Characterization of the co-chaperones of Hsp70 and Hsp90 in *Trypanosoma brucei* and their potential partnerships

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

(Biotechnology)

Biotechnology Innovation Centre

Faculty of Science

RHODES UNIVERSITY

By

Fortunate Mokoena

Feb 2015

Abstract

African Trypanosomiasis, which is caused by *Trypanosoma brucei*, is one of the crippling agents of social and economic development in Africa. *T. brucei* cycles between the cold-blooded insect vector, the tsetse fly (*Glossina spp*), and warm-blooded mammalian hosts. *T. brucei*, *T. cruzi* and *L. major* are mammal infecting kinetoplastid parasites that are collectively referred to as TriTryps. These parasites experience extreme environments as they move between their warm-blooded mammalian hosts and cold-blooded insect vectors which trigger extensive morphological transformations during the life-cycle of the parasite. Molecular chaperones have been implicated in parasite differentiation. TriTryps display significant expansions and diversity in the gene complements encoding molecular chaperones, especially J-proteins. Generally, J-proteins function as co-chaperones of Hsp70s, forming part of vital protein homeostasis processes. Hsp70s show a high degree of conservation, while J-proteins appear to be an extreme case of taxonomic radiation. Although several studies have focused on the molecular and cell biology of Hsp70s in some kinetoplastid parasites, knowledge is still lacking pertaining to J-proteins and their partnerships with Hsp70s. This thesis focused on the classification of kinetoplastid J-proteins into the four types by examining the domain organizations using *T. brucei* as a guide. The potential partnership of J-proteins and Hsp70s were postulated based on predicted subcellular localization. Kinetoplastid parasites, particularly *T. brucei,* have evolved an expanded and specialized J-protein machinery, likely to be a consequence of an evolutionary fitness/trait to adapt to diverse environment present in hosts and vectors. These analyses will yield insight into the process of parasite differentiation as well as provide new leads for chemotherapeutic treatments.

The presence of the STI1 mediated Hsp90 hetero-complex formation has not been confirmed in *T. brucei*. To this end, *in silico* and biochemical techniques were used to characterize the role of TbSTI1, as an adaptor protein of Hsp70 and Hsp90. Through domain architecture analysis, sequence alignments, phylogenetic analysis and three-dimensional structure prediction, TbSTI1 was demonstrated to be the most conserved TPR containing co-chaperone of Hsp70 and Hsp83

in *T. brucei* and also shown to be highly similar to its eukaryotic homologues. Recombinant TbSTI1 was overproduced and purified in *E.coli* cells and subsequently shown to associate with TcHsp70 in a concentration dependent manner and associate weakly with TbHsp70.4. TbSTI1 and TbHsp83 were also demonstrated to be expressed and upregulated upon exposure to heat shock at the bloodstream stage of parasite development. In conclusion, this study is the first to report the interaction of TbSTI1 with a chaperone. Interactions between TbSTI1 and Hsp70s were demonstrated and therefore, the formation of the hetero-complex is predicted based the similarity of TbSTI1 to other STI1 proteins.

Table of Contents

Chapter 1: Literature Review

Chapter 2: *In silico* characterization of Hsp70/J-proteins complex in kinetoplastids

Chapter 3: Analysis of Hsp90 and TPR containing co-chaperones of Hsp70 and Hsp90 in *T. brucei*

Chapter 4: Characterization of TbSTI1 and its interactions with cytosolic Hsp70 and Hsp90

Chapter 5: Concluding remarks and future perspectives

Appendices

Dedication

This thesis is dedicated to my loving grandmother **Mmapelo Rosah Rameetse**, a woman who never got an education but never ceased to push my cousins and I to dream and work hard. Thank you for setting such an amazing example and for everything, Tlou.

Acknowledgments

"*So do not fear, for **I am** with you, do not be dismayed, for **I am** your **God**, **I** will strengthen you and help you, **I** will uphold you with my righteous hand"-(Isaiah 41:10)*

I am deeply grateful for the support and the assistance of the following people, all of whom have played essential roles towards the completion of this research and thesis:

1. My supervisor, Dr Aileen Boshoff, for her constant support, patience, understanding and genuine interest in my work which has been a huge encouragment throughout my studies. Thank you for teaching me to be independent and always strive towards excellence. I am sincerely greatful for the opportunity you granted me to gain such a wealth of experience working in your research group.
2. My co-supervisor, Dr Adrienne Edkins, for all the inputs and fresh ideas. Your passion and excitement for science is truly refreshing and contagious. My heartfelt thank you for being there when I needed it the most.
3. Dr Earl Prinsloo who was always willing to share his knowledge with instruments and data interpretation.
4. My deepest gratitude to Dr Michael Ludewig and Dr Adelle Burger, for their expertise, useful ideas and the support they gave me throughout my studies.
5. My parents, Kenneth Harold and Mapula Yvonne Mokoena, for their unlimited support, constant cheers and unceasing prayers. I am deeply thankful for your selflessness, your faith in me, your patience and your unconditional love.
6. To my friends, my help and my biggest support systems, Dr Lorraine Zvichapera Mutsvunguma, Dr James Njunge and Dr Michael Daniyan, thank you for being friends that stick closer than siblings. I remain deeply humbled by all you have done for me and for helping proofread this thesis.
7. A special thank you to my humans Selby, Mathaabe and Allan, you guys have been phenomenal, especially the one who stayed strong (SS) and peaceful.
8. A big thanks to all present and past members of BioBRU, your positive criticism and shared novel ideas have made this difficult journey enjoyable.
9. I acknowledge the National Research Foundation, Innovation Fellowship funding.

"*Praise be to **God**, who has not rejected my prayers or withheld his love from me"-(*Psalm 66:20).

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 2

Chapter 4

List of Abbreviations

α	Alpha
β	Beta
λ	Lambda
µ	Micro
µg	Microgram(s)
µl	Microlitre(s)
µM	Micromolar
%	Percentage
~	Approximately
>	More than
°C	Degrees Celsius
6 xHis	Hexahistidine tag
A	Absorbance
APS	Ammonium Persulphate
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
Bag	Bcl2-associated athanogene
BiP	Binding protein
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BSA	Bovine Serum Albumin
C-	COOH-terminal
CAAX	Cysteine, A – Aliphatic Residue, X – Any Residue
CHIP	Carboxyl terminus Hsc70 interacting Protein
DNA	Deoxyribonucleic Acid
DnaK	Prokaryotic Hsp70
DOPE	Discrete Optimized Protein Energy
DTT	Dithiothreitol
E. coli	*Escherichia coli*

ER	Endoplasmic reticulum
g	Gram(s)
xg	Gravitational force
GF	Region Glycine-Phenylalanine rich region
H. sapiens	*Homo sapiens*
Hip	Hsc70 interacting protein
Hop	Hsp70/Hsp90 organising protein
HPD motif	Histidine-Proline-Aspartic acid motif
HRP	Horse Radish Peroxidase
Hsc70	70 kDa Heat shock cognate protein
HSF	Heat Shock Factor
Hsp	Heat shock protein
Hsp110	110 kDa heat shock protein
Hsp40	40 kDa heat shock protein
Hsp70	70 kDa heat shock protein
Hsp90	90 kDa heat shock protein
IPTG	Isopropyl-β-D-thiogalactopyranoside
KanR	Kanamycin resistance
Kb	Kilo base pairs
kDa	Kilo Daltons
L	Litre(s)
LB	Luria-Bertani media
Leishmania spp	*Leishmania species*
M	Molar
mAb	Monoclonal antibody
mg	Milligram(s)
ml	Millilitre (s)
mM	Millimolar
mol	Mole(s)
n	Nano
N.D.	Not determined
NaCl	Sodium chloride

NBD	Nucleotide Binding Domain
NEF	Nucleotide exchange factor
Ni-NTA	Nickel nitrilotriacetic acid
ng	Nanogram(s)
N-	NH_2-terminal
P. falciparum	*Plasmodium falciparum*
PBS	Phosphate Buffer Saline
PDB	Protein Data Bank
Pi	Inorganic phosphate
pI	Isoelectric point
PMSF	Phenyl Methyl Sulfonyl Fluoride
polyAb	Polyclonal antibody
rpm	Revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SBD	Substrate Binding Domain
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis
sHsp	Small Heat shock protein
SPR	Surface Plasmon Resonance Spectroscopy
STI1	Stress inducible protein 1
T. brucei	*Trypanosoma brucei*
T. cruzi	*Trypanosoma cruzi*
TBS	Tris-Buffered Saline
TBS-T	Tris Buffer Saline – Tween 20
TM	Transmembrane domain
TEMED	N,N,N',N'-tetramethylethylenediamine
TPR	Tetratricopeptide Repeat
Tris	Tris-2-amino-2-hydroxymethyl-1,3-propanol
U	Unit(s)
UV	Ultraviolet
V	Volts
v/v	Volume per volume

w/v	Weight per volume
WHO	World Health Organization
xg	Gravitational force
YT Yeast-	Tryptone media

List of research outputs

Publication in progress

1. **Mokoena F**, Edkins A.L. and Boshoff A. Structure, function and classification of J-proteins from kinetoplastids kinetoplastids. In preparation for submission to *Cell stress and Chaperone*

Conference proceeding:

1. **Mokoena F**, Ludewig M.H. and Boshoff A (2013). The potential interaction between Hsp70-STI1-Hsp90 in *T. brucei*. VI International Congress on Stress Proteins in Biology and Medicine, August 2013, Sheffield, England.

CHAPTER 1

LITERATURE REVIEW

1.1. Kinetoplastids

Kinetoplastids are an interesting group of eukaryotic micro-organisms that are widely characterized by their rRNA lineage which extends further back than animals, plants and fungi (Simpson et al., 2002; Simpson et al., 2006). Evolutionary studies have indicated that eukaryotic kinetoplastid protists belong to the phylum Euglenozoa, order kinetoplastida and the Trypanosomatidae family (Simpson et al., 2006). Class Kinetoplastida is further subdivided into free-living and parasitic ancient eukaryotes, the former collectively referred to as bodonids (Simpson et al., 2002). Free-living bodonids such as *Bodo* spp play crucial roles in microbial food webs, consumption of bacteria and small eukaryotes (Arndt et al. 2000). Parasitic kinetoplastids are believed to have evolved into four separate clades producing Icthyobodo-Perkinsiella clade, Fish infecting Cryptobia (trypanoplasma) species, the true Cryptobia and trypanosomatids (Simpson et al., 2006).

Kinetoplastids are responsible for a wide range of diseases affecting humans, animals and plants which has a detrimental impact on the environment and the economy. The most prominent members of Trypanosomatids are vertebrate parasites *Leishmania*, *Trypanosoma* and *Endotrypanum* (Maslov et al., 2001), all of which are transmitted by insects and cause major human and veterinary diseases. Major human diseases caused by trypanosomatid parasites include sleeping sickness (*Trypanosoma brucei*), Chaga's disease (*Trypanosoma cruzi*), and leishmaniases. In order to understand the evolution of parasitism, facultative and obligatory kinetoplastids have been studied in relation to their close relatives which are either photo- or phagotrophic (Lukes et al., 2014). Indeed, the hypothesis that parasitism originated with kinetoplastids has been thoroughly supported, also the specific adaptations allowing these flagellates to co-exist in their hosts have been described (Lukes et al., 2014).

The general structure of parasitic kinetoplastids is defined by an elongated and slender morphology; the possession of a single nucleus, unique ultrastructurally organized and separated mitochondrion and a region known as the kinetoplast (Donelson et al., 1999; Simpson et al., 2002) and the glycosome (specialized organelle containing enzymes involved in the metabolic process of glycolysis) (Clayton and Michels 1996; Tielens and Van Hellemond 1998; Faria-e-Silva et al. 2000). Located near the basal body which forms the origin of the single flagellum of a trypanosome, the kinetoplast is defined as an unusual, complex catenated network of mitochondrial DNA organized in a disc-like structure that forms an integral part of the single mitochondrion (Figure 1.1.).

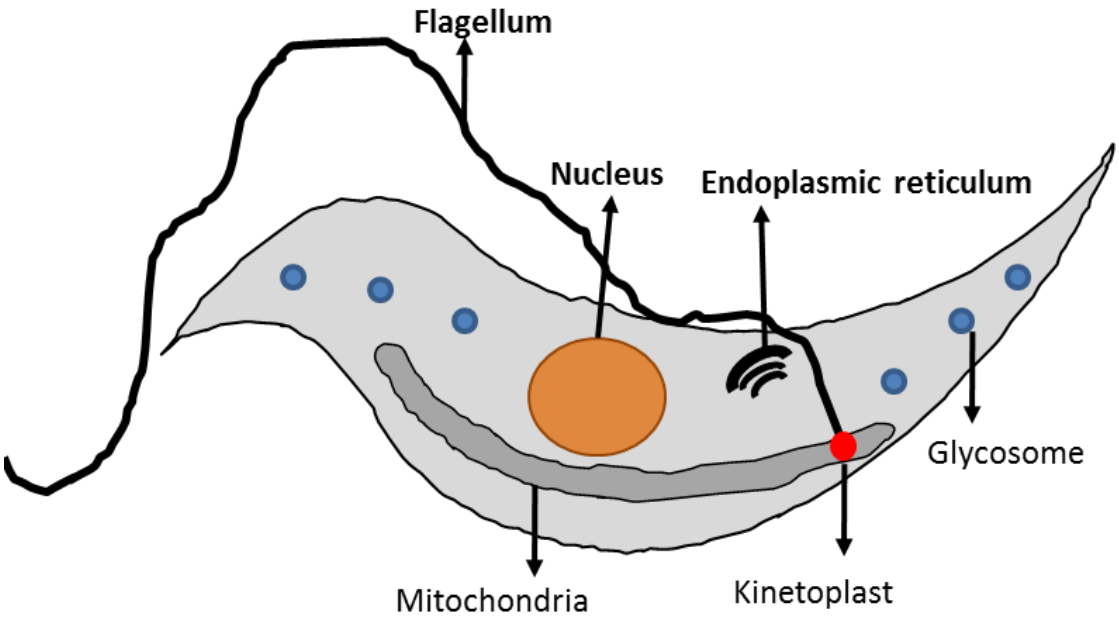


Figure 1. 1: The generalized structure of an African trypanosome in the procyclic trypomastigote stage of the life-cycle. Kinetoplastids typically have one long mitochondria that extends throughout the cell body. The kinetoplast is located inside the mitochondria near the origin of the flagellar basal body. The glycosome, the nucleus and endoplasmic reticulum are integral organelles of trypanosomes and are also indicated.

Indeed, trypanosomes possess some organelles in common with other eukaryotes as well as some unconventional locomotion organelles such as a single motile flagellum with a paraflagellar

rod (Ralston et al., 2009; Portman and Gull, 2010). The genome of kinetoplastids is virtually intronless, their genes are arranged into a massive polycistronic cluster and their nucleotides undergo unprecedented modifications. They are able to trans-splice all mRNA transcripts (Parsons et al., 1984; De Lange et al., 1984; Stuart et al. 2005). Several novel processes have been best studied and were also first described in some kinetoplastids, particularly trypanosomes, these include the intricate and energy consuming mitochondrial RNA editing (Benne et al., 1986), antigenic variation (Namangala, 2011; Rudenko, 2011) and phagocytosis (Simpson et al., 2006).

1.1.1. Trypanosomatids

The obligate, single-celled, flagellate protozoan parasites of the genus *Trypanosoma* are thought to have branched from the metazoan lineage approximately $3x10^9$ years ago (Rout and Field, 2001; Simpson et al., 2002). They are the oldest single-celled organisms to branch from this eukaryotic lineage (Sogin et al., 1986; Fernandes et al., 1993). *Trypanosoma* is taxonomically related to the intracellular, obligate protozoan *Leishmania* and, together they belong to Trypanosomatidae family. The *Leishmania* genus consists of 21 biologically, clinically and epidemiologically diverse species which are transmitted to humans via sandflies and causes a complex of diseases collectively referred to as leishmaniasis (Sibley and Andrews, 2000). Leishmaniasis affects poverty stricken regions of the world, mainly the tropics and subtropics areas (Herwaldt, 1999). The distribution of *Leishmania* is estimated to have put approximately 350 million people at risk (Desjeux, 2004; WHO, 2010a).

Trypanosoma evolved into two types: the sterocoraria trypanosomes (American) and the salivaria trypanosomes (African origin) (Nyalwidhe et al., 2003). In addition to their distribution, their mode of transmission and propagation in the mammalian host are the most fundamental differences between the sterococarian and salivarian trypanosomes. The intracellular sterococarian trypanosome is transmitted to the vertebrate host via the vector's faeces while extracellular salivarian trypanosomes develops in the midgut of the invertebrate vector and is

transmitted to the host following the blood-sucking meal bite (Barrett et al., 2003; Baral, 2010). Salivarian trypanosomes consist of five species which are further subdivided into human infective forms and non-human infective forms. The four non-human infective forms responsible for veterinary diseases are *Trypanosoma vivax*, *T. evansi*, *T. equiperdum* and *T. congolense* which causes the most virulent cases in livestock. *T. brucei* is responsible for causing fatal cases in both humans and livestock (Baral, 2010). The human infections are known as human African trypanosomiasis (HAT; also known as sleeping sickness) while the wild and domestic animal infections are called animal African trypanosomiasis (AAT; also known as Nagana). The most notable sterocoraria trypanosome is *Trypanosoma cruzi* which causes Chagas' disease (also called American trypanosomiasis) that is endemic to approximately 21 countries mainly across Latin America and parts of North America. Around 90 million people are exposed to the parasite and current estimates indicate that 12 million people are infected with *T. cruzi* (WHO 2010b; Rassi Jr. et al., 2010, Rassi Jr. et al., 2012). *T. brucei*, *T. cruzi* and *L. major* are collectively known as the TriTryps (De Souza et al., 2010; Ambit et al., 2011; Langousis and Hill, 2014).

1.2. African Trypanosomiasis

T. brucei, the etiological agent of African trypanosomiasis (AT), belongs to the group of salivarian trypanosomes and is grouped under the subgenus *Trypanozoon*. *T. brucei* consists of three morphologically identical subspecies *T. brucei rhodensiense*, *T. brucei gambiense* and *T. brucei brucei*. The human pathogens *T. b. gambiense* and *T. b. rhodensiense* cause HAT whereas the human non-infective *T. b. brucei* is virulent to a wide range of wild and domestic animals (Pépin and Méda, 2001) causing ATT.

AT is endemic to the poorest regions within sub-Saharan and equatorial Africa (Cecchi et al., 2008). The weak health systems and political instability in these regions have made the processes of disease surveillance and management difficult. As a consequence, current estimates have put

70 million people at risk (Luscher et al., 2007; Kennedy, 2006, WHO, 2012). Although AT was successfully eradicated in certain areas of Africa, the past 100 years have shown long epidemic periods and resurgence (Hide, 1999; Gull, 2002; Garcia et al., 2006). Factors such as population displacements, political problems, civil wars and collapse in the health systems (Gull, 2002) have contributed to the resurgence of AT.

Tsetse refers to blood-sucking flies which serve as reservoirs for *T. brucei* and are vectors for transmitting AT to mammalian hosts. They all belong to the genus *Glossina* (Class: Insecta, Order: Diptera; Suborder: Brachycera; Superfamily: Hippoboscoidea; Family: Muscidae; Subfamily: Glossinidae) (Cross, 2001; Fraumann, 2003; Steverding, 2008). Tsetses are between 8-17 mm in length and morphologically resemble the house fly (Akoy, 2003) and can be distinguished from other biting flies by their forward pointing mouth and characteristic wing venation. The life-cycle of a tsetse is typically between 1-6 months, and once infected, appears to maintain infection (Akoy, 2003). The most fascinating feature about the life-cycle of tsetses is that they do not lay eggs, instead a single larva develops within the female uterus. Unlike mosquitoes, in which the female *Anopheles* transmits the parasite to humans, both the male and female tsetse feed on vertebrate hosts and are thus both bona fide vectors. The overlap of the favoured ecology of the tsetse flies (warm, shady, humid areas) with human activity defines, in part, the spatial pattern of disease (Vickerman, 1997; Steverding, 2008).

Earlier studies indicated that vector competency is determined by two factors: the species of *Glossina* and the type of trypanosome. For instance, *G. morsitans* is a good vector for the etiological agent of AAT, *T. congolense,* while *G. palpalis* is a poor vector. However, *G. palpalis* is the main vector for the etiological agent of HAT, *T. b. gambiense*, which does not reproduce in the mid-gut of *G. morsitans* (Geiger et al., 2005). In light of the recent sequencing of the *G. morsitans* genome, knowledge of the vector-parasite interaction and understanding of the complex relationships between the species will be realized (Watanabe et al., 2014). Data retrieved from genome studies will also contribute to the development of vector control

strategies, understanding of tsetse biology and AT chemotherapeutic targets (Steverding, 2014; Watanabe et al., 2014).

Tsetses are restricted to certain regions of Africa and due to their distribution, exposure and abundance. AT becomes a daily risk to children and adults involved in activities such as farming, hunting, fishing and washing of clothes (Brun et al., 2010). The transmission of AT to cattle and wild animals further heightens the problem of poverty and agricultural development on the continent of Africa. If left untreated, AT is fatal and has been considered a greater cause of mortality and morbidity than HIV/AIDS in parts of Africa most vulnerable to infection (Matthews, 2005). Only 3-4 million of the population most affected has access to medical treatment and surveillance. However, the statistics of AT incidences suffer from gross errors due to under-reporting of new cases and death (Odiit et al., 2005). The problem is increased by lack of effective treatment of the disease as well as financial disincentive for novel drug development and rising resistance to current drug treatments (Barrett, 2003; Barrett et al., 2007).

1.2.1. Distribution

Of the two human infective forms, *T. b. gambiense* causes the majority (>90%) of HAT cases in 24 western and central African countries, with the majority of cases occurring in Angola, Congo, Guinea, Southern Sudan and Northwestern Uganda (Barrett et al., 2003, Picozzi et al, 2005; Baral, 2010) (Figure 1.2). *T. b rhodensiense* is more prevalent in the rift valley and 13 eastern and southern Africa countries with the majority of cases reported in Malawi, Democratic republic of Congo (DRC), Tanzania and Southern and Central Uganda (Barrett et al., 2003, Picozzi et al, 2005; Duffy et al., 2013) (Figure 1.2). Cases of *T. b. gambiense* and *T. b. rhodesiense* infection have also been reported in various animals and HAT caused by *T. b rhodesiense* in east Africa is recognized as a zoonosis (Picozzi et al., 2008). Furthermore, *T. b. gambiense* comprises of two genetically different groups: the first group is less diverse and displays low virulence in rodents (Inoue et al.,

1998) and the second group is more genetically diverse, able to infect rodents and shows biological and genetic similarities to *T. b. brucei* (MacLeod et al., 2001). *T. b. gambiense* strains belonging to the first group possess the smallest of all *T. brucei* genomes, comprising 71-82% of the highest DNA content measured in *T. brucei* (Kanmogne et al., 1997; Symla et al., 2012; Uzureau et al., 2013).

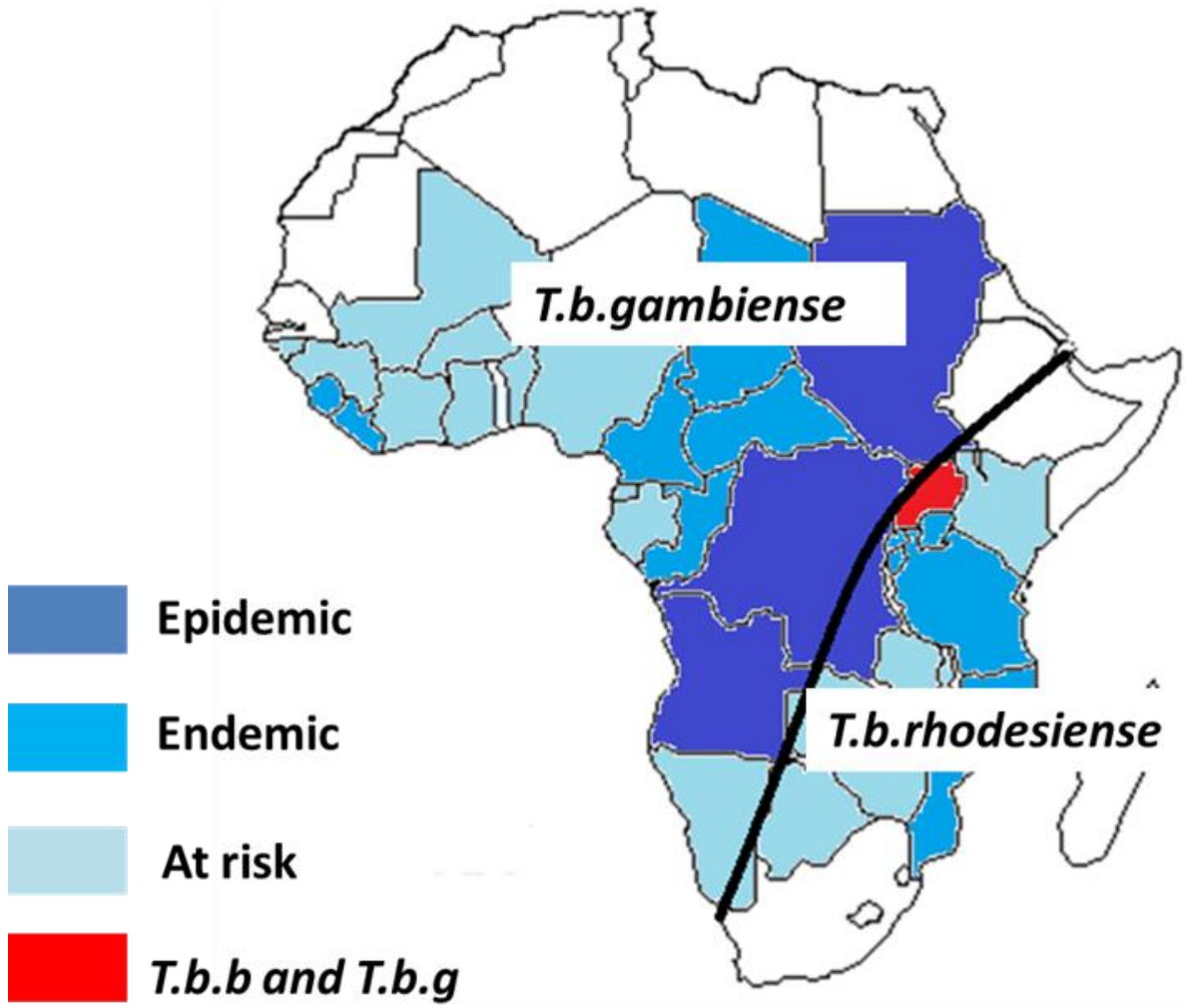


Figure 1. 2: The geographical distribution of HAT by country based on the cases reported in 2012. *T. b. gambiense* is more prevalent in West Africa, while more cases of HAT infections by *T. b. rhodensiense*cases are localized in East and Southern Africa. The emergence of *T. b. rhodensiense* HAT cases in northwest Uganda, annotated in red, is beginning to change the geographical distribution of AT (Adapted from Brun et al., 2010; Kennedy et al., 2013; WHO, 2013).

1.2.2. Clinical manifestations

The clinical manifestations characterizing HAT are the same for the two human infecting trypanosomes, however *T. b. rhodensiense* causes a more acute form of the disease with overt clinical manifestations occurring within days post transmission and ultimately leading to death if left untreated (Brun et al., 2010; Duffy et al., 2013). *T. b. gambiense* on the other hand, is characterized by a chronic progressive course of the disease (Malvy and Chappuis, 2011). HAT signs and symptoms have classically been divided into two phases based on the presence or absence of trypanosomes in the central nervous system (CNS). The first stage, the haemolymphatic phase, is characterized by the confinement of the trypanosome in the blood and lymphatic systems (Brun et al., 2010; Baral, 2010). High circulating parasitemia and the presence of by-product of immune defence against the trypanosome molecules following cell lysis lead to overall debilitation of the host immune system and multiple organ complication (Barret *et al*., 2007; Fields and Carrigan, 2009). Generally, the first stage involves non-specific symptoms such as headaches, fever and joint pain which are difficult to diagnose correctly (Checchi et al., 2008; Kabore et al., 2011; Jamonneau et al., 2012). The second stage, the meningoencephalitic phase, is characterized by the active invasion of trypanosomes into CNS crossing the blood-brain barrier most likely at the choroid plexus but the mechanism by which they accomplish this still remains unclear (Nikolskaia *et al*., 2006; Grab and Kennedy, 2008). Typical symptoms of this phase are serious sleep cycle disruptions (inspired the name of the disease: sleeping sickness), paralysis and progressive mental deterioration, all of which if left untreated are fatal (Checchi et al., 2008; Kabore et al., 2011; Jamonneau et al., 2012).

The progression from the first to the second stages of HAT infections by *T. b. gambiense* can last up to 3 years; contrary to that of *T. b. rhodesiense* which takes from a few weeks up to 6 months (Odiit et al., 1997). An important difference between *T. b. gambiense* and *T. b rhodesiense* is the amount of the trypanosomal chancres that appear after the tsetse bite. By definition, chancres are initial lesions characterized by local erythema, edema, heat, tenderness and a lack of any

suppuration following a tsetse bite (MacLean et al., 2010; Blum et al., 2012, Kennedy, 2013). The chancres from *T. b. gambensie* are rarely visible while approximately 26% of the patients infected with *T. b. rhodensiense* have chancres (MacLean et al., 2010; Blum et al., 2012).

T. b. brucei is non-human infective possibly due to a high-density lipoprotein, called trypanosome lytic factor (TLF) that is found in human serum and lyses the parasite (Hajduk et al., 1992; Smith and Hajduk, 1995; Smith et al., 1995). *T. b. rhodesiense* and *T. b. gambiense* are resistant to lysis by human serum *in vitro*, a phenotype enabling them to infect humans through the action of one or more genes (Turner et al., 2004), such as the serum resistance associated gene (SRA). In *T. b. rhodesiense*, the product of SRA has been determined to be responsible for this phenotype (Oli et al., 2006; Genovese et al., 2013). It is suspected that *T. b. brucei* lacks the human infecting phenotype due to genetic exchange that occurred during the evolution of the three subspecies (Faulkner et al., 2006). The extent to which these subspecies are truly distinct remains a largely debated matter, due to the identification of an intermediate phenotype in *T. b. brucei* which has the ability to develop resistance upon prolonged exposure to human serum (Turner et al., 2004).

1.2.3. Treatment

There are only five drugs registered and available for the treatment of African sleeping sickness, all of which are associated with side effects and increasing rate of treatment failure (Barret et al., 2003; Kennedy, 2004). Pentamidine is a diamidine compound usually used to treat early manifestations of the disease in *T. b. gambiense* (Bacchi, 1993). Pentadimidine has been associated with severe side effects such as nephrotoxicity and pancreatic damage (Paine et al., 2010). *T. b. rhodesiense* early infections are treated with Suramin, which is a colourless analog of the trypanocidal dye trypan blue (Barret et al., 2007). Severe diarrhoea, prolonged high fever, heavy proteinaria and exfoliative dermatitis are some of the side effects that have been observed post Suramin treatments. With the ability to enter the CNS, melarsoprol is thought to be a suitable drug for treating late-stage cases of *T. b. gambiense* and *T. b. rhodesiense* infections. Side

effects such as headaches, tremors, slurring of speech and convulsions have been associated with the usage of Melarsoprol, which also has ~8% fatality rate (Blum et al., 2001; Brun et al., 2010). Eflornithine, a derivative of ornithine, is used to treat both early and late infections of *T. b. gambiense* (Chappuis et al., 2005; Balasegaram et al., 2006; Priotto et al., 2008) and acts by inhibiting the enzyme ornithine caborxylase, which is involved in polyamine synthesis in trypanosomes (Vincent et al., 2012). The most common side effects include diarrhoea, anaemia, leukopenia and convulsions (Steverding and Wang, 2009).

1.3. Biology of *T. brucei*

As single-celled organisms, *T. brucei* and protozoan parasites of the Trypanosomatidae family have typical organelles to that of eukaryotes and also some unique organelles only found in trypanosomatids.

The mitochondrion

Trypanosomes have one large, spindle-shaped mitochondrion that spans the entire length of the cell joining the flagellum via the basal body. The mitochondrion is certainly one of the unique features found trypanosomes; it is subdivided in the cytoplasm of the parasite (Paulin, 1975). During its life-cycle, the trypanosome undergoes structural changes in which the mitochondrion extends and organizes its inner membranes (de Souza et al., 2009). Although the mitochondrial metabolism varies according to different stages of development, some trypanosomes lack functional mitochondria and therefore rely on the glycosome for production of ATP via glycolysis pathways. Indeed *T. brucei* possesses an interesting variation of the mitochondrial metabolism which occurs in the electron transport chain. The procyclic (PRO) form of the parasite has a complete citric acid cycle and respiratory chain (Durieux et al., 1991). Proteins involved in all complexes of *T. brucei* mitochondrial respiratome have been shown to be divergent (Acerstor et

al., 2011). During the bloodstream (BSF) stage of the parasite, mitochondrial metabolism is completely different from the PRO stage as it is based mainly on the glycolytic pathways and the mitochondrion houses only one complex respiratory chain (Chaudhuri et al. 2006).

The kinetoplast

The kinetoplast, also referred to as kinetoplast-DNA or kDNA, resides in the mitochondrion and constitutes 20% of the total DNA in the trypanosome. It forms a physical link with the flagellum through a structure called tripartite attachment complex that transverses the mitochondrial membranes (Gull, 2003). The kinetoplast is arranged into thousands of concatenated minicircles and maxicircles. The highly heterogenous minicircles (~1kb in length) remain in a covalently closed conformation. They encode unique trans-acting guide (g) RNAs, while the homologous maxicircles (~20-40kb in length) code subunits for mitochondrial respiratory complexes and ribosomal RNA genes (Lukes et al., 2005). The gRNAs act as templates to post transcriptionally modify the maxicicle pre-edited and never edited mitochondrial transcripts due to the incorporation or deletions of uridines (Matthews, 2005). The sequences of these gRNAs are especially heterogenous in *T. brucei* and are directly correlated with the extensiveness of RNA editing that takes place to produce mature mitochondrial transcripts (Stuart and Feagin, 1992).

The flagellum

The trypanosome flagellum possess common flagella features such as the presence of axoneme, but also exhibit a rare-axonemal structure, the paraflagellar rod (PFR) (reviewed by Gull, 1999). The motile flagellum of trypanosomes emerges from the flagellar pocket that is directly associated with the kinetoplast (Clayton et al., 1995). The roles of the flagellum include assisting the trypanosome mobility in cases where it is removed from vasculature, and also helping to

maintain a flow of fluid over the parasite's surface. Apart from its role in parasite motility, the flagellum also plays a crucial role in chemotaxis, cell signaling and host cell invasion (especially for the intracellular parasites) (Fridberg et al., 2007; Kohl et al., 2003; Kohl and Bastin, 2005).

1.3.1. Life-cycle

The relatively complex life-cycle of *T. brucei* has four main developmental stages: the epimastigotes (EPI), procyclic forms (PRO), long slender (LS) metacyclic trypomastigotes, and short stumpy (SS) metacyclic trypomastigotes which occur in the tsetse vector and mammalian host (Vickerman, 1985; Sherwin and Gull, 1988, Gull, 2002). The trypanosome is exposed to two contrasting environments in the cold-blooded tsetse vector and warm-blooded mammalian host and thus exhibit differentiated attributes in order to adapt to the changing environments (Barry and McCulloch, 2001, Peacock et al., 2014). The biphasic life-cycle of *T. brucei* involves distinctive morphological and metabolic changes which are characterized by alternate occurrence of proliferating and cell cycle arrested forms.

PRO trypanosomes replicate in the tsetse's mid-gut and migrate to the salivary glands where they transform into infectious stumpy metacyclic forms which are covered by protective variable surface glycoprotein (VSG). The VSG-expressing trypanosome is able to survive transmission after being inoculated into the mammalian host following the tsetse blood meal bite (Brun and Schonenberger, 1981; Czichos et al., 1986). Here, the trypanosome mitochondrion becomes structurally more intricate and plays a vital role in energy generation (Gull, 2002; Matthews, 2005). The kinetoplast also moves further forward from the extreme posterior of the cell (Figure 1.3 A). PRO trypanosomes express a surface coat of GPEET and EP procyclin, GPEET gets down-regulated in later stages of the development (Vasella et al., 2004). Once the cells have increased to a specific population density in the mid-gut, they migrate towards the salivary glands and divide asymmetrically on the way to produce long, slender and short EPI forms (Van Den Abbeele et al., 1999). Short EPI trypanosomes are believed to undergo sexual genetic exchange attaching

to the salivary gland of the tsetse vector (Urwyler et al., 2007). Once in the salivary gland, trypanosomes express a new surface coat of bloodstream alanine-rich protein (BARP) and differentiate into the infective metacyclic forms (Urwyler et al., 2007). This process occurs simultaneously with the acquisition of the VSG coat for evasion of the mammalian immune system (Pays, 2006; Taylor and Rudenko, 2006). PRO trypanosomes can be cultured fairly easily in rich medium in the laboratory and are well studied in comparison to the later EPI and metacyclic insect stages (Hirumi and Hirumi, 1989, Vasella and Boshart, 1996, Vassella et al., 2004).

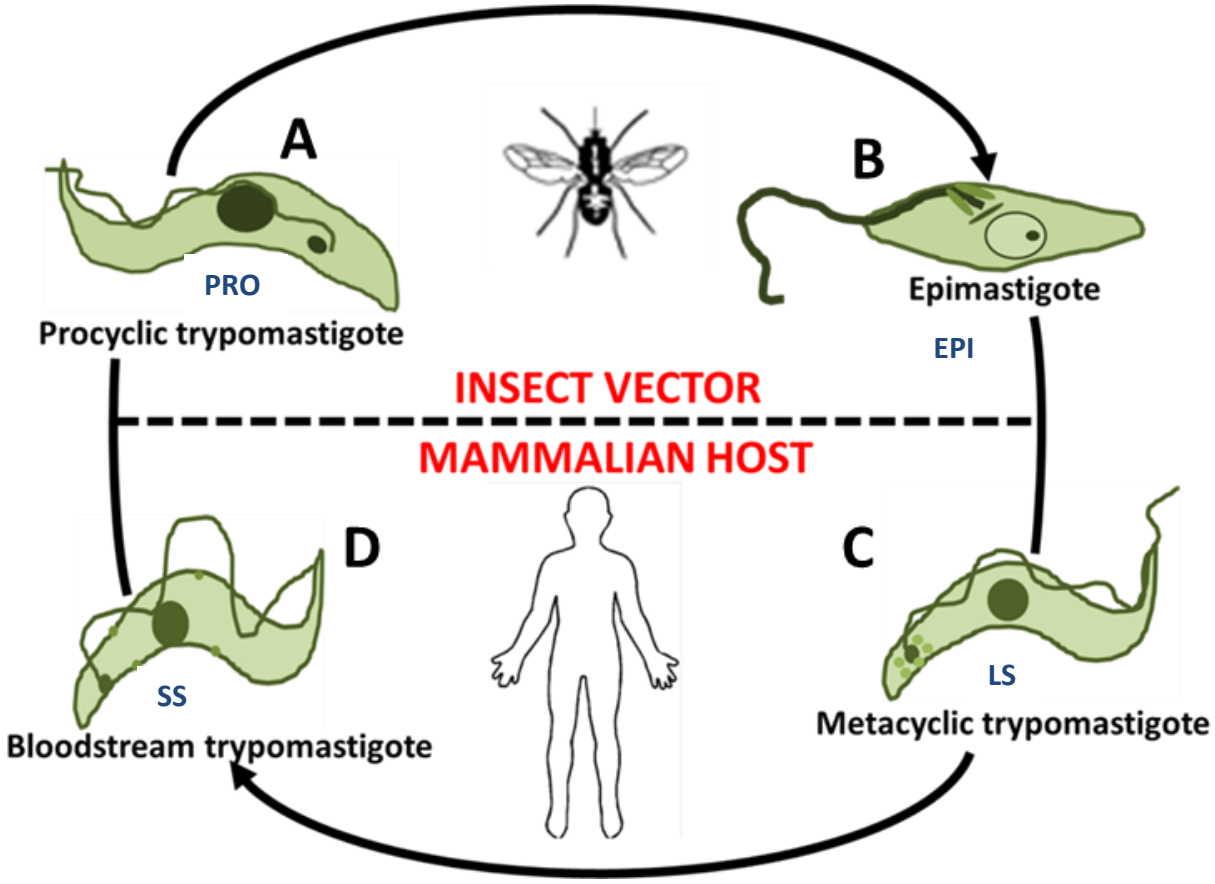


Figure 1. 3: The life-cycle of *T. brucei* as it transitions between the insect vector (A and B) and the human host (C and D). The PRO trypomasigote kinetoplastids situated more anterior into the cells. B, C and D indicating both the trypomastigotes and EPI which are spindle shaped cells with a kinetoplast situated behind the nucleus along with cell flagellum running along the majority of the cell body (Fenn and Matthew, 2007).

The mammalian stage is initiated by an infected tsetse injecting the SS metacyclic trypomastigotes (also known as larval trypanosomes) into the host during a blood meal (Vickerman, 1985). Upon entry into the mammal bloodstream, the SS metacyclic trypomastigotes convert into LS bloodstream trypomastigotes (Figure 1.3 D) (Barry and Emery, 1984; Mhlanga, 1994). The differentiation of non-proliferative SS bloodstream trypomastigotes (Figure 1.3 C) is largely dependent on the population density in the bloodstream and is achieved by moving through a number of intermediate forms as the parasite enters the lymphatic system and passes to the bloodstream (Hirumi and Hirumi, 1994). Although morphologically different, essential functions of the trypanosomes occur in both the LS and SS bloodstream trypomastigotes (Barry and McCulloch 2001, Seed and Sechelski, 2003).

The LS trypanosome cells are covered by bloodstream variant surface glycoproteins (BVSGs), which involves expression of the VSG gene from a distinct genomic location and enables the trypanosome population to evade the immune system (Aksoy et al. 2003; Fenn and Mathews 2007). LS trypanosomes are mainly associated with the early stage of the infection where they metabolize the abundant glucose present in the blood of the mammalian host. The mitochondrion of LS trypanosomes is not actively involved in the process of energy production and most respiratory chain enzymes are missing and therefore, these parasites depend entirely on the glycolytic process occurring in the glycosomes (Herman et al., 2008). Once the cells reach a specific threshold, they begin to differentiate into cell cycle arrested less active SS (Hamm et al., 1990). The morphology of SS trypanosomes differs from LS (Figure 1.3 C) because they are more compact in shape and have a shorter flagellum (Figure 1.3 D). SS trypanosomes are pre-adapted to the tsetse metabolic changes that allow them to switch from a glucose energy source in the mammalian bloodstream to proline energy source in the tsetse midgut (Hendriks et al., 2000). The SS trypanosome has an enlarged mitochondrion (Priest and Hajduk, 1994) (Figure 1.2 D).

1.3.2. Antigenic variation

In order to propagate freely in the bloodstream of the mammalian host, *T. brucei* has adopted a highly sophisticated evasion mechanism known as antigenic variation (McCulloch, 2004; Pays et al., 2004). Generally, antigenic variation is accomplished through the abundance of genes encoding antigenically distinct surface antigens required for host immune defense evasion and are expressed one gene at a time (Van der Woude and Baumler, 2004; Stockdale et al., 2008). As a requirement for antigenic variations, there must be a mechanism to facilitate antigen switching (gene switching) (Stockdale et al., 2008). In order to combat host immune system antibody mediated lysis, the parasite must either express a single antigen or employ the process of VSG switching. Single antigen expression results in gene silencing, which leads to the activation of another VSG gene. The process of antigenic variation has been described in bacteria, *P. falciparum, A. marginale* and *T. brucei*, where it has been extensively reviewed (Barry and McCulloch, 2001; Vanhamme et al., 2001; Donelson, 2003; Pays et al., 2004).

T. brucei achieve antigenic variation by means of sequential expression of antigenically distinct 58 kDa VSGs which are linked to the parasite surface membrane by glycosylphosphatidylinositol (GPI) anchor (Matthews, 2005). The structure of VSG is predominantly comprised of α-helices and a hyper-variable N-terminal sequence. VSGs are monoallelically expressed from a massive archive of genes (Blum et al., 1993; Morrison et al., 2009, Manna et al., 2014). Studies of VSGs have contributed largely to our current understanding of the trypanosome biology and mechanism, early dissection of GPI biosynthesis (Thomas, 1999) and model of intracellular transport studies (Field and Carrington, 2009). VSGs are rapidly exported to the cell surface in a process that involves complex formations in molecular changes which correlates with various post-translational modifications (Seyfang, 1990), rapid endocytosis and recycling pathways serving to carry VSG through the endomembrane system (Engstler et al., 2004). There is an abundance of VSG polypeptides, estimated at 10^7 copies and approximately 90% of cell surface polypeptides. It has also been observed that VSGs are rapidly capped with antibody with implications for immune evasion (Barry, 1979).

What makes the process of antigenic variation in *T. brucei* particularly interesting is the parasite's ability to adopt gene silencing and antigenic switching as a strategy for immune system evasion (Palmer and Brayton, 2007, Stockdale *et al.* 2008). The former involves transcriptional control which results in switching between different VSG expressions while DNA arrangement events slot previously inactive VSG genes into an active VSG expression sites (Rudenko, 1999; Manna et al., 2014). Eventually, the immune system is no longer able to combat the parasite due to the constant change of the VSG (Brun et al., 2010) and antibodies specific to VSGs are raised against the hyper-variable N-terminal domain (Berriman et al., 2005) which is a small part of the VSG molecule (Cross, 1990).

1.3.3. Model organism

There are several advantages to using protozoan parasites as models for animal-based experiments: they are economical to culture in the laboratory, can be used without ethical issues and are abundant, diverse and versatile micro-organisms (Montagnes et al., 2012). The genome of *T. brucei* has been sequenced (Berriman et al., 2005), therefore this organism has become a popular choice for i*n vitro* genetic and metabolic studies (Verner et al., 2010). Also, *T. brucei* can be cultured easily in the laboratory, its organelles such as the nucleus, mitochondria, glycosome and flagella can be purified as well. The cell structure of *T. brucei* is large enough to allow for high resolution immunofluorescence and electron microscopy analyses and it allows for genetic manipulations in which individual gene expression within an inducible system can be reduced or knocked-down using homologous recombinant or RNA inteference which was employed to study the phenotype of 7435 protein coding sequences (Alsford et al., 2011).

Throughout its life-cycle, *T. brucei* is exposed to multiple environmental changes, including temperature, pH and oxidative stress. In order to survive and propagate freely in these environments, the parasites employ, amongst other responses, antigenic variation (Pays et al., 2004) and the heat shock response which are homeostatic mechanisms of adaptation and

protection against the harmful effects of environmental stress (Oslon et al., 1994). The heat shock response is highly conserved and has been studied in many model organisms (Pallavi et al., 2010; Meyer and Shapiro, 2013). Heat shock proteins are induced in response to stress experienced by the cell (Finkelstein and Strausberg, 1983, Fink, 1999, Gross, 2004). This class of proteins plays a much greater role in kinetoplastids as they are essential for the survival, differentiation and pathogenicity of these parasites (Shonhai et al., 2011). A thorough understanding of heat shock proteins is lacking in *T. brucei*.

1.4. Molecular chaperones

Molecular chaperones are widely distributed, present in all organisms from eubacteria to higher eukaryotes and evolutionarily conserved (Ellis, 1987; Lindquist and Craig, 1988; Ellis and van der Vies, 1991). They facilitate the fate of cellular proteins *in vivo* by controlling the correct folding, binding and release of non-native polypeptides, thereby maintain protein homeostasis during normal and stressful conditions (Hartl, 1996; Broadley and Hartl, 2009). A major class of molecular chaperones is the heat shock proteins (Hsp). Molecular chaperones from the Hsp superfamily are involved in protein translocation into different subcellular compartments, the assembly and degradation of multi-protein complexes, cell signaling, protein activation, and facilitating proteolytic degradation (Becker and Craig, 1994; Young et al., 2003; Soti et al., 2005; Edkins and Boshoff, 2014). Other roles of molecular chaperones involve recognizing, binding and therefore facilitating the correct folding of cytosolic, mitochondrial and endoplasmic reticulum proteins (Hendrick and Hartl, 1993, Morimoto et al., 1994). They have been found on the outer plasma membrane and are also secreted into the extracellular environment where they are potentially involved in signaling functions (Henderson, 2010; Tamura et al., 2012), for example Hsp90N (Section 1.5.1).

1.5. Major heat shock proteins

Heat shock proteins are a large family of evolutionarily and structurally conserved proteins, some are induced due to exposure to stress, while others are constitutive and essential for cell growth under normal conditions (Lindquist, 1986; Lindquist and Craig, 1988; Morimoto et al., 1994). Other than heat shock, nutrient deprivation, hypoxia, exposure of tissues to uncouplers of oxidative phosphorylation, inhibitors of the electron transport chain, amino acid analogues, enzyme inhibitors and heavy metal ions are also stress stimuli that lead to the expression of heat shock proteins (Georgopulos, 1992). The induction of heat shock proteins upon exposure to environmental or oxidative stress is a common feature of all organisms from archaebacteria to eubacteria, plants, yeasts, invertebrates and vertebrates (Bonnefoy et al., 1994).

Heat shock proteins originally derived their nomenclature from their molecular size which ranges from 8-150 kDa. The major subfamilies are small heat shock proteins, Hsp40, Hsp60, Hsp70, Hsp90 and Hsp100, although a new updated nomenclature has been proposed (Kampinga et al., 2009). Members of the heat shock protein family that are synthesized constitutively are generally referred to as heat shock cognate (Hsc) proteins, while those induced by stresses are designated as heat shock proteins. Also, heat shock proteins can be classified based on their holding, folding and unfolding roles (Stirling et al., 2003). Holding chaperones are most likely to lack ATP-driven conformational changes as their role is to merely bind, and stabilize non-native proteins for refolding by other chaperones. Certain Hsp40 isoforms are a prime example of holding chaperones as they co-operate with ATP-dependent Hsp70 during the folding process (Mayer et al., 2000a). Small heat shock proteins are also holding chaperones as they bind and trap denatured proteins onto their surfaces (van Monfort et al., 2001). Folding chaperones, such as Hsp70, are capable of capturing and folding nascent polypeptides (Mayer et al., 2000a). The Hsp100/Clp/AAA (ATPases associated with various cellular activities) family is an example of the unfolding chaperone group that is implicated mostly in protein unfolding and disassembly (Stirling et al., 2003).

1.5.1. Hsp90 protein family

Hsp90 is a highly ubiquitous, conserved and essential component of the chaperone network found in bacteria and all eukaryotes (Buchner, 1999; Johnson, 2012). Indeed, Hsp90 accounts for 1-2% of the total soluble protein in the eukaryotic cytoplasm, and therefore represents one of the most abundant proteins even under permissive conditions (Welch and Feramisco, 1982), while levels increase 10-fold during stressful conditions (Lai et al., 1984, Borkovish et al., 1989, Buchner, 1999, Whitesell and Lindquist, 2005). Hsp90 is responsible for essential roles within the cell, some of which include cell-cycle control, signal transduction and hormone responsiveness (Pearl and Prodromou, 2000, Burrows et al., 2004; Seo et al., 2008; Echeverria and Picard, 2010, Echeverria et al., 2010). Hsp90 binds proteins that are almost folded or partially folded and regulates conformation and stability – rather than Hsp70 which is involved in *de novo* folding or stress related refoldings (Taipale et al., 2010). Currently, Hsp90 is known to interact with over 300 client proteins (http://www.picard.ch/downloads/Hsp90interactors.pdf) which are mainly involved in essential signaling transduction pathways, such as the steroid hormone receptors and multiple kinases (Pratt and Toft, 1997; Zhao et al., 2005, Jackson; 2012). Hsp90 clients are involved in numerous focused cell regulatory roles such as controlling cell homeostasis, proliferation, differentiation and cell death under both permissive and stress induced cellular conditions (Pearl and Prodromou, 2006; Taipale, et al., 2010). The activity of Hsp90 is regulated by a wide variety of co-chaperones and ATP hydrolysis (Li et al., 2012).

Bacteria have only one Hsp90 homologue called HtpG (high temperature protein G) whereas the unicellular eukaryotic organism, yeast has two homologues (Hsc82 and Hsp82) and most mammalians possess 4 Hsp90 homologues localizing in different compartments of the cell. In mammals, two distinct genes encode the cytosolic Hsp90 isoforms: Hsp90β is constitutively expressed (Meng et al., 1993) and was previously believed to exist predominantly as a homodimer whereas Hsp90α is inducible in response to cellular stress (Hickey et al., 1989) and was thought to exist mainly as a monomer (Nemoto et al, 1995). However, the current thinking

is that both isoforms of Hsp90 are constitutively dimerized via the C-terminal end (Wayne and Bolon, 2007). The amino acid sequence of human HSP90α is 85% identical to that of HSP90β, an attribute contributing to the varying chaperone activity of the two isoforms which may exhibit differential binding to client proteins (Pepin, 2001, Millson et al., 2007). HtpG is different to the eukaryotic isoforms and does not have homologues of the co-chaperone network present in eukaryotes (Street et al., 2014; Nakamoto et al., 2014).

Found in the endoplasmic reticulum (ER), 94 kDa glucose-regulated protein (GRP94) is required for the maturation and secretion of insulin-like growth factors (Wanderling et al., 2007). The Hsp90 isoform, tumour necrosis factor receptor-associated protein 1 (TRAP-1) serves a crucial role of protecting the cell against oxidative stress in the mitochondria (Felts et al., 2000; Hua et al., 2007). MZB1, which is a B-specific ER-localized protein was recently identified as a co-chaperone of GRP94 (Rosenbaum et al., 2014), while TRAP-1 co-chaperones are yet to be identified. Both GRP94 and TRAP-1 have adenosine triphosphatase (ATPase) activity (Dollins et al., 2007; Frey et al., 2007; Leskovar et al., 2008). A truncated 45kDa Hsp90 has been suggested to be associated with the membrane (Hsp90N). Hsp90N was identified in humans and found to have favorable sequence identity with two cytosolic isoforms of Hsp90 (Schweinfest et al., 1998; Grammatikakis et al., 2002; Zurawska et al., 2008). It remains to be seen whether Hsp90N (membrane associated Hsp90) is a true isoform or an artefact of one particular cell line (Grammatikakis et al., 2002). Hsp90 is also present in the chloroplast of eukaryotic cells and Hsp90C was found to be strongly heat inducible (Willmund and Schroda, 2005).

Hsp90s also belong to the GHLK (gyrase, histidine kinase, MutL) superfamily (Ban et al., 1999; Wigley et al., 1991;Bilwes et al., 2001) of proteins which share an uncommonly shaped ATP binding cleft, referred to as the Bergerat fold, responsible for critical ATPase activity (Dutta and Inouye, 2000). Hsp90 proteins have three domains: the 25 kDa N-terminal domain (NTD), 35 kDa middle domain (MD) and a 12 kDa C-terminal domain (CTD) all of which are highly conserved (Terasawa et al., 2005) (Figure 1.4). The NTD binds ATP, co-chaperones such as p23 and small

molecule inhibitors such as geldanamycin and radicicol (Grenert et al., 1997; Stebbins et al., 1997; Shiau et al., 2006) (Figure 1.4).

The ability of Hsp90 to fold and activate client proteins depends on its prior interaction and/or bridging with a number of different co-chaperone which forms part of the sequential dynamic multi-protein chaperone complex (Pearl and Prodromou, 2006; Pratt and Toft, 2003; Johnson, 2012). The NTD and MD connect through a flexible, charged linker region which is absent in prokaryotic organisms (Hainzl et al., 2009) (Figure 1.4). The MD is a known binding site for client proteins and co-chaperones (Ali et al., 2006). The CTD is crucial for the dimerization of Hsp90 (Nemoto et al., 1995). Positioned at the end of C-terminal domain of Hsp90 is a 4 amino acid motif (EEVD) that binds to tetratricopeptide repeat (TPR) domains of certain co-chaperones (Scheufler et al., 2000; Brinker et al., 2002). The binding of TPR co-chaperones of Hsp90 at the EEVD motif not only regulates the interaction of Hsp90 with client proteins but also its ATPase activity (Scheufler et al., 2000; Brinker et al., 2002).

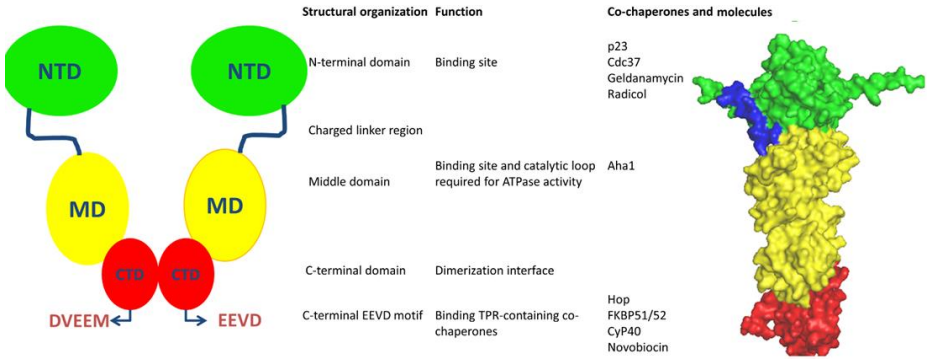


Figure 1. 4: Overview of the structure and domains of Hsp90 (PDB: 2CG9), the domains of the Hsp90 homodimer are annotated on the three-dimensional structure (Far right) generated by PyMol (DeLano, 2002). NTD is the N-terminal domain which binds to ATP, geldanamycin and radicicol and other small molecule inhibitors, together with co-chaperones and potentially client proteins. The NTD is followed by an unstructured highly charged linker region which is not present in the *E. coli* Hsp90 (HtpG). The Middle domain (MD) possesses a catalytic arginine residue which is required for ATPase activity, binds to co-chaperones and is thought to be a major client-protein binding site. CTD is where C-terminal domain which contains the major dimerization interface which makes Hsp90 a constitutive dimer. At the C-terminal end of the CTD, there is a highly conserved EEVD motif which binds to TPR-containing co-chaperones (adapted from Jackson, 2012).

All three domains of Hsp90 are involved in peptide binding (Scheibel et al., 1999; Sato et al., 2000; Terasawa et al., 2005). It has been suggested that the transient dimerization of Hsp90 through the MD and NTD is essential for the ATPase activity of the protein (Prodromou et al., 2000; Meyer et al., 2003; Wegele et al., 2003). It is speculated that Hsp90 acts as molecular "clamp" whose operation is driven by its ATPase activity and regulated through momentary opening and closing events (Prodromou et al., 2000). The CTD is also known to bind small molecule inhibitors such as novobiocin, cisplatin and taxol (Marcu et al., 2000a, Marcu et al., 2000b; Matts et al., 2011). These molecules are believed to bind at the less well characterized second ATP-binding site which opens up when the N-terminal binding site is occupied (Söti et al., 2002). The cohort of co-chaperones present in the Hsp90 multi-protein chaperone complex regulates the ATP dependent opening and closing of Hsp90 (Pearl and Prodromou, 2000; Pratt and Toft, 2003). Hsp90 fulfills its function in the dimeric conformation; its structures circulate between the closed and open state as regulated by co-chaperones, nucleotide binding and client binding and release (Jackson, 2012; Li et al., 2012).

1.5.2. Hsp70 protein family

The heat shock protein (Hsp70) family is ubiquitous and the most highly conserved of all heat shock proteins. These proteins are present in all organisms from archaebacteria to higher eukaryotes, are expressed constitutively and are also upregulated in response to cellular stress (Kumar et al, 1991; Nussinov and Tsai, 2013). The general roles of Hsp70s are the co-ordination of fundamental cellular homeostasis processes such as the folding and assembly of nascent polypeptides, refolding of misfolded and aggregated proteins (Mayer and Bukau, 2005). Hsp70s also translocate secretory proteins, control regulatory proteins and play a role in the degradation of proteins (Cyr et al., 1992; Bukau and Horwich, 1998; Mayer and Bukau, 2005). Hsp70s also share high sequence conservation and homology which suggests a conservation of function and structure.

The role of DnaK, the prokaryotic (*E. coli*) homologue of Hsp70, in folding nascent peptides has been extensively characterized (Hartl and Hayer-Hartl, 2002). There are two other paralogues of DnaK present in *E. coli* known as Hsc62 (Arifuzzaman et al., 2002; Yoshimune et al., 2002) which is involved in the negative regulation of transcription; and Hsc66 (Silberg et al., 1998) which prevents protein aggregation. Unlike prokaryotes, eukaryotes possess multiple copies of the *HSP70* gene localizing in different cellular compartments and displaying even more diverse functions (Desai et al., 2010). Eukaryotic Hsp70s are distributed based on their cytosolic, ER and mitochondrial localizations (Boorstein, et al., 1994; Karlin and Brocchieri, 1998; Shonhai et al., 2007; Kampinga and Craig, 2010). *S. cerevisiae* possesses 6 cytosolic Hsp70s (SSA1, SSA2, SSA3, SSA4, SSB1 and SSB2), 3 genes encode the mitochondrial Hsp70 complement (SSC1, SSQ1/SSC2, and ECM10/SSC3) and SSD1 encodes the single endoplasmic reticulum Hsp70 (Werner-Washburne and Craig, 1989). The human genome encodes 13 members of the human Hsp70 (HSPA) family, as well as many pseudogenes (Kampinga et al., 2009). Although human Hsp70 proteins play a significant role in maintaining cellular homeostasis, they do so in different and separate compartments. The vast majority of Hsp70 proteins are mainly cytosolic and nuclear in location (HSPA1A/B, HSPA1L/HSPA2, HSPA6 and HSPA8), with only a single protein located in the mitochondria (HSPA9) and the ER (HSPA5) (Kampinga et al., 2009).

Structurally, Hsp70 is comprised of three main domains: 45 kDa N-terminal ATPase or nucleotide binding domain (NBD) that is connected to the 15-30 kDa central substrate binding domain (SBD) and 10 kDa C-terminal domain (Mayer et al., 1998, Mayer and Bukau, 2005) (Figure 1.5). The NBD is further divided into four subdomains IA, 1B, IIA and IIB. The SBD of Hsp70 has been classically subdivided into two distinct sub-domains: β-domain and α-helical domain (Strub et al. 2003). The β-domain is where substrates bind in the substrate binding pocket, while the α-helical domain located closest to the C-terminus functions as a lid over the substrate. Positioned at the C-terminal domain of Hsp70 is a 4 amino acid motif (EEVD) that binds a highly conserved TPR binding site (Scheufler et al., 2000; Brinker et al., 2002).The NBD interacts with ATP resulting in conformational changes in the structure of Hsp70 and is also involved in regulating the function of the SBD (Blamowska et al. 2010). The binding of substrate to the SBD stimulates the ATPase

domain (Borges and Ramos, 2008). It has been shown that the linker region, which links the NBD and SBD, is involved in regulating the ATPase activity by binding to ATPase domain (Swain et al. 2008).

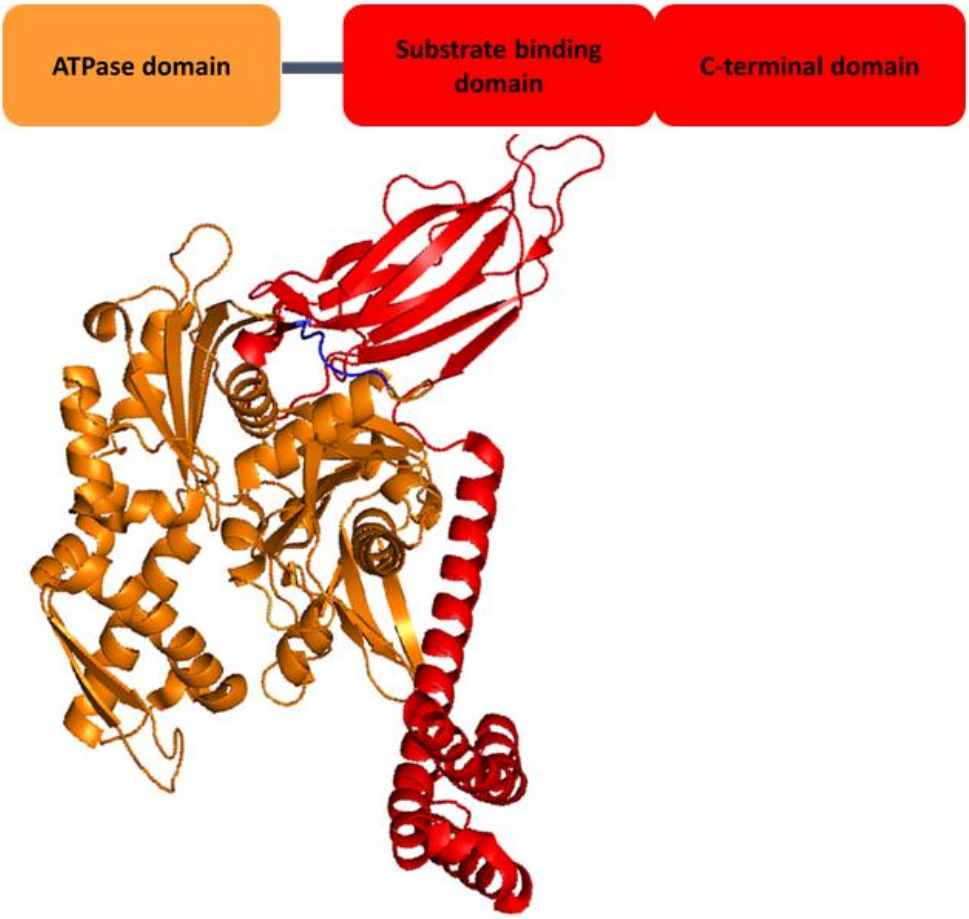


Figure 1. 5: Domain organization (top) and structure of an Hsp70 protein (bottom). Schematic representation of the ATPase domain that is connected to the substrate binding domain by the linker region. Full length structure of *S. cerevisiae* SSE1 (PDB: 3D2F) with the ATPase domain connected to the substrate binding domain that is comprised mainly of β-sheets and C-terminal domain comprised mainly of α-helices.

1.5.2.1. Nucleotide exchange factors

Nucleotide exchange factors (NEFs) are structurally unrelated proteins that interact with Hsp70 at its ATPase domain (Brehmer *et al*., 2001). To date, five main classes of NEFs have been characterized (Dragovic et al., 2006), these proteins interact with Hsp70 in distinct manners. NEFs form physical contacts with Hsp70 by either stimulating the release of ADP by tilting out subdomain IIB of the NBD which results in opening the nucleotide binding cleft, or displacing ADP by wrapping around subdomain IIB of the NBD resulting in a massive destabilization of the subdomains IA and IB (Brehmer *et al*., 2001).

The classes of NEF include GrpE in prokaryotes which acts as a dimer to facilitate nucleotide exchange (Harrison et al., 1997). Hsp110s are described as highly divergent members of the Hsp70 family. The structure of Hsp110 closely resembles that of Hsp70 (Polier et al., 2008). Although Hsp110s have been reported as having similar chaperone properties to Hsp70s (Oh et al., 1997; Santos et al., 1998), differences in their functions have also been noted. Hsp110s have been shown to act as NEFs for cytosolic Hsp70s (Dragovic et al., 2006). Hsp110 is crucial to the ATP cycle required for the chaperone activity of Hsp70. Proteins containing a Bcl2-associated athanogene (Bag) domain serve as NEFs of cytosolic Hsp70s (Höhfeld and Jentsh, 1997; Brehmer et al., 2001). This is important, since no GrpE-like homologue has been found in the eukaryotic cytosol (Sonderman et al., 2001). The interactions between the Bag domain and Hsp70 have been characterized by Sonderman et al. (2001). In terms of mammalian ER-Hsp70s, a BiP associated protein (BAP), has been identified as a NEF (Chung et al., 2002), however, the nature of its interaction with ER-Hsp70 has yet to be determined.

1.5.2.2. J-proteins

The J-proteins also known as the DnaJ, Hsp40 or J-domain protein family exhibits greater protein numbers, heterologous sequences, structural diversity and function per cellular system

compared to Hsp70 (Wittung-stafshede, 2003). J-proteins serve in partnership with Hsp70 as co-chaperones (Cheetham et al., 1994). This is achieved through interaction with both the substrate binding and ATPase domains of Hsp70 (Suh et al, 1998). The Hsp70/J-protein partnership is conserved across most organisms. J-proteins target protein substrates to Hsp70 for folding and then stabilize Hsp70 in a substrate-bound form (Cheetam et al., 1994). J-proteins have been associated with a wide variety of cellular roles such as protein translocation (Jubete et al., 1996), protein degradation (Jiang et al., 1997), clathrin uncoating (Campell et al., 1997; Ma et al., 2002), and viral infection (al-Herran and Ashraf, 1998).

A typical J-protein has a conserved J-domain that is approximately 70 amino acids long consisting of four α-helices and the highly conserved tripeptide signature motif HIS-PRO-ASP (HPD) in the extended loop between the two main helices (II and III). The presence of the J-domain defines all J-proteins, it is critical for their interaction with Hsp70 which is achieved through the HPD motif stimulating the ATPase activity of Hsp70 (Craig et al., 2006; Hageman and Kampinga, 2009). The mechanism by which the J-domain stimulates the ATPase activity of Hsp70 and ensures conformational changes resulting in the stabilization of client proteins remains a largely debated matter. However, it is established that exposed residues of the J-domain form an Hsp70 interaction surface (Greene et al., 1998; Jiang et al 2007).

Despite the presence of the J-domain, J-proteins as a group have a wide variety of additional domains. Historically, J-proteins were divided into three classes: Type I (A), Type II (B) and Type III (C), with type I (A) designated based on motifs/ domains found in *E. coli* DnaJ (Cheetham and Caplan, 1998; Hennessy et al., 2000; Hageman and Kampinga, 2009). Type I consist of an N-terminal J-domain, followed by a GLY/PHE –rich region (GF-domain), four repeats of the cysteine rich sequences type zinc finger (Zn-domain) and a C-terminal domain known to bind substrate proteins (Goffin and Geopoulus, 1998; Lu and Cyr, 1998; Li et al., 2003). The C-terminal domain is composed of two barrel topology domains, CTDI and CTDII (Figure 1.6). Also, type I J-proteins have a hydrophobic pocket in which client proteins are thought to bind, as well as the Zn- domain

extruding from it which is involved in substrate binding (Figure 1.6) (Linke et al., 2003; Kota et al., 2009). J-proteins classified as type II (B) have all domains in common with type I except they lack the Zn- domain. The type III J-proteins only have the J-domain in common with other J-proteins. Type IV have thus far been described in *S. cerevisiae* (Walsh et al., 2004) and *P. falciparum* (Botha et al., 2007), and are defined as J-proteins with an abrogated HPD motif. By this definition, the human DNAJB13, which lack the HPD motif is a putative type IV (Guan and Yuan, 2008; Guan et al., 2009). Type IV J-proteins have more recently been referred to as "pretenders" (Pesce and Blatch, 2014), a lot of work remains to be done fully characterizing the biochemical roles of these proteins.

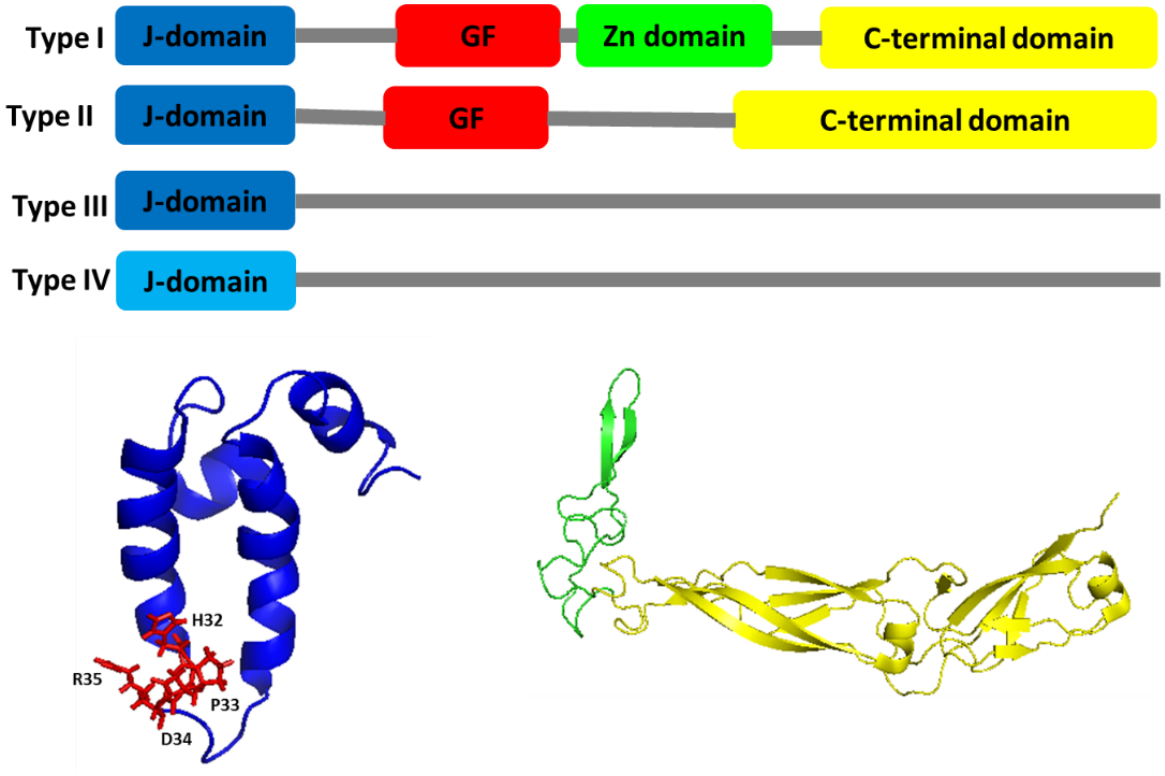


Figure 1. 6: Overview of the domain structures and organization of J-proteins. Top: Type Is consist of the 70-amino acid J-domain, GF rich region, Zn binding domain and the C-terminal domain while type IIs lack the Zn binding domain. Type IIIs possesses only the canonical J-domain and type IVs have a J-domain with a corrupt HPD motif. The structures of the J-domain are shown in blue (PDB code: 1BQ0; Huang et al., 1998) and the zinc binding domain is shown in green and CTD in yellow (PDB code: 1NLT; Li et al., 2003), images were generated by PyMol (DeLano, 2002).

The GF region is thought to be important in the regulation of the substrate binding capabilities of Hsp70 (Pellecchia et al., 1996), potentially through stabilizing the Hsp70-substrate complex (Genevaux et al., 2002; Han and Christen, 2003) and optimizing substrate selection (Kelley, 1999) or facilitation of the transfer of substrate from J-protein to Hsp70 (Rosser and Cyr, 2007). The GF domain has a conserved ASP-ILE-PHE (DIF) motif and is thought to play a significant structural or functional role, possibly in the regulation of the Hsp70 chaperone cycle subsequent to ATP hydrolysis (Wall et al., 1995; Cajo et al., 2006). Type I and Type II J-proteins display differences in the GF domains and it is proposed that these differences may contribute to different chaperone functions. The ZN- domain is characterized by the presence of four cysteine-repeat sequences (CXXCXGXG), capable of co-ordinating the binding of two zinc ions. The possible role of this region in stabilizing the J-protein tertiary structure has been speculated (Greene et al., 1998; Martinez-Yamout et al., 2000). Other studies suggest that the presence of the Zn- domain together with the C-terminal domain is a pre-requisite for substrate binding (Han and Christen, 2003). The C-terminal domain is most variable of all domains and has been implicated in substrate binding (Banecki et al., 1996), it is also involved in dimerization which maybe essential for the activities of some J-proteins (Borges et al., 2005; Wu et al., 2005).

1.5.2.3. Hsp70 chaperone folding cycle

In order to fulfill its function, Hsp70 requires the partnership of co-chaperones such as J-proteins and nucleotide exchange factors to regulate its activity. The Hsp70/J-protein chaperone system is conserved across most organisms and is widely studied. Hsp70 has low basal ATPase activity and requires J-proteins to stimulate the ATPase activity. The affinity of Hsp70 for substrate is determined by the binding and hydrolysis of ATP (Figure 1.7). ATP-bound Hsp70 has lower affinity for substrate while ADP-bound has higher affinity (Hiromura et al, 1998; Mayer and Bukau, 2005). J-proteins bind to the ATPase domain of Hsp70 and stimulate the hydrolysis of ATP to ADP. The ADP-bound Hsp70 changes its conformation thereby resulting in greater affinity for the protein substrate (Figure 1.7). NEFs facilitate the release of ADP from Hsp70 resulting in a conformation change that allows substrate dissociation and primes Hsp70 for another cycle of substrate

binding (Dragovic, et al., 2006). The regeneration of the ATP bound form of Hsp70 by nucleotide exchange factors results in a decrease in affinity for the substrate and hence release of the substrate protein to fold properly or re-enter the folding cycle (Fink, 1999; Mayer et al., 2000).

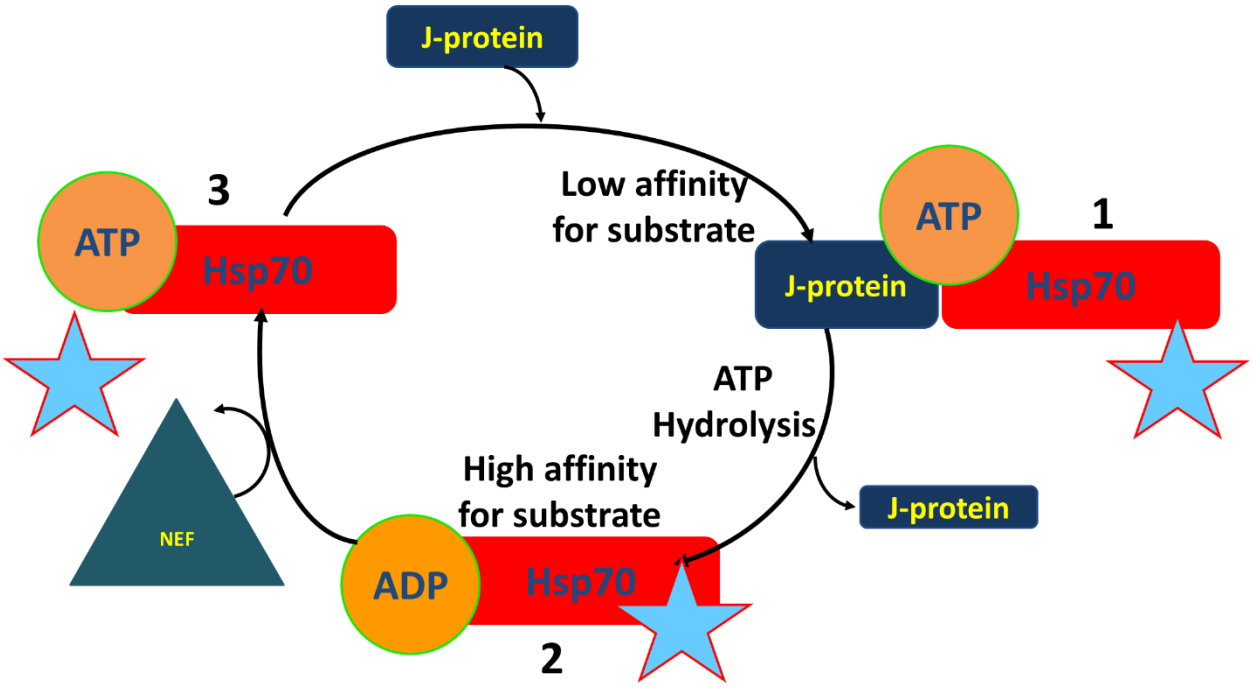


Figure 1. 7: The ATP mediated cyclic interaction of Hsp70, NEF and J-proteins during substrate folding. 1. Low affinity for substrates: J-proteins bind substrates (indicated by the star) and transfer them to Hsp70-ATP and stimulate the ATPase activity of Hsp70 resulting in Hsp70-ADP. J-protein dissociates. 2. In the ADP bound state, Hsp70 has high affinity for substrate, 3. NEF facilitates the release of ADP and the binding of ATP therefore releasing substrates, if necessary the cycle repeats until the substrate is completely folded or the substrate is transferred to another molecular chaperone (adapted from Pesce and Blatch, 2009; Hatherley et al., 2013).

1.6. TPR containing co-chaperones of Hsp70 and Hsp90

Tetratricopeptide (TPR) repeat domains are defined as a degenerate amino acid sequence of protein-protein interaction motifs (Hirano et al., 1990; Sikorski et al., 1990). A typical TPR domain comprises of tandem repeat occurrences of a 34-residue degenerate sequence (Goebl and Yanagida, 1991; Lamb et al., 1995). TPR domains occur in proteins which play diverse roles in the

cell including cell differentiation, transcription and protein transport (Goebl and Yanagida, 1991; Lamb et al., 1995; Blatch and Lässle, 1999; Cliff et al., 2005). Although TPR domains occur in many proteins, a number of Hsp70 and Hsp90 co-chaperones contain these domains, using them as a primary site for interaction with the EEVD motif (Odunuga et al., 2003; Odunuga et al., 2004). The most notable feature of TPR domain containing co-chaperones of Hsp70 and/or Hsp90 is the ability to regulate the activity of the chaperones (Smith, 2004). Co-chaperones such as Hip, STI1, Chip, and Cyp40 play crucial roles in the Hsp70/Hsp90 complexes which are involved in steroid receptor maturation (Smith, 2004).

1.7. Stress inducible protein 1 (STI1)

The 60 kDa stress inducible protein 1 (STI1), otherwise known as Hsp70/Hsp90 organizing protein (Hop) or p60 is found as a single copy in several eukaryotic organisms. It was first isolated and characterized in yeast (Nicolet and Craig, 1989), the system in which it has been most extensively studied (Albanèse et al., 2006). As its name suggests, STI1 is induced significantly in response to stress (Nicolet and Craig, 1989). STI1 is a TPR-containing co-chaperone. The mammalian homologue contains three separate TPR domains that bind to Hsp70 and Hsp90 in a multi-chaperone complex (Hernandez et al., 2002). In collaboration with Hsp70 and J-protein, STI1 also modulates the biogenesis and function of prions in yeast and humans (Jones et al., 2004; Reidy and Masison, 2010; Roffe et al., 2010). Studies of the oligomeric state of the STI1 protein have been somewhat inconsistent, some authors have suggested that this co-chaperone exists as a monomeric protein (Young et al., 1998; Yi et al, 2010; Li et al, 2011), while others have seen the existence of a dimeric species (Prodromou et al., 1999; Flom et al., 2007; Onuoha et al., 2008). One study has shown that STI1 can exist both as a monomer and dimer (van der Spuy et al., 2001). STI1 inhibits the ATPase activity of Hsp90 by blocking access to the nucleotide binding-pocket, while yeast STI1 stimulates the ATPase activity of SSA1 (Wegele et al., 2003). The mammalian homologue of STI1, human Hop, has recently been shown to bind and hydrolyze ATP at the TPR1-DP1-TPR2 regions of the protein (Yamamoto et al., 2014).

Several eukaryotic STI1 proteins have been studied and it has been shown that its localization is widely distributed across the cell. Mouse STI1 (mSTI1) has been found in the cytoplasm (Lässle et al., 1997), Golgi apparatus and vesicles (Honore et al., 1992), cell surface (Martins et al., 1997; Zanata et al., 2002) and within the membrane fraction (Mehrpour et al.,2010). The nuclear localization signal (NLS) of mSTI1 has been demonstrated and is regulated by cell cycle kinases (Longshaw et al., 2003). After exposure to heat shock, mSTI1 localized to the nucleus, it is suspected that it contains two NLS signals (Daniel et al., 2008). Together with other co-chaperones, mSTI1 and human Hop were recruited to stress granules (Lapointe et al., 2009). It is believed that STI1 also localizes externally, primarily in mammalian brain tissue, however some studies have shown the secretion of STI1 by other tissue cultures in mice (Eustace and Jay, 2004; Lima et al., 2007) and ovarian cancer cells in human (Wang et al., 2010).

1.7.1. Structure of STI1

STI1 homologues in various eukaryotic organisms are structurally characterized by the presence of 9 TPR motifs which constitute 3 TPR domains: TPR1, TPR2A and TPR2B (Scheufler et al., 2000; Schmidt et al., 2011; Li and Du, 2013) and two aspartic acid-proline (DP) rich regions. The crystal structures of TPR1, the tightly linked TPR2A and TPR2B and DP regions for human Hop and yeast STI1 have been solved by X-ray crystallography and nuclear magnetic resolution (NMR) (www.pdb.org: Scheufler et al., 2000; Kajander et al.,2009; Schmid et al., 2012). The structure of the glucocorticoid receptor in complex with Hsp90, Hsp70 and STI1 has also been recently resolved by cyro electron microscopy (Kirschke et al., 2014). Although regions separate or adjacent to it may influence binding, the TPR domain is the main mediator of heat shock proteins binding to STI1 (Chen and Smith, 1998; Smith, 2004; Scheufler et al., 2000). The TPR domains are well conserved, structurally and functionally. The N-terminal TPR domain (TPR1) binds to the C-terminal peptide EEVD motif of Hsp70 (Chen et al., 1996; Lässle et al., 1997; Chen and Smith, 1998; van Der Spuy et al., 2000) and Hsp104 (Abbas-Terki et al., 2002) (Figure 1.8). The middle domain TPR2A binds to the C-terminal peptide EEVD on Hsp90, while earlier studies show the

ability of TPR2B to bind the highly acidic pentapeptide EEVD motifs of Hsp70 and Hsp90 (Scheufler et al., 2000; Southworth and Agard, 2011a) and it is also believed to play a role in the dimerization of STI1 (Longshaw et al., 2009). Recent studies have shown that both TPR2A and TPR2B of STI1 bind to the C-terminal and middle domains of Hsp90 respectively (Schmid et al., 2012; Lee et al., 2012). Side chains of certain conserved residues within the TPR domains confer the binding of the TPR domains to the EEVD of Hsp70 and Hsp90 thereby forming the carboxylate clamp (Scheufler et al., 2000). The carboxylate residues interact with the negatively charged side chains of the EEVD motif (Scheufler et al., 2000).

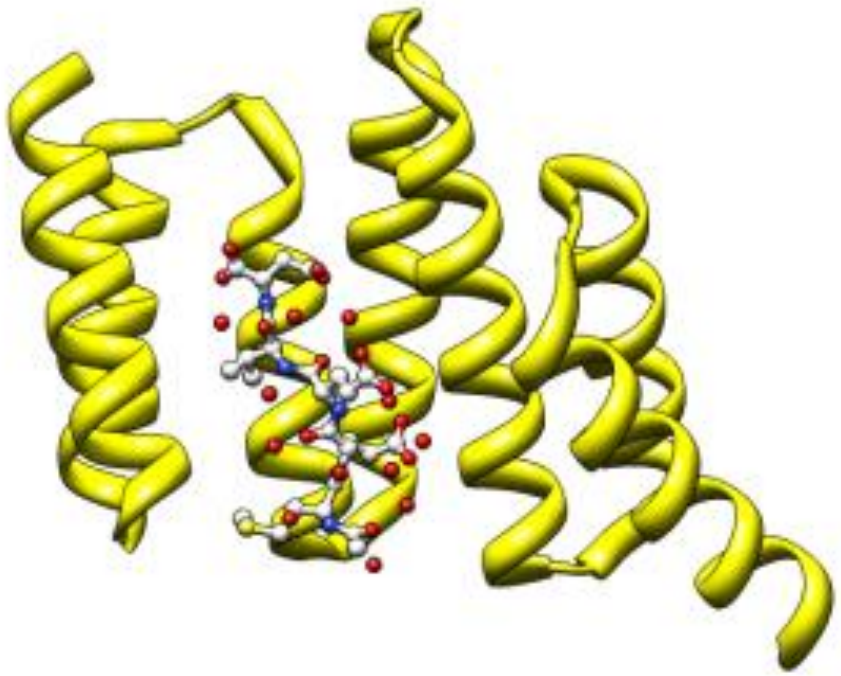


Figure 1. 8: The structure of the TPR1 domain of human STI1 (yellow) (PDB code: 1ELW) in complex with the Hsp70 EEVD motif (ball and stick), generated using PyMol (DeLano, 2002). The TPR domain of STI1 has the classical helix-turn-helix conformation.

The specificity of each TPR domain of STI1 is conferred by hydrophobic amino acids occurring adjacent to the TPR domain groove (Odunuga et al., 2003). Hsp70 and Hsp90 are not associated

with one another in the absence of STI1 (Chen et al., 1996b). The crystal structure of Hsp70 GPTIEEVD motif (PDB code: 1ELW; Scheufler et al., 2000) in complex with TPR1 (Figure 1.8) and the Hsp90 MEEVD motif (PDB code: 1ELR; Scheufler et al., 2000) in complex with TPR2A shows specific interaction between motifs and carboxylate clamp residues (Brinker et al., 2002; Odunuga et al., 2003). Thus the TPR domains facilitate STI1's function as a modulator of Hsp70 and Hsp90 activity. TPR1 and TPR2B-DP2 residues have been identified to be co-evolving with Hsp70, while the residues in TPR1, TPR2A and TPR2B have been identified to co-evolve with Hsp90 (Travers and Fares, 2007).

The DP regions are so-called due to the well conserved residues aspartic acid (D) and proline (P) repeats within the structures (Chen et al., 1996, Odunuga et al., 2004; Schmid et al., 2012). STI1 possesses two structurally similar DP domains which lie between TPR1 and TRP2A (DP1), and TRP2B and the C-terminal end of STI1 (DP2) (Schmid et al., 2012). Studies have shown that yeast STI1 DP domains, more specifically DP2, are important for Hsp90 client processing (Schmid et al., 2012).

1.7.2. Hsp70-STI1-Hsp90 complex

STI1 plays a crucial, but indirect role in the *in vivo* folding of proteins. This is achieved by mediating Hsp70 and Hsp90 interactions. The Hsp90 hetero-complex or multi-chaperone complex has been reviewed extensively in recent years (Southworth and Agard, 2011; Schmid et al., 2012; Li et al., 2012) and in the past wherein the complex composition was obtained through studying the progesterone receptor or the glucocorticoid receptor as stringent Hsp90 substrates (Smith 1993; Hutchison et al. 1994). For the complex to function effectively, coordinated interactions of Hsp70 and Hsp90 in connection with their co-factors and client/substrate proteins is critical, the complex is divided into 3 stages: namely the early, intermediate and mature complexes (Figure 1.9).

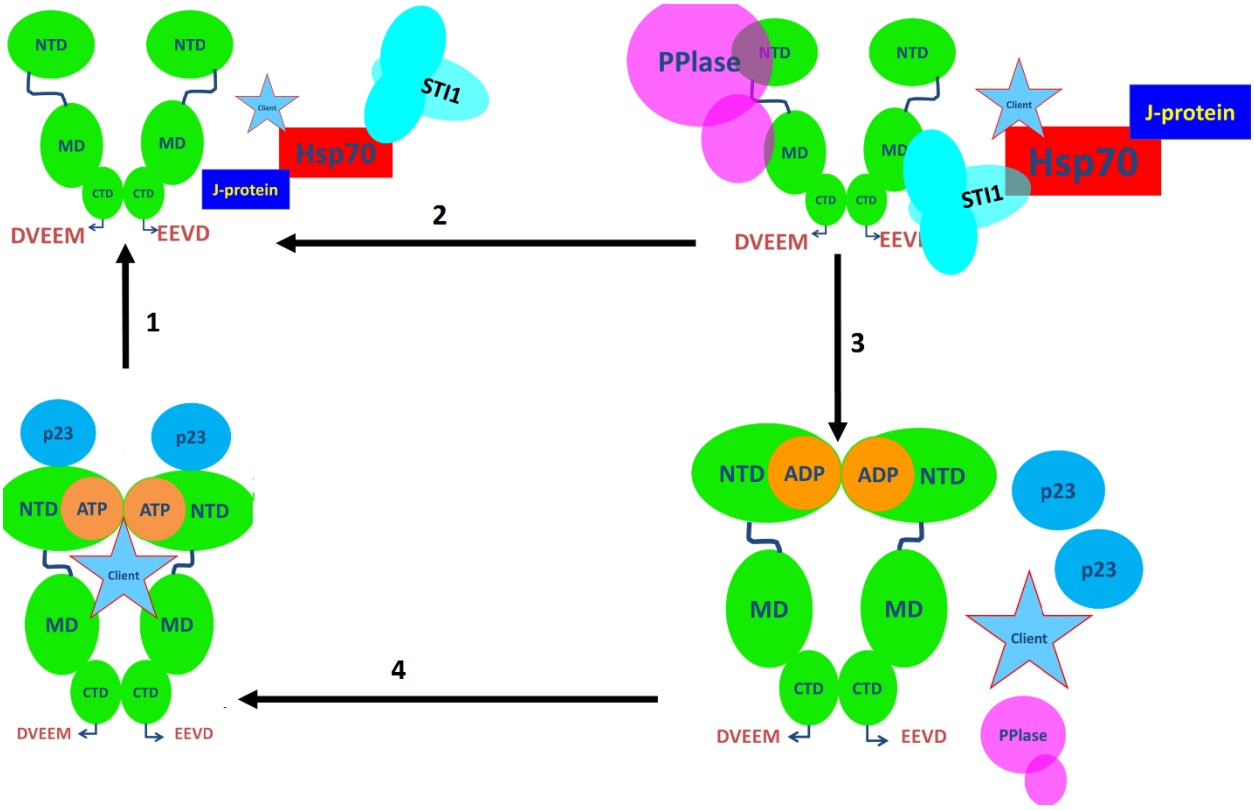


Figure 1. 9: The proposed model for the multi-chaperone complex. 1). Dimeric Hsp90 in its open conformation before association with STI-Hsp70/J-protein, the binding of Hsp70 to a J-protein is facilitated by ATP. 2). The Hsp90 C-terminal EEVD motif binds to STI1, TPR2A domain while the middle domain binds to TPR2B, inhibiting the ATPase activity. 3). Hsp70 is released and the client protein is transferred to the Hsp90's hydrophobic, homodimer cleft 4. STI1 is also released, allowing ATP to bind and processing of the client protein by Hsp90. The cycle is completed after ATP hydrolysis and release of ADP and the activated client occurs, and Hsp90 returns to the relaxed open state.

The early complex occurs prior to the addition of Hsp90 to the assembly of the complex, therefore Hsp70 plays a central role in capturing the client protein together with its co-chaperone J-protein (see Hsp70 chaperone cycle). In the intermediate complex, STI1 stabilizes the client-loading by binding to dimeric Hsp90 in its open conformation at the N-terminal domain, thus inhibiting ATPase activity. It is believed that the concave surface of TPR2A binds to the C-terminal of Hsp90, while the convex surface of TPR2B interacts with the M-domain (Southworth and Agard, 2011) (Figure 1.9). In the progression to the mature complex, Hsp70-coupled substrates bind due to competitive binding. The Hsp70 C-terminal is initially bound to the TPR1 of STI1 and it is thought that DP1 may stabilize the bound client (Schmid et al., 2012) interacting with STI1

and Hsp90. Following this, Hsp70 and the client protein are transferred to TPR2B-DP2 which facilitates the release of the client protein to hydrophobic residues within the Hsp90 interdimer cleft (Southworth and Agard, 2011). The release of STI1 and Hsp70 results in ATP binding to Hsp90. Subsequent to this, the Hsp90 N-terminal domains dimerize forming the closed conformation. The binding of Hsp90 to STI1 reduces the number of Hsp70 binding sites on the STI1 dimer from two sites in the absence of Hsp90 to one site in its presence. STI1 can inhibit the ATP binding and p23 binding activity of Hsp90, which can be reversed if Hsp70 is present in the complex (Hernánde et al., 2002).

The observation that inhibition or depletion of molecular chaperones, particularly Hsp90, results in the induction of the heat shock response (Marchler and Wu, 2001) to compensate for the loss of protein has led to new motivation for pursuing not only inhibitors designed to disrupt the activities of Hsps, but also to inhibit the formation of the Hsp90-STI1-Hsp70 multi-chaperone complex altogether (Horibe et al., 2011). It has been previously demonstrated that a new peptide antagonist interferes with the interaction of Hsp90 and STI1 in a human pancreatic BxPC3 cancer cell line. This particular peptide agonist is designed to mimick TPR2A domain of STI1 and has been shown to disrupt the binding of Hsp90 to TPR2A both *in vitro* and *in vivo* (Horibe et al., 2011). Also, a compound inhibitor of Hsp90 (1, 6-dimethyl-3-propyl pyrimido (5, 4-e) (1, 2, 4) triazine-5, 7-dione) was found to disrupt the Hsp70-STI1-Hsp90 multi-chaperone complex. Further analyses have shown that this compound also inhibits differentiation of breast cancer cells that are highly metastatic (Pimienta et al., 2011). The inhibition is thought to be achieved by disrupting the cancer cells morphology ultimately resulting in prevention of spheroid formation and migration of cancerous cells. Studies have also explored small molecule inhibitors (Yi and Regan, 2008), as well as modified TPR modules which will bind with greater affinity and specificity to Hsp90 and therefore block its interaction STI1 (Cortajarena et al., 2008). Some compounds block the interaction between STI1 and Hsp90 but do not induce Hsp70 (Pimienta et al., 2011). A cell-permeable hybrid TPR peptide specifically inhibits the interaction of Hsp90 with STI1 and affects Hsp90 clients resulting in cytotoxicity to cancer cells (Horibe et al., 2011; Horibe et al., 2012). A

synergistic cytotoxic affect on cancer cells was observed when peptides targeting both Hsp70 and Hsp90 were used on cancer cells (Horibe et al., 2014).

1.7.3. STI1 in parasites

The presence of the Hsp90-STI1-Hsp70 heat shock protein complexes in *P. falciparum* and *T. cruzi* have been confirmed where its canonical modulator role was demonstrated (Gitau et al., 2011, Schmidt et al., 2011). Also, phosphorylation of *Leishmania donovani* (*L. donovani*) LdSTI1 was shown to improve adaptability of the parasite when exposed to stressful conditions (Morales et al., 2010). It was also shown that LdSTI1 is phosphorylated differentially at different stages of the parasite life-cycle, with more prevalence of phosphorylated LdSTI1 being observed at the promastigote stage compared to the amastigote (Morales et al., 2010). LdSTI1 was also shown to be in complex with LdHip, although whether the interaction is direct or indirect remains to be experimentally determined (Morales et al., 2010). Studies of TcSTI1 have shown that heat-induction has no effect on the levels of this protein throughout the parasite life-cycle, while starvation has shown pronounced expression of STI1 at the late epimastigote development stage (Schmidt et al., 2011). PfHop has been shown to localize with PfHsp70 and PfHsp90 in parasite trophozoite stage (Gitau et al., 2011). It was also shown through size exclusion chromatography and immunoprecipitation that PfHop associates with PfHsp70 and PfHsp90, therefore most likely mediating their interaction similar to mammalian cells (Gitau et al., 2011).

1.8. Motivation

The life cycle of *T. brucei* requires both a mammalian host and an insect vector where it undergoes extensive morphological changes and it needs to survive drastic changes in the environment in terms of nutrient status, temperature, pH and metabolic pathways. These transitions are thought to subject the parasite to protein misfolding and aggregation. Ensuring

the stability and functionality of cellular macromolecules is critical and vital for parasite survival. Molecular chaperones help buffer against cellular stress by promoting productive folding of cellular proteins and preventing their misfolding and aggregation (Riezman, 2004). The *T. brucei* genome is comprised of a significant number of heat shock proteins which have been suggested to play a role in protein homeostasis during several stages of development and host-vector transitions (Folguiera and Requena, 2007).

Little work has been conducted on the molecular chaperones of kinetoplastids and even less work has been carried out on *T. brucei*. The sequence of the nuclear 26-megabase *T. b. brucei* genome was obtained in 2005 (Berriman et al., 2005) while the *T. b. gambiense* genome was sequenced in 2010 (Jackson et al., 2010). The *T. brucei* nuclear genome contains 11-megabase-sized chromosomes and encodes for an estimated 9068 genes. Post genomic analyses have also shown that this trypanosome encodes for 12 Hsp90s, 10 of which appear as identical tandem repeats, 12 Hsp70s and 1 STI1, a significant amount of which localizes in the cytoplasm. *T. brucei* also encodes for a significant expansion of diverse J-proteins, consisting mainly of type III J-proteins (Folgueira and Requena, 2007; Louw et al., 2010). The Hsp90 and Hsp70 protein families have been proposed to have conserved domain and structural organization with other well-characterized organisms (Louw et al., 2010; Pizarro et al., 2013). The presence of STI1 on *T. brucei* genome has been confirmed, furthermore this protein was revealed to be essential for parasite survival and fitness in the bloodstream stage (Alsford et al., 2011). Although the presence of these proteins have been reported, the Hsp90-STI1-Hsp70 heat shock protein complex has not been confirmed in *T. brucei*.

Except for the phenotypic knockdown using RNAi (Alsford et al., 2011), very little characterization of *T. brucei* molecular chaperones have been carried out. TbHsp83 has been biochemically characterized and shown to have a 10-fold higher ATPase activity than human Hsp90 (Pizarro et al., 2013) and was also found to be structurally conserved as it was inhibited by well-known Hsp90

inhibitors, geldamycin (Meyer and Shapiro, 2013). A potential partnership between TbHsp70.c and Tbj2 was suggested by Burger et al (2014).

With regards to J-proteins, only Tbj1, a novel type III was heterologously expressed and found to exist as monomer in solution. Although able to assist in the suppression of aggregation, Tbj1 was found to be unable to stimulate the ATPase activity of TcHsp70 and *Medicago* s*avita* (MsHsp70). *T. brucei* J-proteins are yet to be categorized into type I-IV, it is possible that this significant expansion of kinetoplastid J-protein sequences could have novel functions in organelles unique to these organisms or in the development of the parasite throughout its life-cycle. In order to fully appreciate the Hsp90-STI1-Hsp70 heat shock protein complex, co-chaperones of *T. brucei* Hsp70 and Hsp90 would have to be studied in relation to well-characterized organisms. This study will focus on J-protein co-chaperones of Hsp70 and tetratricopeptide repeat (TPR)-containing co-chaperones of Hsp90. It is proposed that a study of the chaperone machinery in *T. brucei* could result in improved knowledge of the parasite infectivity, as well as the roles chaperones play in the development of the parasite. It will also result in improved knowledge of the fundamentals of chaperone interactions within *T. brucei*.

1.9. Research Hypothesis

Hsp70 prevents aggregation and promotes refolding of many different proteins, while Hsp90 serves a discrete set of clients. Both Hsp70 and Hsp90 are regulated by co-chaperones. STI1 is a co-chaperone that functions as an adaptor that connects Hsp90 to Hsp70 for the transfer of client proteins. The existence of Hsp70 and Hsp90 in *T. brucei* suggests the presence of the Hsp90-STI1-Hsp70 pathway. Thus, this study hypotheses that *T. brucei* co-chaperones function in an analogous manner to mammalian counterparts but have structural and mechanistic unique features that could be exploited.

1.10. Aims of the research

1.10.1. Broad aims:

1. *In silico* characterization of the Hsp70/J-proteins complex in kinetoplastids
2. Analysis of Hsp90 and TPR containing co-chaperones of Hsp70 and Hsp90 in *T. brucei*
3. Characterization of TbSTI1 and its interactions with cytosolic Hsp90 and Hsp70s

1.10.2. Specific aims:

1. Generation of an update of Hsp70 and Hsp90 proteins in kinetoplastids post recent genomic sequencing of additional species.
2. Generation of a framework of the multi-chaperone complex in *T. brucei* through *in silico* analysis of Hsp70, J-proteins, Hsp90, TbSTI1 and other TPR containing co-chaperones of Hsp70 and Hsp90 in *T. brucei*.
3. Confirm the presence of TbSTI1 and TbHsp83 in *T. brucei*
4. Establishment of heterologous expression systems for TbSTI1, TbHsp70s and TbHsp83.
5. Purification and biochemical characterization of TbSTI1, TbHsp70s and TbHsp83.
6. Establish the presence of the multi-chaperone complex in *T. brucei* through biochemical interaction studies.
7. Study the canonical features of TbSTI1, binding Hsp70 and Hsp90 in human lysate in relation to mSTI1.

CHAPTER TWO

In silico characterization of the Hsp70/J-proteins complex in kinetoplastids

2.0. Introduction

Kinetoplastid parasites are frequently exposed to stressful conditions, such as extreme temperature fluctuations, as they transition from cold-blooded arthropod vectors to warm-blooded mammalian hosts. Intracellular organisms, including parasites, frequently undergo genome reductions as a consequence of exposure to environmental changes often accompanied by shifts in temperature, starvation and the presence of reactive oxygen species (Vonlaufen et al. 2008; Urményi et al. 2012). This is in contradiction to the presence of an exceptionally large number of genes encoding molecular chaperone proteins, in particular the Hsp70 and J-protein families (Shonhai et al., 2011). Studies of molecular chaperones in kinetoplastids has focused on the intracellular parasites *T. cruzi* and *L. major* (Folgueira and Requena, 2007; Vonlaufen et al., 2008; Shonhai et al., 2011). However, no comprehensive assessment of the Hsp70/J-protein chaperone system has been conducted on either intracellular or extracellular parasites, particularly *T. brucei*.

To date, the only *In silico* analysis on the Hsp70 and J-proteins from kinetoplastids was performed by Tibbetts et al., (1998), who derived the nomenclature and subsequently carried out preliminary characterization of *T. cruzi* J-proteins. Since then, several kinetoplastid genomes have been sequenced to completion and some non-contiguously completed. Post genomic analysis of the molecular chaperone complement from (*T. brucei*, *T. cruzi* and *L. major* (TriTryps) was conducted by Folguiera and Requena (2007). Later, the Hsp70 protein families in the TriTryps were analyzed by Louw and colleagues in 2010 (Louw et al., 2010). Heat shock proteins of intracellular kinetoplastids and other intracellular protozoan parasites (apicomplexans) were reviewed by Shonhai et al., 2011, and, did not include the extracellular parasite *T. brucei*.

The Hsp70 superfamily of most eukaryotic organisms is generally composed of two classes: the HspA, typical Hsp70 proteins and the HspH members, nucleotide exchange factors (Polier et al.,

2008). Hsp70s play a critical role in maintaining cellular homeostasis, as a consequence they are conserved across all organisms (Boorstein et al., 1994). Furthermore, Hsp70 proteins are known to localize in the cytosol, mitochondria and endoplasmic reticulum. Typical Hsp70 proteins are subdivided into induced and constitutive paralogues. Homologues of Hsp70 proteins have been found in kinetoplastids, those in the TriTryps are the best characterized. The *T. cruzi* Hsp70 protein superfamily has been relatively well characterized (Olson et al., 1994). Homologues of Hsp70 proteins that exist in all subcellular compartments have been identified and described in the TriTryps (Folgueira and Requena, 2007; Louw et al., 2010; Shonhai et al., 2011). *T. brucei* possesses 12 Hsp70 proteins and *L. major* possesses 14 (Louw et al., 2010). When discovered initially, *T. cruzi* Hsp70 genes of different subfamilies were believed to occur as tandem repeats on the genome (Requena et al., 1989; Engman et al., 1989, Engman et al., 1992). However earlier studies showed that the CL Brener strain contained duplicate genes, thus leading to discrepancies with regards to the number of Hsp70s in this organism. It was later discovered that the CL Brener strain of *T. cruzi* tends to hybridize (El-Sayed et al., 2005a; de Freitas et al., 2006). Post genomic analysis has suggested that *T. cruzi* consists of 11 Hsp70 genes (El-Sayed et al. 2005), though other studies have noted that, including partial sequences, the total of Hsp70s is 28 (Louw et al., 2010; Shonhai et al., 2011).

Relative to *S. cerevisiae, H. sapiens* and *P. falciparum,* large numbers of J-proteins have been found in kinetoplastid organisms (Folguiera and Requena, 2007). Since the release of the TriTryps genomic sequences, numerous J-proteins were identified, 65 in *T. brucei*, 67 in *T. cruzi* and 66 in *L. Major*. More than one J-protein can interact with a given Hsp70 and confer functional and specificity roles, thus most organisms possess a larger number of J-proteins than Hsp70s (Cyr et al., 1992). The roles of these J-proteins in kinetoplastid organisms is yet to be established, only 6 J-proteins in *T. cruzi* have been biochemically characterized, mitochondrial TcDJ1 (Carreira et al., 1998), cytoplasmic Tcj1-Tcj4 (Tibbetts et al., 1998)and Tcj6 (Salmon et al., 2001). Only 2 type III J-proteins were biochemically characterized in *T. brucei*, Tbj1 and TbjSec63, which is an essential co- and post translational protein (Goldschmidt et al., 2008). Recent studies have also shown that

Tbj2, a type I J-protein is able to suppress the aggregation of the thermally denatured substrate malate dehydrogenase (MDH) and chemically denatured substrate rhodanese (Burger et al., 2014). Tbj2 stimulated the ATPase activity of TbHsp70.c (Burger et al., 2014).

Overall this chapter provides an updated overview of the Hsp70/J-protein machinery in kinetoplastid organisms with a special emphasis on *T. brucei*. All the datasets and annotations of the *T. brucei* genome are available on GeneDB (Hertz-Fowler et al., 2004) and can be found at: www.genedb.org. GeneDB is the current data housing of all pathogenic organisms genome-scale datasets, It was originally created for providing sequenced details of kinetoplastids and *Schizosaccaromyces pombe* as generated by the Pathogen Genomics group, Welcome Trust Sanger Institute (WTSI) (http://www.sanger.ac.uk) (Logan-Klumpler et al., 2012). Genomes within GeneDB are either automatically or manually curated, involving weekly amendments of annotations from both literature and user comments by the curators. Automated curations rely largely on user feedback. The genomes of several kinetoplastid organisms have been manually curated and are listed in Table 2.1. In collaboration with EuPathDB, GenBD sends annotated and curated trypanomastids genome-scale datasets to TriTrypDB (Aslett et al., 2010; http://www.tritrypdb.org).

Table 2. 1: A summary of the current genomes and strains of kinetoplastids, the sequencing method used to assemble the genome and a comparison of the genome sizes, number of chromosomes and genes encoded by each organism.

Parasite Name	Strain	Origin	Sequencing method	Genome size (Mb)	X	Genes	Reference
T.b brucei	TREU 9/427	Tsetse Glossina pallidipes in 1970 in Kiboko, Kenya	Pulsed field gel electrophoresis and sequenced by whole chromosome shortgun approach.	26	11	9068	Berriman et al., 2005
T. b.gambensie	Dal 972	A patient in 1986 in Daloa, Ivory Coast	Whole-genome shotgun strategy	32	11	9068	Jackson et al., 2010
T. vivax	Y486	Bovine in 1976 in Zaria, Nigeria		41	Not known	11 870	
T. congolense	IL3000	Bovine in 1966 in Transmara, Kenya.		34	11	13 496	Bienen et al., 1991
T. evansi	STIB 805	One was isolated from a dog (Colpo et al., 2005) and designated as TeD, and the stock designated as TeH was isolated from a horse in 2009		Not known	Not known	Not known	Duarte et al., 2014
T. cruzi	CL Brener Esmeraldo	Isolated from a Triatoma infestans 1963, collected in Rio Grande do Sul, Encruzilhada, South Brazil	Chromosome by chromosome assembly	67	41	12 000	
T. cruzi	CL Brener Non-esmeraldo		Chromosome by chromosome assembly	67	41	12 000	
T. cruzi	marinkellei	Isolated in São Felipe, Bahia state, Brazil in 1974 and has since then been stored under cryogenic conditions with occasional short periods of in vitro cultivation.	454 and Illumina sequencing	54			Franzén et al., 2012
T. cruzi	Sylvio X10		Whole genome shotgun				Ruvalcaba-Trejo and Strum, 2011
L. major	Friedlin		Chromosome by chromosome using shotgun	32.8	36	8412	Ivens et al., 2005
L. braziliensis	Viannia	Direct culture from a lesion on the right side of the thorax of a man who had been performing survey work in Serra dos Carajás, Brazilian Amazonia.	Whole genome shotgun	32	35	8357	Peacock et al., 2007
L. mexicana	U1103	Ear lesion of a 30-year-old male patient in Gautemala.	Genome assembled do novo using capillary sequencing	32	34	8250	Rogers et al., 2011
L. infatum	JPCM5	The spleen of a naturally infected dog residing in the area in 1998.	Whole genome shotgun	32	36	8241	Peacock et al., 2007
L. donovani	BPK282/0cl4	Visceral leishmaniasis patient in Nepal	Combination of 4S4 Life Science and IIlumina sequencing technologies	32	36	8252	Downing et al., 2011
L. tarentolae	Parrot-TarII		Next generation DNA sequencing technologies	30	32	8201	Raymond et al., 2011

2.1 Objectives

In silico analyses were undertaken to identify all Hsp70 and J-proteins from kinetoplastids, in particular *T. brucei*, for comparison to protozoan, unicellular and multicellular eukaryotes.

The specific objectives were:

1. To identify all Hsp70s and J-proteins in kinetoplastid parasites with special emphasis on *T. brucei* in light of recent genome annotations.
2. To provide chromosomal localization of *T. b. brucei* Hsp70 and J-proteins.
3. To provide protein domain maps, identify homologues and protein features of the Hsp70 superfamily in *T. brucei.*
4. To study the evolutionary relationship of Hsp70 proteins in relation of yeast, humans and *P. falciparum.*
5. To compare cytosolic *T. brucei* Hsp70s to cytosolic Hsp70s of human, yeast and *P. falciparum*, based on domain organization and multiple sequence alignment.
6. To compare the three cytosolic Hsp70s in *T. brucei* based on their predicted three-dimensional structures.
7. To determine the ability of cytosolic Hsp70s in *T. brucei* to interact with STI1, using the three dimensional structure of human Hop, TPR1 domain.
8. To classify J-proteins of the TriTryps into type I, II, III and IV based on *E. coli* DNAJ.
9. To compare the distribution of J-proteins in *T. b. brucei* and *T. b. gambiense* and potentially rule out sequencing errors, identify homologues and domain organization.
10. To characterize J-proteins based on protein features, predicted gene ontology functions and localization (therefore propose potential partner proteins).

2.2. Methods and softwares

2.2.1. Sequence searches and data retrieval

All predicted genomic and amino acid sequences used to study the gene structure, chromosome location and protein features of kinetoplastids Hsp70s and J-proteins were retrieved from TriTryDB version 8 (http://TriTrypdb.org/TriTrypdb/) (Aslett et al., 2010). Generally, sequence retrieval relied on gene IDs obtained from available literature (Folgueira and Requena, 2007; Louw et al., 2010; Shonhai et al., 2011) and scanning the database using keywords such as "Heat shock protein 40", "J-protein", "DnaJ". Hsp70s and J-proteins were also used as queries to acquire a wide range of eukaryotic homologues and orthologues which were retrieved from the National Centre of Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Tatusova et al., 2012) and TriTrypDB respectively. Homologues and orthologues were identified using the standard protein blast (blastp) option in the TriTrypDB and NCBI genome resources and sequences were filtered using the following parameters: >30% sequence identity; E-value ~ 0 and sequence coverage >75% while for TryTrips (*T. cruzi* and *L. major*) a cut-off of sequence identity at 50%, E-value at 0.0 and sequence coverage at 75% was applied. *T. brucei* protein properties were compared to those of two other Tritryps organisms and the protozoan parasite *P. falciparum*, the unicellular eukaryotic organisms *S. cerevisiae* and the mammalian organism *H. sapiens*.

2.2.1.1. *T. brucei* Hsp70 and J-proteins orthologues and homologues

The accession numbers of *T. brucei* Hsp70s were obtained from previous literature (Folgueira and Requena, 2007; Louw et al., 2010) and each Hsp70 homologue and orthologue was obtained through blast searches on TriTrypDB and NCBI. In addition to the sequence retrieval methodology described in section 2.1.1, J-proteins were also identified using the *E. coli* DNAJ, *S. cerevisiae* Sis1 and *H. sapiens* DNAJA1 J-domains as queries against the TriTrypDB databases and sequences

excluded using the paramaters listed in section 2.1.1. The sequences obtained were used for genome comparison of Hsp70s and J-proteins across various organisms.

2.2.2. Phylogenetic analysis

The chaperone properties of *T. brucei* Hsp70s were studied in relation to other kinetoplastids, *H. sapiens*, *S. cerevisiae* and *P. falciparum* by phylogenetic analysis. Molecular Evolutionary Genetics Analysis 6 (mega 6) (Tamura et al., 2013) software was used to generate the cladogram/phylogenetic tree. The primary amino acid sequences of proteins were used as inputs to mega 6, followed by multiple sequence alignment (MSA) using the inbuilt alignment program MUSCLE (Edgar et al., 2004). The MSA is subsequently used to construct a phylogenetic tree based on the best evolutionary model/substitution method which is calculated based on the Bayesian information criterion (BIC) score. Phylogenetic tree construction was carried out using maximum likelihood (Tamura et al., 2011) with the following parameters: bootstrap consensus tree surmised from 1000 replicates and any tree partitions replicated in less 75% of the bootstrap analyzes were collapsed. Based on the calculations, the unrooted tree was generated based on the Poisson substitution model (Zhang et al., 1998). The properties of the Hsp70s assessed to include number of genes on the genome, the amino acid lengths, the molecular masses and the isoelectric points.

2.2.3. Domain organization and subcellular localization prediction

The domain maps of Hsp70s and J-proteins were identified using a combination of 5 softwares. The amino acid of each protein was submitted to each of the programs and categorized into a specific domain based on predictions, related literature and manual confirmation of associated residues. The subcellular localization of proteins was predicted using 9 online servers. Using the protein amino acid sequence as an input, 3 servers (WOLF-PSORT, pTARGET and PredSL) calculate

the subcellular localization through a combination of feature to assess every cellular compartment while Phobius, NucPred, SignalP, MitoPROTII, NLS-mapper and SLP-local predicts organelle-specific localization.

Table 2. 2: Properties of Softwares used for domain identification and subcellular localization predictions.

Software	Website	Prediction	Referrence
Prosite	http://prosite.expasy.org/	Domains	Sigrist, et al., 2010
SMART 7	http://smart.embl-heidelberg.de/	Domains	Letunic, et al., 2012
MEME	http://meme.nbcr.net/	Motifs	Bailey, et al., 2006
TPR pred	http://toolkit.tuebingen.mpg.de/tprpred	TPR motifs	Karpenahalli et al., 2007
Phobius	http://phobius.sbc.su.se/	Cytoplasm	Käll, et al., 2004
NucPred	http://www.sbc.su.se/~maccallr/nucpred/	Nucleus	Brameier et al., 2007
WoLF PSORT	http://wolfpsort.seq.cbrc.jp/	Subcellular compartments	Horton et al., 2007
pTARGET	http://golgi.unmc.edu/ptarget/	Subcellular compartments	Emanuelsson et al., 2000
PredSL	http://hannibal.biol.uoa.gr/PredSL/	Subcellular compartments	Petsalaki et al., 2006
SLP-Local	http://sunflower.kuicr.kyoto-u.ac.jp/~smatsuda/slplocal.html	Chloroplast, mitochondria, secretory pathways and nucleus	Matsuda et al., 2005
SignalP	http://www.cbs.dtu.dk/services/SignalP/	Endoplasmic reticulum	Dyrløv Bendtsen et al., 2004
MitoPROTII	http://ihg.gsf.de/ihg/mitoprot.html	Mitochondria	Claros and Vincens, 1996
NLS mapper	http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi	Nucleus	Kosugi et al., 2009

2.2.4. Multiple sequence alignments

At least four programs were used to generate the multiple sequence alignments of Hsp70 and J-proteins of different organisms namely, T-COFFEE (Notredame et al., 2000), Clustal-omega (Thompson et al., 1994), MUSCLE and Promals3D (Pei et al., 2008). Each of the programs uses a specific algorithm to elucidate the different interrelations between sequences therefore using more than one program eliminates or reduces the chances of receiving biased results. Alignments were carried out for comparison of sequences from Trypanosomes, kinetoplastids and other eukaryotic organisms.

2.2.5. Homology modelling

Homology modelling was carried out following 6 mains steps: 1) The identification of a suitable template, 2) Pairwise alignment of target protein sequence and template structural sequence, 3) Fitting the target protein sequence to the backbone of the template structure, 4) loop and side chain modelling, 5) Model refinement and 6) Model evaluation (di Luccio and Koehl, 2011). Therefore, following the above criteria, models of TbHsp70, TbHsp70.4 and TbHsp70.c were generated using the *E. coli* Hsp70 structure 2KHO and the C-terminal domain of Hsp70 (3L0F) as templates. It has been shown that using more than one template for Hsp70 structural prediction can be detrimental to the accuracy of model predictions as the crystal structures of Hsp70s do not appear in the same orientation in the protein data bank (Hartherley et al., 2013). All three proteins shared >35% sequence identity to 2KHO, thus high-throughput models were generated using modeler 9v12 (Eswar et al., 2008; Shen and Sali, 2006), 100 models were built for each protein. The models were assessed by ProSA (Wiederstein et al., 2007) and MetaMQAPII (Pawlowski et al., 2008) and by calculating their DOPE Z-scores.

2.3. Results and Discussion

2.3.1. Genome comparison of the number of Hsp70 proteins

The Hsp70 superfamily is generally composed of two classes, the HspA and the HspH members and these subdivisions are shown in Figure 2.1. Comparison of the total numbers of Hsp70s (Figure 2.1) show that overall in the TriTryps, *Leishmania* spp.have the highest number of Hsp70s *(n = 14)* relative to *T. brucei (n = 12)* and *T. cruzi* CL Brener Esmeraldo-like *(n = 11)* and marinkellei strain B7 *(n = 12)*. The number of Hsp70s in kinetoplastids is similar to *S. cerevisiae (n = 14)* and smaller than *H. sapiens (n = 17)*. The number of Hsp70s in *T. brucei* is similar to *T. evansi* (*n=12*), *T. congolense* (*n=12*) and *T. vivax* (*n=11*) (Figure 2.1).

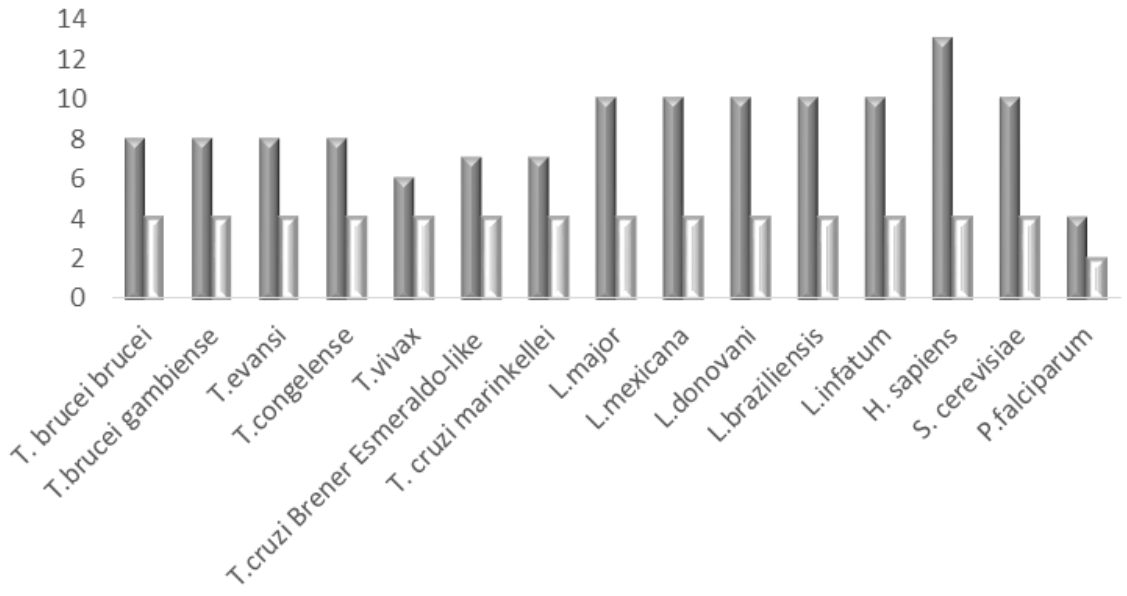


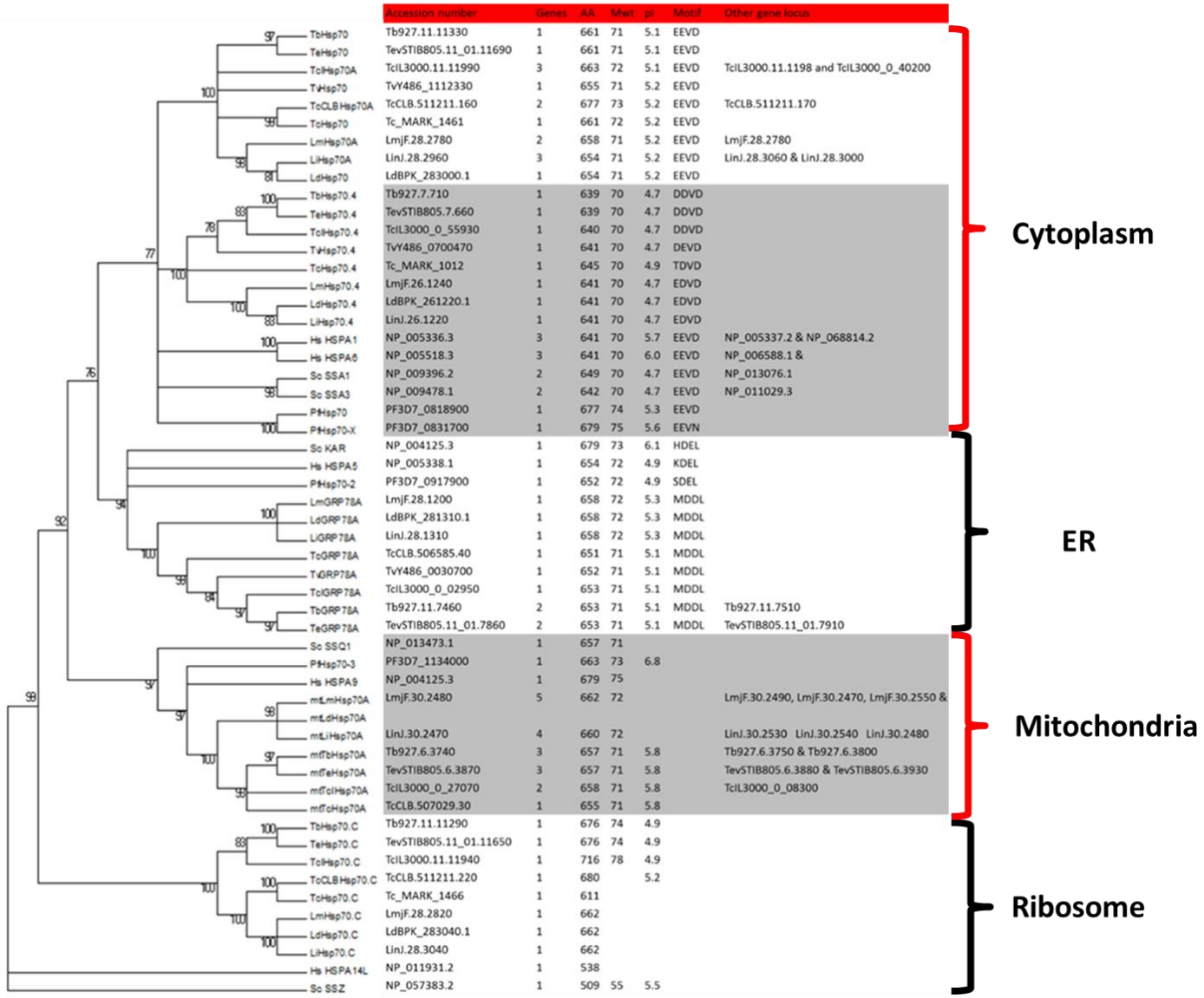
Figure 2. 1: The number of Hsp70s in different genomes. Data for all kinetoplastids obtained from TriTrypDB and for *H. sapiens* (Kampinga et al. 2010); *P. falciparum* (Njunge et al. 2013); *S. cerevisiae* (Walsh et al. 2004; Cherry et al. 1998). The Y-axis represents the numbers of Hsp70s and the X-axis represents the name of each organism. The two divisions of the Hsp70 superfamily are represented as HspA (grey) and HspH (white) respectively.

All the organisms analyzed in this study appear to have the same number (*n=4*) of HspH proteins, with the exception of *P. falciparum* which has only 2 members (Figure 2.1). This potentially suggests a conservation of the roles of the NEFs. Variability was observed in terms of the number of HspA proteins, Hsp70-like proteins across different organisms, where kinetoplastids and yeast consist of between 8-10 homologues while humans have a larger number (*n=13*) (Figure 2.1), possibly due to the abundance of more specialized Hsp70s in multicellular organisms. The results obtained for the number of TriTryps Hsp70s are consistent with previous studies (Folguiera and Requena, 2007; Louw et al., 2010; Shonhai et al., 2011). Earlier studies suggested the presence of 10-tandemly repeated copies of Hsp70s on the *T. cruzi* genome (Requena et al., 1988, Requena et al., 1989, Engman et al., 1989, Engman et al., 1992). However, further post sequencing analysis of the draft genome of *T. cruzi* CL Brener (El-Sayed et al., 2005) showed the presence of 11 Hsp70 protein encoding genes (Louw et al., 2010; Shonhai et al., 2011). This study compared the numbers of Hsp70s in two *T. cruzi* strains CL Brener Esmeraldo-like and marinkellei strain B7 which were revealed to be 11 and 12 respectively (Figure 2.1). An orthologue of the cytosolic Hsp70.4 was found to be lacking in the genome of CL Brener Esmeraldo-like strain. Overall, there is a conservation with regards to the number of Hsp70s across kinetoplastids and yeast and this conservation might infer function.

2.3.2. Phylogenetic analysis, domain organization and homologue identification of *T. brucei* Hsp70 superfamily proteins

The phylogenetic analysis was limited to the HspA members of the Hsp70 superfamily, as the evolutionary relationship of kinetoplastid Hsp70 proteins in relation to human, yeast and *P. falciparum* was the main focus of the analysis (Figure 2.2). As expected the *Leishmania* and *Trypanosoma* proteins group together in the 5 distinct clusters. It was observed that *T. evansi* and *T. brucei* cluster close together in each clade. Overall, the clustering of Hsp70s was based on two main factors: protein conservation and subcellular localization. Interestingly, TbHsp70 and TbHsp70.4 clustered with cytosolic Hsp70s in a similar manner to the well characterized *P.*

falciparum PfHsp70-1 (Joshi et al., 1992; Biswas and Sharma, 1994; Pesce et al., 2008)and the exported, less well characterized PfHsp70-x (Kulzer et al., 2012) (Figure 2.2). The phylogeny and separate homologue identification conducted in this study are complementary to each other with regards to higher eukaryote homologues for *T. brucei* Hsp70s (Figure 2.2 and Figure 2.3). Furthermore, the cladogram generated reveals that TbHsp70.c is homologous to yeast SSB and human HSPA1L as these proteins share a monophyletic clade (Figure 2.2). Also, although predicted to be cytosolic, TbHsp70.c and its orthologues clade is separated from its cytosolic counterparts by ER and mitochondrial Hsp70s, probably a reflection of evolutionary distance.



Accession number	Genes	AA	Mwt	pI	Motif	Other gene locus
TbHsp70	Tb927.11.11330	1	661	71	5.1	EEVD	
TeHsp70	TevSTIB805.11_01.11690	1	661	71	5.1	EEVD	
TcIHsp70A	TcIL3000.11.11990	3	663	72	5.1	EEVD	TcIL3000.11.1198 and TcIL3000_0_40200
TvHsp70	TvY486_1112330	1	655	71	5.2	EEVD	
TcCLBHsp70A	TcCLB.511211.160	2	677	73	5.2	EEVD	TcCLB.511211.170
TcHsp70	Tc_MARK_1461	1	661	72	5.2	EEVD	
LmHsp70A	LmjF.28.2780	2	658	71	5.2	EEVD	LmjF.28.2780
LiHsp70A	LinJ.28.2960	3	654	71	5.2	EEVD	LinJ.28.3060 & LinJ.28.3000
LdHsp70	LdBPK_283000.1	1	654	71	5.2	EEVD	
TbHsp70.4	Tb927.7.710	1	639	70	4.7	DDVD	
TeHsp70.4	TevSTIB805.7.660	1	639	70	4.7	DDVD	
TcIHsp70.4	TcIL3000_0_55930	1	640	70	4.7	DDVD	
TvHsp70.4	TvY486_0700470	1	641	70	4.7	DEVD	
TcHsp70.4	Tc_MARK_1012	1	645	70	4.9	TDVD	
LmHsp70.4	LmjF.26.1240	1	641	70	4.7	EDVD	
LdHsp70.4	LdBPK_261220.1	1	641	70	4.7	EDVD	
LiHsp70.4	LinJ.26.1220	1	641	70	4.7	EDVD	
Hs HSPA1	NP_005336.3	3	641	70	5.7	EEVD	NP_005337.2 & NP_068814.2
Hs HSPA6	NP_005518.3	3	641	70	6.0	EEVD	NP_006588.1 &
Sc SSA1	NP_009396.2	2	649	70	4.7	EEVD	NP_013076.1
Sc SSA3	NP_009478.1	2	642	70	4.7	EEVD	NP_011029.3
PfHsp70	PF3D7_0818900	1	677	74	5.3	EEVD	
PfHsp70-X	PF3D7_0831700	1	679	75	5.6	EEVN	
Sc KAR	NP_004125.3	1	679	73	6.1	HDEL	
Hs HSPA5	NP_005338.1	1	654	72	4.9	KDEL	
PfHsp70-2	PF3D7_0917900	1	652	72	4.9	SDEL	
LmGRP78A	LmjF.28.1200	1	658	72	5.3	MDDL	
LdGRP78A	LdBPK_281310.1	1	658	72	5.3	MDDL	
LiGRP78A	LinJ.28.1310	1	658	72	5.3	MDDL	
TcGRP78A	TcCLB.506585.40	1	651	71	5.1	MDDL	
TvGRP78A	TvY486_0030700	1	652	71	5.1	MDDL	
TcIGRP78A	TcIL3000_0_02950	1	653	71	5.1	MDDL	
TbGRP78A	Tb927.11.7460	2	653	71	5.1	MDDL	Tb927.11.7510
TeGRP78A	TevSTIB805.11_01.7860	2	653	71	5.1	MDDL	TevSTIB805.11_01.7910
Sc SSQ1	NP_013473.1	1	657	71			
PfHsp70-3	PF3D7_1134000	1	663	73	6.8		
Hs HSPA9	NP_004125.3	1	679	75			
mtLmHsp70A	LmjF.30.2480	5	662	72			LmjF.30.2490, LmjF.30.2470, LmjF.30.2550 &
mtLdHsp70A							
mtLiHsp70A	LinJ.30.2470	4	660	72			LinJ.30.2530 LinJ.30.2540 LinJ.30.2480
mtTbHsp70A	Tb927.6.3740	3	657	71	5.8		Tb927.6.3750 & Tb927.6.3800
mtTeHsp70A	TevSTIB805.6.3870	3	657	71	5.8		TevSTIB805.6.3880 & TevSTIB805.6.3930
mtTcIHsp70A	TcIL3000_0_27070	2	658	71	5.8		TcIL3000_0_08300
mtTcHsp70A	TcCLB.507029.30	1	655	71	5.8		
TbHsp70.C	Tb927.11.11290	1	676	74	4.9		
TeHsp70.C	TevSTIB805.11_01.11650	1	676	74	4.9		
TcIHsp70.C	TcIL3000.11.11940	1	716	78	4.9		
TcCLBHsp70.C	TcCLB.511211.220	1	680		5.2		
TcHsp70.C	Tc_MARK_1466	1	611				
LmHsp70.C	LmjF.28.2820	1	662				
LdHsp70.C	LdBPK_283040.1	1	662				
LiHsp70.C	LinJ.28.3040	1	662				
Hs HSPA14L	NP_011931.2	1	538				
Sc SSZ	NP_057383.2	1	509	55	5.5		

Figure 2. 2: Phylogenetic analysis of several kinetoplastid Hsp70s in relation to human, yeast and *P. falciparum*. Analysis based on number of gene copies, amino acid length (AA), molecular mass (Mwt), isoelectric point (pI) and the presence of a C-terminal motif known for binding to co-chaperones or retaining ER Hsp70s in the cellular organelle. The evolutionary relationship was generated using the Maximum likelihood statistical method. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Branches corresponding to partitions replicated in less than 75% of the bootstrap analyzes were collapsed. The evolutionary distances were computed using the Poisson amino acid substitutions model (Please refer to the pdf for better visualization of the accession numbers).

The 12 Hsp70s in *T. brucei* are comprised of classical Hsp70 domains: N-terminal nucleotide binding domain (NBD), also known as the ATPase domain, the highly variable linker region and the substrate binding domain (SBD) (Table 2.3). It is generally accepted that the HspH members of the Hsp70 family are much larger with a molecular mass of approximately ~100kDa, and thus the amino acid sequence length of TbHsp70.A, TbHsp70.B, TbHsp110 and TbGRP170 is at least 160 amino acids larger than the typical HspA subfamily (Table 2.3).

Table 2. 3: Domain organization of the Hsp70s from *T. brucei*. Each protein sequence is represented by an open bar indicated on the left hand side along with predicted homologues from H. sapien (Hs) and S. cerevisiae (Sc). The amino acid sequence length of each protein is indicated on the right hand side of each protein bar along with proteomics on the expression stage of the parasite, predicted localization and the essentiality of the protein following RNAi knockdown. All three domains and linker region of Hsp70s are indicated. Features were identified using Prosite (Sigrist, et al., 2010), SMART 7 (Simple Modular Architecture Research Tool) (Letunic, et al., 2012), and Phobius (Käll, et al., 2004).

Pf	Sc	Hs	*T. b. brucei*				
				Domain map	Stage	RNAi	Localization
PfHsp70-1	SSA1 SSA2	HSPA1 HSPA2	TbHsp70	NBD SBD CTD EEVD 661	BSF & Pro	ALL	Cytoplasm
PfHsp70-x	SSA3 SSA4	HSPA6 HSPA8	TbHsp70.4	NBD SBD CTD DDVD 639	BSF & Pro	NS	Cytoplasm
	SSB1? SSB2?	HSPA1L?	TbHsp70.c	NBD SBD CTD 676	BSF & Pro	DIF	Cytoplasm
PfHsp70-2	KAR2	HSPA5	TbGrp78A TbGrp78B	NBD SBD CTD MDDL 653	BSF & Pro	NS	Endoplasmic reticulum
PfHsp70-3	SSQ EMC10	HSPA9	mtTbHsp70A mtTbHsp70B mtTbHsp70C	NBD SBD CTD 657	BSF & Pro	ALL	Mitochondria
PfHsp70-z	SSE2	HSPA4	TbGrp170	NBD SBD CTD KGDL 720	BSF mem	BSF	ER
PfHsp70-y	SSE2	HSPH	TbHsp110	NBD SBD CTD 811	BSF & Pro	ALL	Cytoplasm
			TbHsp70.A	NBD SBD CTD 833	BSF mem	NS	
			TbHsp70.B	NBD SBD CTD 837	BSF mem	DIF	Cytoplasm

It has been noted that kinetoplastids display stage-specific expression of heat shock proteins (Table 2.3). Based on the mass spectrometry expression data retrieved from TriTrypDB, *T. brucei* Hsp70s are expressed in both the PRO and BSF of the parasite life-cycle, although TbGrp170, TbHsp70.A and TbHsp70.B are expressed mainly in the BSF stage (Table 2.3). Also, Alsford et al. (2011) conducted a high-throughput phenotypic analysis of *T. brucei* genes post RNAi knockdown and their results have shown that most of the Hsp70s are essential for parasite fitness and survival. Except for TbHsp70.4, TbGrp78A, TbGrp78B and TbHsp70.A, all the other Hsp70s in *T. brucei* are essential for parasite survival, mainly in the BSF stage (Table 2.3).

Previous findings reported the presence of an unusual Hsp70 in *T. brucei* (TbHsp70) which possessed a unique C-terminal RRHI motif instead of the canonical EEVD known to interact with the stress inducible protein (STI1) and J-proteins (Louw et al., 2010). Recent updates on the *T. brucei* genome have shown that TbHsp70 does indeed have a similar domain organization to canonical Hsp70s and the C-terminal EEVD to that of its kinetoplastid orthologues and eukaryotic homologues (Figure 2.2 and Table 2.3). Expressed at all stages of the parasite developmental biology, TbHsp70 is also essential for parasite survival and fitness (Jones et al., 2006; Vertommen et al., 2008; Alsford et al., 2011). The mRNA levels of TbHsp70 increased following parasite culturing in conditions mimicking those of the mammalian host (Van Der Ploeg et al., 1985), suggesting that TbHsp70 protein levels may be enhanced when exposed to stress in a similar manner to its *T. cruzi* orthologue (TcHsp70) (Olson et al., 1994; Schmidt et al., 2011). TbHsp70 has no clear nuclear localization signal (NLS), but shares significant sequence homology to TcHsp70 (89% sequence identity, data not shown) which has been shown to change its subcellular localization from the cytoplasm to nucleus when exposed to heat shock (Olson et al., 1994; Schmidt et al., 2011). The *L. tarentolae* orthologue of TbHsp70 has been linked to resistance towards the most commonly used chemotherapeutic drug pentavalent (Brochu et al., 2004), in light of the favourable sequence identity and conservation, it is possible for TbHsp70 to be explored experimentally as a drug target due to its essentiality in parasite pathogenesis. This

study proposes that TbHsp70 is a homologue of human HspA1 and HspA2, yeast SSA1 and SSA2 and PfHsp70-1 in *P. falciparum* (Figure 2.2 and Table 2.3).

TbHsp70.4 on the other hand, possesses a variant DDVD instead of the canonical EEVD motif at the C-terminal end, this modification is also present in PfHsp70-x, which possess a C-terminal EEVN motif, however TbHsp70.4 does not contain the leader sequence that is present at the N-terminus of PfHsp70-x (Kulzer et al., 2012, Hetherley et al., 2013). A variation of the C-terminal motif has also been observed in the kinetoplastid orthologues of TbHsp70.4; TDVD in *T. cruzi*, DEVD in *Leishmania spp* (Table 2.3). It is generally accepted that the TriTryps Hsp70.4 is cytosolic and this is consistent with our prediction analysis (Table 2.3). Following RNAi knockdown, TbHsp70.4 has not led to any loss of fitness of parasite survival (Table 2.3), it is likely to be constitutively expressed, also its mRNA levels decreased upon exposure to heat shock (Kramer et al., 2008). TbHsp70.4 is also highly similar to LmHsp70.4 and TcHsp70.4 which are constitutively expressed. The biological significance of some members of parasitic kinetoplastids expressing stage-specific Hsp70s is yet to be determined but it has been shown that the *T. cruzi* cytosolic Hsp70.4 is highly enriched in amastigotes but undetectable in trypomastigotes, while Hsp70.a was found exclusively in trypomastigotes (Atwood et al. 2005). TbHsp70.c is predicted to localize in the cytosol but lacks the EEVD motif necessary for association with STI1 and J-proteins (Figure 2.2 and Table 2.3). The identification of homologues revealed that ribosome-binding Hsp70s appear to be homologous to this protein, in particular human HspA1L and yeast SSB. One of the most fascinating features of TbHsp70.c is its highly unique SBD (highly charged), this protein has also been shown to be expressed at all stages of the parasite life-cycle and essential for parasite survival and fitness at the differentiation stage (Table 2.3). Biochemical characterization also suggested a potential partnership with a type I J-protein, Tbj2 (Burger et al., 2014). It remains, however, to be experimentally elucidated whether Hsp70.c requires ATP or ATP hydrolysis for functioning as was determined for its putative eukaryotic homologues (Hundley et al., 2002).

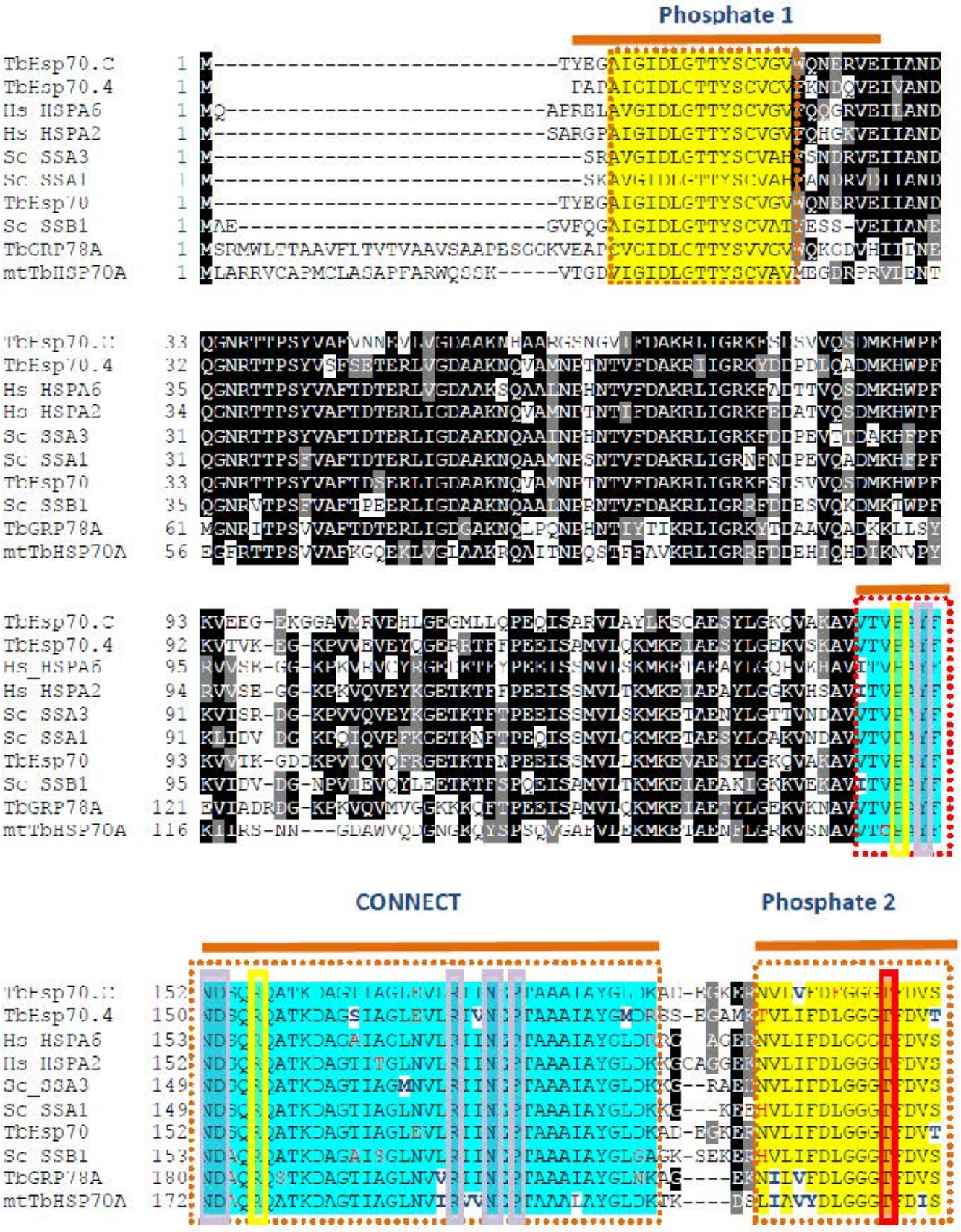
The ER and mitochondrial Hsp70s of *T. brucei* have some very interesting features, for instance they appear in tandem repeats of 2 and 3 respectively. Some studies have explored the ATPase activities of Hsp70 members in kinetoplastids and indeed observed potent activity, *T. cruzi* cytosolic and mitochondrial Hsp70s are 100 times more active than human Hsp70 (Olson et al., 1994) and stimulated by peptides and J-proteins (Edkins et al., 2004). *T. brucei* possesses 3 mitochondrial Hsp70s, while *L. major* possesses 5 proteins arranged in a tandem array (Floguiera and Requena, 2007). *T. cruzi* has a single copy of mitochondrial Hsp70 that possesses autophosphorylation activity (Olson et al., 1994) and remains associated with the kinetoplast even when exposed to stress (Martin et al., 1993). It is possible that mtTcHsp70 is involved in mtDNA replication (Engman et al., 1989). The *L. major* mtHsp70 gene has been isolated and shown to be constitutively expressed at all stages of the parasite life cycle (Searle et al., 1993). Although all three are 100% identical, only mtHsp70A of *T. brucei* has been shown to be essential for survival at all stages of the parasite life-cycle (Alsford et al., 2011). Proteomics data has also indicated that mtTbHsp70A and mtTbHsp70C are expressed at all stages of the parasite life-cycle, while mtTbHsp70B is only expressed in the BSF stage (Simpson et al., 2004; Vertommen et al., 2008). There is possibly a lack of redundancy of function in these particular Hsp70s.

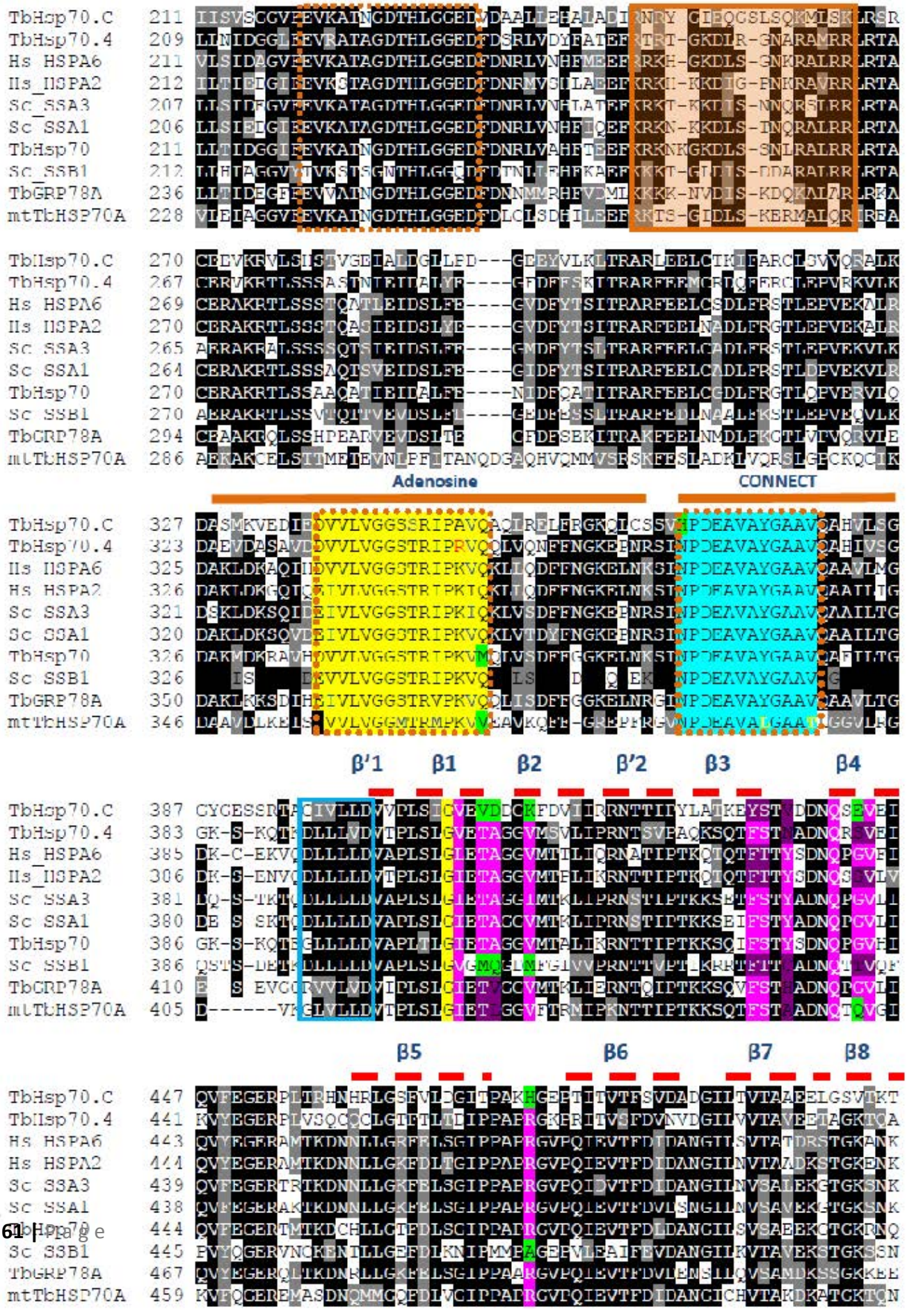
The *T. brucei* genome encodes two TbGrp78 (also known as BiP), both possessing an ER retention sequence, MDDL which differs from the KDEL sequences for mammals (Pidoux and Armstrong, 1992; Bangs et al., 1996). TbGrp78 was the first member of the Hsp70 superfamily to be characterized in TriTryps and has similar features to eukaryotic BiP (Bangs et al., 1993). It localizes in the subpellicular membrane of the parasite around the flagellum, therefore it is likely to be involved in the maintenance of parasite motility (Bangs et al., 1993). Although the two TbGrp78-encoding proteins are expressed at all stages of the parasite life-cycle, pronounced levels of the protein have been detected in the BSF stage which is several fold higher than the levels obtained for the PRO form of the trypanosome (Bangs et al., 1995).

2.3.3. Sequence and structural features of cytosolic Hsp70s in *T. brucei*

For the purposes of this investigation, nucleotide exchange factors or HspH members of the Hsp70 superfamily were excluded from the multiple sequence alignment of cytosolic Hsp70s from *T. brucei* as well as human, yeast and *P. falciparum* homologues (Figure 2.3). The alignment of all *T. brucei* Hsp70s is provided in Appendix F, Figure F1.

As expected the ATPase domains of the three cytosolic Hsp70s in *T. brucei* as well as the other eukaryotic Hsp70s were highly conserved. Early studies used *E. coli* DnaK and DnaJ as models to identify regions that are considered to be essential for Hsp70 function (Mount, 1985; Gässler et al., 1998). Some of the key residues include the proline allosteric switch (Vogel et al., 2006a), which regulates both substrate release and ATP hydrolysis (Han and Christen, 2003). These residues are indeed present in all cytosolic Hsp70s included in the study (Figure 2.3), as well as residues specifically associated with binding to ATP and phosphate (Figure 2.3). Also, residues that are implicated in interdomain communication (Gässler et al., 1998; Jiang et al. 2005) and J-protein interactions are conserved in TbHsp70, TbHsp70.4 and TbHsp70.c. Notably, the linker region of TbHsp70 and TbHsp70.4 is almost identical to its eukaryotic homologues while TbHsp70.c is slightly dissimilar and this is consistent with SSZ1 and HspA14L with are more diverse (Figure 2.3).


Phosphate 1
CONNECT
Phosphate 2


Adenosine
CONNECT
β'1
β1
β2
β'2
β3
β4
β5
β6
β7
β8

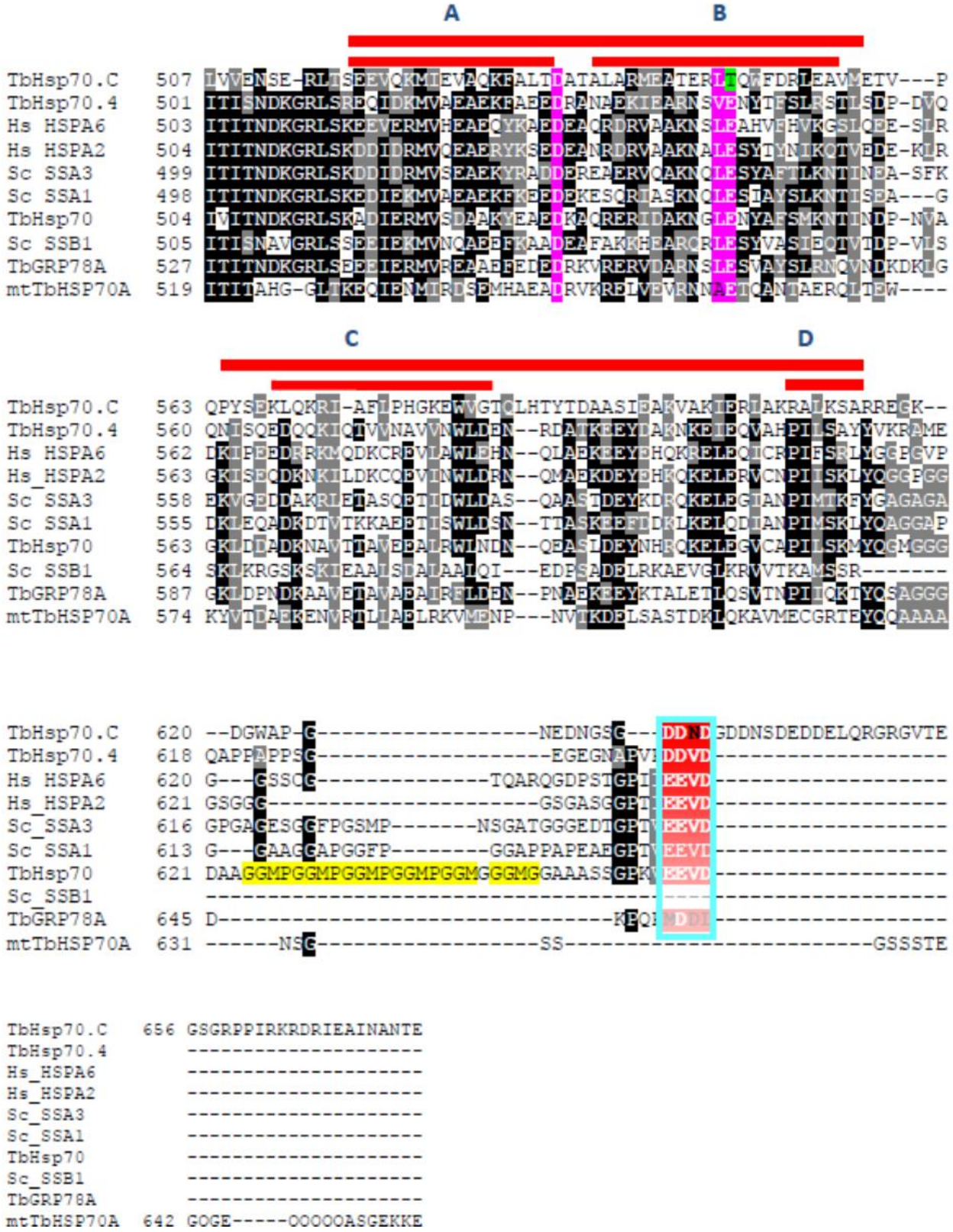


Figure 2.3*:* Amino acid sequence alignment of *T. brucei* cytoplasmic Hsp70s, sequences were compared to *P. falciparum*, *S. cerevisiae* and *H. sapiens*. The multiple sequence alignment was conducted against

full-length Hsp70 sequences which comprises of the ATPase domain, linker region (red box), substrate binding domain (red lines underneath the sequences) and the C-terminal domain (red dashes underneath the sequences). Some of the key residues in each of the domains are highlighted and annotated as follows: The ATPase domain residues associated with heat shock proteins (yellow highlights), Proline (the allosteric switch residue) and Arginine which are involved in interdomain function (Vogel et al., 2006a) are also shown in yellow highlights. Residues involved in J-proteins binding are indicated in a purple box. In *E. coli* DNAK, residues involved in DNAJ binding are Y145, N147, D148, N170 and T173 (Gässler et al., 1998; Suh et al., 1998) and T199 (highlighted in red), a DNAK phosphorylation site (McCarty and Walker, 1991). The three motifs interacting with ATP at the adenosine, phosphate 1 and phosphate 2 are indicated in yellow background which is boxed in red dots and the segments linking nucleotide binding motifs connect 1 and connect 2 are indicated cyan highlights which are boxed in red dashes. The linker regions which connects the two domains in shown in a red box. The least conserved substrate binding domain and C-terminal domain are shown by red lines and red dotted lines beneath the sequences respectively. Residues for *T. brucei* Hsp70 residues which are identical to the *E. coli* DNAK for substrate binding are indicated in magenta highlights, those which differ from the DNAK residues are shown in dark red and green for the substitution which involves a charged residue. With regards to the C-terminal domain the GGMP regions and the EEVD motif at the C-terminal end of cytosolic Hsp70 are shown in green highlights

Structural prediction of *T. brucei* cytosolic Hsp70s shows that these proteins fold into a classic ATPase domain, followed by β-sheet enriched substrate binding domain and the α-helical C-terminus (lid) (Figure 2.4). The binding sites for J-proteins and NEFs, based on the *E. coli* system, in the ATPase domain are also conserved (Figure 2.4).

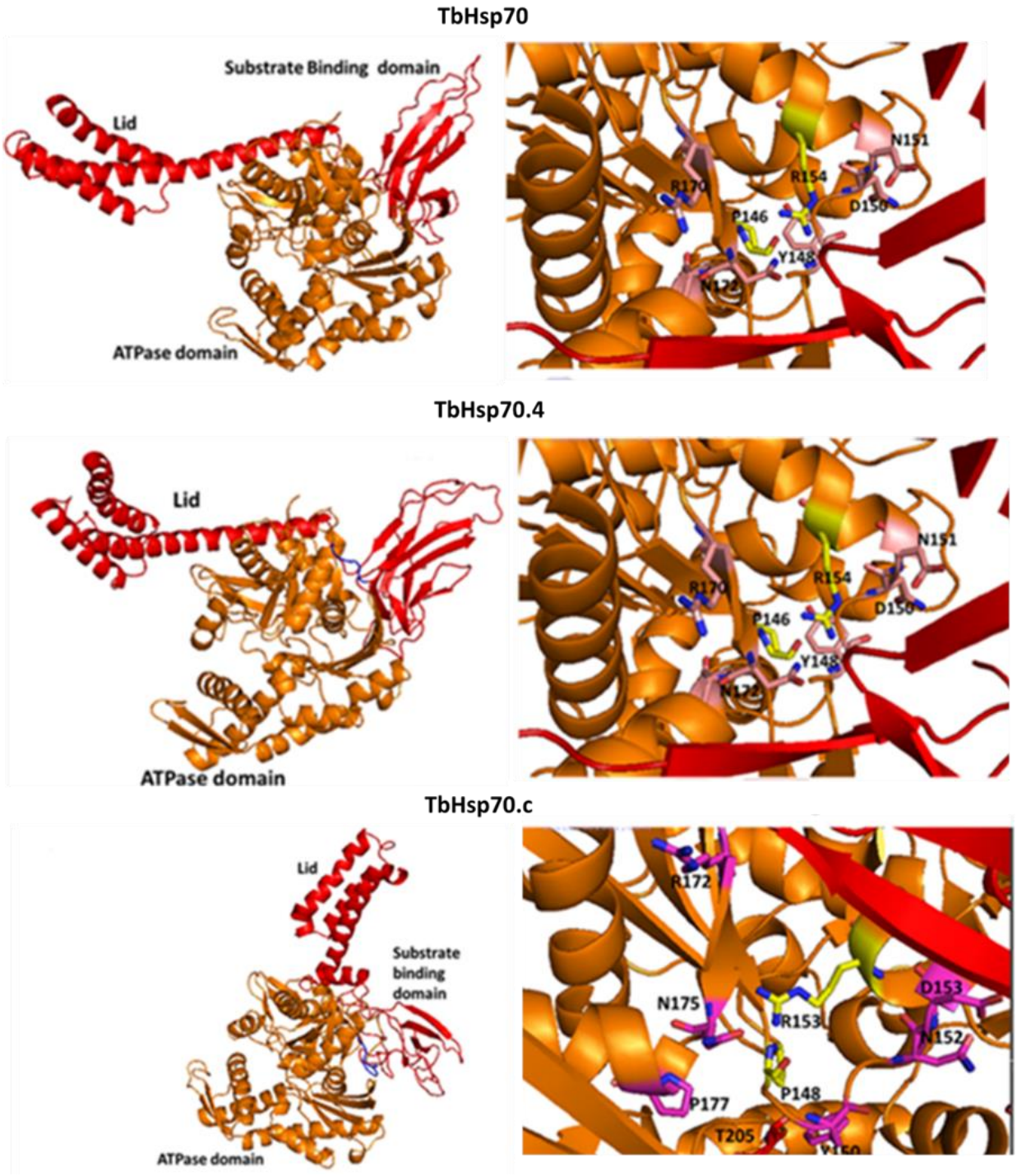


Figure 2. 4: Structural predictions of TbHsp70, TbHsp70.4 and TbHsp70.c which were generated by modeller version 12 (Eswar et al., 2008; Shen and Sali, 2006) using the *E. coli* DnaK (2KHO) as a template for modelling. The left panel shows full-length predictions of cytosolic Hsp70s from T. brucei which comprises of the ATPase domain (gold) followed by the linker (blue) and then the substrate binding domain and lid (red). The right panel indicates a stick representation of the residues known to be important for Hsp70 function and J-protein interaction in the ATPase domain of the Hsp70s. Images generated using PyMol (DeLano, 2002).

The alignment and structures of cytosolic *T. brucei* Hsp70s with other eukaryotes shows some degree of variation in the substrate binding domain (Figure 2.3 and Figure 2.5), especially the ribosome-associated TbHsp70.c. The three-dimensional models of cytosolic Hsp70s in *T. brucei* showing the arch and hydrophobic pocket residues, together with all the other residues that have contact with substrate is shown in Figure 2.5. The substrate binding domain is typically divided into two distinct sub-domains: the β-domain, which contains the substrate binding pocket and the α-helical lid, positioned closest to the C-terminal end of the protein (Stru et al., 2003) (Figure 2.5). The substrate binding domain of cytosolic Hsp70s is more conserved in the β-domain than the α-helical lid, which is highly variable (Figure 2.3 and Figure 2.5). Residues corresponding to those forming physical contact with the substrate during peptide binding (Zhu et al., 1996) are conserved in most of the proteins aligned, variation is only observed in the ribosome-associated TbHsp70.c (Figure 2.3).

With regards to the residues in the hydrophobic arch and the hydrophobic pocket in the substrate binding domain, these are also relatively conserved in all cytosolic Hsp70s (Figure 2.5). TbHsp70 and TbHsp70.4 have a typical Hsp70 arch architecture. Interestingly, the homologues of TbHsp70.c in yeast and human are missing an entire α-helix in the lid of the substrate binding domain (Figure 2.3). It has been observed that Hsp70s prefer acidic and hydrophobic peptide substrates (Rüdiger et al., 1997). The results obtained in this study indicate that TbHsp70 and TbHsp70.4 (Figure 2.5 A and B) possess substrate binding cavities which are rich in hydrophobic residues as is true for typical Hsp70s (Zhu et al., 1996; Mayer et al., 2000), while TbHsp70.c (Figure 2.5 C) has a highly charged substrate binding cavity. Based on these observations, TbHsp70.c will probably favour acidic peptide substrates, a fitting proposal based on the prediction that it associates with the ribosome.

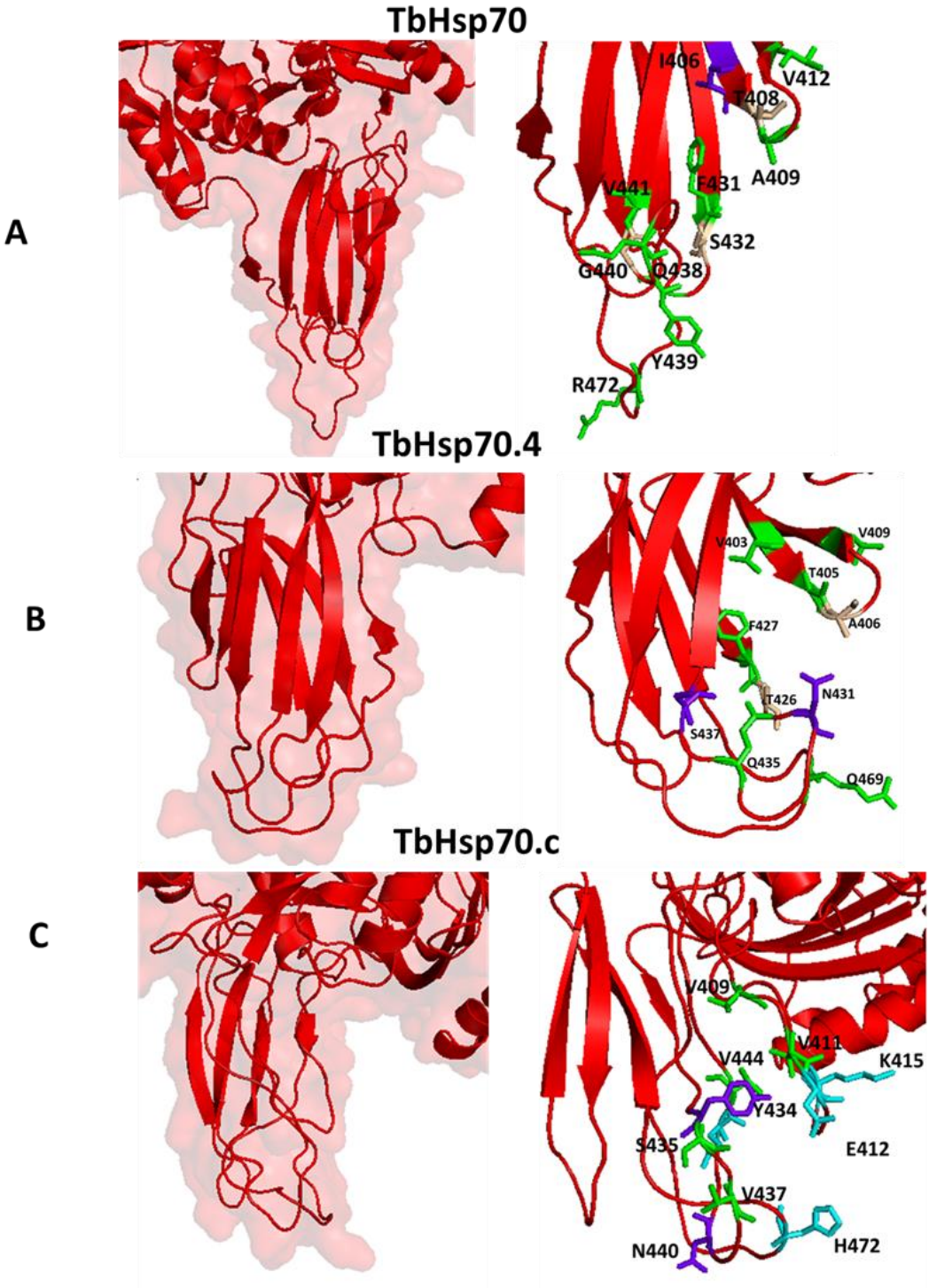


Figure 2. 5: Three-dimensional models of the β-sheet enriched SBD of *T. brucei* cytosolic Hsp70s. Residues proposed to form physical contact with the substrate in the substrate binding cavity are shown. The charged residues of TbHsp70.c are indicated in cyan stick representation while weakly conserved residues are shown in blue purple. The hydrophobic residues are conserved in all three cytosolic Hsp70s (TbHsp70-V441, TbHsp70.4-V439 and TbHsp70.c-V436) and is represented as green sticks.

The structural predictions show that the lid of TbHsp70.c is elevated above the ATPase domain and also in a different conformation (Figure 2.6). TbHsp70 and TbHsp70.4 seem to have adopted canonical Hsp70 three dimensional folds (Figure 2.6).Finally, the C-terminal tetrapeptide end motif of each of the cytosolic Hsp70s (Figure 2.3, red box, green highlights) reveals the presence of the canonical EEVD known for binding to STI1 in all proteins except for PfHsp70-x and TbHsp70.4 which possesses EEVN and DDVD motifs respectively. Based on this alignment, the EEVD motif is totally absent in the slightly larger TbHsp70.c, however the protein possesses a DDND motif which aligns with the EEVD in the *T. brucei* Hsp70 alignment (Appendix F, Figure F1). The C-terminal EEVD motif is crucial for the non-competitive interaction of Hsc70 and its co-chaperones; STI1 and J-proteins (Freeman et al., 1995; Demand et al., 1998). Interestingly, variations are found in the C-terminal tetrapeptide end motifs of all of the *T. brucei* Hsp70s.

The C-terminal end motif of TbHsp70 adopts the same conformation as the PITEEVD (Figure 2.6 A and C) peptide that is in complex with TPR1 of human STI1 (PDB code: 1ELW; Scheufler et al., 2000). The PKVEED found in TbHsp70 (Figure 2.6A) was superimposed on that of PITEEVD found in human Hsc70 (Figure 2.6C) resulting in an Root Mean Square Deviation (RMSD) of 0.915 Å and a structure of the TPR1 domain and peptide complex (Figure 2.6B). The RMSD is used to determine the accuracy of a homology model by comparing it to a theoretical model. The RMSD of bond lengths between the model and template are determined and should be less than 1.5Å. The RMSD bond length also indicates the similarity between two structures, a quantitative measurement between 0 and 1.5 Å shows identical or similar structure in terms of carbon backbone while any value above 1.5Å represents weak structural alignments between structures (Carugo and Pongor, 2001). Thus, the higher the RMSD value the more likely that the structure predicted is inaccurate.

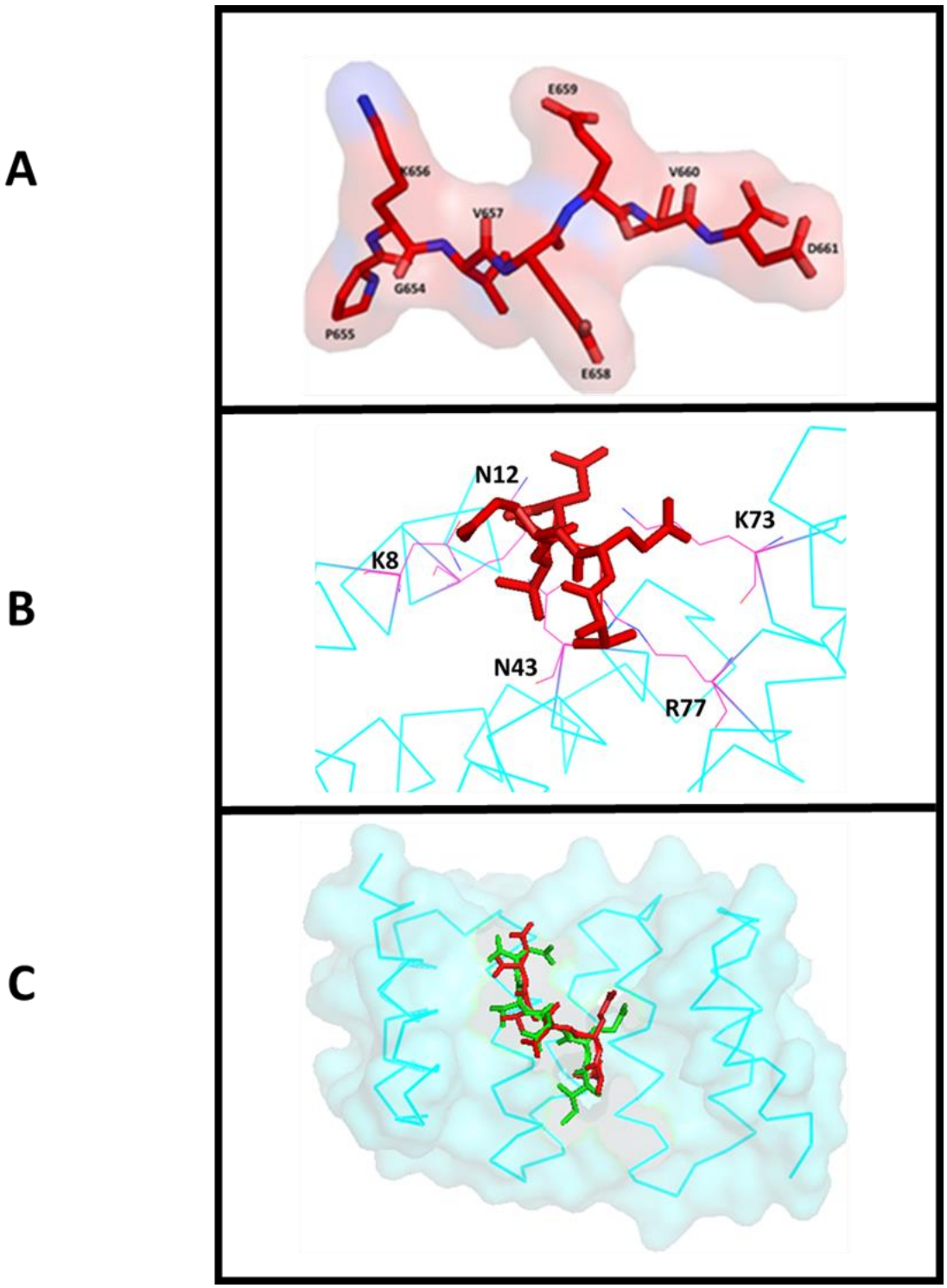


Figure 2. 6: Structural alignments of TbHsp70 C-terminal end contact residues with the three dimensional structure of human TPR1-PITEEVD complex. A) Stick representation of the TbHsp70 C-terminal PKVEEVD peptide. B) The peptide at the C-terminal end of TbHsp70 superimposed on the carboxylate clamp forming residues of TPR1. C) TPR1 domain of STI1 (blue) in complex with Hsc70 C-terminal peptide (green) superimposed on the C-terminal peptide of TbHsp70 (red) resulting in an RMSD of 0.9Å. Images generated using PyMol (DeLano, 2002).

The complex obtained suggests a strong association between TbHsp70 C-terminal tetrapeptide end motif and the TPR1 domain of human STI1. The strength of the association is also revealed by the binding of TbHsp70 PKVEEVD peptide to the STI1 residues known to interact with the C-terminal EEVD motif (Figure 2.6 B and C). The electrostatic interactions between charged residues can be seen in Figure 2.6B, this is not unexpected based on the presence of the EEVD motif in TbHsp70.

TbHsp70.4 has a conformation at the extreme C-terminal end which is different to that of TbHsp70, assessment of the TbHsp70.4 model predicted reveal that this is largely attributed by the presence of two PRO residues in the region (P633 and P635; Figure 2.7A) aswell as the evident orientation of the ASP side chain (the substitution of GLU in TbHsp70 for ASP) (Figure 2.7C). The PVPDDVD peptide of TbHsp70.4 was superimposed on the PITEEVD peptide of human Hsc70 (Figure 2.7C) resulting in an RMSD of 3.515Å, revealing weak carbon backbone alignment. The resulting PVPDDVD contact residue analysis (Figure 2.7 B) suggests that if an interaction occurs at all, TbHsp70.4 is not expected to form a stable complex with STI1 as only N12, N43 and K73 in the STI1 carboxylate clamp forming residues are involved in the association(Figure 2.7B).It is interesting to note that though the C-terminal end that binds to TPR1 of TbHsp70.4 is different to that of Hsc70, this peptide adopts the same conformation as the Hsp90 MEEVD motif found in the human STI1 TPR2A-HSP90 MEEVD complex (PDB code: 1ELR; Scheufler et al., 2000). The association between TPR2A and TbHsp70.4 should be explored.

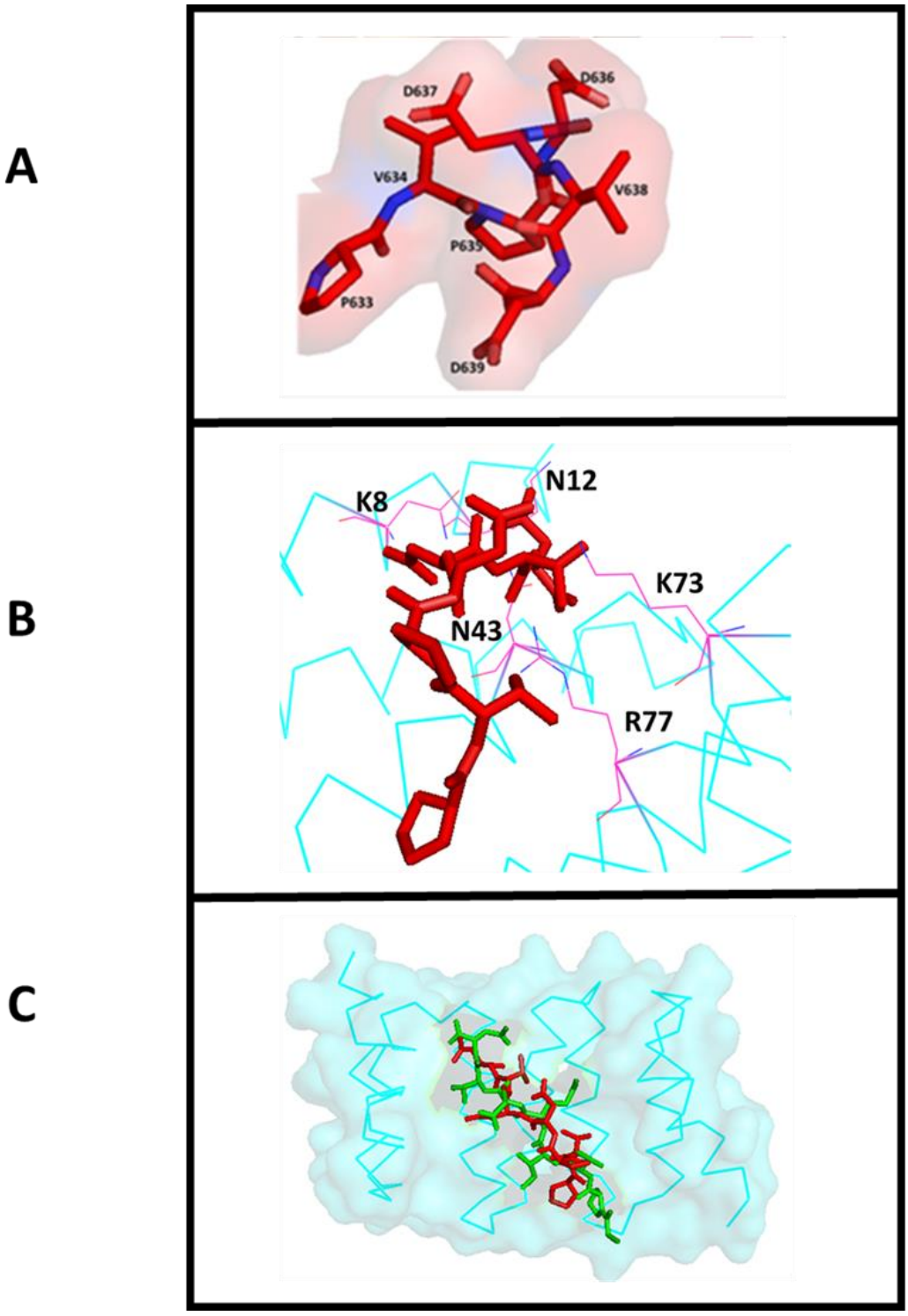


Figure 2. 7: Structural alignments of TbHsp70.4 C-terminal end contact residues with the three dimensional structure of human TPR1-PITEEVD complex. A) Stick representation of the TbHsp70.4 C-terminal PVPDDVD peptide. B) The peptide at the C-terminal end of TbHsp70.4 superimposed on the carboxylate clamp forming residues of TPR1. C) TPR1 domain of STI1 (blue) in complex with Hsc70 C-terminal peptide (green) superimposed on the C-terminal peptide of TbHsp70.4 (red) resulting in an RMSD of 3.515Å.Images generated using PyMol (DeLano, 2002).

The C-terminal peptide of TbHsp70.c that may interact with STI1 (DDND) is different to that of TbHsp70 (EEVD) and it is positioned upstream of the C-terminal end of the protein. The TbHsp70.c peptide lies in the same conformation as that of TbHsp70 which is similar to typical STI1 TPR1 binding C-terminal peptides (Figure 2.8A and Figure 2.6C), this might suggest the possibility of an interaction with STI1. However, the amino acid properties of most residues within the NDGDDND peptide (Figure 2.8A) of TbHsp70.c are functionally different to those present in canonical Hsp70s and the peptide is not located at the C-terminal end. The only interaction between human Hop TPR1 carboxylate clamp forming residues and TbHsp70.c was with K8 which is not sufficient for an interaction (Figure 2.8B).

It is predicted that there will be no interaction between TbHsp70.c and STI1 as structural alignment resulted in an RMSD of 13.678 Å indicating a large degree deviation in the structural identity and similarity. Based on the lack of superimposition of TbHsp70.c with the PITEEEVD motif of Hsc70 (Figure 2.8C) and the lack of interactions with the carboxylate clamp residues, it is unlikely that TbHsp70.c will interact with the TPR1 motif of STI1.

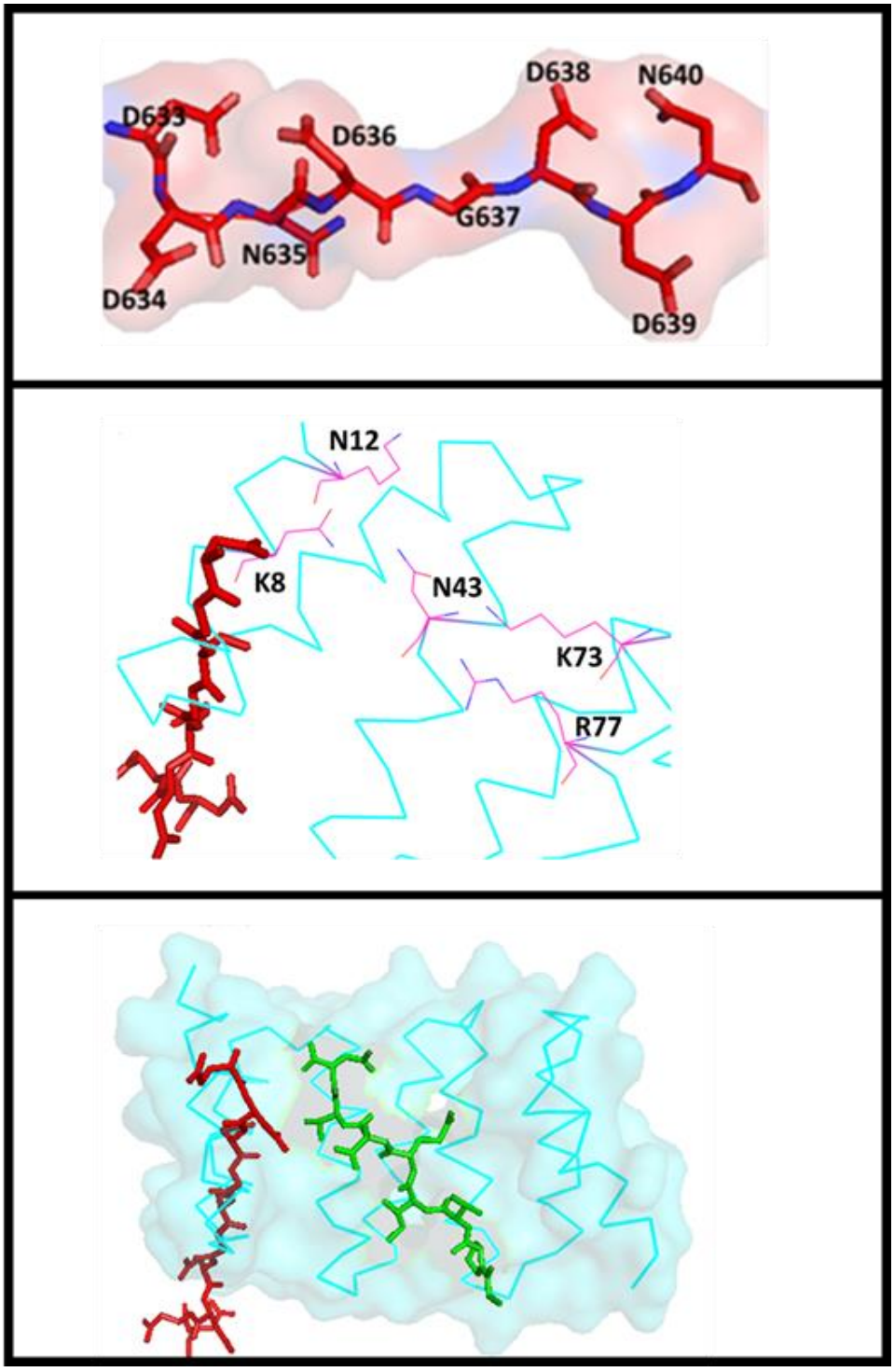


Figure 2. 3: Structural alignments of TbHsp70.c C-terminal potential contact residues with the three dimensional structure of human TPR1-PITEEVD complex. A) Stick representation of the TbHsp70.c C-terminal NDGDDND peptide. B) The peptide upstream of the C-terminal end of TbHsp70.c superimposed on the carboxylate clamp forming residues of TPR1. C) TPR1 domain of STI1 (blue) in complex with Hsc70 C-terminal peptide (green) superimposed on the C-terminal peptide of TbHsp70.c (red) resulting in an RMSD of 13.678 Å, images generated using PyMol (DeLano, 2002).

2.3.4. Genome comparison of the number of J-proteins

Less is known about the kinetoplastid J-proteins than the Hsp70s. A preliminary characterization of *T. cruzi* J-proteins was performed by Tibbetts et al (1998). The results found in this study are not consistent with previous studies, Folguiera and Requena (2007) found 65 J-proteins in *T. brucei* , 67 in *T. cruzi* and 66 in *L. major* while Shonhai and colleagues reported fewer J-proteins in *L. major* (*n=66*)and *T. cruzi* (*n=61*). These results report the latest updates for the numbers of J-proteins in TriTrypDB version 8.0. Sequences were retrieved using a combination of parameters such as querying the database using old accession numbers, carrying out a blind search and blasting the *E. coli* DNAJ J-domain and lastly the presence of a J-domain and the HPD motif were used as criteria for classifying a sequence as a J-protein.

There are notable differences in the number of J-proteins found in kinetoplastids in comparison to higher order eukaryotes (Figure 2.10). In the TriTryps, *T. brucei* appears to possess the highest number of J-proteins (*n = 67*) relative to *L. major* (*n = 60*) and *T. cruzi* CL Brener Esmeraldo-like (*n = 60*). *T. evansi* possesses the same number of J-proteins as *T. brucei*, while the extracellular parasites *T. congolense (n=63)* and *T. vivax* (*n=58*) have fewer J-proteins. There seems to be variability with regards to the number of J-proteins in *Leishmania spp* as *L. donovani*, *L. braziliensis* and *L. infatum* have 61 members while *L. mexicana* has 58 members.The absence of orthologues in *L.* major and *L. mexicana* might suggest the likelihood of species specific J-proteins. Analysis of the two strains of *T. cruzi* shows that Marinkellei has 6 J-proteins less than CL Brener Esmeraldo-like strain. Further assessment for the exact numbers of J-proteins in *T. cruzi* and *Leishmania spp* is requiredas the genomes are constantly being updated and annotated. Despite the variable numbers, kinetoplastids consist of a much larger J-protein complement than humans (*n=43*), yeast (*n=22*) or *P. falciparum* (*n=49*) (Figure 2.10).

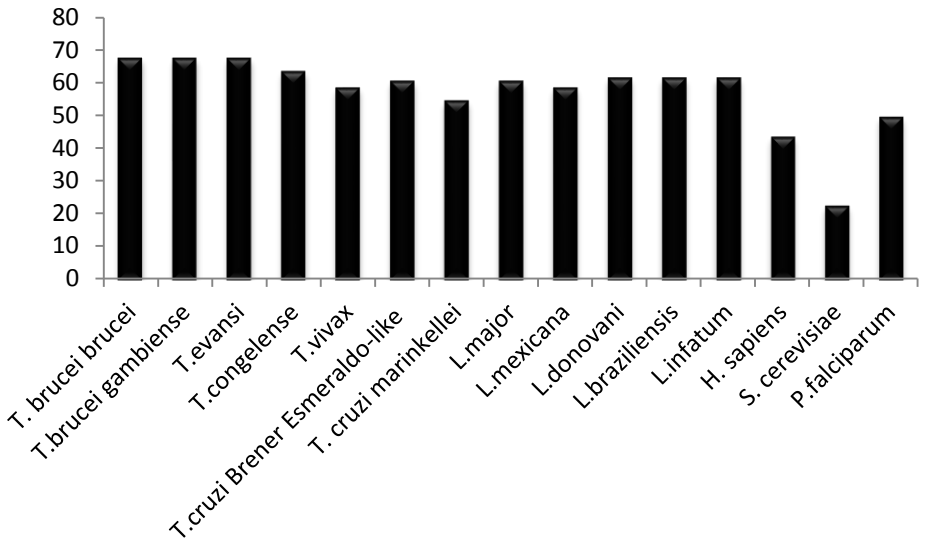


Figure 2. 4: The number of J-proteins in different genomes. Data for all kinetoplastids obtained from TriTrypDB and for *H. sapiens* (Kampinga et al. 2010); *P. falciparum* (Njunge et al. 2013); *S. cerevisiae* (Walsh et al. 2004; Cherry et al. 1998). The Y-axis represents the numbers of Hsp70s and the X-axis represents name of each organism.

2.3.4.1 Chromosomal localization of Hsp70 and J-proteins in *T. brucei*

The nuclear genome of *T. brucei* consists of 11 chromosomes totalling 26 mebabases. J-proteins are distributed on every chromosome while Hsp70s are located on the larger chromosomes (VI-XI). The majority of J-proteins are located on chromosome X and XI, putative type III mitochondrial J-proteins mainly localize on chromosomes IV, VI, VII, VIII, X and XI (Figure 2.11) and interestingly most of these particular J-proteins possess orthologues in other kinetoplastid organisms. Tbj9 and Tbj41 are unique to *T. brucei* while Tbj12 is strictly restricted to trypanosomes with no orthologoues in *Leishmania* and are only expressed in the BSF stage of the parasite life-cycle. Tbj12 and Tbj41 localize on the same chromosome and this suggests a potential complementary role that would be intriguing to investigate as these proteins have no orthologues in *Leishmania* and other eukaryotes (Figure 2.11 and Figure 2.12).

The localization of TPR containing J-proteins; Tbj42, Tbj52 and Tbj67 on the same chromosome (X) as the tandemly repeated TbHsp83 is also noted (Figure 2.11). Also, the majority of J-proteins distributed in different compartments of *T. brucei* localize on chromosome X. It seems that chromosome IV localizes atypical J-proteins, two unique TPR containing J-proteins with no homologues in higher eukaryotes (Tbj51 and Tbj65) and a rather fascinating Tbj19 which is a typical type III J-protein except it is absence in *T. cruzi* and present in *Leishmania.* Determining the function of this J-protein certainly has interesting implications for parasite biology with regards to the roles it plays in both *Leishmania* and *T. brucei* and not in *T. cruzi*.Tbj19 also shares favourable sequence indentities with Tbj36 and Tbj40, which are both predicted to localize in the mitochondria (Figure 2.13), therefore it is unlikely that the absence of this homologues in *T.* cruzi can be compensated by the closely related isoforms. The tandem repeats of mithochondrial Hsp70s are localized to a chromosome that is not dominated by molecular chaperones, while the ER-Hsp70 tandem repeats localize on the same chromosome as TbHsp70 and TbHsp70.c (XI).

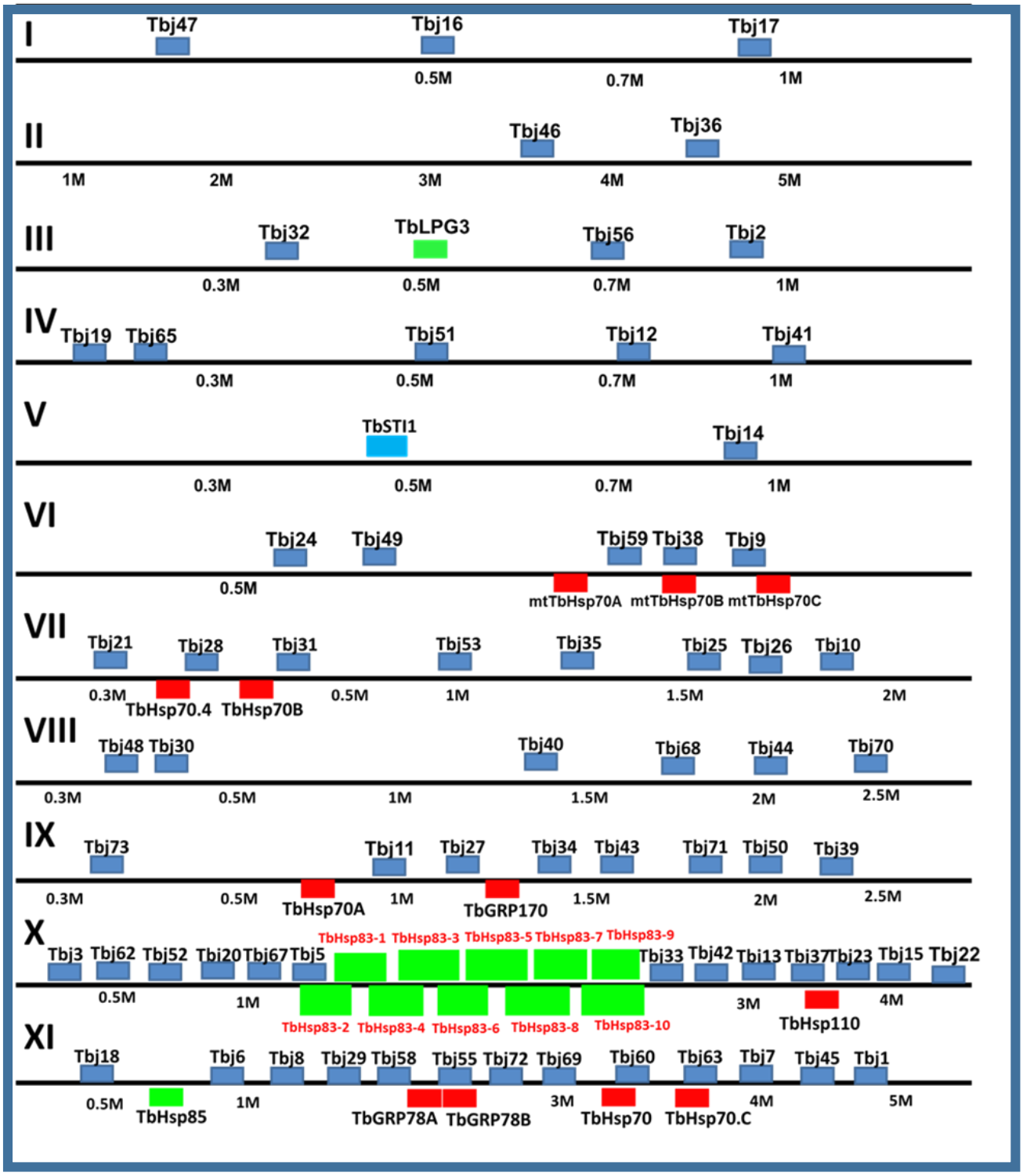


Figure 2. 5: Genomic organization of *T. brucei* Hsp70 and J-protein complements represented in red and blue respectively on this compressed segment of the 11megabased sized chromosomes of *T. brucei*. The chromosome number is represented on the left at the beginning of each chromosome, the Hsp90 complements are represented in green, the Hsp70/Hsp90 co-chaperone, TbSTI1, is represented in light blue.

The sequences of the J-proteins were retrieved from TriTrypDB, analyzed and classified into the four recognized types using the J-protein classification guides (Cheetham and Caplan, 1998; Botha et al., 2007 and Kampinga and Craig, 2010). These classifications were initially based on *E. coli* DnaJ, basically stating that type I J-proteins comprise of the signature J-domain, GF region, Zn binding domain, while type II lack the Zn binding domain and type III possesses only the J-domain. The type IV J-domains were defined and identified in *S. cerevisiae* (Walsh et al., 2004), they have since been identified in *P. falciparum* (Botha et al., 2009; Njunge et al., 2013, Pesce and Blatch, 2014) where they are defined as "pretenders" or J-like proteins with an abrogated HPD motif in the J-domain. J-proteins from *T. brucei* have not been previously classified into these four categories of J-proteins and this study provides the first report. Generally, the composition of J-proteins in kinetoplastids is as follows: type I (*n=6-7*), type II (*n=7*), type III (*n=49-53*) and type IV (*n=2*).

2.3.4.2. Type I J-proteins

According to original bioinformatics analyses, there are six type I J-proteins in *T. brucei*, namely Tbj2, Tbj3, Tbj27, Tbj45, Tbj46, and Tbj50. Very little experimental studies have been conducted on these putative types I J-proteins, two of these J-proteins (Tbj2 and Tbj3) are predicted to be cytosolic and possess orthologues in *T. cruzi* and *Leishmania* (Figure 2.12). The *T. cruzi* orthologues of Tbj2 and Tbj3 were isolated using degenerate oligonucleotides against the highly conserved YHPD sequence found within the J-domain (Tibbetts et al., 1998). Tbj2 shares 65% sequence identity with Tcj2, which has been shown to stimulate the *T. cruzi* cytosolic, heat inducible Hsp70 ATPase activity *in vivo* (Edkins et al., 2004). Additional *in vivo* complementation assays also showed the ability of Tcj2 to complement the *S. cerevisiae* Ydj1 mutant (Edkins et al., 2004) suggesting a conservation of function in these homologues proteins. The mRNA levels of Tcj2 were increased after heat induction (Tibbetts et al., 1998), therefore it is tempting to assume the same for Tbj2 which has been shown to assist in the aggregation-suppression and stimulate the ATPase activity of the novel TbHsp70.c (Burger et al., 2014). Despite the *in vitro*evidence of

this partnership, experimental proof of this chaperone-co-chaperone interaction needs to be shown *in vivo*.

Together with their parasitic kinetoplastid orthologues, Tbj2 and Tbj3 contain a C-terminal CAQQ motif (Figure 2.12 and 2.13) which functions in protein isoprenylation (Clarke, 1992; Zhu et al., 1993). This modification has also been observed in *S. cerevisiae* Ydj1 and the plant *Atriplex nummularia*, Anj1, J-protein (Zhu et al., 1993; Caplan et al, 1992). Tbj2 and Tbj3 are closely related in sequence and are therefore likely to fulfill structurally analogous roles. It is also interesting to note that there is similarity to yeast Ydj1 and Xdj1 (Walsh et al., 2004). Although mass spectrometry data shows that both Tbj2 and Tbj3 are expressed throughout the parasite life-cycle (Figure 2.12), RNAi knockdown studies have indicated that only Tbj2 is essential for all stages (loss of cell viability), while Tbj3 led to loss of fitness in the PRO trypanosomes (Figure 2.12) (Alsford et al., 2011). Tbj2 shares 42% and 43 % sequence identity with *S. cerevisiae* Ydj1 and *H. sapiens* DNAJA1 respectively, as Ydj1 is better characterized than its trypanosomal homologue, it is possible to make some putative inferences about the functions and characteristics of the Tbj2 and Tbj3 proteins. Ydj1 is involved in the translocation of the proteins into the endoplasmic reticulum and mitochondria in *S. cerevisiae* (Caplan and Douglas, 1991; Caplan et al., 1992b, Atencio and Yaffe, 1992; Becker et al., 1996). It is therefore possible that Tbj2 would show similar characteristics in *T. brucei* and based on the level of sequence and structural conservation the same would be expected for parasitic kinetoplastids.

The 2 putative ER type I J-proteins share a high sequence conservation, these J-proteins are present in all kinetoplastid parasites analyzed except *T. congolense* which has no sequence encoding J45 and interestingly *T. cruzi* (CL Brener Esmeraldo-like) has two identical sequences for J46. Phylogenetic analysis (Appendix F, Figure F2) showed that these proteins are homologues of Scj1 (Figure 2.13) which is an ER type I in *S. cerevisiae.* Scj1 regulates the protein folding that is mediated by Hsp70 (Kar2) in the lumen of the ER (Silbertstein et al., 1998). Scj1 is known to be an unusual example of a J-protein, though its domain is evolutionarily conserved to *E. coli* DNAJ,

its substrate binding domain seems to be organized through disulphide bonds rather than the co-ordination of metal ions in the oxidizing environment of the ER (Walsh et al., 2004). It is likely thatkinetoplastid ER type I J-proteins play a redundant role, though it worth noting that although *T. cruzi* appears to have only one ER Hsp70 (TcGRP78), there are three type I J-proteins to assist in its chaperoning functions.

Of the two J-proteins predicted to localize in the mitochondria, Tbj27 has been shown to be essential for parasite survival and fitness at all stages of the life-cycle (Alsford et al., 2011) (Figure 2.13). Sequence analysis of this particular protein from the TriTrypDB database is controversial at first glance as it is annotated to be 334 amino acid residues in length, lacking the J-domain. However, this may be a mis-annotation of the laboratory strain of *T. brucei* 9/427 as the *T. brucei gambiense* entry is a 460 amino acid protein with a J-domain, with favorable sequence identity and alignment to that of its orthologue. Tbj27 also does not possess *H. sapiens*, *S. cerevisiae* or *P. falciparum* homologues (Figure 2.13) making this protein an ideal chemotherapeutic target. Tbj50 has been predicted to be homologous to *H. sapiens* DNAJA3 and *S. cerevisiae* Mdj1 (Figure 2.13) following phylogenetic studies (Appendix F, Figure F2) Mdj1, a mitochondrial type I J-protein in *S. cerevisiae*, is involved in importing proteins into the mitochondrial matrix by binding substrates and stimulating the ATPase activity of the *S. cerevisiae* mitochondrial Hsp70, Sc1 (Neupert, 1997; Walsh et al., 2004). It remains to be determined whether Tbj27 and Tbj50 perform similar co-chaperoning roles with any of the three mitochondrial Hsp70s present in *T. brucei.*

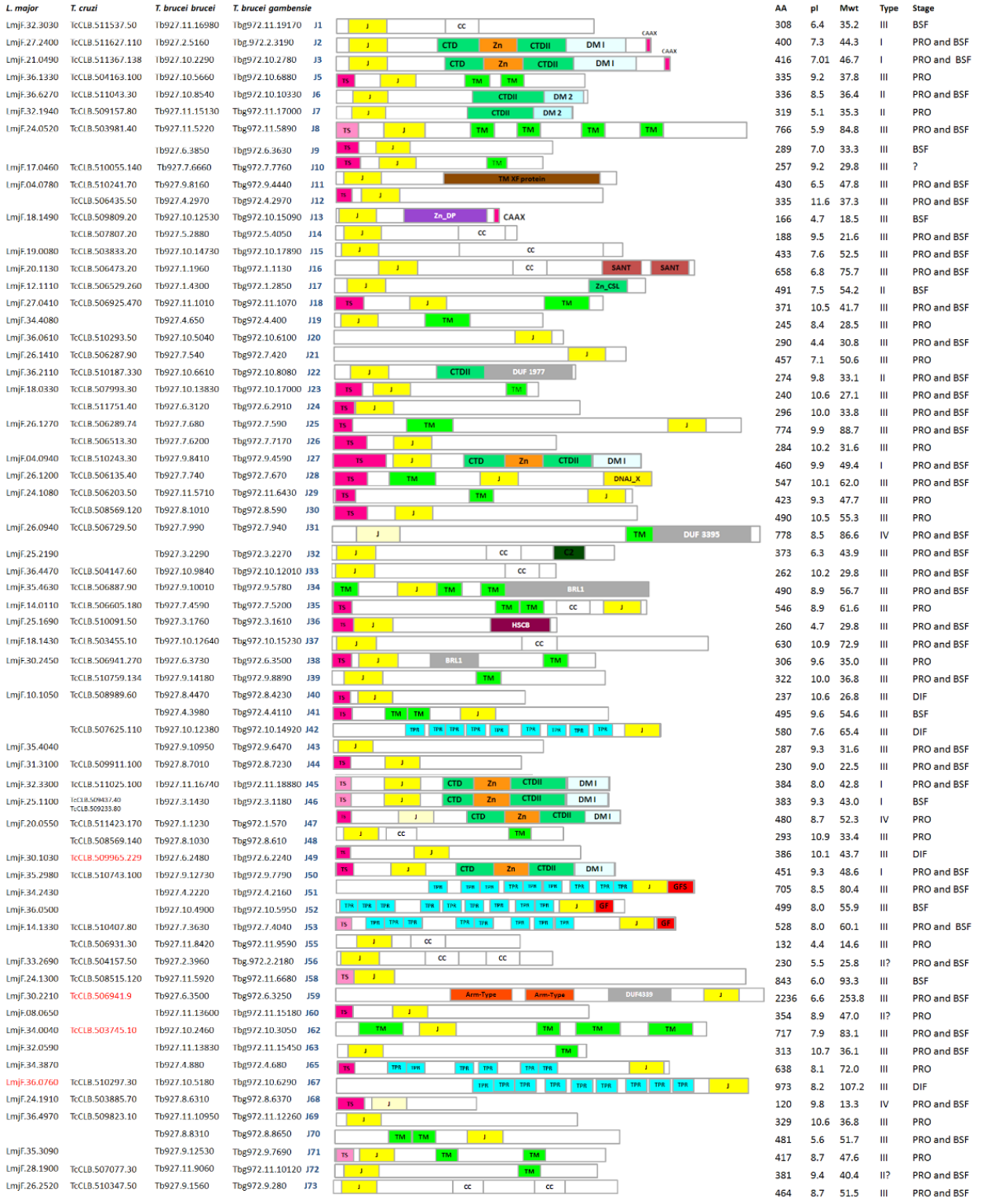


Figure 2. 12: Domain organization *of T. brucei brucei* J-proteins, together with Tritryps orthologue identification based on sequence similiarity. Properties of the proteins listed on the right AA: amino acid length, Pi: Isoelectric point, mwt: molecular weight, Type: indicating the class of J-protein and stage: parasite life -cycle expression for a particular J-protein.

2.3.4.3. Type II J-proteins

In silico predictions suggest that *T. brucei* has 7 type II J-proteins, 5 of which localize in the cytoplasm and nucleus (Tbj6, Tbj7, Tbj17, Tbj22 and Tbj56), 1 in the mitochondria (Tbj60) and the newly identified Tbj72 with an ER targeting signal at the N-terminus. Tbj6 is an orthologue of Tcj6 which has been described and shown to be the homologue of *S. cerevisiae* Sis 1 and *H. sapiens* DNAJB1 (Figure 2.12 and Figure 2.13) required for translation initiation (Salmon et al., 2001). Tcj6 has been shown to be present at constant levels in the epimastogotes and metacyclic trypomastigotes. Additionally, Tcj6 localizes in the cytosol, mostly concentrated around the nucleus and possibly associating with the ER together with the ribosomal subunits, 80S monosomes and smaller polysomes (Salmon et al., 2001). Tcj6 is able to functionally replace yeast Sis1 (Salmon et al., 2001). Based on the high sequence conservation (<65%) of Tbj6 to its *T. cruzi* orthologue, it is tempting to speculate a conserved role for this protein. Predicted to also localize in the cytosol, Tbj7 has been shown to be essential for parasite survival in the blood stream form of the trypanosome (Alsford et al., 2011).

Tbj7 has a *H. sapien* homologue DNAJB5 of unknown function (Appendix F, Figure F2). Both Tbj17 and Tbj22 are homologous to *H. sapien* DNAJB12 and DNAJB14 (Appendix F, Figure F2). It is speculated that these proteins possess a client binding domain for a selection of specific client proteins (Kampinga et al., 2010). It remains to be experimentally determined as to the specificity of these J-proteins; it is however, interesting to note that RNAi knockdown studies have shown that Tbj22, which is expressed at all stages (Figure 2.12) of the parasite, is essential for parasite viability at the PRO stage (Figure 2.13), while Tbj17 is expressed in the BSF trypanosome (Figure 2.12). It is indeed likely that Tbj17 and Tbj22 are complementing each other, probably interacting with the canonical TbHsp70 together with Tbj6 and Tbj7. Tbj72 has not been previously reported as a J-protein, however the sequence is more than 30% identical to *H. sapien* DNAJB4 and a J domain is present. Tbj72 has been predicted to localize in the ER, possessing an N-terminal leader sequence. The role of DNAJB4 has yet to be determined, it has been classified as DnaJ-like with a promiscuous client binding domain (Kampinga et al., 2010).

2.3.4.4. Type III J-proteins

T. brucei consists of 51 type III J-proteins, constituting 72% of the total J-proteins found in this single-celled organism, 43% of these possess a leader sequence for mitochondrial localization, 37% are predicted to localize in the cytoplasm and nucleus, 13% are predicted to be plasma membrane associating proteins and only 5% localizing in the endoplasmic reticulum (Figure 2.12). Additionally, 15 of these type III J-proteins contain a transmembrane domain suggesting association with the membrane, also *T. brucei* contains 6 TPR domain containing J-proteins. The type III J-protein complement as a whole has not been well characterized in any organisms; this is largely attributed to the divergence of these proteins. It has been suggested that the type III J-proteins could serve specialized roles in targeting Hsp70s to suitable organelles where they may be required (Kelley, 1999). Another view is that type IIIs may lack the ability to bind polypeptide substrates, and are therefore less likely to possess chaperone activity on their own (Walsh et al., 2004). In spite of the overall variability of type III J-proteins, the HPD motif is well-conserved throughout (Schlendstedt et al., 1995). However subtle structural differences between J-domains of different J-proteins seems to confer specificity for the substrate and is critical for mediating their interaction with specific Hsp70s (Hartl, 1996). Indeed, sequence alignment of the J-domain of all predicted type IIIs shows some subtle difference with regards to the HPD motif, with some J-proteins possessing an HPDK motif others HPDR and HPDH (Appendix F, Figure F3). There are also some notable differences in the helix III with a significantly small number of trypanosomal J-proteins possessing the KFK motif (Appendix F, Figure F3).

2.3.4.4.1. Cytoplasmic and nuclear J-proteins

Although 19 J-proteins predicted to localize in the cytoplasm and the nucleus were found to be expressed in *T. brucei*, only Tbj1 has been biochemically characterized (Louw et al., 2009) and the only two proteins found to be essential for parasite survival at all stages of the parasite life-cycle are Tbj14 and Tbj33 (Alsford et al., 2011). Interestingly, Tbj33 possesses orthologues in *Leishmania* and *Trypanosoma* while Tbj14 is absent in *Leishmania* (Figure 2.12). Both Tbj14 and

Tbj33 have been predicted to localize in the nucleus and have a typical nuclear localization signal (Appendix F, Table F4) and have homologues in *H. sapiens* and *Plasmodium* (Figure 2.13). Tbj14 and Tbj33 are homologues of DNAJC8 andDNAJC5 respectively, with functions yet to be determined (Figure 2.13). Tbj1 was found to assist *T. cruzi* Hsp70 andMsHsp70 in the suppression of protein aggregation *in vitro*, despite being unable to stimulate their ATPase activities (Louw et al., 2009). This suggests that Tbj1 does not have chaperone function of its own, and it does not possess any eukaryotic homologues.

RNAi knockdown was not performed on Tbj13, a homologue of *H. sapien* DNAJC24, *S. cerevisiae* Jjj3 and *P. falciparum* PFE0135w (Figure 2.13) which are specialized J-proteins generally known to play an essential role in the biosynthesis of diphthamide (DPH), which is a modified histidine residue on translation elongation factor, eFF2 (Thakur et al., 2012; Liu; 2004). Recent studies have shown that DPH4 both recruits Hsp70 to a specific cytoplasmic site and stimulates its ATPase activity in a J-domain dependent manner as well as independent Fe-binding and oligomerization in the cell (Thakur et al., 2012). With greater than 30% identity to its homologues, it remains to be experimentally investigated whether similar roles to those in humans and yeast are found in *T. brucei.* Both Tbj13 and *Plasmodium* PFL0815w contain a CAAX motif (Njunge et al., 2013). Tbj16 has a predicted DNA binding domain at its C-terminal and is a homologue of *S. cerevisiae* Zuo1 and *H. sapien* DNAJC2, which are ribosome associated co-chaperones (Figure 2.13). *S. cerevisiae* has a specialized ribosome-associated Hsp70 (SSB) that independently associates with the 60S subunit. Zuo1 and SSB function as J-proteins and pair with Hsp70 upon the emergence of a nascent polypeptide from the ribosome (Walsh et al., 2004). It is likely that Tbj16 will interact directly, although unlike yeast, not exclusively with the specialized form of Hsp70 in *T. brucei* TbHsp70.c which has been shown to also interact with Tbj2 (Burger et al., 2014). The yeast Jjj1 and human DNAJC20 homologue, Tbj32 contains a zinc finger (C2H2) motif near the C-terminus. Although shown to be insignificant during the parasite life-cycle, together with the constitutively expressed TbHsp70.4, this J-protein might be required during the final stages of ribosome biogenesis. Tbj37, Tbj42, Tbj69 and Tbj70 have also been predicted to be cytosolic J-proteins which are expressed throughout the parasite's life-cycle. These proteins do not have possess any

homologues in human, yeast and *P. falciparum*. It is likely that since more than one J-protein can bind to an Hsp70, that they play a co-chaperone role to the canonical TbHsp70, previously reported to lack an EEVD motif (Louw et al., 2010). Bioinformatic analysis showed that this was likely a mis-annotation as the protein possesses a C-terminal EEVD and the GGMP rich regions characteristic of parasitic canonical Hsp70s (Hartheley, 2013).

2.3.4.4.2. Endoplasmic reticulum J-proteins

Kinetoplastids possess numerous J-proteins, however only 6 of these are predicted to localize in the ER (Figure 2.13). Four (Tbj8, Tbj34, Tbj53 and Tbj58) putative type III J-proteins are predicted to localize in the ER. Tbj53 is a TPR containing J-protein while Tbj34 possesses orthologues in all members of the TriTryps and is homologous to *S. cerevisiae* Sec63, *H. sapien* DNAJC23, and *Plasmodium* PF13_0102 (Figure 2.13). RNAi experiments have implicated ER-chaperones (TbGRP78) and Tbj34 in protein secretion, cell viability and presentation of variant surface glycoproteins (Field et al., 2010), roles that place a unique burden on the parasites' chaperone machinery due to its dependency on VSGs on the cell surface for adaptation and survival in an extracellular environment.

Tbj34 has been shown to be essential for co-and post-translation of proteins in the ER, a function very likely facilitated in partnership with TbGRP78 (Dudek et al., 2009). There is a difference of 117 amino acids between Tbj58 in *T. brucei brucei* (726 aa) and *T. brucei gambiense* (843 aa), the former lacks a J-domain. However, Tbj58 is a typical type III J-protein present in all kinetoplastid organisms and has potentially been mis-annotated (Figure 2.12). Interestingly some of the J-domains found in its trypanosome orthologues (*T. cruzi, T. vivax and T. congolense*) contain two HPD motifs in tandem, the biological significance is yet to be discovered. Tbj58 also shares >30% sequence identity to human DNAJC14 which is a participant in the process of cell surface export. Tbj8, on the other hand, possesses an N-terminal ER targeting sequence and no apparent eukaryotic homologues, however this protein is present in all kinetoplastids (Figure 2.12).

2.3.4.4.3. Plasma membrane J-proteins

This study suggests that 7 transmembrane domain rich J-proteins localize in the plasma membrane or are secreted (Figure 2.12). It is interesting to note that the majority of J-proteins analyzed in the study appear to have a basic (>7) isoelectric point (pI) while most (71%) of the putative secreted J-proteins are acidic (Figure 2.12). Most of the J-proteins predicted to be secreted do not possess eukaryotic homologues (Figure 2.13) and have been shown to be critical for parasite survival, especially at the differential stage of parasite development (Tbj21, Tbj44 and Tbj55). Indeed, none of the trypanosomastid Hsp70s have been predicted to be secreted or localize in the plasma membrane. Thus a study into the mechanistic roles of these important J-proteins in parasite pathogenesis may play a role in the development of chemotherapy drugs. Although shown to be insignificant for parasite fitness and survival, Tbj39 is a strictly trypanosome J-protein with no orthologues in *Leishmania* (Figure 2.12), while Tbj59 is present in all kinetoplastid organisms and has a *H. sapiens* homologue (DNAJC13) which is involved in protein trafficking. The cell surface of trypanomastid parasites is composed of three components: the plasma membrane, the glycocalyx and the subpellicular microtubules. The glycocalyx houses several intergral proteins in *T. brucei* such as glycoproteins, glycolipids and some of the glycosylphosphatydylinositol (GPI)-anchored proteins such as VSGS, the surface mucins of *T. cruzi* and glycoinosotolphospholipids of *Leishmania* (De Souza, 1989; Ferguson, 1997). With regards to *T. brucei*, the GPI-anchored VSG has been extensively studied to investigate processes such as endocytosis and the recycling route (Overath and Engstler, 2004). Further studies remain to investigate the roles of these plasma associated J-proteins in this regard, especially in light of the implication that TbGRP78 is involved in the presentation of the VSG (Field et al., 2010).

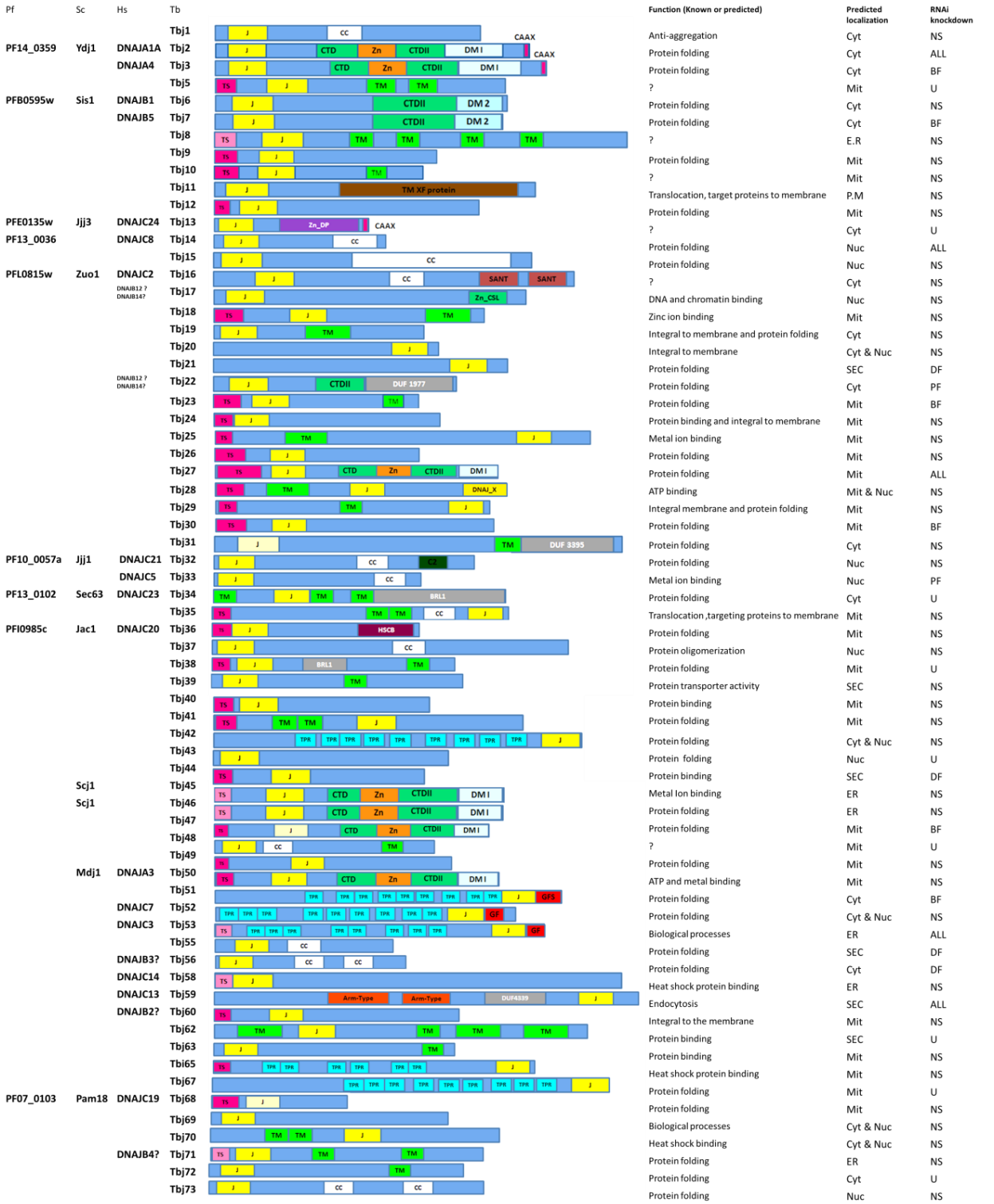


Figure 2. 13: Domain organization of *T. brucei gambiensie* J-proteins. On the right is *P. falciparum*, *S. cerevisiae* and *H. sapiens* homologue based on phylogenetic analysis, sequence identity and secondary structure prediction. The proteins functions were predicted using GO: gene ontology server, predicted localization and targeting sequences for the nucleus (grey), Mitochondria (bold black) and ER (dark Blue).

2.3.4.4.4. Mitochondrial J-proteins

The largest group of J-proteins in kinetoplastids belong to putative type III mitochondrial J-proteins, constituting ~43% of all J-proteins present in the organism. Trypanosomes supposedly have two groups of functional mitochondrial targeting signals. The first being a typical eukaryotic mitochondrial targeting sequence of 15-20 aa in length and the second a shorter sequence of 7-9 amino acid. The occurrence of a shorter peptide signal for mitochondrial import has also been observed in *S. cerevisiae*, but has been found to work inefficiently (Hausler et al., 1997). Therefore, attempts to derive the N-terminal mitochondrial leader sequence were made (Appendix F, Figure F5), however, due to sequence variability no consensus could be derived. A frequent occurrence of the positively charged residues (typical to other eukaryotes); in particular ARG (R) was observed (Appendix F, Figure F5). Common occurrences of VAL (V), LEU (L), ILE (I) and PHE (F) residues in the targeting sequence were also observed. These results agree with the observation that trypanosomes use a divergent cleavable N-terminal peptide for mitochondrial import (Pena-Daiz et al., 2004).

Only Tbj18, Tbj24, Tbj30, Tbj36 are expressed throughout the parasite life-cycle (Figure 2.12) whereas Tbj5, Tbj26, Tbj35, Tbj38 and Tbj65 are expressed at the PRO stage and Tbj40 and Tbj49 are expressed at the DIF stage of the parasite (Figure 2.12). Most if not all of the above J-proteins are not significant for fitness and survival of the parasite as determined by Alsford et al (2011) and do not have eukaryotic homologues (Figure 2.13). Although Tbj9 was not induced in the knockdown studies, a majority of mitochondrial J-proteins have been shown to be insignificant for parasite fitness and survival during the life-cycle. Only Tbj23 and Tbj30 have been shown to be essential for parasite survival at the BSF stage (Figure 2.13), while Tbj36 is proposed to be a homologue of *H. sapien* DNAJC20, *S. cerevisiae* (Jac1) and *P. falciparum* (PF10985c)

2.3.4.4.5 TPR-containing J-proteins

The *T. brucei* TPR-containing J-proteins are predicted to localize in the cytosol (Tbj42, Tbj51 and Tbj52), ER (Tbj53) and mitochondria (Tbj65 and Tbj67). It is suspected that the amount of these J-proteins could be related to the specificity of their function in each cell compartment. When initially proposed, mammalian TPR containing J-proteins were identified as novel binding partners of Hsp90 (Murthy et al., 1996). Later studies showed DNAJC7's ability to interact with cytosolic Hsp70 and Hsp90 (Brychzy et al., 2003). It is able to regulate the functions of Hsp70 through its J-domain and TPR domains and Hsp90 through the TPR domain (Brychzy et al., 2003, Moffatt et al., 2008).

Mammalian genomes typically only have 2 TPR domain containing J-proteins (Kampinga et al., 2010, Rajan and D'Silva, 2009), while kinetoplastids have 6, *Leishmania spp* lack J42 (Figure 2.12). The reasons for the absence of an orthologue of Tbj42 in *Leshmania spp* is yet to be investigated, however this particular J-protein as well as J51, J65 and J67 are novel proteins. The two J-proteins J52 and J53 with mammalian homologues possess a canonical J-domain, three TPR domains and GF rich region. Tbj53 predicted to localize in the ER, has been shown to be essential for cell viability in both the insect and mammalian stages of the parasite life cycle. Tbj53 is a homologue of DNAJC3 (shares 40% sequence identity) which was first discovered in influenza infected cells where it acts as an inhibitor of double stranded RNA-activated protein kinase (Lee et al., 1990, Lee et al., 1992). Further studies showed the ability of DNAJC3 to stimulate the ATPase activity of Grp78 during the unfolded protein response (UPR) caused due to ER stress, where it also helps to restore ER homeostasis (Oyadomari *et al*., 2006, Rutkowski *et al*., 2007, Tao *et al*., 2010; Svärd *et al*., 2011). The J domain of DNAJC3 is able to stimulate the ATPase activity of Grp78 (Tao *et al*., 2010) and no studies have been conducted on the interaction of GRP94 and DNAJC3. *T. brucei*possess two ER Hsp70s (TbGRP78A and TbGRP78b) (Louw et al., 2010) which are likely to interact with Tbj53 in a similar manner to DNAJC3 and GRP78.

The less canonical cytosolic 80.4 kDa Tbj51 has also been shown to be essential in the BSF trypanosome (Figure 2. 11). Instead of the GF region, Tbj51 possess a GFS region. Tbj52 shares 38% sequence identity with DNAJC7 and is therefore a putative homologue.

2.3 4.5. Type IV J-proteins

Initially proposed by Cheetham and Caplan (1998), type IV J-proteins have been defined by the corrupt/abrogated HPD motif in the J-domain. They are also known as extensions of the type I-III class of J-proteins. This class of J-proteins have been well described in *P. falciparum* (Botha et al., 2007; Njunge et al., 2013). Parasitic kinetoplastids possess two putative type IV J-proteins which range in size from 36.3 kDa to 86.6 kDa and the *T. brucei* complements have been predicted to localize in the cytoplasm (Tbj31) and mitochondria (Tbj68) (Figure 2.12).

Predicted to localize in the cytoplasm, Tbj31 has orthologues across all parasitic kinetoplastids. Sequence analysis with other well characterized J-proteins showed that this protein contained the conserved residues TYR7 and LEU10 present in canonical J-domains, but the key LYS/ARG 26 residues was altered to LEU26 residues followed by a GLU27 residue. The HPDK motif commonly seen in type III J-proteins of *T. brucei* was replaced by an HTDK motif in Tbj31 and its kinetoplastid orthologues, also the KFK motif abrogated for an RFQ motif. It was interesting to note that the QKRAA motif in Tbj31 was better conserved with the *E. coli* DnaJ (GKRAA) when compared to other *T. brucei* type III J-proteins. RNAi knockdown studies resulted in no significant loss in the parasite‘s fitness and survival in the absence of Tbj31 (Alsford et al., 2011), the protein has however been detected in the bloodstream plasma membrane fraction and in the PRO stage trypanosome and therefore it is expressed throughout the parasite‘s life cycle (Bridges et al., 2008; Jones et al., 2006).

2.3.4.6. J-like proteins

T. brucei Tbj47 possesses orthologues in all 7 parasitic kinetoplastid species compared in this study, this protein is particularly interesting as RNAi data has shown that its knockdown resulted in defects in the growth of the parasite in the bloodstream form (Subramaniam *et al*., 2006) and therefore this protein is an interesting drug target. No apparent homologues in higher eukaryotic systems were identified, the only hits obtained through these searches were invariably bacterial (less than 20% sequence identity), this therefore raises the possibility that this protein is unique to parasitic kinetoplastids although they also have low sequence identity (<35%). The overall J-domain architecture of Tbj47 was found to be poorly conserved with a <20% sequence identity to the *E. coli* DnaJ. The key LYS/ARG26 residue was not conserved, it was altered to residue ALA25 in Tbj47, ASP26 in *T. cruzi spp* and THR26 in *Leishmania spp*. The HPD motif was abrogated to KDP, LRR and LRE in *T. bruceispp*, *T. cruzi spp* and *Leishmania spp* respectively. The HIS33 residue in the HPD motif is a critical residue required for J-protein and Hsp70 interaction (Mayer *et al*., 1999). Studies in which this HIS residue of the J-domain have been substituted for another amino acid have shown abrogation of J-protein-Hsp70 interaction and subsequent loss of functional interaction between the two proteins (Laufen *et al*., 1999). It will be interesting to investigate how Tbj47 functions without the essential HIS33 as it possess SER in place of HPD.

2.4. Conclusion

It has become increasingly evident that chaperones are essential to the life-cycle of a wide variety of important human and animal pathogens. Kinetoplastids possess an expanded complement of genes encoding J-proteins, possibly due to evolutionary radiation. It is tempting to hypothesize that the large number of J-proteins in kinetoplastids compared to the other organisms could be due to an adaptation related to the pathogenicity of these organisms. It is also likely that these J-proteins may have roles unique to the biology of kinetoplastids which are yet to be determined. Some of these J-proteins could be essential during the parasite transition from insect vector to

the mammalian host, protecting the parasite against environmental stress. However, *P. falciparum* is subjected to the same stress-condition and has two thirds of the number of J-proteins (Botha et al., 2007, Njunge et al., 2013, Blatch and Pesce, 2014). *P. falciparum* exports some of the J-proteins into the infected erythrocyte to interact with host Hsp70s and assist in the remodellng of the erythrocyte by the parasite (Botha et al., 2007). Although this scenario may be likely for extracellular parasitic kinetoplastids (*T. cruzi* and *Leishmania spp*), it is not likely for the extracellular *T. brucei* which possess the highest number of J-proteins compared to all kinetoplastids, this is in contrary to the fact it has the smallest nuclear genome (Berriman et al., 2005). Another noteworthy possibility is that the single-celled protozoan parasites have a greater diversity of substrates that interact with Hsp70 than their eukaryotic counterparts.

Strong evidence has been provided that Hsp70 and J-proteins are essential for parasite development and pathogenesis, in *T. brucei* 50% and 21% of Hsp70 and J-proteins respectively are essential at some or all stages of development. Albeit, the study has made progress in the characterization of *T. brucei* Hsp70s in light of the recent genome updates and classifying J-proteins, predicting their localization to draw connections with Hsp70s. However, despite the progress reported here, there is still much to learn at many levels. For example, the molecular details of the Hsp70/J-proteins chaperone interactions and pathways are yet to be elucidated, some of these pathway may represent emerging drug targets. Several Hsp70s and J-proteins could be explored independently or in a complex as drug targets, for example TbHsp70 based on its high level of identity to the *L. tarentolae* which is a promising drug target (Brochu et al., 2004), also based on its features, TbHsp70 is likely to play a role in *T. brucei* pathogenesis. TbHsp70.c was demonstrated to be a ribosome-associated Hsp70, the motivation for this particular chaperone as a potential drug target lies in its uniqueness. It has been shown that TbHsp70.c does indeed have features in common with canonical Hsp70s as well as a higher sequence similarity to its human and yeast homologues. TbHsp70.c has ~100 residues which are absent in its homologues constituting an entire α-helix in the lid of the substrate binding domain. This region could be explored for chemotherapeutic studies. The uniqueness of Tbj27 also allows this particular J-protein to be pursued independently as a drug target. The diversity of J-proteins,

however provides an appetizing environment for clinical intervention and chemotherapeutic drug exploration as the specificity of Hsp70s has at times been challenging. It was interesting to note that some members of the salivarian trypanosomes (*T. brucei* and *T. evansi*) have the same number of J-proteins. It was also noted that there were some trypanomastid specific J-proteins, J12, J39 and J42 in *T. brucei* and J9 and J41 in *T. evansi* and J4 is specific to the extracellular parasites *Leishmania* and *T. cruzi.* This study also identified three unreported J-proteins and proposes that J66 and J47, which were previously classified as canonical J-proteins (Folguiera and Requena, 2007), should be regrouped as J-like proteins, due to the absence of a J-domain or even abrogated HPD motif in these proteins.

CHAPTER THREE

Analysis of Hsp90 and TPR containing co-chaperones of Hsp70 and Hsp90 in *T. brucei*

3.0. Introduction

Co-chaperones are non-client binding partners of molecular chaperones, they play a crucial role of regulating the activities of proteins such as Hsp70 and Hsp90 (Caplan, 2003). Some co-chaperones may exhibit independent chaperone activity, such as theability to suppress the aggregation of polypeptides and stimulate ATPase activity. Some co-chaperones can bind both the client protein and chaperone simultaneously, while others are unable to bind client proteins (Caplan, 2003). Since the early work on Hsp90, numerous co-chaperones have been isolated and characterized mainly in human and yeast systems (Zuehlke and Johnson, 2012).

There are two main categories of Hsp90 co-chaperones: TPR containing and non-TPR containing proteins. The latter include p23, cell division cycle 37 (Cdc37) and activator of the Hsp90 ATPase (Aha1). Initially discovered in the early 1990s, p23 is a small co-chaperone which binds dimeric ATP-bound Hsp90 at the N-terminal domain (Ali et al., 2006). The main role of p23 is to stabilize the dimeric ATP-bound Hsp90, although it is able to bind weakly to nucleotide-free Hsp90 (apo-Hsp90) (McLaughlin et al., 2006). Aha1 binds to the middle domain on Hsp90, thus promoting ATP hydrolysis and the subsequent release of client proteins (Mayer et al., 2002, Retzlaff et al., 2010). Cdc37 is required for the accurate activation of several cellular kinases, a function which is achieved by inhibiting the ATPase activity of Hsp90 (Grammatikakis et al., 1999, Pearl, 2005).

TPR containing co-chaperones form the majority of Hsp90 associating proteins, they bind at the C-terminal EEVD motif (Scheufler *et al*., 2000, Young *et al*., 2003). STI1, FK506-binding family of immunophilins (FKBP51/52), cyclophilin 40 (Cyp40), protein phosphatase 5 (PP5) and carboxyl terminus of Hsc70 interacting protein (CHIP) are examples of TPR-containing co-chaperones (Smith, 2004).Although the TPR domains are structurally conserved, TPR containing co-chaperones display functional diversity. They bind and regulate the activity of Hsp90 at differentpoints in the heat shock protein complex formed during the Hsp90 cycle. During the formation of the Hsp90 hetero-complex, STI1 is thought to be the first TPR containing co-

chaperone to bind and stabilize the ADP-bound open conformation of Hsp90, thereby blocking its ATPase activity (Young et al., 2001, Schmid et al., 2012). This ensures efficient transfer of client proteins from Hsp70 to Hsp90. The subsequent release of STI1 is followed by the binding of FK506-binding proteins 51 and 52 or Cyp40 which contributes peptidyl-prolyl cis/trans isomeraseactivity (Mayr et al., 2000; Cox et al., 2007), whereas the phosphorylation state of Hsp90 and client protein is achieved by PP5, which competes with STI1 and high molecular weight PPIases for Hsp90 binding (Johnson and Toft 1994; Chen et al. 1996; Cliff et al., 2006; Hinds and Sanchez, 2007). Other TPR containing co-chaperones of Hsp90 are involved in transportation of Hsp90-client complexes or facilitating proteosome-targeted degradation (Cyr *et al*., 2002, Davies and Sanchez, 2005, Cox and Johnson, 2011, Mollapour and Neckers, 2012).CHIP is also involved in negatively regulating the refolding activity of Hsp70 by interfering with its ATPase activity while also assisting in chaperone mediated protein degradation (Connell et al., 2001).More details about Hsp90 co-chaperones can be obtained at: http://www.picard.ch/downloads/Hsp90interactors.pdf.

The functions of most Hsp70s are regulated by their co-chaperones which can be divided into nucleotide exchange factors (NEF), J-proteins and TPR-containing co-chaperones. Generally, J-proteins control the chaperone roles of Hsp70s by passing substrates to Hsp70 and regulating the ATPase cycle (Fan et al., 2003). NEFs and TPR containing co-chaperones have the ability to regulate the function of Hsp70 (Smith, 2004). The general binding of TPR domains to heat shock proteins are conferred by five highly conserved residues in the TPR groove, also known as the carboxylate clamp (Scheufler et al., 2000). Carboxylate clamp residues within TPR1 of STI1 interact with Hsp70 whereas TPR2A and B interact with Hsp90 C-terminal EEVD motifs. There are also other TPR containing co-chaperones of Hsp70 such as Hsp70 interacting protein (Hip) which stabilizes the ADP bound state of Hsp70, which in turn enhances its chaperoning activity (Höhfeld et al., 1995). DNAJC7 and DNAJC3 are human TPR containing J-proteins which have been shown to interact with cytosolic and ER Hsp70 respectively (Melville et al., 1999; Brychzy et al., 2003, Moffatt et al., 2008).

A review of Hsp70 in TriTryps has been published (Louw et al., 2010) and little has been done on Hsp90 (the cytosolic Hsp90s are called Hsp83 in kinetoplastids and will be referred to as such in this chapter) or TPR containing co-chaperones. The presence of the heat shock protein complex has also not been established in *T. brucei*. Typically the genes encoding TbHsp83 give rise to a 704 amino acid residue protein with an estimated size of 85 kDa, which is present at all stages of the parasite life-cycle and commonly referred to as Hsp83 (Dragon et al., 1987). TcHsp83 is present in non-heat shocked cells and its expression is more pronounced under heat shock (Carvalho et al., 1987; Dragon et al., 1987). Unlike mammalian Hsp90, which seems to bind ATP but possesses virtually no ATPase activity, parasitic Hsp90s including PfHsp90 and TeHsp83 (Pallavi et al., 2010), TcHsp83 (Nadeau et al., 1992) and TbHsp83 (Pizarro et al., 2013) have been shown to have high ATPase activity. Little is known about the role of Hsp83 in cellular metabolism, its clients and co-chaperones in kinetoplastids, especially since the trypanosome genome lacks evidence of steroid hormone receptors and tyrosine kinases that are prominent Hsp90 clients in mammalian cells, homologues of other clients are recognizable, including serine/threonine kinases and cell cycle regulators (Parsons et al., 2005; Nett et al., 2009). This chapter is aimed at analyzing Hsp90 across different organisms, with special reference to TbHsp83, as well as finding TPR-containing co-chaperones of Hsp90 and Hsp70. STI1 is the most important TPR-containing co-chaperone as it connects Hsp90 to Hsp70 and the presence of the Hsp90-STI1-Hsp70 complex in *T. brucei* will be investigated using bioinformatics tools.

3.1. Objectives

Original results of Hsp90 and TPR containing co-chaperones of both Hsp90 and Hsp70, particularly STI1, were obtained using various bioinformatics tools. This study focused on comparing proteins from different organisms with emphasis placed on key features present in homologues and orthologues of *T. brucei*.

The specific objectives of this chapter are therefore:

1. To compare genomic distribution of Hsp90 proteins across different organisms.
2. To conduct phylogenetic analysis of Hsp90 proteins in terms of features and taxonomic relationships to different organisms.
3. To predict the 3D structure of Hsp90 using homology modelling
4. To conduct a TPR motif prediction for TbSTI1 and previously uncharacterized *T. brucei* TPR-containing co-chaperones of Hsp70 and Hsp90
5. To analyze the conservation of residues involved in Hsp70 and Hsp90 binding in these co-chaperones
6. To conduct phylogenetic analysis of TbSTI1 proteins in terms of features and taxonomic relationships to different organisms.
7. To analyze the structural conservation and sequence similarity of TbSTI1 in relation to other organisms
8. To predict the 3D structure of TbSTI1 domains using homology modelling

3.2. Methods

3.2.1. Sequence retrieval

All predicted genomic and amino acid sequences used to study the gene structure, chromosome location and protein features of kinetoplastid Hsp90s and TPR containing co chaperones of Hsp90 and Hsp70 were retrieved from TriTryDB. Keyword searches “Heat shock protein 90”, “Hsp90”, “Hsp83” were used to scan the database. Homologues and orthologues of Hsp83 and TPR containing co-chaperones of Hsp90 and Hsp70 were obtained using the same criteria described in section 2.2.1.

3.2.1.1. *T. brucei* Hsp83 proteins orthologues and homologues

Kinetoplastida cytosolic Hsp90s are referred to as Hsp83 and will therefore be termed Hsp83. The primary protein sequence encoding TbHsp83-1 (Accession number: Tb927.10.10890) was obtained by using keyword searches such “Heat shock protein 90”, “Hsp90”, “Hsp83” on TriTrypDB, and sequences were retrieved using the methodology described in section 2.1.1. The sequences obtained were used for comparison of Hsp90 complements across various organisms.

3.2.1.2. TbSTI1 orthologues and homologues

The TbSTI1 (Accession number: Tb927.5.2940) protein sequence was used as a query for pBLAST (Altschul et al., 1997) on NCBI to acquire approximately thirty five STI1 protein sequences from a wide range of eukaryotic organisms which were used for subsequent phylogenetic analyses.

3.2.1.3. TPR containing co-chaperones of Hsp70 and Hsp90

Human TPR containing co-chaperones such as SGT, HIP, PP5, CYP40 and FKBP52 were used as queries for a blast search on TriTrypDB to retrieve homologues of the under-studied co-chaperones of Hsp70 and Hsp90 in *T. brucei*.

3.2.2. Phylogenetic analysis

The chaperone properties of *T. brucei* Hsp83 proteins were studied in relation to kinetoplastids, *H. sapiens*, *S. cerevisiae* and *P. falciparum* while thirty-five organisms containing STI1 were compared by phylogenetic analysis. The cladograms of STI1 and Hsp83s were constructed following the parameters described in section 2.2.2., except the protein substitution model was the Jones-Taylor-Thornton (JTT) evolutional model (Jones et al., 1992).

3.2.3. Domain organization and subcellular localization prediction

Hsp83s and TPR containing co-chaperones of Hsp83 and Hsp70 in *T. brucei* were used as queries in the servers listed in section 2.2.3 to identify the domains, motif and predict the subcellular localization.

3.2.4. Multiple sequence alignments

T-COFFEE, Clustal-omega, MUSCLE and Promals3D were used to generate sequence alignments of the TPR domains and STI1 proteins.

3.2.5. Homology modelling

The modelling process and evaluation steps described in section 2.2.5 were followed to generate 100 models of TbHsp83 using *S. cerevisiae* Hsp82 (PDB code: 2CG9) (Ali et al., 2006) as template. Since the crystal structure of full-length STI1 has not been determined yet, TPR domains and DP domains of TbSTI1 were modelled. High-throughput models of TPR1 domain were generated using human TPR1 (PDB code 1ELW; Scheufler et al., 2000), TPR2A (PDB code: 1ELR; Scheufler et al., 2000) and TPR2B (PDB code: 3UPV; Schmid et al., 2012). The DP domains of TbSTI1 were generated using PDB code: 2llw; Schmid et al., 2012 and PDB code: 2llv, Schmid et al., 2012 for DP1 and DP2 respectively. All targets shared a >30% sequence identity to their templates.

3.3. Results and Discussion

3.3.1. Sequence comparison of *HSP90* genes and proteins

HSP90 genes are arranged in tandem clusters across all kinetoplastid organisms, although the number of genes does not necessarily correlate to the number of Hsp90 isoforms. Figure 3.1 shows the predicted number of *HSP90* genes as well as the number of isoforms.The primary amino acid sequences of the majority of these proteins are identical, *Leishmania* comprises the largest complement of *HSP90* genes (*n=19*) followed by *T. b. brucei* and *T. evansi* (*n=12*) and then *T. cruzi* (*n = 11*)(Figure 3.1), it is also striking to note the difference in the number of genes encoding *HSP90* in the two subspecies of *T. b. brucei* (*n =12)* and *T. b. gambiense* (*n=5*) (Figure 3.1). It is also interesting to note the different distributions of *HSP83* genes across African trypanosomes *T. vivax* (*n=9*) and *T. congolense* (*n=9*) (Figure 3.1). It is possible that the number of *HSP83* genes in African trypanosomes will increase as their genomes get updated in the future. In contrast to humans and *P. falciparum* with 4 Hsp90 isoforms, there are 3 isoforms present in all kinetoplastid organisms used in this study (Figure 3.1).

Hsp90 is a well-validated drug target in most organisms, it would interesting to explore whether the difference in gene distribution between the two *T. brucei* subspecies would have an effect on the parasites 's reaction to typical Hsp90 inhibitors. *T. b. brucei* has been shown to be more selective to geldanamycin structural analogs, 17-AAG and 17-DMAG (Meyer and Shapiro, 2013).

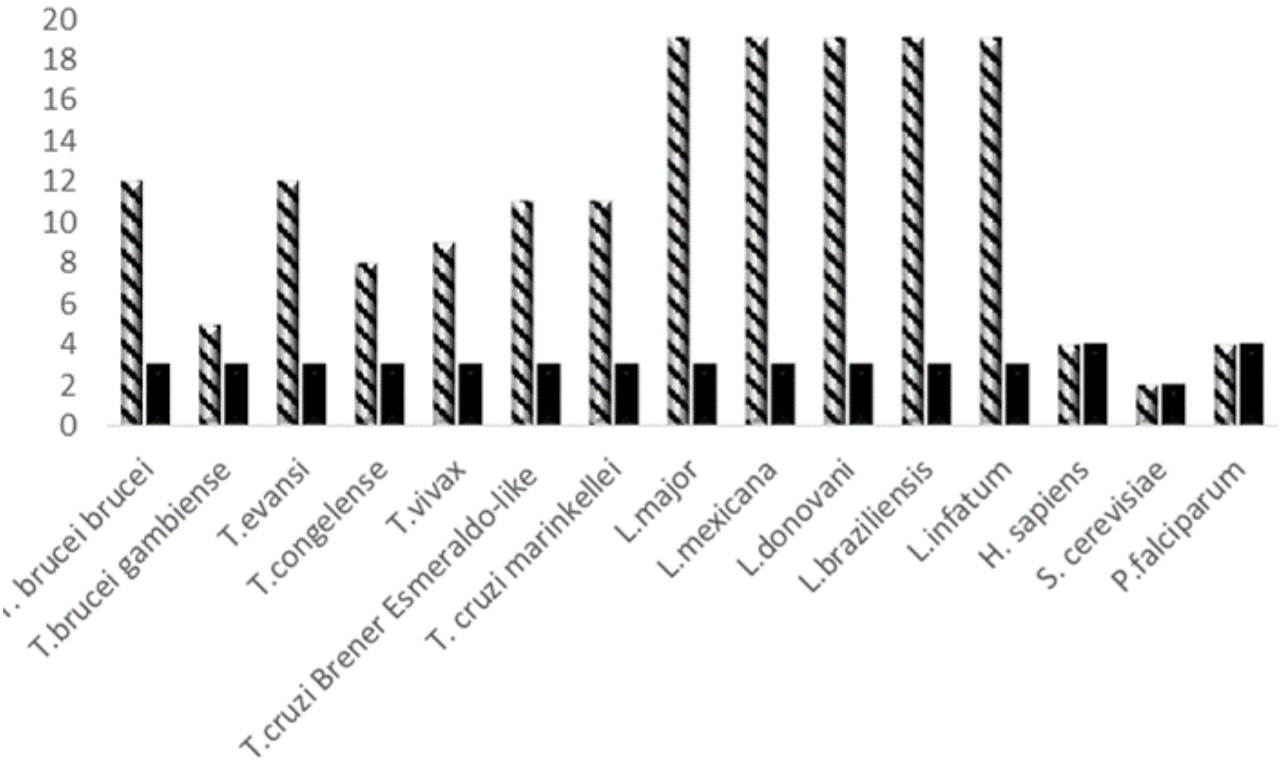


Figure 3. 1: The number of genes encoding Hsp90 as well as the number of isoforms of Hsp90 in different genomes, data for kinetoplastid organisms was obtained at TriTrypDB. The gene complement for *H. sapiens* (Picard, 2002); *P. falciparum* (Acharya et al., 2007); *S. cerevisiae* (Louvion et al., 1996) is also shown. The Y-axis represents the numbers of *HSP90* genes (Striped) and protein isoforms (Black) and the X-axis represents the names of each organism.

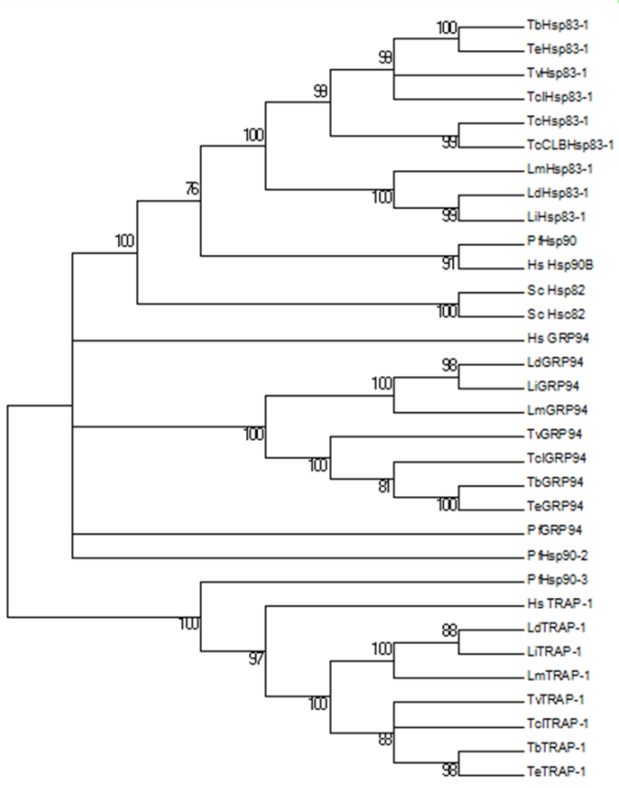
The results obtained in the study for *Leishmania* and *T. b. brucei* are consistent with previous reports (Dragon et al., 1987; Shonhai et al., 2011). Six copies of *HSP90* genes arranged in tandem in *T. cruzi* were determined (Dragon et al., 1987; Shonhai et al., 2011), whilst this study identified 11 copies arranged in tandem. The occurrence of tandem duplication of genes in kinetoplastids has been reported previously, particularly in the case of Hsp70s (Lee et al., 1990; Olson et al., 1994), however the significance of this expansion remains unclear. Some have speculated functional redundancy of encoded proteins, the possibility of abundant gene expression through gene expansion as kinetoplastids lack classical eukaryotic promoter-dependent mechanisms present in of gene regulation has also been suggested (Morimoto et al., 1998; Morales et al., 2005).

3.3.2. Phylogenetic analysis, domain organization and homologue identification of *T. brucei* Hsp83 proteins

Phylogenetic analysis of Hsp90 proteins showed 3 distinct clusters which are subdivided based on cytosolic, ER and mitochondrial subcellular localization (Figure 3.2). This clustering based on localization and homology reflects the level of conservation of Hsp90 proteins in general, for instance TbHsp83 shares 59% and 64% sequence identity with human Hsp90α and Hsp90β respectively with even greater homology in the ATPase domains. With regards to Hsp83s in kinetoplastids, it was observed that typically the ER and mitochondria contain a single Hsp90 homologue with the exception of *T. congolense*. There are two sequences encoding GRP94 in *T. congolense*, TcIL3000_0_48690 appears to be a truncated version of 513 amino acids in length, this sequence is missing the ER leader sequence at the N-terminal domain which is present in TcIL3000_0_48690. It is proposed that this truncated version of TclGRP94 is an artifact of sequencing and is in fact not a bona fide ER-Hsp90 homologue. Similar to the ER resident Hsp70 (GRP78), GRP94 in kinetoplastids lacks the C-terminal tetrapeptide ER retention motif KDEL, instead they possess EGDL (*Leishmania spp*) and AGDL (trypanosomes) (Figure 3.2). Surprisingly, the deduced carboxy-terminal tetrapeptide is a significant divergent of the KDEL motif especially with regards to the trypanosomes, a modification that certainly argues for unusual latitude in the ligand specificity of the trypanosomal receptor (Bangs et al., 1993). It is possible that kinetoplastids use a minimal ER retention signal –DL that is completely independent of the residues in position 1 and 2. With regards to the mitochondrial homologues of TRAP-1 in kinetoplastids, a larger degree of conservation with humans is observed except *Leishmania spp.* are ~114 amino acids smaller than TbTRAP-1.

In contrast to humans and yeast, which have two isoforms for the constitutive and heat-inducible cytosoic Hsp90, it is predited that kinetoplastids contain one Hsp83. The monophyletic cluster of cytosolic Hsp90s suggests a general conservation of function, structure and sequence in Hsp83 homologues (Figure 3.2). It was interesting to note that *Leishmania*Hsp83 possess a C-terminal

MEQVD unlike the common heptapeptide of MEEVD for association with STI1 and TPR containing co-chaperones. However, despite the single point substitution of ASP to GLU, an interaction between LdSTI1 and LdHsp83 was demonstrated (Morales et al., 2010).



Accession number	Genes	AA	Mwt	Pi	Motif	Other gene locus
Tb927.10.10890	10	704	80.7	4.81	EEVD	Tb927.10.10900,Tb927.10.10910, Tb927.10.10920, Tb927.10.10940, Tb927.10.10950, Tb927.10.10960 ,Tb927.10.10970, Tb927.10.10980 & Tb927.10.10930
TevSTIB805.10.11470	10	704	80.7	4.81	EEVD	TevSTIB805.10.11480, TevSTIB805.10.11500, TevSTIB805.10.11520, TevSTIB805.10.11530, TevSTIB805.10.11490, TevSTIB805.10.11550, TevSTIB805.10.11560, TevSTIB805.10.11540 &TevSTIB805.10.11510
TvY486_1010600	7	904	100	5.02	EQVD	TvY486_0003270 TvY486_1010590 TvY486_0023660 TvY486_0043420 TvY486_0035480 TvY486_1010610
TcIL3000_0_26140	6	904	100	5.02	EQVD	TcIL3000_0_26130 TcIL3000_0_48690 TcIL3000_0_26120 TcIL3000_0_27730 TcIL3000_0_58840
Tc_MARK_3581	9	704	80.7	4.81	EQVD	Tc_MARK_9863
TcCLB.509105.140	9	704	80.7	4.81	EQVD	TcCLB.509643.130 TcCLB.507713.30 TcCLB.509105.150 TcCLB.509643.130 TcCLB.509105.140
LmjF.33.0312	17	701	80.5	4.81	EQVD	LmjF.33.0314, LmjF.33.0316, LmjF.33.0318, LmjF.33.0320, LmjF.33.0323, LmjF.33.0326, LmjF.33.0330, LmjF.33.0333, LmjF.33.0336,LmjF.33.0340, LmjF.33.0343, LmjF.33.0346, LmjF.33.0350,LmjF.33.0355,LmjF.33.0360 & LmjF.33.0365
LdBPK_330360.1	17	701	80.5	4.81	EQVD	LdBPK_330370.1 LdBPK_330350.1
LinJ.33.0350	17	701	80.5	4.81	EQVD	LinJ.33.0360 LinJ.33.0370
PF3D7_0708400		745	86.2	4.93	EEVD	
NP_031381.2	1	724	83.2	4.96	EEVD	
NP_015084.1	1	709	81.4	4.84	EEVD	
NP_013911.1	1	705	80.9	4.77	EEVD	
NP_003290.1	1	809	92.4	4.96	KDEL	
LdBPK_290790.1	1	771	90.0	4.67	EGDL	
LinJ.29.0790	1	771	90.0	4.67	EGDL	
LmjF.29.0760	1	771	90.0	4.67	EGDL	
TvY486_0302990	1	767	88.5	5.01	AGDL	
TcIL3000_3_2270	2	758	85.3	5.64	AGDL	TcIL3000_0_48690
Tb927.3.3580	1	771	90.0	4.67	AGDL	
TevSTIB805.3.3780	1	773	90.0	4.67	AGDL	
PF3D7_1222300	1	821	95.2	5.28	SDEL	
PF3D7_1443900	1	927	107	5.38		
PF3D7_1118200	1	930	108.4	6.22		
NP_057376.2	1	704	80.0	8.30		
LdBPK_332520.1	1	634	70.0	8.55		
LinJ.33.2520	1	634	70.0	8.55		
LmjF.33.2390	1	634		8.55		
TvY486_1102710	1	750	75.0	8.00		
TcIL3000.11.2400	1	754	75.0	8.00		
Tb927.11.2650	1	753	75.0	8.00		
TevSTIB805.11_01.2720	1	753	75.0	8.00		

Figure 3. 2*:* Phylogenetic analysis of kinetoplastid Hsp83 proteins in comparison to humans, yeast and *P. falciparum*. Analysis was based on number of gene copies, amino acid length (AA), molecular mass (Mwt), Isoelectric point (pI) and the presence of a C-terminal motif known for binding to co-chaperones or retaining a ER Hsp70s in the cellular organelle. The evolutionary relationship was generated using the Maximum likelihood statistical method. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Branches corresponding to partitions replicated in less than 75% of the bootstrap analyzes were collapsed. The evolutionary distances were computed using the Jones-Taylor- Thornton (JTT) amino acid substitutions model (Please refer to the pdf for better visualization of the accession numbers).

3.3.3. Structural features of *T. brucei* Hsp90s

In common with their homologues, *T. brucei* Hsp90 consists of the N-terminal domain (NTD) which contains the ATPase domain, the middle domain (MD) which is connected to the C-terminal domain (CTD) via the highly charged linker region (Figure 3.3). Conservation at sequence level (data not shown) of the C-terminal domain and the EEVD motif known to associate with TPR containing co-chaperones was also observed for the cytoplasmic TbHsp83.

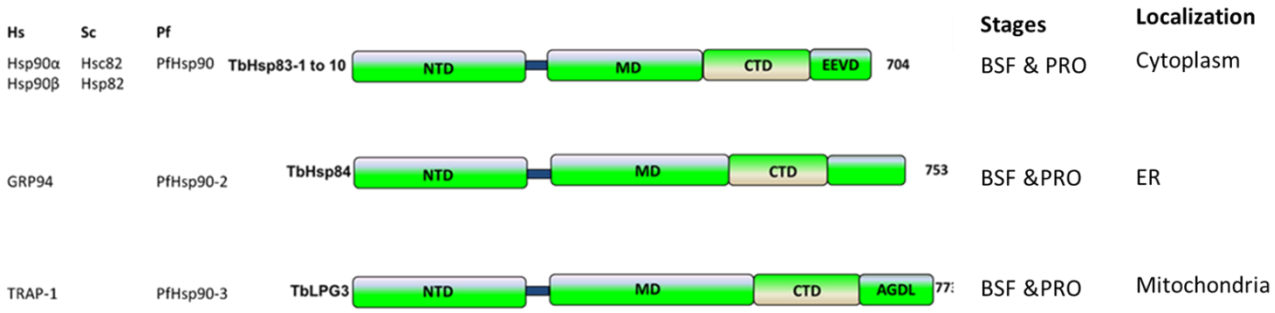


Figure 3. 3: Domain organization of Hsp90 proteins in T. brucei and their homologues in humans, yeast and P. falciparum. NTD is the N-terminal domain, MD is the middle domain and CTD is C-terminal domain with end motifs for cytoplasmic and ER Hsp90s. Proteomics data obtained from TriTrypDB revealed that all 12 *T. brucei* Hsp90 are expressed at all stages of the life-cycle.

A study conducted by Alsford et al (2011) showed that both TbHsp84 (TRAP-1) and TbHsp83 are essential for parasite fitness and survival at the PRO stage. TbHsp83 which has been shown to be essential at all stages of parasite life cycle, has a conserved domain organization with its homologues and ends with a highly unstructured MEEVD at the C-terminus (Figure 3.3). The three-dimensional model of TbHsp83 was found to be very similar to that of the yeast homologue (Figure 3.4).The nucleotide binding domain of TbHsp83 is composed of an α and β sandwich in which a pocket extends from the buried face of the antiparallel β-sheet surface which forms the nucleotide binding site. The key residue (D79 in yeast) (Stebbins et al., 1997; Panaretou et al., 1998) for ATP binding and competitive binding of inhibitors (radical and geldanamycin) is

conserved in TbHsp83. This demonstrates a conservation of the ATPase function in the trypanosome which has been confirmed by Pizarro et al., 2013.

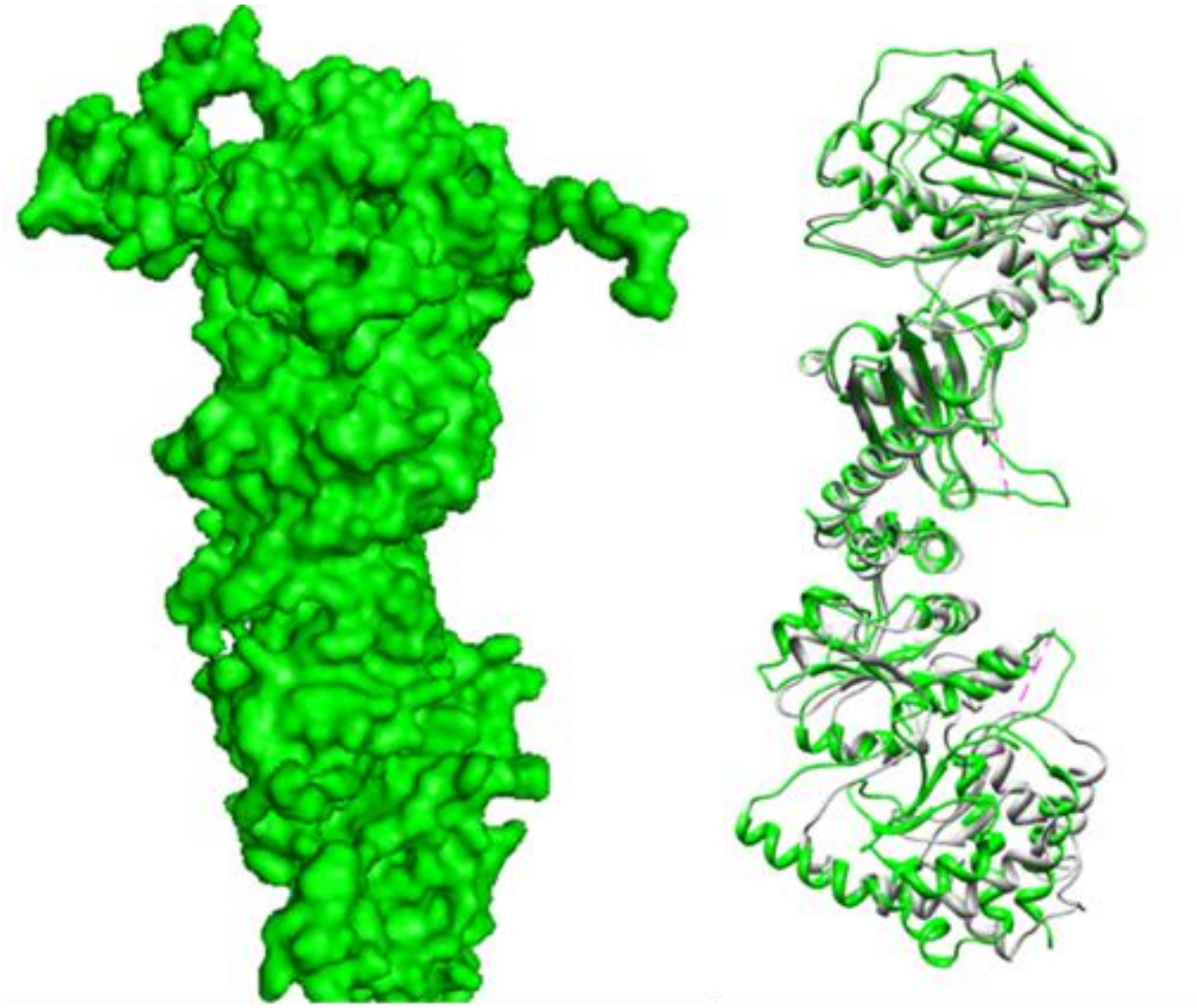


Figure 3. 4: Predictions of the three-dimensional structure of TbHsp83. Yeast Hsp82 was used as a template for homology modelling. The model generated was visualized using PyMol (DeLano, 2002), sphere presentation is shown on the left. TbHsp83 model was also superimposed to its template (grey) resulting in a RMSD of 0.1Å.

It was interesting to find the highly flexible charged linker region connecting the N-terminus and middle domain of cytosolic Hsp90s in most eukaryotic organisms is also present in *T. brucei*. The

linker region is absent in TRAP-1 (Gupta, 1995) and it was certainly not expected to be conserved in TbHsp83 as this region has been shown to play a role in the Hsp90-steroid receptor complex (Cadepond et al., 1993, Dittmar et al., 1997, Kosano et al., 1998) and kinases (Miyata and Yahara, 1995), the evidence for which is lacking on the *T. brucei* genome. Furthermore, mutational studies showed that the absence of the linker does not affect the roles of Hsp90 (Louvion et al., 1996).

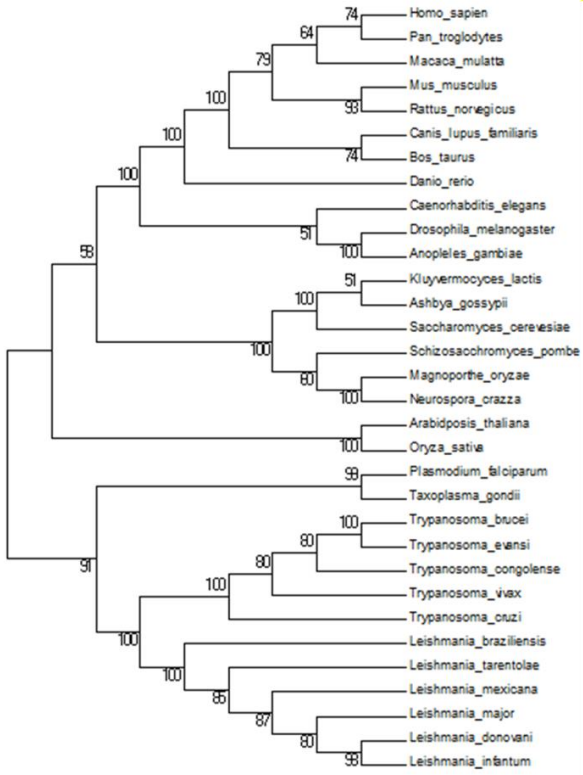
3.3.4. TbSTI1 is a highly conserved Hsp70/Hsp90 associating protein

STI1, the co-chaperone of Hsp70 and Hsp90, is widely studied, distributed and well characterized. Evidence of STI1 modulating the interaction of Hsp70 and Hsp90 has been shown in various higher eukaryotic systems as well as in parasitic organisms. The presence of TbSTI1 (accession number: Tb927.5.2940) has been confirmed previously (Louw et al., 2010), however little has been done to characterize this protein. Stage-specific formation of the Hsp70-STI1-Hsp83 complex has been demonstrated for its orthologue in *L. donovani* (Morales et al., 2010). It has also been shown that TcSTI1 associates with TcHsp70 at the epimastigote stage of the parasite life-cycle (Schmidt et al., 2011). Bioinformatics tools will be used to predict the role of TbSTI1 in relation to other eukaryotic organisms and kinetoplastids in which the presence of the multi-chaperone complex is established.

3.3.4.1. STI1 is found in eukaryotic organisms

A total of 35 sequences for STI1 proteins from different organisms was used to generate the phylogenetic tree. STI1 is a strictly eukaryotic protein and therefore its phylogeny reveals evolutionary relationships. Six distinct clusters, which can be subdivided into mammalia, nematode and insecta, fungae, plantae, apicomplexa and Euglenozoa were generated (Figure 3.5). As expected, the largest cluster contains all the kinetoplastid organisms with branch partitions >80% (Figure 3.5). The properties of STI1 proteins highlight some key differences with

respect to nematode (*C. elegans*) and insecta (*D. melanogaster* and *A. gambiae*) with the amino acid length being between 150-350 residues less than other organisms. With the exception of nematode and insect, most proteins share a >30% sequence identity to TbSTI1 suggesting that this protein is structurally conserved. TbSTI1 and TeSTI1 clustered closely to each other and share 87% sequence identity (Figure 3.5).



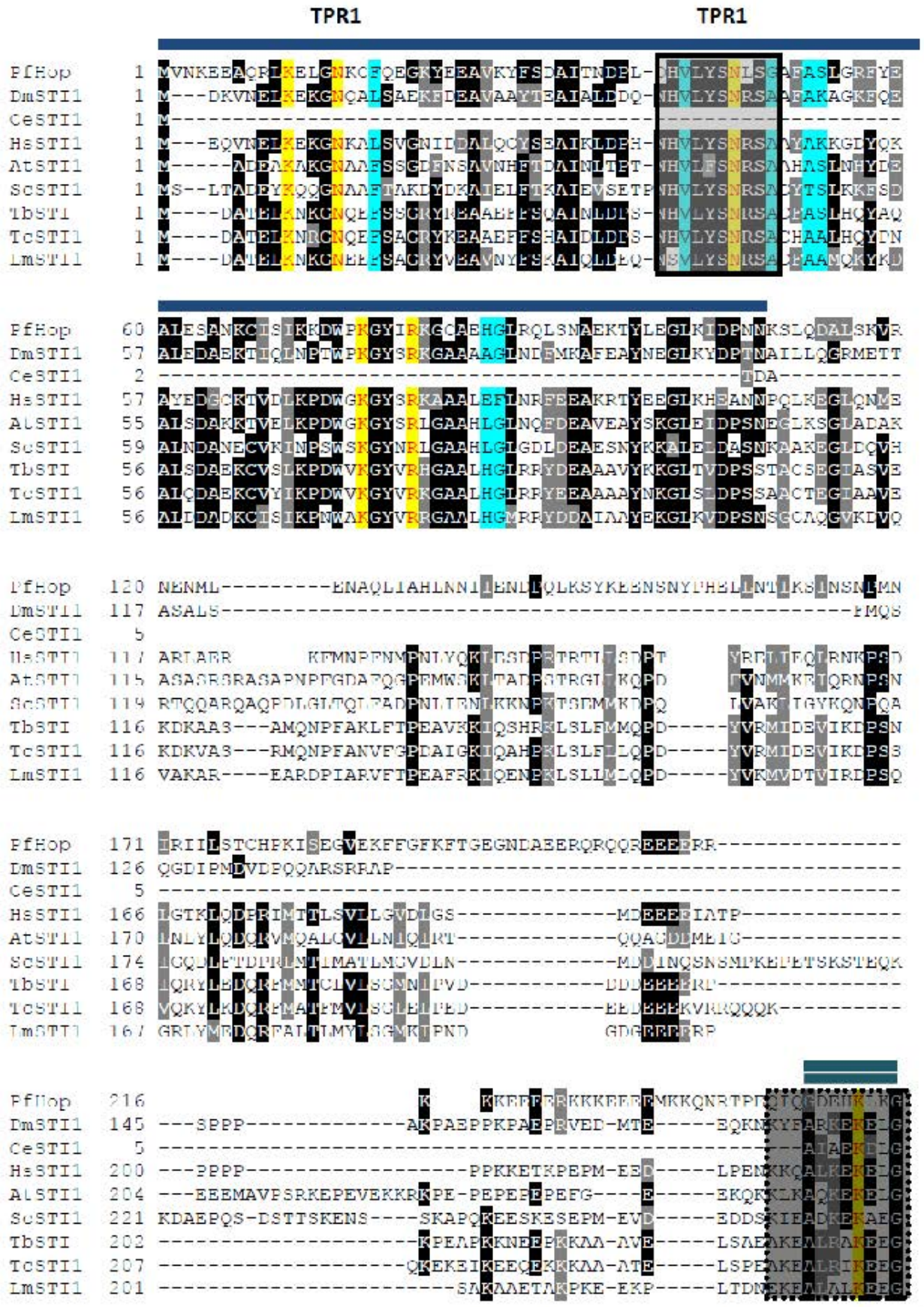
Accession Number	AA	Mwt	Pi	E-value	Seq id	Description
NP_006810.1	543	62.6	6.40	6e-36	41%	Stress-induced-phosphoprotein 1
XP_001163388.1	546	62.6	6.40	6e-36	41%	Stress-induced-phosphoprotein 1 isoform 1
XP_001115412.2	543	62.5	6.40	6e-40	40%	Stress-induced-phosphoprotein 1
NP_058017.1	543	62.5	6.40	1e-36	41%	Stress-induced-phosphoprotein 1
NP_620266.1	543	62.5	6.40	2e-36	41%	Stress-induced-phosphoprotein 1
XP_854960.1	543	62.5	6.40	1e-136	41%	Stress-induced-phosphoprotein 1
NP_001030569.1	543	62.5	6.08	1e-136	41%	Stress-induced-phosphoprotein 1
NP_001007767.1	565	63.4	6.62	9e-110	23%	stress-induced-phosphoprotein 1
NP_503322.1	320	36.9		2e-73	20%	Protein STI-1
NP_477354.1	490	55.6		1e-74	34%	Hsp70/Hsp90 organizing protein homolog
XP_319365.4	325	38.3		4e-80	36%	AGAP010188-PA
XP_451313.1	581	65.1		8e-88	35%	ypothetical protein
NP_984637.1	580			3e-84	36%	AEL224Wp
NP_014670.1	589			1e-96	37%	Stip1
NP_588123.1	591			1e-90	35%	Chaperone activator Sti1
XP_364135.2	503			5e-89	38%	Stress-induced-phosphoprotein 1
XP_964548.1	578	64.5		6e-87	40%	Hypothetical protein NCU00714
NP_172691.1	572	64.6		2e-96	40%	Putative stress-inducible protein
NP_001053429.2	574	64.9		2e-73	42%	Hypothetical protein
PF3D7_1434300	565	63.4		3e-75	35%	Hsp70/Hsp90 organizing protein
	565	63.4	6.63	2e-22	39%	
Tb927.5.2940	550	62.3	6.23	0.0	100%	Stress-induced protein sti1, putative
TevSTIB805.5.3390	500	62.3	6.23	0.0	83%	Stress-induced protein sti1, putative
TcIL3000_5_3210	549	62.5	6.68	0.0	100%	Stress-induced protein sti1, putative
TcIL3000_5_2850	568	64.8	6.53	0.0	99%	Unspecified product
TvY486_0502310	547	61.9	6.38	0.0	79%	Stress-induced protein sti1, putative
TcCLB.506321.290	556	63.1	6.60	0.0	78%	Stress-induced protein sti1, putative
LbrM.08.0880	547	62.4	6.44	0.0	64%	Stress-induced protein sti1
LtaP08.0850	546	62.0	6.53	0.0	65%	Stress-induced protein sti1
LmxM.08.1110	545	62.0	6.24	0.0	66%	Stress-induced protein sti1
LmjF.08.1110	546	62.1	6.04	0.0	65%	Stress-induced protein sti1
LdBPK_081020.1	546	62.2	6.05	0.0	65%	Stress-induced protein sti1
LinJ.08.1020	546	62.2	6.04	0.0	65%	Stress-induced protein sti1

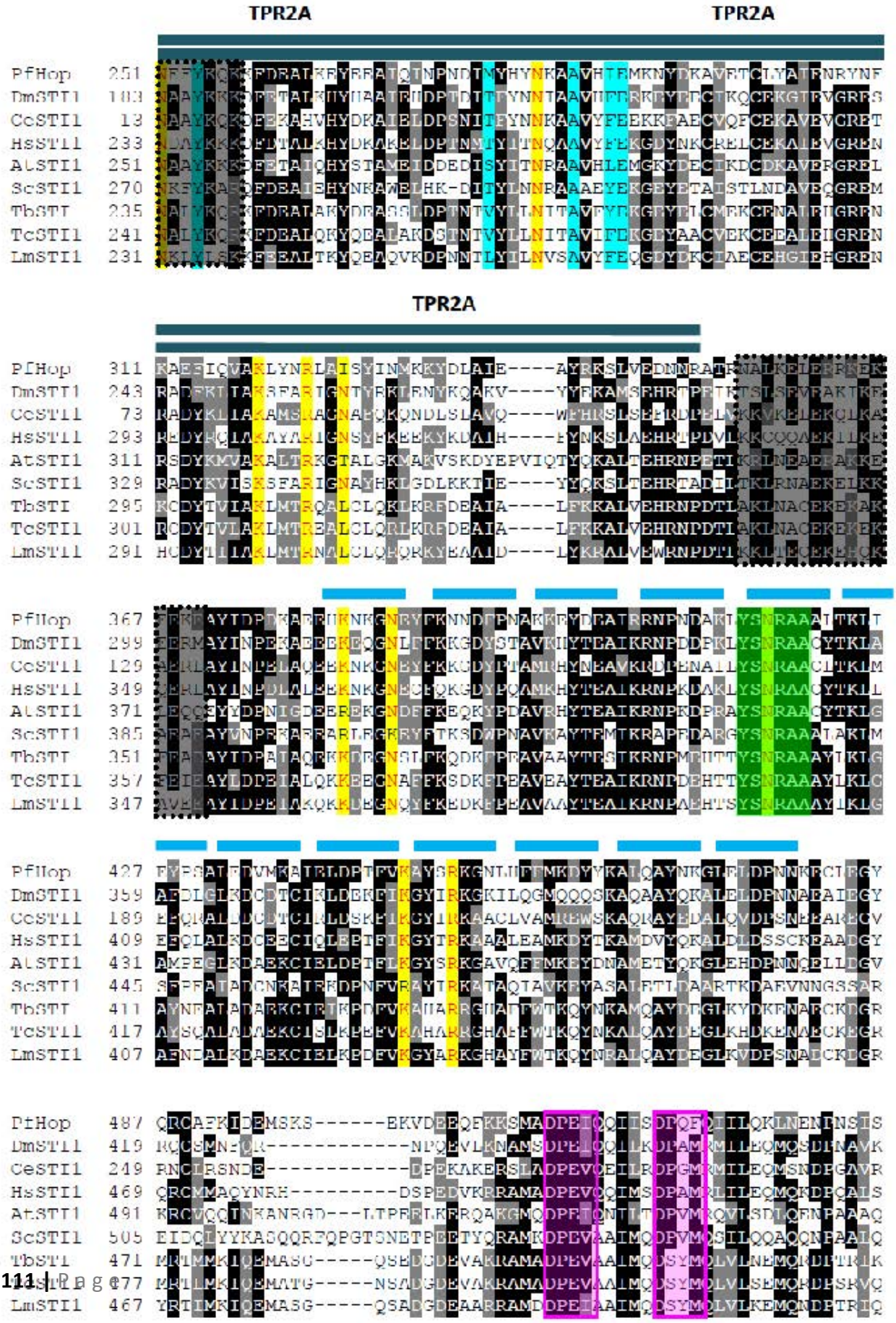
Figure 3. 5: Taxonomic unit of 35 STI1 proteins from different organisms. The evolutionary relationship was generated using the maximum likelihood statistical method. 100 replicates to infer bootstrap consensus tree. Branches corresponding to partitions replicated in less than 50% of the bootstrap analyzes were collapsed.The evolutionary distances were calculated using Poisson amino acid substitution method. STI1 protein properties were analyzed based in their sequence identity to TbSTI1, Isoelectric point (Pi), molecular weight in kDa (Mwt), E-value and amino acid sequence length (Please refer to the pdf for better resolution).

3.3.4.2. Sequence alignment of STI1 proteins

With the exception of *C. elegans*, the alignment revealed three TPR domains in all STI1 homologues (Figure 3.6) which is supplemented by the domain organization prediction (Appendix F, Figure F10). The carboxylate clamp forming residues (K, N and R), which are essential for respective TPR domain interactions with Hsp70 and Hsp90 (Odunuga et al., 2003) are conserved across STI1 homologues (Figure 3.6). Indeed TbSTI1 is demonstrated to possess the residues consistent with Hsp70 and Hsp90 interactions.

Also interestingly, residues determined to confer specificity for either Hsp70 or Hsp90 binding in mSTI1 (Odunuga et al., 2003) were found to be conserved in TbSTI1 (Figure 3.6, cyan highlights) as well as the other STI1 homologues. The Y-motif, which overlaps with TPR2A, consists of a stretch of highly charged amino acids (boxed in red) and was found to be fairly conserved in the STI1 homologues. TbSTI1 and its orthologues (TcSTI1 and LmSTI1) appear to have more negatively charged residues compared to other eukaryotes, the same was also observed for PfHop (Figure 3.6). Using the functional identification of the nuclear localization signal (NLS) in mouse STI1 (Longshaw et al., 2004) as a guide, two potential bipartite NLSs were found in TbSTI1 (NLS 1 and 2) (Figure 3.6, dark grey background). The first NLS overlapped with TPR2A while the second is predicted to be located between TPR2A and TPR2B. Both putative NLSs were found to be conserved except for PfHop which appears to be slightly divergent. Also found to be conserved across STI1 homologues were two motifs, the first is a decamer (NHVLYSNRSA) (Figure 3.6, boxed in black) which is found in TPR1. The hexamer (YSNRAA) (Figure 3.6, yellow background) is positioned on equivalent helix in TPR2B, the same observation was made for mSTI1 (Odunuga et al., 2003) and PfHop (Gitau et al., 2011). Located toward the C-terminal end of the STI1 homologues are the second DP repeats downstream of TPR2B, and these were found to have a substitution of PRO to SER in TriTryps organisms (Figure 3.6, boxed in pink).


TPR1
TPR1
PfHop
DmSTI1
CeSTI1
HsSTI1
AtSTI1
ScSTI1
TbSTI
TcSTI1
LmSTI1


TPR2A
TPR2A
TPR2A
PfHop 251
DmSTI1 183
CeSTI1 13
HsSTI1 233
AtSTI1 251
ScSTI1 270
TbSTI 235
TcSTI1 241
LmSTI1 231
PfHop 311
DmSTI1 243
CeSTI1 73
HsSTI1 293
AtSTI1 311
ScSTI1 329
TbSTI 295
TcSTI1 301
LmSTI1 291
PfHop 367
DmSTI1 299
CeSTI1 129
HsSTI1 349
AtSTI1 371
ScSTI1 385
TbSTI 351
TcSTI1 357
LmSTI1 347
PfHop 427
DmSTI1 359
CeSTI1 189
HsSTI1 409
AtSTI1 431
ScSTI1 445
TbSTI 411
TcSTI1 417
LmSTI1 407
PfHop 487
DmSTI1 419
CeSTI1 249
HsSTI1 469
AtSTI1 491
ScSTI1 505
TbSTI 471
LmSTI1 467

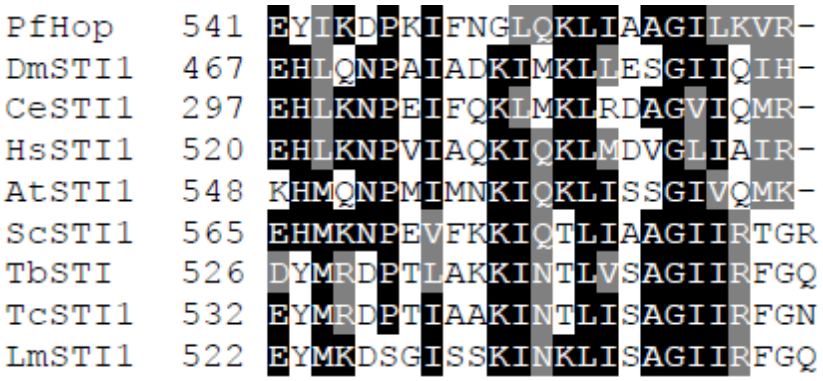


Figure 3. 6: Multiple sequence alignment of STI1 homologues. The TPR domains are highlighted as follows: TPR1 (dark blue line), TPR2A (double green lines) and TPR2B (dashed light blue). The carboxylate clamp forming residues are in yellow highlights (Scheufler et al., 2000)while the amino acid residues that are predicted to be crucial for TPR1-Hsp70 and TPR2A-Hsp90 interaction are indicated in cyan highlights (Odunuga et al., 2003). Residues constituting DP repeats (Carrigan et al., 2005) motifs are represented by magenta highlights. The highly conserved decamer (black box) and hexamer motifs (yellow box) (Odunuga et al., 2003). The nuclear localization signals (NLS1), (NLS2) and Y charged motif are highlighted by black dashed boxes (Cheung-Flynn et al., 2003; Allan et al., 2006).

3.3.4.3. TbSTI1 structural conservation

The presence of 9 TPR motifs constituting 3 functional TPR domains and 2 DP regions structurally defines STI1 homologues. The structural conservation of these features was determined for TbSTI1 (Figure 3.7). Furthermore, homology models of each of the TPR domains of TbSTI1 shows 6 antiparallel α helices and important residues thought to mediate the interaction between TbSTI1 and TbHsp70 and TbHsp83 are surface exposed and appear to project into the grooves (Figure 3.7).

In TbSTI1 there is a substitution of PRO to SER508 in the DP2 region (Figure 3.6), this did not affect the overall three-dimensional structural fold of the region (Figure 3.7), and however, this mutation may have an impact on the overall functional DP regions. The DP repeats are thought to play a role in maintaining the structural integrity of STI1 (Allan *et al*., 2006; Schmid *et al*., 2012). Thus, structural conservation of these domains leads to the expectation that TbSTI1 will bind Hsp70 and Hsp83. The presence of previously identified phosphoTbSTI1 residues (Nett et al.,

2009) SER15, SER221 and SER485 was also confirmed in the multiple sequence alignment (Figure 3.6). Translation initiation, protein folding and protein catabolism are implicated in trypanosomatid differential gene expression (Clayton, 2002), this highlights the importance of protein phosphorylation in the regulation of posttranslational modifications. It is likely that TbSTI1 phosphorylation sites are essential for parasite viability in the same manner as its LdSTI1 orthologues (Morales et al., 2010). The ability to be phosphorylated at the parasite pathogenic stage would suggest that TbSTI1 may be a promising drug target through inhibiting its protein kinases.

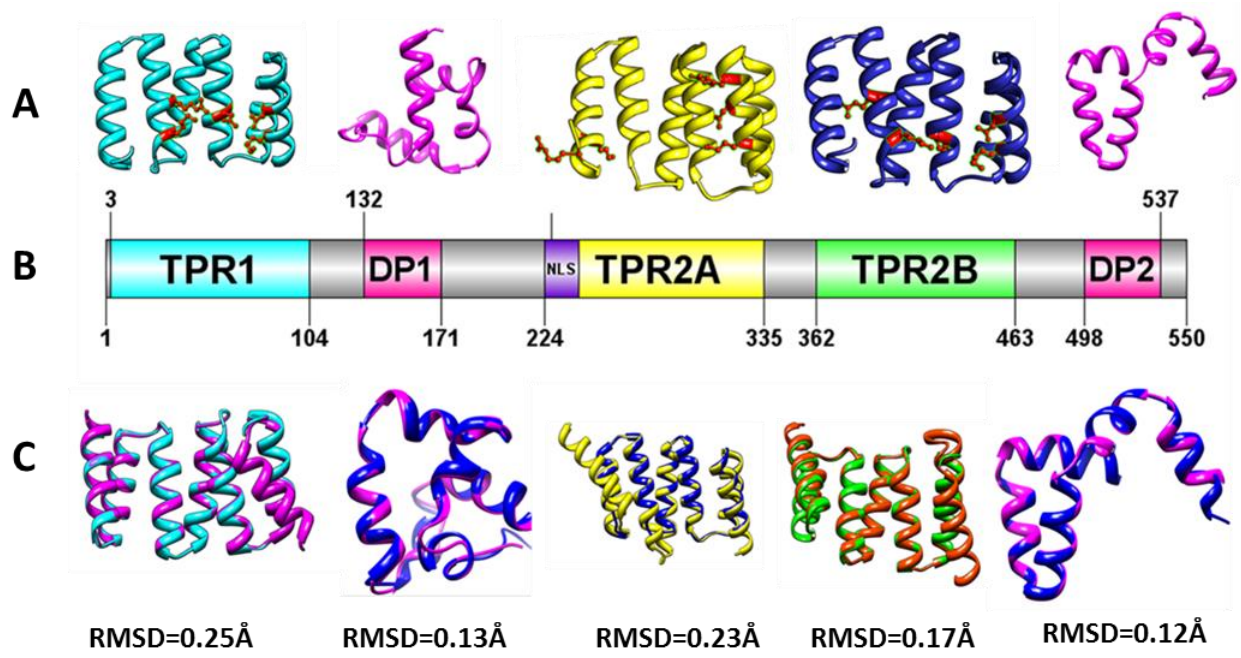


Figure 3. 7: Schematic representations and three dimensional models of TbSTI1 domains. (A) Ribbon presentation of the TPR domains indicating carboxylate clamp forming residues, (B) domains are labelled and highlighted and C) The models of TPR domains and DP regions superimposed against their templates (PDB codes: 1ELW,1ELR,3UPV, 2LLV and 2LLW)and associated RMSD scores.

3.3.5. TPR-containing co-chaperones of Hsp70 and Hsp83 in *T. brucei*

Little has been reported on the TPR containing co-chaperones of Hsp70 and Hsp83 in *T. brucei*, hence bioinformatics tools were used to predict the domain organization and properties of these co-chaperones. Only TbPPP5 has been characterized as a co-chaperone of Hsp83 (Jones et al., 2008). Thus, using the human homologues as queries on TriTrypDB, *T. brucei* TPR-containing co-chaperones of Hsp70and Hsp83 were retrieved and predicted using various programs to increase the possibility of identifying all motifs. *T. brucei* homologues of known TPR containing co-chaperones STI1, protein phosphatase 5 (PP5), immunophillins (CyP40 and FK506), SGT (small gluatamine-rich TPR containing protein) and J-proteins (section 2.3.5.4.5) were identified.

TPR domains mediate protein-protein interactions, they are considered functional when three or more TPR motifs are clustered together in a group (Lamb et al., 1995). It was seen that the number and positions of these motifs differed between proteins (Figure 3.8). TbSTI1, TbPP5, TbCyp40 and TbFK506 follow the same pattern of TPR motif organization as their human homologues (Blatch and Lassle, 1999; Carrello et al., 2004; Vaughat et al., 2008) while an additional TPR motif was predicted for SGT. Whether or not the additional TPR motif in SGT is functional remains the subject of future experimentations.

With regards to the TPR containing co-chaperones, common TPR motif clustering was observed in Tbj42, Tbj51, Tbj52 and Tbj67 while Tbj53 and Tbj65 seem to have a unique TPR motif pattern. The human homologue of Tbj53, DNAJC3, has been predicted to have 8 TPR motifs and thus two functional TPR domains or one incomplete domain (Tao et al., 2010, Svärd et al., 2011). This is consistent with the findings obtained for Tbj53, except it is predicted to have 7 TPR motifs and therefore likely to have 1 functional TPR domain or two incomplete TPR domains (Figure 3.8). In all TPR-containing J-proteins, the J-domain was present at the C-terminus, consistent with DNAJC3 and DNAJC7 (Tsai and Douglas, 1996; Kampinga and Craig, 2010). TbSTI1, TbHip and TbSGT also have the DP domain in common with each other (Figure 3.8). Four TPR-containing co-chaperones are predicted to interact with both Hsp70 and Hsp90 (TbSTI1, Tbj42, Tbj51 and Tbj52).

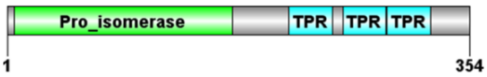
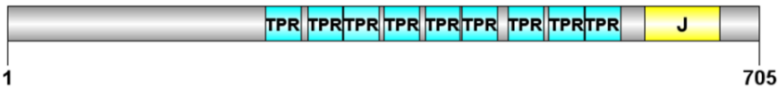
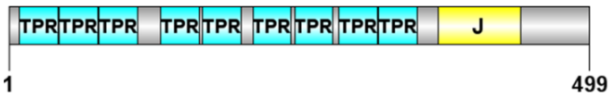
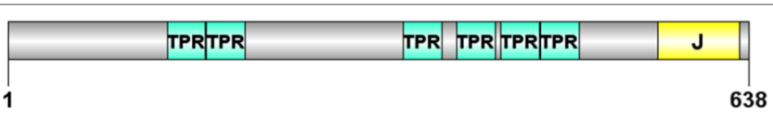
Binding Partner	Accession Number	Protein	Domain organization
Hsp90/ Hsp70	Tb927.5.2940	STI	
Hsp90	Tb927.10.13670	PPP5	
Hsp90	Tb927.9.9780	CyP40	
Hsp90	Tb927.10.16100	FK506	
Hsp70	Tb927.3.5340	Hip	
Hsp90	Tb927.6.4000	SGT	
Hsp90/ Hsp70	Tb927.10.12380	Tbj42	
Hsp90/ Hsp70	Tb927.4.2220	Tbj51	
Hsp90/ Hsp70	Tb927.10.4900	Tbj52	
Hsp70	Tb927.7.3630	Tbj53	
Hsp70	Tb927.4.880	Tbj65	
Hsp70	Tb927.10.5180	Tbj67	

Figure 3. 8: Schematic representation of the domain organizations and motifs for T. brucei TPR containing co-chaperones. TPR motifs were predicted using TPR pred (Karpenahalli et al., 2007) while other domains were predicted using a combination of programs including SMART (Letunic, et al., 2012) and Prosite (Sigrist, et al., 2010). Domain architecture was generated using DOG 2.0 software. Binding partners and TriTrypDB accession numbers are provided on the left, the protein domains are represented from position 1 (N-terminus) and the right indicates the C-terminus.

3.3.6. Sequence analysis of TPR-containing co-chaperones

As described previously, the presence of tandemly repeated TPR motifs structurally defines TPR domains. Although degenerate in nature, the loose 34-amino acid degenerate motif contains a largely conserved pattern of amino acid identity (Blatch and Lassle, 1999). There are 8 particular amino acids at positions 4 (W/L/F), 7 (L/I/M), 8 (G/A/S), 11 (Y/L/F), 20 (A/S/E), 24 (F/Y/L), 27 (A/S/L), and 32 (P/K/E) that have a higher frequency of conservation (D'Andrea and Regan, 2003). Thus, in addition to BLAST retrieval of protein sequences of TPR-containing co-chaperones in *T. brucei*, the occurrence of residues associated with TPR motifs further confirmed the identity of the co-chaperones. It has been observed that for functionally different TPR motifs, conservation of these amino acids is only limited to residues 8, 20, 24 and 27 (Blatch and Lassle, 1999). Furthermore, TPR domains of Hsp70 and/or Hsp90 co-chaperones are known to contain positively charged amino acids (K, N and R) which are usually involved in the formation of a carboxylate clamp with the C-terminal EEVD motif. The sequence analysis and identification of carboxylate clamp residues in *T. brucei* TPR containing co-chaperones of Hsp70 and Hsp83 are described for the first time in this study.

3.3.6.1. TPR1

Putative Hsp70 interacting co-chaperones should bind through the TPR1 domain. TPR1 domains of Tbj67, Tbj53, Tbj51, Tbj42 and TbHip (Figure 3.9) were aligned in relation to TbSTI1 using Promals3D and Boxshade. It was noted that TPR motif associated residues were conserved in all the aligned TPR1 domains (Figure 3.9). Unlike the TPR consensus sequence, carboxylate clamp forming residues were strongly conserved in TbSTI1, Tbj51 and Tbj52. The human homologue of Hip does not appear to have these residues conserved, and as expected neither did TbHip. Interestingly, the residues identified in mSTI1 to confer specificity in terms of Hsp70 binding are not conserved in all TPR-containing co-chaperones analyzed in this study (Figure 3.9). Despite the low conservation, it is likely that these co-chaperones will interact with Hsp70.

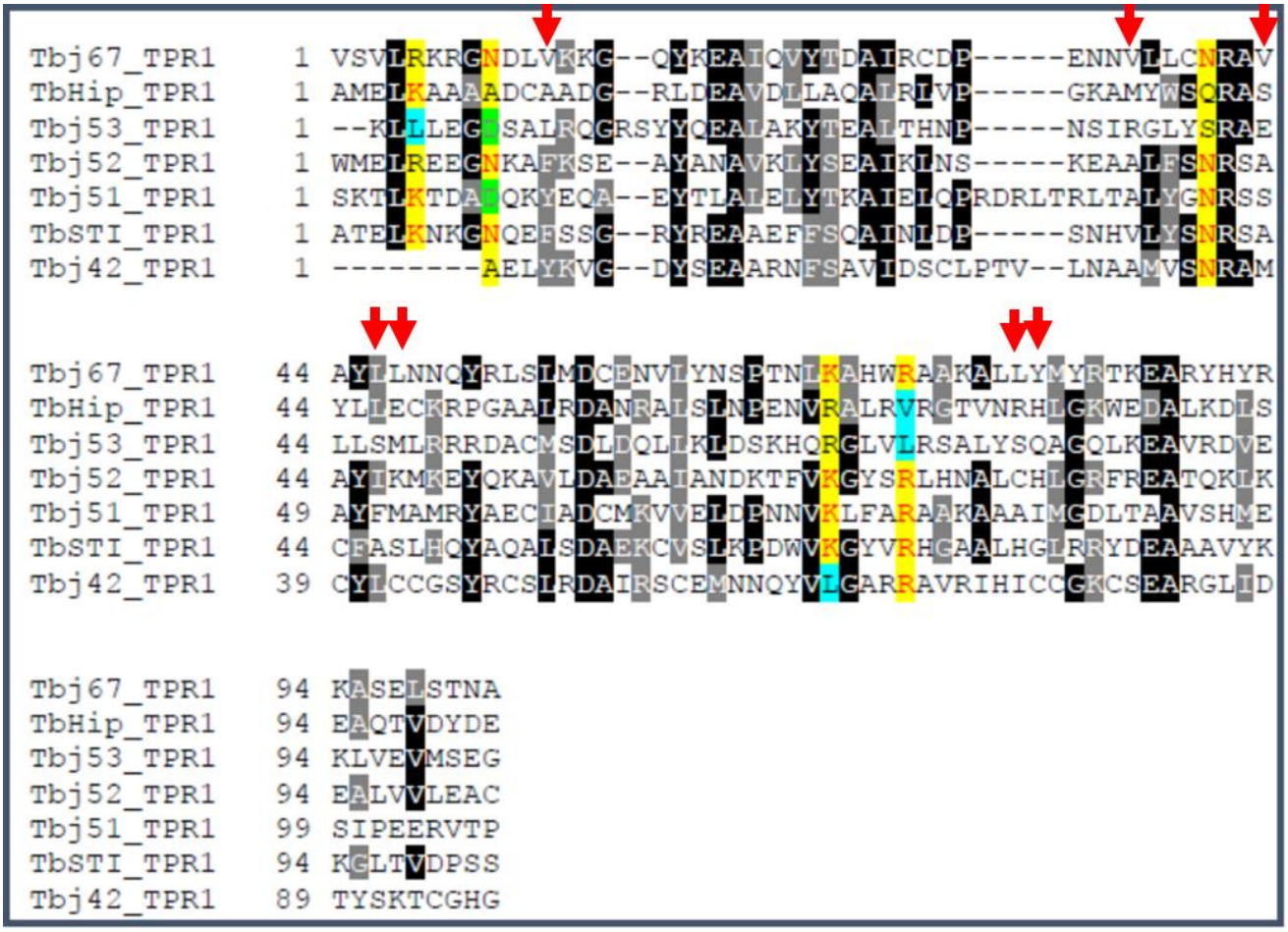


Figure 3. 9: Multiple sequence alignment of TPR1 domains of TPR containing co-chaperones of Hsp70. Amino acid sequences of TPR1 domains involved in the interaction with Hsp70 were isolated and aligned against TbSTI1. The alignment generated was analyzed using Promals3D and Boxshade. Yellow highlights the carboxylate clamp forming residues, conservation of residues is shown in red print or property (black print) while cyan highlights divergent amino acids. Residues involved in the specific interaction of mSTI1 and Hsp70 (Odunuga et al., 2003) are depicted using red arrows.

3.3.6.2. TPR2

The second TPR is the least conserved of all the TPR domains compared in the study. This domain is also referred to as TPR2A in STI1 proteins. TbPPP5, TbCyP40 and TbFK506 only have one putative Hsp90 binding TPR domain and it is referred to as TPR1. Similar to TPR1, a high level of divergence was seen in the individual TPR sequences while residues involved in Hsp90 binding were conserved in most of the co-chaperones.

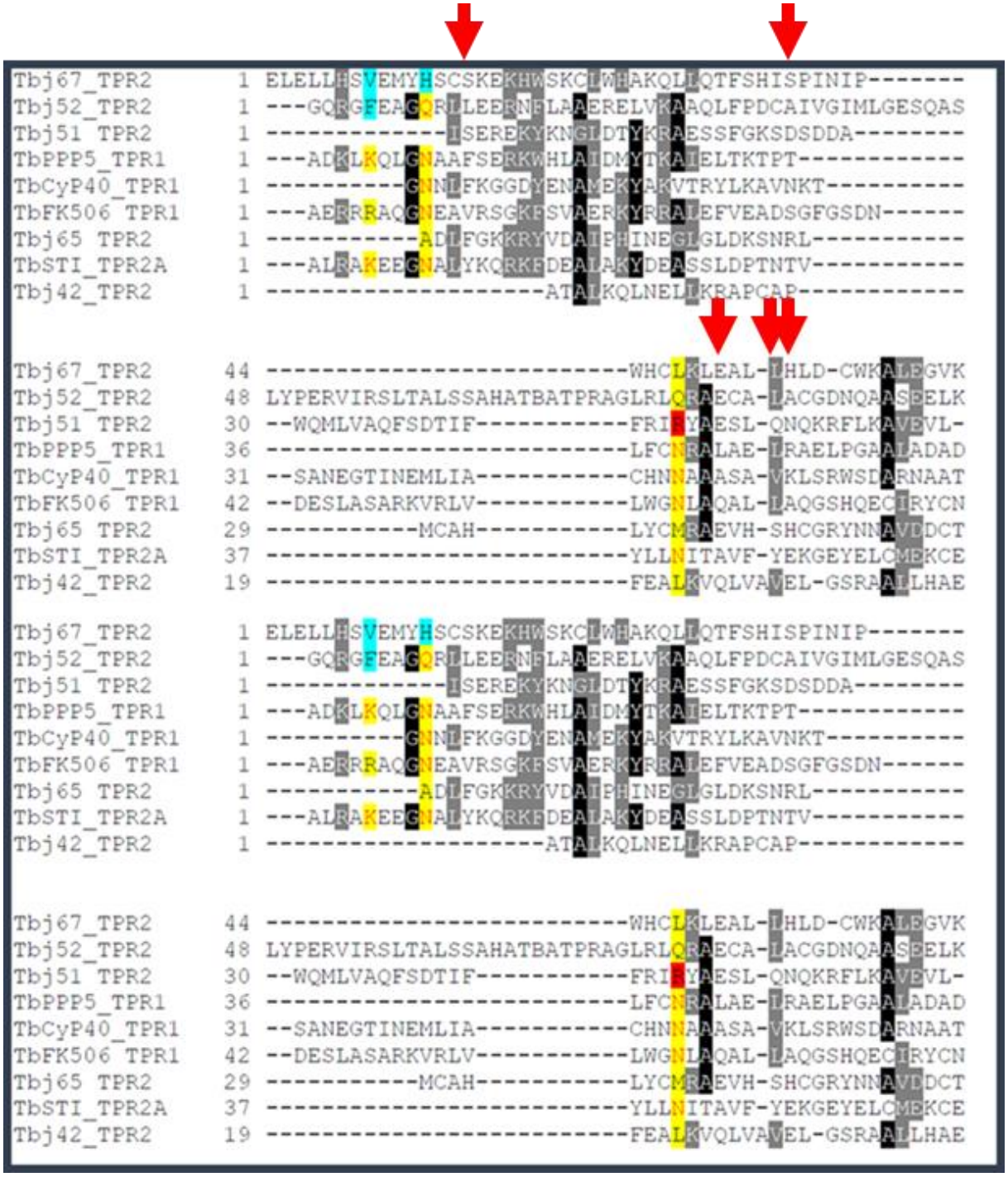


Figure 3. 10: Multiple sequence alignment of TPR2 domains of TPR containing co-chaperones of Hsp90. Amino acid sequence of TPR2 domains involved in the interaction with Hsp90 were isolated and aligned against TbSTI1. The alignment generated was analyzed using Promals3D and Boxshade. Yellow highlights the carboxylate clamp forming residues, a conservation of residues is shown in red print or property (black print) while cyan highlights divergent amino acids and red depicts substitutions of a carboxylate clamp forming residues to a charged residue. Residues involved in the specific interaction of mSTI1 and Hsp90 (Odunuga et al., 2003) are depicted using red arrows.

With regards to carboxylayte clamp forming residues, a high degree of conservation is seen in all the co-chaperones except for Tbj51, Tbj52 and Tbj67 (Figure 3.10). The conservation of these residues in Hsp90 co-chaperones suggest a conservation of the chaperone system. Five specific residues (YVAYE) identified in the TPR2A domain of mSTI1 to be involved in hydrophobic associations with Hsp90 (Odunuga et al., 2003) were found in TbSTI1 but not the other co-chaperones (Figure 3.10). Some of the *T. brucei* TPR co-chaperones have substituted the first residue TYR for another aromatic amino acid PHE, this is seen in TbCyP40 and TbPPP5. Interestingly, except for TbSTI1, most of the *T*. *brucei* TPR co-chaperones of Hsp90 have GLU instead of ALA, third residue in the five specific hydrophobic contact residues. It would be interesting to determine the effect of re-introducing the hydrophobic contact amino acids, especially considering the fact that this particular chaperone co-chaperone interaction is poorly characterized in *T. brucei*.

3.3.6.3. TPR3

TbSGT1 has an addition TPR motif to the TPR domain referred to as TPR2, while the single TPR domains of TbPPP5, TbCyP40 and TbFK506 are also similar to other TPR3 domains. TPR3 is also referred to as TPR2B in STI1 proteins. Unlike TPR2, TPR3 is highly conserved with almost all the carboxylate clamp forming residues conserved in all co-chaperones except for Tbj53 (Figure 3.11). The hexamer found in most STI1s (YSNRAA) (green box) is also present in almost all the TPR containing co-chaperones analyzed in the study, also the occurrence of DP regions are seen in the alignment (pink boxes, Figure 3.11). Until recently, TPR2B was not thought to be involved in Hsp90 binding hence residues contributing to hydrophobic interactions are yet to be identified.

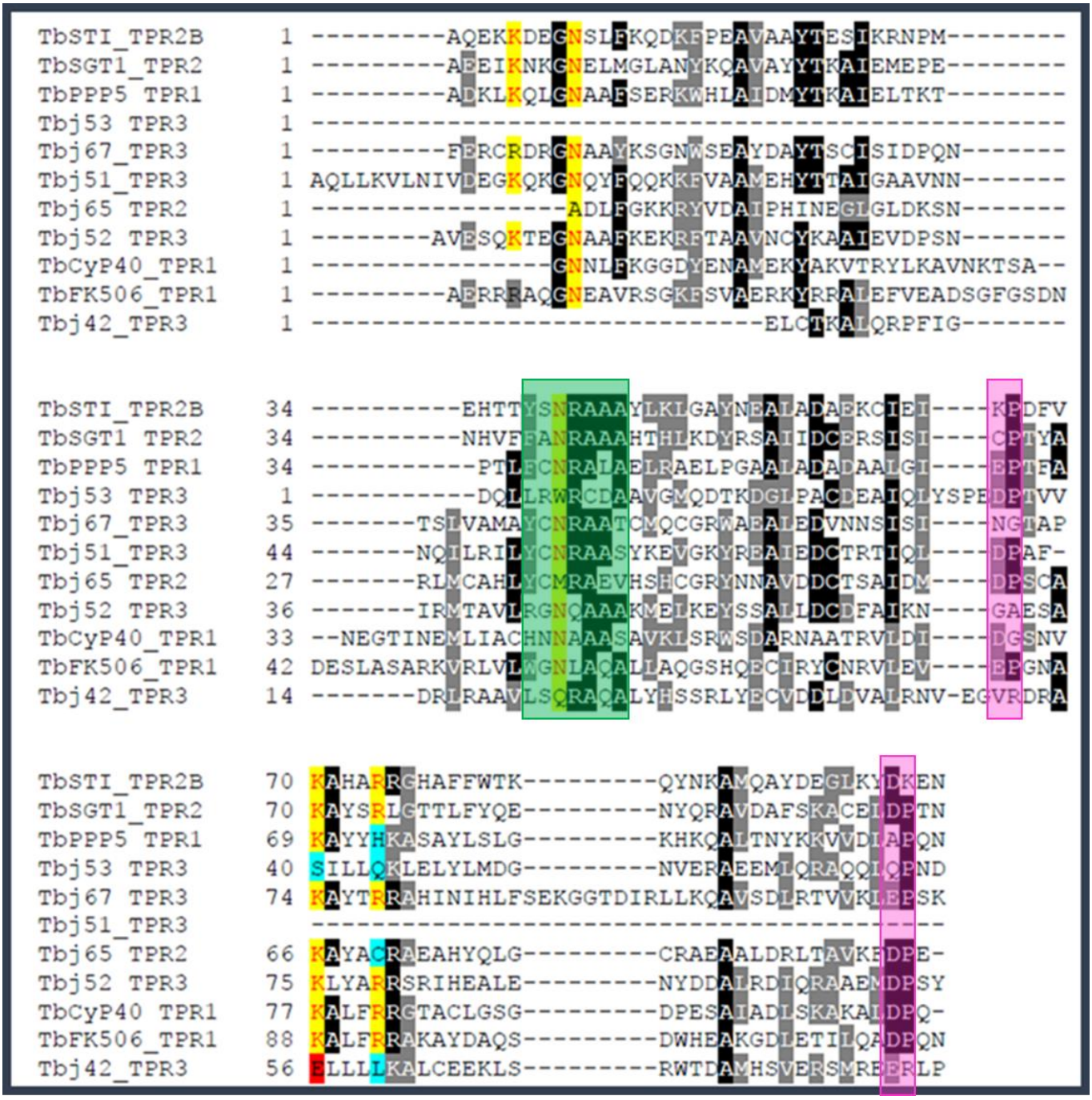


Figure 3. 11: Multiple sequence alignment of TPR3 domains of TPR containing co-chaperones of Hsp90. Amino acid sequence of TPR3 domains involved in the interaction with Hsp90 were isolated and aligned against TbSTI1. The alignment generated was analyzed using Promals3D and Boxshade. Yellow highlights the carboxylate clamp forming residues, a conservation of residues is shown in red print or property (black print) while cyan highlights divergent amino acids. Red depicts substitutions of N of the carboxylate clamp forming residues to a charged residue. Residues involved in the specific interaction of mSTI1 and Hsp90 (Odunuga et al., 2003) are depicted using red arrows. DP and STI1 hexamer shown in pink and green boxes respectively.

Finally, the alignment yields some interesting insights, the binding to Hsp90 by known co-chaperones is certainly conserved in *T. brucei*. This is supported by the conservation of carboxylate clamp forming residues. The TPR domain of known Hsp90 co-chaperones, TbPP5, TbCyP40, TbFK506 and TbSGT1 are indeed more closely related to TPR3 than TPR2 which is certainly more conserved. This observation reveals some interesting features particularly with regards to TPR containing J-proteins as they constitute the majority of the aligned proteins. From this, it can be deduced that TPR3 in most TPR containing J-proteins will bind Hsp90 with the exception of Tbj53.

3.4. Conclusion

This study sought to characterize the *T. brucei* multi-chaperone complex *in silico*. Hsp90 is an important, abundant chaperone that is being pursued as a drug target in other systems. Fundamentally, Hsp90 interacts with specialized client proteins such as kinases (Xu and Lindquist, 1993), steroid hormone receptors (Picard et al., 1990) and transcription factors (Minet et al., 1999). Bioinformatics data generated in this study revealed that of the 12 *HSP90* genes arranged in tandem clusters, 3 isoforms were identified in *T. brucei*. The mitochondrial and cytoplasmic isoforms were determined to be essential for the viability of the parasite (Alsford et al., 2011). The cytoplasmic isoform, TbHsp83, was shown to be a highly conserved member of the Hsp90 protein family, however, important unresolved questions remained to be answered regarding its client proteins as the *T. brucei* genome lack evidence of steroid receptors and kinases.

It is widely established that in order to function optimally or for correct regulation, most chaperones require the presence of a co-chaperone/s. Several co-chaperones of TriTryps heat shock proteins have been identified previously (Folguiera and Requena, 2007), however, only a few have been studied in detail. STI1 has been characterized in *T. cruzi* (Schmidt et al., 2011), *L.major* (Webb et al., 1997) and *L. donovani* (Morales et al., 2010) as a co-chaperone mediating

the interaction between Hsp70 and Hsp90. Also, SGT1 found in *L. donovani* has been defined as atypical (Ommen et al., 2010).

Therefore, using the human system as a guide, TPR containing co-chaperones in *T. brucei* were identified based on their proposed binding to Hsp70 and/or Hsp90. The central role of TbSTI1 in the assembly of the Hsp90 hetero-complex was highlighted in this study. Furthermore, TbSTI1 was demonstrated to be a highly conserved protein and orthologous to *T. cruzi* and *Leishmania* STI1 proteins which have both been demonstrated to display stage-specific formation of the STI1 mediated Hsp90 hetero-complex. It has also been shown that TbSTI1 contains two nuclear localization signals that may play role in the unique regulatory mechanism and interactions of *T. brucei* chaperones and co-chaperones.

Most of the TPR containing co-chaperones were revealed to be atypical, lacking the residues known to interact with the EEVD motifs of Hsp70 and Hsp90. Except for TbPP5, none of the TPR containing co-chaperones identified in this study have been characterized. Only the orthologue of TbSGT in *L. donovani* has been characterized. TcSGT was shown to interact directly or indirectly with LdSTI1, LdHip and LdHsp70. A weak association with LdHsp90 was suggested based on the strong evidence of co-localization using immunofluorescence (Ommen et al., 2010). TbSGT1 has been shown to be a typical Hsp70/Hsp90 interacting co-chaperone. The presence of immunophillins which bind Hsp90 following the release of STI1 was also demonstrated in *T. brucei*, these co-chaperones possess the same domain architecture as their mammalian orthologues.

In conclusion, the identification of conserved features in TPR containing co-chaperones of Hsp70 and Hsp90 suggest the presence of the multi-chaperone complex in *T. brucei*. The most divergent members of these co-chaperones were found to be TPR containing J-proteins; namely Tbj53, Tbj65 and Tbj67. The results obtained for Tbj53 are consistent with its human homologue,

DNAJC3, however the uniqueness of the two novel TPR containing J-proteins has interesting implication for parasite biology. It is proposed that future studies understanding the roles of these co-chaperone may reveal features relating to the multi chaperone complex in *T. brucei*.

CHAPTER FOUR

Characterization of TbSTI1 and its interactions with cytosolic Hsp90 and Hsp70s

4.0. Introduction

Protozoan parasites are exposed to vastly different environments in the mammalian hosts and insect vectors, and therefore undergo significant changes throughout their life-cycles. *T. brucei,* in particular, undergoes rapid and drastic changes wherein the parasite exhibits differences in the types of metabolism preferred in the vector and the host (Van der Ploeg et al., 1985). The *T. brucei* micro-organism losses its protective coat upon entering the tsetse vector and is often exposed to temperature, pH and respiration changes in the mammalian host (Van der Ploeg et al., 1985). The heat shock homeostatic mechanisms have been observed in *T. cruzi*, which synthesizes proteins in response to the marked environmental changes during the parasite life-cycle (Requena et al., 1992). Heat shock proteins are also required to assist the parasite when transitioning from the tsetse vector to the mammalian host whilst remaining viable and infective (Olson et al., 1994). This hypothesis is supported by the observation that *T. brucei* expresses enhanced levels of Hsp70 proteins when cultured in conditions similar to that of the mammalian host. The higher levels of Hsp70 expression is independent of temperature and could potentially be triggered by other stress-related influences (Van der Ploeg et al., 1985).

It has been noted that kinetoplastids, in particular *T. cruzi,* display stage-specific expression of molecular chaperones. Hsp70s are some of the major immunogens detected during *T. cruzi* infections. Their expression is 4-fold more pronounced when cultured in conditions resembling their mammalian host environment (Engman et al., 1990). Furthermore, TriTryps exhibit increased infectivity upon exposure to elevated temperatures (Smejkal et al., 1988); thus molecular chaperones could be related to parasite virulence. Hsp70 transcripts have been found to be more numerous in the BSF trypanosomes (within the mammalian host) than in the PRO trypanosomes (within the tsetse vector) (Lee et al., 1990). This observation suggests that Hsp70s are involved in the protection of the parasite against the stresses encountered upon entering the markedly different environment that its mammalian host presents after development within an insect vector.

Hsp90 appears in tandem repeats across all kinetoplastid genomes (Mottram et al., 1989). The *Leishmania* and *T. cruzi* orthologues of Hsp83 have been shown to play a critical role in temperature sensitive differentiation of the parasites during the transition from the insect vector into the mammalian host (Wiesgigl et al., 2001; Graefe et al., 2002). Hsp83 inhibition effectively inhibits the growth of several kinetoplastids, *T. cruzi* (Wiesgigl et al., 2001; Graefe et al., 2002), *Leishmania* (Li et al., 2009), *T. evansi* (Pallavi et al., 2010) and *T. brucei* (Jones et al., 2008; Alsford et al., 2011; Meyer and Shapiro, 2013) and in mouse models. In contrast to human Hsp90 which has been widely characterized, little is known about the role of Hsp83 in kinetoplastids, its role in cellular metabolism remains to be realized; as well its clients and many of its co-chaperones remain to be identified.

A review on approaches to isolate and characterize molecular chaperones has been published by Nicoll et al. in 2006. The purification of His-tagged Tbj1, Tcj1-2, TcHsp70 and TbHsp70.c has been successfully achieved (Edkins et al., 2004; Louw et al., 2010; Burger et al., 2014). The strategies employed for purification included denaturing methods using 8M urea and native methods, at times in the presence of the detergent sarcosyl. It has been reported that Hsp70 loses its functional integrity following treatment with urea as a result of insufficient refolding; the protein's native state is not always achieved subsequent to denaturation (Kathir *et al.,* 2005). A native nickel-affinity purification protocol for His-tagged Tcj2 and TcHsp70 using competitive binding in the presence of high concentration of imidazole has been described (Louw et al., 2010, Burger et al., 2014). Also the purification of TbHsp83 was also recently achieved using anion exchange and nickel affinity chromatography (Pizarro et al., 2013).

Although STI1 homologues have been identified in other parasitic organisms, very little biochemical characterization has been carried out on this protein in kinetoplastids. The presence of an Hsp70/Hsp90 complex mediated by STI1 has been demonstrated in *P. falciparum,* while no studies have been carried out for *T. brucei*. The presence of TbSTI1 has been reported previously (Louw et al., 2010). However, except for RNAi knockdown studies showing the importance of

TbSTI1 in parasite survival and viability at the BSF stage of parasite development (Alsford et al., 2011), no other characterization has been done. TbSTI1 has been predicted to have a molecular mass of 63 kDa and localize in the cytoplasm (Chapter 3). Furthermore, the homology of TbSTI1 to other eukaryotic STI1 proteins has been proposed, the conservation of carboxylate clamp residues and Hsp70/Hsp90 specific interacting amino acids (hydrophobic contact residues) has been shown. The bioinformatics studies carried out in chapter 3, identified LmSTI1, LdSTI1 and TcSTI1 as orthologues of TbSTI1. LmSTI1 was first described in 1997 as a highly conserved constitutive protein whose mRNA levels were upregulated following a temperature shift from 26°C to 37°C at the promastigote stage of parasite development. Also, stage-specific formation of the heat shock protein complex mediated by STI1 has been described for *L. donovani* (Morales et al., 2010). The study of LdSTI1 also identified phosphorylation sites on this protein and demonstrated that two sites were essential for parasite viability. Furthermore LdSTI1 was demonstrated to associate with ribosomal client proteins in an amastigote-specific manner (Morales et al., 2010). TcSTI1 on the other hand, was shown to co-localize with TcHsp70 and its protein levels remained constant upon induction, enhanced production of this protein was observed when the parasite was subjected to nutritional stress at the epimastoge phase of parasite development (Schmidt et al., 2011).

This study aimed to overproduce and characterize TbSTI1, using the well characterized murine STI1 (mSTI1) as a control. The interactions of TbSTI1 with cytosolic Hsp70 and Hsp90 were also investigated. The heterologous expression and purification of cytosolic Hsp70s was carried out in the *E. coli dnaK* minus strain BB1994 to avoid the effect of contaminating DnaK (prokaryotic Hsp70). TcHsp70 was used as a positive control for Hsp70s due to its canonical features and cytoplasmic localization (Edkins et al., 2004; Louw et al., 2009; Schmidt et al., 2011).

4.1. Objectives

Based on bioinformatics data, TbSTI1 has the features of a typical eukaryotic STI protein, however the protein has never been biochemically characterized, and this study aimed to overproduce and characterize TbSTI1. The ability of TbSTI1 to interact with cytosolic Hsp70s and Hsp90 was also investigated. The presence of the chaperone proteins used in this study, during the bloodstream stage of the lifecycle of *T. brucei*, was probed under permissive and heat shock temperatures.

The specific objectives were:

1. To overproduce and purify His-tagged TbSTI1 and mSTI1, predicted cytosolic TbHsp70.4, TcHsp70 and TbHsp70.c and TbHsp83.1 proteins.
2. To design and test peptide antibodies against TbSTI1 and TbHsp83.1.
3. To determine whether TbSTI1 binds to cytosolic *T. brucei* Hsp70s
4. To determine whether mSTI1 and TbSTI1 can bind human Hsp70 and Hsp90.
5. To detect TbSTI1, cytosolic Hsp70s and Hsp83 in the *T. brucei* parasite lysate and investigate heat inducible expression.

4.2. Materials and methods

4.2.1. Materials

Reagents used were purchased from Sigma Chemicals Co (St. Louis, Mo U.S.A.), Merck Chemicals (Darmstadt, Germany), BioRad (U.S.A.) or Roche Molecular Biochemicals (Indianapolis, IN, U.S.A.). Nickel beads were purchased from Pharmacia Biotech (Uppsala, Sweden). Anti-His antibody was purchased from Amersham Pharmacia Biotech (UK). Anti-TbHsp83 and anti-TbSTI1 were produced and purchased from GenScript (U.S.A). The anti-TbHsp70.c and pQE80-TbHsp70.c were generous gifts from Dr Adelle Burger. The *T. brucei TREU927* strain was a kind donation from Professor George Cross (Rockefeller University, New York, USA). *E. coli* strain BB1994 was kindly provided by Dr. M Mayer (Heidelberg University, Heidelberg, Germany). All the details pertaining to plasmids used in the study are summarized in Table 4.1 below. The TbHsp70.4 antibody was a kind gift from Professor Bangs (University of Winsconsin-Madison, Madison, USA).

Table 4. 1: Properties of plasmids used in this study

Protein	Plasmid	Organism	Plasmid source /Reference
TcHsp70	pET14b-TcHsp70	*T. cruzi*	Edkins et al., 2004
TbHsp70.4	pQE2-TbHsp70.4	*T. brucei*	Dr Michael Ludewig
TbHsp70.c	pQE80-TbHsp70.c	*T. brucei*	Burger et al., 2014
TbHsp83	pQE80-TbHsp83 N-terminal His-tag	*T. brucei*	Dr Michael Ludewig
TbHsp83	pQE60-TbHsp83 C-terminal His-tag	*T. brucei*	Dr Michael Ludewig
TbSTI1	pQE2-TbSTI1	*T. brucei*	Dr Michael Ludewig
mSTI1	pQE2000	*M. musculus*	Van der Spuy et al., 2000

4.2.2. Heterologous production and purification of recombinant TbSTI1 and mSTI1

4.2.2.1. Induction studies for the production of TbSTI1 and mSTI1

The integrity of the pQE2000 plasmid expressing recombinant mSTI1 and pQE2-TbSTI1 for the production of TbSTI1 were confirmed by restriction analysis (Appendix E5) and DNA sequencing (Appendix E6). Subsequent to plasmid verification, competent *E. coli* M15[pREP4] cells transformed with pQE2-TbSTI1 were selected by plating on to 2xYT agar plates containing ampicillin sodium salt (100 µg/ml) and kanamycin sulfate (50 µg/ml) (Appendix E2). A single colony was picked from the YT agar plate and inoculated into 25 ml of YT broth containing 100 µg/ml ampicillin and 50 µg/ml kanamycin for growth at 37°C in a shaking incubator. Cells grown overnight were diluted 10x and incubated at 37°C in a shaking incubator until mid-log phase (A_{600} =0.5-0.6) was reached. At mid log phase, a pre-induction sample was collected and protein production was induced by the addition of 1 mM IPTG. Post induction samples were collected at 1, 2, 3, 4, 5 and 16 hours intervals followed by the determination of A_{600} of the culture at each time point. Each 1 ml sample was harvested by centrifugation (13 000 xg for 2 minutes). The pellets were resuspended in the appropriate volume of PBS buffer, where 150 µl PBS was added for each 0.5 OD_{600nm} absorbance unit. Protein production was determined by SDS-PAGE and western analyses. The detection of TbSTI1 was carried out using mouse anti-His antibody and rabbit anti-TbSTI1 polyclonal antibody (1: 1000, see section 4.2.5.1 for details). The mSTI1 protein overproduction in *E. coli* M15[pREP4] cells was carried out in a similar manner to TbSTI1. Protein production of recombinant TbSTI1 was also induced in *E. coli* XL1-Blue and *E. coli* BB1994 strains.

4.2.2.2. Purification of TbSTI1 and mSTI1

Protein expression of TbSTI1 and mSTI1 was induced using 1 mM IPTG and cells were harvested five and four hours respectively post protein induction. The cells were harvested through centrifugation (5000 xg, 20 min) at 4°C and the pellet obtained was resuspended in 2.5 ml lysis

buffer (10 mM Tris, pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM PMSF), and 1 mg/ml lysozyme. The cells were then frozen at -80°C overnight. The fractions were thawed and cell lysis was achieved by mild sonication at amplitude setting of 50 for seven cycles with 15 seconds pulse and 5 second pause after each cycle. The cell debris were cleared by centrifugation (13 000 xg, 40 min) at 4°C.

The soluble cell extracts were bound to nickel sepharose beads overnight at 4°C. The bead suspension was centrifuged (5000 xg, 1 min) and the beads were washed three times using wash buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 50 mM imidazole and 1 mM PMSF). Bound mSTI1 and TbSTI1 were eluted using elution buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 1 M imidazole and 1 mM PMSF). mSTI1 was dialyzed at 4°C against the storage buffer (100 mM Tris, pH 7.5, 300 mM NaCl, 1 mM DTT, 10% (v/v) glycerol and 1 mM PMSF) and subsequently stored at -80°C for future use. The TbSTI1 was passed through a Sephacryl S-100 column (exclusion limit of 100 kDa) (Amersham Pharmacia Biotech) which was equilibrated using the TbSTI1 dialysis buffer (10 mM Tris–HCl, pH 7.5, 300 mM NaCl, 1 mM DTT and 10% (v/v) glycerol) to remove lower molecular mass contaminating proteins. An alternative approach to removing the contaminating protein was to concentrate the elutions using a vivaspin column (with a membrane cut-off of 10 kDa) (Sartorius Stedim Biotech GmbH, Germany). Purifications of mSTI1 and TbSTI1 were analyzed by SDS-PAGE and western blot using anti-His antibody.

4.2.3. Heterologous production and purification of recombinant TcHsp70, TbHsp70.4, TbHsp70.c and TbHsp83

The plasmid integrities of pET14b-TcHsp70, pQE2-TbHsp70.4 and pQE80-TbHsp70.c were confirmed as described under restriction analysis (Appendix A5) and DNA sequencing (Appendix A6).

4.2.3.1. Heterologous purification TcHsp70

The pET14b-TcHsp70 was transformed into *E. coli* BL21 cells and the production of protein was monitored as described above in section 4.2.2.2, subsequent analysis were carried out by SDS-PAGE and western blot analysis. To purify recombinant TcHsp70, the protein was harvested by centrifugation (5000 xg, 20 min) after 3 hours post-induction growth in *E. coli* BL21 cells. After cell lysis, the soluble TcHsp70 fraction was obtained by centrifugation (13 000 xg, 30 min) at 4°C and allowed to bind to nickel-charged sepharose beads (GE Healthcare, UK) overnight at 4°C with mild agitation. Unbound proteins were removed by centrifugation (5000 xg, 1 min) followed by three washes using wash buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 50 mM Imidazole and 1mM PMSF). Bound TcHsp70 was eluted using the elution buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 500 mM imidazole and 1mM PMSF).

4.2.3.2. Heterologous purification TbHsp70.4 and TbHsp70.c

The overproduction of TbHsp70.4 and TbHsp70.c proteins was carried out as described in section 4.2.2.2 except the cells were grown at 30°C due to the thermal sensitivity of *E. coli* BB1994 cells. Western detection was carried out rabbit anti-TbHsp70.c (1: 5000) and TbHsp70.4 (1: 2500) polyclonal antibodies. HRP-conjugated goat anti-rabbit (1: 5000) (Cell signaling technology, USA) secondary antibody and chemiluminescence-based detection was used for western analysis. To purify recombinant TbHsp70.4, the protein was overproduced for five hours in *E. coli* B1994 cells and harvested by centrifugation (5000 xg, 20 min) at 4°C and the pellet re-suspended in lysis buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 10 mM imidazole, 1 mM PMSF and 1 mg/ml lysozyme. After cell lysis, the soluble TbHsp70.4 was obtained by centrifugation (13 000 xg, 30 min) at 4°C and allowed to bind to sepharose beads overnight at 4°C with mild agitation. Unbound proteins removed by centrifugation (5000 xg, 1 min) followed by three washes using wash buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 50 mM Imidazole and 1 mM PMSF). Bound TbHsp70.4 was eluted using the elution buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 1 M imidazole and 1 mM PMSF).

The same protocol to that of TbHsp70.4 was adopted for the purification of TbHsp70.c. The cells were harvested by centrifugation after 3 hours post-induction growth as described above and the proteins were eluted with elution buffer (10 mM Tris, pH 7.5, 300 mM NaCl, 500 mM imidazole and 1mM PMSF). The proteins were dialysed against the storage buffer (100 mM Tris, pH 7.5, 300 mM NaCl, 1 mM dithiothreitol (DTT), 10% (v/v) glycerol and 1mM PMSF) and frozen at -80^{o}C for future applications.

4.2.4. Heterologous production and purification of recombinant TbHsp83

To purify TbHsp83 (see Appendix A5 for restriction digestion protocol) induction studies as described in section 4.2.2.1 were conducted and subsequent analyses were carried out by SDS-PAGE and western blot analysis. The cells were harvested two hours post-induction using 1 mM IPTG, the purification of TbHsp83 was carried out in a similar manner to TcHsp70 described in section 4.2.3.1 above. Detection of the His-tag was carried out using mouse anti-His antibody and detection of TbHsp83 was carried out using rabbit anti-TbHsp83 polyclonal antibody (1: 1000, see section 4.2.5.2 below for details).

4.2.5. Design and production of peptide directed anti-TbSTI1 and anti-TbHsp83 antibodies

4.2.5.1. Peptide directed anti-TbSTI1 antibody

To identify a unique region on TbSTI1 for antibody production, its amino acid sequence (accession number: Tb927.5.2940) was retrieved from TriTrypDB version 8 (Aslett et al., 2010). A peptide region comprising of 14 amino acids (CKPEAPKKNEEPKK) corresponding to residues 209-222 within the N-terminal of TbSTI1, a few residues before TPR2A, was identified from alignments

using kinetoplastids and eukaryotic proteins. The identified peptide region was queried against local databases to avoid potential cross-reactivity with human and other kinetoplastid proteins.

The protein sequence of TbSTI1 was also subjected to various algorithmic searches for the determination of peptide antigenicity (Hopp and Woods, 1981; Hofmann and Hadge, 1987 andJameson and Wolf, 1988), surface probability(Emini, et al., 1985), hydropathy (Kyte and Doolittle, 1982) and chain flexibility(Karplus and Schulz, 1985) provided on GeneRunner software (version 3.05; Hastings Software Inc) and ProtScale (Http://expasy.org/tools/protscale.html) (Gasteiger et al., 2005). The epitopic region selected (Figure 2.1) was synthesized by GenScript Corporation (Piscataway, New Jersey, USA) using the SC1180 complete affinity-purified polyclonal antibody package. The GenScript Corporation synthesized the peptide and raised antibodies in rabbit after the collection of the pre-immune serum and Anti-TbSTI1 peptide antibody was purified to a yield of 0.847mg/ml. The antibody was subsequently used for TbSTI1 detection from bacterial and *T. brucei* cell lysates.

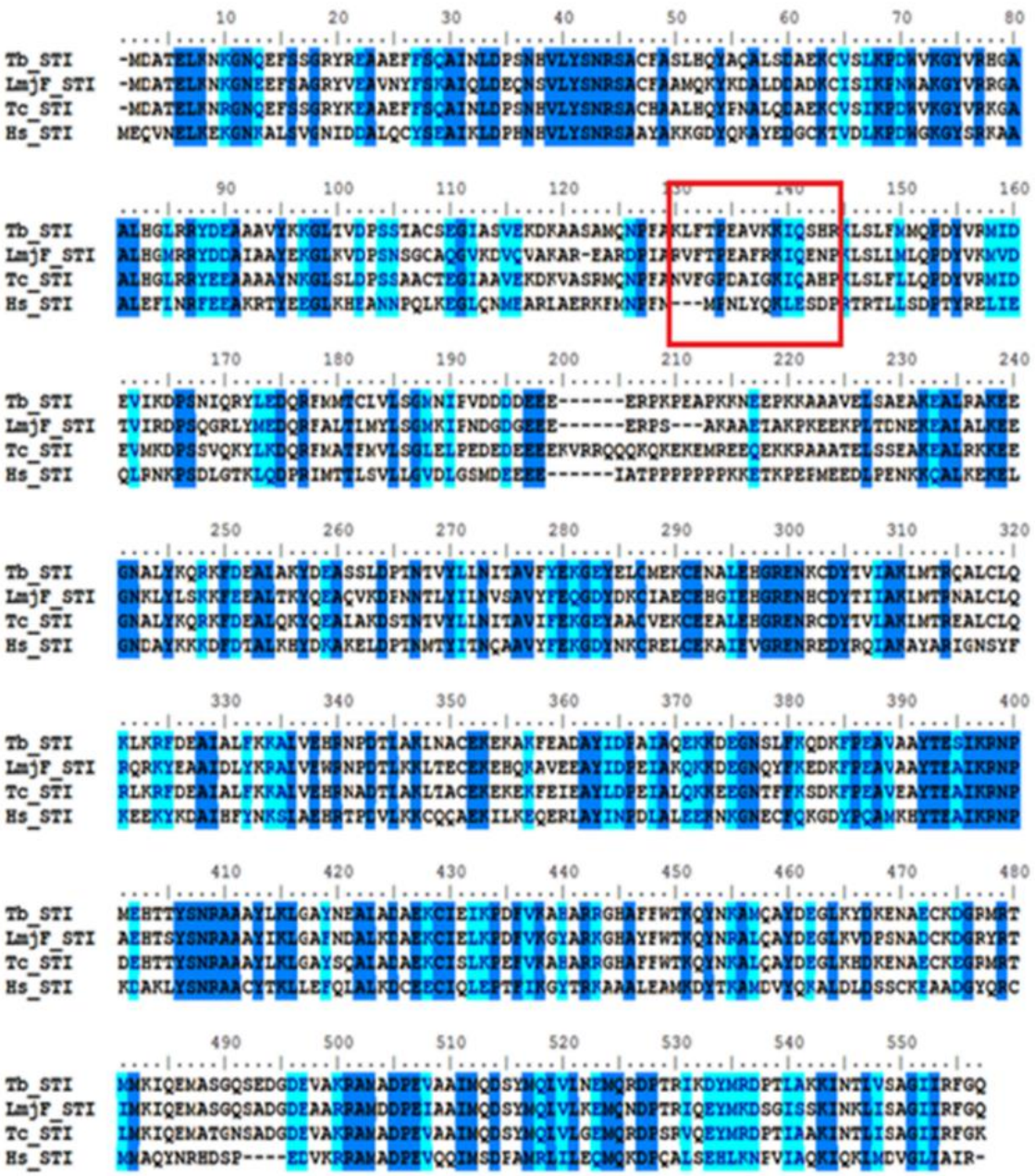


Figure 4.1: Sequence alignment of TriTryps and human STI proteins depicting high level of sequence conservation with all similar residues shown in blue highlights, cyan highlights depicting identical residues and black letters showing variable amino acid sequences. Regions predicted to be exposed to the surface, soluble, possessing high peptide chain flexibility, hydrophilic amino acids stretches and regions of high charge density were assessed using the GeneRunner software (version 3.05; Hasting Software Inc). The peptide region selected (red box) was quered against the NCBI and TriTrpDB blast databases and no hits were obtained.

4.2.5.2. Peptide directed anti-TbHsp83 antibody

Based on the peptide antibody design requirements described in section 4.2.5.2, the epitope region selected for anti-TbHsp83 antibody was designed. Briefly, the amino acid sequence of TbHsp83 (accession number: Tb927.10.10980) was retrieved from TriTrypDB version 8 (Aslett et al., 2010). A peptide region comprising of 14 amino acids (DEPHLRIRVIPDRV) corresponding to residue 57-71 within the slightly varied region in the N-terminal domain ATPase of Hsp90 proteins (Figure 3.1) was selected as the epitopic region. The anti-TbHsp83 antibody was synthesized by GenScript Corporation (Piscataway, New Jersey, USA) using the SC1180 complete affinity-purified polyclonal antibody package. The GenScript Corporation synthesized the peptide and raised antibodies in rabbit after the collection of the pre-immune serum. The anti-TbHsp83 peptide antibody was purified to a yield of 0.41 mg/ml. The antibody was subsequently used for TbHsp83 detection from bacterial and *T. brucei* cell lysates.

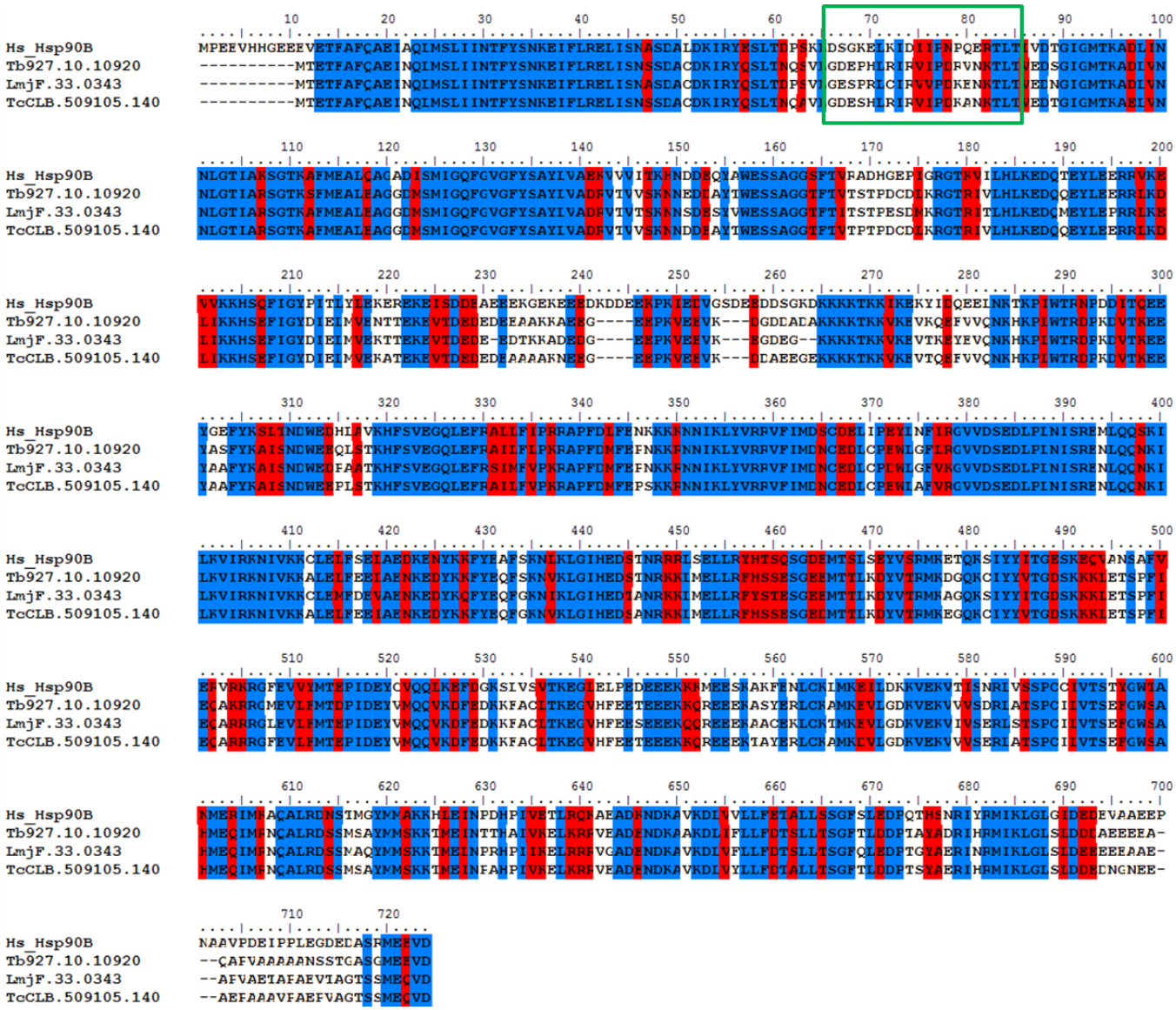


Figure 4.2: Sequence alignment of TriTryps and human Hsp90 proteins depicting high level of sequence conservation, all similar residues shown in blue background, red background depicting identical residues and black letters showing variable amino acid sequences. Regions predicted to be exposed to the surface, soluble, possessing high peptide chain flexibility, hydrophilic amino acids stretches and regions of high charge density were assessed using the GeneRunner software (version 3.05; Hasting Software Inc). The peptide region selected (green box) was queried against the NCBI and TriTrpDB blast databases and no hits were obtained.

4.2.5.3. Dot blots for testing of antibodies

The anti-TbSTI1 antibodies were tested using dot blots with chemiluminescence–based immunodetection for visualization. Recombinant TbSTI1, mSTI1, TbHsp70.4 and TcHsp70 were purified as described previously.Varying concentrations of TbSTI1 and mSTI1 (2µg, 5 µg, 10 µg and 20 µg) were spotted onto 3 Hybond^{TM}–C Extra nitrocellulose membrane in a Bio–Dot ST (BioRad, UK) apparatus connected to a vacuum pump. The mSTI1, TcHsp70 (20 µg) and TbHsp70.4 (20 µg) were used as negative controls. The membranes were blocked with 5% non-fat powdered milk in TBS (50 mM Tris pH 7.5, 150 mM NaCl,) for one hour at room temperature. The membranes were subsequently separately incubated in mouse anti-His antibody (1: 5000), rabbit polyclonal anti-TbSTI1 (1:1000) serum and pre-immune serum for two hours at 4°C. The membranes were washed twice with TBST buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1% Tween–20 (v/v) before incubation with appropriate secondary antibodies, HRP–conjugated goat anti–mouse antibody (1:5000 dilution) and HRP–conjugated goat anti–rabbit antibody (1:5000 dilution) for one hour. The membranes were washed twice with TBST for 15 minutes before incubation with chemiluminescence reagents (ECL, Amersham, UK), and detection of the signal using chemiluminescence–based protein detection. Chemiluminescence–based protein detection was achieved using the ECL^{TM}western blotting kit (GE Healthcare, UK) as per the manufacturer's instructions, and captured with a Chemidoc chemiluminescence imaging system (BioRad, UK).

4.2.6. Presence and heat inducibility of molecular chaperones in *T. brucei* lysate

Whole cell lysates prepared from two strains of *T. b. brucei 427* v.221 and *T.b. brucei 427* v.3 at 37°C and 42°C were kindly donated by Dr Adelle Burger and Dr Michael Ludewig (Rhodes University, Grahamstown, South Africa). The lysates were resuspended in SDS-PAGE loading buffer to a final cell count of 5 x 10^5cells/µl. A total of 5 x 10^6 cells (correlating to approximately

20µl aliquots) were resolved on 10% SDS-PAGE gels and transferred to a nitrocellulose membrane for western analysis. The blots were subsequently incubated with rabbit polyclonal anti-TbSTI1 (1:1000), anti-TbHsp70.4 (1:1000), rabbit polyclonal anti-TbHsp70.c (1:1000) and rabbit polyclonal anti-TbHsp83 (1:1000) primary antibodies and HRP-conjugated goat anti-rabbit (1: 5000) secondary antibody for chemilumiscence-based protein detection.

4.2.7. Biochemical interaction studies with recombinant TbSTI1 and mSTI1

4.2.7.1. Interaction of TbSTI1 by surface plasmon resonance spectroscopy

SPR was used to monitor the biomolecular interactions of TbSTI1 with trypanosomal Hsp70s (TcHsp70, TbHsp70.4 and TbHsp70.4) and Human Hsp90β (cat no.: SPR-102) (Stressmarq, Canada). The kinetic interactions of these chaperone - co-chaperone interactions were carried using the ProteOn ™ XPR36 (Bio-Rad, USA) optical biosensor. All experiments were performed at 25°C using a GLC biosensor chip, with a 2D surface monolayer for protein-protein interaction and < 8000 RU ligand binding density capacity. PBS, pH 7.4 was used as the running buffer and ProteOn Kits (amine coupling, deactivation and post maintenance kits) were obtained from Bio-Rad.

Due to its ability to immobilize His-tagged proteins for protein-protein interactions, the GLC sensor chip was selected (Bronner et al., 2006). Briefly, the GLC sensor chip was initialized using 50% glycerol, the running buffer was set at a 30µl/minute flow rate for 120 seconds while the regeneration and ligand injection steps were maintained at default parameters. Subsequent injections of 0.5% (w/v) SDS and 100 mM HCl were carried out to pre-condition the GLC chip. In order to allow for covalent immobilization of the ligand on the surface of the sensor chip, a 1:1 mixture of 40 mM EDC (N-ethyl-N-(dimethylaminopropyl) carbodiimide) and 10 mM sulfo-NHS (N-hydroxysuccinimide) diluted to 1:5 in degassed water was injected at 30µl/minute flow rate

for 60 seconds. Following sensor chip surface activation, a pre-concentration test to determine the optimum pH and concentration for each ligand to be immobilized was determined based on the protein isoelectric point (pI). The 6 ligands were prepared in 10 mM sodium acetate buffer and were immobilized onto the GLC sensor chip surface and docked into the ProteOn XPR36 instrument as indicated in Table 4.2. The tests were performed by injecting different concentrations of TbSTI1 in a perpendicular orientation at a flow rate of 30 µl/min for 5 minutes. The reference channel (channel 4) was injected with buffer alone and hence no ligand was immobilized on that channel. The channels were subsequently regenerated using 10 mM glycine pH 1.5.

Table 4. 2: A summary of the 6 ligands immobilized on the GLC sensor chip

Channel	Protein	Concentration	pI	pH
L1	TcHsp70	10µg/ml	5.22	4.5
L2	TbHsp70.4	10µg/ml	4.69	4.0
L3	TbHsp70.c	10µg/ml	4.86	4.0
L4	Buffer	10µg/ml		
L5	Hsp90β	10µg/ml	4.89	4.0
L6	TbSTI1	10µg/ml	6.23	5.0

A procedure for washing the chip until the base line was reached was attained through inactivating the remaining active carboxyl groups using 1 M ethanolamine at a flow rate of 30 µl/min for 300 seconds and subjecting the sensor chip to more washes with PBS. Subsequent to the wash steps, the protocol for interactions involving low rate, contact time, dissociation time and regeneration were set-up on ProteOn XPR36 software. The analyte (TbSTI1) was injected at 0.2 µM, 0.4 µM, 0.6 µM, 0.8 µM, and 1 µM concentration at a flow rate of 60 µl/min for 90 seconds, deactivation was subsequently allowed to occur 5 minutes using the assay buffer. The different concentrations of TbSTI1 would allow the study to determine whether the interaction

is concentration dependent which would be useful for kinetic analysis. The reactions were carried out in triplicate on the same channel and double referencing by row or column was adopted for all experiments. Data was analyzed using the *ProteOn Manager™ software* version 3.1.0.6 (Bio-Rad, USA).

4.2.7.2. Pull down assays

4.2.7.2.1. Pull down assay using purified proteins

Reactions of anti-rabbit TbSTI1 polyclonal antibody (20 µl) bound to 20 µl Protein A/G agarose (Santa Cruz, USA) in 1 ml of binding buffer (20 mM HEPES pH 7.2, 50 mM KCl, 5 mM $MgCl_2$, 20 mM Na_2MO_4, 0.5 % (v/v) NP40, 1 mM ATP) at 25°C for 2 hours were set-up. The reaction was centrifuged (5000 xg, 5 minutes) and was subsequently incubated with 20 µg of purified 6 x His tagged TbSTI1. The reaction was allowed to bind at 4°C for 2 hours and following centrifugation (5000 xg, 5 minutes) a total of 20 µg His-tagged purified chaperones (TcHsp70, TbHsp70.4 and TbHsp70.c) were separately added to the reactions which were allowed to bind in binding buffer overnight at 4°C with mild agitation. The reactions were washed three times in binding buffer before being analyzed by SDS-PAGE and anti-His western analysis. Control reactions with protein A/G agarose and purified proteins in the absence of anti-TbSTI1 were carried out in tandem.

4.2.7.2.2. Pull down assay using human cell lysate

A cell lysate from the human breast cancer cell line Hs578T was a kind donation from Dr Lorraine Zvichapera Mutsvunguma (Rhodes University, Grahamstown, South Africa). Purified 6x His tagged TbSTI1 and mSTI1 (20 µg) were prepared in 1 ml PBS. The co-chaperones were incubated with 100 µl of pre-washed nickel-charged sapharose beads at 4°C for 2 hours, with mild agitation. A control of pre-washed nickel-charged sepharose beads (100 µl) in 1 ml PBS was used. The reactions were then washed three times in PBS and centrifuged (5000 xg, 5 minutes). The bound purified proteins were incubated with 500 µl of Hs578T cell lysates overnight at 4°C with mild

agitation. The reaction was washed three times with PBS, centrifuged (5000 xg, 5 minutes) and the pellet was resuspended in 5 x SDS-PAGE sample buffer prior to western analysis for Hsp90 and Hsp70.

5.2.7.3. Interaction of TbSTI1 using dot blot assay

Recombinant TbSTI1, TbHsp70.4, TbHsp70.c and TcHsp70 were purified, dialysed, concentrated and quantified. BSA (50 µg) was used as a negative control while TbSTI1 (20 µg) was a positive control that was spotted alongside 20µg, 50 µg, 100 µg and 200 µg of TbHsp70.4, TbHsp70.c and TcHsp70 on the Hybond^{TM}–C Extra Nitrocellulose membrane in a Bio–Dot ST apparatus connected to a vacuum pump. The membrane was subsequently blocked with 5% (w/v) non-fat powdered milk in TBS for one hour at room temperature and incubated with 100ng/ml TbSTI1 for two hours at 4°C. The membrane was washed twice with TBST for 15 minutes before incubation for one hour in rabbit polyclonal anti-TbSTI1 antibody (1: 1000). The membrane was then washed twice for 15 minutes with TBST before incubation in HRP–conjugated goat anti–rabbit antibody (1:5000 dilution) for an hour. The detection of signal was carried out in the same way was described in section 4.2.5.3.

4.3. Results and Discussion

4.3.1. Heterologous production and purification of STI1 proteins, Hsp70s and Hsp83 proteins

4.3.1.1. Native purification of TbSTI1 and mSTI1

The plasmids used for the expression and purification of TbSTI1 and mSTI1 incorporated an N-terminal 6x His tag to assist purification. The pQE2000 plasmid (Van der Spuy et al., 2000) contained the *mSTI1* coding region inserted between *Sal*I and *Sph*I restriction sites, and the pQE2-TbSTI1 plasmid contained the *TbSTI1* coding region, which had been amplified from *T. brucei* genomic DNA, inserted between the *Not*I and *Nde*I sites. The appropriate enzymes were used to release the inserts (Figure 4.3 A and E). The integrity of the plasmids was also confirmed by restriction digestion, a double digest produced two fragments: TbSTI1 coding region of 1653 bp and the pQE2 vector of 4758 bp (Figure 4.3 B). The double digestion of the pQE2000 plasmid released the *mSTI1* coding region corresponding to 2097 bp and the pQE30 vector backbone (3504 bp) (Figure 4.3 F).

The plasmids expressing STI1 proteins from mouse and *T. brucei* were successfully transformed into *E. coli* M15[pREP4] cells. The level of protein production was analyzed by SDS-PAGE and western blot analysis (Figure 4.3 C and G). Following induction, there was a significant increase in the production of proteins at the expected size, 63 kDa for TbSTI1 and 64 kDa for mSTI1. A high level of protein production was observed 16 hours (O/N) post induction for TbSTI1, whereas mSTI1 was maximally produced after 4 hours. Therefore, these respective time points were used for cell harvesting for protein purifications.

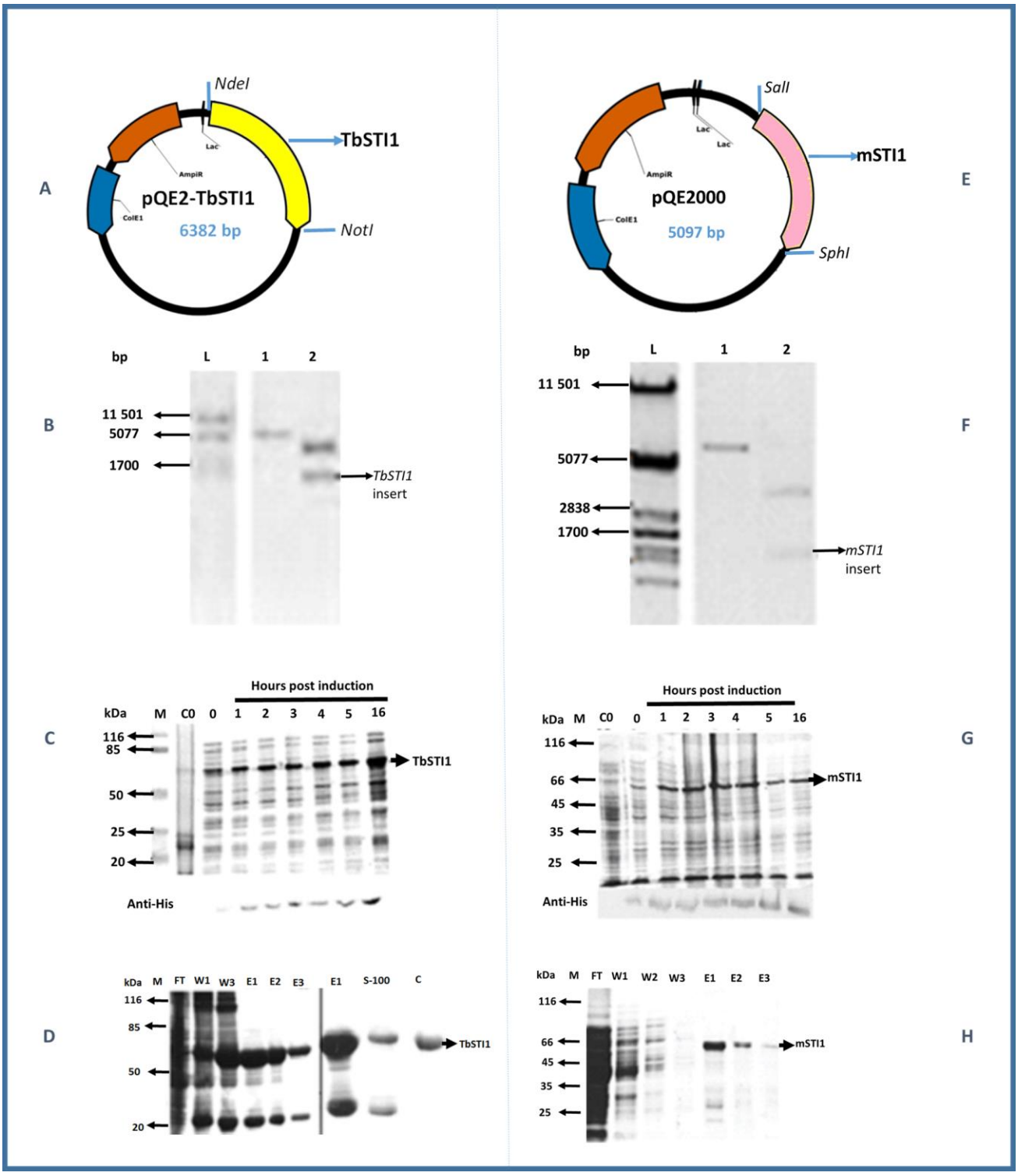


Figure 4. 3*:* Plasmid maps, diagnostic restriction analyses, overproduction and purification of TbSTI1 and mSTI1. Plasmid maps and associated restriction digests of the pQE2-TbSTI1 and pQE2000 plasmids. Both plasmids (A and C) confer resistance to ampicillin when transformed into *E. coli* indicated by the (AmpiR; β–lactamase coding sequence). The 6x His tag segments are upstream the coding sequences. Plasmids also indicate the origin of replication (ColE1 origin) and regions coding for *lac* operator

sequences. Agarose gels (1%) displaying the restriction digestion of pQE2-TbSTI1 (B) and pQE2000 (F), Lane L corresponds to lambda DNA molecular ladder digested with *Pst*I. Lane 1 is a single cut using *Xho*I which cuts both vector backbones and Lane 2 is the double digest to release the inserts corresponding to 2097 bp and 1653 bp for mSTI1 (D) and TbSTI1 (B) respectively. (C and G) 10% SDS-PAGE analysis of the production of 6x His tagged TbSTI1 and mSTI1 proteins. *E. coli* M15[pREP4] cells were transformed individually with pQE2-TbSTI1 (C) and-pQE2000 (G). Lane 1-Precision Plus ProteinTM Standards (BIORAD) marker; Lane C0-uninduced, untransformed *E. coli* M15[pREP4] whole cell extract (negative control), Lane 0-uninduced *E. coli* [pQE2-TbSTI1 (C) pQE2000 (G)] M15[pREP4] cells whole cell extracts; Lane 1-5 whole cells extracts of transformants taken at hourly intervals from 1-5 hours after the addition of 1 mM IPTG and Lane-16 represents whole cell extracts taken after overnight induction. Associated western blots using anti-His antibody are provided below the SDS-PAGE. (D and H) Purification of 6x His-tagged TbSTI1 and mSTI1 by nickel chromatography D). SDS-PAGE (10%) analysis of TbSTI1 purification from *E. coli* M15[pREP4] cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; FT- Flow through; Lane W1-W3, non-denaturing washes (50 mM imidazole, pH 7.5) and Lane E1-E3, Elution 1 to 3 using 1 M imidazole, pH 8.0.. E-showing all the elutions of TbSTI1 pooled together, S100-Fractions obtained after separating the protein on thea Sephacryl S-100 and C- TbSTI1 concentrated using vivaspin column with a cut off of 10 kDa H) SDS-PAGE (10%) analysis of mSTI1 purification from *E. coli* M15[pREP4]cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane FT- Flow through ; Lane W1-W3, non-denaturing washes (50 mM imidazole, pH 7.5) and Lane E1-E3, Elution 1 to 3 using 750 mM imidazole, pH 8.0.

Both recombinant TbSTI1 and mSTI1 proteins were recovered in the supernatant fraction (data not shown) indicating that they were soluble. Therefore, nickel affinity purifications of TbSTI1 and mSTI1 were carried out under native conditions. Both proteins associated strongly with the nickel charged beads resulting in >2 mg/L in the first elution with few contaminants visible in the elution fractions for mSTI1 (Figure 4.3 H). TbSTI1, however, co-purified with a species of approximately 20 kDa. Thus, the protein was subjected to size exclusion chromatography using media with a fractionation range of 1 – 100 kDa. However, the contaminant co-eluted with TbSTI1. The reason for this may have been due to the column length being insufficient to effectively separate the two protein species. These lower molecular weight species were not detected using anti-His antibody showing that it did not have a histidine tag attached, however anti-TbSTI1 antibodies were able to recognize these species (data not shown). It was concluded that the lower molecular weight species are potentially truncated TbSTI1 proteins or the product of incomplete synthesis. However, concentrating the protein by ultrafiltration using a spin column with a molecular weight cut off of 10 kDa, resulted in the removal of the contaminant, TbSTI1 was recovered to a yield of >5 mg/L (Figure 4.3 D).

4.3.1.2. Native purification of TbHsp70.4 and TbHsp70.c

The coding regions of TbHsp70.4 and TbHsp70.c were successfully inserted into the 6x His-tag N-terminal expression vectors pQE80 and pQE2 respectively. The integrity of each plasmid was verified by restriction digestion (Figure 4.4. B and F). The double digest yielded bands corresponding to the expected sizes of the TbHsp70.4 and TbHsp70.c coding regions at 1920 bp and 2031 bp respectively (Figure 4.4 B and F).

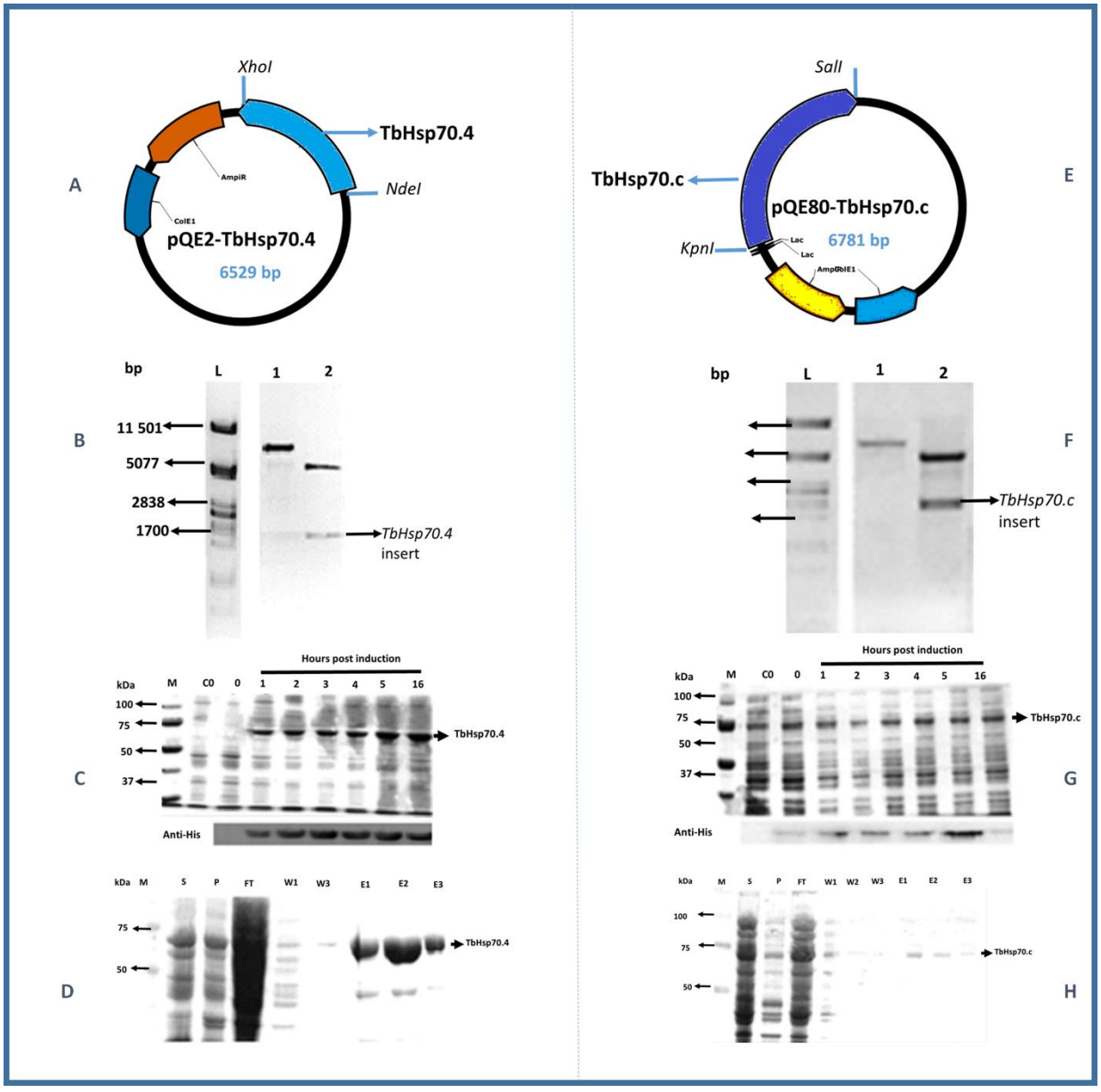


Figure 4. 4: Plasmid maps, diagnostic restriction analyses, overproduction and purification of TbHsp70.4 and TbHsp70.c. Both plasmid maps were generated using the plasma DNA software (Angers-Loustau, et al., 2007), the origin of replication (ColE1 origin) and regions coding for lac operator sequences are shown on each plasmid which confer ampicillin resistance when transformed into *E. coli* cells (depicted by AmpiR; β–lactamase coding sequence) (A and E). The 6xHis-tag segments are upstream the coding regions inserted between *XhoI* and *NdeI* for pQE2-TbHsp70.4 (A) and *SalI* and *KpnI* for pQE80-TbHsp70.c (E). (B and F). Confirmation of the identity of pQE2-TbHsp70.4 and pQE80-TbHsp70.c plasmids by restriction digest using agarose gel (1%) electrophoresis which displays Lane L, lambda DNA molecular ladder digested with *PstI*, Lane 1; linearized plasmids corresponding to 6678 bp yielded by a single cut of the pQE2-TbHsp70.4 using *NdeI* (B) and for pQE80-TbHsp70.c the linearlized plasmid was approximately 6782 bp. Double restriction enzyme analysis (B) Lane 2 corresponding to 1920 bp and 4758 bp for the expected sizes of pQE2 and TbHsp70.4 DNA fragments respectively and (F) Lane 2 corresponding to 2301 bp and 4751 bp for the expected sizes of pQE80 and TbHsp70.c DNA fragments respectively. (C and G): 10% SDS-PAGE analysis of the production of 6x His tag TbHsp70.4 and TbHsp70.c. in *E. coli* BB1994 cells transformed separately with pQE2-TbHsp70.4 (C) and pQE80-TbHsp70.c (G). Lane 1-Precision Plus ProteinTM Standards (BIORAD) marker; Lane C0-uninduced, untransformed *E. coli* BB1994 whole cell extract (negative control), Lane 0-uninduced *E. coli* [pQE2-TbHsp70.4 (C) pQE80-TbHsp70.c (G) untransformed *E. coli* BB1994 cells whole cell extracts; Lane 1-5 whole cells extracts of transformants taken at hourly intervals from 1-5 hours after the addition of 1 mM IPTG and Lane-16 represents *E. coli* BB1994 [(C- pQE2-TbHsp70.4 and G-pQE80-TbHsp70.c)] whole cell extracts taken after overnight induction. Associated western blots are provided below the SDS-PAGE. (D and H): Purification of 6x His-tagged TbHsp70.4 and TbHsp70.c by nickel chromatography. (D) SDS-PAGE (10%) analysis of TbHsp70.4 purification from *E. coli* BB1994 cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane S-Supernatant, Lane P-pellet, Lane U- unbound proteins; Lane W1 and W3, non-denaturing washes (50 mM imidazole, pH 7.5) and Lane E1-E3, Elution 1 to 3 using 750 mM imidazole, pH 8.0. H) SDS-PAGE (10%) analysis of TbHsp70.c purification from *E. coli* BB1994 cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane S-Supernatant, Lane P-pellet, Lane U- unbound proteins; Lane W1-W3, non-denaturing washes (50 mM imidazole, pH 7.5) and Lane E1-E3, Elution 1 to 3 using 500 mM imidazole, pH 8.0.

A gradual increase in the band at approximately 70 kDa corresponded to the molecular mass of His-tagged TbHsp70.4. SDS-PAGE analysis of the TbHsp70.4 induction showed that the whole cell lysate did not produce this protein prior to induction (Figure 4.4C lane C0). High levels of protein production of His-tagged TbHsp70.4 were achieved which was further confirmed by associated western analysis (Figure 4.4.C). Maximum levels of protein production were achieved 5 hours post induction using 1 mM IPTG and maintained for 16 hours (Figure 4.4C 5 and16 hrs). Unlike TbHsp70.4, TbHsp70.c was not overproduced and low basal levels of His-TbHsp70.c were not detected in the uninduced *E. coli* BB1994 whole cell extract (Figure 4.4 G; Lane 0), despite the presence of a band of similar size to TbHsp70.c. According to western analysis, the levels of

protein production of TbHsp70.c increased over 5 hours and then were greatly reduced after 16 hours post induction (Figure 4.4 G). Higher concentrations of TbHsp70.c were obtained using *E. coli* XL1 Blue and *E. coli* M15[pREP4] cells (Appendix G, Figure G1). Although low levels of TbHsp70.c were produced in *E. coli* BB1994, similar to TbHsp70.4, the fifth hour post induction were chosen as points for cell harvesting and protein purification.

Examination of the supernatant and the pellet fractions showed that TbHsp70.c andTbHsp70.4 were partially solubilized as protein corresponding to the sizes of the Hsp70s can be seen in both of these fractions (Figure 4.4 D and H). Subsequent to this observation, nickel affinity purification under native conditions was used to purify the proteins. SDS-PAGE analysis of the unbound proteins showed that TbHsp70.c was present in this fraction (Figure 4.4 H, lane U). For both protein purifications, little protein was lost in the wash steps. Most of the TbHsp70.4 protein eluted in the second elution, (Figure 4.4 D), and a yield of 10 mg/L was obtained. TbHsp70.4 co-purified with a low molecular weight species of ~25kDa (Figure 4.4.D) which was not detected using anti-His antibody or anti-TbHsp70.4 antibodies (data not shown) and therefore considered to be a protein that co-purified with TbHsp70.4. TbHsp70.c was successfully purified to high purity as few contaminants were visible, yielding 3 mg/L, of the protein (Figure 4.4 H).

4.3.1.3. Native purification of TcHsp70

Previous work on the purification and biochemical characterization of TcHsp70 was reported by Olson et al (1994), Edkins et al (2004) and Louw et al (2010). These studies purified TcHsp70 using denaturing methods. Of interest in this study was to over produce and purify TcHsp70 under non-denaturing condition for interaction studies with TbSTI1. It has been demonstrated that TcHsp70 is an orthologue of TbHsp70 (Chapter 2) and that these proteins share 89% sequence identity.

The identity of the pET14b-TcHsp70 plasmid was confirmed by restriction digestion wherein the double digestion resulted in two fragments of 4621 bp and 2084 bp corresponding to the

expected sizes of the pET14b vector and TcHsp70 coding region respectively (Figure 4.5 B, Lane 2). The over-production and detection of His-TcHsp70 in *E. coli* BL21 cells is shown in Figure 4.3 C. Low basal levels of this protein was evident in the uninduced cells (Figure 4.5 C). The maximum levels of the ~71kDa His-tagged TcHsp70 protein were achieved 3 hours post induction.

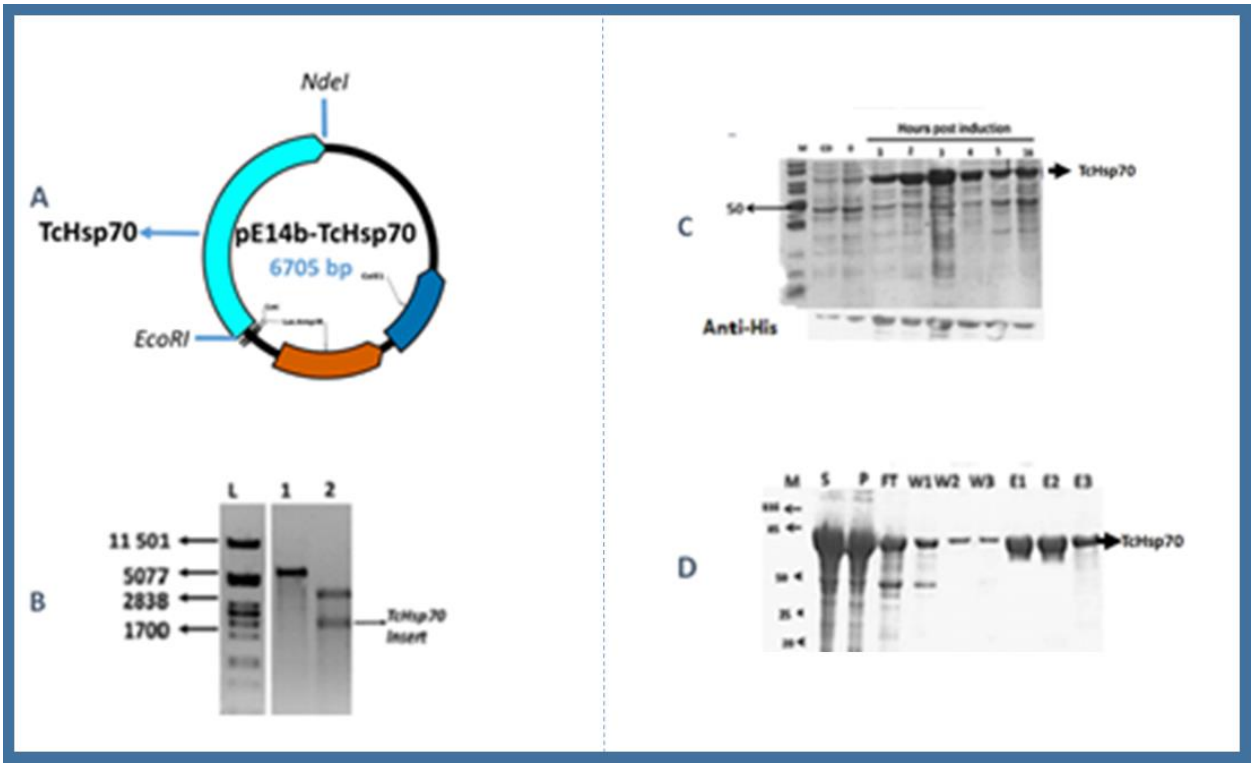


Figure 4. 5*:* Plasmid maps, diagnostic restriction analyses, overproduction and purification of TcHsp70. A) Plasmid map of the pET14b-TcHsp70 showing diagnostic restriction sites, the origin of replication (ColE1 origin) which confer ampicillin resistance when transformed in *E. coli* cells (depicted by AmpiR; β–lactamase coding sequence). B). Confirmation of the identity of the pET14b-TcHsp70 plasmid by a single and double restriction digestion. Lane L *Pst*I digested lambda DNA with corresponding sizes, Lane 1-a single digestion with *Nde*I and Lane 2- double digestion with *Nde*I and *Eco*RI. C) SDS-PAGE and associated western analysis of the 6x His TcHsp70 in *E.coli* BL21 cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane C0- uninduced, untransformed *E coli* BL21 whole cell extract (negative control) Lane 0-5 whole cells extracts of transformants taken at hourly intervals from uninduced (0) to 1-5 hours after the addition of 1mM IPTG and Lane-16 represents *E.coli* BL21 (pET14b-TcHsp70) whole cell extracts taken after overnight induction. D). Purification of His-tagged TcHsp70 by nickel affinity chromatography. SDS-PAGE (10%) analysis of TcHsp70 purification from *E. coli* BL21cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane S-Supernatant, Lane P-pellet, Lane FT-unbound proteins; Lane W1-W3, non-denaturing washes (20 mM imidazole, 10mM ATP and 20% v/v glycerol, pH 7.5) and Lane E1-E3, Elution 1 to 3 using 500 mM imidazole, pH 8.0

His-tagged TcHsp70 was successfully purified by nickel affinity chromatography using native purification procedures (Figure 4.5 D). A large amount of protein did not bind to the beads as they were saturated with 6x His TcHsp70 as indicated by the flow through (FT) fraction and some protein was lost in the washes. Due to the lack of a compatible *E. coli dnaK* minus strain to express proteins from pET vector plasmids, the wash buffer contained 10 mM ATP and 20% (v/v) glycerol to potentially reduce or remove DnaK altogether (Guo et al., 2007; Rial and Ceccarelli, 2002; Burger et al., 2014). High yields of TcHsp70 (2.5mg/L) were purified with a small amount or no DnaK contamination which was confirmed by western analysis using anti- DnaK antibodies (data not shown).

4.3.1.4. Approaches to the purification of TbHsp83

The coding regions of TbHsp83 were inserted into pQE80 and pQE60 expression vectors which resulted in the production of Hsp83 with an N-terminal and C-terminal His tag, respectively. This was done in order to determine if the His tag had an affect on the purification or activity of the protein. Confirmation of the constructs was achieved by restriction digestion analysis wherein the coding sequence of TbHsp83 was inserted between *Bam*HI and *Bgl*II in the pQE60 vector and *Bam*HI and *Sal*I in the pQE80 vector (Figure 4.6. A and F). Both plasmids were linearized using *Bam*HI and the double digest was carried out using appropriate restriction enzymes. The successful confirmation of the 6x His tagged plasmids encoding TbHsp83 allowed for subsequent protein production to be conducted (Figure 4.6 C and G).

The production of N-terminal and C-terminal His-tagged TbHsp83 in *E. coli* M15 [pREP4] cells was monitored over time. The whole cell extract samples were analyzed by SDS-PAGE (Figure 4.6. C and G). The C-terminal 6x His TbHsp83 produced a protein at a molecular weight of ~75 kDa, which is ~10 kDa smaller than the expected size of 85 kDa, while no clear production of protein could be seen using the pQE80 vector. The presence of the C-terminal His- tagged protein produced in *E. coli*M15 [pREP4] cells was confirmed to be TbHsp83 by western analysis using the anti- TbHsp83 antibodies (Appendix G, Figure G3). The truncated His-TbHsp83 was over-

produced, as no protein was expressed prior to induction (Figure 4.6C lane C0). Maximum levels of proteins were obtained 2 hours post induction using 1 mM IPTG and maintained for 16 hours (Figure 4.6C lanes 1 – 5 and 16). Native purification of His-TbHsp83 was attempted using protein expressed using both vectors, although no success was obtained using the N-terminal His tag which was not unexpected as no protein was evident in the induction study. A western analysis should have been conducted to confirm that indeed there was no protein production, and in that case the coding sequence of TbHsp83 should be re-engineered back into the vector. Although successful production of C-terminal His-tag TbHsp83 was achieved, the protein yields (0.003mg/L) were insufficient to carry out *in vitro* characterization and the eluted protein fractions had high levels of contaminating proteins. Full length Hsp90 is very difficult to express and one successful strategy has been to express and purify the individual domains.

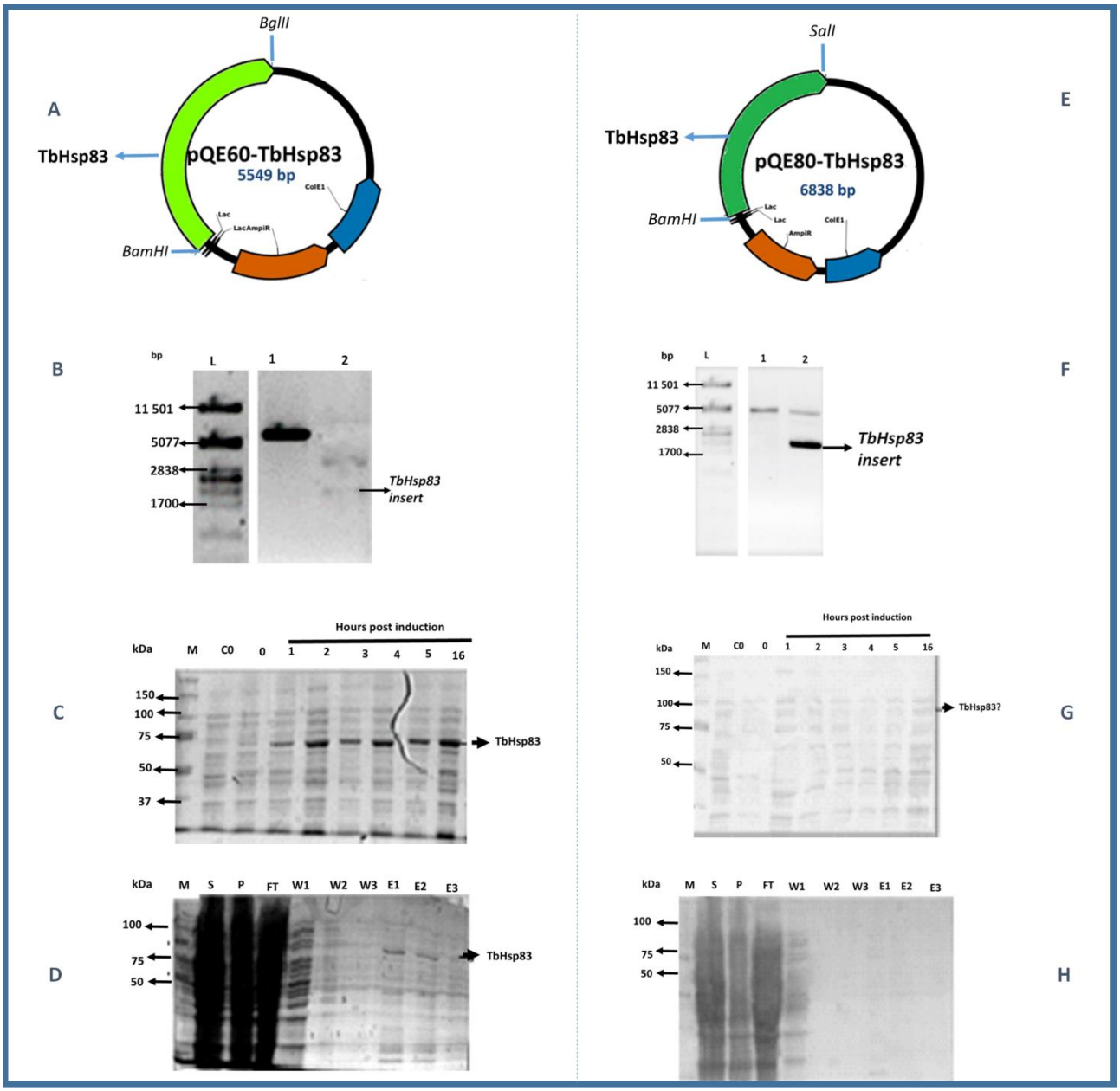


Figure 4. 6: Plasmid maps, diagnostic restriction analyses, overproduction and purification of N-terminal and C-terminal H is tagged TbHsp83.Both plasmid maps were generated using the plasma DNA software, ColE1 origin and regions coding for AmpiR; β–lactamase coding sequence are depicted (A and E). The coding region for TbHsp83 are inserted between *BglII* and *BamHI* for pQE60-TbHsp83 (A) *SalI* and *BamHI* for pQE80-TbHsp83 (E). B and F). Confirmation of the identity of pQE60-TbHsp83 and pQE80-TbHsp83 plasmids by restriction digestion using agarose gel (1%) electrophoresis which displays Lane L, lambda DNA molecular ladder digested with *Pst I*, Lane 1; linearized plasmids corresponding to 5 549 bp yielded by a single cut of the pQE60-TbHsp83 using *Bam HI* (B) and for pQE80-TbHsp83 the linearlized plasmid was approximately 6 838 bp. Double restriction enzyme analysis ([B) Lane 2 corresponding to 2115 bp and 3 439 bp for the expected sizes of pQE60 and TbHsp83 DNA fragment respectively and F). Lane 2 corresponding to 2 115 bp and 4 751 bp for the expected sizes of pQE80 and TbHsp83 DNA

fragment respectively]. C and G: 10% SDS-PAGE analysis of the production of N-terminal and C-terminal 6x His tag TbHsp83 and TbHsp83. *E. coli* M15[pREP4] cells were transformed separately with pQE60-TbHsp83 (C) and-pQE80-TbHsp83 (G). Lane 1-Precision Plus ProteinTM Standards (BIORAD) marker; Lane C0-uninduced, untransformed E. coli M15[pREP4] whole cell extract (negative control), Lane 0-uninduced *E. coli* [pQE60-TbHsp83 (C) pQE80-TbHsp83(G)]M15[pREP4] cells whole cell extracts; Lane 1-5 whole cells extracts of transformants taken at hourly intervals from 1-5 hours after the addition of 1mM IPTG and Lane-16 represent E. coli M15[pREP4] [(C- pQE60-TbHsp83 and G-pQE80-TbHsp83)] whole cell extracts taken after overnight induction. D and H): Purification of N-terminal and C-terminal 6x His-tagged TbHsp83. by nickel affinity chromatography. (D and H) SDS-PAGE (10%) analysis of TbHsp83 purification from *E. coli* M15 [pREP15] cells. Lane 1- Precision Plus Protein TM Standards (BIORAD) marker; Lane S-Supernatant, Lane P-pellet, Lane U/FT- unbound proteins in the flow through; Lane W1-W3, non-denaturing washes (50 mM imidazole, pH 7.5) and Lane E1-E3, Elution 1 up to 3 using 500 mM imidazole, pH 8.0.

4.3.2. TbSTI1 potentially interacts with cytosolic Hsp70s

This experiment sought to study the interaction of TbSTI1 with the cytosolic *Trypanosoma* Hsp70 proteins. As the purification of TbHsp83 was unsuccessful, human Hsp90 was used to study the interaction between TbSTI1 and Hsp90. The interactions were studied using SPR, pull down assays and dot blot assays. To study the biophysical interactions between TbSTI1 with cytosolic Hsp70s and human Hsp90, SPR was employed. TcHsp70 was demonstrated to be a closely related orthologue of TbHsp70 (Chapter 2). Also, an interaction between TcSTI1 (shares 78% sequence identity to TbSTI1) and TcHsp70 has been demonstrated (Schmidt et al., 2011). The Hsp70s and Hsp90 proteins were successfully immobilized via covalent amine coupling to the sensor chip surface. TbSTI1 was also immobilized to study the possibility of it existing as a dimeric species or interacting with itself. Subsequent to the immobilization, the kinetics of the interaction between cytosolic Hsp70s and TbSTI1 was examined using a series of TbSTI1 concentrations injected onto the sensor chip at at a flow rate of 60 µl/min and a contact time of 90s. This was done in order to determine whether the interaction was concentration dependent. It is worth noting that the kinetic profile for the interactions did not fit the Langmuir binding model (O‟Shannessy et al., 1993).

Once human Hsp90, TcHsp70 and TbHsp70.4 were immobilized, varying concentrations of TbSTI1, which was the analyte, were passed over the surface of the sensor chip. For Hsp90, TcHsp70 and TbHsp70.4, higher concentrations of the TbSTI1 analyte resulted in a dose dependent response (Figure 4.7 A, B and D), while with TbHsp70.c immobilized, this was not seen. The data obtained revealed that TbSTI1 does not adsorb strongly to the surface of the chip, as increased dissociation was seen even with higher concentrations injected (Figure 4.7 A-D). The same association-dissociation curve patterns with similar values >6000 RU were observed for human Hsp90 and TcHsp70 (Figure 4.7 A, B and D).

Indeed, high dissociation was observed in all reactions. It is suspected that the concentrations of TbSTI1 used were too high, an observation that is conflicting based on the fact as these were well below saturation. Saturation was determined by plotting the maximum response units for the association phase against the amount of TbSTI1 injected, a plot that did not appear to reach saturation was obtained (Figure 4.7 A, B and D). Lower concentrations of TbSTI1 could have been tried. If saturation was reached, it would imply that the interaction was specific. However, specific binding of TbSTI1 to Hsp90, TbHsp70.4 and TcHsp70 was supported by the fact that TbSTI1 did not interact with all the proteins in the study. In the case of TbHsp70.c, there was no change in the response units with increasing concentrations of TbSTI1.

The main advantage of SPR over the other protein-protein interaction techniques is direct and rapid determination of the interaction by calculating association and dissociation rates of the binding. SPR does not require labelling of protein, and can be conducted using smaller concentrations of protein (Beseničar et al., 2006). It is apparent that low concentrations of TbSTI1 were used as saturation was not reached in each of the plots and as such, further studies are required to examine such interactions quantitatively at higher TbSTI1 protein concentrations. However, the linear relationship between the response and the concentration of TbSTI1 was suggestive of an interaction with Hsp90, TbHsp70.4 and TcHsp70. This is supported by the fact that TbSTI1 did not show a similar trend in binding to TbHsp70.c in the same experiment.

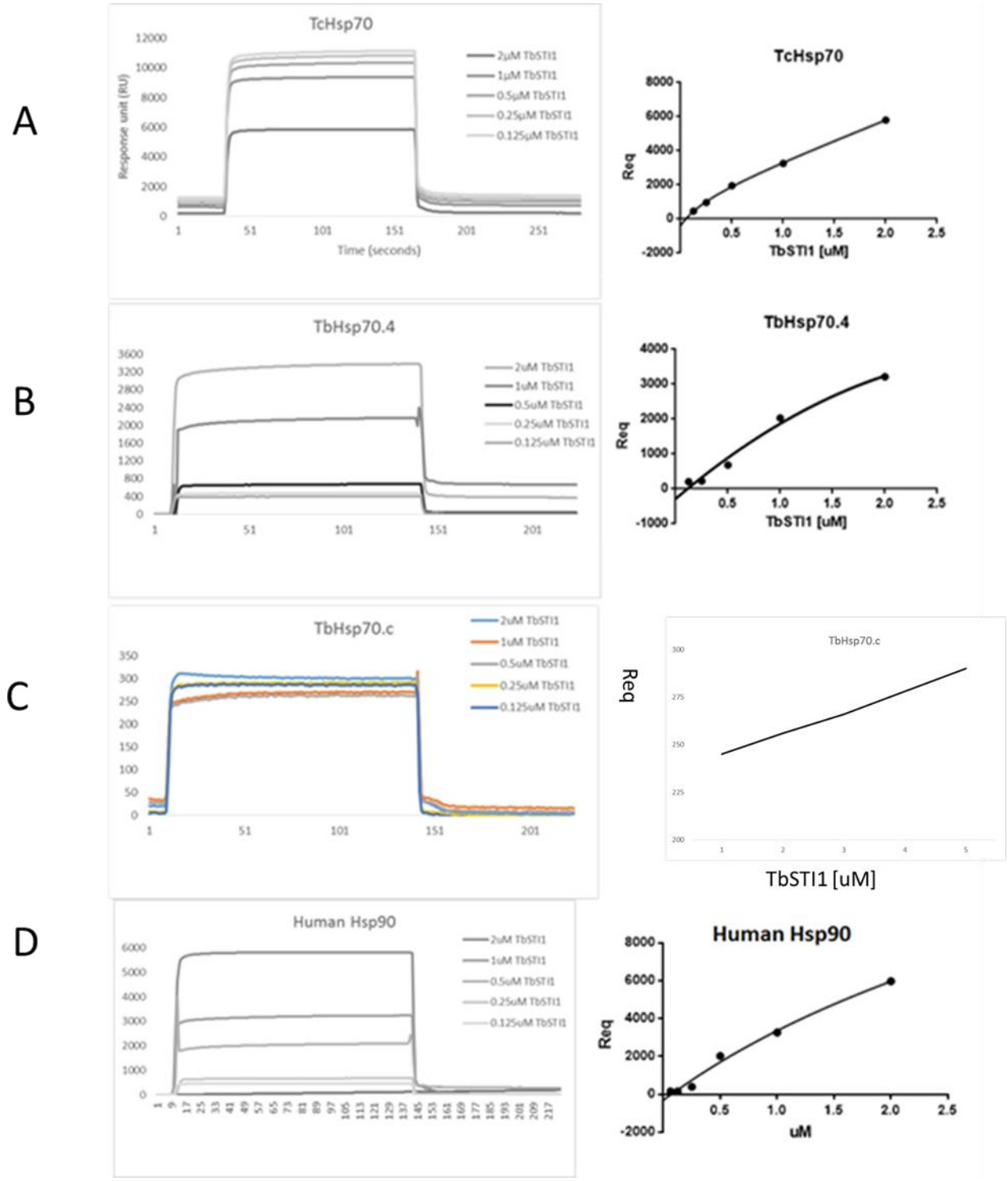


Figure 4. 7: SPR analysis of TbSTI1 interacting with cytosolic Hsp70s at concentrations of 0.0125 µM - 2 µM at a flow rate of 60 µl/min for 90 s followed by a dissociation phase of 5 minutes using the assay buffer, each TbHsp70 protein except TbHsp70.c has an associated saturation plot derived from plotting maximum response units for the association phase obtained in A against the concentration of TbSTI1 injected.

Next, pull-down assays were used to examine the interactions between TbSTI1 and TbHsp70s. A total of 20 µg of purified TbSTI1 was used to pull-down 20 µg TcHsp70, TbHsp70.4 and TbHsp70.c (Appendix G, Figure G4). In this assay, TbSTI1 was not able to interact with the cytosolic Hsp70s. Although bands were observed in the inputs (Appendix, Figure G4), nothing was detected in the test or control. TbSTI1 was observed in the pull-down indicating that the experiment worked, however, it is possible that anti-TbSTI1 antibodies are not appropriate for pull-down. The lack of efficiency in anti-TbSTI1 antibodies for this particular assays is suspected to be as a result of it binding TbSTI1 in such a way that the TPR1 domain binding site for Hsp70s is blocked. The main short-fall for this experiment was all the proteins used in the study were His-tagged, therefore, using a different tag on TbSTI1 might have resulted in positive results. It is possible that the incubation period (overnight) was too long, also, binding proteins could have been mixed and allowed to bind before adding the antibody beads. Detection of binding could have been done by FPLC.

Taken together, this prelimary data would need to be repeated for SPR and pull-down assays. Higher concentrations of up to 10 µM TbSTI1 should be used, as was done in previous studies where successful biophysical binding of STI1 proteins (mSTI1) and Hsc70 was achieved (Odunuga et al., 2003), concentrations up to 8 µM were used. Brinker et al (2002) used between 0-50 µM to show how TPR1 and TPR2A discriminates between Hsp70 and Hsp90 C-terminal ends respectively. For pull-down, engineering a glutation S- transferase tag (Smith and Johnson, 1988), FLAG-tag (Hopp et al., 1988) or Strep-tag (Schmidt and Skerra, 1993) on TbSTI1 will most likely result in binding. It is also possible that TbSTI1 is inactive or folded incorrectly.

Following the observation made from SPR and pull-down studies, higher concentrations of Hsp70s and TbSTI1 were used in a less sensitive method, dot blot assays. It was observed that, consistent with bioinformatics predictions and suggested by the SPR analysis, TbSTI1 bound to TcHsp70 in a concentration dependent manner as indicated by the increasing intensity of the signal (Figure 4.9 A). The signal intensity was plotted against the amount of TbHsp70 resulting in

a typical saturation curve, the saturation point was reached at 100 µg of TcHsp70. The specificity of the interaction was indicated in the associated saturation curve (Figure 4.8A). No binding was seen between TbHsp70.c and TbSTI1 (Figure 4.9 C), this was also supported by the SPR and consistent with bioinformatics predictions as TbHsp70.c lacks an EEVD motif at the end of the protein (Chapter 2). A Kd value of a 3.9 nM was obtained from plotting the pixel density from the signals obtained in the dot blot assays versus concentrations of Hsp70s, suggesting that the interaction between TcHsp70 and TbSTI1 was stronger than between TbHsp70.4 and TbSTI1 (Kd 50.85 nM), which is consistent with bioinformatics predictions that a weak or partial chaperone co-chaperone interaction may exist between these proteins. The interaction between TcHsp70 and TbSTI1 was expected as TcHsp70 has the canonical EEVD motif for interaction with STI1.

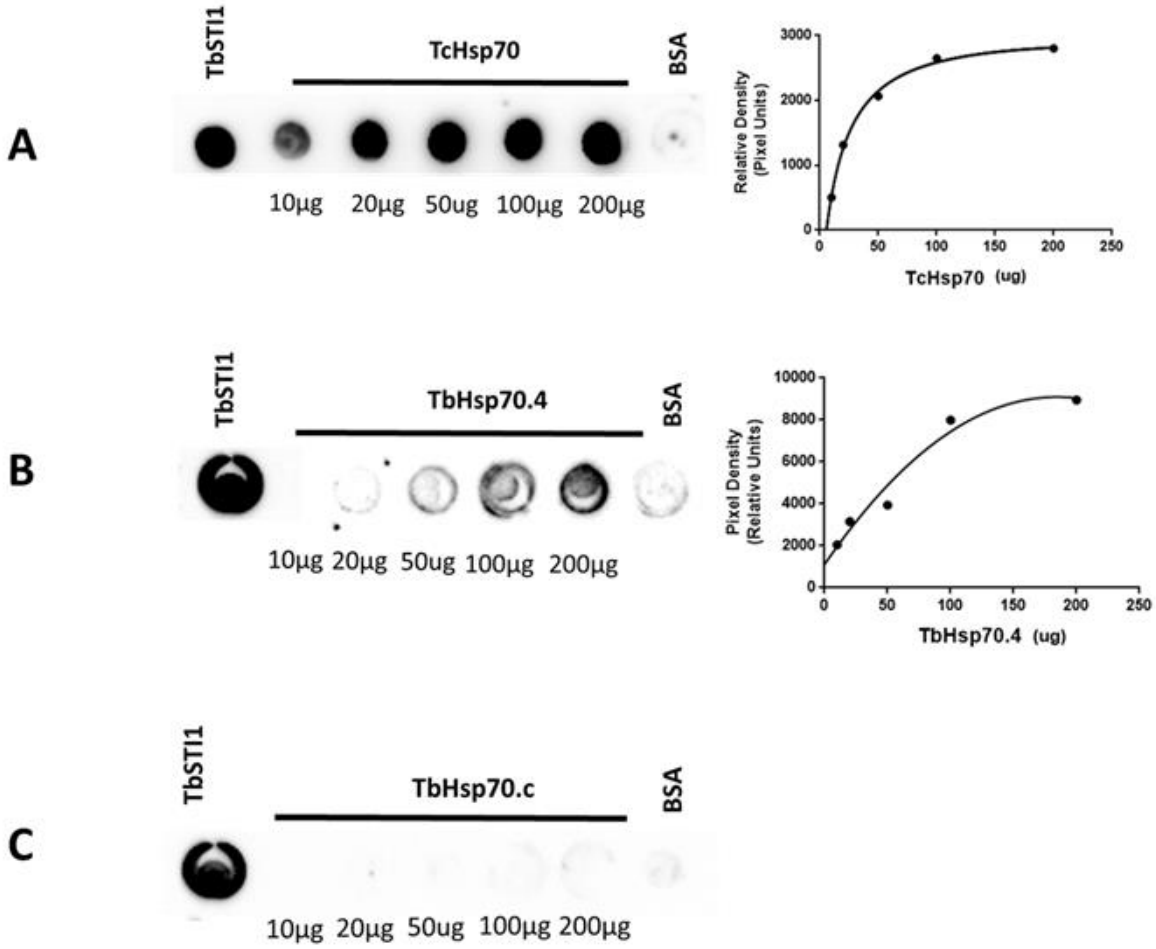


Figure 4. 8: TbSTI1 interactions with cytosolic Hsp70s at increasing concentrations. Dot blot assay; TbSTI1 was used a positive control and varying concentrations of cytosolic Hsp70s were vacuum blotted onto a nitrocellulose membrane, incubated with His-TbSTI1 (100 ng/ml) and bound TbSTI1 was detected by western analysis using anti-TbSTI1 antibody. BSA was blotted as a negative control, each Hsp70 except TbHsp70.c has an associated graphical representation of the dot plot assays generated using graphPad Prism6.

No interaction was seen with an equivalent concentration of BSA which was used as a negative control. The positive control of purified TbSTI1 showed strong antibody specificity (Figure 4.8). The commercial recombinant human Hsp90 was not used as the high concentrations of protein required were not available. Dot blots are usually used for preliminary characterization of protein-protein interaction. Therefore, these preliminary results obtained using dot blot assays provide the necessary background for future studies to use higher concentrations of TbSTI1 in more sensitive methods, such as SPR and pull down assays.

4.3.3. TbSTI1 is unable to bind human Hsp70 and Hsp90 in mammalian cell lysates

Pull-down assays were performed to determine if TbSTI1 was able to interact with human Hsp70 and Hsp90. In this experiment mSTI1 was used as a positive control. Due to the presence of the His-tag on both STI1 proteins, 20 µg of purified TbSTI1 and mSTI1 were bound to nickel beads. Subsequent to this, the STI1 bound beads were incubated with human breast cancer Hs578T cell lysates to determine whether TbSTI1 was able to interact with human Hsp70 and Hsp90. A negative control containing nickel beads alone was also set-up to account for non-specific binding of Hsp70 and Hsp90 to the resin. The reactions were washed separately with PBS and subjected to western analysis. Hs578T cell line was used as it was a readily available source of human Hsp70 and Hsp90. These proteins are abundant in this cancerous cell line which can be cultured easily, thus high concentrations of the lysate can be obtained. Previous research within BioBRU has shown that mSTI1 can interact with human Hsp70 and Hsp90 in this cell line (Lara Contu, Msc thesis).

Both TbSTI1 and mSTI1 were detectable by western blot showing that the nickel resin successfully pulled down the His-tagged protein (Figure 4.9). However, TbSTI1 was unable pull down human Hsp90 or Hsp70 (Figure 4.9). The presence of bands for Hsp70 and Hsp90 in the mSTI1 pull down lysate indicated that mSTI1 could interact with these human chaperones. The negative control (beads) showed no signal for the detection of human Hsp70 and Hsp90 (Figure 4.9). These results reflect that TbSTI1 cannot compete with endogenous human Hop, for the binding of Hsp70 and Hsp90, while mSTI1 was able to interact with Hsp70 and Hsp90 in these cell lysates (Figure 4.9). The results obtained for mSTI1 are consistent with observations made by Contu et al (2014), who showed an interaction between mSTI1 and human Hsp90 and Hsp70.

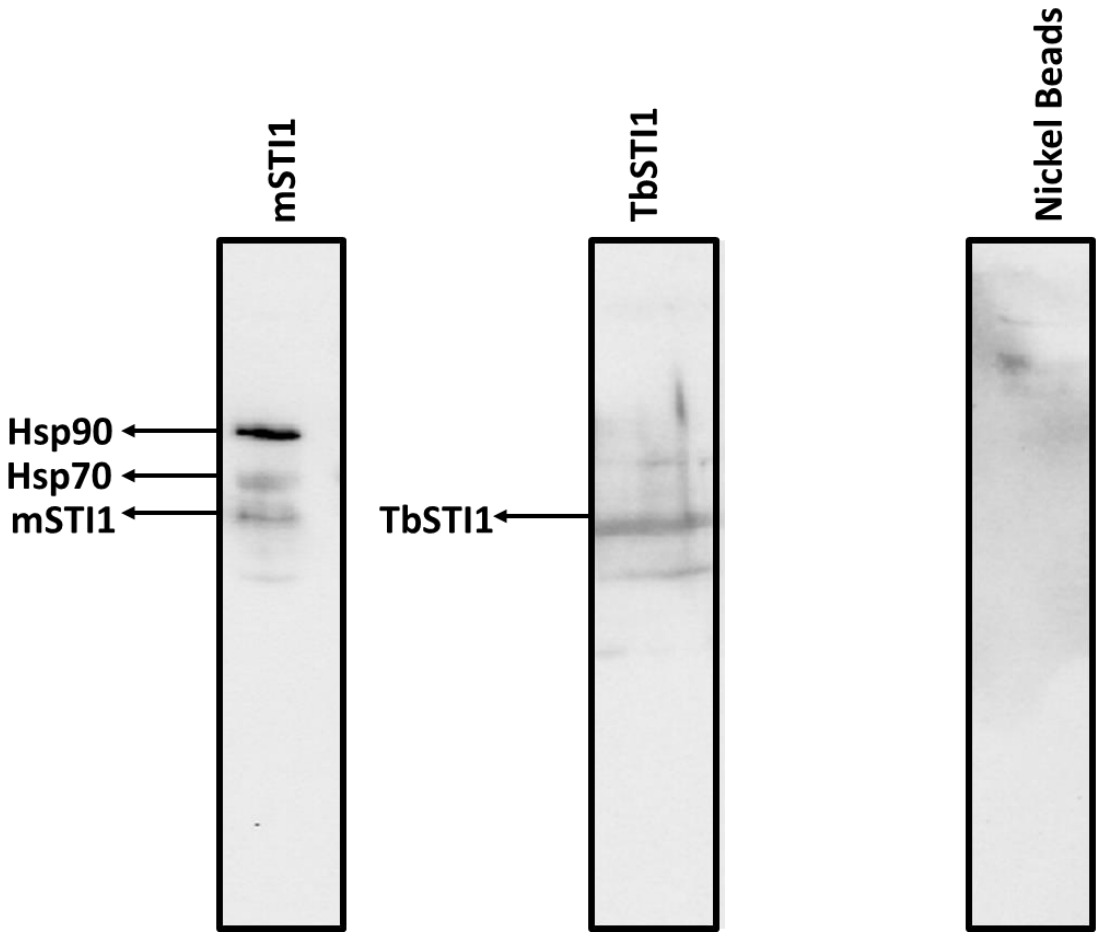


Figure 4. 9: TbSTI1 is unable to pull down human Hsp90 and Hsp70 from Hs578T cell lysate. Pull down assays were conducted by separately incubating His-TbSTI1, His-mSTI1 and sepharose beads with Hs578T cell lysate overnight and analyzed by western blot using anti His, anti-Hsp70 and anti-Hsp90 antibodies.

4.3.4. The detection of TbSTI1, TbHsp70.4, TbHsp70.c and TbHsp83 at the bloodstream stage of parasite development

The aim of this experiment was to determine whether TbHsp83, TbHsp70s and TbSTI1 were expressed during the bloodstream stage of the life-cycle of *T. brucei.* Lysates prepared from *T. brucei* cells were subjected to western analysis using anti-TbSTI1, anti-TbHsp83, anti-TbHsp70.c and anti-TbHsp70.4 antibodies. TbHsp70.c is 73 kDa, while TbHsp70.4 is 70 kDa, the difference in

size should be apparent when these proteins are resolved on SDS-PAGE. The detection of a protein at ~73 kDa confirms previous findings that TbHsp70.c is expressed at the bloodstream stage of the *T. brucei* life-cycle (Burger et al., 2014). TbHsp70.c therefore served as a positive control for the lysate. No protein of 70 kDa was detected, which corresponded to TbHsp70.4, as shown by the arrow (Figure 4.10). According to the bioinformatics data, TbHsp70.4 is predicted to be the constitutively expressed Hsp70 in *T. cruzi*, and thus the protein may have been absent or present at undetectable levels. Another protein was detected at ~85kDa, the expected molecular mass of TbHsp83. Hsp90 is known to be a very abundant protein in the cell and TbHsp83 was present at higher concentrations than the Hsp70s (Figure 4.10). A protein at the expected size of TbSTI1 ~63 kDa was also detected in the *T. brucei* lysate.

The presence of these molecular chaperones during the BSF stage of parasite development suggests the presence of components of the STI1-mediated Hsp90-hetero-complex in *T. brucei*. The same experiment would need to be conducted to detect TbHsp70 at the BSF stage of parasite development as this chaperone plays a crucial role at the intermediate stage of the STI1 mediated Hsp90 hetero-complex. Additional experiments would be to pull down the Hsp70-STI1-Hsp90-complex.

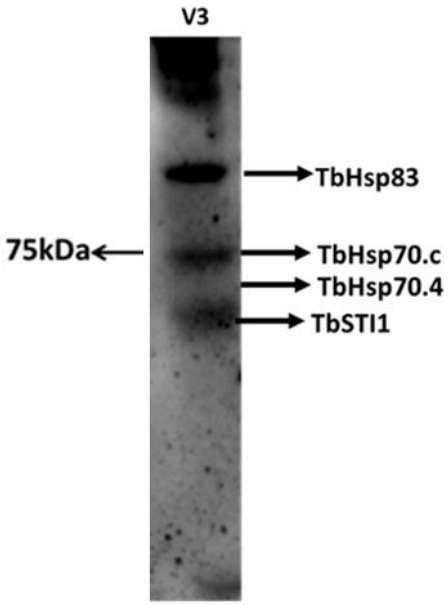


Figure 4. 10*: T. brucei* 427 v.3 bloodsteam stage lysates were probed with antibodies specific for TbHsp83, TbHsp70.4 and TbHsp70.c. and TbSTI1 proteins. Lane M- Precision Plus Protein TM Standards (BIORAD) marker and V3-lysate probed with anti-TbHsp83 (1:1000), anti-TbHsp70.4 (1:1000) ,anti-TbHsp70.c (1:1000) and anti- TbSTI1 (1:1000).

Although the anti-TbHsp70.4 antibody was used to successfully establish the cytoplasmic localization of TbHsp70.4 using immunofluorescence assays (Bangs et al., 1993), additional steps to confirm the reliability of the TbHsp70.4 antibody was carried out by subjecting both *T. brucei* 427 v.221 and *T. brucei* 427 v.3 strains to western analysis, using purified TbHsp70.4 as positive control. A band ~70 kDa, the expected size of TbHsp70.4, was detected using the purified protein and no TbHsp70.4 was detected in lysates prepared from two stains of *T. brucei*. The two strains of lister 427 *T. brucei* have the same genotype and different phenotypes. The absence of TbHsp70.4 in the lysates suggest that it is not expressed during the BSF stage of the parasite life-cycle (Figure 4.11 A and B) it also confirms that the antibodies are detecting TbHsp70.4.

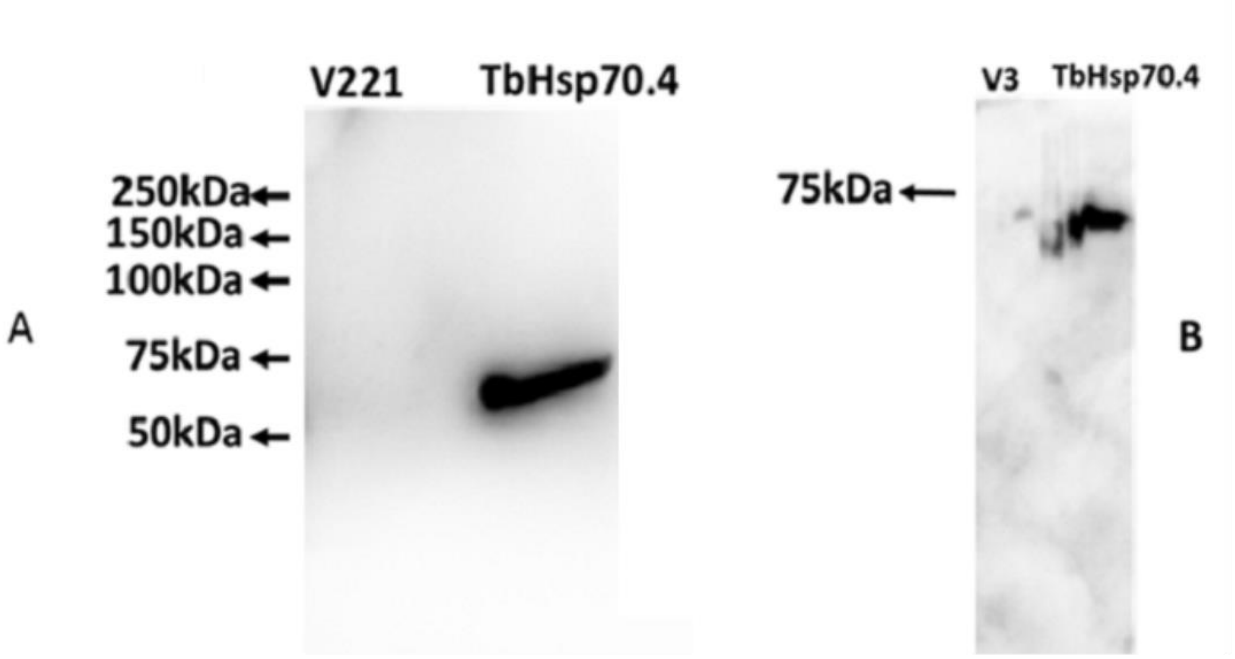


Figure 4. 11:TbHsp70.4 is not expressed during the bloodstream stage of the parasite life-cycle. A-Purified TbHsp70.4 and *T. brucei* 427 v.221 lysates were resolved by SDS-PAGE and then subjected to western analysis of the detection of TbHsp70.4 using anti-TbHsp70.4 B-*T. brucei* 427 v.3 lysate was resolved by SDS-PAGE and then subjected to western analysis using anti-TbHsp70.4.

4.3.5. TbSTI1 and TbHsp83 are heat inducible

Following the detection of TbHsp83 and TbSTI1 in the bloodstream stage of the life-cycle of *T. brucei*, the study sought to determine if the expression of these proteins was upregulated in response to heat shock. Two strains of *T. brucei* were incubated under normal conditions of 37 °C and heat shock at 42°C. Cell lysates of *T. b. brucei 427* v.221 (Figure 4.12 A) and *T. b. brucei 427* v.3 (Figure 4.12 B) grown at both temperatures were subjected to western analysis. The concentration of TbHsp83 and TbSTI1 in both strains were more noticeable after heat shock (Figure 4.12 A and B). The concentrations of TbHsp83 were more obvious than those of TbSTI1. Equivalent numbers of cells were loaded and analysed, however an additional control such as anti-actin, anti-histone or anti-β-tubulin antibody could have been included as a loading control to show that equivalent amounts of cells were added per lane.

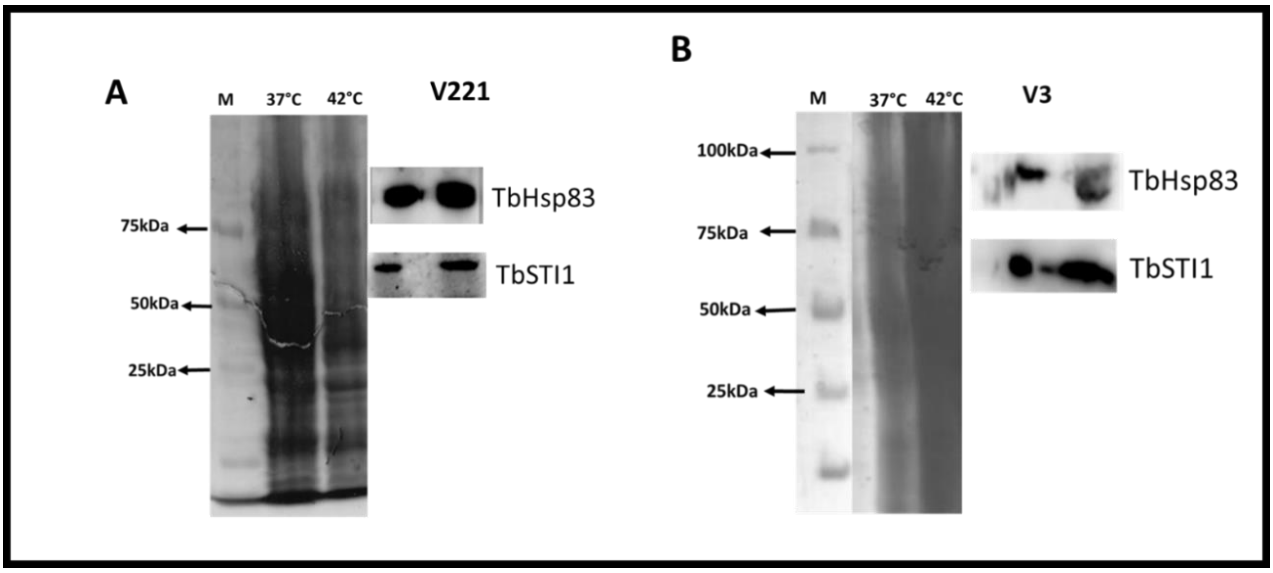


Figure 4. 12: TbHsp83 and TbSTI1 are upregulated following heat shock. SDS PAGE and associated western analysis of the T. brucei 427 v.221 (A) and T. brucei 427 v.3 (B) lysates at 37°C and 42°C. The expressions of TbSTI1 and TbHsp83 during the bloodsream stage were analyzed by western analysis using anti-TbSTI1 and anti-TbHsp83 antibodies.

These findings suggest that TbHsp83 and TbSTI1 are required for protein quality control under both permissive and heat induced conditions. It would have been interesting to check if TbHsp70.4 was induced under heat shock. A quantitative analysis is required to confirm that indeed TbSTI1 and TbHsp83 are significantly increased upon exposure to heat.

4.4. Discussion

The functional characterization of TbSTI1 and TbHsp70.4 has not been previously reported in literature. Recently TbHsp83 was shown to be more selective to inhibitors than its human homologues (Meyer and Shapiro, 2013), and its ATPase activity is 10-fold higher than that of human Hsp90 (Pizarro et al., 2013). This study showed the first successful heterologous production and purification of TbHsp70.4 and TbSTI1 from DnaK minus *E. coli* BB1994 cells. These proteins were purified to sufficient yields without the addition of denaturants. The main

advantage of native purification is the likelihood that the protein would not lose activity (Misra and Ramachandran, 2009). However, using native purification methods, low yields of TbHsp83 were obtained. Therefore no subsequent *in vitro* characterization could be conducted. Hsp90 proteins are typically very difficult to purify despite their abundance in the cell. TbHsp70.c and TcHsp70 purifications have been previously reported (Edkins et al., 2004; Louw et al., 2009 and Burger et al., 2014). Expression of TbHsp83 in yeast, mammalian and insect cells could be attempted to check and see if pure protein in abundant yields would be obtained. It might also be worth attempting the expression and purification of TbHsp83 at lower temperatures such as 18°C and 25°C.

For TbHsp70.c, the protocol was optimized such that contaminating proteins were removed and >80% purify was achieved while contaminating DnaK was removed from TcHsp70. Following the bioinformatics predictions, TbSTI1 was demonstrated to be a TPR containing co-chaperone of Hsp70 and Hsp90. This was also demonstrated by its ability to bind strongly to TcHsp70, a canonical Hsp70 with a C-terminal EEVD motif using dot blots. The binding of TbSTI1 to TbHsp70.4 combined with the identification of phosphorylation residues (Chapter 3) on this protein allows it to be prioritized for pathogenicity studies.

Polyclonal antibodies of TbHsp83 and TbSTI1 were successfully developed, and as a consequence the expression and heat inducibility could be determined. This study has demonstrated the occurrence of TbSTI1 at the bloodstream stage of *T. brucei* development, thus suggesting the conservation of the STI1 mediated Hsp90 hetero-complex in this organism. In common with STI1 homologues from several organisms (Nicolet and Craig 1989, Honoré et al. 1992, Lässle et al. 1997, Heine et al. 1999, Zhang et al. 2003, Song et al. 2009), TbSTI1 was shown to be induced by heat stress. TcSTI1 is not heat inducible as the protein levels did not change following heat shock, increased levels of TcSTI1 were observed at the epimastigote late phase when the parasite was subjected to starvation (Schmidt et al., 2011). Indeed, TbHsp83 was shown to be abundant in the cell, it is tempting to speculate that, similar to its *L. donovani* orthologue, this protein constitutes ~3% of the entire protein content (Brandau et al., 1995), a proposal that would support the

hypothesis that the multiple copies enhances Hsp83 expression. TbHsp83 has been shown to be essential for parasite viability at all stages of the life-cycle (Alsford et al., 2011). Additionally, TbHsp83 inhibition effectively inhibits the growth of the parasite (Meyer and Shapiro, 2013). Stage specific expression of Hsp70s in kinetoplastids has been observed, this is consistent with the findings that TbHsp70.4 is not expressed at the BSF stage of *T. brucei* life-cycle. The orthologue of TbHsp70.4 in *T. cruzi* was undetectable at the trypanomastigote stage of parasite development but highly enriched in the amasitigotes (Atwood et al., 2005). This study also aimed to understand the interaction of cytosolic Hsp70s and TbSTI1 using SPR, pull downs and dot blot assays. Interactions between TbSTI1 and TcHsp70 and TbHsp70.4 but not with TbHsp70.c were suggested by SPR and dot blot analysis.

CHAPTER 5

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

This study successfully addressed numerous knowledge gaps regarding *T. brucei* molecular chaperones, extended the scope and provided novel insight into potential chaperone co-chaperone partnerships. Firstly, an *in silico* analysis of the Hsp70/J-proteins complex in *T. brucei,* coupled with side-by-side comparisons between other kinetoplastid micro-organisms, human, yeast and *P. falciparum,* was conducted. Secondly, an in-depth bioinformatics study of the cytoplasmic molecular chaperones, particularly focusing on Hsp70s, Hsp83 and TPR containing co-chaperones of these heat shock proteins was carried out. Thirdly, the first heterologous overproduction and purification of TbHsp70.4 and TbHsp83 was successfully achieved prior to *in vivo* characterization and interaction studies. TbSTI1 was also extensively characterized using *in silico* techniques, as well as successfully overproduced and purified in an *E. coli* expression system prior to preliminary *in vitro* and *in vivo* biochemical characterization.

5.1. *In silico*

An update of the current status of Hsp70 entries in TriTrypDB was provided in this study, particularly for *T. brucei* which was compared to other kinetoplastids, human, yeast and *P. falciparum*. Most of the results obtained were consistent with previous studies (Folguiera and Requena, 2007, Louw et al., 2010). However, two aspects were added to complement the current knowledge of kinetoplastid Hsp70 systems. Firstly, this is the first report comparing two strains of *T. cruzi* to demonstrate that *HSP70* genes in these micro-organisms are not duplicated. Secondly, it was confirmed that the canonical TbHsp70 is indeed a typical eukaryotic cytosolic Hsp70, possessing the C-terminal end EEVD motif instead of the previously reported RRHI (Louw et al., 2010). A high level of sequence and structural conservation of Hsp70s was shown while a greater emphasis was placed on cytosolic Hsp70 proteins.

TbHsp70.c was proposed to be a ribosome-associated Hsp70, with its homology demonstrated using phylogenetic relationships. The uniqueness of TbHsp70.c however, lies in the fact that it has an extended substrate binding domain different from its homologues, a highly charged SBD

that possess a DDND that aligns with the EEVD motif which would normally be involved in associating with STI1. Further assessment using *in silico* analysis showed that it is unlikely that TbHsp70.c will interact with HsSTI1. The constitutive Hsp70.4 and canonical Hsp70 in *T. brucei* were also characterized in depth and the three-dimensional structures of these proteins were generated. These models were demonstrated to be accurately predicted and the possibility of these Hsp70s binding to STI1 was explored. A weak interaction between TbHsp70.4 and TbSTI1 was predicted due to the absence of the EEVD in this particular protein, although the variant DDVD was demonstrated to be sufficient for associating with STI1 and other TPR containing co-chaperones. The canonical TbHsp70-TbSTI1 system common to other eukaryotes was predicted to be conserved.

A comprehensive assessment of the J-proteins in *T. brucei* had never been conducted previously. A thorough *in silico* analysis classified J-proteins into the four recognized types based on the presence of the J-domain, GF region and Zn binding domain. This study also gave connections and potential partnerships between J-proteins and Hsp70s based on co-localization (predicted) and comparison with homologues. Indeed several features of these proteins were compared to other kinetoplastids. *T. brucei* has an expanded J-protein complement and also several unique members. This study also complemented work by Folguiera and Requena, (2007), who identified 6 type I J-proteins in *T. brucei*. As expected, the domain architecture of type I J-proteins were found to be more conserved than all the other types leading to the identification of higher eukaryotes homologues in almost all 6 J-proteins except for Tbj27. This study is also the first to report the presence of a *bona fide* J-domain in Tbj27 through comparisons of the sequence entries between *T. b. brucei* and *T. b. gambiense*. Two proteins previously defined as type I J-proteins (Tbj47 and Tbj66) are proposed to, in fact, not be J-proteins due to the absence of the J-domain. The presence of all other features defining a type I J-proteins is certainly intriguing and may warrant them to be classified as “type I J-protein-like”. The majority of type II J-proteins in *T. brucei* are predicted to localize in the cytoplasm, while a previously unreported type II J-protein, Tbj72, was identified in this study. A greater conservation was observed for type I and type II J-proteins in *T. brucei* most of which possessed homologues in other eukaryotes. The

conserved nature of these J-proteins may imply a maintance of similar functionality through evolution.

T. brucei and *T. evansi* display an expanded complement of type III J-proteins, most of which were predicted to localize in the mitochondria. As the J-domain is the only defining feature of type III proteins, an alignment was derived to identify residues previously shown to be involved in the maintenance of the structural integrity of the J-domain (Hennesy et al., 2004). It was observed that residues such as LEU10 were more conserved than TYR7 (*E. coli* DnaJ numbering) in *T. brucei* (Appendix F, Figure F.4). Interestingly, the majority of type III J-domains had the TYR7 to LEU substitution (Tbj9, Tbj26, Tbj48, Tbj58, Tbj59 and Tbj69). Substitution to other aromatic amino acids, TRP (Tbj10 and Tbj12) and PHE (Tbj14, Tbj36 and Tbj38) were also observed (Appendix F, Figure F.4), while the newly identified Tbj73 did not have either of the residues implicated in the maintenance of the structural integrity conserved. Even more intriguing was the variability observed in residues shown to be critical for Hsp70/J-protein interactions (Genevaux et al., 2002; Hennessy et al., 2004). The LYS/ARG26 was found to be frequently substituted with a HIS or VAL residue (Appendix F, Figure F.4).

The KFK motif found in helix III was also found to be poorly conserved as was helix IV. Common occurrences of -KF/-FK motifs were seen (Appendix F, Figure F.4). The residue ARG63 which has been shown to be crucial for J-domain function (Hennessy et al., 2005) was found to be largely conserved in *T. brucei* type III J-domains. The identification of these notable differences between the J-domains of *T. brucei* type III J-proteins, compared to canonical J-domains provides a large scope for future studies. The lack of conservation in residues implicated for Hsp70/J-protein specificity potentially serves to define discrete partnerships for these proteins. What role these substitutions play still remains an open question and thus, future studies should employ site-directed mutagenesis to re-introduce canonical J-domain residues to those differing in *T. brucei*. Overall, the large variability observed in Tbj19, Tbj40, Tbj69 and Tbj73, which all lack homologues, allows them to be prioritized for future studies. It was also interesting to note the lack of

conservation of the recognized J-domain features in Tbj12 and Tbj41, since these proteins are only found in extracellular kinetoplastids. Their characterization may reveal some important details regarding the life-cycle. Structural bioinformatics docking the unique J-domains to potentially interacting Hsp70 ATPase domains is also recommended for future study for early screening of functionality and specificity of these proteins. The identification of Hsp70 and J-proteins substrates in *T. brucei* cannot be underestimated and will provide clues regarding the need for an unusually large number of J-proteins in this organism.

Little is still known about type IV J-proteins which have been described to appear in large complement (n = 12) in *P. falciparum* (Botha et al. 2007). This study is the first to identify two type IV J-proteins and an additional two type I J-like proteins in *T. brucei*. The latter is described as proteins containing GF region and Zn binding domain but lacking the signature J-domain. The type IV and J-like proteins have orthologues in other kinetoplastids, except *T. cruzi* which has no orthologue of Tbj66. It is not clear whether these proteins are part of a J-protein specialized mechanism of action or not. However the hypothesis that the J66 protein family is the C-terminal segment of a J-protein and therefore it is likely to interact with a separately transcribed and translated J-domain makes this group of J-proteins of fundamental interest to molecular chaperone research (Michael Ludewig, PhD thesis, 2010). Unlike malaria parasite systems, the *T. brucei* type IV J-proteins do not contain a PEXEL motif. However, it is still noteworthy questioning whether these co-chaperones may be interacting with other ATPases such as actin and myosin but not Hsp70 (Pesce and Blatch, 2014). It is also possible that these J-proteins may interact with Hsp70s from other species which would certainly be interesting, particularly in the case of the extracellular kinetoplastids. Tbj31 and Tbj68 which have been shown to possess the greatest level of similarity to canonical J-domains (Appendix F, Figure F7 and Figure F8) should certainly be pursued for future studies. Tbj47 which has been proposed as a drug target (Chapter 2), the pursuation of Hsp70/J-protein interactions are now considered good chaperone targets in other diseases. Tbj47 has also been shown to be to be essential for *T. brucei* viability *in vivo* (Subramaniam et al., 2006) should be biochemically characterized. Binding assays to determine

whether the lack of the J-domain or the abrogated J-domains are required for the proteins to successfully bind Hsp70 proteins or other ATPases should be pursued.

It is generally accepted that Hsp90 (or Hsp83 in kinetoplastids) associates with numerous co-chaperones that regulate its function in the multi-chaperone complex. To investigate the presence of the STI1-mediated Hsp83 hetero-complex in *T. brucei*, TPR containing co-chaperones of Hsp83 and/ Hsp70 were identified using their eukaryotic homologues. Indeed several of the TPR containing co-chaperones of Hsp70 and Hsp83 were shown to be conserved with respect to the canonical domain architecture found in their homologues. The possibility of TPR containing co-chaperones of *T. brucei* associating with Hsp70 and Hsp90 was supported by the demonstration that carboxylate clamp forming residues were also conserved in these co-chaperones. Only TbSGT was atypical, it has an additional TPR motif while Tbj42, Tbj65 and Tbj67 displayed the most variability with regards to residues involved in Hsp70 and Hsp90 interaction. Unlike other systems where SGT has been shown to associate with both Hsp70 and Hsp90, the orthologue of TbSGT, in *L. dononani* has been shown to interact directly or indirectly with LdSTI1, LdHip and LdHsp70 (Ommen et al., 2009). It is suspected that the additional TPR motif found in kinetoplastids SGT could play a role in this fascinating mechanism of interaction. Future studies exploring the roles of TbSGT in modulating the multi-chaperone complex should be carried out.

5.2. *In vitro*

This study aimed at characterizing TbSTI1 based on its interactions with Hsp70s and Hsp90 with a view to confirming the presence of the STI1 mediated Hsp90 hetero-complex in *T. brucei*. Therefore, this is the first study in which TbSTI1 was expressed and purified successfully. This study also provided the first heterologous expression and purification of TbHsp83 and TbHsp70.4. With regards to TbHsp83, although a small yield of the protein was obtained it is recommended that future studies explore purification techniques amongst other optimizations, the purification of the protein using ion exchange chromatography. Other studies have obtained sufficient yields

of parasitic Hsp90s with few contaminants by employing a purification technique where the Hsp90 was purified by nickel affinity chromatography on a FPLC through His-trap columns followed by ion exchange chromatography (Pallavi et al., 2010, Meyer and Shapiro, 2013). Different cell lines and growing the bacterial cells at lower temperature might also aid the purification of TbHsp83. The bands co-purifying with STI1 and TbHsp70.4 could be identified using mass spectroscopy. The purifications of TbHsp70.c and TcHsp70 were conducted as described by Burger et al (2014).

The interactions of STI1 with Hsp70s were characterized using SPR, pull down assays and dot blots assays. Based on the dot blots assays, a strong interaction between STI1 and TcHsp70 was seen, while a weak interaction was seen for TbHsp70.4. The results obtained from SPR were qualitative; therefore using a range of concentrations (from very low to high) of the proteins would need to be employed in repeated experiments to obtain quantitative data. However, an additional technique to validate SPR results is isothermal titration calorimetry (ITC). This technique can be used to determine the interactions between TbSTI1 and Hsp70s, although it would require even greater concentrations of protein than used in the SPR analyses (Brown, 2009; Pagano et al., 2009). ITC has advantages over SPR and is considered an optimal technique for measuring biomolecular interactions as there are a lack of buffer restrictions or weight limitations and importantly, there is no need for labeling or immobilization as required by SPR, since both the ligand and the analyte are in solution, making it an approach of choice for many researchers (Qian et al., 2002; Ball and Maechling, 2009; Brown, 2009; Magotti et al., 2009).

The study was unable to detect human Hsp70 and Hsp90 using TbSTI1 as a bait protein in Hs578T cancer cell lysate, possibly due to the presence of high concentrations of human Hop in complex with these proteins. Knock-down cells of human Hop could be used in future studies with TbST1 as bait to detect whether this protein interacts with human Hsp70 and Hsp90. Future biochemical characterization of the TbSTI1, TbHsp70 and TbHsp70.4 partnerships can be characterized using assays such as ATPase, substrate aggregation suppression, refolding and complex formation .

5.3. *In vivo*

This study successfully detected TbHsp70.c in the bloodstream stages of *T. brucei* (Alsford et al., 2011). Previous studies also revealed that protein expression of TbHsp70.c was enhanced upon heat shock, suggesting its importance in protein quality control under stressed conditions in the cell (Adelle Burger, 2013, PhD Thesis). TbHsp83 and TbHsp70 are prominent chaperones whose association is predicted to be co-ordinated by STI1 as in other cellular systems. Although the study successfully detected TbSTI1, TbHsp83 and TbHsp70.c at the bloodstream stage of parasite development, the function of TbSTI1 as a mediator of TbHsp70 and TbHsp83 remains to be confirmed for *T. brucei*. The occurrence of these proteins however, gives evidence that the STI1-mediated hetero-complex exists. Hsp83 influences the temperature-sensitive differentiation of insect to mammalian form in *Leishmania* and *T. cruzi* (Graefe et al., 2002; Li et al., 2009). Future studies should involve the design and production of antibodies against TbHsp70, in this way, this chaperone could be detected in the parasite lysate. It will also be established whether its expression is stage-specific or it is found through out the life-cycle of the parasite. Following which, co-immunoprecipation studies will be beneficial especially for providing evidence that the STI1 mediated Hsp90 hetero-complex is present in *T. brucei*. Based on the multi-chaperone complex predicted in this study, Tbj2, a type I J-protein which has been shown to localize in the cytoplasm and its expression enhanced upon exposure to heat shock (Ludewig et al., 2015), is expected to be detected by co-immunoprecipation studies wherein molecular chaperones of STI1 mediated Hsp83 hetero-complex will be detected. The detection of a J-protein in the STI1 mediated Hsp90 hetero-complex has also been seen in *P. falciparum* (Gitau et al., 2011). The localization of these molecular chaperones under normal conditions and temperature, pH and starvation stresses need to be explored in the future. The co-localization especially under stress conditions will also contribute significantly to the biology of the parasite. The *in silico* and *in vivo* analysis conducted in this study has led to the prediction of a model of the multi-chaperone complex (Figure 5.1). The absence of TbHsp70.4 in the bloodstream form *T. brucei* lysates leads to the prediction that this particular chaperone is expressed in the PRO stage of the life-cycle. The same has also been seen for its *L. major* homologues which is detectable at low-levels in the

bloodstream form but highly enriched at the promastigote phase of the life-cycle. Taken together with positive interaction between TbSTI1 and TbHsp70.4 obtained in the *in vitro* studies, it is suspected that the STI1 mediated Hsp83 hetero-complex exist at that PRO stage of parasite life-cycle (Figure 5.1). It would be interesting to check if TbHsp70 is expressed in the PRO stage of *T. brucei* life-cycle.

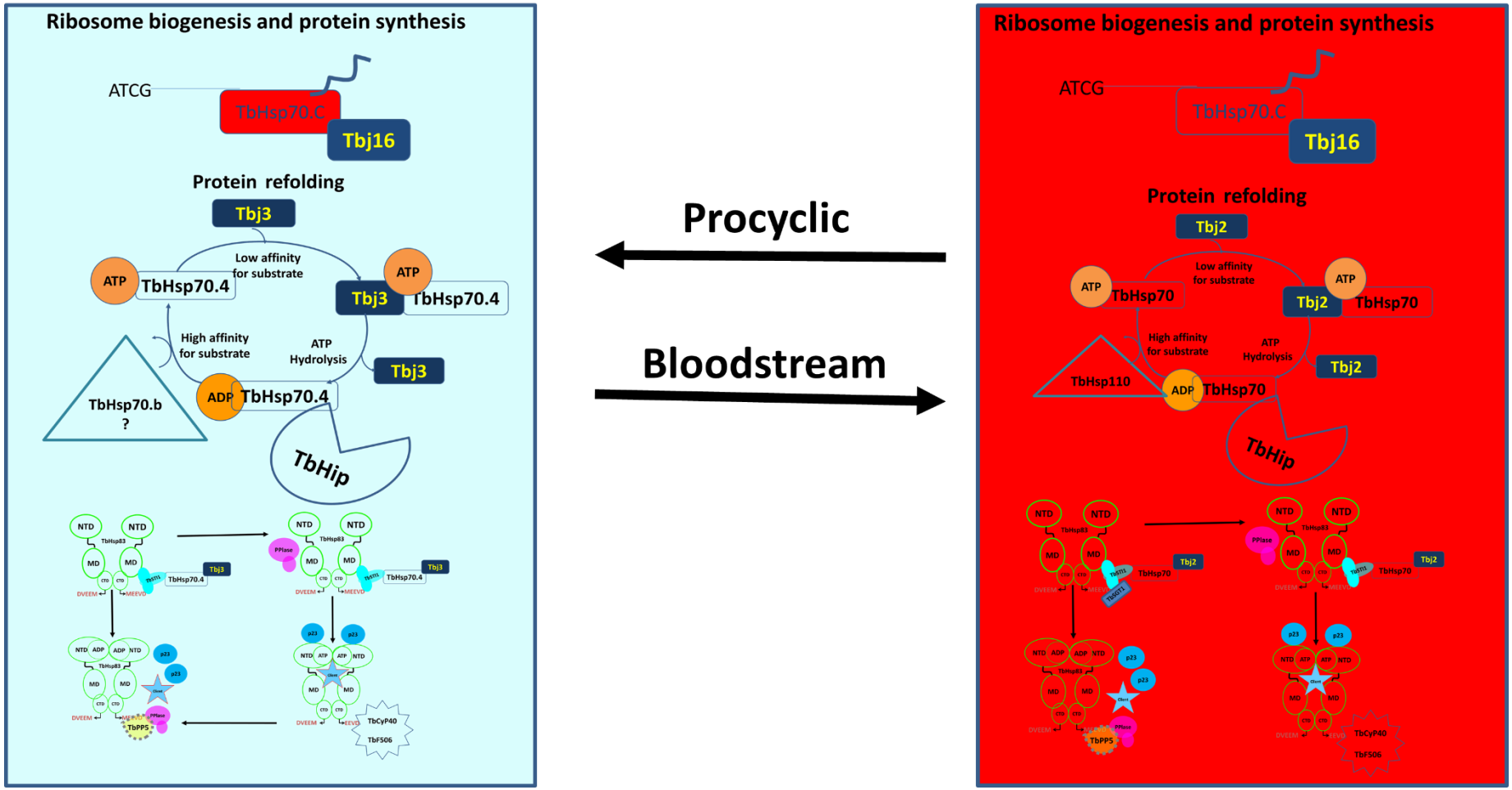


Figure 5. 1: A model for the multi-chaperone complexes taking place in the cytoplasm of *T. brucei* during the procyclic stage (Blue) and the bloodstream stage (Red) of parasite life-cycle.

APPENDICES

Appendix A: AMINO ACIDS AND NUCLEOTIDE NOMENCLATURE

One and three-letter codes were used to represent amino acids, and single letter codes were used to represent nucleotides as set forward by the Joint Commission of Biochemical Nomenclature (JBNC) of IUPAC (International Union of Pure and Applied Chemistry) and the IUBMB (International Union ofBiochemistry and Molecular Biology):

Table A1: Nucleotide single letters

NUCLEOTIDE	SINGLE-LETTER CODE
Adenine	A
Cytosine	C
Guanine	G
Thymine	T
Uracil	U
Any nucleotide	A, C, G, T or U, N

Table A 2: Amino acids representation

AMINO ACID	1-LETTER CODE	3-LETTER CODE	DNA CODONS
Alanine	A	ALA	GCT, GCC, GCA, GCG
Arginine	R	ARG	CGT, CGC, CGA, CGG, AGA, AGG
Asparagine	N	ASN	AAT, AAC
Aspartic acid	D	ASP	GAT, GAC
Cysteine	C	CYS	TGT, TGC
Glutamine	Q	GLN	CAA, CAG
Glutamic acid	E	GLU	GAA, GAG
Glycine	G	GLY	GGT, GGC, GGA, GGG
Histidine	H	HIS	CAT, CAC
Isoleucine	I	ILE	ATT, ATC, ATA

Leucine	L	LEU	CTT, CTC, CTA, CTG, TTA, TTG
Lysine	K	LYS	AAA, AAG
Methionine	M	MET	ATG
Phenylalanine	F	PHE	TTT, TTC
Proline	P	PRO	CCT, CCC, CCA, CCG
Serine	S	SER	TCT, TCC, TCA, TCG, AGT, AGC
Threonine	T	THR	ACT, ACC, ACA, ACG
Tryptophan	W	TRP	TGG
Tyrosine	Y	TYR	TAT, TAC
Valine	V	VAL	GTT, GTC, GTA, GTG
Stop	-	-	TAA, TAG, TGA
Any Amino Acid	X	-	-

Appendix B: Recipes

Yeast-Tryptone (YT) Broth growth medium

Tryptone	16g/L
Yeast Extract	10g/L
NaCl	5g/L

Make up to 1L with water and autoclave (121°C and 119 kPa for 20 minutes)

Yeast-Tryptone (YT) Agar

A similar recipe to that of YT broth was adopted except 15g bacteriological agar was added per liter of broth. Autoclave the solution (121°C and 119 kPa for 20 minutes).

RF1 250 ml (pH 5.8) (Store at 4 °C)

(100 mM KCl, 50 mM $MnCl_2$, 30mM CH_3COOK, 10mM $CaCl_2$, 15% glycerol)

2.45 g or 7.5 ml of 1M CH_3COOK and pH to 5.8 with HCl or acetic acid

Add 37.5ml glycerol and make up to a final volume of 202.5 ml and Autoclave (121 °C, 119 kPa for 15 to 20 minutes). Consequently autoclave, add the following stocks (autoclaved):

25 ml 1M KCl
12.5 ml 1M MnCl2
2.5 ml 1M CaCl2

RF2 150 ml (pH 6.8) (Store at 4 °C)

(10 mM MOPS buffer pH 6.8, 10 mM KCl, 75 mM CaCl2, 15% glycerol)

0.313 g or 1 ml 1M MOPS (pH to 6.8 with KOH)
Add 22.5 ml glycerol

Make up to 150 ml with distilled water and autoclave (121 °C, 119 kPa for 15 to 20 minutes. Consequently autoclave, add the following stocks (autoclaved):

1.5 ml 1M KCl
11.25 ml 1M CaCl2

Appendix C: Organisms

Organism	Strain	Genotype	Referrence
E. coli	JM109	*recA1 supE44 endA1hsdR17gyrA96 relA1thi _(lac-proAB) F'[TraD36 proAB+ lacIq lacZ _M15]*	
E. coli	XL1 Blue	*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac- F' [proAB+ lacIq lacZ _M15 Tn10 (tetr)]*	Bullock et al.1987
E. coli	M15[prep4]	*lac,ara,gal,mtl,recA+,uvr+ [pREP4,lacI,kanar]*	
E. coli	BB1994	*MC4100 dnaK52 sidB1::Tc pDMI,1::CmR KanR*	Dr M Mayer
E. coli	BL21	*F- ompT gal [dcm] [lon] hsdSB λDEs*	Studier et al 1990
T. brucei	V3	SMB, T7RNAP::TETR::NEO	Wirtz et al., 1999
T. brucei	V221	SMB, T7RNAP::TETR::NEO	Wirtz et al., 1999

Appendix D: List of all materials and specialized reagents

Restriction Endonuclease	Supplier
*Kpn*I	Promega, USA
*Nde*I	Thermo Scientific, USA
*Not*I	Thermo Scientific, USA
*Sal*I	Promega, USA
*Xba*I	Thermo Scientific, USA
*Xho*I	Thermo Scientific, USA

Primary antibodies	Supplier
Mouse anti-His monoclonal antibodies	Santa Cruz biotechnology, USA
Rabbit anti-TbHsp70.c polyclonal antibody	Burger et al., 2014
Rabbit anti-TbHsp70.4 monoclonal antibody	Bangs et al., 1993
Mouse anti-Hsp70/Hsc70 monoclonal antibody SC-24	Santa Cruz biotechnology, USA
Mouse anti-Hsp90α/β SC-13	Santa Cruz biotechnology, USA
Secondary antibodies	**Supplier**
HRP-Conjugated goat anti-mouse antibody	Santa Cruz biotechnology, USA
HPR-conjugated goat anti-rabbit antibody	Santa Cruz biotechnology, USA

Appendix E: General experimental procedure

E1: Preparation of competent bacterial cells

A specific strain of bacterial cells was grown overnight at a suitable temperature at 200 rpm in 5ml of 2xYT or LB media supplemented with appropriate antibiotics for selection if required. The overnight culture was then diluted 10x into 50ml 2xYT or LB media to an A600 of 0.1. Growth was then allowed to occur until early log phase (A600 of 0.3 - 0.6). Subsequent to this, cells were harvested by centrifugation (5000 xg, 5 munites) at 4®C. The cells were then resuspended in 50 ml of ice-cold 0.1 M MgCl2 (4°C) and incubated on ice for an hour. Following the incubation, the cells were pelleted by centrifugation (5000 xg,

5 munites) at 4®C and resuspended in 25 ml of ice-cold 0.1 M CaCl2 (4°C). Following incubation at 4°C for one hour the cells were harvested by centrifugation as before and resuspended in 5 ml of 0.1 M CaCl2 and 5 ml of 30% (v/v) glycerol. The competent cells were divided into aliquots and stored for future use at -80°C.

E2: Transformation of competent bacterial cells

An aliquot (100µl) of competent *E. coli* cells were incubated with 50-100ng of the plasmid DNA of interest at 4®C for 30 minutes, followed by heat shock (42°C for 45 seconds) and subsequent cold shock (4°C for 2 minutes). To these cells 900µl of 2xYT or LB media were added, followed by incubation at the appropriate growth temperature for 1 hour at 200 rpm. The bacterial suspension was plated on 2xYT or LB plates containing appropriate antibiotics (100 µg/ml ampicillin for pQE-based plasmid selection; 50 µg/ml kanamycin for *E. coli* M15[pREP4] and BB1994 cells). The plates were incubated in a suitable temperature overnight. Transformation controls included a sterile control with sterile distilled water replacing the plasmid DNA in the incubation mixture and the appropriate plasmid vector without insert.

E3: Isolation of plasmid DNA

Plasmid DNA was isolated using the Thermoscientific GeneJet Plasmid Miniprep kit. Briefly, transformed *E. coli* cells were grown overnight (37®C, 200rpm) in 2x YT (1.6% tryptone, 1% yeast extract, 0.5% NaCl) or LB media (1% tryptone, 0.5% yeast extract, 1% NaCl) supplemented with the appropriate antibiotic for plasmid selection (100 µg/ml) ampicillin for pQE-and pET14b based plasmids). The cells were harvested in a table top microcentrifuge (13 000xg, 1 minute) and the cell pellet resuspended in 250µl of the resuspension solution (2 mg/ml lysozyme, 10 mM EDTA, 50 mM Glucose, 25 mM Tris-HCl, pH 8.0). To the resuspended cells, 250µl of lysis solution % (w/v) SDS, 0.2 M NaOH) was added and mixed 4-6 times by inversion. This was followed by the addition of 350 µl neutralizing solution 1.5 M potassium acetate, 12% (v/v) glacial acetic acid) and mixed immediately by inverting 4-6 times. Cell debri and chromosomal DNA were removed by centrifugation (13 000 xg, 5 minutes). The supernatant was subsequently applied to the GeneJet spin column via pipetting and centrifuged (13 000 xg, 1 minute). The flow through was discarded and the column washed twice with wash solution (70 % ethanol), residual wash solution was then removed by centrifugation (13 000 xg, 1 minute). The pellet was then allowed to dry for 20 minutes at room temperature prior the addition of 50µl of TE buffer (10 mM Tris, 20 µg/ml RNAse A, 1 mM EDTA, pH 8.0). The purified DNA was quantified at 260 nm in a Helios Alpha UV-Vis Spectrophotometer (Thermo Scientific).

E4: Agarose gel electrophoresis (AGE)

Agarose gels were prepared by melting molecular grade agarose (1% (w/v)) in TAE buffer 40 mM Tris, 20 mM acetic acid, 1 mM EDTA; pH 7.6) and supplementing with ethidium bromide to a final concentration of 0.5 µg/ml prior to cooling to casting. Prior loading, DNA samples for electrophoresis were treated with 6x DNA gel loading buffer (0.25 % (w/v) bromophenol blue, 30 % (v/v) glycerol; pH 7.6) in a ration of 5:1 respectively. The samples were loaded with *Pst*I-digested λDNA marker which was prepared by the digestion of 20µl of 526 µg/ml λDNA for two hours at 37°C in a reaction containing 5 U of *Pst*I restriction enzyme (Fermentas), 20 µl of the appropriate 10x restriction enzyme buffer (Fermentas) and distilled water to a final volume of 200 µl.The samples were resolved at 100 V for an 45 minutes and visualised under ultra violet light with a Chemidoc Imaging System (BioRad).

E5: DNA digestion with restriction enzymes

Plasmid DNA was digested with the appropriate restriction endonuclease (s) for two or more hours at the optimal temperature (37°C unless otherwise indicated).The restriction digest reaction was comprised of 200-500ng of plasmid DNA, 2 µl of a compatible restriction buffer, 1-2 U of restriction endonuclease enzyme (s) and distilled water to a final volume of 20µl. The digested DNA was resolved by agarose gel electrophoresis as described in section E4.

E.5.1. Plasmids encoding 6x His-tag TbSTI1 and mSTI1

For the production and purification of 6x His-tagged mSTI1 and TbSTI1, the integrity of the pQE200 plasmid expressing the former recombinant protein was confirmed by restriction digest using *SphI* and *Sal* I. The TbSTI1 insert within pQE2 (Qiagen, U.S.A) was ligated between *Nde* I and *Not* I sites which were used to verify the integrity of the plasmid by restriction digest.

E.5.2. Plasmids encoding 6x His-tag TcHsp70, TbHsp70.4 and TbHsp70.c

For the verification of the pE14b-TcHsp70 plasmid, a restriction digest using *EcoR*I and *Nde*I was carried out and the integrity of the pQE80-TbHsp70.c was confirmed by restriction digests using *Kpn* I and *Sal* I sites for the reaction. The pQE2-TbHsp70.4 was ligated between *Nde* I and *Xho* I sites within the pQE2 expression vector which were then used to release the TbHsp70.4 insert .

E.5.3. Plasmids encoding 6x His-tag TbHsp8

The C-terminal His-tag pQE80 contained TbHsp83 insert between the *Bam* HI and *Sal* I while the TbHsp83 was ligated between *Bam* HI and *Bgl* II site into the N-terminal His-tag pQE60. The appropriate restriction enzymes were used to check the plasmid integrity using restriction digest.

E6: DNA sequencing

Plasmid DNA was isolated for DNA sequencing using the Thermoscientific GeneJet Plasmid Miniprep kit as described in section E1, with the exception of the DNA elution step. A final volume of 50µl of the DNA was eluted using distilled water. DNA sequencing was carried out commercially by Inqaba Biotech (SA). DNA sequencing results were analyzed using either the BioEdit Sequence Alignment Editor (version 7.0.4.1) or Chromas pro software.

E7: Sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE)

Samples for protein production and purification were analysed on by 10% SDS-PAGE as adapted from Laemmli (1970) and (Shapiro, et al., 1967). Protein samples were treated with 5x SDS-PAGE sample buffer (10% glycerol (v/v), 2% SDS (w/v), 5% β-mercaptoethanol (v/v), 0.05% bromophenol blue (w/v), 0.0625 M Tris-HCl (w/v), pH 6.8) in a ratio of 4:1 respectively. Treated protein samples were boiled (100®C, for 10 minutes) and loaded onto a 10% resolving gel (10% (w/v) acrylamide, 0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate (APS), 0.005% (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED), 0.375 M Tris, pH 8.8) and a stacking gel (4% (w/v) acrylamide, 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.005% (v/v) TEMED, 0.125 M Tris, pH 6.8). The gel was run in 1 x SDS running buffer (250 mM Tris, 1.9 M glycine, 10 % (w/v) SDS) and resolved in a Mini ProteanR II system (Bio-Rad) at 150 V for one hour and stained or used for Western analysis (Section E8). The SDS-PAGE gel was stained in Coomassie Blue stain (40% (v/v) methanol, 7% (v/v) acetic acid, 0.25% (w/v) Coomassie Blue R250 in distilled water) for 30 minutes and subsequently destain in destaining solution (40% (v/v) methanol, 7% (v/v) acetic acid in distilled water).

E8: Protein detection by western analysis

The protocol of western blot detection of protein was described by Towbin et al., 1979 and was adapted. Briefly, resolved proteins from SDS-PAGE were transferred onto nitrocellulose membrane in transfer buffer (20% (v/v) methanol, 192 mM glycine, 25 mM Tris) at 100 V for 90 minutes in a Mini ProteanR III Western trans-blot system (Bio-Rad). The transfer of protein was confirmed with Ponceau S stain (0.5 % (w/v) Ponceau S, 1% (v/v) glacial acetic acid). The membrane was then destained in distilled water and subsequently incubated with 5% (w/v) fat-free milk powder in Tris Buffered Saline (TBS; 50 mM Tris, 150 mM NaCl, pH 7.5), blocking solution overnight. The membrane was incubated with the appropriate primary antibody (1:5000 in blocking solution unless stated otherwise in text) for one hour at room temperature and subsequently washed three times with Tris Buffered Saline-Tween buffer (TBS-T; TBS containing 0.1% (v/v) Tween 20). The membrane was similarly incubated with the appropriate horse-radish peroxidise (HRP)-conjugated secondary antibody (1:5000 in blocking solution unless otherwise stated in the text) for one hour at room temperature and washed with TBS-T as before. Chemiluminescence-based protein detection was achieved using the ECLTM Western blotting kit (GE

Healthcare) as per the manufacturer's instructions, and captured with a Chemidoc chemiluminescence imaging system (Bio-Rad).

Appendix F: Bioinformatics supplementary data

Table F.1: Target signals for the nuclear, mitochondrial and ER target sequences as obtained from NLS-MAPPER, MITOPROT II and SignalP

Gene	Target signal
Tbj1	106-VERKRKEDEE-116
Tbj5	Exported to the mitochondria
Tbj8	1-MQLHLSSISMMCLLLVTLQLSFTPKVTNA S-30
Tbj9	**1-MPLFCSSTTFMLKVRRAAVKTWKCFAFIAFVSSSMRYISCGFTAD-46**
Tbj10	**1-MWRVALRRPRVAMPCGCIGHIYRRG-26**
Tbj12	**1-MTTYGHRPEHMLHAWRTLGLRANPT-26**
Tbj14	153-EAARKKQRLEREAKEKAELAAELERQRKEWKD-184
Tbj15	54-LRSPTKRRKYDQE-66
Tbj17	22-ETALRKAFRQKALDLHPDRNPNGADEFKKVNE-53
Tbj18	**1-MRRFVSQQDGKATALWTNLPHARVSSAFFSSLTHESRVVQSVTCVRRVLAPPTPSIYPAFRNTFKSTKRW-71**
Tbj20	380-LGTDSKRRKYD-401
Tbj23	**1-MWRSQCCLLVGGRTCRRLLTPTSSRRMAAGAMGAAPLFVQLRL-44**
Tbj24	**1-MQKNQTAYYRTLGVNREAT-20**
Tbj25	**1-MRRHVSLVWLNAATGWRL-19**
Tbj26	**1-MKPIPVAMNRSVRRAALMTLQPLVCSGTSARCITTGENHSFRPGRVTLFLRSLFVALTPQRLRSRM-67**
Tbj27	**1-MRQRSPLFQQYVRRAIPSTFSLVPSRHFTAPGVRPNTVCASGSIICTAVVSAGLGVSRRFSATNAKKDLYSVLGVARN-79**
Tbj28	**1-MSSGLTARRLILQQLTAVTQHIPLQARRTSGKGWWLSSFVRP-43** 58-AFLKRRCEEE-67
Tbj29	**1-MFFVTSIVRMTVHAPRGIPFQITTG-26**
Tbj30	**1-MSYGKWAGVYSPHFLMSHVTMCRRVAHILWRRPSCSPQRSRS-43**
Tbj32	318-EKKRSCPACKKQFK-331
Tbj33	241- AGKKKKREQ-248
Tbj35	**1-MHADVCIVIHSAVSFQPINHRIGRA-26**

Tbj36	**1-MVVVGMNRRFLATSAWLG-19**
Tbj37	**62-LSNEEKKRKYD-72**
Tbj38	**1-MRSSFFLLSKDPFAVLGLART-22**
Tbj40	**1-MFTYSTFLRSVSVGQACRL-20**
Tbj41	**1-MFVLCTVENMGKSNGCNPSGRKRRR-26**
Tbj42	**538-RALAMKWHPDKWCSASAQEQKEAEEKFKLVKA-569**
Tbj43	**206-RPRVGKRIPPKAPSGCVKRSRA-227**
Tbj45	**1-MRGLAAAHTSNTKLVSFLFVSFACLGFAPVVHS H-34**
Tbj46	**1-MRCTMSSPLFRLTRLSVVITIAFFRNVHAE-30**
Tbj47	**1-MRGITLALAPP-19**
Tbj48	Exported to the mitochondria
Tbj49	**1-MYVFRFHISSFHSFLVNFLQPSRG-25**
Tbj50	**1-MLRFTSVSSIWRRLAAAPPTAATAAFANVSKRQ-34**
Tbj53	**1-MYITGENSAIKCRYNMILKILLLVGAAVGAECS-33**
Tbj58	**1-MLSRSAWVFAMSPRAALTA-19**
Tbj60	**1-MRSWAAFRSVRSLQRRF-18**
Tbj63	Exported to the mitochondria
Tbj65	**1-MSHHIQLPCFAETITSSGTLFPNISGGKPQKKPEMNSVRHPQRPVRTQKKRNKPMSRV-59**
Tbj67	Exported to the mitochondria
Tbj68	**1-MAAPLAALVLLGGAYYIFRLAPRITQRVSMAQGLTCAANRQLRPYRRY-49**
Tbj69	**113-KAGRRRRRV-121**
Tbj70	**309-LRKELKKRRVL-319**
Tbj71	**1-MFKCRVAVLLSSASSSATGS-20**
Tbj73	**417-KAPTKRVVKPVKKAGKKRPRR-437**

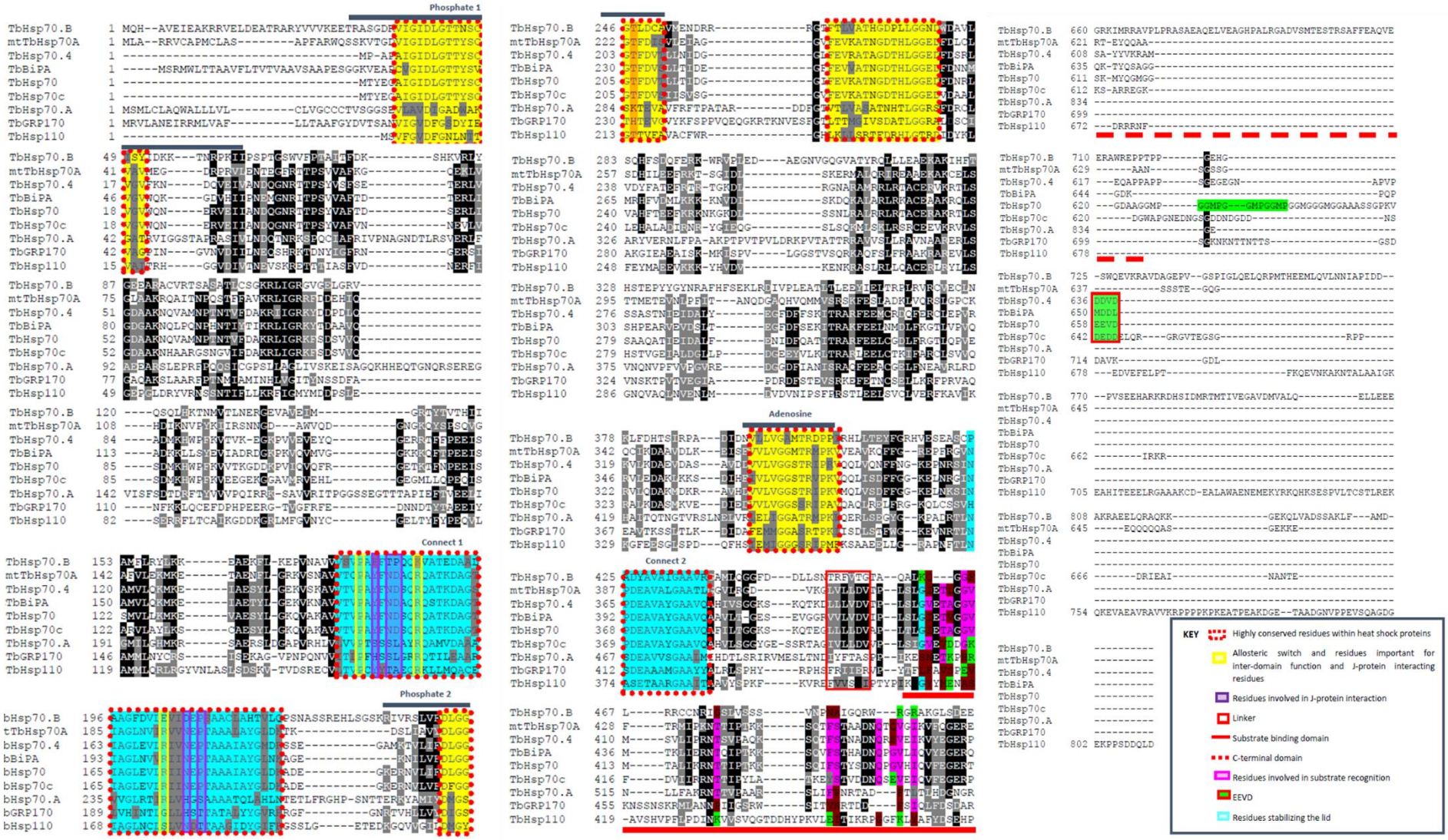

Phosphate 1
Connect 1
Phosphate 2
Adenosine
Connect 2
KEY
Highly conserved residues within heat shock proteins
Allosteric switch and residues important for inter-domain function and J-protein interacting residues
Residues involved in J-protein interaction
Linker
Substrate binding domain
C-terminal domain
Residues involved in substrate recognition
EEVD
Residues stabilizing the lid

Figure F1:Amino acid sequence alignment of *T. brucei* Hsp70s comprising of HspA and HspH members of the Hsp70 superfamily. The multiple sequence alignment was conducted against full-length Hsp70 sequences which comprises of the ATPase domain, linker region (red box), substrate binding domain (red lines underneath the sequences) and the C-terminal domain (red dashes underneath the sequences). Some of the key residues in each of the domains are highlighted and annotated as follows: The ATPase domain residues associated with heat shock proteins (yellow highlights), Proline (the allosteric switch residue) and Arginine which are involved in interdomain function (Vogel et al., 2006a) are also shown in yellow highlights. Residues involved in J-proteins binding are indicated in a purple box. In *E. coli* DNAK, residues involved in DNAJ binding are Y145, N147, D148, N170 and T173 (Gässler et al., 1998; Suh et al., 1998) and T199 (highlighted in red), a DNAK phosphorylation site (McCarty and Walker, 1991). The three motifs interacting with ATP at the adenosine, phosphate 1 and phosphate 2 are indicated in yellow background which is boxed in red dots and the segments linking nucleotide binding motifs connect 1 and connect 2 are indicated cyan highlights which are boxed in red dashes. The linker regions which connects the two domains in shown in a red box. The least conserved substrate binding domain and C-terminal domain are shown by red lines and red dotted lines beneath the sequences respectively. Residues for *T. brucei* Hsp70 residues which are identical to the *E. coli* DNAK for substrate binding are indicated in magenta highlights, those which differ from the DNAK residues are shown in dark red and green for the substitution which involves a charged residue. With regards to the C-terminal domain the GGMP regions and the EEVD motif at the C-terminal end of cytosolic Hsp70 are shown in green highlights

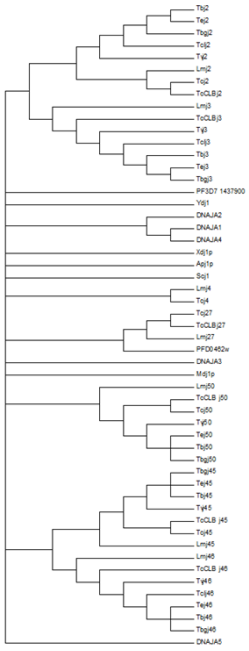


Figure F2: Predicted Type I J-proteins in TriTryps together with type I's from yeast, human and yeast evolutionary relationship was studied by phylogenetic analysis. Three main clusters were observed which can be further subdivided into cytosolic, mitochondrial and ER J-proteins. Homologues in higher eukaryotes were identified.

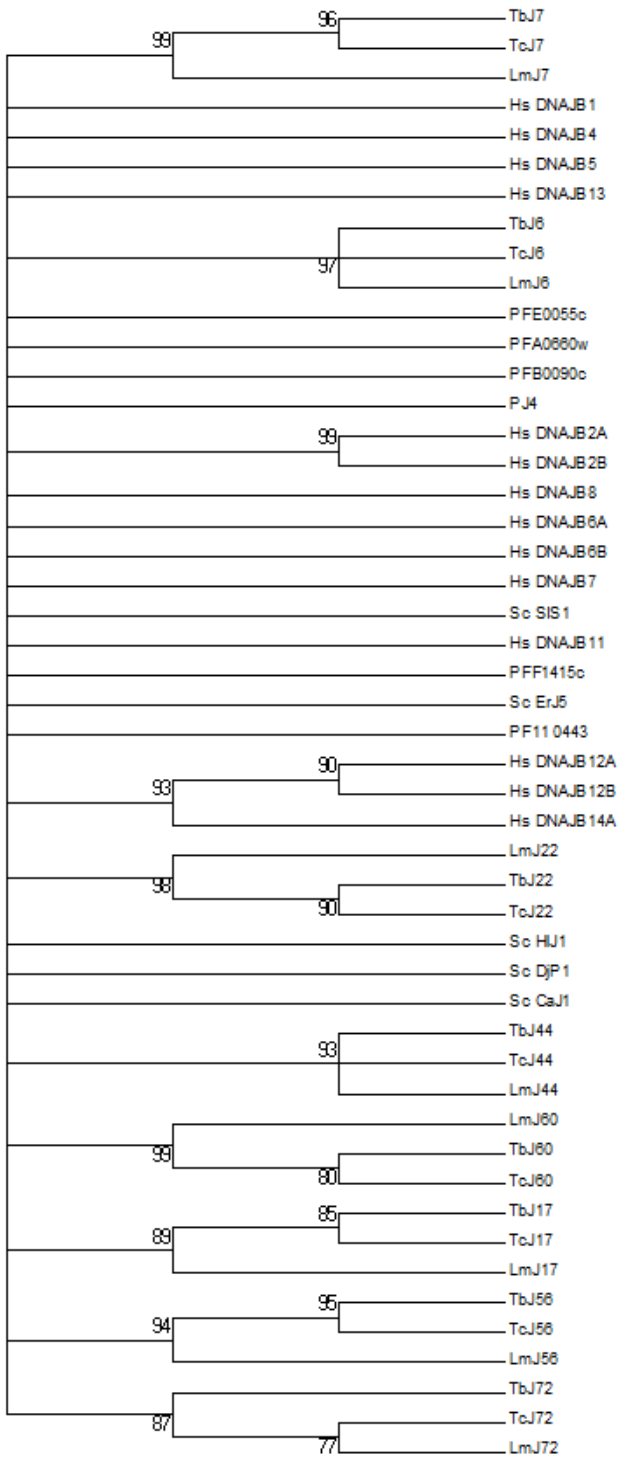
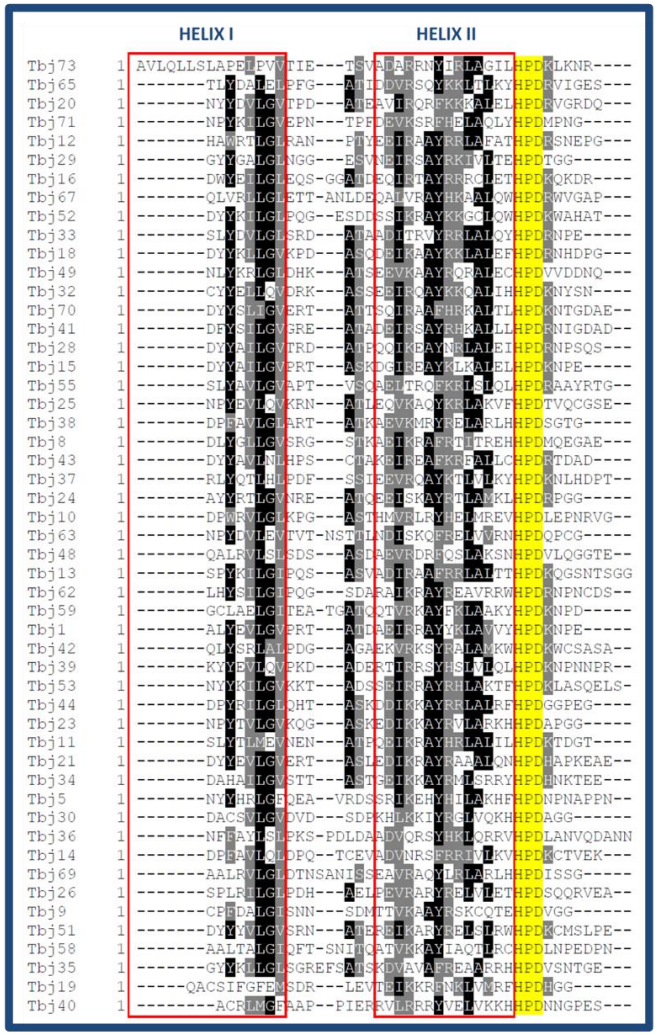


Figure F3: Phylogenetic analysis of putative type II J-proteins in TriTryps in relation to human, yeast and *P. falciparum* type II J-proteins. Eight main cluster of TriTryps J-proteins were generated, although lower sequence identity and human isoforms (tending to form paralogous relationships) resulted in fewer homologues being identified.


HELIX I
HELIX II

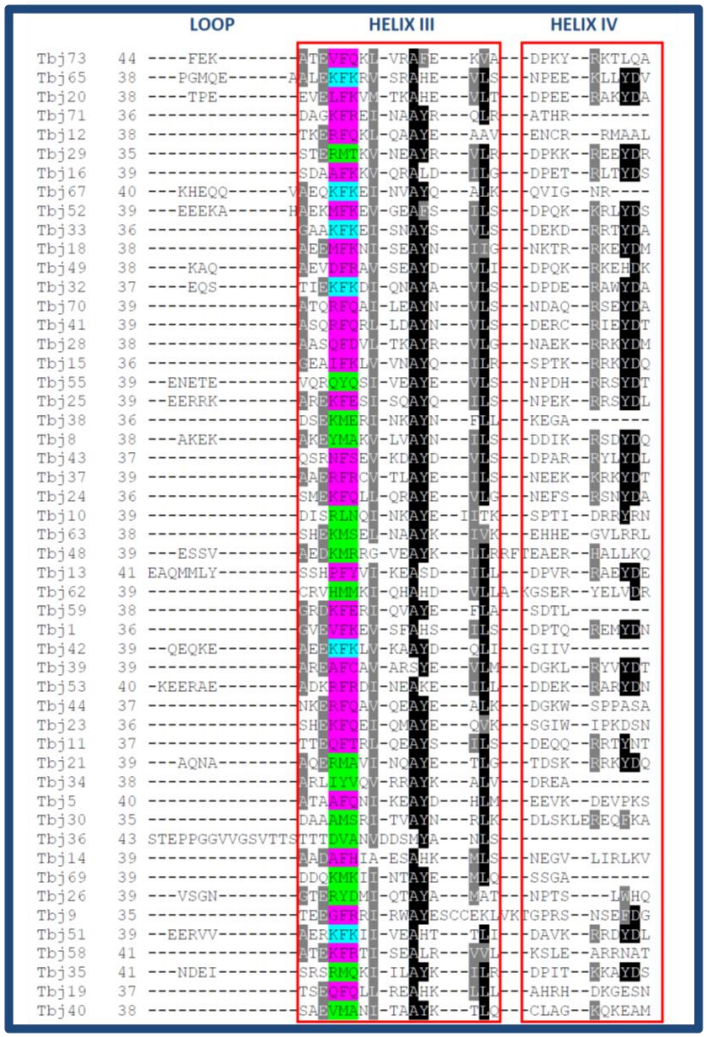

LOOP
HELIX III
HELIX IV

Figure F.4: Amino acid sequence alignment of predicted type III J-domains. helices with the J-domain annotated in red boxes, the conserved HPD motif motif shown in yellow highlights, the lack of conservation of the KFK motif is also shown by highlights, Light blue=KFK, Violet –F-and Green poor sequence conservation

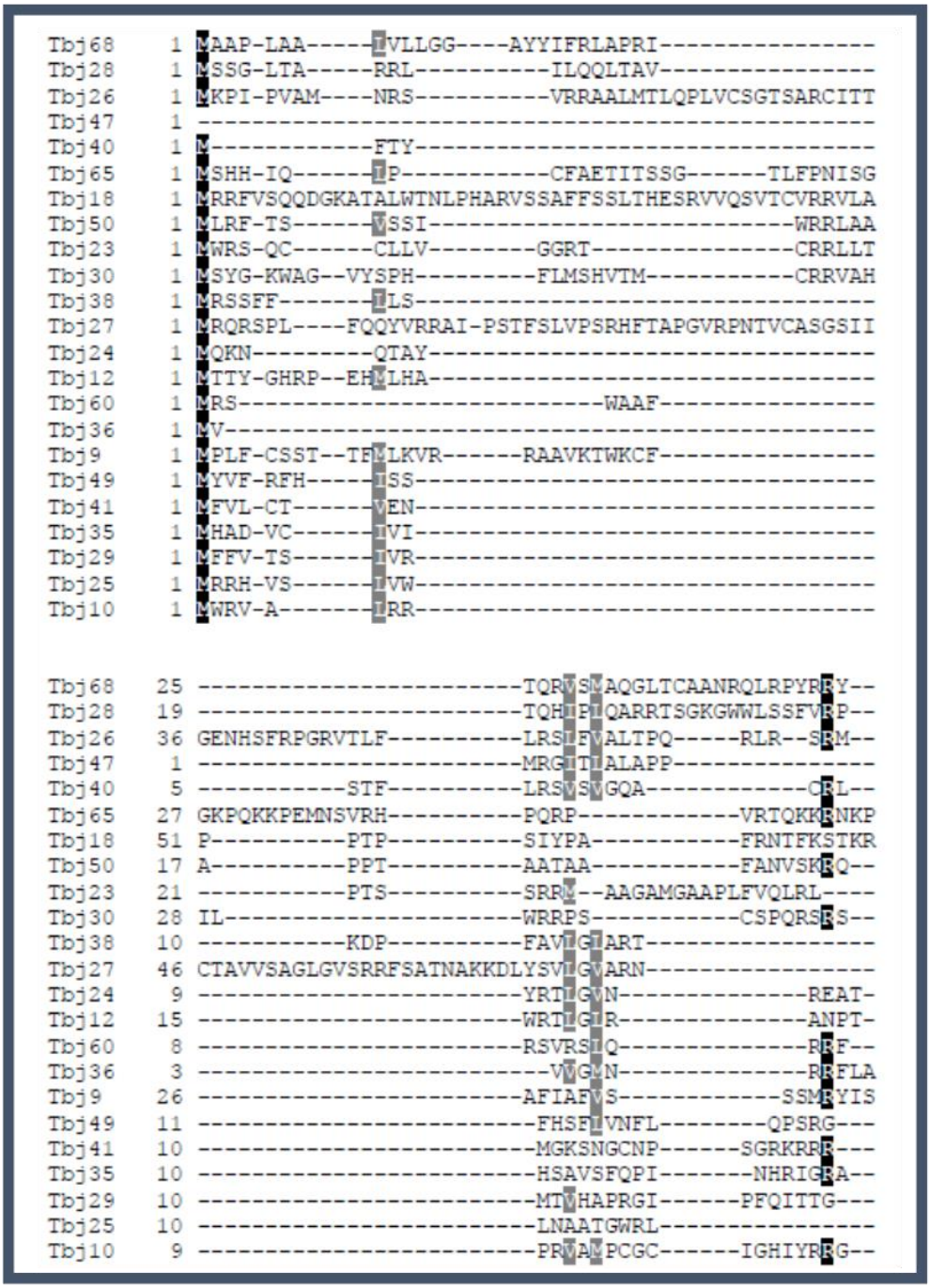


Figure F5: Sequence alignment of the targeting sequences for all J-proteins predicted to localize in the mitochondria. A consensus sequence could not be derived due to sequence variability, however the mitochondrial targeting sequence seems to contain a repetition of positively charged residues especially arginine (R), Valine (V), Leucine (L), Isoleucine (I) and phenylalanine (F) residues.

Figure F6: MEME analysis for type III J-proteins predicted to be secreted and plasma membrane associated. Three motifs were identified, Light blue box corresponding to the J-domain especially residues close to the HPD motif, blue box helix III within the J-domain. The red box corresponds to an interesting motif found in Tbj62, Tbj59, Tbj39 and Tbj21.

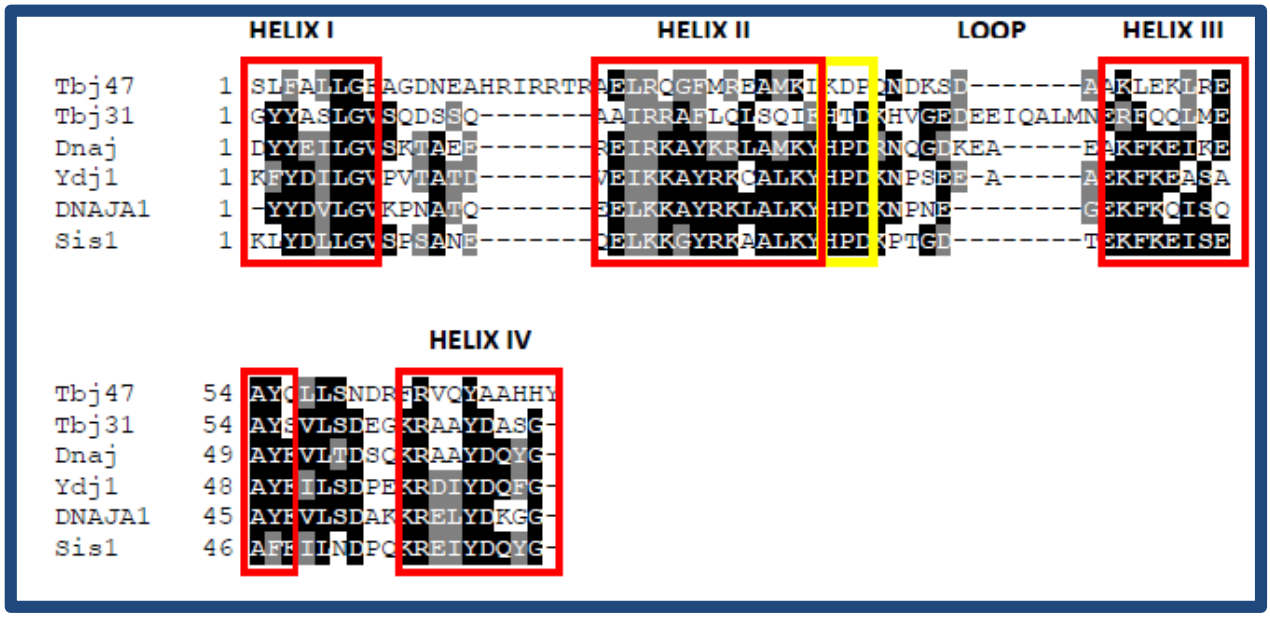


Figure F7: Sequence alignment of J-domains of type IV J-proteins with the canonical J-proteins, *E. coli* DNAJ, *S. cerevisiae* Ydj1 and Sis1 and *H. sapien* DNAJA1

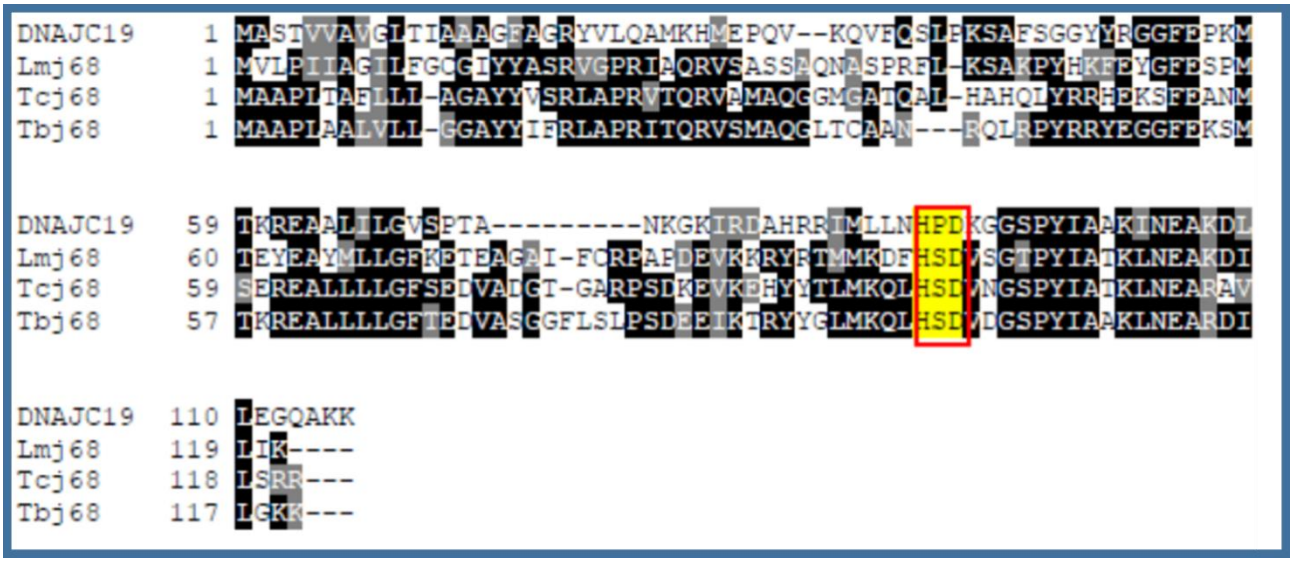


Figure F8: Sequence alignment of Tbj68 with its *L. major* and *T. cruzi* ortholog and putative *H. sapien* homolog DNAJC19.

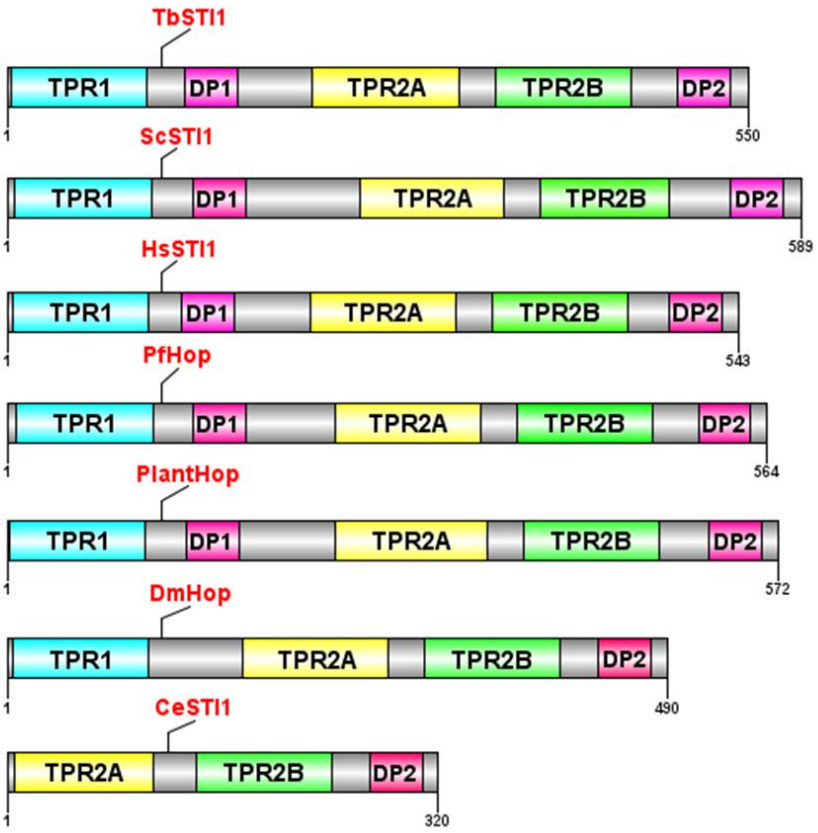


Figure F 10: The domain organization of STI1 proteins showing the TPR1 (cyan), TPR2A (Yellow) and TPR2B (Green) as well as the two DP regions between TPR1 and toward the C-terminus of the proteins

Appendix G: *In vitro* supplementary data

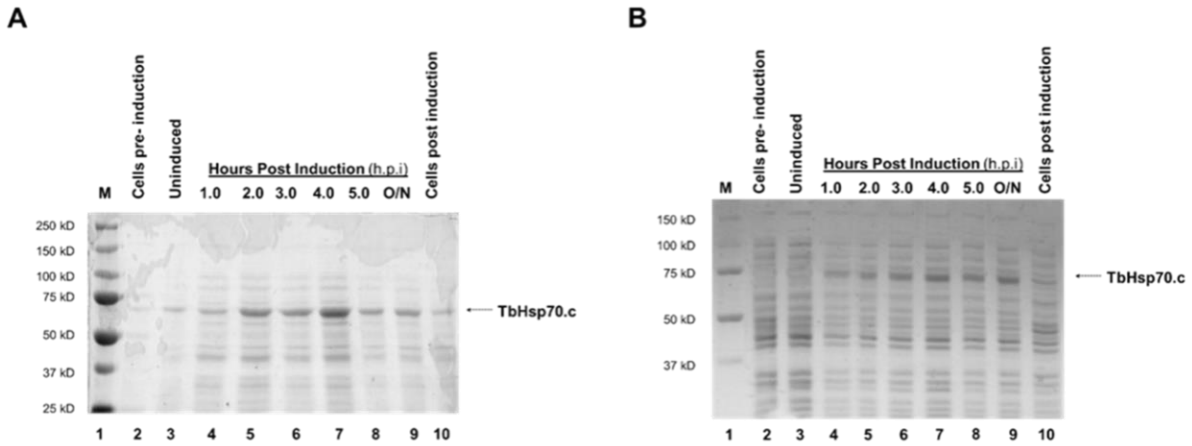


Figure G.1: Expression of TbHsp70.c in *E. coli* XL1 Blue and *E. coli* M15 cells. Samples of *E. coli* XL1 Blue and *E. coli* M15 were transformed separately with N-terminal His tag pQE80/TbHsp70.c . Samples of the transformed *E. coli* XL1 Blue and *E. coli* M15 cells were taken hourly after the addition of 1 mM IPTG. (A) 10 % SDS-PAGE analysis of the expression of TbHsp70.c in *E. coli* XL1 Blue [pQE80/TbHsp70.c]. Lane 1- Precision Plus Protein™ Standards (BIORAD) marker; Lane 2- uninduced, untransformed *E. coli* XL1 Blue whole cell extracts (negative control); Lane 3- uninduced *E. coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract; Lanes 4 to 8- *E.coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract samples taken 1 to 5 hours post induction with 1 mM IPTG; Lane 9- O/N represents the *E.coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract sample take after induction overnight; Lane 10- induced, untransformed *E.coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract sample (negative control). (B) 10 % SDS-PAGE analysis of the expression of TbHsp70.c in *E. coli* XL1 Blue [pQE80/TbHsp70.c]. Lane 1- Precision Plus Protein™ Standards (BIORAD) marker; Lane 2- uninduced, untransformed *E. coli* XL1 Blue whole cell extracts (negative control); Lane 3- uninduced *E. coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract; Lanes 4 to 8- *E.coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract samples taken 1 to 5 hours post induction after the addition of 1 mM IPTG; Lane 9- O/N represents the *E.coli* XL1 Blue [pQE80/TbHsp70.c] whole cell extract sample take after induction overnight; Lane 10- induced, untransformed XL1 Blue [pQE80/TbHsp70.c] whole cell extract sample (negative control).

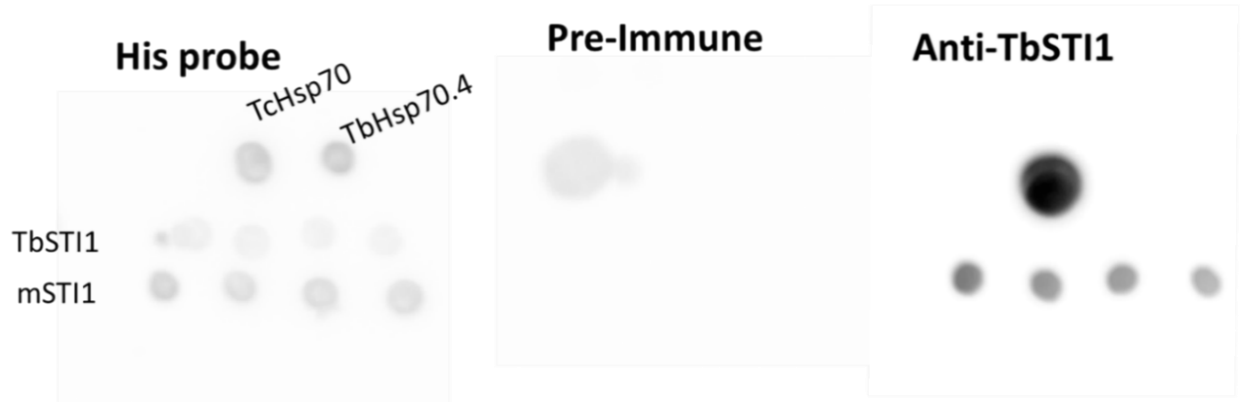


Figure G2: The anti-TbSTI1 was produced and tested using dot blot assays. A series of His-TbSTI1 and His-mSTI1 concentrations (1-20 µg) were spotted on a membrane as well as 20 µg His-TbHsp70.4 and His TcHsp70 and detected using anti-His antibodies. The pre-immune showed no binding and the anti-TbSTI1 serum was also shown to bind specifically to TbSTI1

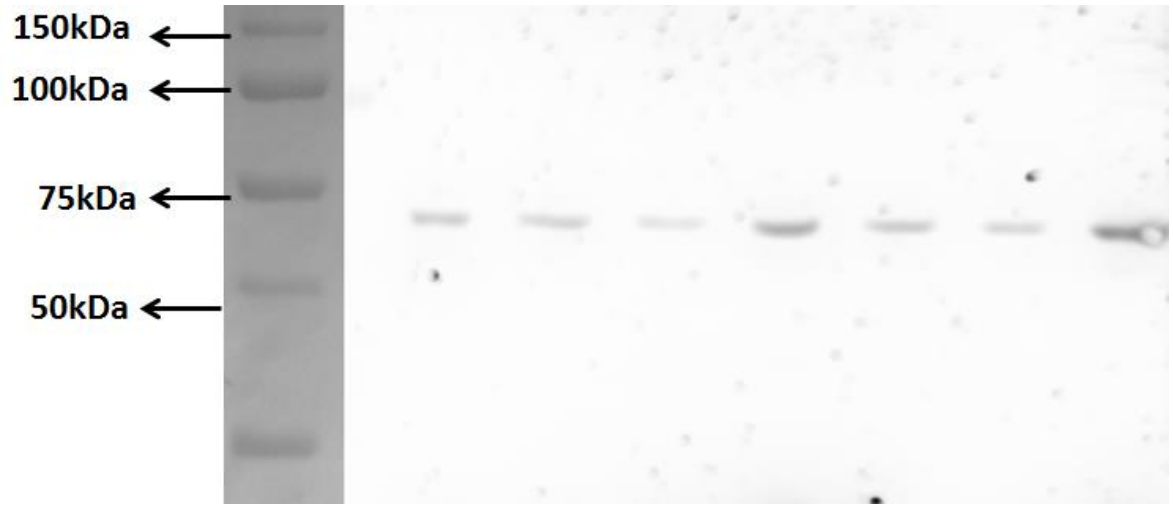


Figure G3: A western detection of TbHsp83 induction using anti-Hsp83 antibodies.

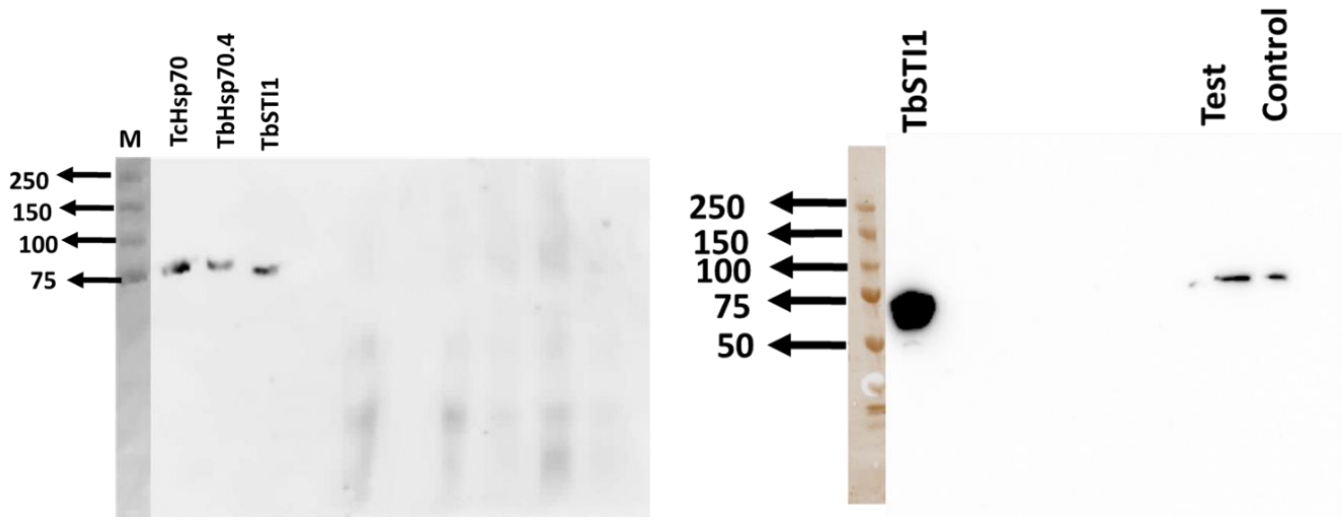


Figure G4: TbSTI1 did not interact directly with TcHsp70, TbHsp70.4 and TbHsp70.c.

REFERENCES

- Abbas-Terki, T., Briand, P.-A., Donzé, O., and Picard, D. (2002). The Hsp90 cochaperones Cdc37 and Sti1 interact physically and genetically. *Biol. Chem.* **383**, 1335-1342.
- Albanèse, V., Yam, A. Y., Baughman, J., Parnot, C., and Frydman, J. (2006). Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell*. **124**, 75-88.
- al-Herran, S., and Ashraf, W. (1998) Physiological consequences of the over-production of *E. coli* truncated molecular chaperone DnaJ. *FEMS Microbiol Lett*. **162**, 117–122.
- Ali, M.M., Roe, S.M., Vaughan, C.K., Meyer, P., Panaretou, B., Piper, P.W., Prodromou, C and Pearl, L.H. (2006). Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nat*. **440**, 1013-1017.
- Alsford, S., Turner, D.J., Obado, S.O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C. and Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res*. **21**, 915-924.
- Arndt H, Dietrich D, Auer B, Cleven E, Grafenham T, Weitere M, Mylnikov AP (2000) Functional Diversity of Heterotrophic Flagellates in Aquatic Ecosystems. In Leadbeater BSC, Green JC (eds) The flagellates. Taylor and Francis, *London*, pp 240–268.
- Aslett,M., Aurrecoechea,C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., Gardner, M.J., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Logan, F.J., Miller, J.A., Mitra, S., Myler, P.J., Nayak, V., Pennington, C., Phan, I., Pinney, D.F., Ramasamy, G., Rogers, M.B., Roos, D.S., Ross, C., Sivam, D., Smith, D.F., Srinivasamoorthy, G., Stoeckert, C.J., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G. and Wang, H. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* **38,** D457–D462.
- Atwood JA, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, Orlando R, Tarleton RL (2005) The *Trypanosoma cruzi* proteome. *Sci.* **309**,473–476.
- Bacchi CJ (1993) Resistance to Clinical Drugs in African Trypanosomes. *Parasitol Today*. **9**,190-193.
- Ban, C., Junop, M and Yang, W. (1999). Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell*. **97**, 85-97.
- 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*. **271**, 14840-14848.
- Bangs, J.D., Brouch, E.M., Ransom, D.M. and Roggy, J.L. (1996) A soluble secretory reporter system in
- Bangs, J.D., Uyetake, L. Brickman, M.J., Balber, A.E. and Boothroyd, J.C. (1993) Molecular cloning and cellular localization of a BiP homologue in *Trypanosoma brucei*. *J. Cell Science*. **105**, 1101-1113.
- Baral, T.N. (2010). Immunobiology of African trypanosomes: need of alternative interventions. *J. Biomed. Biotechnol*. **2010**, 1-24.
- Barrett, M.P., Boykin, D.W., Brun, R. and Tidwell, R.R. (2007). Human African Trypanosomiasis: pharmacological re-engagement with a neglected disease. *British J Pharmacol*. **152**, 1155-1171.
- Barrett, M.P., Burchmore, R.J.S., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J. and Krishna, S (2003). The Trypanosomiases. *The Lanc*. **362**, 1469-1480.
- Barry JD, Hajduk S L, Vickerman K and Le Ray D (1979) Detection of Multiple Variable Antigen Types in Metacyclic Populations of *Trypanosoma brucei*. *Trans R Soc Trop Med Hyg*. **73**,205-208.
- Barry, J.D. and Emery, D.L. (1984). Parasite development and host responses during the establishment of *Trypanosoma brucei* infection transmitted by tsetse fly. *Parasitol*. **88**, 67-84.
- Barry, J.D. and McCulloch, R. (2001). Antigenic variation in trypanosomes: enhanced phenotypic variation in the eukaryotic parasite. *Advanc Parasitol***. 49**, 1-70.
- Benne, R., van den Burg, J., Brakenhoff, J. P. J., Sloof, P., van Boom, J. H., and Tromp, M. C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*. **46**, 819-826.

- Benne, R., van den Burg, J., Brakenhoff, J. P. J., Sloof, P., vanBoom, J. H., and Tromp, M. C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*. 46, 819-826.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Böhme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E. and El-Sayed, N.M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Sci*. **309**, 416–422.
- Beseničar M, Maček P, Lakey JH, Anderluh G, 2006. Surface plasmon resonance in protein– membrane interactions. *Chem. Phys. Lips*.**141**, 169–178.
- Bilwes, A. M., Quezada, C. M., Croal, L. R., Crane, B. R and Simon, M. I. (2001). Nucleotide binding by the histidine kinase CheA. *Nat Struct and Mol Biol*. **8**, 353-360.
- Blatch, G.L. and Lässle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioess*. **21**, 932-939.
- Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J. and Wiley, D. C. (1993). A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nat*. **362**, 603–609.
- Boorstein, W.R., Ziegelhoffer, T. and Craig, E.A. (1994) Molecular evolution of the HSP70 multigene family, *J Mol Evol*, **38**, 1-17.
- Borges, J.C., Fischer, H., Craievich, A.F. and Ramos, C.H.I. (2005) Low resolution structural study of two human Hsp40 chaperones in solution. *J. Biol. Chem*. **280**, 13671-13681.
- Botha, M., Pesce, E-R. And Blatch, GL. (2007). The Hsp40 proteins of *Plasmodium falciparum* and other apicomplexa: regulating chaperone power in the parasite and the host. *Int. J. Biochem. Cell Biol*. **39**, 1781-1803.
- Brandau, S.; Dresel, A.; Clos, J. (1995) High constitutive levels of heat shock proteins in human-pathogenic parasites of the genus *Leishmania*. *Biochem. J*. **310**, 225-232.
- Brehmer, D., Rüdiger, S., Gässler, C. S., Klostermeier, D., Packschies, 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.
- Bridges, D.J., Pitt, A.R., Hanrahan, O., Brennan, K., Voorheis, H.P., Herzyk, P., de Koning, H.P. and Burchmore, R.J. (2008). Characterisation of the plasma membrane subproteome of bloodstream form *Trypanosoma brucei*. Proteomics, 8 (1), 83-99.
- Brinker, A., Scheufler, C., von der Mülbe, F., Fleckenstein, B., Herrmann, C., Jung, G., Moarefi, I., and Hartl, F. U. (2002). Ligand discrimination by TPR domains. Relevance and selectivity of EEVD- recognition in Hsp70÷Hop÷Hsp90 complexes. *J. Biol. Chem*. **277**, 19265-19275.
- Brochu, C.; Haimeur, A.; Ouellette, M. (2004) The heat shock protein HSP70 and heat shock cognate protein HSC70 contribute to antimony tolerance in the protozoan parasite *Leishmania*. *Cell Stress Chap*. **9**, 294-303.
- Brun R and Schonenberger M (1981) Stimulating Effect of Citrate and Cis-Aconitate on the Transformation of *Trypanosoma brucei* Bloodstream Forms to Procyclic Forms *in Vitro*. *Z Parasitenkd* **66**, 17-24.
- Brun, R., Blum, J., Chappuis, F. and Burri, C. (2010). Human African trypanosomiasis. *Lanc*. **375**, 148-159.

- Brychzy, A., Rein, T., Winklhofer, K. F., Hartl, F. U., Young, J. C., and Obermann, W.M. (2003). Cofactor Tpr2 combines two TPR domains and a J domain to regulate the Hsp70/Hsp90 chaperone system. *EMBO J*. **22**, 3613-3623.
- Buchner, J. (1999). Hsp90 & Co. – a holding for folding. *Trends Biochem Sci*. **24**, 136–141.
- Bukau, B. and Horwich, A.L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell*. **92**, 351-366.
- Burger (2013). Purification and characterization of TbHsp70. c , a novel Hsp70 from *Trypanosoma brucei*, PhD Thesis, Rhodes University, Grahamstown, South Africa.
- Burger, A., Ludewig, M. H., and Boshoff, A. (2014). Investigating the Chaperone Properties of a Novel Heat Shock Protein , Hsp70 . c , from *Trypanosoma brucei*. *J.Parasitol.Res*. **172582**, 1-12.
- Cajo, G.C., Horne, B.E., Kelley, W.L., Schwager, F., Georgopoulos, C. and Genevaux, P. (2006). The role of the DIF motif of the DnaJ (Hsp40) co-chaperone in the regulation of the DnaK (Hsp70) chaperone cycle. *J. Biol. Chem*. **281**, 12436-12444.
- Campell, K., Mullane, K., Aksoy, I., Stubdal, H., Zalvide, J., Pipas, J., Silver, P., Roberts, T., Schaffhausen, B., and DeCaprio, J. (1997): DnaJ/Hsp40 chaperone domain of SV40 large T antigen promotes efficient viral DNA replication. *Genes Dev*. **11**, 1098–1110.
- Caplan, A.J., 2003. What is a co-chaperone? *Cell Stress Chap*. **8**, p.105.
- Carrello, A., Allan, R.K., Morgan, S.L., Owen, B.A.L., Mok, D., Ward, B.K., Minchin, R.F., Toft, D.O. and Ratajczak, T. (2004). Interaction of the Hsp90 cochaperone cyclophilin 40 with Hsc70. *Cell Stress Chap*. **9**, 167-181
- Carvalho JF, de Carvalho EF, Rondinelli E, Silva R, de Castro FT (1987) Protein biosynthesis changes in *Trypanosoma cruzi* induced by supra-optimal temperature*. Exp Cell Res*. 168, 338–346
- Chaudhuri M, Ott RD, Hill GC (2006) Trypanosome alternative oxidase: from molecule to function. *Trends Parasitol*. **22**, 485–491
- Checchi, F., Filipe, J. and Barrett, M. (2008). The natural progression of Gambiense sleeping sickness: What is the evidence*? PLOS Neglected Trop Dis*. **2**, e303- 4
- Cheetham, M. E., Jackson, A. P. and Anderton, B. H. (1994). Regulation of 70-kDa heat-shock-protein ATPase activity and substrate binding by human DnaJ-like proteins, HSJ1a and HSJ1b. *Euro J. Biochem*. **226**, 99–107.
- Cheetham, M.E. and Caplan, A.J. (1998). Structure, function and evolution of DnaJ: Conservation and adaptation 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.
- Chen, S., Prapapanich, V., Rimerman, R. A., Honoré, B., and Smith, D. F. (1996). Interactions of p60, a mediator of progesterone receptor assembly, with heat shock proteins Hsp90 and Hsp70. *J. Mol. Endo*. **10**, 682-693.
- Clayton C.E., and Michels P. (1996). Metabolic compartmentation in African trypanosomes. *Parasitol Today* **12**, 465–471.
- Clayton, C., Hausler, T. and Blattner, J. (1995). Protein trafficking in kinetoplastid protozoa. *Microbiol. Rev*. **59**, 325–344.
- Clayton, C.E. (2002). Life without transcriptional control? From fly to man and back again. *EMBO J*. **21**, 1881-1888
- Cliff, M. J., Harris, R., Barford, D., Ladbury, J. E., and Williams, M. A. (2006). Conformational diversity in the TPR domain-mediated interaction of protein phosphatase 5 with Hsp90. *Struct*. **14**, 415-426.
- Cliff, M.J., Williams, M.A., Brooke-Smith, J., Barford, D. and Ladbury, J.E. (2005). Molecular recognition via coupled folding and binding in a TPR domain. *J. Mol. Biol.* **346**, 717-732.
- Connell, P., Ballinger, C. A., Jiang, J., Wu, Y., Thompson, L. J., Hohfeld, J., and Patterson, C. (2001). The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat. Cell Biol.* **3**, 93-96.
- Contu, L. (2014). The effects of extracellular and intracellular Hop on cell migration processes. MSc Thesis, Rhodes University, Grahamstown, South Africa.

- Cortajarena, A. L., Yi, F., and Regan, L. (2008). Designed TPR modules as novel anticancer agents. *ACS Chem. Biol.* **3**, 161-166.
- Cox, M. and Johnson, J., 2011. The Role of p23, Hop, Immunophilins, and Other Co-chaperones in Regulating Hsp90 Function. *Mol. Chap Hum Press*, pp. 45–66.
- Cox, M.B., Riggs, D.L., Hessling, M., Schumacher, F., Buchner, J and Smith, D.F. (2007). FK506-binding protein 52 phosphorylation: a potential mechanism for regulating steroid hormone receptor activity. *Mol. Endoc*. **21**, 2956–2967.
- Craig, E.A., Huang, P., Aron, R. and Andrew, A. (2006). The diverse roles of J-proteins, the obligate Hsp70 cochaperone. *Rev. Physiol. Biochem. Pharmacol*. **156**, 1–21.
- Cyr, D.M., Höhfeld, J. and Patterson, C. (2002). Protein quality control: U-box-containing E3 ubiquitin ligases join the fold. *Trends Biochem Sci*, **27**, pp.368–375.
- Cyr, D.M., Lu, X. and Douglas, M.G. (1992). Regulation of Hsp70 function by a eukaryotic DnaJ homolog. J. Biol. Chem. 267: 20927-20931.
- Czichos J, Nonnengaesser C and Overath P (1986) Trypanosoma brucei: Cis-Aconitate and Temperature Reduction As Triggers of Synchronous Transformation of Bloodstream to Procyclic Trypomastigotes *in Vitro*. *Exp Parasitol*. 62, 283-291.
- D'Andrea, L.D. and Regan, L. (2003). TPR proteins: the versatile helix. *Trends in biochem sci*. **28**, pp.655–62.
- Daniel, S., Bradley, G., Longshaw, V. M., Soti, C., Csermely, P., and Blatch, G. L. (2008). Nucleus translocation of the phosphoprotein Hop (Hsp70/Hsp90 organising protein) occurs under heat shock, and its proposed nucleus localisation signal is involved in Hsp90 binding. *Biochim Biophys Acta*. **1783**, 1003-1014.
- Davies, T.H. and Sánchez, E.R., 2005. FKBP52. *Int.J. biochem cell biol*, **37**, pp.42–47.
- De Lange, T., Michels, P. A., Veerman, H. J., Cornelissen, A. W., and Borst, P. (1984). Many trypanosome messenger RNAs share a common 5' terminal sequence. *Nuc Acids Res*. **12**, 3777-3790.
- De Lange, T., Michels, P. A., Veerman, H. J., Cornelissen, A. W., and Borst, P. (1984). Many trypanosome messenger RNAs share a common 5' terminal sequence. *Nucleic Acids Res*. **12**, 3777-3790.
- de Souza W, Attias M, Rodrigues JCF (2009) Particularities of mitochondrial structure in parasitic protists (Apicomplexa and Kinetoplastida). *Int J Biochem Cell Biol*. **41**, 2069–2080
- Desjeux P (2004) Leishmaniasis: current situation and new perspectives. *Comp Immunol Microbiol Infect Dis* .**27**, 305–381.
- Dollins, D. E., Warren, J. J., Immormino, R. M and Gewirth, D. T. (2007). Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones. *Mol Cell*. **28**, 1–56.
- Donelson, J.E., Gardner, M.J. and El-Sayed, N.M. (1999). More surprises from Kinetoplasts. *Proceed Nat. Acad. Sci U.S.A*. **96**, 2579-2581.
- Dragon, E.A., Sias, S.R., Kato, E.A, Gabe J.D. (1987) The genome of *Trypanosoma cruzi* contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. *Mol Cell Biol*. **7**, 1271–1275
- Dragovic, Z., Broadle, S. A., Shomura, Y., Bracher, A., and Hartl, F.U. (2006) Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. *EMBO J*. **25**, 2519-2528.
- Durieux, P.O., Schütz, P ., Brun, R., and Kohler, P. (1991) Alterations in Krebs cycle enzyme activities and carbohydrate catabolism in two strains of *Trypanosoma brucei* during *in vitro* differentiation of their bloodstream to procyclic stages. *Mol Biochem Parasitol*. **45**,19–27
- Dutta, R and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Tr Biochem Sci*. **25**, 24–28.
- Echeverría, P. C., and Picard, D. (2010). Molecular chaperones, essential partners of steroid hormone receptors for activity and mobility. *Biochim. Biophys. Acta*. **1803**, 641-649.
- Edkins AL and Boshoff A. (2014) General structural and functional features of heat shock proteins. Book Chapter in: Heat shock proteins of malaria. Editors: Addmore Shonhai and Greg Blatch. Springer, ISBN 978-94-007-7437-7; pp5 -45"

- Edkins, A.L., Ludewig, M.H. and Blatch, G.L. (2004). A *Trypanosoma cruzi* heat shock protein 40 is able to stimulate the adenosine triphosphate hydrolysis activity of heat shock protein 70 and can substitute for a yeast heat shock protein 40. *Int. 'J. Biochem and Cell Biol*. **36**, 1585-1598.
- Ellis, J. (1987) Proteins as molecular chaperones. *Nat*. **328**, 378-379.
- Ellis, R.J. and van der Vies, S.M. (1991) Molecular Chaperones. *Annu Rev Biochem.* **60**, 321-347.
- Emini, E.A., Hughes, J.V., Perlow, D.S. and Boger, J. (1985) Induction of hepatitis A virus-neutralising antibody by a virus-specific synthetic peptide. *J. Virol*. **55**, 836-839.
- Engman, D.M., Dragon, E.A. and Donelson, J.E. (1990). Human humoral immunity to hsp70
- Engman, D.M., Fehr, S.C. and Donelson, J.E. (1992) Specific functional domains of mitochondrial hsp70s suggested by sequence comparison of the trypanosome and yeast proteins. *Mol Biochem Parasitol*/ **51**, 153–155.
- Engman, D.M., Kirchhoff, L.V. and Donelson, J.E. (1989a). Molecular cloning of mtp70, a mitochondrial member of the hsp70 family. *Mol Cell Biol*. **9**, 5163–5168
- Engman, D.M., Sias, S.R., Gabe, J.D., Donelson, J.E. and Dragon, E.A. (1989b). Comparison of HSP70 genes from two strains of Trypanosoma cruzi. *Mol Biochem Parasitol.* **37**, 285–287
- Engstler, M., Thilo, L., Weise, F., Grünfelder, C. G., Schwarz, H., Boshart, M. and Overath, P. (2004). Kinetics of endocytosis and recycling of the GPI anchored variant surface glycoprotein in *Trypanosoma brucei*. *J Cell Sci*. **117**, 1105–1115.
- Eustace, B. K. and Jay, D. G. (2004). Extracellular roles for the molecular chaperone, hsp90. *Cell Cycle*. **3**, 1098-1100.
- Fan, C. Y., Lee, S., Ren, H. Y., and Cyr, D. M. (2004) The Type I Hsp40 Zinc Finger-like region is required for Hsp70 to capture non-native polypeptides from Ydj1. *Mol. Biol. Cell*. **15**, 761–773
- Faria-e-Silva PM, Attias M, de Souza W (2000) Biochemical and ultrastructural changes in *Herpetomonas roitmani* related to the energy metabolism. *Biol Cell*. **92**, 39–47.
- Felts, S. J., Owen, B. A., Nguyen, P., Trepel, J., Donner, D. B., and Toft, D. O. (2000). The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties. *J. Biol. Chem*. **275**, 3305-3312.
- Fenn K and Matthews K R (2007) The Cell Biology of *Trypanosoma brucei* Differentiation. *Curr Opin Microbiol*. **10**, 539-546.
- Fernandes, A. P., Nelson, K., and Beverley, S. M. (1993). Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. *Proc. Natl. Acad. Sci. USA*. **90**, 11608-11612.
- Fink, A.L. (1999). Chaperone-mediated protein folding. *Physiol Rev*. **79**, 425-449.
- Flom, G., Behal, R. H., Rosen, L., Cole, D. G and Johnson, J. L. (2007). Definition of the minimal fragments of Sti1 required for dimerization, interaction with Hsp70 and Hsp90 and in vivo functions. *Biochem J*. 404, 159-167.
- Folgueira, C. and Requena, J.M. (2007) A postgenomic view of the heat shock proteins in kinetoplastids. *FEMS Microbiol Rev*. **4**, 359-377.
- Fraumann, R. (2003). "*Glossina morsitans*" (On-line), Animal Diversity Web. Accessed September 17, 2013 at http://animaldiversity.ummz.umich.edu/site/accounts/information/Glossina_morsitans.html.
- Frey, S., Leskovar, A., Reinstein, J and Buchner, J. (2007). The ATPase cycle of the endoplasmic chaperone Grp94. *J. Biol Chem*. **282**, 35612-35620.
- Fridberg, A., Buchanan, K.T. and Engman, D.M. (2007) Flagellar membrane trafficking in kinetoplastids. *Parasitol Res*. **100**, 205-212.
- Garcia, A., Courtin, D., Solano, P., Koffi, M. and Jamonneau, V. (2006). Human African Trypanosomiasis: connecting parasite and host genetics. *Tr Parasitol*. **22**, 405-409
- Genevaux, P., Schwager, F., Georgopoulos, C., and Kelley, W. L. (2002). Scanning mutagenesis identifies amino acid residues essential for the in vivo activity of the Escherichia coli DnaJ (Hsp40) J-domain. *Gen*. **162**, 1045–1053.

- Giordano, A., Whyte, P., Harlow, E., Jr., B. R. F., Beach, D., and Draetta, G. (1989). A 60 kd cdc2-associated polypeptide complexes with the E1A proteins in adenovirus-infected cells. *Cell*. **58**, 981 – 990.
- Gitau, G. W., Mandal, P., Blatch, G. L., Przyborski, J., and Shonhai, A. (2011): Characterisation of the *Plasmodium falciparum* Hsp70-Hsp90 organising protein (PfHop). *Cell Stress Chaperon* **17**, 191–202.
- Goebl, M. and Yanagida, M. (1991). The TPR snap helix: a novel protein repeat motif from mitosis to transcription. *Tr. Biochem Sci.* 16, 173-177.
- Goffin, L. and Georgopoulos, C. (1998). Genetic and biochemical characterisation of mutations affecting the carboxy-terminal domain of the *Escherichia coli* molecular chaperone DnaJ. *Mol Microbiol*. 30, 329-340.
- Graefe, S.E.B., Wiesgigle, M., Gaworski, I., MacDonald, A. and Clos, J. (2002). Inhibition of Hsp90 in Trypanosoma cruzi induces a stress response but no stage differentiation. *Eukaryot. Cell*. **1**, 936-943.
- Grammatikakis, N., Lin, J. H., Grammatikakis, A., Tsichlis, P. N., and Cochran, B. H. (1999). p50 (cdc37) acting in concert with Hsp90 is required for Raf-1 function. *Mol Cell Biol*. **19**, 1661-1672.
- Grammatikakis, N., Vultur, A., Ramana, C.V., Siganou, A., Schweinfest, C.W., Watson, D.K and Raptis, L. (2002). The role of HSP90N, a new member of the Hsp90 family, in signal transduction and neoplastic transformation. *J. Biol Chem*. **277**, 8312 – 8320.
- Greene, M.K., Maskos, K. and Landry, S.J. (1998). Role of the J-domain in the cooperation of Hsp40 with Hsp70. *PNAS*. **95**, 6108–6113.
- Grenert, J.P., Sullivan, W.P., Fadden, P., Haystead, T.A.J., Clark, J.,Mimnaugh, E., Krutzsch, H., Ochel, H.J., Schulte, T.W., Sausville, E., Neckers, L.M and Toft, D.O. (1997). The amino-terminal domain of heat shock protein 90 (hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates Hsp90 conformation. *J. Biol. Chem*. **272**, 2384 -23850.
- Gull, K. (1999). The cytoskeleton of trypanosomatid parasites. *Annu. Rev. Microbiol*. **53**, 629–655.
- Gull, K. (2002) The Cell Biology of Parasitism in *Trypanosoma brucei*: insights and drug targets from genomic approaches. *Curr Pharm Des*. **8**, 241-256.
- Gull, K. (2003). Host-parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies. *Curr. Opin. Microbiol*. **6**, 365–370.
- Hageman, J. and Kampinga, H.H. (2009). Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. *Cell Stress Chaperon*. **14**, 1–21.
- Hainzl O., Lapina, M.C., Buchner, J and Richter, K. (2009). The charged linker region is an important regulator of Hsp90 function. *J. Biol. Chem*. **284**, 22559–22567.
- Han, W. and Christen, P. (2003). Mechanism of the targeting action of DnaJ in the DnaK molecular chaperone system. *J. Biol Chem*. **278**, 19038-19043.
- Harrison, C. J., Hayer-Hartle, M., Di Liberto, M., Hartl, F. U., and Kuriyan, J. (1997) Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Sci*. **276**, 431-435.
- Hartl, F.U. (1996) Molecular chaperones in cellular protein folding. *Nat*. **381**, 571-580.
- Hartl, F.U. and Hayer-Hartl, M. (2002) Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein. *Sci*. **295**, 1852-1858.
- Hatherley, R., Blatch, G. L., and TastanBishop, Ö. T. (2013). Plasmodium falciparum Hsp70-x : A Heat Shock Protein at the Host – Parasite.
- Hendrick, J. P., and Hartl, F. U. (1993) Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem*. **62**, 349-384
- Hennessy, F. (2004). Characterisation of the J-domain amino acid residues important for the interaction of DnaJ-like proteins with Hsp70 chaperones. Phd Thesis, Rhodes University, Grahamstown, Republic of South Africa.
- Hennessy, F., Boshoff, A. and Blatch, G.L. (2005). Rational mutagenesis of a 40 kDa heat shock protein from *Agrobacterium tumefaciens* identifies amino acid residues critical to its *in vivo* function. *Int J Biochem. Cell Biol*. **37**, 177-191.

- Hennessy, F., Cheetham, M.E., Dirr, H.W. and Blatch, G.L. (2000). Analysis of the levels of conservation of the J domain among the various types of DnaJ-like proteins. *Cell Stress and Chaperon*. **5**, 347-358.
- Herman M, Pérez-Morga D, Schtickzelle N and Michels, P. A. (2008) Turnover of glycosomes during differentiation of *Trypanosoma brucei* . *Aut*. **4**, 294–308.
- Hernández, M.P., Sullivan, W.P., and Toft, D.O. (2002). The assembly and intermolecular properties of the hsp70–Hop–hsp90 molecular chaperone complex. *J Biol Chem*. **277**, 38294–38304.
- Hickey, E., Brandon, S. E., Smale, G., Lloyd, D and Weber, L. A. (1989). Sequence and regulation of a gene encoding a human 89-kilodalton heat shock protein. *Mol Cell Biol*. **9**, 2615-2626.
- Hinds TD, Sanchez ER (2008) Protein phosphatase 5. *Int J Biochem Cell Biol*. **40**, 2358–2362
- Hirano, T., Kinoshita, N., Morikawa, K. and Yanagida, M. (1990). Snap helix with knob and hole: essential repeats in S. pombe nuclear protein nuc2+. *Cell*. **60**, 319-328.
- Hirumi, H. and Hirumi, K. (1994) Axenic Culture of African Trypaonsome Bloodstream Forms. *Parasitol Today*. **10**, 80-84.
- Hofmann, H.J. and Hadge, D. (1987). On the theoretical prediction of protein antigenic determinants from amino acid sequences, *Biomed Biochim Acta*, **46**, 855-866.
- Höhfeld, J. and Jentsch, S. (1997). GrpE-like regulation of the Hsc70 chaperone by the anti-apoptotic protein BAG-1. *EMBO J*. **16**, 6209-6216.
- Höhfeld, J., Minami, Y. and Hartl, F-U. (1995). Hip, a novel cochaperone involved in the eukaryotic Hsc/Hsp40 reaction cycle. *Cell*. **83**, 589-598
- Honoré, B., Leffers, H., Madsen, P., Rasmussen, H. H., Vandekerckhove, J., and Celis, J. E. (1992). Molecular cloning and expression of a transformingsensitive human protein containing the TPR motif and sharing identity to the stress-inducible yeast protein STI1. *J Biol Chem*. **267**, 8485-8491.
- Hopp, T. P., Pricket, K. S., Price, V. L., Libby, R. T., March, C. J., Ceretti, D. P., Urdal, D. L., and Conlon, P. J. (1988): A short polypeptide marker sequence useful for recombinant protein identification and purification. *Biotech*. **6**, 1204 – 1210.
- Hopp, T.P. and Woods, K.R. (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A*. **78**, 3824-3828.
- Horibe, T., Kawamoto, M., Kohno, M., and Kawakami, K. (2012). Cytotoxic activity to acute myeloid leukemia cells by Antp-TPR hybrid peptide targeting Hsp90. *J. Biosci. Bioeng*. **114**, 96-103.
- Horibe, T., Kohno, M., Haramoto, M., Ohara, K., and Kawakami, K. (2011). Designed hybrid TPR peptide targeting Hsp90 as a novel anticancer agent. *J Transl Med*. **9**, 8-12.
- Horibe, T., Torisawa, A., Kohno, M., and Kawakami, K. (2014). Synergetic cytotoxic activity toward breast cancer cells enhanced by the combination of Antp-TPR hybrid peptide targeting Hsp90 and Hsp70- targeted peptide. *BMC Cancer.* **14**, 615.
- Hua, G., Zhang, Q and Fan, Z. (2007). Heat shock protein 75 (TRAP1) antagonizes reactive oxygen species generation and protects cells from granzyme M-mediated apoptosis. *J. Biol Chem*. **282**, 20553–20560.
- Huang, K., Flanagan, J.M. and Prestegard, J.H. (1998). The influence of C-terminal extension on the structure of the J-domain in *E. coli* DnaJ. *Prot Sci*. **8**, 203-214.
- Jackson, S. E. (2012). Hsp90: Structure and function. *Top. Curr. Chem*. **328**, 155-240.
- Jameson, B.A. and Wolf, H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants, *Comput Appl Biosci*. **4**, 181-186.
- Jiang, J., Maes, E.G., Taylor, A.B., Wang, L., Hinck, A.P., Lafer, E.M. and Sousa, R. (2007). Structural basis of the J cochaperone binding and regulation of Hsp70. *Mol. Cell*. **28**, 422-433.
- Jiang, R. F., Greener, T., Barouch, W., Greene, L., and Eisenberg, E. (1997): Interaction of auxilin with the molecular chaperone, Hsc70. *J. Biol. Chem*. **272**, 6141–6145.
- Johnson, J. L. (2012). Evolution and function of diverse Hsp90 homologs and co chaperone proteins. *Biochim. Biophys. Acta*. **1823**, 607-613.

- Johnson, J. L., and Toft, D. O. (1994). A novel chaperone complex for steroid receptors involving heat shock proteins, immunophilins, and p23. *J. Biol. Chem*. **269**, 24989-24993.
- Jones, A., Faldas, A., Foucher, A., Hunt, E., Tait, A., Wastling, J.M. and Turner, C.M. (2006). Visualisation and analysis of proteomic data from the procyclic form of *Trypanosoma brucei*. *Proteom*. **6**, 259-267.
- Jones, C., Anderson, S., Singha, U.K. and Chaudhuri, M. (2008). Protein phosphatise 5 is required for Hsp90 function during proteotoxic stresses in *Trypanosoma brucei*. *Parasitol. Res*. **102**, 835-844.
- Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *CABIOS*. **8**, pp.275–282.
- Jones, G., Song, Y., Chung, S., and Masison, D. C. (2004). Propagation of *Saccharomyces cerevisiae* [PSI+] prion is impaired by factors that regulate Hsp70 substrate binding. *Mol. Cell. Biol*. **24**, 3928-3937.
- Jubete, Y., Maurizi, M. R., and Gottesman, S. (1996): Role of the heat shock protein DnaJ in the lon-dependent degradation of naturally unstable proteins. *J. Biol. Chem*. **271**, 30798–30803.
- Kajander, T., Sachs, J.N., Goldman, A., and Regan, L. (2009). Electrostatic interactions of Hsp-organising protein tetratricopeptide domains with Hsp70 and Hsp90: computational analysis and protein engineering. *J Biol Chem*. **11**, 25364-25374.
- Käll, L., Krogh, A. and Sonnhammer, E.L.L. (2004). A combined transmembrane topology and signal peptide prediction method, *J Mol Biol*, **338**, 1027-1036.
- Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., Cheetham, M. E., Chen, B., and Hightower, L. E. (2009): Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress and Chaperon*. **14**, 105–111.
- Kampinga, H.H. and Craig, E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol*. **11**, 579-592.
- Kampinga, H.H., Hageman, J., Vos, M.J., Kubota, H., Tanguay, R.M., Bruford, E.A., Cheetham, M.E., Chen, B. and Hightower, L.E. (2009) Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress and Chaperon*. **14**, 105-111.
- Karlin, S. and Brocchieri, L. (1998) Heat shock protein 70 family: multiple sequence comparisons, function, and evolution. *J Mol Evol*. **47**, 565-577.
- Karplus, P.A. and Schulz, G.E. (1985) Prediction of Chain Flexibility in Proteins - a Tool for the Selection of Peptide Antigens, *Naturwissenschaften*. **72**, 212-213.
- Kathir, K. M., Kumar, T. K. S., Rajalingam, D. and Yu, C. (2005). Time-dependent changes in the denatured state(s) influence the folding mechanism of an all sheet protein. *J. Biol. Chem*. **280**, 29682-29688.
- Kelley, W.L. (1999) Molecular Chaperones: how J-domains turn on Hsp70s. *Curr. Biol*. **9**, R305-R314.
- Kennedy, P.G.E. (2006). Diagnostic and neuropathogenesis issues in human African trypanosomiasis. *Int. J. Parasitol*. **36**, 505-512
- Kirschke, E., Goswami, D., Southworth, D., Griffin, P. R., and Agard, D. A. (2014). Glucocorticoid receptor function regulated by coordinated action of the hsp90 and hsp70 chaperone cycles. *Cell*. **157**, 1685-1697.
- Kohl, L. and Bastin, P. (2005). The flagellum of trypanosomes. *Int. Rev. Cytol*. **244**; 227–285.
- Kohl, L., Robinson, D. and Bastin, P. (2003). Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes. *EMBO J*. **22**, 5336–5346.
- Kota P, Summers DW, Ren HY, Cyr DM, Dokholyan NV. (2009) Identification of a consensus motif in substrates bound by a Type I Hsp40. *Proc Natl Acad Sci USA*. **106**, 11073–8.
- Kumar, N., Zhao, Y., Graves, P., Perez Folgar, J., Maloy, L. and Zheng (1991) Induction and localization of *Plasmodium falciparum* stress proteins related to the heat shock protein 70 family, *Mol Biochem Parasitol*. **48**, 47-58.
- Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. **157**, 105-132.

- Lai, B.-T., Chin, N. W., Stanek, A. E., Keh, W., and Lanks, K. W. (1984). Quantitation and intracellular localization of the 85K heat shock protein by using monoclonal and polyclonal antibodies. *Mol. Cell. Biol*. **4**, 2802-2810.
- Lamb, J.R., Tugendreich, S. and Hieter, P. (1995). Tetratrico peptide repeat interactions: to TPR or not to TPR ? *Tr Biochem Sci*. **20**, 257-259.
- Lässle, M., Blatch, G.L., Kundra, V., Takatori, T and Zetter, B.R. (1997). Stress-inducible, murine protein mSTI1. Characterization of binding domains for heat shock proteins and *in vitro* phosphorylation by different kinases. *J. Biol. Chem*. **3**, 1876-1884.
- 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. Proceedings of the National Academy of Sciences, 96, 5452-5457.
- Lee, C.-T., Graf, C., Mayer, F. J., Richter, S. M., and Mayer, M. P. (2012). Dynamics of the regulation of Hsp90 by the co-chaperone Sti1. *EMBO J*. **90**, 1–11.
- Lee, M.G. and van der Ploeg, L.H. (1990). Transcription of the heat shock 70 locus in *Trypanosoma brucei*. *Mol. Biochem. Parasitol*. **41**, 221-231
- Lee, M.G-S., Polvere, R.I. and van der Ploeg, L.H.T. (1990). Evidence for segmental gene conversion between a cognate Hsp70 gene and the temperature-sensitively transcribed Hsp70 genes of *Trypanosoma brucei*. *Mol Biochem Parasitol*. **41**, 213-220
- Leskovar, A., Wegele, H., Werbeck, N. D., Buchner, J and Reinstein, J. (2008). The ATPase cycle of the mitochondrial Hsp90 analog Trap1. *J.Biol Chem*. **283**, 11677–11688.
- Letunic, I., Doerks, T. and Bork, P. (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res*. **40**, D302-D305.
- Li, H.B., and Du, Y.Z. (2013). Molecular cloning and characterisation of an Hsp90/70 organising protein gene from Frankliniella occidentalis (Insecta: Thysanoptera, Thripidae). *Gene*. **15**, 148-155.
- Li, J., Qian, X. and Sha, B. (2003) The Crystal Structure of the Yeast Hsp40 Ydj1 Complexed with Its Peptide Substrate. *Struct*. **11**, 1475-1483.
- Li, J., Qian, X. and Sha, B. (2009). Heat shock protein 40: structural studies and their functional implications. *Protein Pept. Lett*. **16**, 606–612.
- Li, J., Richter, K and Buchner, J. (2011). Mixed Hsp90-cochaperone complexes are important for the progression of the reaction cycle. *Nat Struct Mol Biol*. **18**, 61–66.
- Li, J., Soroka, J., and Buchner, J. (2012a). The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. *Biochim. Biophys. Acta*. **1823**, 624-635.
- Liedberg B, Nylander C, Lundstorm I. (1983). Surface plasmon resonance for gas detection. *Sensors and Actuators*. **4**, 299-304.
- Lima, F. R., Arantes, C. P., Muras, A. G., Nomizo, R., Brentani, R. R., and Martins, V. R. (2007). Cellular prion protein expression in astrocytes modulates neuronal survival and differentiation. *J. Neurochem*. **103**, 2164-2176.
- Lindquist, S. (1986). The Heat-Shock Response. *Annu Rev Biochem.* **55**, 1151-1191.
- Lindquist, S. and Craig, E.A. (1988) The Heat-Shock Proteins. *Annu Rev Genet*. **22**, 631-677.
- Lindquist, S. and Craig, E.A. (1988). The Heat-Shock Proteins. *Annu Rev Genet*. **22**, 631-677.
- Linke K, Wolfram T, Bussemer J, Jakob U. (2003). The roles of the two zinc binding sites in DnaJ. *J Biol Chem*. **278**, 44457–66.
- Longshaw, V. M., Baxter, M., Prewitz, M., and Blatch, G. L. (2009). Knockdown of the co-chaperone Hop promotes extranuclear accumulation of Stat3 in mouse embryonic stem cells. *Eur. J. Cell Biol*. **88**, 153-166.
- Louw, C.A., Ludewig, M.H. and Blatch, G.L. (2009). Overproduction, purification and characterization of Tbj1, a novel Type III Hsp40 from *Trypanosoma brucei*, the African sleeping sickness parasite. *Prot Exp Purific*. **69**,168-177.

- Louw, C.A., Ludewig, M.H., Mayer, J. and Blatch, G.L. (2010). The Hsp70 chaperones of the Tritryps are characterized by unusual features and novel members. *Parasitol. Int*. **59**; 497-505.
- Lu, Z. And Cyr, D.M. (1998). The conserved carboxyl terminus and zinc finger-like domain of the co-chaperone YDJ1 assist Hsp70 in protein folding. *J. Biol. Chem*. **273**, 5970-5978.
- Ludewig, M.H. (2009). The characterization of trypanosomal Type I DnaJ-like proteins. PhD Thesis, Rhodes University, Grahamstown, South Africa.
- Ludewig, MH, Boshoff, A, Horn, D and Blatch, GL. (2015). *Trypanosoma brucei* J protein 2 is a stress inducible and essential Hsp40. *Int. J. Biochem Cell Biol*. **1**, pp. 93-98.
- Lukes , J., Skalicky´, T., Ty´c, J, Voty´, J. and Yurchenkoa, V. (2014). Evolution of parasitism in kinetoplastid flagellates. *Mol and Biochem parasitol.* **195**, 115–122.
- Lukes, J., Hashimi, H. and Zikova, A. (2005). Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. *Curr Genet*. **48**, 277-299.
- Luscher A, Onal P, Schweingruber A.M., Maser, P. (2007). Adenosine kinase of *Trypanosoma brucei* and its role in susceptibility to adenosine antimetabolites. *Antimicrob Agents Chemother*. **51**, 3895–3901
- Ma, Y., Greener, T., Pacold, M. E., Kaushal, S., Greene, L. E., and Eisenberg, E. (2002): Identification of domain required for catalytic activity of auxilin in supporting clathrin uncoating by Hsc70. *J. Biol Chem*. **277**, 49267–49274.
- Macleod, A., Tait, A. and Turner, C. M. (2001). The population genetics of *Trypanosoma brucei* and the origin of human infectivity. *Philos Trans R Soc Lond, B, Biol Sci.* **356**, 1035–1044.
- Manna, P. T., Kelly, S., and Field, M. C. (2014). Molecular Phylogenetics and Evolution Adaptin evolution in kinetoplastids and emergence of the variant surface glycoprotein coat in African trypanosomatids. *Mol. Phylogenet. Evolut*. **67**, 123–128.
- Marchler, G., and Wu, C. (2001). Modulation of Drosophila heat shock transcription factor activity by the molecular chaperone DROJ1. *EMBO J*. **20**, 499-509.
- Marcu, M.G., Chadli, A., Bouhouche, I., Catelli, M and Neckers, L.M. (2000a). The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP binding domain in the carboxyl terminus of the chaperone. *J. Biol. Chem*. **275**, 37181–37186.
- Marcu, M.G., Schulte, T.W and Neckers L. (2000b). Novobiocin and related coumarins and depletion of heat shock protein 90-dependent signaling proteins. *J. Nat. Cancer Inst*. **92**, 242–248.
- Martinez-Yamout, M., Legge, G.B., Zhang, O., Wright, P.E. and Dyson, H.J. (2000) Solution structure of the cysteine-rich domain of the *Escherichia coli* chaperone protein DnaJ. *J.Mol. Biol*. **300**, 805-818.
- Martins, V. R., Graner, E., Garcia-Abreu, J., de Souza, S. J., Mercadante, A. F., Veiga, S. S., Zanata, S. M., Neto, V. M., and Brentani, R. R. (1997). Complementary hydropathy identifies a cellular prion protein receptor. *Nat. Med*. **3**, 1376-1382.
- Matthews, K.R. (2005). The developmental cell biology of *Trypanosoma brucei*. *J. Cell. Sci*. **118**, 283-290.
- Matts, R. L., Dixit, A., Peterson, L. B., Sun, L., Voruganti, S., Kalyanaraman, P., Hartson, S.D., Verkhivker, G.M and Blagg, B. S. (2011). Elucidation of the Hsp90 C-Terminal Inhibitor Binding Site. *ACS Chem Biol*. **6**, 800-807.
- Mayer, M. P., Nikolay, R., and Bukau, B. (2002). Aha, another regulator for Hsp90 chaperones. *Mol. Cell*. **10**, 1255-1256.
- Mayer, M.P. and Bukau, B. (2005) Hsp70 chaperones: cellular functions and molecular mechanism, *Cell Mol Life Sci*. **62**, 670-684.
- Mayer, M.P., Laufen, T., Paal, K., McCarty, J.S. and Bukau, B. (1998) Investigation of the interaction between DnaK and DnaJ by surface plasmon resonance spectroscopy. *J. Mol. Biol*. **289**, 1131-1144.
- Mayer, M.P., Rudiger, S. and Bukau, B. (2000a) Molecular basis for interactions of the DnaK chaperone with substrates. *Biol Chem*. **381**, 877-885.

- Mayr, C., Richter, K., Lilie, H and Buchner, J. (2000). Cpr6 and Cpr7, two closely related Hsp90-associated immunophilins from *Saccharomyces cerevisiae*, differ in their functional properties. *J. Biol. Chem*. **275**, 34140–34146.
- McCulloch, R. (2004). Antigenic variation in African trypanosomes: monitoring progress. *Tr. Parasitol*. **20**, 117-121.
- McLaughlin, S. H., Sobott, F., Yao, Z. P., Zhang, W., Nielsen, P. R., Grossmann, J. G., Laue, E. D., Robinson, C. V., and Jackson, S. E. (2006). The co-chaperone p23 arrests the Hsp90 ATPase cycle to trap client proteins. *J. Mol. Biol*. **356**, 746-758.
- Mehrpour, M., Esclatine, A., Beau, I. and Codogno, P. (2010). Overview of macroautophagy regulation in mammalian cells. *Cell Res*. **20**, 748-62.
- Melville, M.W., Tan, S.L., Wambach, M., Song, J., Morimoto, R.I. and Katze, M.G. (1999). The cellular inhibitor of the PKR protein kinase, P58(IPK), is an influenza virus-activated cochaperone that modulates heat shock protein 70 activity. *J. Biol Chem*. **274**, 3797-3803
- Meng, X., Jérôme ,V., Devin, J., Baulieu, E.E and Catelli M.G. (1993). Cloning of chicken hsp90 beta: the only vertebrate hsp90 insensitive to heat shock. *Biochem. Biophysic Res Comm*. **190**, 630-636.
- Meyer, K. J., and Shapiro, T. A. (2013). Potent antitrypanosomal activities of heat shock protein 90 inhibitors *in vitro and in vivo*. *J. Infect. Dis*. **208**, 489-499.
- Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S.M., Panaretou, B., Piper, P.W and Pearl, L.H. (2003). Structural and functional analysis of the middle segment of Hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. *Mol Cell*. **11**, 647 –58.
- Mhlanga, J.D.M. (1994). Antigenic variation in *Trypanosoma brucei*, a relationship with poly ADP-ribose polymerase. Thesis: University of Sussex, UK.
- Millson, S. H., Truman, A. W., Rácz, A., Hu, B., Panaretou, B., Nuttall, J., Mollapour, M., Söti, C., and Piper, P. W. (2007). Expressed as the sole Hsp90 of yeast, the α and β isoforms of human Hsp90 differ with regard to their capacities for activation of certain client proteins, whereas only Hsp90β generates sensitivity to the Hsp90 inhibitor radicicol. *FEBS J*. **274**, 4453-4463.
- Minet, E., Mottet, D., Michel, G., Roland, I., Raes, M., Remacle, J., and Michiels, C. (1999). Hypoxia-induced activation of HIF-1: role of HIF-1a-Hsp90 interaction. *FEBS Lett*. **460**, 251-256.
- Misra, G. and Ramachandran, R. (2009) Hsp70-1 from *Plasmodium falciparum*: protein stability, domain analysis and chaperone activity, *Biophys Chem*. **142**, 55-64.
- Moffatt, N. S., Bruinsma, E., Uhl, C., Obermann, W. M., and Toft, D. (2008). Role of the cochaperone Tpr2 in Hsp90 chaperoning. *Biochem*. **47**, 8203-8213.
- Mollapour, M. and Neckers, L., 2012. Post-translational modifications of Hsp90 and their contributions to chaperone regulation. *Biochimica et biophysica acta*. **1823**, pp.648–55.
- Montagnes, D., Roberts, E., Lukeš, J. and Lowe, C. (2012). The rise of model protozoa. *Tr Microbiol*. **20**, 184-191.
- Morales, M. A., Watanabe, R., Dacher, M., Chafey, P., Osorio y Fortea, J., Scott, D. A., Beverley, S. M., Ommen, G., Clos, J., Hem, S. (2010). Phosphoproteome dynamics reveal heat-shock protein complexes specific to the Leishmania donovani infectious stage. *Proc Natl Acad Sci USA*. **107**, 8381-8386.
- Morrison, L. J., Marcello, L. and McCulloch, R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cell Microbiol.* **11**, 1724–1734.
- Nadeau K, Sullivan MA, Bradley M, Engman DM, Walsh, C.,T. (1992) 83-kilodalton heat shock proteins of trypanosomes are potent peptide-stimulated ATPases. *Protein Sci.* **1**, 970–979.
- Nakamoto, H., Fujita, K., Ohtaki, A., Watanabe, S., Narumi, S., Maruyama, T., Suenaga, E., Misono, T. S., Kumar, P. K., Goloubinoff, P. et al. (2014). Physical interaction between bacterial heat shock protein (Hsp) 90 and Hsp70 chaperones mediates their cooperative action to refold denatured proteins. *J. Biol. Chem*. **289**, 6110-6119.

- Namangala, B., Noël, W., de Baetselier, P., Brys, L. and Beschin, A. (2011). Relative contribution of interferon-gamma and interleukin-10 to resistance to murine African trypanosomosis. *J Infect Dis*. **183**, 1794–1800.
- Nemoto, T., Ohara-Nemoto, Y., Ota, M., Takagi, T., and Yokoyama, K. (1995). Mechanism of dimer formation of the 90-kDa heat-shock protein. *Eur. J. Biochem*. **233**, 1-8.
- Nett IRE, Martin DMA, Miranda-Saavedra D, Lamont D, Barber JD, Mehlert A, Ferguson MAJ (2009) The phosphoproteome of bloodstream form *Trypanosoma brucei* , causative agent of African sleeping sickness. *Mol Cell Proteomics*. **8**, 1527–1538.
- Nicolet, C.M. and Craig, E.A. (1989). Isolation and characterisation of STI1, a stress inducible gene from *Saccharomyces cerivisiae*. *Mol Cell Biol*. **9**, 3638-3646.
- Njunge, J. M., Ludewig, M. H., Boshoff, A., Pesce, E., and Blatch, G. L. (2013): Hsp70s and J proteins of *Plasmodium* parasites infecting rodents and primates : Structure , function , clinical relevance , and drug targets. *Curr. Pharm. Des*. **19**, 387-403.
- Nyalwidhe, J., Maier, U-G. and Lingelbach, K. (2003). Intracellular parasitism: cell biological adaptations of parasitic protozoa to a life inside cells. *Zool*. **106**, 341-348.
- Odiit M, Coleman P G, Liu W C, McDermott J J, Fevre E M, Welburn S C and Woolhouse M E (2005) Quantifying the Level of Under-Detection of *Trypanosoma brucei rhodesiense* Sleeping Sickness Cases. *Trop Med Int Health*. 10, 840-849.
- Odunuga, O.O., Hornby, J.A., Bies, C., Zimmerman, R., Pugh, D.J. and Blatch, G.L. (2003). Tetratricopeptide repeat motif-mediated Hsc70-mSTI1 interaction: Molecular characterisation of the criticial contacts for successful binding and specificity. *J. Biol. Chem.* **278**, 6896-6904.
- Odunuga, O.O., Longshaw, V.M. and Blatch, G.L. (2004). Hop: more than an Hsp70/Hsp90 adaptor protein. *Bioessays*. **26**, 1058-1068.
- 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
- Ommen, G.; Chrobak, M.; Clos, J. (2010). The co-chaperone SGT of *Leishmania donovani* is essential for the parasite's viability. *Cell Stress Chaperon*. **15**, 443-455.
- Onuoha, S. C., Coulstock, E. T., Grossmann, J. G and Jackson, S. E. (2008). Structural studies on the cochaperone Hop and its complexes with Hsp90. *J. Mol. Biol*. **379**, 732-744.
- O'Shannessy, D.J., Brigham-Burke, M., Soneson, K.K., Hensley, P., and Brooks, I. (1993). Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods. *Anal Biochem.* **212**, 457-468.
- Pallavi, R., Roy, N., Nageshan, R. K., Talukdar, P., Pavithra, S. R., Reddy, R., Venketesh, S., Kumar, R., Gupta, A. K., Singh, R. K. (2010). Heat shock protein 90 as a drug target against protozoan infections. *J. Biol. Chem*. **285**, 37964-37975.
- Palmer,G.H., Brayton,K.A. (2007) Gene conversion is a convergent strategy for pathogen antigenic variation. *Trends Parasitol*. **23**, 408-413.
- Parsons M, Worthey EA, Ward PN, Mottram JC (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major* , *Trypanosoma brucei* and *Trypanosoma cruzi* . *BMC Genomics*. **6**, 127.
- Parsons, M., Nelson, R. G., Watkins, K. P. and Agabian, N. (1984). Trypanosome mRNAs share a common 5' spliced leader sequence. *Cell*. **38**, 309–316.
- Parsons, M., Nelson, R. G., Watkins, K. P. and Agabian, N. (1984). Trypanosome mRNAs share a common 5' spliced leader sequence. *Cell*. **38**, 309–316.
- Pays, E., Vanhamme, L., and Perez-Morga, D. (2004) Antigenic variation in *Trypanosoma brucei*: Facts, Challenges and mysteries. *Curr Opin Microbiol*. **7**, 369-374.

- Peacock, L., Ferris, V., Bailey, M., and Gibson, W. (2014). Mating compatibility in the parasitic protist *Trypanosoma brucei*. *Parasit .Vect*. **7**, 1–10.
- Pearl, L. H. (2005). Hsp90 and Cdc37 -- a chaperone cancer conspiracy. *Curr. Opin. Genet. Dev*. **15**, 55-61.
- Pearl, L. H., and Prodromou, C. (2000). Structure and *in vivo* function of Hsp90. *Curr. Opin. Struct. Biol*. **10**, 46-51.
- Pearl, L. H., and Prodromou, C. (2006). Structure and mechanism of the hsp90 molecular chaperone machinery. *Annu. Rev. Biochem*. **75**, 271-294.
- Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C., and Wüthrich, 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.
- Pepin, K., Momose, F., Ishida, N and Nagata, K. J. (2001). Molecular cloning of horse Hsp90 cDNA and its comparative analysis with other vertebrate Hsp90 sequences. *J.Vet. Med. Sci*. **63**, 115–124.
- Picard, D., Khursheed, B., Garabedian, M. J., Fortin, M. G., Lindquist, S., and Yamamoto, K. R. (1990). Reduced levels of hsp90 compromise steroid receptor action in vivo. *Nat*. **348**, 166-168.
- Pimienta, G., Herbert, K.M., and Regan, L. (2011). A compound that inhibits the HOP-Hsp90 complex formation and has unique killing effects in breast cancer cell lines. *Mol Pharmacol.* **8**, 2252–2261.
- Pizarro, J. C., Hills, T., Senisterra, G., Wernimont, A. K., Mackenzie, C., Norcross, N. R., Ferguson, M. A., Wyatt, P. G., Gilbert, I. H., and Hui, R. (2013). Exploring the *Trypanosoma brucei* Hsp83 potential as a target for structure guided drug design. *PLoS Negl. Trop. Dis*. **7**, e2492.
- Pratt, W. B., and Toft, D. O. (2003). Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. *Exp. Biol. Med*. **228**, 111-133.
- Pratt, W.B. and Toft, D.O. (1997). Steroid Receptor Interactions with Heat Shock Protein and Immunophilin Chaperones. *Endoc Rev*. **18**, 306-360.
- Prodromou, C., Panaretou, B., Chohan, S., Siligardi, G., O'Brien, R., Ladbury, J. E., Roe, S. M., Piper, P. W., and Pearl, L. H. (2000). The ATPase cycle of Hsp90 drives a molecular 'clamp' via transient dimerization of the N-terminal domains. *EMBO J*. **19**, 4383-4392.
- Prodromou, C., Siligardi, G., O'Brien, R., Woolfson, D.N., Regan, L., Panaretou, B., Ladbury, J.E., Piper, P.W and Pearl, L.H. (1999). Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR)-domain cochaperones. *EMBO J*. **18**, 754–762.
- Rassi A Jr, Rassi A, Marin-Neto JA (2010) Chagas disease. *Lanc*. **375**, 1388–1402
- Rassi A Jr, Rassi A, Rezende JM (2012) American trypanosomiasis (Chagas disease). *Infect Dis Clin N Am.* **26**, 275–291
- Reidy, M., and Masison, D. C. (2010). Sti1 regulation of Hsp70 and Hsp90 is critical for curing of *Saccharomyces cerevisiae* [PSI+] prions by Hsp104. *Mol. Cell. Biol*. **30**, 3542-3552.
- Requena JM, Lopez MC, Jimenez-Ruiz A, Morales G, Alonso C (1989) Complete nucleotide sequence of the hsp70 gene of *Trypanosoma cruzi* . *Nucleic Acids Res.* **17**, 797–797.
- Requena, J., Jiminez-Ruiz, A., Soto, M., Assiego, R., Santaren, J.F., Lopez, M.C., Patarroyo, M.E. and Alonso, C. (1992). Regulation of hsp70 expression in *Trypanosoma cruzi* by temperature and growth phase. *Mol. Biochem. Parasitol*. **53**, 201-212.
- Requena, J.M., Lopez, M.C., Jiminez-Ruiz, A., de la Torre, J.C. and Alonso, C. (1988). A head-to-tail tandem organisation of hsp70 genes in *Trypanosoma cruzi*. *Nucleic Acids Res*. **16**, 1393-1406
- Retzlaff, M., Hagn, F., Mitschke, L., Hessling, M., Gugel, F., Kessler, H., Richter, K., and Buchner, J. (2010). Asymmetric activation of the Hsp90 dimer by Its cochaperone Aha1. *Mol. Cell*. **37**, 344-354.
- Rich RL, Myszka DG. 2000. Advances in surface plasmon resonance biosensor analysis. *Curr. Opin.in Biotechnol*. **11**, 54-61.
- Rich RL, Myszka DG. 2004. Why you should be using more SPR biosensor technology.Drug Discov. Today. 1:301–308.
- Riezman, H. (2004) Why do cells require heat shock proteins to survive heat stress?, *Cell Cycle*. **3**, 60-62.

- Roffé, M., Beraldo, F. H., Bester, R., Nunziante, M., Bach, C., Mancini, G., Gilch, S., Vorberg, I., Castilho, B. A., Martins, V. R. (2010). Prion protein interaction with stress-inducible protein 1 enhances neuronal protein synthesis via mTOR. *Proc. Natl. Acad. Sci. USA* . **107**, 13147-13152.
- Rosser, M.F.N. and Cyr, D.M. (2007). Do Hsp40s Act as Chaperones or Co-Chaperones? In BLATCH, G.L. (ed.),Networking of Chaperones by Co-Chaperones. Georgetown, *Landes Biosci*. **1**, 38-51.
- Rudenko, G. (2011) African trypanosomes: the genome and adaptations for immune evasion. *Essays Biochem*. **51**, 47–62.
- Sato, S., Fujita, N., and Tsuruo, T. (2000). Modulation of Akt kinase activity by binding to Hsp90. *Proc. Natl.Acad. Sci. USA*. **97**, 10832-10837.
- Scheibel, T., Siegmund, H. I., Jaenicke, R., Ganz, P., Lilie, H., and Buchner, J. (1999a). The charged region of Hsp90 modulates the function of the N-terminal domain. *Proc. Natl. Acad. Sci. USA*. **96**, 1297-1302.
- Scheufler, C., Brinker, A., Bourenkov, G., Pegararo, S., Moroder, L., Bartunik, H., Hartl, F. U and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell*. **101**, 199-210.
- Schmid, A. B., Lagleder, S., Grawert, M. A., Rohl, A., Hagn, F., Wandinger, S. K., Cox, M. B., Demmer, O., Richter, K., Groll, M. (2012). The architecture of functional modules in the Hsp90 co-chaperone Sti1/Hop. *EMBO J*. **31**, 1506-1517.
- Schmid, A. B., Lagleder, S., Gräwert, M. A., Röhl, A., Hagn, F., Wandinger, S. K., Cox, M. B., et al. (2012). The architecture of functional modules in the Hsp90 co-chaperone Sti1/Hop. *EMBO J*. **31**, 1506–1517.
- Schmidt, J.C., Soares, M.J., Samuel Goldenberg, S., Pavoni, D.P., and Krieger, M.A. (2011). Characterisation of TcSTI-1, a homologue of stress induced protein-1, in *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz*. **106**, 70–77.
- Schmidt, T. G. M., and Skerra, A. (1993): The random peptide libraryassisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment. *Protein Eng*. **6**, 109 – 122.
- Schweinfest, C.W., Graber, M.W., Henderson, K.W., Papas, T.S., Baron, P.L and Watson, D.K. (1998). Cloning and sequence analysis of Hsp89alpha DeltaN, a new member of the Hsp90 gene family. *Biochimica et Biophysica Acta*. **1398**, 18–24.
- Shiau, A. K., Harris, S. F., Southworth, D. R and Agard, D. A. (2006). Structural Analysis of E. coli Hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements. *Cell*. 1**27**, 329-340.
- Shonhai A, Maier AG, Przyborski JM, Blatch GL (2011) Intracellular protozoan parasites of humans: the role of molecular chaperones in development and pathogenesis. *Protein Pept Lett*. **18**, 143–15.
- Sibley, L.D. and Andrews, N.W. (2000). Cell invasion by un-palatable parasites. *Traffic*. **2**, 100-106.
- Sigrist, C.J.A*., et al.* (2010) PROSITE, a protein domain database for functional characterization and annotation, *Nucleic Acids Res*. **38**, D161-D166.
- Sikorski, R.S., Boguski, M.S., Goebl, M. and Hieter, P. (1990). A repeating amino acid motif in CDC23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. *Cell*. **60**, 307-317.
- Silberg, J. J., Hoff, K. G., and Vickery, L. E. (1998): The Hsc66-Hsc20 chaperone system in Escherichia coli: chaperone activity and interactions with the DnaK-DnaJ-grpE system. *J. Bacteriol*. **180**, 6617–24.
- Simpson, A.G., Stevens, J. and Lukes, J. (2006). The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol*. **22**, 168-174.
- Simpson, A.G.B., Gill, E.E., Callahan, H.A., Litaker, R.W. and Roger, A.J. (2004). Early evolution within kinetoplastids (euglenozoa), and the late emergence of trypanosomatids. *Protist*. **155**, 407–422.
- Smejkal, R.M., Wolff, R. and Olenich, J.G. (1988). *Experimental Parasitol*. **65**, 1-9.
- Smith, D. B., and Johnson, K. S. (1988): Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*. **67**, 31 – 40.

- Smith, D.F. (1993). Dynamics of heat shock protein 90–progesterone receptor binding and the disactivation loop model for steroid receptor complexes. *Mol Endocrinol*. **7**, 1418–1429.
- Smith, D.F. (2004). Tetratricopeptide repeat cochaperones in steroid receptor complexes. *Cell Stress and Chaperon*. **9**, 109-121.
- Sondermann, H., Scheuffler, C., Schneider, C., Hohfeld, J., Hartl, F. U. and Moarefi, I. (2001). Structure of a Bag/Hsc70 Complex: Convergent Functional Evolution of Hsp70 Nucleotide Exchange Factors. *Sci*. **291**, 1553-7.
- Song, H. O., Lee, W., An, K., Lee, H. S., Cho, J. H., Park, Z. Y., and Ahnn, J. (2009). *C. elegans* STI-1, the homolog of Sti1/Hop, is involved in aging and stress response. *J. Mol. Biol*. **390**, 604-617.
- Sóti, C., Rácz, A., and Csermely, P. (2002). A nucleotide-dependent molecular switch controls ATP binding at the C- terminal domain of Hsp90: N-terminal nucleotide binding unmasks a C- terminal binding pocket. *J. Biol. Chem*. **277**, 7066-7075.
- Southworth, D.R., and Agard, D.A. (2011). Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90: Hop complex. *Mol Cell*. **42**, 771-781.
- Stebbins, C. E., Russo, A. A., Schneider, C., Rosen, N., Hartl, F. U and Pavletich, N. P. (1997). Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. *Cell*. **89**, 239-250
- Steverding, D. (2008). The history of African trypanosomiasis. Parasites and Vectors, 12 (1),3
- Steverding, D. (2014). Visible spectral distribution of shadows explains why blue targets with a high reflectivity at 460 nm are attractive to tsetse flies. *Parasit. Vect*. **6**, 285-286.
- Stirling, P.C., Lundlin, V.F. and Leroux, M.R. (2003) Getting a grip on non-native proteins. European Molecular Biology Organisation (EMBO) Journal 4(6): 565-570.
- Street, T. O., Zeng, X., Pellarin, R., Bonomi, M., Sali, A., Kelly, M. J., Chu, F., and Agard, D. A. (2014). Elucidating the mechanism of substrate recognition by the bacterial Hsp90 molecular chaperone. *J. Mol. Biol*. **426**, 2393-2404.
- Stuart KD, Andersson B (2005) The genome sequence of *Trypanosoma cruzi* , etiologic agent of Chagas' disease. *Sci*. **309**, 409–415.
- Subramaniam, C., Veazey, P., Redmond, S., Hayes-Sinclair, J., Chambers, E., Carrington, M., Gull, K., Matthews, K., Horn, D. and Field, M.C. (2006). Chromosomewide analysis of gene function by RNA interference in the African Trypanosome. Eukaryotic *Cell*. **5**, 1539-1549.
- Suh, W.-C. C., Burkholder, W. F., Lu, C. Z., Zhao, X., Gottesman, M. E., and Gross, C. A. (1998): Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proceed Nat Acad.Sci*. **95**, 15223–8.
- Svärd, M. et al., 2011. The Crystal Structure of the Human Co-Chaperone P58 IPK. *PloS one*. *6*.
- Swain, J., Dinler, G., Sivendran, R, Montgomery, D., Stotz, M. and Gierasch, L. (2008) Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated bu the interdomain linker. *Mol. Cell*. **26**, 27-39.
- Taipale, M., Jarosz, D. F., and Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell*. *Biol*. **11**, 515-528.
- Tao, J. et al., 2010. Crystal structure of P58(IPK) TPR fragment reveals the mechanism for its molecular chaperone activity in UPR. *J. Mol. Biol*. **397**, pp.1307–1315.
- Terasawa, K., Minami, M and Minami, Y. (2005). Constantly updated knowledge of Hsp90. *J. Biochem*. **137**, 443-447.

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. Biochem. Parasitol*. **91**, 319-326.

- Tielens AG, Van Hellemond JJ (1998) Differences in energy metabolism between trypanosomatidae. *Parasitol Today* **14**, 265–272.
- Travers, S. A., and Fares, M. A. (2007). Functional coevolutionary networks of the Hsp70-Hop-Hsp90 system revealed through computational analyses. *Mol. Biol. Evol*. **24**, 1032-1044.

- Urményi TP, Rodrigues DC, Silva R, Rondinelli E (2012) The stress response of *Trypanosoma cruzi* . In: Requena JM (ed) Stress response in microbiology. Caister Academic Press, Norwich, pp 345–373.
- Urwyler S, Studer E, Renggli C K and Roditi I (2007) A Family of Stage-Specific Alanine-Rich Proteins on the Surface of Epimastigote Forms of *Trypanosoma brucei. Mol Microbiol*. **63**,218-228.
- Van Den Abbeele J, Claes Y, van Bockstaele D, Le Ray D and Coosemans M (1999) *Trypanosoma brucei Spp*. Development in the Tsetse Fly: Characterization of the Post-Mesocyclic Stages in the Foregut and Proboscis. *Parasitol*. **118**, 469-478.
- Van der Ploeg, L.H.T., Giannini, S.H. and Cantor, C.R. (1985). Heat shock genes: regulatory role for differentiation in parasitic protozoa. *Sci*. **228**, 1443-1446.
- van der Spuy, J., Cheetham M.E., Dirr H.W and Blatch G.L. (2001). The cochaperone murine stress-inducible protein 1: overexpression, purification, and characterization. *Protein Express Purific*. **21**, 462 –469.
- van Der Spuy, J., Kana, B. D., Dirr, H. W., and Blatch, G. L. (2000). Heat shock cognate protein 70 chaperone-binding site in the co- chaperone murine stress inducible protein 1 maps to within three consecutive tetratricopeptide repeat motifs. *Biochem. J*. **345**, 645-651.
- Van der Woude, M. W. and Bäumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**, 581–611.
- Vanhamme, L., Poelvoorde, P., Pays, A., Tebabi, P., Van Xong, H. and Pays, E. (2001). Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol Microbiol*. **36**, 328–340.
- Verner, Z., Paris, Z. and Lukeš, J. (2010). Mitochondrial membrane potential-based genome-wide RNAi screen of *Trypanosoma brucei*. *Parasitol. Res*. **106**, 1241-1244.
- Vertommen, D., Van Roy, J., Szikora, J.P., Rider, M.H., Michels, P.A. and Opperdoes, F.R. (2008). Differential expression of glycosomal and mitochondrial proteins in the two major life-cycle stages of *Trypanosoma brucei*. *Mol. Biochem Parasitol*. **158**, 189-201
- Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. *Br. J. Med. Bull*. **41**, 105–114.
- Vickerman, K. (1997). Landmarks in trypanosome research. In: Hide, G., Mottram, J.C., Coombs, G.H. and Holmes, P.H. (Eds.) Trypanosomiasis and Leishmaniasis. CAB International, Oxford, pp. 1-37
- Vonlaufen N, Kanzok SM, Wek RC, Sullivan WJ (2008) Stress response pathways in protozoan parasites. *Cell Microbiol.* **10**, 2387–2399.
- Wall, D., Zylicz, M. and Georgopoulos, C. (1995). The conserved G/F motif of the DnaJ chaperone is necessary for the activation of the substrate binding properties of the DnaK chaperone. *J. Biol. Chem*. **270**, 2139-2144.
- Walsh, P., Bursac, D., Law, Y.C., Cyr, D. and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. *Euro Mol Biol Org*. **23**, 567–571.
- Wanderling, S., Simen, B. B., Ostrovsky, O., Ahmed, N. T., Vogen, S. M., Gidalevitz, T and Argon, Y. (2007). GRP94 is essential for mesoderm induction and muscle development because it regulates insulin-like growth factor secretion. *Mol. Biol. Cell* **18**, 3764-3775.
- Wang, T.-H., Chao, A., Tsai, C.-L., Chang, C.-L., Chen, S.-H., Lee, Y.-S., Chen, J.-K., et al. (2010). Stress-induced phosphoprotein 1 as a secreted biomarker for human ovarian cancer promotes cancer cell proliferation. *Mol. Cell. Proteomic*. **9**, 1873–84.
- Wayne, N., and Bolon, D. N. (2007). Dimerization of Hsp90 is required for *in vivo* function: design and analysis of monomers and dimers. *J. Biol. Chem***. 282**, 35386-35395.
- Wegele, H., Haslbeck, M., Reinstein, J., and Buchner, J. (2003). Sti1 is a novel activator of the Ssa proteins. *J. Biol. Chem*. **278**, 25970-25976.
- Wegele, H., Muschler, P., Bunck, M., Reinstein, J., and Buchner, J. (2003). Dissection of the contribution of individual domains to the ATPase mechanism of Hsp90. *J. Biol. Chem*. **278**, 39303-39310.

- Werner-Washburne, M. and Craig, E.A. (1989). Expression of members of the *Saccharomyces cerevisiae* hsp70 multigene family. *Genome*. **31**, 684-689.
- Whitesell, L and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. *Nat. Rev. Canc*. **5,** 761-772.
- Wiesgigl M, and Clos J (2001) Heat shock protein 90 homeostasis controls stage differentiation in Leishmania donovani . *Mol Biol Cell* **12**, 3307–3316.
- Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A and Dodson, G. (1991). Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nat*. **351**, 624 -629.
- Wirtz, E., Leal, S., Ochatt, C. and Cross, G.A. (1999). A tightly regulated inducible system for the conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei. Mol. Biochem. Parasitol*. **99**, 89-101.
- Wittung-Stafshede, P., Guidry, J., Horne, B.E. and Landry, S.J. (2003). The J-Domain of Hsp40 Couples ATP Hydrolysis to Substrate Capture in Hsp70. *Biochem*. **42**, 4937-4944.
- World Health Organization (2010a) Control of the leishmaniases. *World Health Organ Tech Rep Ser* **949**, 186.
- World Health Organization (2010b) Chagas disease (American trypanosomiasis) fact sheet (revised in June 2010). *Wkly Epidemiol Rec* **85**, 334–336.
- World Health Organization (2012) Accelerating work to overcome the global impact of neglected tropical diseases: a roadmap for implementation. Geneva (WHO/HTM/NTD/2012.1).
- Wu, Y., Li, J., Jin, Z., Fu, Z. and Sha, B. (2005). The crystal structure of the C-terminal fragment of yeast Hsp40 Ydj1 reveals novel dimerisation motif for Hsp40. *J. Mol. Biol*. **346**, 1005-1011.
- Xu, Y., and Lindquist, S. (1993). Heat-shock protein hsp90 governs the activity of pp60v-src kinase. *Proc. Natl. Acad. Sci. USA* **90**, 7074-7078.
- Yamamoto, S., Subedi, G. P., Hanashima, S., Satoh, T., Otaka, M., Wakui, H., Sawada, K., Yokota, S., Yamaguchi, Y., Kubota, H. et al. (2014). ATPase activity and ATP-dependent conformational change in the co-chaperone Hsp70/Hsp90-organizing protein (HOP). *J. Biol. Chem*. **289**, 9880-9886.
- Yi, F., and Regan, L. (2008). A novel class of small molecule inhibitors of Hsp90. *ACS Chem. Biol*. **3**, 645-656.
- Yi, F., Doudevski, I and Regan, L. (2010). HOP is a monomer: Investigation of the oligomeric state of the cochaperone HOP. *Protein Sci*. **19**, 19-25.
- Young, J. C., Barral, J. M and Ulrich Hartl, F. (2003). More than folding: localized functions of cytosolic chaperones. *Trends Biochem Sci*. **28**, 541-547.
- Young, J. C., Obermann, W. M and Hartl, F. U. (1998). Specific binding of tetratricopeptide repeat proteins to the C-terminal 12-kDa domain of hsp90. *J. Biol. Chem*. **273**, 18007-18010.
- Young, J.C., Moarefi, I and Hartl, F.U. (2001). Hsp90: a specialized but essential protein folding tool. *J. Cell Biol*. **154**, 267-273.
- Zanata, S. M., Lopes, M. H., Mercadante, A. F., Hajj, G. N., Chiarini, L. B., Nomizo, R., Freitas, A. R., Cabral, A. L., Lee, K. S., Juliano, M. A. et al. (2002). Stressinducible protein 1 is a cell surface ligand for cellular prion that triggers neuroprotection. *EMBO J*. **21**, 3307-3316.
- Zhang, Z., Quick, M. K., Kanelakis, K. C., Gijzen, M., and Krishna, P. (2003). Characterization of a plant homolog of hop, a cochaperone of hsp90. *Plant Physiol*. **131**, 525-535.
- Zhao, R., Davey, M., Hsu, Y.C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., Boone, C., Emili, A and Houry, W.A. (2005). Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the Hsp90 chaperone. *Cell*. **120**, 715–727.
- Zuehlke, A. D., and Johnson, J. L. (2012). Chaperoning the chaperone: a role for the co-chaperone Cpr7 in modulating Hsp90 function in *Saccharomyces cerevisiae*. *Genetic* 191, 805-814.
- Zurawska, A., Urbanski, J and Bieganowski P. (2008). Hsp90N - An accidental product of a fortuitous chromosomal translocation rather than a regular Hsp90 family member of human proteome. *Biochimica et Biophysica Acta*. **1784**, 1844-1846.