

https://dspace.org/

The ligandin activity of Schistosoma 26-kDa and 28-kDa glutathione transferases towards 17-Hydroxyandrost-4-ene-3-one from a biophysical perspective

Makumbe, Hattie Hope

2023
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42555
Makumbe, Hattie Hope. (2023). The ligandin activity of Schistosoma 26-kDa and 28-kDa glutathione transferases towards 17-Hydroxyandrost-4-ene-3-one from a biophysical perspective. [PhD thesis, University of the Witwatersrand, Johannesburg]. https://hdl.handle.net/10539/42555

Downloaded from a DSpace repository. https://dspace.org/



The ligandin activity of Schistosoma 26-kDa and 28-kDa glutathione transferases towards 17β-Hydroxyandrost-4-ene-3-one from a biophysical perspective

by

Hattie Hope Makumbe 2418629

Thesis

Submitted in fulfilment of the requirements for the degree

for

Doctor of Philosophy degree

in

Molecular and Cell Biology

In the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Dr. Ikechukwu Achilonu

27 October 2023

Declaration

I, Hattie Hope Makumbe (2418629), am a student registered for the degree of Doctor of Philosophy (PhD) in the academic year 2023.

I hereby declare the following:

✓✓I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.

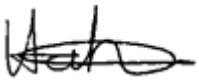
✓✓I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.

✓✓I have not submitted this work before for any other degree or examination at this, or any other University.

✓✓The information used in the Thesis has not been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.

✓✓I have followed the required conventions in referencing the thoughts and ideas of others.

✓✓ I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature:  **Date:** 27 October 2023

Dedication

To the most-high God who bypasses procedures

A vision don’t bleed !!! Here it is.

He raiseth the poor out of the dunghill to set them as princes.

This is the Lord's doing; it is marvelous in our eyes.

Abstract

Schistosomiasis, caused by helminth worms, ranks second amongst parasitic diseases and accounts for over 220 million fatalities globally. Statistics show that in South Africa, schistosomiasis (bilharzia) has infected approximately 4 million individuals. Currently, there are parasite resistance challenges with the sole available remedy. The World Health Organisation (WHO) acknowledges the need for new effective drugs. The 26-kDa *Schistosoma bovis/haematobium* (Sbh26GST) and 28-kDa *Schistosoma haematobium* (Sh28GST) are parasite Glutathione S-transferases (GSTs) which consist of two identical subunits that perform a vital role in mitigating the adverse effects of harmful electrophilic substances within the parasite since the parasite is devoid of the neutralizing cytochrome P-450. This automatically renders these parasite GSTs as potential therapeutic targets for schistosomiasis. Testosterone, the major hormone responsible for sexual characteristics and growth in males, can be repurposed as a drug target against schistosomiasis. In this study, we examined the structural, stability and functional interactions between the parasite GSTs and testosterone. After confirmation of inhibition, IC_{50} experiments were performed. The enzymes were overexpressed in *Escherichia coli* (*E.coli*) and then purified through a single-step nickel ion-immobilized metal affinity chromatography (IMAC). Extrinsic fluorescence spectroscopy was also done to provide evidence for the binding of the recombinant GSTs with testosterone. The GST activity was measured by employing 1-chloro-2,4-dinitrobenzene (CDNB) as the substrate. Additionally, we investigated if the enzyme activity was influenced by the presence of testosterone. To analyse the stability of the enzymes, a SYPRO Orange-based thermal shift assay was used in the presence and absence of testosterone. In addition to empirical investigations, computational modelling, molecular docking, and molecular dynamic simulations were used to provide complementary insights to show binding affinities, prediction of binding modes and stability of the GST-testosterone complex. The secondary structural composition was found to be predominantly alpha-helical. Insights into tertiary structure analysis revealed the presence of buried solvent exposed tryptophan residues. The findings from spectroscopy with 8-anilino-1-naphthalenesulfonate (ANS) indicated that both Human GST-mu and parasite GSTs bound to ANS. Enzyme kinetic studies show that testosterone is a potent inhibitor of the parasite GSTs, with a specific activity that decreases from 16 μmol $min^{-1}mg^{-1}$ to 0.03 μmol $min^{-1}mg^{-1}$ and IC_{50} in the nanomolar range of 20 μM for Sh28GST. Sbh26GST exhibited a specific activity that decreased from 20 μmol $min^{-1}mg^{-1}$ to 0.14 μmol $min^{-1}mg^{-1}$, and a testosterone IC_{50} of 23 μM. The thermal stability assay confirmed Sh28GST to be more stable than Sbh26GST, and this stability of Sh28GST intensified when the enzyme bound to testosterone and GSH. Steady state kinetics towards glutathione (GSH) revealed a K_m of 4.2mM and 6.6 mM for Sh28GST and Sbh26GST respectively. The present study has practical implications for novel application of the enzymes to serve as a basis for future studies aimed at development of inhibitors with potential therapeutic benefits through rational drug design.

Acknowledgements

To the Most High, one and only God, the creator of the universe who has deemed it fit for this degree to be. Thank you, Lord, for your unending mercies. We have experienced the supernatural in our day. Handina kumbobvira ndarota zvamanga makandirongera Ishe wangu.

Thank you to my supervisor, Dr. I. Achilonu, for your continuous support and guidance throughout my degree. Your enthusiasm for science was always inspirational and facilitated my curiosity for wanting to understand the fundamental principles of science.

Thank you to the NRF for your financial support throughout my studies.

Heartfelt thanks to my mother who was a wellspring of emotional and spiritual support throughout the doctoral studies. Your unfeigned love and faith kept me going throughout my studies.

My sincere gratitude goes to my king, my ruler, my better half, my life companion. Thank you, Marrion my God-given partner who initiated and spearheaded the Doctorate studies. Thanks, honey, for believing in me.

Listening to the message of the hour whilst writing this script gave me exceptional strength and kept me going.

Patience my God given adopted sister, only God knows how to repay you for being used in such a tremendous way.

Thank you, Dad, Michael, Solomon and Pastor Hannes, for your support. The prayers made it possible.

Contents

Dedication ..... **iii**
Abstract ..... **iv**
Acknowledgements ..... **v**
Contents ..... **vi**
Table of figures ..... **x**
List of Tables ..... **xii**
Abbreviations ..... **xiii**
Chapter 1 Introduction **15**
1.1 Problem statement 15
1.2 Rationale 5
1.3 Novelty 6
1.4 Aim and Objectives 9
1.5 Outline of the thesis 10
Chapter 2 Literature Review **11**
2.1 Brief background on schistosomiasis 11
2.2 Schistosoma classification 14
2.1.1 Schistosomiasis disease burden 15
2.1.2 Schistosoma haematobium life cycle 17
2.1.3 Disease prevalence in Southern Africa 17
2.2 Treatment of schistosomiasis 19
2.3 Structure and function of GST 20
2.3.1 Classification and structure 21
2.3.2 Catalytic ability 22
2.3.3 GST as a potential druggable enzyme in Schistosoma 22
2.4 Testosterone as a therapeutic agent 23
2.5 Computational Modelling 26
Chapter 3 Expression, purification and basiccharacterization of Sbh26GST-, Sh28GST- and

Human GST-mu- **27**
3.1 Introduction 27
3.2 Materials 28
3.3 Methods 29
3.3.1 Expression of the GST proteins 29
3.3.2 Purification of parasite and Human GST-mu proteins 29
3.3.3 Qualitative analysis of the purified recombinant GSTs 30
3.3.4 Secondary structure analysis of recombinant GSTs 31
3.3.4.1 Far UV circular dichroism (CD) spectroscopy 31
Equation 3.5 32
3.3.5 Tertiary structure analysis of recombinant GSTs 32
3.3.5.1 Intrinsic tryptophan fluorescence spectroscopy 32
3.3.5.2 Extrinsic ANS-fluorescence spectropolarimetry 33
3.3.5.3 Extrinsic fluorescence: ANS binding studies 34
3.4 Results 34
3.4.3 Expression of the GST proteins 34
3.4.4 Purification of the GST proteins 35
3.4.5 Secondary structural analysis of the Human GST-mu and parasite GSTs 37
3.4.6 Tertiary structure analysis of the parasite GSTs and Human GST-mu 39
3.5 Discussion 41
Chapter 4 ..... **43**
Effect of testosterone on the functionand stability of GSTs 43
4.1 Introduction 43
4.1.1 Thermal shift assay 43
Equation 4.2 44
4.1.2 Specific activity of the parasite and Human GST-mu and IC_{50} analysis of testosterone 44
4.2 Method 44
4.2.1 Thermal shift assay of the parasite and Human GST 44
4.2.2 Steady state kinetics of the parasite and Human GST 45
4.2.2.1 Specific activity and IC50 determination of the parasite and Human GST 45
4.3 Results 46
4.3.1 Thermal shift assay 46

4.3.2 Steady state kinetics 46
4.3.3 Specific activity and IC_{50} determination of the parasite and Human GST-mu 47
4.4 Discussion 52
Chapter 5 Computational modelling of 53
GST-mu 53
Abstract 53
5.1 Introduction 53
5.1.1 Computational protein structure prediction 54
5.1.2 Protein-ligand interaction computational simulation 56
5.1.3 Induced-fit molecular docking (IFMD) simulations 59
5.1.4 Molecular dynamics (MD) simulation 60
5.1.5 Ligand motion 60
5.2 Methods 60
5.2.1 Induced-fit model docking simulations 60
5.2.2 Molecular dynamics simulations 61
5.3 Results 62
5.4 Discussion 70
5.4.2 Generation and validation of Sbh26GST model 70
5.4.3 Simulation of Testosterone docking on GSTs 71
5.4.4 Molecular dynamics of testosterone-GST binding 72
Chapter 6 74
Abstract 74
6.1 Introduction 75
6.2 Materials 78
6.3 Methods 78
6.3.1 Protein expression and purification 78
6.3.2 Protein characterization 79
Equation 6.2 81
6.3.3 Molecular modelling studies 81
6.4 Results 83
6.4.1 Expression and purification 83
6.4.2 Sh28GST enzyme activity assay 84

6.4.3 Far-UV CD 84
Equation 6.3 84
6.4.4 Fluorescence spectroscopy 86
6.4.5 X-ray crystallography 86
6.4.6 B-factors analysis 96
6.4.7 Cavity analysis 96
6.4.8 Molecular dynamic simulation 97
6.5. Discussion 98
Chapter 7 103
General discussion, conclusion andfuture work 103
References 108

Table of figures

Figure 1.1 Sub-Saharan African (SSA) map showing the predicted burden of schistosomiasis 3
Figure 1. 2 Prevalence of schistosomiasis in South Africa from 2011 to 2018 7

Figure 2. 1 World distribution of schistosomiasis. 11
Figure 2.2 Inter relationship between bioinformatics and machine learning 13
Figure 2. 3 A framework for schistosomiasis-targeted interventions. 15
Figure 2. 4 Schistosomiasis life cycle. 17
Figure 2. 5 The process of GST detoxification. 20
Figure 2. 6 Structure of a typical GST 21
Figure 2.7 Schematic of the formation of human steroid hormones from cholesterol 24
Figure 2.8 Steroid action in cells. 25

Figure 3.1 Structure of ANS dye 33
Figure 3.2 SDS-PAGE analysis of the expression and purification profiles of the GSTs. 35
Figure 3.3 Specific activity profiles of parasite and Human GST-mu 36
Figure 3.4 Secondary structural analysis of the GSTs 38
Figure 3.5 Intrinsic fluorescence emission spectra of the GSTs 39
Figure 3.6 ANS spectroscopy. 40

Figure 4.1 The IC_{50} values of testosterone 49
Figure 4.2 Steady state kinetic velocity plot of Sbh26GST and Sh28GST versus GSH concentration in the presence and absence of testosterone 49
Figure 4.3 Steady state kinetic velocity plot of Sbh26GST and Sh28GST versus CDNB concentration in the presence and absence of testosterone 50
Figure 4.4. Thermal stability profile for Human GST-mu, Sh28GST and Sbh26GST in the presence and absence of GSH and testosterone 51

Figure 5.1 Basic principles of the 3 experimental methods of protein structure determination 55
Figure 5.2 The 3 approaches to prediction of protein structure 56
Figure 5.3 Machine learning methods workflow for protein-ligand interaction prediction. 59
Figure 5.4 Schematic of the Induced Fit Dock (IFD) and Average MD Structure of the interaction between testosterone and the GSTs 63
Figure 5.5 Motions of testosterone 64
Figure 5.6 Schematics of testosterone-GST contact and RMSF values for amino acid residues 65
Figure 5.7 Bar graphs of the proportion of interaction of each of the interacting residues 66
Figure 5.8 Plots of the RMSD and radius of gyration. 67
Figure 5.9 Plots of C^{α}-RMSF relative to the respective apo targets. 68
Figure 5.10 SASA of testosterone 69

Figure 6.1 A 12.5% (w/v) polyacrylamide gel of the non-liganded Sh28GST..... 83
Figure 6.2 Structural characterization studies for Sh28GST..... 85
Figure 6.3 Sh28GST crystals 87
Figure 6.4 Overall Sh28GST structure..... 87
Figure 6.5 Cross section of two structures of the 28-kDa GST dimer 89
Figure 6.6 Molecular surface of the free and bound Sh28GST..... 90
Figure 6.7 Cavities in the A bound and B free Sh28GST 91
Figure 6.8 The crystal contacts made at the crystal-packing interface..... 92
Figure 6.9 Sequence analysis of 8ALS and 1OE8 93
Figure 6.9 Post dynamic analysis of the molecular dynamic simulation of 8ALS and 1OE8. 94

List of Tables

Table 4.1 Kinetic parameters for Sbh26GST and Sh28GST 49

Table 6.1 Data collection and refinement statistics 96
Table 6.2 Cavity analysis of the 28-kDa GST structures 97

Abbreviations

A_{260}	absorbance at 260 nm
A_{280}	absorbance at 280 nm
A_{340}	absorbance at 340 nm
ANS	8-Anilino-1-naphthalene sulfonate
BLASTp	Basic Local Alignment Search Tool for proteins
CD	Circular dichroism
CDNB	1-Chloro-2,4-dinitrobenzene
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
G-site	Glutathione binding site
GSH	Glutathione
GST	Glutathione transferase
GTX	S-hexylglutathione
hGSTμ	Human μ class glutathione transferase
hGSTP1	Human pi class glutathione transferase
H-site	Electrophilic substrate binding site
IC_{50}	Half-maximal inhibitory concentration
IFD	Induced fit docking
IPTG	Isopropyl β-D-1-thiogalactopyranoside

ITC	Isothermal titration calorimetry
MD	Molecular docking and molecular dynamics
MDS	Molecular dynamic simulations
MRW	Mean residue weight
MWM	Molecular weight
OD	Optical density
ORF	open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PEG	polyethylene glycol
PDB	Protein Data Bank
PMSF	phenylmethylsulphonyl fluoride
RMSD	root-mean-square standard deviation
RMSF	root-mean-square fluctuations
Rg	Radius of gyration
SSA	Sub-Saharan Africa
SASA	Solvent accessible surface area
TEMED	tetramethylethylenediamine
TCEP	tris (2-carboxyethyl) phosphine-hydrogen chloride
µM	Micromolar

Chapter 1
Introduction

1.1 Problem statement

Schistosomiasis, (snail fever), is a persistent and crippling disease caused by a parasite helminth. The disease commonly manifests as fever and abdominal pains although the symptoms are largely dependent on the type of worm involved (Nour, 2010). This disease plagues over 200 million people globally (WHO, 2020) with approximately 800 million people exposed to the disease in 78 countries (Ogongo *et al.,* 2022). Life-threatening conditions are experienced in approximately 10% of the *Schistosoma* sufferers (Johnson *et al.,* 2003). Notwithstanding, the diagnosis of schistosomiasis cases is often over- looked, resulting in under reported figures. It is hypothesized that there is a 1:1 ratio in respect of an egg- negative and egg-positive infected individual, thus extrapolating the real estimates of *Schistosoma* sufferers to be higher (around 600 million) than the 200 million mark pegged by the WHO (Rashika *et al.,* 2013).

The global burden of schistosomiasis far outweighs that of malaria and tuberculosis (King and Dangerfield, 2008). Furthermore, an estimated 3.8 million disability-adjusted life-years are attributed to the disease (Toor *et al.,* 2020). The schistosomiasis burden is mostly experienced on the African continent, with the school-going age group being the most vulnerable (Saathof *et al.,* 2004). It is believed that more than ninety percent of schistosomiasis cases are found in sub-Saharan Africa (Adenowo *et al.,* 2015). Tragically, it is believed that in Africa alone,
over 500 000 fatalities per annum occur due to this debilitating disease (Adenowo *et al.,* 2015). A research investigation carried out in South Africa revealed that there were approximately 4.5 million schistosomiasis cases per year, and more than 25.7 million people were at risk of exposure to the pathogen (Wolmarans and De Kock, 2009). A more recent study (2011 to 2018), however, revealed that the average of microscopically confirmed schistosomiasis cases was 36 people per 100 000 per year. Males who were aged from 5-19 years mostly suffered with the disease across the 4 provinces (Figure 1.1) (De Boni *et al.,* 2021).

Although this overlooked parasitic disease exerts tremendous pressure on the socio-economic and healthcare arena, (as livelihoods and governments grapple under the negative impact of the incapacitating disease), only 10% of the health care budget is allocated towards research and development of potential schistosomiasis drugs (Commission on Health Research for Development, 1990). While there remains a substantial research gap for neglected diseases today, there have been positive shifts in both the health research landscape and the allocation of the global disease burden since 1990. According to G-Finder, (the primary source of neglected disease R&D funding data for the WHO Global Observatory on Health R&D and Donor Tracker) research and development funding for neglected diseases exceeded $3 billion in 2011. The primary focus on three 'top tier' diseases: HIV/AIDS (33.8 percent), TB (17.3 percent), and malaria (18.4 percent). This leaves just over 30 percent of the funding for neglected diseases available for research on all other neglected diseases (Moran *et al.,* 2012). The development of new medicines receives a significantly larger share of funding compared to vaccines. Reports indicate that less than 5 percent of the total annual R&D budget allocated for neglected diseases is dedicated to diagnostics (BIO Ventures, 2011).

Schistosomiasis vaccine development has suffered from major setbacks (Eyayu *et al.,* 2022; Molehin, 2022). The current WHO-approved drug (Praziquantel) is threatened by parasite resistance (WHO, 2015). Praziquantel is a bespoke drug for the treatment of schistosomiasis. The drug executes a highly effective mission in eradicating adult worms in comparison to the juvenile worms (Vale *et al.,* 2017). However, Praziquantel resistance has been reported in *Schistosoma* sufferers, whereby numerous praziquantel doses could not completely cure the disease (Bottros and Bennett, 2007; Liang *et al.,* 2003; Wang *et al.,* 2012). Further, adverse reactions such as hives, fever, bloody diarrhea, and dizziness, among others, have been reported (Lee, 2011). There is a need to design efficacious drugs which can eradicate all stages of the worm's life cycle, instead of relying on drugs (such as Praziquantel) which have such a selective therapeutic regime.

Efforts to control and eradicate schistosomiasis hugely concentrate on snail regulation and surveillance (King and Bertsch, 2015). Praziquantel is also mass administered to communities in a bid to eradicate schistosomiasis. Health education on schistosomiasis, and provision of clean water and sanitation facilities is a major step in preventing the debilitating disease. The latter has greatly lowered the prevalence of schistosomiasis and morbidity rates in Eastern Mediterranean countries such as Oman (Al Abaidani *et al.,* 2016). The WHO has indicated the need for novel

schistosomiasis vaccines and therapies to ameliorate the need for the limited Praziquantel (WHO, 2015). The successful vaccine or therapy should preferably consider strategies which result in

parasite annihilation and subsequent fertility reduction (Molehin, 2020). This introduction will demonstrate that GST activity is inhibited after binding to testosterone. The resulting effect of such binding is lowered *Schistosoma* fertility.

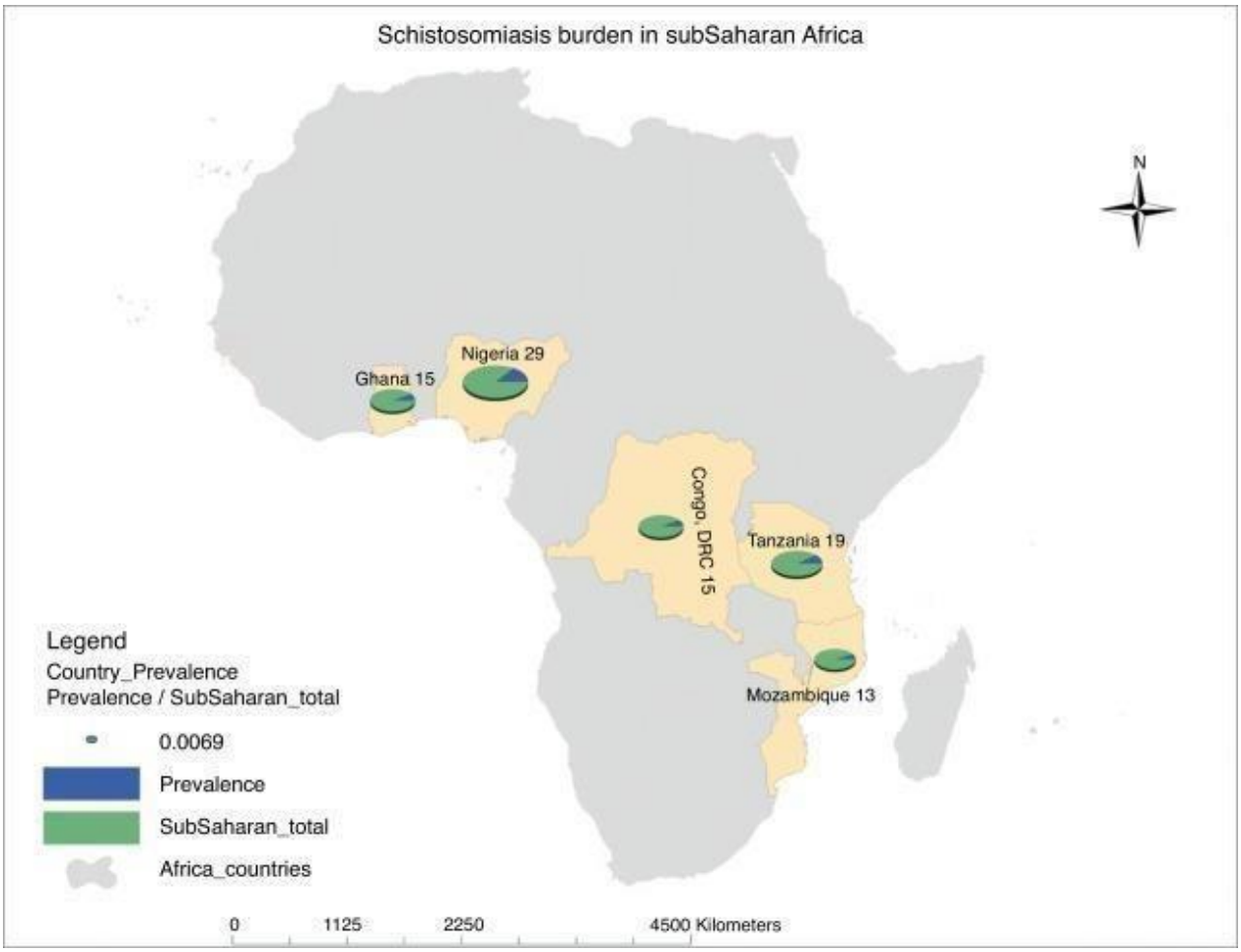


Figure 1.1 Sub-Saharan African (SSA) map showing the predicted burden of schistosomiasis.

Prevalence of the disease is estimated to be around 192 million (alarmingly this is already 93% of the prevalence of schistosomiasis on a global scale). Demographics in this figure show that a cumulative sum of 29 million individuals are plagued by this debilitating disease in Nigeria, followed by Tanzania which has 19 million sufferers. A total of 15 million sufferers hail from Ghana and the Democratic Republic of Congo whilst Mozambique consists of 13 million sufferers. The total disease burden is represented as pie chart inserts. The data in the pie chart was obtained by utilizing the fraction of the total disease burden in SSA and normalizing the figure by a factor of 0.006. Adapted from Adenowo *et al.*, 2015.

1.2 Rationale

The function and properties of protein-ligand interactions play a crucial role in comprehending the intricate biological regulations that occur within living systems and cells. Enzymes remain principal targets for rational drug design protocols since altering enzyme activity possesses well defined and rapid effects. This study has the potential to facilitate the discovery of testosterone that can be used to treat schistosomiasis.

One such drug target, the GST family, utilises GSH in various reactions involved in the modification of many compounds with carcinogenic and therapeutic potential. This family contains a plethora of enzymes and exists in abundance in the testes (Hemachand and Shaha, 2003) The major function of GSTs is to detoxify xenobiotics through facilitating the attack by GSH on electrophiles, resulting in the suppression of their interaction with important cellular proteins and nucleic acids. Hence, GSH is an indispensable hepatic metabolic molecule for human beings and animals.

To date, efforts to treat schistosomiasis have all centered on Praziquantel, which remains the best antidote since its development by Bayer Pharma (Berlin, Germany). Notwithstanding, Praziquantel drug resistance has been well documented (Botros and Bennet, 2007; Molehin *et al.,* 2020; Wang *et al.,* 2012) and is attributed to the immunity of the host (Wilson *et al.,* 2008). Further, the inability of Praziquantel to target juvenile worms remains a major drawback. Another downside is that Praziquantel has no capability of preventing the re-infection of schistosomiasis, thus prompting the need for a vaccine (Aruleba *et al.,* 2019). The vaccine development process for schistosomiasis is currently underway and it is in various phases of human clinical trials (Fonseca *et al.,* 2015; Molehin *et al.,* 2016; Siddiqui and Siddiqui, 2017) where problems resulting from intricate interplay between the host's immune system and the life cycle of the parasite have been encountered. It is generally believed that bioavailability of Praziquantel within the circulation is minimized due to swift metabolism, which leads to minimal efficiency against juvenile schistosome worms (Tayo, *et al.*, 2021). Lipid-based nano-systems have proved to aide in the lipid affinity of the drug thus enabling interaction with hydrophobic nuclei of the worm (Botros and Bennet, 2007).

GSTs are primary detoxifying dimeric enzymes which are essential in schistosome reproduction and viability during infection in the human host infection. All 3 major *Schistosoma* parasites possess 2 major alleles, the 26 and 28 kDa-GSTs for this detoxification. GSTs serve as major pathogenic factors since they aid in the context of schistosomes surviving inside their human hosts.

The 2 isoforms are versatile enzymes which take part in parasite and host interactions, thus posing as lucrative candidates for the schistosomiasis vaccine. Each allele shows contrasting substrate idiosyncrasies and cellular positions. On one hand, studies have showed that if the 28-kDa homolog is inactivated, there will be a reduction in egg quality emanating from decrease in fecundity due to the redox reaction imbalance (Huyse *et al.*, 2009). On the other hand, the 26-kDa homolog exhibits somewhat destructive tendencies in storing and transporting toxins which eventually leads to blockage of the worms' bowels (Toh *et al.*, 2010; Zhan *et al.*, 2010). Although the 2 isoforms work antagonistically, they have been considered excellent drug targets since they can bind to ligands.

The use of hormones as therapeutic agents is well documented (Hyman and Kelner, 2007). Mammalian GST can bind sexual hormones to influence their physiological activity, metabolism, transportation and parasitic infections levels (Johansson and Mannervik, 2001). In particular, the ability of testosterone to attach to the 28-kDa-GST enzyme found in the *Schistosoma haematobium* fluke has been well documented where a tight binding (Kd $=2.57 \times 10^{-7}$ M) to the parasite 28-kDa-GST protein was recorded (Remoue *et al*, 2002). In another study, GSTM1enzyme was able to bind to testosterone and estradiol (Onaran *et al.*, 2007). Use of GSTs in drug discovery often involves targeting both alleles (the 26 and 28-kDa isoforms of Sh28GST) by the inhibitor under investigation.

The primary objective of the present study was to further explore testosterone's role as a potent inhibitor. Given the disparity which exists between the two alleles, there was need to differentiate them based on their structure and function through stability studies. Molecular modelling studies were used to back up the empirical data.

1.3 Novelty

Bilharzia remains a pertinent disease based on the reported high morbidity and mortality rates. Further, lack of vaccines and robust anti-schistosomal drug regimens has prompted the need for novel drugs to enter the bio-pharmaceutical arena. Testosterone binds and is known to inhibit the enzyme activityof Sh28GST by affecting the transferase functionality (Remoue and coworkers, 2002). Notwithstanding, no study has explored the relationship and potential inhibition of parasite GSTs using testosterone. In the quest for novel schistosomiasis therapy, testosterone poses as a

potentialdrug candidate. The sex hormone is the predominant male hormone accountable for sex distinction, thus leading to peculiar male characteristics and production of sperms. Some diseases have been shown to be influenced by sex hormones, particularly testosterone (Lasrado *et al.*, 2020). Experiments done in rodents proved that maturation of schistosomula (a development stage of *Schistosoma* parasites) was dependent on gender of the infected host. Male rodents showed decreased schistosomula growth due to testosterone presence (Abdul-Ghani and Hassan, 2010). Testosterone negatively affects the growth of the worm. A study conducted on humans highlighted a strong link between *S. haematobium* or *S. mansoni* infections with low levels of testosterone (Jatsa *et al.*, 2022). Throughout the course of human bilharzia infection, testosterone could possibly have an important function in the parasite fertility by carrying out anti-schistosome resistance and directly inhibiting the enzyme activity of the GST (Remoue *et al.*, 2000).

When human beings are infected with the much-dreaded schistosomiasis disease, binding of steroid hormones to Sh28GST is potentially a pivotal function in anti-parasite resistance (Aguilar-Diaz *et al.*, 2015). Consequentially, anti-schistosome immunity is negatively affected since there is reduction in the fertility rate. The findings from Remoue *et al,* (2000) revealed that testosterone has a higher tendency to hinder the activity of Sh28GST enzyme in a dose-dependent manner. This suggests that the sex hormone may be directly linked to infertility. These steroidal sex hormones affect parasite infection levels. Studies done by Rijal *et al.* (2001) explored the levels of various parasite infections due to gender selectivity after the onset of puberty, well supported by the notion that anti-parasite immunity depends on regulation by sex hormones (Roberts *et al.*, 2001). A plausible explanation about the effect of testosterone on parasitic worm maturation could be linked to the fact that schistosomes themselves exhibit a receptor which shows congruency to the sex steroids receptor (de Mendoca *et al.*, 2000). The available evidence suggests that steroid hormones play a role in shaping the progression and outcome of parasitic infections (Romano *et al.*, 2015).

Testosterone acts as an indicator molecule by triggering steroid hormone receptors (Bennet *et al.*, 2010), and as a regulatory molecule of GSTs (Coecke *et al.*, 2000). GST enzymes exhibit important duties in steroid hormone biosynthesis, binding of steroids and isomerization in the predominant cells (Robaire and Hamzeh, 2011). In this current investigation, the effects of testosterone on the *Schistosoma* enzymes were evaluated. Experiments such as molecular modelling, inhibition, kinetic, stability and fluorescence studies were performed to analyse the binding and inhibition mechanisms by testosterone.

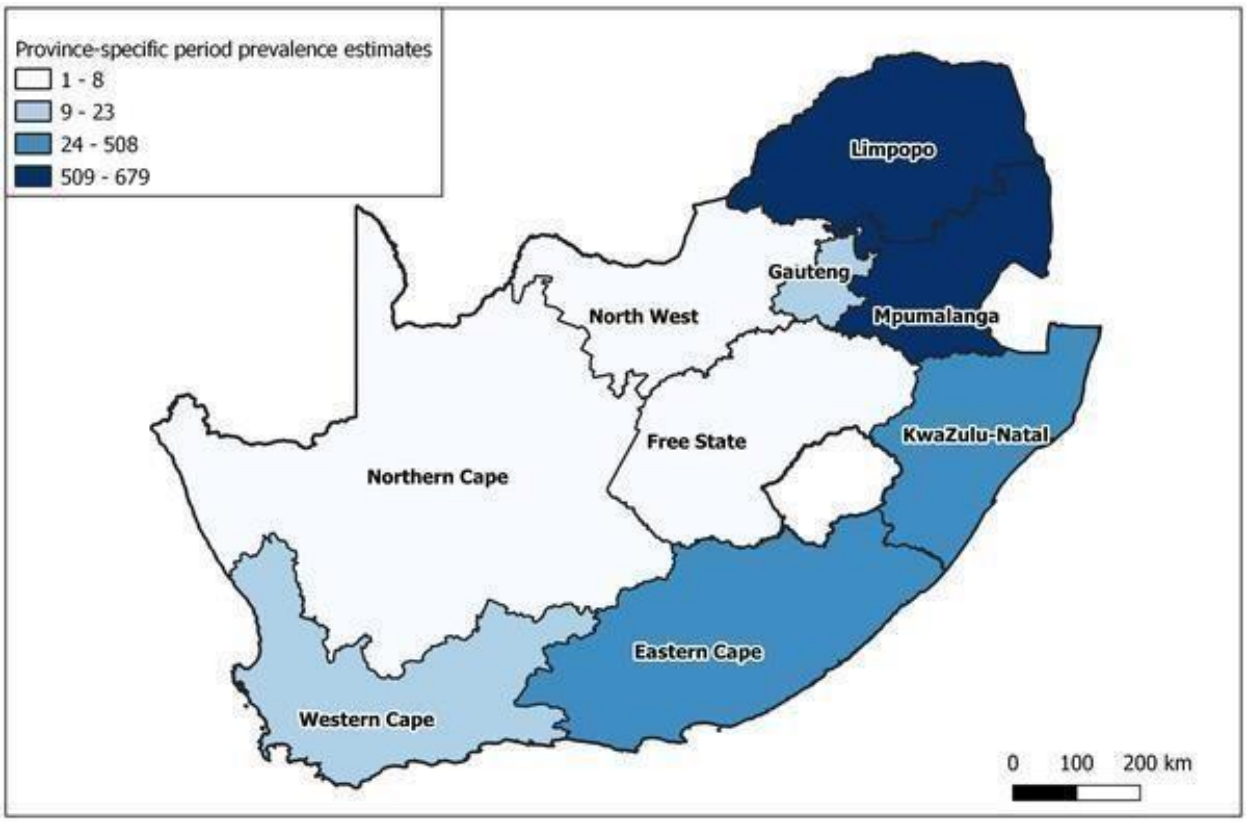


Figure 1. 2 Prevalence of schistosomiasis in South Africa from 2011 to 2018.

Approximation of microscopically confirmed *S. haematobium* cases within South Africa's various provinces. The estimates were given for each 100 000 people by use of the Jenks natural breaks method. Esri ArcGIS 10.2, was used for plotting the map by use of the 2016 provincial boundary shapefile from OCHA ROSEA and CC BY- IGO) license. With reference to epidemiological information, Limpopo and Mpumalanga provinces were greatly affected. Adapted from De Boni *et al.*, 2021.

1.4 Aim and Objectives

This study aims to explore the biophysical properties of Sbh26GST and Sh28GST, to uncover the underlying mechanism through which testosterone exerts its inhibitory impact on these parasite GSTs. Furthermore, this work contributes to the broadening of our comprehension regarding the interplay between parasite GSTs and testosterone, enhancing our understanding of their structure - function relationship.

To achieve this aim, the following methodologies were employed;

- To heterologously express recombinant Sh28GST and Human GST-mu in pET-11a and Sbh26GST in pET-28a, expression vector using the *E. coli* expression system and purification of the recombinant proteins by use of immobilised metal affinity chromatography (IMAC).
- To establish the specific activity of the recombinant Sbh26GST, Sh28GST and Human GST-mu by use of the GSH and the CDNB conjugation assay.
- To perform structural characterisation of the recombinant GST proteins using far-UV CD and fluorescence spectroscopy (intrinsic and extrinsic).
- To determine the IC_{50} of testosterone on Sbh26GST, Sh28GST and Human GST-mu activity.
- To establish the K_m and V_{max} values of Sbh26GST, Sh28GST and Human GST-mu with or without testosterone by Michaelis-Menten kinetics.
- To thermally analyse the stability of Sbh26GST, Sh28GST and Human GST-mu in the presence and absence of testosterone using a thermal shift assay.
- To computationally analyse the dynamics of Sbh26GST, Sh28GST and Human GST-mu and testosterone through molecular modelling studies, induced-fit docking experiments, and molecular dynamic simulation.

1.5 Outline of the thesis

The report began with a review of the literature (Chapter 2) that provided information about schistosomiasis, the innovative use of testosterone with GSTs, and various biophysical techniques used in this study. The third chapter dealt with the production, measurement, and qualitative analysis of recombinant proteins for their further use in downstream applications. The fourth chapter focused on the functional characterization of recombinant proteins, including enzyme activity and thermal stability. The enzyme kinetics were also determined in this chapter. In the fifth chapter, the report discussed the use of computational modeling, molecular docking, and molecular dynamic simulations to complement the study by demonstrating the binding affinities, binding modes, and stability of the GST-testosterone complex. The 6th chapter focused on crystallographic studies on the Sh28GST. Finally, chapter seven presents a general discussion of the study's results, including the concluding remarks and future studies.

Chapter 2
Literature Review

2.1 Brief background on schistosomiasis

Schistosomiasis is an under researched tropical water disease which ranks as the second cause of death (after malaria) amongst parasitic diseases. The five main species responsible for causing schistosomiasis comprise; *Schistosoma mekongi, Schistosoma haematobium, Schistosoma guinesis, Schistosoma japonicum, Schistosoma intercalatum* and *Schistosoma mansoni* (Adenowo *et al.,* 2015). Among these species, the three major species which affect people are *S. japonicum S. mansoni,* and *S. haematobium* (Colley, 2014). The schistosomiasis disease burden is mostly ascribed to *S. mansoni* and *S. haematobium* (Utzinger *et al.,* 2009). Whilst the *Schistosoma haematobium* species is prevalent in Africa (Figure 2.1), the *Schistosoma japonicum* species is predominantly found in Asia. The unacknowledged impact on poor communities of living with schistosomiasis is of great concern since the disease causes socio - economic ripple effects emanating from the chronic illness (Colley *et al.,* 2014).

This parasite undergoes a life cycle that involves two different hosts, namely snails and mammals. Schistosomiasis can be categorized into two types - acute and chronic. The acute form may cause symptoms such as fever and headache, while the chronic form may lead to dysuria and hyperplasia. This infection can affect various parts of the body such as the bile ducts, intestine, and bladder (Gray *et al.,* 2011).

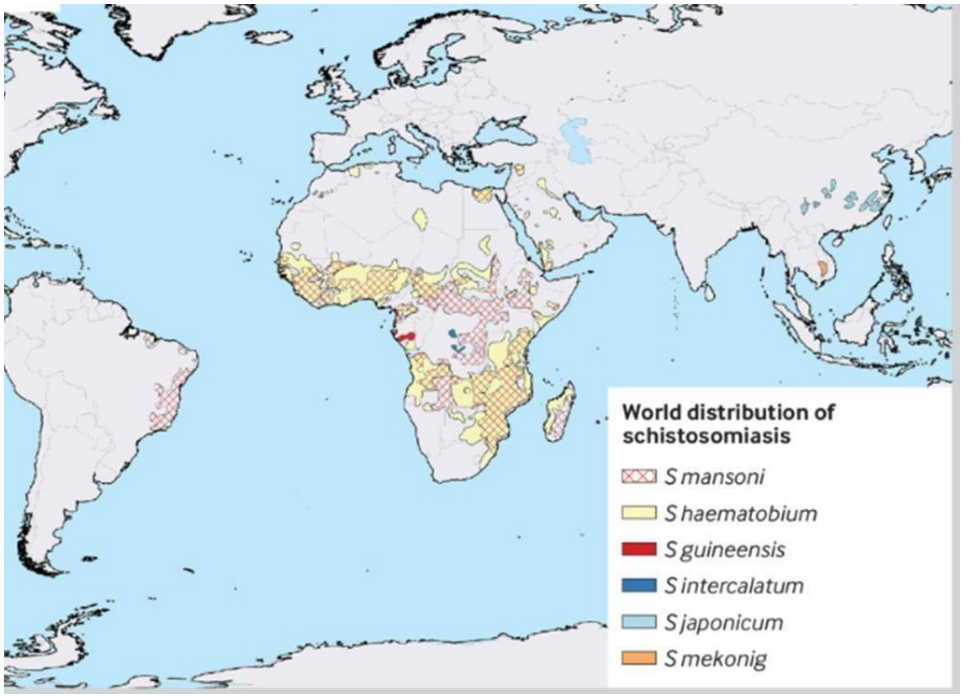


Figure 2. 1 World distribution of schistosomiasis.

Schistosomiasis cases are predominantly found on the African continent, where the majority of cases are due to *S. mansoni* and *S. haematobium* (Taken from De Leo *et al.*, 2020).

The schistosome life cycle rotates between the human host and the freshwater snail intermediate host. The life cycle (Figure 2.4) starts when the snails release cercariae which is responsible for causing schistosomiasis. When people enter contaminated water bodies when they are swimming, wading, bathing or fishing, cercariae pierce through the skin. These *Schistosoma* cercariae enter the hepatic portal vein before lodging in the liver as adult worms and proceeding to target various body organs; for example, *S. haematobium* damages the bladder, whilst *S. mansoni* targets the intestines. Failure of the schistosome eggs to be fully excreted leads to these eggs permanently abiding within the liver and intestines.

The longevity of *Schistosoma* is aided by GST enzymes that prevent oxidative damage and detoxify peroxidation products within the parasite. These enzymes, specifically two prominent 26 and 28 kDa-GSTs, have distinct substrate specificity due to variations in their H-site. Together, they prevent toxin accumulation in the worm by amalgamating glutathione to lipid peroxidation products. The active site of the enzyme contains both a hydrophobic substrate-binding H-site and a G-site. However, various chemical compounds have been found to inhibit the activity of GSTsin schistosomes, which presents an opportunity to develop vaccines or drugs targeting these enzymes if the binding mechanisms are understood.

Since schistosomes are parasites, drugs targeting schistosomal GSTs will also affect the host GSTs. In the case of humans, the α, μ and π GST classes are actively involved in drug metabolism and are also antioxidant in nature. Contemporary experimental and computational analyses facilitate the drug discovery and development by investigating, predicting and interpreting clinical , molecular, genetic and multifaceted scientific data (Wishart, 2005). The most popular drug discovery initiatives usually revolve around enzyme conformation, systemic or ligand interactions (Schaduangrat, *et al.,* 2020). Bioinformatic techniques are indispensable tools used in drug discovery to search for enzyme inhibitors against disease-causing organisms (Xia, 2017).

The literature review will focus on use of experimental and bioinformatic methods which aid in identifying potential enzyme inhibitors for a robust drug design protocol (Nikolic *et al.*, 2016). It is necessary to study enzyme inhibition mechanisms since the mode of action of various drugs utilise enzyme inhibition mechanisms (Ramsay and Tipton, 2017). The basis for inhibitor selection relies on the contrasting functional attributes rendered by the same type of enzymes, albeit executing the same roles, especially in human beings (Serrano *et al.,* 2013). In this present study, the knowledge of the structure-function of Sbh26GST, Sh28GST and Human GST-mu will be expanded in addition to the available solved GST structures. X-ray crystallographic studies will enable elucidation of the ligand-bound Sbh26GST, Sh28GST and Human GST-mu complexes.

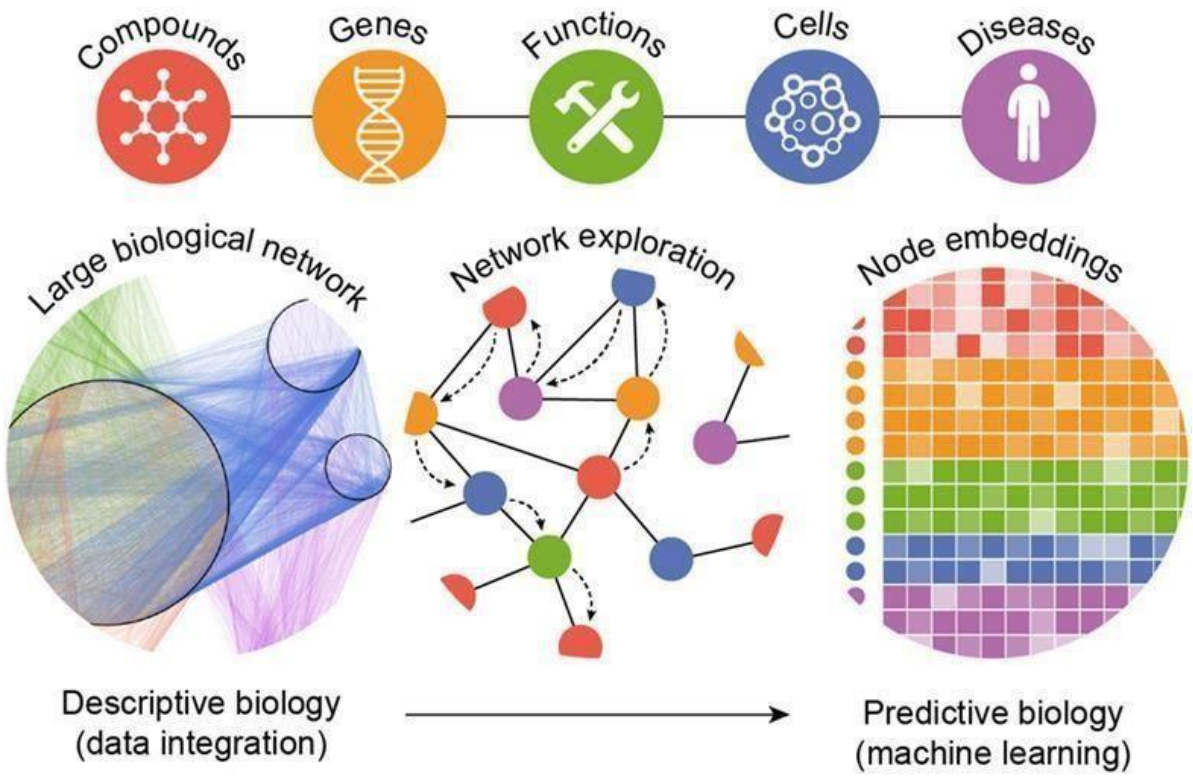


Figure 2.2 Interrelationship between bioinformatics and machine learning.

Data is initially integrated from large biological networks to suggest drug targets by means of machine learning. Extensive databases containing small molecules and metabolites, have enabled more accurate protein-ligand docking experiments and more insightful virtual screening. Adapted from Dara *et al.*, 2022.

2.2 Schistosoma classification

Schistosomatidae family members (*Schistosomatinae*, *Gigantobilharziinae* and *Bilharziellinae*) are dioecious trematode parasites of birds and mammals (Brant and Loker, 2013). It is believed schistosome diversity might be greater (Brant *et al.*, 2006; Morgan *et al.*, 2003). They are believed to belong to a distinct lineage of trematodes with a 2-host life cycle where males and females live separately in the intermediate gastropod host and the definitive host's vascular system (Brant and Locker, 2013). Compared to other trematodes, mature schistosomes occupy the mesenteric veins of their definitive hosts where they have access to nutrients from the blood (Loker and Brant, 2006).

In mammals, 4 genera and 30 schistosome species have been described. All known mammalian schistosomes have freshwater-based life cycles where they use the pulmonated freshwater snails (Pomatiopsidae, Lymnaeidae and Planorbidae) as intermediate hosts (Brant and Locker, 2009; Brant *et al.*,2013).

While mammalian schistosomes infect about eight orders of mammals, they differ in their definitive host range, with *S. japonicum*, for example, being an extreme generalist, while *S. haematobium*, is largely confined to humans (Brant and Loker, 2013).

2.1.1 Schistosomiasis disease burden

Schistosoma, a parasite, is responsible for causing schistosomiasis (bilharzia) in humans. Despite the availability of treatments, this disease continues to persist and has even re- emerged in regions where it was previously controlled (Knopp *et al.*, 2019; Wiegand *et al.*, 2017). Schistosomiasis is transmitted in a cycle that involves water sources becoming contaminated with excreta, intermediate hosts of freshwater snails, and human exposure to the polluted water. Worldwide, approximately 200 million people are infected with schistosomiasis. During human infection, schistosomes extract nutrients from the blood and globulins anaerobically, and they regurgitate the debris back into the blood. *S. mansoni* Biomphalaria snails transmit the type of schistosomiasis that results in hepatic and intestinal symptoms, while Oncomelania snails transmit *S. japonicum*, which causes schistosomiasis in the intestines and liver. Although humans are the primary hosts of *S. mansoni*, this parasite can also infect rodents and primates (Gryseels *et al.*, 2006).

Penetration of the skin by schistosome cercariae can result in a long-lived rash or swimmer's itch that can last for many days (Bottieau *et al.*, 2006; Horak and Kolarova, 2005). If the itch is serious, it leads to a hypersensitive response to migrating schistosomula, characterized by fever, pain, coughing, fatigue, which can deteriorate to weight loss, diarrhea, blood poisoning, rash and stomach pain. Infection of the urinal organs can cause dysuria and haematuria (Paul *et al.*, 2002). Intestinal schistosomiasis caused by eggs migrating across intestinal walls results in inflammation of the mucosa, micro ulceration and bleeding (Chen, 2014).

Hepatic schistosomiasis causes hepatomegaly and fibrosis (Shaker *et al.*, 2014). These symptoms are caused by schistosome eggs secreting proteolytic enzymes in host tissue, inducing inflammation. Symptoms vary between individuals due to differences in immune responses and infection intensity. Ectopic schistosomiasis includes pulmonary and genital schistosomiasis, which affects the lungs and reproductive organs respectively, and causes hypertrophic and ulcerative lesions that lead to bronchial symptoms and infertility. Neuro-schistosomiasis affects cerebral and spinal veins which can lead to spinal, or brain pathologies (Ferrari, 2004).

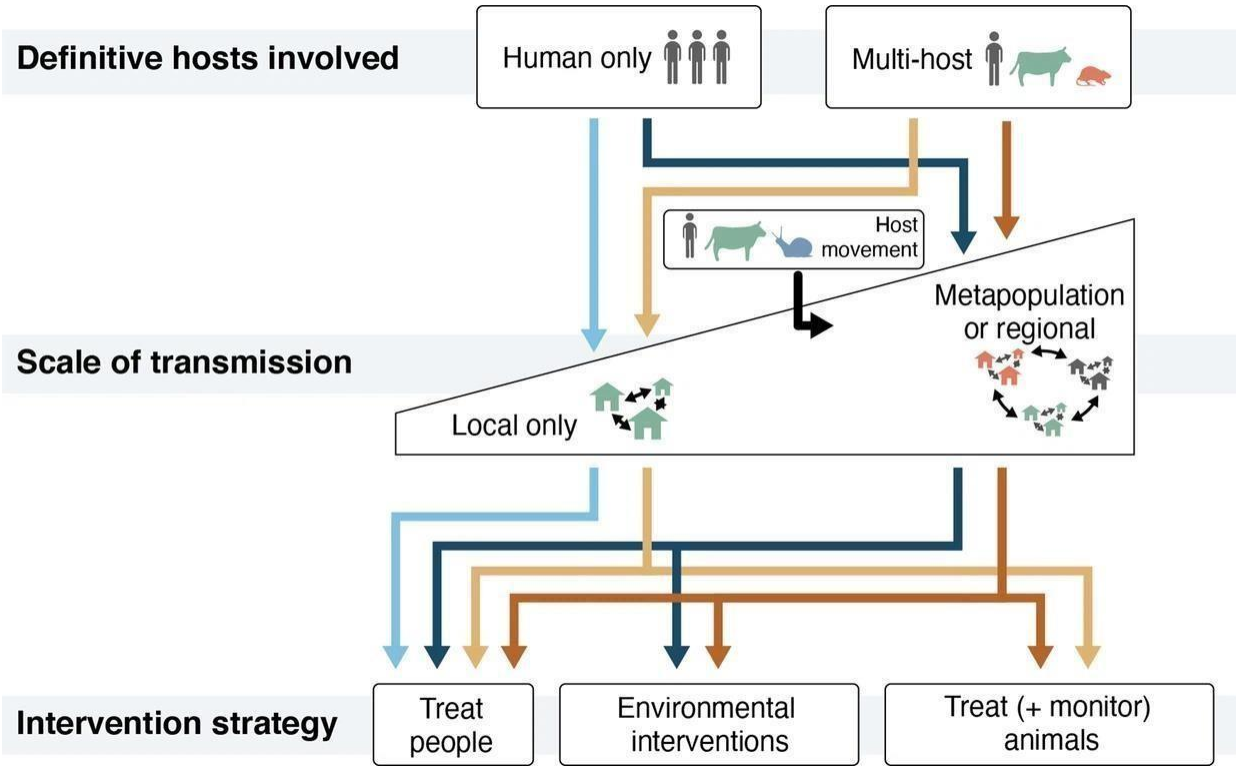


Figure 2. 3 A framework for schistosomiasis-targeted interventions.

Intervention strategies including treatment, managing the environment and monitoring/treating animals are deployed to mitigate against local and regional spread of infection among human and non-human hosts (Lund *et al.*, 2022).

According to the WHO, schistosomiasis affects 221 million people globally, with the majority belonging to poor communities. About 90% of the global cohort of schistosomiasis-infected people are found in the Africa. The disease is spread through poor sanitary practices such as open defecation and urination; and infection occurs when the parasitic larvae penetrate human skin exposed to contaminated water. Economic and social activities linked to water (such as fishing and swimming) increase the vulnerability of at-risk communities to infection. In addition to the symptoms discussed in section 2 above, schistosomiasis can result in anemia, stunted growth and reduced ability to learn or to be economically productive. Urogenital schistosomiasis also increases the risk of infection with STIs such as HIV due to damaged genital epithelia (WHO, 2013).

Indirect morbidity is a pertinent concern since it can include physical and cognitive impairment, it is difficult to dissociate it from other socio-economic factors and it is challenging to quantify (King *et al.,* 2005). Schistosomiasis is counted among the ten topmost causes of disability in Africa, with approximately school-going children being the most affected (Gabaake *et al.,* 2022). The control measures against schistosomiasis include improving sanitation and access to safe water, better environmental management and snail control, preventive chemotherapy, and health education (WHO, 2013).

2.1.2 *Schistosoma haematobium* life cycle

S. haematobium is endemic to Africa and the Middle East, and mainly affects the genitourinary system, and additionally the colon and lungs. It has a significant effect on morbidity and mortality. The main hosts for parasitic trematode *S. haematobium* are humans. Fertilised *S. haematobium* eggs hatch in freshwater aquatic environments and the miracidia infest a specific intermediate host which is the *Bulinus* freshwater snail. Schistosomulae can traverse the host through the circulatory system and will mature to adult worms in 6 weeks. They will mate in the venules of the pelvis where they lodge, and lay eggs. The adult worms may produce eggs for years as they persist in the host within the bladder, rectum, ureter and other pelvic organs. Some of the eggs may be excreted through the bladder or intestines, and some eggs cause micro- mucosal perforation as they migrate in the pelvic tissue. The urothelial micro-mucosal perforations lead to haematuria, and unsecreted eggs become encapsulated in fibrous granulomas, leading to the specific pathologies associated with urogenital schistosomiasis (Bichler *et al.,* 2006; Gryseels *et al.,* 2006; Odegaard and Hsieh, 2014; Shebel *etal.,* 2012).

2.1.3 Disease prevalence in Southern Africa

The South African government recognizes schistosomiasis as a major neglected health problem (NICD, 2018). In South Africa, 4 million people are at risk and in some endemic locations in the northern and eastern provinces, prevalence of the disease in children is 95% (De Boni *et al.,* 2021).

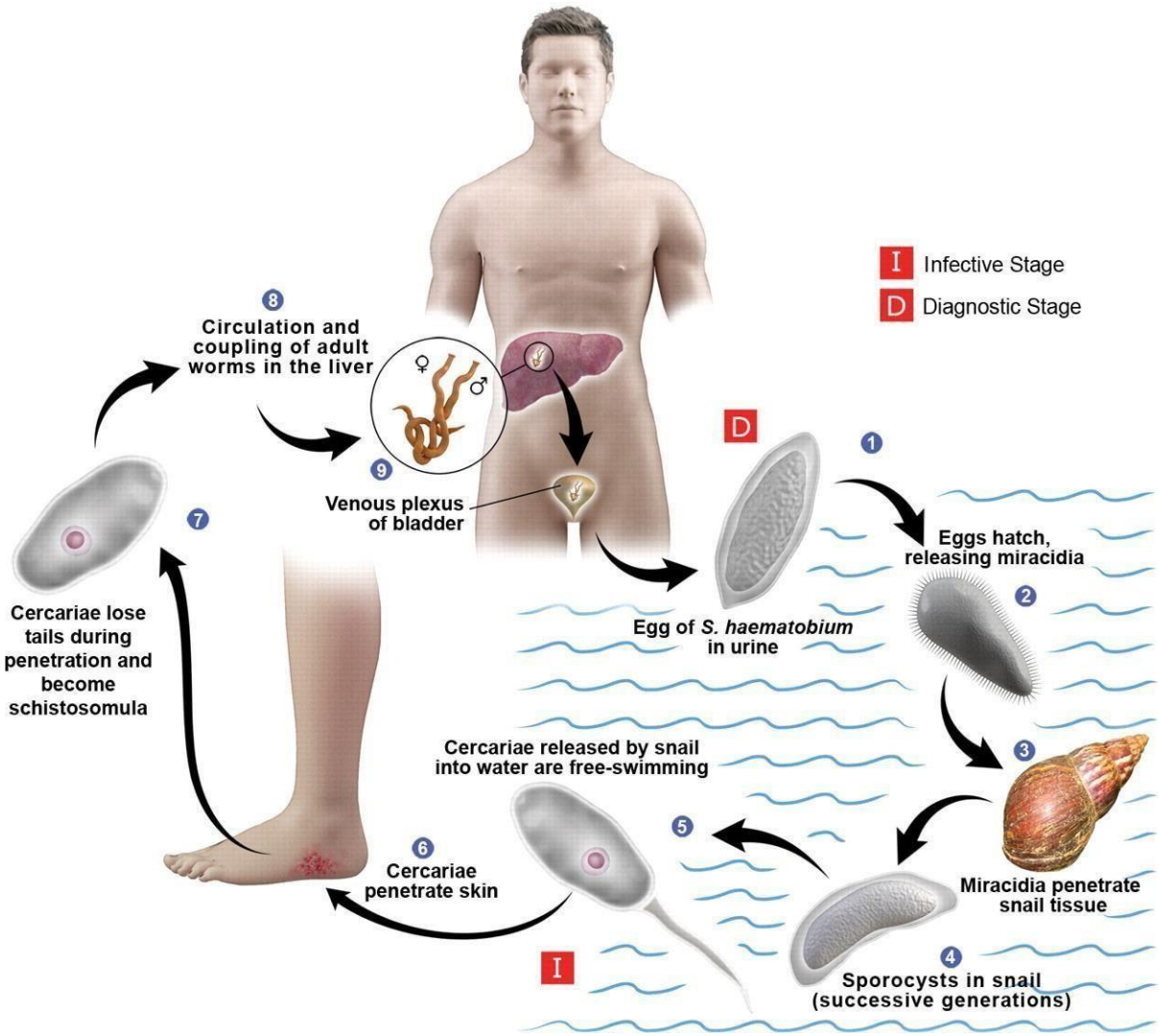


Figure 2. 4 Schistosomiasis life cycle. Schistosomal cercariae enter the skin and navigate through the lungs into thehepatoportal system, where the adult developmental stages mature and engage in mating upon contact with contaminated fresh water. After mating, adult females migrate to either the gastrointestinal system (*S. mansoni* and *S. japonicum*) or the urogenital system (*S. haematobium*), where they lay thousands of eggs. These eggs are then excreted in feces or urine into freshwater. Subsequently, they hatch, giving rise to miracidia. This stage, capable of free movement, infects the intermediate snail host and produces successive generations of sporocysts. Cercariae are eventually removed after a span of several weeks**.** Adapted from Shebel *et al.*, 2012.

In the Limpopo province, co-infection with *S. haematobium* and *S. mansoni* has been recorded in up to 73% of school children (Wolmarans *et al.*, 2005). In Lesotho, schistosomiasis is considered non-endemic and had aprevalence rate of 8.3% in 2015 (Lai *et al.*, 2015). The national prevalence of the disease in Swaziland, where it is endemic, was 22.6% in 2012 (Rollinson *et al.*, 2013).

In Zimbabwe, schistosomiasis prevalence in primary school children was 22.7% in 96% of districts (Midzi *et al.,* 2014). Botswana bore a 38% prevalence of *Schistosoma* (Gabaake *et al.,* 2022). Schistosomiasis was also endemic in the northern regions of Namibia, with a prevalence of 9% (Mupakeleni *et al.,* 2017; Sousa-Figueiredo, 2015). The approximate average prevalence of *S. haematobium* in Mozambique was recorded to be 47%, nevertheless in coastal areas, prevalence may be as high as 93% (Augusto *et al.,* 2009; Phillips *et al.,* 2018). The prevalence of *S. haematobium* among school going children and adults was recorded at 32.2% and 54% respectively in Zambia (Kalinda *et al.,* 2019). Malawi exhibited a national *S haematobium* prevalence rate of 6.9%, although the rate could have possibly been as high as 43% in certain regions (Makaula *et al.,* 2014).

2.2 Treatment of schistosomiasis

While stool and urine examination under a microscope remains the standard diagnostic method for schistosomiasis, other techniques such as PCR-based diagnostic assays, rectal or bladder biopsies, and antibody detection have also been useful in detecting schistosomiasis infections. The most used drug against schistosomes is Praziquantel, which is a pyrazinoisoquinolone derivative. Praziquantel acts on the calcium ion channels of schistosomes, inducing tetanic contractions and the formation of tegumental vacuoles, which leads to detachment from vein walls and eventual death. Studies have demonstrated that a daily dose of 20 mg/kg, administered orally two or three times, is effective, and if eosinophilia and high antibody levels persist, a repeat treatment after four weeks is suggested (Cioli and Pica-Mattoccia 2003; Doenhoff *et al.,* 2008). The side effects of Praziquantel are believed to be caused by the death of the schistosomes and include nausea, dizziness, rash, and pruritus (Cioli and Pica-Mattoccia, 2003). However, resistance to Praziquantel (*S. haematobium* and S. *mansoni*) was reported in certain African regions (Ross *et al.,* 2002). Therefore, research into new drugs for treating schistosomiasis has become necessary.

Apart from Praziquantel, corticosteroids, such as prednisone, may be used to treat symptoms (such as allergic reactions and inflammation) associated with schistosomiasis (Ross *et al.,* 2007). For treating cerebral schistosomiasis (induced seizures), anticonvulsants and surgery may be employed (Ferrari *et al.,* 2004). Oxamniquine and metrifonate have also been used to a lesser extent to treat *S. mansoni* and *S. haematobium* (Gryseels *et al.,* 2006). Arthemether has been used as a prophylactic but cannot be used in regions of malaria endemicity due to the risk of selecting for artemether-resistant Plasmodium (Xiao, 2005).

2.3 Structure and function of GST

GST enzymes belong to a supergene family of dimeric proteins that conjugate glutathione to detoxify electrophiles such as by-products of reactive oxygen species (ROS) and polycyclic aromatic hydrocarbons (Strange *et al.*, 2001). These GSTs are responsible for Phase II detoxification. They attach GSH to various xenobiotics, which helps to protect cells from environmental and oxidative stress. However, this ability can also lead to resistance to drugs. Recent advances in molecular biology have shown that GSTs have a wider range of functions than previously thought. They are involved in prostaglandin, steroid, and leukotriene biosynthesis and metabolism, managing toxic byproducts of lipid oxidation and S-glutathiolated proteins caused by oxidative stress, and even in the development of resistance to chemotherapeutic agents. It is noteworthy that GSTs achieve a concentration of 0.5-0.8 mM within cells, which allows them to function in vivo even when the concentration of xenobiotic molecules is much lower than that of GST (Fabrini *et al.*, 2011).

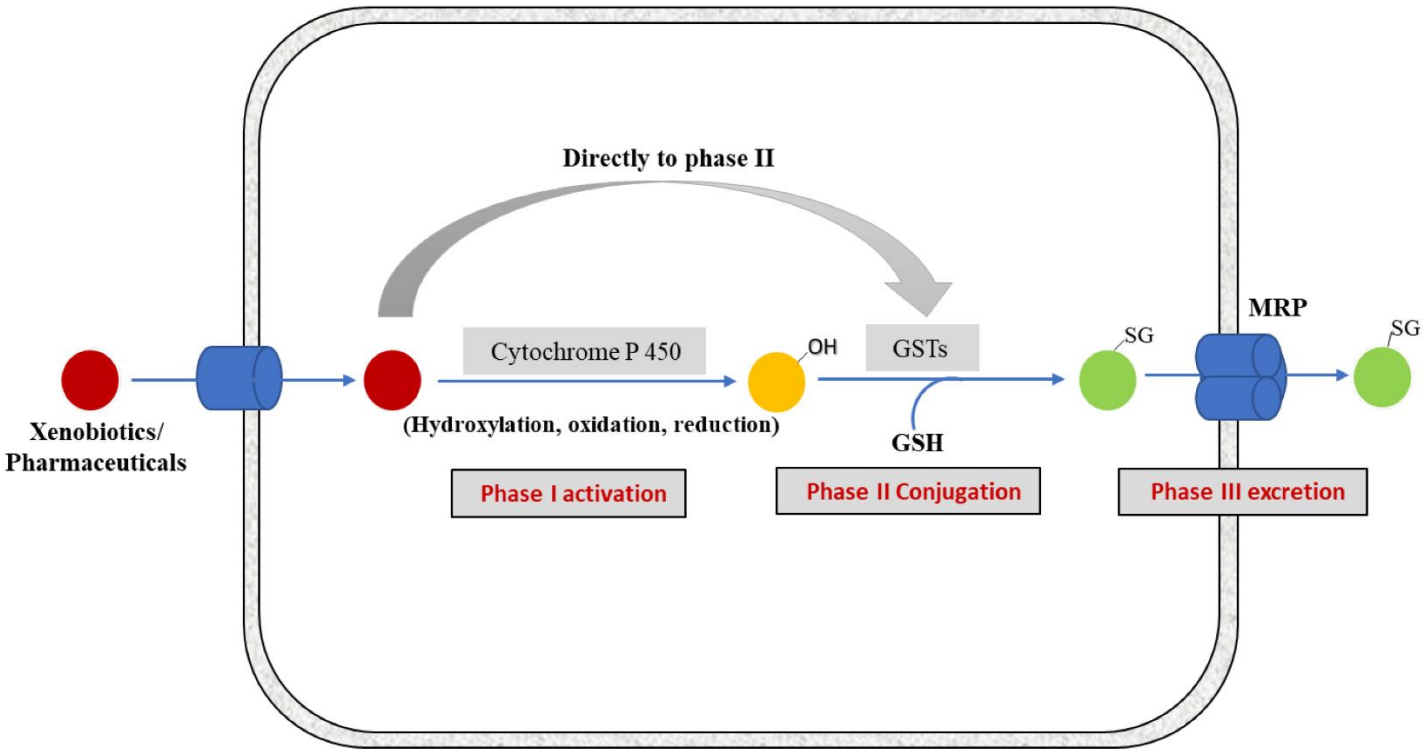


Figure 2. 5 The process of GST detoxification.

Toxic compounds (including endogenous, exogenous, and drugs) inside the cell are transformed by enzymes based on their chemical properties. Phase I enzymes, such as the Cytochrome P450 family, are responsible for bio-transforming lipophilic compounds, while more polar compounds undergo Phase II reactions catalyzed by enzymes like GSTs (Mazari *et al.*, 2023).

2.3.1 Classification and structure

Since their discovery about 3 decades ago seven classes (Omega, Mu, Alpha, Zeta, Pi, Theta and Sigma) of GSTs currently exist in both animals and human beings. GSTs can be divided into 3 groups based on their cellular localisation: cytosolic, mitochondrial or membrane, that is, soluble and insoluble (Hayes and Strange, 2000). All GSTs found in the cytosol are made up of two protein subunits and have comparable three-dimensional structures, despite having low similarity in their amino acid sequences (Figure 2.6). Each subunit has a specific area where it binds to glutathione (G-site) as well as another area (H-site) where it binds to hydrophobic toxic substances (Chronopoulou *et al.*, 2012). The gene families of mammalian soluble GSTs are divided into several classifications, including o (omega), (alpha), p (pi), s (sigma), y (theta), m (mu), and z (zeta), each of which is localized to specific chromosomes. In humans, the μ gene family is located on Chromosome 1, where five μ class genes are situated in tandem (Strange *et al.*, 2001). The κ class is composed of mitochondrial GSTs. There are four subgroups of membrane GSTs, also known as membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG), termed I, II, III, and IV (Jakobsson *et al.*, 2000).

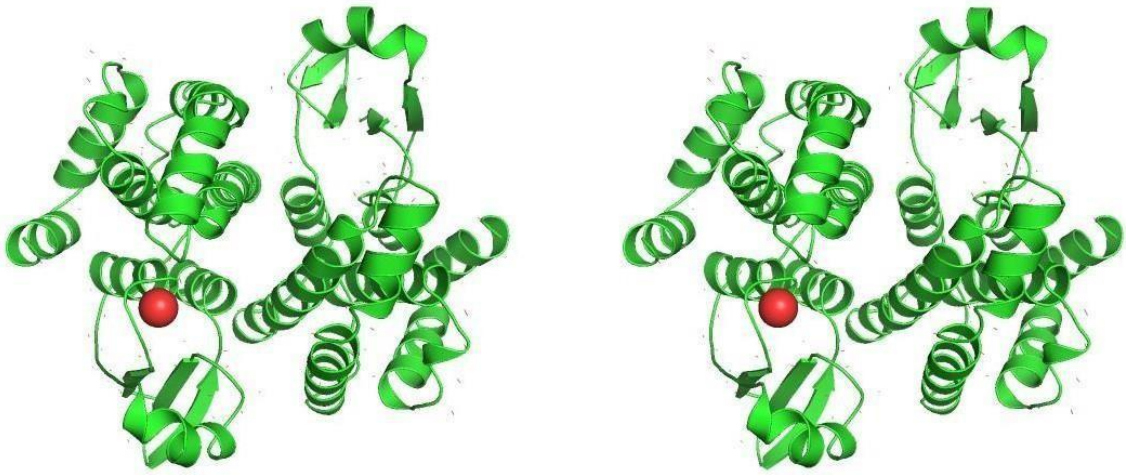


Figure 2. 6 Structure of Sh28GST

The structure of Sh28GST is presented here, show casing the GST dimer with a ribbon representation of the complex, revealing two monomers in green and a red sodium atom. The PDB ID for this structure is 8ALS. Figure taken from Makumbe *et al*. (unpublished data) and created using The PyMOL Molecular Graphics System DeLano Scientific, LLC, USA. http://www.pymol.org

GSTs can be found in diverse aerobic living beings, which encompass fish, bacteria, fungi, plants, yeasts and insects (Mazzetti *et al.,* 2015). Cytosolic GSTs are typically homodimeric or heterodimeric and act independently (Oakley, 2011). GSTs are further classified based on the amino acid they use to activate GSH, which can either be a tyrosine-type (Y-GSTs) or serine/cysteine-type (S/C-GSTs) (Atkinson and Babitt, 2009).

2.3.2 Catalytic ability

GSTs are phase-II enzymes which recognize a broad range of substrates which are structurally diverse, mostly hydrophobic and contain an electrophilic center (Sheehan *et al.,* 2001). They catalyse the conjugation of GSH with a hydrophobic co-substrate that contains an electrophilic center. GSH conjugation with the co-substrates renders the co-substrates more suitable for cellular export (Oakley, 2011). Most cytosolic GSTs recognize drugs and xenobiotics as substrates (Oakley, 2011). GSTs may also utilize GSH as a cofactor or act in hormone synthesis, amino acid degradation, peroxide breakdown and signaling inhibition (Hayes *et al.,* 2005). With regards to cancer, pharmaceutical research is concentrating on exploring novel chemotherapeutic agents and developing pro-drug compounds.

2.3.3 GST as a potential druggable enzyme in Schistosoma

Schistosomes possess two primary GST isoforms, specifically the 26-kDa and the 28-kDa variants, each exhibiting distinct substrate specificities. These GSTs have been sequenced and the regulation of their transcription is known. Their 3D constructs have also been determined using X-ray crystallography and in complex with various substrates (Smyth and Martin, 2000). The schistosomal GSTs work in protection of the schistosomes from attack by the hosts natural defense and to passively detoxify anti-schistosomal drugs (Scott and McManus, 2000). Inhibition of GST activity exposes the schistosomes to accumulation of toxic electrophiles and immune system agents (Shu-Hua *et al.,* 2002). Anti-GST antibodies have been shown to inhibit female worm fecundity and egg viability, although the exact mechanisms for this are unknown (You, 2018). Immunisation of *S. haematobium*-infected monkeys with Sh28GST reduced worm fecundity (Boulanger *et al.,* 1999). Taken together, the existence of substrates with inhibitory or anti-GST activity potentiates the targeting of GST for drug design. The varied functional repertoire of schistosomal GSTs leads to a multiplicity of GST inhibitors that can be employed in rational drug design strategies (Mahajan and Atkins, 2005).

2.4 Testosterone as a therapeutic agent

The intricate procedure of generating steroid hormones, such as progesterone and testosterone, entails a succession of oxidation and isomerization processes originating from cholesterol, which itself arises from a convoluted metabolic pathway that commences from the mevalonate pathway (Buhaescu and Izzedine, 2007). An essential stage in this route is the conversion of androst-5-ene-3,17-dione (Δ5-AD) to androst-4-ene-3,17-dione (Δ4-AD) through a reaction of double bond isomerization. While bacteria employ α-ketosteroid isomerase (KSI) to carry out this process (Pollack, 2004), mammalian tissue relies on glutathione S-transferase (GST) A3-3 as its close analogue. Apart from their recognized role in detoxification, it has been revealed that glutathione transferases (GSTs) possess other biological functions, including the ability to act as ketosteroid isomerases, which seems to assist in the production of steroid hormones in the tissues of mammals. Steroid hormone production, which is critical for regulating a variety of physiological processes, depends largely on the conversion of cholesterol into several end products through a sequence of biochemical transformations. These end products then activate receptors on the cell membrane or within the nucleus, initiating their biological effects.

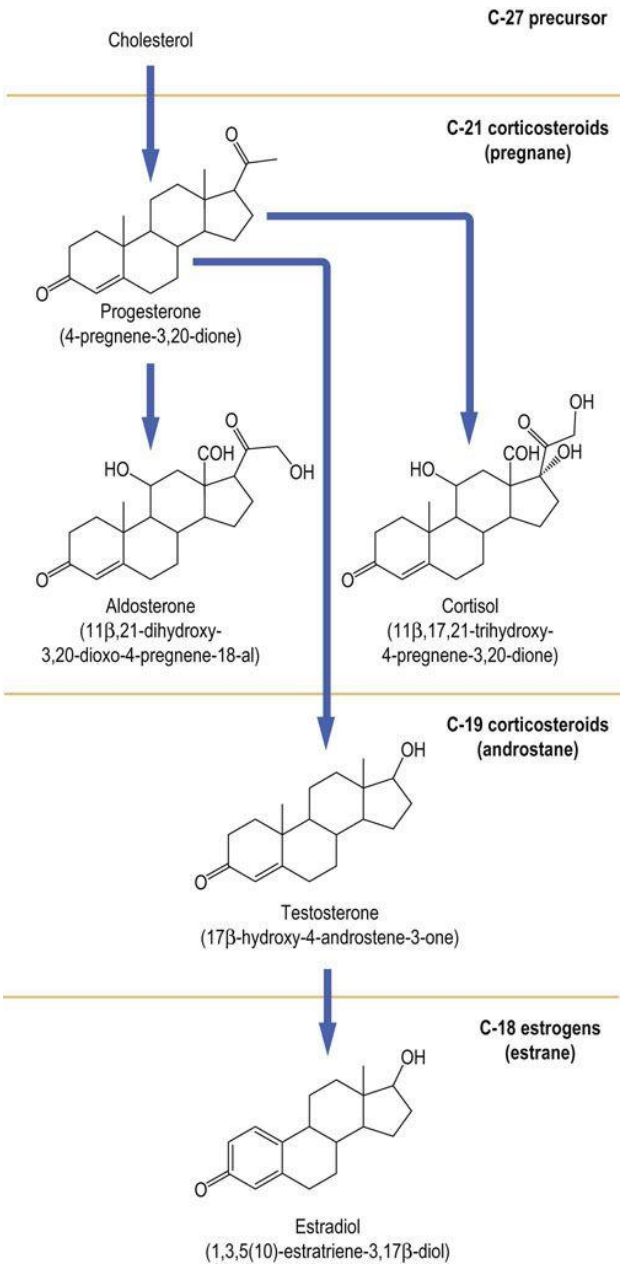


Figure 2.7 Schematic of the formation of human steroid hormones from cholesterol.

Steroid hormones are derived from cholesterol: progestins (which have an active role in pregnancy) and androgens (which promote male sexual characteristics (Baynes and Dominizack *et al.*, 2015).

The biosynthetic pathway to steroid catabolism is mediated by GST in humans (Johansson and Mannervik, 2001). The ketosteroid isomerase function displayed by GSTs illustrates how detoxification enzymes have developed new roles in the steroid biochemistry and played a part in

broadening the functional range of glutathione transferases. The metabolism of testosterone involves the participation of the human GST enzyme, and the pathway responsible for testosterone's action can be inhibited at multiple sites, indicating that testosterone could be a promising drug candidate (Johansson and Mannervik, 2001).

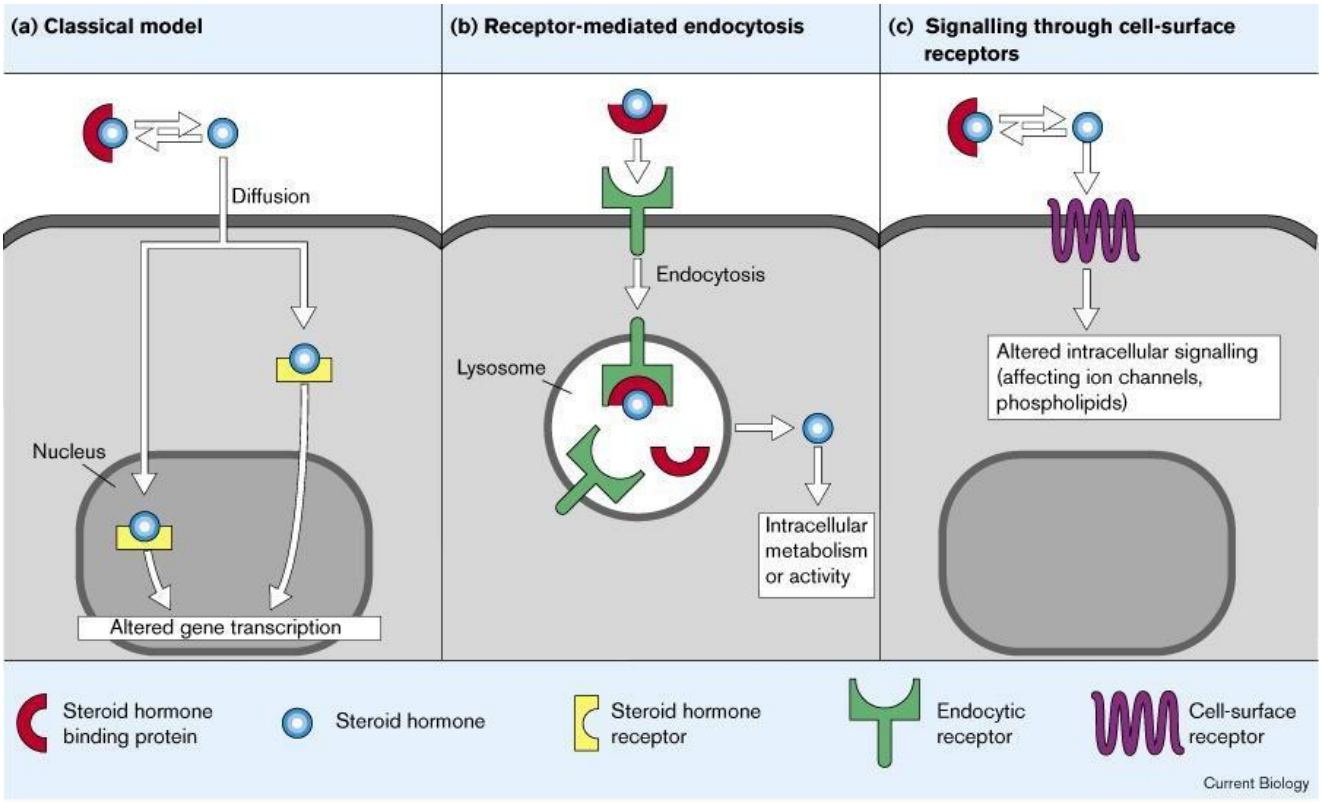


Figure 2.8 Steroid action in cells. (a) The steroid hormone dissociates from its plasma carrier protein, traverses thecell membrane, and upon entry into the cell, binds to an intracellular receptor, leading to modifications in gene transcription. (b) The steroid hormone, while bound to its plasma carrier protein, is internalized through a cell-surface receptor. (c) The unbound steroid hormone influences intracellular signaling by binding directly to these receptors (taken from Chen and Farese, 1999).

The extent of parasitic infection is known to be influenced by sex hormones as gender-dependent patterns have been observed. Sex steroids such as testosterone regulate antiparasitic immunity and therefore pose as good targets for development of therapeutic agents (Gomez *et al.*, 2013).

The suppression of the 28-kDa from *S. haematobium* by testosterone seems to rely on applied dosage and entails the amino acid residues 24-43, which are found at the active site (Remoué *et al.,* 2002). Testosterone can potentially be exploited as a therapeutic agent targeting GSTs for inhibition and reducing fecundity of the *Schistosoma*.

2.5 Computational Modelling

Development of new drugs is a costly and extremely lengthy process which can run into decades (Dimasi *et al.,* 2010). Drug discovery is often initiated by scientists to interfere with enzyme activity to make certain biological processes defective (Mohs and Greig, 2017). Some technical or scientific feasibility issues may also prolong or even deter the drug discovery process altogether (Sancenon *et al.,* 2015). Computational methods involving bioinformatics, machine- learning and artificial intelligence aid in drug discovery (Wishart, 2007). Drug-target associations can be predicted and tested computationally (Chen *et al.,* 2016). Virtual screening of molecules with therapeutic target affinities can be conducted by analyzing bioactive conformation of drug- target complexes (Barrett and Langdon, 2006). Machine learning has been adopted rapidly for drug discovery because of its capability to discern patterns in big data sets (Elbadawi *et al.*, 2021). High quality models of drug-target interactions are now possible due to the combination of solved 3D protein structures, bioinformatics and molecular modelling (Lounnas *et al.,* 2013).

The objective of this study is to explore the therapeutic impact of testosterone on both parasite and Human GST-mu. The goal is to determine if testosterone can be used as a pro drug to treat schistosomiasis. This will be accomplished through a comprehensive analysis of the structure, function, kinetics, and molecular modelling studies.

Chapter 3
Expression, purification and basic characterization of Sbh26GST-, Sh28GST- and Human GST-mu-testosterone interaction

3.1 Introduction

Structural and functional understanding of membrane proteins is crucial for rational drug design. Membrane proteins are accountable for interaction with approximately 70% of the available medicinal drugs. Howbeit, within the field of structural biology, the most complex systems for functional studies belong to the membrane proteins. Obtaining pure recombinant proteins is a challenging process which often relies on the use of genetically encrypted affinity tags to facilitate the purification process. From a technical perspective, the challenges can be overcome by the isolation and expression of an ear marked gene within heterologous organisms, thereby enabling functional characterization studies (Calvete, 2017). Expression of functional proteins within heterologous hosts has been a breakthrough in contemporary biochemistry since this technique conveys the desired DNA fragments to hosts (Gustafsson *et al.,* 2004).

The manufacture of heterologous proteins mainly involves the recognition and segregation of an earmarked gene by co-opting a specific vector and its' assembly into a recombinant network. In general, vectors consist of an origin of replication, a multiple cloning site (where the desired gene is placed), a selectable marker and endonuclease restriction sites (Wagner *et al.,* 2007). Among the host organisms exploited for protein heterologous expression, *E. coli* is popularly the organism of choice to perform bacterial modus operandi (Makino *et al.,* 2011). Rosani and Ceccarelli (2014) articulate the numerous advantages that the advances in working with recombinant proteins in *E.coli* have proffered to the biotechnology arena. These bacterial systems offer unparalleled efficacy due to their ability to be exploited, their swiftness, low cost and high yield (Terpe, 2006).

Protein solubility or insolubility has profound effects on the biophysical characterization and functional properties. Various protein characteristics are analyzed by use of different techniques which include: *in silico*, *in vitro* and *in vivo* methods. Howbeit, these analytical methods need stringent quality control for desired results. In this study, we expressed, purified and performed structural characterization in a bid to make better inferences with regards to protein-protein interactions and for discrimination between the structures of the parasite protein isoforms. Further, insights into the inhibition levels were also investigated by performing half-maximal inhibitory concentration (IC_{50}). This IC_{50} is a very useful approach for evaluating drug potency in high throughput drug discovery processes.

3.2 Materials

The Uniprot KB database was used to obtain the nucleotide sequences of Sbh26GST, Sh28GST, and Human GST-mu. These sequences were then aligned using Clustal Omega analysis, which confirmed their accurate and efficient alignment. An additional N-terminal 6×His-tag was added upstream of the GSTs open reading frame (ORF) sequence. Sequencing, oligonucleotide primer synthesis and confirmation of the plasmid sequence was done at Inqaba Biotec (South Africa). The vectors encoding the entire DNA sequence of the GSTs were acquired from Genscript. The following items were obtained from Sigma-Aldrich (USA): ampicillin, chloramphenicol, kanamycin, GSH, testosterone, IPTG, reduced L-glutathione, protein ladder, 2xYT medium, Coomassie Blue-G250 stain, glutathione-agarose resin, glutathione-agarose resin, 1-Chloro-2,4-Dinitrobenzene and ANS. SnakeSkin™ dialysis tubing and Pierce Protein Concentrators PES were procured from Thermo Fisher Scientific™, Massachusetts, USA.

3.3 Methods

3.3.1 Expression of the GST proteins

The *E. coli* glycerol T7 Express^{TM} containing the parasite GSTs were cultured overnight at 30°C at 230 revolutions per minute (rpm) in 2YT medium. The Sh28 and Human GST-mu vectors contained ampicillin resistance genes. Consequently, the presence of antibiotics during the expression of recombinant proteins allowed for bacterial cell selection with the pertinent plasmid (Depardieu *et al.,* 2007). The following day, the cultures were diluted in freshly prepared 2xYT medium before incubation at 220 rpm at a temperature of 37°C. A photographic density of 600 $AUcm^{-1}$ was achieved within 2 hours before the flasks were placed on icy water for approximately 10 minutes, followed by a 6-hour expression at 30°C at a speed of 230 rpm. The culture was centrifuged for 15 minutes at 5000x*g* to obtain a pellet weighing approximately 4 grams. An SDS-PAGE consisting of reducing sample buffer (12.5% (v/v) was used to analyse the over expression of the GSTs as shown in Figure 3.1. The pellet was resuspended in 1x PBS, pH 7.5, and supplemented with lysozyme (10 µg/mL) before subsequent storage at -20°C. The function of lysozyme was to lyse the cell through bacterial cell wall disruption (Swaminathan *et al.,* 2011).

3.3.2 Purification of parasite and Human GST-mu proteins

Obtaining pure protein is dependent on isolating a protein from disparate cell components by use of suitable high-resolving fractionation techniques such as affinity chromatography. The main principle underlying affinity chromatography depends on the ability of the molecule to bind to a ligand with high affinity. All impurities are subsequently eluted through the chromatographic column (Sheehan, 2009). Frozen cell suspensions of the GST proteins were defrosted at room temperature (20°C) in water and mixed for approximately 30 minutes on a rotor. Sonication in an icy water bath followed for 30 second intervals at 50 amps for 8 times with an ultrasonic processor. Centrifugation at 18000x*g* for 30 minutes at 4°C took place with an aim of obtaining a clear supernatant. The resulting soluble cell fractions were filter-sterilized to prevent clogging.

The principle underlying immobilized metal affinity chromatography (IMAC) lies in the coalition between a metal ion such as nickel on a stationary matrix and the amino acid side chains (Bornhorst and Falke, 2000). Histidine will exhibit a high affinity towards the metal ions due to the presence of the imidazole ring. As a result, the hexahistidine-tagged proteins are confined to the matrices prior to elution. Such elution is triggered by the presence of imidazole (Spriestersbach *et al.,* 2015). The GST supernatants were introduced into the column and permitted to pass through it. The liquid that passed through the column was gathered, and the column was subsequently rinsed using 20 column volumes of an equilibration buffer. This ensured that any commonly adhered protein was thoroughly removed. The bound proteins were eluted in a single elution step. After purification, the column was stored in 20% (v/v) ethanol. The supernatant, lysate, pellet, and purified protein fractions were separated on a glycine SDS-PAGE gel to analyze them (Fig 3.2).

Further, eluted fractions which contained pure GST proteins were blended and concentrated by spin columns [Pierce Protein Concentrators PES (polyether sulfone from Thermo Scientific, with a molecular weight cut off (MWCO) of 10K] to a final concentration of approximately 150 µM. Thereafter, the concentrated GST proteins were placed in a dialysis tube with a MWCO of 10 kDa. Dialysis was done against 1x PBS which contained 1.5 mM EDTA and 0.02% (w/v) sodium azide, pH 7.5 at 4 °C, overnight (16 hours). The dialysed GST protein samples were stored at 4 °C on ice and used for ensuing assays. The purification of proteins constitutes a fundamental aspect of protein research. The exploration of protein function, structure, and interactions is greatly dependent on the level of purity and quality achieved in isolating the target protein.

3.3.3 Qualitative analysis of the purified recombinant GSTs

The protein concentration was determined using a Jasco V-630 UV-VIS spectrophotometer (Jasco Inc., Tokyo, Japan) and applying the Beer-Lambert law:

$$A = \varepsilon cl \qquad \textbf{Equation 3.1}$$

In the formula, A represents the absorbance value of the sample, ε is the molar extinction coefficient (MEC) ($M^{-1}.cm^{-1}$), c represents the concentration of the sample in the solution (M), and l represents the path length of the light passing through the sample through the solution in cm's.

An MEC of 39 880 $M^{-1}cm^{-1}$ was used for the three GST proteins. To determine the concentrations, a linear regression was used on at least five points obtained froma serial dilution. The buffer-only signal was taken into account for all readings. The qualitative analysis of GSTs was carried out through SDS-PAGE, a widely recognized technique for separating proteins based on their ability to undergo electrophoresis. SDS, a detergent, helps to unfold proteins by breaking non-covalent bonds. Subsequently, the proteins move based on their size in an electrically charged field (Righetti, 2005).

3.3.4 Secondary structure analysis of recombinant GSTs

3.3.4.1 Far UV circular dichroism (CD) spectroscopy

CD has been extensively utilized for gaining valuable insights into the protein's structure, as well as the extent of changes in both ligand binding and structure. The secondary structure of both Human GST-mu and parasite GSTs was investigated using CD. This secondary structure arises from the harmonious hydrogen bonding among the atoms of the amino hydrogen of the polypeptide backbone and the carbonyl oxygen. In the far-UV region, the CD spectra of proteins are mainly concentrated around 200 nm to 240 nm and is influenced by the $\pi_o \rightarrow \pi^*$ (185-200 nm) and $n \rightarrow \pi^*$ (at 215-230 nm) energy transitions of peptide bonds (Kelly and Price, 2000). Incidence of circularly polarized light at a certain wavelength permeates through an asymmetric chromophore, resulting in the right and left sections of light being absorbed contrastingly. Therefore, far-UV CD is an outstanding tool that is used for protein secondary structure analysis due to the inherent ability of light interaction with chiral protein molecules.

The overall major secondary structural features of the Human GST-mu and parasite GSTs' far-UV CD was established spectrophotometrically by use of the Jasco J-1500 CD spectrometer (Jasco Inc., Germany). The parameters included 15 accumulations, with a one second response time, bandwidth of 5nm, data pitch of 1nm and a scanning speed of 200 nm/minute at 20 °C. The GST samples were previously dialysed in PBS at pH 7.5, at a final concentration of 5 µM before subsequent analysis in the presence and absence of testosterone. Filtered 1x PBS buffer was used for conducting the analysis. The plots were examined for secondary structure features as depicted by the troughs obtained in Figure 3.6. The GSTs secondary structure was evaluated in the absence and presence of testosterone.

Mean residue ellipticity values were obtained from spectra from non-enzymatic rates according to the following formula:

$$[\theta] = \frac{100 \times \theta}{cnl} \qquad \textbf{Equation 3.5}$$

In this equation, the ellipticity signal (measured in millidegrees) is denoted as θ, the protein concentration (in millimoles per liter) is represented as c, the number of amino acid residues in the protein chain is denoted as n, and the path length (in centimeters) is represented as l.

3.3.5 Tertiary structure analysis of recombinant GSTs

3.3.5.1 Intrinsic tryptophan fluorescence spectroscopy

The GSTs tertiary structure was characterized spectroscopically by means of intrinsic tryptophan fluorescence. Fluorescence analyses involve irradiation excitation and emission of protein molecules at specific wavelengths. Protein molecules contain aromatic amino acids, namely tyrosine, phenylalanine, and tryptophan, that are outstanding probes to ascertain the tertiary assembly of proteins due to their aromatic fluorescing properties (Lakowicz, 2006). Consequently, the underlying basis of fluorescence relies on the methods' functionality as a structural analysis tool owing to the transformation in the environment (whether it will be hydrophobic or hydrophilic), protein binding or folding. Hidden tryptophan's within a hydrophobic setting reveal an increased quantum yield emanating in elevated fluorescence intensity readings. In this study, the three-dimensional structure of the parasite GSTs and Human GST-mu was thereby characterized. The Sh28GST, Sbh26GST, and Human GST-mu proteins' intrinsic tryptophan fluorescence spectra was obtained. The spectrophotometer was installed with Spectra Manager software v1.5.00, and the measurements were taken at a temperature of 20°C using a quartz cuvette with a path length of 10mm. The experiment was conducted thrice with an average of ten accumulations, using a continuous scanning mode of 100 nm per minute. The protein was in 1x PBS (w/v) NaN_3.To analyse the outcome of testosterone on Sh28GST, Sbh26GST and Human GST- mu, 50 μM oftestosterone was incubated with 5 μM of the recombinant proteins for a period of half an hour. The acquired emission spectra were buffer corrected, slit widths of 5 nm were used for excitation, 2.5 nm used for emission, excitation wavelength at 295 nm and a collection range between 280 nm and 350 nm.

3.3.5.2 Extrinsic ANS-fluorescence Spectro polarimetry

Hydrophobic binding cavities were characterized by ANS (Figure 3.1) which is an anionic fluorescent dye with lowered fluorescence intensity within polar conditions. Notwithstanding, the presence of solvent-exposed hydrophobic clefts, displaces the dye and leads to changes in polarity resulting in the quantum yield ($\Delta\Phi$) increase. Such displacement provides important inferences into the binding capabilities of ligands. Electrons of ANS molecules were excited at approximately 395 nm. When ANS binds in a non-covalent fashion and is exposed to hydrophobic conditions, the fluorescence is increased tremendously, and the fluorescence emission wavelength experiences a blue shift of approximately 500 nm and 480 nm wavelengths (Gasymov and Glasgow, 2007). The applications of extrinsic ANS-fluorescence spectroscopy include analysis of enzyme active sites and exposure of protein aggregates, protein folding and investigation of surface hydrophobicity (Hawe *et al.,* 2008).

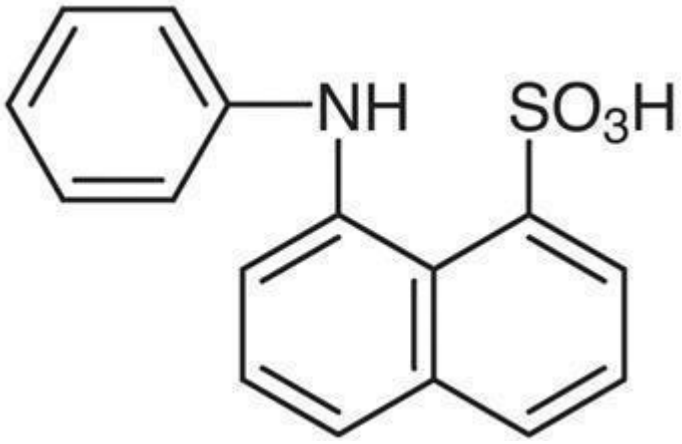


Figure 3.1 Structure of ANS dye.

The illustration depicts both the naphthalene ring and the sulfonate group. The naphthalene ring's reactivity is influenced by the polarity of its surroundings (Singh *et al.*, 2019).

To prepare ANS stock solutions with a concentration of 10 mM, 1x PBS buffer with a pH of 7.5, 1 mM EDTA, and 0.02% (w/v) NaN_3 were used. The concentration of ANS was determined using the Beer-Lambert law (Equation 3.1) by measuring its absorbance at 340 nm and using a theoretical molar extinction coefficient of 5000 $M^{-1}.cm^{-1}$. The ANS curves are displayed in Figure 3.7. The parasite GSTs and Human GST-mu samples (with a concentration of 5μM) were incubated in a fresh ANS solution for an hour, while wrapped in foil to prevent photolysis, to promote ANS binding to the GSTs. The spectra with or without testosterone was obtained. Signal monitoring was observed at wavelengths ranging from 400 to 650 nm at 20 ºC by use of a Jasco FP-6300 spectrophotometer that was linked to a Peltier temperature controller (Peltier from Jasco., Inc Germany). All obtained spectra was averaged and buffer corrected.

3.3.5.3 Extrinsic fluorescence: ANS binding studies

Samples of 5 μM of Human GST-mu, Sbh26GST and Sh28GST were incubated at 20℃ with freshly prepared ANS (100 μM) for sixty minutes. Subsequently, various testosterone concentrations (0- 100 μM) were titrated against the GST samples to investigate the ANS binding studies. The resulting emission spectra at 493 nm was fitted against various concentrations of testosterone (Figure 3.7).

3.4 Results

3.4.3 Expression of the GST proteins

In this study, parasite and Human GST-mu were expressed optimally within *E. coli* cells. Initially, sequences of these GSTs (previously inserted in pET-11a for Sh28GST and Human GST-mu and in pET-28a for Sbh26GST) substantiated the hexahistidine tag enriched nucleic acid arrangement. The dimeric molecular weight was found to be approximately 50 kilodalton (kDa) which agrees with the theoretical molecular weight. The pellet samples showed lower amounts of the GST amounts in relation to the supernatant samples which showed copious GST amounts. These results imply the lucrative over-expression of hexahistidine-tagged Sbh26GST, Sh28GST and Human GST-mus within the soluble fraction. Figure 3.1 shows the over-expression profile of the GSTs. The wet biomass obtained was approximately5 g per liter from each of the pellets of the proteins.

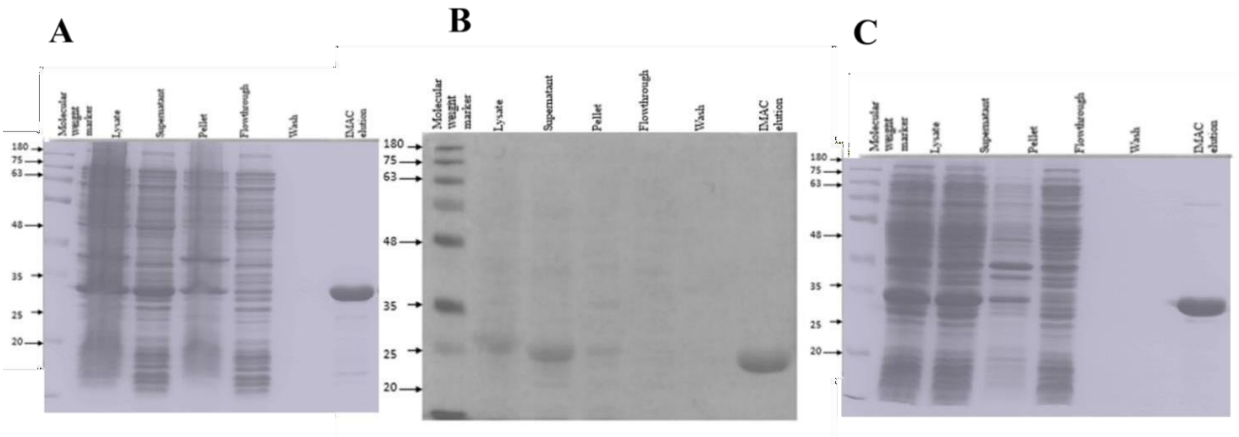


Figure 3.2 Expression and purification profiles of the GSTs using SDS-PAGE

A: Sbh26GST, B: Sh28GST and C: Human GST-mu. Gel electrophoresis on a 12.5% tris-glycine polyacrylamide gel. The first lane of each gel contains a molecular weight marker, lanes two to four shows the expression profiles whereby crude lysates of *E. coli* containing over expressed parasite and Human GST-mu, whilst the last lane in all GSTs depicts the respective purified protein. The polyacrylamide gels were stained with Coomassie brilliant blue dye.

3.4.4 Purification of the GST proteins

The soluble, recombinant Human GST-mu and parasite GSTs were purified by means of IMAC as previously described by Hornby *et al.*, 2000. All three GSTs had an additional hexa-histidine tag, (on the C-terminal) which shows a remarkable interconnectivity with the matrix-bound Ni^{2+} (Bornhorst and Falke, 2000). The integrity of Sh28GST, Human GST-mu and Sbh26GST was assessed by means of molecular weight determination, purity and homogenous assessments using SDS-PAGE analysis (Walker, 2002). All GSTs emerged as solitary bands with a molar mass of approximately 25 kDa on a 12.5% polyacrylamide gel.

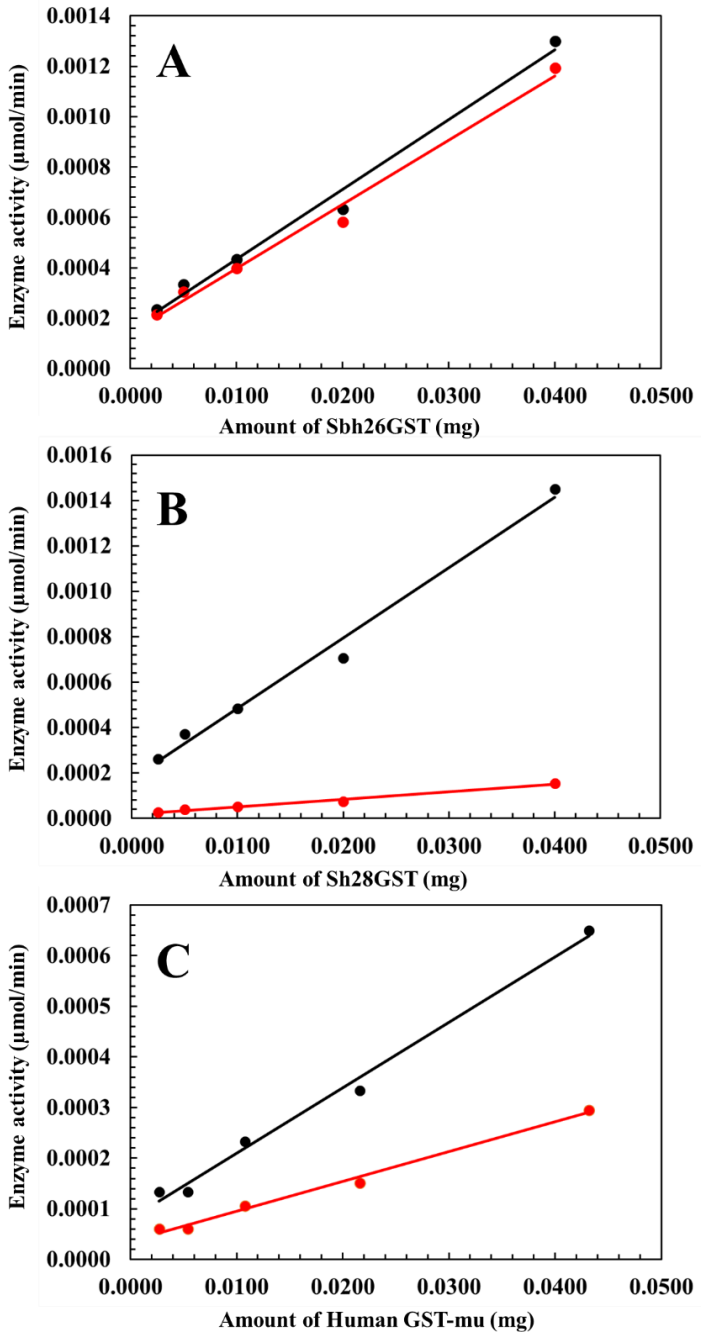


Figure 3.3 Specific activity profiles of parasite and Human GST-mu.

A: varying concentrations of Sbh26GST were analysed with (red) presence of testosterone and (black) without testosterone and slopes of 20 µmol $min^{-1}mg^{-1}$ & 0.14 µmol $min^{-1}mg^{-1}$ respectively. B: varying concentrations of Sh28GST were determined in the presence of testosterone (red) and absence (black) to obtain slopes of 16 µmol min-1mg-1 and 0.03 (µmol $min^{-1}mg^{-1}$) respectively C: Lastly, specific activity for Human GST-mu in the presence (red) and absence (black) was determined as 21 µmol $min^{-1}mg^{-1}$ and 20.5 (µmol $min^{-1}mg^{-1}$) respectively. All experiments were conducted in triplicate and the data was buffer corrected.

3.4.5 Secondary structural analysis of the Human GST-mu and parasite GSTs

Secondary structure assessment of the GSTs was done spectrophotometrically. The spectra (Figure 3.6) were collected between wavelengths of 180 nm and 250 nm courtesy of a temperature controller set at 20℃. Positive maxima and negative minima were observed at 208 nm and 222 nm for Sbh26GST (± testosterone), 206 nm positive maxima and 222 negative minima for Sh28GST (± testosterone) and positive maxima of 206 and a negative-minima of 224 nm for Human GST-mu (± testosterone) respectively. These results are typically indicative of the attributes of α-helical proteins.

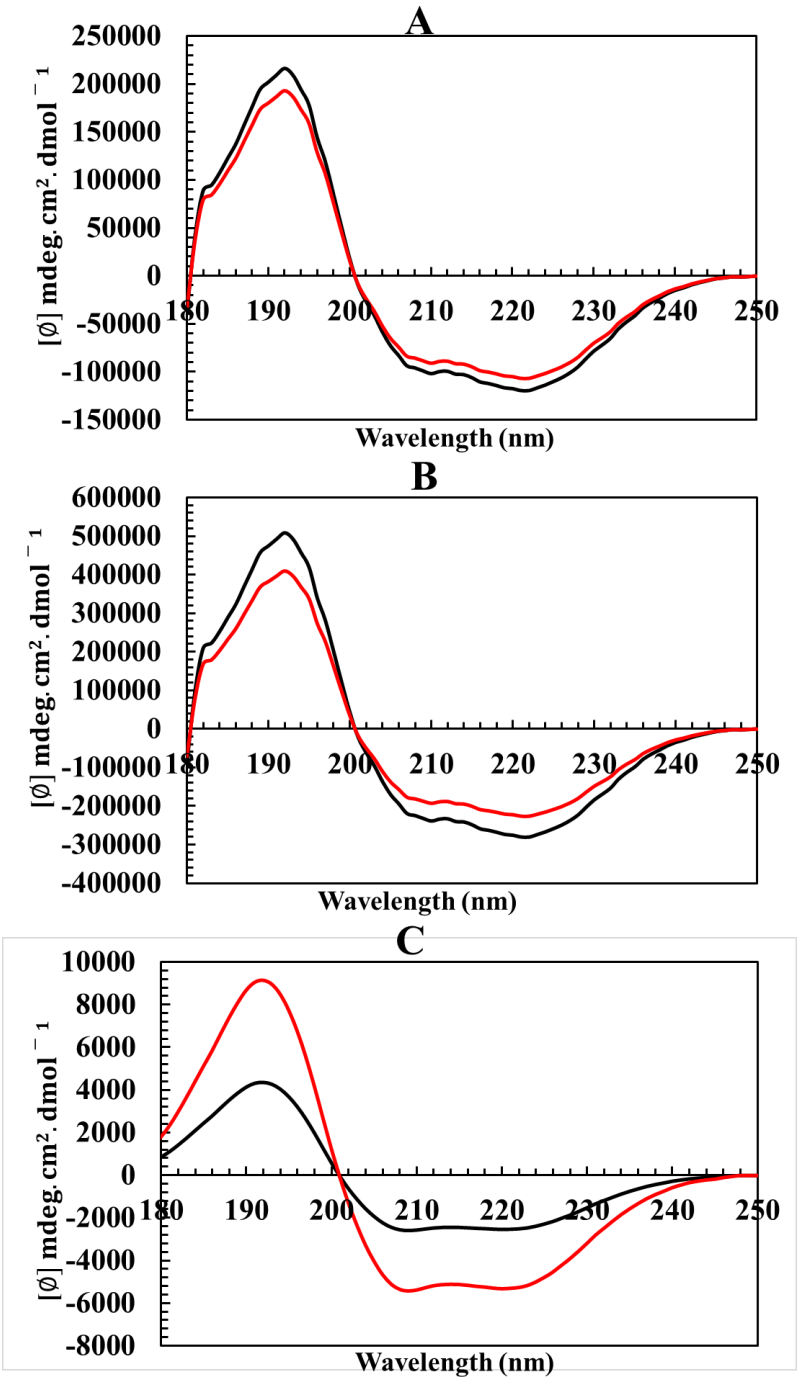


Figure 3.4 Secondary structural analysis of the GSTs.

A: Sbh26GST, B: Sh28GST and C: Human GST-mu, were resolved at 20ºC by far-UV CD spectrophotometry using 5 µM of ear marked protein in storage buffer. The presence of testosterone is denoted by red, whilst the absence of testosterone is denoted by black. All spectra were measured in triplicate and the CD spectra for all three GSTs revealed two eminent troughs which ranged from 206 nm and 226nm. Such troughs are characteristic of α-helical proteins.

3.4.6 Tertiary structure analysis of the parasite GSTs and Human GST-mu

The tertiary structure of the parasite GSTs and Human GST-mu was assayed by means of intrinsic fluorescence spectro polarimetry with or without testosterone. This assay was done in relation to the inherent tryptophan environment. The peculiarity of tryptophan's indole band enables excitation and emission at wavelengths of 295 nm and 350 nm respectively liable to the conversions brought about by a polar environment. The obtained spectra suggests that certain tryptophan molecules are exposed to the solvent, as evidenced by the peak emission wavelength of Sbh26GST at 341.5 nm. Likewise, Sh28GST mu showed increased intensity levels at 340nm. Lastly Human GST-mu recorded intensity readings at 341nm. There was no affirmation of nucleic acid adulteration nor aggregation with the protein concentrations that were used. However, when testosterone bound to all the GSTs the fluorescence emission spectrum was slightly lower (Figure 3.5).

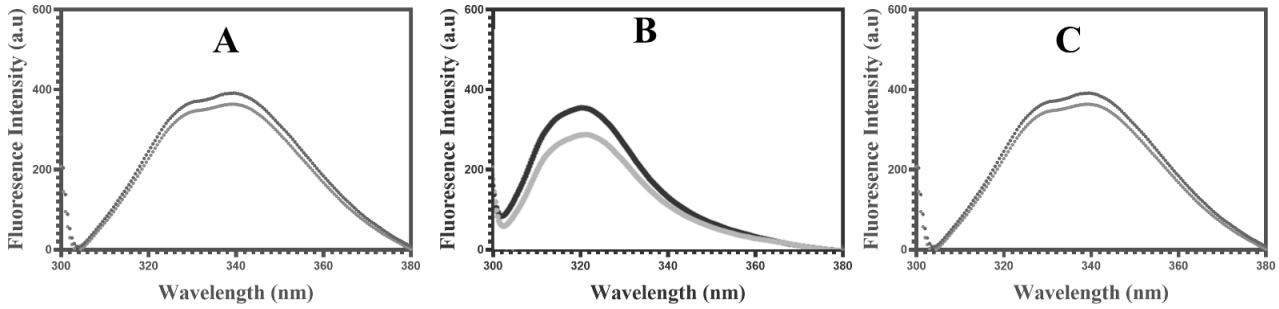


Figure 3.5 Intrinsic fluorescence emission spectra of the GSTs.

A: Sbh26GST B: Sh28GST C: Human GST-mu. The darker line illustrates the emission spectrum of the protein alone, whereas the lighter color tone denotes testosterone plus the protein. The experiments were conducted with 5 µM of the GST proteins in 1x PBS at a pH of 7.5.

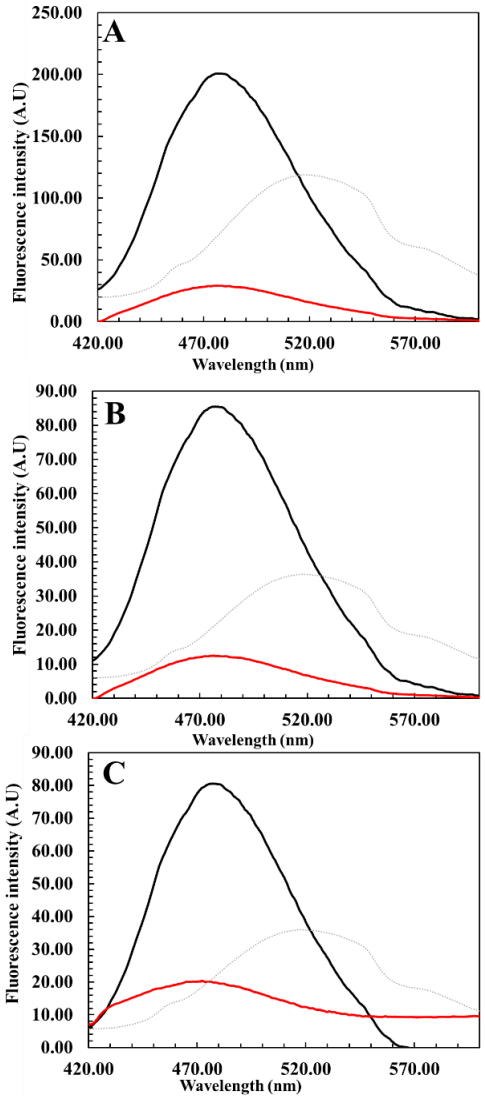


Figure 3.6 ANS spectroscopy.

A: Sbh26GST B: Sh28GST C: Human GST-mu. The black line illustrates the emission spectrum of the protein alone, whereas the grey color denotes testosterone plus the protein. Red represents free ANS. The experiments were conducted with 5 µM of the GST in 1x PBS at a pH of 7.5.

3.5 Discussion

Successful recombinant expression of parasite and Human GST-mu was accomplished. Considering the extensive research done on GSTs, the induction trials were not rigorous. A highly soluble protein was obtained. Acquiring a pure protein was extremely necessary for downstream functional characterization. Purification of the GSTs was ameliorated by charging the column with Ni^{2+} and attaching additional fusion hexa-histidine tags on the N-termini to facilitate retention of the GSTs (Bell *et al.*, 2013). Theoretical molecular weight and actual molecular weight on SDS analysis was contrasted due to several reasons. The first may include differences in sample composition, the presence of impurities, and the accuracy of the analytical methods used. Further, methods used in preparing the samples may also affect the results thereby giving inaccurate molecular weight of the sample. The GSTs were deemed stable since protein aggregation did not occur at high concentrations. Purified proteins with excellent yield and functionality are necessary for functional characterization studies such as the ability of proteins to absorb UV light. The capacity of proteins to assimilate light within the UV spectrum lies in the inherent structure of proteins, chemicals and surroundings. Concentration of GSTs were in accordance with the Beer-Lambert law as previously described in Equation 3.1.

The GSTs secondary structure was determined to ascertain if the purified proteins were appropriately folded. The obtained results of a positive maxima and negative minima of 208 nm and 222 nm for Sbh26GST (± testosterone), 206 nm positive maxima and 222 negative minima for Sh28GST (± testosterone) and positive maxima of 206 and a negative minimum of 224 nm for Human GST-mu (± testosterone) were anticipated since the obtained troughs (in the presence and absence of testosterone) are all characteristic of alpha helical proteins. Such far-UV CD profiles show that the cohesion of the polypeptide backbone of GSTs was not altered by testosterone after it was successfully recovered from the soluble fraction. Hence, secondary structural integrity was not disturbed by the interaction between testosterone and the GSTs.

Inference into the GSTs tertiary structure was possible through intrinsic tryptophan fluorescence. The underlying principle in tryptophan fluorescence relies on tryptophan's ability to fluoresce when light is incident on its indole ring. The wavelength of incident light ranges from around 280 to 295 nm. The tryptophan molecules assimilate this light and become excited to an increased energy state. Relaxation to the ground state ensues, resulting in emission of light at wavelengths of around 305-350 nm. The characteristic troughs observed are characteristic of tryptophan residues and can be

attributed to the four tryptophan residues belonging to Sbh26GST, of which Trp^{206} lies within the active site precinct. The principle behind tryptophan fluorescence is that, since tryptophan has an indole ring system, excitation by light of a particular wavelength, (typically in the range of 280-295 nm), is absorbed by the tryptophan molecule which gains energy and subsequently moves to a higher energy level. Upon return to the original state, light with a longer wavelength, (approximately 350 nm) is released. The magnitude of the spectra will depend on phenomena such as subjection to the solvent, existence of ligands, and adjacency to amino acid residues.

By analyzing the properties of the tryptophan fluorescence emission spectrum, researchers can obtain valuable information about protein structure and function. The successful binding of testosterone reduced the magnitude of the fluorescence emission spectrum for all GSTs. The reported emission maxima revealed that tryptophan residues exhibited solvent exposure tendencies (Lakowicz, 2006). Notwithstanding, such tendencies do not alter the proteins' global structure. Additionally, probing of hydrophobic patches within the GSTs was enabled via ANS spectroscopy. The modus operandi of this type of spectroscopy relies on the fact that ANS attaches to hydrophobic binding patches within proteins. ANS is an indispensable tool in exploring the changes in protein configuration. The spectra obtained for bound ANS showed reduced emission maxima. Results showed that these ANS clefts resulted in a blue shift due to testosterone binding to the parasite GSTs. This implies that the ANS binding is not entirely driven by hydrophobic interactions. Testosterone introduction into the GST-ANS solutions results in a significant reduction in the observed quantum yield. This decrease indicates that testosterone displaces ANS from its binding site (s).

In this study, the inhibition of the GST enzymes by testosterone suggests that testosterone reduces the capacity of the parasites to detoxify and thrive in the presence of xenobiotics. The susceptibility of the parasites is then increased, thus rendering them vulnerable to treatments with drug regimens.

Chapter 4
Effect of testosterone on the function and stability of GSTs

4.1 Introduction

The thermal shift assay evaluates the melting point, also known as the melting temperature (Tm), of a protein by determining the temperature at which 50% of its structure is denatured. In this chapter, it is explained how fluorescent dyes like Sypro Orange can interact with the hydrophobic pockets on proteins, causing a change in the dye's fluorescence properties. The efficiency at which this dye produces a quantum yield increases when it is attached to the uncovered hydrophobic regions of a protein that has been thermally denatured. By analyzing the fluorescent changes that occur in the presence and absence of testosterone, we gained insight into the stability and interaction of this steroidal hormone with parasite and Human GST-mu's.

4.1.1 Thermal shift assay

A thermal shift assay can be employed to gauge a protein's stability, which entails monitoring alterations in its temperature for denaturation. Various factors like changes in pH level, drug dosage, buffer composition, ionic concentration or genetic mutations can influence the situation. The study observed the GSTs activity using functional characterization techniques such as enzyme kinetic assays. Michaelis-Menten kinetics is one such well-known enzyme kinetic model within the field of biochemistry. This model is expressed through an equation that describes the rate of enzymatic reactions by linking it with the reaction rate:

$$y = y_{min} + \frac{y_{max} - y_{min}}{1 + \left(\frac{log\,[testosterone]}{IC_{50}}\right) - n}$$ **Equation 4.1**

In the context of this statement, y refers to the parasite and Human GST-mu activity (%), y_{max} denotes the highest level of enzyme activity, y_{min} indicates the lowest level of enzyme activity, IC_{50} represents the concentration of a substance that causes 50% inhibition of enzyme activity, and n represents the slope at the point of maximum curvature on the sigmoidal curve.

The Michaelis-Menten plot was used to determine the V_{max} and K_m.

$$v_o = \frac{(V_{max}[S])}{(K_M + [S])}$$

Equation 4.2

The symbol V_0 represents the starting rate of reaction in µmol/min, while [S] represents the concentration of the substrate in mM. This method involves using inhibitors to assess the protein's functionality by measuring its activity and kinetic parameters.

4.1.2 Specific activity of the parasite and Human GST-mu and IC_{50} analysis of testosterone

Both specific activity and IC_{50} was determined for each of the GSTs by means of the GSH-CDNB colorimetric assay, with or without testosterone at a concentration of 10 µM. A 3 ml quartz cuvette was used to contain 100ul of CDNB (in ethanol) and 2.9 ml of reaction buffer (0.001M EDTA, 0.1M NaH_2PO_4, 0.005M DTT, 0.02% (w/v) NaN_3, 0.001M GSH). Variable concentrations of GSTs were used (6.5 nM - 50 nM). The experiments were done in triplicate and observed for a minute at A_{340}. Protein activity can be observed through functional identification techniques such as IC_{50} analysis. Usually with such analyses, the inhibitor affects the activity of the enzyme. Specifically, the IC_{50} ascertains the strength of a drug in inhibiting a specific biochemical action. Herein, the efficiency of the drug is determined by quantification of the concentration of the inhibitor which is required to inhibit half (50%) of a given biochemical process (Neubign *et al.*, 2003). The IC_{50} determination of testosterone on parasite GSTs was done by assessing the enzyme's function when exposed to testosterone using the GSH-CDNB analysis. Noteworthy is the fact that CDNB is an excellent and universal GST enzyme activity probe.

4.2 Method

4.2.1 Thermal shift assay of the parasite and Human GST

The study examined how testosterone interacts with parasite and Human GST-mu, and how stable the interaction was. This was done using the thermal shift assay, which measures the melting temperature (T_m) of individual GSTs using SYPRO orange dye. The experiment was conducted using the Bio-Rad CFX96 Touch machine, which is an instantaneous PCR identification system. The protein sample was exposed to dialysis using a 10 kDa MWCO tube against 1x PBS at 4 °C for 18 hours (overnight). The 96 well plates comprised of 25 µL in total, which comprised of 20

μM of the GSTs, 0.5mM GSH, 10x SYPRO Orange dye and 25 μM testosterone (IC_{50}). The ratio of protein concentration to dye concentration in the mixture was 20 μM:10x. The plate was then sealed with a sealing film from Bio-Rad, vortexed, and centrifuged for 30 seconds before being analyzed using a Real-Time PCR Detection System.

The process involved measuring the melting curve betwixt 20°C to 80°C with temperature accruals of 0.5°C for 10 seconds. The samples were tested three times, and the results were averaged and adjusted to remove any signals from buffer and dye fluorescence. The resulting data, including relative fluorescence unit (RFU) and the derivative of the fluorescence signal with respect to temperature (-d(RFU)/dT), were exported to excel from CFX Maestro software. To obtain accurate results, the average RFU or -d(RFU)/dT was adjusted by subtracting the average blank. The degree to which the substance was unfolded was determined using the corrected average of the RFU, and this was plotted against the temperature. Similarly, the corrected average of (-d(RFU)/dT) was also plotted against temperature.

4.2.2 Steady state kinetics of the parasite and Human GST

The steady state kinetics experiment involved the analysis of GST samples (at constant protein concentration) in triplicate both with and without testosterone (ranging from 0 to 15 μM). Each data spot is the mean value obtained from conducting the experiment three times. The spectra were buffer-corrected before conducting any experiments, the pH level was modified to 6.5. Data fitting was done using Graphpad prism version 9.5.1(https://www.graphpad.com/).

4.2.2.1 Specific activity and IC_{50} determination of the parasite and Human GST

The IC_{50} was performed on the 2 parasite GSTs. Human GST-mu acted as a negative control. The experiment was done in triplicate using 50 μM of the parasite GSTs which had been previously incubated for 60 minutes in different concentrations of the steroid inhibitor. The resulting progress curves were observed by a spectrophotometer at 340 nm for one minute. Standard deviations and blank corrected progress curves were computed for every data point. The IC_{50} was acquired by plotting the activity values (expressed as a percentage) against the log testosterone concentration.

4.3 Results

4.3.1 Thermal shift assay

The thermal stability assay confirmed Human GST-mu to be the most stable amongst all the three proteins, followed by Sh28GST which was more stable in relation to Sbh26GST. Thermal stability intensified when the enzyme was bound to both GSH and testosterone. Figure 4.1 illustrates the combined unfolding curve of Human GST-mu, Sh28GST and Sbh26GST during thermal denaturation. At a temperature of 20°C, the GST proteins maintained their stability. However, as the temperature rises, the GSTs unfolded, as indicated by a sole changeover that occurred at around 58°C for all the three GSTs. As the hydrophobic regions become exposed, more of the SYPRO Orange dye will bind to these areas. Consequentially, there is an increase in fluorescence until the dye becomes saturated. Resultingly, the dye becomes quenched due to aggregation, resulting in a decrease in fluorescence, which causes the plot to become uneven. The presence of testosterone and GSH enhanced the stability of the GSTs as shown by a slight increase in Tm from 47°C to 48.5°C, from 51°C to 55.5°C and from 53°C to 57°C for Human GST-mu, Sh28GSTand Sbh26GST respectively.

4.3.2 Steady state kinetics

The K_m of Sh28GST and Sbh26GST towards GSH was investigated and found to be 4.2mM and 6.6 mM, with a V_{max} of 0.64 µmol $min^{-1}mg^{-1}$ and 0.42 µmol $min^{-1}mg^{-1}$ respectively also towards GSH (Table 4.1). These values are somewhat inflated as compared to other studies which found a K_m value of 1.42 mM for Lactuca sativa GST for CDNB (Park *et al.,* 2005). Our elevated K_m values in relation to other previous studies indicate a diminished affinity of the parasite GST enzymes towards CDNB. In the presence of testosterone, there was a 50% and 57% reduction observed in the K_m, V_{max}, and IC_{50} values against GSH (Figure 4.1), which indicates uncompetitive inhibition of activity. Unfortunately, the Michaelis-Menten steady state kinetic parameters with regards to CDNB in the presence of testosterone were recorded but not deemed accurate since the GSTs did not reach saturation at the analysed higher CNDB concentrations.

Table 4.1 Kinetic parameters for Sbh26GST and Sh28GST

	Enzyme	
	Sbh26GST	Sh28GST
Specific activity ($\mu mol\ min^{-1}mg^{-1}$):		
CDNB	16	20
Varied GSH concentration		
K_m (mM)	6,6	4,2
V_{max} (µmol/min)	0,42	0,64
Varied CDNB concentration		
K_m (mM)	1,26	2,13
V_{max} (µmol/min)	0,22	0,18

4.3.3 Specific activity and IC_{50} determination of the parasite and Human GST-mu

Individual slopes for all GST reaction curves provided inference into the initial velocity which was plotted versus the concentration of the enzyme. The specific activity (Figure 3.3) was obtained from calculating the activity of the individual enzyme concentrations of the parasite and HumanGST-mu against the GSTs' mass (mg). Generally, all GSTs showed remarkable levels of enzymatic activity; however, the distinction was vivid when the sex hormone was introduced to the parasite GSTs. Sh28GST demonstrated an enzyme activity rate of 20 µmol $min^{-1}mg^{-1}$ in the absence of the testosterone and a sharp decline in the presence of the inhibitor (0.03 µmol $min^{-1}mg^{-1}$).

Sbh26GST exhibited a rate of 16 μmol $min^{-1}mg^{-1}$ when the inhibitor was absent and 0.14 μmol $min^{-1}mg^{-1}$ when present. Human GST-mu exhibited 21 μmol $min^{-1}mg^{-1}$ in the absence of testosterone and barely changed (20.5 μmol $min^{-1}mg^{-1}$) when the inhibitor was introduced. The function of the parasitic GSTs was almost hindered when 0.5 mM testosterone was introduced. According to enzyme kinetics, testosterone is a powerful inhibitor of Sh28GST parasite, resulting in a specific function reduction from 16 μmol $min^{-1}mg^{-1}$ to 0.03 μmol $min^{-1}mg^{-1}$ and a testosterone IC_{50} of 20 μM. Similarly, Sbh26GST demonstrated a specific activity decrease from 20 μmol $min^{-1}mg^{-1}$ to 0.14 μmol $min^{-1}mg^{-1}$, and a testosterone IC_{50} of 23 μM.

The IC_{50} of testosterone towards parasite GSTs activity was established by use of the CDNB-GSH conjugation assay. The acquired IC_{50} value was within the μM span, with Sh28GST exhibiting a much smaller IC_{50} value of 20 μM (Figure 3.5). Sbh26GST displayed an IC_{50} of 23 μM. Here, analysis of the parasite GSTs (at a 50 μM concentration) was done at different testosterone concentrations. Data fitting was done by use of a regression analysis model (without linearity) which possessed a variable slope with four parameters.

Contrastingly, Human GST-mu scarcely inhibited the CDNB enzyme activity inhibition (Figure 4.3). The feasible explanation for such behavior, is due to identification of molecular targets such as Sbh26GST and Sh28GST in parasites whilst concurrently sparing the human host. Potency of the results are comparable to a similar study by Padi *et al.* (2021) where a potency of 13 μmol $min^{-1}mg^{-1}$ was found when bromosulfophthalein (BSP) was used as the inhibitor. By employing GSH-CDNB conjugation as a prototypical GST-mediated reaction, we illustrate that testosterone is an effective inhibitor of the parasite GST enzymes with IC_{50} values ranging from 16-22 μmol $min^{-1}mg^{-1}$. Obtaining the IC_{50} value at the outset was crucial, as it provided the foundation for planning subsequent studies on inhibition kinetics and thermodynamics for the testosterone-GST interaction.

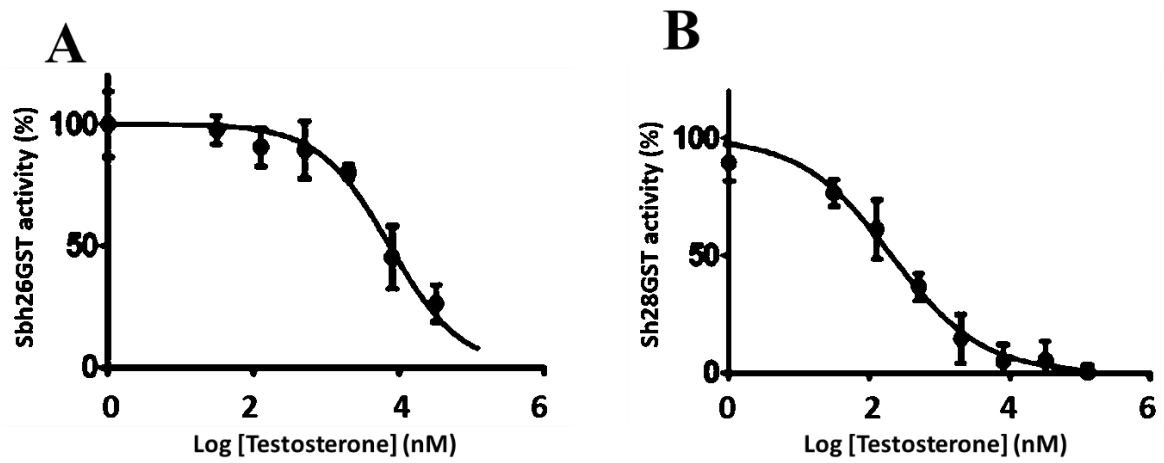


Figure 4.1 The IC_{50} values of testosterone

Inhibitory effect on A: Sbh26GST and B: Sh28GST were determined. Testosterone concentrations were varied and incubated with a 50 Nm concentration of the GSTs (for 30 min) in 1x PBS. Triplicate data of the protein activity was monitored up to a minute at an absorbance reading of 340 nm. The obtained log concentration of testosterone was plotted against the activity (expressed as a percentage).

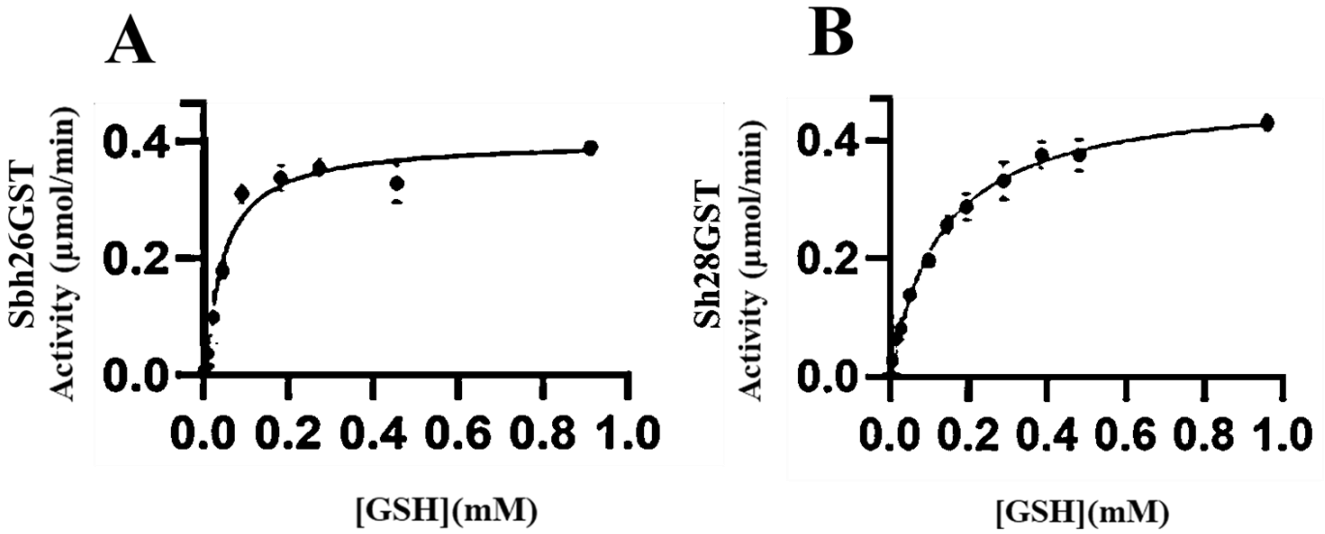


Figure 4.2 Steady state kinetic velocity plot of Sbh26GST and Sh28GST versus GSH concentration

(A) The kinetics of Sbh26GST-CDNB conjugation (B) Kinetics of Sh28GST-CDNB conjugation. All reactions were monitored at 340nm in the presence of 1 mM GSH and 1mM CDNB for 60 seconds as shown by the hyperbolic curve which eventually flattens out. The K_m and V_{max} values were obtained from the corresponding curves. The error bars depict the standard deviation from the replicates.

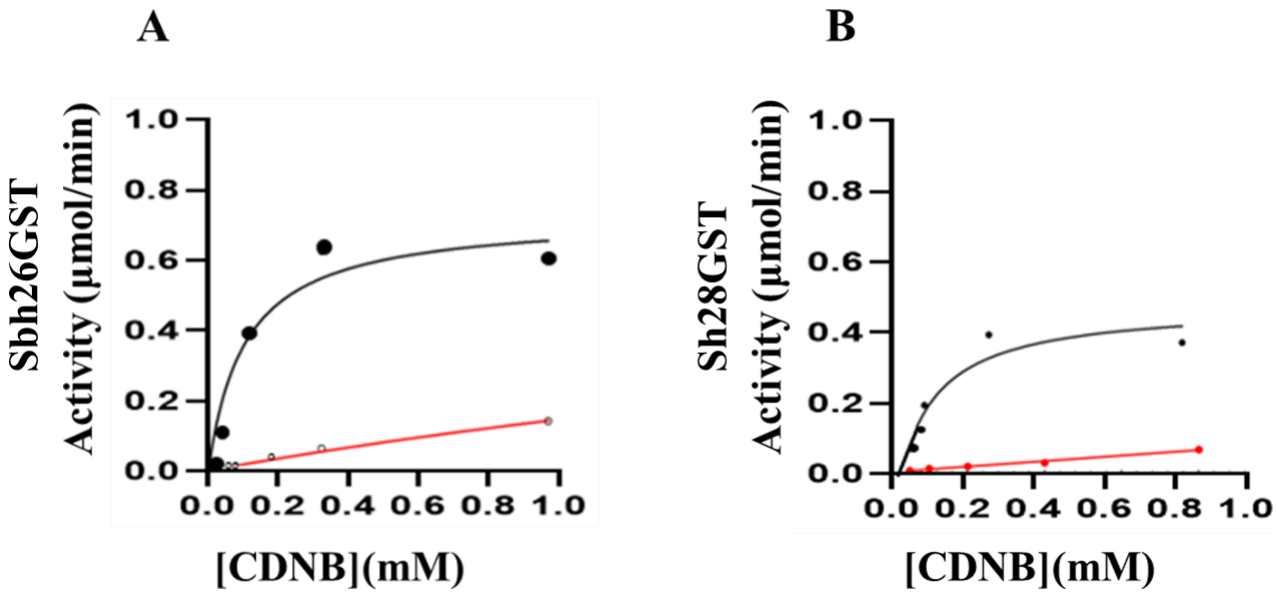


Figure 4.3 Steady state kinetic velocity plot of Sbh26GST and Sh28GST versus CDNB concentration in the presence and absence of testosterone

Hyperbolic curve for A. Sbh26GST and B. Sh28GST. The equation of the varied CDNB could not be determined as the experiment did not reach saturation. This experiment was performed in triplicate. All spectra were buffer corrected. Data fitting was done using Graphpad prism version 9.5.1 (https://www.graphpad.com/).

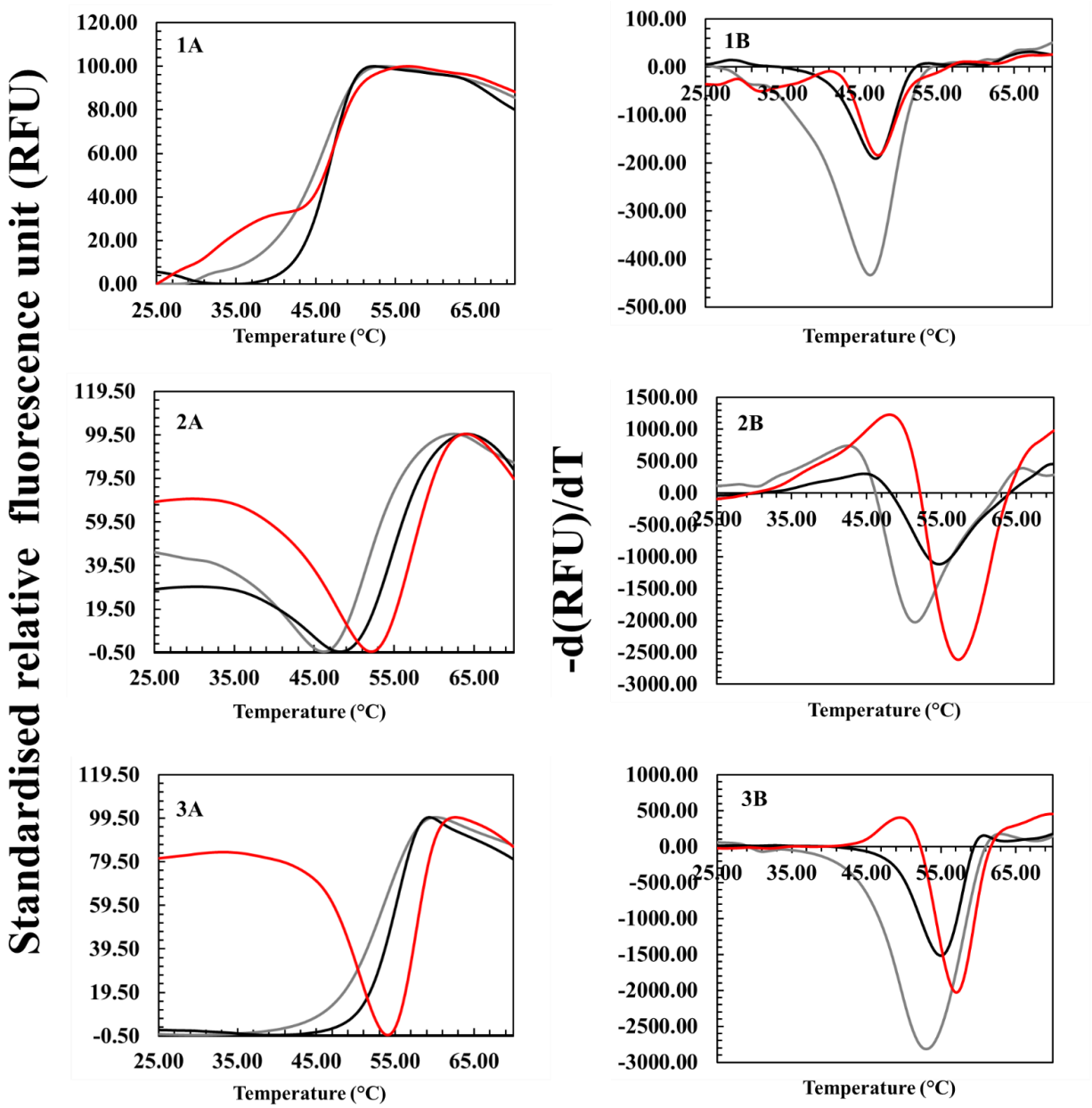


Figure 4.4. Thermal stability profile for Human GST-mu and the parasite GSTs.

Standardized RFU vs. temperature (◦C), are shown in 1A, 2A and 3A for Human GST-mu, Sh28GST and Sbh26GST respectively, whilst the obtained -d(RFU)/Tvs. temperature (◦C) are shown in 1B, 2B and 3B for Human GST-mu, Sh28GST and Sbh26GST respectively. The Tm (◦C) of each profile is shown by the minimum point. The grey profile represents the protein alone, the red profile represents testosterone and lastly the black profile represents the GSTs and testosterone combined. Data fitting was done using Graphpad prism version 9.5.1 (https://www.graphpad.com/).

4.4 Discussion

The fluorescence thermal shift assay used SYPRO Orange dye to monitor the denaturation of Sbh26GST, Sh28GST and Human GST mu as the temperature increased. The dye was chosen because it tends to bind to hydrophobic regions, which causes an increase in the dyes' quantum yield. The findings show that that testosterone had a stabilizing effect on the GSTs used in this study, particularly when GSH was present. These results indicate that testosterone could be manipulated to become a viable option for drug development, given the observed changes. These results may impact the enzyme's functionality through changes in conformational dynamics.

To ascertain the inhibitory strength of testosterone on the activity of Sh28GST, Sbh26GST and Human GST-mu, the IC_{50} was determined. The obtained results shows that testosterone is an ideal inhibitor. This IC_{50} value is an indispensable tool in drug discovery regimes since it infers the measure of potency to drug candidates. These results (23 μM for Sbh26GST and 20 μM for Sh28GST) are comparable to other published data but are somewhat elevated than the 10.8 μmol $min^{-1}mg^{-1}$ obtained by Kaplan *et al.* (1997). The elevated IC_{50} obtained from our study means our drug was less potent (in comparison to that obtained by Kaplan and co-workers) since a decreased IC_{50} value denotes a more powerful drug (Meyer *et al.*, 2019). Generally, pharmaceutical industries stipulate reduced IC_{50} values as more effective than their elevated counterparts. The obtained IC_{50} values of 20-23 μM prove that testosterone has an increased inhibitory effect on parasites GSTs, thereby posing as plausible anti-schistosomiasis drugtargets. Howbeit, IC_{50} data is not used in isolation during drug screening processes.

The K_m values were determined for both GSH and CDNB by using different concentrations of their respective substrates. Harwaldt *et al.* (2002) reported a significantly lower K_m value of 0.16 mM, implying a higher substrate affinity compared to the K_m values (4.2mM and 6.6 mM for Sh28GST and Sbh26GST respectively), obtained in this study. The K_m value for CDNB at varied concentrations could not be accurately determined as the reaction did not reach saturation. Harwaldt *et al.* (2002) estimated the K_m value to be >2 mM, but due to CDNB's strong absorption nature, concentrations higher than 1.5 mM cannot be tested.

Chapter 5
Computational modelling of Sbh26GST, Sh28GST and Human GST-mu

Abstract

Schistosomes primarily rely on Glutathione transferases (GSTs) as their main detoxification enzymes. These parasitic enzymes tend to undergo upregulation in response to drug treatment. Hence, Sbh26GST and Sh28GST could potentially be identified as therapeutic targets for schistosomiasis treatment. Molecular dynamic simulations were conducted to assess the stability of protein-ligand complexes, offering valuable insights into their molecular interactions. Our Molecular Dynamics Simulation (MDS) results corroborate our Induced Fit Docking (IFD) study, as the ligand remains securely bound within the receptors of the parasite GST, whereas it drifts out of the receptor pocket in Human GST-mu. In general, the Cα RMSD values of the enzymes indicate that their dynamic behavior is affected in the presence of water. This suggests that the binding of testosterone to parasite GST is allosteric, as none of the substrate binding sites are obstructed by testosterone. Instead, testosterone binds at the L-site, which seems to influence the enzyme's dynamic behavior crucial for GSH-CDNB conjugation. Current endeavors are directed towards the creation of a schistosomiasis vaccine, proposed as a more effective approach when used independently or in conjunction with Praziquantel to manage and eradicate this disease. This review aims to underscore these efforts.

Keywords: Computational modelling**,** Glutathione transferases, Praziquantel, schistosomiasis

5.1 Introduction

When crystal structures are not available, the contribution of bioinformatics and computational chemistry in the lingering battle against an effective and logical drug discovery is warranted. This chapter articulates the merits that bioinformatic studies offer as indispensable tools that are used in validating other biophysical and biochemical analyses. Computational modelling of proteins exploits homogeneous biological information such as sequences, structures and genomic

information from different sources of information. Families or groups of proteins tend to possess similar phylogenies (Marmier *et al.,* 2019). Structure and function of proteins can be inferred from the patterns within a protein family (Lin *et al.,* 2022).

The arrangement of amino acids in proteins is distinctive and defines the secondary, tertiary, and quaternary structures of proteins. Features such as the hydrophobicity, inter- and intramolecular interactions, and amino acid composition play an important role in determining protein structure (Hu *et al.,* 2021). The secondary structures include alpha helix, beta sheet and coiled coil (Rehman *et al.,* 2022). Online repositories such as the Protein Data Bank (PDB) (Armstrong *et al.,* 2019), UniProt (UniProt Consortium, 2019), SWISS-PROT (Boeckman *et al.,* 2003), and TrEMBL (Möller *et al.,*1999) provide unrestricted access to protein sequences. Experimental techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), and cryo-electron microscopy (cryo-EM) (Masrati *et al.,* 2021) are utilized to determine protein structure.

5.1.1 Computational protein structure prediction

A major challenge for computational modelling of proteins is to arrange amino acids in a low-energy conformation. Traditionally, an experimentally determined structure was used to fix the amino acid backbone (Kortemme and Baker, 2004). Recent advances now allow backbone formation by perturbation of the torsion angles (Xue *et al.,* 2022), loop conformation alteration (Kuhlman *et al.,* 2002), homology modelling, fold recognition and threading (Chen and Siu, 2020), and *de novo* fold design by coiled-coil topology parameterization or prediction of protein structure from first principles (Kuhlman *et al.,* 2003). With the vast addition of easily accessible protein data over the past decades, computational models involving proteins have developed rapidly to facilitate large-scale, high-throughput prediction as well as deep learning for problem solving (Bordin *et al.,* 2021; Hu *et al.,* 2021).

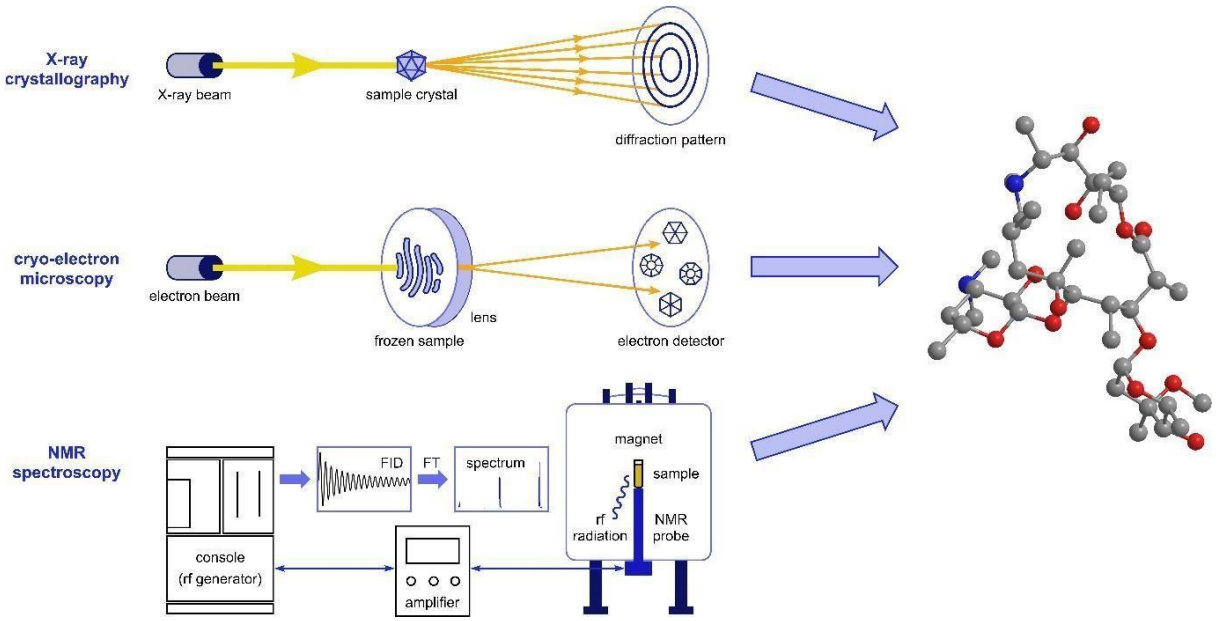


Figure 5.1 Basic principles of the 3 experimental methods of protein structure determination.

X-ray crystallography directs a beam of X-rays at the crystalline structure, causing them to scatter in various specific directions. By quantifying the angles and strengths of these scattered beams, a crystallographer is able to generate a three-dimensional representation of the electron density distribution within the crystal. Cryo-electron microscopy enables the examination of a specimen of interest at extremely low temperatures, known as cryogenicconditions. NMR spectroscopy observes the local magnetic fields surrounding atomic nuclei. (Jednačak *et al.*, 2020).

Multiple groups are able to use deep learning methods to model generation of fold-level accuracy of proteins lacking homologs in the PDB from sequence alignments (Abriata *et al.*, 2019; Yang *et al.*, 2019). AlphaFold and RaptorX are able to distance between amino acid residues by deep learning in residue-coevolution data (Evans *et al.*, 2018; Xu *et al.*, 2019). AlphaFold uses artificial intelligence to predict protein structure (Varadi *et al.*, 2022). In terms of health and disease, forecasting protein structures is crucial for comprehending which diseases affect an organism and for identifying new potential drug targets. It has been argued that simplified protein structure prediction can help "to understand and treat neglected tropical diseases, where research is currently under-resourced" (Stroe, 2023).

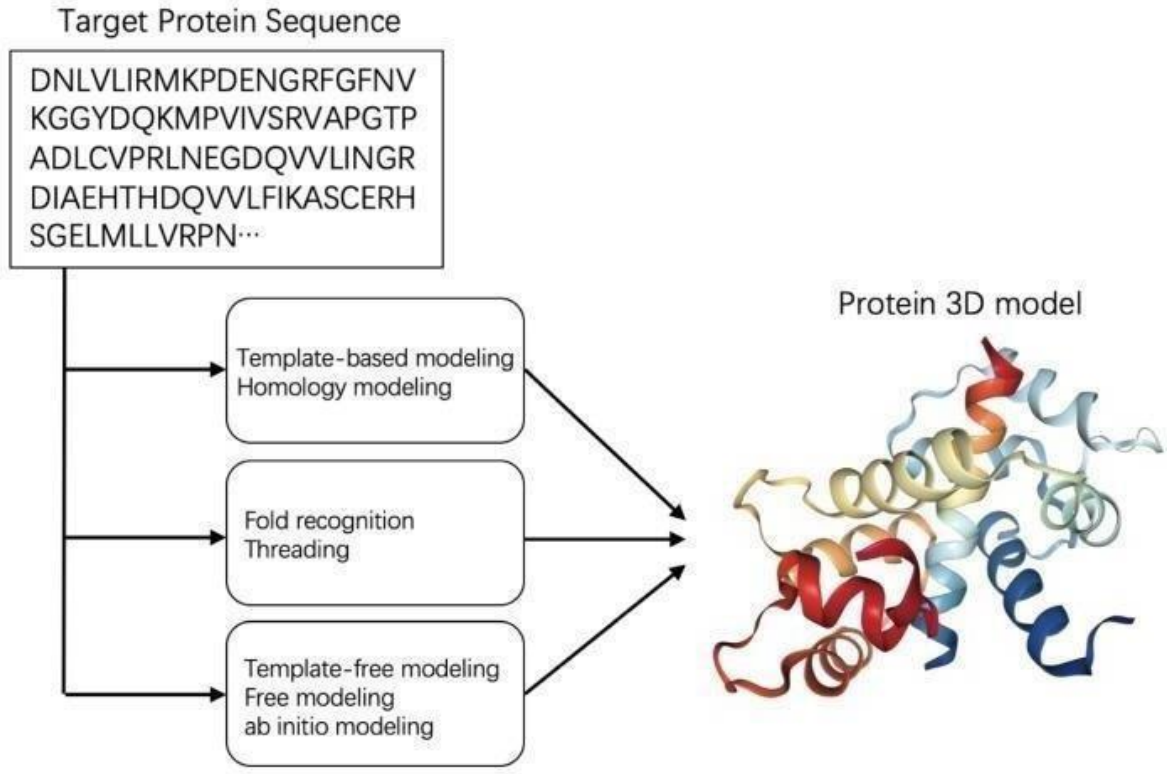


Figure 5.2 The 3 approaches to prediction of protein structure.
Template-based methods utilizes proteins that are homologous to the target protein and have known 3D structures as templates. Fold recognition is a protein modeling technique employed to predict the structure of proteinsthat share the same fold as proteins with known structures ,although lacking homologous proteins with confirmed structural information. Template-free modeling is a computational approach utilized to forecast the three-dimensional arrangement of proteins, independent of any reliance on templates derived from previously elucidated protein structures. (Chen *et al.,* 2020).

5.1.2 Protein-ligand interaction computational simulation

The true value of knowing protein structure is in understanding the functional characteristics of these proteins. Proteins are known to interact with and bind to ligands, and this requires a 3-dimensional conformation and sequence-specific molecular recognition. The binding site of a protein can attach to a ligand, the latter can be a molecule, ion, or protein. Functional characterization of proteins depends solely on the interaction with the ligand at the binding site.

Predicting ligand binding properties is the first step in drug discovery (Kandel *et al.,* 2021; Rube *et al.,* 2022) as it bypasses expensive and tedious *in vitro* investigations for screening potential

drug targets. At its core, drug discovery involves drug-target interaction, drug-target binding affinity, drug-target interaction sites, and drug bioactivity on proteins (Zhao *et al.,* 2022). Multiple methods have been used for binding site prediction. The traditional methods use geometry. Modern methods employ machine learning, deep learning, probabilistic mixture models and high-dimensional embedding (Alipanahi *et al.,* 2015; Toivonen *et al.,* 2018; Yuan *et al.,* 2019). The traditional methods include the geometry-based Fpocket which uses Voronoi partitions and solvent-accessible surface spheres (Le Guilloux *et al.,* 2009), Cartesian grid-based LIGSITE (Hendlich *et al.,* 1997) and POCKET (Levitt and Banaszak, 1992) which use solvent-accessible grid points, energy-based Alignments of Small Molecules Interface Finder and SITEHOUND (Ghersi and Sanchez, 2009) which exploit molecular interaction fields, geometry-based ConCavity (Capra *et al.,* 2009) which combines evolutionary sequence conservation, and template-based COACH where a support vector machine is used to predict the binding pocket (Yang *et al.,* 2013). Machine learning-based methods such as P2Rank (Krivák and Hoksza, 2018) and ProBound, and deep learning methods such as DeepSite (Jiménez *et al.,* 2017), kalasanty (Stepniewska-Dziubinska *et al.,* 2020), DeepSurf (Mylonos *et al.,* 2021) and DeepPocket (Aggarwal *et al.,* 2021) make use of algorithms and neural networks respectively.

Machine learning and deep learning for drug target discovery routinely make use of protein-ligand interaction (PLI) datasets. The typical workflow involves retrieval of protein-ligand pairs from PLI databases, training of machine learning using PLI feature vectors/matrix from biological, topological and physiochemical information, and finally testing of the trained model.

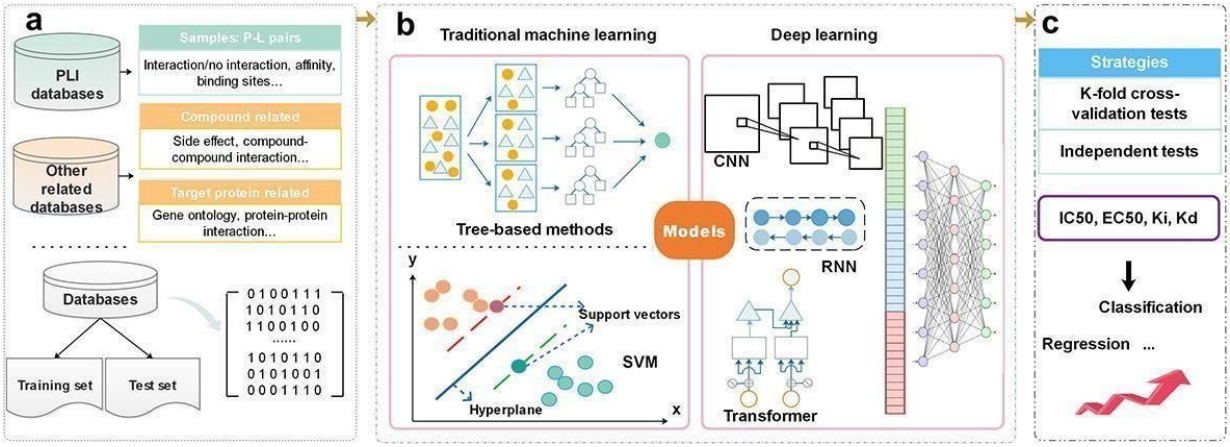


Figure 5.3 Machine learning methods workflow for protein-ligand interaction prediction.

Benchmark data is collected and preprocessed. The PLI framework is built, and the model is trained before evaluation of the model (Zhao *et al.*, 2022).

There are four main categories into which computational protein-ligand interaction methods can be classified: ligand-based, structural, network-based, and feature-based. Ligand-based methods assume that chemically similar ligands have similar biological activities and protein targets; and therefore, they need known protein ligands to compare candidate molecules (Ahneman *et al.*, 2018; Balakin *et al.*, 2002). Structural methods simulate PLIs by molecular docking using the 3-D structure of proteins and ligands to score the conformation (Levin *et al.*, 2017; Zong *et al.*, 2017). Network methods use biological networks and graph theory to compute the relationship model between proteins and ligands as a bipartite network (Ding *et al.*, 2019; Yildirim *et al.*, 2007). Feature-based methods use a machine learning framework to train the model from raw data (Bijral *et al.*, 2021; You *et al.*, 2019).

5.1.3 Induced-fit molecular docking (IFMD) simulations

The induced-fit model hypothesizes that proteins undergo conformational change upon binding. Molecular docking using the induced-fit model is popular because it exploits the conformational flexibility of the binding site and generates a protein-ligand complex in an optimized 3-dimensional conformation from which mode of binding can be inferred.

5.1.4 Molecular dynamics (MD) simulation

The aim of molecular dynamics (MD) simulations is to anticipate protein movements by utilizing Newtonian physics approximations for simulating atomic motions. After building a protein-ligand binding model, the estimation of forces acting on the atoms in a system is achieved through parametrization that mimics the motion behavior of real molecules using force field parameters. To represent the chemical bonds and atomic angles in a system, a common approach is to use simple springs, non-bonded forces, and electrostatic interactions, as explained by Durrant and McCammon in 2011.

5.1.5 Ligand motion

The protein that is approached by a ligand is in constant motion in solution. The ligand itself may also induce conformation changes in the protein for maximum accommodation in the binding pocket. Therefore, binding cannot be simulated with static structures (De Vivo *et al.,* 2016). The movement of atoms across time can be computed using Newton's equations of motions for predicting each atomic spatial position as a function of time, that is,

$$\frac{d2ri(t)}{dt2} = \frac{Fi(t)}{mi} \qquad \textbf{Equation 3.5}$$

In this context, the symbol Fi (t) represents the force that acts upon atom, i at a given time t, while ri (t) is the vector position of atom i at that same time t. Additionally, mi refers to the mass of atom i. In essence, a 3-dimensional trajectory describing the atomic position and motion at each timepoint can be created and controlled (Hollingsworth and Dror, 2018). Hence, molecular dynamics is valuable for investigating the conformational dynamics of ligands and their target proteins (Salo-Ahen *et al.,* 2021).

5.2 Methods

5.2.1 Induced-fit model docking simulations

Docking using the induced-fit model was executed in Maestro v12.2 (Schrödinger Release 2023-1: Maestro, Schrödinger, LLC, New York, NY, 2021). The molecular configuration of testosterone (compound CID: 6013) was acquired from PubChem (https://pubchem.ncbi.nlm.nih.gov/; Kim *et al.,* 2023) and underwent energy minimization through the LigPrep tool using the OPLS_2005 force field. An algorithm was programmed to produce likely pH 7.0±2 conditions for testosterone.

The algorithm implemented the Epik module to forecast the pKa at the set pH, and the testosterone underwent desalination at pH 7.0 ±2, with some chiral centers being modified to yield structures with low-energy states.

The Sh28GST (ID: 8ALS) was sourced from PDB (https://www.rcsb.org/; Herman *et al.*, 2000). The proteins were pre-processed using the Protein Preparation Wizard tool. The process involved in assigning bond orders and hydrogen atoms to a protein structure also entailed assigning zero bond orders to disulfides and metal groups and removing any water molecules found within a 5 Å distance between heteroatoms. Ionizable residues were optimized for charged states using the PROPKA tool (Li *et al.*, 2005) with the OPLS_2005 force field, with an RMSD cut-off of 0.3 Å by water molecule alignment sampling. The IFMD simulation for each protein utilized the GST and testosterone structures, with each protein dimer serving as a binding site. The best docked structure was selected based on the pose with the lowest IFMD score (i.e., lowest glide score and glide emodel), with low glide scores indicating higher affinity binding predictions.

5.2.2 Molecular dynamics simulations

To characterize the biophysical behavior of recombinant Sbh26GST, Sh28GST, and Human GST-mu, both empirical studies and computational modeling were utilized. In addition to empirical studies, molecular dynamic (MD) simulation was used as a complementary approach to further understand the dynamics of the GSTs in the presence and absence of testosterone. The MD simulation technique was also used to investigate the dynamics of ligands within the receptor site and substrate dynamics. The assessment involved measuring specific parameters, such as the root mean square deviation (RMSD), protein alpha carbon atoms (Cα), Root-mean-square fluctuations (RMSF), solvent accessible surface area (SASA) and radius of gyration (Rg) of the ligands.

The codes for the best docked structure of testosterone on Human GST-mu, Sh28GST and Sbh26GST, were written in Ubuntu/Linux operating system and submitted for MD simulations in Maestro v12.2. The MD simulations were conducted on a GPU-enabled *Schrödinger* Desmond machine.

The complexes were solvated with the OPLS_2005, neutralized by computationally conditioning to 0.15M NaCl, and equilibrated for potential and kinetic energy forces. To simulate solvation, the explicit solvent model TIP3P was employed along with the OPLS_2005 force field. The testosterone/GST complex was positioned in a water box that had an orthorhombic symmetry, with a minimum distance of 10 Å between the outermost atoms of the complex and the edge of the box. Each complex was centralized in its box and the box volume was minimized. Ionisation with Mg^{2+} was conducted to physiologically condition the complex for ion binding. The solvated and ionized complexes were submitted for MD production phase (Achilonu *et al.,* 2020; Pooe *et al.,* 2021).

Dynamic analysis and visualization were performed in Maestro and PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). MD simulation was performed by detection of parameters in a solvated complex. Conversely, molecular dynamics simulation of the system was carried out under various conditions. First, an NVT simulation was performed at -263°C with restraints on the solute-heavy atoms using 2 ps parameters of 100 and 12. Next, an NPT simulation was conducted at -263°C with restraints on the heavy atoms, followed by a simulation where pocket solvation was excluded for 12 ps. Another NPT simulation was then run at -263°C without any restraints on the heavy atoms. Finally, a long-range simulation was performed at a constant temperature of 26.85°C for 500 ns (Achilonu *et al.,* 2020; Pooe *et al.,* 2021).

Maestro v12.2 was used to perform a post-dynamic trajectory analysis, and the Simulation Quality Analysis algorithm was utilized to assess the simulation's quality. The values for temperature (°C), mean, standard deviation, and slope (ps-1), potential and total energy (kcal/mol), and volume (Å3) were obtained.

5.2 Results

According to the 2D molecular interaction plot of the IFD shown in Figure 5.4, the interaction between the ligands and amino acid side chains was mainly stabilized by van der Waals and hydrogen bonding interactions.

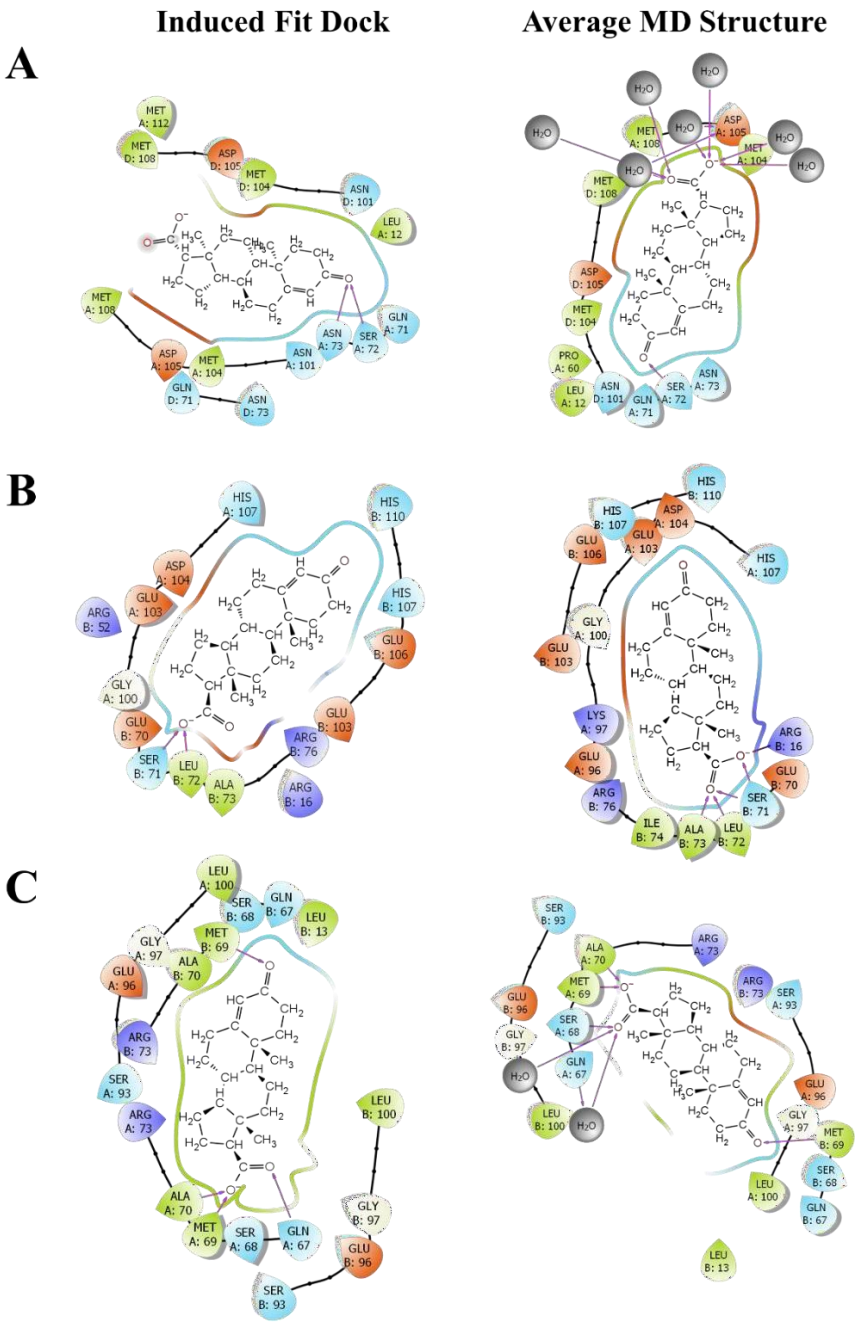


Figure 5.4 Schematic of the Induced Fit Dock (IFD) and Average MD Structure of testosterone and the GSTs.

(A) Sbh26GST (B) Sh28GST and (C) Human GST-mu. The amino acid residues are color-coded based on charge: charged (orange), polar (blue) or nonpolar (green). H-bond interactions are shown by the purple arrows. Average MD Structures show the water molecules and the ligand groups with which they interact, since MD simulations contain explicit solvent molecules. Ligand interactions were modelled in Maestro v12.2.

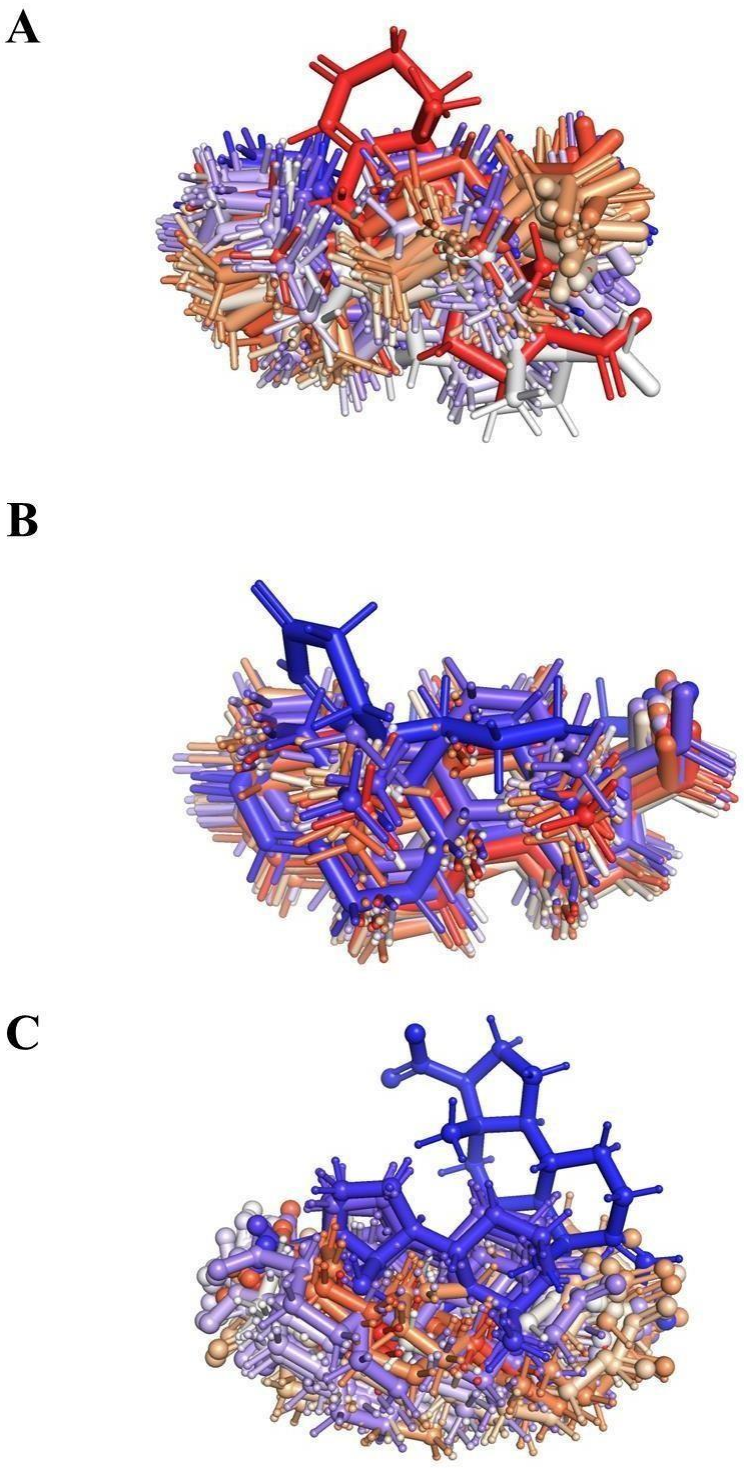


Figure 5.5 Motions of testosterone.

(A) Sh28GST, (B) Sbh26GST and (C) Human GST-mu in the MD simulation. The conformations of testosteroneexhibit the least deviation (RMS) in the Sbh26GST system, with both the Sh28GST and Human GST-mu systemsexhibiting greater deviation of roughly equal magnitude.

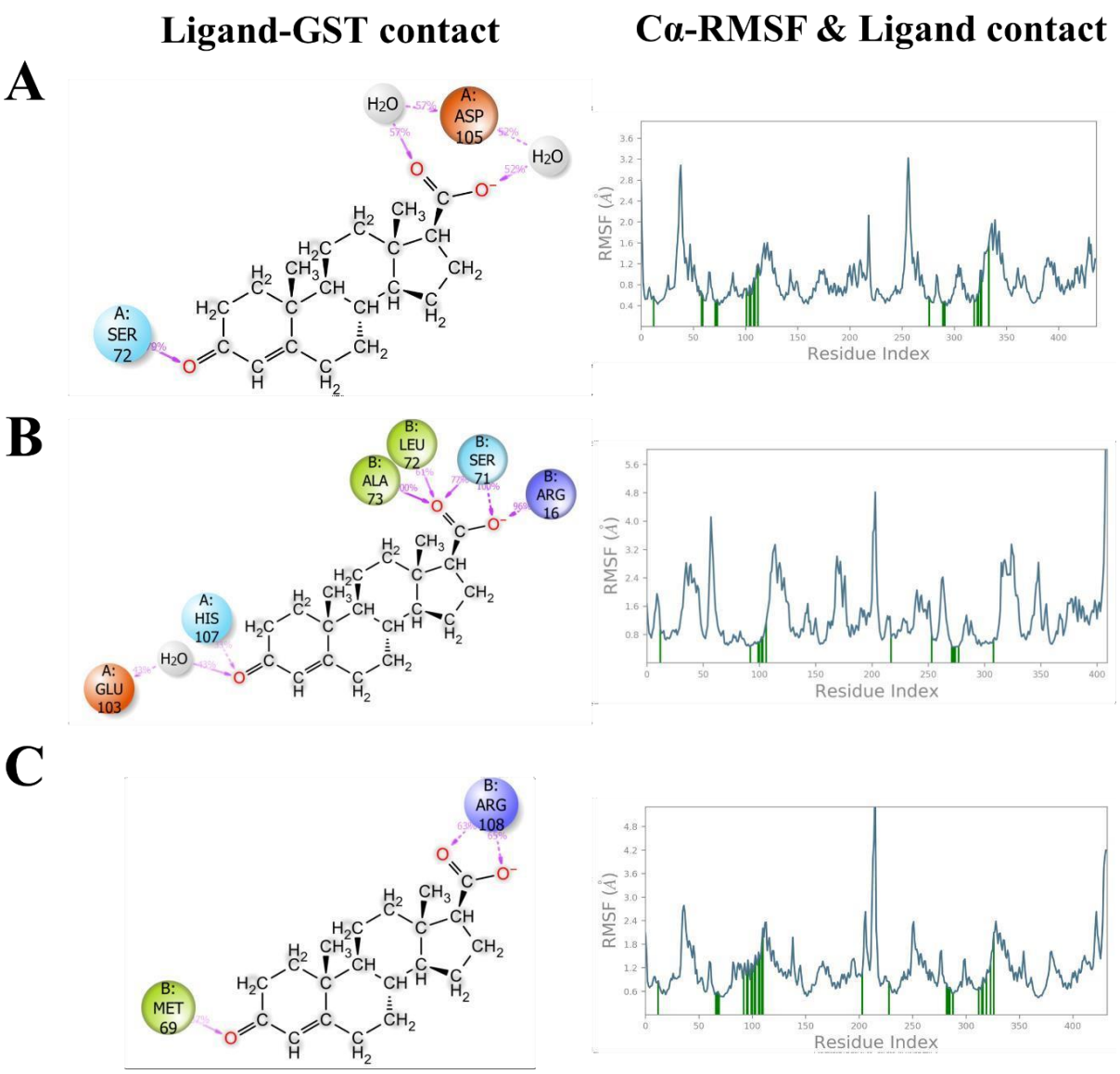


Figure 5.6 Schematics of testosterone-GST contact and RMSF values for amino acid residues.

(A) Sh28GST, (B) Sbh26GST and (C) Human GST-mu. The testosterone-GST contact figures show the most important interactions present between the GSTs and testosterone. The frequencies show the percentage appearance of the interaction throughout the simulation. The right-hand-side shows RMSF plots based on α-carbon atoms for each GST. The green lines depict the positions of residues involved in the interactions.

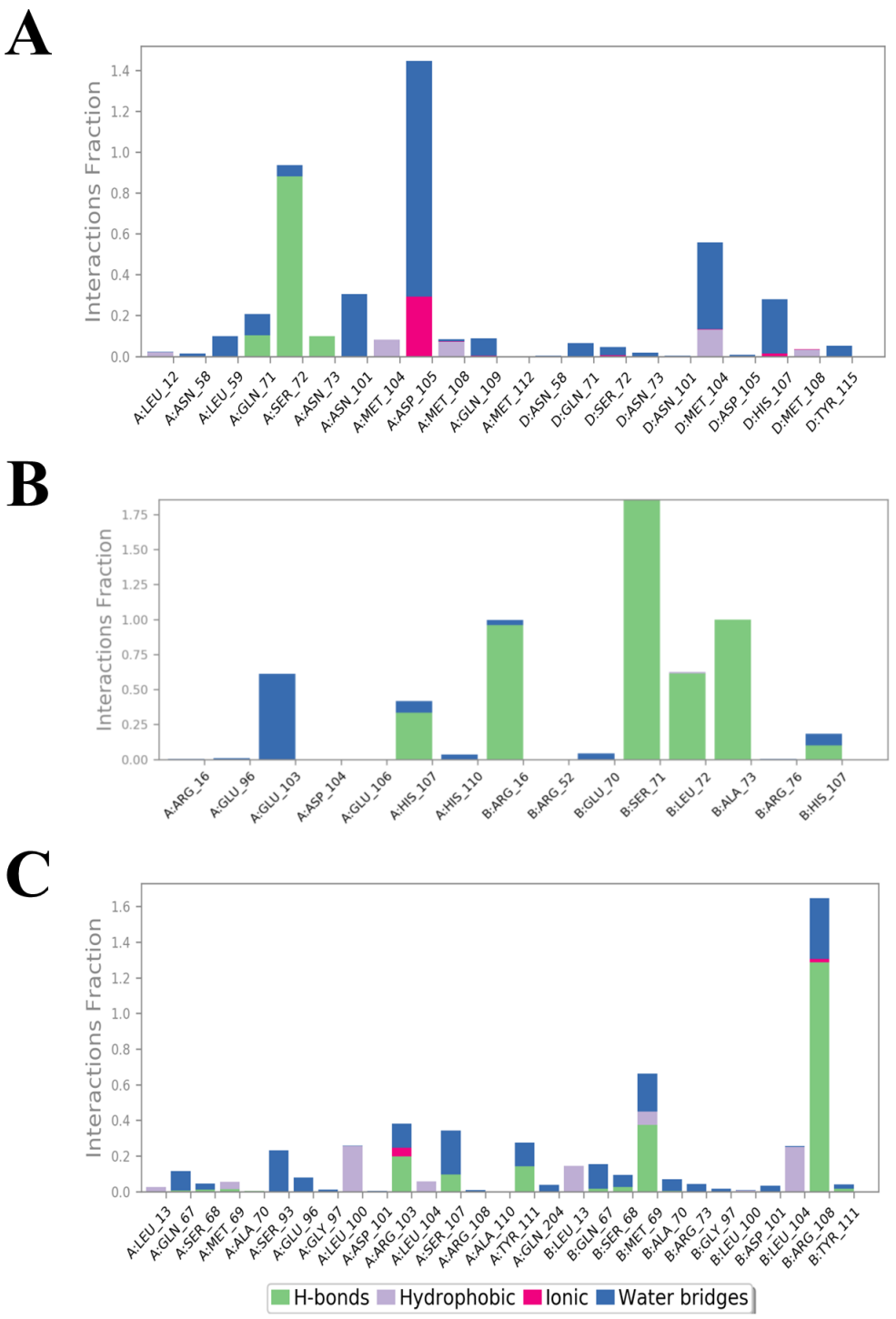


Figure 5.7 Bar graphs of the proportion of interaction of each of the interacting residues.

(A) Sh28GST, (B) Sbh26GST and (C) Human GST-mu. The interaction proportions are color-coded to signify the type of interaction present. All the data was extracted from SID™ from the Desmond™ (Schrödinger Release 2023-1)

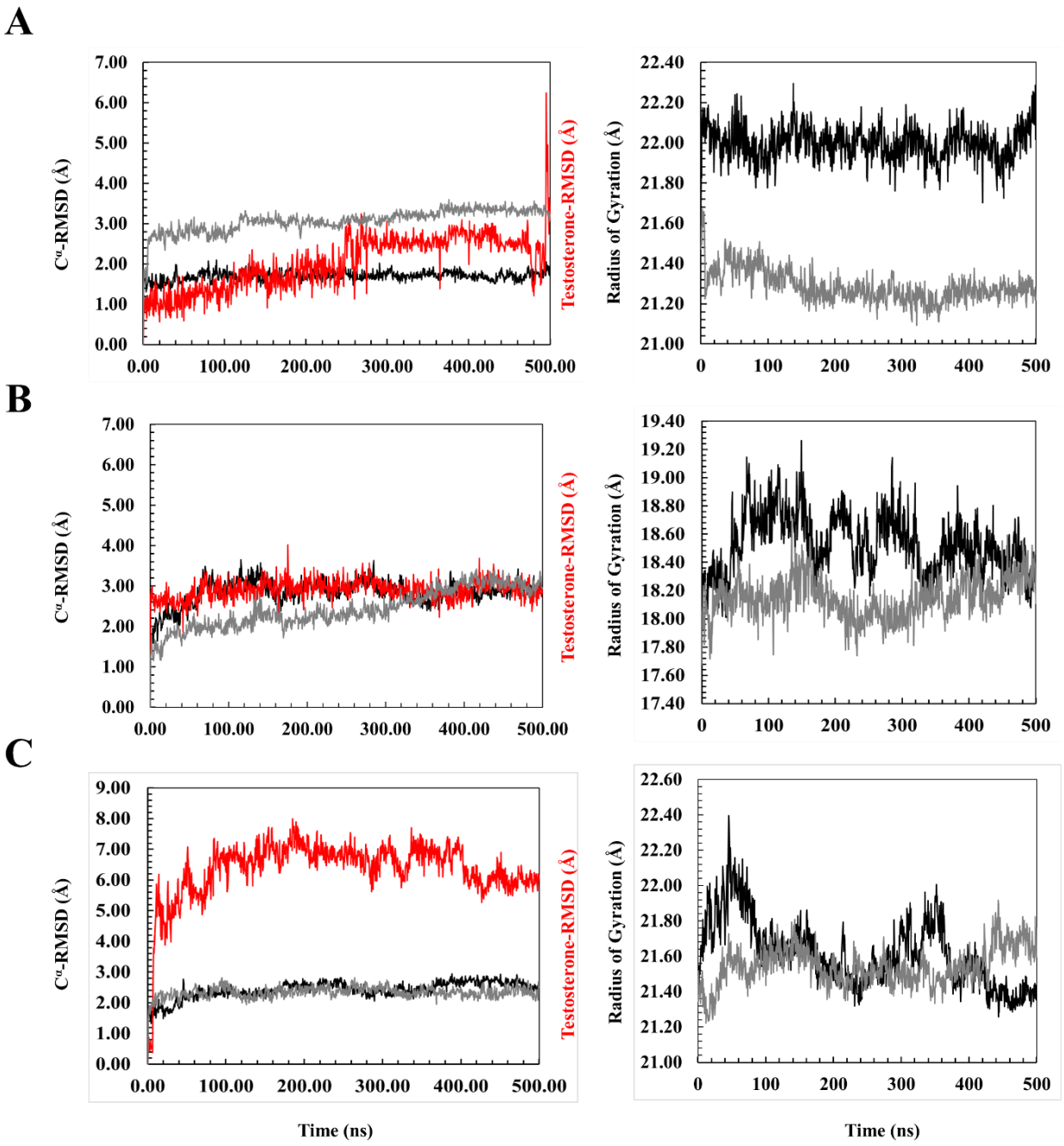


Figure 5.8 Plots of the RMSD and radius of gyration.

(A) Sh28GST (B) Sbh26GST and (C) Human GST-mu testosterone complexes. The testosterone RMSD is shown in red, the α-carbon RMSD of the target is shown in black, and grey is the apo target. The radius of gyration indicates how compact the target structure is. Black (complexed) is compared to apo (grey). All the data was extracted from SID™ from the Desmond™ (Schrödinger Release 2023-1)

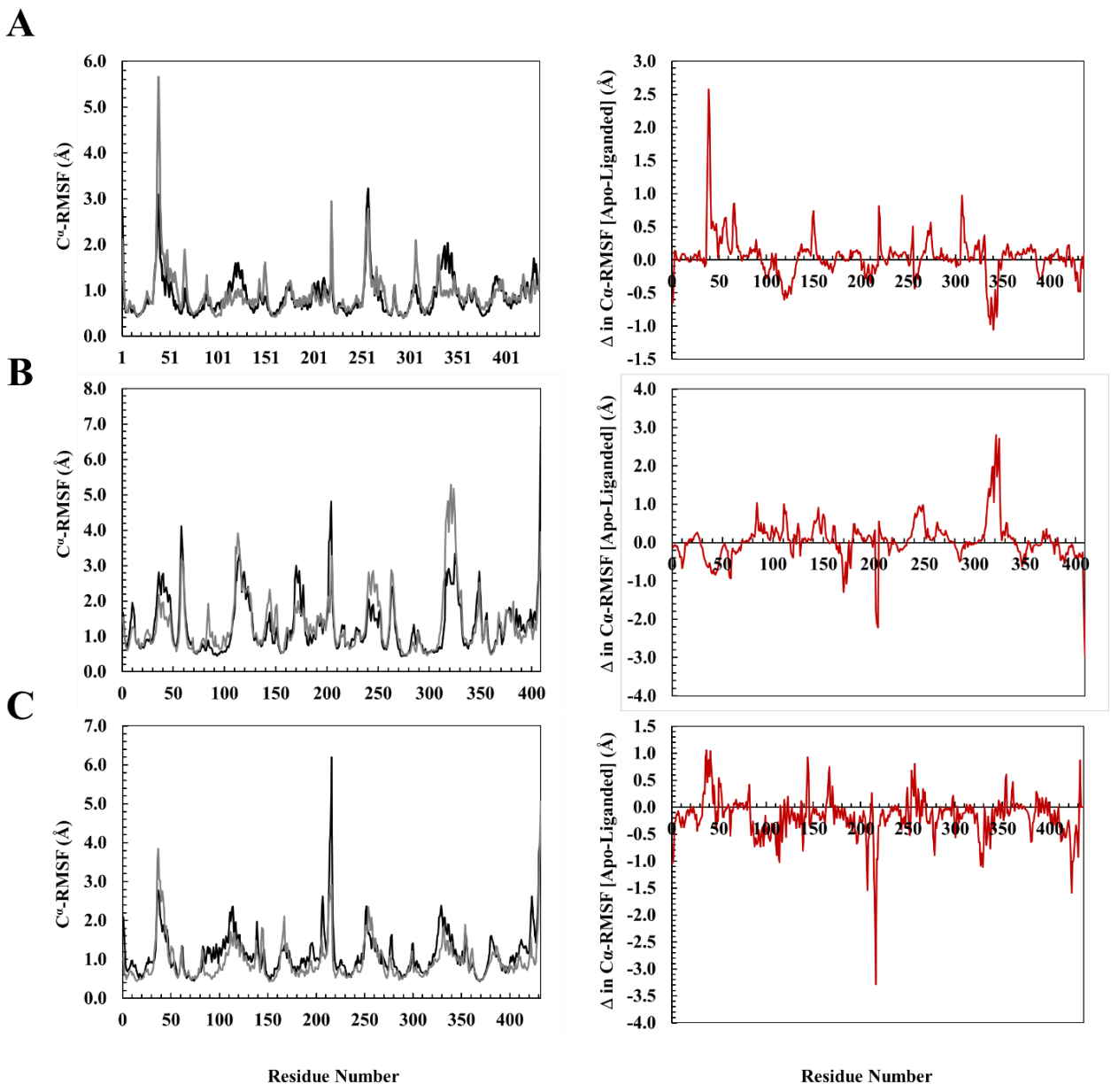


Figure 5.9 Plots of C^{α}-RMSF relative to the respective apo targets.

A: Sh28GST had residue 41 showing the greatest flexibility (right), whilst residue 340 shows the greatest rigidity (left). B: Sbh26GST had residue 330 showing the greatest flexibility, whilst residue 200 shows the greatest rigidity (left) and (C) Human GST-mu had residue 220 showing the greatest flexibility, whilst residue 200 shows the greatest rigidity (left). All the data was extracted from SID™ from the Desmond™ (Schrödinger Release 2023- 1).

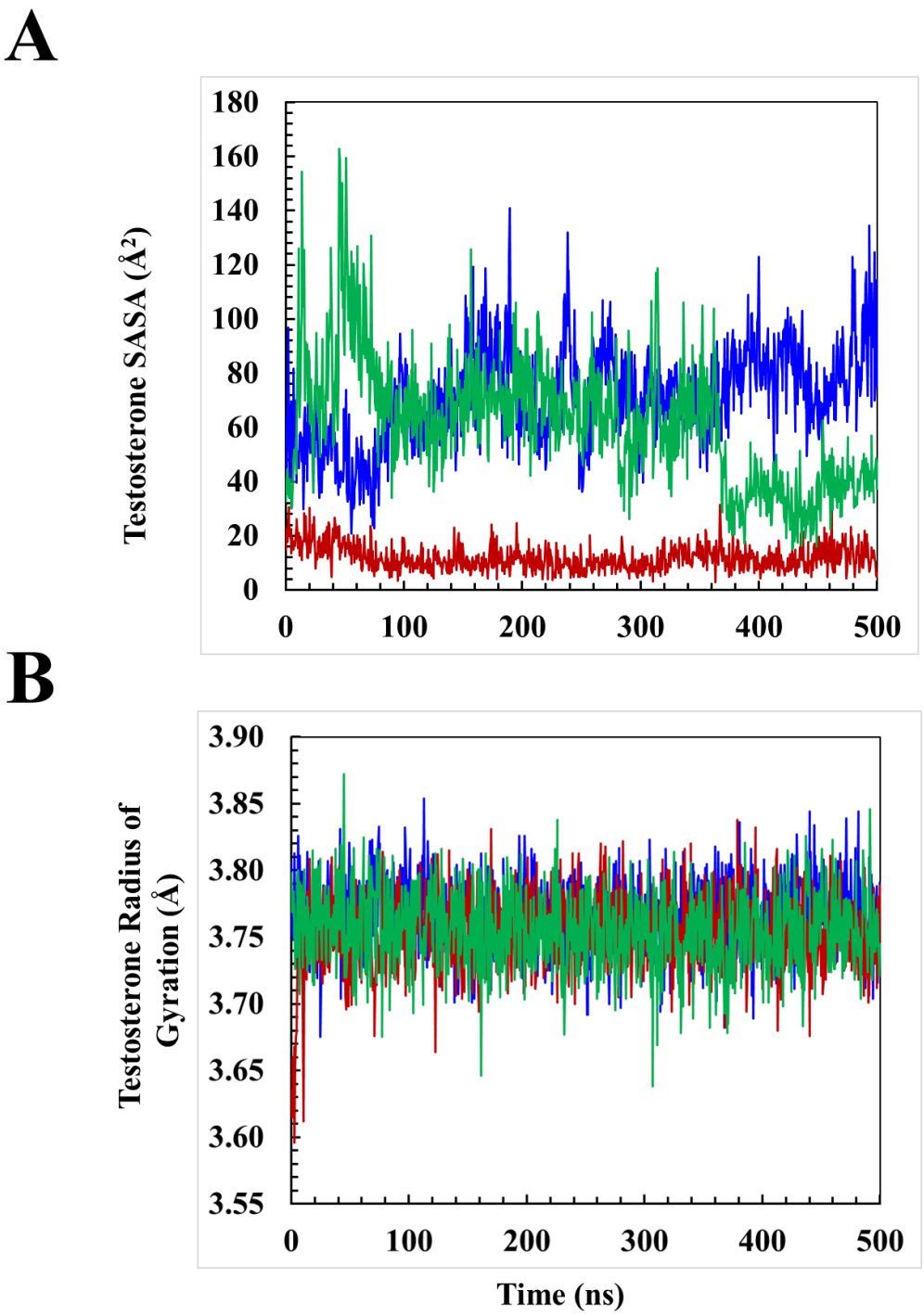


Figure 5.10 SASA of testosterone.

Sh28GST (red), Sbh26GST (green) and Human GST-mu (blue), as well as the radius of gyration of testosterone. The SASA offers inference to exposure of the ligand during binding to the protein. All the data was extracted from SID™ from the Desmond™ (Schrödinger Release 2023-1).

5.3 Discussion

Computational modeling now plays a significant role in the biochemical sciences, especially with regard to protein chemistry and structural biology. This is primarily due to the ever-improving level of prediction because of greater sophistication of the algorithms employed in *in silico* studies. The advances on computing frameworks has made it possible for computationally- intensive calculations such as MDS and meta dynamics to be run on GPU platforms. The deductions made from computational studies can be used to inform experimental realities. In the context of this study, the structure prediction algorithm was implemented in a SWISS MODEL Homology modelling engine, in combination with IFD and Desmond MDS, to gain a better understanding of how testosterone and the GSTs interact dynamically with each other. Thus, this theoretical perspective is meant to affirm and validate some of the empirical data derived in this study.

Given our observation that testosterone interacts with parasite GSTs, and based on the spectroscopic studies of the conjugation reaction by the 2 parasite GSTs, but not Human GST-mu, we hypothesized that testosterone may be forming a tight interaction with Sbh26GST and Sh28GST in comparison with Human GST-mu, and that this strong interaction may compromise the dynamic behavior of both enzymes that may be crucial in their catalytic function. Furthermore, the TSA results, which both suggest tighter binding of testosterone to Sbh26GST and Sh28GST may be validated using computational modeling.

5.3.2 Generation and validation of Sbh26GST model

In the scarcity of a Sbh26GST crystal structure and a full complement of Sbh26GST ORF, we used a reconstructed pseudo-26GST sequence to derive a homology model of Sbh26GST (Padi *et al.*, 2021). It is unlikely that the parasite does not express Sbh26GST because the sequence data suggest that the parasite encodes the enzyme in its genome. The SWISS Model generated a model using Sj26GST as the template. This is acceptable because Sj26GST has a sequence identity of 99.9 % with the pseudo Sbh26GST that was used in this study. Further validation of the theoretical Sbh26GST model using Ramachandran analysis implemented in the ProCheck algorithm indicates that the model is acceptable based on the Ramachandran statistics. Hence, any deduction made from this model can be plausibly accurate.

5.3.3 Simulation of Testosterone docking on GSTs

IFD was employed to simulate the interaction between testosterone and the 3 GST models. This docking stimulation has been shown to outperform most docking algorithms such as the widely used AutoDock Vina (Pagadala *et al.,* 2017; Wang *et al.,* 2016). We postulated that for a ligand to dock into a receptor with (i) lower number of docking poses, and (ii) lower glide emodel energy is suggestive of the strong interaction between the ligand and the receptor, a low number of poses is indicative of restricted degree of freedom for the ligand within the receptor. Thus, the tighter the binding, the lower the number of poses and the more negative the glide emodel energy. We selected the entire dimer search space in order not to cause bias to the system.

For the IFD (Figure 5.4), testosterone bound to Sh28GST and Sbh26GST using the Asn^{73} and Ser^{72} residues to the C=O group and Ser^{71} and Leu^{72} residues to the $C\text{-}O^-$ group respectively. Human GST-mu bound to testosterone with Ala^{70}, Met^{69}, Gln^{67} and Met^{69} residues to the C=O and $C\text{-}O^-$ groups. The AMDS (Figure 5.4) which takes into account solvent structure, shows that the testosterone-Sh28GST complex interacts with 3 H_2O molecules via H-bonds with the C=O and $C\text{-}O^-$ groups at 2 amino acid residues, $Ser^{72}$and Asp^{105}. The testosterone-Sbh26GST does not interact with H_2O molecules at all, and binding occurs at the Ala^{73}, Leu^{72}, Ser^{71} and Glu^{70} with the C=O and $C\text{-}O^-$. For Human GST-mu, testosterone interacts with 2 H_2O molecules and 4 amino acid residues, Ala^{70}, Met^{69}, Ser^{68} and Gln^{67}, using the C=O and $C\text{-}O^-$ groups. Our results affirm that testosterone binds tighter to the parasite GSTs compared to Human GST- mu and that the binding of testosterone to these parasite GSTs appears to be restricted to the dimer interface in a position relatively similar to the interaction between ellagic acid and Sj26GST (Akumadu *et al.,* 2020). This was not the case for Human GST-mu, which showed that testosterone binds at 2 sites, which are the L-site (dimer interface) and the H-site.

In the Sbh26GST system, the conformations of testosterone show the smallest RMSD. Such low RMSD values offer a more dependable measure of distinction, showing increased protein stability, a reduced occurrence of structural reconfigurations and minimal conformational alterations within the binding site residues. (Katari *et al.,* 2016). Meanwhile, both the Sh28GST and Human GST-mu systems display a slightly larger but roughly equivalent deviation (Figure 5.5). Ligand docking studies are not totally reliable in the absence of MDS.

5.3.4 Molecular dynamics of testosterone-GST binding

In the context of this study, MDS is critical because it serves as a validation of IFD. This is because IFD and most other docking studies are done *in vacuo* in an implicit solvent system. Hence, the contribution of water in such interactions is not accounted for. In addition, with an extensive simulation time such as the 500 ns used in this study, it is possible to observe dynamic behavior of the ligand with the receptor site. Average MD structures show that the presence of water molecules changes the amino acid residue profile for H-bonding with testosterone. Thus, the interaction between testosterone and Human GST-mu is not restricted to the dimer interface only.The IFD and MSD studies therefore substantiated the spectrometric activity and TSA studies.

Testosterone-GST contact simulation shows that Sh28GST binds testosterone strongly with percentage contact values greater than 50% for Asp^{105} and 9% for Ser^{72} (Figure 5.7). For Sbh26GST, residuesLeu^{72}, Ser^{71}, Arg^{16} bind testosterone with 61-100% contact. Human GST-mu binds testosterone with residues Arg^{108} and Met^{69} with 63 and 65% contact respectively. The fluctuation of residues during simulation indicates the residue flexibility and can be used to assess RMSF values to provide an indication of the amino acid residues that are highly represented in the molecular motions associated with binding to testosterone. The Cα RMSF shows high fluctuation of amino acid residues 90-120, 320-340 and 270-290 during testosterone binding for all the GSTs (Figure 5.6). A more comprehensive examination of the connections between structural stability was obtained through RMSF analyses. It was observed that the regions associated with ligand binding and catalytic activity exhibited the highest degrees of structural fluctuation.

The GST-testosterone interaction profile study revealed high representation of strong H-bond (H) and water bridge (W) interactions between the GST and testosterone (Figure 5.7). Ionic (I) and hydrophobic (Hy) interactions were poorly represented. For Sh28GST, Gln^{71} (H/W), Ser^{72} (H), Asn^{73} (H), Asn^{101} (W), Met^{104} (Hy/W), Asp^{105} (H/W/I) and His^{107} (W) significantly interacted with testosterone. Sbh26GST only had a H- bond and water bridge interactions at residues Glu^{103} (W), His^{107} (H/W), Arg^{16} (H/W), Ser^{71}(H), Leu^{72}(H), Ala^{73}(H) and His^{107}(H/W). The Human GST- mu-testosterone complex showed significant interactions with Ser^{93} (W), Leu^{100} (Hy), Arg^{103}(H/W/I), Ser^{107} (H/W), Tyr^{111} (H/W), Met^{67} (H/Hy/W) and Arg^{108} (H/I/W). Of the three complexes, Sh28GST had a high representation of water bridge interactions, whereas Sbh26GST had a high representation of H-bond interactions. The high number of interactions for each of these 3

complexes indicates that they are stable complexes. Testosterone is highly hydrophobic with minimal propensity to form H-bonding within the binding pocket. A hydropathy plot of α4-α5 shows that the majority of residues within the 2° structural elements are mostly hydrophilic (or hydrophobic) and hence it is expected that the interaction within the receptor should be driven largely by interaction. Our results comprehensibly affirm this postulate.

*RMSF measurement can simulate atomic fluctuations between GSTs and bound testosterone during simulation. This study shows that the regions identified by protein interaction study have less fluctuation, suggesting minor residue deviation and greater stability. *SASA was used to evaluate testosterone-water interactions during simulation. SASA values can indicate whether an amino acid residue is hidden from or exposed to the solvent and thus suggest steric availability of the atom. Sh28GST exhibited the lowest SASA value when compared to the other two proteins (Figure 5.10A). A low SASA value is more desirable in drug applications since it denotes protein stability.

One key parameter derived from the trajectory of an MD simulation of a ligand-receptor complex is the RMSD (as implemented in Desmond). RMSD measurement aims to determine the stability of the testosterone-GST complex. The radius of gyration can be used to evaluate structural compression change during the simulation time. The sudden peaks in radius of gyration for the 3 testosterone-GST complexes indicate structural transformation during the 500ns observation time, particularly for the Sbh26GST and the Human GST-mu complexes (Figure 5.8). The most stable testosterone-GST complex is the Sh28GST complex due to the moderate conformational changes recorded during simulation, showing intact testosterone-Sh28GST complex for the stimulation period. Our MDS results further affirm our IFD study since the ligand remains bound within the parasite GST receptors while the ligand drifted out of the receptor pocket in Human GST-mu (Figure 5.9). Overall, the C^{α} RMSD of the enzymes predicts that the dynamic behavior of the proteins are altered in the presence of water. This implies that binding of testosterone to parasite GST is allosteric because none of the substrate binding sites are blocked by testosterone. Rather, testosterone binds at the L-site, which appears to perturb the dynamic behavior of the enzyme critical for GSH-CDNB conjugation.

Chapter 6
The apo-crystal structure of a polymorphic form of the 28-kDa *Schistosoma haematobium* glutathione transferase

Abstract

GST is a potential therapeutic drug target against helminthic diseases. This essential enzyme prevents oxidative stress by inhibiting toxic build-up of xenobiotics in the worm during human host infection. The *Schistosoma haematobium* schistosome manifests two GST isoforms, namely the 26- and 28-kDa GST, which exhibit contrasting substrate specificities within the worm. This study was designed to explore the structural and functional relationships of Sh28GST through crystallographic studies. The CDNB experiment revealed that the enzyme was catalytically active. The Sh28GST enzyme has an mostly alpha-helical structure, as confirmed by the secondary structural content. The use of extrinsic ANS substitution revealed that the enzyme has buried hydrophobic clefts. IMAC was used to purify Sh28GST. The crystal structures of the un-ligated enzyme were determined, with an average B-factor of 23.59 Å2 and a resolution of 2.3 Å and an R-factor of 20.65%. The crystals belonged to the orthorhombic space group P_{21}, and there were two biological molecules in the asymmetric unit. The unit-cell parameters were a = 54.133, b = 77.086, c = 54.015 Å. Computational modelling studies suggested that 8ALS may have a different catalytic propensity from 1OE8, possibly due to differences in the molecular dynamic trajectory, particularly at the dimer interface of this enzyme, which is obligately dimeric.

6.1 Introduction

Bilharzia is a disease caused by trematode parasites of the *Schistosoma* genus. It is a tropical and poverty-associated disease that is both harmful and neglected. The disease is responsible for the debilitating health conditions of over 240 million infected people (Raj *et al.,* 2022). The human and freshwater snail intermediate host maintains the schistosome life cycle. *Schistosoma haematobium* infects humans and animals from contaminated water sources (Adenowo *et al.,* 2015). The pupae proliferate within the skin of the human host, thereby entering the circulation via the lymph nodes and capillaries (Torres-Rivera *et al.,* 2008). The schistosome parasite exhibits different life stages exhibited within the host where the parasite acquires host immune system evasion strategies. The three major *Schistosoma* parasites (*Schistosoma mansoni*, *Schistosoma japonicum*, and *S. haematobium*) all consist of two major GSTs, namely the 26 and 28-kDa GSTs (Sbh26GST and Sh28GST). The main effects of *Schistosoma* infection include anemia, malnutrition, and retarded growth among many other debilitating effects (King & Dangerfield- Cha., 2008).

The GST superfamily is divided into three broad categories: membrane- associated proteins, mitochondrial GSTs and cytosolic GSTs, with the latter being the most abundant and pharmacologically significant (Jancova *et al.,* 2010). The "C" terminal of GSTs is predominantly alpha-helical, while the selective "N" terminal exhibits an alternating α-helix and β-strand topology. The active site consists of a hydrophobic co-substrate, the "H-site" and a GSH binding site "G-site" (Hayes *et al.,* 2005). This binding site forms the interface of the GST dimer and is commonly implicated in drug binding (Figure 6.5). Differences amongst GSTs include variation in structure and chemistry of the "H" site between the various classes (Allocati *et al.,* 2018), and such variability makes ligand binding possible. To date, the alpha (σ), mu (μ) and pi (π) kappa, omega, theta and zeta GST classes are well categorised amongst cytosolic GSTs in mammals (Vaish *et al.,* 2020). The Sh28GST under investigation in this study originates from the sigma class. The sequences of various GST classes exhibit a low similarity index despite exhibiting identical folding in virtually all the recorded three-dimensional GST structures (Sheehan *et al.,* 2001).

The primary function of GSTs is to detoxify pathogenic cells of helminths that cause schistosomiasis (Sheehan *et al.,* 2001). Detoxification is achieved by catalysing the reduced glutathione. In addition, GSTs possess ligandin capabilities and bind to numerous electrophiles, posing as plausible schistosomiasis drug and vaccine targets (McManus and Loukas, 2008). It is assumed that there is dormant detoxification which GSTs induce in anti-helminthic drugs (Townsend *et al.,* 2003). Drug design efforts have focused on Praziquantel, whose mode of action is poorly understood (Vale *et al.,* 2017). The GSTs also function in chemotherapy resistance (Oakley, 2011), thus making them potential drug targets for chemotherapy (Sheehan *et al.,* 2001). The ligandin functionality was named after the promiscuous attachment of different ligands that are that can bind and prevent catalytic activity (Kolobe *et al.,* 2004). Future promising drug targets against schistosomiasis should inhibit the catalytic activity of GSTs, thereby detoxifying the electrophilic substrates (Mahajan and Atkins, 2005).

X-ray crystallography is an indispensable tool in highlighting the importance of structural and functional relationships of the pharmacologically important GST superfamily to aid the pharmacology industry with novel drug design. The current study investigates the three-dimensional structure determination of the 28-kDa *Schistosoma haematobium* glutathione transferase (Sh28GST) to understand the protein's functional and structural relationship. An outstanding characteristic of Sh28GST is that the catalytic tyrosine at position 10 occupies two well-defined inner and outer positions (Angelucci *et al.,* 2005). This phenomenon might provide a basis for anti-schistosomiasis rational drug design.

Previous crystallisation and structure determination reports for Sh28GST have always included enzyme inhibitors. This study reports the first case of a true apo-crystal form of the Sh28GST enzyme. The importance of characterising an enzyme's apo state constitutes the basis for designing small molecule soaking and co-crystallisation experiments (Newman *et al.,* 2011). Ligand binding is known to induce numerous structural changes in a protein. This postulate is supported by several studies on GSTs in solution, which has revealed massive conformational changes (Narayanan *et al.,* 2018). These changes may range from subtle loop or side chain shifts to more severe changes, such as protein unfolding and hinge bending, resulting in the ligand's expulsion (Brylinski & Skolnick, 2008). The work of Brylinski & Skolnick, (2008) epitomises analysis of the extent to which protein structural changes from ligand binding occur.

Brylinski and Skolnick (2008) agree with Woolf *et al.* (2000) that the existence of a ligand does

not significantly alter the overall assembly of a protein. They further explain that numerous prediction algorithms overlook big cavities and such oversights result in structurally disfigured molecules (Brylinski & Skolnick, 2008). The innate discrepancies in liganded and un-liganded protein structures are initially determined as a basis. This basis is then contrasted to the configuration changes from the apo to the liganded state to probe induced agility (Clark *et al.*, 2019). Analysis of apo and holo-enzyme forms is crucial in understanding protein binding functionality.

Moreover, liganded proteins help to elucidate the function of enzyme-catalysed reactions with respect to binding site accommodation (Umemoto *et al.*, 2003). They further argue that protein structure is affected by an array of features such as but not limited to hydrogen bonding, electrostatic, disulphide, and hydrophobic forces. Conversely, studies by Gibrat *et al.* (2012) accentuate the stability of liganded proteins over unliganded ones. A study by Chrysina *et al*. (2012) about structural analyses between the liganded and non-liganded protein structures dispels the existence of major conformational changes. Their study rather assumes a more forgiving approach of revealing slight conformational variations between the apo and holo enzymes in the presence of calcium.

Temperature or atomic displacement parameters, referred to as B-factors (Debye-Waller factor), are calculated variables in protein-ligand complex crystallography structures (Johnson *et al.*, 2018). They highlight the reflection of the disarray of atoms from their theoretical equilibrium localities due to the position of disorder and the inherent thermal movement. Such movement is essential for protein-structure functions. These B-factors shed light on the dynamic movement of proteins in the crystal state, thus providing a benchmark for motion in the solute state across various crystal structures. Precise recognition of cavities is crucial in protein structure and function analysis, ligand binding capability, design and stability (Stank *et al.*, 2016).

In this current study, unliganded crystal structures of the 28-kDa *Schistosoma* glutathione transferases were solved at a molecular level to interpret the proteins' interaction, function and structure. Dealing with compounds at the molecular level is essential for revealing the intricacies of drug discovery. This present study aims to enclose the existing knowledge deficit in the structure and function of the 28-kDa allele from *S. haematobium*.

6.2 Materials

Vector synthesis was done by GenScript (NJ, USA). Crystal screening buffers, reagents and other apparatus were purchased from Hampton Research (CA, USA). Purchases from Sigma-Aldrich (MO, USA) included reagents such as Coomassie Blue-G250 dye, antibiotics (chloramphenicol and ampicillin), 2xYT, Glutathione- agarose resin, IPTG, protein ladder, BSP, L-glutathione and CDNB. In contrast, the ANS was procured from the local Sigma-Aldrich (South Africa). The 15-well EasyXtal® plates were purchased from Qiagen.

6.3 Methods

6.3.1 Protein expression and purification

The Pet-11a (+) vector that was used included an ampicillin resistance gene as a selectable marker. Cloning was done by restriction enzymes. The 5′ sequencing primer consisted of T7, whilst the 3′ sequencing primer consisted of a T7-terminal. The vector also possesses an f1 origin of replication for sequencing and mutagenesis purposes. The C-and N- terminal hexahistidine tag facilitated metal chelation purification. Expression and purification of the GST was performed by a method previously described by (Johnson *et al.,* 2003). To assess how pure and how homogenous the GST protein sample was, SDS-PAGE was conducted. The concentration of Sh28GST was estimated using the Beer-Lambert law through spectroscopic analysis:

$$A_{280} = \varepsilon . c . l \qquad \textbf{Equation 6.1}$$

The concentration of Sh28GST was determined using the Beer-Lambert law, which involves measuring the absorption value at 280 nm (A_{280}), determining the molar extinction coefficient (ε) at 280 nm using the Expasy ProtParam tool (Gasteiger *et al.,* 2005), and calculating the concentration value (c) with the known cell pathlength (l) in centimeters.

6.3.2 Protein characterization

Specific activity assay for Sh28GST

The specific activity of Sh28GST was determined by the CDNB glutathione conjugation activity assay. The absorbance change at 340 nm was measured to analyze the association between CDNB and GSH. The assay involved using a reaction mixture consisting of 100 mM NaH_2PO_4 with 0.02% (w/v) NaN_3, 1 mM EDTA, 2 mM DTT, and 1 mM GSH, pH 6.5, with concentrations of 10-50 nM of Sh28GST in three replicates. CDNB was added at a concentration of 1 mM, and the product's appearance was measured using the Jasco V-550 UV/Vis spectrophotometer. The reaction rate was determined by calculating the slope from the linear regression curve. The product formed was 1-(s-glutathionyl)-2,4-dinitrobenzene.

Circular dichroism (CD) spectroscopy

The Sh28GSTs secondary structure content was analyzed using CD Spectro polarimetry with a Jasco J-1500 spectropolarimeter and spectra manager software v1.5.00 (Jasco Inc., Tokyo, Japan). The protein samples, with a concentration of 2μM, were prepared in 20 Mm NaH_2PO_4, pH 7.5, and 0.02% (w/v) NaN_3. The spectra measurements were taken at 20ºC, with a bandwidth of 5 nm, a scan speed of 200 nm/min, a data pitch of 1 nm, and a 1-second response time, with ten accumulations between 185 and 245 nm. The data was adjusted from millidegrees to mean residue ellipticity [θ] ($deg.cm^2.dmol^{-1}$) and buffer-corrected.

Intrinsic fluorescence

The 3-dimensional structure of Sh28GST was analysed through intrinsic tryptophan fluorescence by either including or excluding GSH. A Jasco FP-6300 spectrofluorometer with spectra manager software v1.5.00 (Jasco Inc., Tokyo, Japan) was used to measure the spectra at 20ºC in triplicate. The samples either included or excluded GSH with a 10μM Sh28GST concentration. Further, incubation in NaH_2PO_4, pH 7.5, 1 mM DTT, 1 mM EDTA and 0.02% (w/v) NaN_3 occurred over a 30-minute time frame by including or excluding GSH at 100 μM.

Extrinsic ANS fluorescence

Hydrophobic clefts were probed using extrinsic ANS fluorescence. The signal was monitored from 395 to 650 nm with fifteen accumulations at 1 cm path length, excitation bandwidth of 5 nm and emission of 5 nm, data pitch of 1 nm, scanning speed of 200 nm/min, and excitation at 395 nm. Spectra from the replicates were averaged and buffer corrected.

Protein crystallisation

The purified Sh28GST was dialysed against 25 mM NaH_2PO_4 buffer. Crystals of Sh28GST (apo form) were produced in hanging drops and set up in 15-well Xtal Easy crystallisation plates. The mother liquor solution consisted of 2.1M $(NH_4)_2SO_4$, 0.1M Tris ($C_4H_{11}NO_3$) (pH 7.5), and 0.005 MB-mercaptoethanol. A hanging drop (5 μL) was created for each experiment by combining the stock protein solution [231 μM- 5.3 mg/mL] in NaH_2PO_4 with the reservoir solution in a 1:2 ratio. The protein crystals were harvested by using the CrystalCap™ SPINE HT goniometer base. The crystals were then immersed thrice in 100% (v/v) Parabar 10312 (Paratone) and mounted on a cryo-loop. Large diffractive cubic crystals grew within two days.

2.5 X-ray crystallography

The Sh28GSTs X-ray crystallographic data was obtained using a Bruker D8 Venture Bio PHOTON III area detector diffractometer. To cool the Sh28GST crystals, they were exposed to a cold nitrogen stream and chilled at -173 ºC with the reservoir solution serving as a cryoprotectant. PROTEUM[4] software was utilized to process the diffraction data, and structural refinement was carried out through the Molecular replacement method, employing the PHASER program in PHENIX software (McCoy *et al.,* 2007). Glutathione saturated Sh28GST, PDB ID: 1OE8 served as the starting model for the refinement. Finally, the model was built utilizing the WinCoot Program (Emsley *et al.,* 2010).

B-factors analysis

Normalised B-factors which are expressions of the B-factor (with respect to the standard deviation around the mean value) were calculated using the following equation:

$$B' = \frac{(B-\mu B)}{\theta B}$$ **Equation 6.2**

The normalised B-factors were computed with UCSF Chimera. The protein files were inputted as PDB files.

6.3.3 Molecular modelling studies

Computer hardware

Two high-performance gaming computing units were utilized to conduct molecular modeling studies. The first unit, which ran the Maestro algorithm on a Windows OS PC, consisted of an AMD RYZEN Threadripper 1950X Processor, an Asus Rog Strix X399-E Gaming Ryzen AMD motherboard with 40MB of Cache Memory, a 4TB drive 6GB/s 3D NAND internal SSD with 560 MB/s read and 520 MB/s write speeds, and a MSI GeForce RTX 2080 Ti 11 GB GDDR6 4352 CUDA core graphics card. Additionally, it had a 4.0 GHz Precision Boost (up to 4.2 GHz) X399 chipset and Quad-Channel DDR4, as well as 64GB (16GB x 4) 3200 MHz DDR4 desktop gaming RAM.

The second unit, which ran the Desmond molecular dynamics simulation algorithm on an Ubuntu OS PC, was equipped with an AMD Threadripper 3990X 4.3 GHz GT 1030 2 GB PRO high-performance workstation with 288 MB Cache, 64x Cores, and 4.3GHz Turbo. The unit also had an MSI TRX40 PRO 10G AMD Ryzen Threadripper motherboard, a GeForce RTX 2070 8GB GDDR6 graphics card, a 3200MHz 64GB gaming RAM, a 1TB M.2 SSD with speeds of up to 3.5 GB/s, and a 4TB HDD.

Preparation of molecular systems

The crystal structure coordinates of *S. haematobium* GST bound to GSH (1OE8) were obtained from RCSB-PDB, and an in-house variant of *S. haematobium* GST (RCSB- PDB ID: 8ALS) was subjected to energy minimization using OPLS_2005, which was implemented in the protein preparation wizard algorithm of Meastro v13.0. Two molecular systems, namely Apo-8ALS and Apo-1OE8, were prepared for molecular dynamic simulation calculation by creating a solvated system containing counter charges in a 0.15 M NaCl system using the SystemBuilder tool of Maestro v13.0. To build each molecular system, the following parameters were used with the OPLS_2005 force field: (i) TIP3P solvent model, which assigned charges and Lennard-Jones parameters to each three atoms of a rigid water molecule, (ii) an orthorhombic periodic boundary box with box conditions specified as distances from the box boundary (Å): a =10 × b = 10 × c = 10 and angles: $\alpha = 90° \times \beta = 90° \times \gamma = 90°$, and the volume was minimized, (iii) counter ions (Na^+ or Cl^-) were added based on the overall charge of each molecular system, and (iv) finally, the systems were conditioned to have 0.15 M NaCl. Each system was saved as a .mae file and submitted to the Desmond molecular dynamics simulation engine for molecular dynamic simulation.

Molecular dynamic simulation

To accelerate the simulation of the molecular systems with different time scales, the RESPA integrator was utilized, with the following separation of time scales: bonded forces at 2 fs, near-range nonbonded forces at 2 fs, and far-range non-bonded forces at 6 fs. The Nosé-Hoover chain thermostat method was used to keep the temperature at 26.85 °C with a relaxation time of 1 ps. The Martyna-Tobias-Klein barostat method was employed to maintain the pressure at ~ 1 atm with a relaxation time of 2 ps and isotropic coupling style. Short-range Coulombic interactions were dealt with through the cut off method, where this cutoff radius was 9 Å. The molecular dynamics simulation had a production phase of 500 ns and was run using the Desmond algorithm's default relaxation protocol. The MDS phase was composed of eight distinct stages, with the first seven being equilibration stages that involved short simulation steps. The solute heavy atoms were restrained in a 100 ps Brownian Dynamics simulation under NVT conditions at 10 K during stage 2. The final stage was the long-range 500 ns simulation stage

Post-dynamic analysis

After conducting the MD simulation, we performed post-dynamic analyses on the trajectories obtained from the simulation using Maestro v13.0. The aim was to compare the conformational dynamics of the two molecular systems. First, we evaluated the quality of the simulation by analyzing the average energy, pressure, temperature, and volume of each simulated system using the Simulation Quality Analysis feature available in Maestro v13.0. Secondly, we calculated the root-mean-square deviation (RMSD) of the Cα atoms, and the RMSF of the residues.

6.4 Results

6.4.1 Expression and purification

The sequences were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Recombinant Sh28GST was overexpressed at 30°C for 6 hours (0.5 mM IPTG) and yielded a wet cell pellet of around 4.5 g/L of culture. Analysis of Figure 6.1 from SDS-PAGE shows an excellent yield of soluble protein. Immobilized Ni^{2+}-affinity chromatography was used to purify the protein. The protein concentration was determined to be approximately ~5.3 mg/mL using spectrophotometry. The molecular weight of the Sh28GST monomer was approximately 25 kDa, which is consistent with theoretical expectations.

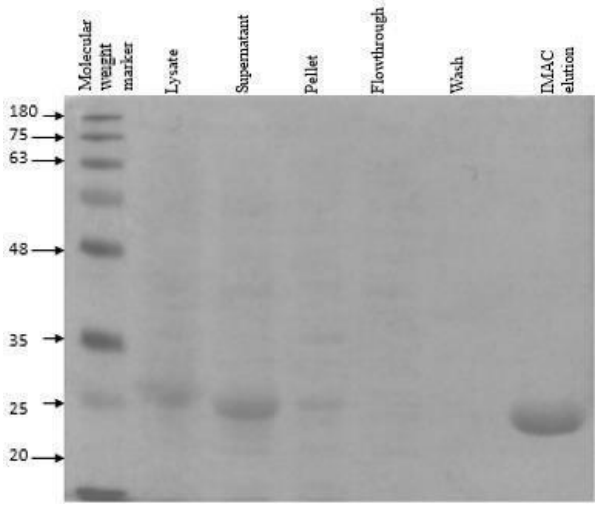


Figure 6.1 Non-liganded Sh28GST perceivable with Coomassie brilliant blue staining.

The purified Sh28GST was analyzed using SDS–PAGE. Initially, a crude lysate of *E. coli*, which harbored the overexpressed Sh28GST, underwent centrifugation to separate the soluble fraction (supernatant) from the insoluble fraction (pellet). The solubilized fraction was then subjected to purification through Ni^{2+} IMAC and subsequently eluted using 350 mM imidazole. Subsequently, all the obtained fractions were loaded onto a 12.5% (w/v) polyacrylamide gel containing 10%(w/v) SDS, followed by visualization through Coomassie brilliant blue staining.

6.4.2 Sh28GST enzyme activity assay

After successfully expressing and purifying the recombinant Sh28GST, its specific activity was evaluated through spectrophotometric analysis. The enzyme's specific activity was determined by using an extinction coefficient of 9600 $M^{-1}.cm^{-1}$ for 1-(S-glutathionyl)-2,4-dinitrobenzene in the assay. The reaction rate was measured in the absence of the substrate GSH, but the results were insignificant. The Jasco V-550 UV/Vis spectrophotometer was employed to measure the product's linear increase spectroscopically over a minute. The specific activity of Sh28GST was subsequently reported to be 14 μmol $min^{-1}mg^{-1}$ with an R^2 value of 0,9755. Sh28GST exhibited high transferase activity with CDNB. The error bar shows the standard deviation. The experimental data was buffer corrected.

6.4.3 Far-UV CD

The ability of Sh28GSTs peptide bonds to polarize light in the far-UV region indicated the proteins' function as a chromophore. To evaluate the content of Sh28GSTs secondary structure in the presence or absence of GSH, far-UV CD was utilized. Spectral measurements were taken at 20°C using 2 μM of Sh28GST in 20mM NaH_2PO_4 (pH 7.5) and 0.02% (w/v) NaN_3. The measurement parameters included a 5 nm bandwidth, 0.5 nm data pitch, a 200 nm/min scanning speed, and fifteen accumulations between 180 and 250 nm. The data was normalized and buffer corrected to mean residue ellipticity, denoted by [θ] ($deg.cm^2.dmol^{-1}$), using a specific equation.

$$[\theta] = \frac{100 \times \theta}{cnl}$$

Equation 6.3

The CD signal (in millidegrees) denoted by θ, was calculated by multiplying the protein concentration in mM (c) with the number of residues (n) and the path length (l) in cm. The resulting CD spectra exhibited a positive ellipticity peak at 195 ± 0.5 nm and two negative ellipticity troughs at 222 ± 0.5 nm and 208 ± 0.5 nm (Figure 6B), indicating a protein profile that is mainly composed of alpha-helical structures. The CD spectra obtained in the presence of GSH did not differ from those obtained without GSH. The secondary structure content was determined by deconvoluting the CD spectra data using the DichroWeb analysis program (Whitmore and Wallace, 2004).

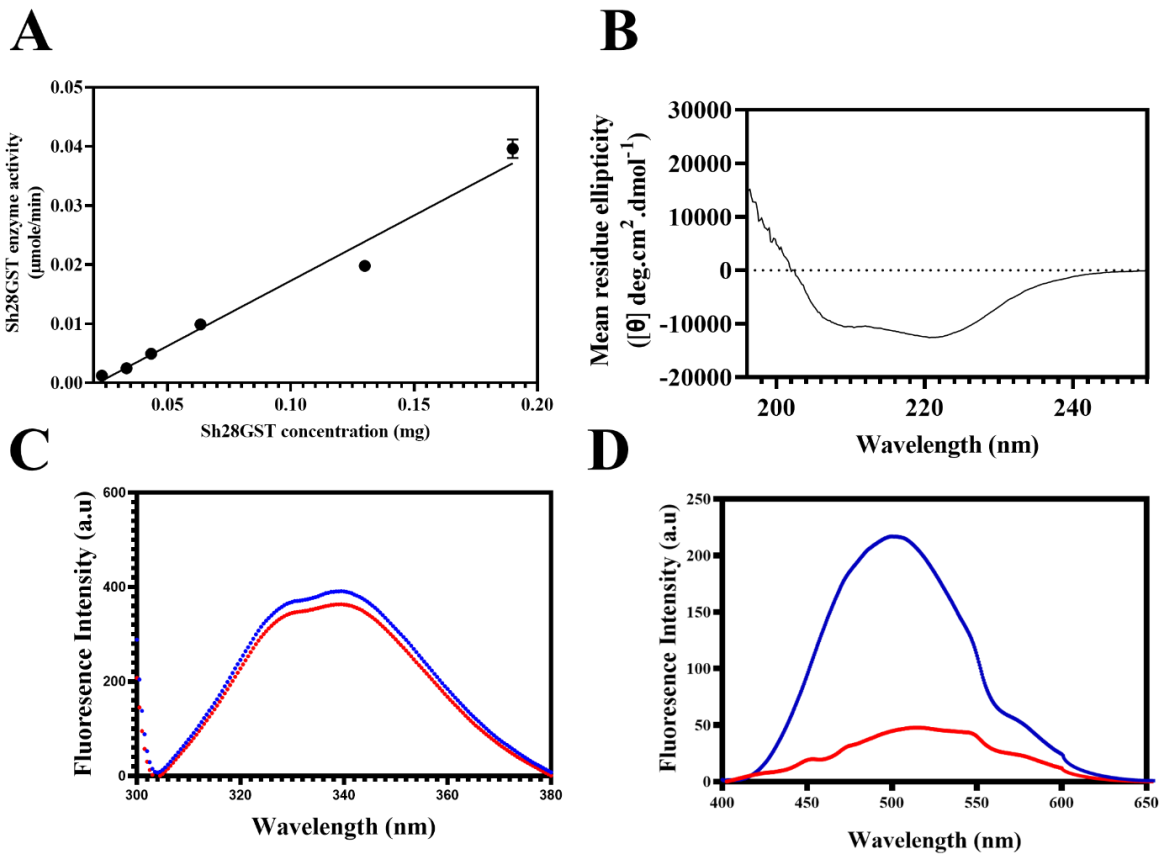


Figure 6.2 Structural characterization studies for Sh28GST.

(A) Excellent enzymatic capability was demonstrated by the specific activity profile of Sh28GST at varying concentrations. B) Sh28GST exhibited a protein structure that consisted mainly of alpha helices. C) Sh28GST showed intrinsic tryptophan fluorescence when bound with 1 mM GSH (blue) and in the presence of BSP (red). D) ANS fluorescence of Sh28GST was observed both in unbound form (red line) and when bound to 2 µM Sh28GST (blue line).

6.4.4 Fluorescence spectroscopy

In this study, intrinsic fluorescence spectroscopy was employed to assess the three-dimensional structure of Sh28GST. This analysis was conducted both in the presence and absence of GSH and BSP, with a focus on the local environment surrounding the tryptophan residues within the enzyme. The intrinsic fluorescence spectra of Sh28GST with and without GSH exhibited only a minor alteration, manifesting as a negligible 1nm shift towards longer wavelengths (depicted in Figure 6.2C). When ANS bound to concealed hydrophobic clefts on proteins, it led to a shift towards shorter wavelengths (hypsochromic shift) and an augmented fluorescent intensity compared to free ANS in solution. This outcome suggests that ANS bound to hydrophobic regions on the protein's surface. However, in the presence of BSP, a notable reduction in the quantum yield was observed. This suggests that BSP displaced ANS from the protein, as evidenced by the shift towards shorter (blue) wavelengths of the emission peaks, transitioning from 515 nm to 502 nm (as illustrated in Fig. 6.2D).

6.4.5 X-ray crystallography

Using the hanging drop vapor diffusion method, crystals of Sh28GST were obtained (Figure 6.3A). The atomic structure of Sh28GST was determined through the molecular replacement method, employing the 1OE8 coordinate as the replacement model. The PHASER program, which incorporates the PHENIX algorithm, was utilized for molecular replacement (McCoy *et al.,* 2007). The presence of the sodium atom (Figure 6.4A) was expected since the crystallization buffer contained sodium phosphate in the precipitation buffer. The final R_{factor} of the structure refinement was 18.7% at 1.53 Å resolution, with a solvent content of 48.1% (Table 1). The Matthews coefficient (V_m) demonstrated a direct correlation with the amount of the unit cell that was occupied by the solvent. The Sh28GST (PDB ID 8ALS) exhibited eight hydrogen bonds at the crystal packing site, compared to six hydrogen bonds in PDB ID 1OE8 (Fig 6.8). The displacement of the ligand molecule is significant in terms of crystal contact (Luo *et al.,* 2015).

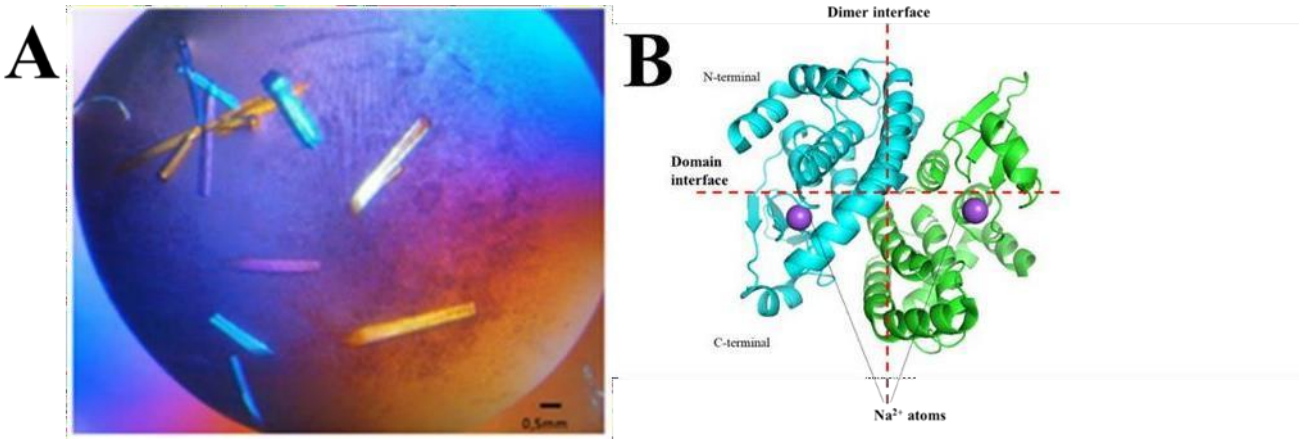


Figure 6.3 Sh28GST crystals

(A) Using the hanging-drop method, crystals of Sh28GST without ligand were grown in 15-well tissue culture plates containing a well solution of 2.1 M $(NH_4)_2SO_4$, 0.1 M Tris, and 0.005 M B-mercaptoethanol. The CrystalCap™ SPINE HT goniometer base was employed to harvest the protein crystals, which were subsequently immersed in 100% (v/v) Parabar 10312 (Paratone) two to three times before being mounted on a cryo-loop.
(B) The ribbon diagram of non-liganded Sh28GST illustrates two subunits (green) and a red sodium atom. The PyMol molecular graphics program was used to create the homology model based on the template of Sh28GST PDB ID 1OE8.

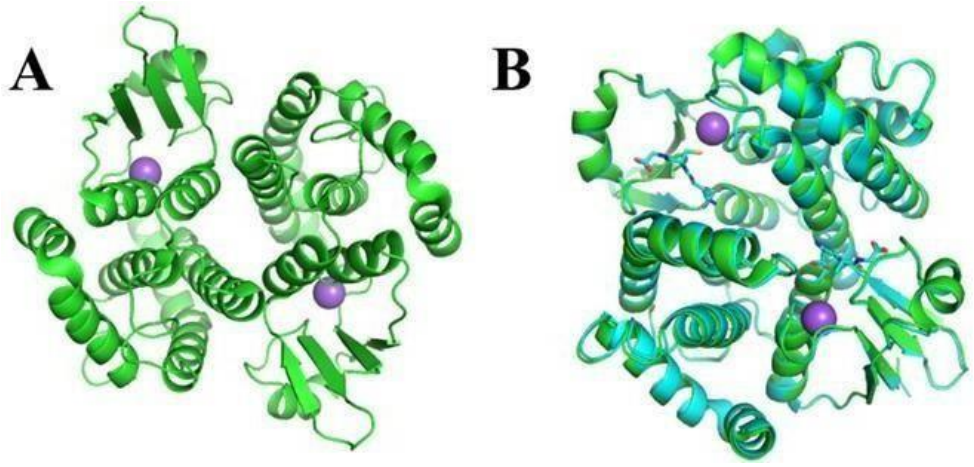


Figure 6.4 Overall Sh28GST structure.

(A) Stereo view of Sh28GST (PDB ID 8ALS) superimposed with PDB ID: 1OE8), a sodium atom is shown in purple. (B) Stereo view of Sh28GST (PDB ID 8ALS) superimposed with PDB ID 1OE8). Sodium atom is shown in purple. GSH is rainbow colored. The images were generated using PyMol. Taken from Makumbe *et al*. (unpublished data).

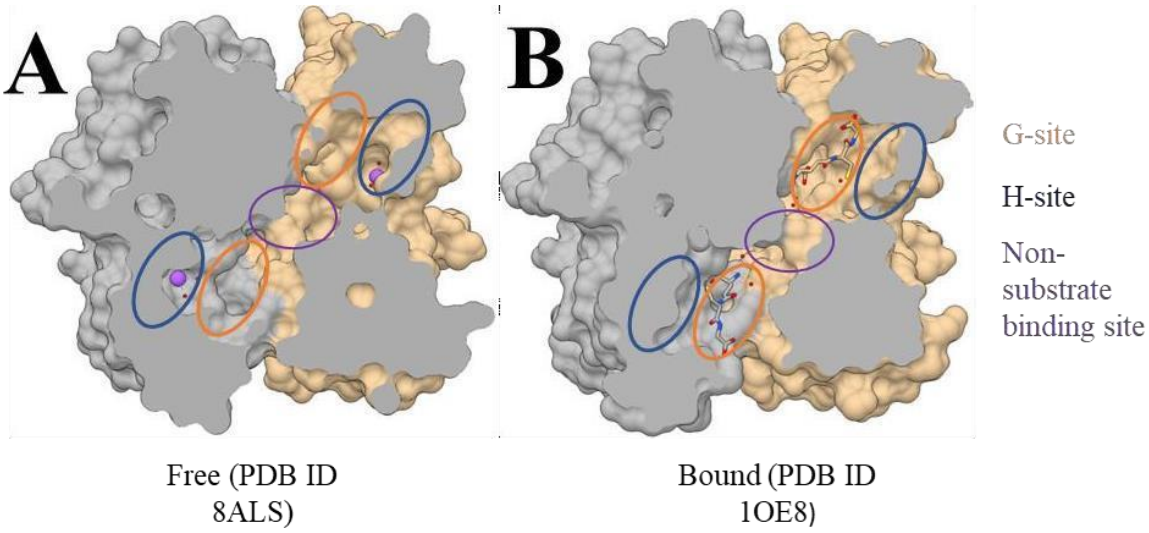


Figure 6.5 Cross section of two structures of the 28-kDa GST dimer.

The two figures were obtained from *S. haematobium,* A: Free structure and B: Bound structure, showing the various binding sites. The boundstructure (1OE8) is shown with glutathione in both G-sites. Figures generated with Chimera1.15.

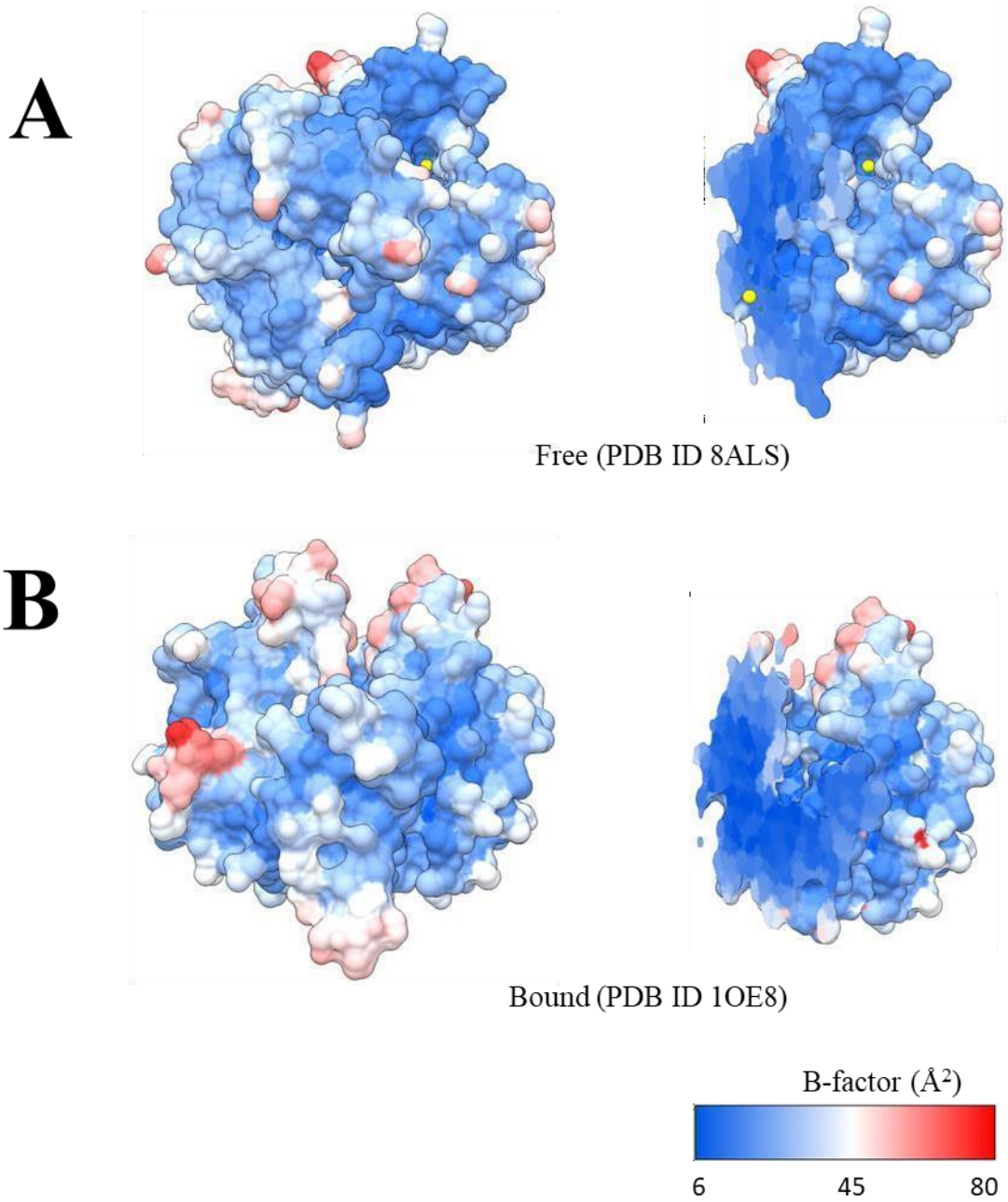


Figure 6.6 Molecular surface of the free and bound Sh28GST.

The bound and the unbound structures were colored by the B-factor. Cross-section view is depicted across the dimer interface plane.

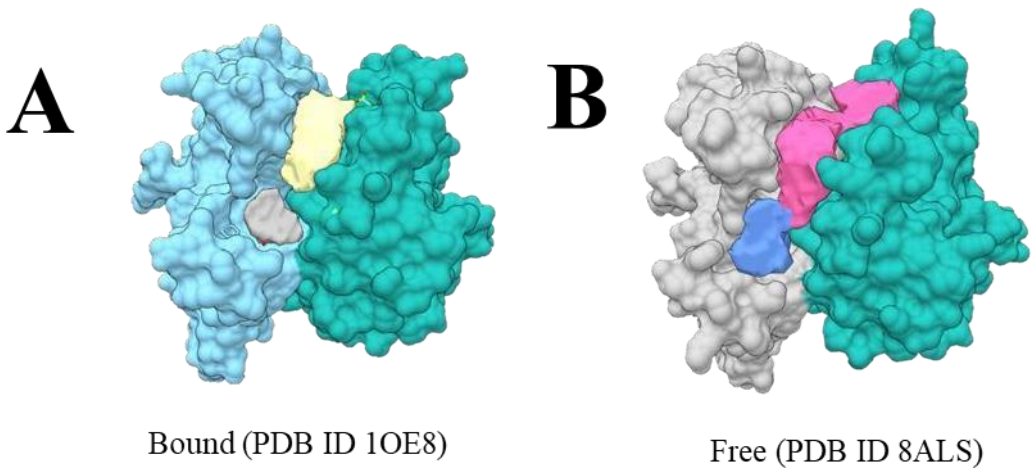


Figure 6.7 Cavities in the A bound and B free Sh28GST.

Residues that are in closest proximity to each cavity are shown. **A:** 1OE8: Yellow cavity: A: Arg^{16}, Arg^{52}, Ser^{71}; B: Asp^{104}. Grey cavity: A: Asp^{10}; B: Arg^{16}, Tyr^{10}, Lys^{45}, Trp^{41}, Glu^{70}, Arg^{52}. **B:** 8ALS: Blue cavity: A: Arg^{52}, Pro^{42}, Pro^{116}. Pink A: Glu^{119}, Lys^{122}, Glu^{126}; B: Lys^{115}, Met^{114}, Lys^{111}.

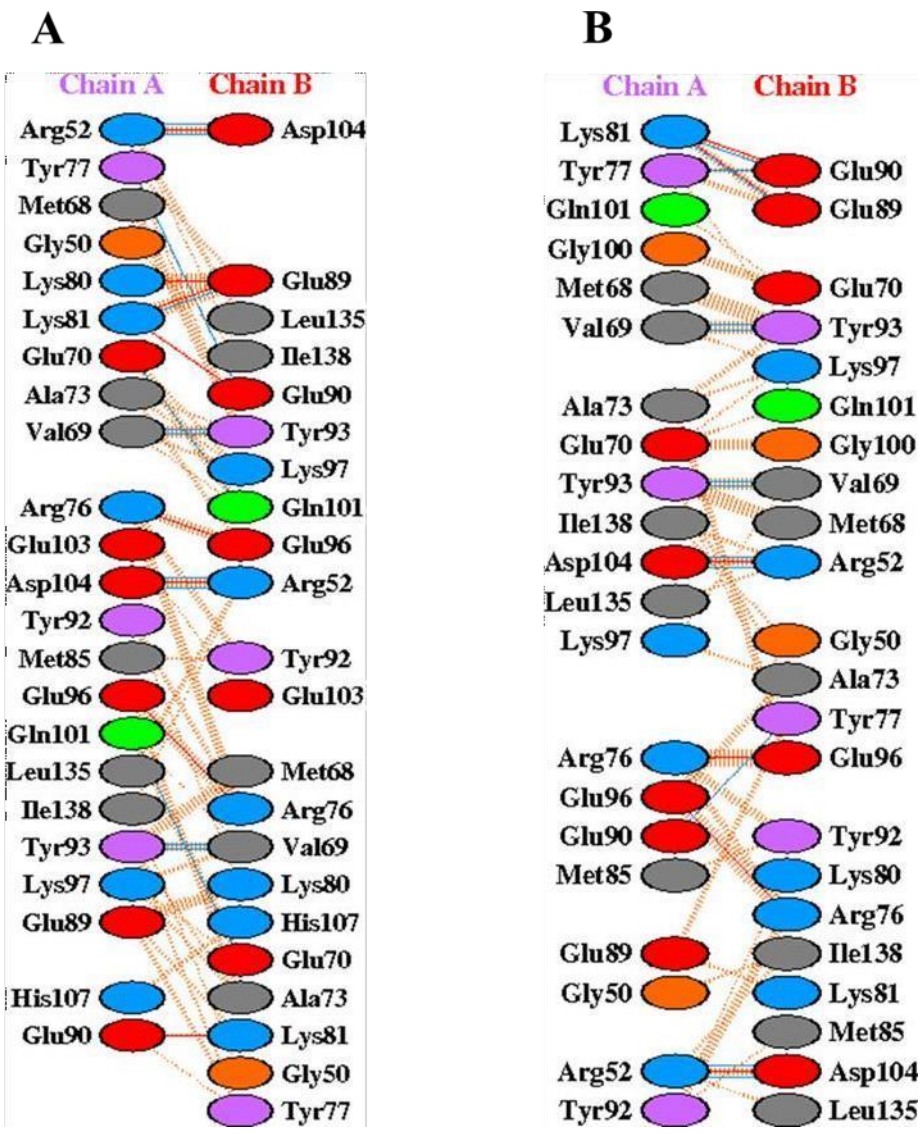


Figure 6.8 The crystal contacts made at the crystal-packing interface.

(A) Sh28GST (PDB ID:8ALS) formed between chain-A and chain-B. Residues directly involved in the crystal contacts at this site were Lys^{80}, Glu^{89}, Arg^{76}, Glu^{96}, Asp^{104}, Arg^{52}, Met^{68}, Lys^{81}, Glu^{90} were directly involved in crystal-packing contacts. (B) Sh28GST (PDB ID: 1OE8). Residues at this crystal contact site included: Lys^{80}, Glu^{89}, Arg^{76}, Glu^{96}, Asp^{104}, Arg^{52}, Met^{68}, Lys81, Glu^{90}. All crystal contacts were generated by using the protein-protein interaction tool in PDB sum (Laskowski, 2001).

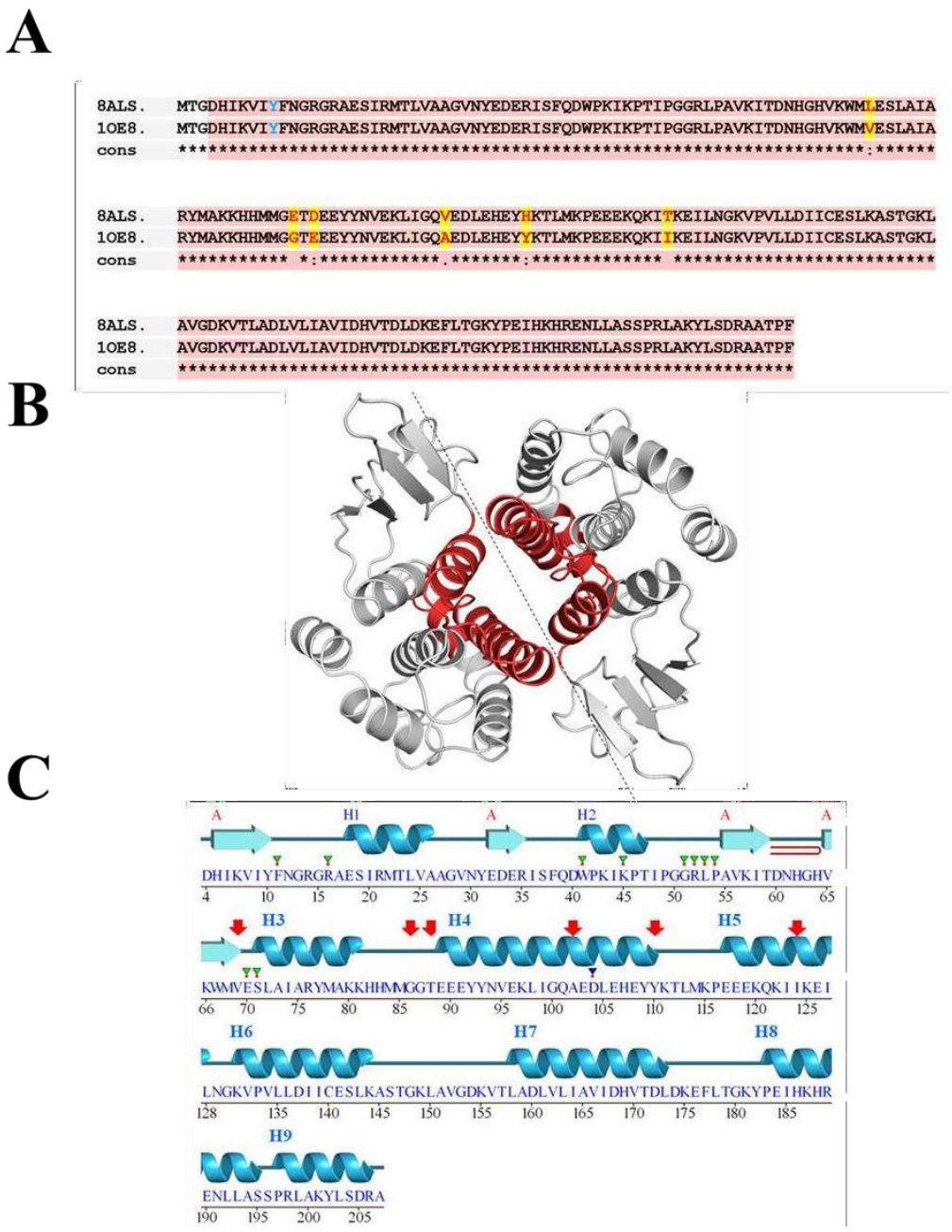


Figure 6.9 Sequence analysis of 8ALS and 1OE8 showing the positions of the variant amino acids.

(A) is a structural alignment of 8ALS and 1OE8 using Expresso^{TM}, which indicates the structural similarity between both protein models. (*B*) ribbon representation of 8ALS showing helix 5 and helix 6 (in red), which largely contributes the dimer interface topography. (*C*) is the primary sequence analysis of 8ALS showing the position of the variant amino acids with respect to the secondary structural element

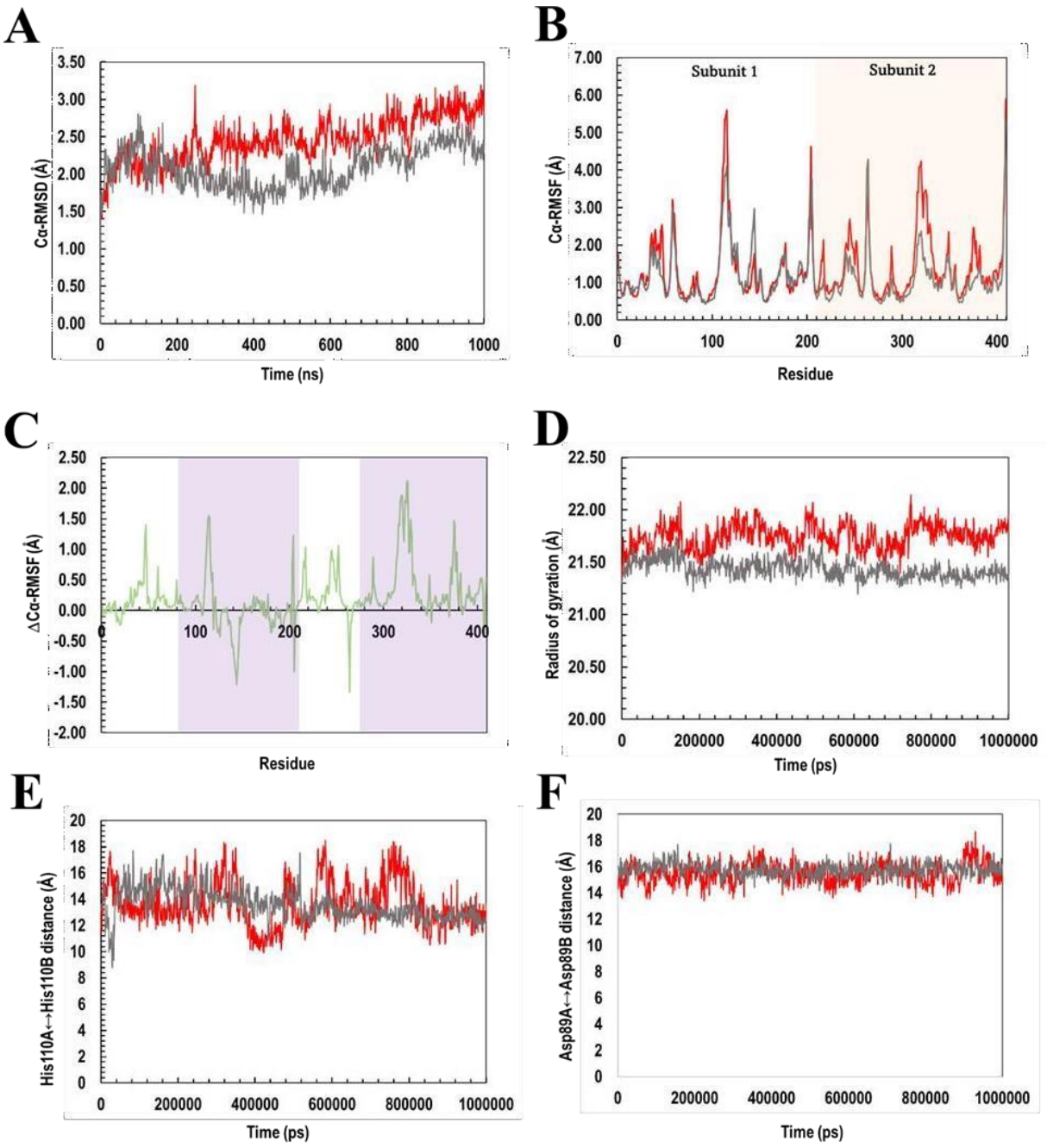


Figure 6.10 Post dynamic analysis of the molecular dynamic simulation of 8ALS and 1OE8.

8ALS (red profile) and 1OE8 (grey profile). Trajectories of the C^{α} showing the RMSD (***A***), RMSF (***B***), and changes in RMSF of 8ALS with reference to 1OE8 (***C***), Rg (***D***) and interatomic distances (***E*** and ***F***) the data points were collected and graphed based on their relationship to the simulation time or residue numbers.

Table 6.1 Statistics related to the collection and refinement of data.

Data collection	
Light source	Cu Kα
Wavelength (Å)	1.5418
	Bruker Venture Bio PHOTON
Detector	III area detector
Temperature (°C)	-173
Space group	P121 1
Unit-cell parameters	
a, b, c (Å)	54.133, 77.086, 54.015
α, β, γ (°)	90.00, 93.15, 90.00
Resolution range (Å)	24.68 -2.30 (2.38-2.30) *
No. of observed reflections	158575 (10629) *
No. of unique reflections	19612 (1752) *
Completeness (%)	98.8 (89.9) *
I/σ(I)	11.4 (3.4) *
R_{merge}	0.132 (0.600) *
Multiplicity	8.1 (6.1) *
Structure refinement	
Reflection used	911 (4.65%)
Resolution range	24.68 – 2.3
Final overall R_{factor} (%)	20.95
R_{work} (%) / R_{free} (%)	20.65 / 27.20
No. of protein atoms	3801
Average B factor value (Å^2)	25.0
RMSD in bond length (Å)	0.002
RMSD in bond angles (°)	0.447
Ramachandran statistics	
Favoured; Allowed; Outliers (%)	97.28, 2.22, 0.42
Matthew's coefficient V_M (Å^3 Da^{-1})	2.37
Solvent content (%)	48.1

a $R_{merge} = \Sigma\ hkl\Sigma\ i|I\ i(hkl)\ F\ \langle(hlk)\rangle\ |/\Sigma\ hkl\Sigma\ i|I\ i(hkl)F$,
In the formula, I(hkl) represents the intensity of reflection hkl, Σ hkl represents the sum over all reflections, and Σ i represents the sum over i measurements of reflection hkl. The value of b R_{free} is obtained by calculating a random selection of 5% of reflections that were not used for structure refinement, while R_{work} is calculated for the remaining reflections. The values in brackets refer to the outermost shell

6.4.6 B-factors analysis

Mobility of the protein structure is captured by the B-factor (Figure 6.6), where 8ALS indicated a noteworthy distinction with a statistically significant level of confidence (p-value < 0.01)., with an all-atom average of 23.59 $Å^2$ compared to 33.46 $Å^2$ of 1OE8. The surface residues of 1OE8 are identified as the primary source of this variation. Furthermore, we observe few residues of high B-factor in both structures. This includes corresponding Ala residues in each chain of 1OE8 and an Asp, Asn and His in 8ALS.

6.4.7 Cavity analysis

Cavity analysis (Table 2) was performed on the 1OE8 and 8ALS structures. Both structures exhibited two corresponding cavities each, as identified by spherical probes bearing radii bounds of 3 – 10 Å. The smaller cavity (cavity 1) is at the dimer interface, with cavity 2 existing within the adjoining glutathione and non-substrate binding sites (Figure 6.7). Overall, the cavities on the free structure were of more extensive volume and surface area.

Table 6.2 Cavity analysis of the 28-kDa GST structures.

Cavity 1 exists within the dimer interface core of the enzyme, and cavity 2 occupies the adjoining G/H binding region.

Cavity	Parameter	8ALS	1OE8
1	Volume ($Å^3$)	507	321
	Surface area ($Å^2$)	393	254
	Effective radius (Å)	3.87	3.79
2	Volume ($Å^3$)	1690	708
	Surface area ($Å^2$)	1073	495
	Effective radius (Å)	4.72	4.29

6.4.8 Molecular dynamic simulation

Although 8ALS and 1OE8 share strong structural homology (Figure 6.9), they vary in amino acid sequence at six positions. The six mutations (using 1OE8 as the wild type) are V69L, G87E, E89D, A102V, Y110H and I124T, which indicates that 8ALS have extra charges compared to 1OE8 due to the presence of extra Glu^{87} and His^{110}. Only one (V69L) of the variations occurred at the N-terminus thioredoxin domain containing the βαβαββα secondary structure architecture, the remaining five mutations are in the C-terminus all-alpha helical domain. None of the six mutations is within 4 Å of the catalytic Tyr10. Following energy minimisation and systems preparations, the waters and the GSH ligand (in 1OE8) were stripped from the overall structures. Both systems were fully solvated using the TIP3P water model, and the net charge after solvation was zero because the charges were neutralised with counter ions (equivalent Na^+ or Cl^-). In addition, both systems were fully equilibrated for the MD simulation production. Hence, inferences can be made from both systems without bias.

The trajectory of the C^{α}, based on the RMSD (Figure 6.10A), shows that the protein models exhibit different dynamics within the 500 ns simulation time. The 8ALS model appears to be slightly more dynamic compared to 1OE8, this indicates that the average root-mean-square deviation of the Cα atoms was 2.95 Å, with a standard deviation included. (±0.32) and 2.09 Å (±0.17) for 8ALS and 1OE8, respectively. The C^{α}-RMSF trajectory substantiates this observation (Figure 6.10B) derived within the 500 ns simulation time. Our results suggest that the backbone of 8ALS fluctuated marginally more than the backbone of 1OE8 with an average (±SD) of 1.65 Å (±0.86) and 1.14 Å (±0.39) for 8ALS and 1OE8, respectively. Furthermore, the Cα-RMSF trajectory revealed that backbone fluctuation is not precisely identical across the two identical subunits. This variation in the backbone fluctuation is more pronounced at the C-terminus all alpha-helical domain (Figure 6.10C) of this obligate homodimeric enzyme.

The Rg reflects the hydrodynamic volume of a globular protein, showing that 8ALS is less compact in comparison with 1OE8. This further confirms the variation in overall conformational dynamics as judged by the Cα-RMSD trajectory. The overall average Rg (±SD) for 8ALS and 1OE8 is 22.75Å (±0.42) and 21.44 Å (±0.09), respectively. To further support our postulate that 8ALS is more dynamic than the 1EO8 model, we measured the distance between two atoms in helix4 (α4), which forms part of the dimer interface topology. Using His110Cα and Asp89 C^{α} as

reference atoms, we observed that the distance between the Cα of the two His110 (of the two subunits) fluctuated more in 8ALS than 1OE8; however, this was not the case for Asp89C^{α} of the two subunits.

6.5. Discussion

Schistosomiasis is a tropical disease of poverty which ranks as the second cause of death among parasitic diseases (Adenowo *et al.*, 2015). The present study aimed to evaluate the unbound crystal assembly of the 28-kDa isoform of GST from *S. haematobium*. In this regard, protein crystal structures are essential since they highlight mechanisms of biological activities, thereby facilitating drug development to exploit protein function. In this research, the crystal structures of unbound Sh28GST were solved to provide insight into the proteins' structural conformations. Analysing the un-ligated form of Sh28GST facilitates the ability to selectively inhibit the activity of Sh28GST can be achieved through the design of molecular compounds.

The complete DNA sequence of Sh28GST was inserted into the pET-11a vector to express the recombinant protein in the soluble fraction of *E. coli* cells. The purity of the protein was assessed by SDS-PAGE analysis according to Section 3.1. The purified protein displayed a single band corresponding to a subunit molecular weight of approximately 25 kDa on a 12.5% SDS-PAGE gel (Figure 6.1). The Protparam bioinformatics tool was utilized to calculate the molecular weight, which yielded a theoretical value of 23.4 kDa based on the net isotopic mass of each amino acid (Gasteiger *et al.*, 2005). Discrepancies between the two molecular weights are due to the actual number of amino acids present in the Sh28GST sequence as opposed to assumption of residues.

Sh28GST contains two ligand binding sites per subunit, namely the GSH-binding site (G-site) located in the N-terminal domain and the hydrophobic xenobiotic binding site (H-site) found in the C-terminal domain (Mohsenzadeh *et al.*, 2011). To evaluate the enzymes' catalytic activity, its ability to catalyze the covalent conjugation of GSH to CDNB and produce 1-(S-glutathionyl)-2, 4-dinitrobenzene (the product) was monitored. The results indicate that Sh28GST is enzymatically active. The specific activity was calculated by determining the slope of a linear plot from the experimental data (Figure 6.2A).

The secondary structural content of Sh28GST agrees with literature which categorises cytosolic GSTs as predominantly alpha-helical (Gildenhuys *et al.*, 2011). The 222 nm/208 nm ratio of 8ALS

to 1OE8 was reported to be 0.73 and 0.82, respectively, indicating that the secondary structure of Sh28GST comprised more coiled-coil α-helices. These differences in ellipticity amongst predominantly alpha-helical proteins depict the natural variation of the helices, thereby classifying them into coiled-coil helices or random coils if the 222 nm/208 nm ratio exists as ≥ 1 or ≤ 0.86, respectively Truebeisten *et al.,* 2016. This result confirmed that the integrity of the secondary structure of the enzyme was not compromised during recovery from the soluble *E. coli* cell fraction.

Predictably, the endogenous tryptophan fluorescence intensity of Sh28GST was remarkably low during the experiment, the local tryptophan environment was altered in the presence of GSH, indicating that the natural tryptophan residues were affected by the solvent (Lacowicz, 2006). The binding of ANS resulted in a significantly higher fluorescence quantum yield compared to unbound ANS, which is consistent with previous fluorescence studies (Sayed *et al.,* 2002). To compete with ANS for binding pockets, GSH was added in the ANS binding studies to block the "G" site. However, extrinsic fluorescence spectra remained unchanged, confirming that ANS binds to Sh28GST without affecting the protein's conformational stability.

There are vast differences between the conformational changes of ligated and un-ligated GSTs in solution. Ligands can confer stability to proteins (Hassell *et al.,* 2007). Both the Sh28GST (PDB 8ALS) and the Sh28GST (PDB 1OE8) are within the same space group of P1 21 1. In this situation, it is anticipated that the ligand will not cause any major changes in the proteins' conformation. Therefore, obtaining crystals of the unliganded form of Sh28GST was deemed relevant for BSP binding studies. The examination of the crystal packing parameters showed that the crystal lattice contains two biological molecules in the asymmetric unit, with a Matthew's coefficient of 2.37 Å Da^{-1} and a solvent content of 48.1%. Additional details on data collection can be found in Table 6.1.

The conformation of the polypeptide backbone of the refined Sh28GST model was assessed to ensure its structural accuracy. Analysis of the main chain angles showed that the majority, 97.28%, are within the preferred region of the Ramachandran plot. The remaining 2.2% are in the allowed region, while 0.42% occupy the less favorable generously allowed region. The protein structure has been deposited as entry 8ALS in the Protein Data Bank. The Sh28GST structure is dimeric in solution. This dimeric feature is also exhibited by other GST family members (Fabrini *et al.,* 2007).

The model exhibits common GST features such as the N and C terminals on the domain interface, with an additional sodium ion (Figure 6.4). Monomers A and B comprised of 204 residues. The two monomers look similar, howbeit, they exhibit very small similarities (Figure 6.3B). Sh28GST encodes a Tyrosine residue at position 10, within the activating conformation position (Johnson *et al.,* 2003). The enzyme displayed canonical GST characteristics. Anti- schistosomal drug design can be made possible by switching the position of the catalytic Tyrosine10 with inhibitors such as BSP (Valli and Achilonu, 2022). A total of 151 bonded contacts (dominated by non-polar interactions) containing eight salt links were present. These are 2 extra salt bridges in Sh28GST PDB ID 8ALS as compared to its Sh28GST (PDB 10E8) counterpart. Salt bridges result in increased thermal stability of the enzyme (Lee *et al.*, 2014).

Normalised B-factors have been widely used as benchmarks for protein structure analysis to analyse motion in varying sections of the protein. For example, better insight into a proteins active site, secondary structure or buried clefts can be gained. Here, despite similar B-factor data in the protein core region, we observed a statistically significant differences between the structures due to the surface residue displacements of 1OE8. The cavity analysis showed two corresponding cavities on both the bound and free structures, including one cavity biding the active site and another within the dimer interface core. The free 8ALS structure exhibited larger areas and volumes of both cavities. This observation is likely the result of minor active site modifications as the bound form has adjusted to accommodate the glutathione ligand, thereby reducing the cavity size.

Having observed the differences in amino acid composition between our model (8ALS) and the molecular replacement template (1OE8), we hypothesised that these six amino acid variations might impact the global structure of 8ALS. This implies that there may be variations in the structure and catalytic turnover rate of 8ALS in comparison with 1OE8. Several studies have strongly correlated amino acid mutation in the form of substitution, and insertions to variations in structure, dynamics, and kinetics of enzymes (Radkiewicz and Brooks, 2000; Schnell *et al.,* 2004 and Agarwal, 2006). Five of the six substitutions are on the predominantly alpha-helical C-terminus domain (Figure 6.9). In contrast, the only substitution on the N-terminus domain appears to be withdrawn from the hydrophobic xenobiotic and glutathione binding site. The substitutions occurred mostly between α4 and α5, which are the two main helices of the dimer interface, which indicates that the variation may affect the hydrodynamic volume of this globular enzyme.

Although the variation may not seem to visibly induce any conformational changes in 8ALS compared to 1OE8, from the static structure point of view, it was critical that we used computational approach to unravel the structural landscape of the two proteins from an atomistic perspective. Therefore, it is crucial in this study's context to gain an atomistic interpretation of the variance in amino acid composition between what we postulate as isoenzymes of the 28-kDa *S. haematobium* GST. To support our theory, A 500 ns molecular dynamics simulation was conducted to investigate the overall conformation of 8ALS in relation to 1OE8 (Figure 5.9), using Cα-RMSD (alpha carbon), Cα-RMSF, the Rg and the distance between corresponding catalytic atoms of the two subunits at the dimer interface. We supposed that 500 ns trajectory time is enough to explain our hypothesis based on the C^{α}-RMSD, C^{α}-RMSF, Rg and interdomain distance measurement. Molecular dynamic simulation of biological systems has shown relevance in interpreting empirical data. Although the crystallographic resolution may not be the same between these two model structures, we could still make reasonable extrapolations from MD simulation data. Our results suggest that there may be variation in the overall dynamics and hydrodynamic volume of 8ALS with respect to 1OE8. This may imply that the two proteins may have different kinetics, but not in function. Hence, we can justifiably denote to 8ALS as a variant of 1OE8, not just based on amino acid composition, but on enzyme kinetics.

We expected that this variation might also alter the dimer interface topology and the conformational stability of 8ALS. We measured the distance between the two dimers based on His^{110} and Asp^{89} inter-subunit distance as a function of the 500 ns trajectory period. The results show that the distance between the dimers vary across the two proteins, suggesting that the variation in amino acid sequence does alter the topology of the dimer interface. This is corroborated by the dimer interface analysis results (Figure 6.10), which show that the volume of the 8ALS dimer interface is larger than that of the 1OE8. It is important to mention that the dimer interface of GSTs plays a critical role in the structure and function of the enzyme, resulting in an altered hydrodynamic volume. GSTs are obligate dimers, except for *Plasmodium falciparum* GST, which shows tetramerisation when GSH is missing, but dimerises when GSH is available. Hence, any major perturbation at the dimer interface, such as alteration in the distance between the

subunits, may affect the kinetic propensity of the protein, such as turnover rate. It suffices to say that the variation in the amino acid across the two enzymes does not entirely alter the function but could affect the kinetics as well as stability of the enzyme. An empirical approach, such as chemical-induced equilibrium denaturation, would be needed to affirm this postulate.

To sum up, the study accomplished the overexpression and purification of Sh28GST, which allowed for an examination of its structural and functional characteristics, including crystal structure determination. The acquired data is beneficial for potential applications of the Sh28GST enzyme without ligands. Moreover, the crystal structures obtained provide a foundation for the development of inhibitors against schistosomiasis with therapeutic benefits using structure-based drug design. Further X-ray crystallographic analysis of liganded-Sh28GST will confirm the binding site for BSP.

Chapter 7
General discussion, conclusion and future work

Schistosomiasis, an often-overlooked tropical illness caused by parasitic flatworms, is a significant public health concern in numerous underdeveloped African countries and is frequently linked to inadequate sanitation and poverty. Addressing neglected diseases like schistosomiasis can have significant economic benefits by improving the health and productivity of affected populations, which in turn can contribute to the overall economic growth of endemic regions. Identification of potential substrate(s) or inhibitor(s) for parasite GSTs may contribute to the development of more effective and targeted treatments for this neglected disease.

Currently, schistosomiasis can only be treated with Praziquantel, which is more effective against adult worms than schistosomulae. However, in regions where the disease is prevalent, some parasites have developed resistance to Praziquantel, and there are no alternative drugs available. The lack of an alternative drug has become a significant concern for the public health sector, and A pressing requirement exists for the creation of a novel medication. In parasitic helminths, GSTs, which are phase II detoxification enzymes, have been identified as potential therapeutic drug targets.

In eukaryotic organisms, GSTs have various functions, including protecting cells from oxidative damage by detoxifying harmful substances, participating in sterol synthesis and metabolism, and regulating signaling pathways. These enzymes, which belong to a family of multiple genes encoding isoenzymes, are present in most living organisms and are classified as Phase II enzymes are accountable for the catalysis of GSH conjugation with various xenobiotics. Although this detoxification process provides cellular protection against oxidative and environmental stress, it can also contribute to drug resistance. Recent molecular biology advancements have revealed additional functions of GSTs, such as their involvement in the biosynthesis and metabolism of prostaglandins, steroids, and leukotrienes, as well as their management of toxic byproducts generated by oxidative stress, such as S-glutathiolated proteins. Moreover, GSTs play a part in the emergence of resistance to chemotherapeutic drugs.

Sex hormones play a vital role in determining the level of infection based on age and gender in parasitic diseases such as schistosomiasis. In mammals, GSTs have been shown to bind to hormones, particularly sexual steroids, which can affect their transport, metabolism, and physiological functions. This suggests that the regulatory functions of sex steroids on antiparasitic immunity maybe influenced by GST activity. Notably, studies have found that female hormones can enhance antibody production response against specific antigens, making women more resistant to several parasitic infections (Remoue, *et al.,* 2002). In addition, high levels of testosterone after puberty act as important immune modulators, reducing susceptibility to infection with age. The binding of sex steroids to GST can directly impact parasite fecundity by acting on both worm metabolism and parasite reproductive organs. This functional binding could influence the level of infection based on age and sex in human schistosomiasis.

In any protein biochemistry research, the expression and purification of a protein are often the rate-limiting or bottleneck steps. To overcome this challenge, pet-11a and pet-28a vectors were synthesized by incorporating a non-cleavable 6×his-tag. Successful expression and purification of the protein was achieved, resulting in a high concentration of purified protein. This is advantageous for conducting both structural and functional studies on the protein. The purified Sbh26GST, Sh28GST and Human GST-mu were deemed active based on the GSH-CDNB conjugation enzyme activity assay. Furthermore, the activity of the protein was found to decrease in the presence of testosterone, confirming that testosterone does binds to and inhibits these GSTs.

To ensure accurate downstream analysis in protein structural-functional research, it is crucial to obtain high yields of recombinant protein and samples that are highly purified. To achieve this, IMAC was utilized in this study to purify clarified supernatants and obtain highly pure samples of Sbh26GST, Sh28GSTand Human GST-mu. To confirm that the purified Sbh26GST, Sh28GST and Human GST-mu were folded correctly according to their expected secondary structure, far-UV circular dichroism was utilized as a structural probe. The typical secondary structure of cytosolic GST proteins consists of predominantly alpha helices. The Far-UV CD profile of Sbh26GST, Sh28GST and Human GST-mu's showed characteristic troughs between 206 and 224 nm for all the 3 GSTs which positively confirmed that the GSTs were predominantly alpha-helical in structure (Greenfield, 2006). This indicates that the secondary structure of the enzyme was not compromised during its recovery from the soluble fraction.

The spectra obtained suggested that some of the tryptophan's are partially exposed to the solvent, with the highest emission wavelength observed at 340 nm. The intensity of the fluorescence emission spectrum decreased upon testosterone binding to the Sbh26GST, Sh28GST and Human GST-mu which indicated that the tryptophan's local environment was more polar or solvent-exposed. This change does not necessarily indicate a shift in the protein's overall structure, as evidenced by the circular dichroism data that showed minimal changes in the secondary structure of the GSTs that were used in this study.

Fluorescence spectroscopy, using the extrinsic fluorophore ANS, was used to probe binding sites of Sbh26GST, Sh28GST and Human GST-mu and in addition monitor the binding of testosterone to these GSTs. Changes in the surroundings of the fluorophore causes changes in the fluorescent intensity and shifts in the maximum wavelength emission (Rodrigues *et al.,* 2023). When ANS binds to parasite GSTs and Human GST-mu in the absence of testosterone, it results in an increase in fluorescent intensity and a blue shift in the maximum emission wavelength, indicating a decrease in the polarity of the fluorophore's environment (Rodrigues *et al.,* 2023). This indicates that ANS binds to the hydrophobic patches on the surface of these GSTs. However, ANS alone exhibits a reduction in fluorescent intensity and a red shift due to an increase in solvent polarity surrounding the fluorophore (Möller and Denicola, 2002)

To determine the effectiveness of testosterone in inhibiting Sbh26GST, Sh28GST, and Human GST-mu activity, we assessed the IC_{50} value. However, it is essential to determine the type of inhibition that occurs during inhibitor-enzyme interactions as the IC_{50} value alone can be misleading. Non-linear regression in enzyme kinetics was used to accurately determine kinetic parameters such as K_m and V_{max}, both in the presence and absence of testosterone. This method eliminated the need for linear reciprocal plots, which tend to be unreliable and inaccurate, such as the Lineweaver-Burke plot, providing a more comprehensive understanding. Our study revealed that testosterone acted as an uncompetitive inhibitor, leading to a decrease in both K_m and V_{max} values in the presence of the inhibitor. The observed IC_{50} value suggested a high affinity of testosterone towards Sbh26GST and Sh28GST, but it should be noted that in uncompetitive inhibition, the IC_{50} value is substrate concentration dependent.

However, our attempt to investigate the kinetic effects of testosterone on parasite GSTs by varying the concentration of CDNB in the presence of testosterone was unsuccessful, as the curves did not reach saturation even at the highest assayed CDNB concentration. As a result, the kinetic implications of testosterone on parasite GSTs in this context could not be accurately determined.

In the field of bioinformatics, machine learning has become a crucial instrument for scientists to process and comprehend massive and intricate biological data sets. Bioinformatics encompasses computational techniques utilized for the analysis and interpretation of biological data. In contrast, machine learning, a division of artificial intelligence, involves the implementation of algorithms and statistical models to enable computers to learn from data and make forecasts or decisions (Xia, 2017). By employing IFMD and molecular docking simulations, the computational analysis demonstrated that the interaction of testosterone to Sbh26GST and Sh28GST enhanced the dynamics of the protein, as observed in the simulation. The simulations of the three systems were well-equilibrated during the 500 ns simulations. To describe the quality of the simulated systems, the simulation quality analysis algorithm was employed. Our findings demonstrate that the binding affinity of testosterone to parasite GSTs is stronger than its binding to human GST-mu. It appears that testosterone binds specifically to the dimer interface of parasite GSTs in a similar fashion as how ellagic acid interacts with Sj26GST (Akumadu *et al.,* 2020), whereas in human GST-mu, testosterone binds to two different sites; the L-site (dimer interface) and the H-site. It's important to note that ligand docking studies may not be completely dependable without molecular dynamics simulations. The RMSD of the alpha carbons for all three systems were comparable, indicating that the simulated structures were suitable for downstream analysis.

In summary, this research aims to explore the development of new anti-*Schistosoma* drugs using a rational approach, which typically involves identifying a drug-target within the organism. The target identified in this study is the Sbh26GST and Sh28GST. The study does not claim that testosterone is a potential anti-Schistosoma drug, but rather it suggests that it could be used as a starting point for drug discovery. The study involved expressing, purifying, and quantifying the GST enzyme, and conducting various structural, functional, thermodynamic, and computational studies to evaluate its potential inhibition by testosterone. This study may offer insights into the mechanisms underlying the parasites' resistance to current treatments and may lead to the development of new drugs to combat schistosomiasis. The results of the study provide further

knowledge about the binding and inhibitory modes of testosterone on schistosome GSTs, which may be a potential candidate for the development of novel anti- schistosomal drugs to alleviate the debilitating effects of schistosomiasis.

In future, it would be beneficial to explore additional computational techniques such as PCA,drug binding studies, principal and per-residue energy decomposition calculations, and MM/GBSA free binding energy calculations, to gain valuable insights for the rational discovery, design, and development of drugs. It would also be advantageous to utilize surface plasmon resonance (SPR) and hydrogen deuterium exchange mass spectrometry (HDX-MS) techniques tofurther investigate the decrease in fluorescent intensity caused by testosterone displacing ANS.

References

Abdul-Ghani, R.A., Hassan, A.A. Murine schistosomiasis as a model for human schistosomiasis mansoni: similarities and discrepancies. *Parasitol Res* **107**, 1–8 (2010). https://doi.org/10.1007/s00436-010-1855-5

Adekiya, T. A., Kumar, P., Kondiah, P. P.D., Pillay, V., & Choonara, Y. E. (2021). Synthesis andTherapeutic Delivery Approaches for Praziquantel: A Patent Review (2010-Present). *ExpertOpinion on Therapeutic Patents*, *31*(9), 851-865. https://doi.org/10.1080/13543776.2021.1915292.

Adenowo, A. F., Oyinloye, B. E., Ogunyinka, B. I., & Kappo, A. P. (2015). Impact of human schistosomiasis in sub-Saharan Africa. *The Brazilian journal of infectious diseases : an officialpublication of the Brazilian Society of Infectious Diseases*, *19*(2), 196–205. https://doi.org/10.1016/j.bjid.2014.11.004.

Agarwal P. K. (2006). Enzymes: An integrated view of structure, dynamics and function. *Microbialcell factories*, 5(2). https://doi.org/10.1186/1475-2859-5-2.

Aguilar-Díaz, Hugo & Ke, Nava. (2015). Endocrine Immune Interactions in the Host-Parasite Relationship: Steroid Hormones as Immune Regulators in Parasite Infections. Journal of Steroids & Hormonal Science. https://doi.org/06.10.472/2157-7536.1000165.

Al Abaidani, I., Al-Abri, S., Shaban, M., Ghugey, S. L., Al Kathery, S., Al-Mashikhi, K., Garba, A.,& Gabrielli, A. F.(2016). Decline in transmission of schistosomiasis mansoni in Oman. *Infectious diseases of poverty*, *5*(1), 112. https://doi.org/10.1186/s40249-016-0210-1.

Alexander J., & Stimson W.H. (1988) Sex Hormones and the Course of Parasitic Infection. *Parasitology Today 49*(7), 189–93. https://doi.org/10.1016/0169-4758(88)90077-4.

Allocati, N., Masulli, M., Di Ilio, C., & Federici, L. (2018). Glutathione transferases: substrates,inihibitors and pro-drugs in cancer and neurodegenerative diseases. *Oncogenesis*, *7*(1), 8. https://doi.org/10.1038/s41389-017-0025-3.

Andrews, P. (1985). Praziquantel: Mechanisms of Anti-*Schistosoma*l Activity. *Pharmacology &Therapeutics*, *29*(1), 129–156. https://doi.org/10.1016/0163-7258(85)90020-8.

Angelucci, F., Baiocco, P., Brunori, M., Gourlay, L., Morea, V., & Bellelli, A. (2005). Insights intothe catalytic mechanism of glutathione S-transferase: the lesson from *Schistosoma* haematobium. *Structure (London, England : 1993)*, *13*(9), 1241–1246. https://doi.org/10.1016/j.str.2005.06.007.

Aruleba, R. T., Adekiya, T. A., Oyinloye, B. E., Masamba, P., Mbatha, L. S., Pretorius, A., & Kappo, A. P. (2019). PZQ Therapy: How Close are we in the Development of Effective Alternative Anti-schistosomal Drugs?. *Infectious disorders drug targets*, *19*(4), 337–349. https://doi.org/10.2174/1871526519666181231153139

Ballatori, N., Krance, S. M., Notenboom, S., Shi, S., Tieu, K., & Hammond, C. L. (2009). Glutathionedysregulation and the etiology and progression of human diseases. *Biological chemistry*, *390*(3), 191–214. https://doi.org/10.1515/BC.2009.033.

Bell, M. R., Engleka, M. J., Malik, A., & Strickler, J. E. (2013). To fuse or not to fuse: what is your purpose? *Protein science : a publication of the Protein Society*, *22*(11), 1466–1477. https://doi.org/10.1002/pro.2356.

Bennett, N. C., Gardiner, R. A., Hooper, J. D., Johnson, D. W., & Gobe, G. C. (2010). Molecular CellBiology of Androgen Receptor Signaling. *The International Journal of Biochemistry & Cell Biology*, *42*(6), 813-827. https://doi.org/10.1016/j.biocel.2009.11.013.

Bennett P. M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British journal of pharmacology*, *153*(Suppl 1), S347–S357. https://doi.org/10.1038/sj.bjp.0707607.

Beuckmann, C. T., Fujimori, K., Urade, Y., Hayaishi, O., & Gobe, G. C. (2000). Identification of Mu-Class Glutathione Transferases M2-2 and M3-3 as Cytosolic Prostaglandin E Synthases in the Human Brain. *Neurochemical Research, 25*(5), 733-738. https://doi.org/10.1023/a:1007579507804.

BIO Ventures for Global Health, South African Technology Innovation Agency (2011). South Africa becomes first government to use the Pool for Open Innovation to stimulate neglected disease drug research and development. Pool for Open Innovation against neglected tropical diseases. Accessed 2 September 2023. Available: http://ntdpool.org/news/releases/southafrica-becomes-first-government-use-pool-ope.

Birdsell, D. C., & Cota-Robles, E. H. (1967). Production and Ultrastructure of Lysozyme and Ethylenediaminetetraacetate-Lysozyme Spheroplasts of Escherichia Coli. *Journal of Bacteriology, 93*(1), 427-437. https://doi.org/10.1128/jb.93.1.427-437.1967.

Bocedi, A., Fabrini, R., Lai, O. *et al.* Erythrocyte glutathione transferase: a general probe for chemical contaminations in mammals. *Cell Death Discovery* **2**, 16029 (2016). https://doi.org/10.1038/cddiscovery.2016.29

Bornhorst, J. A., & Falke, J. J. (2000). Purification of proteins using polyhistidine affinity tags. *Methods in enzymology*, *326*, 245–254. https://doi.org/10.1016/s0076-6879(00)26058-8.

Botros, S. S., & Bennett, J. L. (2007). Praziquantel Resistance. *Expert Opinion on Drug Discovery,2*(1), 35-40. https://doi.org/10.1517/17460441.2.S1.S35.

Brylinski, M., & Skolnick, J. (2008). What is the relationship between the global structures of apo and holo proteins?. *Proteins*, *70*(2), 363–377. https://doi.org/10.1002/prot.21510.

Buhaescu, I., & Izzedine, H. (2007). Mevalonate pathway: a review of clinical and therapeuticalimplications. *Clinical biochemistry*, *40*(9-10), 575–584. https://doi.org/10.1016/j.clinbiochem.2007.03.016.

Bundy, D. A. (1988). Gender-Dependent Patterns of Infections and Disease. *Parasitology Today(Personal Ed.), 4*(7), 186–189. https://doi.org/10.1016/0169-4758(88)90076-2.

Bryant , J. H., & Harrison, P. F. (1996). *Global Health in Transition: A Synthesis: Perspectives from International Organizations*. National Library of Medicine. Retrieved September 9, 2022, from https://www.ncbi.nlm.nih.gov/books/NBK209096/.

Caffrey, C. R., Rohwer, A., Oellien, F., Marhöfer, R. J., Braschi, S., Oliveira, G., McKerrow, J. H., &Selzer, P. M.(2009). A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, *Schistosoma* mansoni. *PloS one*, *4*(2), e4413. https://doi.org/10.1371/journal.pone.0004413.

Calvete J. J. (2017). Venomics: integrative venom proteomics and beyond. The Biochemical journal, 474(5), 611 – 634. https://doi.org/10.1042/BCJ20160577.

Centres for Disease Control and Prevention (2020). *About Schistosomiasis*. Retrieved September 9,2022, from https://www.cdc.gov/parasites/schistosomiasis/gen_info/faqs.html.

Chen, H. C., & Farese, R. V. (1999). Steroid hormones: Interactions with membrane-bound receptors. *Current biology : CB*, *9*(13), R478–R481. https://doi.org/10.1016/s0960-9822(99)80300-5

Chen M. G. (2014). Assessment of morbidity due to Schistosoma japonicum infection in China. *Infectious diseases of poverty*, 3 (1), 6. https://doi.org/10.1186/2049-9957-3-6.

Chronopoulou, E., Madesis, P., Asimakopoulou, B., Platis, D., Tsaftaris, A., & Labrou, N. E. (2012). Catalytic and structural diversity of the fluazifop-inducible glutathione transferases from Phaseolus vulgaris. Planta, 235(6), 1253–1269. https://doi.org/10.1007/s00425-011-1572-z

Chrysina, E. D., Brew, K., & Acharya, K. R. (2000). Crystal structures of apo- and holo-bovine alpha-lactalbumin at 2. 2-A resolution reveal an effect of calcium on inter-lobe interactions. *The Journal of biological chemistry*, *275*(47), 37021–37029. https://doi.org/10.1074/jbc.M004752200.

Clark, J. J., Benson, M. L., Smith, R. D., & Carlson, H. A. (2019). Inherent versus induced protein flexibility: Comparisons within and between apo and holo structures. *PLoS computational biology*, *15*(1), e1006705. https://doi.org/10.1371/journal.pcbi.1006705.

Coates, A. R. M., Hu, Y., Holt, J., & Yeh, P. (2020). Antibiotic Combination Therapy against Resistant Bacterial Infections: Synergy, Rejuvenation and Resistance Reduction. *Expert Review of Anti-Infective Therapy*, *18*(1), 5-15. https://doi.org/10.1080/14787210.2020.1705155.

Coecke, S., Vanhaecke, T., Foriers, A., Phillips, I. R., Vercruysse, A., Shephard, E. A., & Rogiers, V. (2000). Hormonal regulation of glutathione S-transferase expression in co-cultured adult rat hepatocytes. *The Journal of endocrinology*, 166 (2), 363–371. https://doi.org/10.1677/joe.0.1660363.

Colley, D. G., Bustinduy, A. L., Secor, W. E., & King, C. H. (2014). Human schistosomiasis. *Lancet (London, England)*, *383*(9936), 2253–2264. https://doi.org/10.1016/S0140-6736(13)61949-2.

Cooper, A. J. L., & Hanigan, M. H. (2018). Metabolism of Glutathione S-Conjugates: Multiple Pathways. Comprehensive Toxicology, 363–406. https://doi.org/10.1016/B978-0-12-801238-3.01973-5.

Cui, J., Li, G., Yin, J., Li, L., Tan, Y., Wei, H., Liu, B., Deng, L., Tang, J., Chen, Y., & Yi, L. (2020). GSTP1 and cancer: Expression, methylation, polymorphisms and signaling (Review). *International Journal of Oncology*, *56*(4), 867-878. https://doi.org/10.3892/ijo.2020.4979.

Daikos, G. L., Tsaousi, S., Tzouvelekis, L. S., Anyfantis, I., Psichogiou, M., Argyropoulou, A., Stefanou, I., Sypsa, V.,Miriagou, V., Nepka, M., Georgiadou, S., Markogiannakis, A., Goukos, D., & Skoutelis, A. (2014). Carbapenemase-producing Klebsiella pneumoniae blood stream infections: lowering mortality by antibiotic combination schemes and the role of carbapenems. *Antimicrobial agentsand chemotherapy*, *58*(4), 2322–2328. https://doi.org/10.1128/AAC.02166-13.

Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. M., & Ahsan, M. J. (2022). Machine Learning in Drug Discovery: A Review. *Artificial intelligence review*, *55*(3), 1947–1999. https://doi.org/10.1007/s10462-021-10058-4

De Boni, L., Msimang, V., De Voux, A., & Frean, J. (2021). Trends in the prevalence of microscopically-confirmed schistosomiasis in the South African public health sector, 2011-2018. *PLoS neglected tropical diseases*, *15*(9), e0009669. https://doi.org/10.1371/journal.pntd.0009669.

De Leo G A, Stensgaard A, Sokolow S H, Nâ€™Goran E K, Chamberlin A J, Yang G et al. Schistosomiasis and climate change BMJ 2020; 371: m4324 doi:10.1136/bmj.m4324

de Mendonça, R. L., Escrivá, H., Bouton, D., Laudet, V., & Pierce, R. J. (2000). Hormones and nuclear receptors in schistosome development. *Parasitology today (Personal ed.)*, *16*(6), 233–240. https://doi.org/10.1016/s0169-4758(00)01641-0.

Dessì, M., Noce, A., Dawood, K. F., Galli, F., Taccone-Gallucci, M., Fabrini, R., Bocedi, A., Massoud, R., Fucci, G., Pastore, A., Manca di Villahermosa, S., Zingaretti, V., Federici, G., &Ricci, G. (2012). Erythrocyte glutathione transferase: a potential new biomarker in chronic kidney diseases which correlates with plasma homocysteine. *Amino acids*, *43*(1), 347–354. https://doi.org/10.1007/s00726-011-1085-x.

Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., & Courvalin, P. (2007). Modes and modulations of antibiotic resistance gene expression. *Clinical microbiology reviews*, 20(1), 79–114. https://doi.org/10.1128/CMR.00015-06.

Dirr, H. W., Little, T., Kuhnert, D. C., & Sayed, Y. (2005). A conserved N-capping motif contributes significantly to the stabilization and dynamics of the C-terminal region of class Alpha glutathione S-transferases. *The Journal of biological chemistry*, *280*(20), 19480–19487. https://doi.org/10.1074/jbc.M413608200.

Dirr, H. W., Reinemer, P., & Huber, R. (1994). X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. *European journal of biochemistry*, *220*(3), 645–661. https://doi.org/10.1111/j.1432-1033. 1994.tb18666. x.

Dominiczak, M. H., Beastall, G., & Wallace, A. M. (2014). Biosynthesis of cholesterol and steroids. Medical Biochemistry E-Book, 200, 379-387.

El Ridi, R. A., & Tallima, H. A. (2013). Novel therapeutic and prevention approaches for schistosomiasis: review. *Journal of advanced research*, *4*(5), 467–478. https://doi.org/10.1016/j.jare.2012.05.002.

Eloi-Santos, S., Olsen, N. J., Correa-Oliveira, R., & Colley, D. G. (1992). *Schistosoma* mansoni:mortality, pathophysiology, and susceptibility differences in male and female mice. *Experimental parasitology*, *75*(2), 168–175. https://doi.org/10.1016/0014-4894(92)90176-b.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 4), 486–501. https://doi.org/10.1107/S0907444910007493.

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G., Lo Bello, M., & Ricci, G. (2009). Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. Biochemistry, 48(43), 10473–10482. https://doi.org/10.1021/bi901238t.

Fabrini, R., Bocedi, A., Dawood, K. F., Turella, P., Stella, L., Parker, M. W., Pedersen, J. Z., Federici,G., Antonini, G., & Ricci, G. (2011). The extended catalysis of glutathione transferase. *FEBS letters*, *585*(2), 341–345. https://doi.org/10.1016/j.febslet.2010.12.009.

Fonseca, C. T., Oliveira, S. C., & Alves, C. C. (2015). Eliminating Schistosomes through Vaccination: What are the Best Immune Weapons?. *Frontiers in immunology*, *6*, 95. https://doi.org/10.3389/fimmu.2015.00095.

Gasymov, O. K., & Glasgow, B. J. (2007). ANS fluorescence: potential to augment the identification of the external binding sites of proteins. *Biochimica et biophysica acta*, *1774*(3), 403–411. https://doi.org/10.1016/j.bbapap.2007.01.002

Gibrat, G., Assairi, L., Craescu, C. T., Hui Bon Hoa, G., Loew, D., Lombard, B., Blouquit, L., & Bellissent-Funel, M. C. (2012). Use of SANS and biophysical techniques to reveal subtle conformational differences between native apo-calmodulin and its unfolded states. *Biochimica et biophysica acta*, *1824*(10), 1097–1106. https://doi.org/10.1016/j.bbapap.2012.06.001.

Gildenhuys, S., Dobreva, M., Kinsley, N., Sayed, Y., Burke, J., Pelly, S., Gordon, G. P., Sayed, M., Sewell, T., & Dirr, H. W. (2010). Arginine 15 stabilizes an S(N)Ar reaction transition state and the binding of anionic ligands at the active site of human glutathione transferase A1-1. Biophysical chemistry, 146(2-3), 118–125. https://doi.org/10.1016/j.bpc.2009.11.003.

García-Gómez, E., González-Pedrajo, B., & Camacho-Arroyo, I. (2013). Role of sex steroid hormones in bacterial-host interactions. *BioMed research international*, *2013*, 928290. https://doi.org/10.1155/2013/928290.

Greenfield N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols*, *1*(6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.

Gustafsson, C., Govindarajan, S., & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in biotechnology*, *22*(7), 346–353. https://doi.org/10.1016/j.tibtech.2004.04.006.

Habig, W. H., Pabst, M. J., & Jakoby, W. B. (1974). Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *The Journal of biological chemistry*, *249*(22), 7130–7139.

Harwaldt, P., Rahlfs, S., & Becker, K. (2002). Glutathione S-transferase of the malarial parasite Plasmodium falciparum: characterization of a potential drug target. *Biological chemistry*, *383*(5),821–830. https://doi.org/10.1515/BC.2002.086.

Hassan, S., Karpova, Y., Baiz, D., Yancey, D., Pullikuth, A., Flores, A., Register, T., Cline, J. M.,D'Agostino, R., Jr, Danial, N., Datta, S. R., & Kulik, G. (2013). Behavioral stress accelerates prostate cancer development in mice. *The Journal of clinical investigation*, *123*(2), 874–886.https://doi.org/10.1172/JCI63324.

Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., 3rd, Deng, S. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., & Shewchuk, L. M. (2007). Crystallization of protein-ligand complexes. *Acta crystallographica. Section D, Biological crystallography*, *63*(Pt 1), 72–79.https://doi.org/10.1107/S0907444906047020.

Hawe, A., Sutter, M., & Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. *Pharmaceutical research*, *25*(7), 1487–1499. https://doi.org/10.1007/s11095-007-9516-9.

Hayes, J. D., Flanagan, J. U., & Jowsey, I. R. (2005). Glutathione transferases. *Annual review of pharmacology and toxicology*, *45*, 51–88. https://doi.org/10.1146/annurev.pharmtox.45.120403.095857.

Hemachand, T. and Shaha, C. (2003) Functional role of sperm surface glutathione S-transferases and extracellular glutathione in the haploid spermatozoa under oxidative stress. FEBS Letters. 2003 Mar;538(1-3):14-18. https://doi.org/10.1016/s0014-5793(03)00103-0.

Hilborn, E., Stål, O., & Jansson, A. (2017). Estrogen and androgen-converting enzymes 17β- hydroxysteroid dehydrogenase and their involvement in cancer: with a special focus on 17β- hydroxysteroid dehydrogenase type 1, 2, and breast cancer. *Oncotarget*, *8*(18), 30552–30562.https://doi.org/10.18632/oncotarget.15547.

Hitchens, T. K., Mannervik, B., & Rule, G. S. (2001). Disorder-to-order transition of the active site of human class Pi glutathione transferase, GST P1-1. *Biochemistry*, *40*(39), 11660–11669. https://doi.org/10.1021/bi010909+.

Hotez, P. J., Alvarado, M., Basáñez, M. G., Bolliger, I., Bourne, R., Boussinesq, M., Brooker, S. J.,Brown, A. S., Buckle, G., Budke, C. M., Carabin, H., Coffeng, L. E., Fèvre, E. M., Fürst, T., Halasa, Y. A., Jasrasaria, R., Johns, N. E., Keiser, J., King, C. H., Lozano, R., … Naghavi, M.(2014). The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases. *PLoS neglected tropical diseases*, *8*(7),e2865. https://doi.org/10.1371/journal.pntd.0002865.

Hotez, P. J., & Kamath, A. (2009). Neglected tropical diseases in sub-saharan Africa: review of their prevalence, distribution, and disease burden. *PLoS neglected tropical diseases*, *3*(8), e412. https://doi.org/10.1371/journal.pntd.0000412.

Huyse, T., Webster, B. L., Geldof, S., Stothard, J. R., Diaw, O. T., Polman, K., & Rollinson, D. (2009). Bidirectional introgressive hybridization between a cattle and human schistosome species. *PLoS pathogens*, *5*(9), e1000571. https://doi.org/10.1371/journal.ppat.1000571.

Hyman, P., & Kelner, P. (2007). Pharmacotherapeutic uses of hormones. *The Nursing clinics of North America*, *42*(1), 1–v. https://doi.org/10.1016/j.cnur.2006.11.002.

Jakobsson, P. J., Morgenstern, R., Mancini, J., Ford-Hutchinson, A., & Persson, B. (2000). Membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG). A widespread protein superfamily. *American journal of respiratory and critical care medicine*, *161*(2 Pt 2), S20–S24. https://doi.org/10.1164/ajrccm.161.supplement_1.ltta-5.

Jancova, P., Anzenbacher, P., & Anzenbacherova, E. (2010). Phase II drug metabolizing enzymes. Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia, 154(2), 103–116. https://doi.org/10.5507/bp.2010.017

Johansson, A. S., & Mannervik, B. (2001). Human glutathione transferase A3-3, a highly efficient catalyst of double-bond isomerization in the biosynthetic pathway of steroid hormones. *The Journal of biological chemistry*, *276*(35), 33061–33065. https://doi.org/10.1074/jbc.M104539200.

Johnson, K. A., Angelucci, F., Bellelli, A., Hervé, M., Fontaine, J., Tsernoglou, D., Capron, A., Trottein, F., & Brunori, M. (2003). Crystal structure of the 28 kDa glutathione S-transferase from *Schistosoma* haematobium. *Biochemistry*, *42*(34), 10084–10094. https://doi.org/10.1021/bi034449r.

Johnson, T. W., Gallego, R. A., Brooun, A., Gehlhaar, D., & McTigue, M. (2018). Reviving B-Factors: Retrospective Normalized B-Factor Analysis of c-ros Oncogene 1 Receptor Tyrosine Kinase and Anaplastic Lymphoma Kinase L1196M with Crizotinib and Lorlatinib. ACS medicinal chemistry letters, 9(9), 878–883. https://doi.org/10.1021/acsmedchemlett.8b00147

Kaplan, W., Hüsler, P., Klump, H., Erhardt, J., Sluis-Cremer, N., & Dirr, H. (1997). Conformational stability of pGEX-expressed *Schistosoma* japonicum glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. *Protein science : a publication of the Protein Society*, *6*(2), 399–406. https://doi.org/10.1002/pro.5560060216.

Katari S.K., Natarajan P., Swargam S., Kanipakam H., Pasala C., Umamaheswari A. Inhibitor design against JNK1 through e- pharmacophore modeling docking and molecular dynamics simulations. *J. Recept. Signal Transduct*. 2016; 36:558– 571.https://doi: 10.3109/10799893.2016.1141955.

Kelly, S. M., & Price, N. C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current protein & peptide science*, *1*(4), 349–384. https://doi.org/10.2174/1389203003381315

Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et biophysica acta*, *1751*(2), 119–139. https://doi.org/10.1016/j.bbapap.2005.06.005.

King, C. H., & Dangerfield-Cha, M. (2008). The unacknowledged impact of chronic schistosomiasis. *Chronic illness*, *4*(1), 65–79. https://doi.org/10.1177/1742395307084407.

Kolobe, D., Sayed, Y., & Dirr, H. W. (2004). Characterization of bromosulphophthalein binding to human glutathione S-transferase A1-1: thermodynamics and inhibition kinetics. *The Biochemicaljournal*, *382*(Pt 2), 703–709. https://doi.org/10.1042/BJ20040056.

Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy* (2nd ed.). Springer Science +Business Media,LLC. http://dx.doi.org/10.1007/978-0-387-46312-4.

Ladbury, J. E. (2001). Isothermal titration calorimetry: Application to structure based drug design. *Thermochimica Acta*, *380*.

Laemmli U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680–685. https://doi.org/10.1038/227680a0.

Laskowski, R. A., Luscombe, N. M., Swindells, M. B., & Thornton, J. M. (1996). Protein clefts in molecular recognitionand function. *Protein science : a publication of the Protein Society*, *5*(12),2438–2452. https://doi.org/10.1002/pro.5560051206.

Lasrado, N., Jia, T., Massilamany, C., Franco, R., Illes, Z., & Reddy, J. (2020). Mechanisms of hormones in autoimmunity: focus on EAE. *Biology of sex differences*, *11*(1), 50. https://doi.org/10.1186/s13293-020-00325-4.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K., Wertheim, H. F., Sumpradit, N., Vlieghe, E.,Hara, G. L., Gould, I. M., Goossens, H., Greko, C., So, A. D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A. Q., Qamar, F. N., Mir, F., Kariuki, S., … Cars, O.(2013). Antibiotic resistance-the need for global solutions. *The Lancet. Infectious diseases*, *13*(12), 1057–1098. https://doi.org/10.1016/S1473-3099(13)70318-9.

Lee, C. W., Wang, H. J., Hwang, J. K., & Tseng, C. P. (2014). Protein thermal stability enhancement by designing salt bridges: a combined computational and experimental study. *PloS one*, *9*(11), e112751. https://doi.org/10.1371/journal.pone.0112751.

Lee, J. M., Lim, H. S., & Hong, S. T. (2011). Hypersensitive reaction to praziquantel in aclonorchiasispatient. *The Korean journal of parasitology*, *49*(3), 273–275. https://doi.org/10.3347/kjp.2011.49.3.273.

Li, S., Ying, Z., Xue, Y., Sun, Z., Liu, J., & Liu, Q. (2022). Effects of different drugs and hormone treatment on Toxoplasma gondii glutathione S-transferase 2. *Parasites & vectors*, *15*(1), 461.https://doi.org/10.1186/s13071-022-05589-w.

Liang, Y. S., Dai, J. R., Zhu, Y. C., Coles, G. C., & Doenhoff, M. J. (2003). Genetic analysis of praziquantel resistance in Schistosoma mansoni. The Southeast Asian journal of tropical medicine and public health, 34(2), 274–280.

Listowsky, I., Abramovitz, M., Homma, H., & Niitsu, Y. (1988). Intracellular binding and transport of hormones and xenobiotics by glutathione-S-transferases. *Drug metabolism reviews*, *19*(3-4), 305–318. https://doi.org/10.3109/03602538808994138.

Luo, J., Liu, Z., Guo, Y., & Li, M. (2015). A structural dissection of large protein-protein crystal packing contacts.

Scientific reports, *5*, 14214. https://doi.org/10.1038/srep14214.

Mahajan, S., & Atkins, W. M. (2005). The chemistry and biology of inhibitors and pro-drugs targeted to glutathione S-transferases. *Cellular and molecular life sciences : CMLS*, *62*(11), 1221–1233. https://doi.org/10.1007/s00018-005-4524-6.

Mazari, A. M. A., Zhang, L., Ye, Z. W., Zhang, J., Tew, K. D., & Townsend, D. M. (2023). The Multifaceted Role of Glutathione S-Transferases in Health and Disease. *Biomolecules*, *13*(4), 688. https://doi.org/10.3390/biom13040688.

Mazzetti, A. P., Fiorile, M. C., Primavera, A., & Lo Bello, M. (2015). Glutathione transferases and neurodegenerative diseases. *Neurochemistry international*, *82*, 10–18. https://doi.org/10.1016/j.neuint.2015.01.008.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, *40*(Pt 4), 658–674. https://doi.org/10.1107/S0021889807021206.

McManus, D. P., & Loukas, A. (2008). Current status of vaccines for schistosomiasis. Clinical microbiology reviews, 21(1), 225–242. https://doi.org/10.1128/CMR.00046-07.

McTigue, M. A., Williams, D. R., & Tainer, J. A. (1995). Crystal structures of a *Schistosoma*l drug and vaccine target: glutathione S-transferase from *Schistosoma* japonica and its complex with the leading anti*Schistosoma*l drug praziquantel. *Journal of molecular biology*, *246*(1), 21–27. https://doi.org/10.1006/jmbi.1994.0061.

Mohsenzadeh, Sasan & Esmaeili, Maryam & Moosavi, Fatameh & Shahrtash, Maryam & Saffari, Babak & Mohabatkar, Hassan. (2011). Plant glutathione S-transferase classification, structure and evolution. African Journal of Biotechnology. https://doi.org/10. 8160-8165. 10.5897/AJB11.1024.

Molehin A. J. (2020). Schistosomiasis vaccine development: update on human clinical trials. *Journal of biomedicalscience*, *27*(1), 28. https://doi.org/10.1186/s12929-020-0621-y.

Molehin, A. J., Rojo, J. U., Siddiqui, S. Z., Gray, S. A., Carter, D., & Siddiqui, A. A. (2016). Development of a schistosomiasis vaccine. *Expert review of vaccines*, *15*(5), 619–627. https://doi.org/10.1586/14760584.2016.1131127.

Möller, M., & Denicola, A. (2002). Study of protein-ligand binding by fluorescence. *Biochemistry and Molecular Biology Education*. https://doi.org/10.1002/bmb.2002.494030050089

Moran, M., Guzman, J., Henderson, K., Liyanage, R., Wu, L., Chin, E., Chapman, N., Abela-Oversteegen, L., Gouglas, D. & Kwong, D. (2012). Neglected disease research and development: A five-year review. https://doi.org/10.1371/journal.pmed.1000030

Mukherjee, S. B., Aravinda, S., Gopalakrishnan, B., Nagpal, S., Salunke, D. M., & Shaha, C. (1999). Secretion of glutathione S-transferase isoforms in the seminiferous tubular fluid, tissue distribution and sex steroid binding by rat GSTM1. The Biochemical journal, 340 (Pt 1)(Pt 1), 309–320.

Narayanan, C., Bernard, D. N., Bafna, K., Gagné, D., Agarwal, P. K., & Doucet, N. (2018). Ligand-Induced Variations in Structural and Dynamical Properties Within an Enzyme Superfamily. Frontiers in molecular biosciences, 5, 54. https://doi.org/10.3389/fmolb.2018.00054

Neubig, R. R., Spedding, M., Kenakin, T., Christopoulos, A., & International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification (2003). International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. *Pharmacological reviews*, *55*(4), 597–606. https://doi.org/10.1124/pr.55.4.4.

Newman, J., Pearce, L., Lesburg, C. A., Strickland, C., & Peat, T. S. (2011). Crystallization of an apo form of human

arginase: using all the tools in the toolbox simultaneously. *Acta crystallographica. Section F, Structural biology and crystallization communications*, *67*(Pt 1), 90–93. https://doi.org/10.1107/S1744309110046208.

Nikolic, K., Mavridis, L., Djikic, T., Vucicevic, J., Agbaba, D., Yelekci, K., & Mitchell, J. B. (2016). Drug Design for CNS Diseases: Polypharmacological Profiling of Compounds Using Cheminformatic, 3D-QSAR and Virtual Screening Methodologies. *Frontiers in neuroscience*, *10*, 265. https://doi.org/10.3389/fnins.2016.00265.

Nour N. M. (2010). Schistosomiasis: health effects on women. *Reviews in obstetrics &gynecology*, *3*(1), 28–32.

Oakley A. (2011). Glutathione transferases: a structural perspective. *Drug metabolism reviews*, *43*(2),138–151. https://doi.org/10.3109/03602532.2011.558093.

Oakley, A. J., Lo Bello, M., Battistoni, A., Ricci, G., Rossjohn, J., Villar, H. O., & Parker, M. W. (1997). The structures of human glutathione transferase P1-1 in complex with glutathione and various inhibitors at high resolution. *Journal of molecular biology*, *274*(1), 84–100. https://doi.org/10.1006/jmbi.1997.1364.

Oakley, A. J., Lo Bello, M., Nuccetelli, M., Mazzetti, A. P., & Parker, M. W. (1999). The ligandin (non-substrate) binding site of human Pi class glutathione transferase is located in the electrophile binding site (H-site). Journal of molecular biology, 291(4), 913–926. https://doi.org/10.1006/jmbi.1999.3029.

Park, H. J., Cho, H. Y., & Kong, K. H. (2005). Purification and biochemical properties of glutathioneS-transferase from Lactuca sativa. *Journal of biochemistry and molecular biology, 38(2), 232– 237.* https://doi.org/10.5483/bmbrep.2005.38.2.232.

Paul, J.F., Verma, S and Berry K. 2002. Urinary schistosomiasis. *Emergency Medicine Journal*. ;19:483-484

Platis, M., Vlachakis, D., Foudah, A. I., Muharram, M. M., Alqarni, M. H., Papageorgiou, A. C., &Labrou, N. E. (2021). The Interaction of *Schistosoma Japonicum* Glutathione Transferase with Cibacron Blue 3GA and its Fragments. *Medicinal chemistry (Shariqah (United Arab Emirates))*, *17*(4), 332–343. https://doi.org/10.2174/1573406416666200403074742.

Pollack R. M. (2004). Enzymatic mechanisms for catalysis of enolization: ketosteroid isomerase. *Bioorganic chemistry*, *32*(5), 341–353. https://doi.org/10.1016/j.bioorg.2004.06.005.

Radkiewicz, J. L., & Brooks, C. L. (2000). Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. *Journal of the American Chemical Society*, (122(2), 225-231.

Raj, E., Calvo-Urbano, B., Heffernan, C., Halder, J., & Webster, J. P. (2022). Systematic review to evaluate a potential association between helminth infection and physical stunting in children. *Parasites & vectors*, *15*(1), 135, 1-19. https://doi.org/10.1186/s13071-022-05235-5.

Rai, M., & Padh, H. (2001). Expression systems for production of heterologous proteins. *CurrentScience*, *80*(9), 1121–1128. http://www.jstor.org/stable/24105768.

Ramsamy, Y., Essack, S. Y., Sartorius, B., Patel, M., & Mlisana, K. P. (2018). Antibiotic resistance trends of ESKAPE pathogens in Kwazulu-Natal, South Africa: A five-year retrospective analysis. *African journal of laboratory medicine*, *7*(2), 887. https://doi.org/10.4102/ajlm.v7i2.887.

Reece, J. B., & Campbell, N. A. (2002). *Biology* (pp. 119-175). San Francisco: Benjamin Cummings.

Rehman, I., Farooq, M., & Botelho, S. (2022). Biochemistry, Secondary Protein Structure. In *StatPearls*. StatPearls Publishing.

Remoué, F., Rogerie, F., Gallissot, M. C., Guyatt, H. L., Neyrinck, J. L., Diakkhate, M. M., Niang, M., Butterworth, A. E., Auriault, C., Capron, A., & Riveau, G. (2000). Sex-dependent neutralizing humoral response to Schistosoma mansoni 28GST antigen in infected human populations. *The Journal of infectious diseases*, 181(5), 1855–1859. https://doi.org/10.1086/315454.

Remoué, F., Mani, J. C., Pugnière, M., Schacht, A. M., Capron, A., & Riveau, G. (2002). Functional specific binding of testosterone to Schistosoma haematobium 28-kilodalton glutathione S-transferase. *Infection and immunity*, *70*(2), 601–605. https://doi.org/10.1128/IAI.70.2.601-605.2002.

Ricci, G., Caccuri, A. M., Lo Bello, M., Rosato, N., Mei, G., Nicotra, M., Chiessi, E., Mazzetti, A. P.,& Federici, G. (1996). Structural flexibility modulates the activity of human glutathione transferase P1-1. Role of helix 2 flexibilityin the catalytic mechanism. *The Journal of biologicalchemistry*, *271*(27), 16187–16192. https://doi.org/10.1074/jbc.271.27.16187.

Righetti P. G. (2005). Electrophoresis: the march of pennies, the march of dimes. Journal of chromatography. A, 1079(1 - 2), 24–40. https://doi.org/10.1016/j.chroma.2005.01.018.

Rijal, B., Oda, Y., Basnet, R., Rijal, B., Parajuli, K., Gurung, C. K., & Sherchand, J. B. (2001). Gender variations in the prevalence of parasitic infections and the level of awareness in adolescents in rural Nepal. *The Southeast Asian journal of tropical medicine and public health*, *32*(3), 575–580.

Robaire, B., & Hamzeh, M. (2011). Androgen action in the epididymis. *Journal of andrology*, *32*(6),592–599. https://doi.org/10.2164/jandrol.111.014266.

Roberts, C. W., Satoskar, A., & Alexander, J. (1996). Sex steroids, pregnancy-associated hormones and immunity to parasitic infection. *Parasitology today (Personal ed.)*, *12*(10), 382–388. https://doi.org/10.1016/0169-4758(96)10060-0.

Dos Santos Rodrigues, F. H., Delgado, G. G., Santana da Costa, T., & Tasic, L. (2023). Applications of fluorescence Spectroscopy in protein conformational changes and intermolecular contacts. *BBA advances*, *3*, 100091. https://doi.org/10.1016/j.bbadva.2023.100091

Romano, M. C., Jiménez, P., Miranda-Brito, C., & Valdez, R. A. (2015). Parasites and steroid hormones: corticosteroid and sex steroid synthesis, their role in the parasite physiology and development. Frontiers in neuroscience, 9, 224. https://doi.org/10.3389/fnins.2015.00224.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. *Frontiers in Microbiology*, *172*. https://doi.org/10.3389/fmicb.2014.00172.

Ruxana Begum, S. K., & Kedam, T. R. (2010). Effect of acrylamide on chick embryo liver GSTs. *Med J Nutr and Met*, *3*(1), 31-38. https://doi/10.3233/s12349-009-0070-3.

Saathoff, E., Olsen, A., Magnussen, P., Kvalsvig, J. D., Becker, W., & Appleton, C. C. (2004).Patterns of *Schistosoma* haematobium infection, impact of praziquantel treatment and re-infection after treatment in a cohort of schoolchildren from rural KwaZulu-Natal/South Africa. *BMC infectious diseases*, *4*, 40. https://doi.org/10.1186/1471-2334-4-40.

Sato K. (1989). Glutathione transferases as markers of preneoplasia and neoplasia. *Advances in cancer research*, *52*, 205–255. https://doi.org/10.1016/s0065-230x(08)60214-6.

Sayed, Y., Hornby, J. A., Lopez, M., & Dirr, H. (2002). Thermodynamics of the ligandin function of human class Alpha glutathione transferase A1-1: energetics of organic anion ligand binding. *The Biochemical journal*, *363*(Pt 2), 341–

346. https://doi.org/10.1042/0264-6021:3630341

Schaduangrat, N., Lampa, S., Simeon, S., Gleeson, M. P., Spjuth, O., & Nantasenamat, C. (2020).Towards reproducible computational drug discovery. *Journal of cheminformatics*, *12*(1), 9. https://doi.org/10.1186/s13321-020-0408-x.

Schmid, F. (2005). Spectroscopic techniques to study protein folding and stability. *Protein folding handbook*, 22-44.

Schnell, J. R., Dyson, H. J., & Wright, P. E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annual review of biophysics and biomolecular structure*, *33*, 119–140. https://doi.org/10.1146/annurev.biophys.33.110502.133613.

Schuurs, A. H., & Verheul, H. A. (1990). Effects of gender and sex steroids on the immune response. *Journal of steroid biochemistry*, *35*(2), 157–172. https://doi.org/10.1016/0022-4731(90)90270-3.

Serrano, A., Ferreira, P., Martínez-Júlvez, M., & Medina, M. (2013). The prokaryotic FAD synthetase family: a potential drug target. *Current pharmaceutical design*, *19*(14), 2637–2648. https://doi.org/10.2174/1381612811319140013.

Shaker, Y., Samy, N., & Ashour, E. (2014). Hepatobiliary Schistosomiasis. *Journal of clinical and translational hepatology*, *2*(3), 212–216. https://doi.org/10.14218/JCTH.2014.00018.

Sheehan, D., Meade, G., Foley, V. M., & Dowd, C. A. (2001). Structure, function and evolution ofglutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily.*The Biochemical journal*, *360*(Pt 1), 1–16. https://doi.org/10.1042/0264-6021:3600001.

Sheikh, B. R., & Kedam, T. (2010). Effect of acrylamide on chick embryonic liver glutathione S-transferases. *Mediterranean Journal of Nutrition and Metabolism*, *3*, 31-38. https://doi.org/10.1007/s12349-009-0070-3.

Siddiqui, A. A., & Siddiqui, S. Z. (2017). Sm-p80-Based Schistosomiasis Vaccine: Preparation forHuman Clinical Trials. *Trends in parasitology*, *33*(3), 194–201. https://doi.org/10.1016/j.pt.2016.10.010.

Skelly, P. J., Secor, W. E., Reis, M. G., Ramos, E. A., Carmo, T. M., Peixoto, E. E., & Harn, D. A.(1994). Failure of schistosomiasis to significantly decrease testosterone levels in Brazilian men. *The American journal of tropical medicine and hygiene*, *51*(1), 40–44. https://doi.org/10.4269/ajtmh.1994.51.40.

Smyth, M. S., & Martin, J. H. (2000). x ray crystallography. Molecular pathology : MP, 53(1), 8–14. https://doi.org/10.1136/mp.53.1.8.

Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., & Maertens, B. (2015). Purification of His-Tagged Proteins. *Methods in enzymology*, *559*, 1–15. https://doi.org/10.1016/bs.mie.2014.11.003.

Stank, A., Kokh, D. B., Fuller, J. C., & Wade, R. C. (2016). Protein Binding Pocket Dynamics. Accounts of chemical research, 49(5), 809–815. https://doi.org/10.1021/acs.accounts.5b00516.

Strickland E. H. (1974). Aromatic contributions to circular dichroism spectra of proteins. *CRC critical reviews in biochemistry*, *2*(1), 113–175. https://doi.org/10.3109/10409237409105445.

Swaminathan,R Vijay Kumar Ravi, Satish Kumar, Mattaparthi Venkata Satish Kumar, Nividh Chandra, Lysozyme: A model protein for amyloid research, Editor(s): Rossen Donev, Advances in Protein Chemistry and Structural Biology, Academic Press, Volume 84,2011, Pages 63-111, https://doi.org/10.1016/B978-0-12-386483-3.00003-3.

Terpe K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, *72*(2), 211–222. https://doi.org/10.1007/s00253-006-0465-8.

Tew K. D. (2016). Glutathione-Associated Enzymes In Anticancer Drug Resistance. *Cancer research*, *76*(1), 7–9. https://doi.org/10.1158/0008-5472.CAN-15-3143.

Toor, J., Alsallaq, R., Truscott, J. E., Turner, H. C., Werkman, M., Gurarie, D., King, C. H., & Anderson, R. M. (2018). Are We on Our Way to Achieving the 2020 Goals for Schistosomiasis Morbidity Control Using Current World Health Organization Guidelines?. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 66(suppl_4), S245–S252. https://doi.org/10.1093/cid/ciy001.

Torres-Rivera, A., & Landa, A. (2008). Glutathione transferases from parasites: a biochemical view. *Acta tropica*, *105*(2), 99–112. https://doi.org/10.1016/j.actatropica.2007.08.005.

Truebestein, L., & Leonard, T. A. (2016). Coiled-coils: The long and short of it. *BioEssays : news and reviews in molecular, cellular and developmental biology*, *38*(9), 903–916. https://doi.org/10.1002/bies.201600062.

Townsend, D. M., & Tew, K. D. (2003). The role of glutathione-S-transferase in anti-cancer drug resistance. *Oncogene*, *22*(47), 7369–7375. https://doi.org/10.1038/sj.onc.1206940.

Tsuchida, S. (2000). *Glutathione Transferase. Experimental Protocols for Reactive Oxygen and Nitrogen Species* (pp. 83-85). Oxford University Press.

Umemoto, K., Leffler, H., Venot, A., Valafar, H., & Prestegard, J. H. (2003). Conformational differences in liganded and unliganded states of Galectin-3. *Biochemistry*, *42*(13), 3688–3695.https://doi.org/10.1021/bi026671m

Utzinger, J., Raso, G., Brooker, S., De Savigny, D., Tanner, M., Ornbjerg, N., Singer, B. H., & N'goran, E. K. (2009). Schistosomiasis and neglected tropical diseases: towards integrated and sustainable control and a word of caution. *Parasitology*, *136*(13), 1859–1874. https://doi.org/10.1017/S0031182009991600.

Vale, N., Gouveia, M. J., Rinaldi, G., Brindley, P. J., Gärtner, F., & Correia da Costa, J. M. (2017). Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance. *Antimicrobial agents and chemotherapy*, *61*(5), e02582-16. https://doi.org/10.1128/AAC.02582-16.

Vaish, S., Gupta, D., Mehrotra, R., Mehrotra, S., & Basantani, M. K. (2020). Glutathione S- transferase: a versatile protein family. *3 Biotech*, *10*(7), 321. https://doi.org/10.1007/s13205-020-02312-3.

Valli, A., & Achilonu, I. (2022). Comparative structural analysis of the human and *Schistosoma* glutathione transferase dimer interface using selective binding of bromosulfophthalein. *Proteins*, *90*(8), 1561–1569. https://doi.org/10.1002/prot.26338.

Vibanco-Pérez, N., & Landa-Piedra, A. (1998). Glutathione S-transferase in helminth parasites. *Revista latinoamericana de microbiologia*, *40*(1-2), 73–85.

Voss, N. R., & Gerstein, M. (2010). 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic acids research*, *38*(Web Server issue), W555–W562. https://doi.org/10.1093/nar/gkq395.

Walker, J., Crowley, P., Moreman, A. D., & Barrett, J. (1993). Biochemical properties of cloned glutathione S-transferases from *Schistosoma* mansoni and *Schistosoma* japonicum. *Molecular and biochemical parasitology*, *61*(2), 255–264. https://doi.org/10.1016/0166-6851(93)90071-5.

Walker JM. The Protein Protocols Handbook. Humana Press; 2002. https://doi.org/10.1007/978-1-59745-198-7.

Wang, C. H., Hsieh, Y. H., Powers, Z. M., & Kao, C. Y. (2020). Defeating Antibiotic-Resistant Bacteria: Exploring Alternative Therapies for a Post-Antibiotic Era. *International journal of molecular sciences*, *21*(3), 1061. https://doi.org/10.3390/ijms21031061.

Wang, W., Wang, L., & Liang, Y. S. (2012). Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitology research*, *111*(5), 1871–1877. https://doi.org/10.1007/s00436-012-3151-z.

Wagner, E. K., Hewlett, M. J., Bloom, D. C., & Camerini, D. (2007). *Basic Virology*. Wiley-Blackwell, New Jersey.

Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses fromcircular dichroism spectroscopic data. *Nucleic acids research*, *32* (WebServer issue), W668–W673 https://doi.org/10.1093/nar/gkh371.

Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., &Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. *Cell*, *16*(4), 777–785. https://doi.org/10.1016/0092-8674(79)90093-x.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., &Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)*, *112*, 531–552. https://doi.org/10.1385/1-59259-584-7:531.

Wilson, R. A., Langermans, J. A., van Dam, G. J., Vervenne, R. A., Hall, S. L., Borges, W. C., Dillon, G. P., Thomas, A. W., & Coulson, P. S. (2008). Elimination of *Schistosoma* mansoni AdultWorms by Rhesus Macaques: Basis for a Therapeutic Vaccine?. *PLoS neglected tropical diseases*, *2*(9), e290. https://doi.org/10.1371/journal.pntd.0000290.

Wishart D. S. (2005). Bioinformatics in drug development and assessment. *Drug metabolism reviews*, *37*(2), 279–310. https://doi.org/10.1081/dmr-55225.

Wolmarans, C. T., & De Kock, K. N. (2009). The influence of health education on the prevalence, intensity and morbidity of *Schistosoma* haematobium infections in children over a two-year period in the Limpopo Province, South Africa. *Southern African Journal of Epidemiology and Infection*, *24*(1), 13-17. https://doi.org/10.1080/10158782.2009.11441334.

Woolf, T. B., Grossfield, A., & Tychko, M. (2000). Differences between apo and three holo forms of the intestinal fatty acid binding protein seen by molecular dynamics computer calculations. *Biophysical journal*, *78*(2), 608–625. https://doi.org/10.1016/S0006-3495(00)76621-5.

Xia X. (2017). Bioinformatics and Drug Discovery. *Current topics in medicinal chemistry*, *17*(15),1709–1726. https://doi.org/10.2174/1568026617666161116143440.

Xiao, S. H., You, J. Q., Gao, H. F., Mei, J. Y., Jiao, P. Y., Chollet, J., Tanner, M., & Utzinger, J. (2002). *Schistosoma* japonicum: effect of artemether on glutathione S-transferase and superoxide dismutase. *Experimental parasitology*, *102*(1), 38–45. https://doi.org/10.1016/s0014-4894(02)00145-5.

Xue, K., Nimerovsky, E., Tekwani Movellan, K. A., Becker, S., & Andreas, L. B. (2022). Backbone Torsion Angle

Determination Using Proton Detected Magic-Angle Spinning Nuclear Magnetic Resonance. *The journal of physical chemistry letters*, *13*(1), 18–24. https://doi.org/10.1021/acs.jpclett.1c03267.

You, H., Cai, P., Tebeje, B. M., Li, Y., & McManus, D. P. (2018). Schistosome Vaccines for Domestic Animals. Tropical medicine and infectious disease, 3(2), 68. https://doi.org/10.3390/tropicalmed3020068.

Zhan, B., Perally, S., Brophy, P. M., Xue, J., Goud, G., Liu, S., Deumic, V., de Oliveira, L. M., Bethony, J., Bottazzi, M. E., Jiang, D., Gillespie, P., Xiao, S. H., Gupta, R., Loukas, A., Ranjit, N., Lustigman, S., Oksov, Y., & Hotez, P. (2010). Molecular cloning, biochemical characterization, and partial protective immunity of the heme-binding glutathione S-transferases from the human hookworm Necator americanus. *Infection and immunity*, *78*(4), 1552–1563. https://doi.org/10.1128/IAI.00848-09.