

Biochemical and biophysical characterization of PFF1010c, a *Plasmodium falciparum* heat shock protein

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

Master of Science (MSc) in Biochemistry

Department of Biochemistry and Microbiology

University of Venda

Thohoyandou, Limpopo

South Africa

By

Thabo Rolffy Masoga

Student number: 22021309

Supervisor: Prof. A. Shonhai

February 2025

Abstract

Since time immemorial malaria remains a health concern despite the extensive interventions to eliminate the disease. *Plasmodium falciparum*, the agent for the most lethal form of the disease, survives insults imposed by the human host response and the adverse effects of antimalarial drugs, thereby building resistance against this intervention. Amongst other means, the parasite employs the heat shock protein (HSP) machinery in response to the imposed hostilities. Compared to other malaria parasites, *P. falciparum* expresses a large complement of an HSP family known as the J domain proteins (JDPs). Hsp70 is a prominent molecular chaperone whose chaperone function is coordinated by JDPs. While several JDPs have been characterized, the role of PFF1010c, a JDP of *P. falciparum* remains unknown. This study explored the structural and functional features of PFF1010c using *in silico approaches* (BLASTP, PSIPRED, CLUSTALW, MEGA11, SOPMA, CDD/SPARCLE ESPRIPT, Alphafold, GalaxyHomomer, ProtParam, PlasmoDB, and STRING). A plasmid construct expressing recombinant PFF1010c was expressed using *E. coli* XL1 Blue. The secondary structure of the protein was investigated using circular dichroism spectroscopy. PFF1010c thermostability was assessed by ANS probe based extrinsic fluorescence spectroscopy. The functional cooperation of PFF1010c with PfHsp70-1 was investigated by monitoring refolding of heat-denatured luciferase *in vitro*. PFF1010c appeared to assume a fold constituted by two domains: the SVN motif-containing J domain and the C-terminal domain. PFF1010c was found to be thermostable. Biochemical and *in silico* studies suggested that the protein self-associates through the C-terminal domain. Also, PFF1010c abrogated the refolding activity of PfHsp70-1. Although PFF1010c appeared to interact with PfHsp70-1, it was incapable of stimulating PfHsp70-1 chaperone function. This suggests that PFF1010c may interact with PfHsp70-1 possibly to improve the holdase chaperone function of PfHsp70-1 as well as to conserve the energy economy of the parasite cell.

Keywords: Hsp40, Hsp70, J domain, JDP, Malaria, PFF1010c, *P. falciparum*.

Declaration

I, Thabo Rolffy Masoga (Student number: 22021309), declare that I am the author of this thesis and submit it to the Department of Biochemistry and Microbiology, University of Venda, for the degree of MSc in Biochemistry. This work was not reproduced from any other study, either national or international papers. To the best of my knowledge, I certify that this work will not infringe upon anyone's copyright and that any ideas, techniques, or other material from the work of other people included in this thesis, published or otherwise, are fully acknowledged.

Signature



Date: 2025/01/24

Dedication

I dedicate this thesis to my late uncle, Mr Papamogolo Hendrick Masoga, for believing in me and making me work hard in my childhood, preparing me for the world and the future.

Acknowledgements

I thank my supervisor, Prof Addmore Shonhai, for taking me in as his student. His leadership is paramount, and it is a great honour to be under his supervision. Prof Shonhai's contribution to my literacy journey is invaluable.

I wish to acknowledge Ms Petunia Mudau for mentoring me, walking me through all the basic techniques used in the Protein Biochemistry and Malaria Research (ProBioM) group and advising me throughout this project. I am forever grateful for her unwavering support.

Prof Ntakadzeni Madala, Dr Lutendo Mathomu, and Dr Adelle Burger's offices were always open for me. I appreciate the support.

Much thanks to the DSI-HSRC internship programme for funding.

Finally, I would like to acknowledge all my laboratory mates in the ProBioM at the University of Venda.

Table of Contents

Abstract.....	i
Declaration.....	ii
Dedication	iii
Acknowledgements	iv
List of Figures	xi
List of tables	xii
List of appendices.....	xiii
List of symbols	xiv
Outputs	xv
1. Chapter 1: Introduction.....	1
1.1. Malaria.....	1
1.1.1. Malaria history.....	1
1.1.2. Agents of malaria: <i>Plasmodium</i> species	2
1.1.3. <i>Plasmodium</i> species life cycle.....	2
1.2. Heat shock proteins.....	7
1.2.1. Malaria heat shock proteins	8
1.2.2. <i>P. falciparum</i> heat shock protein 70	8
1.2.3. <i>P. falciparum</i> heat shock protein 100	12
1.2.4. <i>P. falciparum</i> heat shock protein 90	12

1.2.5.	<i>P. falciparum</i> heat shock protein 60	13
1.2.6.	<i>P. falciparum</i> small heat shock proteins	13
1.2.7.	<i>P. falciparum</i> J domain proteins	13
1.3.	Interventions against malaria.....	16
1.3.1.	Diagnosis	16
1.3.2.	Artemisinin combination therapy	17
1.3.3.	The Mosquirix vaccine	17
1.4.	Study rationale	17
1.5.	Hypothesis	18
1.6.	Aim	18
1.7.	Objectives	18
1.7.1.	In silico analyses to establish the structural features of PFF1010c.....	18
1.7.2.	Expression and purification of recombinant PFF1010c.....	18
1.7.3.	Biophysical and biochemical characterization.....	18
2.	Chapter 2: Materials and Methods	20
2.1.	Ethics concern.....	20
2.2.	Materials.....	20
2.2.1.	Equipment.....	20
2.2.2.	General reagents and chemicals.....	20
2.2.3.	Stock solutions, buffers, and growth media	21

2.2.4.	Kits	21
2.2.5.	Proteins, enzymes, and standards	21
2.2.6.	Plasmid constructs and bacterial strains.....	21
2.3.	Bioinformatic analyses.....	22
2.3.1.	PFF1010c homologue search	22
2.3.2.	Multiple sequence alignment.....	23
2.3.3.	Homologues phylogenetic tree	23
2.3.4.	Domain identification on PFF1010c.....	23
2.3.5.	Prediction of PFF1010c secondary and tertiary structures	24
2.3.6.	Prediction of PFF1010c homo-oligomerization	24
2.3.7.	Prediction of PFF1010c localisation	24
2.3.8.	Prediction of PFF1010c cellular protein associates.....	25
2.3.9.	Prediction of PFF1010c post-translational modification sites	25
2.3.10.	Prediction of physicochemical properties of PFF1010c	25
2.4.	Experimental analysis	26
2.4.1.	Preparation of competent cells	26
2.4.2.	Transforming <i>E. coli</i> competent cells	27
2.4.3.	Plasmid DNA extraction	27
2.4.4.	Restriction digest of plasmid DNA	27
2.4.5.	Agarose Gel Electrophoresis.....	28

2.4.6.	Expression of recombinant protein forms of PFF1010c, PfHsp40, and PfHsp70-1	29
2.4.7.	Protein solubility assessment	29
2.4.8.	Protein purification	30
2.4.9.	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)	30
2.4.10.	Western blot analysis	31
2.4.11.	Liquid Chromatography - Mass Spectrometry (LC-MS)	32
2.4.12.	Biophysical characterization	33
2.4.13.	Biochemical characterization.....	34
3.	Chapter 3: Results	36
3.1.	<i>In silico</i> data	36
3.1.1.	BLASTP Homologues of PFF1010c	36
3.1.2.	Multiple sequence alignment.....	39
3.1.3.	PFF1010c phylogeny	41
3.1.4.	Protein domains of PFF1010c.....	43
3.1.5.	Predicted secondary and tertiary structures of PFF1010c	43
3.1.6.	Homo-oligomerization of PFF1010c	47
3.1.7.	Predicted PFF1010c cellular localization	49
3.1.8.	STRING tool predicted the interactome of PFF1010c	50
3.1.9.	Predicted posttranslational modification (PTM) sites	50
3.1.10.	Physicochemical properties.....	51

3.2.	Biochemical and biophysical assays.....	52
3.2.1.	pQE30/PFF1010c plasmid confirmation.....	52
3.2.2.	PFF1010c expression and solubility in <i>E. coli</i> XL1 Blue.....	53
3.2.3.	PFF1010c solubilization and purification.....	54
3.2.4.	LC-MS analysis of purified PFF1010c	56
3.2.5.	Biophysical characterization.....	57
3.2.6.	Biochemical characterization	60
4.	Chapter 4: Discussion and Conclusions	62
4.1.	Discussion	62
4.2.	Conclusions	66
5.	References.....	68
	Appendices.....	A
	Appendix A: Equipment, expendables, prepared solutions, and reagents.....	A
	Appendix B: Copyright attributions.....	D
	Appendix C: Bioinformatics supplementary data.....	E
	Appendix D: Experiments supplementary data	G
	Appendix D1. Plasmid confirmations of <i>pQE30</i> , <i>pQE30/PfHsp70-1</i> , and <i>pQE30/PfHsp40</i>	G
	Appendix D2. SDS-PAGE and Western blot for overexpression and purification of <i>pQE30/PfHsp70-1</i> , and <i>pQE30/PfHsp40</i>	I
	Appendix D3. SDS-PAGE for purification of PFF1010c.....	J

Appendix D4. BeStSel data analysis of PFF1010c CD spectra.....	K
--	---

List of Figures

Chapter 1

- Figure 1.1 The phylogenetic relationship among the 5 *Plasmodium* species responsible for causing malaria in humans
- Figure 1.2 Malaria parasite life cycle
- Figure 1.3 Gametocyte stages of *P. falciparum*
- Figure 1.4 PfHsp70 domain organization
- Figure 1.5 Functional cycle of PfHsp70-1 with PfHsp70-z and JDPs)
- Figure 1.6 PfHsp70-1 bound to ATP, a client protein and complexed with JDP, PfHop and PfHsp90
- Figure 1.7 Subfamilies of PfJDPs divided by the domain architecture and conservation

Chapter 3

- Figure 3.1 Multiple sequence alignment of PFF1010c with its BLASTP possible homologues
- Figure 3.2 Multiple sequence alignment of PFF1010c with its HHPRED possible homologues
- Figure 3.3 Phylogenetic trees of BLASTP(A) and HHPRED(B) protein hits
- Figure 3.4 The predicted secondary and tertiary structures of PFF1010c
- Figure 3.5 The tertiary structure of PFF1010c from Alphafold
- Figure 3.6 Five Predicted PFF1010c oligomer models, including the positive control
- Figure 3.7 STRING Predicted functional associates of PFF1010c
- Figure 3.8 Restriction endonuclease cleaving of pQE30/PFF1010c
- Figure 3.9 Expression and solubility assessments of PFF1010c
- Figure 3.10 Solubilization and purification of PFF1010c using 8 M urea
- Figure 3.11 SDS-PAGE gel cut out for LC-MS analysis and LC-MS results
- Figure 3.12 ANS extrinsic fluorescence emission of PFF1010c at different temperatures (°C)
- Figure 3.13 Luciferase refolding activity of PfHsp70-1, PfHsp40, and PFF1010c separately and in complex with each other

List of tables

Chapter 1

Table 1.1 Gametocytogenesis stages and their description

Table 1.2 HSP superfamily's function

Table 1.3 PfHsp70s localization and function.

Table 1.4 Localization

Chapter 2

Table 2.1 Plasmids and bacterial strains

Table 2.2 Restriction endonuclease reaction reagents and their volumes

Table 2.3 Reagent amounts used for preparing 12 % SDS-PAGE gel

Chapter 3

Table 3.1 PFF1010c protein hits from BLASTP (top 5)

Table 3.2 PFF1010c protein hits from HHPRED (top 5)

Table 3.3 PFF1010c PSIPRED predicted secondary structure content

Table 3.4 PFF1010c predicted secondary structure difference and mean values between two predictors PSIPRED and SOPMA

Table 3.5 Number of subunits, interface area and docking score for each of the five models of PFF1010c predicted oligomers and for controls

Table 3.6 Expasy's Protparam computed PFF1010c physicochemical properties

Table 3.7 LC-MS identified proteins

Table 3.8 PFF1010c CD spectroscopy estimated secondary structure content (%)

Table 3.9 The closest single CATH domains in a Protein Data Bank subset filtered for structures with 90 % sequence homology

List of appendices

Appendix A: Equipment, expendables, prepared solutions, and reagents

Appendix B: Copyright attributions

Appendix C: Bioinformatics supplementary data

Appendix D: Experiments supplementary data

List of symbols

Abbreviations of units	Symbol Interpretation
%	Percent
A ₂₆₀	Absorbance at 260 nanometres
A ₂₈₀	Absorbance at 280 nanometres
A ₅₉₅	Absorbance at 595 nanometres
A ₆₀₀	Absorbance at 600 nanometres
bp	Base pairs
Kbp	Kilobase pairs
kDa	Kilodaltons
μM	Micromolar
nM	Nanomolar
M	Molar
°C	Degrees Celsius
μl	Microlitres
ml	Millilitres
l	Litres
w/v	Weight per volume
v/v	volume per volume
μg	Micrograms
mg	Milligrams
g	Grams
α	Alpha
β	Beta

Outputs

Conference oral presentation

Thabo Rolffy Masoga, Thendo Pertunia Mudau, Hygon Mutavhatsindi, Tawanda Zininga, Addmore Shonhai. Characterization of PFF1010c, a type IV J domain protein of *Plasmodium falciparum*. The South African Society for Biochemistry and Molecular Biology (SASBMB) 28th congress, 7 - 10 July 2024, Protea Hotel Ranch Resort, Polokwane, South Africa.

Chapter 1

1. Chapter 1: Introduction

1.1. Malaria

Malaria persists as one of the most common and lethal maladies in the world, especially in Africa (Makam & Matsa, 2021; WHO, 2023). The WHO (World Health Organisation) World Malaria Report 2023 records 249 million cases and 608 000 deaths due to malaria by 2022 (WHO, 2023). Although the African region had mortality due to malaria decreased by 36 % from the year 2000 (839 720 deaths) to the year 2019 (534 201 deaths), the mortality increased by 11 % from 2019 to 2020 (602 020 deaths), owing to the impediments of ongoing malaria intervention services such as the COVID-19 pandemic (Hussein *et al.*, 2024; WHO, 2023).

In 2021, malaria cases increased by 2 million (to 247 million cases), followed by a further increase in 2022 (to 249 million cases) (WHO, 2023). However, it should be noted that malaria mortality decreased by 11 thousand deaths (1.7 %) in 2022. These statistics reflect the success of current malaria interventions. In South Africa, malaria is endemic and concentrated in Limpopo, Mpumalanga, and Kwazulu-Natal provinces (Noeth *et al.*, 2024). Besides South Africa being malaria endemic, the country has achieved over 40 % reduction in incidence in 2020 (Shin & Lee, 2021; WHO, 2023).

1.1.1. Malaria history

It is important to stress that malaria has existed from millennia BC, and the disease was acknowledged as early as the 1880s (Cox, 2010; Talapko *et al.*, 2019). The name of the ailment derives from old beliefs. “*Malaria*” is a Latin morpheme that translates to “*bad air*”. The “*mal-*” prefix translates to the word “*bad*”, and the suffix “*-aria*” translates to “*air*”. The “*bad air*” originated from the belief that the air of dead organic matter causes malaria, but others believed that malaria comes from water and other non-living elements of nature (Cox, 2010; Hempelmann & Krafts, 2013; O’Neill *et al.*, 2015). The belief, dubbed the miasma theory, was primarily supported until the advent of Louis Pasteur’s germ theory and Antoni van Leeuwenhoek’s discovery of microbes as causal agents of diseases (Cox, 2010). The biomedical characterization of malaria began in the 1880s when Alphonse Laveran isolated the pathogen from a patient who had died of malaria (Cox, 2010; Hempelmann & Krafts, 2013; Nye, 2002).

1.1.2. Agents of malaria: *Plasmodium* species

Malaria is caused by species of the *Plasmodium* genus (Sato, 2021). Five *Plasmodium* species are causative agents of malaria, namely, *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, and *P. knowlesi*. The *P. ovale* branches into *P. ovale curtisi*, and *P. ovale wallikeri*. Although most literature points out five malaria pathogens, the sixth species come from individually recognising the *P. ovale* subspecies, the *P. ovale curtisi* and *P. ovale wallikeri* subspecies. As such, other studies count six species responsible for causing the malaria disease (Ahmed & Cox-Singh, 2015; Calderaro *et al.*, 2013). Figure 1.1 below depicts the phylogenetic relationship between these *Plasmodium* species visualized using PhyloT version 2 and iTOL version 5 online tools (Letunic & Bork, 2021) (Appendix B, Table B1). The deadliest of the five *Plasmodium* species is *P. falciparum* (WHO, 2023). *P. falciparum* is unique from the other *Plasmodium* species as reflected by the phylogenetic tree diagram (Figure 1.1) (Escalante *et al.*, 2022).

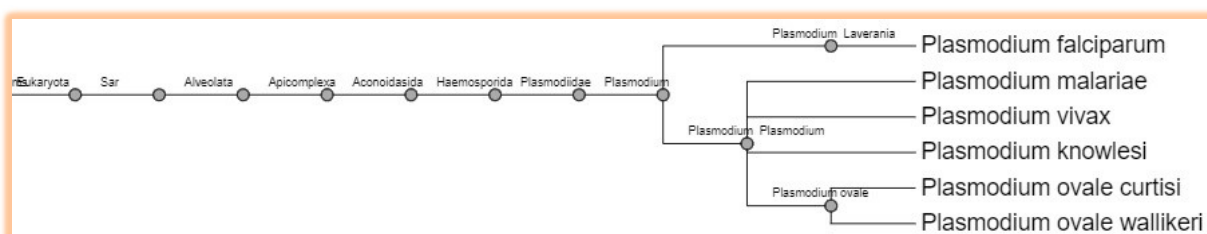


Figure 1.1 The phylogenetic relationship among the 5 *Plasmodium* species responsible for causing malaria in humans

P. falciparum is unique from the three closely related *P. malariae*, *P. vivax* and *P. knowlesi* which form a paraphyletic group. *P. ovale* delineated into *P. ovale wallikeri* and *P. ovale curtisi*. Adapted from (Letunic & Bork, 2021) (Appendix B, Table B1).

1.1.3. *Plasmodium* species life cycle

The life cycle of the malaria parasites is separated into two phases, the sexual phase occurring in the *Anopheles* mosquitoes and the asexual phase occurring in the human (Figure 1.2) (Geneva *et al.*, 2019; Nureye & Assefa, 2020; Shapiro *et al.*, 2017; Walter *et al.*, 2016). *Anopheles* mosquitoes are thus the definitive hosts whereas humans are the intermediate hosts of *Plasmodium* species. *Plasmodium* species go through five main morphological states throughout their life cycle (Cox, 2010; Milner, 2018). These morphological states represents the five stages of malaria life cycle, namely, the sporozoite stage, hepatic exoerythrocytic schizogony stage, erythrocytic schizogony stage, mosquito gametocyte ingestion stage, mosquito sporogony stage and sporozoite production stage (Cox, 2010).

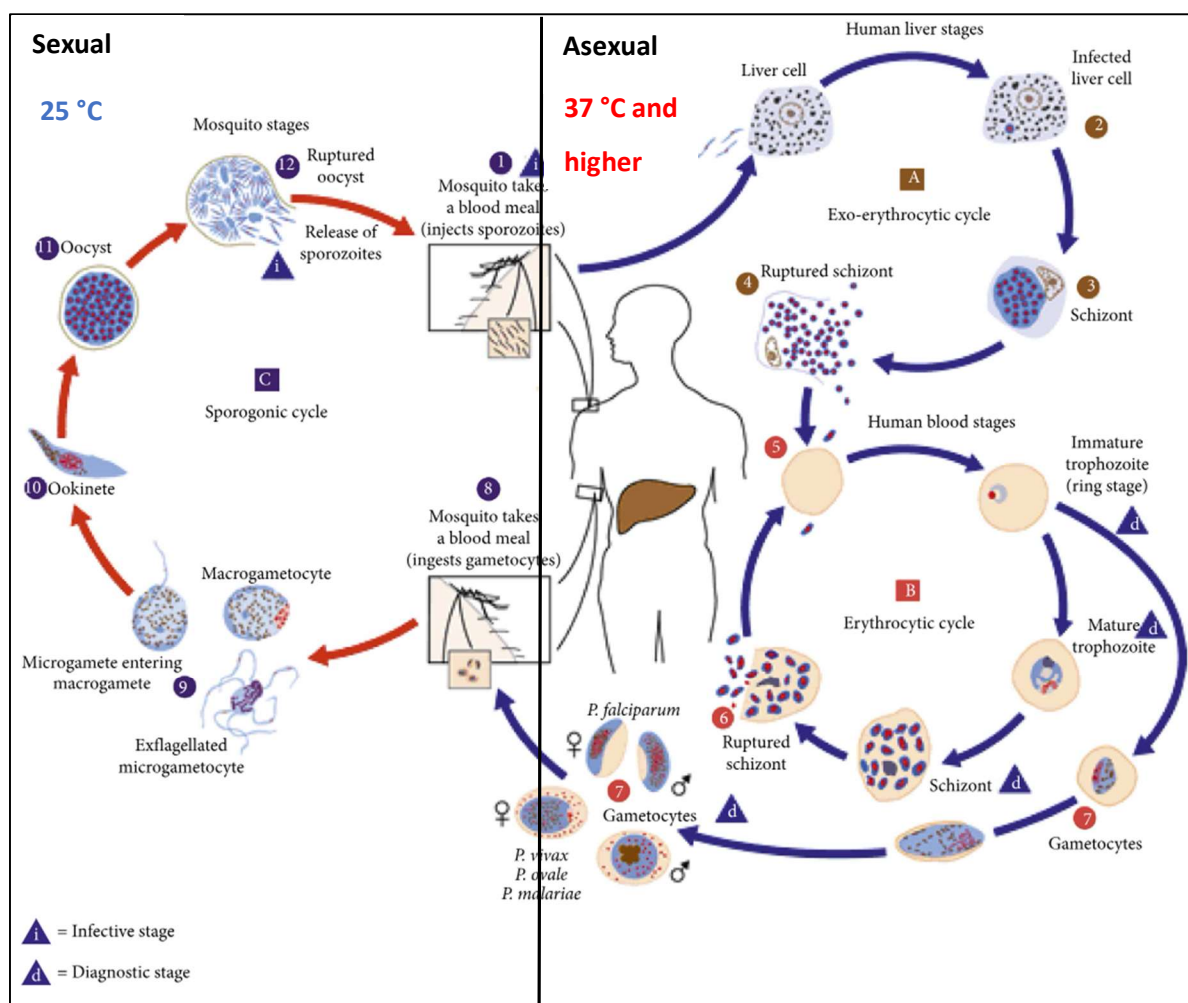


Figure 1.2. Malaria parasite life cycle

The sexual stage occurring at room temperature in the *Anopheles* mosquito is separated from the asexual stage occurring at around 37 °C and higher temperatures associated with febrile conditions in the human body. The temperature is denoted in blue font for normal temperatures and red for higher temperatures to reflect these conditions. (A) denotes the exo-erythrocytic cycle, (B) the erythrocytic cycle and (C) the sporogonic cycle. This Figure was adapted from (Nureye & Assefa, 2020) (Appendix B, Table B1).

1.1.3.1. The asexual phase of the malaria parasite in humans

The first stage (denoted "1"; Figure 1.2) is when the female *Anopheles* mosquitoes inject the sporozoites into the human host (Cox, 2010). The sporozoite stage is the morphological state of the parasite that shifts the parasite from the definitive to the intermediate hosts. Sporozoites are injected in large numbers, an average of one hundred and fifty sporozoites (Gomes *et al.*, 2016). This large number increases the probability of escaping the host immunity while in the host circulatory system (Gomes *et al.*, 2016). The second stage (denoted "2"; Figure 1.2) is the hepatic exoerythrocytic schizogony stage where the sporozoite has reached and infected the hepatocytes (Cox, 2010). The

sporozoites asexually replicate and differentiate into erythrocytes-infecting merozoites (Frischknecht & Matuschewski, 2017). The asexual reproduction and differentiation processes are collectively called schizogony. The merozoites are specialized to infect erythrocytes (Goh *et al.*, 2021). They employ the Merozoite Surface Protein-1 (MSP1) to recognize and infect the erythrocytes.

In the third and fourth stages, merozoites are released as schizonts from the initially sporozoite-infected hepatocytes which infect erythrocytes (Figure 1.2). This part is called the erythrocytic schizogony or cycle. Frischknecht and Matuschewski (2017) points out that there are taxa closely related to *Plasmodium* such as *Hepatocystis* and *Nicteria* that do not asexually replicate but rather undergo gametocytogenesis only (Frischknecht & Matuschewski, 2017). However, *Plasmodium* merozoites do undergo asexual reproduction followed by gametocytogenesis (Cox, 2010; Frischknecht & Matuschewski, 2017). The asexual progeny continues to infect more erythrocytes constituting the clinical malaria disease. The malaria parasite is adapted to survive the high temperatures assumed during febrile conditions associated with the disease (Walter *et al.*, 2016).

1.1.3.2. The sexual phase of the malaria parasite in *Anopheles* mosquito

The human phase of the malaria parasite life cycle (from “2” to “7” in Figure 1.2) terminates when the mosquitoes ingest the infected host blood meal containing the parasite gametes (Cox, 2010). The sexual reproduction phase (from “8” to “12” in Figure 1.2) of the *Plasmodium* species then ensues. The ingested microgametes (males) and macrogametes (females) differentiate from merozoites (Fujioka & Aikawa, 1999). The midgut environment of the *Anopheles* mosquitoes allows the macrogametes to germinate out of the ingested erythrocytes and develop into mature macrogametes with flagella (Fujioka & Aikawa, 1999). The resulting zygote undergoes sporogony, producing sporozoites that will initiate infection when the female *Anopheles* mosquitoes go for a blood meal (Cox, 2010; Fujioka & Aikawa, 1999). The sexual reproduction phase is less complex compared to the human phase. This highlights that the mosquitoes are only required for reproduction and the most important phase is the complexed and multistage human phase (Nureye & Assefa, 2020).

1.1.3.3. The *P. falciparum* gametocytogenesis

Gametocytogenesis of the *P. falciparum* life cycle occurs as an escape pathway from the erythrocytic cycle (stages denoted “5” and “6”; Figure 1.2) ensuring transmission (Dyer & Day, 2000). Figure 1.3 below depicts the gametocytogenesis stages (Carter & Miller, 1979). Gametocytogenesis initiates in trophozoites that have committed to sexual differentiation from the merozoites stage (Talman *et al.*,

2004). Sexual differentiation is activated by the transcriptional activator PfAP2-G (PF3D7_1222600) that binds to a large number of sites (hundreds) across the *P. falciparum* genome (Josling *et al.*, 2020). Gametocytogenesis takes place in the bone marrow and spleen and the mature gametocytes enter the bloodstream during their final stage of development (Usui & Williamson, 2021).

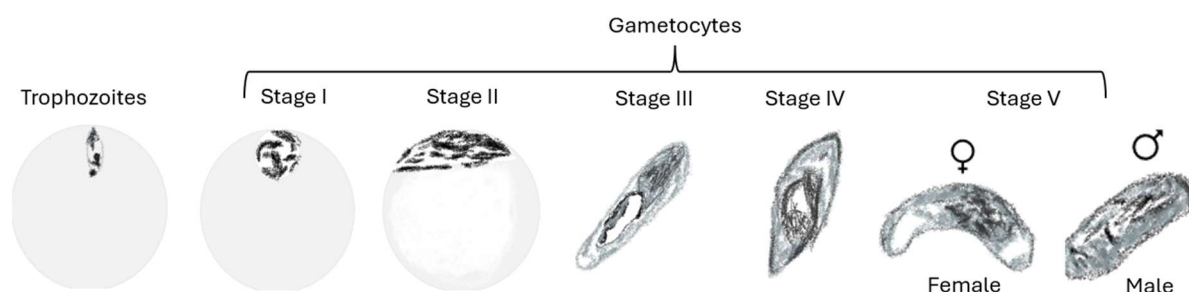
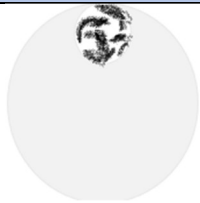
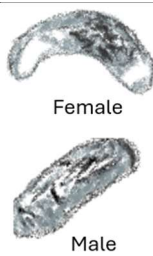


Figure 1.3. Gametocyte stages of *P. falciparum*

The parasite goes through 5 stages of gametogenesis from stage I to stage V. The first stage of gametocytogenesis initiates in trophozoites resulting in the mature gametocytes in peripheral blood. Adapted from (Carter & Miller, 1979) (Appendix B, Table B1).

Gametocytogenesis five stages (stage I–V) are based on the morphologies of the developing gametocyte (Table 1.1) (Carter & Miller, 1979; Sinden, 1983; Usui & Williamson, 2021). Stage I and II constitute the G_1 (first growth) phase, while the remaining stages are in the G_2 (second growth) phase of the cell cycle. The G_1 phase is when the cell physically grows and synthesizes essential enzymes while the G_2 phase is when the cell continues to grow and rearranges the cell organelles in preparation for cell division. Sinden (1983) added that the metabolic rate is high in stage IV and low in the matured stage V gametocytes (Sinden, 1983). The down regulation of gene expression and low metabolic rate is for the parasite to undergo latency and ensure less resources are utilized during this period.

Table 1.1. Gametocytogenesis stages and their description

Gametocyte stages (cell cycle phase)	Morphological state	Features description	Reference
Stage I (G ₁ Phase)		Growth and synthesis of essential enzymes. Parasite rounded and almost clear cytoplasm. Pigmentation initiating.	(Carter & Miller, 1979; Sinden, 1983)
Stage II (G ₁ Phase)		Growth and synthesis of essential enzymes. Assumes half-moon shape and occupies 50 % of the host erythrocyte. Pigments spread out.	(Carter & Miller, 1979; Sinden, 1983)
Stage III (G ₂ Phase)		Continued growth and rearrangement of organelles. Parasite extended with blunted ends. Haemoglobin lost; host erythrocyte no longer visible.	(Carter & Miller, 1979; Sinden, 1983)
Stage IV (G ₂ Phase)		Continued growth and rearrangement of organelles. Parasite cytoplasm and nucleus begin to stain. Cannot distinguish sex.	(Carter & Miller, 1979; Sinden, 1983)
Stage V (G ₂ Phase)	 Female Male	Continued growth and rearrangement of organelles. Fully morphological maturity. Centralization of chromatin.	(Carter & Miller, 1979; Sinden, 1983)

Several genes including the transcriptional activator *PfAP2-G* are overexpressed during the parasite trophozoite stage committed to gametocytogenesis (Prajapati *et al.*, 2020). The *PFF1010c* gene was not one of them. However, *PFF1010c* (*PF3D7_0620700*) is expressed at the gametocyte stage of the *P. falciparum* based on gene profiling studies (Le Roch *et al.*, 2003; Njunge *et al.*, 2013). Zhang *et al.* (2018) mutagenesis aided screen for essential genes found *PFF1010c* to be essential (Zhang *et al.*, 2018). In essence, *PFF1010c* must play a crucial role in the *P. falciparum* gametocytes.

1.2. Heat shock proteins

As *P. falciparum* switches between its endothermal definitive and ectothermal intermediate hosts, there is a temperature shift from 37 °C human normal temperature to 30 °C *Anopheles* normal temperatures (Beck-Johnson *et al.*, 2017; Geneva *et al.*, 2019; Shapiro *et al.*, 2017). To add, an infected human assumes febrile conditions in response to the infection. Body temperatures could rise above 40 °C during malaria fever episodes (Walter *et al.*, 2016). *E. coli* elicits a global response upon damage to its DNA called the SOS response (Zdravko & Darja-Žgur, 2021). The survival and development of the malaria parasite could depend on a system that mimics the bacterial SOS response, which could assist it to adapt to the temperature shifts. It is proposed that *P. falciparum* employs heat shock proteins in order to circumvent physiological stress (Daniyan *et al.*, 2019).

All life forms possess the heat shock response system (Alagar-Boopathy *et al.*, 2022; Hu *et al.*, 2022; Lang *et al.*, 2021; Lindquist & Craig, 1988). HSPs are responsible for cytoprotection and cell proteostasis (Alagar-Boopathy *et al.*, 2022; Lang *et al.*, 2021). The HSPs narrative dates back to 60 years ago when Ritossa (1962) published the *Drosophila*'s polytene chromosome “puffing” pattern in response to heat shock (Ritossa, 1962). The puffing pattern occurs when newly synthesized ribonucleic acid strands protrude off the chromosome in large quantities. The removal of heat shock reduced the puffing's of transcription. The stressors were not limited to heat. Chemical inducers were also reported (Ritossa, 1962). The stimulated transcription was of a set of proteins now called HSPs (Hu *et al.*, 2022).

HSPs are categorized based on their molecular weights and named accordingly (Hagymasi *et al.*, 2022). The superfamilies are Hsp100s, Hsp90s, Hsp70s, Hsp60s, Hsp40s and small Hsps (Dubrez *et al.*, 2020; Hagymasi *et al.*, 2022; Hu *et al.*, 2022). Table 1.2 summarizes HSPs function in most organisms. The Hsp40s have recently been noted as J domain proteins (JDPs) due to their characteristic J domain (Marszalek *et al.*, 2023; Piette *et al.*, 2021). Hsp70s carry out the refolding of unfolded proteins and JDPs assist them by stimulating their ATPase activity, a well characterized chaperone-co-chaperone relationship in multiple model organisms (Chand *et al.*, 2021; Faust *et al.*, 2020; Liu *et al.*, 2020). Hsp90s are involved in protein folding of newly synthesized proteins and refolding of unfolded proteins (Hoter *et al.*, 2018; Hu *et al.*, 2022). Hsp60s are implicated in protein assembly assisted by sHsps Hsp100s are mostly disaggregases (Kędzińska-Mieszkowska & Zolkiewski, 2021). Hsp100s also promote proteolysis of proteins marked for degradation (Lee *et al.*, 2022).

Table 1.2. HSP superfamily's functions

Heat shock protein superfamily (size in kDa)	Function	References
Small Hsps (~<40)	Prevent protein aggregation	(Haslbeck, 2002; Sugino <i>et al.</i> , 2009)
JDP (~40-50)	Binds denatured proteins and present to Hsp70, stimulate ATPase activity of Hsp70	(Botha <i>et al.</i> , 2007; Dutta, Pesce, <i>et al.</i> , 2021; Kampinga <i>et al.</i> , 2019; Njunge <i>et al.</i> , 2013)
Hsp60 (~60)	Protein folding and assembly	(Rodriguez, 2022; Tang <i>et al.</i> , 2022)
Hsp70 (~70)	Protein folding, refolding, trafficking, and degradation	(Chand <i>et al.</i> , 2021; Faust <i>et al.</i> , 2020; Liu <i>et al.</i> , 2020)
Hsp90 (~90)	Protein folding and maturation, prevent protein aggregation	(Chiosis <i>et al.</i> , 2023; Noddings <i>et al.</i> , 2022)
Hsp100 (~100)	Disaggregases, proteolysis promotion and regulation of transcription.	(Kędzierska-Mieszkowska & Zolkiewski, 2021; Lee <i>et al.</i> , 2022)

1.2.1. Malaria heat shock proteins

HSPs of *P. falciparum* have gained research interest considering their role in imparting stress tolerance to this deadly parasite (Makhoba *et al.*, 2023; Shonhai & Blatch, 2021; Tintó-Font & Cortés, 2022). The HSP superfamilies in *P. falciparum* are assigned the *Pf* abbreviation prefix like PfHsp40, PfHsp90 etc. The notorious *P. falciparum* survives the febrile human body and drug induced stress through activation of its stress response mechanisms, including the expression of its expanded HSP machinery (Zhang *et al.*, 2021). *P. falciparum* HSP arsenal includes all the superfamilies of HSPs, namely the PfHsp70s, PfHsp90s, PfHsp60s, PfHsp40s/PfJDPs, PfsHsps and PfHsp100s (Acharya *et al.*, 2007; Chakafana *et al.*, 2021; Kampinga *et al.*, 2019; Shonhai, 2014). The *P. falciparum* genome expresses 92 HSPs in total, taking up 2 % of the parasite genome (Acharya *et al.*, 2007).

1.2.2. *P. falciparum* heat shock protein 70

Six members of the Hsp70 superfamily are expressed in *P. falciparum* (Nyakundi & Boshoff, 2021; Winer *et al.*, 2020). The protein products are, namely, PfHsp70-1 (PF3D7_0818900), PfHsp70-2 (PF3D7_0917900), PfHsp70-3 (PF3D7_1134000), PfHsp70-x (PF3D7_0831700), PfHsp70-y (PF3D7_1344200), and PfHsp70-z (PF3D7_0708800) (Aurrecochea *et al.*, 2009). PfHsp70-x, PfHsp70-y, and PfHsp70-z constitute a subclass of PfHsp70s named PfHsp110s, owing to their molecular weight around 110 kDa (Aurrecochea *et al.*, 2009; Bracher & Verghese, 2022; Chakafana & Shonhai, 2021). The canonical subclass of PfHsp70s, weighing around 70 kDa, are PfHsp70-1, PfHsp70-2, and PfHsp70-

3. PfHsp70s have an ATPase domain at their N termini and a substrate binding domain at their C-termini as illustrated below (Figure 1.4) (Zininga *et al.*, 2016). The two domains are joined by a conserved linker allowing allosteric interaction between the domains. PfHsp70s assemble and fold newly formed proteins, refold improperly folded proteins, and disassembles protein aggregates (Shonhai & Blatch, 2014). Some PfHsp70s participate in the signal transduction pathway, DNA replication, and translocation of secretory and organelle proteins (Shonhai & Blatch, 2014; Shonhai *et al.*, 2007).



Figure 1.4: PfHsp70 domain organization

N represents the N-terminus of the protein, NBD represents the nucleotide binding domain, SBD represents the substrate binding domain, and C represents the C-terminus of the protein.

1.2.2.1. Localization and functions of heat shock proteins 70

Table 1.3 summarizes the subcellular localizations and functions of PfHsp70s. The individual PfHsp70s specialize in protein folding, protein trafficking, exchanging nucleotides for other PfHsp70s, suppress aggregation, and disaggregate aggregated proteins (Table 1.3). While PfHsp70-1 localizes in the cytosol, it also moves into the nucleus in response to heat shock (Acharya *et al.*, 2007; Gitau *et al.*, 2012). The endoplasmic reticulum resident PfHsp70-2 is involved in the export of proteins out of the parasite cell and into the infected erythrocyte (Cortés *et al.*, 2020). PfHsp70-3 localizes in the mitochondria and suppresses aggregation (Barth *et al.*, 2022; Nyakundi & Boshoff, 2021). PfHsp70-x is exported from the malaria parasite cell into the infected erythrocyte (Dutta, Singh, *et al.*, 2021; Ramatsui *et al.*, 2023). The endoplasmic reticulum resident PfHsp70-y and cytosol resident PfHsp70-z function as nucleotide exchanging factors of PfHsp70-2 and PfHsp70-1, respectively (Shonhai *et al.*, 2007; Zininga *et al.*, 2016). Nucleotide exchange factors replace the ADP co-factor bound to Hsp70s by another ATP. The localization of PfHsp70s allows the chaperone to efficiently specialize in different organelles (Shonhai *et al.*, 2007). However, the above mentioned PfHsp70s have not been shown to interact with the putative PFF1010c protein (Njunge *et al.*, 2013). PFF1010c could be interacting with either of the PfHsp70 members, regulating their function.

Table 1.3. PfHsp70s localization and function

Hsp70s of <i>P. falciparum</i> (accession number)	Molecular weight in kDa	Localization	Function	References
PfHsp70-1 (PF3D7_0818900)	74	Cytosol, nucleus	Protein folding and translocation	(Acharya <i>et al.</i> , 2007; Anas <i>et al.</i> , 2020; Gitau <i>et al.</i> , 2012; Shonhai, 2014)
PfHsp70-2 (PF3D7_0917900)	73	Endoplasmic reticulum	Protein exportation from <i>P. falciparum</i> to the infected erythrocyte; protein unfolding inhibition	(Cortés <i>et al.</i> , 2020; Ramya <i>et al.</i> , 2006)
PfHsp70-3 (PF3D7_1134000)	73	Mitochondria	Protein aggregation suppressor	(Barth <i>et al.</i> , 2022; Nyakundi & Boshoff, 2021; Nyakundi <i>et al.</i> , 2016)
PfHsp70-x (PF3D7_0831700)	76	The parasitophorous vacuole and the cytoplasm of the infected erythrocyte	Protein trafficking, aggregation suppressor	(Charnaud <i>et al.</i> , 2017; Daniyan <i>et al.</i> , 2016; Grover <i>et al.</i> , 2013)
PfHsp70-y (PF3D7_1344200)	108	Cytosol, nucleus, endoplasmic reticulum	Exchange nucleotides for PfHsp70-2	(Kudyba <i>et al.</i> , 2019)
PfHsp70-z (PF3D7_0708800)	100	Cytosol, nucleus, chloroplast, endoplasmic reticulum	Exchange nucleotides for PfHsp70-1; suppress aggregation; Stabilization of asparagine rich proteins	(Muralidharan <i>et al.</i> , 2012; Shonhai, 2014; Zininga <i>et al.</i> , 2016)

1.2.2.2. Co-chaperones of *P. falciparum* heat shock protein 70-1

PfHsp70s are known to complex with other HSPs and form a multicomplex machinery for organization, efficient coordination, and synergy (Basquin *et al.*, 2016; Lubkowska *et al.*, 2021; Shiber *et al.*, 2018; Wickramaratne *et al.*, 2023). The functional cycle of PfHsp70-1 regulated by JDPs and nucleotides exchanged by nucleotide exchange factor PfHsp70-z is illustrated below (Figure 1.5) (Njunge *et al.*, 2013; Njunge *et al.*, 2015; Shonhai *et al.*, 2007). The J domain proteins regulate the PfHsp70-1 activity (Blatch, 2022; Zhang *et al.*, 2018). PfHsp70-z is a nucleotide exchange factor for PfHsp70-1 as mentioned above (Table 1.3).

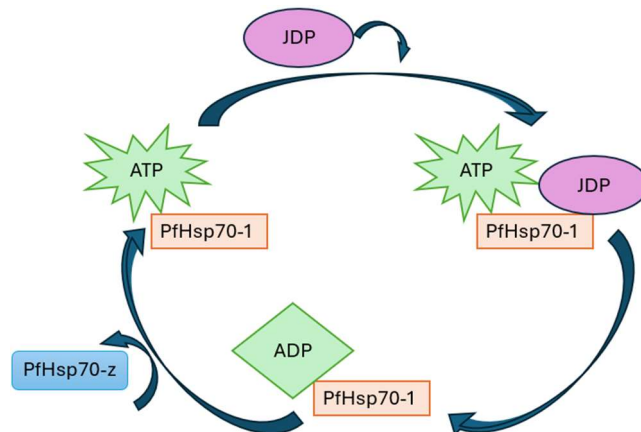


Figure 1.5. Functional cycle of Hsp70-1 as regulated by PfHsp70-z and JDP

Hsp70-1 ATPase activity is regulated by JDPs while PfHsp70-z serves as a nucleotide exchange factor.

In addition to JDPs and the nucleotide exchange factors, PfHsp70-1 complexes with PfHsp90, aided by the *P. falciparum* Hsp70/Hsp90 organizing protein (PfHop, PF3D7_1434300) (Gitau *et al.*, 2012). PfHsp70-1 binds unfolded proteins and passes the native-like substrate to a PfHsp90 member for further folding (Wegele *et al.*, 2006). In essence, PfHop regulates the PfHsp70 – PfHsp90 substrate transfer (Gitau *et al.*, 2012).

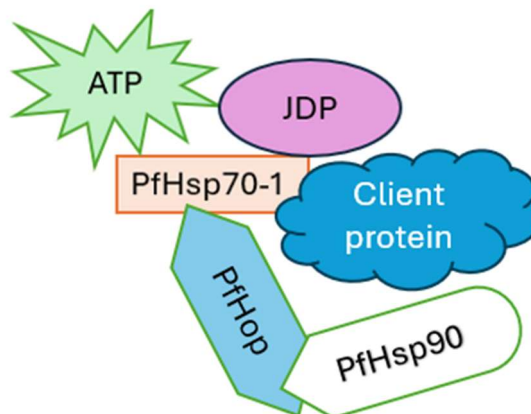


Figure 1.6. PfHsp70-1 bound to ATP, a client protein and complexed with JDP, PfHop and PfHsp90

The JDP, PfHop and PfHsp90 proteins come together to form a complex structure that refolds an unfolded client protein.

1.2.3. *P. falciparum* heat shock protein 100

Hsp100s are caseinolytic proteases (Clps) and chaperones (Bakkouri *et al.*, 2013). They are capable of the dual functions, folding and protein breakdown. The proteins dimerize to execute their function and are ATPases (Bakkouri *et al.*, 2010). *P. falciparum* expresses four isoforms of Hsp100s, namely, PfClpB1 (PF3D7_0816600), PfClpB2 (PF3D7_1116800), PfClpC (PF3D7_1406600) and PfClpM (PF3D7_API03600) (Bakkouri *et al.*, 2013; Ngansop *et al.*, 2013). PfClpB2 appears in the parasitophorous vacuole and found to function in protein export from the parasitophorous vacuole to the cytoplasm of the infected erythrocyte (Bakkouri *et al.*, 2010). PfClpB1, PfClpM and PfClpC are apicoplast residents (Bakkouri *et al.*, 2013). The apicoplast is the well-known plastid of the Apicomplexa phylum responsible for lipid anabolism (Gleeson, 2000). Bakkouri *et al.* (2010) found that the recombinant PfClpB1 facilitated recovery of a firefly luciferase protein activity in *E. coli* (Bakkouri *et al.*, 2010). This finding suggests that PfClpB1 facilitates the survival of *P. falciparum* during stressful conditions, by reactivating deactivated proteins. The study also discovered that PfClpB1 does not have a human homolog, making it a desirable antimalarial target.

1.2.4. *P. falciparum* heat shock protein 90

Hsp90s are ATPases and are capable of forming higher order species such as dimers (Lepvrier *et al.*, 2019; Silva *et al.*, 2020). Hsp90s are also highly conserved and expressed in all organisms including plants and humans (Silva *et al.*, 2020; Stofberg *et al.*, 2021; Zainordin *et al.*, 2021). *P. falciparum* expresses four Hsp90 isoforms and each isoform localizes to an organelle in the parasite (Stofberg *et al.*, 2021). Isoform PfGrp94 (PF3D7_1222300) is found in the endoplasmic reticulum, PfHsp90_A (PF3D7_1443900) localizes in the apicoplast, PfHsp90_M (PF3D7_1118200) in the mitochondria, while the PfHsp90_C (PF3D7_0708400) isoform occurs in the cytosol. A property that makes PfHsp90 members stand out from the other heat shock proteins is their dimer state ability to assemble with other PfHsp90 dimers, forming hexamer superstructures dubbed “cosy nests” (Lepvrier *et al.*, 2019). PfHsp90 cosy nests exhibit a chaperone activity, although the mechanism is not fully understood. PfHsp90 assembles without nucleotide binding but the presence of metal oxyanions or divalent cations and heat shock (Lepvrier *et al.*, 2019). PfHsp90 is a target for malaria treatment (Stofberg *et al.*, 2021).

1.2.5. *P. falciparum* heat shock protein 60

Two PfHsp60s are known to be expressed by *P. falciparum*. The PfHsp60 (PF3D7_1015600) and PfCpn60 (PFH3D7_1232100) are parasite resident proteins, with the initial localizing to the mitochondria and the latter to the apicoplast (Liu & Houry, 2014). PfHsp60s are classified as chaperonins, they fold a nascent protein while inside V-shaped tetradecameric superstructure formed by PfHsp60 subunits (Liu & Houry, 2014). The chaperonins form a complex with *P. falciparum* small Hsps which assist chaperonins in their refolding function (Liu & Houry, 2014).

1.2.6. *P. falciparum* small heat shock proteins

The *P. falciparum* small Hsps family (PfsHsps) represents one of the most distinct set of molecular chaperones (Shonhai, 2010). SHsps are small heat shock proteins weighing around 15 to 30 kDa (Buchner *et al.*, 1998). *P. falciparum* sHsps isoforms are 6, namely, PfsHsp20-a (PF3D7_1304500), PfsHsp20-b (PF3D7_0816500), PfsHsp20-c (PF3D7_1211100), PfsHsp20-d (PF3D7_0708500), PfsHsp20-e (PF3D7_0808800), and PfsHsp20-f (PF3D7_10198100) (Timothy & Zininga, 2023). The PfsHsp20-a, PfsHsp20-b, and PfsHsp20-c localize to the mitochondria, the PfsHsp20-d and PfsHsp20-f to the nucleus and PfsHsp20-e to the apicoplast of the *P. falciparum* (Timothy & Zininga, 2023). The PfsHsp20s are holdases; they 'hold' an unfolded protein and prevent aggregation (Timothy & Zininga, 2023). Generally, the absence of sHsps leads to protein aggregation, which would ultimately lead to neurodegenerative diseases and cancer in mammals (Carra *et al.*, 2019). However, sHsps' ubiquitous distribution in the parasite cell and involvement in disaggregation makes them ideal drug candidates (Carra *et al.*, 2019; Timothy & Zininga, 2023; Wyciskiewicz *et al.*, 2019).

1.2.7. *P. falciparum* J domain proteins

Among the malaria parasites, *P. falciparum* expresses the highest complement of JDPs (Njunge *et al.*, 2013). The JDPs were recently noted as Hsp40s until Kampinga *et al.* (2019) coined them J domain proteins (JDPs) based on this protein domain present in *E. coli* DnaJ (Botha *et al.*, 2007; Kampinga *et al.*, 2019). The PfJDPs are divided into four classes based on their conservation of the DnaJ domains (Figure 1.7) (Cheetham & Caplan, 1998). *E. coli* DnaJ possesses four domains: the J domain, the Glycine or Phenylalanine -rich domain, the zinc finger domain, and the C-terminal domain (Dutta *et al.*, 2021). PfJDP type I contains all the domains of their ancestor protein DnaJ, PfJDP type II contains all the domains but the zinc finger domain, and type III PfJDPs only contain the J domain with the highly conserved HPD tripeptide within the J domain (Cheetham & Caplan, 1998; Kampinga *et al.*, 2019).

While Cheetham & Caplan (1998) demonstrated the domain conservation and the classification of PfJDPs into type I, type II, and type III relative to DnaJ domains, Botha *et al.* (2007) added the fourth subfamily, PfJDP type IV, after the realisation that the proteins conserve a J domain with a different motif substituting the canonical HPD (Botha *et al.*, 2007; Cheetham & Caplan, 1998).

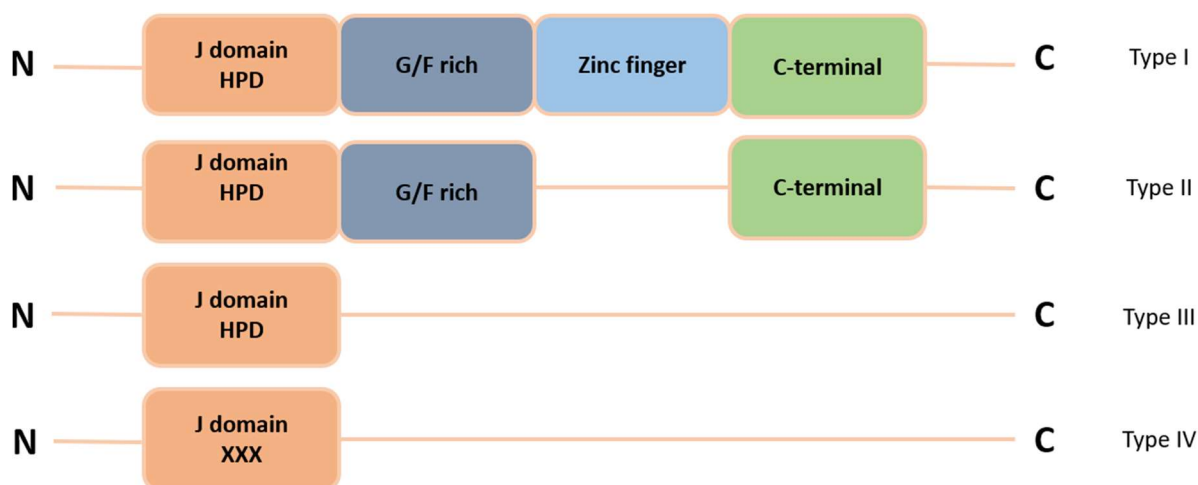


Figure 1.7. Subfamilies of PfJDPs divided by the domain architecture and conservation

The rectangles are domain sections along the polypeptides labelled from the N-terminal (N) to the C-terminal (C). The J domain is the orange rectangle with the tripeptide motif within it. The Glycine or phenylalanine rich domain is reflected by the blue-grey rectangle, the zinc finger by the blue rectangle and the C-terminal domain by the green rectangle. XXX symbolizes varied residues in place of the HPD motif. The type IV J domain does not possess the canonical HPD motif.

P. falciparum expresses 2 type I JDPs, 9 type II JDPs, 25 type III JDPs, and 13 type IV JDPs (Table 1.4) (Botha *et al.*, 2007; Njunge *et al.*, 2013). Table 1.4 summarises the 49 PfJDPs localization, essentiality and function. Type III subclass has the largest PfJDP complement. Out of the 49 PfJDPs, 18 are exported from the parasite cell and into the host erythrocyte. Most of these exported proteins conserve the export motif PEXEL and participate in the infected erythrocyte remodelling (Dutta *et al.*, 2021; Gabriela *et al.*, 2022). The PEXEL motif is displayed by proteins exported through the *Plasmodium* translocon of the exported (PTE) proteins channel (Gabriela *et al.*, 2022). Proteins lacking the PEXEL motif are exported through an alternative, poorly characterised process (Kreutzfeld *et al.*, 2021). The two type 1 PfJDP, the PfHsp40 (PF3D7_1437900) and Pfj1 (Table 1.4), interact with the cytosolic PfHsp70-1 (Dutta *et al.*, 2021). Type IV PfJDPs are 13, with 11 exported and two parasite residents. PFF1010c is a type IV JDP that is yet to be characterised (Botha *et al.*, 2007; Njunge *et al.*, 2013; Zhang *et al.*, 2018).

Table 1.4. Localization and functional features of *P. falciparum* JDPs

Common name/ other accession number	PlasmoDB accession number	PfJDP subclass	Localization	Essential / not essential	Function
Pfj1	PF3D7_0409400	Type I	Parasite resident	Not essential	Housekeeping
PfHsp40	PF3D7_1437900	Type I	Parasite resident	Essential	Housekeeping
PFA0660w	PF3D7_0113700	Type II	Host erythrocyte	Not essential	pathogenic
PFE0055c	PF3D7_0501100	Type II	Host erythrocyte	Essential	Pathogenic
KAHsp40	F3D7_0201800	Type II	Host erythrocyte	Not essential	pathogenic
PfSis1	PF3D7_0213100	Type II	Parasite resident	Essential	Housekeeping
PFF1415c	PF3D7_0629200	Type II	Parasite resident	Not essential	Housekeeping
PfJ4	PF3D7_1211400	Type II	Parasite resident	Essential	Housekeeping
PF14_0137	PF3D7_1413900	Type II	Parasite resident	Essential	Uncharacterized
Pfj2	PF3D7_1108700	Type II	Parasite resident	Essential	Housekeeping
MAL13P1.277	PF3D7_1356700	Type II	Parasite resident	Essential	Uncharacterized
PFB0920w	PF3D7_0220100	Type III	Host erythrocyte	Essential	Pathogenic
PfJjj3	PF3D7_0502800	Type III	Parasite resident	Not essential	Housekeeping
PFE1170w	PF3D7_0523400	Type III	Parasite resident	Essential	Uncharacterized
PFI0935w	PF3D7_0919100	Type III	Parasite resident	Essential	Uncharacterized
PFI0855w	PF3D7_0917500	Type III	Parasite resident	Essential	Uncharacterized
PfJac1	PF3D7_0920100	Type III	Parasite resident	Essential	Housekeeping
PFL0055c	PF3D7_1201100	Type III	Host erythrocyte	Not essential	Uncharacterized
PfZuo1	PF3D7_1216900	Type III	Parasite resident	Essential	Housekeeping
PfPam18/Tim14	PF3D7_0724400	Type III	Parasite resident	Essential	Housekeeping
PF08_0032	PF3D7_0823800	Type III	Parasite resident	Essential	Housekeeping
PF08_0115	PF3D7_0806500	Type III	Parasite resident	Essential	Uncharacterized
PF10_0032	PF3D7_1002800	Type III	Parasite resident	Not essential	Uncharacterized
PfJjj1	PF3D7_1005600	Type III	Parasite resident	Not essential	Housekeeping
PfJ3	PF3D7_1038800	Type III	Host erythrocyte	Not essential	Uncharacterized
PF11_0273	PF3D7_1126300	Type III	Parasite resident	Essential	Uncharacterized
PF11_0380	PF3D7_1136800	Type III	Parasite resident	Essential	Uncharacterized
PF11_0433	PF3D7_1142100	Type III	Parasite resident	Not essential	Uncharacterized
PF11_0513	PF3D7_1149600	Type III	Host erythrocyte	Not essential	Uncharacterized
PF13_0036	PF3D7_1307200	Type III	Parasite resident	Essential	Uncharacterized
PfSec63	PF3D7_1318800	Type III	Parasite resident	Essential	Housekeeping and pathogenic
PF14_0111	PF3D7_1411300	Type III	Parasite resident	Essential	Uncharacterized
PF14_0213	PF3D7_1422300	Type III	Parasite resident	Essential	Uncharacterized
PF14_0700	PF3D7_1473200	Type III	Parasite resident	Essential	Uncharacterized
MAL13P1.162	PF3D7_1330300	Type III	Parasite resident	Not essential	Uncharacterized

MAL8P1.204	PF3D7_0831200	Type III	Parasite resident	Not essential	Uncharacterized
RESA	PF3D7_0102200	Type IV	Host erythrocyte	Not essential	Pathogenic
PFA0675w	PF3D7_0114000	Type IV	Host erythrocyte	Not essential	Uncharacterized
PFB0085c	PF3D7_0201700	Type IV	Host erythrocyte	Essential	Uncharacterized
PFB0925w	PF3D7_0220400	Type IV	Host erythrocyte	Not essential	Uncharacterized
MESA/PfEMP2	PF3D7_0500800	Type IV	Host erythrocyte	Not essential	Pathogenic
PFF1010c	PF3D7_0620700	Type IV	Parasite resident	Possibly essential	Uncharacterized
PfGECO	PF3D7_1253000	Type IV	Host erythrocyte	Not essential	Pathogenic
PF10_0381	PF3D7_1039100	Type IV	Host erythrocyte	Not essential	Pathogenic
PF11_0034	PF3D7_1102200	Type IV	Host erythrocyte	Not essential	Uncharacterized
PF11_0443	PF3D7_1143200	Type IV	Parasite resident	Not essential	Uncharacterized
RESA3	PF3D7_1149200	Type IV	Host erythrocyte	Not essential	Pathogenic
RESA2	PF3D7_1149500	Type IV	Host erythrocyte	Not essential	Pathogenic
PF14_0013	PF3D7_1401100	Type IV	Host erythrocyte	Essential	Uncharacterized
References	(Botha <i>et al.</i> , 2007; Dutta, Pesce, <i>et al.</i> , 2021; Njunge <i>et al.</i> , 2013; Zhang <i>et al.</i> , 2018)				

PFF1010c (Table 1.4, highlighted in red) is an essential type IV PfJDP lacking human homologues, rendering it an ideal drug targets (Pesce *et al.*, 2010). The HPD motif varied in type IV PfJDP has been shown to directly interact with the Hsp70 ATP-binding domain to regulate the Hsp70 ATPase activity (Dutta *et al.*, 2021; Kityk *et al.*, 2018; Tsai & Douglas, 1996). Actually, HPD mutants of *Mycobacterium tuberculosis* DnaJ1 and DnaJ2 inhibited the *M. tuberculosis* DnaK ATPase activity (Lupoli *et al.*, 2016). However, the binding activity of *M. tuberculosis* DnaJ1 and DnaJ2 HPD mutants to *M. tuberculosis* DnaK was retained. PFF1010c could be binding to and inhibiting *P. falciparum* Hsp70 homologues ATPase activity.

1.3. Interventions against malaria

1.3.1. Diagnosis

There are three main ways to diagnose malaria (Mace *et al.*, 2018; WHO, 2023). The first is blood film microscopy, the second is the Polymerase Chain Reaction (PCR), and the third is the Rapid Diagnostic Test (RDT). Blood film microscopy visualizes the patients' erythrocytes and the appearance of parasite particles in the erythrocytes. Malaria parasites affect the infected erythrocytes' shape and form ring-like structures within the infected erythrocyte (Depond *et al.*, 2020). The PCR method detects the presence of malaria parasites' signature nucleic acids (Sazed *et al.*, 2021). The RDT test detects antigens of the parasite. Danwang *et al.* (2021) review hailed the sensitivity of malaria RDT (Danwang *et al.*, 2021). These are the diagnosis that have led to the determination of the recent prevalence.

1.3.2. Artemisinin combination therapy

A protective efficacy of 86 % against *P. falciparum* using artemisinin derivatives and amodiaquine combination therapy was deduced after a controlled randomized trial (Thera *et al.*, 2018). Artemisinin-based combination therapy remains the gold standard for malaria treatment and prevention (Cohee *et al.*, 2020; Haldar *et al.*, 2018). Artemisinin and its derivatives target the ring stage of the trophozoites and the early gametocyte stage, alkylating the parasite proteome and ultimately subjecting the parasite to cell death (Suresh & Haldar, 2018). The Artemisinin molecule contains an endoperoxide functional group that breaks down and produces reactive oxygen radicals (Ismail *et al.*, 2016). The oxygen radicals are involved in the alkylation of the parasite proteome. However, *P. falciparum* has developed resistance against artemisinin (Suresh & Haldar, 2018). The gene responsible for artemisinin resistance is called *k13* or *pfkelch13*. The *k13* product is implicated in the upregulation of pathways regulating parasite oxidative stress and proteostasis (Suresh & Haldar, 2018).

1.3.3. The Mosquirix vaccine

By 2020, at least 17 vaccine candidates were in preclinical and clinical trials (Duffy & Gorres, 2020). Most of the vaccine candidates were subunit vaccines, with a few being attenuated whole sporozoite or attenuated whole blood stage. While all the vaccine candidates were at preclinical, phase 1 or phase 2 of clinical trials, RTS,S/AS01 (also known as Mosquirix) was already in phase 4, standing out as the most promising vaccine candidate. The malaria subunit vaccine RTS,S/AS01, aimed to protect children under five years in malaria endemic regions by aiding their underdeveloped immune system (Laurens, 2020). RTS,S/AS01 phase 4 clinical trials did not produce concerning signals (Hammershaimb & Berry, 2024). WHO recommended RTS,S/AS01 for use in children over 5 months old and residing in moderate to high malaria transmission regions (WHO, 2023).

1.4. Study rationale

Despite the current interventions, malaria continues to be a deadly disease. *P. falciparum*, the deadliest malaria parasite, overcomes various forms of stress, including the host's immune response. One way *P. falciparum* survives adverse conditions is by activating the heat shock response. The HSP arsenal provides cytoprotection and maintains proteostasis. JDPs are one of the prominent families of HSP in *P. falciparum*. Notably, *P. falciparum* expresses the most diverse and the largest contingent family of HSPs compared to other *Plasmodium* species. Most JDPs are essential. Canonical JDPs are

known to harness the Hsp70's refolding activity by partnering and regulating Hsp70 ATPase activity. However, *P. falciparum* JDPs are poorly characterized, with most archived as hypothetical proteins. PFF1010c is a JDP, one of the essential and poorly characterized JDP proteins. The annotated genome data of *P. falciparum* suggest that PFF1010c is expressed during the gametocyte stage of *P. falciparum* and is essential for the parasite's survival. For these reasons, PFF1010c is proposed to perform a significant proteostatic function during the gametocyte stage. As such, this study characterized the recombinant form of PFF1010c using *in silico*, biochemical and biophysical approaches towards elucidating the structure–function features of this medically important protein.

1.5. Hypothesis

PFF1010c is a co-chaperone whose structural and functional features make it amenable to its essential role in the development of the malaria parasite.

1.6. Aim

To characterize PFF1010c structural and functional features.

1.7. Objectives

1.7.1. In silico analyses to establish the structural features of PFF1010c

Approaches: Structural, functional, localization and interactors prediction tools (BLASTP, PSIPRED, CLUSTALW, MEGA11, SOPMA, CDD/SPARCLE ESPRIPT, AlphaFold, GalaxyHomomer, ProtParam, PlasmoDB, and STRING).

1.7.2. Expression and purification of recombinant PFF1010c

Approaches: (a) Confirmation of the integrity of the plasmid construct expressing PFF1010c through restriction endonuclease digest and agarose gel electrophoresis; (b) recombinant expression of PFF1010c using *E. coli* XL1 Blue; (c) purification using 8 M urea and immobilized metal (Nickel) affinity chromatography (IMAC) and (d) protein confirmation using sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE), Western blotting and liquid chromatography-mass spectrometry (LC-MS).

1.7.3. Biophysical and biochemical characterization

Approaches: (a) PFF1010c secondary structure analysis through circular dichroism (CD) spectroscopy, and (b) exploring the interaction of PFF1010c with PfHsp70-1 and investigating

the role of PFF1010c on the refolding activity of PfHsp70-1 using luciferase as a model substrate.

Chapter 2

2. Chapter 2: Materials and Methods

This chapter outlines the materials used in this study and the methods used to achieve the objectives presented in chapter 1. Chapter 2 begins by reporting the ethical adherence of this study, continues by cataloguing the materials used and finalises with methods followed to achieve the objectives.

2.1. Ethics concern

It is important in scientific research to fulfil ethical requirements (Freedman, 1987). Ethical clearance was obtained from the Faculty Research Ethics Committee of the Faculty of Science, Engineering and Agriculture, University of Venda, Thohoyandou, South Africa. Ethical clearance no: FSEA/22/MBY/36/1707. The approval period is from June 2023 to June 2025. The risk of the study is categorized under Category 1 - minimal risk to humans, animals, or the environment.

2.2. Materials

2.2.1. Equipment

The equipment used to conduct this study is housed at the University of Venda Biochemistry and Microbiology department laboratories unless stated otherwise. Table A1 in appendix A summarises the equipment with their model names and manufacturers. Table A2 in appendix A shows the expendable items that aided in carrying out this project.

2.2.2. General reagents and chemicals

The reagents were purchased from Merck (Darmstadt, Germany), Whitehead Scientific (Cape Town, South Africa), B&M Scientific (Cape Town, South Africa), Thermo Fisher Scientific (Massachusetts, USA) and VWR Life Science (Pennsylvania, USA). Table A3 in appendix A lists the reagents and their trading companies.

2.2.3. Stock solutions, buffers, and growth media

Solutions and buffers that could not with-stand autoclaving were filter-sterilized using 0,2 µm filters. 1 M HCL and 1 M NaOH were used to bring the solutions to the desired pH. The rest of the reagents including growth media were sterilized by autoclaving at 121°C and 103 KPa for 20 minutes. Table A4 in Appendix A summarises the solutions and growth media, their usage, and how they were prepared.

2.2.4. Kits

Two kits were used in this project for DNA extraction and Western blot visualisation. GeneJET Plasmid Miniprep Kit purchased from Thermo Fischer Scientific (Massachusetts, USA) was used to extract *pQE30/PFF1010c* from *E. coli* JM109 used for *pQE30/PFF1010c* plasmid enrichment. Pierce™ Fast Western Blot Kit, also purchased from Thermo Fischer Scientific (Massachusetts, USA), was used to resolve Western blots to validate protein expression and purification.

2.2.5. Proteins, enzymes, and standards

The primary antibodies used in the current study were rabbit-raised. These included anti-PFF1010c (Mutavhatsindi, 2016), anti-PfHsp70-1 previously described by Gitau *et al.* (2012), and anti-Histidine (Thermo Fischer Scientific, USA). Secondary antibodies were mouse-raised, and Horseradish peroxidase (HRP) was conjugated for visualization (Thermo Fischer Scientific, USA). Luciferase enzyme used for protein aggregation and refolding studies was purchased from EMD Millipore (Massachusetts, USA). Bovine serum albumin and lysozyme were purchased from Thermo Fischer Scientific (Massachusetts, USA). The molecular weight markers used were the GeneRuler 1 kb DNA ladder and the Prestained Protein Molecular Weight Marker (20 – 120 kDa) from Thermo Fischer Scientific (Massachusetts, USA).

2.2.6. Plasmid constructs and bacterial strains

The plasmids and *E. coli* cells used for the expression and purification of recombinant proteins are listed in Table 2.1.

Table 2.1. Plasmids and bacterial strains

Plasmid construct	Description	Supplier
<i>pQE30/PFF010c</i>	<i>pQE30</i> expression vector with <i>PFF010c</i> codon harmonized insert and an ampicillin selection marker, amp ^R	Mutavhatsindi, 2016
<i>pQE30/PfHsp70-1</i>	<i>pQE30</i> expression vector with <i>PfHsp70-1</i> codon harmonized insert and an ampicillin selection marker	Chakafana <i>et al</i> , 2021
<i>pQE30/PfHsp40</i>	<i>pQE30</i> expression vector with <i>PfHsp40</i> codon harmonized insert and an ampicillin selection marker	Lebepe, 2018
<i>pQE30</i>	Bacterial vector for expressing 6xHis-tagged proteins, amp ^R selection marker	Qiagen (Hilden, Germany)
Bacterial Strains	Description	Supplier
<i>E. coli</i> JM109	K strain, <i>recA</i> ⁻ and <i>endA</i> ⁻ minimize recombination, improve the quality of plasmid DNA, allows blue/white screening	Thermo Fischer Scientific (Massachusetts, USA)
<i>E. coli</i> XL1 Blue	K strain, <i>recA</i> ⁻ and <i>endA</i> ⁻ minimize recombination, improve the quality of plasmid DNA, allows blue/white screening,	Thermo Fischer Scientific (Massachusetts, USA)

2.3. Bioinformatic analyses

A blueprint for characterising the biophysics and the biochemistry of a hypothetical or putative protein is *in silico* analysis (Bastiaan-Net *et al.*, 2020; Gasteiger *et al.*, 2005; Gaudermann *et al.*, 2006). Protein primary structures used in this study were retrieved from the public domain databases, PlasmoDB 63 ([PlasmoDB](#)) and NCBI ([National Center for Biotechnology Information \(nih.gov\)](#)) in FASTA format (Aurrecoechea *et al.*, 2008; Schoch *et al.*, 2020).

2.3.1. PFF1010c homologue search

Homology can be defined as similarities due to common ancestry (Lewis *et al.*, 2012; Pearson, 2013). A homology search was conducted to interrogate the phylogeny of PFF1010c (PF3D7_0620700). PFF1010c putative homologs were searched using BLASTP (accessed at [Protein BLAST: search protein databases using a protein query \(nih.gov\)](#)) (Pearson, 2013). Protein BLAST is arguably a leading sequence search tool (Fatlawi *et al.*, 2023). Uncharacterised and hypothetical protein hits were excluded due to insufficient information about their actuality. The issue of the conventional BLASTP's limit to detect remote homologues can be addressed by using a follow-up homologue search tool suitable for detecting distant homologues (Gabler *et al.*, 2020). HHPRED (accessed at [Bioinformatics Toolkit \(mpg.de\)](#)) (Söding *et al.*, 2005), a tool reported to be more sensitive for distant homologue identification was used as a follow up homologue search tool (Gabler *et al.*, 2020). Most studies use BLASTP only (Bitacura & Santos, 2022; Dishnica *et al.*, 2023; Lopez *et al.*, 2021; Mooi *et al.*, 2023; Xiao

et al., 2021). One study coupled normal BLASTP to PSI-BLAST (Ang'ang'o *et al.*, 2023). HHPRED is similar to PSI-BLAST but more sensitive (Hildebrand *et al.*, 2009). As such, HHPRED was used in this study instead of PSI-BLAST. Protein hits are considered homologous when the E-value is greater than 0.001, percent identity greater than 30 % and a bit score greater than 50 (Pearson, 2013).

2.3.2. Multiple sequence alignment

Multiple sequence alignment reveals conserved regions in the query protein and its aligned homologues. The data can be used to infer homology and predict the structure and function of the protein in the study (Edgar & Batzoglou, 2006). The selected homologues from the homology searches were aligned using a multiple sequence alignment tool called ClustalW (Sievers & Higgins, 2021) on the Bioedit Sequence Alignment Editor software (Hall *et al.*, 2011; Hall, 1999) as an accessory application. The alignments were transferred to ESPript (accessed at [ESPrpt 3 \(ibcp.fr\)](http://ESPrpt3.ibcp.fr)) (Gouet *et al.*, 1999) for alignment annotation with the inclusion of the query protein secondary structure. ESPript was used to highlight conserved motifs and residues.

2.3.3. Homologues phylogenetic tree

Homology assessment is incomplete without assessing the phylogenetic relationship between the query protein and the similar homologue proteins (Ochoterena *et al.*, 2019). MEGA11 is commonly used for phylogenetic tree construction and analysis (Ma *et al.*, 2022; Rosales *et al.*, 2023). Tree diagrams were constructed using MEGA11 software for the selected BLASTP and HHPRED protein hits (Tamura *et al.*, 2021). Evolutionary history was inferred using the neighbour-joining method with 500 (default) bootstrap replications.

2.3.4. Domain identification on PFF1010c

Protein domains are structural subunits of proteins that fold independently and confer a function to the protein (Schultz, 2006; Wang *et al.*, 2021). Since protein domains are classified into families based on the fold and function, a domain search on PFF1010c can provide protein families PFF1010c clusters to and eventually the protein family function (Morris *et al.*, 2022). PFF1010c domains were recognised using the CDD/SPARCLE (accessed at [NCBI Conserved Domain Search \(nih.gov\)](http://NCBI Conserved Domain Search (nih.gov))). This online database contains a classification of proteins through protein subfamily architectures (Marchler-Bauer *et al.*, 2016).

2.3.5. Prediction of PFF1010c secondary and tertiary structures

A protein fold determines its function (Kesheri *et al.*, 2015). Secondary structure prediction was rendered using PSI-BLAST based secondary structure prediction (PSIPRED, accessed at [UCL-CS Bioinformatics Introduction](#)) (Buchan & Jones, 2019; Jones, 1999). PSIPRED is an online tool for multiple bioinformatic predictions including secondary structure of a protein. PSIPRED sustains over 76 % secondary structure prediction accuracy (Torrissi *et al.*, 2020). However, none of the tools was capable of correctly predicting the whole secondary structure of a protein on its own (Angamuthu & Piramanayagam, 2017). Self-optimized prediction method with alignment (SOPMA) was also used to predict the secondary structure of PFF1010c. SOPMA correctly predicts 69.5 % of the protein sequence (Geourjon & Deleage, 1995). SOPMA aligns homologues from SWISSPROT using ClustalW and predicts the secondary structure based on the homologues multiple sequence alignment. Angamuthu and Piramanayagam (2017) reported that SOPMA predicts sheets more accurately than helices (Angamuthu & Piramanayagam, 2017). PFF1010c tertiary structure was searched for on AlphaFold (accessed at [AlphaFold Protein Structure Database](#)), a database of predicted tertiary structures of proteins (David *et al.*, 2022). AlphaFold predicted structures have been shown to be competitively reliable against experimental structures (Jumper *et al.*, 2021). This makes AlphaFold an ideal tool for tertiary structure prediction.

2.3.6. Prediction of PFF1010c homo-oligomerization

GalaxyWEB is a server providing programs for modelling protein structures and interactions (Shin *et al.*, 2014). Oligomerization of PFF1010c was predicted using the GalaxyWEB server's GalaxyHomomer tool (Baek *et al.*, 2017). This tool uses the GalaxyGemini program that outperformed all the other programs in the CASP (critical assessment of techniques for protein structure prediction) competition (Shin *et al.*, 2014). PFF1010c amino acid sequence was submitted on the server's GalaxyHomomer tool to predict homo-oligomerization. PfHsp40 (PF3D7_1437900) was used as a positive control since the protein is a canonical DnaJ protein of the *P. falciparum* 3D7 strain and forms dimers (Anas *et al.*, 2020; Botha *et al.*, 2011). PfHop (PF3D7_1434300) was used as a negative control since PfHop is reported to occur as a monomer (Makumire *et al.*, 2020).

2.3.7. Prediction of PFF1010c localisation

Finding the localisation of a putative protein aids in finding possible interactors and, eventually, the possible function. The online server, PSORT II (accessed at [WoLF PSORT: Protein Subcellular](#)

[Localization Prediction \(hgc.jp\)](#)), was used to predict the subcellular localisation of PFF1010c based on various sequence features in the protein linked to protein sorting signals (Horton *et al.*, 2007). Other servers used for protein localisation included SCLpred-EMS (Subcellular prediction – endomembrane system, accessed at [SCLpred-EMS \(ucd.ie\)](#)) and BUSCA (Bologna Unified Subcellular Component Annotator, accessed at [BUSCA - Bologna Biocomputing Group \(unibo.it\)](#)).

2.3.8. Prediction of PFF1010c cellular protein associates

The possible role of a protein can be deduced from its functional associates (Farooq *et al.*, 2021). The STRING protein-protein interaction database archives predicted protein interactions based on experiments, co-expression, cooccurrence and other *in silico* studies suggesting interaction between proteins (Szkłarczyk *et al.*, 2023). PFF1010c was searched on the STRING website using the keyword 'PFF1010c'. The hit was limited to 10 possible interactors.

2.3.9. Prediction of PFF1010c post-translational modification sites

Some proteins are modified post translation to channel them to a certain organelle or to perform a certain function in eukaryotes (Ramazi & Zahiri, 2021). Possible posttranslational modifications of PFF1010c were predicted using MusiteDeep (accessed at [MusiteDeep](#)), a web server for predicting protein post-translational modification sites (Wang *et al.*, 2020).

2.3.10. Prediction of physicochemical properties of PFF1010c

Lack of background knowledge about the behaviour of putative proteins poses a challenge during protein expression. An attempt to circumvent this is to theoretically determine the protein physical and chemical properties. The Expasy server's ProtParam tool accessed at ([Expasy - ProtParam tool](#)) computes a protein's physicochemical properties from its amino acid composition (Gasteiger *et al.*, 2019). PFF1010c (PlasmoDB accession number, PF3D7_0620700) amino acid sequence obtained from the PlasmoDB website was submitted on the ProtParam tool query section for computation. ProtParam calculated the molecular weight, theoretical pI (Bjellqvist *et al.*, 1993), atomic composition, extinction coefficient (Gill & Hippel, 1989), half-life (Bachmair *et al.*, 1986), instability index (Guruprasad *et al.*, 1990), aliphatic index (Ikai, 1980) and grand average of hydropathicity (GRAVY) (Kyte & Doolittle, 1982).

2.4. Experimental analysis

Bioinformatic analysis and wet lab analysis can corroborate and result in a deeper understanding of PFF1010c (Laub *et al.*, 2023; Xu *et al.*, 2022). As a follow-up, biochemical and biophysical studies were conducted to further elucidate the structure-function features of PFF1010c.

2.4.1. Preparation of competent cells

A DNA plasmid constructed from the gene of *PFF1010c* inserted into the *pQE30* vector was transformed into *E. coli* JM109 and *E. coli* XL1 Blue for DNA plasmid enrichment and protein expression, respectively. The Hanahan method was used to make cells competent at taking up plasmid DNA from their environment, as described below (Hanahan, 1983). The materials used are in Appendix A, Table A1 to A3.

E. coli JM109 and *E. coli* XL1 Blue cells (Thermo Fischer Scientific, USA) were streaked overnight on double strength (2x) Yeast-Tryptone (YT) agar plates while the control plate was streaked with dH₂O and incubated overnight (16 hours). The overnight single colonies from the respective plates were separately inoculated into 5 mL of fresh 2x YT liquid broth and incubated overnight, shaking in a 37 °C incubator. The control plate did not have colonies. The 5 mL overnight cell suspension was diluted 10 times by transferring into a fresh 45 mL 2x YT broth. The resulting cell suspensions (both *E. coli* JM109 and *E. coli* XL1 Blue) were allowed to cultivate by placing the cell culture into the incubator for 1 - 3 hours. The cells were allowed to reach optical density (OD) between 0.4 and 0.6 (mid-exponential growth) (Collins *et al.*, 2013). OD readings were measured at 600 nm for cell density, hence denoted as OD₆₀₀ (Myers *et al.*, 2013). The cells were pelleted into a 50 mL falcon tube by centrifuging the 50 mL cell cultures at 5000 × g for 10 minutes at 4 °C to separate them from the spent medium. The pelleted cells were resuspended in 10 mL of 0.1 M ice cold MgCl₂ and left on ice for 30 minutes. This allowed Magnesium (II) cations to attach to the cell's surface. The MgCl₂ cell suspensions were centrifuged for 10 minutes at 4000 ×g at 4 °C to separate the cells from the MgCl₂ solution. The Mg²⁺ cells pelleted were gently resuspended in 10 ml ice cold 0.1 M CaCl₂, followed by incubation on ice for 4 hours to attach Ca²⁺ on the cells surface. Then, centrifugation was carried out at 4000 ×g at 4 °C for 10 minutes to separate the cells from the CaCl₂ solution. The Ca²⁺ cells pelleted were aliquoted with 5 ml of sterile 15 % glycerol and 5 mL of 0.1 M CaCl₂. The *E. coli* JM109 and *E. coli* XL1 Blue competent cells were left to sit for 10 minutes on ice, thereafter 200 µl were aliquoted into correctly labelled microcentrifuge tubes and stored at -80 °C.

2.4.2. Transforming *E. coli* competent cells

Transformation in genetic engineering is the taking up of naked DNA by a bacterium from its environment (Hasegawa *et al.*, 2018). While some bacteria take up DNA naturally (natural transformation), there are artificial ways to improve transformation. Heat shock transformation is the most common artificial transformation, whereby heat is used to open up cells in order to take up plasmid DNA (Ren *et al.*, 2019). DNA plasmid constructs were transformed into the prepared competent cells through heat shock transformation for plasmid DNA enrichment and protein expression.

Plasmid DNA of 1 µl (~10 ng) was mixed with 50 µl aliquot of competent cells. The DNA and cells mix was incubated on ice for 30 minutes and heat shocked at 42°C for 45 seconds afterwards. The heat shocked cells were moved to ice for 10 minutes. The 2xYT broth of 900 µl was added and the resulting mix was incubated at 37°C for 1 hour. The 1 hour spent media was removed by centrifugation, leaving 100 µl for culturing the transformed cells. The 100 µl of transformed cells were streaked onto the 2xYT plate containing the ampicillin (100 µg/ml) to select for transformants. The streaked plates were incubated at 37 °C overnight.

2.4.3. Plasmid DNA extraction

E. coli JM109 was transformed with *pQE30/PFF1010c*, *pQE30/PfHsp70-1* and *pQE30/PfHsp40* DNA plasmids for plasmid enrichment. Agar plates with transformants were inoculated in 5 ml of 2xYT broth and cultured overnight. The enriched plasmids were extracted using the GeneJET Plasmid Miniprep Kit from Thermo Fischer Scientific (Massachusetts, USA), following the manufacturer's instructions. The extracted DNA was quantified using NanoPhotometer® (Implen, Inc., USA), along with purity assessment.

2.4.4. Restriction digest of plasmid DNA

The extracted plasmids needed to be confirmed to know that the correct DNA was extracted. Plasmid constructs can be confirmed by digesting them with restriction endonucleases. Restriction endonucleases are enzymes that recognize and cut DNA at a specific site (Green & Sambrook, 2021). DNA plasmids *pQE30/PFF1010c*, *pQE30/PfHsp70-1*, and *pQE30/PfHsp40* were constructed by inserting the genes in pQE30 vector at the *Bam*HI and *Hind*III cut sites (Chakafana *et al.*, 2021; Mutavhatsindi, 2016). FastDigest restriction endonucleases *Bam*HI and *Hind*III (Thermo Fischer

Scientific, USA) were used to cut the plasmid DNA constructs once and twice to get the linear fragment and cut out the gene from the vector. Table 2.2. summarised the reagents and their volumes used to conduct the experiment. Four reactions were conducted with the first reaction serving as the control tube. No enzyme was added for the control tube. The following two reactions were single digest by either *Bam*HI and *Hind*III restriction endonucleases, and the last reaction was double digestion using both enzymes *Bam*HI and *Hind*III. The reaction samples were incubated at 37°C for 1 hour, permitting the digestion of the nucleic acids by the *Bam*HI and *Hind*III restriction endonucleases.

Table 2.2. Restriction endonuclease reaction reagents and their volumes

Reagents	Reaction tube 1	Reaction tube 2	Reaction tube 3	Reaction tube 4
Nuclease free water	16 µl	15 µl	14 µl	13 µl
Green buffer	2 µl	2 µl	2 µl	2 µl
<i>Bam</i> HI	-	1 µl	-	1 µl
<i>Hind</i> III	-	-	2 µl	2 µl
DNA	2 µl (~ 2 µg)	2 µl (~ 2 µg)	2 µl (~ 2 µg)	2 µl (~2 µg)
Total	20 µl	20 µl	20 µl	20 µl

2.4.5. Agarose Gel Electrophoresis

The digested DNA plasmids need to be separated by an analytical technique to resolve it by size. Agarose gel electrophoresis is a biochemical technique used to separate DNA based on length (Martin, 2020). 0.8 % agarose gel used to resolve DNA of sizes lengths between 12 to 0.8 kilobase pairs (Kbps) (Ogden & Adams, 1987), was prepared by dissolving 1.2 g of agarose with 1X Tris Acetate EDTA (TAE) buffer in 150 mL Erlenmeyer flask. Because agarose has low solubility constant, the flask with its content was microwaved until the agarose was fully dissolved. After the solution has cooled, ethidium bromide (EtBr) was added to the final concentration of 5 µg/mL. The solution was poured into the casting tray. The comb was placed in the gel to create wells, and the tray was left to sit for 30 minutes at room temperature to solidify. The gel electrophoresis unit (Merck, Darmstadt, Germany) was filled with 1X TAE buffer until the gel placed inside was completely covered. The molecular weight marker was loaded into the first lane. The digested DNA samples including the control were loaded into the gel wells. The gel unit was plugged to the power source to run at 100 Volts for 1 hour. The gel was visualized using ChemiDoc Imaging Systems from Bio-Rad Laboratories, Inc. (California, USA).

2.4.6. Expression of recombinant protein forms of PFF1010c, PfHsp40, and PfHsp70-1

E. coli XL1 blue strain is routinely used in the protein laboratory for heterologous protein expression (Wong & Tee, 2020). However, the ampicillin resistant gene on the DNA vector sequence was used as a selective marker instead of the blue-white screening. The extracted plasmid DNAs (section 3.2.1 and appendix C1) were transformed into *E. coli* XL1 blue competent cells.

Single colonies of transformants were inoculated into 5 mL 2xYT broth and grown overnight to make starter cultures. The starter cultures were transferred to fresh 45 mL 2xYT broth flasks and grown for 10 to 30 minutes in the 37 °C incubator until the culture reaches the OD₆₀₀ between 0.4 and 0.6. Protein expression was induced using Isopropyl β-d-1-thiogalactopyranoside (IPTG), an allolactose mimic since the genes of interest were under the control of the *lac* operator (Faust *et al.*, 2015). Before returning the flask back into incubation, the 0th hour sample was taken for analysis. The culture OD₆₀₀ reading was taken and sampled every hour until the 6th. Samples were 1 ml aliquots after taking OD₆₀₀ reading. The spent media in the hourly samples was removed by centrifugation at 5000 x g for 5 minutes and discarding the supernatant. The cell pellets were resuspended in phosphate buffered saline (PBS, composition: 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.5), in a volume equivalent to OD₆₀₀ x 200 µl and kept at -20 °C for analysis.

2.4.7. Protein solubility assessment

Protein solubility affects the choice of purification method to employ (Terol *et al.*, 2021; Wong & Tee, 2020). Proteins were overexpressed using *E. coli* XL1 Blue for 6 hours and the 50 ml culture was pelleted at 6000 g for 20 minutes at 4 °C. The pellet was resuspended in 5 ml lysis buffer (Tris buffered saline (TBS), composition: 10 mM Tris, pH 8.0, 300 mM NaCl and 10 mM imidazole). 1 mM of protease inhibitor phenylmethylsulfonyl fluoride (PMSF) and 1 mg/ml of cell wall enzyme Lysozyme were added. PMSF protects proteins from proteases after cell lysis. The suspension was incubated for 30 minutes at room temperature to allow lysozyme action. Lysozyme produced a spheroplast, easing the further lysis process. Afterwards, the cells were sonicated at an amplitude of 30 % for 12 cycles with 15 seconds pulse and 30 seconds pause between pulses. Cell lysates were sampled before centrifuging at 6000 x g for 10 minutes at 4 °C to separate the soluble fraction (supernatant) from the insoluble fraction (pellet). The supernatant was sampled and decanted into a sterile falcon tube. The pellet was resuspended in 5 ml PBS (pH 8.0) sampled as the insoluble fraction.

2.4.8. Protein purification

The overexpressed proteins needed to be purified for further analysis. The pQE30 vector was designed to introduce an N-terminal polyhistidine (at least 6 histidine residues) tag (Loughran *et al.*, 2006). Polyhistidine tag has affinity for Nickel (II) metal cation. The recombinant proteins were purified using immobilized metal affinity chromatography (IMAC), a technique for purifying polyhistidine tagged proteins (Riguero *et al.*, 2020).

Pellets from expression cultures were resuspended with lysis buffer. For insoluble proteins, 8 M urea was supplemented to the lysis buffer to solubilize the proteins. PMSF of 1 mM was always supplemented to the lysis buffer to protect proteins. The cells were further lysed by sonication at an amplitude of 30 % for 12 cycles with 15 seconds pulse and 30 seconds pause between pulses. Cell lysates were centrifuged at 6000 x g for 20 minutes to bring up the soluble proteins and separate them from the cell debris in the pellet. The supernatant containing soluble proteins, including the expressed His-tagged proteins, was added to the IMAC column containing HisPur™ nickel-nitrilotriacetic acid (Ni-NTA) Resin (Thermo Fischer Scientific, USA) and incubated for 1 hour rocking at 150 revolutions per minute (rpm) on ice. The column outlet was opened for the solution to flow through. The flow through was sampled. The polyhistidine tag chelates the metal on the resin, and the non-tagged proteins were washed off the resin twice with imidazole containing TBS buffer with low imidazole concentration. The TBS composition for wash buffers included 50 mM Tris pH 8.0, 150 mM NaCl. Wash 1 contained 5 mM imidazole and wash 2 contained 80 mM Imidazole. The washes were sampled. The bound protein was eluted using high imidazole concentrations in TBS used for washes. Elution 1 contained 250 mM Imidazole and elution 2 contained 350 mM Imidazole. The elutions were collected and sampled. All samples were temporarily stored at -20 °C. Purified proteins were quantified using the NanoPhotometer® (Implen, Inc., USA).

2.4.9. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Analysis of protein expression, solubility and purification required an analytical technique for the separation and visualization of the proteins. SDS-PAGE allows the separation of proteins based on size (Gandhi *et al.*, 2022). Laemmli method was followed for SDS-PAGE analysis in this study (He, 2011). Protein samples were treated with SDS-PAGE sample loading buffer (composition: 250 mM Tris-Cl, pH 6.8), 8 % SDS, 0.1 % Bromophenol blue, 40 % glycerol, 100 mM Dithiothreitol, diluted in dH₂O) (Kielkopf *et al.*, 2021). The samples were heated at 70-100 °C for 10 minutes and analysed by 12 % SDS-PAGE

gel. 12 % SDS-PAGE gel was prepared as described below, including the 5 % stacking gel (Table 2.2) (Gallagher, 2012; Patel *et al.*, 2016).

Table 2.3. Reagent amounts used for preparing 12 % SDS-PAGE gel

Reagent	12 % Separating gel (μl)	5 % Stacking gel (μl)
30 % acrylamide/Bis solution	2080	235
1.5 M Tris pH 8.8	1250	0
0.5 M Tris pH 6.8	0	438
10 % SDS	50	17
dH ₂ O	1580	1050
10 % APS	25	85
TEMED	20	20

The gel was transferred into the electrophoresis tank and SDS-PAGE running buffer (composition: 25 mM Tris pH 8.3, 250 mM Glycine and 0.1 % SDS) was added. The heated samples were loaded in respective wells and pre-stained protein molecular weight marker was also loaded in a separate well. The electrophoresis gel was run at 120 volts and current set at 300 mA for 1 hour and 30 minutes. The gels were removed from the apparatus and stained with Coomassie blue stain (composition: 0,1 % Coomassie blue R250, 40 % ethanol, 10 % acetic acid in dH₂O) for at least 2 hours rocking at 150 rpm at room temperature. The gels were destained with the destain solution (composition: 40 % ethanol, 10 % acetic acid in dH₂O) for 30 minutes, 3 times rocking at 150 rpm at room temperature. The clear gels with visible resolved proteins were imaged using ChemiDoc Imaging Systems (Bio-Rad Laboratories Inc., USA).

2.4.10. Western blot analysis

To confirm that the correct protein was expressed and purified, Western blot technique was used (Kurien, 2021). An SDS-PAGE gel was run. The gel was soaked in the Western transfer buffer (Towbin buffer, composition: 25 mM Tris, 192 mM glycine, pH 8.3 and 20 % percent methanol) along with 4 filter papers and the nitrocellulose membrane for 5 minutes. The Western blot transfer stack was prepared on a Power Blotter–semi-dry transfer system (Thermo Fischer Scientific, USA) in the following order: 2 filter papers, nitrocellulose membrane, SDS-PAGE gel, and 2 filter papers. The stack was placed on the anode, rolled with a roller to remove bubbles, and covered with the cathode on top. The transfer ran at 25 volts and 1.3 A for 7 minutes using the Power Blotter. Afterwards, the membrane was moved into a container with 5 % non-fat milk in a TBS buffer (pH 7.5) and incubated on ice for 1 hour rocking at 100 rpm. The milk blocked part of the membrane that did not have proteins

transferred. The membrane was then washed 3 times with TBS buffer supplemented with 0,1 % TBS-Tween, pH 7.4 in 5-minute intervals. Washing was followed by addition of primary antibodies (antibodies against target proteins, composition: 1:2000 dilution in 5% non-fat milk prepared in TBS) incubated on ice and rocking at 60 rpm for 1 hour. Non-bound antibodies were removed by washing the membrane the same way as the first wash step. Secondary antibodies (antibodies recognizing the primary antibodies) were added and incubated under the same conditions as the primary antibodies. The membrane was washed as the initial washes. The bands were then visualized using Pierce™ Fast Western Blot Kit's ECL substrate (Thermo Fischer Scientific, USA) and ChemiDoc Imaging Systems (Bio-Rad Laboratories Inc., USA).

2.4.11. Liquid Chromatography - Mass Spectrometry (LC-MS)

Western blot is a sensitive and reliable technique for identifying a given protein (Moritz, 2020). In this study, purification SDS-PAGE gel, and Western blot analysis of PFF1010c returned results suggestive of oligomerization. The possible oligomers were cut out of the SDS-PAGE gel and subjected to LC-MS. LC-MS is another sensitive technique that couples the separation of peptides by liquid chromatography to the identification of the peptides by mass spectrometry (Kulyyassov *et al.*, 2021). 10 ml of purified PFF1010c was dialyzed and concentrated to 1 ml to intensify the faint bands in the background and cut out the bands for trypsin in-gel digestion and LC-MS analysis. LC-MS analysis was conducted in Dr Previn Naicker's laboratory at CSIR, Pretoria, South Africa.

2.4.11.1. Trypsin in-gel digestion

In-gel digestion was conducted according to Shevchenko *et al.*, (2006). Briefly, the proteins were reduced in gel with 10 mM DTT in 25 mM sodium bicarbonate (NH_4HCO_3) for one hour at 60 °C. Samples were allowed to cool to room temperature, then 100 % acetonitrile was added and incubated for ten minutes at room temperature. The supernatant was discarded and 55 mM iodoacetamide (IAA) in 25 mM was added to the gel cuts. The reaction proceeded in the dark for twenty minutes at room temperature. The supernatant was discarded, and gels were dehydrated with 25 mM NH_4HCO_3 in 50 % acetonitrile, vortexed and the supernatant was removed. The gel pieces were dried to completeness and freshly prepared trypsin was added, protein digestion was allowed to proceed overnight at 37 °C. The digestion was quenched by adding final 0.1 % formic acid (FA) and the samples were dried under vacuum. Dried samples were re-suspended in 2 % acetonitrile and 0.2 % FA for mass spectrometry analysis.

2.4.11.2. LC-MS analysis

Tryptic peptides from each gel fraction were analysed using a Dionex Ultimate 3000 RSLC system (Thermo Fischer Scientific, USA) coupled to an AB Sciex 6600 Triple TOF mass spectrometer (SCIEX, USA). Injected peptides were inline de-salted using an Acclaim PepMap C18 trap column ((Thermo Fischer Scientific, USA; 75 μm $\times$ 2 cm; 2 min at 5 $\mu\text{l}/\text{min}$ 1 using 2 % ACN/0.2 % formic acid). Trapped peptides were gradient eluted and separated on a Waters nanoEase CSH C18 column (Waters, USA; 75 μm $\times$ 25 cm, 1.7 μm particle size) at a flow rate of 0.3 $\mu\text{l}/\text{min}$ with a gradient of 6-40 % B over 15 min (A: 0.1 % FA; B: 80 % ACN/0.1 % FA). The 5600 Triple TOF mass spectrometer (SCIEX, USA) was operated in positive ion mode. Data-dependent acquisition (DDA) was employed; precursor (MS) scans were acquired from m/z 400-1500 (2+ - 5+ charge states) using an accumulation time of 100 milliseconds (ms) followed by 40 fragment ion (MS/MS) scans, acquired from m/z 100-1800 with 20 ms accumulation time each. Raw data files (.wiff) were searched with Protein Pilot V5.0 software (SCIEX), using a protein database containing available *E. coli* sequences downloaded from UniProt (on 27 November 2023), common contaminants and sequences of recombinant proteins of interest. Trypsin was set as the digestion enzyme, cysteine alkylation (iodoacetamide) was allowed as a fixed modification in the search parameters.

2.4.12. Biophysical characterization

2.4.12.1. Circular dichroism (CD) spectroscopy analysis

CD spectroscopy, a technique that measures the interaction of biomolecules with plane polarized light, is commonly used to estimate the secondary structure of a protein (Kumari & Ramakrishnan, 2023; Miles *et al.*, 2021). The experimental estimation of PFF1010c secondary structure was carried out according to the protocol described by (Greenfield, 2006) with few modifications.

PFF1010c was purified and dialyzed with PBS buffer. PFF1010c concentration was adjusted to 2 μM and CD spectra were obtained on the J-1500 CD Spectrophotometer (Jasco, Japan). The CD spectrophotometer machine was equilibrated with the PBS buffer without proteins. The CD spectra were measured from 178 to 280 nm with 0,5 nm steps. The temperature was 25 $^{\circ}\text{C}$ and the cell path length was 1 cm. The obtained PFF1010c CD spectra were uploaded on the B-Structural Selection (BeStSel) webserver, accessed at <https://bestsel.elte.hu/>, for fitting and analysis. Fitting on BeStSel required the protein concentration, number of amino acid residues and the cell path length used to measure the CD spectra. A Normalized Root Mean Square Deviation (NRMSD) value was generated upon analysis. A suitable NRMSD value should be less or equal to 0.1. BeStSel webserver is said to

overperform other similar webserver by providing 8 secondary structure modules and matching the submitted protein domains to CATH domain database domains with 90 % sequence homology (Micsonai *et al.*, 2023).

2.4.12.2. 1-Anilino-8-Naphthalene Sulfonate (ANS) extrinsic fluorescence

Initial *in silico* analysis of PFF1010c suggested that the protein is thermostable. In this regard, a protein is deemed to be thermostable when it retains its native structure at temperatures above 37 °C (Hussian & Leong, 2023). Thermostability could be inferred by evaluating the conformational changes the protein undergoes in response to heat stress.

Conformational changes of PFF1010c were assessed using ANS (Merk, Germany) fluorescence as previously described by Lebepe *et al.* (2020) with few modifications. ANS has affinity for hydrophobic patches of proteins, hence it is used as a probe to detect protein denaturation as represented by increasing hydrophobicity (Ota *et al.*, 2021). ANS (200 mM) was prepared using 1 % dimethyl sulfoxide (DMSO, Merck, Germany) and temporarily stored at 25 °C in the dark. PFF1010c was serially diluted 5 times from 2 µM by a dilution factor of 2. The varying amounts of PFF1010c were mixed with 200 µM final concentration of ANS. ANS was allowed to bind to protein by incubating the protein-ANS mix at room temperature. The extrinsic emission spectra were generated by exciting the protein-ANS sample at 375 nm and emission spectra was measured from 400 to 600 nm using the JASCO FP-6300 spectrofluorometer (JASCO, Spain). The ANS sample in PBS buffer without protein was used as a blank. PFF1010c concentration that produced a defined peak was heated at different temperatures while mixed with ANS and the emission spectra was measured as described above. The assay was repeated three times using independent batches of the protein.

2.4.13. Biochemical characterization

2.4.13.1. Luciferase refolding assay

The luciferase refolding assay has been extensively used to study the chaperoning activity of heat shock proteins (Chakafana *et al.*, 2021; Jung & Ahn, 2022; Manukhov *et al.*, 1999; Zavgelsky *et al.*, 2020). Firefly luciferase (Promega, USA) enzyme catalyses a luciferin oxidation reaction that luminesce (Gould & Subramani, 1988). As refoldases, Hsp70s co-operate with co-chaperones such as JDPs that regulate the ATPase function of the Hsp70 chaperones (Chakafana *et al.*, 2021; Lebepe *et al.*, 2020).

PFF1010c was predicted to interact with PfHsp70-1 and PfHsp40 (PF3D7_1437900, section 3.1.7). PfHsp70-1 was mixed with PFF1010c, to see if PFF1010c affects PfHsp70-1 refolding activity. A luciferase refolding assay was conducted as described by (Chakafana *et al.*, 2021), towards evaluating the refolding activity of PfHsp70-1 in the presence of PFF1010c. PfHsp40 (PF3D7_1437900) served as a non refoldase control. PfHsp40 (PF3D7_1437900) also served as a positive control for co-chaperone assisted refolding activity only when coupled with PfHsp70-1. Stock of luciferase (1 uM) was denatured in the luciferase assay buffer (Promega, USA) at 42 °C for 20 minutes. The denatured luciferase (0,1 uM final concentration) was mixed with 3 uM PfHsp70-1, 0.5 uM PfHsp40 (PF3D7_1437900), and 0.5 uM PFF1010c separately and in combination to refold the denatured enzyme. Refolding was allowed for 30 minutes at 25 °C. The chaperone – luciferase mix was mixed with 25 parts (volume) of the 100 ug/ml (working solution) luciferin reagent. Luminescence was measured at 526 nm using SpectraMax M3 spectrometer (Molecular Devices, USA). Positive control was a native luciferase, and negative control was denatured luciferase.

Chapter 3

3. Chapter 3: Results

The aim of this study was to characterize the structure-function features of PFF1010c. Here, findings of the study are reported. *In silico* analyses were conducted prior to biochemical and biophysical analyses.

3.1. *In silico* data

3.1.1. BLASTP Homologues of PFF1010c

Table 3.1 below contains the first 5 selected and possible homologues of PFF1010c, with the first protein in the table being PFF1010c, the query protein. The rest of the table is attached in appendix B as Table B1. The protein hit table includes the protein description with the protein organism in the brackets. Additional information includes the query coverage percentage, percent identity, protein sequence length, the accession number of the protein and the statistical E-value indicating that the hit was due to chance. A higher E-value suggests that the query and hit proteins alignment was random while a lower E-value indicate a significant alignment and biological relationship (Dash *et al.*, 2021). The protein hits can be considered as PFF1010c homologues since the E-value is less than 0.001, percent identity greater than 30 % and a bit score greater than 50. With the E-value being greater than 0.001, the protein hits were significant.

Table 3.1. PFF1010c protein hits from BLASTP (top 5)

Description	JDP subclass	Total score	Query cover (%)	E value	Percent identity (%)	Amino acid length	Accession number
1. DnaJ protein, putative [<i>Plasmodium falciparum</i> 3D7] (Query)	Type IV	716	100	0.0	100	362	XP_966193.1
2. DnaJ domain containing protein [<i>Plasmodium gonderi</i>]	Type IV	437	96	2e-149	64.76	377	XP_028544409.1
3. DnaJ domain containing protein [<i>Plasmodium cynomolgi</i> strain B]	Type IV	424	89	9e-145	65.23	362	XP_004223295.1
4. DnaJ domain containing protein [<i>Plasmodium vivax</i>]	Type IV	403	88	1e-136	61.99	358	XP_001616220.1
5. DnaJ domain containing protein [<i>Theileria equi</i> strain WA]	Type III	119	70	4e-27	31.13	279	XP_004829319.1
6. DNA RNA-binding Alba [<i>Babesia bovis</i>]	Type III	121	69	7e-26	29.37	847	GFE54964.1

The 15 protein hits (including the query protein) belonged to 9 genera. Protein hits most similar to PFF1010c were of the *Plasmodium*, *Theileria*, and the *Babesia* genera. Protein hits of the *Plasmodium* species were JDPs. However, the *Babesia bovis* protein was nucleic acid binding Alba protein. Alba proteins are an ancient family of proteins that bind nucleic acids to serve many functions, with stress adaptation amongst them (Wai *et al.*, 2021). The protein hits with less similarity to PFF1010c were members of the JDP subfamily C in the *Octopus* species. Subfamily/ class C, previously described as type III Hsp40s, are JDPs producing only the J domain with a HPD motif and the absence of glycine/phenylalanine rich region (Harm H. Kampinga *et al.*, 2019). Type IV JDPs are variations of type III forms with the distinction that the type IVs have a J domain lacking the highly conserved HPD motif. As such, type III JDPs are more likely to be part of the most similar PFF1010c (type IV) proteins than type I and type II JDPs. Overall, closest proteins were of the apicomplexan parasites. The other proteins belonged to the Molluscs and Cnidarian phyla.

HHPRED takes a different approach from BLASTP on predicting homologues. HHPRED server pulls crystalized or NMR structures of remote homologues (Söding *et al.*, 2005). Table 3.2. contains few (6) HHPRED protein hits. HHPRED uses parameters used by the BLASTP with few additional parameters. The probability measure highlights chances that the hit occurred by chance. The Table with raw data is attached under appendix B as Table B2. The higher the probability, the higher the likelihood that the hit occurred by chance. The E-value and the P-value are statistical measures also signifying how the hit

occurred by chance and the lesser the value, the greater the likelihood that the hit occurred by chance. The score or similarity score is the same as the bit score used by BLASTP. The bit score is high (>50) when there is an alignment between a pair of sequences. The SS score is the secondary structure similarity score. Cols is the number of columns matching between the alignment of the query protein and the protein hit. Query is the query coverage measure and HMM is the Hidden Markov Model. HMM's are statistical models used in machine learning, in this case HHPRED, to predict remote homologs.

Table 3.2. PFF1010c protein hits from HHPRED (top 5)

Description	JDP subclass	Probability	E value	P value	Score	SS	Cols	Query	HMM (amino acid length)
1. 6U3S_A DnaJ homolog subfamily [<i>Homo sapiens</i>]	Type II	99,5	4.2e-14	6.4e-19	109.0	9.2	69	21-89	2-72 (190)
2. 3U03_B J-type co-chaperone JAC [<i>Saccharomyces cerevisiae</i>]	Type II	99.5	6.6e-13	1e-17	100.2	13	124	15-140	4-156 (181)
3. 4IT5_D Co-chaperone protein Hs [<i>Vibrio cholerae</i>]	Type II	99.5	9e-13	1.4e-17	98.1	13.5	123	20-142	2-153 (174)
4. 1FPO_B Chaperone protein HSCB [<i>Escherichia coli</i>]	Type II	99.5	1.6e-12	2.5e-17	98.1	13.0	110	22-131	1-131 (171)
5. 7VPX_J DnaJ homolog subfamily [<i>Homo sapiens</i>]	Type II	99.4	3.2e-11	4.9e-16	95.2	15.2	125	6-130	40-182 (253)

HHPRED pulled JDPs with a wide variety of functions from co-chaperone functions to transport, structural and biogenesis functions. This reflects JDPs involvement in many processes besides chaperone activities. hits were of eukaryotic organisms such as *Homo sapiens* and *Saccharomyces cerevisiae*. It is also notable that some potential homologues originated from thermophilic organisms such as *Thermomyces lanuginosus* and *Tetrahymena thermophila*. This agrees with the computed physicochemical properties reported later in this chapter (section 3.1.10) suggesting that PFF1010c is thermostable. It is noteworthy also that both BLASTP and HHPRED did not pull orthologs. PFF1010c is thus unique in *P. falciparum* because close homologues are outside this species.

3.1.2. Multiple sequence alignment

PFF1010c multiple sequence alignment with protein hits from BLASTP sequence revealed that proteins with high sequence similarity to PFF1010c have an SVH motif instead of the typical HPD motif found in proteins with low sequence similarity, probably distant relatives of PFF1010c (Figure 3.1) (Gouet *et al.*, 2005; Hall *et al.*, 2011; Thompson *et al.*, 2003). The conservation of varying motifs of PFF1010c and its homologues are shown by two horizontal red lines and shading of the motifs (Figure 3.1). The SVN/H motif is conserved within the *Plasmodium* species. On the other hand, the SDH motif is conserved

among the *Octopus* and *Sepia* genera. *Octopus* and *Sepia* genera are not parasitic while *Plasmodium* species are (Tedesco *et al.*, 2020). Proteins bearing the SDH motif are more distant to the PFF1010c sequence in comparison to proteins bearing the HPD motif. The HPD motif appears to be widely and highly conserved as opposed to SDH and SVN/H motifs. HPD is present in many different genera. It is intriguing that PFF1010c displays an SVN motif instead of SVH. The residues adjacent to the SVN/H motif are highly conserved. One of them is the lysine residue flanking the conserved tripeptides. The SVN/H conservation could be suggesting a separately evolving group of proteins from the proteins bearing the HPD motif. Residues beyond the J domain showed a high level of sequence similarity. This is observed beyond residue 175 of the PFF1010c sequence. This alignment and conservation are shown in Figure 3.1. second panel below.

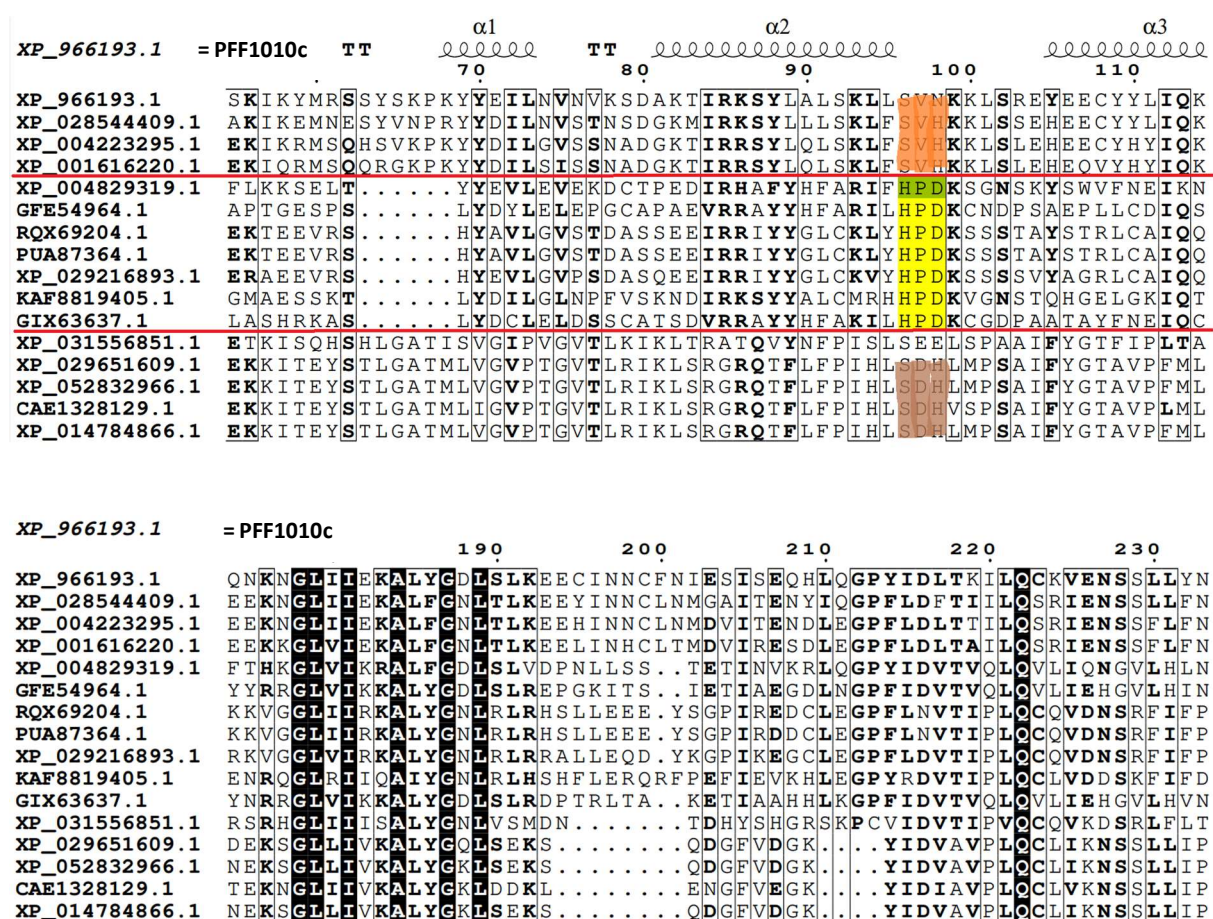


Figure 3.1. Multiple sequence alignment of PFF1010c with its BLASTP possible homologues

Multiple sequence alignment snippets with greater sequence similarity. Multiple sequence alignment was produced using clustalW omega tool on Bioedit and annotated using ESPript 3.0 online tool (Hall *et al.*, 2011; Robert & Gouet, 2014). ESPript 3.0 includes the predicted PFF1010c (Q-query protein) secondary structure. "TT" denotes strict β -turn and α 1 is the first α -helix in the protein predicted secondary structure. Digits above the alignment denotes positions the amino acids are at polypeptide sequence. A percent similarity is calculated using physicochemical properties and the shading is black and white. Solid black represents high similarity while white is no similarity. Similarity percentage cut off was 70%. sequence similarity

greater than 70% is boxed. The HPD tripeptide is yellow shaded, SVN/H tripeptide is orange shaded and SHD tripeptide is brown shaded.

PFF1010c multiple sequence alignment with HHPRED proteins (Figure 3.2) aligned the SVN motif in PFF1010c with the HPD motif in all the other proteins. No other sequence similarity as strong as that of the HPD motif was observed within or outside the J domain with the HHPRED protein hits alignment.

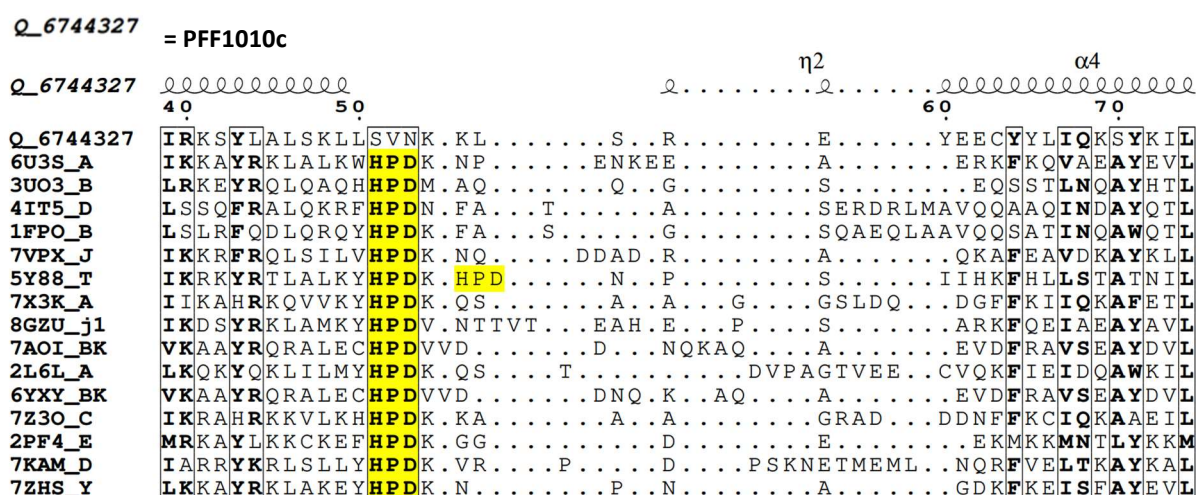


Figure 3.2. Multiple sequence alignment of PFF1010c with its HHPRED possible homologues

Multiple sequence alignment was produced in the same way as that of BLASTP protein hits in Figure 3.1 HPD motif is highly conserved within the J domain of this proteins.

3.1.3. PFF1010c phylogeny

Phylogenetic trees generated for both homologue searches (BLASTP and HHPRED) are depicted (Figure 3.3) (Felsenstein, 1985; Tamura *et al.*, 2021). PFF1010c appears to have evolved from the proteins bearing the HPD tripeptide in their J domain and assumed the SVN tripeptide over the evolutionary time. The SDH tripeptide evolved before the HPD and the SVN/H. This looks like reverse evolution because SVN/H is more like SDH than HPD. As seen with multiple sequence alignment (Section 3.1.2), the proteins with a domain containing the SVN/H tripeptide appears to be a separately evolving clade restricted to the *Plasmodium* genus. PFF1010c predates all the other SVN/H tripeptide JDP members. The immediate PFF1010c relative in this tree diagram is a JDP from *Plasmodium gonderi*. Besides causing malaria in monkeys, *P gonderi* harbours genes involved in vertebrate host interaction and undergoing fast evolution (Cepeda *et al.*, 2024). This could explain the incorporation of SVN/H motif

from HPD. This also allows the assumption that *P. falciparum* could be producing similar characteristics like *P. gonderi*, leading to *PFF1010c* incorporating the SVN/H motif.

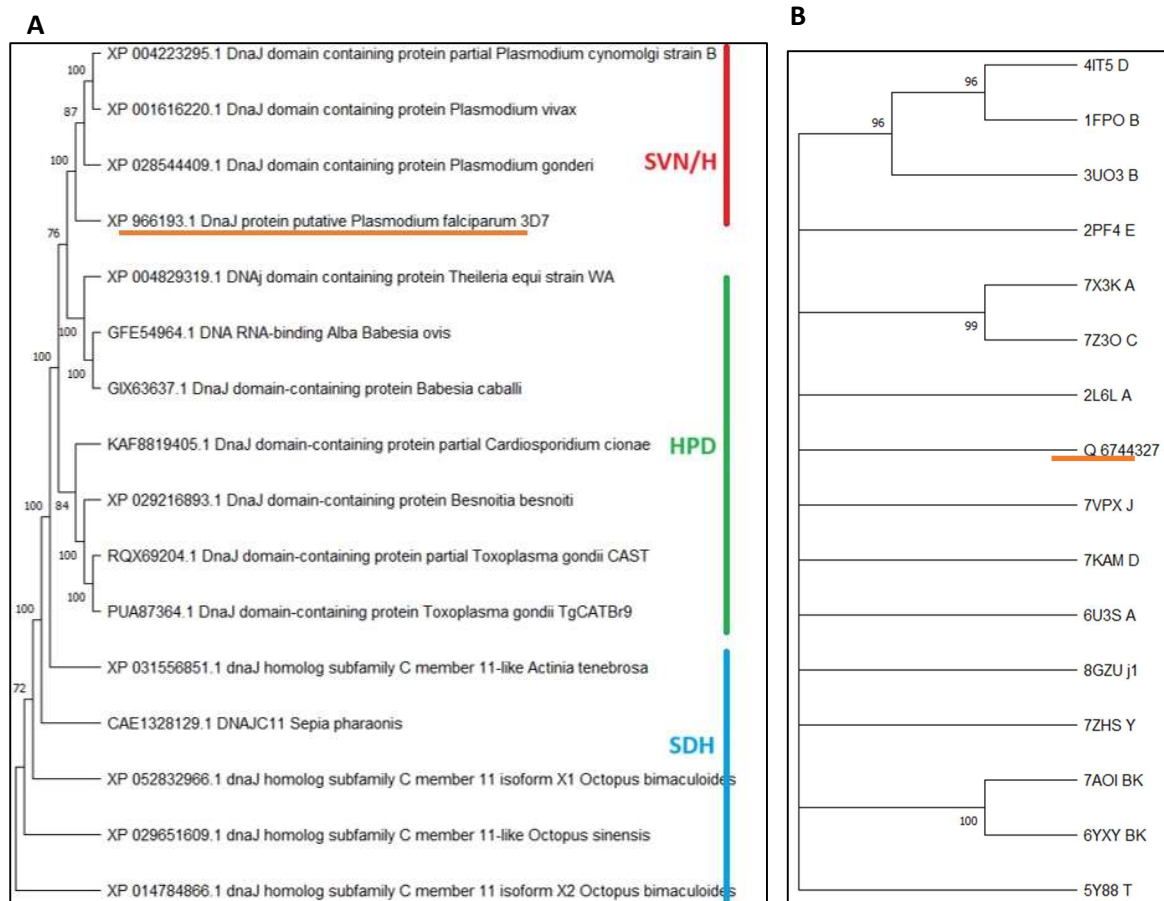


Figure 3.3. Phylogenetic trees of BLASTP(A) and HHPRED(B) protein hits.

The values on a branch represent the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates). Proteins bearing the SVN/H tripeptide in their J domain labelled in red, the HPD ones in green and SDH proteins in turquoise. In both trees, PFF1010c was underlined with an orange line. PFF1010c appeared to have evolved from the JDPs bearing the HPD tripeptide BLASTP proteins phylogenetic tree.

The HHPRED protein hits are too remote to produce clear ancestry lineages and to infer a phylogeny among the protein hits and PFF1010c (Figure 3.3.B). For BLASTP hits phylogenetic tree, the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) was over 72 % (Figure 3.3.A) (Felsenstein, 1985; Tamura *et al.*, 2021). The BLASTP hits phylogenetic tree 72 % bootstrap support value, suggest that the tree is 72 % reliable. Zhao *et al.* (2020) obtained bootstrap support values lower than 72 % when characterizing insulin and glucagon genes from Pygmy Sperm Whales (*Kogia breviceps*) and reported them as higher (Zhao *et al.*, 2020). A bootstrap support

value threshold of 70 % was used when characterizing a novel HIV-1 recombinant protein CRF91_cpx (Topcu *et al.*, 2022). Based on these two studies, The phylogenetic tree generated in this study is sufficiently reliable.

3.1.4. Protein domains of PFF1010c

The Conserved Domain Database (CDD) recognised two domains, namely, the J domain and an unknown domain. The J domain, known for modulation of the ATPase activity stimulation of Hsp70 proteins, spans 64 residues (68 to 132) on the N-terminal of PFF1010c. The J domain is currently recognized as DUF3395, and the accession number is pfam11875 on the CDD database. The unknown domain spans 162 residues (134 to 296) and is located on the C-terminus. The unknown C-terminal domain found in PFF1010c has the accession number, smart00271. According to the CDD database, this domain is not functionally characterized.

3.1.5. Predicted secondary and tertiary structures of PFF1010c

3.1.5.1. Predicted PFF1010c secondary structure

PFF1010c secondary structure predicted by PSIPRED constituted mostly of α -helices and coils (

Figure 3.4. A). The β -strands are not as prevalent as α -helices and coils and are restricted to the carboxy terminus. From the PSIPRED prediction, PFF1010c contains 35 % α -helices, 16 % β -sheets and 49 percent random coils (Table 3.3.). SOPMA produced the percent fractions of each secondary structure conformational states in the protein. The SOPMA predicted PFF1010c secondary structure is constituted by 40.61 % α -helices, 19,6 % extended strands, 7.73 % β -strands and 32.6 % random coils (Figure 3.4. C).

Secondary structure predictions generated using SOPMA and PSIPRED are closely related. This validates the reliability of these two tools. However, minor differences between the structures are shown through the cartoon structures (Figure 3.4.). The difference between the predicted secondary structure conformational states were 5.61 % α -helices, 11.33 % β -sheets and 16.4 % random coils (Table 3.4.). The mean values between the two predictors produced 37.8 % α -helices, 21.6 % β -sheets and 40.8 % random coils. Random coils are the most accessible secondary structure modules as opposed to the others (Lins *et al.*, 2003). This suggests that PFF1010c folds into a loosened and less compact structure.

Table 3.4. PFF1010c predicted secondary structure difference and mean values between the two predictors PSIPRED and SOPMA

Secondary structure	PSIPRED and SOPMA secondary structure content difference (%)	PSIPRED and SOPMA secondary structure content mean values (%)
α -helices	5.61	37.8
β -sheets	11.33	21.6
Random coils	16.4	40.8

3.1.5.2. Predicted PFF1010c tertiary structure

Based on the predicted tertiary structure of PFF1010c, the J domain and the C-terminal domain formed a cleft between the two domains which could possibly be the binding pocket of the protein (Figure 3.5) (Varadi *et al.*, 2022). While the two PFF1010c domains are far from each other on the extended structure of the protein, the predicted fold brings the domains together in space, forming the suspected protein cleft.

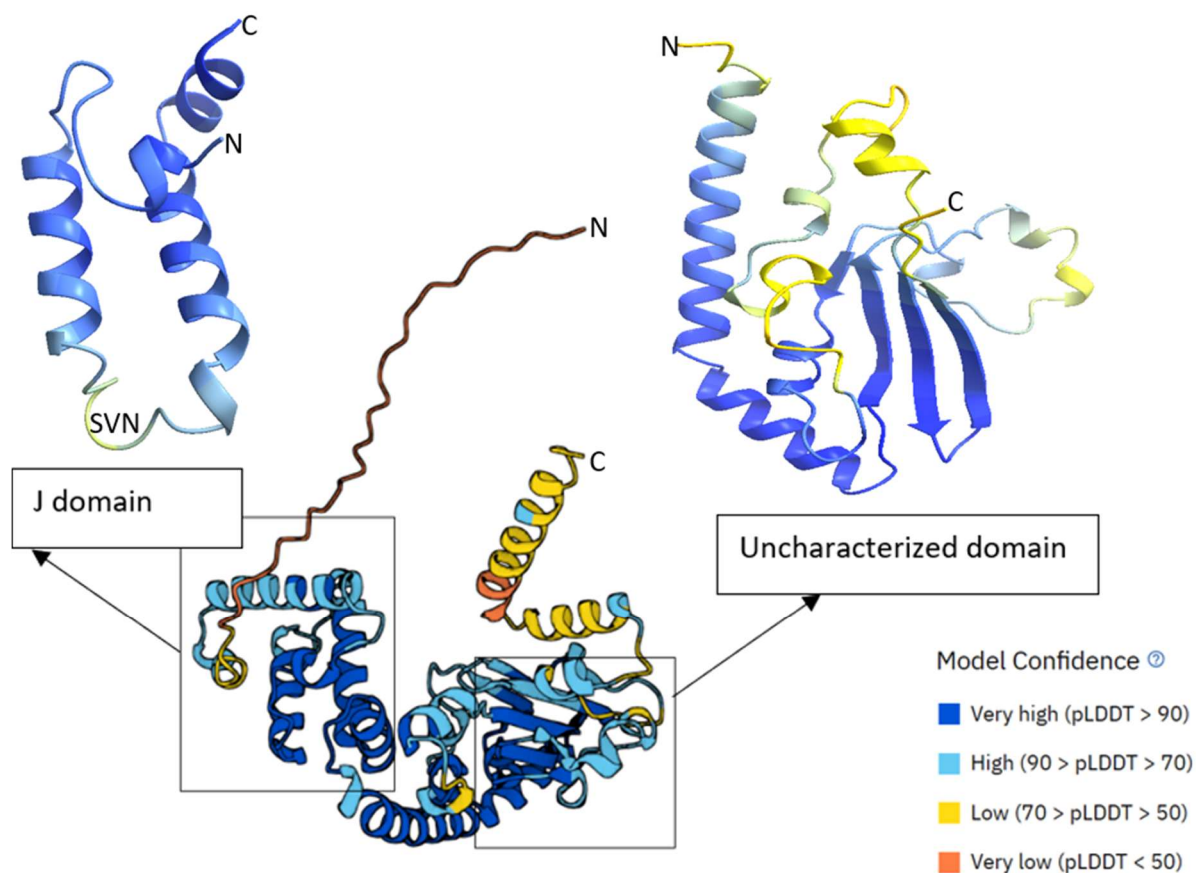


Figure 3.5. The tertiary structure of PFF1010c from AlphaFold

The J domain and the uncharacterized domain are boxed and zoomed in to reflect their position and their fold. The PFF1010c J domain is defined by the 4 α -helices and the uncharacterized C-terminal domain at the carboxy terminus is defined by the β -strands. Peptide start and end are denoted with N for the amino terminus and C for the carboxyl terminus. SVN is a motif on the protein loop behind it. AlphaFold produces a per-residue model confidence score (pLDDT) between 0 and 100 as shown in the model confidence key.

The J domain is constituted by 64 residues from proline positioned at 68 to asparagine positioned at 132 as in section 3.1.4. The position of the SVN motif in the J domain was reflected by placing the 'SVN' on the protein loop. The predicted J domain of PFF1010c was similar to the crystalized *E. coli* DnaJ, placing the SVN motif on the loop that would have had the HPD tripeptide in the J domain of DnaJ (Barends *et al.*, 2013). The C-terminal, uncharacterized domain spanned 162 residues, from leucine positioned at 134 to tyrosine positioned at 296 (section 3.1.4, Figure 3.5). AlphaFold structure is more confident with the folds of inner residues than the terminal residues (Ruff & Pappu, 2021). The overall predicted PFF1010c structure had a high confidence score (pLDDT), except for some residues at the extremities. The SVN residues were poorly predicted, however, situated at the expected position in the J domain relative to *E. coli* DnaJ.

3.1.6. Homo-oligomerization of PFF1010c

Amino acid sequence of the proteins PFF1010c (PF3D7_0620700), PfHsp40 (PF3D7_1437900) and PfHop (PF3D7_1434300) were submitted to the GalaxyHomomer server. The server started with sequence similarity search, structure similarity modelling, and ab initio docking until 5 models are generated. For model 1, automatically detected less reliable loop or terminus regions (called ULR) were re-modelled, and the overall oligomer structure relaxed using GalaxyRefineComplex (Baek *et al.*, 2017). The five predicted PFF1010c homo-oligomers, including the negative and positive controls are shown below (Figure 3.6).

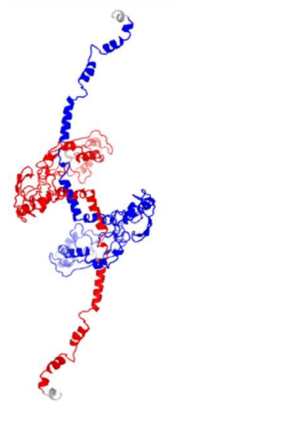
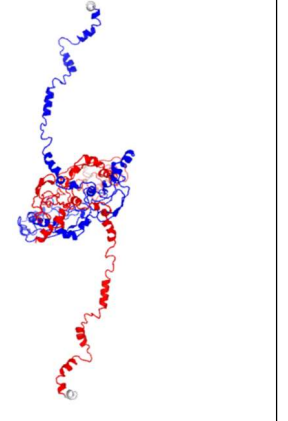
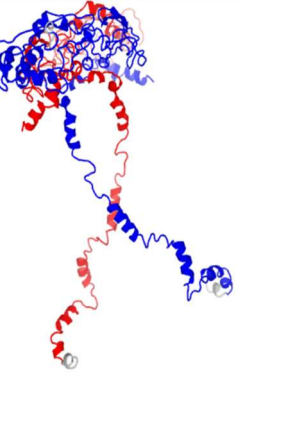
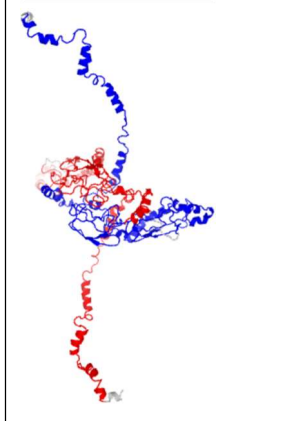
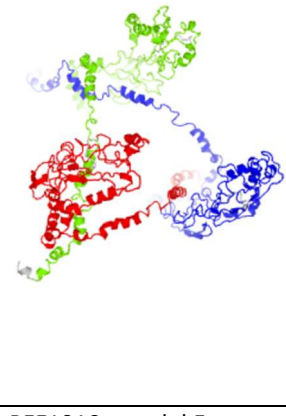
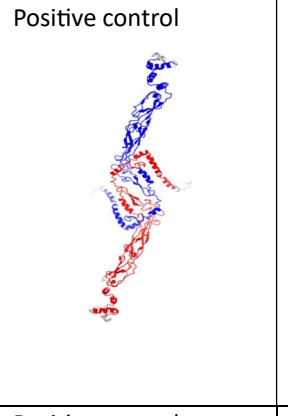
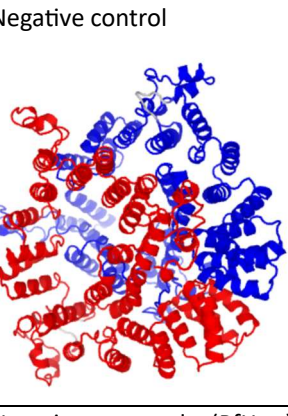
			
PFF1010c model 1	PFF1010c model 2	PFF1010c model 3	PFF1010c model 4
	Positive control 	Negative control 	
PFF1010c model 5	Positive control (PfHsp40) model 1	Negative control (PfHop) model 1	

Figure 3.6. Five predicted models of PFF1010c oligomers, including the positive control (PfHsp40) and the negative control (PfHop)

The blue region represents one chain while red represents another chain of PFF1010c in the homodimer. The light green represents the third chain in the homotrimer. Below the modelled oligomers are the name of the proteins predicted to oligomerize by GalaxyHomomer.

PFF1010c predicted homodimer models (4 models) seemed to be all highly likely to occur since their docking score is greater than the positive control. The greater the docking score, the higher the chances for the homodimer to occur (Table 3.5). Model 5 produced a lower docking score, even lower than the negative control, suggesting that a homotrimer of PFF1010c is highly unlikely to occur. The negative control produced a low docking score, relative to the positive control. The positive control high docking score relative to the negative control, concurred with the experimental X-ray light scattering by Anas *et al.* (2020) on PfHsp40 (PF3D7_1437900) dimerization (Anas *et al.*, 2020). As such, the incorporation of controls validated the results and the predicting tool. GalaxyHomomer agreed

with experimental findings that PfHsp40 (PF3D7_1437900) forms a dimer while PfHop does not (Anas *et al.*, 2020; Makumire *et al.*, 2020; Shin *et al.*, 2014).

Table 3.5. Number of subunits, interface area and docking score for each of the 5 models of predicted oligomers of PFF1010c

Protein name	Model no.	Number of subunits	Interface area	Docking score
PFF1010c	1	2-mer	3469.3	2172.493
	2	2-mer	3772.7	1980.282
	3	2-mer	3085.1	1919.280
	4	2-mer	2919.5	1818.585
	5	3-mer	3820.3	1153.085
Positive control (PfHsp40)	1	2-mer	4354.2	1563.541
Negative control (PfHop)	1	2-mer	4759.0	1309.478

Of note with the oligomerization prediction is that the PFF1010c structure obtained from the homo-oligomer server is an entirely different structure from the AlphaFold (Section 3.1.5.2). The difference could result from the torsional strain the structure experiences during oligomerization (Liu, 2015). However, the GalaxyHomomer protein tertiary structures cannot be superimposed on the AlphaFold predicted structure.

3.1.7. Predicted PFF1010c cellular localization

On the PlasmoDB website, PFF1010c was archived as a protein associated with the mitochondria and predicted to be involved in cristae formation. PSORT II server predicts that PFF1010c is likely to localize to the cytoplasm (16,5 %), nucleoplasm (13 %), nucleus (6,5 %), extracellular (6 %), cytoskeleton (2 %) and plasma membrane (1 %) in the chronological order of decreasing likelihood. The percentage in the brackets is the localization score. SCLpred-EMS was 88 % confident that PFF1010c does not localize to part of the endomembrane system or the secretory pathway. BUSCA subcellular location predictor produced the most reliable score of 70 % that PFF1010c localizes to the cytoplasm. The alternative localization to that of the cytoplasm with a score of 30 % was the nucleus. From this data, PFF1010c would localize to the cytoplasm.

3.1.8. STRING tool predicted the interactome of PFF1010c

Out of the 10 possible associators, STRING predicted 7 hypothetical proteins and 3 characterized proteins (Figure 3.7). With regards to known proteins, PFF1010c is likely to be associated with the ER resident PfHsp70-2 (PFI0875w) than to the cytoplasm resident PfHsp70-1 (PF08_0054). These two possible associates are Hsp70s and the interaction likelihood score was almost the same (Figure 3.7). PFF1010c was predicted to localize to the cytoplasm (section 3.1.7). As such, PFF1010c is likely to interact with PfHsp70-1 than PfHsp70-2. Interestingly, the last characterized possible associate was the PfHsp40 (PF3D7_1437900) type 1, a DnaJ homologue of PFF1010c. This suggests possible interaction of PfHsp40 (PF3D7_1437900) with PFF1010c. PfHsp40 (PF3D7_1437900) is a known cytosolic protein (Rug & Maier, 2011). PFF1010c predicted interaction with PfHsp40 (PF3D7_1437900) produced a lower interaction likelihood score than PfHsp70-1, suggesting that PFF1010c interacts with PfHsp70-1.

Your Input:

PFF1010c Conserved hypothetical protein; DNAJ protein, putative (362 aa)

Predicted Functional Partners:

	Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
PFE1105c Conserved hypothetical protein; Uncharacterized protein				●					0.783
PF07_0033 Hypothetical protein; Cg4 protein; Belongs to the heat shock protein 70 family	●			●	●	●			0.766
PFD0487c Conserved hypothetical protein; Actin-like protein, putative; Belongs to the actin family				●	●	●			0.740
PFD0530c Conserved hypothetical protein; Uncharacterized protein	●			●	●				0.727
PFI0875w Endoplasmic reticulum chaperone b1p; Chaperone DnaK; 78 kDa glucose-regulated protein homolog; Belongs to the heat sh...	●			●	●	●			0.727
PF14_0160 Conserved hypothetical protein; Uncharacterized protein				●	●	●			0.724
PF08_0054 Heat shock 70kda protein 1/2/6/8; Belongs to the heat shock protein 70 family	●			●	●	●			0.723
PF14_0359 DnaJ homolog subfamily a member 2; Uncharacterized protein; HSP40, subfamily A, putative				●	●	●			0.721
PFI1560c Conserved hypothetical protein; Uncharacterized protein	●			●					0.720
PF11_0243 Conserved hypothetical protein; Leucine-rich repeat protein 11				●		●			0.715

Figure 3.7. Predicted functional associates of PFF1010c using STRING

The predicted partners are given using their identifiers and a short description of the protein. The partners association with PFF1010c is shown on the far right. A dot represents how the protein association was determined and the dot scale (greyish to black, with black indicating greater scoring attached to that type of association). The scoring indicates level of confidence and not the strength or specificity of the interaction.

3.1.9. Predicted posttranslational modification (PTM) sites

MusiteDeep only predicted the phosphorylation type of PTM on PFF1010c. After applying a cutoff value 0,5 (default cut-off), four phosphorylation events were predicted along the entire PFF1010c protein. Two phosphorylation events are predicted at the N-terminal residues threonine positioned at 3 (PTM score = 57,8 %) and serine positioned at 22 (PTM score = 60 %). The last two phosphorylation events were predicted with high likelihood at the C-terminal serine positioned at 308 (PTM score = 85 %) and serine positioned at 312 (PTM score = 87,7 %). When taking down the PTM score cutoff value,

more threonine and serine residues are likely to be phosphorylated. No other PTM was predicted besides phosphorylation.

3.1.10. Physicochemical properties

The physicochemical properties of PFF1010c are summarised below (Table 3.6). With an estimated isoelectric point of 8.66, the 362 amino acid residue protein would migrate to the basic region of the pH gradient and around 43 kDa during 2D gel electrophoresis. Prior knowledge of the protein isoelectric point is required when preparing buffers for the protein to migrate in the gel (Magdeldin *et al.*, 2014). The half-life of PFF1010c in an *E. coli* cell was predicted to be 10 hours. Hence, allowing expression of PFF1010c beyond 20 hours in an *E. coli* expression system would possibly lead to low yield. PFF1010c was predicted to be unstable *in vitro* (Instability index, 57.46). Proteins with an instability index greater than 40 are generally regarded as unstable (Guruprasad *et al.*, 1990). Nonetheless, PFF1010c's higher content of aliphatic residues (aliphatic index 94.23) renders it thermostable (Ikai, 1980). Ikai (1980) related the aliphatic index to globulin thermostability and reported that a high aliphatic index indicates globular protein thermostability. Surprisingly, PFF1010c was estimated to be slightly hydrophilic (Gravy value -0.470), despite its apparently high aliphatic index. However, PFF1010c might assume a conformation that renders it highly polar or non-polar, concealing specific amino acid residues beside the individual residue's polarity and position in the protein's primary structure.

Table 3.6. Expasy's Protparam computed PFF1010c physicochemical properties

Parameters	Computed value
Number of amino acids	362
Molecular weight	42792.30
Theoretical pI	8.66
Half-life (<i>E. coli</i> , in vivo)	10 hours
Instability Index	57.46
Aliphatic index	94.23
Grand average of hydropathicity (GRAVY)	-0.470

Notably, *in silico* data is generally predictive but not necessarily decisive in value. However, *in silico* data provides the basis for experimental analysis of the protein.

3.2. Biochemical and biophysical assays

3.2.1. pQE30/PFF1010c plasmid confirmation

A sample of plasmid DNA after extraction assumes generally three distinct conformations that can be resolved using agarose gel electrophoresis with lower percentage than 0.8 % (Figure 3.8.A) (Maayah *et al.*, 2022). The three conformations are namely the closed circular plasmid, open- circular plasmid and the linear plasmid. The DNA sample in the uncut lane was not as intense as the single digests because the plasmid DNA sample was distributed into the above-mentioned conformations. A single restriction endonuclease cleavage linearized the whole fraction of plasmids, resulting in the intense bands which migrated at 4500 bp as expected. The double digest releasing the *PFF1010c* gene out of the *pQE30* plasmid left a fraction of the intact recombinant plasmid *pQE30/PFF1010c*. Despite this, the band representing *PFF1010c* insert was confirmed as it corresponds to the size of the *PFF1010c* DNA inserted in the *pQE30* vector in Figure 3.8.B. The plasmid map (Figure 3.8.B) reflects the size and location of *PFF1010c* inserted, including the selection markers (*AmpR*, ampicillin resistance gene and *CmR*, chloramphenicol resistance gene) and the origin of replication (*ori*). *PFF1010c* was inserted within the *Bam*HI and *Hind*III site located in the multiple cloning site of the *pQE30* vector. The *pQE30/PFF1010c* harbouring transformants were selected using ampicillin in growth media.

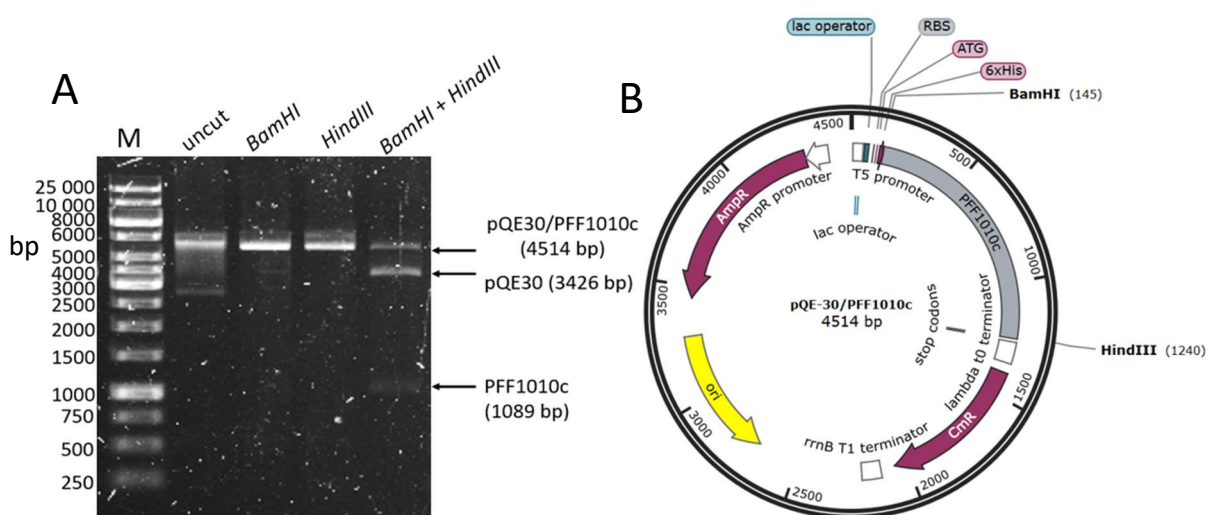


Figure 3.8. Restriction endonuclease cleaving of *pQE30/PFF1010c*

(A) 0.8 % agarose electrophoresis gel depicting the migration and separation of DNA fragments generated from digesting *pQE30/PFF1010c*. Lane M is the molecular marker lane, uncut lane is the *pQE30/PFF1010c* uncut, *Bam*HI is single digest with the enzyme, *Hind*III is digestion with *Hind*III, and lane *Bam*HI + *Hind*III represents DNA digested with both enzymes. Plasmid DNA fragments labelled using arrows and the estimated sizes in brackets. The molecular marker is labelled in base pairs (bp). (B) The *pQE30/PFF1010c* plasmid map highlighting the restriction sites and the plasmid features.

3.2.2. PFF1010c expression and solubility in *E. coli* XL1 Blue

The successful expression of recombinant PFF1010c in *E. coli* XL1 Blue is shown (Figure 3.9.A). The expression appeared to peak after 3 hours of expression. The decline after 3 hours of expression was backed up by the predicted half-life (Table 3.4). The expressed protein was confirmed by Western blot analysis (Figure 3.9.A, lower panel), using α -PFF1010c antibodies. PFF1010c occurred as an insoluble complex in *E. coli* cells (Figure 3.9.B). PFF1010c thus had to be solubilized using a detergent to recover it through the IMAC method of purification.

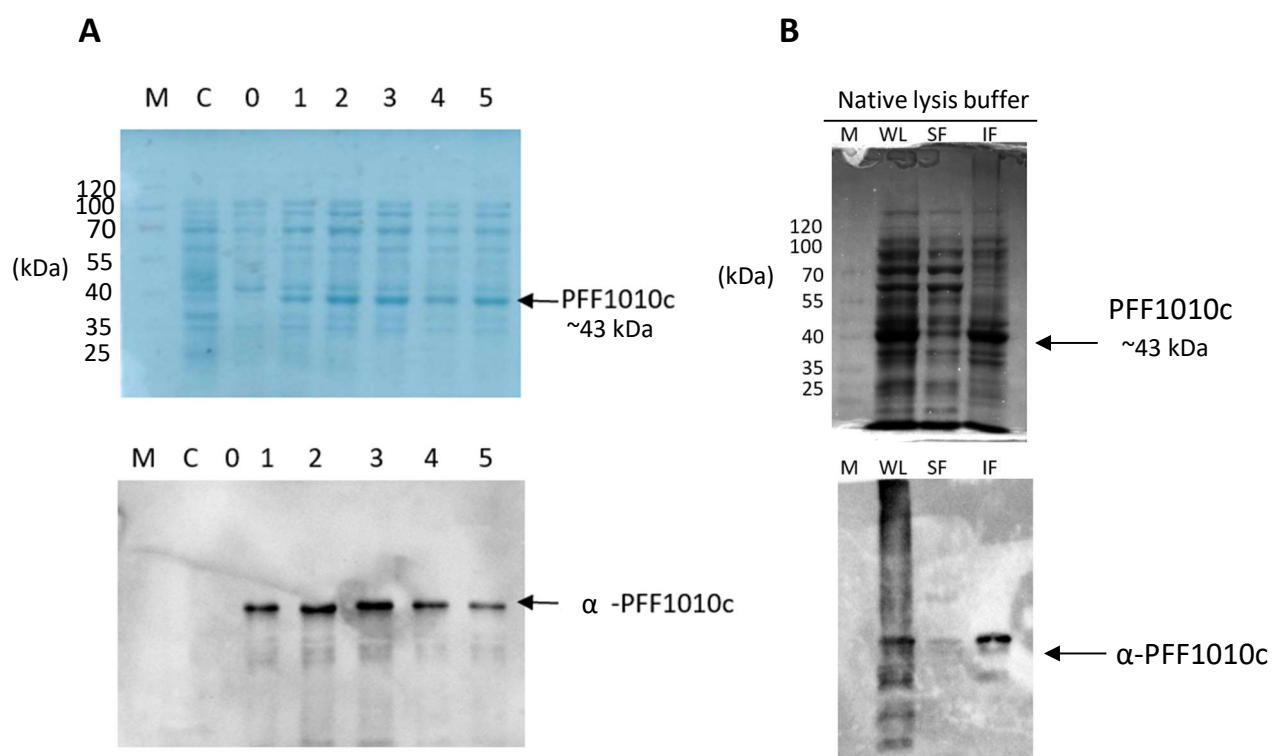


Figure 3.9. Expression and solubility assessments of PFF1010c

(A) Expression of PFF1010c. The top panel is the Coomassie Brilliant blue stained SDS-PAGE gel loaded with hourly expressions (0 hour to 5 hours) of PFF1010c by transformed *E. coli* XL1 Blue cells. The control sample was XL1 Blue cells transformed with the *pQE30* plasmid used as a vector for the *PFF1010c* gene. The bottom panel depicts the Western blot obtained using α -PFF1010c antibodies. The Protein marker used in A was PageRuler Prestained Protein Ladder, catalogue number 26616, and 2 μ L was loaded. (B) Solubility assessment of PFF1010c. The top panel depicts the SDS-PAGE gel loaded with M- the molecular marker, WL- the whole lysate of *E. coli* XL1 Blue cells used to express PFF1010c, SF- the soluble fraction and IF- the insoluble fraction. The gel was stained with Coomassie Brilliant blue. The bottom panel is the accompanying Western blot for the upper gel to identify the protein PFF1010c.

3.2.3. PFF1010c solubilization and purification

The 8 M Urea detergent solubilized much of the protein (Figure 3.10). However, not much of the protein was recovered during purification. The protein yield was only 0,447 mg/ml. A significant amount of the protein was lost in the flow through and the washes (Figure 3.10.B). PFF1010c appeared to be not binding strongly to the IMAC column. Using less stringent buffers such as wash buffers with less than 80 mM imidazole retained contaminants that appeared with the eluted protein (Appendix C3, Figure C3). Panel C shows the Western blot membrane obtained from the purification SDS-PAGE gel of the samples used in B. Upper faint bands indicate PFF1010c was detected in regions around 70 kDa (Figure 3.10.C). This suggested that PFF1010c may have still be interacting with another protein to form a higher order structure. The suspected PFF1010c complex resists denaturing SDS-PAGE conditions and migrate to the 70 kDa region of the SDS-PAGE gel. Another possible explanation for detecting PFF1010c around the 70 kDa band is dimerization of PFF1010c as predicted in section 3.1.6. Dimerization or PFF1010c-protein complexes, both justify the low concentration yield of the PFF1010c. The protein concentration could have been buffered by the formation of either oligomers or PFF1010c-protein complexes, sequestering the polyhistidine tag.

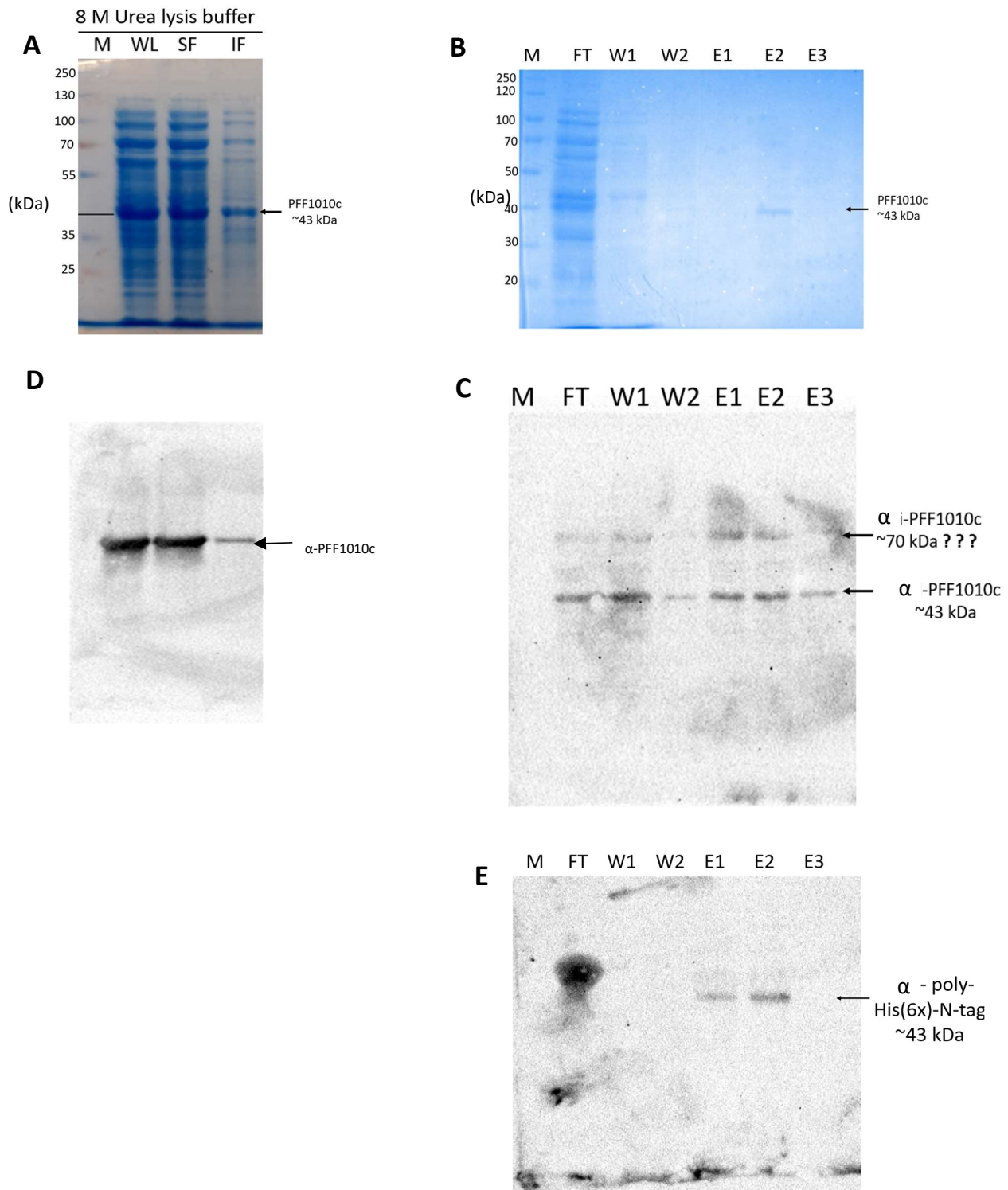


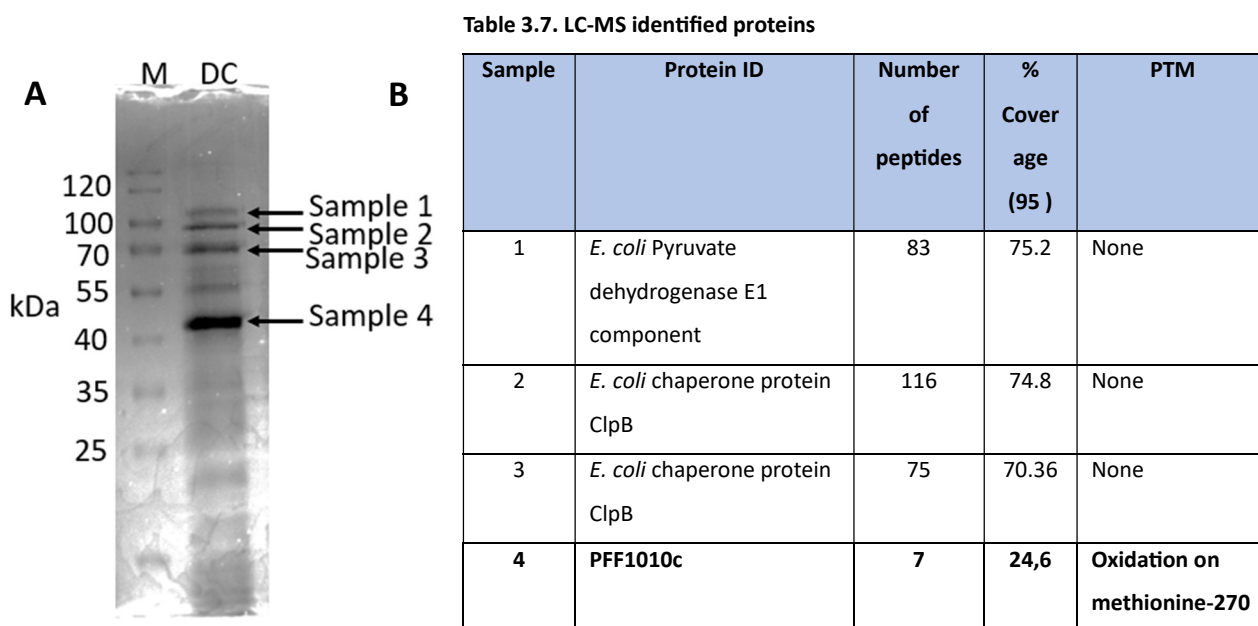
Figure 3.10. Solubilization of PFF1010c by 8 M urea and its purification by Nickel affinity chromatography

(A) SDS-PAGE gel showing solubilization of PFF1010c when 8 M urea was added to the lysis buffer. The labels follow the same protocol as the labels used for solubility assessment in Figure 3.8. (B) SDS-PAGE gel for PFF1010c purification. FT is the flow through of the lysate on the IMAC column, W1 and W2 are column washes and E1, E2, and E3 are the elutions of the protein from the IMAC column. Washes and elutions contained increasing concentration of imidazole as stipulated in the method.

Western blot analysis for PFF1010c solubilization and purification using (C-D) using α -PFF1010c antibodies and (E) α -polyhistidine antibodies.

3.2.4. LC-MS analysis of purified PFF1010c

The SDS-PAGE gel of the purified, concentrated, and dialyzed PFF1010c is shown below (Figure 3.11). Multiple bands became visible, suggesting that either PFF1010c eluted with other proteins or formed oligomers. LC-MS was conducted to identify proteins represented by the various bands observed on the SDS-PAGE that associated with PFF1010c. Each of the prominent bands were cut out and sampled (sample 1-4). PFF1010c was identified only in sample 4 (Table 3.5). This means PFF1010c was not identified in the other bands that eluted with PFF1010c during purification. PFF1010c could have been present in the background at low concentrations that could not be detected by LC-MS. 7 tryptic peptides of PFF1010c were identified at 95 % confidence level in sample 4. One of the peptides showed oxidation on the methionine-270 residue. Methionine is oxidized to methionine sulfoxide (Liang *et al.*, 2012). PFF1010c was not detected in other samples (SDS-PAGE gel bands). These results suggest that PFF1010c did not form or retain dimers after denaturation during SDS-PAGE gel electrophoresis.



The most abundant protein was included in Table 3.5 for sample 1, 2, and 3 since the samples did not contain PFF1010c. Interestingly, PFF1010c co-eluted with two ClpB proteins and one pyruvate dehydrogenase E1 component of the *E. coli* XL1 Blue cells. As described earlier (section 1.2.3.), ClpB proteins are heat shock proteins and belong to the Hsp100 superfamily. While PFF1010c is expected to interact with Hsp70 homologues (section 3.1.8), These results suggest that PFF1010c is more likely to interact with Hsp100 homologues than Hsp70 homologues. Also, PFF1010c was predicted to be involved in cristae formation of the mitochondria (section 3.1.7). Pyruvate is oxidized in the mitochondria in eukaryotic cells by pyruvate dehydrogenase (Zangari *et al.*, 2020). Whether PFF1010c and pyruvate dehydrogenase interact within the parasite cell remains unknown. It is also possible that the co-elution of PFF1010c with some of the *E. coli* proteins reported here was random and may not mirror functional pathways that exists within the parasite.

3.2.5. Biophysical characterization

3.2.5.1. Secondary structure analysis of PFF1010c

The secondary structure estimated modules of PFF010c, and their respective percentage content are shown below (Table 3.6). BeStSel webserver groups 3_{10} helices in the ‘others’ conformational state group. The ‘others’ contain conformational states that BeStSel does not recognize as well-defined β -sheets, α -helices and turns. Since no regular α -helices and distorted α -helices were recognized from the PFF1010c CD spectra, the helices detected could all be 3_{10} helices. Typical α -helices characteristics include 3.6 residues per turn with consecutive residues making an angle of 100° around the helical axis, a helical rise per residue of $1.5 \text{ \AA}$, and a helical pitch of $5.4 \text{ \AA}$ (Vieira-Pires & Morais-Cabral, 2010). On the contrary, 3_{10} helices have residues per turn with 120° between consecutive residues, a helical rise per residue of $1.93\text{--}2.0 \text{ \AA}$, and a helical pitch of $5.8\text{--}6 \text{ \AA}$ (Vieira-Pires & Morais-Cabral, 2010). The purified PFF1010c was constituted mostly of β -sheets which BeStSel subdivided into antiparallel relaxed (3.3 %), antiparallel right twisted (22.6 %) and parallel (15.6 %) β -sheets. The 3_{10} helices and other random coils were grouped as ‘other’ secondary structure conformational states and occupied 16.1 % of the protein. The fitted and raw data chart is attached in Appendix D4. The NRMSD value (0.188) was slightly greater than 0.1.

Table 3.8. PFF1010c CD spectroscopy estimated secondary structure content (%)

Secondary structure conformational state	Content (%)
Regular α -helices	0.0
Distorted α -helices	0.0
Antiparallel left-twisted β -sheets	0.0
Antiparallel relaxed β -sheets	3.3
Antiparallel right-twisted β -sheets	22.6
Parallel β -sheets	15.6
Turns	0.0
Others	16.1

BeStSel webserver fold recognition algorithm generates a list of domains on CATH domain database and limit the list to 10 domains with 90 % sequence homology to the generated CD spectra estimated secondary structure. The predicted domain structure can be accessed on PDB and CATH using the PDB identity (ID) (Table 3.). The letter A in the table denotes the architecture, and T the topology of the domain based on the CATH classification system. Interestingly, BeStSel generated the Hsp40/DnaJ peptide binding domain two times (highlighted) with the sandwich architecture and under the mainly beta class, not the signature DnaJ domain. This could be that the purified PFF1010c folded well around the domain constituted by β -sheets.

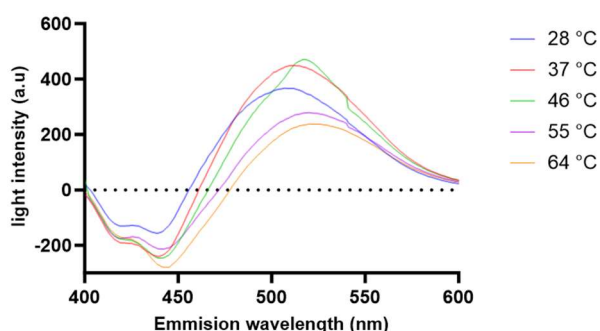
Table 3.9. Closest single CATH domains in a Protein Data Bank subset filtered for structures with 90 % sequence homology

PDB ID	Chain	C	A	T	Class	Architecture	Topology	Length
1. 2BW2	A	3	10	20	A-Beta	Roll	Ubiquitin-like (UB roll)	55
2. 2Q2G	A	2	60	260	Mainly Beta	Sandwich	Hsp40/DnaJ peptide binding domain	81
3.1FEU	A	2	40	240	Mainly Beta	B-Barrel	Ribosomal Protein L25, Chain P	91
4.3ITW	A	3	30	720	A-Beta	2-Layer Sandwich	Signal recognition particle alu RNA binding heterodimer, srp9/1	56
5.3LZ8	B	2	60	260	Mainly Beta	Sandwich	Hsp40/DnaJ peptide-binding domain	89
6.1EU3	A	2	40	50	Mainly Beta	B-Barrel	OB fold (Dihydrolipoamide Acetyltransferase, E2P)	80
7.2PMA	A	2	40	70	Mainly Beta	B-Barrel	Cathepsin D, subunit A, domain 1	121
8.2J4X	A	3	40	50	Mainly Beta	B-Barrel	OB fold (Dihydrolipoamide Acetyltransferase, E2P)	72
9.3SKI	A	2	30	720	A-Beta	2-Layer Sandwich	Signal recognition particle alu RNA binding heterodimer, srp9/1	58
10. 1NYC	B	3	40	310	Mainly Beta	B-Barrel	Staphostatins	111

3.2.5.2. PFF1010c thermostability

ANS extrinsic fluorescence spectra of PFF1010c at different temperatures is shown below (Figure 3.12). The increased fluorescence intensity from 28 °C to 46 °C reflects unfolding of PFF1010c with increasing temperature. That is the protein increasingly exposed its hydrophobic regions which were bound by the ANS probe thus modulating the fluorescence signal (Figure 3.12.A). The fluorescence intensity shifted by 50 % to low intensity when temperature was increased to 55 °C. This could be associated with aggregation, suggesting that the protein may have denatured to form aggregates exhibiting different fluorescence intensity opposed to the folded version as observed at normal temperatures (Zhang *et al.*, 2009). The conformational shifts of PFF1010c at various temperatures are mirrored by Figure 3.12.B. The use of fluoresce spectroscopy in monitoring protein conformational changes has been previously advocated (Baker *et al.*, 2004). From 28 °C to 46 °C, PFF1010c red shifted (emission moved to longer wavelength). However, the red shift signal slowed down from 37 °C to 46 °C, suggesting an approach to a critical temperature where the protein would lose its folded structure and completely denature. The protein fluorescence blue shifted from 46 °C, exposing its hydrophobic patches, a phenomenon associated with denaturation (Kumar *et al.*, 2023). These results (Figure 3.12.) combined with the mentioned definition of a thermostable protein (section 2.4.12.2), suggest that PFF1010c denatures at temperatures higher than 46 °C, rendering PFF1010c thermostable. Also, if PFF1010c was heat denatured as the results suggested, PFF1010c was probably folded on purification.

A ANS extrinsic fluorescence emission spectra of PFF1010c at different temperatures



B ANS extrinsic fluorescence emission peaks of PFF1010c at different

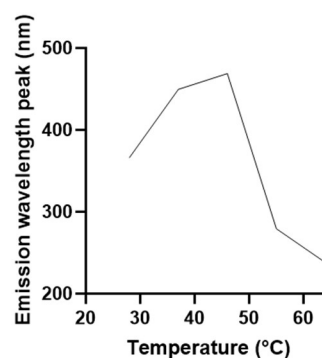


Figure 3.12. ANS extrinsic fluorescence emission of PFF1010c at different temperatures (°C)

(A) ANS extrinsic fluorescence, emission intensity spectra of PFF1010c at different temperatures. The JASCO FP-6300 spectrofluorometer (JASCO, Spain) was blanked with the ANS in PBS buffer without proteins. Light intensity was shown in arbitrary units and emission wavelength in nanometres (nm). Colour labelling is shown on the right of the diagram. PFF1010c

ANS fluorescence increased from 28 °C to 46 °C and went down from 55 °C to 64 °C. (B) Plot of the PFF1010c ANS fluorescence peak wavelengths (nm) against temperature (°C). PFF1010c loses its fold at 46 °C.

3.2.6. Biochemical characterization

3.2.6.1. Luciferase refolding assay

The data obtained from luciferase refolding assay are depicted below (Figure 3.13). PfHsp70-1 sufficiently refolded the denatured luciferase even in the presence of PfHsp40 (PF3D7_1437900). However, PFF1010c appears to decrease the chaperone activity of PfHsp70-1. PfHsp40 (PF3D7_1437900) decreased the chaperone activity of PfHsp70-1, but not as greatly as PFF1010c did. PFF1010c's presence in the refolding mix and separately did not refold the denatured luciferase to produce over 50 % of the luciferase activity. PFF1010c significantly reduced the refolding capacity of PfHsp70-1 when combined with PfHsp40 (PF3D7_1437900), suggesting that PFF1010c inactivates the PfHsp70-1 and PfHsp40 (PF3D7_1437900). The denatured luciferase appeared to have not fully denatured, which explains why PfHsp40 (PF3D7_1437900) and PFF1010c have shown a level of refolding even though the proteins are not refoldases. Also, the native luciferase shown a lower luciferase activity as opposed to PfHsp40 (PF3D7_1437900) and PfHsp70-1 refolded luciferase activity. These could be indeterminate errors accompanying the assay and the results should not be final.

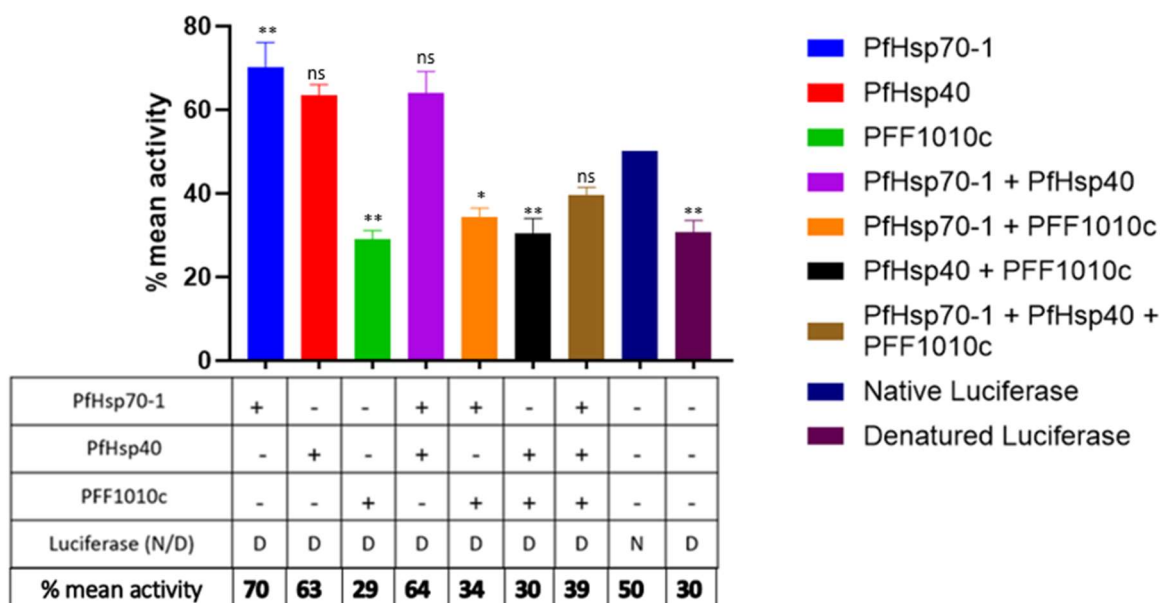


Figure 3.13. Luciferase refolding activity of PfHsp70-1 in the presence of PfHsp40, and PFF1010c

The bar graph depicts the refolding of heat-denatured luciferase. The table below the bar graph indicates the presence (+) and absence of the protein in the row (-). N represents the native luciferase and D for denatured luciferase. The native luciferase

% activity was set at 100. The asteriks indicate the level of significant between the group bar and the native luciferase bar, where *** for p value < 0.001, ** when p value is < 0.01, * when p value is < 0.05, and n.s when p value > 0.05.

Chapter 4

4. Chapter 4: Discussion and Conclusions

4.1. Discussion

The deadliest malaria parasite, *P. falciparum*, expresses the highest complement of JDPs relative to other malaria parasites (Njunge *et al.*, 2013; Phelan *et al.*, 2023). Only two of the thirteen type IV JDPs (PFF1010c and PF11_0443) are not exported to the infected erythrocyte. These two parasite-resident type IV JDPs are putative. A putative protein is annotated to be expressed based on RNA-based expression studies and awaits experimental characterization (Alam & Dwivedi, 2016; Uttarotai *et al.*, 2022). Dutta, Pesce *et al.* (2021) wrote on Zhang *et al.* (2018) saturation-scale mutagenesis screening of *P. falciparum* essential genes, including PFF1010c. PF11_0443 is not essential. This is the first study to report the secondary structural features of PFF1010c and to explore its heat stability.

Homology defined PFF1010c as a member of the JDP family. Based on the HHPRED and BLASTP homologue search, nucleic acid binding and heat stress buffering constitute some of the predicted possible cellular functions of PFF1010c. BLASTP detected immediate homologues with fewer functional differences, allowing the immediate prediction of the PFF1010c function. Based on BLASTP, a probable function can be drawn that PFF1010c is a JDP superfamily heat shock protein involved in a nucleic acid binding function during heat-induced stress. *Plasmodium* female gametocytes carry pools of mRNAs assembled and translationally suppressed by an RNA binding protein complex that is poorly understood (Mair *et al.*, 2010). Perhaps PFF1010c forms part of this complex.

BLASTP also linked PFF1010c to proteins of the apicomplexan parasites, suggesting that Type IV JDPs like PFF1010c play a role in apicomplexan parasite virulence and pathogenicity. Most studies use BLASTP only (Bitacura & Santos, 2022; Dishnica *et al.*, 2023; Lopez *et al.*, 2021; Mooi *et al.*, 2023; Xiao *et al.*, 2021). One study coupled normal BLASTP to PSI-BLAST (Ang'ang'o *et al.*, 2023). HHPRED is similar to PSI-BLAST but more sensitive (Hildebrand *et al.*, 2009). As such, HHPRED was used in this study instead of PSI-BLAST. However, HHPRED did not appear to be the right tool to pull homologues and deduce a function of an uncharacterized protein. HHPRED homologues produced JDPs with varying functions and a wide range of organisms bearing those proteins.

Multiple sequence alignment differentiated PFF1010c from its JDP relatives by a single motif within the J domain. PFF1010c clustered with proteins bearing the SVN/H motif and closely related to proteins bearing the HPD. The HPD motif is well known for modulating the JDP stimulation of Hsp70s ATPase activity (Dutta, Pesce, *et al.*, 2021; Faust *et al.*, 2020; Tomiczek *et al.*, 2020). The SVN/H and SDH role in the J domain remains a mystery. Also, the roles of J domains harbouring these SVN/H and SDH motifs in place of the HPD motif are not known (Botha *et al.*, 2007; Njunge *et al.*, 2013; Shonhai & Blatch, 2021). The SVN/H group appeared to be a separately evolving clade, corroborated by the phylogenetic tree (Section 3.1.3, Figure 3.3).

Multiple sequence alignment results also revealed a highly conserved module at the C-terminus of the BLASTP proteins. The conservation is found in all BLASTP aligned proteins, suggesting a domain equivalent to the J domain in terms of domain definition. The CDD, dubbed the domain as an 'uncharacterized domain' (Section 3.1.4). Cheetham and Caplan (1998) review paper mentioned that the fourth domain at the C-terminus is less conserved and not well understood (Cheetham & Caplan, 1998). A C-terminal fragment of a yeast type I JDP was crystalized and found to be a dimerization module (Wu *et al.*, 2005). Also, the C-terminal domain mutation in a type I human JDP, DjA1, compromised dimerization (Borges *et al.*, 2005). Type 1 DjA1 dimer was formed by the C-terminal facing the N-terminal of the other monomer, while the human type II DjB4 dimer was formed by the C-terminal facing the C-terminal of the other monomer. Borges *et al.* (2005) reflected the dimerization differences between the two subfamilies, type I and type II JDPs. PFF1010c is a type IV JDP with a C-terminal, possibly a dimerization domain. PFF1010c could form dimers predicted in section 3.1.6 and require dimerization for interacting with Hsp70s or substrates like DnaJ and JDPs (Kampinga *et al.*, 2019; Ruger-Herreros *et al.*, 2024; Shi *et al.*, 2005).

The association prediction, STRING, predicted PFF1010c to interact with PfHsp70-2 more strongly than the second possible interactor, PfHsp70-1. The third and least possible associate was found to be PfHsp40 (PF3D7_1437900). PfHsp70-1 and PfHsp40 (PF3D7_1437900) are cytosolic, while PfHsp70-2 is an endoplasmic reticulum (ER) localized protein (Anas *et al.*, 2020; Gladys T. Cortés *et al.*, 2020; Lebepe *et al.*, 2020). Recently, (Mathews *et al.*, 2021) found that 48.7 % of Hsp40 localized to the ER. Perhaps a fraction of PFF1010c localizes to the ER and interacts with ER residents PfHsp70-2 and PfHsp40 (PF3D7_1437900). Concurrently, the cytosolic PFF1010c would interact with the cytosol resident PfHsp70-1 and PfHsp40 (PF3D7_1437900). Matthews *et al.* (2021) pointed out that farnesylation inhibition decreased the ER-localized PfHsp40 (PF3D7_1437900), suggesting that PfHsp40 (PF3D7_1437900) farnesylation channels the protein to the ER (Mathews *et al.*, 2021).

PFF1010c posttranslational modification prediction did not produce a farnesylation (Section 3.1.9). PFF1010c was predicted to not localize to any part of the endomembrane system, including ER. PFF1010c seemed to localize to the cytosol. PlasmoDB website archived the putative bioinformatics of PFF1010c and records PFF1010c as a mitochondrial protein involved in cristae formation. This function and localization do not match the previous online tools used to predict the localization.

When the PFF1010c recombinant protein was overexpressed from the confirmed plasmid pQE30/PFF1010c in *E. coli* XL1 Blue cells, it displayed an insolubility character. PFF1010c migrated to the insoluble (pellet) fraction (Figure 3.9). However, 8 M Urea (Singhvi & Panda, 2022) could solubilize the protein and move it to the soluble fraction of the centrifuged cell lysate overexpressing PFF1010c (Figure 3.10). A Western blot that was generated using the α -PFF1010c antibodies, which were supposed to detect only PFF1010c, detected a species with a band that migrated around 70 kDa. The entity was suggestive of a PFF1010c dimer. While analyses that followed purification used a pure protein, the consistent bands around 70 kDa and 150 kDa after concentrating the protein are not ignorable (Figure 3.11). The SDS-PAGE sample loading buffer contains denaturants that would denature the supposedly homo-oligomers or hetero-oligomers. LC-MS could not resolve the Western blot detected PFF1010c in the other bands besides the 40 kDa SDS-PAGE migration point. Notably, LC-MS relies on reference proteins present in the database (Li & Zhu, 2020), and this constitutes part of its limitations. This may account for the misalignment of findings obtained by Western blot versus the LC-MS data.

Using LC-MS, however, produced a posttranslationally modified species of PFF1010c in the form of an oxidised methionine-270 residue (Table 3.5., Figure 3.11). LC-MS is a reliable tool for the detection of posttranslational modifications (Mann & Jensen, 2003; Virág *et al.*, 2020; Zolg *et al.*, 2018). The resulting methionine sulfoxide alters the function of the protein bearing it (Moskovitz & Smith, 2021). Methionine residue oxidation could be a reactive oxygen species sensor (Chandran & Binniger, 2023). Notably, *E. coli* does not possess posttranslational machinery. However, the oxidation of the methionine residue may have occurred spontaneously and as such the LC-MS finding in this respect does not reflect the complete posttranslational possibilities for this protein. Also, methionine is the most oxidised residue in *E. coli* (Brown *et al.*, 2017). It would have been interesting to study the protein within the native host, *P. falciparum*, cell settings. It is possible that the observed oxidation occurs because PFF1010c acts as a sensor of oxidation. Indeed, heat shock proteins are induced by oxidative stress (Jeyachandran *et al.*, 2023; Lan *et al.*, 2020; Mulyani *et al.*, 2020; Szyller & Bil-Lula, 2021).

Circular dichroism results of PFF1010c did not overlap with the previously predicted structures of PFF1010c. However, experimental results supersede predicted *in silico* data (Petrini *et al.*, 2022; Rajpal *et al.*, 2023). β -sheets are associated with an unfolded segment of a protein. The unfolding of tau proteins in Alzheimer's disease involves the adoption of β -sheets from random coils, favouring protein aggregation (Bergen *et al.*, 2005). Relating tau's aggregation to the β -sheet dominated secondary structure of PFF1010c, perhaps the protein was not sufficiently refolded before the circular dichroism spectroscopy analysis. β -sheets are prerequisites of protein aggregation, specifically amyloid-type protein aggregation (Willbold *et al.*, 2021). Amyloid-type protein aggregation is typical of prions (Willbold *et al.*, 2021). PFF1010c displayed aggregation activity upon thermostability analysis (section 3.4.5.2). It would make sense that PFF1010c binds to denatured proteins during stress and aggregates with them, ultimately leading to cell death.

The distinguishing feature of BeStSel from other circular dichroism spectroscopy analysing tools is the algorithm that relates detected domains from the analysed structure to the characterized domains database. A J domain hit would signal that the dialysis protocol successfully refolded the analysed protein following denaturing purification. The J domain was not detected (section 3.2.5.1). However, two out of the top ten closest CATH domain matches were Hsp40/DnaJ peptide binding domains. The domains were both mainly β -sheets. The β -sheets in PFF1010c are found in the uncharacterized domain. Based on circular dichroism spectra of PFF1010c, the β -sheets constitute a structure like the characterized DnaJ peptide binding domains. It was then speculated in this study that perhaps the PFF1010c maintained a correct fold around the β -sheets. Otherwise, the β -sheets tend to retain their structure during the denaturation step of purification. So far, the reported peptide binding domain was the only experimentally determined structure that suggests that PFF1010c is, indeed a JDP. Originally, β -sheets-rich proteins denature into α -helices upon detergent treatment and are insufficiently refolded (Rasmussen *et al.*, 2020). PFF1010c was denatured and successfully refolded upon detergent removal.

ANS extrinsic fluorescence spectroscopy-based study suggested that PFF1010c is a thermostable molecule. The protein only aggregated at high temperatures (>55 °C, Figure 3.12.A). The aggregation activity concurs with the circular dichroism results, indicating the domination of β -sheets in PFF1010c secondary structure suspected of mediating aggregation, as discussed above. Also, the displayed thermostability agrees with the predicted thermostability (Section 3.1.10). *P. falciparum* gametocytes mature in the bone marrow and spleen cool environments and later enter the warm circulatory system (Casas *et al.*, 2016; Draghici & Taylor, 2021; Usui & Williamson, 2021). It makes sense for the

gametocyte to be armed with a thermostable proteome, including PFF1010c. Typically, the activity of luciferase drops at temperatures above 25 °C (Kargar *et al.*, 2020). Most luciferase refolding assay protocols denature luciferase enzyme by heating it at 42 °C (Hu *et al.*, 2023; Li *et al.*, 2021; Morrow *et al.*, 2006). PFF1010c maintained its fold at 42 °C and only started losing it at 46 °C. The thermotolerance investigation in this study was an endpoint study where fluorescence was measured at different temperature points. It is intriguing what would appear between 46 °C and 55 °C temperatures. Monitoring the fluorescence peaks between 46 °C and 55 °C could indicate that PFF1010c starts to denature at a higher temperature than the determined 46 °C.

Luciferase refolding assay was conducted to explore the effect of PFF1010c on the refolding activity of PfHsp70-1. STRING analysis predicted PFF1010c to associate with PfHsp40 (PF3D7_1437900), PfHsp70-1, and PfHsp70-2. However, the direct interaction of PfHsp70-2 could not be experimentally ascertained due to technical limitations. PFF1010c appeared to abrogate the chaperone activity of PfHsp70-1 (section 3.2.6.1). PFF1010c alone did not display chaperone activity. The results were inconclusive since the denatured luciferase retained much of its activity after heat denaturation. The 42 °C used to denature the native luciferase was not adequate. However, temperatures above 42 °C can denature luciferase beyond a possibility of renaturation by the chaperones in the study. As such, the luciferase refolding assay protocol needs to be revised. Using a similar approach, Chakafana *et al.* (2021) demonstrated the refolding activity of PfHsp70-1 in cooperation with PfHsp70-z and PfHsp40 (PF3D7_1437900) (Chakafana *et al.*, 2021). Notably, in the study by Chakafana *et al.* (2021), it was demonstrated that PfHsp40 (PF3D7_1437900) and PfHsp70-1 independently refolded the denatured luciferase to less than 10 % activity. However, in the current study, it appears that the setback was that luciferase was not fully inactivated as it only attained 30 % denaturation (Figure 3.13). A follow up refolding study is therefore important to conduct.

4.2. Conclusions

This study confirmed that PFF1010c is a JDP heat shock protein. The utility of the predictive *in silico* tools was demonstrated with the sensitivity of normal BLASTP at detecting recently archived JDP homologues. To a certain extent, PFF1010c domains were accurately predicted, with the online tools being in concordance with each other. Furthermore, experimental studies were used to explore the structure-function features of PFF1010c. In this regard, PFF1010c was shown to be thermostable and appeared to form dimers or readily aggregate through the β -sheet-dominated modules. However, PFF1010c appeared to lack independent chaperone activity. Further studies need to confirm and

characterize the tertiary structure of PFF1010c through X-ray crystallography. Also, further biochemical characterization of PFF1010c can shed light on the biochemistry and function of SVN/H containing JDPs. Studies could go further to investigate the essential function PFF1010c confers to *P. falciparum* during the gametocyte stage.

5. References

1. Acharya, P., Kumar, R., & Tatu, U. (2007). Chaperoning a cellular upheaval in malaria: Heat shock proteins in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, 153(2), 85-94.
2. Ahmed, M., & Cox-Singh, J. (2015). *Plasmodium knowlesi*—an emerging pathogen. ISBT science series, 10(S1), 134-140.
3. Al-Fatlawi, A., Menzel, M., & Schroeder, M. (2023). Is Protein BLAST a thing of the past? *Nature Communications*, 14(1), 8195.
4. Alagar Boopathy, L. R., Jacob-Tomas, S., Alecki, C., & Vera, M. (2022). Mechanisms tailoring the expression of heat shock proteins to proteostasis challenges. *Journal of Biological Chemistry*, 298(5), 101796.
5. Alam, S. I., & Dwivedi, P. (2016). Putative function of hypothetical proteins expressed by *Clostridium perfringens* type A strains and their protective efficacy in mouse model. *Infect Genet Evol*, 44, 147-156.
6. Anas, M., Shukla, A., Tripathi, A., Kumari, V., Prakash, C., Nag, P., Kumar, L. S., Sharma, S. K., Ramachandran, R., & Kumar, N. (2020). Structural–functional diversity of malaria parasite's PfHSP70-1 and PfHSP40 chaperone pair gives an edge over human orthologs in chaperone-assisted protein folding. *Biochemical Journal*, 477(18), 3625-3643.
7. Ang'ang'o, L. M., Herren, J. K., & Tastan Bishop, Ö. (2023). Structural and Functional Annotation of Hypothetical Proteins from the Microsporidia Species *Vittaforma corneae* ATCC 50505 Using in silico Approaches. *International Journal of Molecular Sciences*, 24(4), 3507.
8. Angamuthu, K., & Piramanayagam, S. (2017). Evaluation of in silico protein secondary structure prediction methods by employing statistical techniques. *Biomedical and Biotechnology Research Journal (BBRJ)*, 1(1), 29-36.
9. Aurecochea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., & Harb, O. S. (2009). PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research*, 37(suppl_1), D539-D543.
10. Aurecochea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Miller, J. A., Nayak, V., Pennington, C., Wang, H. (2008). PlasmoDB: a

- functional genomic database for malaria parasites. *Nucleic Acids Research*, 37(suppl_1), D539-D543.
11. Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, 234(4773), 179-186.
 12. Baek, M., Park, T., Heo, L., Park, C., & Seok, C. (2017). GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure. *Nucleic Acids Research*, 45(W1), W320-W324.
 13. Baker, S. N., McCleskey, T. M., Pandey, S., & Baker, G. A. (2004). Fluorescence studies of protein thermostability in ionic liquids. *Chemical Communications*(8), 940-941.
 14. Barends, T. R. M., Brosi, R. W. W., Steinmetz, A., Scherer, A., Hartmann, E., Eschenbach, J., Lorenz, T., Seidel, R., Shoeman, R. L., Zimmermann, S., Bittl, R., Schlichting, I., & Reinstein, J. (2013). Combining crystallography and EPR: crystal and solution structures of the multidomain cochaperone DnaJ. *Acta Crystallographica Section D*, 69(8), 1540-1552.
 15. Barth, J., Schach, T., & Przyborski, J. M. (2022). HSP70 and their co-chaperones in the human malaria parasite *P. falciparum* and their potential as drug targets [Mini Review]. *Frontiers in Molecular Biosciences*, 9.
 16. Basquin, J., Taschner, M., & Lorentzen, E. (2016). Complex Reconstitution from Individual Protein Modules. In M. C. Vega (Ed.), *Advanced Technologies for Protein Complex Production and Characterization* (pp. 305-314). *Springer International Publishing*.
 17. Bastiaan-Net, S., Pina-Pérez, M. C., Dekkers, B. J. W., Westphal, A. H., America, A. H. P., Ariëns, R. M. C., de Jong, N. W., Wichers, H. J., & Mes, J. J. (2020). Identification and in silico bioinformatics analysis of PR10 proteins in cashew nut. *Protein Sci*, 29(7), 1581-1595.
 18. Beck-Johnson, L. M., Nelson, W. A., Paaijmans, K. P., Read, A. F., Thomas, M. B., & Bjørnstad, O. N. (2017). The importance of temperature fluctuations in understanding mosquito population dynamics and malaria risk. *Royal Society Open Science*, 4(3), 160969.
 19. Bitacura, J. G., & Santos, M. D. (2022). In Silico Structural Analysis, Classification, and Functional Annotation of an Uncharacterized Protein from an Aquatic Fungus *Lindgomyces ingoldianus*. *Genetics of Aquatic Organisms*, 7(1).
 20. Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., Frutiger, S., & Hochstrasser, D. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*, 14(10), 1023-1031.

21. Blatch, G. L. (2022). *Plasmodium falciparum* Molecular Chaperones: Guardians of the Malaria Parasite Proteome and Renovators of the Host Proteome [Mini Review]. *Frontiers in Cell and Developmental Biology*, 10.
22. Borges, J. C., Fischer, H., Craievich, A. F., & Ramos, C. H. (2005). Low resolution structural study of two human Hsp40 chaperones in solution: DjA1 from subfamily A and DjB4 from subfamily B have different quaternary structures. *Journal of Biological Chemistry*, 280(14), 13671-13681.
23. Botha, M., Chiang, A. N., Needham, P. G., Stephens, L. L., Hoppe, H. C., Külzer, S., Przyborski, J. M., Lingelbach, K., Wipf, P., Brodsky, J. L., Shonhai, A., & Blatch, G. L. (2011). *Plasmodium falciparum* encodes a single cytosolic type I Hsp40 that functionally interacts with Hsp70 and is upregulated by heat shock. *Cell Stress and Chaperones*, 16(4), 389-401.
24. Botha, M., Pesce, E.-R., & Blatch, G. L. (2007). The Hsp40 proteins of *Plasmodium falciparum* and other Apicomplexa: regulating chaperone power in the parasite and the host. *The International Journal of Biochemistry & Cell Biology*, 39(10), 1781-1803.
25. Bracher, A., & Verghese, J. (2022). Nucleotide Exchange Factors for Hsp70 Molecular Chaperones: GrpE, Hsp110/Grp170, HspBP1/Sil1, and BAG Domain Proteins. *The Networking of Chaperones by Co-Chaperones*, 1-39.
26. Brown, C. W., Sridhara, V., Boutz, D. R., Person, M. D., Marcotte, E. M., Barrick, J. E., & Wilke, C. O. (2017). Large-scale analysis of post-translational modifications in *E. coli* under glucose-limiting conditions. *Bmc Genomics*, 18, 1-21.
27. Buchan, D. W., & Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Research*, 47(W1), W402-W407.
28. Buchner, J., Grallert, H., & Jakob, U. (1998). [27] Analysis of chaperone function using citrate synthase as nonnative substrate protein. *Methods In Enzymology* (Vol. 290, pp. 323-338). Elsevier.
29. Calderaro, A., Piccolo, G., Gorrini, C., Rossi, S., Montecchini, S., Dell'Anna, M. L., De Conto, F., Medici, M. C., Chezzi, C., & Arcangeletti, M. C. (2013). Accurate identification of the six human *Plasmodium* spp. causing imported malaria, including *Plasmodium ovale wallikeri* and *Plasmodium knowlesi*. *Malaria Journal*, 12, 1-6.
30. Carra, S., Alberti, S., Benesch, J. L., Boelens, W., Buchner, J., Carver, J. A., Cecconi, C., Ecroyd, H., Gusev, N., & Hightower, L. E. (2019). Small heat shock proteins: multifaceted proteins with important implications for life. *Cell Stress and Chaperones*, 24, 295-308.

31. Carter, R., & Miller, L. H. (1979). Evidence for environmental modulation of gametocytogenesis in *Plasmodium falciparum* in continuous culture. *Bulletin of the World Health Organization*, 57, 37-52.
32. Casas, A. L. F., Santos, G. M. D., Chiochetti, N. B., & de Andrade, M. (2016). Effects of temperature variation on the human cardiovascular system: a systematic review. *Climate Change and Health: Improving Resilience and Reducing Risks*, 73-87.
33. Cepeda, A. S., Mello, B., Pacheco, M. A., Luo, Z., Sullivan, S. A., Carlton, J. M., & Escalante, A. A. (2024). The genome of *Plasmodium gonderi*: insights into the evolution of human malaria parasites. *Genome Biology and Evolution*, 16(2), evae027.
34. Chakafana, G., Mudau, P. T., Zininga, T., & Shonhai, A. (2021). Characterisation of a unique linker segment of the *Plasmodium falciparum* cytosol localised Hsp110 chaperone. *International Journal of Biological Macromolecules*, 180, 272-285.
35. Chakafana, G., & Shonhai, A. (2021). The Role of Non-Canonical Hsp70s (Hsp110/Grp170) in Cancer. *Cells*, 10(2), 254.
36. Chand, K., Iyer, K., & Mitra, D. (2021a). Comparative analysis of differential gene expression of HSP40 and HSP70 family isoforms during heat stress and HIV-1 infection in T-cells. *Cell Stress and Chaperones*, 26, 403-416.
37. Chand, K., Iyer, K., & Mitra, D. (2021b). Comparative analysis of differential gene expression of HSP40 and HSP70 family isoforms during heat stress and HIV-1 infection in T-cells. *Cell Stress and Chaperones*, 26(2), 403-416.
38. Chandran, S., & Binninger, D. (2023). Role of Oxidative Stress, Methionine Oxidation and Methionine Sulfoxide Reductases (MSR) in Alzheimer's Disease. *Antioxidants*, 13(1), 21.
39. Charnaud, S. C., Dixon, M. W., Nie, C. Q., Chappell, L., Sanders, P. R., Nebl, T., Hanssen, E., Berriman, M., Chan, J.-A., & Blanch, A. J. (2017). The exported chaperone Hsp70-x supports virulence functions for *Plasmodium falciparum* blood stage parasites. *PloS One*, 12(7), e0181656.
40. Che Hussian, C. H. A., & Leong, W. Y. (2023). Thermostable enzyme research advances: a bibliometric analysis. *Journal of Genetic Engineering and Biotechnology*, 21(1), 37.
41. Cheetham, M. E., & Caplan, A. J. (1998). Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. *Cell stress & chaperones*, 3(1), 28.
42. Chiosis, G., Digwal, C. S., Trepel, J. B., & Neckers, L. (2023). Structural and functional complexity of HSP90 in cellular homeostasis and disease. *Nature Reviews Molecular Cell Biology*, 24(11), 797-815.

43. Cohee, L. M., Opondo, C., Clarke, S. E., Halliday, K. E., Cano, J., Shipper, A. G., Barger-Kamate, B., Djimde, A., Diarra, S., Dokras, A., Kanya, M. R., Lutumba, P., Ly, A. B., Nankabirwa, J. I., Njagi, J. K., Maiga, H., Maiteki-Sebuguzi, C., Matangila, J., Okello, G., Chico, R. M. (2020). Preventive malaria treatment among school-aged children in sub-Saharan Africa: a systematic review and meta-analyses. *The Lancet Global Health*, 8(12), e1499-e1511.
44. Collins, T., Azevedo-Silva, J., da Costa, A., Branca, F., Machado, R., & Casal, M. (2013). Batch production of a silk-elastin-like protein in *E. coli* BL21(DE3): key parameters for optimisation. *Microbial Cell Factories*, 12(1), 21.
45. Cortés, G. T., Wiser, M. F., & Gómez-Alegría, C. J. (2020). Identification of *Plasmodium falciparum* HSP70-2 as a resident of the Plasmodium export compartment. *Heliyon*, 6(6), e04037.
46. Cox, F. E. (2010). History of the discovery of the malaria parasites and their vectors. *Parasites & Vectors*, 3, 1-9.
47. Daniyan, M. O., Boshoff, A., Prinsloo, E., Pesce, E.-R., & Blatch, G. L. (2016). The malarial exported PFA0660w is an Hsp40 co-chaperone of PfHsp70-x. *PloS One*, 11(2), e0148517.
48. Daniyan, M. O., Przyborski, J. M., & Shonhai, A. (2019). Partners in mischief: functional networks of heat shock proteins of *Plasmodium falciparum* and their influence on parasite virulence. *Biomolecules*, 9(7), 295.
49. Danwang, C., Kirakoya-Samadoulougou, F., & Samadoulougou, S. (2021). Assessing field performance of ultrasensitive rapid diagnostic tests for malaria: a systematic review and meta-analysis. *Malaria Journal*, 20(1), 245.
50. Dash, S., Rahman, S. R., Hines, H. M., & Feng, W.-c. (2021). iBLAST: Incremental BLAST of new sequences via automated e-value correction. *PloS One*, 16(4), e0249410.
51. David, A., Islam, S., Tankhilevich, E., & Sternberg, M. J. (2022). The AlphaFold database of protein structures: a biologist's guide. *Journal Of Molecular Biology*, 434(2), 167336.
52. Depond, M., Henry, B., Buffet, P., & Ndour, P. A. (2020). Methods to investigate the deformability of RBC during malaria. *Frontiers In Physiology*, 10, 1613.
53. Dishnica, K., Piubelli, C., Manfredi, M., Kondaveeti, R. T., Longoni, S. S., Degani, M., Buonfrate, D., Giorgetti, A., & Tiberti, N. (2023). Novel insights into the somatic proteome of *Strongyloides stercoralis* infective third-stage larvae. *Parasites & Vectors*, 16(1), 45.
54. Draghici, A. E., & Taylor, J. A. (2021). Mechanisms of bone blood flow regulation in humans. *Journal of Applied Physiology*, 130(3), 772-780.

55. Dubrez, L., Causse, S., Borges Bonan, N., Dumétier, B., & Garrido, C. (2020). Heat-shock proteins: chaperoning DNA repair. *Oncogene*, 39(3), 516-529.
56. Duffy, P. E., & Patrick Gorres, J. (2020). Malaria vaccines since 2000: progress, priorities, products. *NPJ Vaccines*, 5(1), 48.
57. Dutta, T., Pesce, E.-R., Maier, A. G., & Blatch, G. L. (2021). Role of the J Domain Protein Family in the Survival and Pathogenesis of *Plasmodium falciparum*. In A. Shonhai, D. Picard, & G. L. Blatch (Eds.), *Heat Shock Proteins of Malaria* (pp. 97-123). *Springer International Publishing*.
58. Dutta, T., Singh, H., Gestwicki, J. E., & Blatch, G. L. (2021). Exported plasmodial J domain protein, PFE0055c, and PfHsp70-x form a specific co-chaperone-chaperone partnership. *Cell Stress and Chaperones*, 26(2), 355-366.
59. Dyer, M., & Day, K. P. (2000). Commitment to gametocytogenesis in *Plasmodium falciparum*. *Parasitology Today*, 16(3), 102-107.
60. Edgar, R. C., & Batzoglou, S. (2006). Multiple sequence alignment. *Current Opinion in Structural Biology*, 16(3), 368-373.
61. El Bakkouri, M., Pow, A., Mulichak, A., Cheung, K. L., Artz, J. D., Amani, M., Fell, S., de Koning-Ward, T. F., Goodman, C. D., & McFadden, G. I. (2010). The Clp chaperones and proteases of the human malaria parasite *Plasmodium falciparum*. *Journal of Molecular Biology*, 404(3), 456-477.
62. El Bakkouri, M., Pow, A., Mulichak, A., Cheung, K. L. Y., Artz, J. D., Amani, M., Fell, S., de Koning-Ward, T. F., Goodman, C. D., McFadden, G. I., Ortega, J., Hui, R., & Houry, W. A. (2010). The Clp Chaperones and Proteases of the Human Malaria Parasite *Plasmodium falciparum*. *Journal of Molecular Biology*, 404(3), 456-477.
63. El Bakkouri, M., Rathore, S., Calmettes, C., Wernimont, A. K., Liu, K., Sinha, D., Asad, M., Jung, P., Hui, R., & Mohmmmed, A. (2013). Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of *Plasmodium falciparum*. *Journal of Biological Chemistry*, 288(2), 1022-1031.
64. Escalante, A. A., Cepeda, A. S., & Pacheco, M. A. (2022). Why *Plasmodium vivax* and *Plasmodium falciparum* are so different? A tale of two clades and their species diversities. *Malaria Journal*, 21(1), 139.
65. Farooq, Q. U. A., Shaukat, Z., Aiman, S., & Li, C. H. (2021). Protein-protein interactions: Methods, databases, and applications in virus-host study. *World J Virol*, 10(6), 288-300.

66. Faust, G., Stand, A., & Weuster-Botz, D. (2015). IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured *Escherichia coli*. *Engineering in Life Sciences*, 15(8), 824-829.
67. Faust, O., Abayev-Avraham, M., Wentink, A. S., Maurer, M., Nillegoda, N. B., London, N., Bukau, B., & Rosenzweig, R. (2020). HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature*, 587(7834), 489-494.
68. Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783-791.
69. Freedman, B. (1987). Scientific value and validity as ethical requirements for research: a proposed explication. *IRB: Ethics & Human Research*, 9(6), 7-10.
70. Frischknecht, F., & Matuschewski, K. (2017). Plasmodium sporozoite biology. *Cold Spring Harbor Perspectives in Medicine*, 7(5), a025478.
71. Fujioka, H., & Aikawa, M. (1999). The Malaria Parasite And Its Lifecycle (Vol. First edition.). *Harwood Academic Publishers*.
72. Gabler, F., Nam, S. Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A. N., & Alva, V. (2020). Protein sequence analysis using the MPI bioinformatics toolkit. *Current Protocols in Bioinformatics*, 72(1), e108.
73. Gallagher, S. R. (2012). SDS-polyacrylamide gel electrophoresis (SDS-PAGE). *Current Protocols Essential Laboratory Techniques*, 6(1), 7.3. 1-7.3. 28.
74. Gandhi, K., Sharma, N., Gautam, P. B., Sharma, R., Mann, B., & Pandey, V. (2022). Polyacrylamide Gel Electrophoresis. In *Advanced Analytical Techniques in Dairy Chemistry* (pp. 103-120). *Springer*.
75. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., & Bairoch, A. (2019). Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, 571-607.
76. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). *Humana Press*.
77. Gaudermann, P., Vogl, I., Zientz, E., Silva, F. J., Moya, A., Gross, R., & Dandekar, T. (2006). Analysis of and function predictions for previously conserved hypothetical or putative proteins in *Blochmannia floridanus*. *BMC Microbiology*, 6(1), 1.
78. Geneva, I. I., Cuzzo, B., Fazili, T., & Javaid, W. (2019). Normal Body Temperature: A Systematic Review. *Open Forum Infectious Diseases*, 6(4).

79. Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, 11(6), 681-684.
80. Gill, S. C., & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, 182(2), 319-326.
81. Gitau, G. W., Mandal, P., Blatch, G. L., Przyborski, J., & Shonhai, A. (2012). Characterisation of the *Plasmodium falciparum* Hsp70–Hsp90 organising protein (PfHop). *Cell Stress and Chaperones*, 17(2), 191-202.
82. Gleeson, M. T. (2000). The plastid in Apicomplexa: what use is it? *International Journal for Parasitology*, 30(10), 1053-1070.
83. Goh, X. T., Lim, Y. A. L., Lee, P. C., Nissapatorn, V., & Chua, K. H. (2021). Diversity and natural selection of Merozoite surface Protein-1 in three species of human malaria parasites: Contribution from South-East Asian isolates. *Molecular and Biochemical Parasitology*, 244, 111390.
84. Gomes, P. S., Bhardwaj, J., Rivera-Correa, J., Freire-De-Lima, C. G., & Morrot, A. (2016). Immune escape strategies of malaria parasites. *Frontiers In Microbiology*, 7, 1617.
85. Gouet, P., Courcelle, E., Stuart, D. I., & MV© toz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics (Oxford, England)*, 15(4), 305-308.
86. Gouet, P., Robert, X., & Courcelle, E. (2005). ESPript/ENDscript: sequence and 3D information from protein structures. *Acta Crystallogr*, 61, 42-43.
87. Gould, S. J., & Subramani, S. (1988). Firefly luciferase as a tool in molecular and cell biology. *Analytical Biochemistry*, 175(1), 5-13.
88. Green, M. R., & Sambrook, J. (2021). Restriction enzymes. *Cold Spring Harbor Protocols*, 2021(4), pdb. top101360.
89. Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*, 1(6), 2876-2890.
90. Grover, M., Chaubey, S., Ranade, S., & Tatu, U. (2013). Identification of an exported heat shock protein 70 in *Plasmodium falciparum*. *Parasite*, 20.
91. Guruprasad, K., Reddy, B. V., & Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng*, 4(2), 155-161.
92. Hagymasi, A. T., Dempsey, J. P., & Srivastava, P. K. (2022). Heat-shock proteins. *Current Protocols*, 2(11), e592.

93. Haldar, K., Bhattacharjee, S., & Safeukui, I. (2018). Drug resistance in Plasmodium. *Nat Rev Microbiol*, 16(3), 156-170.
94. Hall, T., Biosciences, I., & Carlsbad, C. (2011). BioEdit: an important software for molecular biology. *GERF Bull Biosci*, 2(1), 60-61.
95. Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*,
96. Hammershaimb, E. A., & Berry, A. A. (2024). Pre-erythrocytic malaria vaccines: RTS, S, R21, and beyond. *Expert Review of Vaccines*, 23(1), 49-52.
97. Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal Of Molecular Biology*, 166(4), 557-580.
98. Hasegawa, H., Suzuki, E., & Maeda, S. (2018). Horizontal Plasmid Transfer by Transformation in *Escherichia coli*: Environmental Factors and Possible Mechanisms [Mini Review]. *Frontiers In Microbiology*, 9.
99. Haslbeck, M. (2002). sHsps and their role in the chaperone network. *Cellular and Molecular Life Sciences CMLS*, 59, 1649-1657.
100. He, F. (2011). Laemmli-sds-page. *Bio-Protocol*, e80-e80.
101. Hempelmann, E., & Krafts, K. (2013). Bad air, amulets and mosquitoes: 2,000 years of changing perspectives on malaria. *Malaria Journal*, 12, 1-14.
102. Hildebrand, A., Remmert, M., Biegert, A., & Söding, J. (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins: Structure, Function, and Bioinformatics*, 77(S9), 128-132.
103. Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C., & Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic acids research*, 35(suppl_2), W585-W587.
104. Hoter, A., El-Sabban, M. E., & Naim, H. Y. (2018). The HSP90 Family: Structure, Regulation, Function, and Implications in Health and Disease. *Int J Mol Sci*, 19(9).
105. Hu, C., Yang, J., Qi, Z., Wu, H., Wang, B., Zou, F., Mei, H., Liu, J., Wang, W., & Liu, Q. (2022). Heat shock proteins: Biological functions, pathological roles, and therapeutic opportunities. *MedComm* (2020), 3(3), e161.
106. Hu, L., Sun, C., Kidd, J. M., Han, J., Fang, X., Li, H., Liu, Q., May, A. E., Li, Q., & Zhou, L. (2023). A first-in-class inhibitor of Hsp110 molecular chaperones of pathogenic fungi. *Nature Communications*, 14(1), 2745.
107. Hussein, M. F., Kyei-Arthur, F., Saleeb, M., Kyei-Gyamfi, S., Abutima, T., Sakada, I. G., & Ghazy, R. M. (2024). Hesitancy towards R21/Matrix-M malaria vaccine among Ghanaian

- parents and attitudes towards immunizing non-eligible children: a cross-sectional survey. *Malaria Journal*, 23(1), 142.
108. Ikai, A. (1980). Thermostability and Aliphatic Index of Globular Proteins. *The Journal of Biochemistry*, 88(6), 1895-1898.
 109. Ismail, H. M., Barton, V., Phanchana, M., Charoensutthivarakul, S., Wong, M. H. L., Hemingway, J., Biagini, G. A., O'Neill, P. M., & Ward, S. A. (2016). Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites *Plasmodium falciparum* 3D7. *Proceedings of the National Academy of Sciences*, 113(8), 2080-2085.
 110. Jeyachandran, S., Chellapandian, H., Park, K., & Kwak, I.-S. (2023). A review on the involvement of heat shock proteins (extrinsic chaperones) in response to stress conditions in aquatic organisms. *Antioxidants*, 12(7), 1444.
 111. Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal Of Molecular Biology*, 292(2), 195-202.
 112. Josling, G. A., Russell, T. J., Venezia, J., Orchard, L., van Biljon, R., Painter, H. J., & Llinás, M. (2020). Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nature Communications*, 11(1), 1503.
 113. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., & Potapenko, A. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
 114. Jung, M., & Ahn, Y.-J. (2022). Growth-enhancing effect of bacterial and plant heat shock proteins in *Escherichia coli*. *Biocatalysis and Agricultural Biotechnology*, 46, 102545.
 115. Kampinga, Andreasson, C., Barducci, A., Cheetham, M. E., Cyr, D., Emanuelsson, C., Genevaux, P., Gestwicki, J. E., Goloubinoff, P., & Huerta-Cepas, J. (2019). Function, evolution, and structure of J-domain proteins. *Cell Stress and Chaperones*, 24, 7-15.
 116. Kampinga, H. H., Andreasson, C., Barducci, A., Cheetham, M. E., Cyr, D., Emanuelsson, C., Genevaux, P., Gestwicki, J. E., Goloubinoff, P., & Huerta-Cepas, J. (2019). Function, evolution, and structure of J-domain proteins. *Cell Stress and Chaperones*, 24, 7-15.
 117. Kampinga, H. H., Andreasson, C., Barducci, A., Cheetham, M. E., Cyr, D., Emanuelsson, C., Genevaux, P., Gestwicki, J. E., Goloubinoff, P., Huerta-Cepas, J., Kirstein, J., Liberek, K., Mayer, M. P., Nagata, K., Nillegoda, N. B., Pulido, P., Ramos, C., De los Rios, P., Rospert, S., Marszalek, J. (2019). Function, evolution, and structure of J-domain proteins. *Cell Stress and Chaperones*, 24(1), 7-15.

118. Kargar, F., Mortazavi, M., Torkzadeh-Mahani, M., Lotfi, S., & Shakeri, S. (2020). Evaluation of Luciferase Thermal Stability by Arginine Saturation in the Flexible Loops. *Current Proteomics*, 17(1), 30-39.
119. Kędzierska-Mieszkowska, S., & Zolkiewski, M. (2021). Hsp100 Molecular Chaperone ClpB and Its Role in Virulence of Bacterial Pathogens. *International Journal of Molecular Sciences*, 22(10), 5319.
120. Kesheri, M., Kanchan, S., Chowdhury, S., & Sinha, R. P. (2015). Secondary and Tertiary Structure Prediction of Proteins: A Bioinformatic Approach. In Q. Zhu & A. T. Azar (Eds.), *Complex System Modelling and Control Through Intelligent Soft Computations* (pp. 541-569). *Springer International Publishing*.
121. Kielkopf, C. L., Bauer, W., & Urbatsch, I. L. (2021). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of proteins. *Cold Spring Harbor Protocols*, 2021(12), pdb. prot102228.
122. Kudyba, H. M., Cobb, D. W., Fierro, M. A., Florentin, A., Ljolje, D., Singh, B., Lucchi, N. W., & Muralidharan, V. (2019). The endoplasmic reticulum chaperone PfGRP170 is essential for asexual development and is linked to stress response in malaria parasites. *Cellular Microbiology*, 21(9), e13042.
123. Kulyyassov, A., Fresnais, M., & Longuespée, R. (2021). Targeted liquid chromatography-tandem mass spectrometry analysis of proteins: basic principles, applications, and perspectives. *Proteomics*, 21(23-24), 2100153.
124. Kumar, R., Tripathi, T., & Saudagar, P. (2023). Fluorescence Spectroscopy-Based Methods to Study Protein Folding Dynamics. In P. Saudagar & T. Tripathi (Eds.), *Protein Folding Dynamics and Stability: Experimental and Computational Methods* (pp. 25-36). *Springer Nature Singapore*.
125. Kumari, K., & Ramakrishnan, V. (2023). Circular Dichroism (CD) Spectroscopy. In *Biophysical Characterization of Functional Peptides* (pp. 45-50). *Springer*.
126. Kurien, B. T. (2021). Western Blotting for The Non-Expert. *Springer international publishing*.
127. Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157(1), 105-132.
128. Lan, Y., Wang, Y., Huang, K., & Zeng, Q. (2020). Heat shock protein 22 attenuates doxorubicin-induced cardiotoxicity via regulating inflammation and apoptosis. *Frontiers In Pharmacology*, 11, 257.

129. Lang, B. J., Guerrero, M. E., Prince, T. L., Okusha, Y., Bonorino, C., & Calderwood, S. K. (2021). The functions and regulation of heat shock proteins; key orchestrators of proteostasis and the heat shock response. *Archives of Toxicology*, 95(6), 1943-1970.
130. Laub, V., Devraj, K., Elias, L., & Schulte, D. (2023). Bioinformatics for wet-lab scientists: practical application in sequencing analysis. *BMC Genomics*, 24(1), 382.
131. Laurens, M. B. (2020). RTS, S/AS01 vaccine (Mosquirix™): an overview. *Human vaccines & Immunotherapeutics*, 16(3), 480-489.
132. Le Roch, K. G., Zhou, Y., Blair, P. L., Grainger, M., Moch, J. K., Haynes, J. D., De la Vega, P., Holder, A. A., Batalov, S., & Carucci, D. J. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*, 301(5639), 1503-1508.
133. Lebepe, C. M., Matambanadzo, P. R., Makhoba, X. H., Achilonu, I., Zininga, T., & Shonhai, A. (2020). Comparative Characterization of *Plasmodium falciparum* Hsp70-1 Relative to *E. coli* DnaK Reveals the Functional Specificity of the Parasite Chaperone. *Biomolecules*, 10(6).
134. Lee, G., Kim, R. S., Lee, S. B., Lee, S., & Tsai, F. T. (2022). Deciphering the mechanism and function of Hsp100 unfoldases from protein structure. *Biochemical Society Transactions*, 50(6), 1725-1736.
135. Lepvrier, E., Thomas, D., & Garnier, C. (2019). Hsp90 quaternary structures and the chaperone cycle: Highly flexible dimeric and oligomeric structures and their regulation by co-chaperones. *Current Proteomics*, 16(1), 5-11.
136. Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293-W296.
137. Lewis, A. C., Jones, N. S., Porter, M. A., & Deane, C. M. (2012). What evidence is there for the homology of protein-protein interactions? *PLOS Computational Biology* 8(9): e1002645.
138. Li, J., & Zhu, H.-J. (2020). Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)-Based Proteomics of Drug-Metabolizing Enzymes and Transporters. *Molecules*, 25(11), 2718.
139. Li, K. P., Wong, C. H., Cheng, C. C., Cheng, S. S., Li, M. W., Mansveld, S., Bergsma, A., Huang, T., van Eijk, M. J., & Lam, H. M. (2021). GmDNJ1, a type-I heat shock protein 40 (HSP40), is responsible for both Growth and heat tolerance in soybean. *Plant Direct*, 5(1), e00298.
140. Liang, X., Kaya, A., Zhang, Y., Le, D. T., Hua, D., & Gladyshev, V. N. (2012). Characterization of methionine oxidation and methionine sulfoxide reduction using methionine-rich cysteine-free proteins. *BMC Biochemistry*, 13(1), 21.

141. Lindquist, S., & Craig, E. A. (1988). The heat-shock proteins. *Annual Review of Genetics*, 22(1), 631-677.
142. Lins, L., Thomas, A., & Brasseur, R. (2003). Analysis of accessible surface of residues in proteins. *Protein science*, 12(7), 1406-1417.
143. Liu, K., & Houry, W. A. (2014). Chaperones and proteases of *Plasmodium falciparum*. Heat Shock Proteins of Malaria, *Springer*. 161-187.
144. Liu, Q., Liang, C., & Zhou, L. (2020). Structural and functional analysis of the Hsp70/Hsp40 chaperone system. *Protein Science*, 29(2), 378-390.
145. Liu, S. (2015). A review on protein oligomerization process. *International Journal of Precision Engineering and Manufacturing*, 16(13), 2731-2760.
146. Lopez, J., Le, N. H., Moon, K. H., Salomon, D., Bosis, E., & Feldman, M. F. (2021). Formylglycine-Generating Enzyme-Like Proteins Constitute a Novel Family of Widespread Type VI Secretion System Immunity Proteins. *J Bacteriol*, 203(21), e0028121.
147. Loughran, S. T., Loughran, N. B., Ryan, B. J., D'Souza, B. N., & Walls, D. (2006). Modified His-tag fusion vector for enhanced protein purification by immobilized metal affinity chromatography. *Analytical Biochemistry*, 355(1), 148-150.
148. Lozano Terol, G., Gallego-Jara, J., Sola Martinez, R. A., Martinez Vivancos, A., Cánovas Díaz, M., & de Diego Puente, T. (2021). Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. *Frontiers In Microbiology*, 12, 682001.
149. Lubkowska, A., Pluta, W., Strońska, A., & Lalko, A. (2021). Role of heat shock proteins (HSP70 and HSP90) in viral infection. *International Journal of Molecular Sciences*, 22(17), 9366.
150. Lupoli, T. J., Fay, A., Adura, C., Glickman, M. S., & Nathan, C. F. (2016). Reconstitution of a *Mycobacterium tuberculosis* proteostasis network highlights essential cofactor interactions with chaperone DnaK. *Proceedings of the National Academy of Sciences*, 113(49), E7947-E7956.
151. M Njunge, J., H Ludewig, M., Boshoff, A., Pesce, E.-R., & L Blatch, G. (2013). Hsp70s and J proteins of *Plasmodium* parasites infecting rodents and primates: structure, function, clinical relevance, and drug targets. *Current Pharmaceutical Design*, 19(3), 387-403.
152. Ma, Y., Lou, F., Yin, X., Cong, B., Liu, S., Zhao, L., & Zheng, L. (2022). Whole-genome survey and phylogenetic analysis of *Gadus macrocephalus*. *Bioscience Reports*, 42(7), BSR20221037.

153. Maayah, Y., Nusrat, H., Pang, G., & Tambasco, M. (2022). Assessing the DNA Damaging Effectiveness of Ionizing Radiation Using Plasmid DNA. *International Journal of Molecular Sciences*, 23(20), 12459.
154. Mace, K. E., Arguin, P. M., & Tan, K. R. (2018). Malaria Surveillance - United States, 2015. Morbidity and mortality weekly report. *Surveillance Summaries (Washington, D.C.: 2002)*, 67(7), 1-28.
155. Magdeldin, S., Enany, S., Yoshida, Y., Xu, B., Zhang, Y., Zureena, Z., Lokamani, I., Yaoita, E., & Yamamoto, T. (2014). Basics and recent advances of two dimensional- polyacrylamide gel electrophoresis. *Clinical Proteomics*, 11(1), 16.
156. Mair, G. R., Lasonder, E., Garver, L. S., Franke-Fayard, B. M. D., Carret, C. K., Wiegant, J. C. A. G., Dirks, R. W., Dimopoulos, G., Janse, C. J., & Waters, A. P. (2010). Universal features of post-transcriptional gene regulation are critical for *Plasmodium* zygote development. *PLoS Pathogens*, 6(2), e1000767.
157. Makam, P., & Matsa, R. (2021). "Big Three" infectious diseases: tuberculosis, malaria and HIV/AIDS. *Current Topics in Medicinal Chemistry*, 21(31), 2779-2799.
158. Makhoba, X. H., Godlo, S., Maxam, N., Mthembu, Y., Mpumza, A., Makumire, S., Mkwetshana, N., & Govender, K. K. (2023). The Computational Analysis of *Plasmodium falciparum* Heat Shock Proteins Reveals an Interplay with Polyamines. *Biorxiv*, 2023.2005.2011.540332.
159. Makumire, S., Zininga, T., Vahokoski, J., Kursula, I., & Shonhai, A. (2020). Biophysical analysis of *Plasmodium falciparum* Hsp70-Hsp90 organising protein (PfHop) reveals a monomer that is characterised by folded segments connected by flexible linkers. *PLOS ONE*, 15(4), e0226657.
160. Mann, M., & Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nature biotechnology*, 21(3), 255-261.
161. Manukhov, I. V., Eroshnikov, G. E., Vyssokikh, M. Y., & Zavilgelsky, G. B. (1999). Folding and refolding of thermolabile and thermostable bacterial luciferases: the role of DnaKJ heat-shock proteins. *FEBS letters*, 448(2-3), 265-268.
162. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Bryant, S. H. (2016). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, 45(D1), D200-D203.

163. Marszalek, J., De Los Rio, P., Cyr, D., Mayer, M. P., Adupa, V., Andréasson, C., Blatch, G. L., Braun, J., Brodsky, J. L., & Bukau, B. (2023). J-Domain Proteins: from molecular mechanisms to diseases. *Cell Stress and Chaperones*.
164. Martin, R. (2020). Gel electrophoresis: nucleic acids. *Garland Science*.
165. Micsonai, A., Moussong, É., Wien, F., Boros, E., Vadász, H., Murvai, N., Lee, Y.-H., Molnár, T., Réfrégiers, M., & Goto, Y. (2023). Bestsel: Updated webserver for secondary structure and fold prediction for protein CD spectroscopy. *Biophysical Journal*, 122(3), 179a.
166. Miles, A., Janes, R. W., & Wallace, B. A. (2021). Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review. *Chemical Society Reviews*, 50(15), 8400-8413.
167. Milner, D. A. (2018). Malaria pathogenesis. *Cold Spring Harbor Perspectives in Medicine*, 8(1), a025569.
168. Mooi, N., Roy, S. W., & Connor, E. F. (2023). Cytokinin biosynthesis in Hexapoda and Insecta: a bioinformatic analysis. *Arthropod-Plant Interactions*.
169. Moritz, C. P. (2020). 40 years Western blotting: A scientific birthday toast. *Journal of proteomics*, 212, 103575.
170. Morris, R., Black, K. A., & Stollar, E. J. (2022). Uncovering protein function: from classification to complexes. *Essays in Biochemistry*, 66(3), 255-285.
171. Morrow, G., Heikkilä, J. J., & Tanguay, R. M. (2006). Differences in the chaperone-like activities of the four main small heat shock proteins of *Drosophila melanogaster*. *Cell stress & chaperones*, 11(1), 51.
172. Moskovitz, J., & Smith, A. (2021). Methionine sulfoxide and the methionine sulfoxide reductase system as modulators of signal transduction pathways: a review. *Amino Acids*, 53(7), 1011-1020.
173. Mulyani, W. R. W., Sanjiwani, M. I. D., Sandra, Prabawa, I. P. Y., Lestari, A. A. W., Wihandani, D. M., Suastika, K., Saraswati, M. R., Bhargah, A., & Manuaba, I. B. A. P. (2020). Chaperone-based therapeutic target innovation: Heat shock protein 70 (HSP70) for Type 2 diabetes mellitus. *Diabetes, Metabolic Syndrome and Obesity*, 559-568.
174. Muralidharan, V., Oksman, A., Pal, P., Lindquist, S., & Goldberg, D. E. (2012). *Plasmodium falciparum* heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nature Communications*, 3(1), 1310.
175. Mutavhatsindi, H. (2016). Characterization of PFF1010c, a type IV *Plasmodium falciparum* heat shock protein 40. *University of Venda*. (Master's thesis).

176. Myers, J. A., Curtis, B. S., & Curtis, W. R. (2013). Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophysics*, 6(1), 4.
177. Ngansop, F., Li, H., Zolkiewska, A., & Zolkiewski, M. (2013). Biochemical characterization of the apicoplast-targeted AAA+ ATPase ClpB from *Plasmodium falciparum*. *Biochemical and Biophysical Research Communications*, 439(2), 191-195.
178. Njunge, J., H Ludewig, M., Boshoff, A., Pesce, E.-R., & L Blatch, G. (2013). Hsp70s and J proteins of Plasmodium parasites infecting rodents and primates: structure, function, clinical relevance, and drug targets. *Current Pharmaceutical Design*, 19(3), 387-403.
179. Njunge, J. M., Mandal, P., Przyborski, J. M., Boshoff, A., Pesce, E.-R., & Blatch, G. L. (2015). PFB0595w is a *Plasmodium falciparum* J protein that co-localizes with PfHsp70-1 and can stimulate its in vitro ATP hydrolysis activity. *The International Journal of Biochemistry & Cell Biology*, 62, 47-53.
180. Noddings, C. M., Wang, R. Y.-R., Johnson, J. L., & Agard, D. A. (2022). Structure of Hsp90–p23–GR reveals the Hsp90 client-remodelling mechanism. *Nature*, 601(7893), 465-469.
181. Noeth, K. P., Kaiser, M. L., Mashatola, T., Dahan-Moss, Y. L., Matamba, P. A., Spillings, B., Christian, R., Erlank, E., Tshikae, B. P., Jamesboy, E., Sibambo, S., Nkosi, B. G., Silawu, B. T., Mkhabela, L. J., Ndlovu, F. S., Mgwenya, T. P., Coetzee, M., Brooke, B. D., Koekemoer, L. L., Oliver, S. V. (2024). A review of historical trends in *Anopheles gambiae* Giles (Diptera: Culicidae) complex composition, collection trends and environmental effects from 2009 to 2021 in Mpumalanga province, South Africa. *Medical and Veterinary Entomology*, 1-12.
182. Nureye, D., & Assefa, S. (2020). Old and Recent Advances in Life Cycle, Pathogenesis, Diagnosis, Prevention, and Treatment of Malaria Including Perspectives in Ethiopia. *The Scientific World Journal*, 2020, 1295381.
183. Nyakundi, D. O., & Boshoff, A. (2021). Biological Characterization of *Plasmodium falciparum* Mitochondrial Heat Shock Protein PfHsp70-3: Possible Involvement in Malaria Pathogenesis. *Tanzania Journal of Science*, 47(4), 1338-1351.
184. Nyakundi, D. O., Vuko, L. A., Bentley, S. J., Hoppe, H., Blatch, G. L., & Boshoff, A. (2016). *Plasmodium falciparum* Hep1 is required to prevent the self-aggregation of PfHsp70-3. *PloS One*, 11(6), e0156446.
185. Nye, E. R. (2002). Alphonse Laveran (1845–1922): discoverer of the malarial parasite and Nobel laureate, 1907. *Journal of medical Biography*, 10(2), 81-87.

186. O'Neill, S., Gryseels, C., Dierickx, S., Mwesigwa, J., Okebe, J., d'Alessandro, U., & Grietens, K. P. (2015). Foul wind, spirits and witchcraft: illness conceptions and health-seeking behaviour for malaria in the Gambia. *Malaria Journal*, 14, 1-10.
187. Ochoterena, H., Vrijdaghs, A., Smets, E., & Claßen-Bockhoff, R. (2019). The Search for Common Origin: Homology Revisited. *Syst Biol*, 68(5), 767-780.
188. Ogden, R. C., & Adams, D. A. (1987). [8] Electrophoresis in agarose and acrylamide gels. *Methods In Enzymology* (Vol. 152, pp. 61-87). *Academic Press*.
189. WHO. (2023). *WHO guidelines for malaria*.
190. Ota, C., Tanaka, S.-i., & Takano, K. (2021). Revisiting the Rate-Limiting Step of the ANS-Protein Binding at the Protein Surface and Inside the Hydrophobic Cavity. *Molecules*, 26(2), 420.
191. Patel, A., Sharp, S. Y., Hall, K., Lewis, W., Stevens, M. F. G., Workman, P., & Moody, C. J. (2016). Fused imidazoles as potential chemical scaffolds for inhibition of heat shock protein 70 and induction of apoptosis. Synthesis and biological evaluation of phenanthro[9,10-d]imidazoles and imidazo[4,5-f][1,10]phenanthrolines [10.1039/C6OB00471G]. *Organic & Biomolecular Chemistry*, 14(16), 3889-3905.
192. Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Current Protocols In Bioinformatics*, 42(1), 3.1. 1-3.1. 8.
193. Pesce, E. R., Cockburn, I. L., Goble, J. L., Stephens, L. L., & Blatch, G. L. (2010). Malaria heat shock proteins: drug targets that chaperone other drug targets. *Infect Disord Drug Targets*, 10(3), 147-157.
194. Petrini, L., Pennati, G., & Fotiadis, D. I. (2022). Editorial: Verification and Validation of in silico Models for Biomedical Implantable Devices [Editorial]. *Frontiers in Medical Technology*, 4.
195. Phelan, J. E., Turkiewicz, A., Manko, E., Thorpe, J., Vanheer, L. N., van de Vegte-Bolmer, M., Ngoc, N. T. H., Binh, N. T. H., Thieu, N. Q., Gitaka, J., Nolder, D., Beshir, K. B., Dombrowski, J. G., Di Santi, S. M., Bousema, T., Sutherland, C. J., Campino, S., & Clark, T. G. (2023). Rapid profiling of Plasmodium parasites from genome sequences to assist malaria control. *Genome Medicine*, 15(1), 96.
196. Piette, B. L., Alerasool, N., Lin, Z.-Y., Lacoste, J., Lam, M. H. Y., Qian, W. W., Tran, S., Larsen, B., Campos, E., & Peng, J. (2021). Comprehensive interactome profiling of the human Hsp70 network highlights functional differentiation of J domains. *Molecular Cell*, 81(12), 2549-2565. e2548.

197. Prajapati, S. K., Ayanful-Torgby, R., Pava, Z., Barbeau, M. C., Acquah, F. K., Cudjoe, E., Kakaney, C., Amponsah, J. A., Obboh, E., Ahmed, A. E., Abuaku, B. K., McCarthy, J. S., Amoah, L. E., & Williamson, K. C. (2020). The transcriptome of circulating sexually committed *Plasmodium falciparum* ring stage parasites forecasts malaria transmission potential. *Nature Communications*, 11(1), 6159.
198. Rajpal, S., Mishra, P., & Mizaikoff, B. (2023). Rational In Silico Design of Molecularly Imprinted Polymers: Current Challenges and Future Potential. *International Journal of Molecular Sciences*, 24(7), 6785.
199. Ramatsui, L., Dongola, T. H., Zininga, T., Multhoff, G., & Shonhai, A. (2023). Human granzyme B binds *Plasmodium falciparum* Hsp70-x and mediates antiplasmodial activity in vitro. *Cell Stress and Chaperones*, 28(3), 321-331.
200. Ramazi, S., & Zahiri, J. (2021). Post-translational modifications in proteins: resources, tools and prediction methods. *Database*, 2021, baab012.
201. Ramya, T., Surolia, N., & Surolia, A. (2006). 15-Deoxyspergualin modulates *Plasmodium falciparum* heat shock protein function. *Biochemical and Biophysical Research Communications*, 348(2), 585-592.
202. Rasmussen, H. Ø., Enghild, J. J., Otzen, D. E., & Pedersen, J. S. (2020). Unfolding and partial refolding of a cellulase from the SDS-denatured state: From β -sheet to α -helix and back. *Biochimica et Biophysica Acta (BBA). General Subjects*, 1864(1), 129434.
203. Ren, J., Karna, S., Lee, H.-M., Yoo, S. M., & Na, D. (2019). Artificial transformation methodologies for improving the efficiency of plasmid DNA transformation and simplifying its use. *Applied Microbiology and Biotechnology*, 103(23), 9205-9215.
204. Riguero, V., Clifford, R., Dawley, M., Dickson, M., Gastfriend, B., Thompson, C., Wang, S.-C., & O'Connor, E. (2020). Immobilized metal affinity chromatography optimization for poly-histidine tagged proteins. *Journal of Chromatography A*, 1629, 461505.
205. Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in *Drosophila*. *Experientia*, 18(12), 571-573.
206. Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), W320-W324.
207. Rodriguez, A. (2022). Structural and Biochemical Studies on Disease Associated Proteins: Heat Shock Protein 60 and HIV Capsid. *The University of Texas at El Paso*. (Doctoral dissertation).

208. Rosales, M. E. T., Tejeda, M. G. S., Tejeda, J. F. S., & Ruiz, J. F. S. (2023). Phylogenetic Analysis by Molecular Sequence of Various Human Interleukins. *Journal of Pharmaceutical Research International*, 35(21), 28-35.
209. Ruff, K. M., & Pappu, R. V. (2021). AlphaFold and implications for intrinsically disordered proteins. *Journal of Molecular Biology*, 433(20), 167208.
210. Rug, M., & Maier, A. G. (2011). The heat shock protein 40 family of the malaria parasite *Plasmodium falciparum*. *IUBMB Life*, 63(12), 1081-1086.
211. Ruger-Herreros, C., Svoboda, L., Mogk, A., & Bukau, B. (2024). Role of J-domain proteins in yeast physiology and protein quality control. *Journal of molecular biology*, 168484.
212. Sato, S. (2021). Plasmodium—a brief introduction to the parasites causing human malaria and their basic biology. *Journal of Physiological Anthropology*, 40(1), 1.
213. Sazed, S. A., Kibria, M. G., & Alam, M. S. (2021). An optimized real-time qPCR method for the effective detection of human malaria infections. *Diagnostics*, 11(5), 736.
214. Schoch, C. L., Ciufo, S., Domrachev, M., Hottton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., & Karsch-Mizrachi, I. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*, 2020.
215. Schultz, J. (2006). Protein Domains. In *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*. Springer Berlin Heidelberg. (pp. 1490-1495).
216. Shapiro, L. L. M., Whitehead, S. A., & Thomas, M. B. (2017). Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. *PLOS Biology*, 15(10), e2003489.
217. Shi, Y.-y., Hong, X.-g., & Wang, C.-c. (2005). The C-terminal (331–376) sequence of *Escherichia coli* DnaJ is essential for dimerization and chaperone activity: a small angle X-ray scattering study in solution. *Journal of Biological Chemistry*, 280(24), 22761-22768.
218. Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., & Bukau, B. (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature*, 561(7722), 268-272.
219. Shin Hyun-Il, L. H.-I. (2021). 2020 World Malaria Report. *Public Health Weekly Report*, 14(17), 997-1012.
220. Shin, W.-H., Lee, G. R., Heo, L., Lee, H., & Seok, C. (2014). Prediction of protein structure and interaction by GALAXY protein modeling programs. *Bio Design*, 2(1), 1-11.
221. Shonhai, A. (2010). Plasmodial heat shock proteins: targets for chemotherapy. *FEMS Immunology & Medical Microbiology*, 58(1), 61-74.

222. Shonhai, A., & Blatch, G. L. (2014). Heat shock proteins of malaria. *Springer*.
223. Shonhai, A., Blatch, G.L. (2021). Heat Shock Proteins of Malaria: Highlights and Future Prospects. In: Shonhai, A., Picard, D., Blatch, G.L. (eds) Heat Shock Proteins of Malaria. Advances in Experimental Medicine and Biology, vol 1340. *Springer, Cham*.
224. Shonhai, A., Boshoff, A., & Blatch, G. L. (2007). The structural and functional diversity of Hsp70 proteins from *Plasmodium falciparum*. *Protein Science*, 16(9), 1803-1818.
225. Sievers, F., & Higgins, D. G. (2021). The clustal omega multiple alignment package. Multiple sequence alignment: *Methods and Protocols*, 3-16.
226. Silva, N. S. M., Torricillas, M. S., Minari, K., Barbosa, L. R. S., Seraphim, T. V., & Borges, J. C. (2020). Solution structure of *Plasmodium falciparum* Hsp90 indicates a high flexible dimer. *Archives of Biochemistry and Biophysics*, 690, 108468.
227. Sinden, R. (1983). Sexual development of malarial parasites. *Advances in parasitology*, 22, 153-216.
228. Singhvi, P., & Panda, A. K. (2022). Solubilization Solubilization and Refolding of Inclusion Body Proteins. In *Insoluble Proteins: Methods and Protocols* (pp. 371-387). Springer.
229. Söding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, 33(Web Server issue), W244-248.
230. Stofberg, M. L., Caillet, C., de Villiers, M., & Zininga, T. (2021). Inhibitors of the *Plasmodium falciparum* Hsp90 towards Selective Antimalarial Drug Design: *The Past, Present and Future. Cells*, 10(11), 2849.
231. Sugino, C., Hirose, M., Tohda, H., Yoshinari, Y., Abe, T., Giga-Hama, Y., Iizuka, R., Shimizu, M., Kidokoro, S. i., & Ishii, N. (2009). Characterization of a sHsp of *Schizosaccharomyces pombe*, SpHsp15. 8, and the implication of its functional mechanism by comparison with another sHsp, SpHsp16. 0. *Proteins: Structure, Function, and Bioinformatics*, 74(1), 6-17.
232. Suresh, N., & Haldar, K. (2018). Mechanisms of artemisinin resistance in *Plasmodium falciparum* malaria. *Current Opinion in Pharmacology*, 42, 46-54.
233. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., & Pyysalo, S. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 51(D1), D638-D646.
234. Szyller, J., & Bil-Lula, I. (2021). Heat Shock Proteins in Oxidative Stress and Ischemia/Reperfusion Injury and Benefits from Physical Exercises: A Review to the Current Knowledge. *Oxidative Medicine and Cellular Longevity*, 2021, 6678457.

235. Talapko, J., Škrlec, I., Alebić, T., Jukić, M., & Včev, A. (2019). Malaria: the past and the present. *Microorganisms*, 7(6), 179.
236. Talman, A. M., Domarle, O., McKenzie, F. E., Arie, F., & Robert, V. (2004). Gametocytogenesis : the puberty of *Plasmodium falciparum*. *Malaria Journal*, 3(1), 24.
237. Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7), 3022-3027.
238. Tang, Y., Zhou, Y., Fan, S., & Wen, Q. (2022). The multiple roles and therapeutic potential of HSP60 in cancer. *Biochemical Pharmacology*, 201, 115096.
239. Tedesco, P., Bevilacqua, S., Fiorito, G., & Terlizzi, A. (2020). Global patterns of parasite diversity in cephalopods. *Scientific Reports*, 10(1), 11303.
240. Thera, M. A., Kone, A. K., Tangara, B., Diarra, E., Niare, S., Dembele, A., Sissoko, M. S., & Doumbo, O. K. (2018). School-aged children based seasonal malaria chemoprevention using artesunate-amodiaquine in Mali. *Parasite Epidemiology and Control*, 3(2), 96-105.
241. Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*(1), 2.3. 1-2.3. 22.
242. Timothy, F. M., & Zininga, T. (2023). Small heat shock proteins as modulators of cell death in *Plasmodium falciparum* parasites and its human host. *Frontiers in Cell Death*, 2, 1322780.
243. Tintó-Font, E., & Cortés, A. (2022). Malaria parasites do respond to heat. *Trends in Parasitology*, 38(6), 435-449.
244. Tomiczek, B., Delewski, W., Nierzwicki, L., Stolarska, M., Grochowina, I., Schilke, B., Dutkiewicz, R., Uzarska, M. A., Ciesielski, S. J., & Czub, J. (2020). Two-step mechanism of J-domain action in driving Hsp70 function. *PLoS Computational Biology*, 16(6), e1007913.
245. Topcu, C., Georgiou, V., Rodosthenous, J. H., Demetriades, I., Foley, B. T., & G. Kostrikis, L. (2022). Characterization of a novel HIV-1 circulating recombinant form, CRF91_cpx, comprising CRF02_AG, G, J, and U, mostly among men who have sex with men. *Virulence*, 13(1), 1331-1348.
246. Torrisi, M., Pollastri, G., & Le, Q. (2020). Deep learning methods in protein structure prediction. *Computational and Structural Biotechnology Journal*, 18, 1301-1310.
247. Usui, M., & Williamson, K. C. (2021). Stressed out about *Plasmodium falciparum* gametocytogenesis. *Frontiers in Cellular and Infection Microbiology*, 11, 790067.
248. Uttarotai, T., Mukjang, N., Chaisoung, N., Pathom-Aree, W., Pekkoh, J., Pumas, C., & Sattayawat, P. (2022). Putative Protein Discovery from Microalgal Genomes as a Synthetic Biology Protein Library for Heavy Metal Bio-Removal. *Biology*, 11(8), 1226.

249. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., & Laydon, A. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, 50(D1), D439-D444.
250. Vieira-Pires, R. S., & Morais-Cabral, J. H. (2010). 310 helices in channels and other membrane proteins. *Journal of General Physiology*, 136(6), 585-592.
251. Virág, D., Dalmadi-Kiss, B., Vékey, K., Drahos, L., Klebovich, I., Antal, I., & Ludányi, K. (2020). Current trends in the analysis of post-translational modifications. *Chromatographia*, 83, 1-10.
252. von Bergen, M., Barghorn, S., Biernat, J., Mandelkow, E.-M., & Mandelkow, E. (2005). Tau aggregation is driven by a transition from random coil to beta sheet structure. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1739(2), 158-166.
253. Wai, A. H., Cho, L.-H., Peng, X., Waseem, M., Lee, D.-j., Lee, J.-M., Kim, C.-K., & Chung, M.-Y. (2021). Genome-wide identification and expression profiling of *Alba* gene family members in response to abiotic stress in tomato (*Solanum lycopersicum* L.). *BMC Plant Biology*, 21(1), 530.
254. Walter, E. J., Hanna-Jumma, S., Carraretto, M., & Forni, L. (2016). The pathophysiological basis and consequences of fever. *Critical Care*, 20(1), 200.
255. Wang, D., Liu, D., Yuchi, J., He, F., Jiang, Y., Cai, S., Li, J., & Xu, D. (2020). MusiteDeep: a deep learning based webserver for protein post-translational modification site prediction and visualization. *Nucleic Acids Research*, 48(W1), W140-W146.
256. Wang, Y., Zhang, H., Zhong, H., & Xue, Z. (2021). Protein domain identification methods and online resources. *Computational and Structural Biotechnology Journal*, 19, 1145-1153.
257. Wegele, H., Wandinger, S. K., Schmid, A. B., Reinstein, J., & Buchner, J. (2006). Substrate Transfer from the Chaperone Hsp70 to Hsp90. *Journal of Molecular Biology*, 356(3), 802-811.
258. WHO. (2023). World malaria report 2023. *World malaria report 2023*.
259. Wickramaratne, A. C., Liao, J.-Y., Doyle, S. M., Hoskins, J. R., Puller, G., Scott, M. L., Alao, J. P., Obaseki, I., Dinan, J. C., & Maity, T. K. (2023). J-domain Proteins form Binary Complexes with Hsp90 and Ternary Complexes with Hsp90 and Hsp70. *Journal of Molecular Biology*, 435(17), 168184.

260. Willbold, D., Strodel, B., Schröder, G. F., Hoyer, W., & Heise, H. (2021). Amyloid-type protein aggregation and prion-like properties of amyloids. *Chemical Reviews*, 121(13), 8285-8307.
261. Winer, B., Edgel, K. A., Zou, X., Sellau, J., Hadiwidjojo, S., Garver, L. S., McDonough, C. E., Kelleher, N. L., Thomas, P. M., Villasante, E., Ploss, A., & Gerbasi, V. R. (2020). Identification of *Plasmodium falciparum* proteoforms from liver stage models. *Malaria Journal*, 19(1), 10.
262. Wong, T. S., & Tee, K. L. (2020). Protein Expression Hosts and Expression Plasmids. In T. S. Wong & K. L. Tee (Eds.), *A Practical Guide to Protein Engineering* (pp. 39-61). *Springer International Publishing*.
263. Wu, Y., Li, J., Jin, Z., Fu, Z., & Sha, B. (2005). The Crystal Structure of the C-terminal Fragment of Yeast Hsp40 Ydj1 Reveals Novel Dimerization Motif for Hsp40. *Journal of Molecular Biology*, 346(4), 1005-1011.
264. Wyciszkievicz, A., Kalinowska-Lyszczarz, A., Nowakowski, B., Kaźmierczak, K., Osztynowicz, K., & Michalak, S. (2019). Expression of small heat shock proteins in exosomes from patients with gynecologic cancers. *Scientific Reports*, 9(1), 9817.
265. Xiao, Y., Liu, Y.-D., Yuan, G., Mao, R.-Q., & Li, G. (2021). An uncharacterized protein from the metagenome with no obvious homology to known lipases shows excellent alkaline lipase properties and potential applications in the detergent industry. *Biotechnology Letters*, 43(12), 2311-2325.
266. Xu, X., Zhu, Y., Yue, C., Yang, Q., & Zhang, Z. (2022). Comprehensive Bioinformatics Analysis Combined with Wet-Lab Experiments to Find Target Proteins of Chinese Medicine Monomer. *Molecules*, 27(18), 6105.
267. Zainordin, Z. F., San, C. T., & Ahmad, A. (2021). Isolation, molecular characterisation and expression level of cytosolic heat-shock protein 90 (hsp90) genes of *Cryptocoryne ciliata*, a halophyte plant. *Journal of Sustainability Science and Management*, 16(7), 31-45.
268. Zangari, J., Petrelli, F., Maillot, B., & Martinou, J.-C. (2020). The Multifaceted Pyruvate Metabolism: Role of the Mitochondrial Pyruvate Carrier. *Biomolecules*, 10(7), 1068.
269. Zavilgelsky, G., Gnuchikh, E. Y., & Melkina, O. (2020). Thermostability and Refolding of Proteins in Bacteria Is Determined by the Activity of Two Different ATP-Dependent Chaperone Groups. *Molecular Biology*, 54, 262-268.
270. Zdravko, P., & Darja Žgur, B. (2021). The *Escherichia coli* SOS Response: Much More than DNA Damage Repair. E. Marjanca Starčič (Ed.), *Escherichia coli* (pp. Ch. 3). *IntechOpen*.

271. Zhang, C., Liu, M.-S., & Xing, X.-H. (2009). Temperature influence on fluorescence intensity and enzyme activity of the fusion protein of GFP and hyperthermophilic xylanase. *Applied Microbiology and Biotechnology*, 84(3), 511-517.
272. Zhang, M., Wang, C., Oberstaller, J., Thomas, P., Otto, T. D., Casandra, D., Boyapalle, S., Adapa, S. R., Xu, S., & Button-Simons, K. (2021). The apicoplast link to fever-survival and artemisinin-resistance in the malaria parasite. *Nature Communications*, 12(1), 4563.
273. Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D., & Mayho, M. (2018). Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science*, 360(6388), eaap7847.
274. Zhao, L., Wang, L., Aierken, R., Wang, W., Wang, X., & Li, M. (2020). Characterization of Insulin and Glucagon Genes and Their Producing Endocrine Cells From Pygmy Sperm Whale (*Kogia breviceps*) [Original Research]. *Frontiers in Endocrinology*, 11.
275. Zininga, T., Achilonu, I., Hoppe, H., Prinsloo, E., Dirr, H. W., & Shonhai, A. (2016). *Plasmodium falciparum* Hsp70-z, an Hsp110 homologue, exhibits independent chaperone activity and interacts with Hsp70-1 in a nucleotide-dependent fashion. *Cell Stress and Chaperones*, 21(3), 499-513.
276. Zolg, D. P., Wilhelm, M., Schmidt, T., Médard, G., Zerweck, J., Knaute, T., Wenschuh, H., Reimer, U., Schnatbaum, K., & Kuster, B. (2018). ProteomeTools: Systematic characterization of 21 post-translational protein modifications by liquid chromatography tandem mass spectrometry (LC-MS/MS) using synthetic peptides. *Molecular & Cellular Proteomics*, 17(9), 1850-1863.

Appendices

Appendix A: Equipment, expendables, prepared solutions, and reagents

Table A1. List of equipment and their suppliers

Equipment	Model name	Manufacturer
Water purification system	Direct-Pure EDI	Microsep (Johannesburg, South Africa)
Centrifuges	Z216-MK Refrigerated Microcentrifuge	Hermle Centrifuges (New Jersey, USA)
	Legend XTR refrigerated centrifuge	Thermo Fischer Scientific (Massachusetts, USA)
Spectrophotometers	NanoPhotometer®	Implen, Inc. (California, USA)
	Bibby Scientific 73 Series Spectrophotometers	Keison Products (Essex, United Kingdom)
Incubator	Benchmark Incu-Shaker Mini CO2 Shaking Incubator	Parco Scientific (Michigan, USA)
Shaker	SHK-O0220II Orbital Shaker	Infitek (Shandong, China)
	SHK-R0525 Rocker Shaker	
Gel imager	ChemiDoc XRS+ System	Bio-Rad Laboratories Inc. (Cape Town, South Africa)
Protein vertical electrophoresis Cell	Mini-PROTEAN® Tetra Cell	Bio-Rad Laboratories (Cape Town, South Africa)
	Bio-Rad's PowerPac™	
DNA horizontal electrophoresis cell	wide Mini-Sub cell GT DNA electrophoresis systems	
-80 Freezer	TSX Series Ultra-Low Freezers	Thermo Fisher Scientific (Massachusetts, USA)
-20 Freezer	KIC KCG 570/2 WH	Makro (Limpopo, South Africa)
Sonicator	QSonica Sonicators mod. 500 ultrasonic disintegrator w. probe	QSonica (Connecticut, USA)
Mass balance	GR Series Analytical Electronic Balances	A&D Company (Tokyo, Japan)
pH meter	Dedicated PH/ORP Meter - HI2002	Hanna instruments (Cape Town, South Africa)

Table A2: Expendable items and their suppliers

Expendables	Supplier
Agar plates	Merck, Germany
Eppendorf tubes	Sigma Thermo Scientific, USA
Gas canisters	Whitehead scientific, South Africa
Gloves	Merck, Germany
Lighter	Shoprite, South Africa
Microcentrifuge tubes	Thermo Fischer Scientific, South Africa
Parafilm	Ideal medical solutions, South Africa
PCR tubes	Thermo Fischer Scientific, South Africa

Table A3: Reagents and their suppliers

Reagents	Supplier
Acetic acid	Merck, Germany
Adenosine triphosphate	Sigma Thermo Scientific, USA
Agarose	Whitehead scientific, South Africa
Ammonium molybdate	Merck, Germany
Ammonium persulphate	Merck, Germany
Ampicillin	Sigma Thermo Scientific, USA
Ascorbic acid	B & M Scientific, South Africa
Bovine serum albumin	Sigma Thermo Scientific, USA
Bromophenol blue	Sigma Thermo Scientific, USA
Calcium chloride	Merck, Germany
Chloramphenicol	Sigma Thermo Scientific, USA
Coomassie brilliant blue R250	Merck, Germany
Dithiothreitol	Sigma Thermo Scientific, USA
Ethidium bromide	Sigma Thermo Scientific, USA
Glacial acetic acid	Merck, Germany
Glycerol	Merck, Germany
Glycine	Merck, Germany
Guanidine hydrochloride	Merck, Germany
HEPES	Merck, Germany
Imidazole	Sigma Thermo Scientific, USA
Isopropyl-1-thio-D-galacopyranoside	Sigma Thermo Scientific, USA
Lysozyme	Merck, Germany
Magnesium chloride	Merck, Germany
Methanol	Merck, Germany
Ni-NTA resin	Sigma Thermo Scientific, USA
Nitrocellulose membrane	Invitrogen, USA
PageRuler Prestained Protein Ladder	Thermo Fischer Scientific, USA
Peptone	Merck, Germany
Phenylmethylsulfonyl fluoride	Sigma Thermo Scientific, USA
Polyacrylamide	Merck, Germany
Potassium chloride	Merck, Germany
Potassium dihydrogen phosphate	Merck, Germany
Restriction enzymes	Thermo Scientific, USA
Sodium chloride	Merck, Germany
Sodium dodecyl sulphate	Merck, Germany
Sodium hydroxide	Merck, Germany
TEMED	VWR Life Science, USA
Tris	Merck, Germany
Triton X-100	Merck, Germany
Tryptone	Merck, Germany
Tween 20	Merck, Germany
Urea	Merck, Germany
Yeast extract powder	Merck, Germany
β -mercaptoethanol	Sigma Thermo Scientific, USA

Appendix B: Copyright attributions

Table B1: Copyright attributions for reprinted or adapted figures

Figure	Source information	Copyright statement
Figure 1.1. The phylogenetic relationship among the 5 Plasmodium species responsible for causing malaria in humans	From Letunic, I. and Bork, P., 2021. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic acids research, 49(W1), pp.W293-W296.	This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact moc.puo@snoissimrep.slanruoj
Figure 1.2. Malaria parasite life cycle	Adapted from Nureye, D. and Assefa, S., 2020. Old and recent advances in life cycle, pathogenesis, diagnosis, prevention, and treatment of malaria including perspectives in Ethiopia. The Scientific World Journal, 2020, pp.1-17.	Copyright © 2020 Dejen Nureye and Solomon Assefa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Figure 1.3. Gametocyte stages of <i>P. falciparum</i>	Adapted from Carter, R. and Miller, L.H., 1979. Evidence for environmental modulation of gametocytogenesis in <i>Plasmodium falciparum</i> in continuous culture. Bulletin of the World Health Organization, 57, pp.37-52.	Articles and other material in Europe PMC usually include an explicit copyright statement. In the absence of a copyright statement, users should assume that standard copyright protection applies, unless the article contains an explicit statement to the contrary. In case of doubt, contact the journal publisher to verify the copyright status of an article.

Appendix C: Bioinformatics supplementary data

Table C1: Homologue Protein hits from BLASTP

Description	JDP subclass	Total score	Query cover (%)	E value	Percent identity (%)	Amino acid length	Accession number
1. DnaJ protein, putative [<i>Plasmodium falciparum</i> 3D7] (Query)	Type IV	716	100	0.0	100	362	XP_966193.1
2. DnaJ domain containing protein [<i>Plasmodium gonderi</i>]	Type IV	437	96	2e-149	64.76	377	XP_028544409.1
3. DnaJ domain containing protein [<i>Plasmodium cynomolgi</i> strain B]	Type IV	424	89	9e-145	65.23	362	XP_004223295.1
4. DnaJ domain containing protein [<i>Plasmodium vivax</i>]	Type IV	403	88	1e-136	61.99	358	XP_001616220.1
5. DnaJ domain containing protein [<i>Theileria equi</i> strain WA]	Type III	119	70	4e-27	31.13	279	XP_004829319.1
6. DNA RNA-binding Alba [<i>Babesia ovis</i>]	Type III	121	69	7e-26	29.37	847	GFE54964.1
7. DnaJ domain-containing protein [<i>Toxoplasma gondii</i> CAST]	Type III	114	71	7e-25	30.19	333	RQX69204.1
8. DnaJ domain-containing protein [<i>Toxoplasma gondii</i> TgCATBr9]	Type III	114	71	8e-24	29.81	442	PUA87364.1
9. DnaJ domain-containing protein [<i>Besnoitia besnoiti</i>]	Type III	110	69	8e-23	28,96	433	XP_029216893.1
10. DnaJ domain-containing protein, partial [<i>Cardiosporidium cionae</i>]	Type III	105	68	1e-20	30,35	531	KAF8819405.1
11. DnaJ domain-containing protein [<i>Babesia caballi</i>]	Type III	63.2	46	1e-06	30.2	552	GIX63637.1
12. DnaJ homolog subfamily C member 11-like [<i>Actinia tenebrosa</i>]	Type IV	63,2	46	1-06	30,23	552	XP_031556851.1
12. DnaJ homolog subfamily C member 11-like [<i>Octopus sinensis</i>]	Type IV	62	44	3e-06	32.14	562	XP_029651609.1
13. DnaJ homolog subfamily C member 11 isoform X1 [<i>Octopus bimaculoides</i>]	Type IV	61.2	42	6e-06	33.33	558	XP_052832966.1
14. DNAJC11 [<i>Sepia pharaonis</i>]	Type IV	59.3	47	1e-05	31.28	339	CAE1328129.1
15. dnaJ homolog subfamily C member 11 isoform X2 [<i>Octopus bimaculoides</i>]	Type IV	58,9	47	4e-05	32,92	553	XP_014784866.1

Table C2: Homologue Protein hits from HHPRED

Description	JDP subclass	Prob	E value	P value	Score	SS	Cols	Query	HMM
1. 6U3S_A DnaJ homolog subfamily [<i>Homo sapiens</i>]	Type II	99,5	4.2e-14	6.4e-19	109.0	9.2	69	21-89	2-72 (190)
2. 3U03_B J-type co-chaperone JAC [<i>Saccharomyces cerevisiae</i>]	Type II	99.5	6.6e-13	1e-17	100.2	13	124	15-140	4-156 (181)
3. 4IT5_D Co-chaperone protein Hs [<i>Vibrio cholerae</i>]	Type II	99.5	9e-13	1.4e-17	98.1	13.5	123	20-142	2-153 (174)
4. 1FPO_B Chaperone protein HSCB [<i>Escherichia coli</i>]	Type II	99.5	1.6e-12	2.5e-17	98.1	13.0	110	22-131	1-131 (171)
5. 7VPX_J DnaJ homolog subfamily [<i>Homo sapiens</i>]	Type II	99.4	3.2e-11	4.9e-16	95.2	15.2	125	6-130	40-182 (253)
6. 5Y88_T Pre-mRNA-splicing factor [<i>Saccharomyces cerevisiae</i>]	Type II	99.4	3.2e-12	4.8e-17	104.8	9.4	90	19-108	11-108 (283)
7. 7X3K_A Zuotin; RAC, co-translation [<i>Saccharomyces cerevisiae</i>]	Type II	99.3	2.7e-12	4.2e-17	113.6	7.5	72	20-91	94-172 (433)
8. 8GZU_j1 DnaJ domain protein; E [<i>Tetrahymena thermophila</i>]	Type II	99.3	1.1e-11	1.7e-16	104.8	10.8	90	1-91	3-97 (317)
9. 7AOI_BK Chaperone protein DnaJ [<i>Trypanosoma brucei</i>]	Type II	99.3	1.1e-10	1.6e-15	96.4	15.3	111	21-131	1-166 (273)
10. 2L6L_A DnaJ homolog subfamily [<i>Homo sapiens</i>]	Type III	99.3	2.8e-11	4.3e-16	85.5	9.0	73	15-87	3-82 (155)
11. 6YXY_BK mL77; mitoribosome, as [<i>Trypanosoma brucei</i>]	Not characterized	99.3	4.8e-11	7.3e-16	103.5	11.2	114	16-129	79-247 (386)
12. 7Z3O_C Putative ribosome association [<i>Thermochaetoides thermophila</i>]	Not characterized	99.3	9.9e-12	1.5e-16	109.1	7.1	72	20-91	99-176 (446)
13. 2PF4_E Small T antigen; PP2A, [<i>Mus musculus</i>]	Not characterized	99.3	7.1e-12	1.1e-16	96.6	4.8	67	15-84	4-72 (174)
14. 7KAM_D Transport protein HSCB; [<i>Thermomyces lanuginosus</i>]	Not characterized	99.2	3.9e-11	5.9e-16	109.1	9.3	73	17-89	115-195 (719)
15. 7ZHS_Y Ubiquitin-like protein [<i>Saccharomyces cerevisiae</i>]	Type I	99.2	6e-11	9.2e-16	102.2	9.2	83	3-88	97-179 (497)

Appendix D: Experiments supplementary data

Appendix D1. Plasmid confirmations of *pQE30*, *pQE30/PfHsp70-1*, and *pQE30/PfHsp40*

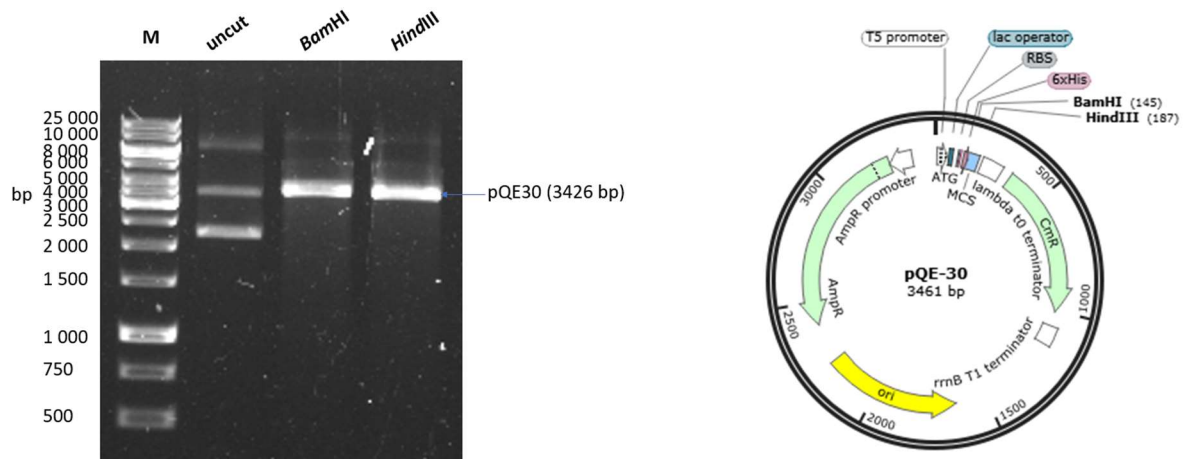


Figure D1.1. Restriction endonuclease cleaving of pQE30.

(A) 0.8 % agarose electrophoresis gel depicting the migration and separation of DNA fragments generated from digesting pQE30. Lane M is the molecular marker lane, uncut lane is the pQE30 uncut, BamHI is single digest with the enzyme, and HindIII is digestion with HindIII. Plasmid DNA fragments labelled using arrows and the estimated sizes in brackets. The molecular marker is labelled in base pairs (bp). (B) The pQE30 plasmid highlighting the cut sites and size markers.

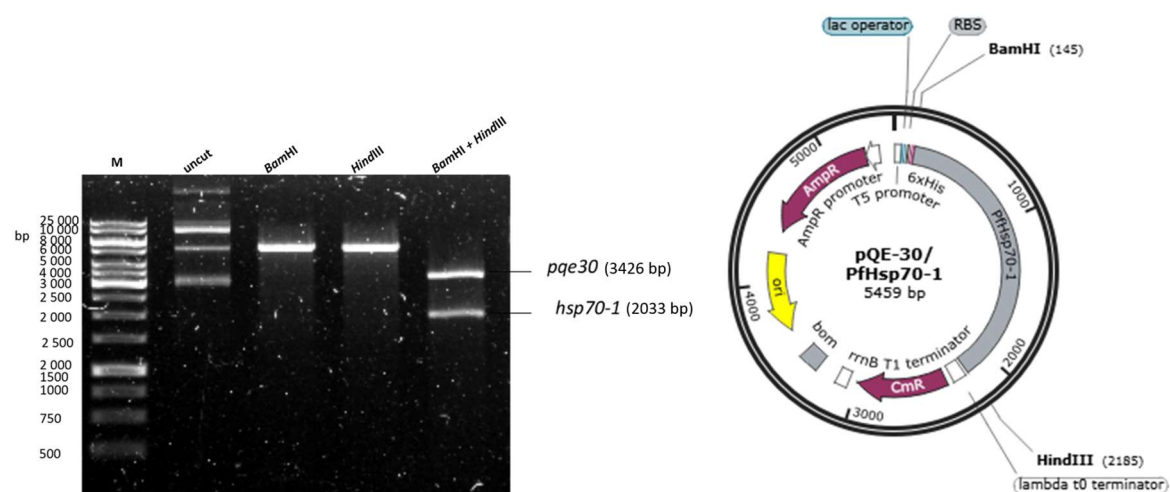


Figure D1.2. Restriction endonuclease cleaving of pQE30/PfHsp70-1.

(A) 0.8 % agarose electrophoresis gel depicting the migration and separation of DNA fragments generated from digesting pQE30/PfHsp70-1. Lane M is the molecular marker lane, uncut lane is the pQE30/PfHsp70-1 uncut, BamHI is single digest with the enzyme, HindIII is digestion with HindIII, and lane BamHI + HindIII is double digest with both enzymes. Plasmid DNA fragments labelled using arrows and the estimated sizes in brackets. The molecular marker is labelled in base pairs (bp). (B) The pQE30/PfHsp70-1 plasmid construct highlighting the cut sites and size markers.

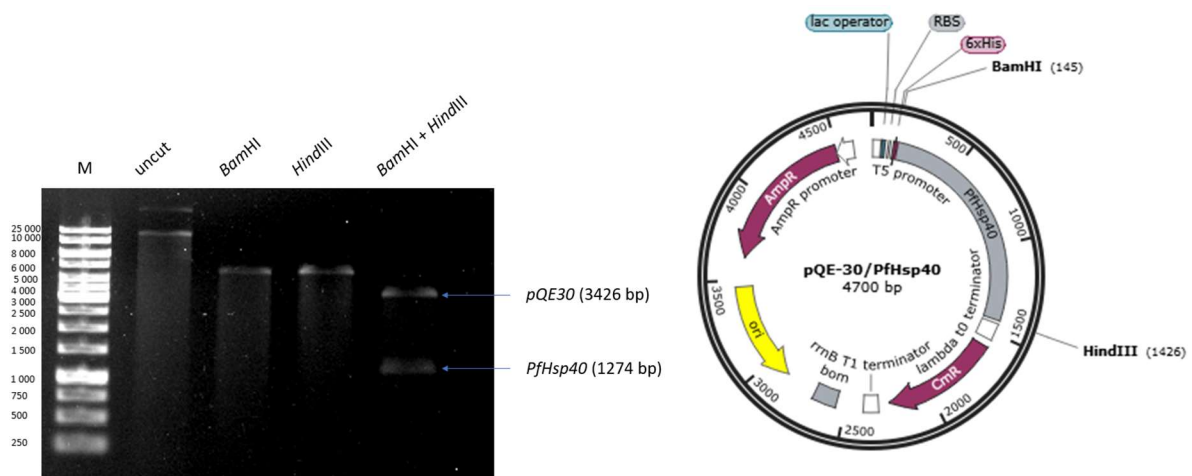


Figure D1.3. Restriction endonuclease cleaving of pQE30/PfHsp40.

(A) 0.8 % agarose electrophoresis gel depicting the migration and separation of DNA fragments generated from digesting pQE30/PfHsp40. Lane M is the molecular marker lane, uncut lane is the pQE30/PfHsp40 uncut, BamHI is single digest with the enzyme, HindIII is digestion with HindIII, and lane BamHI + HindIII is double digest with both enzymes. Plasmid DNA fragments labelled using arrows and the estimated sizes in brackets. The molecular marker is labelled in base pairs (bp). (B) The pQE30/PfHsp40 plasmid construct highlighting the cut sites and size markers.

Appendix D2. SDS-PAGE and Western blot for overexpression and purification of *pQE30/PfHsp70-1*, and *pQE30/PfHsp40*

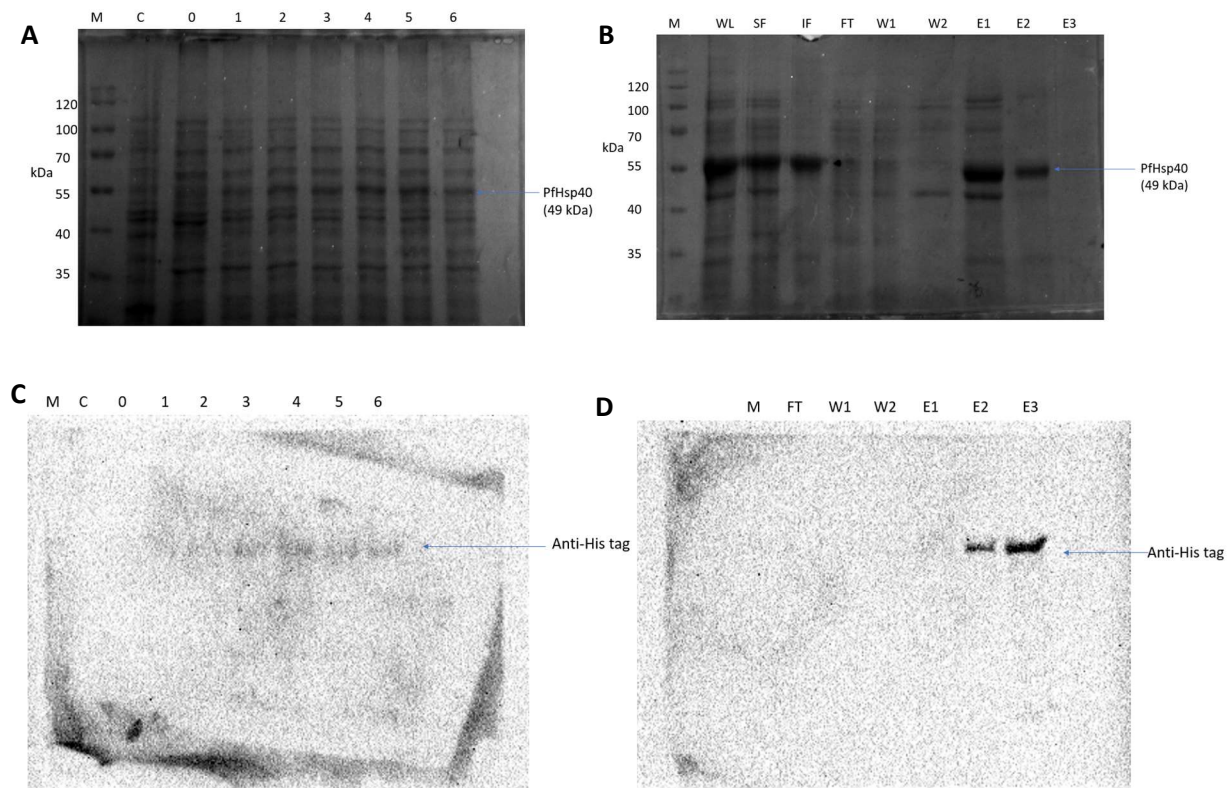


Figure D2.1. SDS-PAGE and Western blots for the expression and purification of PfHsp40.

(A) SDS-PAGE for the expression of PfHsp40 (PF3D7_1437900). Values 0 to 6 are hourly expressions (0 hour to 6 hours) of PfHsp40 by *E. coli* XL1 Blue cells. M represent the molecular marker and C the control cells transformed with the plasmid vector pQE30 carrying the PfHsp40 (PF3D7_1437900) gene (B) SDS-PAGE for the purification of PfHsp40 (PF3D7_1437900). FT is the flow through, W for washes and E for elutions. Coomassie Brilliant blue stain was used for SDS-PAGE (C) Western blot for PfHsp40 expression using α -polyhistidine tag primary antibodies. (D) Western blot for PfHsp40 (PF3D7_1437900) purification using α -polyhistidine primary antibodies. The primary antibodies were rabbit-raised while secondary, HRP-conjugated antibodies were mouse-raised. The molecular weights are in kilodaltons. The identified protein was labelled with the estimated molecular weight.

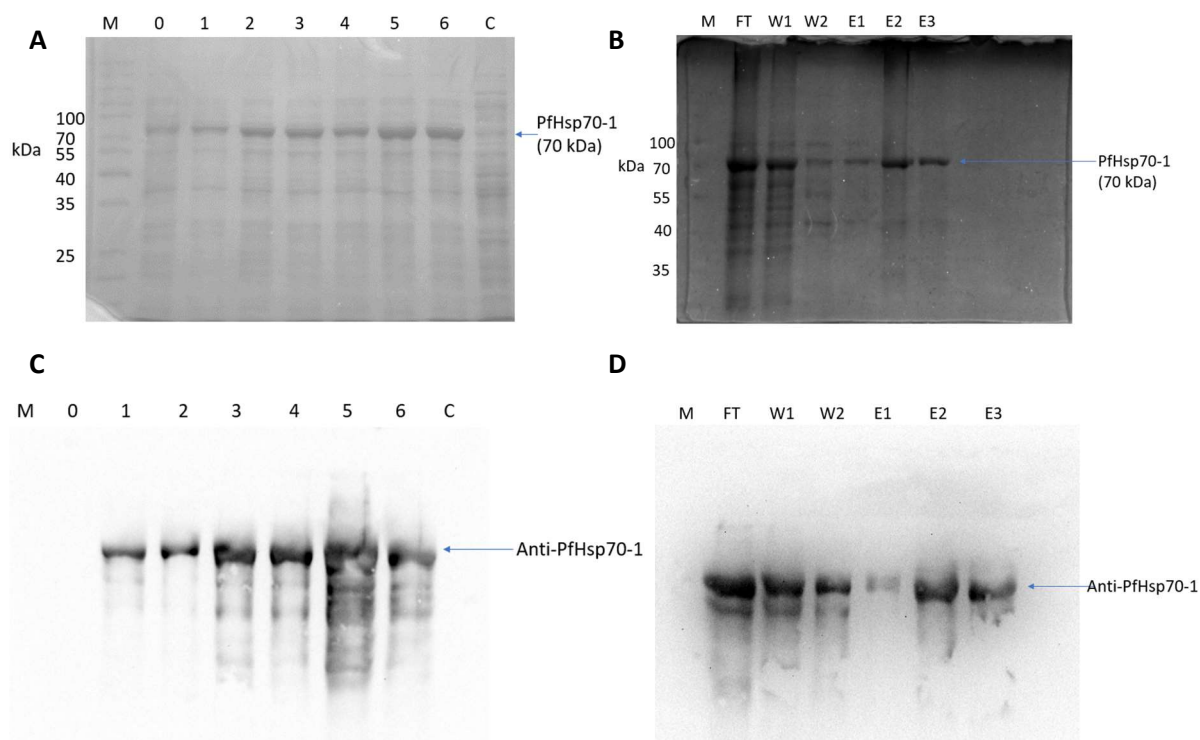


Figure D2.2. SDS-PAGE and Western blots for the expression and purification of PfHsp70-1.

(A) SDS-PAGE for the expression of PfHsp70-1. Values 0 to 6 are hourly expressions (0 hour to 6 hours) of PfHsp70-1 by *E. coli* XL1 Blue cells. M represent the molecular marker and C the control cells transformed with the plasmid vector pQE30 carrying the PfHsp70-1 gene (B) SDS-PAGE for the purification of PfHsp70-1. FT is the flow through, W for washes and E for elutions. Coomassie Brilliant blue stain was used for SDS-PAGE (C) Western blot for PfHsp70-1 expression using α -PfHsp70-1 primary antibodies. (D) Western blot for PfHsp70-1 purification using α -PfHsp70-1 primary antibodies. The primary antibodies were rabbit-raised while secondary, HRP-conjugated antibodies were mouse-raised. The molecular weights are in kilodaltons. The identified protein was labelled with the estimated molecular weight.

Appendix D3. SDS-PAGE for purification of PFF1010c

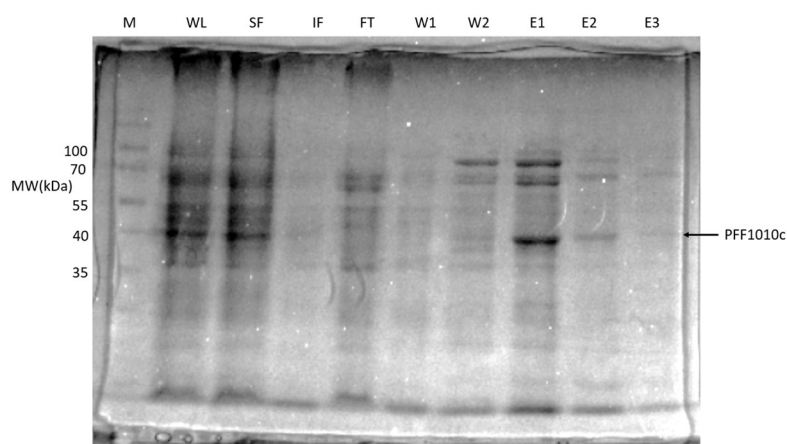


Figure D3. SDS-PAGE for purification of PFF1010c.

M represent the molecular marker and C the control cells transformed with the plasmid vector pQE30 carrying the PfHsp70-1 gene (B) SDS-PAGE for the purification of PfHsp70-1. FT is the flow through, W for washes and E for elutions. Coomassie Brilliant blue stain was used for SDS-PAGE.

Appendix D4. BeStSel data analysis of PFF1010c CD spectra

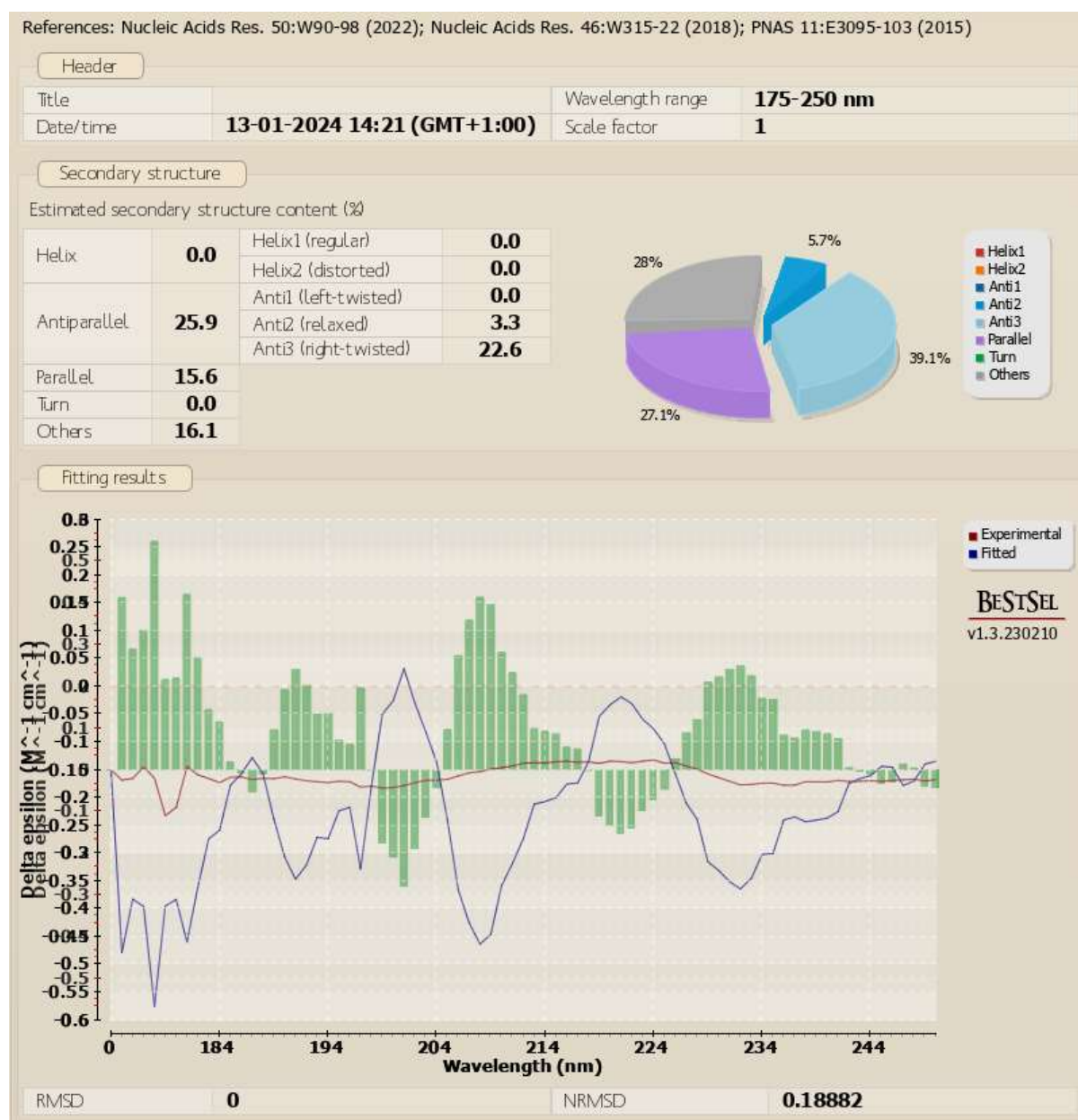


Figure D4. BeStSel analysis chart of PFF1010c CD spectra.

Raw CD spectra (red) data was fitted (blue) and PFF1010c secondary structure was estimated based on the fitted data. Secondary structure conformational states content is shown.