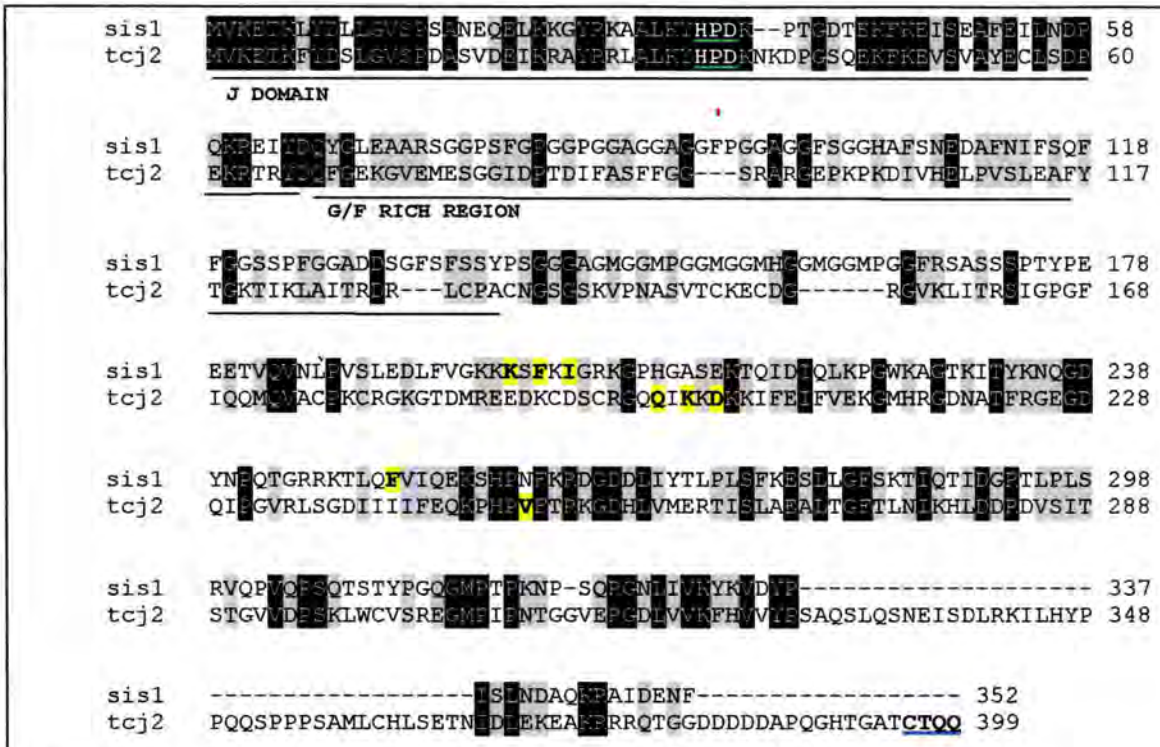


Figure 4.14: Alignment of Primary Amino Acid Sequences of Tcj2 and Sis1

The primary sequence alignment of Tcj2 and Sis1 reveals that the proteins share a high degree of similarity in the N terminal region containing the J domain (residues 1-70). The regions of the protein sequence comprising the J domain and G/F rich regions are indicated (underlined). The highly conserved HPD motif is underlined in green. The similarity decreases in the C terminal (substrate binding) regions of the two proteins. Residues highlighted in black are identical and those in grey are similar. The C-terminal prenylation site of Tcj2 is underlined in blue. Residues highlighted in yellow indicate those in Sis1 and the corresponding residues in Tcj2, shown to be essential for polypeptide binding by Sis1 (Lee et al., 2002)

The difference between the C-termini of Tcj2 and Sis1 is not surprising, since Sis1 was a Type II J protein, containing only the J domain and G/F region. Tcj2 by contrast was a Type I Hsp40, containing the J domain, G/F rich region and cysteine repeat region, making the two proteins structurally and, potentially functionally, distinct. In addition, work by another group has shown that Tcj6, a Type II Hsp40 from *T. cruzi*, can complement for the absence of Sis1 (Salmon et al., 2001).

4.4 CONCLUSION

The basal ATPase activity of His-TcHsp70 was determined as ~40 nmol/min/mg. Maximum stimulation of this ATPase activity was observed in the presence of His-Tcj2 and RCMLA (~60 nmol/min/mg). His-TcHsp70 was shown to form discrete species of different mobilities in the presence of RCMLA, His-Tcj2 and Tcj1-His, which may be indicative of His-TcHsp70 oligomers or complexes with these proteins. Tcj2, but not Tcj3, was able to function as a Ydj1 equivalent in *S. cerevisiae* JJ160. Tcj2 could not function as a Sis1 equivalent in *S. cerevisiae* WY26.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

5.1 CONCLUSIONS

5.1.1 Proposed Model for Potential Roles of Trypanosomal Heat Shock Proteins in *T. cruzi*

Based on the conclusions reached in this study, we have suggested a potential model for the interactions of the trypanosomal Hsps within the parasite (Figure 5.1). The potential roles of the trypanosomal Hsps in the cell and the possible interactions occurring have been suggested. This model is speculative, partly based on literature and on the data generated in this study, and will be used to design future experiments to further characterise these proteins and to fully elucidate their potential functions.

5.1.2 Bioinformatic Analysis of Trypanosomal Heat Shock Proteins

The bioinformatic analysis of TcHsp70 primary amino acid sequence revealed that it contained the conserved canonical domains defined by all Hsp70s. There was little deviation from the consensus sequence in the N-terminal ATPase domain of TcHsp70 when compared to that of other Hsp70s (Bork et al., 1992). In addition, TcHsp70 contained the generally conserved residues in its substrate binding regions. The C-terminus sequence of TcHsp70, proposed to be involved in interactions with the protein STI1, showed a deviation from the consensus sequence. Where most of Hsp70s contain the sequence GTPIEEVD, TcHsp70 contained the sequence GTEVEEVD at its C-terminus. The TcSTI1 sequence also showed a change in the residues governing the specificity of the Hsp70 – STI1 interaction (Brinker et al., 2001). These deviations in sequence in the two proteins may suggest an alternate specificity of binding between TcHsp70 and TcSTI1. This could be relevant to the formation of the TcHsp70/TcSTI1/TcHsp90 complexes in the parasite. Experimental investigation of the interaction of TcHsp70 with STI1 proteins containing the consensus sequence and TcSTI1 would be required to support this suggestion.

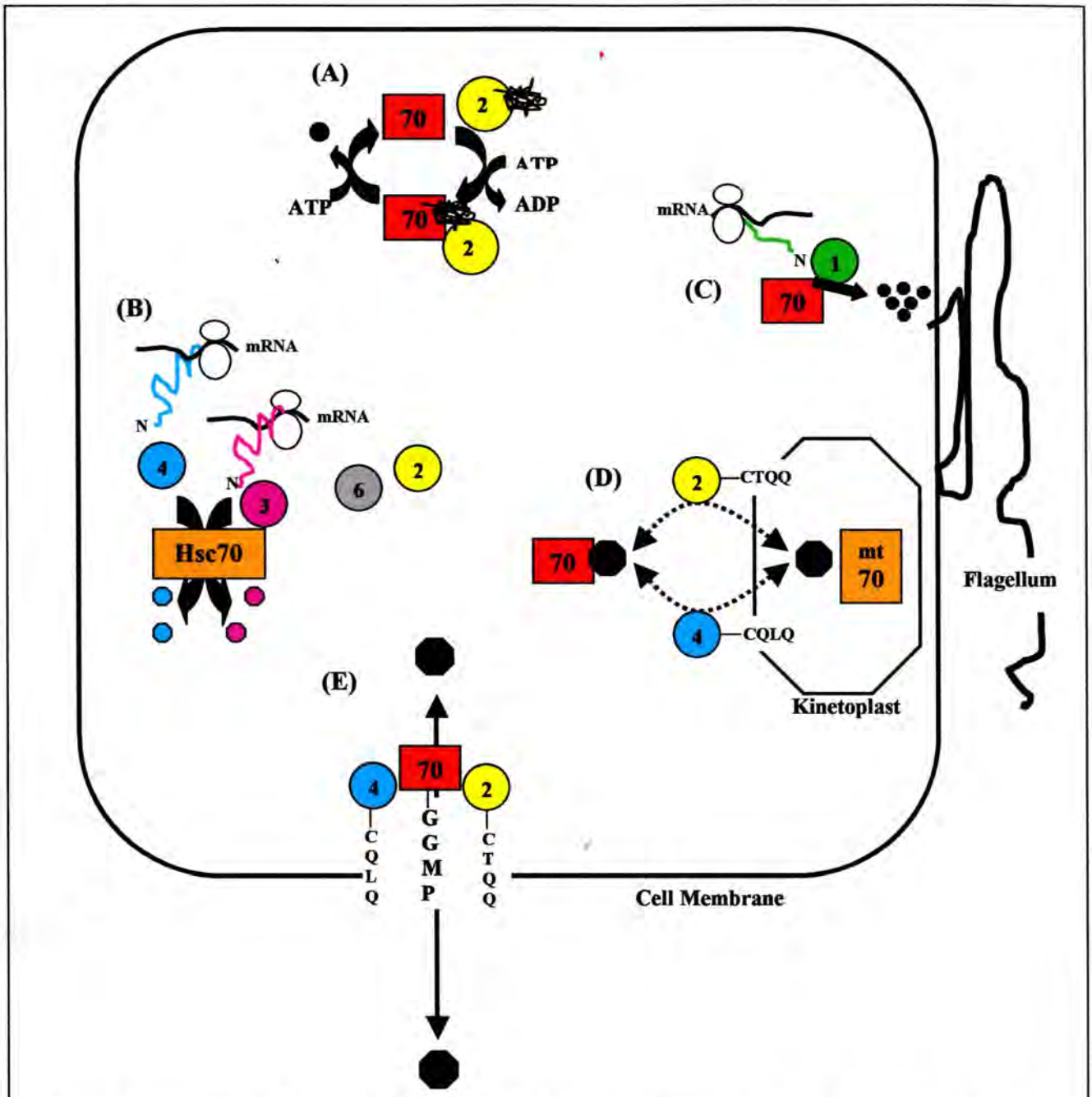


Figure 5.1: Proposed Model of Potential Roles of Trypanosomal Heat Shock Proteins in *T. cruzi*

The figure above represents the parasitic cell and the potential roles of TcHsp70, Tcj1, Tcj2, Tcj3 and Tcj4. This is a speculative model, based partly on proposals in literature and partly on the conclusions reached in this study, and will be used to design future experiments to characterise and test the potential functions of these trypanosomal Hsps.
(A) TcHsp70/Tcj2 function as molecular chaperones in the ATP/ADP regulated process of assisted folding of denatured trypansomal proteins. (B) TcHsp70 forms different partnerships with other Tcjs, such as Tcj3 or Tcj4, for the assisted folding of specific, distinct trypanosomal proteins. (C) TcHsp70 and Tcj1 interact to fufill a specific function in the cell, such as flagellar assembly. (D) TcHsp70 interacts with membrane associated Tcj2 and/or Tcj4 for translocation of trypanosomal proteins into different organelles. (E) Membrane associated TcHsp70 interacts with membrane associated Tcj2 and/or Tcj4 for the processing and translocation of trypanosomal and host proteins in and out of the parasitic cell. Circles indicates Tcj proteins (number indicates the respective one, Tcj1- 1, Tcj2- 2 etc.) The squares represents trypanosomal Hsp70s. Unless labelled otherwise, 70 signifies TcHsp70.

In addition, TcHsp70 contained an extended GGMP repeat in its C-terminal region. This region has been shown to be highly antigenic and predicted to be a potential transmembrane like region (Kumar and Zheng, 1998). If this GGMP region was capable of associating with the cellular membrane, it may be exposed to the host cell environment on the surface of the parasite. This could suggest the reason for it being antigenic, as it would come into contact with the host immune system. This association of TcHsp70 with the membrane may be required for the translocation and processing of host and/or trypanosomal proteins (Polla, 1988; Arispe and DeMaio, 2000) (Figure 5.1 [D] and [E])

Tcj2, Tcj3, and Tcj4 were classified as Type I Hsp40s and Tcj1 as a Type III Hsp40. All of the proteins show sequence similarity in their J domains. Tcj3 was the only of the three Type I proteins to contain the consensus sequence in its substrate binding domain. This suggested that Tcj3 may have a certain substrate specificity, which was different from the substrate specificities of Tcj2 and Tcj4. Tcj2 and Tcj4 contained deviations in their substrate binding motifs, suggesting that the structure of their substrate binding domains may be different from that of Hsp40s containing the consensus sequence. This potential difference in substrate binding domains may affect the substrate specificities of Tcj2 and Tcj4. It could be possible that the potential change in structure of their substrate binding domains would be suited to the types of substrate with which Tcj2 and Tcj4 associate. The potential different substrate specificities may be required in the parasitic system for the chaperoning of certain proteins that may have specific functions. Tcj2 and Tcj4 also contained the C-terminal prenylation motif CaaX. This suggested the potential of these proteins to be modified to allow association with cellular membranes (Marshall, 1993; Glomset et al., 1990). This potential association could be required for the translocation of certain trypanosomal proteins across cellular membranes (Figure 5.1 [D] and [E]).

5.1.3 Over-production and Purification of 6xHis-tagged Trypanosomal Heat Shock Proteins

The trypansomal Hsps were recombinantly produced as 6xHis fusion proteins in *E. coli*. His-TcHsp70, Tcj1-His and His-Tcj2 were successfully over-produced and purified by Ni-affinity chromatography in a sufficiently native state to allow function analysis. The problems encountered with the purification of Tcj3-His and His-Tcj4 made functional analyses of these

proteins not possible. The purification of Tcj3-His and His-Tcj4 could be attempted in alternate host systems, such as *S. cerevisiae*, or with alternative fusion tags, such as the GST affinity tag, in an attempt to purify the proteins for functional studies.

5.1.4 Analysis of Chaperone Activity of Trypanosomal Heat Shock Proteins

His-TcHsp70 was analysed in an attempt to characterise the protein as a molecular chaperone in terms of its ATPase activity and substrate binding capabilities. The basal ATPase activity of His-Tchsp70 was determined as 40 nmol/min/mg, higher than that reported for most other Hsp70s (Table 4.1). The basal ATPase activity of GST-TcHsp70 was previously determined to be 520 nmol/min/mg by Olson et. al. (1994). This value is approximately ten fold higher than the values determined for the basal ATPase activity of His-TcHsp70 in this study (Table 4.1). In the same study, Olson et. al. (1994) determined the basal ATPase activity of human GST-Hsp70 to be 5.2 nmol/min/mg. This value is ten fold higher than that quoted for untagged human Hsp70, and corresponds well with the ATPase activity of stimulated human Hsp70, in previous studies (Chamberlain and Burgoyne, 1997). It could be suggested that the presence of the GST tag is responsible for the observed ten fold increase in the ATPase activities of GST-TcHsp70 and human GST-Hsp70. It could be that this potential increase in ATPase activity was a possible result of GST acting as a substrate or the location of GST at the N-terminal region containing the ATPase domains of the Hsps (Table 4.1).

The high potential ATPase activity of TcHsp70 suggested that TcHsp70 might have the ability to bind non-native substrates in the absence of stimulation by an Hsp40. This suggestion would mean that it would be more efficient at promoting the refolding of denatured proteins, which could be advantageous in the parasite. The efficient chaperoning of proteins may be required in the parasitic cell within its mammalian hosts, as the elevated temperature of the host may provide a continuous heat shock, will resultant protein denaturation. The recent investigation of the specialized *E. coli* Hsp70, HscC, showed that this protein also had a higher ATPase activity than that of other reported Hsp70s and was reported as being independent of nucleotide exchange (Kluck et al., 2002). If the high basal ATPase activity of HscC was due to the fact that it did not require a nucleotide exchange factor, then it could be suggested that the same principle applies to TcHsp70, as it also exhibited a high basal ATPase activity. However, further analysis of the ATPase cycle of TcHsp70 would be required to make conclusions in this regard.

The basal ATPase activity of His-TcHsp70 was stimulated approximately 1.5 times by Tcj2 in the presence of RCMLA (60 nmol/min/mg). The stimulation of the ATPase activity of TcHsp70 in the presence of a potential co-chaperone and a substrate was consistent with the findings of other Hsp70 – Hsp40 interactions. If the suggestion that the high ATPase activity of TcHsp70 is sufficient for the chaperoning of trypanosomal proteins was well founded, then stimulation of the high basal ATPase activity of TcHsp70 may be required in the parasite in times of enhanced stress, such as during invasion of the host or during differentiation of the parasite within the host cells. Tcj1-His was not seen to significantly stimulate the basal ATPase activity of TcHsp70 at any of the concentrations tested. This did not mean that Tcj1-His and TcHsp70 do not interact. It is possible that they may interact to fulfil a specific function in the cell, for which the basal ATPase activity of TcHsp70 is sufficient. It has been proposed that the flagellar protein is the substrate for Tcj1 and thus, the interaction between TcHsp70 and Tcj1 may occur during flagellar assembly (Dr D Engman, personal communication) (Figure 5.1[C]). This would be consistent with the suggested role of trypanosomal Hsps during differentiation of the parasite (Polla, 1988).

His-TcHsp70 was shown to form discrete species in the presence of RCMLA during native PAGE analysis of the ability of His-TcHsp70 to recognise and bind non-native protein conformations. The presence of the trypanosomal Hsp40s, Tcj1-His and His-Tcj2, appeared to influence the formation of these different species. The identity of the species could not be confirmed conclusively due to the lack of specific antibodies to each of the species. The species observed could have potentially been one of the following – TcHsp70 interactions with RCMLA, TcHsp70 interactions with the Tcj proteins or complexes of TcHsp70, the Tcjs and RCMLA. The formation of these complexes would be consistent with the conclusion that TcHsp70 functions as a molecular chaperone. The preliminary native PAGE analysis and the ATPase activity assays suggested that TcHsp70 is capable of functioning as a chaperone and that Tcj2 was potentially involved in this interaction (Figure 5.1[A]). This does not mean that Tcj2 is the only partner for TcHsp70. The lack of purified Tcj3 and Tcj4 negated the study of the interaction of TcHsp70 with these Hsp40s. Tcj3 and Tcj4 would have to be purified to allow the interactions to be studied, but they may interact with TcHsp70, or the constitutive TcHsc70, to fulfil specialised functions in the cell (Figure 5.1[B]). Tcj6 was shown in a previous study to be involved in the initiation of translation in *T. cruzi* (Salmon et al., 2001) and therefore, may also interact with TcHsc70 or potentially TcHsp70.

5.2 FUTURE WORK

Preliminary analysis has suggested that TcHsp70 and Tcj2 may interact in the ATP/ADP regulated process of chaperone-assisted protein folding. In order to characterise this interaction, further analyses need to be performed. The stimulation of the ATPase activity of TcHsp70 by Tcj2 in the presence of a model protein substrate should be investigated using a different assay and a range of protein substrates. The ability of TcHsp70 to promote the refolding of proteins such as luciferase would be required to confirm its ability to function as a molecular chaperone (Figure 5.1 [A]). In addition, the interaction of TcHsp70 with the Tcj3 Tcj4 and potentially Tcj6 would require that the proteins be purified and tested (Figure 5.1[B]). In addition, the isolation of the constitutive form of TcHsp70, if present in *T. cruzi*, and the study of its interactions with the Tcj proteins would be required. Co-precipitation interaction assays could be used to determine the interaction between the trypanosomal Hsps and other proteins (Figure 5.1[B]). This would necessitate the generation of untagged or differently tagged trypanosomal Hsps.

Localisation studies or cell fractionation studies would be required to determine whether TcHsp70, Tcj2 and Tcj4 associate with membranes. The isolation of specific cell organelle fractions and the identification of the trypanosomal Hsps in those fractions would confirm the suggestions in Figure 5.1 [D] and [E]. This analysis could be performed by cell localisation studies using a marker such as green fluorescent protein or by immunostaining.

In vivo complementation systems could be used to further analyse the ability of the Tcjs to complement for the absence of both eukaryotic Hsp40s (such as the *S. cerevisiae* Hsp40s described in this study) and prokaryotic Hsp40s (*E. coli* DnaJ) Delouche et al., 1997; Johnson and Craig, 2001).

Once interactions of the trypanosomal proteins have been confirmed, rational protein design could be used to map the sites of interaction and to examine whether the sequence differences noted on bioinformatic analysis are indeed indicative of different functioning and activity. The complete elucidation of the interactions of TcHsp70 with its co-chaperones in *T. cruzi* will provide a clear picture of the chaperone functioning in the parasite and will then allow the proteins to become potential targets for anti-trypanosomal drugs.

CHAPTER 6

APPENDICES

6.1 Chemicals and Reagents

All general buffers and reagents were from Saarchem, RSA or Sigma, USA. Growth media was from Biolab, UK or Difco, USA. A list of specialised reagents is detailed below.

REAGENT	SUPPLIER
Expand™ High Fidelity PCR System	Boehringer Mannheim, Germany
Frozen-EZ Yeast Transformation II™ Kit	Zymo Research, USA
5-Fluoroorotic Acid	Sigma, USA
L-Tryptophan	Sigma, USA
L-Leucine	Sigma, USA
L-Histidine	Sigma, USA
Yeast Synthetic Dropout Media Supplement	Sigma, USA
Yeast Nitrogen Base without amino acids or ammonium sulphate	Difco, USA
PCR Nucleotide Mix	Roche, Germany
Restriction Enzymes and Buffers (all)	Amersham Pharmacia Biotech, USA
High Pure Plasmid Isolation Kit	Roche, Germany
Nucleospin® Extract 2 in 1 Kit	Macherey – Nagel, Germany
T4 ligase	Promega, USA
T4 DNA polymerase	Amersham Pharmacia Biotech, USA
Cheniluminescence Western Blotting Kit (Mouse/Rabbit)	Boehringer Mannheim, Germany
SDS-PAGE Molecular Weight Standards Broad Range (myosin-200 kDa; β-galactosidase-116 kDa; phosphorylase b-97.4 kDa; ovalbumin-45 kDa; carbonic anhydrase- 29 kDa; trypsin inhibitor- 21 kDa; lysozyme- 14.4 kDa; aprotinin- 6.5 kDa)	Bio-Rad, USA
Silver Stain Molecular Weight Markers	Sigma, USA
Hybond™-c extra Nitrocellulose Membrane	Amersham Pharmacia Biotech, USA
Hyperfilm™ MP Autoradiography Film	Amersham Pharmacia Biotech, USA
Shrimp Alkaline Phosphatase (SAP)	Roche, Germany
pGEM®-T Easy Vector System I	Promega, USA
Uracil	Sigma, USA
Anti-His Antibody (mouse)	Amersham Pharmacia Biotech, USA
Chelating sepharose	Amersham Pharmacia Biotech, USA
IPTG	Roche, Germany

6.2 Protein Standard Curve for Bradford's Assay

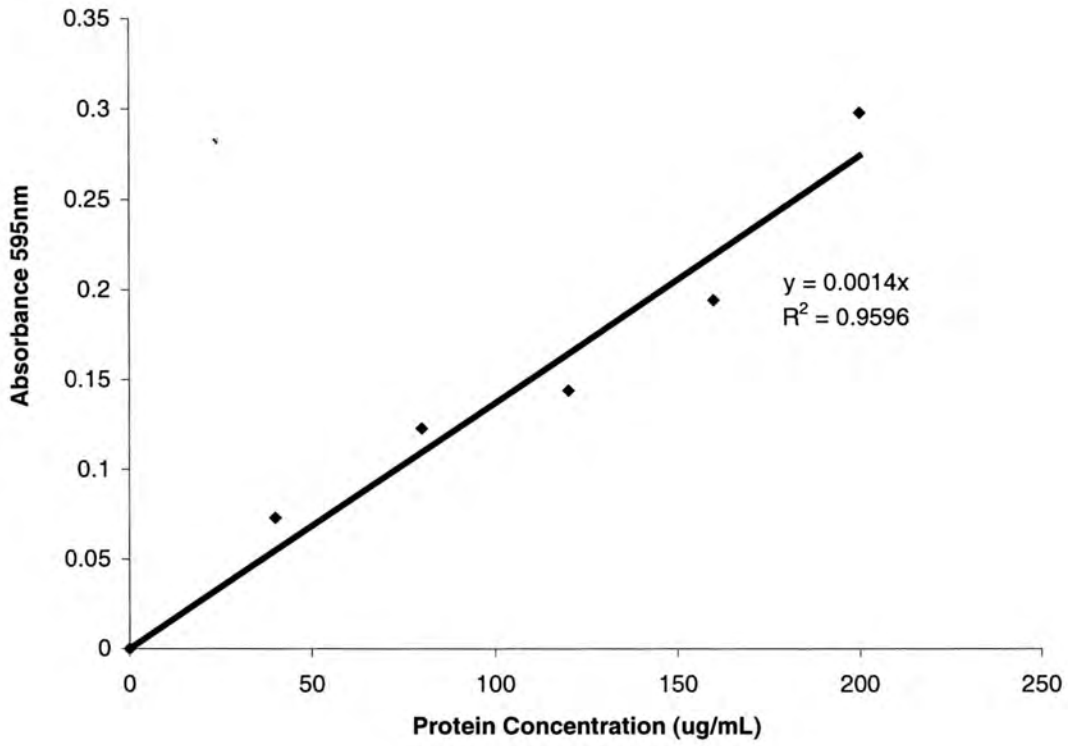


Figure 6.1: BSA Standard Curve for Protein Concentration determination by Bradford's Assay.

6.3 Phosphate Standard Curve For ATPase Assays

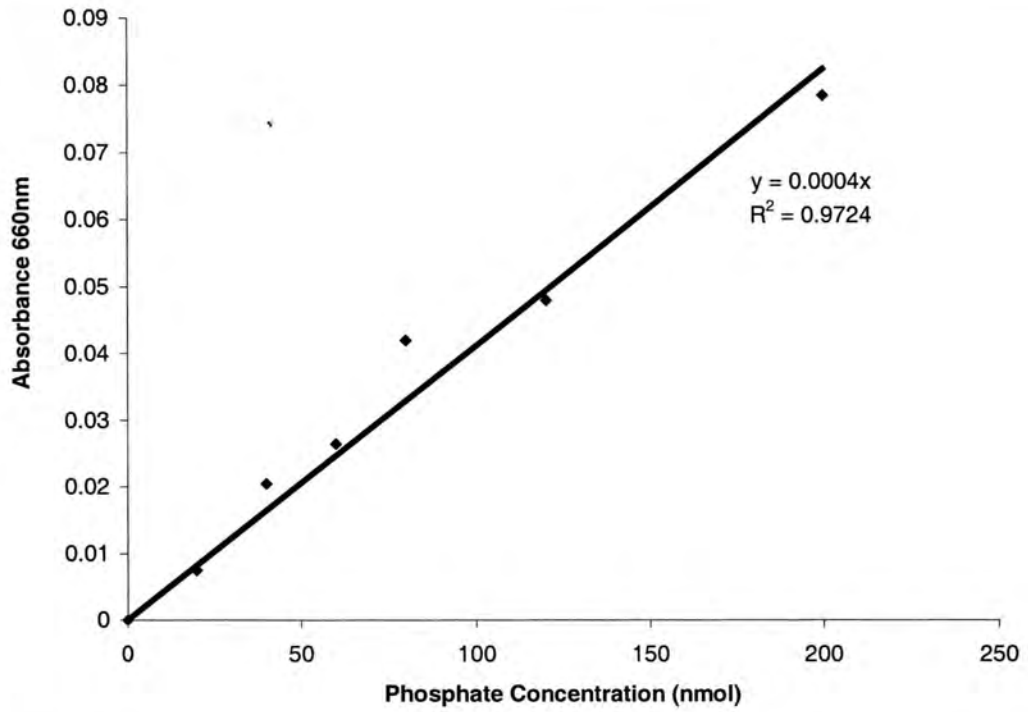


Figure 6.2: Potassium phosphate Standard Curve for determination of phosphate concentrations in ATPase Assays

REFERENCES

Ahmed, A.A., Gao, L., Davis, C.D. and Helm, D.S. (1999) Antibody response to heat shock proteins and histopathology in mice infected with *Trypanosoma cruzi* and maintained at elevated temperature. J. Parasitol. **85**, 1089-1099

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

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389-3402

Arispe, N. and De Maio, A. (2000) ATP and ADP modulate a cation channel formed by Hsc70 in acidic phospholipid membrane. J.Biol.Chem. **275**,30839-30843

Balbas, P. (2001) Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. Mol.Biotechnol. **19**, 251-267

Banecki, B., Liberek, K., Wall, D., Wawrzynow,A., Georgopoulos,C., Bertoli,E., Tanfani,F. and Zylicz,M. (1996) Structure-function analysis of the zinc finger region of the dnaJ molecular chaperone. J.Biol.Chem. **25**, 14840-14848

Benaroudji, N., Triniolles, F. and Ladjimi, M.M. (1996) Effect of nucleotides, peptides, and unfolded proteins on the self-association of the molecular chaperone Hsc70. J.Biol.Chem.**271**, 18471-18476

Birnboim, H and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acid.Res. **7**, 1513-1523

Blond-Elguindi, S., Fourie, A.M., Sambrook, J.F. and Gething, M-J.H. (1993) Peptide-dependent stimulation of the ATPase activity of the molecular chaperone Bip is the result of the conversion of oligomers to active monomers. J. Biol. Chem. **268**, 12730-12735

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

Bork, P., Sander, C. and Valencia, A. (1992) An ATPase domain common to prokaryote cell cycle proteins, sugar kinases, actin and Hsp70 heat shock proteins. PNAS. **89**, 7290-7294

Brehemer, D., Rudiger, S., Gassler, C.G., Klostermeier, D., Pakschies, L., Reinstein, J., Mayer, M.P. and Bukau, B. (2001). Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. Nat.Struct.Biol. **8**, 427-432

Brinker, A., Scheufler, C., Von der Mulbe, F., Fleckenstein, B., Herrmann, C., Jung, G., Moarefi, I. And Hartl, F.U. (2002) Ligand discrimination by TPR domains. J.Biol.Chem. **277**, 19265-19275

Buchberger, A., Schroder, H., Buttner, M., Valencia, A. and Bukau, B.(1994) A conserved loop in the ATPase domain if the DnaK chaperone is essential for stable binding of GrpE. Struct.Nat.Biol. **1**, 95-101

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

Caplan, A., Cyr, D.M. and Douglas, M.G. (1993) Eukaryotic homologues of *Escherichia coli* DnaJ: A diverse protein family that functions with Hsp70 stress proteins. Molecular Biology of the Cell. **4**, 555-563

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

Chen, S. and Smith, D.F. (1998) Hop as an adaptor in the heat shock protein 70 (Hsp70) and Hsp90 chaperone machinery. J.Biol.Chem. **273**, 35194-35200

Chou, P.Y. and Fasman, G.D. (1978) Prediction of protein secondary structure of proteins from their amino acid sequence. Advances inEnz. **47**, 4-147

Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. Nucl.Acids.Res. **16**, 10881-10890

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

Dagert, M. and Ehrlich, S.D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene. **6**, 23-28

Deloche, O., Kelley, W.L. and Georgopoulos, C. (1997) Strucutre-function analyses of the Ssc1p, Mdj1p and Mge1p *Saccharomyces cerevisiae* mitochondrial proteins in *Escherichia coli*. J.Bact. **179**, 6066-6075

Dobson, C.M. (1999) Protein misfolding, evolution and disease. TIBS. **24**, 329-332

Emini, E., Hughes, J.V., Perlow, D.S. and Boger, J. (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide.J.Virol. **55**, 836-839

Engelman, D.M., Steitz, T.A. and Goldman, A.(1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Ann.Rev.Biophys.Biophys.Chem. **15**, 321-351

Farr, C.D. and Witt, S.N. (1999) ATP lowers the activation enthalpy barriers to DnaK-peptide complex formation and dissociation. Cell.Stress.Chap. **4**, 77-85

Feifel, B., Schonfield, H.J. and Christen, P. (1998) D-peptide ligands for the co-chaperone DnaJ. J.Biol.Chem. **273**, 11999-12002

Fewell, S.W., Pipas, J.M. and Brodsky, J.L. (2002) Mutagenesis of a functional chimeric gene in yeast that identifies mutations in the simian virus 40 large T antigen J domain. PNAS. **99**, 2002-2007

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

Fink, A.L. (1999) Chaperone-mediated protein folding. Physiol. Rev. **79**, 425-449

Flaherty, K.M., Deluca-Flaherty, C. and McKay, D.B. (1990) Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature. **346**, 623-629

Freeman,B.C., Michels,A., Song,J., Kampinga,H.H. and Morimoto,R.I. (1998) in "Methods in Molecular Biology (99) Stress Response: Methods and Protocols, Humana Press Inc. Totowa NJ

Freeman,B.C., Myers,M.P., Schumacher,R. and Morimoto,R.I.(1995) Indentification of a regulatory motif in Hsp70 that affects ATPase activity, substrate binding and interaction with HDJ-1. EMBO J. **14**, 2281-2292

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

Garnier,J., Osguthorpe,D.J. and Robson,B.(1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**, 97-120

Gassler,C.S., Wiederkehr,T., Brehmer,D., Bukau,B. and Mayer,M. (2001) Bag-1M accelerates nucleotide release for human Hsc70 and Hsp70 and can act concentration-dependent as positive and negative cofactor. J.Biol.Chem. **276**,32538-32544

Glomset,J.A., Gelb,M.H. and Farnsworth,C.C.(1990) Prenyl proteins in eukaryotic cells: a new type of membrane anchor. TIBS. **15**, 139-142

Hartl,F.U. and Martin,J. (1992) Protein folding in the cell: the role of molecular chaperones hsp60 and hsp70. Annu. Rev. Biophys. Biomol.Struct. **21**,293-322

Hendrik, J.P. and Hartl, F.U. (1993) molecular chaperone functions of heat shock proteins. Annu. Rev. Biochem. **62**, 349-384

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

Hiromura,M., Yano,M., Mori,H., Inoue,M., and Kido,H. (1998) Intrinsic ADP-ATP exchange activity is a novel function of the molecular chaperone Hsp70. Journ. Biol. Chem. **273**, 5435-5438

Hopp,T.P. and Woods,K.R.(1981) Prediction of protein antigenic determinants from amino acid sequences. PNAS.**78**, 3824-3828

Horton,L.E., James,P., Craig,E.A. and Hensold,J.O.(2001) The yeast hsp70 homologue Ssa is required for translation and interaction with Sis1 and Pab1 on translating ribosomes. J.Biol.Chem. **276**, 14426-14433

Huang,S-P., Tsai,M-Y., Tzou,Y-M., Wu,W-G. and Wang,C. (1993) Aspartyl residue 10 is essential for ATPase activity of rat Hsc70. J.Biol.Chem. **268**, 2063-2068

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

Jameson,B.A. and Wolf,H.(1988) The antigenic index: a novel algorithm for predicting antigenic determinants. CABIOS. **4**.

Janin,J., Wodak,S., Levitt,M. and Maigret,B.(1978) Conformation of amino acid side-chains in proteins. J.Mol.Bio.**125**, 357-386

Johnson,J.L. and Craig,E.A.(2001)An essential role for the substrate-binding region of Hsp40s in *Saccharomyces cerevisiae*. J. Cell. Biol. **152**,851-856

Johnson,B.D., Schumacher,R.J., Ross,E.D. and Toft,D.O. (1998) Hop modulates hsp70/hsp90 in Protein Folding. J.Biol.Chem. **273**, 3679-3686

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

Karplus,P.A. and Schultz,G.E. (1988) Prediction of chain flexibility in proteins. Naturwissenschaften. **72**, 212-213

Kelley, W.L. (1999) Molecular chaperones: how J domains turn on Hsp70s. Curr. Biol. **9**, R305-R308

Kelley, W.L. and 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. PNAS.**94**, 3679-3684

Kiang, J.G. and Tsokos, G.C.(1998) Heat shock protein 70kDa: molecular biology, biochemistry, and Physiology. Pharmacol.Ther. **80**, 183-201

King,C., Eisenberg,E. and Greene,L.E. (1999) Interaction between Hsc70 and DnaJ homologues: Relationship between Hsc70 Polymerisation and ATPase Activity. Biochemistry. **38**, 12452-12459

Kluck,C.J., Patzelt,H., Genevaux,P., Brehmer,D., Rist,W., Schneider-Mergener,J., Bukau,B. and Mayer,M.P. (2002) Structure-function analysis of HscC, the *Escherichia* coli member of a novel subfamily of specialised Hsp70 chaperones. J. Biol. Chem. Published online ahead of print jbc.M206520200

Kraulis,P (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. App. Cryst. **24**, 946-950

Krautz,G.M., Peterson,J.D., Godsel,L.M., Krettli,A.U. and Engman,D.M. (1998) Human antibody repsonse to *Trypanosoma cruzi* 70-kD heat-shock proteins. Am. J. Trop. Med. Hyg. **58**, 137-143

Kumar,N and Zheng,H (1998) Evidence for epitope-specific thymus independent response against a repeat sequence in a protein antigen. Immunol. **94**, 28-34

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

Laufen , T., Mayer,M.P., Beisel,C., Klostermeier,D., Mogk,A., Reinstein,J. and Bukau,B. (1999) Mechanism of regulation of Hsp70 chaperones by DnaJ co-chaperones. PNAS. **96**, 5452-5457

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature. **227**, 680-685

Lee,S., Fan,C.Y., Younger,J.M., Ren,H. and Cyr,D.M.(2002) Identification of essential residues in the type II hsp40 sis1 that function in polypeptide binding. J.Biol.Chem. **277**, 21675-21682

Lu,Z. and Cyr,D.M.(1998) Protein Folding Activity of Hsp70 is modified differentially by the Hsp40 Co-chaperones Sis1 and Ydj1. J.Biol.Chem. **273**, 27824-27830

Maranon,C., Planelles,L., Alonson,C. and Lopez,M.C. (2000) HSP70 from *Trypanosoma cruzi* is endowed with specific cell proliferation potential leading to apoptosis. Internat.Immunol. **12**, 1685-1693

Marshall,C.J.(1993) Protein prenylation: a mediator of protein-protein interactions. Science. **259**, 1865-1866

Martinez-Yamout,M., Legge,G.B., Zhang,O., Wright,P.E. and Dyson,J.(2000) Solution strucutre of the cysteine-rich domain of the *Escherichia coli* chaperone protein dnaj. J.Mol.Biol. **300**, 805-818

Mayer,M.P. and Bukau,B. (1998) Hsp70 chaperone systems: diversity of cellular functions and mechanism of action. Biol.Chem. **379**, 261-268

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

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

McCarty, J.S. and Walker, G.C. (1991) DnaK as a thermometer: threonine-199 is the site of autophosphorylation and is critical for ATPase Activity. PNAS. **88**, 9513-9517

Minami, Y., Hohfield, J., Ohtsuka, K.and Hartl, U. (1996) Regulation of the heat-shock protein 70 reaction cycle by the mammalian DnaJ homolog Hsp40. J.Biol.Chem. **271**, 19617-19624

Montgomery,D.L., Morimoto,R.I. and Gierasch,L.M.(1999) Mutations in the substrate binding domain of the *Escherichia coli* 70kda molecular chaperone, dnak, which alter substrate affinity or interdomain coupling. J. Mol. Biol.**286**, 915-932

Muchowski,P.J., Schaffar,G., Sittler,A., Wanker,E.E., Hayer-Hartl,M.K. and Hartl,F.U.(2000) Hsp70 and Hsp40 chapeornes can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. PNAS. **97**, 7841-7846

Multoff, G., Botzler,C. and Issels,R. (1998) The role of heat shock proteins in the stimulation of an immune response. Biol. Chem. **379**, 295-300

Nandan, D., Daubenberger, C., Mpimbaza, G. and Pearson, T.W. (1994) A rapid, single-step purification method for immunogenic members of the hsp70 family: validation and application. J. Immunol. Meth. **176**, 255-263

O,Brien,M.C., Flaherty,K.M. and McKay,D.B.(1996) Lysine 71 of the chaperone protein hsc70 is essential for ATP hydrolysis. J.Biol.Chem. **271**,15874-15878

Olson,C.L., Nadeau,K.C., Sullivan,M.A., Winquist,A.G., Donelson,J.E., Walsh,C.T. and Engman,D.M. (1994) Molecular and biochemical comparison of the 70-kDa heat shock proteins of *Trypanosoma cruzi*. J. Biol. Chem. **269**, 3868-3874

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

Peitsch,M.C.(1995)Protein Database Quarterly Newsletter.**72**, 4

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

Planelles, L., Thomas,M.C., Alonso,C. and Lopez,M.C. (2001) DNA immunisation with *Trypanosoma cruzi* HSP70 fused to the KMP11 protein elicits a cytotoxic and humoral immune response against the antigen and leads to protection. Infect.Immun. **69**, 6558-6563

Polla, B.S.(1988) A role for heat shock proteins in inflammation? Immunol. Today. **9**, 134-137

Porath,J., Carlsson,J., Olsson,I. And Belfrage,G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. Nature. **258**, 598-599

Rajapandi, T., Greene, L.E. and Eisengerg,E.(2000) The molecular chaperones hsp90 and hsc70 are both necessary and sufficient to activate binding by glucocorticoid receptor. J.Biol.Chem. **275**, 22597-22604

Richarme,G. and Kohiyama,M. (1993) Specificity of the *Escherichia coli* chaperone DnaK (70-kDa Heat Shock Protein) for hydrophobic amino acids. J. Biol. Chem. **268**, 24074-24077

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

Rudiger,S., Buchberger,A. and Bukau, B. (1997) Interaction of Hsp70 chaperones with substrates. Nature. Struct. Biol. **4**, 342-349

Salmon,D., Montero-Lomeli,M. and Goldenberg,S.(2001) A DnaJ-like protein homologous to the yeast co-chaperone sis1(tcj6p) is involved in initiation of translation in *trypanosoma cruzi*. J.Biol.Chem.

Sayle, R.A. and Milner-White, E.J. (1995) Rasmol: Biomolecular Graphics for all. TIBS. **20**, 374-376

Sharma,Y.D. (1992) Structure and possible function of heat-shock proteins in *Falciparum* malaria. Comp. Biochem. Physiol. **102B**, 437-444

Shen,Y., Meunier,L. and Hendershot,L.M.(2002) indetification and characterisation of a novel endoplasmic reticulum (ER) dnaJ homologue, which stimulates ATPase activity of Bip *in vitro* and is induced by ER stress.J.Biol.Chem **277**, 15947-15956

Skeiky,Y.A.W., Benson,D.R., Guderian,J.A., Whittle,J.A., Bacelar,O., Carvalho,E.M. and Reed,S.G. (1995) Immune responses of Leishmanias patients to heat shock proteins of *Leishmania* species and humans. Infect.Immun. **63**, 4105-4114

Slavotinek,A.M. and Biesecker,L.G. (2001) Unfolding the role of chaperones and chaperonins in human disease. Trends. Genet. **17**, 528-535

Smith,D.B. and Johnson,K.S. (1988) Single-step purification of peptides expressed in *Escherichia coli* as fusions with glutathione-*S*-transferase. Gene. **67**, 31-40

Smith,D.F., Whitesell,L. and Katsanis,E. (1998) Molecular chaperones: biology and prospects for pharmacological intervention. Pharm. Rev. **50**, 493-510

Sullivan,C.S., Tremblay,J.D., Fewell,S.W., Lewis,J.A., Brodsky,J.L. and Pipas,J.M. (2000) Species-specific elements in the large T antigen J domain are required for cellular transformation and DNA replication by simian virus 40. Mol.Cell.Biol. **20**, 5749-5757

Szabo,A., Korszun,R., Hartl,F.U. and Flanagan,J. (1996) A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrates. EMBO. **15**, 408-417

Taufari, W.L. (1999) Immunopathology of Chaga's disease- a historical overview. Mem. Inst. Oswaldo Cruz. **94**, 247-248

Tibbetts,R.S., Jensen,J.L., Olson,C.L., Wang,F.D. and Engman,D.M. (1998) The DnaJ family of protein chaperones in *Trypanosoma cruzi*. Mol. Biol. Parasit. **91**, 319-326

Wawrzynow,A., Banecki,B., Wall,D., Liberek,K., Georgopoulos,C. and Zylicz,M.(1995) ATP hydrolysis is required for the DnaJ-dependent activation of DnaK chaperone for binding both native and denatured protein Substrates. J.Biol.Chem. **270**, 19307-19311

Whitley,D., Goldberg,S.P. and Jordan,W.D.(1999) Heat shock proteins: a review of the molecular chaperones. J.Vasc. Surg. **29**, 748-751

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