

Title: The generation of Plasmodium falciparum Hsp90 selective inhibitors using rational drug design

T Matlhodi



orcid.org0000-0001-9763-0567

Dissertation accepted in fulfilment of the requirements for the
degree master of science in biochemistry at the
North West University

Supervisor: Dr F Mokoena

Co-supervisor: Dr N Gumede

Graduation ceremony: May 2024

DECLARATION

I, Thato Matlhodi, hereby declare that this dissertation, titled: “The generation of Plasmodium falciparum Hsp90 selective inhibitors using rational drug design”, is the result of my research and academic work undertaken towards the completion of my Master of Science (MSc) degree at North-West University.

I confirm that this dissertation represents my original work and that it has not been previously submitted for any other degree, diploma, or academic qualification. Any sources of information used in the dissertation have been duly acknowledged through proper citation and referencing, in accordance with the guidelines provided by my university.

I further declare that all the primary data and findings presented in this dissertation are authentic, and any contributions made by others to this research have been duly recognized through acknowledgment or co-authorship, as applicable.

Throughout the process of conducting this research, I have adhered to the highest ethical standards. I have obtained the necessary permissions, where required, to access and use any data, materials, or resources employed in this study. I have also maintained confidentiality and ensured the privacy of any individuals or organizations involved in this research.

I declare that my dissertation supervisors, Dr Fortunate Mokoena and Dr Njabulo Gumede, have provided guidance and support throughout the research process. Their expertise and valuable feedback have significantly contributed to the successful completion of this dissertation.

I understand that any attempt to misrepresent the content or veracity of this dissertation may result in the imposition of penalties, including the withholding of my degree.

By submitting this dissertation, I affirm that I have read, understood, and complied with all the relevant regulations and requirements set forth by my university and the department responsible for my MSc program.



Thato Matlhodi

27325903

19/05/2023

DEDICATION

I dedicate this dissertation to my beloved baby girl, Letlotlo Bokgabo Mongale. You have been my constant source of inspiration and motivation throughout this academic journey. As I worked tirelessly to complete this MSc degree, I was driven by the desire to show you that with determination and perseverance, dreams can become a reality. You have filled my life with boundless love and joy, and I dedicate this achievement to you, my precious daughter. May this serve as a testament to the power of hard work and the infinite possibilities that lie ahead for you.

I would also like to dedicate this dissertation to my parents, Khumo and Puoetsile Matlhodi. Your unwavering faith in my abilities and unwavering support have been invaluable throughout my academic pursuits. You have instilled in me a strong work ethic and taught me the importance of never giving up on my dreams. Your constant encouragement and belief in my potential have been the guiding lights that have led me to this point. This achievement is as much yours as it is mine, and I am forever grateful for your love, guidance, and sacrifices.

To Letlotlo Bokgabo Mongale and my parents, Khumo and Puoetsile Matlhodi, I dedicate this dissertation. May it serve as a testament to the power of love, determination, and unwavering support. Your presence in my life has made all the difference, and I am honored to have you by my side on this remarkable journey.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the National Research Foundation (NRF) and the NWU Masters' bursary for providing the financial support that made it possible for me to pursue this degree. Their funding has been instrumental in enabling me to focus on my studies and research, and I am deeply appreciative of their investment in my academic journey.

I extend my heartfelt appreciation to my supervisor, Dr Fortunate Mokoena, for her invaluable guidance, support, and belief in my abilities. Her expertise, dedication, and motivation have been essential in shaping the direction and success of my MSc research. I am truly grateful for the opportunities provided and the wisdom shared throughout this process.

I would also like to acknowledge and thank my co-supervisor, Dr. Njabulo Gumede, for his contributions and assistance as a co-supervisor. His input and expertise have enriched my research and broadened my perspective, and I am thankful for his valuable insights and guidance.

To my parents, I am forever grateful for your unwavering support, love, and faith in me. Your encouragement and belief in my abilities have been the driving force behind my accomplishments. Your constant presence and reassurance have provided the foundation upon which I have built my academic journey. Thank you for always being there for me.

I extend my appreciation to my lab mates, Tlhalefo Ntseane and Ofentse Mafethe. We have endured challenges together, celebrated successes, and shared countless hours of hard work and dedication in the Parasite Drug Discovery Lab. Your friendship, collaboration, and mutual support have made this journey more meaningful and enjoyable. I hope that we all continue to thrive and succeed in our future endeavors.

To my partner, Reabetswe Paballo Mongale, thank you for being my rock and constant source of inspiration. Your unwavering support, love, and understanding have sustained me throughout this demanding period of my life. You have been my pillar of strength, lifting my spirits when they were low and believing in me when I doubted myself. I am grateful to have you by my side.

I would like to acknowledge my best friends, Tlotlo Montwedi and Kgalalelo Letimela, for their support, encouragement, and friendship. Your presence and belief in my abilities have been invaluable. Thank you for standing by me and cheering me on every step of the way.

To all those mentioned above and to the countless others who have supported me along this journey, thank you from the bottom of my heart. Your contributions, encouragement, and belief in me have made this accomplishment possible. I am forever grateful for your presence in my life.

OUTPUTS

The findings of this study were showcased at the:

South African Society of Biochemistry and Molecular Biology SASBMB conference in January 2022, receiving recognition for their contribution to the field of bioinformatics and biochemistry.

- **Matlhodi Thato**, Gumede Njabulo, Choene Mpho and Mokoena Fortunate. the use of machine learning for generation of *Plasmodium falciparum* Hsp90 selective inhibitors. 27th congress of South African biochemistry and molecular biology. Virtual event (23-26 January 2022)

Furthermore, the study featured a poster presentation at an H3D symposium held in October 2022.

- **Matlhodi Thato**, Gumede Njabulo, Choene Mpho and Mokoena Fortunate. the use of machine learning for generation of *Plasmodium falciparum* Hsp90 selective inhibitors. 27th congress of South African biochemistry and molecular biology. Webersburg estate, Stellenbosch (25-28 October 2022).

The work is being prepared for publication

- **Thato Matlhodi**, Lisema Patrick Makatsela, Tendamudzimu Harmfree Dongola, Addmore Shonhai, Njabulo Joyfull Gumede and Fortunate Mokoena. Auto QSAR-based Active learning docking for hit identification of potential inhibitors of *P. falciparum* Hsp90 as antimalarial agents.

ABSTRACT

Malaria, which is mainly caused by *Plasmodium falciparum* parasite, remains a devastating disease that poses a significant burden on the public health systems and socio-economic development of countries globally. This is particularly true in regions with limited resources, such as sub-Saharan Africa. Drug-resistant strains of the *P. falciparum* parasite have significantly restricted effective treatment options, necessitating the development of new and effective antimalarial drugs that can overcome drug resistance. The molecular chaperone, heat shock protein 90 (Hsp90), plays a critical role in parasite survival and proliferation. Therefore, it is a promising target for the development of new antimalarial drugs. The generation of selective inhibitors is important for target specificity, therapeutic efficacy, safety, and drug tolerability. Using rational drug design strategies, this study aimed to generate selective inhibitors of *P. falciparum* Hsp90 (PfHsp90) through molecular modelling, facilitated by an active learning approach, followed by biochemical and biophysical assays. This study used a combination of computational techniques, namely reaction-based enumeration, active learning, induced fit docking, molecular dynamics simulations, absorption, distribution, metabolism, excretion, and toxicity (ADME/T) predictions, and molecular mechanics with generalized born surface area (MM-GBSA) to generate a total of 10 000 drug-like design ideas. Furthermore, a large compound library from chEMBL was used to identify compounds with high affinity and selectivity for the target protein. A total of 13 models were built using active learning autoQSAR, which achieved good statistical results of $R^2 = 0.69$, $Q^2 = 0.61$, and $RMSE = 0.44$, on average for all models. The current study then assessed the ability of a subset of the purchased compound to inhibit the ATPase activity of PfHsp90 using an ATPase assay and protein-ligand interaction measured using surface plasmon resonance (SPR). Compounds FTN-T4 and FTN-T10 brought the most robust inhibition of PfHsp90's ATPase activity, greater than that of the positive control harmine, which is in correlation with the *in-silico* data, while the majority of the compounds displayed a similar inhibition of the ATPase activity to harmine. In the biophysical study, compound FTN-T2 displayed moderate binding with a K_D of 7 μ M, while FTN-T10 and FTN-T5 displayed binding affinities comparable to that of ADP, at K_D values of 13 μ M and 19 μ M, respectively. The results presented in this study demonstrate the great potential of active learning models to generate novel *P. falciparum* Hsp90 inhibitors. The identified compounds were found to disrupt the ATPase activity of PfHsp90, form sufficient interactions between the protein and ligand, possess a good safety profile and selectivity, and represent promising candidates for further lead optimization. By targeting PfHsp90, these inhibitors offer potential avenues for overcoming drug resistance and improving the efficacy of antimalarial therapies.

Keywords: Malaria; *Plasmodium falciparum*; Heat shock protein 90; active learning

TABLE OF CONTENTS

DECLARATION	II
DEDICATION	III
ACKNOWLEDGEMENTS	IV
OUTPUTS	V
ABSTRACT	VI

CHAPTER 1 16

1. INTRODUCTION	16
Hypothesis	19
Aim and objectives.....	19
1.1 Aim:	19
1.2 Objectives:.....	19

CHAPTER 2 20

2. LITERATURE REVIEW.....	20
2.1 Malaria	20
2.2 <i>Plasmodium falciparum</i> life cycle	21
2.3 Uncomplicated and severe form of malaria.....	23
2.4 Past and present antimalarial treatments	24
2.5 Molecular chaperones as drug targets	27
2.6 Hsp90	29
2.7 Mechanism of Hsp90 ATPase activity.....	30
2.8 Small molecule inhibitors of Hsp90	31
2.9 <i>P. falciparum</i> Hsp90 (PfHsp90) as a drug target	32
2.10 Drug discovery	33
2.10.1 Rational drug design.....	34
2.10.2 Application of molecular docking in drug discovery.....	35

2.10.3 Molecular dynamic simulation	36
2.10.4 Molecular Mechanics/Generalized Born Surface Area (MMGBSA)	37
2.10.5 Machine learning (ML) in drug discovery	38
2.10.6 Quantitative structure activity relationship (QSAR)	38
2.11 ATPase activity assay	39
2.12 Biophysical methods in early drug discovery	40
2.12.1 Surface Plasmon resonance	40

CHAPTER 3 42

3. MATERIALS AND METHODS	42
3.1. <i>In silico</i> methods	42
3.1.1 Computational software programme	42
3.1.2 Protein retrieval and preparation	44
3.1.3Ligand preparation.....	44
3.1.4 Induced fit docking (IFD).....	44
3.1.5 Reaction-based enumeration.....	45
3.1.6 Classical glide SP ligand docking	46
3.1.7Receptor grid generation	47
3.1.8 Active learning model generation.....	47
3.1.9 <i>In silico</i> ADME predictions	48
3.1.10 Free binding energy calculations	49
3.1.11 Molecular dynamic simulation	50
3.2 Biochemical methods.....	50
3.2.1 Transformation	50
3.2.2 Plasmid extraction	50
3.2.3 Restriction enzyme digest.....	51
3.2.4 Protein expression	51
3.2.5 Protein purification	51
3.2.6 Protein concentration determination	52
3.2.7 ATPase activity assay.....	53
3.3. Biophysical methods.....	53
3.3.1 Surface plasmon resonance (SPR).....	53

CHAPTER 4 55

4.1 RATIONAL DESIGN RESULTS AND DISCUSSIONS.....	55
4.1.1 Reaction-based enumeration.....	55
4.1.2 Active learning model enhanced by ligand docking.	57
4.1.4 Induced fit docking.....	60
4.1.5 ADME/T property predictions	65
4.1.6 Molecular dynamics simulations	69
4.1.3 Binding free energy calculations.....	76

CHAPTER 5 78

5.1 BIOCHEMISTRY RESULTS AND DISCUSSIONS.....	78
5.1.1 Confirming the integrity of the pET28b (+)-PfHsp90-NTD construct.....	78
5.1.2 Recombinant PfHsp90 expression	78
5.1.3 Protein purification.....	79
5.1.4 ATPase assay	81

CHAPTER 6 84

6.1 BIOPHYSICAL ANALYSIS	84
6.1 Surface plasmon resonance (SPR)	84

CHAPTER 7 88

CONCLUSION	88
Future recommendations.....	89
BIBLIOGRAPHY	90
APPENDIX.....	102
1.1.3 Reagents.....	102
1.1.4 Plasmid	103
1.1.5 Transformation	103
1.1.6 Pure-yield plasmid extraction kit	103

1.1.7 Restriction digest	105
1.1.8 Agarose gel electrophoresis	106
1.1.9 Sodium dodecyl-polyacrylamide gel electrophoresis (SDS-PAGE)	106
1.1.10 Western blotting	108

RESULTS 109

2.1 Molecular dynamic simulations.....	111
---	------------

112

LIST OF TABLES

Table 1: Pathfinder generated reaction pathways -Schrödinger Suite 2021-3	45
Table 2: IFD results of the reference compound 7 represented as CP-7 and compound 10 as CP-10.....	55
Table 3: The best Auto QSAR models were used to train the listed design concepts to predict their docking scores.....	58
Table 4: induced fit docking results of the purchased compounds showing Docking score, glide energy, glide emodel of Induced fit docked complexes.	61
Table 5: in silico physio-chemical property predictions by Qikprop	66
Table 6: Binding free energy calculation for investigated compounds using Prime MM- GBSA.....	76
Table 7: Kinetics constants measured by SPR for the interaction between tested compounds and immobilized PfHsp90.....	84

LIST OF FIGURES

Figure 1: Life cycle of <i>P. falciparum</i>	22
Figure 2: Chemical structures of the past antimalarials.....	25
Figure 3: Diagrammatic sketch of the full length <i>P. falciparum</i> Hsp90.....	29
Figure 4: Model of the Hsp90 ATPase cycle showing the obligatory sequential hydrolysis and conformational switching.	30
Figure 5: <i>in silico</i> methodology flow diagram	43
Figure 6: Oxadiazole-1 selected pathway.	46
Figure 7: reference compounds.....	55
Figure 8: Details the ligand interaction diagrams for compound 10 and 7 in the PfHsp90 N-terminal domain ATP binding site..	56
Figure 9: Auto QSAR active learning scatter plot for model 1 showing the performance of the QSAR KPLS model's predicting activity for experimental binding affinity for the test set.	59
Figure 10: Auto QSAR active learning plots for model 2	59
Figure 11: Auto QSAR active learning plots for model 13 with an R^2 of 0.7460 and a Q^2 of 0.6182.....	60
Figure 12: PfHsp90 (PDBcode:3K60) interaction with six selected compounds.	64
Figure 13: Compound structures analysed by Qikprop for the <i>in silico physio-chemical</i> <i>property predictions</i>	69
Figure 14: Protein-ligand RMSD plots of all the frames in the trajectory for PfHsp90-NTD and Compounds complexes at a 50ns MDS trajectory and FTN-T2 at 150ns.....	Error! Bookmark not defined.
Figure 15: 2D Protein-ligand interaction diagrams of all the frames in the trajectory for PfHsp90-NTD and Compounds complexes at a 50ns MDS trajectory and FTN-T2 at 150ns.....	74

Figure 16: Restriction enzyme digest analysis for pET28b (+)-PfHsp90-NTD plasmid construct, ...	78
Figure 17: PfHsp90-NTD expression on BL21(DE3) E. coli cells analysis.....	79
Figure 18: SDS-PAGE analysis of PfHsp90 protein purification samples.	80
Figure 19: ATPase activity of PfHsp90 against our inhibitors. Compounds were added in increasing concentrations (0.25–20 μ M) to a fixed amount of PfHsp90 (2 μ M) to investigate its effect on the ATPase activity of PfHsp90 with Harmine as our positive control.	83
Figure 20: Sansogram from the Analysis of the interaction of FTN-T2, FTN-T5, FTN-T10, geldanamycin and Harmine with PfHsp90-NTD by surface plasmon resonance spectroscopy.	86

LIST OF ABBREVIATIONS

• 2D	Two dimensional
• 3D	Three dimensional
• ACTs	Artemisinin-based combination therapies
• ADME	Absorption, distribution, metabolism, and excretion
• ADP	Adenosine diphosphate
• AL	Active learning
• ATP	Adenosine triphosphate
• BSA	Bovine serum albumin
• CD	Circular dichroism
• CDC	Centre for Disease Control
• CTD	C-terminal domain
• DMSO	Dimethyl sulfoxide
• DNA	Deoxyribonucleic acid
• DTT	Dithiothreitol
• EC ₅₀	Half maximal effective concentration
• EGFR	Epidermal growth factor receptors
• GA	Geldanamycin

• PDGFRA	Platelet Derived Growth Factor Receptor Alpha
• PfCRT	Chloroquine resistance transporter
• PfHsp90	<i>P. falciparum</i> Heat shock protein 90
• Pi	Inorganic phosphate
• PMSF	Phenylmethylsulfonyl fluoride
• QSAR	Quantitative structure-activity relationship
• RDD	Rational drug design
• RNA	Ribonucleic acid
• ROS	Reactive oxygen species
• SASBMB	South African society of biochemistry and molecular biology
• SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
• SMARTS	Smiles arbitrary target specification
• SP	Standard precision
• SPR	Surface plasmon resonance
• TPR	Tetratricopeptide repeat
• TPSA	Topological polar surface area
• WHO	World health organization
• XP	Extra precision
• YT	Yeast tryptone

CHAPTER 1

1. INTRODUCTION

Malaria is an infectious disease that is transmitted to humans by female *Anopheles* mosquito vectors. The population most at risk of malaria infection is mainly children under the age of 5 and pregnant women (Steketee *et al.*, 2001; Gontie *et al.*, 2020), with an estimated 619 000 individuals succumbing to the deadly disease in 2021 (WHO, 2022). Most of the fatalities associated with malaria are attributed to what is considered the deadliest malarial species, *Plasmodium falciparum* (Ismail *et al.*, 2002; Battle *et al.*, 2019; Weiss *et al.*, 2019). The emergence of drug-resistant parasitic strains has greatly hindered the fight against malaria, limiting the effectiveness of the current treatment options (Malaguarnera and Musumeci, 2002).

Current treatment options for malaria include vaccines, insecticides, and therapeutic drugs. However, each strategy has its own drawbacks. For instance, the most advanced malaria vaccine is Mosquirix (RTS) S/AS01 (Nadeem *et al.*, 2022), provides partial protection against *P. falciparum*. However, its efficacy is limited, and multiple doses are required to achieve a moderate level of protection. Furthermore, the duration of protection offered by vaccination is short (Laurens, 2020; Syed, 2022). Common strategies for controlling malaria-transmitting mosquitoes include insecticide-treated bed nets and indoor residual spraying (Okumu and Moore, 2011; Pryce *et al.*, 2022). While these interventions have proven effective in reducing malaria transmission, they face challenges, such as insecticide resistance in mosquitoes. Over time, mosquitoes have developed resistance to commonly used insecticides, which limits the effectiveness of this approach (Okumu and Moore, 2011). Drugs such as artemisinin-based combination therapies (ACTs) as the first-line treatment for uncomplicated malaria (WHO, 2022), face the major challenge of drug resistance, which has significantly hindered the effectiveness of these drugs, leading to treatment failure and increased mortality rates of *P. falciparum* parasites. Although vaccines, insecticides, and drugs are important tools in the fight against malaria, their limitations, and challenges, such as limited efficacy, insecticide resistance, and drug resistance, highlight the need for continuous research and development of novel and more effective strategies for malaria prevention and treatment.

Environmental stress caused by the cyclic transition from the mosquito vector to the human host of the *P. falciparum* parasite, leads to the expression of an arsenal of molecular chaperones, such as heat shock protein 90 (Hsp90), Hsp70, Hsp40 and others, to ensure parasite survival and

proliferation. These environmental stresses include changes in the pH, temperature, and nutrient supply (Pavithra *et al.*, 2004; Uwimana *et al.*, 2020). Hsp90 is one of the most important and highly expressed members of the molecular chaperones, constituting approximately 2% of the total proteome (Fink, 1999; Garnier *et al.*, 2002; Hunter *et al.*, 2014). Hsp90 is a key regulator of vital biological processes, including stress control, protein folding, DNA repair, development, immunological response, and signalling pathways (Aligue *et al.*, 1994; Pratt, 1998; Pearl, 2005; Prodromou, 2016). Hsp90 interacts with a wide range of proteins referred to as client proteins. Some of the Hsp90 client proteins are transcription factors, growth receptors, and kinases (Prodromou, 2016), making Hsp90s a desirable pharmacological target for a variety of illnesses, including cancer, neurological disorders, and infectious diseases such as malaria, owing to their crucial role in cell survival.

P. falciparum Hsp90 (PfHsp90) plays a critical role in assisting with the folding of newly synthesized proteins, regulating the actions of transcription factors and protein kinases, and acts as a buffer to mitigate the cellular toxicity caused by misfolded or aggregated proteins in response to stress, growth, and development during its host-vector cyclic infection stages (Banumathy *et al.*, 2003; Gitau *et al.*, 2012). Targeted gene mutations of some client proteins of the molecular chaperone have been shown to give rise to the development of drug resistance. For example, the Shihanas study (2013) showed that PfHsp90 interacts directly with *the chloroquine resistance transporter (crt)* (Shaninas *et al.*, 2013; Stoffberg *et al.*, 2021). Furthermore, multidrug resistance genes, such as *crt*, are located on the same gene cluster as Hsp90, which plays a central role in protein quality control (Zininga, 2015). Ashly *et al.* (2015) also linked artemisinin resistance to an enhanced stress response, implicating PfHsp90. Therefore, PfHsp90 is an attractive drug target for the development of effective and dependable antimalarial drugs that inhibit *P. falciparum* infection and growth while reducing the possibility of developing resistance to treatment.

Hsp90 is a dimeric protomer composed of three domains: the N-terminal domain (NTD), which binds ATP, connected to the middle domain (MD) through a charged linker region, and the C-terminal domain (CTD) (Jackson, 2013). NTD inhibitors of Hsp90 act by binding to the ATP-binding pocket, thereby preventing the binding of ATP to Hsp90 and inhibiting ATPase activity. This, in turn, prevents conformational changes required for Hsp90 to interact with its client proteins and promote their folding and stability. Most of the Hsp90 inhibitors target the NTD by competing with ATP for binding and therefore, abrogating the enzymatic activities of the molecular chaperone (Chène, 2002; Huang *et al.*, 2017; Wang *et al.*, 2019; Shin *et al.*, 2020; Banerjee *et al.*, 2021).

Using known human Hsp90 inhibitors to investigate the function and essentiality of PfHsp90 over the years, several studies have validated PfHsp90 as a drug target for the development of novel antimalarial drugs that would most likely circumvent resistance mechanisms (Kumar *et al.*, 2003; Banumathy *et al.*, 2003; Pallavi *et al.*, 2010; Shahinas *et al.*, 2013). Critical findings have been obtained using these known inhibitors, including differences in the binding affinities of the inhibitors to human Hsp90 and PfHsp90. Geldanamycin (GA) was found to have a higher affinity for PfHsp90 than for human Hsp90 (Kumar *et al.*, 2003). Secondly, PfHsp90 has higher ATPase activity than its human homolog, making it more susceptible to inhibition (Pallavi *et al.*, 2010; Shahinas *et al.*, 2013). Thirdly, GA was also able to halt the development of *P. falciparum* cultures in the intraerythrocytic cycle of the parasite between the ring and trophozoite stages (Banumathy *et al.*, 2003), drastically reducing the number of merozoites released into the bloodstream. The elucidation of the PfHsp90 crystal structure by Corbett and Berger in 2010 has been an important breakthrough in validating PfHsp90 as a drug target by providing more avenues to identify unique conformational differences between PfHsp90 and Hsp90s from other organisms.

The identification of this drug target offers an advantage in the implementation of rational drug design strategies, such as those documented by Egan (2005), Plouffe *et al.* (2008), Sharman *et al.* (2012), Wang *et al.* (2016), Everson *et al.* (2021), and Tamari (2022), over traditional phenotypic screens. Using this approach, inhibitors with high potency and modest affinity have been identified (Wang *et al.*, 2016; Everson *et al.*, 2021). Rational drug design strategies form the basis of multidisciplinary approaches required for the drug development process (Mandal *et al.*, 2009), overcoming the challenges encountered with traditional screening strategies, including very high costs, long turnaround times, and specificity of target proteins (Reddy and Parrill, 1999). Selective targeting is essential for drug discovery because of its increased efficacy and reduced toxicity. Broadly, rational drug design strategies can be classified into two categories: firstly, small-molecule development with desired qualities for targets, biomolecules (proteins or nucleic acids) with established functional functions in cellular processes, and 3D structural information. The second is the development of small molecules with predefined target properties (Mandal *et al.*, 2009). In recent years, to accelerate and improve drug discovery and development, machine learning (ML) techniques have been increasingly integrated into rational drug design. Therefore, this study shows the synergy between active learning-based Glide docking tools, and Pathfinder reaction-based enumeration tools, highlighting their combined effectiveness in the context of drug discovery.

The present study implemented a machine learning approach, active learning (AL), which is a classification under supervised ML, to design novel inhibitors with a starting point of previously described molecules with binding affinity towards PfHsp90 (Everson *et al.*, 2021). The active

learning approach helps develop highly accurate quantitative structure-activity relationship (QSAR) models, effectively exploring wide chemical spaces with docking and deep learning (Dixon *et al.*, 2016; Yang *et al.*, 2021), followed by validation of novel inhibitors through biochemical and biophysical methods.

Research question

Can active learning-based Glide docking be utilized to generate novel inhibitors with distinct chemical scaffolds and enhanced affinity for PfHsp90, resulting in the disruption of PfHsp90's enzymatic properties and demonstrate notable binding affinity towards it? Aim and objectives

1.1 Aim:

The aim of this study was to generate small molecules with selective inhibition of PfHsp90 using a rational drug design coupled with biochemical and biophysical validation techniques.

1.2 Objectives:

The key objectives of this study were to:

- generate inhibitors of PfHsp90 using Pathfinder reaction-based enumeration from known inhibitors taken from literature,
- identify compounds that bind to PfHsp90 through active learning-based Glide docking and Auto QSAR techniques,
- assess the generated inhibitors ADME (absorption, distribution, metabolism, and excretion) properties and safety profile *in silico*,
- evaluate the conformational stability of the interactions between PfHsp90 and the generated compounds by molecular dynamics simulations,
- conduct heterologous overexpression and purification of PfHsp90-NTD,
- evaluate the inhibitory capabilities of generated compounds using ATPase assays,
- assess the biophysical interactions between selected inhibitors and PfHsp90 using surface plasmon resonance (SPR).

CHAPTER 2

2. LITERATURE REVIEW

2.1 Malaria

Recent estimates indicate that approximately 247 million people were affected and 619 000 succumbed to the deadly vector-borne parasitic disease of malaria globally in the year 2021 (WHO, 2022). Children under the age of five and expectant mothers are the most severely affected population members in areas with high transmission rates, such as sub-Saharan Africa (Steketee *et al.*, 2001; van Eijk *et al.*, 2015; Gontie *et al.*, 2020). In the malaria-endemic regions of Africa, where an estimated 30 million women become pregnant (WHO, 2003), malaria infection increases the risk of maternal anaemia and delivers a low birth weight (LBW) baby (<2,500g), which in turn, is a significant risk factor for infant mortality (Guyatt and Snoe, 2004).

The main etiological agents of malaria are *Plasmodium* species that are transmitted to the human host via a bite from an infected female *Anopheles* mosquito (Sandosham, 1984; Beck-Johnson *et al.*, 2013). Human malaria is mainly caused by six *Plasmodium* species, including *P. falciparum*, *P. vivax*, *P. malariae*, *P. knowlesi*, *P. ovale* (*spp crustisi* and *wallikeri*) (Schofield and Grau, 2005; Ashley *et al.*, 2018), with the greatest threat posed by *P. falciparum* and *P. vivax*. Globally, approximately 95% of all infections are caused by *P. falciparum* and *P. vivax* (Battle *et al.*, 2019; Weiss *et al.*, 2019). *P. falciparum* and *P. vivax* are quite distinct organisms based on their morphological characteristics, ecology, and biology. *P. falciparum* has a high frequency in all tropical areas of the world (Manske *et al.*, 2012), with the highest prevalence in sub-Saharan Africa, whereas *P. vivax* is mostly found in Asia, the Middle East, and South America, but rarely in select regions of Africa (Rougeron *et al.*, 2021). Additionally, these two parasites exhibit biological differences, including features of various life cycle stages, morphology, mean parasitemia, and duration, among which the two are particularly noticeable. *P. vivax* preferentially invades young erythrocytes and can present a quiescent form in the liver that can cause relapses (dormant hypnozoites), which is not the case for *P. falciparum* (Bourgard *et al.*, 2018). Virulent *P. falciparum* can lead to life-threatening complications such as cerebral malaria, severe anaemia, and organ failure (Mohanty *et al.*, 2003; Kochar *et al.*, 2006; Doumbo *et al.*, 2010). The symptoms of uncomplicated *P. falciparum* malaria include high fever, chills, headache, muscle pain, fatigue, nausea, vomiting, and diarrhoea (Centers for Disease Control and Prevention (CDC), 2022). In severe cases, patients may experience seizures, coma, and respiratory distress (Idro *et al.*, 2010). In contrast, *P. vivax* malaria is usually less severe and dormant for a long period (Dayanand *et*

al., 2018). However, *P. vivax* malaria can also cause relapse of the disease, which is uncommon in *P. falciparum* (White, 2011).

Malaria risk has always been geographically specific, predominantly occurring in tropical and subtropical regions (Gallup and Sachs, 2001), because temperature and aridity are favourable for the spread of *Anopheles* vectors (Guerra *et al.*, 2008). Factors such as environmental stability, climate, altitude, vegetation, and poor implementation of control measures, play a major role in disease incidence (Autino *et al.*, 2012). Several studies employing global climate change projections and population-based studies have identified climate change as one of the major contributing factors to the distribution of malaria in high-risk areas such as sub-Saharan Africa (IPCC, 2014; Watts *et al.*, 2020; Sweileh, 2020). Environmental factors such as the rise in temperature, precipitation, and extreme weather events, including floods, heat waves, and storms, are influenced by climate change, thereby having a direct impact on the epidemiology of malaria, as they are favourable for the survival and transmission of the mosquito vectors of malaria (Kulkarni *et al.*, 2022). Therefore, urgent efforts are needed to identify effective, novel, and affordable antimalarials with alternative modes of action to control the spread of mosquito vectors.

2.2 *Plasmodium falciparum* life cycle

The life cycle of *P. falciparum* involves two hosts, as infection occurs between the human host and mosquito vector (Summarized in Figure 1). Human host infection begins when an infected female *Anopheles* mosquito consumes a blood meal by biting the host. In this way, sporozoites already present on the salivary glands of the mosquito are inoculated into the bloodstream of humans (Figure 1). Approximately 45 minutes after the bite, sporozoites travel through the blood circulation and enter liver parenchymal cells (hepatocytes) (Figure 1). Each sporozoite starts an asexual reproductive phase inside the hepatocytes (pre-erythrocytic liver stage), which takes approximately 1-2 weeks resulting in the formation of schizonts that contain thousands of merozoites (Figure 1). Merozoites are released into the bloodstream when mature schizonts burst. For *P. falciparum*, the hepatic phase of parasite development (hepatic schizogony) lasts for approximately five days. Merozoites quickly infiltrate red blood cells after entering the bloodstream to begin asexual reproduction (erythrocytic schizogony) (Figure 1). Once red blood cells are invaded, the *P. falciparum* parasite enters the ring stage for a period of 20-24 hours, where merozoites develop inside the red blood cells into trophozoites and schizonts (Sherman, 2005; Bannister and Mitchell, 2003), which then burst to release the next generation of merozoites, followed by infiltration of additional red blood cells and completion of the erythrocytic cycle

(Bannister and Mitchell, 2003; Kaussis *et al.*, 2009). When a schizont ruptures, parasites and erythrocytic material are released into the bloodstream, initiating the pathophysiology of malaria and the onset of symptoms.

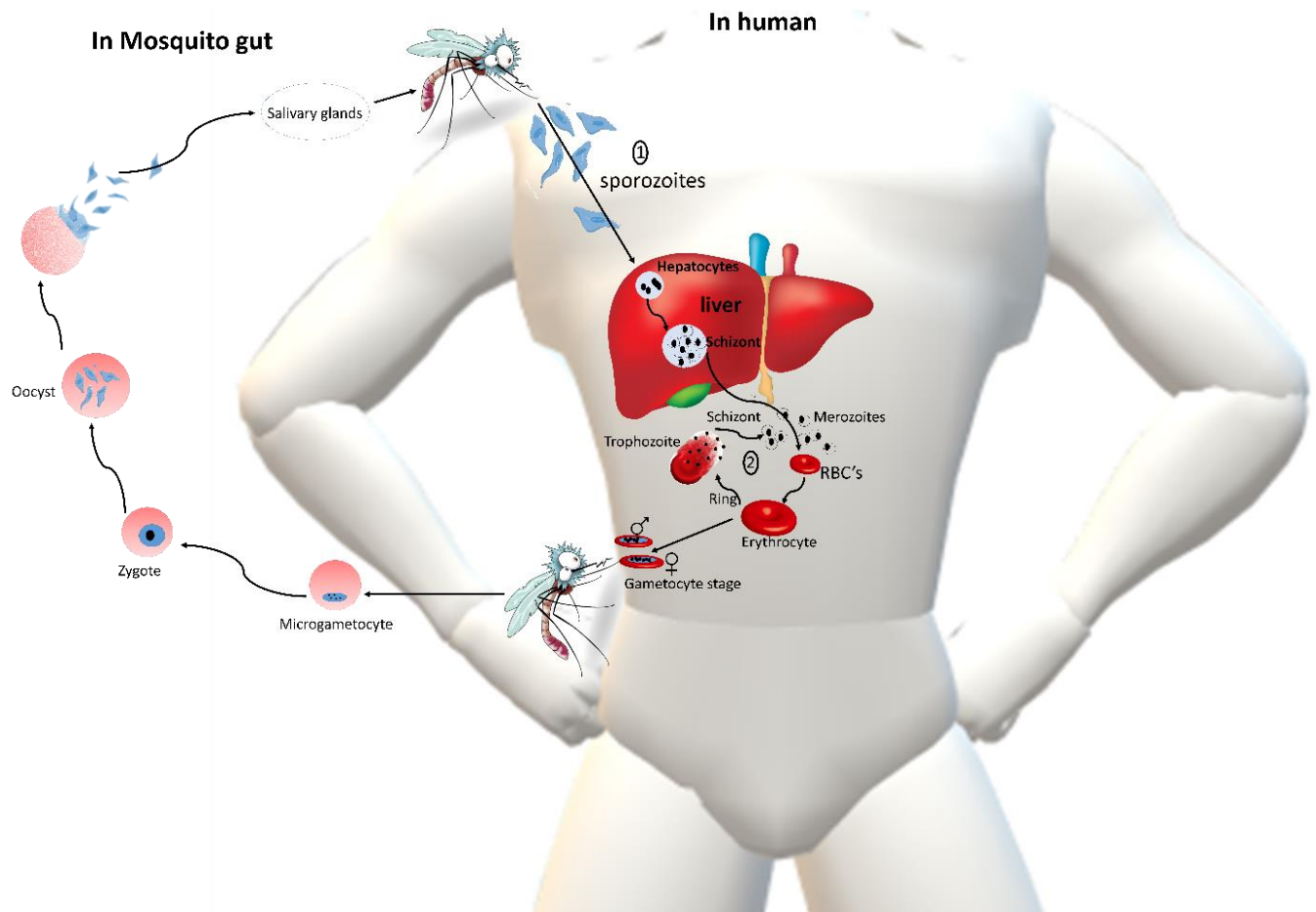


Figure 1: Life cycle of *P. falciparum*. 1, migration of sporozoites through the blood stream after a bite from the infected female *Anopheles* mosquito to invade hepatocytes. 2, the sporozoites differentiate and mitotically divide into thousands of merozoites which are subsequently released and enter the intra-erythrocytic cycle (asexual blood stage), while a sub-population of the merozoites switch to sexual development gametocytes. 3, the gametocytes are ingested by female *Anopheles* mosquito and undergo development, which results in sporozoites in the salivary glands of the mosquito. Figure adapted from Tankeshwar., 2021.

Male and female gametes are produced by a subpopulation of the merozoites that would have switched to sexual development (Josling and Llinas, 2015; Bousem and Drakely, 2016). Five morphologically recognizable stages take place for the formation and maturation of gametocytes. These distinctive transitional stages transmit malaria to the mosquito via a blood meal. Once ingested by mosquitoes, each individual gametocyte forms 1 female macrogamete or up to 8 male microgametes. In the mosquito midgut, the fusion of gametes results in the formation of a zygote that develops into a motile ookinete that can penetrate the midgut wall to form oocysts. The

oocysts enlarge over time and burst to release sporozoites that migrate to the mosquito salivary gland, rendering the mosquito infectious to human beings (Figure 1) (Bousem and Drakely, 2016).

2.3 Uncomplicated and severe form of malaria

The onset of symptoms greatly depends on factors such as the degree of immunity, type of infecting parasite species and density of parasite inocula, all having a great impact on the duration of the incubation period (Bartoloni & Zammarchi, 2012). For *P. falciparum*, the incubation period lasts from 9 to 30 days and uncomplicated malaria initial symptoms are non-specific and common to all fever infections. These include flu-like manifestations, often rendering clinical diagnosis unreliable (Bartoloni & Zammarchi, 2012). It is precisely for this reason that sensitive molecular diagnostic tools such as Polymerase chain reaction (PCR), should be employed as crucial detection tools of malaria (Erdman and Kain, 2008). The most common sign of malaria infection is high fever (CDC, 2022). Prodromal symptoms, such as chills, body-aches, headache, cough, diarrhoea, and vomiting, may appear two days prior to the onset of fever (Warrell, 2017). The fever generally manifests irregularly and fluctuates over time accompanied by shaking and light chills (Mayo clinic, 2023). The fever in *P. falciparum* malaria may occur every 48 hours, but is usually irregular, showing no distinct periodicity (Crutcher and Hoffman, 1996).

In the case of severe malaria, symptoms affect the central nervous system (cerebral malaria), pulmonary system (respiratory failure), renal system (acute renal failure), resulting in severe anaemia (Bunnang, 1988). Any malaria patient must be evaluated and get the correct treatment as the development of these complications can be swift, and because the illness can be deadly at this phase (Day and Dondorp, 2007). Frequent observations are necessary to assess for the first indications of systemic problems. On the other hand, cerebral malaria presents as acute lung injury, which can develop into breathing difficulties (25% of patients), acidosis, typically presenting as acute tubular necrosis, and acute kidney injury (Taylor *et al.*, 2012).

The advancement of the most prevalent fatal consequences of malaria (cerebral malaria, respiratory failure and acute renal failure) and development of severe malaria from uncomplicated malaria are greatly dependent on immunological and epidemiological factors. Factors such as lack of prompt response to the early symptoms in children, lack of knowledge of the early malaria symptoms, use of wrong malaria medication and dosage, closeness and accessibility of health centres (Anumudu *et al.*, 2007). In the absence of immune compromise and parasite resistance to available drugs, the frequency of malaria infection can be reduced, and the development of severe malaria and its consequences monitored, if caregivers are aware of the symptoms and treat them promptly with appropriate therapies.

2.4 Past and present antimalarial treatments

Even though malaria infection numbers have reduced from 245 million in 2000 to 230 million in 2015 across the 108 countries that were malaria endemic in 2000, recent reports have shown that malaria cases have surged since 2016, with the largest yearly increase of 13 million cases reported between 2019 and 2020, during the first year of the COVID-19 pandemic (WHO, 2022). Emphasizing the importance of avoiding complacency with current treatment and preventive techniques (Cheney, 2018; Mwangi, 2018; Tse *et al.*, 2019). Antimalarial drugs may also result in a prolonged incubation period, due to their effects on the parasite's multiplication rate, although, ineffective for treating malaria (Bartoloni & Zammarchi, 2012).

The treatment of malaria depends on several factors, including the severity of the infection, the parasite strain responsible for the infection, and whether the patient presents with symptoms of severe or uncomplicated malaria (Bartoloni and Zammarchi, 2012). WHO has recommended artemisinin-based combination therapies (ACTs) as first-line treatment option for *P. falciparum* malaria, as well as supportive care measures such as hydration and management of fever and other symptoms (WHO, 2022). Severe or complicated malaria treatment may need hospitalization and reinforcement through administration of intravenous fluids (such as normal saline, dextrose saline, dextrose, Hartmann's solution, Ringer's lactate, but not including human albumin or plasma substitutes such as gelatin) and oxygen therapy (Hodgson and Angus, 2016). The risk of parasitic drug resistance is a growing concern in the treatment of malaria. WHO has identified drug resistance as one of the key challenges to malaria control and elimination efforts (WHO, 2020). Gene mutations in the *P. falciparum* parasite leads to the development of drug resistance if antimalarial medications are taken improperly or often, which might result in treatment failure (Shibeshi *et al.*, 2020). It is challenging to control and eradicate malaria in afflicted areas because of this resistance's fast expansion (White, 2004; Okeke *et al.*, 2005; Buabeng *et al.*, 2007; Shibeshi *et al.*, 2020)

Quinine was the first chemically pure alkaloid successful therapy for malaria in the 1820's that comes from the bark of the cinchona tree (*Cinchona pubescens*), introduced to Europe in the 17th century from Peru (Greenwood, 1992), and since then additional organic and synthetic chemicals have been developed [Figure 2], (Tse *et al.*, 2019). Quinine works by inhibiting haemoglobin breakdown in *Plasmodium* parasites, which is a comparable mode of action to chloroquine (Blasco *et al.*, 2017). Regrettably, mutations hindering proper drug transport have rendered this treatment less effective over time (Capela *et al.*, 2019), as parasite developed resistance to these medications, decreasing their efficacy. As a result, they are no longer in use.

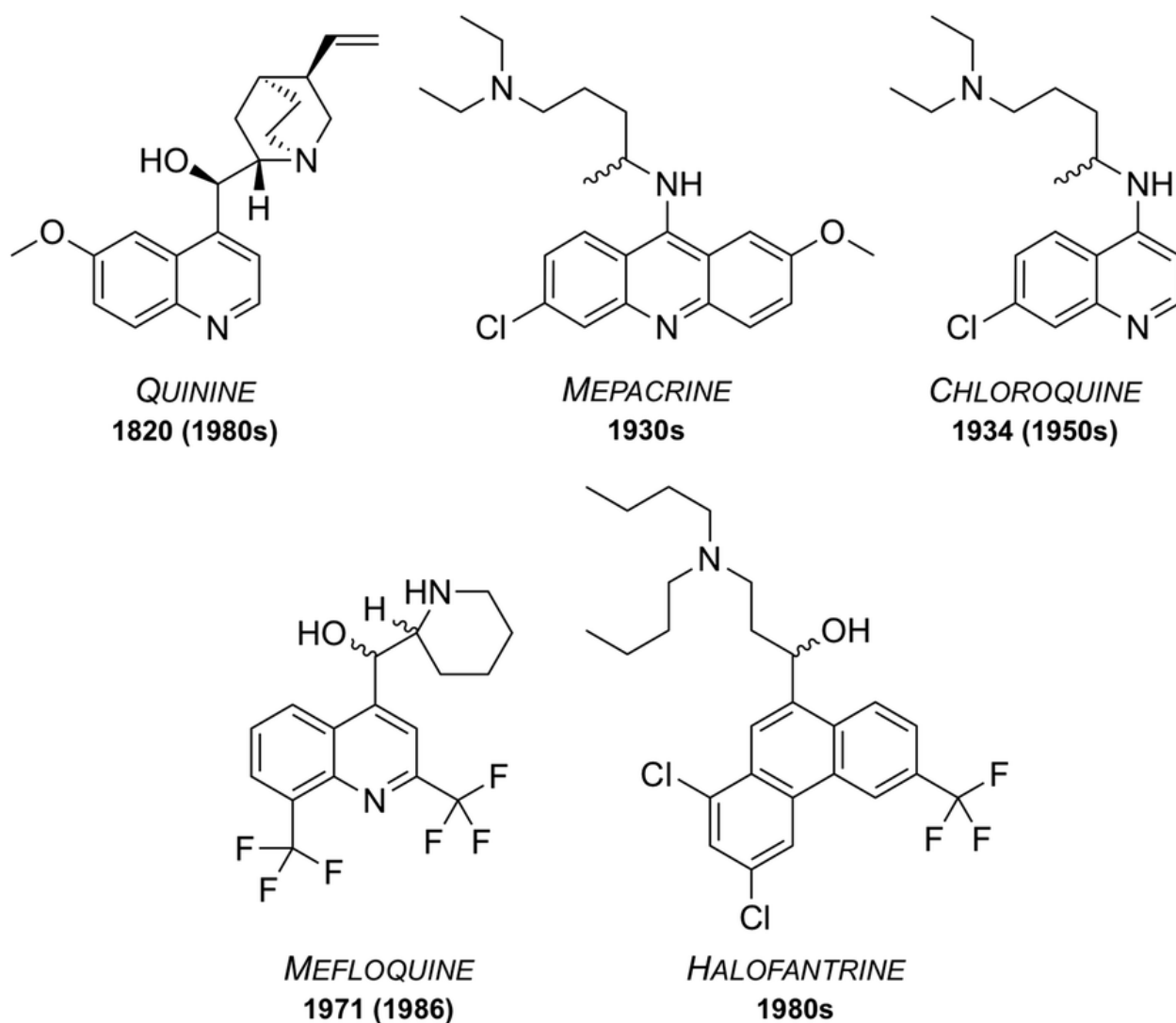


Figure 2: Chemical structures of the past antimalarials that were used.

Since its discovery in the 1820's, quinine was one of the most effective antimalarial treatments (Reviewed by Achan *et al.*, 2011). The first case of resistance was reported in the 1980's (Bunnag *et al.*, 1996), and as of 2006, quinine was no longer used as the first line therapy for malaria.

During the 1930s, the antimalarial chloroquine was developed and became widely used as an antimalarial during the 1960's, having fewer side effects (Loeb *et al.*, 1946). Both quinine and chloroquine are quinoline containing drugs, targeting the heme degradation pathway of the parasite. The compounds use a weak base mechanism to kill the malaria parasite by inhibiting haemoglobin degradation from occurring in the acidic food vacuoles of the intraerythrocytic trophozoites (Slater, 1993; Fitch 2004; Bray *et al.*, 2005). The first cases of chloroquine resistance arose in South-East Asia and South America within a decade of its introduction.

Ever since then, it has expanded slowly but unabatedly throughout endemic regions, eventually reaching Africa in the late 1970s (Bray *et al.*, 2005; Mita *et al.*, 2009; Packard, 2014). Past studies on *P. falciparum*, the most dangerous form of human-infecting malaria parasite, have revealed that point mutations in the chloroquine resistance transporter (PfCRT) protein are crucial for the development of chloroquine resistance (White, 2004; Bray *et al.*, 2005). Usage of chloroquine was discontinued from the treatment of *P. falciparum* infection in almost all endemic regions due to the global spread of resistance (Balikagala *et al.*, 2020).

Artesunate is an antimalarial drug that belongs to the artemisinin class of compounds. It is one of the active ingredients in artemisinin-based combination therapies (ACTs), which are the recommended first-line treatment for uncomplicated malaria caused by *P. falciparum*. The mechanism of action of artesunate include the formation of reactive oxygen species (ROS) in the malaria parasite (Gopalakrishnan and Kumar, 2015). Artesunate is triggered by heme, which is released from the parasite's hemoglobin during red blood cell digestion (Tilley *et al.*, 2016). The activated artesunate then combines with the heme to form reactive oxygen species (ROS), which damage the parasite's cell membrane, proteins, and DNA, ultimately leading to death (Gopalakrishnan and Kumar, 2015). Artemisinins are currently the established antimalarial agents with a good safety profile. However, artesunate needs to be complimented by other drugs to be effective, as the key objective when treating malaria is not only eliminating parasites but also ensuring complete clearance within a reasonable period without any chances of recurrence. Artesunate is fast acting and possess a short half-life. Therefore, alone, artesunate cannot provide long-lasting immunity against malarial parasites (Borrmann *et al.*, 2003), additional medications are often necessary for enhanced efficacy duration and prevention of reoccurrence post treatment cycles completion. Clinical trials with artemisinin monotherapies have shown that these compounds are taking considerably longer to clear malaria infections in Southeast Asia typically twice as long as observed a decade ago.

Drugs like mefloquine or lumefantrine have been identified as viable options to facilitate prolonged action while inhibiting the emergence of malarial drug resistance due to their long half-life (Ruwezhi *et al.*, 2022). Artesunate targets intra-erythrocytic asexual blood-stage malaria parasites, and generally partner drugs in the ACT complemented by targeting multiple stages of the parasite's life cycle, increasing the likelihood of complete parasite clearance, and lowering the risk of recurrence (Tilley *et al.*, 2019). This is significant because the *Plasmodium* parasite has a complicated life cycle that includes several stages, including the liver, blood, and transmission phases (Nosten and White, 2007). WHO advised the use of ACTs as the first-line therapy for *P. falciparum* malaria in all endemic countries in 2005 (Nosten and White, 2007; WHO, 2015). Malaria-related morbidity and death have decreased because of the use of artemisinin-based

combination therapy in place of the inefficient, failed treatments (Ashley *et al.*, 2014). There are about 67 countries that are affected by *P. falciparum* malaria, 41 of which are in Africa, and have adopted the ACTs as a treatment option (Bosman and Mendis, 2007).

Extensively used ACTs include artesunate-amodiaquine, artemether-lumefantrine, dihydroartemisinin-piperaquine, artesunate-mefloquine and arteunate plus sulfadoxine-pyrimethamine (Ouji *et al.*, 2018). The WHO (2015) also provided evidence-based guidelines for the treatment of severe malaria. Individuals with severe malaria must be administered artesunate, either intravenously or intramuscularly until they are able to take the medication orally (Esu *et al.*, 2014). There have been issues with the solubility of artesunate, hence it is mainly taken as a tablet (Blasco *et al.*, 2017). Then the patient can be given ACTs treatment, which must be completed within three days. However, clinical trials with artemisinin monotherapies have shown that these compounds are taking considerably longer to clear malaria infections in Southeast Asia typically twice as long as observed a decade ago (Corey *et al.*, 2016). Improper usage or incompleteness of the ACT is likely to result in the resistance to ACTs by the *P. falciparum* parasite (Onyango *et al.*, 2012). The fundamental reason that ACTs may contribute to the establishment of resistance is that artemisinin derivatives have a short half-life, which means they are quickly cleared from the body. Some parasites may be able to survive the initial artemisinin treatment and develop resistance to the drug because of this. If these resistant parasites are allowed to proliferate, they can spread and become dominant, resulting in treatment failure and the establishment of artemisinin-resistant malaria parasite strains (Nsanjabana, 2019). Although ACTs have been efficient in tackling malaria, there is already emerging evidence of increasing treatment failures in parts of south Asia (Ashley *et al.*, 2018) and Rwanda (Rosenthal, 2021). In addition, concerns regarding the cost and sustainability of ACTs in Africa have been raised (Oladipo *et al.*, 2022). In this perspective, it is critical to rapidly identify, develop and validate targets and molecules, which exhibit different structural and mechanistic modalities to the current front-line therapies. By targeting molecular chaperones as they are critical for the survival of the parasite, it may be possible to develop therapies that are less prone to resistance than existing front-line therapies, such as ACTs. Therefore, the identification and validation of molecular chaperones as a drug target represents an important step in the development of new, effective antimalarial drugs.

2.5 Molecular chaperones as drug targets

Molecular chaperones are proteins capable of interacting with and stabilizing non-native protein structures or newly formed polypeptides, preventing aggregation while promoting the formation of correct and functional conformations (Hendrik and Hartl, 1993; Ellis, 2006). These wide range

of proteins act as catalysts in the refolding of stress denatured or newly formed polypeptides, controlling the assembly of oligomeric proteins, transporting proteins and facilitating protein degradation, and protein–protein interactions (Fink, 1999; Edkins and Boshoff, 2014).

Heat shock proteins (Hsp) are members of the molecular chaperone family, functioning to overcome the effects brought about by cellular stress including elevated temperatures or a crowded environment (Edkins and Blatch, 2012). Eukaryotic Hsps are both constitutively expressed under normal conditions and heat inducible (Donnelly *et al.*, 2008). Hsp90 alpha in human is the house-keeping gene, therefore, constitutively expressed under normal cellular homeostasis. However, Hsp90 beta is the stress protein, expressed following exposure to stress ie. It is nicknamed the cancer Hsp90. During cellular stress, the expression of these Hsps is upregulated, preventing the aggregation of the stress denatured proteins, assisting them to re-fold, therefore, maintaining a functional proteome (Derry, 2018). All eukaryotic organisms possess the same ability to respond to stress, and this capability entails the use of molecular chaperones to help the organism cope with the impacts of stress. Adapting to the differences in host and vector conditions is particularly very important for parasitic organisms as it can be very stressful for the organism (Kumar *et al.*, 2007). Hsps are classified and named according to their function and molecular weight on kilo Daltons (kDa), for instance Hsp90 has a molecular size of approximately 90 kDa (Zininga and Shonhai, 2019). The diverse family of Hsps include small heat shock proteins (Hsp), Hsp40, Hsp70, Hsp90 and Hsp110 (Lindquist and Craig, 1988; Kim *et al.*, 2013). Hsps are highly desirable therapeutic targets because of their roles in human disease and cellular stress responses, particularly Hsp90 and Hsp70.

Since chaperones are typically overexpressed in cancer cells and have also been linked to malignant cell maintenance (Calderwood *et al.*, 2006), Hsp90, for instance, is well established as a therapeutic target in cancer and neurological illnesses. It is currently the subject of research in the context of parasites. The crucial role played by these proteins in the survival and development of these parasites in different physiological habitats (human host and mosquito vector), is what makes them such ideal drug targets. In *P. falciparum*, Hsps are responsible for pathogenesis, regulation of infection, thermal protection of the parasite and additionally, the exportation of parasite proteins to the red blood cells (Shonhai, 2010). The major types of Hsps found in *P. falciparum* include, PfHsp40, PfHsp70 and PfHsp90. With PfHsp90 and PfHsp70 being the most prominently expressed, regulating transcription factors and protein kinases while also helping to fold newly synthesized proteins (Gitau *et al.*, 2012; Kim *et al.*, 2013).

2.6 Hsp90

The most widely occurring molecular chaperones in eukaryotic cells is the Hsp90 (Picard, 2002). The ATP dependent molecular chaperone of Hsp90 plays critical roles in fundamental processes of cell signalling and, cell cycle control. Hsp90s are the central hub of proteostasis (Bieble *et al.* 2019), which are conserved across all organisms, displaying >50% sequence similarity between *E. coli* and human Hsp90s (Chen *et al.*, 2006).

Hsp90 isoforms are located in the cytosol (Hsp90 α and Hsp90 β), the mitochondria (Hsp75/ TRAP-1), and the endoplasmic reticulum (GRP94/GRP96). There are specialised forms of Hsp90s found in different cellular organelle with specialised sets of client proteins. In humans, Hsp90 has also been found to localize in plasma membrane or extracellular matrix (Edkins and Blatch, 2012). *P. falciparum* harbours an apicoplast localizing the isoform of Hsp90 (Acharya *et al.*, 2007). The cytosolic Hsp90 isoform is responsible for properly folding over 300 protein substrates, termed clients, including protein kinases, transcription factors, and receptors critical for maintaining protein homeostasis and regulating vital cellular processes (Acharya *et al.*, 2007). All members of the Hsp90 family are structurally configured as homodimers comprising of three main domains (Figure 3), the N-terminal domain (NTD) which is the primary binding site for ATP/ADP (Figure 3 in red), responsible for the provision of requisite energy for client folding (Banerji, 2009). A charged linker region that connects the NTD with the middle domain (MD) (Figure 3 in green), which varies in amino acid sequence across different organisms (Jackson, 2013). The MD functions as a discriminator of different client proteins (Hawle *et al.*, 2006). Lastly, the C-terminal domain (CTD), which is adjacent to the MD, is the site of dimerization (Figure 3 in blue). Cytosolic Hsp90s possess an end C-terminal EEVD motif which acts as a binding site for tetratricopeptide repeat (TPR) containing co-chaperones (Prodromou *et al.*, 2000).



Figure 3: Diagrammatic sketch of the full length *P. falciparum* Hsp90, with the N-terminal (NTD) ATP binding site (Amino acid 1-275) shown in Red rectangle, the Middle domain (MD) from amino acid 275 – 444 depicted in green circle and then the C-terminal domain (CTD) from amino acid 444-677 in blue circle. The 4 vertical lines on the NTD represents the ATP binding residues. Adapted from Kabakov, 2021.

2.7 Mechanism of Hsp90 ATPase activity

The ATP-dependent Hsp90 possess an NTD that serves as the ATP binding site and an ATPase cycle as depicted in Figure 4. ATP binding and hydrolysis is critical for the functioning of Hsp90 (Obermann *et al.*, 1998; Panaretou *et al.*, 1998). ATP binding also influences conformation of Hsp90, promoting the rearrangement of the client binding region (Csermely *et al.*, 1993; Grenert *et al.*, 1997). A study by Weikl *et al.* (2000) provided evidence that conformational changes occur in the Hsp90 molecule upon binding of ATP. Upon client protein binding, Hsp90 would be trapped in the ATP bound state. The ATP molecule will subsequently undergo hydrolysis (Figure 4, 1st hydrolysis), releasing ADP and P_i. ADP induces a conformational change of the Hsp90 molecule from a widely open V-shaped molecule to the NTD dimerized closed state, while the conversion to the compact ADP state will displace the bound client, followed by reopening to reset the chaperone (Figure 4) (Elnatan *et al.*, 2017).

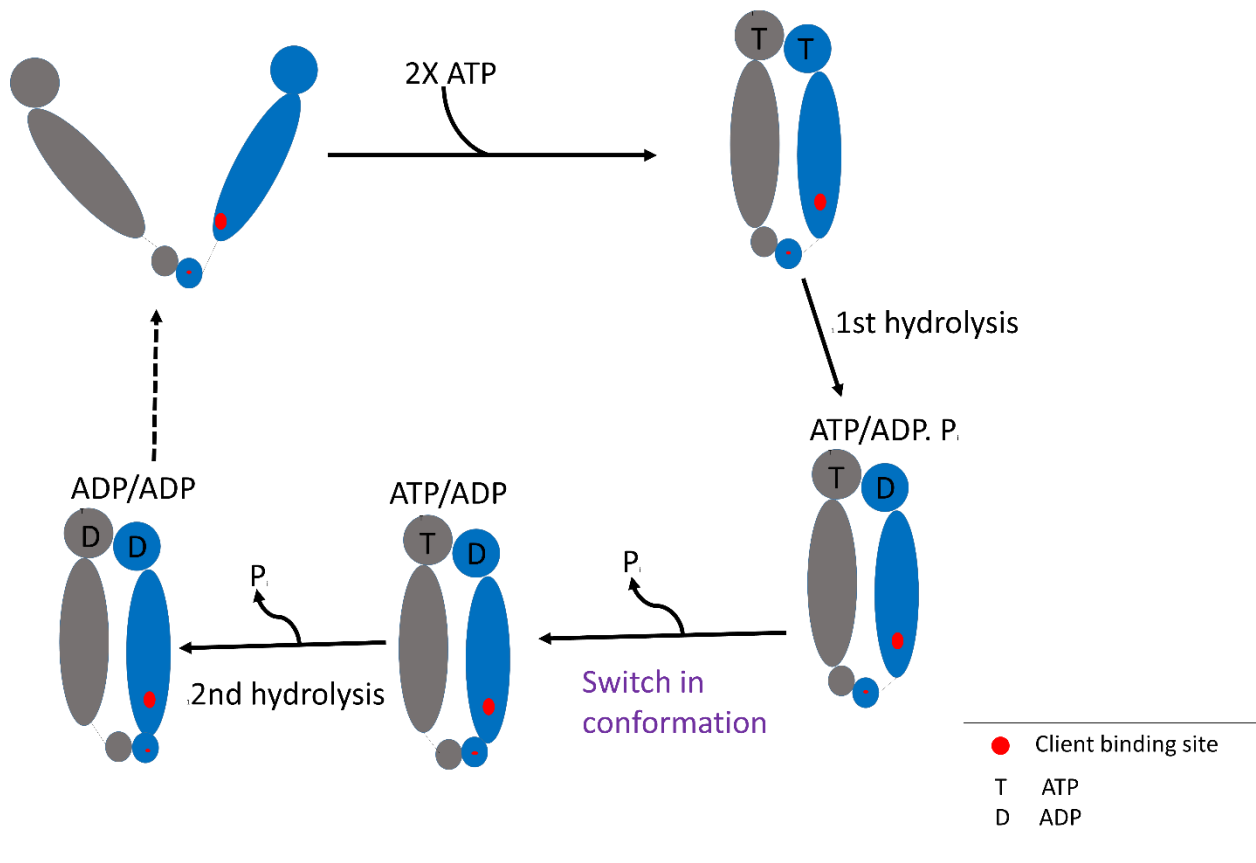


Figure 4: Model of the Hsp90 ATPase cycle showing the obligatory sequential hydrolysis and conformational switching. Coloured in grey are the protomers. 2 ATP molecules bind to an Hsp90 molecule apo state (top left), which induces a dimerized closed state of the NTD. During the closed state, the buckled protomer hydrolyses the 1st ATP molecule releasing P_i, which then drives a conformational switch of the straight protomer bound to ATP, to a buckled conformation, while the previously buckled protomer (now ADP), straightens. This rearrangement facilitates the remodelling of the client. As a result, the now buckled promoter hydrolyses the second ATP. Finally, the ADP/ADP dimer re-opens, releasing nucleotides and resetting TRAP1 to the apo state.

In the past decade, Hsp90 has become a major therapeutic target for cancer, there has also been increasing interest in it as a therapeutic target in neurodegenerative disorders, and in the development of anti-virals and anti-protozoan infections (Jackson, 2012). There has been great success with regards to targeting Hsp90 in cancer research, where it has already been studied extensively (Kim *et al.*, 2009; Edkins, 2016; Park *et al.*, 2020). Currently there are 18 compounds/inhibitors of Hsp90 at various phases of cancer clinical trials. AUY922 (also known as Luminespib) is a clinically approved Hsp90 inhibitor that has been demonstrated to have anticancer efficacy in preclinical and clinical investigations (Jensen *et al.*, 2008). AUY922 is a small chemical inhibitor that binds to Hsp90's ATP-binding site, reducing its activity and encouraging the degradation of client proteins that rely on Hsp90 for stability (Jensen *et al.*, 2008). AUY922 has showed encouraging outcomes in clinical studies for treating several forms of cancer, including gastric cancer (Park *et al.*, 2018). Another example is jesselhy® (pimipib) an orally administered small molecule inhibitor that specifically targets the α and β isoforms of Hsp90 to treat gastrointestinal stromal tumours (GIST), which gained its first approval in June 2022 in Japan (Hoy, 2022). Inhibition of Hsp90 leads to antitumor effect by destabilization and reduction of proteins involved in cancer growth and survival, including receptor tyrosine kinase (KIT), platelet derived growth factor receptor alpha (PDGFRA), human epidermal growth factor receptor 2 (HER2), and epidermal growth factor receptors (EGFR) (Thaiho Pharma, 2022).

The laudable efforts of Hsp90 targeting in eukaryotic organisms has inspired repurposing of the compounds for parasitic disease. However, several lessons have been learned from currently available Hsp90 inhibitors, especially from cancer research. For instance, geldanamycin (GA), although highly effective and with selective affinity towards cancerous cells with minimal effect on normal cells, has been reported to have hepatotoxic side effects (Neckers *et al.*, 1999). This has limited its pharmaceutical application. Hence the molecular structure of geldanamycin has been used as a scaffold to develop new smaller derivatives, which are able to retain the potency towards Hsp90, without (or with fewer) adverse effects. Some of the derivatives currently in anticancer clinical trials include, tenespimycin (17-AAG) in phase 3 (Selleckchem, 2023; [https://www.selleckchem.com/products/17-AAG\(Geldanamycin\).html](https://www.selleckchem.com/products/17-AAG(Geldanamycin).html)) and Alvepipimycin (17-DMAG) HCL in phase 2 ([https://www.selleckchem.com/products/17-AAG\(Geldanamycin\).html](https://www.selleckchem.com/products/17-AAG(Geldanamycin).html)).

2.8 Small molecule inhibitors of Hsp90

Previous studies have shown that inhibiting the ATPase activity at the N-terminal domain of the Hsp90 is an effective strategy for disrupting the functions and interactions with other client proteins (Usmani and Chiosis, 2012). A study by Theoderaki and Caplan (2011), demonstrated that the inhibition of Hsp90 with N-terminal inhibitors leads to the degradation of client proteins.

To date, a number of synthetic and natural Hsp90 targeting inhibitors have been evaluated (Neckers 2003; Drysdale *et al.*, 2006; Taldone *et al.*, 2008), including the most common GA and its derivatives such as 17-allylamino-17-demethoxygeldanamycin (17-AAG), 17-dimethylamino-ethylamino-17-demethoxygeldanamycin (17-DMAG) (Pallavi *et al.*, 2010), which are already in phase 3 clinical trials despite the fact that it can only be administered intravenously because of solubility concerns (Koca *et al.*, 2016).

With regards to targeting PfHsp90, these inhibitors have been found to be cytotoxic to human cells as they are unable to be selective of PfHsp90 over the human Hsp90 (Shonhai, 2010), rendering them ineffective as antimalarials. This calls for other alternative small molecule inhibitors with high specificity for PfHsp90 over the human Hsp90. The objective of the current study was to generate novel small inhibitors that selectively target PfHsp90 to serve as antimalarial agents.

2.9 *P. falciparum* Hsp90 (PfHsp90) as a drug target

PfHsp90 is an attractive drug target for the development of novel antimalarial drugs as the *P. falciparum* parasite's ability to survive and proliferate is greatly dependent on the molecular chaperone. Molecular chaperones are expressed by the parasite to effectively cope with severe environmental stresses, including changes in pH, temperature, and nutrient supply, during its cyclic transition from the mosquito vector to the human host (Pavithra *et al.*, 2004; Uwimana *et al.*, 2020). The cellular stress induces the abundant expression of PfHsp90 in the parasite (Shahinas *et al.*, 2013). Functional interruption of PfHsp90 using the NTD inhibitors from anti-cancer agents have previously been explored (Kumar *et al.*, 2003; Banumathy *et al.*, 2003; Pallavi *et al.*, 2010; Shahinas *et al.*, 2013; Murillo-Solani *et al.*, 2017; Posfai *et al.*, 2018). These studies demonstrated the essentiality of PfHsp90 in the survival and proliferation of the *P. falciparum* and have provided evidence to investigate PfHsp90 as a drug target. A study by Banumathy *et al.* (2003) showed that geldanamycin can inhibit PfHsp90 in blood stage parasites, thereby preventing growth and transition from the ring to the trophozoite stage of the parasite. PfHsp90 was validated as a drug target with pre-clinical studies wherein 17-AAG, PUH-71 and harmine were able to cure *P. berghei* infected mice in a pre-clinical trial (Pallavi *et al.*, 2010). Another study that aimed to show the effect of geldanamycin on *P. falciparum* replication and morphology by Kumar *et al.*, (2003) gave evidence that treatment with geldanamycin resulted in a more robust inhibition to PfHsp90 than its human counterpart, the human Hsp90, where the parasite's growth was strongly inhibited by GA with an IC₅₀ of 200 nM in human erythrocyte culture.

Harmine is capable of selectively binding PfHsp90 over human Hsp90 (Shihanas *et al.*, 2012), providing the evidence that selective inhibitors of PfHsp90 can be developed. The challenge with regards to designing selective inhibitors of PfHsp90 over human Hsp90 is the highly conservative nature of the molecular chaperones (sequence homology of about 69%) (Corbett and Berger, 2010). Biochemically, PfHsp90 has been demonstrated to possess a higher ATPase activity than its human homolog, making it more susceptible to inhibition (Pallavi *et al.*, 2010; Shahinas *et al.*, 2013). Overall, the observed difference suggest that the two chaperones have distinct enzymatic properties and these, can be exploited biochemically for inhibitor design. As PfHsp90 is a validated drug target in malaria, the process of identifying and developing its inhibitors involve hit discovery, hit-to-lead optimization, and preclinical testing.

2.10 Drug discovery

Drug discovery is the process involving a confluence of scientific disciplines, including chemistry, biochemistry, biology, pharmacology, toxicology, and computational science to identify and develop new therapeutic entities that can effectively treat or cure diseases by targeting specific molecular pathways or biological processes (Zhou and Zhong, 2017).

In target-based drug discovery, the drug discovery pipeline typically starts off by the identification and validation of a potential therapeutic target (protein or enzyme) that plays key role in the development or mechanism of action of the disease (Hughes *et al.*, 2011). Followed by identification of a small potent compounds can selectively bind with high affinity. Hit identification employ various strategies to screen enormous libraries of chemicals for molecules that interacts with the target (Schenone *et al.*, 2013). These compounds are then optimized through a process of iterative testing and modification to improve their potency, selectivity resulting in a lead candidate. Once a lead molecule has been identified, it must go through a series of *in vitro* and *in vivo* preclinical testing in animal models to determine its safety and effectiveness. If the results are good, the molecule will undergo safety and efficacy testing in increasingly larger groups of patients in human clinical trials (Hughes *et al.*, 2011).

There are other different approaches for drug discovery including traditional drug design, which is based on screening thousands of natural and synthetic compounds for activity (Reddy and Parrill, 1999). This approach includes the oldest known technique, the serendipity method (Ban, 2022). Examples in the history of pharmacy for drugs discovered by the serendipity method include the most popular story about penicillin, a drug that saved millions of lives during the Second World War and for which Fleming, Florey and Chain received the Nobel Prize in 1945.

Once a potential drug compound was selected by this process of screening, medicinal chemists would then synthesize hundreds of related compounds to develop the safest, most effective drug for patients use (Reddy and Parrill, 1999). However, the costs and risks associated with this process are enormous. This phenotypic screening technique's shortcoming is that it does not indicate why a compound is active or inert, or how it may be improved. It also does not guarantee that an active compound is specific for a particular human target protein (Reddy and Parrill, 1999). A lack of such specificity can be a major cause of unpleasant side effects, potentially halting a drug's clinical development. Other approaches include, chemical modification of established drugs or natural compounds (Guo, 2017), screening of databases, virtually or by high throughput (HTS) assay to discovering new drugs (Smith, 2002; Shoichet, 2004; Butkiewicz *et al.*, 2017).

Nowadays, the most sophisticated method for drug discovery is rational drug design, which is both the smartest and least expensive form of drug discovery (Flower, 2002; Young, 2009; Doytchinova, 2022). The drug discovery process is a long and complex one, often taking more than a decade and costing billions of dollars. However, successful drug discovery can lead to life-changing treatments and cures for a wide range of diseases.

2.10.1 Rational drug design

Rational drug design (RDD) is a process in which drugs are developed by using knowledge of a target molecule's structure and function (Batool *et al.*, 2019). The process involves using information about the molecular target, such as its 3D structure, to design a molecule that interacts specifically with that target to produce a therapeutic effect (Kholodenko, 2000). RDD aims to create more effective and specific drugs with fewer side effects, compared to traditional phenotypic screening approaches (Mahapatra and Karuppasamy, 2022). The molecular structure of the target is used to predict how a drug molecule will interact with it, and the design process can involve computer simulations, molecular modelling and synthesizing new compounds to test their efficacy (Sharp, 2002).

The success of RDD depends on the availability of detailed information about the target molecule, when this information is available, the process can lead to the development of novel and highly effective drugs (Wang and Gainza, 2003). In the absence of a crystal structure or experimental data for the target protein, homology modelling can provide valuable insights into the protein's structure and aid in the design of potential drugs (Vyas *et al.*, 2012). RDD strategies form part of the different approaches required for the process of drug discovery, with the objective of overcoming the challenges around the traditional drug design process (Mahapatra and Karuppasamy, 2022). The process is challenging as it is complex, expensive and time-

consuming, involving numerous factors that must be taken into consideration (Reddy *et al.*, 1999). Biomolecules play crucial roles in the development of a disease by interacting with one another through protein-protein or protein nucleic acid interactions, spreading signalling events or changes in the metabolic processes, therefore, altering the bioactivities carried out by these biomolecules (Baltoumas *et al.*, 2021). Targeting these biomolecules could be advantageous. This can be done by either by binding to and inhibiting their activity using small molecules, whose competitive binding affinities would be higher than that of their natural ligands that bind on the active site on the biomolecule, or by inhibiting the biomolecular interactions with each other, preventing crosstalk between the molecules (Fuller *et al.*, 2009). RDD has significantly advanced in recent years due to the advent of computational tools and computer aided methodologies such as virtual screening of libraries and machine learning approaches (Henry, 2001; Miller 2009). The successful application of rational drug design strategies is seen in the development of drugs such as saquinavir (Invirase) and ritonavir (Norvir) in HIV treatment, that work by targeting the protease enzyme involved in the replication of HIV, inhibiting its activity, and preventing the production of new viruses (Santos *et al.*, 2015).

2.10.2 Application of molecular docking in drug discovery

Molecular docking is a computational technique that simulates the interaction between a ligand and a receptor to predict their binding mode and affinity (Meng *et al.*, 2011). Molecular docking has become an indispensable tool in early-stage drug discovery for hit identification, lead optimization, and virtual screening of large compound libraries (Morris *et al.*, 2009). Predictions of the binding modes and binding affinities by docking is attributed to docking algorithms that essentially identify favourable orientation of a ligand in the binding site of the protein. Glide being one of the most widely used estimate binding affinities (Friesner *et al.*, 2004). Glide estimates binding affinities through a combination of molecular docking and scoring methods. It utilizes a grid-based algorithm to predict the binding conformation of a ligand within the active site. The glide scoring function considers various factors such as steric clashes, hydrogen bonding, hydrophobic interactions, and electrostatic interactions to evaluate the binding affinity of the ligand (Friesner *et al.*, 2004). The starting point for glide scoring is the empirically based ChemScore function of Eldridge *et al.* (1997) which can be written as:

$$\Delta G_{bind} = C_0 + C_{lipo} \sum f(r_{lr}) + C_{hbond} \sum g(\Delta r)h(\Delta a) + C_{metal} \sum f(r_{lm}) + C_{roth}H_{roth} \quad (1)$$

Chemscore defines the summation in the second term including all ligand-atom/receptor-atom as lipophilic, whereas the summation in the third term includes all ligand-receptor hydrogen bonding interactions. Functions such as f, g, and h (equation 1) produce a complete score 1.00 for lengths

or angles that fall within conventional boundaries and a half score 1.000.00 for distances or angles that are beyond those limitations but fall within greater threshold values (Friesner *et al.*, 2004).

While docking is generally successful in reproducing experimentally determined ligand poses, accuracy can be affected by factors such as protein flexibility and ligand conformational variability (Kitchen *et al.*, 2004). Therefore, to improve docking accuracy computational softwares like the Schrödinger's induced fit docking (Halgren *et al.*, 2004; Friesner *et al.*, 2004; Friesner *et al.*, 2006) and molecular dynamics simulations (Amber *et al.*, 2005) have been designed to take into consideration the flexibility of the protein and ligand.

Schrödinger's Induced fit docking approach is taken to address the issue of receptor flexibility during ligand binding by combining docking techniques. Using glide to account for ligand flexibility and the prime refinement module accounting for receptor flexibility where minimization is used to sample side-chain degrees of freedom in the receptor while enabling modest backbone movements (Sherman *et al.*, 2006). Therefore, the methodology compensates for both modest backbone relaxations and substantial side-chain conformational changes in the receptor structure.

Docking has been successfully applied in various stages of drug discovery for instance in virtual screening, enabling the screening of large compound libraries to identify potential lead compounds. Virtual screening methods, such as high throughput docking, enable the prioritization of compounds based on their predicted binding affinities (Ferreira *et al.*, 2015). Docking plays a crucial role in lead optimization by guiding chemical modifications of compounds to improve binding affinity and selectivity. Structure-based design approaches, including fragment-based docking and *de novo* design, aid in the design of potent and specific ligands (Hajduk *et al.*, 2007).

2.10.3 Molecular dynamic simulation

Molecular dynamic simulation (MDS) is an important application in drug discovery studies particularly in ligand docking applications as it enables the study of conformations of a molecule or complex that are thermally accessible. Furthermore, MDS gives insights into the natural dynamics of biomolecular interactions in a solution at different timescales (Hansson *et al.*, 2002). MDS can be applied in the study of protein ligand complexes, mixed solvents, organic solids and synthetic macromolecular complexes to analyse molecular and condensed-phase systems that are dynamic in nature. Coupled with other computational tools MDS seeks to address challenges faced in drug discovery studies where the dynamic nature of the proteins cannot be ignored, such as ligand induced conformational changes in the receptors active site, by capturing dynamic

events of scientific interest, making it a fundamental computational tool in drug discovery (Schrödinger, 2022).

The principle behind MDS is that given the positions of all atoms in a biomolecular system, the force exerted on each atom by all other atoms can be calculated. Thus, making it possible to predict each atom's spatial position as a function of time using Newton's equations of motion. To be more specific, one calculates the forces on each atom periodically while stepping through time, updating each atom's location and velocity using those forces generating a trajectory. In essence, the generated trajectory is a three-dimensional movie that depicts the system's atomic-level configuration at each point over the simulated time period (Hollingsworth and Dror, 2018).

2.10.4 Molecular Mechanics/Generalized Born Surface Area (MMGBSA)

MMGBSA is a computational approach used to estimate binding free energies and analyze protein-ligand interactions (Genheden and Ryde, 2015). The overall approach includes energy calculations for each snapshot, molecular dynamics simulations to sample the conformational space, and statistical analysis to determine the average binding free energy. By breaking down the free energy of binding into several energy components such as van der Waals interactions, electrostatic interactions, solvation energies, and entropy contributions, MMGBSA is able to compute the free energy of binding (Genheden and Ryde, 2015). MMGBSA can be applied to a wide range of studies including drug binding affinity Prediction, protein-ligand binding free energy calculations and protein-protein interaction studies (Wang *et al.*, 2018).

The strengths of MMGBSA contributing to its value drug discovery and protein-ligand interaction studies includes its computational efficiency, making it less demanding compared to more rigorous methods such as free energy perturbation or thermodynamic integration (Virtanen *et al.*, 2015). This efficiency allows for the analysis of large datasets and enables rapid screening of potential ligands, saving valuable time and computational resources. Additionally, MMGBSA has demonstrated reasonable accuracy and good correlation with experimental binding affinities (Wang *et al.*, 2019; Forouzesh and Mishra, 2021; Taylor and Ho, 2023).

Hybrid approaches that combine MMGBSA with other methods, such as quantum mechanics calculations or enhanced sampling techniques, have shown promise in improving accuracy and accounting for specific phenomena (Massova and Kollman, 2000; Jerome, 2015). Enhanced sampling techniques, such as molecular dynamics or metadynamics, can enhance conformational sampling and provide more accurate free energy estimates. Additionally, the integration of machine learning approaches has been explored to improve MMGBSA predictions by capturing complex relationships between molecular features and binding affinities (Dong *et al.*, 2021; Ji *et al.*, 2021).

2.10.5 Machine learning (ML) in drug discovery

Machine learning (ML) is playing an increasingly important role in drug discovery. ML algorithms analyze large amounts of data to identify patterns and make predictions, which can be used to inform drug discovery efforts (Hansen, 2016). For example, ML can be used to predict the activity of potential drug compounds, to prioritize compounds for further testing, and to predict the toxicity and side effects of drugs (Chen *et al.*, 2018). Drug discovery aims at finding novel molecules with chemical characteristics for the treatment of disease. With that objective in mind, recent advances in technology have enabled the use of ML tools and techniques in drug discovery to the research process while reducing risk and expenditure during clinical trials (Dera *et al.*, 2021). One popular approach in drug discovery is to use ML algorithms to analyze high-throughput screening data, where the activity of many compounds is measured against a particular target (Chen *et al.*, 2019).

Another approach is to use ML to analyze chemical and biological data to predict drug-target interactions and guide the design of new compounds (Xu *et al.*, 2020). The application of ML in drug discovery is rapidly advancing, and it is expected to have a significant impact on the speed and efficiency of drug discovery efforts in the coming years (Kohavi *et al.*, 2018). These techniques are aimed at improving the decision-making process regarding pharmaceutical data with applications such as quantitative structure activity relationship (QSAR) generating predictive models, *de novo* drug design and hit discovery, retrieving more accurate outcomes. ML is a category that falls under artificial intelligence tools, utilizing numerous algorithms to interpret and attain knowledge from data (Chandra and Hareendran, 2021).

2.10.6 Quantitative structure activity relationship

Through mathematical systems, QSAR possesses the ability to predict with a high level of accuracy, the physicochemical and biological fate properties that a *de novo* compound will have based on the knowledge of its chemical structure and from previous experimental studies by numerically linking the chemical structures with their biological activities (Carracedo-Reboredo *et al.*, 2021). To theoretically anticipate the biological activity, QSAR models integrate computational and statistical methodologies, permitting the development of potential novel molecules without having to go through trial-and-error process of organic synthesis (Carracedo-Reboredo *et al.*, 2021). This method is cheaper and faster as it exists in a virtual environment, relating chemical structure to biological activity, without the need for certain resources such as equipment, reagents, instruments, and kits (Carracedo-Reboredo *et al.*, 2021). To be able to carry out a QSAR study, one needs molecular structures of different compounds with a common mechanism of action, biological activity of each ligand in the study and physicochemical properties (Coronin and Schultz, 2003).

For the QSAR models to be able to predict the biological activities of ligands that do not yet exist, is by generating equations to calculate descriptors with a high probability of pharmacological success (Carracedo-Reboredo *et al.*, 2021). There are different QSAR methods that can be used in drug discovery to train models that predict the activity or chemical properties of compounds. An overview of some popular QSAR techniques is provided below:

- Molecular descriptors, which are numerical representations of molecular attributes including size, shape, and chemical properties, are frequently used in QSAR models. Based on chemical structures, these descriptors may be computed and utilized as input features in machine learning models (Khan, 2016).
- Multiple Linear Regression (MLR), is a well-known QSAR technique that uses linear equations to connect chemical descriptors with activity data. In order to build a predictive model, it makes the assumption that there is a linear relationship between the descriptors and the activity (Santiago *et al.*, 2018).
- An ensemble learning technique called Random Forest (RF) mixes many decision trees to produce a prediction model. The final forecast is achieved by averaging the predictions made by each individual tree, each of which is trained on a distinct sample of the data (Svetnik *et al.*, 2003; Haffar *et al.*, 2023).
- Neural networks: Convolutional and recurrent neural networks, as well as other deep neural networks, can be used for QSAR modeling. These networks may capture intricate connections between molecular characteristics and activities because they are made up of several layers of linked nodes (neurons) (Alzubaidi *et al.*, 2021).
- Active learning QSAR is a methodology that involves the selection of informative data points for annotation by an oracle (expert) to train a predictive model. With the goal of reducing the amount of labeled data required while maintaining or improving model performance (Reker and Schneider, 2015; Tu *et al.*, 2023).

2.11 ATPase activity assay

ATPase activity assays are biochemical methods used to measure the hydrolysis of adenosine triphosphate (ATP) by ATPase enzymes (Bartolommei *et al.*, 2013; Rule *et al.*, 2016). ATPase activity assays measures the release of inorganic phosphate (Pi) during ATP hydrolysis by ATPase enzymes (Bartolommei *et al.*, 2013). These assays have been widely employed in the study of Hsp90, to investigate the role of Hsp90 (Rowlands *et al.*, 2010), also modulation of Hsp90 by monitoring ATP hydrolysis in the presence or absence of modulators, the impact on Hsp90 function can be assessed, providing insights into its allosteric regulation and potential therapeutic

targeting (Obermann *et al.*, 1998; Dixit and Verkhivker, 2012). Even in High-throughput screening campaigns to identify small molecule modulators of Hsp90 ATPase activity (Rowlands *et al.*, 2004). Over the years these assays have been adapted for automation, allowing the rapid evaluation of compound libraries for potential Hsp90 inhibitors and activators.

2.12 Biophysical methods in early drug discovery

Biophysical methods are an integration of techniques from physics, chemistry and biology, used in the study of function, structure, and interactions of biological molecules, forming part of the key components in early drug discovery (Holdgate *et al.*, 2019). The most used biophysical methods include, X-ray crystallography, which is used as a starting point for many drug discovery studies by determining the 3D structure of molecules by crystallizing the molecule of interest and then exposing it to X-ray radiation, allowing the study of specific interactions of a protein target and ligand (Smyth and Martin, 2000; Carvalho *et al.*, 2009). Isothermal titration calorimetry (ITC), is a method most used in drug design to provide a comprehensive thermodynamic and kinetic profile of ligands binding, by measuring interactions between protein and small molecules to determine the binding affinity, enthalpy, and entropy (Linkuvienė *et al.*, 2016; Su and Xu, 2018). Nuclear magnetic resonance (NMR) spectroscopy is an essential screening method in structure-based drug discovery for studying ligand binding to protein targets, offering crucial structural information on protein-ligand interactions, which can be used to help optimize weak-binding hits into high-affinity leads (Pellecchia *et al.*, 2002; Gossert and Jahnke, 2016). Surface plasmon resonance Spectroscopy (SPR), SPR is a real-time optical method for studying biomolecular interactions. It monitors binding events by measuring changes in the refractive index of a medium near a metal surface without the requirement for labelling (Narayan and Carroll, 2017). Protein-protein interactions, small molecule binding, and DNA-protein interactions can be studied using SPR (Olaru *et al.*, 2015). Circular dichroism (CD) spectroscopy, is extensively used in drug discovery for evaluating the stereochemistry of chiral molecules and proteins, by providing information on protein secondary structure, folding, and conformational changes (Greenfield, 2006). There are many other biophysical techniques which can be employed in early drug discovery, but since our study aimed to investigate small molecule binding, hence SPR was employed.

2.12.1 Surface Plasmon resonance

SPR is a powerful biophysical method extensively employed in early drug discovery to investigate biomolecular interactions. The principle of SPR is based on the interaction between light and free electrons at the interface of a metal film and a dielectric medium. The phenomenon occurs when the incident light's wavelength matches the resonance condition, resulting in the generation of

surface plasmons (Deng *et al.*, 2017; Prabowo *et al.*, 2018). These surface plasmons can be detected by monitoring the changes in the refractive index of the medium near the metal surface, which is indicative of molecular interactions (Hinman *et al.*, 2018). SPR can be used in the study of protein-protein interactions (Douzi, 2017), small molecule interactions (Frostell *et al.*, 2013), DNA and RNA Interactions (Licatalosi, and Jankowsky, 2020). SPR can provide valuable information about conformational changes in biomolecules upon ligand binding (Paynter and Russell, 2002).

CHAPTER 3

3. MATERIALS AND METHODS

3.1. *In silico* methods

3.1.1 Computational software programme

The following tasks were performed using the graphical user interface (GUI) Maestro v12.9 in Schrödinger Suite 2021-3. All molecular modelling calculations were performed on various modules available in Maestro such as Pathfinder, Ligprep, protein preparation wizard, Glide, AutoQSAR, IFD, Desmond, MM-GB/SA and Qikprop. The process flow diagram depicting the *in silico* experimental procedures followed in this dissertation is presented in Figure 5.

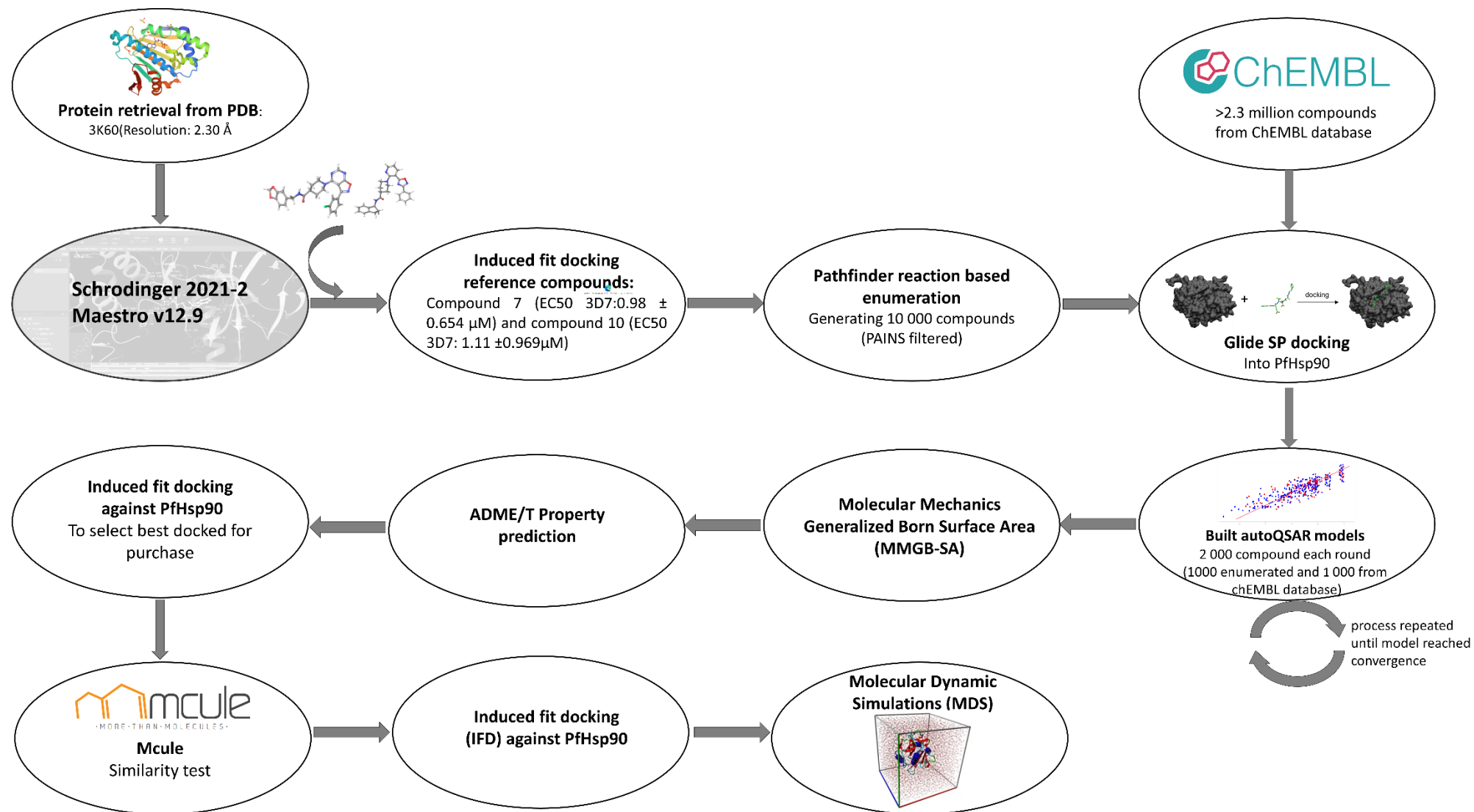


Figure 5: *in silico* methodology flow diagram

3.1.2 Protein retrieval and preparation

The crystal structure of the N-terminal domain of PfHsp90 bound to ADP with a resolution of 2.30Å [PDB code: 3K60] (Corbett and Berger, 2010), was retrieved from the protein data bank (www.rcsb.org) and prepared by implementing the protein preparation wizard module in Maestro (v12.9). The pre-processing procedure was carried out to repair structural defects and add missing information. The protein was refined and waters beyond 3Å removed, while generating het states using Epik: pH = 7.4 +/- 2.0 (Sastry *et al.*, 2013; Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021). Hydrogen bond networks were optimized using PROPKA at pH = 7.4. Energy minimization of the protein structure were done by converging heavy atoms to RMSD value of 0.30Å using an optimized potential for liquid simulations 4 (OPLS4) force field.

3.1.3 Ligand preparation

In a previous study conducted by Everson and colleagues (2021), compounds 7 (EC₅₀ *P. falciparum* 3D7a 0.98 ± 0.654 µM) and compound 10 (EC₅₀ 3 *P. falciparum* D7a 1.11 ± 0.969 µM) were shown to be near sub-molar potency towards chloroquine sensitive *P. falciparum* strain. These compounds were used as reference compounds for this study. To understand the binding mode and nature of interaction that these compounds have towards PfHsp90, the 2D structures of compounds 7 and 10 were drawn using 2D sketcher on maestro and converted to 3D structures using Ligprep module (Ligprep, Schrödinger, LLC, New York, NY 2021), generating low energy structures with Three-dimensional (3D) coordinates which has tautomeric states. Additionally, the study used two groups of ligands, approximately >2 million compounds from chEMBL database (<https://www.ebi.ac.uk/chembl/g/#browse/compounds>) and 10 000 enumerated products generated in this study. Low energy 3D structures with tautomeric states of the compounds generated, from 2D structures of both the enumerated products and the ligands downloaded from the chEMBL database using the LigPrep module (Ligprep, Schrödinger, LLC, New York, NY 2021). The input structures were optimized using OPLS4 force field, generating possible ionization states at a target pH of 7.4 +/- 2.0 using EPIK. Stereoisomers were computed to retain specified chiralities at most generating 32 per ligand (LigPrep, Schrödinger, LLC, New York, NY, 2021).

3.1.4 Induced fit docking (IFD)

Induced fit docking (IFD) was implemented to predict the binding modes and conformational changes introduced into PfHsp90 upon binding different ligand. The IFD flexible docking protocol

relies on glide and refinement module in prime for accurate scoring. The IFD protocol on maestro uses glide to initially produce multiple poses of the ligand within the receptor (Halgren *et al.*, 2004; Friesner *et al.*, 2004; Friesner *et al.*, 2006: Glide, Schrödinger, LLC, New York, NY, 2021) and prime to refine them using an automated process from a python script. The whole process is repeated several times producing a final list of poses ranked according to prime energy (Jacobson *et al.*, 2002; Jacobson *et al.*, 2004; Prime, Schrödinger, LLC, New York, NY, 2021) and glide XP empirical scoring functions, accurately predicting the ligand binding modes and the resulting structural changes of the receptor. To dock compounds 7 and 10, a grid box of the binding site for the prepared structure of PfHsp90-NTD was generated considering the co-crystallized ligand ADP as a centroid. Followed by the removal of ADP from ATP binding region residues to provide more room for ligand docking. To refine the side chains of residues located within a 5Å distance from the ligand, the prime refinement step was employed. After the initial docking of each ligand, up to 20 poses were selected for further refinement using the XP mode.

3.1.5 Reaction-based enumeration

The 2D sketcher on Maestro was used to draw the structures of compound 10 which was minimized using the macromodel module, followed by displaying of breakable bonds on compound 10. Pathfinder-driven enumeration was performed on compound 10 with a maximum depth of 1.

Table 1: Pathfinder generated reaction pathways -Schrödinger Suite 2021-3

Pathway	Cross-coupling method
Path 1	Amide_coupling-1
Path 2	Amide_coupling-2
Path 3	Amination-1
Path 4	Hayima-1
Path 5	Negishi
Path 6	Oxadiazole-1
Path 7	Stille
Path 8	Suzuki
Path 9	Suzuki-2

Pathfinder analysis of compound 10 yielded a total of 9 pathways (Table 1) that could be potentially used for synthetic route as a starting point. Oxadiazole-1 (Figure 6), pathway 6 was selected as pathway of interest as it maintains critical contacts with the receptor. Using

Oxadiazole-1 pathway, the reactants were defined, with reactant 2 retained as the original reactant while reactant 1 was selected to be the variable reactant. The ligands containing reactive functional groups were eliminated using smiles arbitrary target specification (SMARTS), and Pan assay Interfering Structures (PAINS) offenders that generally react with anything were filtered. The core scaffold (reactant 2) was kept while different functional groups were incorporated in place of reactant 1, generating up to 10 000 products using enumeration. Default physiochemical parameters such as a MW between 150 and 575, LogP between -1.50 and 5.0, a TPSA between 30 and 150, HBA between 0 and 12, and HBD between 0 and 5, and a maximum number of rotatable bonds less than 10 were retained for design ideas.

Pathway 6 which is oxadiazole-1 coupling was chosen as the pathway of interest shown in Figure 6 below.

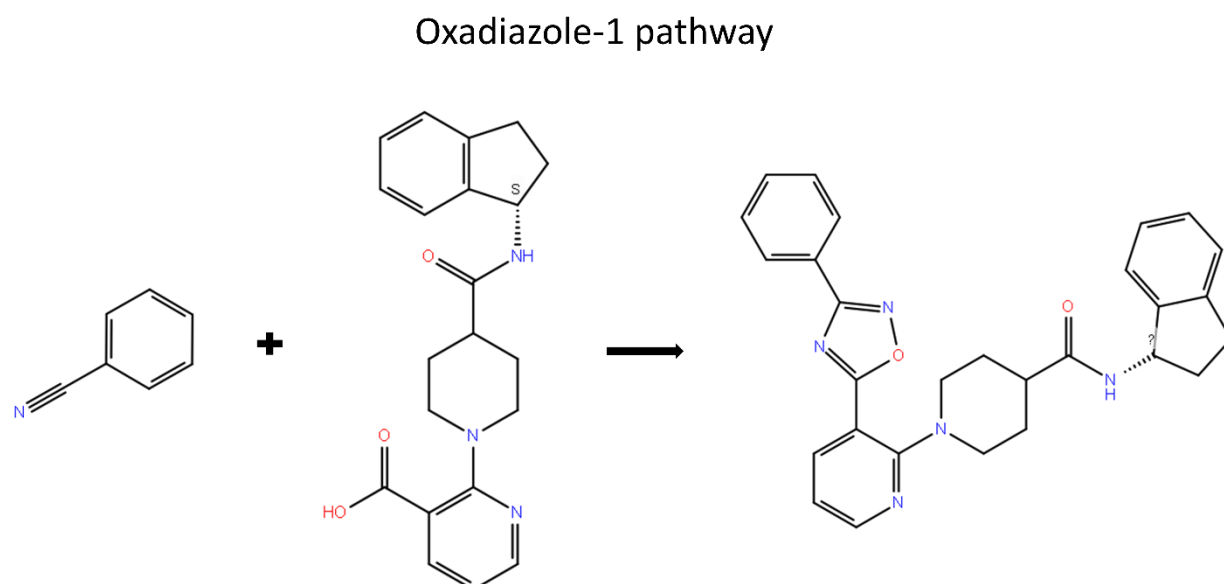


Figure 6: Oxadiazole-1 selected pathway.

3.1.6 Classical glide SP ligand docking

Classical ligand docking, a glide workflow offering speed and accuracy for virtually screening a large number of ligands against a target receptor, offering SP (standard precision) for reliably docking large number of ligands with high accuracy, or XP (extra precision) mode which further eliminates false positives by extensive sampling and advanced scoring, resulting in high enrichment (Halgren *et al.*, 2004; Friesner *et al.*, 2004; Friesner *et al.*, 2006; Farid *et al.*, 2006; Sherman *et al.*, 2006; Glide, Schrödinger, LLC, New York, NY, 2021). Glide explores favourable interactions, linking the receptor and the ligand. The two sets of ligands (enumerated and chEMBL database ligands) were subjected to ligand docking studies to select ligands exhibiting

favourable binding affinity (<-5 kcal/mol) towards PfHsp90. The ligands exhibiting favourable binding energies were then used to build the active learning models. The glide ligand docking protocol from Maestro (v12.9) software Schrödinger suite 2021-3 was used for the calculations. As receptor grid calculations are required for running ligand docking jobs, the prepared receptor of PfHsp90 complexed with ADP (PDB code: 3K60) were generated using the receptor grid generation panel in maestro (Schrödinger 2021-3).

3.1.7 Receptor grid generation

A receptor grid was generated based on ligand binding residues to specify the position and size of the receptors active site for glide SP ligand docking, utilizing the receptor grid generation tool in Maestro (v12.9) Schrödinger 2021-2. With PfHsp90 positional constraints were defined as ASN37, ARG98 and ILE173 as they are residues found to be specific to PfHsp90 (Wang *et al.*, 2014). As the structure of PfHsp90 is complexed with ADP the ligand was selected to be excluded from the grid generation.

3.1.8 Active learning model generation

The study employed active learning protocol enhanced by ligand docking utilizing AutoQSAR. AutoQSAR an automated tool facilitating the generation of predictive QSAR models using a best practice workflow which has the ability for descriptor generation, feature selection, create many QSAR models with varying training/test sets splits from various machine learning models (Dixon *et al.*, 2016; AutoQSAR, Schrödinger, LLC, New York, NY, 2021). The AutoQSAR module available in the Maestro software (v12.9) was used to generate the active learning models enhanced by ligand docking. AutoQSAR is a best practice protocol for automated generation of quantitative structure activity relationship (QSAR) models with limited user input and understanding. It also builds categorical or numerical models based on informative descriptors and fingerprints, where a given model is trained against a particular random subset of input structures.

The docked poses of the top scoring compounds from the glide SP ligand docking were selected to construct the QSAR model. The build model task was chosen on the AutoQSAR module, the model was then built from the project table selected entries (top scoring) and the property to be tested was the docking score. The data set was divided into 70% training and 30% test sets. Upon completion, the view model and make predictions task was chosen to view the output. The model report was then viewed to select the best model based on the scoring (R^2 Q^2 RMSE and SD) which was model 1. This model was then used to predict the second set of compounds.

Approximately 2000 ligands were randomly selected from both the enumerated products and ligands from the chEMBL to make a prediction, the job was given a particular name. This gave the predicted activity of the ligands. The output was incorporated into the working directory and then subjected to glide SP ligand docking. The top scoring docked poses were then selected and added to the first set of compounds to build the second model (Model 2). The same procedure was repeated to build model 3 to 14, testing almost 40 000 ligands. The top scoring compounds from the active learning model enhanced by ligand docking were selected and subjected to IFD.

3.1.9 *In silico* ADME predictions

QikProp was used to efficiently evaluate pharmaceutically relevant properties such as absorption, distribution, metabolism, and excretion (QikProp, Schrödinger, LLC, New York, NY, 2021). Qikprop module which part of the Schrödinger suite 2021-3 was used to calculate the physicochemical parameters of the compounds. The acute, easy-to-use absorption, distribution, metabolism, and excretion (ADME) predictions of the top scoring compounds from the induced fit docking. When conducting a Qikprop analysis, atom types and charges, rotor counts, and sample molecule's volume and surface area are all evaluated. Qikprop then employs this information in regression equations, together with the derived physical descriptors. Consequently, the pharmacologically relevant features of a molecule would then be accurately predicted. Qikprop analysis can be conducted using either normal mode or fast mode, in which certain time-consuming calculations are disregarded. Using the ligand based ADME/T predictions task on Maestro physiochemical properties of the ligands of interest were predicted.

Physicochemical properties investigated Include molecular Weight, Lipophilicity (LogP), Hydrophilicity (LogD), pKa and Lipinski's Rule of Five. Molecular Weight influences the drug's solubility, permeability, and absorption in the body with its maximum threshold being 500g/mol, usually, drugs with molecular weights above 500g/mol may face challenges in crossing biological barriers. LogP is the drug's ability to dissolve in lipid or fat-based environments. It affects the drug's distribution, tissue penetration, and membrane permeability. Typically, a LogP value of up to 5 is considered desirable. LogD represents the drug's ability to dissolve in water or aqueous environments. It influences drug solubility, formulation, and elimination a LogD value close to zero is preferred for good water solubility. pKa indicates the ionization state of a compound at a specific pH. It influences drug absorption, distribution, and interactions with target receptors. Ideally, a drug's pKa should be within a suitable range for optimal absorption and activity. Lastly Lipinski's Rule of Five is a widely used guideline in drug discovery that helps assess a compound's drug-likeness and oral bioavailability. According to Lipinski's rule, for a drug to have a higher chance of success:

- The molecular weight should be below 500g/mol.
- The LogP value should be less than 5.
- The number of hydrogen bond donors should be less than 5.
- The number of hydrogen bond acceptors should be less than 10.

These guidelines serve as a general indication, but exceptions can be made depending on the specific context and therapeutic target. Drug candidates that violate one or more of Lipinski's rule may still be successful if they possess other favorable properties.

3.1.10 Free binding energy calculations

The molecular mechanics with generalized born surface area (MM-GBSA) was performed to calculate the free binding energies determining the stability of the protein-ligand complexes from docking conformations, using the prime module in Maestro (v13.2) from the Schrödinger suite 2022-2. MMGBSA estimates the binding free energy of a ligand-receptor complex by combining molecular mechanics force fields, which describe the intermolecular interactions, with a solvation model based on the Generalized Born (GB) theory (Genheden and Ryde, 2015; Forouzesh and Mishra, 2021). The calculations for the docked complexes were on OPLS4 force field, VSGB solvation model and the minimized the sampling.

$$\Delta G_{bind} = \Delta E + \Delta G_{solv} + \Delta G_{SA} \quad (2)$$

Where ΔE = Difference in minimized complex energy and Σ energies of unbounded receptor and ligand; ΔG_{solv} = Difference in GBSA solvation energy of receptor-ligand complex and Σ solvation energies of unbounded receptor and ligand; ΔG_{SA} = Difference in surface area energy of IFD complex and Σ surface area energies of unbounded receptor and ligand.

$$\Delta E = E_{complex} - E_{protein} - E_{ligand} \quad (3)$$

Where $E_{complex}$, $E_{protein}$, and E_{ligand} are the minimized energies for protein-inhibitor complex, protein, and inhibitor, respectively.

$$\Delta G_{solv} = \Delta G_{solv}(complex) - \Delta G_{solv}(protein) - \Delta G_{solv}(ligand) \quad (4)$$

$$\Delta G_{SA} = \Delta G_{SA}(complex) - \Delta G_{SA}(protein) - \Delta G_{SA}(ligand) \quad (5)$$

Where, ΔG_{SA} is the nonpolar contribution to the solvent energy of the surface zone. G_{SA} (complex), G_{SA} (protein), and G_{SA} (ligand) are the surface energies of complex, protein, and ligand, respectively.

3.1.11 Molecular dynamic simulation

Molecular dynamics simulations (MDS) on the Desmond module to investigate the dynamic behavior of molecules, explore their conformational changes, and provide insights into their interactions and properties (Ivanova *et al.*, 2018). Selected docking poses were subjected to molecular dynamic (MD) simulations using the Desmond module from Schrödinger suite 2021-3, to evaluate the conformational stability of interactions in a similar physiological condition to the human body in a particular time range. Performing an MD simulations procedure involves several steps such as simulation set up, relaxation of the system, running the simulation, viewing the trajectories and analysis of the results. System builder was utilized to generate a solvated system for simulation. The prepared complexes were solvated using the transferable intermolecular potential with 3 points (TIP3P) model in an orthorhombic box and minimized the box volume. Na^+ ions and Cl^- ions were added at a salt concentration of 0.15 M to neutralize the system using OPLS4 force field (Jorgensen *et al.*, 1996). For each system, simulation was carried out at 50-100 ns, with a simulation trajectory of 50 ps and 500 approximate number of frames. NTP ensemble class was used at a temperature of 300.0K and pressure of 1.01325bar.

3.2 Biochemical methods

3.2.1 Transformation

The nucleotide sequences coding for the N-terminal domains of PfHsp90 (Pasma DB ID: PF3D7_0708400) in vector pET28b (+) was codon optimized for expression in *Escherichia coli* (*E. coli*) and synthesized by Genescript (USA). The pET28b (+)-PfHsp90-NTD was used to transform in *E. coli* BL21 (DE3) cells (Appendix 1.1.4).

3.2.2 Plasmid extraction

A colony from transformed *E. coli* BL21 (+) cells containing the pET28b (+)-PfHsp90-NTD was inoculated into 5 mL of 2 X YT broth (16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl) with kanamycin (50 μ g/mL) (glentham life sciences, corsham, UK) and incubated overnight at 37°C with shaking at a speed of 160 rpm. The following day the plasmid was extracted using the pure Yield™ plasmid miniprep system (Promega, Madison, USA) as per manufactures instructions (Appendix 1.1.5). The integrity of the extracted plasmid was analysed by running a 1% agarose

gel electrophoresis (Appendix 1.1.7) and the purity and concentration of the extracted Plasmid was estimated using the Thermofisher™ Nanodrop instrument.

3.2.3 Restriction enzyme digest

The integrity of the plasmid was confirmed by restriction digest. The Promega restriction enzymes *Bam*HI and *Nde*I were used to digest the plasmid pET28b (+)-PfHsp90-NTD (Appendix 1.1.6).

3.2.4 Protein expression

A single colony from transformed *E. coli* BL21 (+) cells was inoculated into 25mL 2x YT broth containing 50µg/mL kanamycin and then incubated at 37°C overnight with shaking. The starter cultures were then sub-cultured into 225mL of 2x YT broth containing 50µg/mL kanamycin, making 250mL of culture. The cultures were grown to a density of OD₆₀₀ ranging from 0.4 to 0.6. Once the desired OD₆₀₀ was reached, protein expression was induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was then incubated at 25°C with shaking for 24 hours while measuring OD₆₀₀ after every hour until time point 6 (T6) and collecting 1 mL of samples hourly into 1.5 mL sterile Eppendorf tubes. The overnight sample and OD₆₀₀ reading were taken in the morning.

The 1 mL hourly samples taken were configured at 14000xg for 1 minute and supernatant discarded. The remaining pellet was resuspended in 1X phosphate buffer saline (PBS ;137 mM NaCl, 2.7 mM KCl 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.5). To ensure that equal number of bacterial cells were used throughout the protein expression studies, a lysate was created with the following PBS amounts:

$$\frac{ABS(OD600)}{0,5} \times 150 \quad (6)$$

The obtained samples were kept at -20°C for SDS-PAGE analysis. For this study un-induced cell cultures were used as controls.

3.2.5 Protein purification

Expressed PfHsp90-NTD was purified under native conditions. In brief, transformed *E. coli* were grown in 2x YT broth for 4 hours, the cells were then harvested by centrifugation. The pellet was re-suspended in 5mL of lysis buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 20 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF)). The lysate was then stored overnight at -80°C. The

following day the lysate was thawed on ice, while adding 10mg/mL lysozyme after 10 minutes until completely thawed. The lysate was clarified by sonication for 30 seconds for 4 cycles. This was followed by centrifugation at 4000xg for 15 minutes removing the cell debris. Pellet and supernatant fractions of approximately 80µl were taken and added into a 1.5 mL tube for SDS-PAGE analysis. Protein purification was initiated by addition of supernatant onto pierce™ high-capacity Ni-IMAC resin (Thermo Scientific™, Germany). Binding was allowed to occur overnight on ice with shaking. The mixture was then centrifuged at 4000xg for 5 minutes, a flow through sample was collected. The remaining unbound proteins were discarded, the beads were then washed three times with 2 mL of wash buffer (10 mM Tris-HCl pH 8, 300 mM NaCl, 50 mM imidazole and 1 mM PMSF) to remove the remainder of weakly bound proteins. PfHsp90-NTD was then eluted three times by adding 1 mL of the elution buffer (10 mM Tris-HCl pH 8, 300 mM NaCl, 750 mM imidazole and 1 mM PMSF). The eluted purified protein was stored at 4°C.

3.2.6 Protein concentration determination

The Bradford assay is used to detect the amount of protein in an unknown protein concentration solution based on the complex formation between the brilliant blue G-250 dye and the proteins in a solution. This results in the red form of the dye turning blue when it binds to protein. Thus, absorbance readings can be used to calculate the amount of protein present in a solution. A 20mg/mL bovine serum albumin (BSA). Stock solution of serial dilutions of BSA in Tris-HCl, pH8 was used to construct a standard curve. A BSA stock solution of 20mg/mL and 0,10M of tris-HCl pH8 were firstly prepared. Dilution series of 1mg/mL, 0.75mg/mL, 0.5mg/mL, 0.25mg/mL and 0.25mg/mL BSA concentration were prepared. For the protein samples 3 sets of dilutions of the unknown protein concentration was prepared, as follows: 1x sample = 20mL of sample protein, 2x dilution = 10mL to 10mL Tris-HCl, pH8 and 4x dilution = 5 mL to 15 mL Tris-HCl, pH8. Eppendorf tubes containing the standards and unknown protein dilutions were labeled. 1mL of Bradford reagent was added to the tubes containing 20mL of the samples. The samples were vortexed after the addition of Bradford Reagent, to ensure the samples are properly mixed. Followed by incubation of the samples for 5 minutes at room temp and then vortexed again. 250µL of each sample was then transferred to a 96 well microplate in triplicates. Blanks containing Tris-HCl, pH8 only and Tris-HCl + Bradford reagent were also added to the plate as controls. The absorbance was measured at 595nm, using the Glowmax explorer (Promega, USA). A standard curve of BSA concentration vs Absorbance was prepared, using Microsoft excel. The unknown protein concentration was calculated from the standard curve equation, where X-intercepts is the unknown protein concentration.

3.2.7 ATPase activity assay

The colorimetric ATPase assay was used to investigate the amount of phosphate released by PfHsp90-NTD when ATP is being hydrolysed to ADP. The ATPase activity of PfHsp90-NTD was determined by using a previously described assay by Matambo *et al.* (2004). In brief, the basal ATPase activity of the molecular chaperones was determined by setting up a reaction with assay buffer containing 10% SDS, 0.5M KCl, 0.5M HEPES pH 7.4, 10mM Dithiothreitol (DTT), and 0.5M MgCl₂. The master mix of assay buffer was prepared in Pi-free glassware and dispensed 130µl into labelled Eppendorf tubes and varying concentrations of PfHsp90-NTD were then added to the assay buffer. The volume of the amount of chaperone to be added was calculated and added, then filled up to 1mL with distilled H₂O. The reaction was initiated by addition of 2mM ATP as previously established (Pallavi *et al.*, 2010). Samples were analysed at 30 minutes after the addition of ATP. About 50µl of 10% SDS was added to a 96 well plate, followed by the addition of 50µl of the sample reaction to the SDS containing wells stopping the reaction. This was done in triplicates and repeated at every time point. Once the samples were taken at all time points, the reaction was developed by adding 50µl ascorbic acid and 50µl of 1% ammonium molybdate prepared in HCl. 125µl sodium citrate (2%) acetic acid (2%) was added to each well and the plate incubated for 10 minutes. The release of inorganic phosphate was quantified by spectrophotometrically reading the plate at 850nm (end point read) using a 96 well plate reader (Glowmax® Explorer multimode microplate reader, from Promega). The reaction was repeated in the presence of 0-15µM. To generate a phosphate standard curve, phosphate standards (0.02, 0.14, 0.6, 0.8 and 1nmol/µl) from a stock of 500nmol/µl (0.5M). The 0.5M stock was added to 3.2mL distilled H₂O making 100nmol/µl stock, then added it to 3.2mL of dH₂O to make up a final concentration of 20nmol/µl. 10µl of each standard dilution was added to the 96 well plate in triplicates and added 40µl assay buffer to keep the volumes the same as in the sample wells. 50µl of 10% SDS buffer was then added to each standard, and lastly developed the standards in the same manner and time as the ATPase sample reactions.

3.3. Biophysical methods

3.3.1 Surface plasmon resonance (SPR)

Surface Plasmon Resonance (SPR) is a powerful analytical tool for studying molecular interactions in real-time, label-free, and non-invasive manner. It is widely used in various fields including biochemistry, pharmacology, and nanotechnology. A typical SPR setup includes a metal film (usually gold or silver) deposited on a transparent substrate. Light is shone onto the metal film and the resulting evanescent wave penetrates the surrounding medium. If a biomolecular interaction occurs at the metal surface, it will result in a change in the refractive index, causing a

shift in the resonant angle of the SPR. This shift is proportional to the amount of mass change and can be used to determine the binding kinetics, affinity, and specificity of the interaction (Nie and Emory 1996; El-Sayed, 2010; Felgner, 2015; Dukor, 2012).

In this study for the surface preparation, the sensor chip was cleaned using 2M NaCl/0.1M NaOH for a period of 10 minutes. The Carboxymethyl dextran coated sensor slide (BioNavis SPR102-CMD-3D) was used to immobilize the protein of interest PfHsp90-NTD, followed by activation of the chip with 0.1M EDC/0.05M NHS. The protein was coupled with sodium acetate (5mM) pH 5. To remove excess NHS-EDC, the chip was deactivated with 1M ethanolamine HCl, and then finally clean the sensor using NaCl/NaOH unspecific binding molecules. Serial dilutions of the compounds were performed and analysed. Five dilution series of the compound (5000nM-312.5nM and 2000nM-125nM) were measured at 25°C. Samples in PBS (pH=7.4) [phosphate buffered saline] containing 10% DMSO to enhance drug solubility. For Sample injection, different concentrations of the compounds were introduced into the flow cell and allowed to bind to the surface. The SPR instrument Bionavis 420A SPR system collects the reflectance data as a function of the angle of incidence. The data is processed to determine the resonance angle, which is related to the mass change at the surface. The kinetics and affinity of the interaction were analysed using TraceDrawer™ for MP-SPR Navi™ (BioNavis Ltd. Finland.).

CHAPTER 4

4.1 RATIONAL DESIGN RESULTS AND DISCUSSIONS

4.1.1 Reaction-based enumeration

The PfHsp90 shares a 69% sequence similarity with the human Hsp90, therefore making it challenging to design novel small molecule inhibitors with selectivity towards PfHsp90. Targeting PfHsp90 without affecting human Hsp90 is challenging because inhibiting human Hsp90 may lead to significant adverse effects due to its critical roles in normal cellular processes. There are two studies that demonstrated that it is possible to design selective inhibitors of PfHsp90, the studies by Wang *et al.* 2016 and Everson *et al.* 2021 which demonstrated the benefit of rational drug design strategies to obtain selective inhibitors of PfHsp90. This study used compound 7 and compound 10 from the Everson *et al.*, (2021) study (Figure 7), as the reference compounds due to their proven activity against PfHsp90 and known low EC₅₀ values.

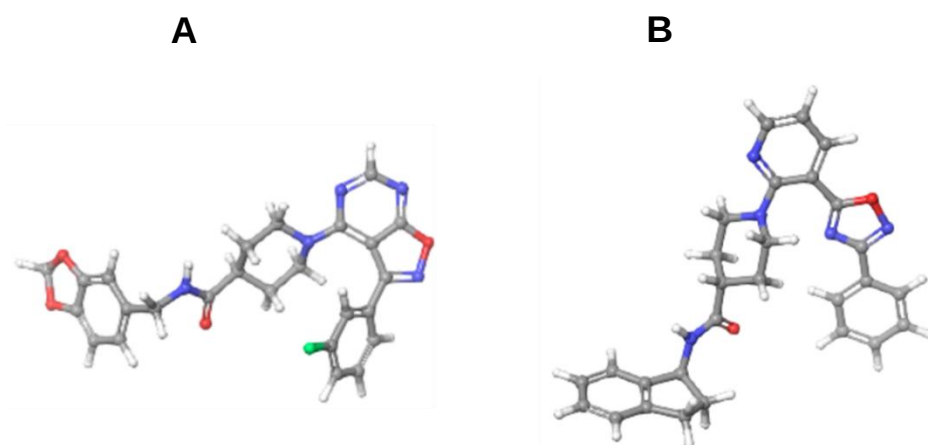


Figure 7: Reference compounds. With A being, compound 7 and B compound 10. In the molecules hydrogen atoms are indicated in white, aromatic rings and carbon atoms in grey, Oxygen atoms in red, nitrogen atoms in blue and Fluorine atom in green.

The reference compounds were subjected to IFD using Glide XP, with the end goal of selecting the compound with the greatest binding affinity towards PfHsp90 and subject it to reaction-based enumeration.

Table 2: IFD results of the reference compound 7 represented as CP-7 and compound 10 as CP-10.

Title	docking score	glide emodel	XP GScore	IFD Score	EC ₅₀ (uM) ^a
CP-7	-8.031	-79.054	-8.031	-534.01	0.98 ±0.654
CP-10	-10.485	-76.156	-10.485	-532.21	1.11 ±0.969

^a The two compounds were tested for antiproliferative activity against the chloroquine-sensitive *P. falciparum* 3D7 strain (MRA-102) in human O-positive erythrocytes (Everson *et al.*, 2021).

The IFD results showed that compound 10 had a better fitting into the ATP binding pocket of PfHsp90, evidenced by a docking score of -10.485 kcal/mol (Table 2). Compound 10 made several hydrogen water-mediated bond interactions with amino acid residues such as LYS44 and ARG98, respectively and pi-pi interactions with PHE124 and TRP148 (Figure 8A). Compound 7, on the other hand, had a slightly lower docking score of -8.031 kcal/mol (Table 2) also making fewer contacts with the PfHsp90 binding pocket, only two interactions were observed, a hydrogen bond with PHE124 and a water mediated hydrogen bond with LYS44 (Figure 8B).

Based on its high binding affinity towards PfHsp90, compound 10 was subjected to reaction-based enumeration. The top scoring docking pose of PfHsp90-compound 10 was examined and after visual inspection of the functional groups forming key interactions (Fischer *et al.*, 2021), it was seen that the outer benzene ring did not participate in hydrogen bonding or pi-pi interaction. Therefore, the bonds forming the outer benzene bond were broken (Figure 8), leaving the core (reactant 2) subjected to R-group enumeration which introduction of different chemical structures in place of the benzene. Here, the intention was to generate compounds with better affinity for PfHsp90 than compound 10 without much alteration of the core structure.

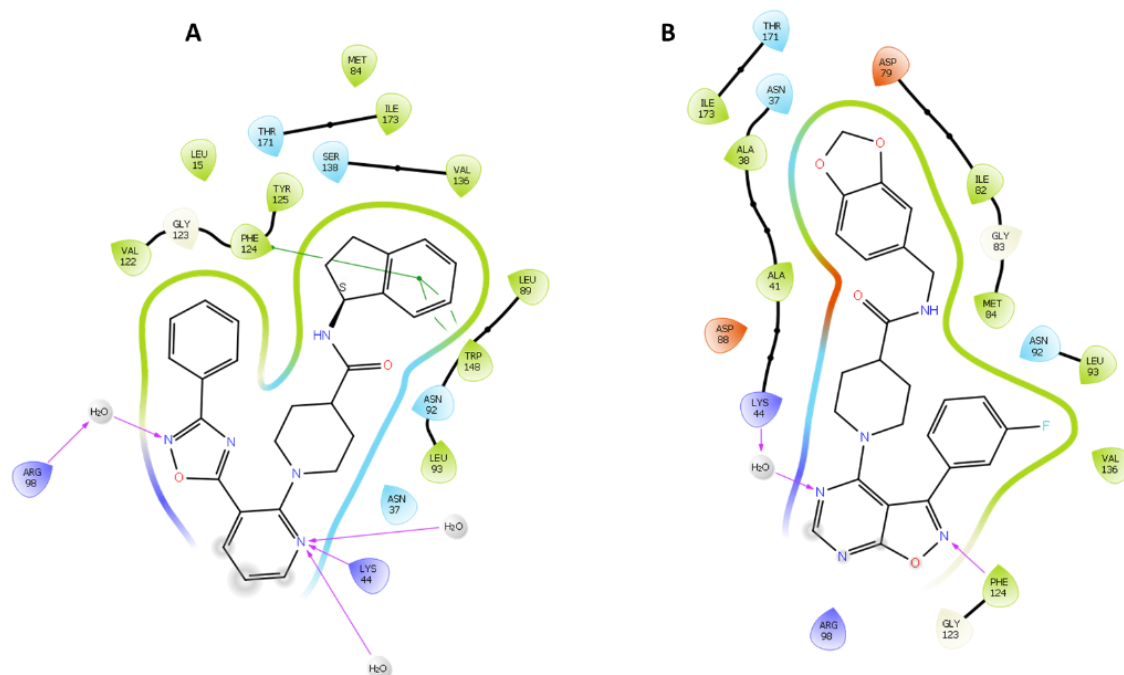


Figure 8: Details the ligand interaction diagrams for compound 10 and 7 in the PfHsp90 N-terminal domain ATP binding site. (A) Compound 10 making pi stacking interactions with PHE124 and TRP148, a water mediated hydrogen bond with ARG98 and lastly a direct hydrogen bond with LYS44. (B) Compound 7 making a water mediated hydrogen bond with LYS44 and a direct hydrogen bond with PHE124. Positively charged amino acids indicated in blue, negatively charged amino acids in orange, polar in light blue, non-polar in green, Pi-Pi stacking interactions within a 3Å in dark green, all hydrogen bond interactions radius in purple.

Pathfinder is a retrosynthetic analysis application that incorporates more than 100 reactions, allowing for the rapid exploration of synthetically tractable ligands. Reaction-based enumeration was carried to create a large library of synthetically tractable compounds *in silico* (Konze *et al.*, 2019). This study implemented a pathfinder reaction-based enumeration for the conversion of compound 10 (Everson *et al.*, 2021). The result of this was 10 000 design ideas that were chemically stable, synthetically tractable, and as unique drug like ligands of compound 10. In the process of generating the design ideas SMARTS filters were employed to remove pan assay interfering compounds (PAINS). This removed all ligands that will generally react with anything.

4.1.2 Active learning model enhanced by ligand docking

Initially, 2000 design ideas (1000 from enumerated products and 1000 from chEMBL) were randomly selected from all sets of pathways were prepared using Ligprep to conduct Glide SP docking using PfHsp90 as a receptor. The binding affinities (i.e. docking score) of these ligands towards PfHsp90 were calculated and used to select the compounds for the Auto QSAR models to be trained using enumerated products and those retrieved from the chEMBL database. Using the active learning strategy, models were developed to filter the large number of compounds and choose the ligands that would be further docked with a scoring function with more in-depth analysis, such IFD. The chEMBL database was explored to enable the discovery of more potent hit compound and unlock new areas of a chemical space in the active learning model.

Building the initial model, ligands with docking scores ranging from -5.0kcal/mol to -7.6kcal/mol were retrieved, totalling 123 ligands that satisfy the criteria, including the two reference compounds. AutoQSAR divided the selected ligands into the training and test sets, with a random training set at 75% (92 compounds) and test set at 25% (31 compounds). The resulting best model was selected based on the values of the ranking score of the model (score), the standard deviation of the model (SD), the training set accuracy (R^2), root-mean square error of the test set predictions (RMSE), test set accuracy (Q^2) and lastly the Q^2 value for the model in which the property is simply fit to the molecular weight (Q^2 MW). With a value of 1 in R^2 and Q^2 indicating a perfect prediction. The approach of adding 2000 compounds, followed by GlideSP docking was reiteratively followed until the model reached convergences, meaning even with addition of more compounds, the statistical parameters would not be altered.

Table 3 below details the 13 built models using the active learning approach including their statistical parameters. As it can be seen, on the table all the models were generated using kernel-based partial least square regression (KPLS) but they differed in terms of binary fingerprints this is due to the compounds being chemically and structurally diverse. It can be noted that models 2

to model 8 lost some correlations (represented by overall score). This can be explained by the random variation or statistical fluctuations (Aggarwal, and Ranganathan, 2016). The observed decrease in correlation does not necessarily indicate a significant decline in model performance but rather reflect the inherent variability in the data.

Table 3: The best Auto QSAR models were used to train the listed design concepts to predict their docking scores, implementing an active learning technique to assess ability to bind to PfHsp90

Model	Model code	Score	S.D	R ²	RMSE	Q ²	Q ² MW (Null hypothesis)
1	Kpls_linear_36	0,7476	0,3669	0,7698	0,3576	0,6687	-0,0020
2	Kpls_molprint2D_31	0,6413	0,4500	0,6341	0,4164	0,5368	-0,0083
3	Kpls_radial_34	0.6519	0.4601	0.6767	0.4276	0.5864	-0.0038
4	Kpls_molprint2D_42	0.5781	0.4414	0.6787	0.4650	0.5321	0.1204
5	Kpls_radial_37	0.6652	0.4556	0.6680	0.4417	0.5986	0.0876
6	Kpls_desc_37	0.6029	0.4828	0.5883	0.4467	0.5937	0.0659
7	Kpls_desc_44	0.5362	0.4713	0.5919	0.4792	0.5271	0.0026
8	Kpls_molprint2D_27	0.5910	0.5213	0.5770	0.4929	0.5973	-0.0013
9	Kpls_desc_33	0.5898	0.4543	0.7140	0.4999	0.6009	-0.0014
10	Kpls_radial_33	0.6255	0.4596	0.7076	0.4855	0.6167	-0.0009
11	Kpls_radial_36	0.5836	0.5227	0.6331	0.5345	0.5659	-0.0013
12	Kpls_radial_42	0.5951	0.5498	0.6446	0.5623	0.5755	0.0010
13	Kpls_desc_23	0,5563	0,4756	0,7460	0,5600	0,6182	0,0727

For the first model built, KPLS_linear_36 (Figure 9) was shown to be the best model, which is a QSAR model generated by KPLS with linear fingerprints. The model had a score of 0,7476, SD of 0,3669, R² of 0,7698, RMSE of 0,3576 Q² of 0,6687 and Q² MW (Null hypothesis) of -0,0020. Model 1 scatter plot gives an indication that in order for the trend line to be more linear towards the more active compounds, more compounds need to be included in the high activity range. At this point, there are highly active compounds which the model is unable to predict because as it is trained to predict with the highest level of accuracy the range from -5kcal/mol to -7kcal/mol. The model was trained with these set of compounds, also having similar chemical characteristics hence they are clustered around that region alone.

The first model was then used to predict the activity of the second set of compounds.

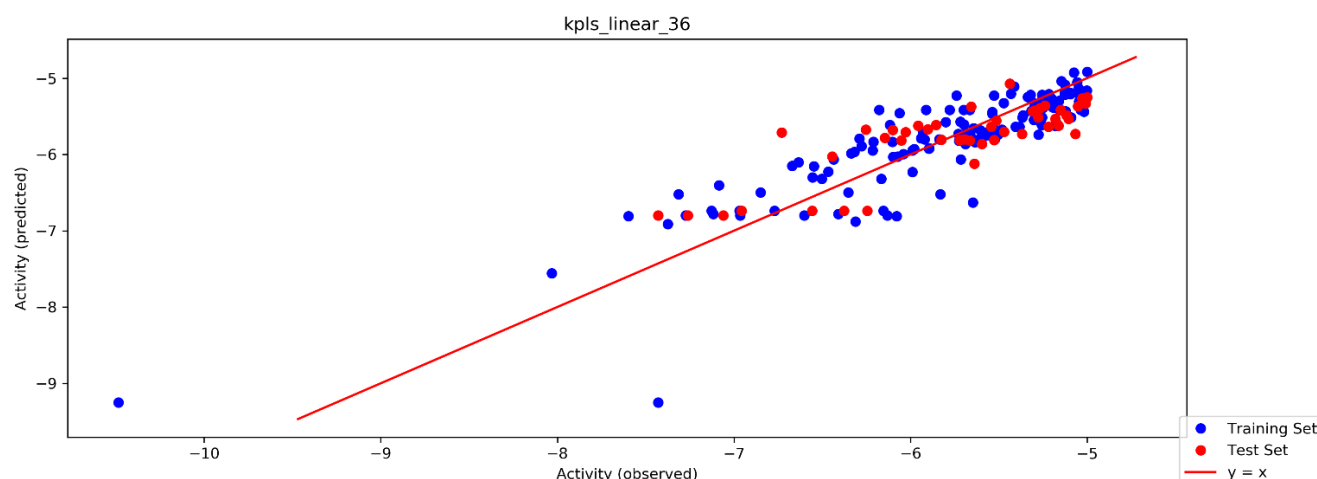


Figure 9: Auto QSAR active learning scatter plot for model 1 showing the performance of the QSAR KPLS model's predicting activity for experimental binding affinity for the test set. Model trained 123 ligands with docking scores ranging from -5.0kcal/mol to -7.6kcal/mol,. AutoQSAR randomly divided the selected ligands into the training at 75% (Blue dots) and test sets at 25% (Red dots).

A total of 2000 ligands (1000 enumerated and 1000 from the chEMBL database) were selected to make a prediction, giving the predicted activity and subjected to Glide SP ligand docking. Docking yielded 1731 compounds and the best compounds with a docking score ranging from -5.5kcal/mol to -7.3kcal/mol were selected and added to the first 123 to build model 2.

Model 2 report showed slightly lower values for the Score, S.D, R^2 , RMSE, Q^2 and Q^2 MW (Null hypothesis), this is due to the addition of more compounds to the first model having different chemical characteristics, as it can be seen from the scatter plot (figure 10).

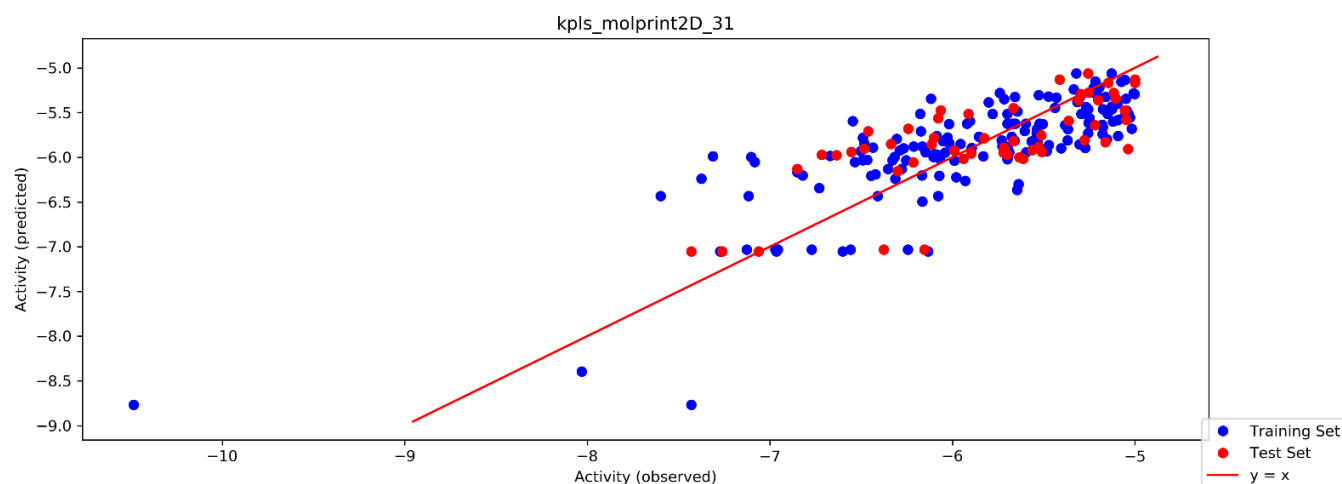
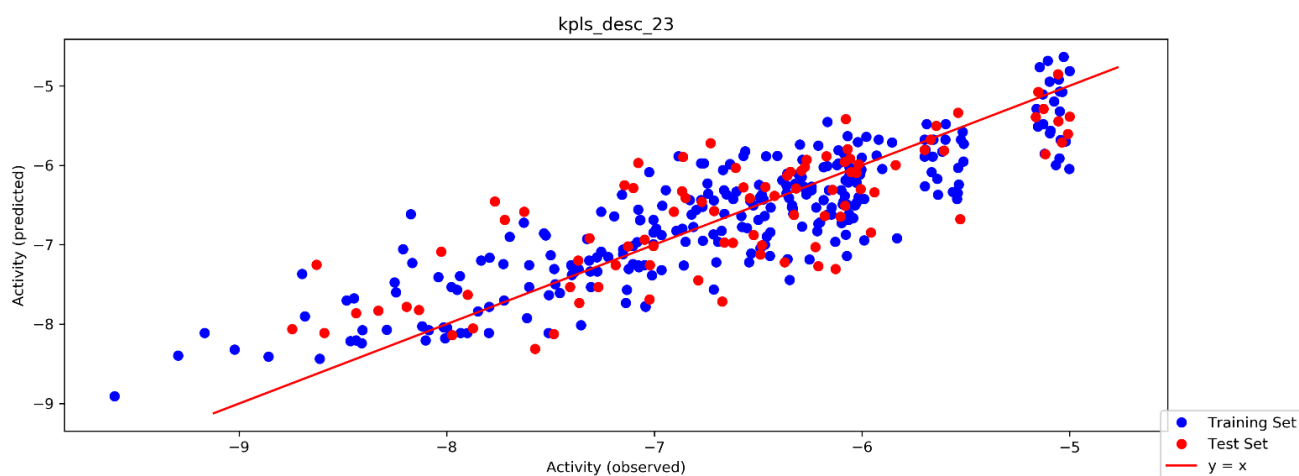


Figure 10: Auto QSAR active learning plots for model 2. Model trained 188 ligands with docking scores ranging from -5.0kcal/mol to -7.6kcal/mol,. AutoQSAR randomly divided the selected ligands into the training at 75% (Blue dots) and test sets at 25% (Red dots).

The same procedure was repeated until the model reached convergence also changing the activity data range to include compound with higher binding affinities towards PfHsp90-NTD. As the procedure is repeated more and more compounds are added getting into the high activity region. The trend line gets linear towards the observed high activity region. Model 13 (Figure 11), which is where the model attained its convergence, denotes that the model is now able to recognize compounds it could not initially, as it can be seen from the previous models (Appendix).

Figure 11: Auto QSAR active learning plots for model 13 with an R^2 of 0.7460 and a Q^2 of 0.6182. Model trained 456 ligands with docking scores ranging from -5.0kcal/mol to -9.294kcal/mol,. AutoQSAR randomly divided the selected ligands into the training at 75% (Blue dots) and test sets at 25% (Red dots).



Initially there were fewer compounds in the high activity region, which is between -8kcal/mol and -10kcal/mol, in model 13, more compounds can be observed in that region indicating that the model is more sensitive towards that region. The model was able to learn how to predict the activities of the highly active compounds and compounds in different chemical spaces.

4.1.4 Induced fit docking

The active learning model enhanced by classical GlideSP ligand docking yielded 236 best docked compounds, with the ability to bind to the receptor PfHsp90 N-terminal domain ATP binding site having docking scores of -10kcal/mol to -6kcal/mol, from a total of 43097 docked compounds. The 236 compounds were then subjected to IFD against receptor PfHsp90 N-terminal domain. The extra precision (XP) IFD resulted in 113 top-scoring compounds which were selected based on the docking score, e-model score, IFD score and a visual inspection of the binding mode. It was interesting to note that, even though autoQSAR models were only trained on compounds with docking scores between -5 and -10kcal/mol, IFD, a more robust scoring function, was able to calculate compounds with docking scores between >-5kcal/mol and >-13kcal/mol. IFD typically utilizes a more extensive sampling of ligand conformations and protein-ligand interactions

compared to GlideSP. This increased conformational sampling can improve the accuracy of predicting ligand binding modes and binding affinities (Elokely and Doerksen, 2013). Approximately 62 compounds (docking score shown in Appendix, tabel A11) were intended to be purchased from the 113, this selection was based on the compounds complying with Lipinski's rule of five which entails: molecular weight range of up to 500g/mol (Tice, 2001; Nendza and Müller, 2010). However, due to most of the compounds not being purchasable, Mcule subjected the compounds to a secondary structure similarity test and the compounds with a similarity of >90% were purchased. A total of 18 compounds were purchased, see Table 4. The activity of the purchased compounds was predicted using model 13 and further analysed with IFD. The IFD results are presented in Table 4 below.

Table 4: induced fit docking results of the purchased compounds showing Docking score, glide energy, glide emodel of Induced fit docked complexes.

Title	docking	glide emodel	XP GScore	IFDScore
FTN-T1	-13.486	-83.951	-13.486	-516.39
FTN-T10	-13.198	-103.809	-13.213	-526.65
FTN-T2	-13.098	-66.551	-13.098	-517.56
FTN-T6	-12.046	-63.464	-12.046	-519.19
FTN-M1	-11.933	-82.746	-11.933	-521.37
FTN-T4	-11.811	-89.537	-11.811	-521.42
FTN-T5	-11.750	-70.276	-11.750	-514.41
FTN-M6	-11.622	-78.526	-11.622	-521.80
FTN-M4	-11.530	-74.353	-11.566	-523.65
FTN-T9	-10.857	-60.374	-10.857	-518.39
FTN-M7	-10.749	-56.581	-10.749	-514.49
FTN-M2	-10.451	-81.977	-12.896	-522.04
FTN-M3	-10.307	-57.524	-10.307	-514.95
FTN-B1	-8.636	-73.969	-8.636	-518.12
FTN-B4	-6.363	-55.654	-6.363	-518.01
Harmine	-8.292	-41.862	-8.406	-517.44

IFD (Table 4) revealed that compounds FTN-T1, FTN-T10, FTN-T2, FTN-T6, FTN-M1, FTN-T4, FTN-T5, FTN-M6, FTN-M4, FTN-T9 and FTN-M7 (docking scores range of -13.486 to -10.749kcal/mol) exhibited better and comparable binding energies to compound 10 (reference compound) (docking score of -10.485kcal/mol). The highest scoring compound FTN-T1 exhibited better binding energies of 13.486kcal/mol a -3.00kcal/mol margin to the reference compound complementing autoQSAR model predictive abilities. High-affinity ligands typically exhibit lower docking scores (the more negative, the better the binding), indicating a more favourable ligand-protein complex formation while ligands with low binding affinities will exhibit higher docking scores, indicating weaker interactions with the target protein. In the case of ligands with low

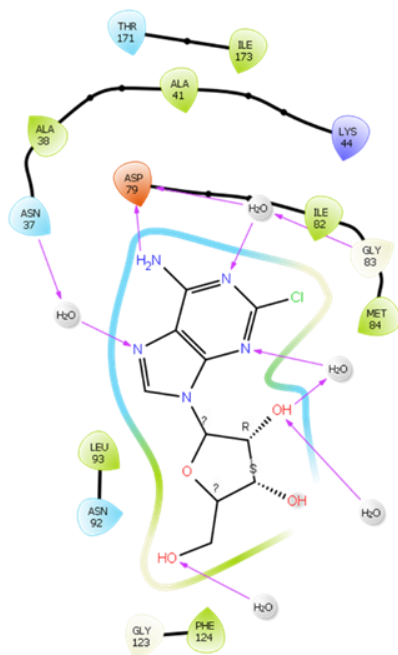
binding affinities, IFD results reveal insufficient complementarity between the ligand and the protein's active site. This lack of complementarity can lead to reduced van der Waals interactions, hydrogen bonding, or electrostatic interactions, resulting in a weaker binding affinity. Additionally, steric clashes or unfavourable conformational changes induced by the ligand binding may contribute to the lower affinity.

The 2D interaction diagrams indicate that majority of the ligands are well accommodated within the ATP binding site of the PfHsp90. From a visual inspection it can be observed that compounds in the T and M clusters share a similar scaffold and only differ in the outer functional groups this is a critical observation as the chemical structure of a compound has a significant effect on the affinity and selectivity of the compound. Compounds FTN-T1, FTN-T2, FTN-T4, FTN-T6 FTN-T9 FTN-T10 and FTN-M1 (Table 4; Figure 12) contains an 7H-Purine moiety, like ADP/ATP. The strong complementarity of these compound to PfHsp90 could be explained by due to their resemblance of its natural ligand. It is proposed that these ligands might outcompete ADP/ATP given their high docking.

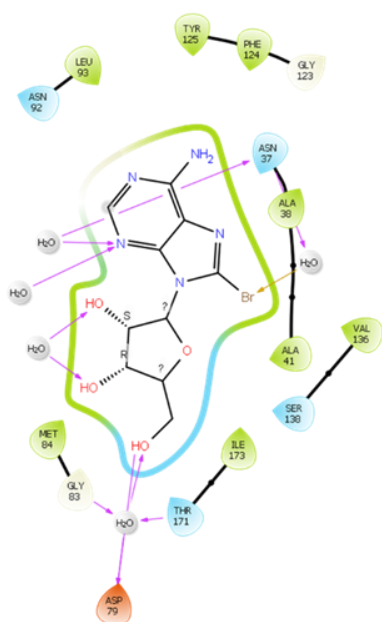
Generally, purine containing compounds, such as ATP and its analogs, are known to have a high affinity to the ATP-binding pocket Hsp90 as the purine ring system allows for the formation of multiple hydrogen bonds and hydrophobic interactions with the binding pocket residues, resulting in a higher affinity towards the protein. The results obtained in this study is consistent with that observation (Tabel 4). While FTN-T5, FTN- M4 and FTN-M6 all have a pyrimidine moiety, which likely played a role in their overall lower binding affinity towards PfHsp90. These compounds are smaller and less polar, limiting their ability to form strong hydrogen bonds and hydrophobic interactions with the binding pocket (Taldone *et al.*, 2015).

All the ligands represented in Figure 12 below make prominent interactions with ATP binding residues such as ASN37, ASP79, GLY83, ASN92, LYS44, PHE124, and ALA38 (Corbet and Burger, 2010). It can also be seen that ALA38 is donating a water mediated hydrogen bond to compounds FTN-T10 (Figure 12- FTN-T10). FTN1 and FTN10 appear to be analogues of each other. Therefore, binding similar ATP binding residues Notably, ASN37, ASP79, GLY83, ASN92 contributing to hydrogen bonds, with a difference only with the outer functional group. ARG 98 is found in the binding interface, However, FNT-T10 is interacting with ALA38 the *P. falciparum* specific pocket It is being predicted that *in vitro* activity will show potent inhibition and selectivity

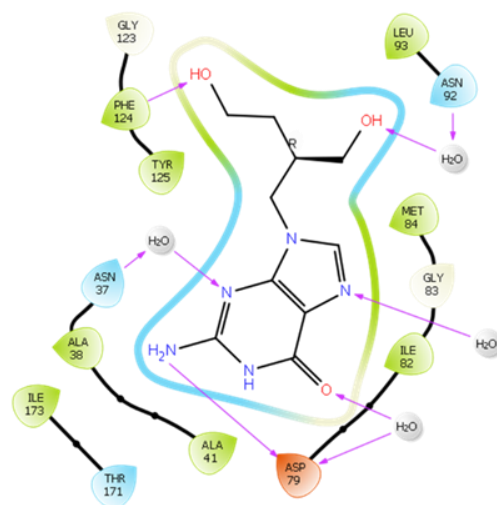
FTN-T1



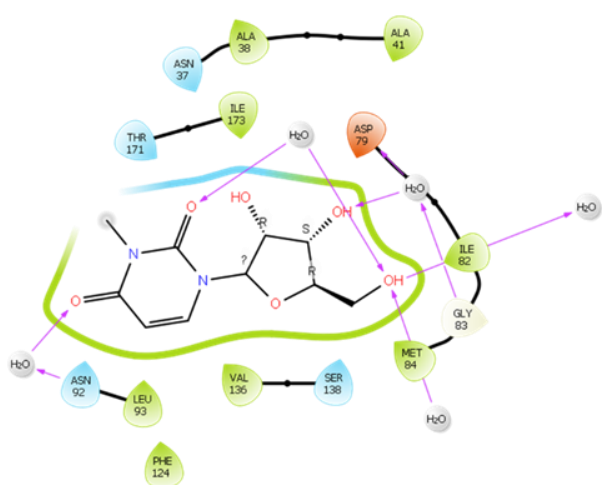
FTN-T2



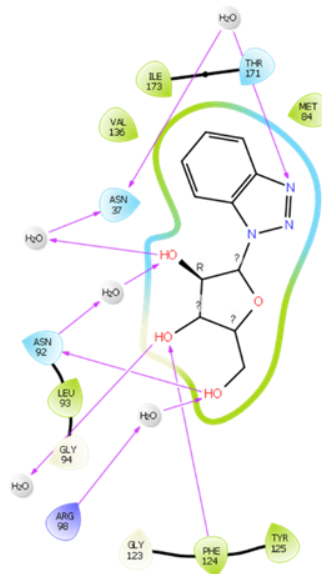
FTN-T4



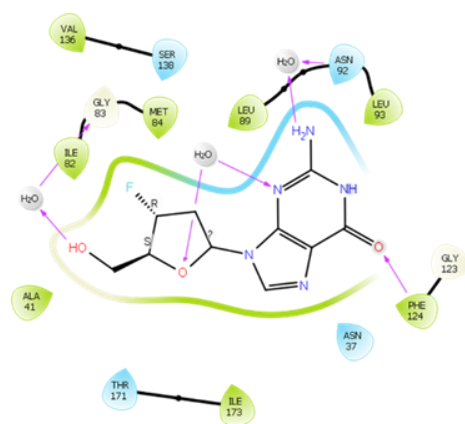
FTN-T5



FTN-T6



FTN-T9



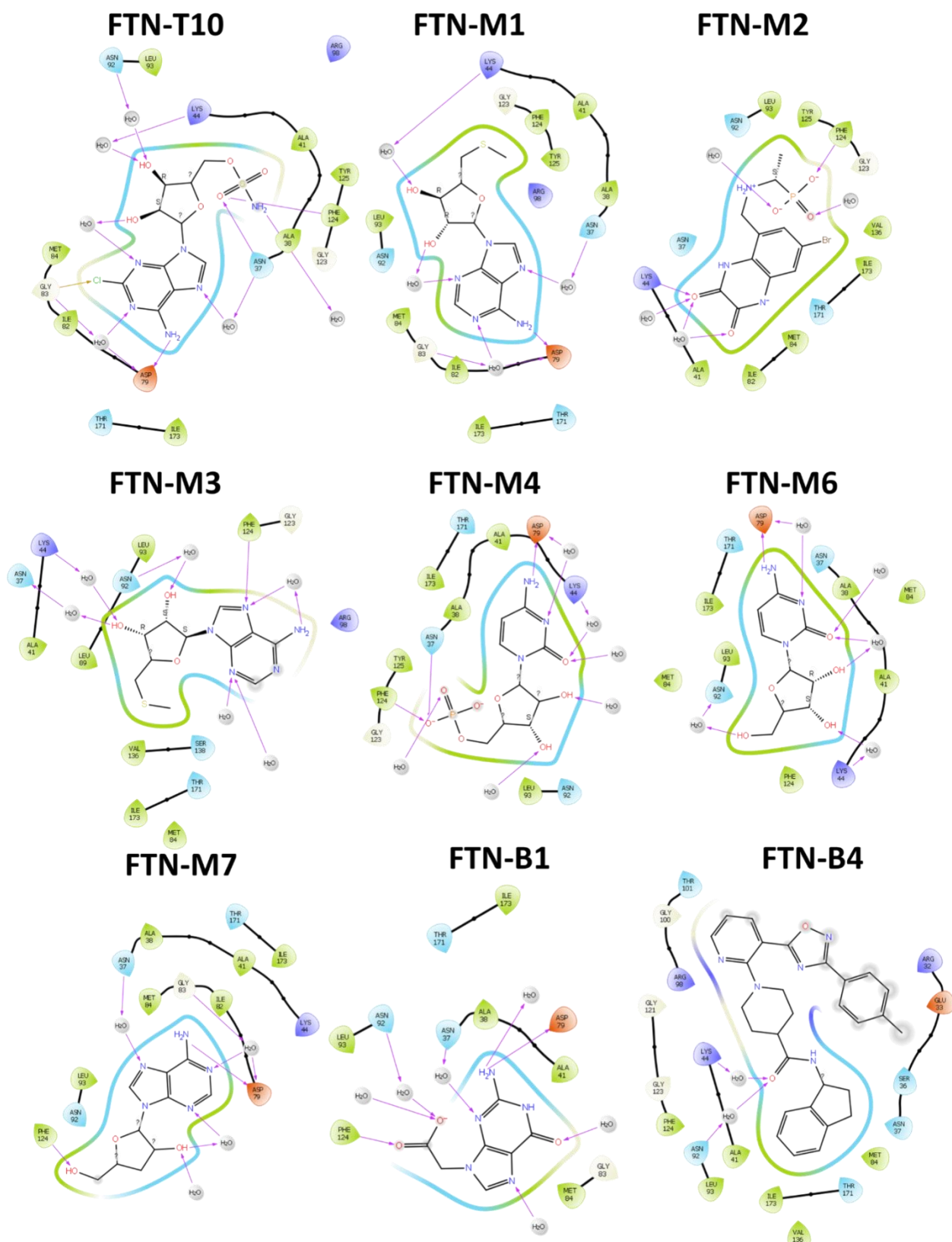


Figure 12: PfHsp90 (PDBcode:3K60) interaction with 15 selected compounds. The hydrogen bonds and pi-cation are indicated in purple and red lines, respectively. Annotations of amino acid residues showing respective polar, non-polar, negatively charged and positively charged cyan, green, orange and blue. A represents compound FTN-T1, B FTN-T10, C FTN-M4, D FTN-M6, E FTN-B1 and lastly F FTN-B4

for compound which appear to be interacting with both ATP and *P. falciparum* specific pockets residues. FTN-B4 derived from the enumerated compounds, extends to the outside of the binding pocket and making two hydrogen bonds with LYS44 and ASN92. Therefore, expected to be unstable within the binding pocket and display very weak binding from the invitro assay. Through the observation of these IFD results residues involved in binding interactions were identified, providing valuable information for further analysis and optimization of ligands with improved affinity and specificity.

4.1.5 ADME/T property predictions

The physio-chemical property predictions of the compounds (Figure 13) were carried out using Qikprop module to assess the drug likeness of the compounds based on Lipinski's rule of 5 such as, ($\text{mol_MW} < 500$, $\text{QPlogPo/w} < 5$, $\text{donorHB} \leq 5$, $\text{acceptorHB} \leq 10$). Compounds that satisfy these rules are considered drug-like compounds. (The "five" refers to the limits, which are multiples of 5), molecular weight of the compound, logP, logS, logD, hydrogen bond acceptors, Hydrogen bond donors and Topological polar surface area ($\text{\AA}^2$) (Table 5).

The majority of the compounds comply with the Lipinski rule of five, except for FTN-B2, FTN-M6, and FTN-B4 in violation of 1 rule. No further violations were observed for the number of hydrogen bonds that would be accepted or donated by the solute from water molecules in an aqueous solution (HBA and HBD). This is important because hydrogen-bond interactions play a crucial role in biological recognition processes, specifically in the interaction between a drug molecule and its surrounding environment (Bissantz *et al.*, 2010). Potency (Williams, and Ladbury, 2003), permeability (Toulmin *et al.*, 2008), and solubility (Stanton *et al.*, 2009), among others are therefore greatly impacted by hydrogen bonding characteristics of a drug molecule.

Table 5: *in silico* physio-chemical property predictions by Qikprop

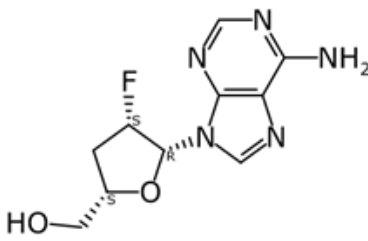
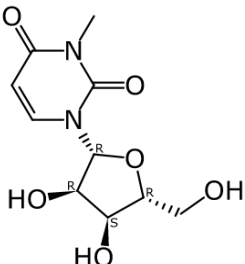
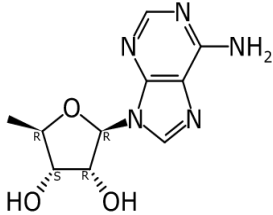
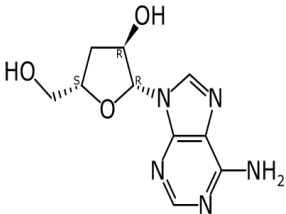
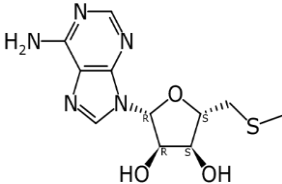
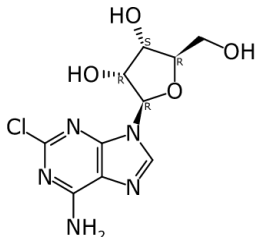
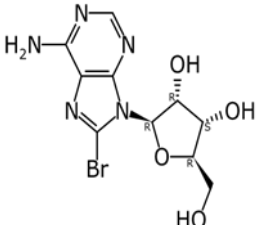
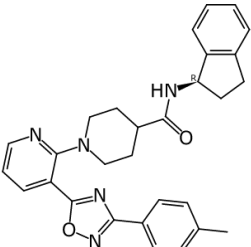
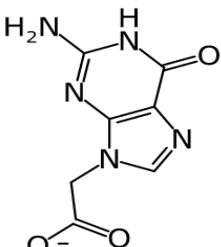
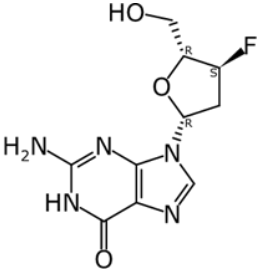
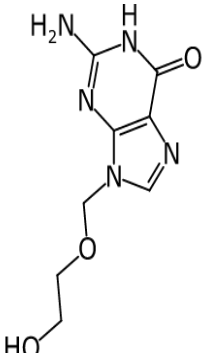
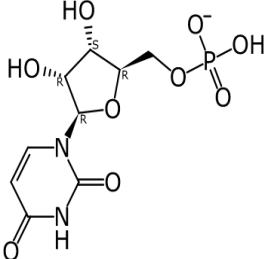
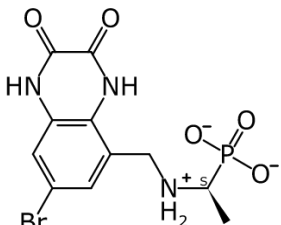
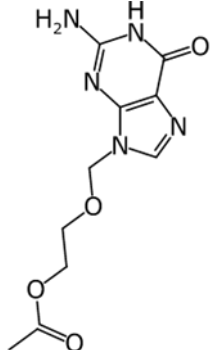
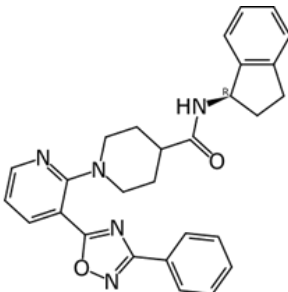
Title	MW (g/mol)	log P	HBD	HBA	Lipinski violations	log D	log S	TPSA (Å ²)
FTN-M10	253,237	-0,39	3	7	0	-0,39	-1,12	99,08
FTN-T5	258,23	-1,61	3	8	0	-1,78	-0,24	113,92
FTN-T8	251,246	-0,66	4	8	0	-0,66	-0,91	119,31
FTN-M7	251,246	-0,79	4	8	0	-0,79	-0,81	119,31
FTN-M3	297,34	0,25	4	8	0	0,25	-1,88	119,31
FTN-T1	301,69	-0,41	5	9	0	-0,5	-1,43	139,54
FTN-T2	346,141	-0,17	5	9	0	-0,71	-1,9	139,54
FTN-B4	479,584	6,28	1	7	1	5,95	-7,55	84,15
FTN-B1	209,165	-1,63	4	8	0	-2,34	0,1	126,89
FTN-T9	269,236	-0,69	4	8	0	-0,69	-1	119,05
FTN-B7	225,208	-1,44	4	8	0	-1,44	-0,15	119,05
FTN-M6	324,182	-1,56	5	11	1	-3,12	-0,71	171,31
FTN-M2	378,119	-1,76	5	8	0	-2,45	-0,92	135,28
FTN-B6	267,245	-0,98	3	9	0	-0,98	-0,76	125,12
FTN-B2	465,557	5,88	1	7	1	5,46	-7,16	84,15
FTN-T4	253,262	-0,9	5	8	0	-0,9	-0,73	130,05

^aTable 5 shows the *in silico* physicochemical property predictions by Qikprop, where MW is Molecular weight of the molecule. logP is Predicted octanol/water partition coefficient.(−2.0 – 6.5). LogS is Predicted aqueous solubility (−6.5 – 0.5). logD is Predicted the octanol/water distribution coefficient.(−2.0 – 3) HBD estimates number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution (0.0 – 6.0). HBA Estimates number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution (2.0 – 20.0). Lipinski violations are Number of violations of Lipinski's rule of five . The rules are: mol_MW < 500, QPlogPo/w < 5, donorHB ≤ 5, acpptHB ≤ 10. Compounds that satisfy these rules are considered druglike. (The “five” refers to the limits, which are multiples of 5.) Topological polar surface area (Å²) is the Van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms (7.0 – 200.0)

Physicochemical property predictions are essential in evaluating the drug-likeness and potential risks associated with candidate compounds by giving an insight into factors such as lipophilicity, solubility, permeability, and distribution of the compounds (Leeson *et al.*, 2021). The intention therefore was to assess whether the compounds generated in this study possessed a good safety profile. Absorption and distribution of a drug within the human body is highly influenced by the binding to plasma proteins and its ability to cross cell membranes. Therefore, physicochemical property predictions such as logS, logP and logD defines if the drug will be able to be taken up by the cells. Lipophilicity is represented by the descriptors such as logP and logD, and is used, for example, to help predict *in-vivo* permeability of active compounds (Bhal, 2019).

The results obtained indicated Table 5 no violations of these properties were observed in all compounds, which is in accordance with guidelines established for absorption for human use, except for compounds FTN-B4 with a logD of 5.95 and logS of -7.55 while FTN-B2 logD of 5.46 and logS of -7.61, which is expected as the intermolecular forces of the compound were observed to be weak in the IFD results. Both intramolecular forces in the solid state and intermolecular interactions between the compound and the surrounding intestinal fluid affect the compound's solubility in water. The solubility will therefore be low if the drug molecules prefer their solid state to dissolving into the aqueous fluid because it is more energetically advantageous for them to attach to each other than to the water molecules. Poor solubility, nevertheless, may also be brought about by unfavourable interactions between the molecules of water and compounds (Norinder and Bergström, 2006). Several studies have shown that logD or logP values greater than 3 lead to promiscuity and an increased likelihood of severe *in vivo* toxicological consequences (Leeson and Springthorpe, 2007; Hughes *et al.*, 2008; Waring, 2009). Therefore, with LogP values ranging from -1.76 to 6.28, LogD values from -3.12 to 0.25, logS values from -1 to 1, the generated compounds appear possess a good safety profile. Although *in vitro* assessment needs to be undertaken to validate our compounds as drug candidates.

The molecular descriptor topological polar surface area ($\text{\AA}^2$) giving an insight into drug transport properties such as intestinal absorption by assessing Van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms (Li *et al.*, 2005) no violations were observed, where the generated compounds exhibited PSA in the range of 84.15 – 171.31 suggesting that the compounds are highly polar and can easily travel through membranes. By considering these physicochemical properties, although some of the compounds did not meet all the physiological property parameters at this stage, the study was prioritizing activity. With the idea in mind that informed decisions regarding the future development and optimization of the compounds by structure activity relationship to improve the affinity, solubility, lipophilicity, and other pharmacokinetic parameters, ultimately improving their drug-like properties and increasing the likelihood of successful translation into effective and safe antimalarial therapeutics. Most of the compound a very small in size providing room for further optimization.

FTN-M10 	FTN-T5 	FTN-T8 
FTN-M7 	FTN-M3 	FTN-T1 
FTN-T2 	FTN-B4 	FTN-B1 
FTN-T9 	FTN-B7 	FTN-M6 
FTN-M2 	FTN-B6 	FTN-B2 

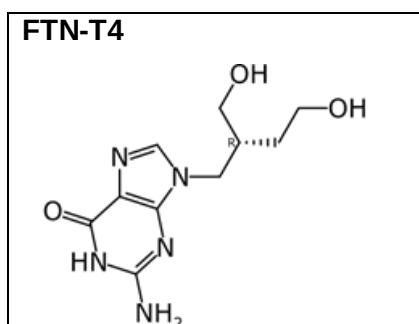


Figure 13: Compound structures analysed by Qikprop for the *in silico* physio-chemical property predictions.

4.1.6 Molecular dynamics simulations

The coupling of IFD with the more precise molecular dynamic simulation techniques, IFD findings can be greatly enhanced (Santos *et al.*, 2019).. Therefore, MDS was used as a post IFD processing technique to evaluate the conformational stability and time-dependent binding capacity of ligands in the PfHsp90-NTD ATP binding pocket, using the best-docked poses obtained from the IFD results. The resulting MDS trajectory was utilized to assess the dynamics' equilibrium as a function of time. The root mean square deviations (RMSD) measures the average difference in displacement between a group of atoms for a certain frame and a reference frame. Metric of the starting structure and the average simulated structure of all MDS trajectory frames are used to gain insight into the convergence of simulated protein-ligand complexes. It is calculated for all frames in the trajectory. The RMSD for frame x is:

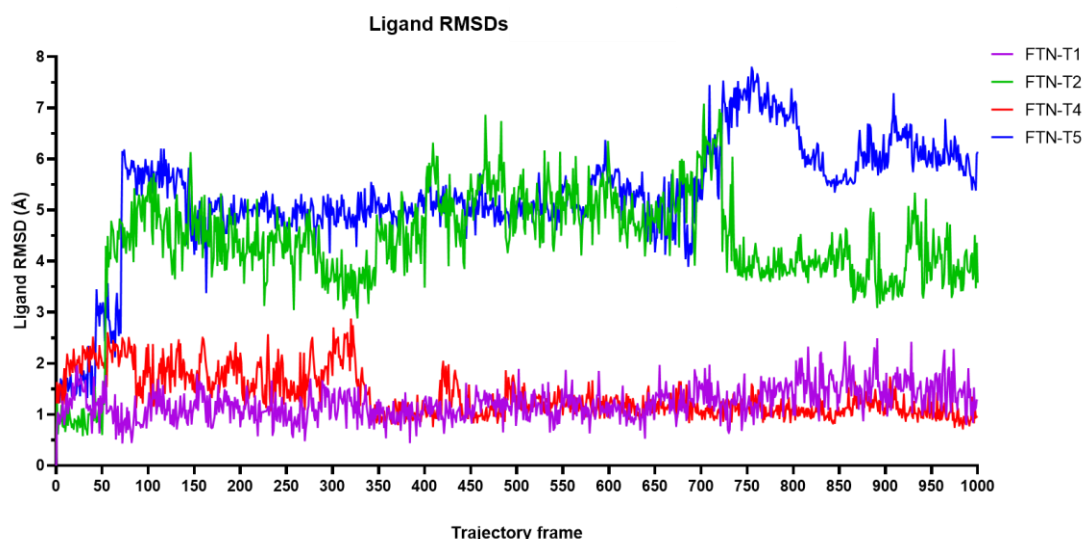
$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i^t(t_x) - r_i(t_{ref}))^2} \quad (7)$$

Where N is the number of atoms in the atom selection; t_{ref} is the reference time, (typically the first frame is used as the reference and it is regarded as time $t=0$); and r' is the position of the selected atoms in frame x after superimposing on the reference frame, where frame x is recorded at time t_x . The procedure is repeated for every frame in the simulation trajectory.

The Ligand RMSD plots (Figure 14) for all compounds shows stabilization throughout the captured trajectory frame (with ligand RMSD values ranging from 0.50 to 6Å) (Figure 14) with minor fluctuations observed from the ligand at the beginning of the simulation, due to the conformational changes of the ligand adjusting itself into the binding site, then reaching equilibration from 300 frames until the end of the simulation. This is with the exception of FTN-T2 and FTN-T5 which show signs of binding instability when evaluating their ligand RMSD over simulation trajectory frame (Figure 14), which shows fluctuates until the end of the simulation

window, suggesting that the ligands adopted incorrect binding modes and are undergoing conformational changes throughout the simulation window. The observed RMSD values of these two ligands are significantly larger than the RMSD of the protein (Figure 15), indicating that the ligand has diffused away from its initial binding site. Typically, ligands that require significant binding adjustments from their initial docked conformation are more likely to dissociate from the binding site during the course of the interaction (Guterres and Im, 2020).

A



B

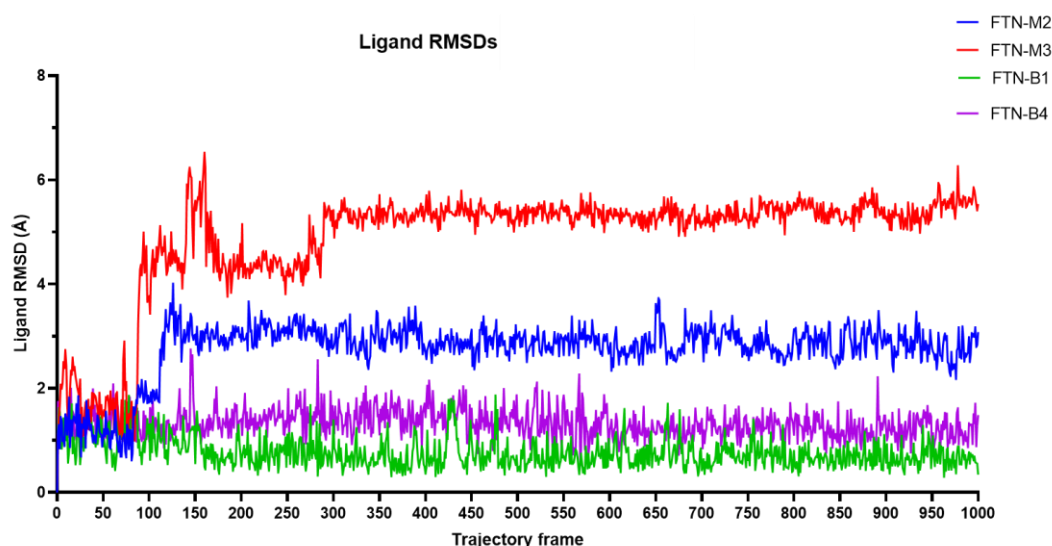
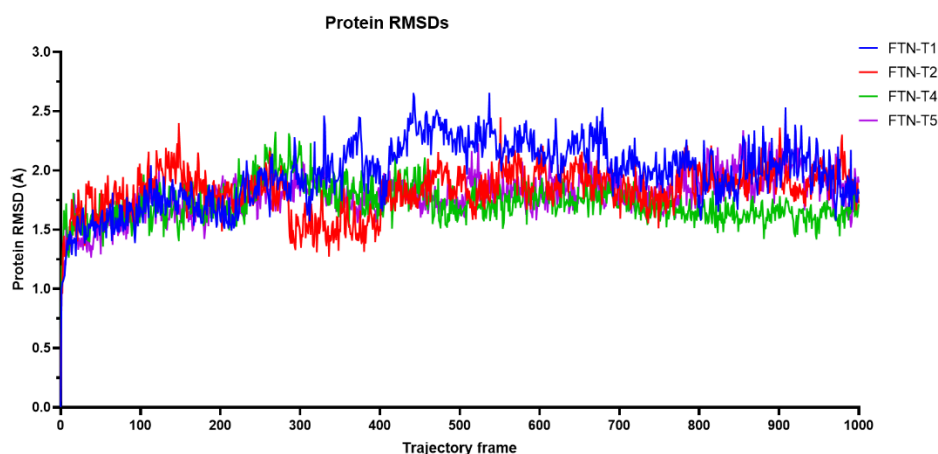


Figure 14: MDS results of the stability assessment of bound ligands using Ligand RMSD plots of all the frames in the trajectory for PfHsp90-NTD and Compounds FTN-T1, FTN-T4, FTN-T5, FTN-M2, FTN-M3, FTN-B1 and FTN-B4 complexes’.

The protein RMSD plots (Figure 15) provides valuable insights into its structural conformation throughout the simulation. By aligning all frames on the reference backbone and calculating the RMSD based on atom selection, stability of the protein's structure can be monitored. A well-

equilibrated simulation is characterized by RMSD fluctuations around a thermal average structure towards the end of the simulation, typically in the range of 1-3 Å for small, globular proteins. The PfHsp90 protein is observed to be stable in its interaction with all the tested ligands, averaging RMSD values of 2Å for all complexes (Figure 15). Larger changes in RMSD suggest that the protein is undergoing significant conformational changes during the simulation, which is not the case for all the simulated complexes. This is considered a reliable analysis as simulations are observed to have equilibrated and reached convergence, meaning that the RMSD values stabilize around a fixed value 1.7Å. If the RMSD is still increasing or decreasing on average at the end of the simulation, it indicates that the system has not fully equilibrated. The RMSD analysis serves as a key indicator of the simulation's quality, providing essential information about the protein-ligand dynamics and stability.

A



B

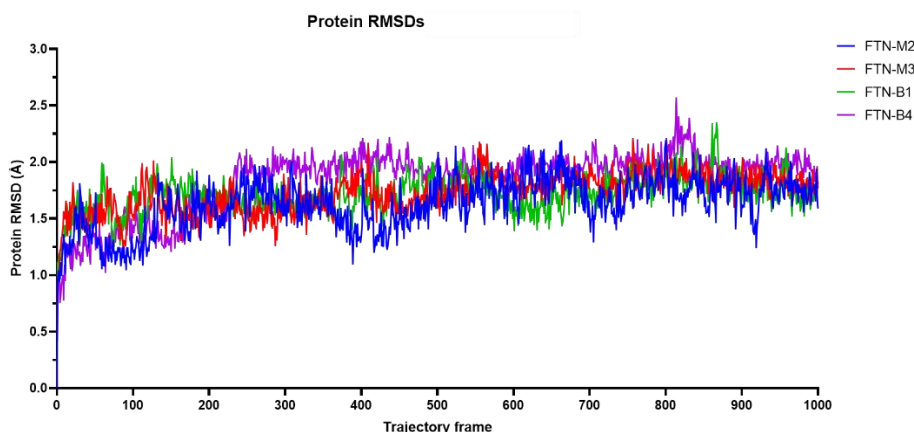


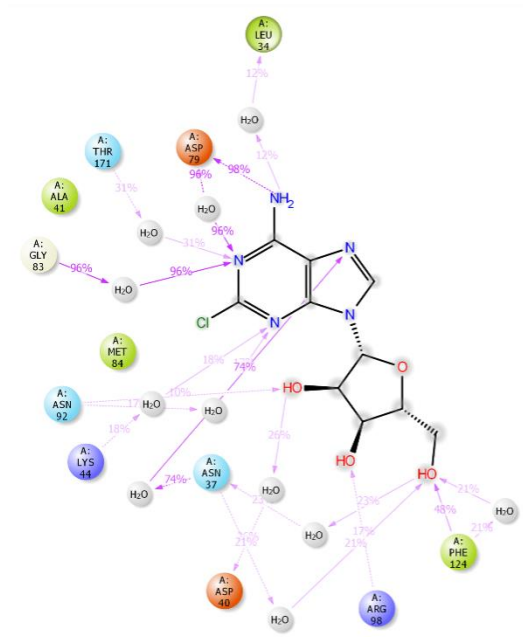
Figure 15: MDS protein RMSD plots of all the frames in the trajectory for PfHsp90-NTD and Compounds FTN-T1,FTN-T4, FTN-T5, FTN-M2, FTN-M3, FTN-B1 and FTN-B4 complexes'.

FTN-T1 and FTN-T4 are observed to be the most stable compounds in the PfHsp90 binding pocket. This exceptional stability of compound FTN-T1 and FTN-T4 is most likely explained by the direct strong hydrogen bond with the NH2 group with ASP79 remaining stable for 99% and

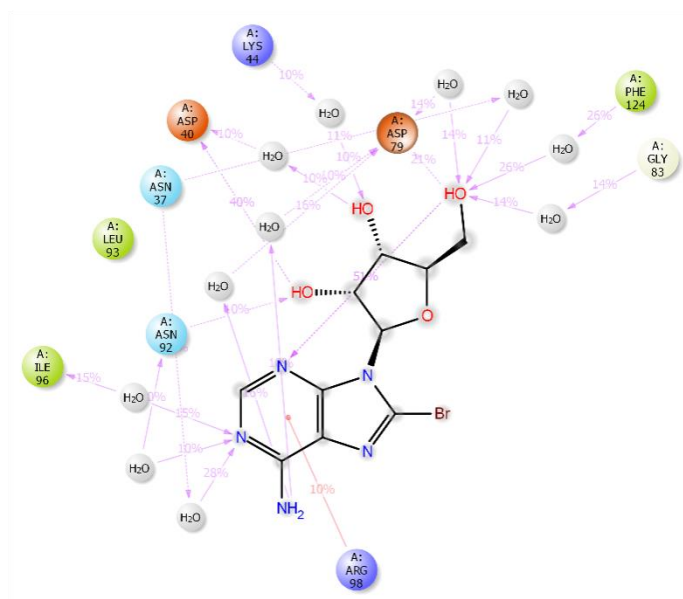
92% respectively (Figure 16), of the simulation time. The amino acid PHE124 hydrogen bonds with the OH group, which remained stable for 84% and 60%, the whole simulation duration, which is not observed in other complex simulation of the T cluster.

Generally, when compared to T cluster molecules, the protein-ligand RMSD plots for the M cluster compounds (FTN-M2 and FTN-M3) indicate stability throughout the course of simulation periods at somewhat higher RMSD values for the ligand (Figure 14 : FTN-M2 in blue and FTN-M3 in red). FTN-M2 and FTN-M3 (Figure 14) RMSD values indicated that the ligand-protein complexes were stable enough, with FTN-M2 ligand RMSD value of 3.0Å and FTN-M3 ligand RMSD value of 5.6 Å. The stability of compound FTN-M2 is most likely given by the direct strong hydrogen bond with the ARG98 at >150% of the simulation time (Figure 16), which also contributes to specificity binding. The value over 100% is possible the protein residue making multiple contacts of the same subtype with the ligand. The hydrogen bond of the NH₂⁺ group with ASP37 which remained stable for 84% of the simulation time seems to also be contributing greatly to the stability of the ligand within the pocket (Figure 16). With regards to the FTN-M3 PfHsp90-NTD complex, a pi-cation interaction is made with ARG98 at 13% of the simulation time (Figure 16). The S group doesn't seem to be doing much to contribute to the stability of the ligand as was expected, as it is the main difference between these compounds this can also be seen with the Br group on FTN-M2's stability (figure 16)

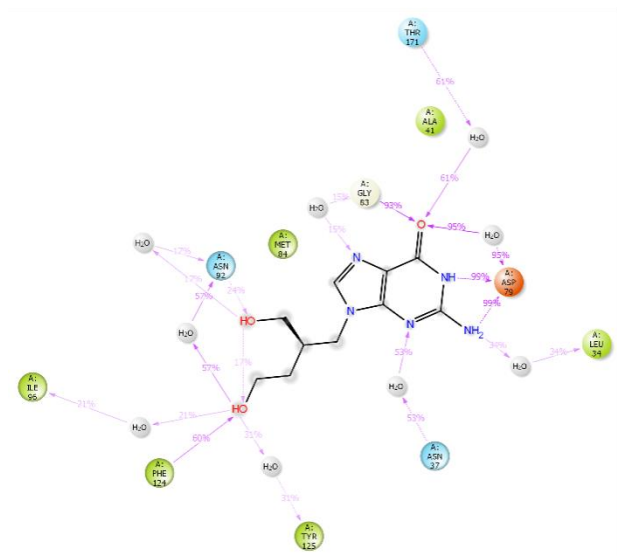
FTN-T1



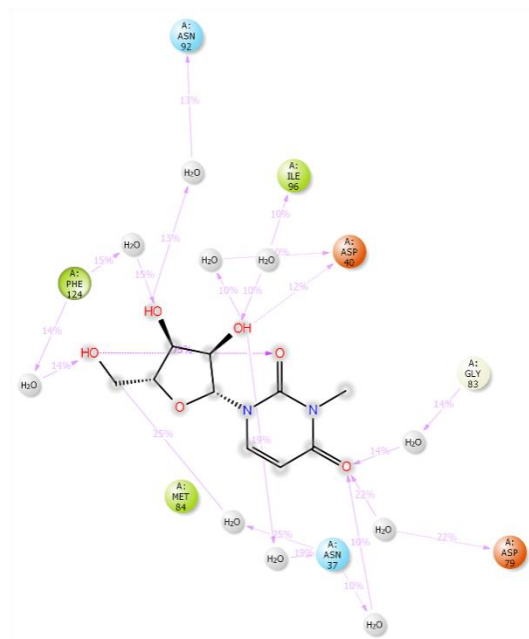
FTN-T2



FTN-T4



FTN-T5



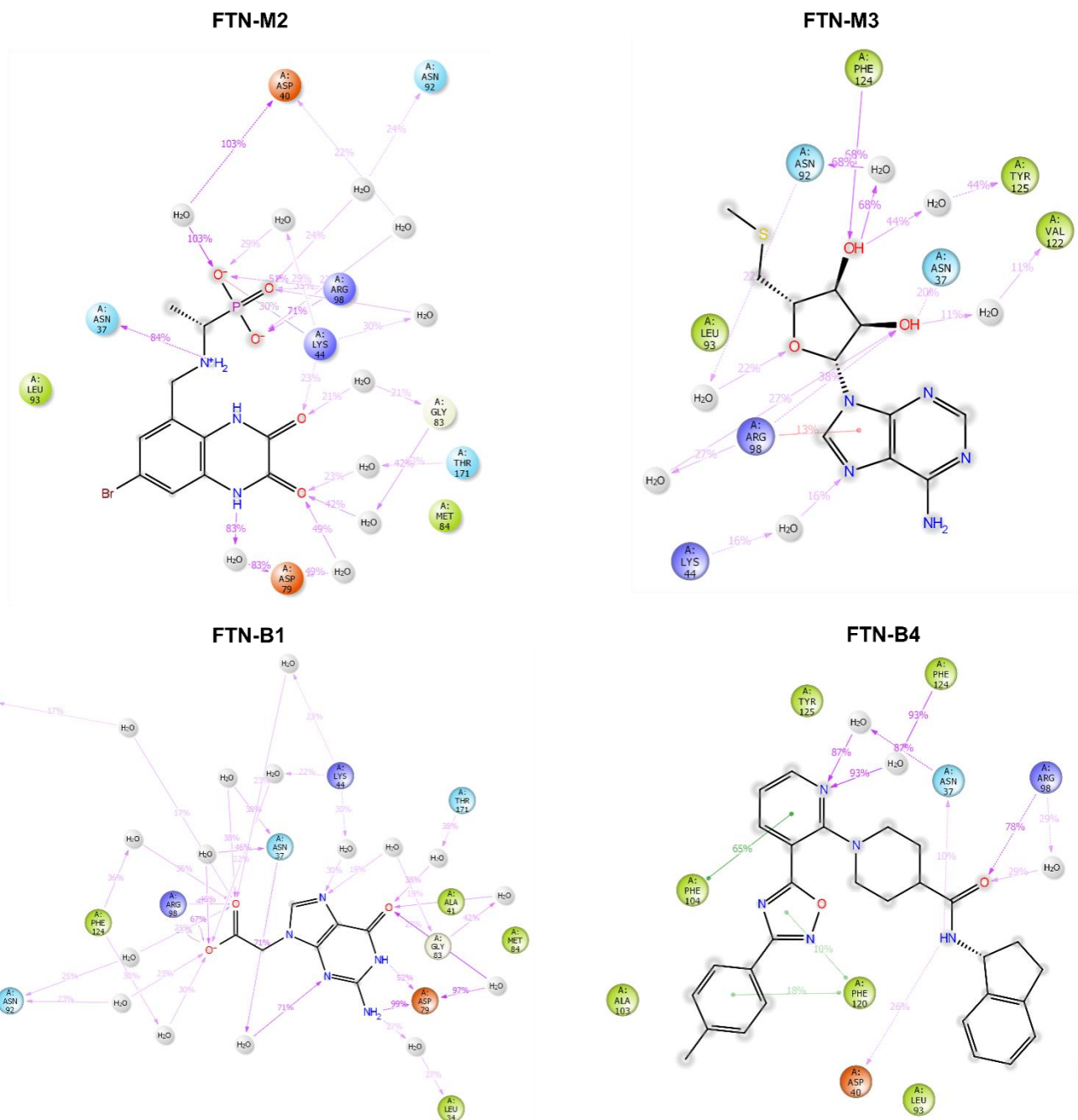
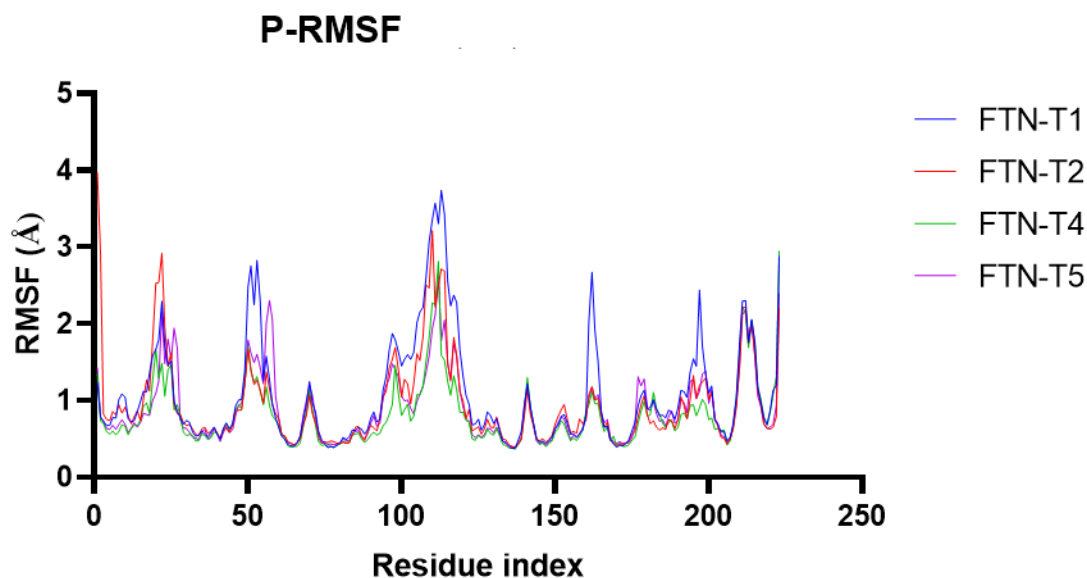


Figure 16: 2D Protein-ligand interaction diagrams of all the frames in the trajectory for PfHsp90-NTD and Compounds FTN-T1, FTN-T4, FTN-T5, FTN-M2, FTN-M3, FTN-B1 and FTN-B4 complexes at a 50ns MDS trajectory and FTN-T2 at 150ns.

The Root Mean Square Fluctuation (RMSF) is valuable in assessing local variations along the protein chain during the simulation. The RMSF plots for all the simulated protein ligand complexes show that residue index window at positions 50-70 and 100-125 had the highest RMSF protein fluctuations, this is due to the loops in that region of the protein, as they are more unstructured therefore less rigid leading to fluctuations.

A



B

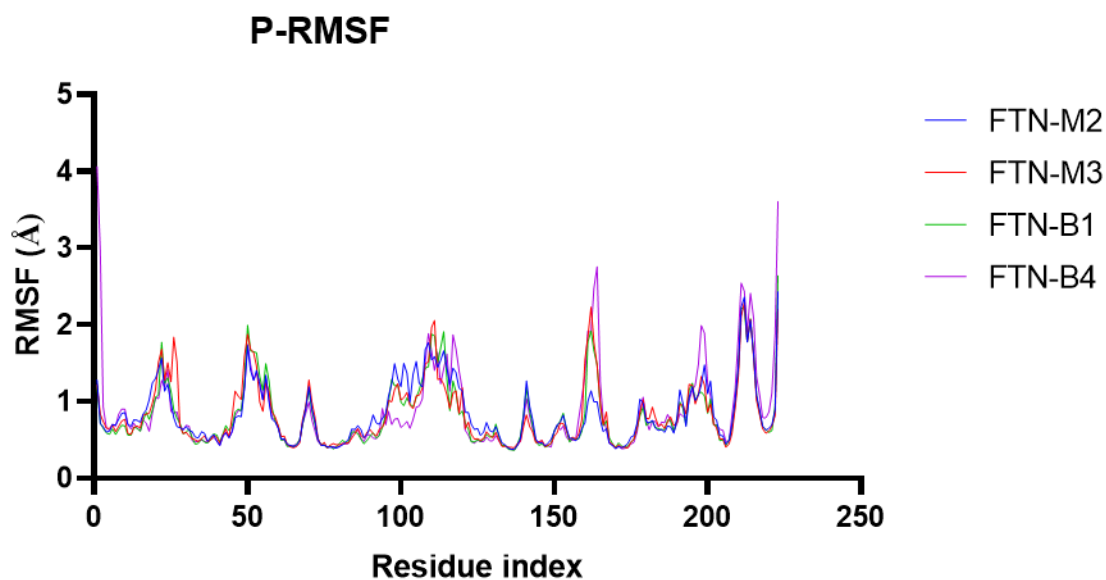


Figure 17: MDS protein RMSF plots for PfHsp90 interaction with FTN-T1, FTN-T2, FTN-T4, FTN-T5, FTN-M2, FTN-M3, FTN-B1 and FTN-B4.

4.1.7 Binding free energy calculations

The ranking of prospective leads is enhanced by the binding free-energy calculations obtained using the MM-GBSA approach, which are more accurate and reliable than the Glide XP docking (Lyne *et al.*, 2006; Das *et al.*, 2009; Yang *et al.*, 2009). Furthermore, MMGBSA serves as a refinement and validation step following the docking procedure by identifying potential errors or inaccuracies in the docking predictions (Poli *et al.*, 2020).

Table 6: Binding free energy calculation for investigated compounds using Prime MM-GBSA

Title	ΔG_{Bind}	ΔG_{Bind} Coulomb	ΔG_{Bind} Covalent	ΔG_{Bind} Hbond	ΔG_{Bind} Lipo	ΔG_{Bind} Packing	ΔG_{Bind} Solv GB	ΔG_{Bind} vdW
FTN-T10	-78.77	-57.40	2.61	-8.56	-7.06	-1.63	36.74	-43.47
FTN-T1	-67.31	-40.70	2.48	-6.28	-6.80	-1.67	25.98	-40.31
FTN-M1	-63.09	-39.63	1.95	-5.67	-5.94	-1.60	29.50	-41.71
FTN-M3	-63.09	-39.63	1.95	-5.67	-5.94	-1.60	29.50	-41.71
FTN-T2	-58.66	-40.74	2.88	-6.32	-5.05	-1.39	27.27	-35.31
FTN-M7	-57.56	-36.59	2.02	-5.60	-5.64	-1.56	26.10	-36.29
FTN-M4	-44.05	17.50	8.77	-7.82	-6.31	-0.47	-22.28	-33.43
FTN-T4	-41.60	-35.74	5.60	-5.37	-5.99	-1.40	25.52	-24.21
FTN-M6	-40.90	-29.52	3.11	-4.61	-6.14	-0.52	26.52	-29.74
FTN-T6	-35.61	-25.42	3.75	-3.73	-6.63	-1.63	29.80	-31.76
FTN-T5	-34.72	-12.96	1.94	-3.33	-8.04	-1.44	19.60	-30.48
FTN-T9	-25.08	-31.65	6.60	-3.25	-8.47	-3.02	38.98	-24.27
FTN-B1	-18.17	39.81	2.26	-3.74	-3.22	-1.51	-24.23	-27.53
FTN-M6	-16.91	13.67	5.00	-3.86	-5.58	-3.03	3.76	-26.88
FTN-B4	-14.71	-11.45	1.86	-3.29	-3.18	-3.35	31.64	-26.96
Comp 10	-28.47	-16.36	4.24	-3.67	-7.57	-4.04	33.01	-34.08
Harmine	-27.52	-3.23	1.65	-1.65	-11.23	-2.70	18.39	-28.76

^a ΔG_{Bind} Coulomb: Coulomb energy, ΔG_{Bind} Covalent: Covalent binding, ΔG_{Bind} Hbond: hydrogen bonding, ΔG_{Bind} Lipo: the lipophilic binding binding, ΔG_{Bind} Packing: π - π packing interactions, and ΔG_{Bind} vdW: the van der waals energy (vdW).

The results of the binding free energy calculations for the compounds are shown in Table 6 above. All the tested compounds outperformed the reference compound (Comp 10) (-28.47kcal/mol) and Harmine (-27.52kcal/mol) the known inhibitor of PfHsp90, suggesting stronger binding affinity towards PfHsp90, except for four compounds, FTN-T9 (-25.08kcal/mol), FTN-B1 (-18.17kcal/mol), FTN-M6 (-16.91kcal/mol) and FTN-B4 (-14.71kcal/mol). The more negative values indicate stronger binding (Zhang *et al.*, 2017). The compound that showed the highest

binding affinity free energy calculated is compound FTN-T10 at -78.77kcal/mol with PfHsp90, followed by FTN-T1 at -67.31kcal/mol. The least favourable free binding energy calculation was observed in compound FTN-B4 at -14.71 kcal/mol. The major energy contributors were identified as coulombs interactions ($\Delta G_{\text{Bind Coulomb}}$), covalent ($\Delta G_{\text{Bind Covalent}}$), hydrogen bonding ($\Delta G_{\text{Bind Hbond}}$) packing ($\Delta G_{\text{Bind Packing}}$), lipophilic energy ($\Delta G_{\text{Bind Lipo}}$), and the van der waals $\Delta G_{\text{Bind vdW}}$, enhancing the binding affinity of the compounds towards PfHsp90's ATP binding pocket.

4.2 Chapter summery

This chapter aimed to address the challenge of designing selective inhibitors for PfHsp90 while avoiding interference with human Hsp90, crucial for normal cellular processes due to their 69% sequence similarity. Leveraging the promising outcomes of Wang *et al.* (2016) and Everson *et al.* (2021), compound 10 from the latter study was identified as having superior binding affinity toward PfHsp90. Employing reaction-based enumeration based on compound 10, 10 000 design ideas were generated, and an active learning approach combining GlideSP docking and AutoQSAR modeling was applied, yielding 236 compounds for IFD analysis. IFD identified 62 compounds with strong binding potential, and 18 were purchased for further investigation. Notably, compounds FTN-T1, FTN-T10, and others exhibited comparable or improved binding energies compared to compound 10. Physicochemical property predictions revealed drug-likeness in most compounds, although solubility issues were noted for a few. Molecular dynamics simulations confirmed the stability of several compounds within the PfHsp90 binding site, notably FTN-T1 and FTN-T4, showcasing strong hydrogen bonding interactions. MM-GBSA binding free energy calculations further validated the high binding affinity of most compounds towards PfHsp90, with FTN-T10 showing the strongest binding. These findings emphasize the potential of these compounds as selective PfHsp90 inhibitors, warranting further in vitro validation and optimization for antimalarial therapeutics

CHAPTER 5

5.1 BIOCHEMISTRY RESULTS AND DISCUSSIONS

5.1.1 Confirming the integrity of the pET28b (+)-PfHsp90-NTD construct

The integrity of the pET28b (+)-PfHsp90-NTD plasmid was confirmed by restriction enzyme digest reactions. Agarose gel electrophoresis showed two fragments confirming the double digest using both *Bam*HI and *Nde*I yielded bands corresponding to the size of pET28b (+) plasmid vector at ~6000bp and PfHsp90-NTD insert just under 1000bp as the original size of the insert is 620bp (Figure 18B).

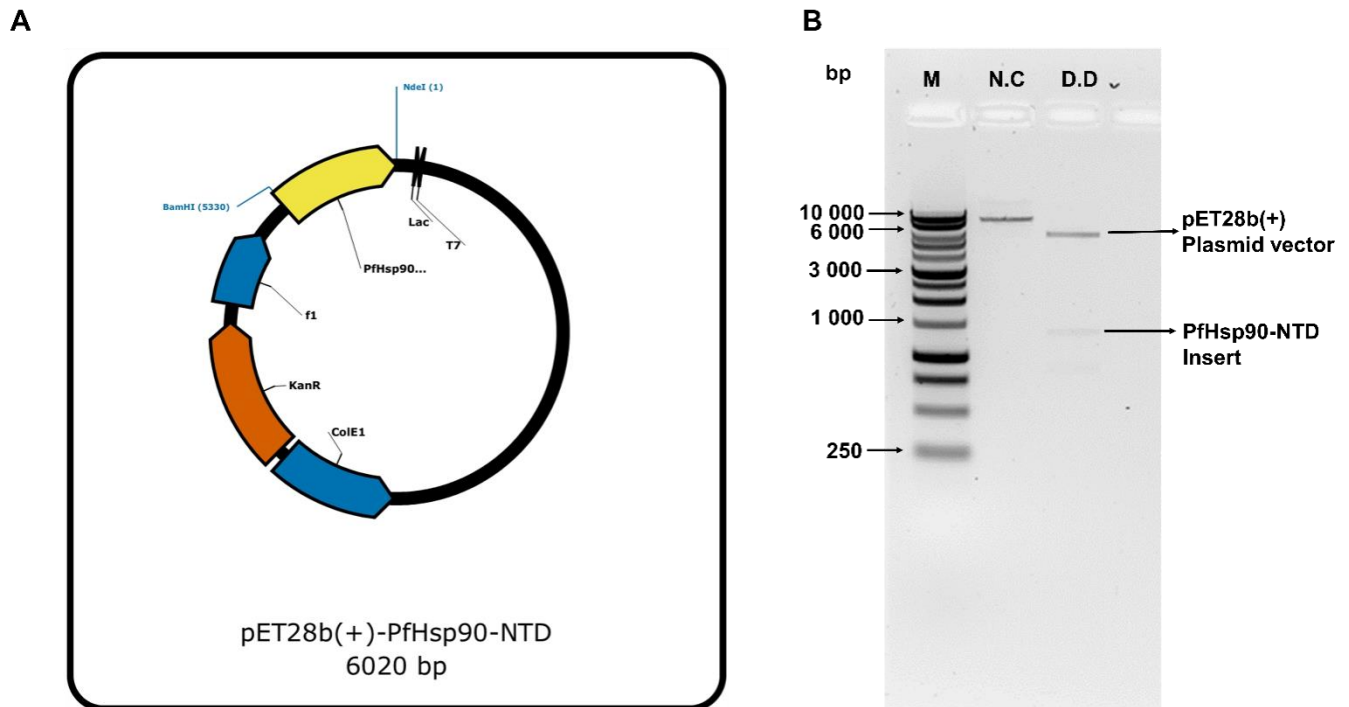


Figure 18: Restriction enzyme digest analysis for pET28b (+)-PfHsp90-NTD plasmid construct, (A) plasmid map of pET28b(+)-PfHsp90-NTD, showing the PfHsp90-NTD insert (yellow), BamHI and NdeI restriction sites, KanR the kanamycin resistance site (Orange) and the T7 promoter region. (B) Agarose gel electrophoresis of pET28b(+)-PfHsp90-NTD plasmid: lane M, it's the DNA molecular weight marker in Bp's; lane NC, is the uncut pET28b(+)-PfHsp90-NTD plasmid; lane SD, is the single digest fragment, digested with BamHI; lane DD, is the double digest of pET28b(+)-PfHsp90-NTD plasmid using both restriction enzymes, BamHI and NdeI.

5.1.2 Recombinant PfHsp90 expression

The plasmids expressing PfHsp90-NTD protein from *P. falciparum* was successfully transformed into *E. coli* BL21 cells. The level of protein production was analysed by SDS-PAGE and Western blot (Figure 19). Following the induction, there was a gradual increase in the production of protein

at approximately 29kDa, the expected size of PfHsp90-NTD (Figure 19A). Protein production did not occur for an hour following induction with 1mM IPTG. However, after 2 hours significant increase in the production of PfHsp90-NTD can be seen (Figure 19A). Maximum amount of protein produced after 24 hours (Figure 19A). The Western blot confirms that No expression occurred in the uninduced cells (Figure 19- J.C). Although there seems to be some unspecific binding, which might be attributed to the polyclonal primary antibody used (Figure 19B).

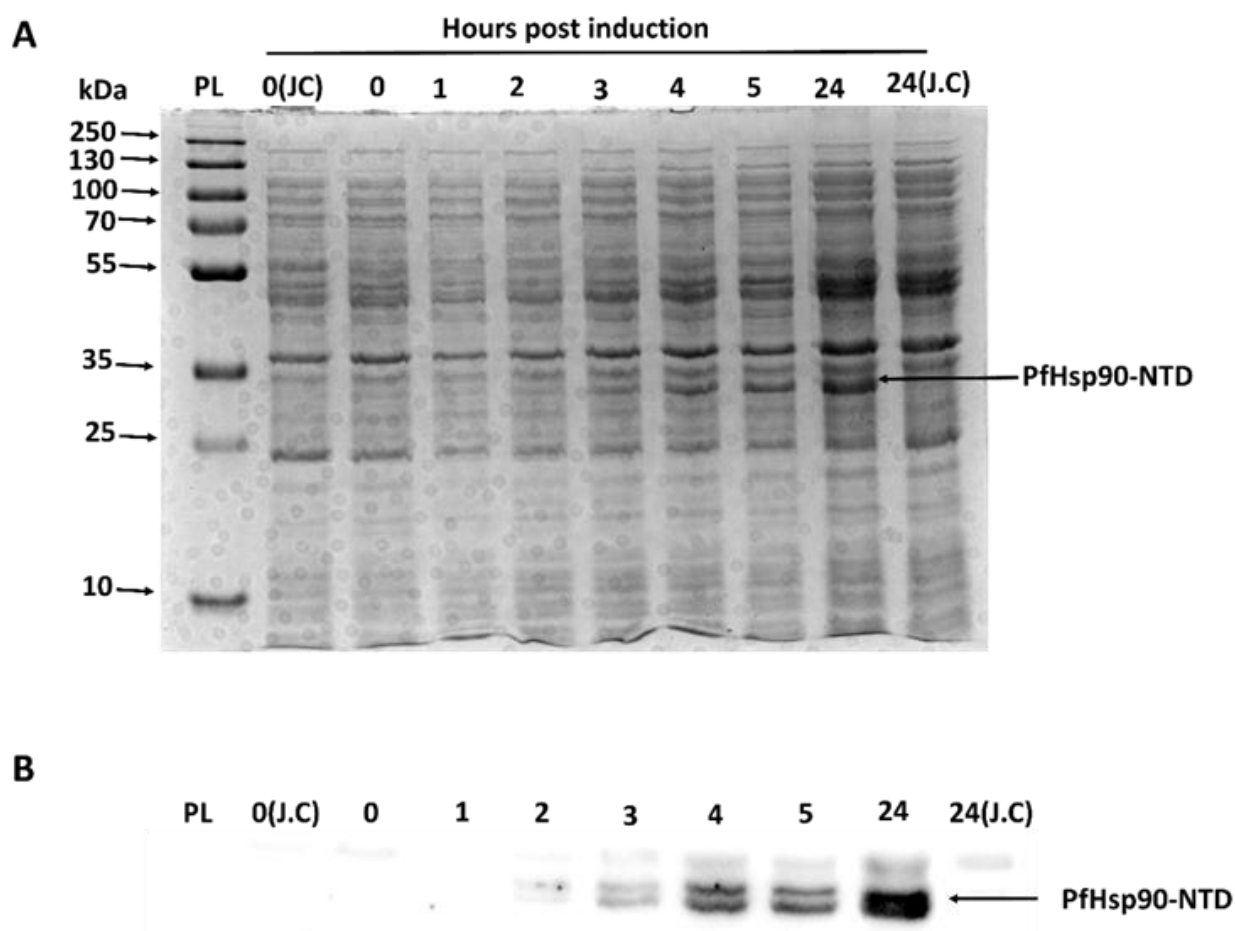


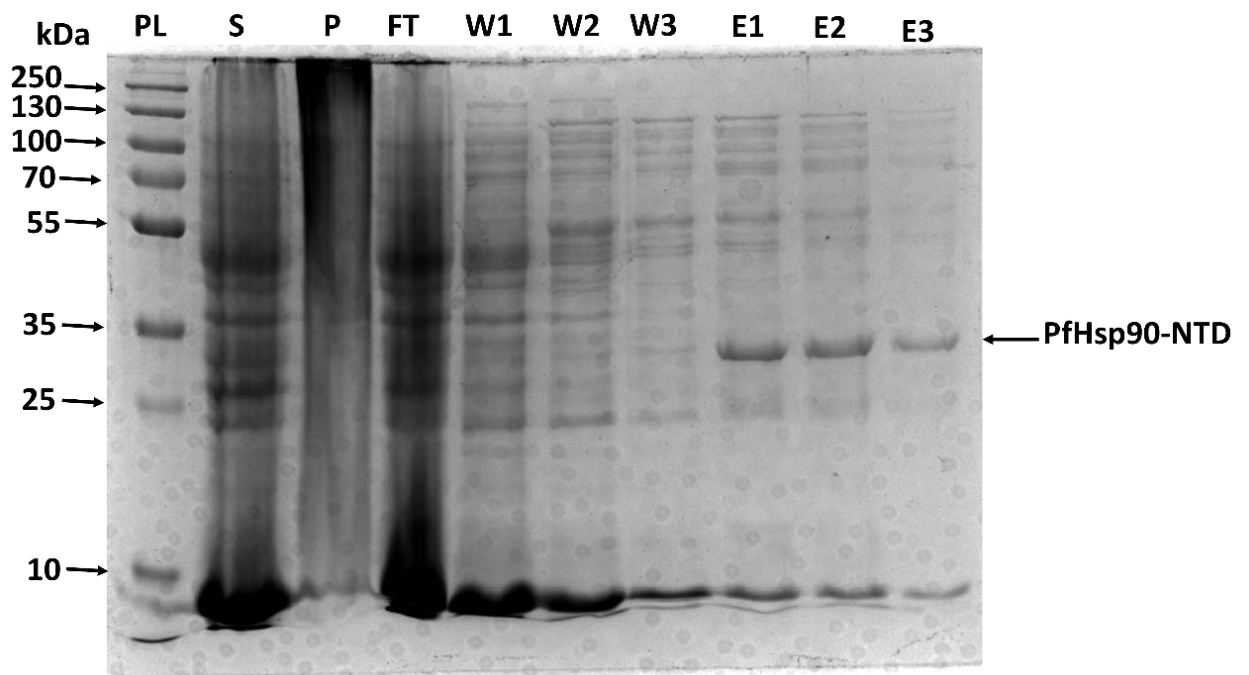
Figure 19: PfHsp90-NTD expression on BL21(DE3) E. coli cells analysis. (A) SDS-PAGE stained with Coomassie blue. (B) Western blot analysis using His-tagged antibodies. Lane PL represents protein ladder, lane 0(J.C) uninduced cells at time point zero, lane 0 total extract of cells transformed prior to IPTG induction; lane 1 – 5 and 24 are hourly samples and overnight samples respectively, after induction with IPTG; 24(J.C) are uninduced overnight sample.

5.1.3 Protein purification

Coomassie-stained 12% SDS-PAGE was used to analyse the purification samples (Figure 20 A), also the proteins were trans-blotted onto nitrocellulose membranes, stained with Ponceau Stain to confirm the protein bands, and then probed with a mouse anti-His polyclonal antibody to find the purification of recombinant proteins using Western blot analysis (Figure 20 B). The PfHsp90-

NTD protein was therefore successfully purified from the soluble fractions. As it can be seen from Figure 20 was successfully bound to the beads as none of the protein was detected in the flow through (Figure 20; lane FT). Additionally, during the washes, no detectable PfHsp90-NTD protein was washed away, therefore the protein was eluted successfully, although some contaminating proteins were observed just above our recombinant protein. To avoid confounding results in this study the contaminants were further eliminated using protein dialysis. The amounts of protein produced was quantified using Bradford assay, and the results showed on average >0.5mg/mL yields were obtained, representing sufficient amounts of protein to carry out subsequent biochemical and biophysical assays.

A



B

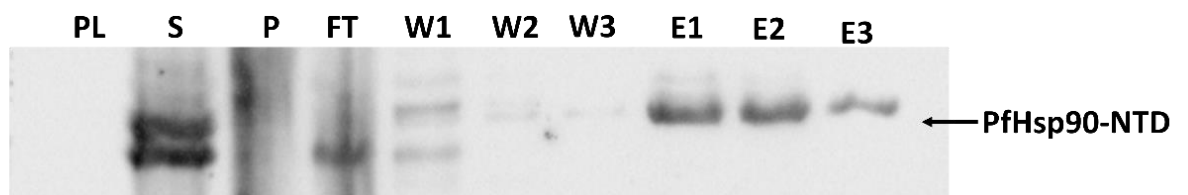
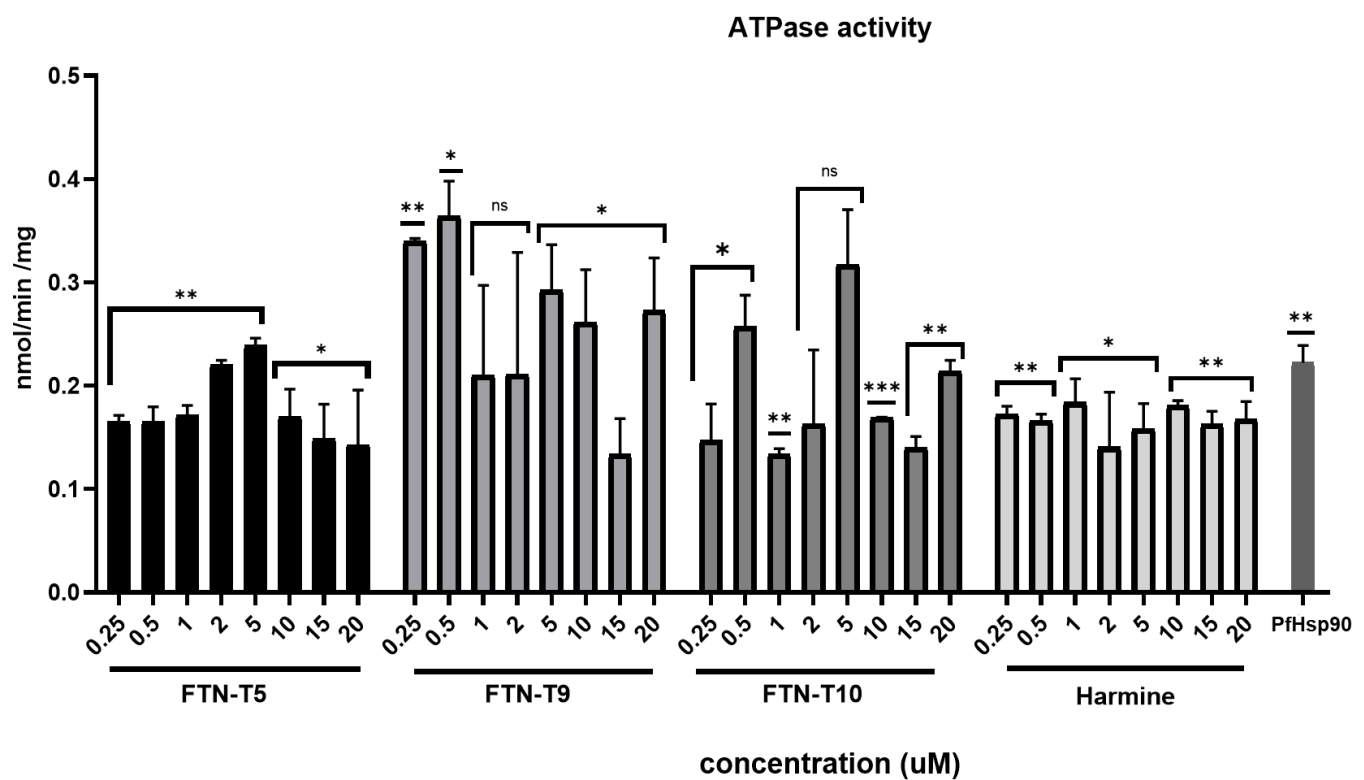
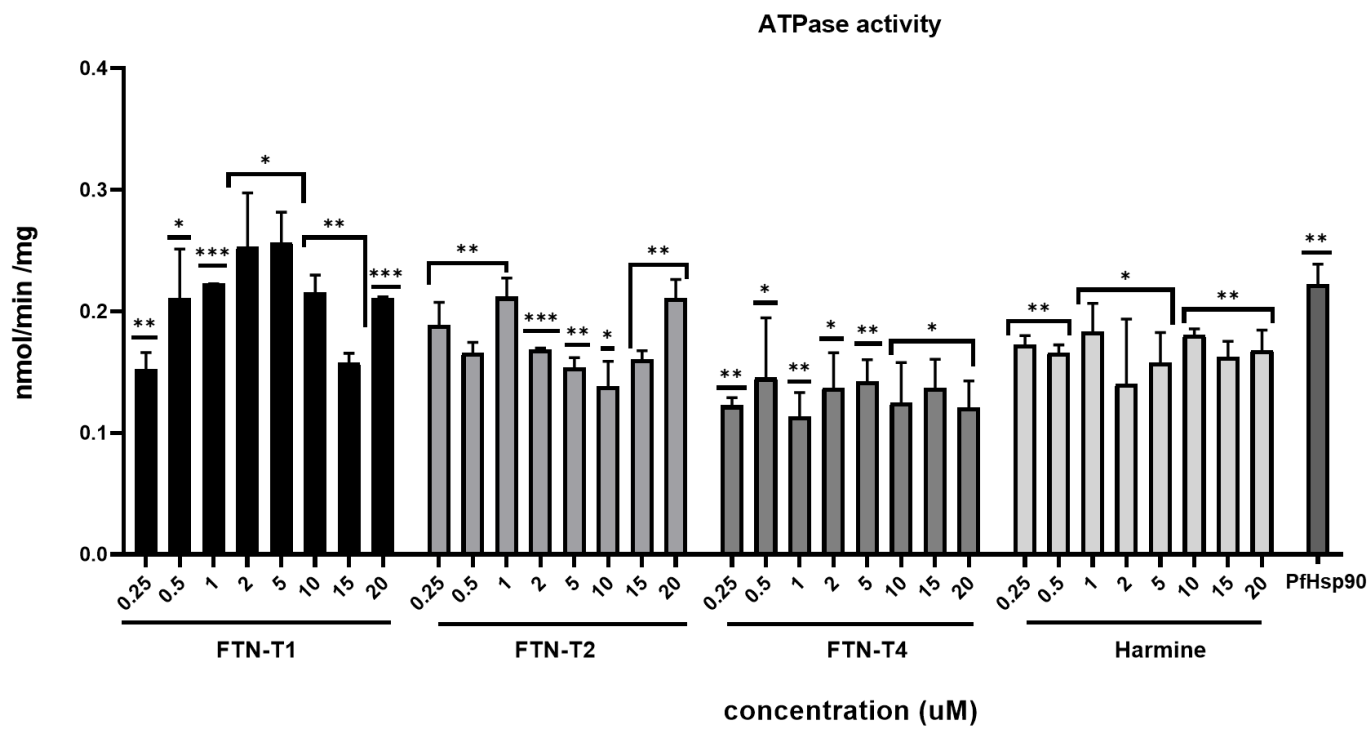


Figure 20: SDS-PAGE analysis of PfHsp90 protein purification samples. The purification samples were analysed by SDS-PAGE on a 12 % gel. PL: Protein ladder; Lane 2: Filtered lysate (soluble supernatant); Lane 3: insoluble fraction (pellet); Lane 4: Sample flow-through; Lane 5: Wash 1; Lane 6: Wash 2; Lane 7: wash 3; Lane 8: Elution1; Lanes 9: elution 2 and Lane 10: elution 3.

5.1.4 ATPase assay

The generated inhibitors compete with ATP for binding site of the molecular chaperone, therefore, there is expected to be competitive binding for the ATPase domain. If the inhibitors have stronger affinity for PfHsp90-NTD, they will displace ATP from the binding pocket, resulting in reduced phosphate release. The ATPase assay was used to assess the binding effect of generated inhibitors on the ATPase activity of PfHsp90 following the *in-silico* studies on the interactions between our inhibitors and PfHsp90 was investigated. ATPase activity of PfHsp90-NTD at a concentration of 2 μ M ATPase was determined in the presence of our inhibitors at different concentrations (0.25, 0.5, 1, 2, 5, 10, 15 and 20 μ M) including a positive control of Harmine. Figure 21 displays the results of the ATPase activity assay for the compounds. FTN-T4 and FTN-T10 brought about the most reduction in the ATPase activity of PfHsp90 even lower than the positive control harmine, suggesting that they have a higher potency than harmine, correlating with the *in-silico* data obtained. With FTN-T4, the greatest reduction in the ATPase activity was observed at a compound concentration of 1 μ M followed by 10 μ M, while with FTN-T10, concentrations of 0.25 μ M, 1 μ M and 15 μ M produced the greatest reduction, with observed outliers at concentrations of 0.5 μ M and 5 μ M as they don't follow a similar inhibition pattern to the rest of the concentrations, which is due to human error, not necessarily indicating that the ligand promotes ATPase activity of the protein at that concentration. FTN-T2, FTN-T5, FTN-M2, FTN-M3 and FTN-B1 produced a similar reduction to the ATPase activity of PfHsp90 as the positive control harmine. The compounds, FTN-T1, FTN-T9 and FTN-B4 resulted in the increase of the ATPase activity of PfHsp90, which was very puzzling particularly with FTN-T1 which showed the highest binding affinity towards PfHsp90 in the *in silico* data indicating the complexity of their interactions and potential modes of action. However, Hsp90 proteins can assume a variety of conformations, including open and closed states, which may impact the accessibility of the ATP binding site and, as a result, the inhibitors' ability to inhibit the protein (Chiosis *et al.*, 2013). Therefore, conformational flexibility of PfHsp90 and the potential for induced fit effects should be considered, as these may influence the binding interactions and subsequent ATPase activity inhibition. The inhibitors displaying low ATPase activity inhibition may have limited binding affinity or suboptimal binding modes, resulting in weaker inhibition (Zheng and Ramirez, 2000). In contrast, inhibitors with substantial ATPase activity inhibition are expected to have higher binding affinity and favourable interactions with key ATP hydrolysis residues (Chou *et al.*, 2014).

Validating the observed inhibitory effects of the compounds by additional assays such as fluorescence polarization assay (FPA) binding studies (biophysical analysis) and cellular assays would be crucial to confirm the potency and selectivity of these inhibitors against PfHsp90.



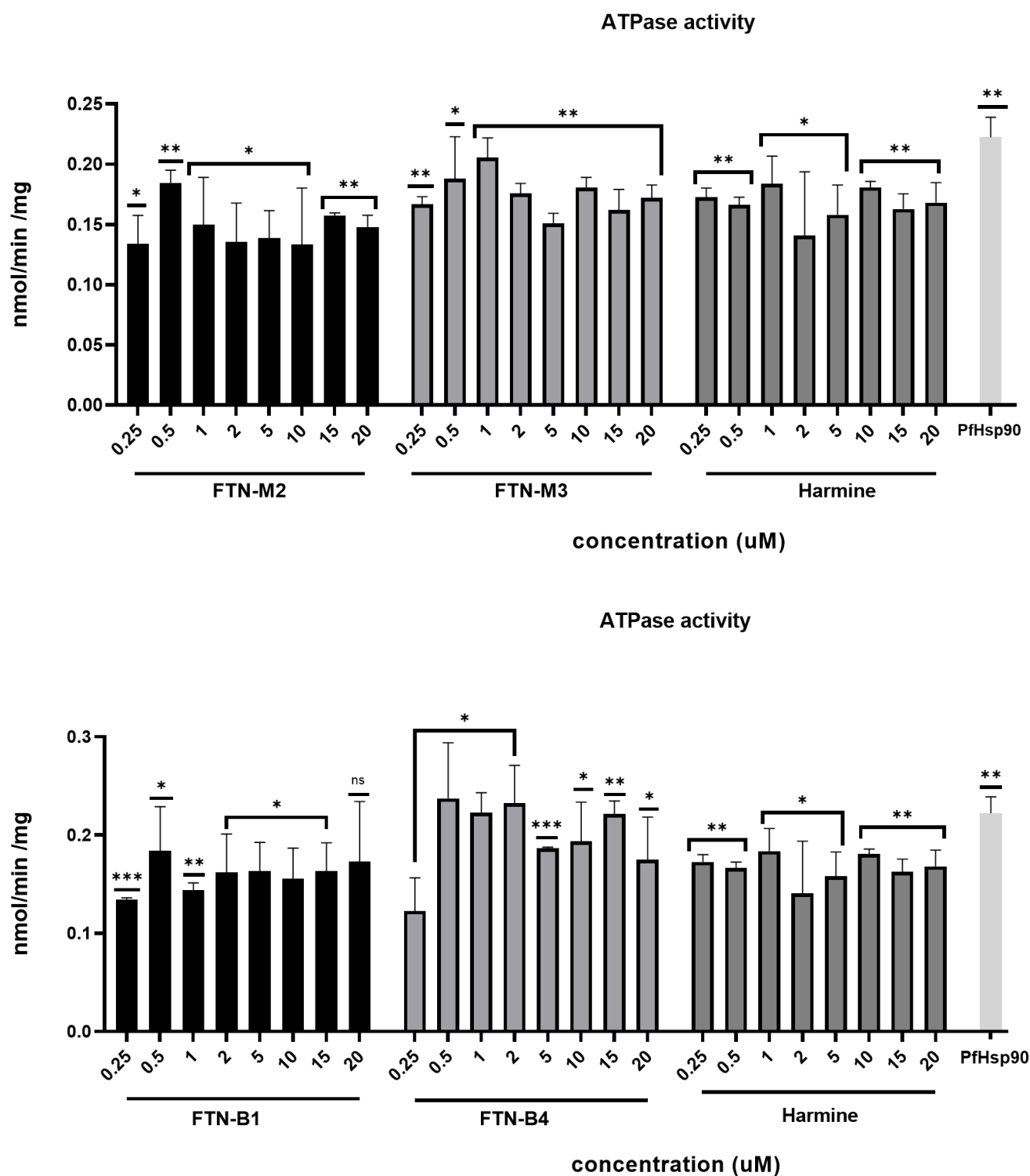


Figure 21: ATPase activity of PfHsp90 against our inhibitors. Compounds were added in increasing concentrations (0.25–20 μ M) to a fixed amount of PfHsp90 (2 μ M) to investigate its effect on the ATPase activity of PfHsp90 with Harmine as our positive control. Data are reported as the residual ATPase activity (nmol/min/mg) compared to that observed for an untreated sample (PfHsp90). Data are the mean of independent experiments performed in duplicates analysed by Graphpad prism. The error bar represents the standard deviation of measurements, while * indicates significance at $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ and ns > 0 .

5.2 BIOPHYSICAL ANALYSIS

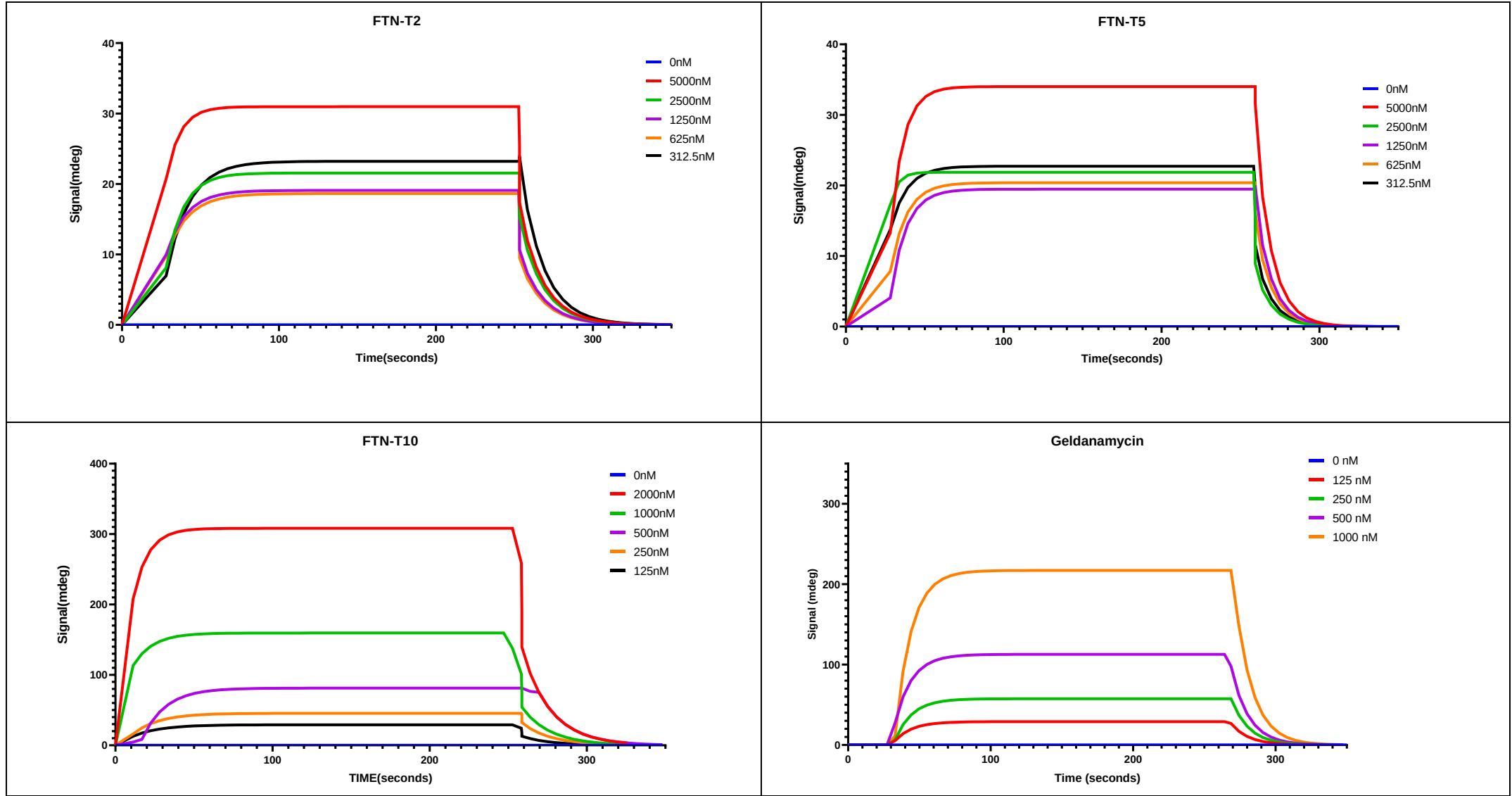
5.2.1 Surface plasmon resonance (SPR)

Surface plasmon resonance (SPR) is a powerful biophysical technique used to study protein-protein, protein-ligand, and other biomolecular interactions in real-time. The results obtained from SPR experiments provide valuable information on the binding affinity, specificity, kinetics, and thermodynamics of these interactions (Cooper, 2003). The SPR results for this study are presented in Figure 22 as sensorgrams, which are graphs that show the changes in the refractive index of the sensor surface as a function of time. A dose dependent response was observed for all three of the tested compounds interacting with the immobilized PfHsp90-NTD with clearly identifiable exponential curves throughout the attachment and dissociation phases (Figure 22). Moderate binding affinity was observed from compound FTN-T2 with an equilibrium constant (K_D) of 7,11 μM and FTN-T10 at K_D value of 1.37 μM (Table 7) comparable to that of geldanamycin (8.39 μM), while FTN-T5 exhibited K_D values of 19.9 μM (Table 7), which are comparable to ADP (20.6 μM) (Murillo-Solano *et al.*, 2017). High binding affinity is indicated by a low K_D value. Generally the observed K_D values were comparable to that of ADP, but FTN-T2 and FTN-T10 are binding more strongly than ADP although with moderate affinity but comparable to geldanamycin in terms of binding affinity.

Table 7: Kinetics constants measured by SPR for the interaction between tested compounds and immobilized PfHsp90. The association rate constant is represented by K_a (1/MS), dissociation rate constant K_d (1/s), and the equilibrium constant denoting affinity K_D (μM).

Ligand	Analyte	K_a (1/MS)	K_d (1/s)	K_D (μM)	Chi ²	U-value
FTN-T2	PfHsp90	9,44($\pm$ 1.12)e ³	6,71($\pm$ 1.19)e ⁻⁵	7,11	3,47	4,4
FTN-T5		4,84($\pm$ 7.40)e ²	9,63($\pm$ 7.33)e ⁻⁶	19,9	3,07	5,2
FTN-T10		3,07($\pm$ 2.86)e ²	4,22($\pm$ 1.98)e ⁻⁵	1,37	0,02	3,5
Harmine		1.76($\pm$,49)e ²	3,20 ($\pm$ 4,82)e ⁻⁶	18,2	0	11,7
Geldanamycin		1.46 ($\pm$ 2.11)e ⁴	1,23 ($\pm$ 2.83)e ⁻⁴	8,39	13,5	8,7

^a K_a and K_d values provided insights into the kinetics of the interaction, with high K_a indicating rapid association and a high K_d indicating rapid dissociation. The Chi² and U-value gave information about the goodness of fit for the graphs and if it was within the acceptable ranges, and also if all the tested compounds were within the acceptable ranges (Chi² ranged from 0 to 11.32 and U-value ranges from 3.5 to 11.7).



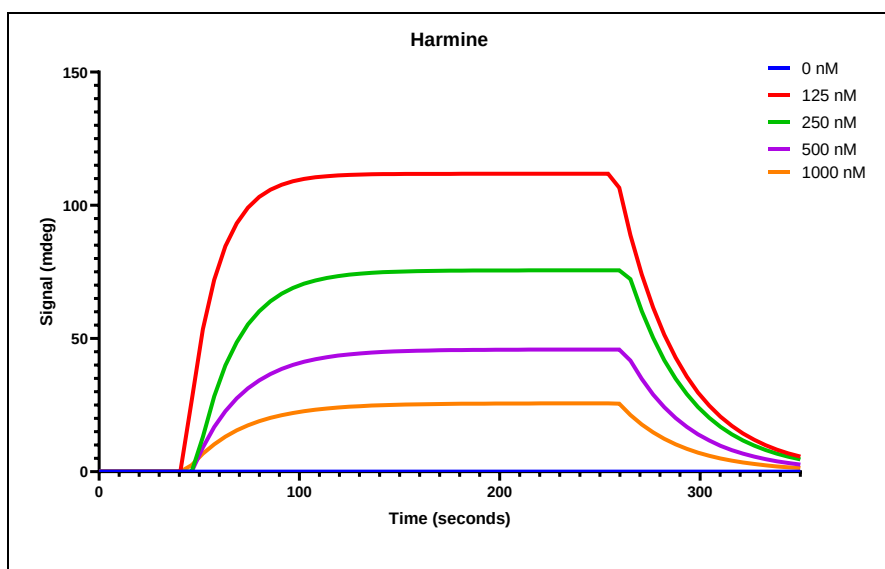


Figure 22: Sansogram from the Analysis of the interaction of FTN-T2, FTN-T5, FTN-T10, geldanamycin and Harmine with PfHsp90-NTD by surface plasmon resonance spectroscopy.

The affinity determined by SPR is in agreement with the docking scores obtained through IFD and the binding free energy ΔG calculated by MM-GB/SA method. The positive correlations indicate that molecules with higher docking scores and more favourable binding free energy values tend to exhibit stronger affinity as measured by SPR. This agreement and correlation between experimental and computational methods provide confidence in the accuracy and reliability of the predictions made by the docking and MM-GB/SA approaches.

5.3 Chapter summary

In this chapter, the successful construction and confirmation of the pET28b (+)-PfHsp90-NTD plasmid via restriction enzyme digest reactions were detailed. The expression of the PfHsp90-NTD protein in *E. coli* BL21 cells following transformation showed a gradual increase in protein production after induction with 1mM IPTG, peaking at 24 hours, corroborated by SDS-PAGE and Western blot analyses. Protein purification was achieved via Coomassie-stained SDS-PAGE and Western blotting, confirming successful binding to beads and elution, yielding >0.5mg/mL of purified protein. An ATPase assay was conducted to evaluate the impact of generated inhibitors on PfHsp90-NTD's ATPase activity, revealing FTN-T4 and FTN-T10 as potent inhibitors with reduced ATPase activity compared to the positive control harmine. The study highlighted inconsistencies in some inhibitor effects on ATPase activity, suggesting diverse interaction modes and potential conformational influences on protein activity. Recommendations were made for additional assays like fluorescence polarization and cellular studies. The chapter also discussed

surface plasmon resonance (SPR) findings, indicating moderate to high binding affinities of the tested compounds with PfHsp90-NTD, in agreement with computational predictions, thereby affirming the reliability of the *in silico* approaches employed. Overall, the chapter showcased successful protein expression, purification, inhibitor evaluation, and binding affinity assessment of PfHsp90-NTD, setting the stage for further validation and characterization of these inhibitors through additional biophysical and cellular assays.

CHAPTER 6

CONCLUSION

In conclusion, this study aimed to generate *P. falciparum* inhibitors targeting the molecular chaperone PfHsp90, which plays a crucial role in the survival of the parasite, to address the urgent need for novel and potent antimalarial drugs in the face of drug-resistant strains of *P. falciparum*, using rational drug design strategies and validating them through biochemical and biophysical assays. The study sought to overcome the limitations of current antimalarial therapies and pave the way for more effective treatment options. Through a comprehensive analysis of the binding interactions with PfHsp90, small molecule inhibitors of PfHsp90 were identified.

The pathfinder reaction-based enumeration was employed to enumerate analogues using a reference compound from literature with proven affinity towards PfHsp90. Using this retrosynthetic analysis, we enumerated 10 000 design ideas after PAINS filters, and physicochemical filters. This in turn, enabled the construction of active learning models to transverse the large chemical space of the enumerated compounds of >2 million compounds from the chEMBL database and predict their activities. Exploring the chEMBL database enabled the discovery of more potent inhibitors while also unlocking new areas of a chemical space using active learning models. A total of 13 models were built using active learning QSAR, the final model (model 13) where convergence was attained, achieved good statistical results such as an $R^2 = 0.74$, $Q^2 = 0.62$, STD Dev = 0.48, and RMSE = 0.56. Additional computational techniques, including IFD, MM-GB/SA, Qikprop and MDS, further filtered the compounds to identify drug-like compounds having good drug-like properties with a modest affinity for the PfHsp90 while also investigating the conformational stability of the protein ligand complexes. Through the *in-silico* study, compounds FTN-T1, FTN-T2 and FTN-T10 were seen to have the highest affinity towards PfHsp90 among the compounds studied.

The validation of the designed inhibitors through *in vitro* biochemical and biophysical studies further demonstrated the potential of these inhibitors as antimalarial agents. FTN-T4, FTN-T5 and FTN-T10 produced the greatest reduction in the chaperones ATPase activity at varying concentrations. Moreover, SPR shed significant insights into the binding kinetics of the identified inhibitors. The insights gained from this study, contribute extraordinarily to our knowledge of the structure-function relationship of PfHsp90. The generation of PfHsp90 selective inhibitors using machine-learning techniques represents a significant advancement in the field of antimalarial drug development. The findings presented in this dissertation

contribute to the broader goal of eradicating malaria and improving global health outcomes in malaria-endemic regions. A subsequent study will be further undertaken for the development and optimization of structure-activity relationships (SAR) of the identified inhibitors. These SAR studies will provide valuable insights into the chemical moieties necessary for elevated inhibition while minimizing off-target effects, contributing to the overall success of the rational drug design process.

This study shows how ML applications can serve as a democratization tool to African researchers, empowering them to actively contribute to field of drug discovery and overcome traditional barriers. Thanks to the availability of high processing capacity through cloud computing platforms, cost-effective infrastructure and the increasing availability of genomic and clinical data, African scientists can develop ML models and make predictions relevant to drug discovery facilitating complex calculations, simulations, and data analysis. Additionally, collaboration is encouraged since ML promotes international cooperation among academics, allowing African scientists to work with specialists across the world, share expertise, and gain insight from other viewpoints. African researchers may gain machine learning competence through capacity development initiatives, fostering inclusivity and scientific creativity in drug discovery, and ultimately enabling Africa to have a significant stake in the sector.

Future recommendations

Conducting detailed SAR studies is crucial to optimize the potency and selectivity of the PfHsp90 generated inhibitors. Modifying the lead compounds' chemical structure and evaluating the impact of these modifications on their activity against PfHsp90 will guide the design of more potent and selective inhibitors. In the future it will be important validate the therapeutic potential of the generated PfHsp90 inhibitors through in vivo efficacy studies using animal models infected with *P. falciparum* parasite. The effectiveness, safety, and pharmacokinetics of these inhibitors can be better understood through these investigations. Resistance Studies should also be undertaken, as with any antimalarial drug, the potential for the development of resistance by the Plasmodium parasite must be monitored. This could involve combination therapy with other antimalarial drugs or the development of multi-targeted inhibitors. If the PfHsp90 inhibitors show promising preclinical results, the next step would be to translate the research findings into clinical trials. Collaborating with pharmaceutical companies, regulatory bodies, and funding agencies can facilitate the progression of the inhibitors from bench to the markets. Designing and conducting rigorous clinical trials to evaluate the safety, efficacy, and optimal dosage of the inhibitors in humans is crucial for their potential use as novel antimalarial agents.

BIBLIOGRAPHY

- Achan, J., Talisuna, A.O., Erhart, A., Yeka, A., Tibenderana, J.K., Baliraine, F.N., Rosenthal, P.J. and D'Alessandro, U., 2011. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria journal*, 10(1), pp.1-12.
- Aggarwal, R. and Ranganathan, P., 2016. Common pitfalls in statistical analysis: The use of correlation techniques. *Perspectives in clinical research*, 7(4), p.187.
- Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L., 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, 8, pp.1-74.
- Ashley, E.A., Dhorda, M., Fairhurst, R.M., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Anderson, J.M., Mao, S., Sam, B. and Sopha, C., 2014. Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *New England Journal of Medicine*, 371(5), pp.411-423.
- Balikagala, B., Sakurai-Yatsushiro, M., Tachibana, SI. 2020. Recovery and stable persistence of chloroquine sensitivity in *Plasmodium falciparum* parasites after its discontinued use in Northern Uganda. *Malar J* 19.
- Ban, T.A., 2022. The role of serendipity in drug discovery. *Dialogues in clinical neuroscience*.
Doytchinova, I., 2022. Drug design—past, present, future. *Molecules*, 27(5), p.1496.
- Bartolommei, G., Moncelli, M.R. and Tadini-Buoninsegni, F., 2013. A method to measure hydrolytic activity of adenosinetriphosphatases (ATPases). *PLOS one*, 8(3), p.e58615.
- Bartoloni, A. and Zammarchi, L., 2012. Clinical aspects of uncomplicated and severe malaria. *Mediterranean journal of hematology and infectious diseases*, 4(1).
- Batool, M., Ahmad, B. and Choi, S., 2019. A structure-based drug discovery paradigm. *International journal of molecular sciences*, 20(11), p.2783.
- Battle, K.E., Lucas, T.C., Nguyen, M., Howes, R.E., Nandi, A.K., Twohig, K.A., Pfeffer, D.A., Cameron, E., Rao, P.C., Casey, D. and Gibson, H.S., 2019. Mapping the global endemicity and clinical burden of *Plasmodium vivax*, 2000–17: a spatial and temporal modelling study. *The Lancet*, 394(10195), pp.332-343.
- Bissantz, C., Kuhn, B. and Stahl, M., 2010. A medicinal chemist's guide to molecular interactions. *Journal of medicinal chemistry*, 53(14), pp.5061-5084.
- Blasco, B., Leroy, D. and Fidock, D.A., 2017. Antimalarial drug resistance: linking *Plasmodium falciparum* parasite biology to the clinic. *Nature medicine*, 23(8), pp.917-928.

- Blasco, B., Leroy, D. and Fidock, D.A., 2017. Antimalarial drug resistance: linking *Plasmodium falciparum* parasite biology to the clinic. *Nature medicine*, 23(8), pp.917-928.
- Borrmann, S., Adegnik, A.A., Missinou, M.A., Binder, R.K., Issifou, S., Schindler, A., Matsiegui, P.B., Kun, J.F., Krishna, S., Lell, B. and Kremsner, P.G., 2003. Short-course artesunate treatment of uncomplicated *Plasmodium falciparum* malaria in Gabon. *Antimicrobial agents and chemotherapy*, 47(3), pp.901-904.
- Bourgard, C., Albrecht, L., Kayano, A.C., Sunnerhagen, P. and Costa, F.T., 2018. *Plasmodium vivax* biology: insights provided by genomics, transcriptomics and proteomics. *Frontiers in cellular and infection microbiology*, 8, p.34.
- Bray, P.G., Martin, R.E., Tilley, L., Ward, S.A., Kirk, K. and Fidock, D.A., 2005. Defining the role of PfCRT in *Plasmodium falciparum* chloroquine resistance. *Molecular microbiology*, 56(2), pp.323-333.
- Buabeng, K.O., Duwiewua, M., Dodoo, A.N., Matowe, L.K. and Enlund, H., 2007. Self-reported use of anti-malarial drugs and health facility management of malaria in Ghana. *Malaria journal*, 6, pp.1-6.
- Butkiewicz, M., Wang, Y., Bryant, S.H., Lowe Jr, E.W., Weaver, D.C. and Meiler, J., 2017. High-throughput screening assay datasets from the pubchem database. *Chemical informatics (Wilmington, Del.)*, 3(1).
- Capela, R., Moreira, R. and Lopes, F., 2019. An overview of drug resistance in protozoal diseases. *International journal of molecular sciences*, 20(22), p.5748.
- Chen, L., Li, X., & Liu, Y. 2019. Machine learning for drug discovery. *Frontiers in Chemistry*, 7, 176.
- Hansen, K. 2016. Machine learning in drug discovery. *Drug Discovery Today*, 21(10), 1480-1489.
- Kohavi, R., Long, P., & Warmuth, M. K. (2018). Automated machine learning: Methods, systems, challenges. *Journal of Machine Learning Research*, 19(81), 1-5.
- Wang, Y., Zhang, Y., & Lu, Y. 2018. Machine learning in drug discovery: Opportunities and challenges. *Current Opinion in Chemical Biology*, 44, 9-16.
- Xu, Y., Song, Q., & Liu, J. (2020). Machine learning in drug discovery: Applications and future perspectives. *Journal of Chemical Information and Modeling*, 60(9), 4396-4407.
- Cheney C. Fight against malaria stalling and could reverse, warns 2017 World Malaria Report. <http://theconversation.com/global-malaria-report-reveals-africas-hits-and-misses-heres-what-to-do-88577>. Accessed 29 Oct 2022
- Chiosis, G., Dickey, C.A. and Johnson, J.L., 2013. A global view of Hsp90 functions. *Nature structural & molecular biology*, 20(1), pp.1-4.

- Cibulskis, R.E., Alonso, P., Aponte, J., Aregawi, M., Barrette, A., Bergeron, L., Fergus, C.A., Knox, T., Lynch, M., Patouillard, E. and Schwarte, S., 2016. Malaria: global progress 2000–2015 and future challenges. *Infectious diseases of poverty*, 5, pp.1-8.
- Colón-González, F.J., Sewe, M.O., Tompkins, A.M., Sjödin, H., Casallas, A., Rocklöv, J., Caminade, C. and Lowe, R., 2021. Projecting the risk of mosquito-borne diseases in a warmer and more populated world: a multi-model, multi-scenario intercomparison modelling study. *The Lancet Planetary Health*, 5(7), pp.e404-e414.
- D. D. D. D. Dukor. 2012. "Surface Plasmon Resonance Biosensors: Applications in Drug Discovery and Development," *Journal of Biomolecular Screening*, vol. 17, pp. 997-1009.
- Das, D., Koh, Y., Tojo, Y., GhoshMitsuya, A. K. H., and Mitsuya, H., 2009. Prediction of Potency of Protease Inhibitors Using Free Energy Simulations with Polarizable Quantum Mechanics-Based Ligand Charges and a Hybrid Water Model. *J. Chem. Inf. Model.* 49 (12), 2851–2862.
- Dekel, E. and Alon, U., 2005. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050), pp.588-592.
- Dixit, A. and Verkhivker, G.M., 2012. Probing molecular mechanisms of the Hsp90 chaperone: Biophysical modeling identifies key regulators of functional dynamics. *PloS one*, 7(5), p.e37605.
- Doumbo, O.K., Thera, M.A., Koné, A.K., Raza, A., Tempest, L.J., Lyke, K.E., Plowe, C.V. and Rowe, J.A., 2009. High levels of Plasmodium falciparum rosetting in all clinical forms of severe malaria in African children. *The American journal of tropical medicine and hygiene*, 81(6), p.987.
- Doytchinova, I., 2022. Drug design—past, present, future. *Molecules*, 27(5), p.1496.
- Elokely, K. M., & Doerksen, R. J. 2013. Docking challenge: Protein sampling and molecular docking performance. *Journal of chemical information and modeling*, 53(8), 1934.
- Fischer, A., Smiesko, M., Sellner, M. and Lill, M.A., 2021. Decision making in structure-based drug discovery: visual inspection of docking results. *Journal of Medicinal Chemistry*, 64(5), pp.2489-2500.
- Fitch, C.D., 2004. Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs. *Life sciences*, 74(16), pp.1957-1972.
- Flower, D.R. ed., 2002. Drug design: cutting edge approaches (Vol. 279). Royal Society of Chemistry.

- Forouzesh, N. and Mishra, N., 2021. An effective MM/GBSA protocol for absolute binding free energy calculations: A case study on SARS-CoV-2 spike protein and the human ACE2 receptor. *Molecules*, 26(8), p.2383.
- Forouzesh, N. and Mishra, N., 2021. An effective MM/GBSA protocol for absolute binding free energy calculations: A case study on SARS-CoV-2 spike protein and the human ACE2 receptor. *Molecules*, 26(8), p.2383.
- Garnier, C., Lafitte, D., Tsvetkov, P.O., Barbier, P., Leclerc-Devin, J., Millot, J.M., Briand, C., Makarov, A.A., Catelli, M.G. and Peyrot, V., 2002. Binding of ATP to heat shock protein 90: evidence for an ATP-binding site in the C-terminal domain. *Journal of Biological Chemistry*, 277(14), pp.12208-12214.
- Genheden, S. and Ryde, U., 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, 10(5), pp.449-461
- Genheden, S. and Ryde, U., 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, 10(5), pp.449-461.
- Gilles, H.M. and Warrell, D.A., 1996. *Bruce-Chwatt's essential malariology* (No. Ed. 3). Edward Arnold (Publisher) Ltd.
- Gontie, G.B., Wolde, H.F. and Baraki, A.G., 2020. Prevalence and associated factors of malaria among pregnant women in Sherkole district, Benishangul Gumuz regional state, West Ethiopia. *BMC Infectious Diseases*, 20(1), pp.1-8.
- Gopalakrishnan, A.M. and Kumar, N., 2015. Antimalarial action of artesunate involves DNA damage mediated by reactive oxygen species. *Antimicrobial agents and chemotherapy*, 59(1), pp.317-325.
- Gopalakrishnan, A.M. and Kumar, N., 2015. Antimalarial action of artesunate involves DNA damage mediated by reactive oxygen species. *Antimicrobial agents and chemotherapy*, 59(1), pp.317-325.
- Greenwood, D., 1992. The quinine connection. *Journal of Antimicrobial Chemotherapy*, 30(4), pp.417-427.
- Guo, Z., 2017. The modification of natural products for medical use. *Acta Pharmaceutica Sinica B*, 7(2), pp.119-136.
- Guterres, H. and Im, W., 2020. Improving protein-ligand docking results with high-throughput molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 60(4), pp.2189-2198.

- Guyatt, H.L. and Snow, R.W., 2004. Impact of malaria during pregnancy on low birth weight in sub-Saharan Africa. *Clinical microbiology reviews*, 17(4), pp.760-769.
- Harinasuta, T. and Bunnang, D., 1988. Malaria, principles and practice of malariology. First edition. Edinburg: Churchill Livingstone.
- Hoy, S.M., 2022. Pimitepib: first approval. *Drugs*, 82(13), pp.1413-1418.
- Hughes, J.D., Blagg, J., Price, D.A., Bailey, S., DeCrescenzo, G.A., Devraj, R.V., Ellsworth, E., Fobian, Y.M., Gibbs, M.E., Gilles, R.W. and Greene, N., 2008. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorganic & medicinal chemistry letters*, 18(17), pp.4872-4875.
- Hughes, J.P., Rees, S., Kalindjian, S.B. and Philpott, K.L., 2011. Principles of early drug discovery. *British journal of pharmacology*, 162(6), pp.1239-1249.
- Hunter, M.C., O'Hagan, K.L., Kenyon, A., Dhanani, K.C., Prinsloo, E. and Edkins, A.L., 2014. Hsp90 binds directly to fibronectin (FN) and inhibition reduces the extracellular fibronectin matrix in breast cancer cells. *PloS one*, 9(1), p.e86842.
- Idro, R., Marsh, K., John, C.C. and Newton, C.R., 2010. Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatric research*, 68(4), pp.267-274.
- IPCC. Climate change 2014: Synthesis Report, Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland; 2014. 151 pp
- Ismail, M. R., J. Ordi, C. Menendez, P. J. Ventura, J. J. Aponte, E. Kahigwa, R. Hirt, A. Cardesa, and A. P. L. Alonso. 2000. Placental pathology in malaria: a histological, immunohistochemical, and quantitative study. *Hum. Pathol.* 31:85-93.
- Jackson, S.E., 2013. Hsp90: structure and function. *Molecular chaperones*, pp.155-240.
- Jensen, M.R., Schoepfer, J., Radimerski, T., Massey, A., Guy, C.T., Brueggen, J., Quadt, C., Buckler, A., Cozens, R., Drysdale, M.J. and Garcia-Echeverria, C., 2008. NVP-AUY922: a small molecule HSP90 inhibitor with potent antitumor activity in preclinical breast cancer models. *Breast Cancer Research*, 10(2), pp.1-12.
- Khan, A.U., 2016. Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug discovery today*, 21(8), pp.1291-1302
- Kholodenko, B. N. 2000. Signaling networks and cell cycle control. *Nature*, 407(6801), 725-732.

- Kochar, D.K., Kochar, S.K., Agrawal, R.P., Sabir, M., Nayak, K.C., Agrawal, T.D., Purohit, V.P. and Gupta, R.P., 2006. The changing spectrum of severe falciparum malaria: a clinical study from Bikaner (northwest India). *Journal of vector borne diseases*, 43(3), p.104.
- Kochar, D.K., Saxena, V., Singh, N., Kochar, S.K., Kumar, S.V. and Das, A., 2005. Plasmodium vivax malaria. *Emerging infectious diseases*, 11(1), p.132.
- Kulkarni, M.A., Duguay, C. & Ost, K., 2022. Charting the evidence for climate change impacts on the global spread of malaria and dengue and adaptive responses: a scoping review of reviews. *Global Health* 18, 1
- Laurens, M.B., 2020. RTS, S/AS01 vaccine (Mosquirix™): an overview. *Human vaccines & immunotherapeutics*, 16(3), pp.480-489.
- Leeson, P.D. and Springthorpe, B., 2007. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature reviews Drug discovery*, 6(11), pp.881-890.
- Leeson, P.D., Bento, A.P., Gaulton, A., Hersey, A., Manners, E.J., Radoux, C.J. and Leach, A.R., 2021. Target-based evaluation of “drug-like” properties and ligand efficiencies. *Journal of medicinal chemistry*, 64(11), pp.7210-7230.
- Li, S., He, H., Parthiban, L.J., Yin, H. and Serajuddin, A.T., 2005. IV-IVC considerations in the development of immediate-release oral dosage form. *Journal of pharmaceutical sciences*, 94(7), pp.1396-1417.
- Loeb, F., Clark, W.M., Coatney, G.R., Coggeshall, L.T., Dieuaide, F.R., Dochez, A.R., Hakansson, E.G., Marshall, E.K., Marvel, C.S., McCoy, O.R. and Saper, J.J., 1946. Activity of a new antimalarial agent, chloroquine (SN 7618): Statement approved by the Board for coordination of malarial studies. *Journal of the American Medical Association*, 130(16), pp.1069-1070.
- Lyne, P. D., Lamb, M. L., and Saeh, J. C., 2006. Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J. Med. Chem.* 49 (16), 4805–4808.
- M. A. El-Sayed, 2010. "Surface plasmon resonance-based sensing for biomedical and environmental applications," *Chemical Society Reviews*, vol. 39, pp. 2989-3005.
- Mahapatra, M.K. and Karuppasamy, M., 2022. Fundamental considerations in drug design. In *Computer Aided Drug Design (CADD): From Ligand-Based Methods to Structure-Based Approaches* (pp. 17-55). Elsevier.
- Mita, T., Tanabe, K. and Kita, K., 2009. Spread and evolution of Plasmodium falciparum drug resistance. *Parasitology international*, 58(3), pp.201-209.

- Mohanty, S., Mishra, S.K., Pati, S.S., Pattnaik, J. and Das, B.S., 2003. Complications and mortality patterns due to *Plasmodium falciparum* malaria in hospitalized adults and children, Rourkela, Orissa, India. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 97(1), pp.69-70.
- Murillo-Solano, C., Dong, C., Sanchez, C.G. and Pizarro, J.C., 2017. Identification and characterization of the antiplasmodial activity of Hsp90 inhibitors. *Malaria journal*, 16, pp.1-14.
- Mwangi T. Global malaria report reveals Africa's hits and missed: here's what to do. <https://www.devex.com/news/fight-against-malaria-stalling-and-could-reverse-warns-2017-world-malaria-report-91636>. Accessed 29 Oct 2022.
- Nadeem, A.Y., Shehzad, A., Islam, S.U., Al-Suhaimi, E.A. and Lee, Y.S., 2022. Mosquirix™ RTS, S/AS01 vaccine development, immunogenicity, and efficacy. *Vaccines*, 10(5), p.713.
- Neckers, L., Schulte, T.W. and Mimnaugh, E., 1999. Geldanamycin as a potential anti-cancer agent: its molecular target and biochemical activity. *Investigational new drugs*, 17, pp.361-373.
- Nendza, M. and Müller, M., 2010. Screening for low aquatic bioaccumulation (1): Lipinski's 'Rule of 5' and molecular size. *SAR and QSAR in Environmental Research*, 21(5-6), pp.495-512.
- Norinder, U. and Bergström, C.A., 2006. Prediction of ADMET properties. *ChemMedChem: Chemistry Enabling Drug Discovery*, 1(9), pp.920-937.
- Nsanzabana, C., 2019. Resistance to artemisinin combination therapies (ACTs): do not forget the partner drug!. *Tropical medicine and infectious disease*, 4(1), p.26.
- Obermann, W.M., Sondermann, H., Russo, A.A., Pavletich, N.P. and Hartl, F.U., 1998. In vivo function of Hsp90 is dependent on ATP binding and ATP hydrolysis. *The Journal of cell biology*, 143(4), pp.901-910.
- Okeke, I.N., Klugman, K.P., Bhutta, Z.A., Duse, A.G., Jenkins, P., O'Brien, T.F., Pablos-Mendez, A. and Laxminarayan, R., 2005. Antimicrobial resistance in developing countries. Part II: strategies for containment. *The Lancet infectious diseases*, 5(9), pp.568-580.
- Okumu, F.O. and Moore, S.J., 2011. Combining indoor residual spraying and insecticide-treated nets for malaria control in Africa: a review of possible outcomes and an outline of suggestions for the future. *Malaria journal*, 10(1), pp.1-13.
- Oladipo, H.J., Tajudeen, Y.A., Oladunjoye, I.O., Yusuff, S.I., Yusuf, R.O., Oluwaseyi, E.M., AbdulBasit, M.O., Adebisi, Y.A. and El-Sherbini, M.S., 2022. Increasing challenges of malaria control in sub-Saharan Africa: Priorities for public health research and policymakers. *Annals of Medicine and Surgery*, 81, p.104366.

- P. Felgner, 2015. "Plasmonic nanobiosensors: A review of surface plasmon resonance-based optical biosensors," *Sensors and Actuators B: Chemical*, vol. 217, pp. 1-14.
- Park, H.K., Yoon, N.G., Lee, J.E., Hu, S., Yoon, S., Kim, S.Y., Hong, J.H., Nam, D., Chae, Y.C., Park, J.B. and Kang, B.H., 2020. Unleashing the full potential of Hsp90 inhibitors as cancer therapeutics through simultaneous inactivation of Hsp90, Grp94, and TRAP1. *Experimental & molecular medicine*, 52(1), pp.79-91.
- Park, K.S., Hong, Y.S., Choi, J., Yoon, S., Kang, J., Kim, D., Lee, K.P., Im, H.S., Lee, C.H., Seo, S. and Kim, S.W., 2018. HSP90 inhibitor, AUY922, debilitates intrinsic and acquired lapatinib-resistant HER2-positive gastric cancer cells. *BMB reports*, 51(12), p.660.
- Pavithra, S.R., Banumathy, G., Joy, O., Singh, V. and Tatu, U., 2004. Recurrent fever promotes *Plasmodium falciparum* development in human erythrocytes. *Journal of Biological Chemistry*, 279(45), pp.46692-46699.
- Pearl, L.H., 2005. Hsp90 and Cdc37—a chaperone cancer conspiracy. *Current opinion in genetics & development*, 15(1), pp.55-61.
- Poli, G., Granchi, C., Rizzolio, F. and Tuccinardi, T., 2020. Application of MM-PBSA methods in virtual screening. *Molecules*, 25(8), p.1971.
- Posfai, D., Eubanks, A.L., Keim, A.I., Lu, K.Y., Wang, G.Z., Hughes, P.F., Kato, N., Haystead, T.A. and Derbyshire, E.R., 2018. Identification of Hsp90 inhibitors with anti-*Plasmodium* activity. *Antimicrobial agents and chemotherapy*, 62(4), pp.e01799-17.
- Pratt, W.B., 1998. The hsp90-based chaperone system: involvement in signal transduction from a variety of hormone and growth factor receptors. *Proceedings of the Society for Experimental Biology and Medicine*, 217(4), pp.420-434.
- Prodromou, C., 2016. Mechanisms of Hsp90 regulation. *Biochemical Journal*, 473(16), pp.2439-2452.
- Pryce, J., Medley, N. and Choi, L., 2022. Indoor residual spraying for preventing malaria in communities using insecticide-treated nets. *Cochrane Database of Systematic Reviews*, (1).
- Reker, D. and Schneider, G., 2015. Active-learning strategies in computer-assisted drug discovery. *Drug discovery today*, 20(4), pp.458-465.
- Rowlands, M., McAndrew, C., Prodromou, C., Pearl, L., Kalusa, A., Jones, K., Workman, P. and Aherne, W., 2010. Detection of the ATPase activity of the molecular chaperones Hsp90 and Hsp72 using the transcreener™ ADP assay kit. *Journal of biomolecular screening*, 15(3), pp.279-286.

- Rowlands, M.G., Newbatt, Y.M., Prodromou, C., Pearl, L.H., Workman, P. and Aherne, W., 2004. High-throughput screening assay for inhibitors of heat-shock protein 90 ATPase activity. *Analytical biochemistry*, 327(2), pp.176-183.
- Rule, C.S., Patrick, M. and Sandkvist, M., 2016. Measuring in vitro ATPase activity for enzymatic characterization. *JoVE (Journal of Visualized Experiments)*, (114), p.e54305.
- Ruwizhi, N., Maseko, R.B. and Aderibigbe, B.A., 2022. Recent Advances in the Therapeutic Efficacy of Artesunate. *Pharmaceutics*, 14(3), p.504.
- S. M. Nie and S. R. Emory, 1996. "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *Science*, vol. 271, pp. 1465-1469.
- Santiago, C.B., Guo, J.Y. and Sigman, M.S., 2018. Predictive and mechanistic multivariate linear regression models for reaction development. *Chemical science*, 9(9), pp.2398-2412.
- Santos, L.H., Ferreira, R.S. and Caffarena, E.R., 2015. Computational drug design strategies applied to the modelling of human immunodeficiency virus-1 reverse transcriptase inhibitors. *Memórias do Instituto Oswaldo Cruz*, 110, pp.847-864.
- Saunders, D.L., Vanachayangkul, P. and Lon, C., 2014. Dihydroartemisinin–piperaquine failure in Cambodia. *New England Journal of Medicine*, 371(5), pp.484-485.
- Schenone, M., Dančík, V., Wagner, B.K. and Clemons, P.A., 2013. Target identification and mechanism of action in chemical biology and drug discovery. *Nature chemical biology*, 9(4), pp.232-240.
- Shahinas, D., MacMullin, G., Benedict, C., Crandall, I. and Pillai, D.R., 2012. Harmine is a potent antimalarial targeting Hsp90 and synergizes with chloroquine and artemisinin. *Antimicrobial agents and chemotherapy*, 56(8), pp.4207-4213.
- Shahinas, D., Folefoc, A., Taldone, T., Chiosis, G., Crandall, I. and Pillai, D.R., 2013. A purine analog synergizes with chloroquine (CQ) by targeting Plasmodium falciparum Hsp90 (PfHsp90). *PLoS One*, 8(9), p.e75446.
- Sharp, P. A. 2002. RNA interference: nature's own RNAi. *Trends in Biochemical Sciences*, 27(4), 166-170.
- Shibeshi, M.A., Kifle, Z.D. and Atnafie, S.A., 2020. Antimalarial drug resistance and novel targets for antimalarial drug discovery. *Infection and Drug Resistance*, pp.4047-4060.
- Shoichet, B.K., 2004. Virtual screening of chemical libraries. *Nature*, 432(7019), pp.862-865.
- Shonhai, A., 2010. Plasmodial heat shock proteins: targets for chemotherapy. *FEMS Immunology & Medical Microbiology*, 58(1), pp.61-74.

- Slater, A.F., 1993. Chloroquine: mechanism of drug action and resistance in *Plasmodium falciparum*. *Pharmacology & therapeutics*, 57(2-3), pp.203-235.
- Smith, A., 2002. Screening for drug discovery: the leading question. *Nature*, 418(6896), pp.453-455.
- Stanton, M.K., Tufekcic, S., Morgan, C. and Bak, A., 2009. Drug substance and former structure property relationships in 15 diverse pharmaceutical co-crystals. *Crystal Growth and Design*, 9(3), pp.1344-1352.
- Steketee, R.W., Nahlen, B.L., Parise, M.E. and Menendez, C., 2001. The burden of malaria in pregnancy in malaria-endemic areas. The Intolerable Burden of Malaria: A New Look at the Numbers: Supplement to Volume 64 (1) of the *American Journal of Tropical Medicine and Hygiene*.
- Stofberg, M.L., Caillet, C., de Villiers, M. and Zininga, T., 2021. Inhibitors of the *Plasmodium falciparum* hsp90 towards selective antimalarial drug design: the past, present and future. *Cells*, 10(11), p.2849. .
- Su, X.Z., Kirkman, L.A., Fujioka, H. and Wellems, T.E., 1997. Complex polymorphisms in an ~ 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell*, 91(5), pp.593-603.
- Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P. and Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. *Journal of chemical information and computer sciences*, 43(6), pp.1947-1958.
- Sweileh, W.M., 2020. Bibliometric analysis of peer-reviewed literature on climate change and human health with an emphasis on infectious diseases. *Globalization and Health*, 16(1), pp.1-17.
- Tice, C.M., 2001. Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals?. *Pest Management Science: formerly Pesticide Science*, 57(1), pp.3-16.
- Tilley, L., Straimer, J., Gnädig, N.F., Ralph, S.A. and Fidock, D.A., 2016. Artemisinin action and resistance in *Plasmodium falciparum*. *Trends in parasitology*, 32(9), pp.682-696.
- Toulmin, A., Wood, J.M. and Kenny, P.W., 2008. Toward prediction of alkane/water partition coefficients. *Journal of medicinal chemistry*, 51(13), pp.3720-3730.
- Tse, E.G., Korsik, M. and Todd, M.H., 2019. The past, present and future of anti-malarial medicines. *Malaria journal*, 18(1), pp.1-21.
- Tu, Z., Stuyver, T. and Coley, C.W., 2023. Predictive chemistry: machine learning for reaction deployment, reaction development, and reaction discovery. *Chemical Science*.

- Uwimana, A., Legrand, E., Stokes, B.H., Ndikumana, J.L.M., Warsame, M., Umulisa, N., Ngamije, D., Munyaneza, T., Mazarati, J.B., Munguti, K. and Campagne, P., 2020. Emergence and clonal expansion of in vitro artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nature medicine*, 26(10), pp.1602-1608.
- van Eijk, A.M., Hill, J., Noor, A.M., Snow, R.W. and ter Kuile, F.O., 2015. Prevalence of malaria infection in pregnant women compared with children for tracking malaria transmission in sub-Saharan Africa: a systematic review and meta-analysis. *The Lancet Global Health*, 3(10), pp.e617-e628.
- Vyas, V.K., Ukawala, R.D., Ghate, M. and Chintha, C., 2012. Homology modeling a fast tool for drug discovery: current perspectives. *Indian journal of pharmaceutical sciences*, 74(1), p.1.
- Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J.Z. and Hou, T., 2019. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chemical reviews*, 119(16), pp.9478-9508.
- Wang, Y., & Gainza, P. 2003. Molecular docking studies of protein–ligand interactions. *Journal of Molecular Graphics and Modelling*, 21(6), pp.397-407.
- Waring, M.J., 2009. Defining optimum lipophilicity and molecular weight ranges for drug candidates—molecular weight dependent lower log D limits based on permeability. *Bioorganic & medicinal chemistry letters*, 19(10), pp.2844-2851.
- Warrell, D.A. and Gilles, H.M., 2017. *Essential malariology*. CRC Press.
- Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Beagley J, Belesova K., 2020 report of The *Lancet* Countdown on health and climate change: responding to converging crises. *Lancet*. 2021;397(10269), pp.129–70.
- Weiss, D.J., Lucas, T.C., Nguyen, M., Nandi, A.K., Bisanzio, D., Battle, K.E., Cameron, E., Twohig, K.A., Pfeffer, D.A., Rozier, J.A. and Gibson, H.S., 2019. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. *The Lancet*, 394(10195), pp.322-331.
- White, N.J., 2004. Antimalarial drug resistance. *The Journal of clinical investigation*, 113(8), pp.1084-1092.
- White, N.J., 2011. Determinants of relapse periodicity in *Plasmodium vivax* malaria. *Malaria journal*, 10(1), pp.1-36.
- Williams, M.A. and Ladbury, J.E., 2003. Hydrogen bonds in protein-ligand complexes. *Protein-ligand interactions: from molecular recognition to drug design*, pp.137-161.

- World Health Organization, 2015. *Guidelines for the treatment of malaria*. World Health Organization.
- Yang, C. Y., Sun, H., Chen, J., Nikolovska-Coleska, Z., and Wang, S., 2009. Importance of Ligand Reorganization Free Energy in Protein-Ligand Binding-Affinity Prediction. *J. Am. Chem. Soc.* 131 (38), 13709–13721.
- Young, D.C., 2009. *Computational drug design: a guide for computational and medicinal chemists*. John Wiley & Sons.
- Zhang, X., Perez-Sanchez, H. and C Lightstone, F., 2017. A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin. *Current topics in medicinal chemistry*, 17(14), pp.1631-1639.
- Zhou, S.F. and Zhong, W.Z., 2017. Drug design and discovery: principles and applications. *Molecules*, 22(2), p.279.
- Zininga T. Doctoral Dissertation. University of Venda; Venda, South Africa: 2015. [(accessed on 19 April 2023)]. Characterization of heat shock protein 70-z (PfHsp70-z) from *Plasmodium falciparum*. Available online: <https://univendspace.univen.ac.za/handle/11602/619>

APPENDIX

Methods and materials

A list of all significant compounds, reagents, and tools utilized in this investigation are included in this chapter, along with a brief description of each procedure performed.

1.1 Materials

1.1.1 List of equipment's

- Benchtop centrifuge – Tomos, multistar 21 from CJ labs (south Africa)
- Shaker – Orbi-Blotter™ Benchmark
- Spectrophotometer – UV-21 ONDA
- Heat block – Digital dry bath, Bio-Rad (Hercules, California, United States)
- Mini-PROTEAN® Tetra system Bio-Rad
- PowerPack™ Basic Bio-Rad
- Vortex mixer, Lasec
- Sonicator – OMNI Sonic Ruptor 400, OMNI international
- Incubator – Scientific
- Centrifuge – Rotofix 32A, Hettich zentrifugen, Labotec
- MiniCentrifuge – WhiteSCI
- NanoDrop™ One/One^C microvolume UV-Vis Spectrophotometer
- pH meter – pH.ISE, HANNA
- ChemiDoc XRS+ Gel imaging system, Bio-Rad
- Glowmax® Explorer multimode microplate reader, from Promega

1.1.2 Kits

- Pierce™ Reversible protein stain kit- fro Nitrocellulose membranes, Thermo Scientific
- PureYield™ Plasmid miniprep system, Promega
- Vivaspin® 2 50kDa MWCO, Cytiva

1.1.3 Reagents

- BenchMark™ Protein ladder, Invitrogen, by Thermo Fisher Scientific
- Blue/Orange 6x DNA loading dye, Promega
- 1KB DNA loading dye, Promega
- Restriction enzymes, Promega

- Primary antibody – 6x-His Tag Monoclonal Antibody, Invitrogen
- Goat anti-Mouse IgG (H+L) secondary antibody
- Pierce™ High capacity Ni-IMAC Resin, EDTA compatible, Thermo scientific
- Pre-Cut Nitrocellulose membranes, Thermo scientific
- N,N,N',N'-Tetramethyl-ethylenediamine (TEMED) – BioReagent, stable for electrophoresis, SIGMA.
- Commassie plus – the better Bradford assay™ reagent, Thermo scientific.

30% Acrylamide/Bis Solution 29:1, Bio-Rad.

1.1.4 Plasmid

	pET28b(+)-PfHsp90-NTD
Insert	PfHsp90-NTD
Resistance	Kanamycin
Length	6020bp
Tag	His-tag
Expression level	High

1.1.5 Transformation

To transform the E.coli cells, 1µl of the plasmid DNA was added into 50µl of the competent cells in a sterile 1.5ml Eppendorf tube and then mixed by flicking the tube 3 times. The tube was then incubated on ice for 30 minutes and subsequently incubated in a heat block instrument for 20 seconds to induce heat Shock, and then returned on ice for 2 minutes. 450µl of 2x YT broth was added to the mixture in the tube and incubated at 37°C for an hour. Since both our plasmids have kanamycin resistance 50ml of the sample was plated on to 2x YT agar plate containing kanamycin (50mg/ml) and incubated the agar plate at 37°C overnight.

1.1.6 Pure-yield plasmid extraction kit

Perform the following procedure at room temperature.

1. Transfer 600µl of bacterial culture grown in LB medium to a 1.5ml microcentrifuge tube.

Note: If you wish to process larger volumes of bacterial culture (up to 3.0ml), use the protocol provided in Section 4.C.

2. Add 100µl of Cell Lysis Buffer, and mix by inverting the tube 6 times.

The solution should change from opaque to clear blue, indicating complete lysis.

Note: Proceed to Step 3 within 2 minutes. Excessive lysis can result in denatured plasmid DNA. If

processing a large number of samples, process samples in groups of ten or less. Continue with the next set of

ten samples after the first set has been neutralized and mixed thoroughly.

3. Add 350µl of cold (4–8°C) Neutralization Solution, and mix thoroughly by inverting the tube.

The sample will turn yellow when neutralization is complete, and a yellow precipitate will form.

Invert the

sample an additional 3 times to ensure complete neutralization.

4. Centrifuge at maximum speed in a microcentrifuge for 3 minutes.

5. Transfer the supernatant (~900µl) to a PureYield™ Minicolumn.

Do not disturb the cell debris pellet. For maximum yield, transfer the supernatant with a pipette.

6. Place the minicolumn into a PureYield™ Collection Tube, and centrifuge at maximum speed in a microcentrifuge

for 15 seconds.

7. Discard the flowthrough, and place the minicolumn into the same PureYield™ Collection Tube.

8. Add 200µl of Endotoxin Removal Wash to the minicolumn. Centrifuge at maximum speed in a microcentrifuge

for 15 seconds. It is not necessary to empty the PureYield™ Collection Tube.

9. Add 400µl of Column Wash Solution to the minicolumn. Centrifuge at maximum speed in a microcentrifuge for

30 seconds.

10. Transfer the minicolumn to a clean 1.5ml microcentrifuge tube, then add 30µl of Elution Buffer directly to the

minicolumn matrix. Let stand for 1 minute at room temperature.

Notes:

1. Nuclease-free water at neutral pH can also be used to elute DNA.

2. For large plasmids (>10kb), warm the Elution Buffer to 50°C prior to elution, and increase elution volume

to 50µl. Also incubate the column at room temperature (22–25°C) for 5–10 minutes before proceeding to

Step 11.

11. Centrifuge at maximum speed in a microcentrifuge for 15 seconds to elute the plasmid DNA. Cap the

microcentrifuge tube, and store eluted plasmid DNA at -20°C .

1.1.7 Restriction digest

The restriction enzyme digest was carried out in a volume of 20 μl reaction, with 200-500ng/ μl of the plasmid DNA and 1 μl of restriction enzyme for a single digest. For double digest 1 μl of each restriction enzyme, making it 2 μl of restriction enzymes. For the plasmid Pet28b(+)-PfHsp90-NTD the restriction enzymes were BamHI and indel while for the Pet28a(+)-HsHsp90-NTD the HindIII and XhoI restriction enzymes were used.

In a sterile 1.5ml Eppendorf tube the components were assembled in the order listed in the table 3 for Pet28a(+)-HsHsp90-NTD and table 4 for Pet28b(+)-PfHsp90-NTD.

Table A8: pET28a(+)-HsHsp90-NTD restriction enzyme digest components.

components	Single digest volume	Double digest volume
Sterile H ₂ O	13.3 μl	12.3 μl
10x buffer restriction enzyme	2 μl	2 μl
Acetylated BSA, 10mg/ml	0,2 μl	0.2 μl
Plasmid DNA mg/ml	4,0 μl	4.0 μl
Mixed by pipetting up and down Then added:		
Restriction enzyme	1 μl of HindIII or xhoI	1 μl of HinIII and 1 μl of XhoI= 2 μl
Find volume	20 μl	20 μl

Table A9: pET28b(+)-PfHsp90-NTD restriction enzyme digest components.

components	Single digest volume	Double digest volume
Sterile H ₂ O	13.3 μl	12.3 μl
10x buffer restriction enzyme	2 μl	2 μl
Acetylated BSA, 10mg/ml	0,2 μl	0.2 μl
Plasmid DNA mg/ml	4,0 μl	4.0 μl
Mixed by pipetting up and down Then added:		
Restriction enzyme	1 μl of BamHI or indel	1 μl of BamHI and 1 μl of indel=2 μl
Find volume	20 μl	20 μl

The components were gently mixed by pipetting up and down and centrifuged of 15 seconds. The reaction was then incubated at 37°C for 2 hours. Finally the reaction was stopped by adding 10x loading dye (10 µl). For experimental control the uncut plasmid were used as control

1.1.8 Agarose gel electrophoresis

To confirm the integrity of our plasmids subsequent to extraction and restriction enzyme digest, agarose gel electrophoresis was performed.

For the the preparation of 1%agarose gel, 0.5g of Seakem LE agarose was dissolved in 50ml of 1x TAE buffer(40mM Tris,20mM acetic acid and 1mM EDTA). The mixture was heated in a microwave until it became clear and then left to cool down before the addition of ethidium bromide. The gel was then poured into a casting tray, inserted combs and left to solidify. Once solidified the combs were removed and the gel placed in geltank containing TAE buffer. The gel was ran at 100V for 45 minutes. To visualise the DNA and capture images a Bio-Rad ChemiDoc XRS+ Gel Imaging System was used.

1.1.9 Sodium dodecyl-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a technique which works by the principles that charged molecule will migrate to an electronic possessing an opposite courage in an electric field, with charged smaller molecules migrating faster. Therefore, SDS-PAGE separates protein based on their molecule weight.

In this study since our proteins of interest (PfHsp90-NTD and HsHsp90-NTD) are 25KDa, a 12% revolving gel was used. The glass plates and spacers were assembled in the Bio-Rad Mini-protean casting apparatus, the components of the resolving gel and stacking where mixed according to table 5 and 6 respectively.

Table A10: resolving gel components

Resolving gel (RG) (12%)	1 gel
H ₂ O	1,4ml
RG buffer (1,5M Tris-HCl pH 8,8)	1,3ml
30% Acrylamide	1,2ml
20% SDS	100ml
10% Ammonium persulphate	10ml

TEMED	4ml
-------	-----

Stacking gel (SG)	2 gels
H ₂ O	3,9ml
SG buffer (1M Tris-HCl pH 6.8)	500µl
30% Acrylamide	500µl
20% SDS	100µl
10% Amonium persulphate	16µl
TEMED	8µl

The resolving gel was firstly prepared and poured into the glass plates to a level 2cm below the top and left to solidify for 30 minutes before preparing and adding the stacking gel. Following the addition of the stacking gel combs were inserted and the stacking was left to solidify for 1 hour at room temperature.

For sample preparation 20µl of the 4x sample buffer (125mM Tris-HCl, pH 6.8, 20% [v/v] Glycerol, 4% [w/v] SD, 10% [v/v] β-mecaptoethanol, 0.5mg/ml bromophenol blue) was added to 100µl of sample in a sterile 1.5ml Eppendorf tube and incubated at 100°C for 5 minutes, before centrifuging at 12 000xg for 30 seconds.

To run the gel the combs were firstly removed from the gel and the cast gel was assembled into a Bio-Rad mini-protein gel chamber containing 1x running buffer. 10µl of the protein ladder was loaded onto the gel and then added the prepared samples from well 2. The gel was ran at 100V until the dye front migrated into the running gel (~15 minutes), and Increased to 200 V until the dye front reached the bottom of the gel (~45 min).

The gel was then removed from the glass plates sand placed inside a small container, containing comassie blue staining solution (45% [v/v] water, 45% [v/v] methanol, 10% [v/v] acetic acid and 0.5 g/L Coomassie blue R-250) making sure to cover the entire gel and left overnight on a shaker. The following day the gel was de-stained using a de-staining solution (10% [v/v] acetic acid, 45% [v/v] methanol, 45%[v/v] water) until the gel became clear to visualise and capture images using the Bio-Rad ChemiDoc XRS+ Gel Imaging System.

1.1.10 Western blotting

The samples were firstly ran on an SDS-PAGE, the gels were then removed from the glass plates packed into a sandwich with a nitrocellulose membrane, being careful not to trap any air bubbles inside to ensure maximum and even protein transfer. The sandwich was inserted into a chamber filled with cold Western blot transfer buffer (25Mm Tris-HCl, 2.5M glycine and 20% [v/v] methanol).the transfer was ran at 100V for 90 minutes. To detect successful protein transfer onto the nitrocellulose membrane, the MemCode™ reversible protein stain kit was used according to the manufacturer's instructions. To prevent nonspecific antibody reactions the membrane was blocked with 5% [w/v] skim milk in 100ml of TBS (50mM Tris, 150mM NaCl) overnight. The blocking solution was then washed with distilled water and added the primary antibody in a 1:1000 ration in the blocking solution and incubated the membrane in it for 1 hour. The membrane was then washed 3 times with TBS-T (0.1%) Tween 20 in TBS. The membrane was incubated for 1 hour with secondary antibody in the blocking solution. Finally the membrane was washed with TBS-T and visualised using Bio-Rad ChemiDoc XRS+ Gel Imaging System.

RESULTS

2.1 Induce fit docking

Table A11: induced fit docking results of the top 62 compounds against the PfHsp90 N-terminal domain receptor

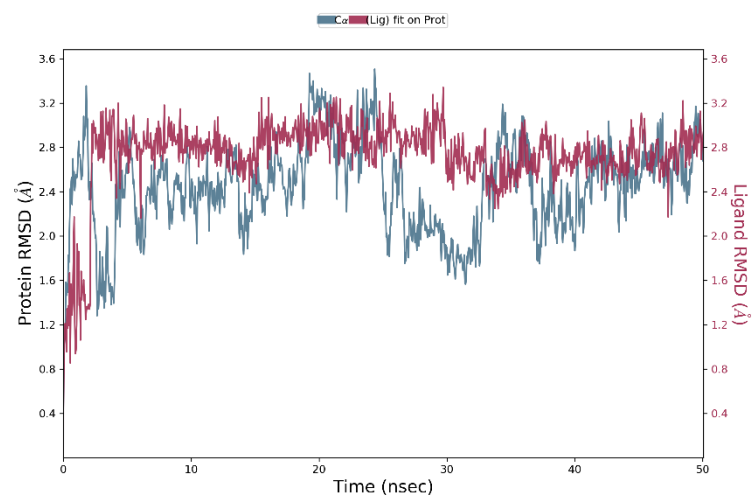
Title	docking score	glide gscore	glide emodel	XP GScore	IFDScore
FTN40281232	-12.032	-12.058	-106.600	-12.058	-526.97
FTN 2	-12.351	-12.351	-126.680	-12.351	-526.86
FTN315699906	-10.501	-10.746	-113.012	-10.746	-526.14
FTN315452139	-10.396	-10.396	-117.889	-10.396	-525.99
FTN48606212	-10.796	-10.796	-92.753	-10.796	-524.35
FTN 3	-10.380	-10.457	-101.975	-10.457	-523.50
FTN5	-11.216	-11.216	-103.704	-11.216	-523.34
FTN39249872	-10.178	-10.431	-65.285	-10.431	-523.26
FTN41168581	-11.209	-11.224	-84.635	-11.224	-522.53
FTN313761635	-10.603	-10.611	-86.084	-10.611	-520.29
FTN4919	-13.161	-13.162	-108.122	-13.162	-533.02
FTN717	-12.399	-12.435	-85.494	-12.435	-528.05
FTN1424	-11.975	-11.976	-107.794	-11.976	-526.61
FTN1841	-11.396	-11.885	-105.884	-11.885	-526.20
FTN1424	-12.446	-12.447	-101.188	-12.447	-526.03
FTN 1	-14.746	-14.794	-120.018	-14.794	-525.66
FTN3967	-11.880	-12.442	-111.528	-12.442	-525.22
FTN4079	-13.233	-14.746	-116.205	-14.746	-524.84
FTN1133	-11.006	-11.171	-107.716	-11.171	-524.79
FTN 4	-14.313	-14.313	-105.919	-14.313	-524.73

FTN1133	-10.175	-11.013	-106.508	-11.013	-524.43
FTN1841	-11.894	-12.235	-86.219	-12.235	-524.04
FTN1715	-10.790	-13.133	-102.921	-13.133	-523.90
FTN3630	-13.914	-13.914	-84.466	-13.914	-523.58
FTN1813	-11.941	-12.423	-101.299	-12.423	-523.44
FTN4994	-10.301	-10.301	-101.453	-10.301	-523.43
FTN3885	-13.873	-13.874	-100.826	-13.874	-522.27
FTN4184	-10.790	-10.955	-102.957	-10.955	-522.18
FTN4363	-12.400	-12.400	-90.507	-12.400	-522.02
FTN3782	-10.461	-10.461	-103.342	-10.461	-522.00
FTN1261	-12.330	-12.332	-98.280	-12.332	-521.62
FTN3885	-12.396	-12.397	-105.667	-12.397	-521.51
FTN499	-12.397	-12.759	-89.158	-12.759	-521.49
FTN4363	-12.281	-12.281	-92.724	-12.281	-521.47
FTN1715	-13.770	-13.781	-112.874	-13.781	-521.33
FTN4843	-11.544	-11.544	-85.764	-11.544	-521.27
FTN499	-11.470	-11.933	-115.374	-11.933	-521.21
FTN3885	-11.725	-11.726	-105.179	-11.726	-521.15
FTN3885	-11.995	-11.996	-108.696	-11.996	-521.07
FTN2075	-11.412	-11.983	-117.573	-11.983	-520.93
FTN2925	-10.489	-11.022	-119.488	-11.022	-520.92
FTN3057	-11.430	-11.430	-78.680	-11.430	-520.64
FTN2701	-10.676	-10.774	-104.290	-10.774	-520.06
FTN4031	-10.535	-10.535	-91.366	-10.535	-519.80
FTN2474	-12.242	-12.242	-89.123	-12.242	-519.66
FTN0	-11.056	-11.056	-77.271	-11.056	-519.33
FTN2110	-12.134	-12.153	-76.051	-12.153	-519.33
FTN1589	-10.151	-10.151	-79.481	-10.151	-518.91
FTN2211	-10.254	-10.264	-79.071	-10.264	-518.87

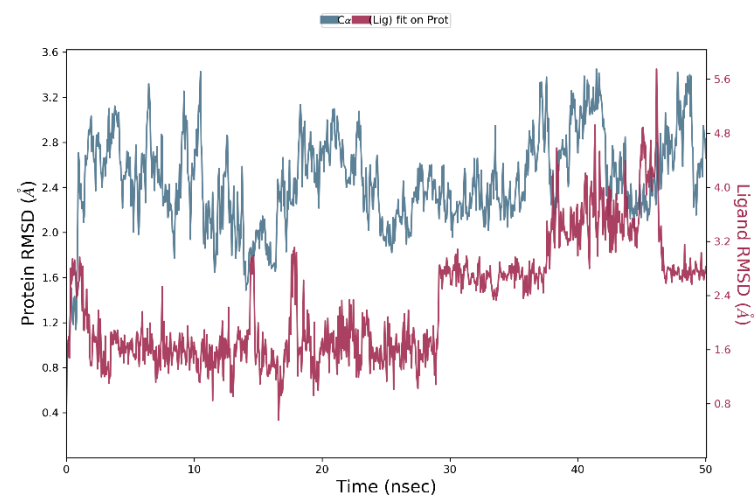
FTN1388	-10.186	-10.188	-73.541	-10.188	-517.80
FTN1454	-10.579	-11.025	-80.285	-11.025	-517.77
FTN3550	-11.025	-11.025	-52.640	-11.025	-517.74
FTN3550	-11.156	-11.156	-75.256	-11.156	-517.20
FTN3967	-11.988	-12.278	-89.206	-12.278	-516.97
FTN3885	-14.593	-14.594	-96.499	-14.594	-516.89
FTN3550	-10.190	-10.190	-54.918	-10.190	-516.65
FTN1956	-12.433	-12.433	-56.229	-12.433	-516.10
FTN4541	-10.567	-10.567	-82.188	-10.567	-515.48
FTN2403	-14.162	-14.162	-100.693	-14.162	-515.14
FTN3816	-12.238	-12.238	-71.732	-12.238	-514.82
FTN95	-12.502	-12.502	-88.717	-12.502	-514.19
FTN 6	-14.176	-14.176	-77.866	-14.176	-512.26

2.2 Molecular dynamic simulations

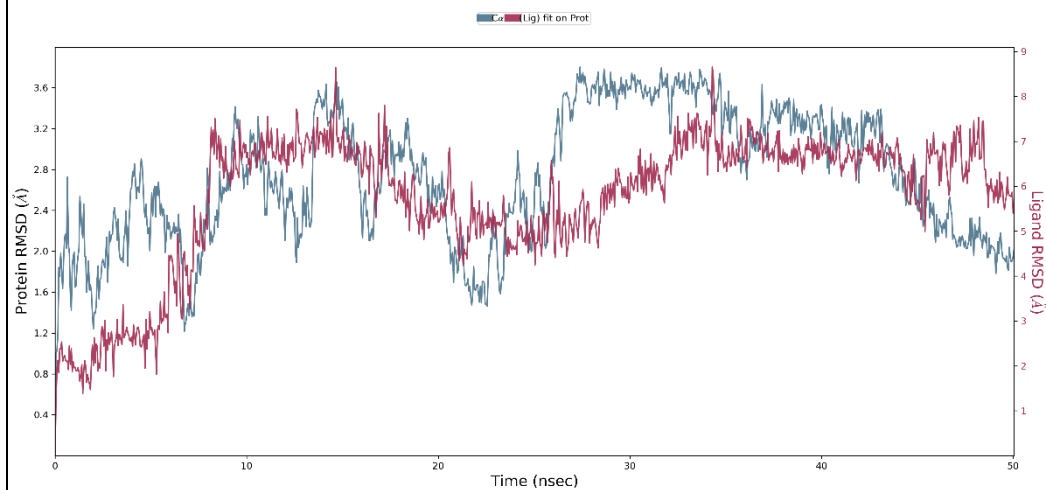
FTN-T1



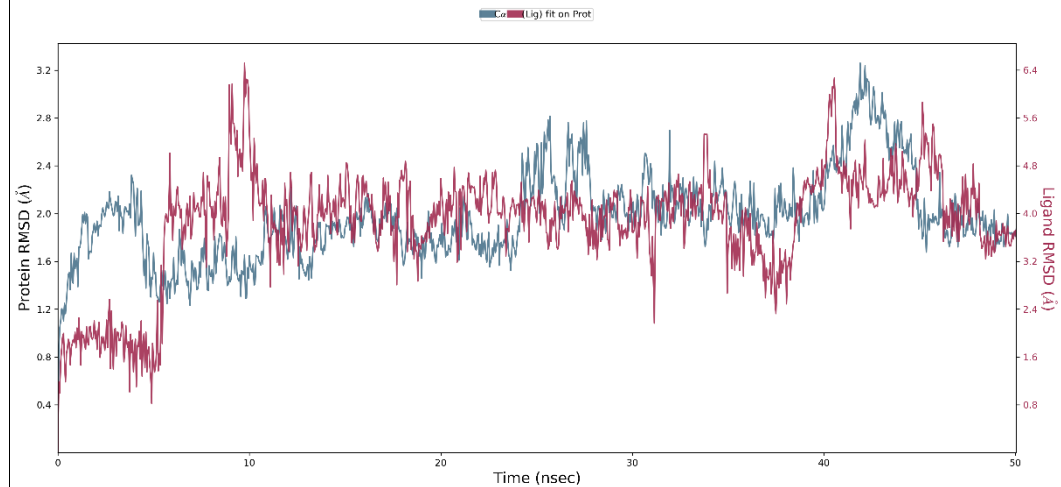
FTN-T2



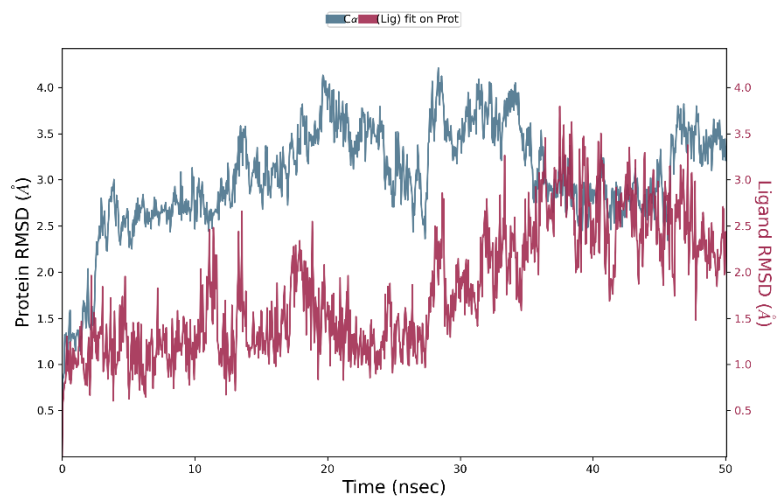
FTN-T4



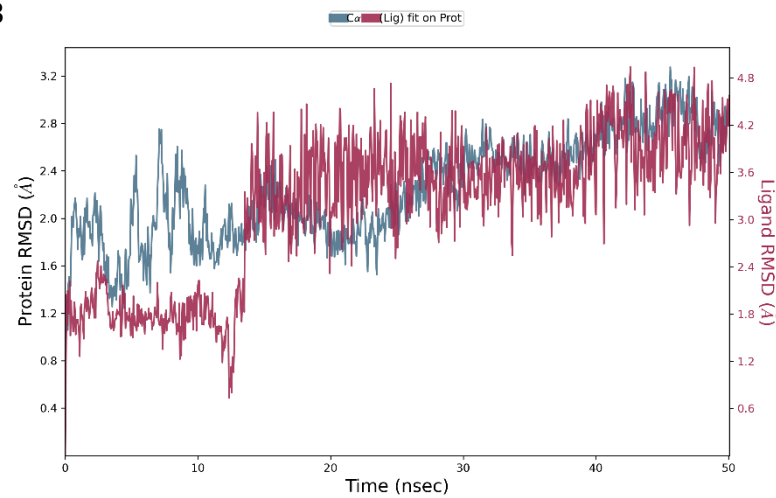
FTN-T5



FTN-M2



FTN-M3



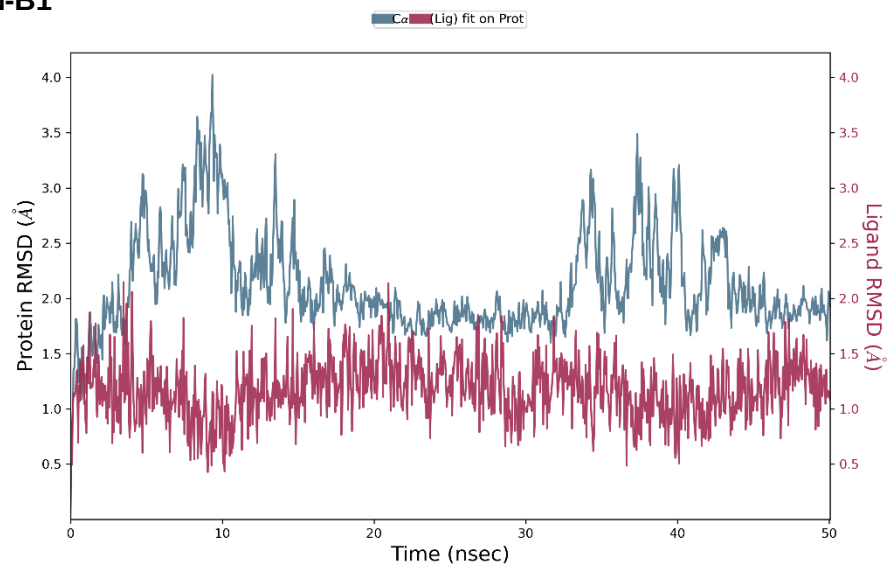
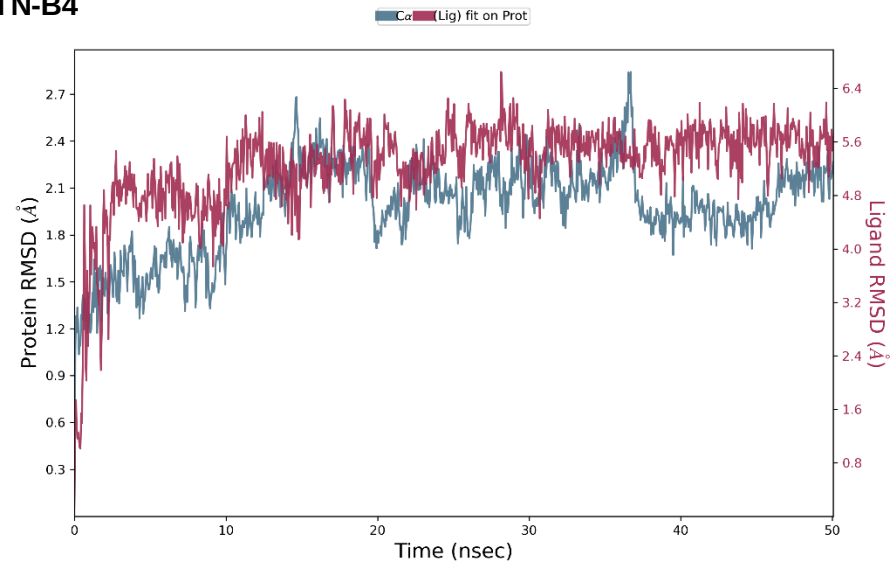
FTN-B1**FTN-B4**

Figure A23: Protein-ligand RMSD plots of all the frames in the trajectory for human Hsp90-NTD and Compounds FTN-T1,FTN-T4, FTN-T5, FTN-M2, FTN-M3, FTN-B1 and FTN-B4 complexes at a 50ns MDS trajectory and FTN-T2 at 150ns. Red plots indicate the ligand RMSD values and Blue plots indicates the protein RMSD.