Elucidating Conformational Dynamics and Ligand Specificity of *Schistosoma bovis* 28 kDa Glutathione Transferases: A Structural Investigation.

By

Sizamile Mbalenhle Mfeka

Submitted in fulfilment of the requirement for the degree of

Doctor of Philosophy

in

Biochemistry

in the

College of Agriculture, Engineering and Sciences at the

University of KwaZulu-Natal

Supervised by Dr Thandeka Khoza & Professor Ikechukwu Achilonu

Preface

The experimental work described in this thesis was carried out at the PSFRL at WITS University and the School of Life Science, University of KwaZulu-Natal, Pietermaritzburg from May 2021 to May 2024, under the supervision of Dr. Thandeka Khoza and Professor Ikechukwu Achilonu. The studies represent original work by the author, and none of this work has been submitted for the award of any degree or examination at any university. All authors of data and any other information have been acknowledged accordingly by reference.

Student: Sizamile Mbalenhle Mfeka (213511702)

Signature: [redacted] Date: 28 May 2025

As the candidate's supervisor, I agree to the submission of this dissertation.

Supervisor: Dr. Thandeka Khoza

Signature: [redacted] Date: 28 May 2025

Co-supervisor: Professor Ikechukwu Achilonu.

Signature: [redacted] Date: 28 May 2025

Declaration - Plagiarism

I, Sizamile Mbalenhle Mfeka, declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original research.
2. This dissertation has not been submitted for any degree or examination at any other university.
3. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a. Their words have been re-written but the general information attributed to them has been referenced.
   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the references sections.

Student Name: Sizamile Mbalenhle Mfeka

Signature: ___  ___ Date: ___28 May 2025___

Declaration - Publications

Results in Chemistry, Published: 30 April 2025

Crystal enigma: Understanding Diverse Protein Conformational Dynamics, Ligand Selectivity and Interactions in Multi-Space Group Crystals Using Computational Modelling

Mbalenhle Mfeka [a b], Olalekan Onisuru [b], Ramesh Pandian [b], Yasien Sayed [b], Thandeka Khoza [a], Ikechukwu Achilonu [b]

[a] Department of Biochemistry, School of Life Sciences, College of Agriculture, Engineering & Science, University of KwaZulu-Natal, Pietermaritzburg, 3209, South Africa

[b] Protein Structure-Function and Research Laboratory, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg 2000, South Africa

*** Corresponding authors' email:** Ikechukwu.Achilonu@wits.ac.za

Author contributions:

Mbalenhle Mfeka: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Olalekan Onisuru:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Ramesh Pandian:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Yasien Sayed:** Writing – review & editing, Project administration, Funding acquisition. **Thandeka Khoza:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Ikechukwu Achilonu:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Student: [redacted] Date: 28 May 2025

Supervisor: _____ Date: 28 May 2025

Co-supervisor: [redacted] Date: 28 May 2025

Current Protein and Peptide Science: Under Review

The Apo-Crystal Structure of a Polymorphic Form of The 28-kDa *Schistosoma bovis* Glutathione Transferase in Orthorhombic Form

Mbalenhle Mfeka [ab], Olalekan Onisuru[b], Hattie Makumbe[b], Ramesh Pandian [b], Yasien Sayed [b], Thandeka Khoza [a], and Ikechukwu Achilonu [b*]

[a] Department of Biochemistry, School of Life Sciences, College of Agriculture, Engineering & Science, University of KwaZulu-Natal, Pietermaritzburg, 3209, South Africa

[b] Protein Structure-Function and Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg 2000, South Africa

***Correspondence to Ikechukwu Achilonu (Ikechukwu.Achilonu@wits.ac.za)**

Author contributions:

Mbalenhle Mfeka: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Olalekan Onisuru:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Ramesh Pandian:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Yasien Sayed:** Writing – review & editing, Project administration, Funding acquisition. **Thandeka Khoza:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Ikechukwu Achilonu:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Student: __ [redacted] __ Date: ___28 May 2025___

Supervisor: _____ Date: __28 May 2025___

Co-supervisor: __ [redacted] _ Date: __28 May 2025___

Abstract

Schistosomiasis, caused by *Schistosoma* trematodes is the second most devastating neglected tropical disease. *Schistosoma bovis* largely affects ruminant hosts, significantly crippling the economic output of livestock farmers and food security. A promising anti-schistosomal drug target is the multifunctional detoxification enzyme *Schistosoma bovis* 28 kDa glutathione transferase (Sb28GST). While empirical structural techniques are insightful and used in drug design, they fall short in providing the dynamics of the protein of interest. Computer aided drug design (CADD) offers a time and cost-efficient workflow to screen and analyze drug candidates. This method has limitations, whose misinterpretation can hinder drug design and development. In this study we caution the careful selection of the crystal structure space groups used for CADD as isolated analyses influence the conformational states visualized and simulated as we search for Sb28GST drugs for Schistosomiasis intervention. Upon the successful expression and purification of Sb28GST, crystal conditions were manipulated to obtain the protein's crystal structure at 2.4 Å in an orthorhombic crystal system. A 500 ns molecular dynamic simulation consisting of eight stages was performed on the orthorhombic 8BHZ to compare to the monoclinic Sb28GST 8ALS. Two distinct trajectories were undertaken by the apo proteins, revealing significant differences in the dynamic behaviour of the protein systems. A library of flavonoid compounds was used to perform high-throughput virtual screening for the identification of potential Sb28GST ligands, whose outputs were filtered using standard precision and extra precision mode. The screening process showed a diverse selection of flavonoid compounds between the different space groups with strong favourable interactions implied from the low free binding. The only common ligand selected between the protein polymorphs was apigenin 7-O-(2G-Rhamnosyl) Gentiobioside however, quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside was the best performing ligand overall, with the lowest glide score of -15.6645 kcal/mol. The molecular dynamic simulations of the protein polymorphs with the selected ligands additionally highlighted different molecular interactions in which the proteins used to stabilize the proteins bond to the ligands. Concluding that the combined assessment of protein polymorphs provides more detailed insights into the conformation dynamics of the protein which would not be otherwise identified in a single space group analysis.

To test the inhibitory potency of the selected ligands a CDNB-GSH assay showed that both apigenin and quercetin significantly reduced the activity of Sb28GST. The IC_{50} of the confirmed inhibitors showed that apigenin was a more potent inhibitor (IC_{50} = 0.130 mM) than quercetin (IC_{50} = 0.120 mM). The extrinsic fluorescence studies suggest that apigenin and quercetin both bind to the hydrophobic H-site and allosteric L-site at the dimer interface. Analysis of the thermal stability of Sb28GST showed that the presence of the ligands reduces the thermal stability of the protein in a very insignificant way. This study presents the first empirical validation of computationally selected lead flavonoid compounds as inhibitors of Sb28GST for schistosomiasis intervention. From this studies' findings, future research would entail validating the lead flavonoid inhibitors through kinetic and mechanistic studies. In addition, *in vitro* and *in vivo* testing of the flavonoid inhibitors would be expanded to other Schistosome GST isoforms to guide clinical candidate development.

Acknowledgements

First and foremost, I would like to thank my loving and caring family. My fathers Siyanda Mfeka and Thabani Thwala encouraging my pursuit of change through the fulfilment of my educational goals. My late mom Jabu Mthembu, for her unwavering and unconditional support throughout my studies, for persistently and patiently believing in my dreams, even when I couldn't. My youngest brother, Bandile Mfeka, thank you for your encouragement throughout my academic pursuits and tremendous help on the long and rather bumpy road of research.

I express my sincere gratitude to the National Research Foundation, Professor Niesler Dr Khoza and Professor Achilonu for their financial assistance.

My deepest gratitude to my supervisors, Dr Thandeka Khoza and Professor Ikechukwu Achilonu for having allowed me to pursue my postgraduate studies in the Discipline of Biochemistry. Thank you for the delicate guidance throughout my studies. I have an unending gratitude for the never-ending positive attitude, for patience, and for helping me to overcome many crises throughout my studies and the completion of this dissertation

To Dr Olalekan Onisuru and all my co-authors, thank you for sharing your expertise and affording me the chance to learn and master various skills in the field of bioinformatics and molecular biology at large.

I am also appreciative of the postgraduate students in Lab 39 and the Protein Structure-Function Research Laboratory for the hours we spent in the laboratory together. You have made my working environment extremely pleasant with your stimulating discussions and lightening the workload with much needed laughter.

To my loving partner Gugulethu Ncube, and caring friends, Nontobeko Cele and Samkelisiwe Ndlela thank you is the last word I would say to express my appreciation and unbelievable support throughout this time.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

∆H˚	Change in enthalpy
∆S˚	Change in entropy
∆G˚	Change in Gibbs free energy
°C	Degrees Celsius
3D	Three dimensional
µg	Microgram
µg/mL	Microgram per millilitre
µL	Microliter
µM	Micromolar
x g	Relative centrifugal force
α	Alpha
ε	Molar extinction coefficient
λ	wavelength
γ	Gamma
γ-GCS	γ-Glutamyl cysteine synthetase
ANS	8-Anilino-1-naphthalene sulfonate
AU	Arbitrary unit
β	Beta
BSP	Bromosulfophthalein
Cα	Alpha-carbon atom of an amino acid residue
CA	California
CADD	Computer aided drug design
CDC	Centres for Disease Control and Prevention
CDNB	1-chloro-2, 4-dinitrobenzene
CLIC	Chloride intracellular channel
cm	Centimetres

C-terminal	Carboxyl terminal
DNA	Deoxyribonuecleic acid
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra-acetic acid
EMBL	European molecular biology laboratory
FT	Flow through
GPX	Glutathione peroxidase
GPU	Graphics processing unit
GR	Glutathione reductase
GS	Oxidized glutathione
GSH	Reduced glutathione
G-site	Glutathione binding site
GSSG	Oxidized glutathione disulphide
GST	Glutathione transferase
HTVS	High-throughput virtual screening
h	Hour/s
hGSTA	Human glutathione transferase alpha class
hGSTO	Human glutathione transferase omega class
hGSTPi	Human glutathione transferase pi class
H-site	Hydrophobic electrophilic binding site
IC_{50}	Half-maximal inhibitory concentration
IMAC	Immobilized metal affinity chromatography
IntraHB	Intramolecular hydrogen bonds
IPTG	Isopropyl-β-D-thio-galactoside
ITC	Isothermal titration calorimetry
K	Kelvin
Kcal/mol	Kilocalorie per mole

K_d	Dissociation constant
kDa	Kilodaltons
LigRMSD	Ligand root mean square deviation
LigRMSF	Ligand root mean square fluctuation
LigPrep	Ligand preparation
L-site	Ligandin binding site
M	Molar
MAPEG	Membrane associated proteins involved in eicosanoid and glutathione metabolism
MolSA	Molecular surface area
MD	Molecular dynamics
MDS	Molecular dynamic simulation
min	Minute/s
mg/mL	Milligram per millilitre
mL	Millilitre
mM	Millimolar
MM/GBSA	Molecular mechanics, generalized born surface area
MWM	Molecular weight marker
n	Stoichiometry
ng/µL	Nanograms per microliter
nM	Nanomolar
nm	Nanometres
nm/min	Nanometres per minute
NPT	Number of molecules/atoms in assembly Pressure and Temperature
ns	Nanoseconds
N-terminal	Amino terminal
NVT	Number of molecules/atoms in assembly Volume and Temperature

OD	Optical density
OPLS	Optimized potentials for liquid simulations
p	Pellet
PBS	Phosphate-buffered saline
ps	Picoseconds
PAGE	Polyacrylamide gel electrophoresis
PfGST	*Plasmodium falciparum* glutathione transferase
PZQ	Praziquantel
RAM	Random access memory
RFU	Relative fluorescence unit
Rg	Radius of gyration
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Rotations per minute
s	Supernatant
SASA	Solvent accessible surface area
Sb28GST	*Schistosoma bovis* 28 kda Glutathione transferase
Sj26GST	*Schistosoma japonicum* 26 kda Glutathione transferase
SDS	Sodium dodecyl sulphate
SP	Standard precision
spp	Species
TCEP	tris (2-carboxyethyl) phosphine-hydrogen chloride
TEMED	N',N',N',- tetramethylethylenediamine
TGR	Thioredoxin glutathine reductase
T_H2	Specific T helper cells
Tm	Melting temperature
Tr1	Type 1 regulatory T cells

Tris	2-amino-2-(hydroxymethyl)-1.3-propandiol
UK	United kingdom
USA	United States of America
UV	Ultraviolet
UV/Vis	Ultraviolet-visible
v/v	Volume per volume
WHO	World Health Organization
w/v	Weight per volume
XP	Extra precision
YT	Yeast and tryptone

Amino acids are named according to their IUPAC three letter codes

Chapter 1

Introduction and Literature Review

1.1. Global Burden of Schistosomiasis

The global human population increases at a rate of approximately 1% each year. This growth is accompanied with a proportionately increased need for natural resources whose turnover remains unchanged despite the growing demand (Strydom et al., 2023). Meat consumption in particular is projected to have increased by 14% from the average consumption rate recorded from 2018-2020 (Use, 2019). To combat this imbalance, beef production is expected to grow by 5.9% to sustain this ever-growing market (Outlook, 2021).

The top beef and veal producers globally turned over upwards of 58 184 000 metric tons of beef carcass weight in 2022 constituting 74% of the global supply of beef products (Strydom et al., 2023). With commodities of this magnitude, any minor changes experienced in production efficiency would ultimately leave the industry vulnerable to massive economic losses (Mustafa et al., 2024). Unfortunately, beef production or more specifically, cattle operations do in fact face many challenges. One being parasite-borne diseases which directly affect cattle and consequently, farming efficiency (Vercruysse et al., 2018). The disruptions brought through animal health issues within this industry is a cause of great concern. These disruptions have the potential to reduce the world production of food animals by upwards of 20%, thus further contributing to the deficient surplus food supply (Charlier et al., 2015). Helminth-associated diseases in particular have been millennia-old scourges on cattle and the livelihood of farmers. Their impact has been catastrophic largely due to their undetectable insidious nature. They have been able to achieve this level of stealth better than their bacterial and viral counterparts due to chronic diseases they cause (Vercruysse et al., 2018). These diseases heavily impact animal health by way of nutrient absorption, growth, mortality rates, carcass weight, fertility, and milk yield. This negatively impacts farming efficiency and ultimately economic profitability (Charlier et al., 2020).

1.1.1. Schistosomiasis as a Burden of Disease

One of the major helminth diseases that is a global burden not only to wild and domestic animals, but humans as well is schistosomiasis (Utzinger et al., 2009). Schistosomiasis is considered the most important water-based disease from a global health perspective and is listed as a neglected tropical disease by the World Health Organization (Organization, Steinmann et al., 2006). It is estimated by the WHO that schistosomiasis and soil helminths account for over 40% of the global disease burden caused by all tropical diseases, with the exclusion of malaria (Adenowo et al., 2015). Schistosomiasis in particular is ranked the third most devastating tropical disease after malaria and intestinal helminthiases (Olveda et al., 2013).

As depicted in Figure 1.1, schistosomiasis is predominantly found in developing countries in Africa, the Middle East, Asia, South America and the Caribbean (Utzinger et al., 2009). In Africa, Asia and South America specifically, there are over 250 million people infected with schistosomiasis with approximately 732 million more exposed to possible infection (Adenowo et al., 2015). It is estimated that 90% of the 250 million persons infected are concentrated in Sub-Saharan Africa (Van der Werf et al., 2003). In South Africa alone studies showed that 4.5 million schistosomiasis cases arise each year with over 25 million people exposed to possible infection (Wolmarans et al., 2009). Additionally, the burden of disease attributable to schistosomiasis species ranges between 24 - 29 million disability adjusted life years (King et al., 2005).

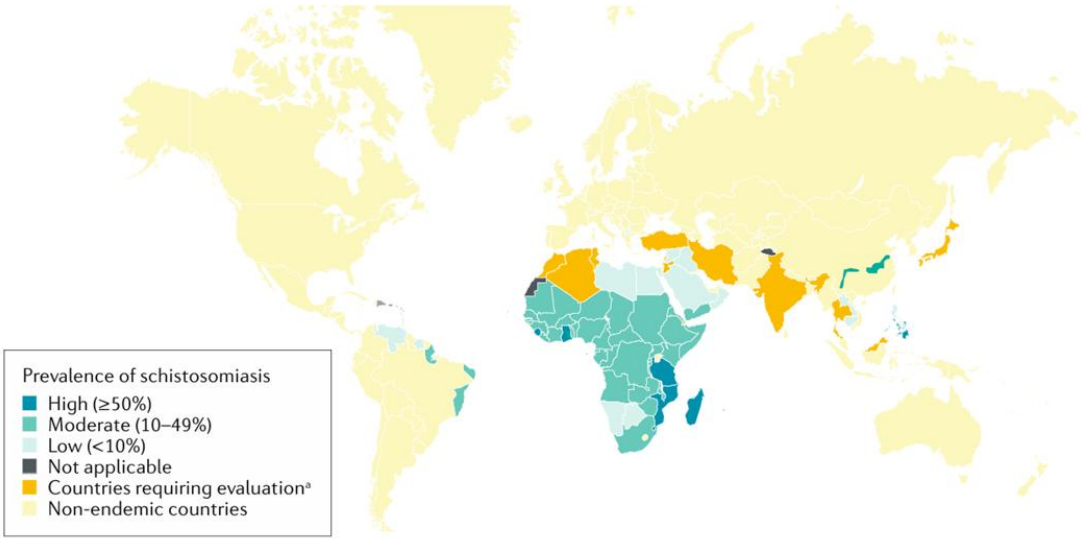


Figure 1. 1: Global distribution of schistosomiasis as adapted from (McManus et al., 2018)

Despite the risk that schistosomiasis poses to public health, its resolution and mitigation has largely been neglected due to several factors: (Adenowo et al.,2015) (i) Schistosomiasis cases are typically concentrated in the worlds most disadvantaged tropical and sub-tropical areas whose rural and urban communities are largely marginalized. The plight of these communities is often overlooked and excluded from national political agendas (Utzinger et al., 2009, Van der Werf et al., 2003). (ii) Schistosomiasis is not as salient as the “Big Three” namely Malaria, HIV/AIDS and tuberculosis, and is once again overlooked by public health officials, researchers and funders for inclusion in national strategic plans and global health initiatives (Adenowo et al., 2015, King and Dangerfield-Cha, 2008). (iii) Due to the insufficient resources allocated to schistosomiasis research, the global health burden as expressed by the disability-adjusted life years lost is reported to be severely underestimated (Olveda et al., 2013, King and Dangerfield-Cha, 2008).

The world’s largest producers of livestock and beef products are the United States of America (Jonsson et al.), Brazil, China and the European Union (Strydom et al., 2023). In areas endemic to schistosomiasis infection, Ethiopia leads in the continent of Africa with over 70 million heads of cattle produced each year (Molla et al., 2022). This commodity, like in many parts of Sub Saharan Africa and Asia provides a source of protein, manure for farmland and household energy, export commodities and food security in times of crop failure, all while allowing for a means of wealth accumulation (Strydom et al., 2023). The livestock sector in Sub-Saharan Africa contributes an average of 35% to the agricultural gross domestic product (GDP) of their nations (Molla et al., 2022). The total GDP contributed in schistosomiasis endemic areas such as Ethiopia, Tanzania and Somalia are 19%, 7.4% and 40% respectively with up to 50% contributing to national foreign exchange (Demlew and Tessma, 2020, Molla et al., 2022). Unfortunately, this sector is not fully exploited and additionally suffers massive economic losses each year due to inadequately treated diseases.

In addition to the economic impact that schistosomiasis poses to afflicted areas, it also negatively affects the health and well fare of farmed cattle. Cattle suffering with schistosomiasis do not meet the World Organization for Animal Health’s status of “good animal welfare” which is defined as “healthy, comfortable, well nourished, safe, is not suffering from unpleasant states such as pain, fear and distress, and is able to express behaviours that are important for its physical and mental state. Good animal

welfare requires disease prevention and appropriate veterinary care, shelter, management and nutrition" (Escobar et al., 2018). This shortfall is due to the interferences that schistosomiasis causes in cattle feeding, rumination, resting and standing behaviours, negatively impacting cattle welfare (King and Dangerfield-Cha, 2008). The cyclic nature of schistosomiasis-causing organisms negatively contributes to cattle welfare through the transmission of disease-causing organism.

The additional impact that schistosomiasis has on global warming is another devastating burden that the disease causes. Reports have shown that cattle with schistosomiasis and gastrointestinal diseases produce higher greenhouse gas emissions than their uninfected counterparts (Jonsson et al., 2022). This excessive production acts in direct opposition of many government's emission reduction targets. Thus, compelling cattle producers to reduce their carbon footprint through growing societal and political pressures (Capper, 2023).

The enormity of the global impact that schistosomiasis has on finance, livelihood, animal welfare and greenhouse gas emissions is more than enough to justify urgent control interventions of this disease. That however, requires the examination of the micro-organism responsible for the global disease burden of schistosomiasis.

1.1.2. Biology of *Schistosoma bovis*

Schistosomiasis is caused by parasitic dioecious, dimorphic trematodes of the genus *Schistosoma*, which belong to the Schistosomatidae family. These organisms are classified as belonging to the Animalia kingdom, Platyhelminthes phylum under the class and sub-class Trematoda and Digenea respectively (Angora et al., 2020). There is a total of 19 described *Schistosoma* species most of which can affect many organisms from humans to wild and domestic animals (Demlew and Tessma, 2020).

The three main species responsible for human schistosomiasis are *Schistosoma japonicum, S. mansoni* and *S. haematobium (Tsuji, 2020)*. *Schistosoma haematobium* specifically is most prevalent in Africa and responsible for over two thirds of all schistosomiasis cases within that region (Molla et al., 2022). The threat that *S. haematobium* poses to basic health also extends to animals, with the zoonotic nature of the parasite allowing five different *S. haematobium* species to cause schistosomiasis in livestock (Pennance et al., 2021). While there are a total of ten *Schistosoma* species known to naturally infect cattle, six are recognized to have

relevant veterinary significance; *S bovis, S. indicum, S. japonicum, S. matthei, S. intercalatum, S. nasale* and *S. rodhoni* (Molla et al., 2022). Of the mentioned, *S. bovis* is the most prevalent, owing to it's pathogenicity and being the most geographically distributed throughout *Schistosoma* endemic areas. Additionally, *S. haematobium* and *S. bovis* are closely related phylogenetically. This has slowly allowed interspecies mating, giving rise to introgression and hybridization between the two species (Angora et al., 2020). In recent years, there has been an increasing number of these hybridization cases in humans and ruminants, with both infected children and snail intermediate hosts shedding *S. haematobium* - *S. bovis* hybrids (Léger et al., 2020). Allowed to continue, this evolutionary zoonotic adaptation will threaten control interventions, disease dynamics and transmission. A controlled study by (Leger and Webster, 2017) has already shown that these hybrids exhibit increased virulence, expanded snail host spectrum, maturation and egg production.

1.1.2.1. Life Cycle of *Schistosoma bovis*

Schistosomiasis causing parasites have a complex life cycle which involves freshwater snails as intermediate hosts and vertebrates as definitive hosts (Escobar et al., 2018). Interestingly, the schistosomiasis symptoms presented in definitive hosts differ in accordance to the snail intermediate hosts. *Biomphalaria* snail transmitted *Schistosomes* results in hepatic and intestinal symptoms while *Oncomelania* snail transmitted *Schistosomes* cause intestinal and liver targeted symptoms (Léger et al., 2020). Upon *schistosoma* infection, the overall symptoms of infection are characterized with stomach pain, low-grade fever, coughing and fatigue. If untreated, these symptoms escalate to profuse watery diarrhoea, emaciation, blood poisoning, a persistent rash and abdominal cramps (Pérez-Sánchez et al., 2006). These symptoms heavily disrupt animal and ultimately food product production. Once again affecting economic profitability, further crippling the farmers that make a living from livestock while leaving the global market deficient of meat products.

Much like all *Schistosoma* spp, *S. bovis* is dependent on two hosts to complete their life cycle. The snail hosts required for *S. bovis* propagation belong to the genus *Bulinus*, *Indoplanorbis* and *Planorbis*. In Africa however, *Bulinus afrocanus, B. truncates* and *B. abysinicus* play the role of *S. bovis* intermediate host (Demlew and Tessma, 2020). Upon transdermal infection of the definitive host through the skin of the leg or perioral skin, the cercariae remain in the skin for approximately three days.

At this stage the male and female cercariae co-populate within the lumen of the definitive hosts veins. During this process the cercariae lose their tail and glycocalyx and transform into *Schistosoma* larvae. Thereafter the adult females migrate against the venous blood flow into small venules where approximately 200 – 2000 eggs are deposited per day. During migration, the ova tend to pass through the blood vessel walls to adjacent tissues such as the liver, heart and the portal vein of definitive host but 40-80% of the ova laid are ultimately retained in the mesenteric vessels. At this point, pairing and sexual maturation of the ova occurs. While a large amount of the ova migrates to maturation, some of the ova are expelled through to the definitive hosts faeces and urine. This expulsion is discharged in bodies of water, resulting in the ova hatching to release miracidia. The miracidia then proceed to invade the intermediate hosts which are freshwater snails where they continue to develop into cercariae. Once the cercariae fully mature, they are released by the snail hosts back into freshwater bodies. The cercariae can once again penetrate their definitive hosts through the skin or mucous membranes. Thereafter, the cycle restarts with the cercariae maturing into schistosomula and recirculating through the definitive hosts lymphatic system and blood vessels to live in the mesenteric veins for years, causing autoinfection as shown in Figure 1.2 (Adenowo et al., 2015, de la Torre-Escudero et al., 2017, Olveda et al., 2013).

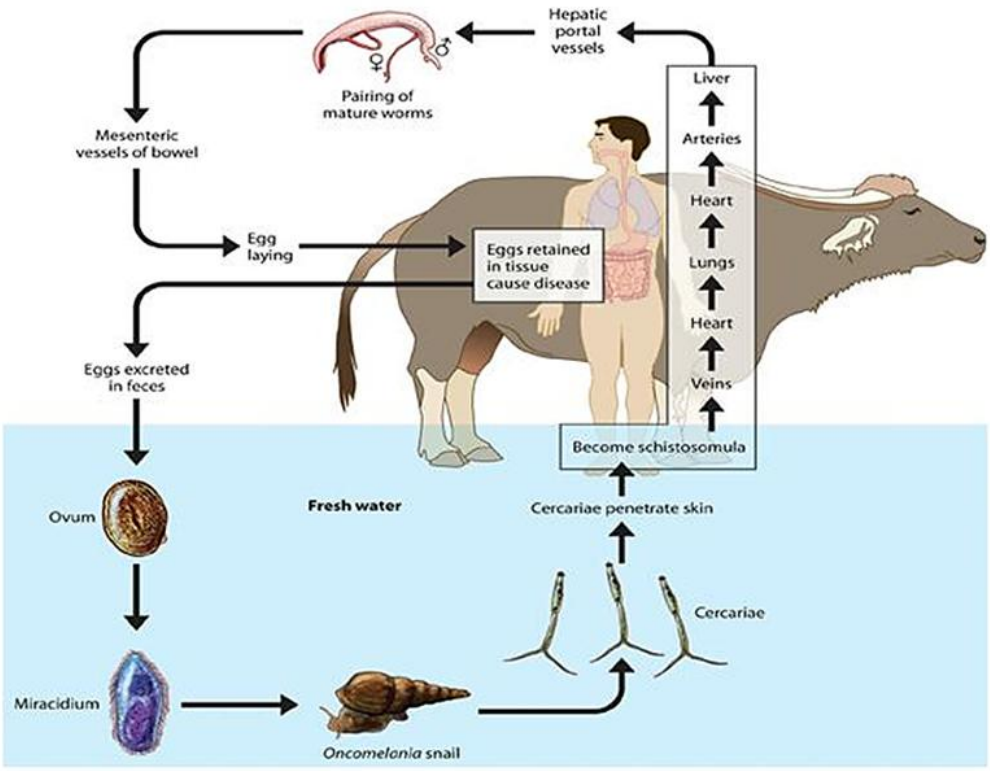


Figure 1. 2: Life cycle of *Schistosoma bovis* (Demlew and Tessma, 2020)

1.1.3. Schistosomiasis Diagnosis

Schistosomiasis infections are typically diagnosed using a range of techniques, such as clinical assessments, serology, laboratory methods, molecular assays and antigen detection techniques (Payne et al., 2023). Serology tests detect the presence of anti-*Schistosoma* antibodies in blood samples. However, the results obtained from these tests do not differentiate between past or active infections, as antibodies typically develop 6-8 weeks after initial infection. Serology techniques are, unfortunately, limited by their inability to distinguish between *Schistosoma* species (Hoekstra et al., 2021, Weerakoon et al., 2015). This is resolved by using positive serology results in combination with clinical assessments, where clinicians further test urine for haematuria or faecal matter for calprotectin. The gold standard for schistosomiasis diagnostic techniques is laboratory microscopy-based methods, where *Schistosoma* eggs are visualised directly from stool and urine samples (Utzinger et al., 2015). These visualisation techniques, such as the Kato-Katz assay and the Formalin Ethyl acetate concentration technique allow for the visualisation of the egg size and morphology which is vastly different from one *Schistosoma* species to another (Lamberton et al., 2014). The use of these techniques for instance, allows diagnosticians to clearly

distinguish between *S. haematobium* eggs possessing terminal spines from *S. mansoni* eggs with lateral spines as shown in Figure 1.3. This then allows for a full report on the extent and type of *Schistosoma* infection present (Payne et al., 2023).

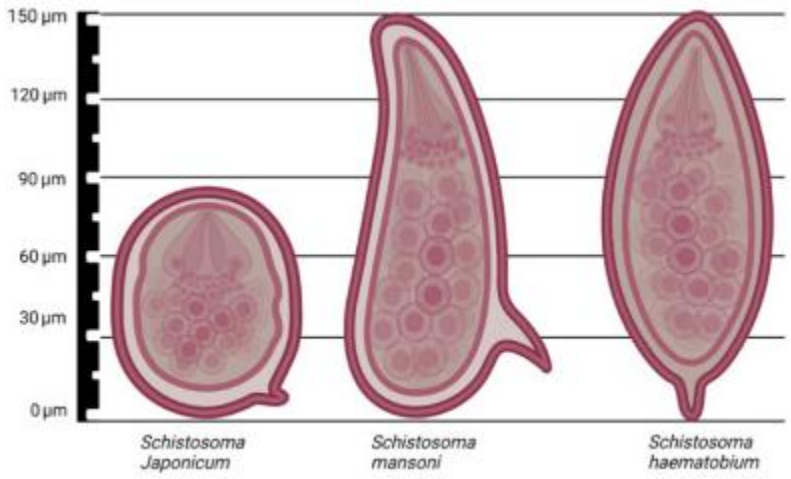


Figure 1. 3: Morphology of *Schistosome* eggs seen using light microscopy (Payne et al., 2023)

1.1.4. Pathogenesis of *Schistosoma bovis*

The course of schistosomiasis infection often occurs in three phases: migratory, acute and chronic (Molla et al., 2022). The infection is initiated with the migratory phase. While this phase may be asymptomatic, in other cattle, slight dermatitis can be observed at the point of infection. Additionally, the movement of the ova through the bovine vessels can cause mechanical damage, pulmonary lesions and pneumonitis. The ova released into the local tissue during the migration phase can also release toxins and enzymes, which induce a T_H2 mediated immune response. The inflammatory and granulomatous reaction that occurs to destruct the eggs can cause fibrosis and scar the affected tissue (Demlew and Tessma, 2020, King and Dangerfield-Cha, 2008). Following the migratory phase is the acute phase of the disease where more extensive haemorrhagic lesions are apparent in the mucosa of the intestine. This progresses to the physical deterioration of the whole intestine with it adopting a grey appearance with apparent thickening and swelling. The unmitigated development of this phase leads to its advancement into the chronic phase, where organ damage and dysfunction are observed (Molla et al., 2022).

Despite the hosts intense anti-*Schistosome* immune response to eliminate parasitic invasion, the adult *S. bovis* worms successfully manage to live in their bovine hosts

for years. This monumentally impairs global initiatives placed to combat the ever-growing socio-economic demand for efficient and sustainable production of livestock to meet the world's protein needs (de la Torre-Escudero et al., 2017).

1.1.5. Current Anti-*Schistosome* Treatments

Several anti-schistosomal treatments such as oxamiquin, niridazole, hyacanthone and metrifonate have been deployed in the past to combat *Schistosoma* infections. They unfortunately, were not as effective as intended due to the toxicity accumulation they produced before treatment courses could be completed (Olveda et al., 2013). Their use contaminated meat products, resulting in farmers failing to meet food demands and additionally increasing chemical pollutant risk in the environment (Vercruysse et al., 2018). These treatments were replaced with a pyrazinoisoquinolone derivative in 1979 called Praziquantel (PZQ) shown in Figure 1.4, which is now the primary chemotherapeutic agent for schistosomiasis (Léger et al., 2020). While the exact mechanism of the drug is not fully understood despite being in use for decades, studies have shown that the indirect molecular target of PZQ is the *Schistosome* calcium ion channel. It is hypothesised that PZQ attaches onto the *Schistosome* membrane where it induces the rapid intake of calcium ions, resulting in the complete destabilization of calcium ion homeostasis within the parasite. This results in the destabilization of the parasites membrane, causing its tegumental vacuoles to contract convulsively. This ultimately forces the *Schistosomes* latched onto the hosts venous walls to detach and die (Adenowo et al., 2015). The tegumental disruption critically impairs the parasite's ability to absorb nutrients, excrete waste and simultaneously exposes *Schistosome* surface antigens, increasing susceptibility to the host's immune responses (Vale et al., 2017). Praziquantel's known interaction with voltage-gated calcium channels is considered another contributor to its deregulatory effect. This is due to PZQs agonist role to the *Schistosome* transient receptor potential melastatin channel (Park et al., 2019). Other PZQ molecular interactions have been speculated to affect the myosin regulatory light chain and glutathione *S*-transferase, directly modulating *Schistosomes*' muscle contraction and detoxification pathways (You et al., 2021). Praziquantel additionally aids the hosts' immune response by stimulating T cell differentiation and reducing inflammation, allowing the host to better clear parasitic infection (Doenhoff et al., 2008). Recent studies have also shown that PZQ has antifibrotic effects, including the downregulation of fibrotic biomarkers and

proinflammatory cytokines, which could alleviate schistosomiasis-associated tissue fibrosis (Vale et al., 2020).

While administering PZQ causes a range of interactions as shown in Figure 1.5 and Table 1.1 leading to decreased *Schistosome* related morbidities, high re-infection and transmission rates are typically maintained. It has been reported that PZQ administration results in poor cure rates overall and treatment failure in some cases (Utzinger et al., 2009). This may be due to PZQ's inability to target all stages of the parasites life cycle, as its efficacy is only efficient in adult *Schistosome* worms (Olveda et al., 2013). The selective efficacy of the drug results in re-administration of PZQ which raises concerns of the emergence of drug resistance. Artemisinin and its analogues artesunate and artemether are anti-malarial drugs which have been reported to have anti-schistosomal properties. Their efficacy is particularly successful against juvenile *Schistosome* larvae (Adenowo et al., 2015). While they could be administered alongside PZQ, their use is also frowned upon in malaria-endemic areas due to the fear of *Plasmodium* drug resistance (King and Dangerfield-Cha, 2008). The threat to drug resistance that the continued and extensive use of current anti-schistosomal drugs poses, coupled with the partial control that the drugs provide, necessitates research into new anti-schistosomal drugs/therapies.

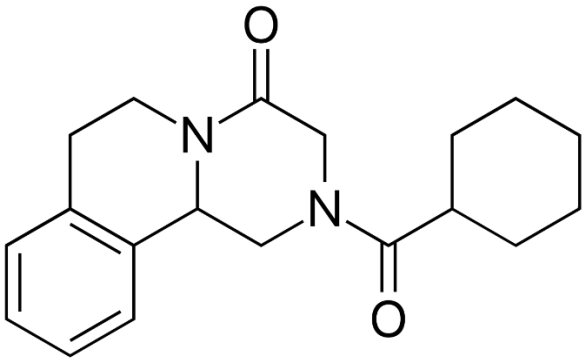


Figure 1. 4: 2D structure of Praziquantel

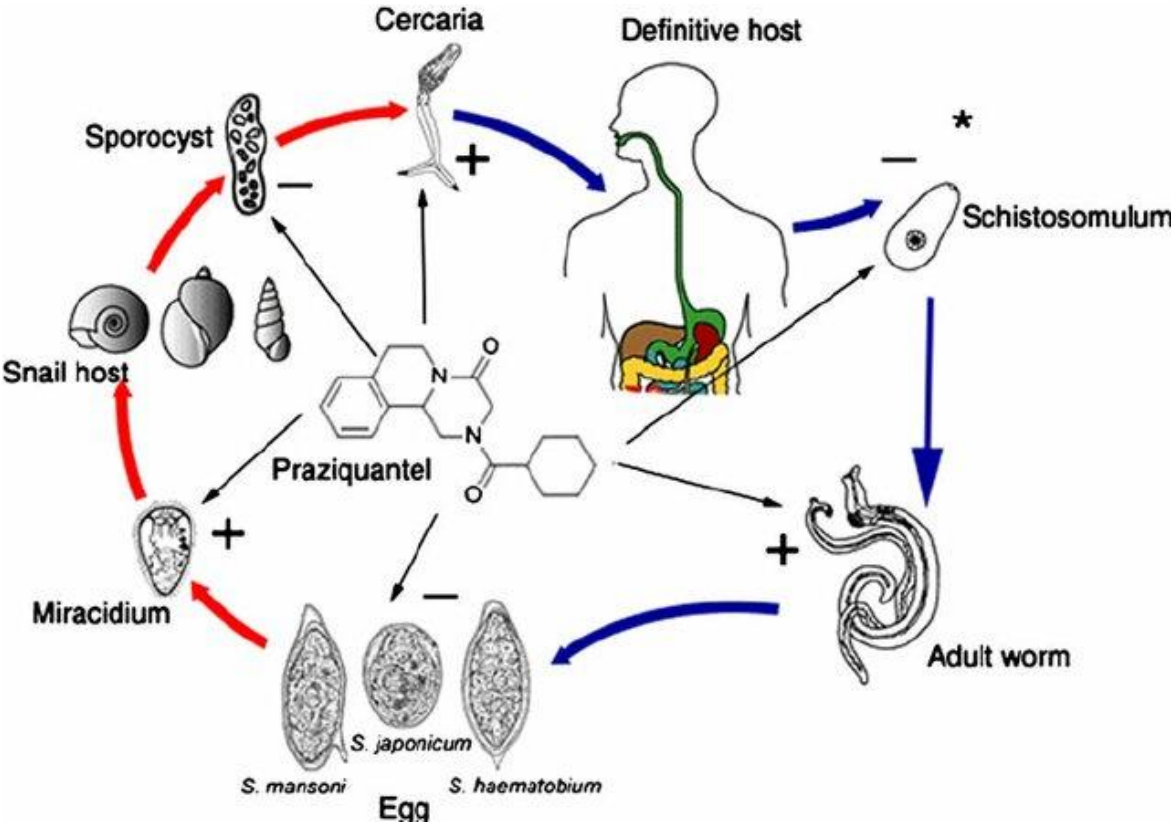


Figure 1. 5: The effect of Praziquantel at different stages of the *Schistosome* life cycle (Wu et al., 2011)

Table 1. 1: Range of mechanisms disrupted by Praziquantel interaction

Mechanism	Description	References
Calcium Ion Influx	Disrupts calcium homeostasis, causing paralysis and tegumental damage	(Babes et al., 2017, Greenberg, 2007, Thomas and Timson, 2020)
Tegumental Damage	Damages the outer surface, exposing antigens to the immune system	(Pearson et al., 2021, Xiao et al., 2008)
Interaction with Molecules	Interacts with myosin light chain, GST	(Babes et al., 2017, Nogueira et al., 2022)
Immune Modulation	Promotes Tr1 cell differentiation, reduces inflammation	(Nogueira et al., 2022)
Resistance Concerns	Reduced susceptibility observed, potential for resistance	(Doenhoff et al., 2008, Wang et al., 2012)
Antifibrotic Effects	Reduces fibrosis by decreasing fibrotic biomarkers and cytokines	(Niu et al., 2022, Pereira et al., 2021)

1.2. Redox Metabolism in Parasites and Anti-Oxidative Stress Proteins

Many parasites have adapted distinctive thiol-based redox mechanisms to survive their hosts oxidative immune response (Flohé, 2012, Kehr et al., 2010). Parasites such as *Trypanosoma* and *Leishmania* for instance, have unique trypanothione-centred thiol-based redox mechanisms which are essential for detoxification, drug resistance, and maintaining a redox balance within the parasites (Ariyanayagam et al., 2001). The maintenance of this redox state is crucial as it ensures the prevention of cellular damage caused by reactive oxygen species, thus allowing the parasites to survive. A survival adaptation that many parasites possess to counteract oxidative stress is the deployment of antioxidant systems, often specific for their mammalian hosts (Flohé, 2012, Müller et al., 2003). *Plasmodium falciparum* for instance, possesses a sophisticated redox system which includes peroxiredoxins, thioredoxins, and glutathione-dependent pathways (Rahbari et al., 2015).

Much like many other parasites, *Schistosomes* defence mechanisms have evolved to be able to consistently circumvent internal and external stresses such as UV radiation, drugs, free radicals as well as the host's immune response (Angelucci et al., 2005). Interestingly, *Schistosomes* manage to maintain their level of parasitic dominance on their hosts despite the absence of vital phase 1 detoxification enzymes. In their place, *Schistosomes* possess glutathione transferases (GST's) as their sole defence mechanism against endogenously produced toxins and their host's immune responses (Akumadu et al., 2020).

Many organisms under the Animalia kingdom from snails to humans use oxidative stress for the neutralisation of xenobiotic foreign compounds and intracellular parasites. However, parasitic protozoa make use of the antioxidant system to evade oxidative stress and consequently, oxidative killing carried out by the host's immune effector cells (Torres-Rivera and Landa, 2008). This highly sophisticated antioxidant system uses xenobiotic-metabolising enzymes to facilitate the metabolism, detoxification and expulsion of toxic compounds in three major steps (Angelucci et al., 2005).

The first step in this antioxidative system utilises phase I reactions which typically convert lipophilic xenobiotic compounds into less toxic and more hydrophilic metabolites. This is carried out by phase I enzymes such as aldo-keto reductase,

carboxylesterases and cytochrome P450 (Findlay et al., 2006). Through hydrolysis, reduction or oxidation reactions, these enzymes modify the functional groups of their targeted xenobiotic compounds. This then provides sites for phase II conjugation reactions to occur (Croom and science, 2012). The next step is carried out by phase II enzymes such as GST, glutathione peroxidase, N-acetyltransferase and sulfotransferase. These enzymes interact with phase I metabolites to increase their hydrophilicity through conjugation reactions. Finally, the conjugated metabolites are bound to phase III enzymes to be transported out of the cells using the antiporter system (Findlay et al., 2006). The genome analysis of *S. bovis* interestingly shows the absence of vital phase I detoxification enzymes. In their place *S. bovis* expresses the phase II detoxification enzyme GST as its primary defence mechanism against oxidative stress and host immune responses, allowing it to continuously strain its hosts (Makumbe et al., 2024). This highlights the crucial role that the *S. bovis* GST plays for the survival and pathogenicity of the parasite.

1.2.1. Role of Glutathione in Parasite Redox Metabolism

Glutathione (GSH) plays a critical and multifaceted role in parasite redox metabolism, underpinning their survival, pathogenicity, and resistance to environmental stress (Figure 1.6). Central to maintaining cellular redox homeostasis, GSH provides reducing equivalents to essential enzymes and protects parasites against oxidative stress, which is particularly important for organisms like *Plasmodium falciparum* that lack catalase and glutathione peroxidase, relying heavily on the thioredoxin system for redox control (Jortzik and Becker, 2012, Müller et al., 2003, Sies and medicine, 1999). Through its antioxidant properties, GSH scavenges free radicals and detoxifies reactive oxygen species (ROS) generated from the parasites' intense metabolic activities (Müller et al., 2003, Franco et al., 2009, Meister and Anderson, 1983). Beyond its antioxidant role, GSH contributes significantly to drug resistance mechanisms; for example, it modulates parasite sensitivity to antimalarials like chloroquine and artemisinin, and inhibition of GSH biosynthesis can sensitize parasites to antifolate drugs, offering promising therapeutic strategies (Becker et al., 2004, Müller and Kappes, 2007). Glutathione also detoxifies antimalarial compounds by conjugation via glutathione transferase, further aiding parasite survival (Müller, 2004). Metabolically, GSH supports rapid parasite growth by donating electrons for deoxyribonucleotide synthesis, detoxifies toxic heme derived from haemoglobin digestion, and functions as a coenzyme in the glyoxalase system to neutralize methylglyoxal, a harmful byproduct of glycolysis (Müller, 2004).

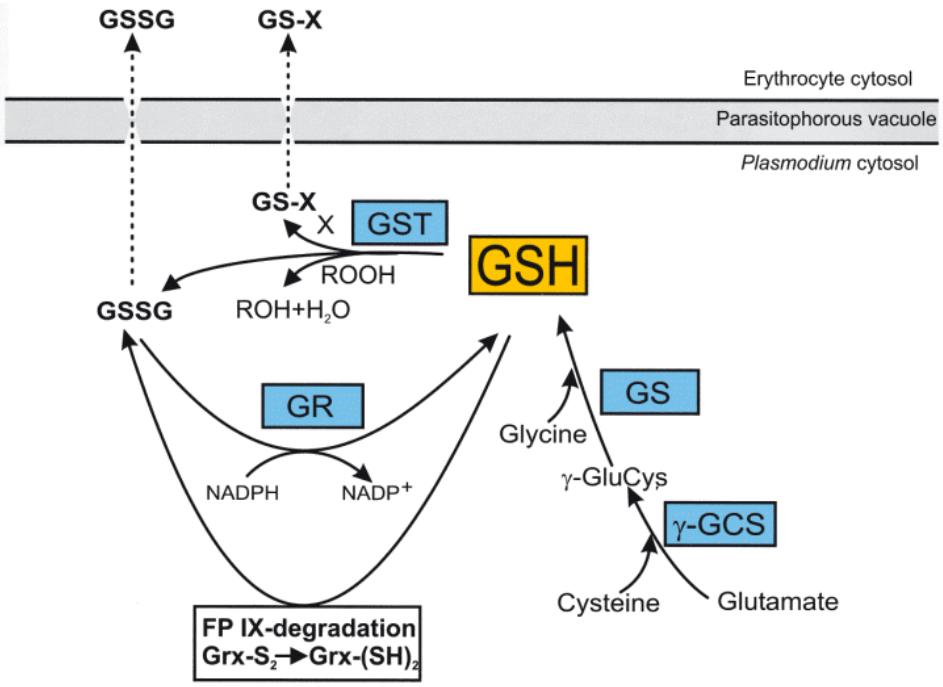


Figure 1. 6: Parasitic control of intracellular glutathione levels (Müller, 2004)

The enzymatic network maintaining GSH functionality is equally vital to parasites. Glutathione reductase (GR) reduces oxidized glutathione disulfide (GSSG) back to GSH, preserving redox balance, and has been shown to be critical for the mosquito oocyst stage of *Plasmodium* (Becker et al., 2003). Some parasites also utilize thioredoxin-glutathione reductase (TGR), a multifunctional enzyme supplying reducing equivalents to both thioredoxin and glutathione systems, compensating for the absence of conventional reductases (Müller, 2004, Sies and medicine, 1999). These pathways present attractive therapeutic opportunities: targeting enzymes involved in GSH metabolism, such as γ-glutamyl cysteine synthetase (γ-GCS) and GR, holds promise for novel antiparasitic strategies (Ginsburg and Golenser, 2003, Müller and Kappes, 2007). Notably, methylene blue, an inhibitor of *P. falciparum* GR, has shown synergistic effects with chloroquine, highlighting the feasibility of targeting GSH metabolism in drug development (Müller, 2004). Overall, glutathione is indispensable to parasite redox metabolism, influencing survival, growth, and drug resistance and a deeper understanding of its roles and mechanisms opens exciting avenues for targeted therapeutic interventions against parasitic diseases.

1.2.2. Glutathione Transferases

Glutathione transferases [EC 2.5.2.18] are a family of multifunctional enzymes responsible for phase II detoxification of xenobiotic compounds. These enzymes primarily catalyse the conjugation of the tripeptide glutathione (GSH, γ-glutamyl–cysteinyl–glycine) to the electrophilic centres of various xenobiotic compounds (Sheehan et al., 2001). These endogenous and exogenous compounds include, but are not limited to; alkyl or aryl halides, activated alkenes, epoxides, olefins, sulphate esters and quinones (Torres-Rivera and Landa, 2008). The resulting GSH-conjugates would then be subjected to bioactivation through other metabolic pathways or excreted in various ways. The role of GSTs is not limited to detoxification as their conjugation to electrophilic compounds aids in the following: the catabolism of aromatic amino acids, transportation of organic peroxides, the biosynthesis of prostaglandin and leukotriene, isomerization, the modulation and development of ion channels and the regulation of cell signalling transduction pathways which are responsible for cell survival and apoptosis through protein-protein interactions (Oakley, 2005).

1.2.2.1. Glutathione Transferase Superfamily

Glutathione transferases are divided into three sub-families based on their cellular locations. These GST sub-families are cytosolic GSTs, mitochondrial GSTs and microsomal GSTs, otherwise known as membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG) (Sheehan et al., 2001). Within these sub-families, the enzymes are further divided into different classes. Glutathione transferases are allocated into their classes per their amino acid sequence identities. There is no specific criterion for amino acid sequence identity established for GST classification into their respective classes (Torres-Rivera and Landa, 2008). However, it is widely accepted that GSTs share more than 60% sequence identity within a class. This is aided by the observation that GST sequence identity within classes is typically high, at around 70%, while the amino acid sequence identity of GSTs between classes drops to as low as 10% (Oakley, 2005). The enzymes grouped within each class share the same immunological cross reactivity, electrophilic substrate specificity and inhibitor sensitivity.

1.2.2.2. Glutathione Transferase Structure

Glutathione transferases are composed of 23-28 kDa monomers comprised of ~ 220 amino acids. These monomers adopt the same tertiary and quaternary dimeric features (Sheehan et al., 2001). Glutathione transferase dimers may be comprised of monomers from the same class (homodimer) as depicted in Figure 1.7, or different classes (heterodimer). A monomer of GST is composed of two domains, namely the N-terminal and C-terminal domain (Oakley, 2005). The topology of the N-terminal domain is highly conserved across all GST classes. This conserved domain resembles a thioredoxin like fold, which consists of four β-sheets and three α-helices arranged in a βαβ–ββα motif linked by an α-helix. A short linker sequence connects the N-terminal domain to the C-terminal domain (Johnson et al., 2003). The topology of the C-terminal domain is vastly different from GST class to class. This domain is exclusively α-helical, with the number of helices ranging from 4-7 depending on the GST class (Sheehan et al., 2001).

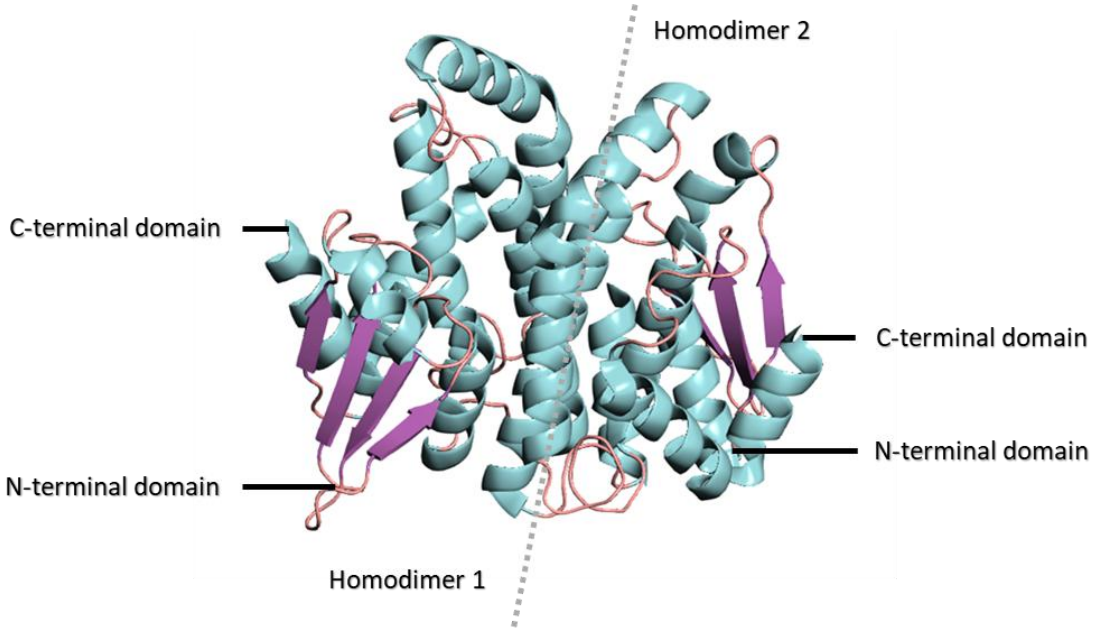


Figure 1. 7: Homodimer of *S. bovis* 28 kDa GST PDB: 8ALS showing secondary structure elements generated using Pymol (Makumbe et al., 2024)

1.2.2.3. Glutathione Transferase Binding Sites

Each GST monomer consists of two independent catalytic binding sites. The N-terminal domain houses the GSH binding site (G-site), while the C-terminal domain houses the hydrophobic substrate binding site (H-site). The conserved nature of the N-terminal domain is owed to the high specificity of the G-site (Torres-Rivera and

Landa, 2008). In this region, the GSH sulphur is activated for nucleophilic attack. The tripeptide is anchored in position through a network of hydrogen bonds, facilitating electrostatic interactions (Angelucci et al., 2005). The vital hydroxyl groups of tyrosine or serine, typically located in the GSH active site act as hydrogen bond donors to the GSH thiol group (Croom and science, 2012). This leads to the formation and stabilisation of a highly reactive thiolate anion, which is the target for nucleophilic attack of an electrophilic substrate. The H-site is considerably less conserved in amino acid sequence and topology between GST classes. The diversity observed in the C-terminal domain is largely responsible for the vast number of electrophilic substrate specificities that each GST possess (Armstrong and Biology, 1993). The H-site interacts directly with the hydrophobic moieties of the electrophilic substrate, orienting the substrate for nucleophilic attack by the GSH thiol group. Although GST monomers possess these binding sites, catalytic functionality is only achieved through dimerization of two GST monomers, as one subunit is dependent on the other for secondary structural elements (Angelucci et al., 2005). The dimerization of GST monomers produces an additional non-substrate binding site in the dimer interface known as the allosteric binding site (L-site). The L-site binds small non-substrate hydrophobic molecules which can induce conformational changes to the G-site and thus modulate its affinity for GSH, ultimately affecting the activity and dynamics of the enzyme (Sheehan et al., 2001).

1.2.2.4. Structural Variations Between Human and Parasite GSTs

Analysis of GST structures belonging to parasites and those of humans shows a wide range of differences which can be exploited for therapy intervention (Hiller et al., 2006). While all GSTs possess a thioredoxin like fold, there are multiple variable regions which are both constructed and fold differently, revealing unique functionalities between human and parasitic GSTs (Torres-Rivera and Landa, 2008). The H-site is where most variability exists between GST species. In most GSTs, the loop joining β1 and α1, the C-terminal part of the helix α4 together with the residues after α8 are the constituents that form the H-site (Hiller et al., 2006). The H-site is generally shielded by the C-terminal region from surrounding solvents in various ways depending on the GST class (Frova, 2006). The human alpha (α) class uses its large α-9 present at the C-terminus while Mu (μ) and Pi (π) classes possess wall-like structures. The μ class

has approximately 10 additional amino acid residues located between β2 and α1 called the μ loop, which is also said to assist in reducing solvent accessibility (Board and Menon, 2013, Mannervik et al., 2005). On the other hand, parasitic GSTs such as the *Plasmodium falciparum* GST (Pf-GST1) also have a μ loop which is typically truncated, containing only 5 amino acids and thus cannot form a wall. This results in an expanded hydrophobic substrate active site that accommodates larger substrates, which is a property that can be specifically targeted for selective inhibitor design in parasites (Torres-Rivera and Landa, 2008).This suggests that the H-site of PfGST allows for the binding of a broader range of substrates that include amphiphilic compounds (Sheehan et al., 2001). Similarly, *Schistosoma* and mammalian GSTs are vastly different, sharing an amino acid sequence identity of approximately 24%. This low sequence identity is indicative of the fact that the different GSTs will have dissimilar immunological cross-reactivity and specificity towards electrophilic substrates and sensitivity to inhibitors (Oakley, 2005). Thus, the underscoring Sb28GST's viability as a therapeutic target that will allow the design of selective inhibitors with minimal host toxicity.

1.2.3. Glutathione Transferase as a Therapeutic Drug Target

The absence of phase I and other detoxification enzymes from *Schistosomes'* survival arsenal highlights just how significant a role GSTs play within these parasites. The multifunctional aspect of schistosomal GSTs makes them a satisfactory primary defence mechanism against biochemical toxins and the host's immune response (Scott and McManus, 2000). This is corroborated by studies showing the increased expression of schistosomal GSTs in response to chemotherapeutic interventions (Sheehan et al., 2001). Additionally, there have been growing claims reporting schistosomal resistance to PZQ (Adenowo et al., 2015). Further investigation into these claims showed that GST was responsible for this growing resistance. (Barycki et al., 1997) showed that PZQ molecules bind to the dimer interface (L-site) of the enzyme 9° from the tyrosine active site, suggesting that drug binding is the cause of the drug resistance as opposed to the enzyme catalyzed GSH conjugation. Research from (Tang et al., 2019) showed that silencing schistosomal genes that code for the 28 kDa isoform of GST inhibited female worm fecundity. This ultimately causes a redox reaction imbalance within the parasite and terminates the development of the *Schistosomes* at the egg stage of their life cycle (Torres-Rivera and Landa, 2008). The

ubiquitousness of GSTs across different organisms might raise concerns about inhibitor sensitivity as *S. bovis* 28 kDa GST is within the same class as host GSTs. However, Sb28GSTs and helminth GSTs in general contain sufficient regions of structural differences from their host GSTs to be affected by different inhibitors (Grevelding et al., 2018, Islam et al., 2011). The inhibition of this detoxification function could expose the *Schistosome* parasite to the toxic products generated by ruminant immune attack and hinder further propagation (Huang et al., 2012). Additionally, the ligandin function that the enzyme possesses has implications in GST activity as ligand binding can inhibit catalysis effectively aiding in immune response thus making them promising drug targets for the treatment of schistosomiasis (Yassin et al., 2004). A *Schistosoma japonicum* GST (Sj26GST) enzyme was hypothesised by (McTigue et al., 1995) to be the molecular target of antiparasitic PZQ. To critically assess this theory, researchers tested whether PZQ could inhibit the enzyme activity of Sj26GST using various model substrates, or disrupt the binding of physiologically relevant non-substrates (Vale et al., 2017). The enzymatic activity assays showed that PZQ did not inhibit the enzyme even at high concentrations of up to 500 μM when tested with standard GST substrates such as 1-chloro-2,4-dinitrobenzene, 3,4-dichloronitrobenzene and ethacrynic acid. When assayed in the presence of bulkier GST substrates such as 5-androsten-3,17-dione and sulfobromophthalein, no enzyme activity was observed, which indicated the inability of Sj26GST to act on them in the presence of PZQ. The inhibitory effects of the G-site specific S-alkyl-glutathione conjugates were not altered in the presence of PZQ, showing non-interference with glutathione binding at the G-site. Lastly, PZQ had no effect on bilirubin-mediated inhibition and other non-substrate ligands, but instead offered Sj26GST non-specific partial protection against hematin-mediated inhibition (Vale et al., 2017). (McTigue et al., 1995) hypothesis was dispelled, highlighting the need for better targeted inhibitors with GSTs as a promising adjunct anti-*Schistosome* therapy.

1.3. Computer Aided Drug Design

While the interest to establish GSTs as therapeutic drug targets against schistosomiasis has been highlighted, the route to drug discovery and development is often long and costly. The pipeline that various lead compounds travel from in-vitro and in-vivo studies through to their final development into small molecule drugs can take upwards of ten years and cost a minimum of $ 2 - 5 billion (Figure 1.8) (Zhang et

al., 2022). This is largely attributed to the repetitive optimisation of lead compounds and late-stage failure of drug candidates (Sánchez-Linares et al., 2012). The inclusion of computer aided drug design (CADD) methods into the drug discovery and development process however, has massively lowered both the timeline and expenditure previously experienced (Pooe et al., 2021). Many researchers now consider it an essential part of the pipeline that is to be systematically integrated with empirical methods.

Figure 1. 8: Drug design and development adapted from (Zhang et al., 2022)

This increasingly popular rationale approach to drug discovery uses high power-atomic level computers with powerful Graphics Processing Units (GPUs) (Carmena-Bargueño et al., 2023). This advanced machinery is used in conjunction with sophisticated algorithms that incorporate Newtonian physics to simulate the quantum mechanical motions of biomolecules such as proteins with high accuracy (Hollingsworth and Dror, 2018). This essentially allows for the prediction of promising lead compounds whose molecular functions and intermolecular dynamic interactions can be extrapolated (Zhang et al., 2022). This approximation is achieved through the calculation of the forces exerted by bonded and non-bonded atoms relative to their spatial positions (Hospital et al., 2015).

These computational tools are initially used in combination with empirical structural biology techniques such as X-ray crystallography, cryo-electron microscopy, nuclear magnetic resonance and electron paramagnetic resonance (Hollingsworth and Dror, 2018). The structural data from the mentioned techniques typically provides a starting point from which a computer model of the molecular system of interest is built (Durrant and McCammon, 2011). Admittedly, the 3D snapshots offered by these techniques are undoubtedly essential for drug design, providing invaluable insights into the biomolecules' macromolecular structure (Maveyraud and Mourey, 2020). They

unfortunately provide a partial static view which excludes the constant motion and ultimately, the dynamic properties of the protein being studied (Karplus and McCammon, 2002). The two main tools used in CADD are high-throughput virtual screening (HTVS) and molecular dynamic simulations (MDS) (Pearson et al.,2021).

1.3.1. High Throughput Virtual Screening

High throughput screening of a wide selection of compounds has traditionally been considered the best method for the identification of lead compounds in ligand-based drug discovery (Sánchez-Linares et al., 2012). However, this method requires high costs to procure and biochemically test hundreds of thousands to millions of potential drug candidates. The computational adaptation of this drug discovery technique not only allows for a cost-effective way to screen compounds, but also allows for the expansion of drug candidates through the exploration of non-natural ligands (Pyzer-Knapp et al., 2015).

The efficiency and precision of HTVS hits are heavily dependent on how the receptor binding pocket/s are described empirically (Yang et al., 2011). This description is typically translated from the biomolecular crystal structure bound to the ligands through molecular docking techniques (Zhang et al., 2022). A library of druggable compounds and the protein drug targets are initially prepared for high throughput docking in multiple poses (Zhou and Caflisch, 2010). Thereafter, a scoring system is implemented to calculate the binding affinities between the molecular drug targets and library compounds (Tripathi and Bandyopadhyay, 2022, Zhang et al., 2014). The scoring functions use molecular mechanical models to either represent water as a homogenous dielectric or simulate individual water molecules within the prepared system (Pyzer-Knapp et al., 2015). Alternative scoring functions use Generalized Born Surface Area models to account for the electrostatic interactions parametrised on the Poisson-Boltzmann equation (Kuhn et al., 2005, Massova et al., 2000). To arrive at a final score, which is the measure of the binding free energy, these scoring models are used together. This allows for a well-rounded calculation of the binding score with influences from solvation, entropy and receptor flexibility taken into consideration (Yang et al., 2011). The most popularly used scoring function combinations are the molecular mechanics/ Poisson-Boltzmann surface area and the molecular mechanics models (Zhang et al., 2014).

1.3.2. Molecular Dynamic Simulations

While HTVS allows for the quantitative identification of protein binding pockets and the determination of binding affinities between ligand-protein complexes, it cannot predict the quantum mechanical motions of proteins/complexes as a function of time (Durrant and McCammon, 2011). Proteins are highly dynamic molecules which undergo conformational changes to perform their molecular functions (Hollingsworth and Dror, 2018). In drug design specifically, determining how a ligand might bind to a protein target is key to elucidating its function. A frame by frame depiction of proteins in motion could reveal loop or domain closures, which could isolate the active site consequently changing the chemical environment around substates. Conversely, conformational changes could also initiate catalytic events, bringing essential partners together (Hospital et al., 2015).

Much like HTVS, experimentally obtained structures are used to build a model from which the molecular system can be simulated. Thereafter, the molecular structure is used to determine the potential energy of the model through the calculation of the forces that act on each of the atoms, known as force fields (MacKerell Jr et al., 1998, Ott and Meyer, 1996). The mathematical models used in force field equations to approximate the behaviour of the molecular interactions that occur over a specified time are simplified using the analogues; springs, periodic functions, Lennard Jones potentials and Coulomb's law (Hospital et al., 2015). These equations allow for the simplification of molecular interaction calculations for even the largest molecular systems. For instance, springs are used for bond lengths and angles where the energy is determined from how the bonds stretch and compress from equilibrium. The periodic functions are simplified for bond rotations, representing the periodic nature of the torsional energy barriers. Lennard Jones potentials for Van de Waals interactions where atoms attract and repel each other based on their proximity (Wormald et al., 2002). Columbs laws for electrostatic interactions where forces between charged particles are calculated based on their charge and distance (Perez et al., 2008, Rueda et al., 2007, Wormald et al., 2002). Once the individual forces of the atoms are obtained, Newtonian laws of motion are used to approximate accelerations and velocities and consistently updated for the position of the atoms (Hollingsworth and Dror, 2018). This allows for a full depiction of the protein's molecular dynamic activity

and ligand bioactivity in an explicit solvent model to be extrapolated for drug design (Durrant and McCammon, 2011).

1.4. Protein Structure and Dynamics in Different Protein Crystal Space Groups

As powerful and invaluable as these computational tools are in the simplification of drug design and discovery, their haphazard utilization can lead researchers astray. More specifically, researchers have not considered a more cautious approach to the implementation of the empirical data from which the computational studies are built, and consequently, the extrapolations from the resulting data. The blind assumption that proteins crystallised in different space groups would behave similarly in *in-silico* modelling has often led to inaccurate predictions of lead compound performances when validated empirically (Beran, 2016). Principles in molecular biology however, dictate that varying symmetries and arrangements of protein molecules in the crystal lattice could induce various discrepancies between computational predictions and empirical results (Yeates and Kent, 2012).

1.4.1. Protein Crystallography

Elucidating the atomic structures of proteins is a profound gateway to understanding their function and assists in structure-based drug design (Maveyraud and Mourey, 2020). The leading technique and gold standard for obtaining the 3D structure of proteins is X-ray crystallography (Yeates and Kent, 2012). In this technique, protein crystals are grown using highly specific components and conditions that promote the protein's self-assembly into a crystalline state from solution (Ilari et al., 2008). The formation of this crystalline state requires controlled protein precipitation from a saturated solution. This allows for the periodic assembly of the protein molecules into highly sophisticated patterns that extend in three dimensions, building the crystal lattice (McPherson and Gavira, 2014). The individual structural element that makes up the crystal lattice is called the unit cell. The unit cell, in a 3D box formation, is repeated in all directions, eventually creating the entire crystal (Nespolo et al., 2018). Thereafter, an X-ray from any range of sources is illuminated onto the protein crystal, where the scattered and diffracted beam’s direction and intensity are measured (Yeates and Kent, 2012). Subsequently, information regarding the electron density is extrapolated to determine the structure of the molecules within the unit cells (Ilari et al., 2008).

1.4.2. Protein Crystal Space Groups

During protein precipitation, protein molecules can assemble in a wide variety of arrangements into a crystalline form (Yeates and Kent, 2012). The molecular arrangement in which the unit cells organize is guided by protein packing at its lowest free energy state and most stable form (Beran, 2016). The unit cell is defined by three axes: *a*, *b* and *c*, which describe the edges of the unit cell, and three angles α (angle between *b* and *c*), β (angle between *a* and *c*) and γ (angle between *a* and *b*) (Nespolo et al., 2018). Additionally, crystals exhibit symmetry, which is defined through unit cell inversion, reflections, translational and rotational operations (Yeates and Kent, 2012). The culmination of the unit cell parameters, arrangement of the unit cell within the crystal lattice and the variable combination of the symmetry operations gives rise to 230 space groups. Protein crystals, however, can only assemble into 65 different space groups due to their chirality (Wukovitz and Yeates, 1995). These space groups are divided into seven crystal systems which are namely triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal and cubic shown in Figure 1.9 (Gaur, 2021). The classification into crystal systems is based on the geometry of the unit cells and primarily defined by the relationship between the unit cell axes (*a*, *b*, *c* and the angles between them (α, β and γ), shown in Table 1.2.

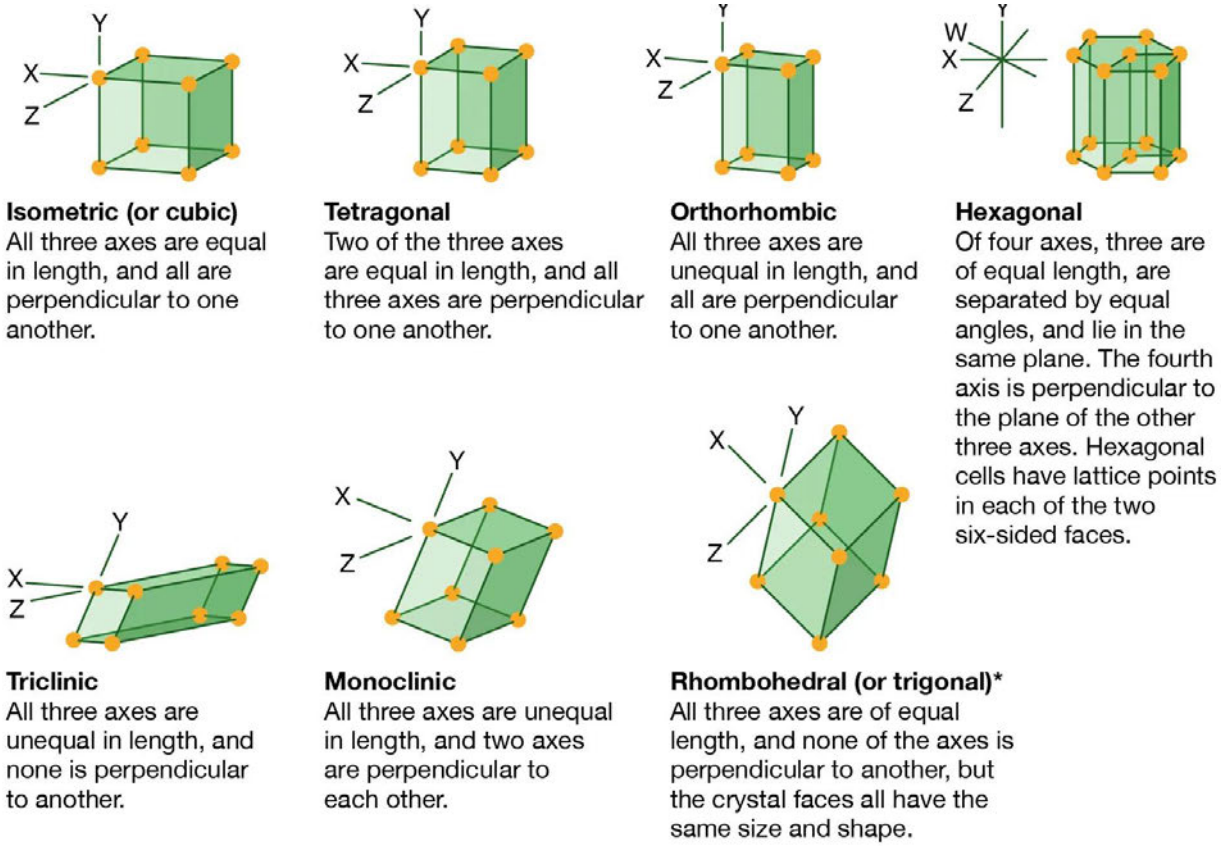


Figure 1. 9: The seven primitive crystal systems (The Editors of Encyclopaedia Britannica, 2017)

Table 1. 2: Geometrical characteristics of the seven primitive crystal systems

Crystal system	Unit cell axes	Unit cell angle	Space group
Triclinic	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90^{\circ}$	P1
Monoclinic	$a \neq b \neq c$	$\alpha = \gamma = 90^{\circ}, \beta \neq 90^{\circ}$	P2, $P2_1$,C2
Orthorhombic	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^{\circ}$	P222, $P222_1$, $P2_12_12$, $P2_12_12_1$, $C222_1$, C222, F222, I222, $I2_12_12_1$
Tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90^{\circ}$	P4, $P4_1$, $P4_2$, $P4_3$, I4, $I4_1$, P422, $P42_12$, $P4_122$, $P4_12_12$, $P4_222$, $P4_22_12$, $P4_322$, $P4_32_12$, I422, $I4_122$
Trigonal	$a = b = c$	$\alpha = \beta = \gamma \neq 90^{\circ}$	P3, $P3_1$, $P3_2$, R3, P312, P321, $P3_112$, $P3_121$, $P3_212$, $P3_221$, R32
Hexagonal	$a = b \neq c$	$\alpha = \beta = 90^{\circ}, \gamma = 120^{\circ}$	P6, $P6_1$, $P6_5$, $P6_2$, $P6_4$, $P6_3$, P622, $P6_122$, $P6_522$, $P6_222$, $P6_422$, $P6_322$
Cubic	$a = b = c$	$\alpha = \beta = \gamma = 90^{\circ}$	P23, F23, I23, $P2_13$, $I2_13$, P432, $P4_232$, F432, $F4_132$, I432, $P4_332$, $P4_132$, $I4_132$

1.4.3. Protein Crystal Polymorphs

During protein precipitation, the crystal packing order can occur in any number of ways. Variation in the molecular assembly of the unit cells can thus alter the unit cell parameters and symmetry operations (Nespolo et al., 2018). This in turn allows proteins to be formed in different crystallographic phases or space groups in a phenomenon commonly referred to as crystal polymorphism (Beran, 2016). While the intrinsic structure determined from crystal polymorphs remains unchanged, there still exists fundamental molecular differences observed in the crystal packing, symmetries and interactions within the crystal lattice between polymorphs (Yeates and Kent, 2012).

Protein crystal polymorphism is not an uncommon phenomenon, as proteins can be naturally inclined to form structural polymorphs based on a multitude of factors. These factors can range from the protein's properties to crystallisation conditions, which ultimately affect the crystal packing forces (Ilari et al., 2008). For instance, when a dynamic protein with flexible regions approaches supersaturation, reaching its lowest free energy state, the conformation in which nucleation begins will ultimately dictate the packing forces utilized to build the protein crystal (Beran, 2016). Thus, allowing the protein crystal to form in different space groups. In a similar fashion, the surface properties of a dynamic protein in solution, governed by the distribution of hydrophobic, polar, and charged amino acid residues, can engage in different intermolecular interactions (Du et al., 2016, Nangia, 2008). This can influence the formation of the crystal lattice in different crystallographic phases, allowing different space group crystals to emerge (Nespolo et al., 2018). Changes in crystallization conditions can similarly induce the formation of different space group crystals. Understandably so, as the crystal environment plays a significant role in stabilizing conformations. For instance, changes in pH, and the addition of different precipitants at varied concentrations could alter protein-protein interactions which again alters crystal packing forces and consequently influence the formation of different space groups. Changes in temperature as well as the addition of ligands and co-factors affect the molecular dynamics and the free energy state of proteins, effectively inducing different crystal packing forces yet again (Ilari et al., 2008). Additionally, the inclusion of ligands favours specific conformations and intermolecular interactions, inducing different crystal packing forces yet again (Yeates and Kent, 2012).

A consequence to be expected from varied packing assemblies in the formation of protein crystals is different packing densities, with high protein density leading to reduced conformational freedom of a protein. Conversely, low packing density in turn potentially allows increased flexibility and conformational variability (Weathers et al., 2007). Furthermore, the solvent content of crystal polymorphs may reveal drastically different solvent channels due to the crystal packing forces utilised (Hospital et al., 2015). These solvent-filled holes and channels can occupy anywhere from 40 – 60 % of the crystal volume (Ilari et al., 2008). Differences in crystal packing forces additionally influence symmetry operations and consequently symmetry constraints. These constraints alter the functionality of flexible regions of proteins such as loops and amino acid side chains (Nangia, 2008). Most significantly, variations in crystal packing events may promote varying intermolecular interactions such as hydrogen bonds and Van de Waals interactions with neighbouring protein molecules (Charlton, 1998).

1.4.3.1. Polymorphism In Glutathione Transferases

The frequent nature of crystal polymorph formation is reflected through the abundance of 3D structural data deposited in many reputable databases such as Protein Data Bank (PDB) and RCSB (Bank, 1971). The analysis of the 38 *Schistosoma* GSTs deposited in PDB shows a great number of 3D structures obtained from polymorphic crystals. From the 29 GST structures of *S. japonicum* deposited on PDB, 26 were derived from crystal polymorphs with their distribution highlighted on Figure 1.10. Similarly, from a total of nine 3D structures of *S. haematobium* deposited on PDB, six were derived from polymorphic crystals. A broader analysis of GSTs with a larger sample size shows that human GSTs share the same polymorphic propensity. Human GST Alpha 1-1 (HGSTA1-1), Human GST Pi 1-1 (HGSTPi1-1) and Human GST Omega 1 (HGSTO1) each with a total of 47, 43 and 15 3D structures deposited into PDB respectively similarly had 46, 40 and 14 structures derived from polymorphic crystals (Berman et al., 2000). The conformational adaptivity that GSTs possess to accommodate the large diversity of ligands they neutralize, permits the proteins to be malleable to various crystal packing assemblies and consequently, crystal group formations.

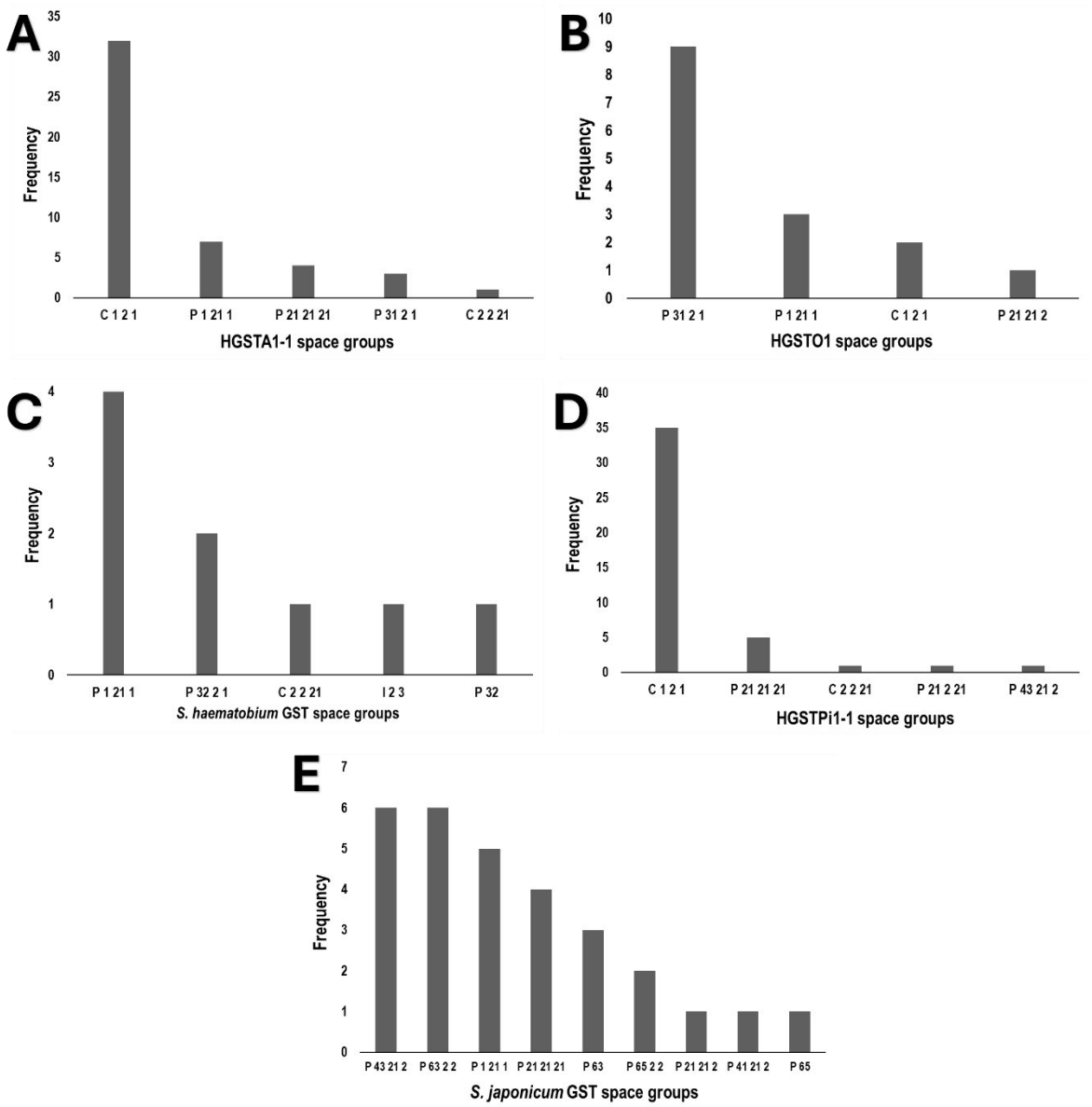


Figure 1. 10: Frequency and distribution of GST crystal polymorphs populated from PDB

1.4.4. Crystal Polymorph Implications in Computational Studies

As the addition of computational tools becomes an increasingly pertinent step in the drug design and development pipeline, the scrutinization 3D structure space groups should be considered a more critical priority. It is typically assumed that the selection criteria to build reliable molecular systems is 3D structure resolution (Maveyraud and Mourey, 2020). However, the impact of neglecting crystal space groups and crystal systems influence over the predicted data could lead to incomplete and misleading

data extrapolation (Du et al., 2016). The consequence of which would be the selection of false positives as lead compounds to be investigated and optimised.

The shortcomings of haphazardly selecting 3D structures solely on resolution and not multiplicities include but are not limited to: (i), Oversimplification. Data from a single space group crystal is translated into a model, which is fixed in a particular conformation (Maveyraud and Mourey, 2020). The data translated computationally will not show the full range of conformational states that the protein can adopt. As a result, the force fields calculated will be oversimplified, computing very particular conformational flexibilities and molecular interactions that fail to show complex interactions and lack nuance (Weathers et al., 2007, Yang et al., 2011). This could potentially lead to inaccuracies when modelling highly dynamic systems such as GSTs. (ii), Alteration of binding site. One space group model could exhibit varied unit cell geometry from the next, resulting in the accessibility of the ligand binding site being slightly altered. This could lead to the exclusion of important molecular interactions and steric clashes (Wormald et al., 2002). The unfortunate implications of this would be the additional exclusion of promising lead compounds, due to the biased selection of compounds better fit for the protein conformation inputted. (iii). Thermodynamic and kinetic inconsistencies. Polymorphic crystals possess different crystal lattice energies. This creates differences in the energy landscape of the polymorphs, with varying free energy levels computed for ligand binding (Beran, 2016, Du et al., 2016). Consequently, the use of one space group model could produce suboptimal docking poses and misleading binding affinity predictions and scores (Zhang et al., 2014). Therefore, to improve the predictive accuracy and reliability of computationally determined outputs, the incorporation of crystal multiplicities should be implemented in CADD.

1.5. Rationale and Significance of This Study

Schistosomiasis is a global burden not only to wild and domestic animals, but human as well. From a global health perspective, it is considered the most important and devastating water-based disease and as such, is listed by the WHO as a NTD (Organization, 2022). While the global burden of schistosomiasis far outweighs that of many diseases, it is still overlooked by policymakers, who prioritise the likes of malaria and TB, ultimately limiting resources for schistosomiasis research (Organization, 2013). Fortunately, this oversight is slowly being corrected with the WHO proposing a

revised roadmap to the control and elimination of schistosomiasis by the year 2030 using a One Health approach. With the One Health approach aiming to optimise the health of people, animals and the ecosystem, therapies against *S. bovis*, which economically and socially threaten the livelihoods of traditional and commercial farmers, food security, animal welfare and the environment, would greatly help this cause (de la Torre-Escudero et al., 2017, Organization, 2020).

While there exist anti-schistosomal drugs such as PZQ, the partial control they provide, coupled with the emerging resistance to them, allows the parasite to continue plaguing much of the world's human and animal population (King and Dangerfield-Cha, 2008). However, the development of new generation anti-schistosomal drugs as aimed in this study, would aid in recovery measures taken to propel anti-schistosomal strategies. Targeting *S. bovis'* sole defence mechanism against toxin build up and host immune responses appears to be the most promising approach to aid in the elimination of the parasite's pathogenicity and transmission potential. The indispensable role of Sb28GST to the parasites' redox metabolism, combined with its structural divergence from mammalian GSTs, presents the enzyme as an attractive and strategic selective target for anti-schistosomal drug development (Scott and McManus, 2000, Sheehan et al., 2001).

The approach to identifying new generation anti-schistosomal drugs while addressing a critical gap in CADD serves to caution researchers of the pitfalls of haphazardly using tools simply based on their popularity and ease of use. Most CADD methods rely on structural data from crystal structures obtained experimentally as a basis for the construction of simulation systems (Hollingsworth and Dror, 2018). However, empirical structural methods do not fully capture the dynamic nature of proteins (Maveyraud and Mourey, 2020). Therefore, an investigation into the influence that protein crystal polymorphs have on the conformational dynamics of simulated protein systems and their subsequent impact on ligand selectivity could reveal improved predictive accuracy and reliability of computationally determined lead compounds.

1.6. Study Hypothesis

This study hypothesises that the use of protein models derived from different crystal space groups is crucial for a more comprehensive understanding of protein behaviour, leading to the prediction of reliable lead compounds in CADD. In so doing, the

acknowledgement of the impact protein polymorphism has in CADD will allow the successful identification of effective inhibitors of parasitic enzymes such as Sb28GST, to advance the development of new generation anti-schistosomal drugs.

1.7. Aims and Objectives

This study aims to investigate the influence that protein crystal polymorphs have on the conformational dynamics of simulated protein systems and their subsequent impact on ligand selectivity in CADD. Furthermore, this work contributes to the development of new generation anti-*Schistosomes* through the identification empirical confirmation of Sb28GST inhibitors.

The following objective were set to achieve this aim:

- Overexpress and purify recombinant Sb28GST using an *Escherichia coli* (*E. coli*) system and affinity chromatography respectively.
- Crystallize Sb28GST in a monoclinic alternative space group
- Solve the structure of Sb28GST and determine conformational features
- Perform MD simulations of Sb28GST from different space group models
- Evaluate and compare ligand selectivity from different space group models
- Identify lead compounds and determine molecular interactions with Sb28GST using MD simulations
- Characterize the ligandin properties of Sb28GST with selected lead compounds

1.8. Novelty of The Study

While schistosomal GSTs from species like *S. japonicum* and *S. haematobium* have been explored as drug targets for years, their host spectrum is primarily limited to humans (Tsuji, 2020). The extent of their global burden does not exert tremendous socio-economic pressure on the livelihood of farmers, the food export trade of 3rd world countries, compromise the environment through excessive production of greenhouse gases, nor do they threaten food security (Pennance et al., 2021, Strydom et al., 2023). *Schistosoma bovis* however, is guilty of the aforementioned, whilst receiving limited attention for its control. This study reports the first investigation of synthesizable flavonoids as therapeutic agents against the parasitic Sb28GST protein. Flavonoids, which are natural polyphenolic compounds found in abundance in fruits and

vegetables, have been reported to have antiviral, antibacterial, antioxidant and antiradical properties (Procházková et al., 2011). Studies have reported the modulation of various drug metabolising enzymes by flavonoids. Not only do these compounds reduce the production of GSTs within cells and tissues, but studies have also shown their ability to directly inhibit the catalytic activity of GSTs (Chaudhary et al., 2010, Wiegand et al., 2009). Therefore, in the quest to develop novel anti-schistosomal drugs, synthesizable flavonoids strongly present as promising therapeutic agents. Addressing the pursuit of developing novel anti-schistosomal drugs, this study successfully expressed, purified and crystallised catalytically active recombinant Sb28GST. This led to the first successful empirical validation of computationally selected lead flavonoid compounds as inhibitors of Sb28GST.

This study additionally addresses a significant gap in CADD methodology through investigation of the impact of protein crystal polymorphism on computational ligand binding and effectively, ligand selectivity. Through the successful crystallisation of Sb28GST in an orthorhombic crystal system, the comparative computational studies carried out showed the repercussions of neglecting multi-space group crystal systems in CADD. The study showed that proteins crystallised in different crystal systems exhibit distinct cavity constructions, dynamic behaviours, and ligand interactions in molecular dynamic simulations (Mfeka et al., 2025). This suggests the importance of the collective assessment of multi-space group crystals to obtain a more comprehensive, multi-dimensional understanding of protein dynamics and ligand interactions.

1.9. References

ADENOWO, A. F., OYINLOYE, B. E., OGUNYINKA, B. I. & KAPPO, A. P. J. B. J. O. I. D. 2015. Impact of human schistosomiasis in sub-Saharan Africa. 19**,** 196-205.

AKUMADU, B. O., PANDIAN, R., OLFSEN, J., WORTH, R., THULO, M., MENTOR, T., FANUCCHI, S., SAYED, Y., DIRR, H. W., ACHILONU, I. J. M. & PARASITOLOGY, B. 2020. Molecular basis of inhibition of Schistosoma japonicum glutathione transferase by ellagic acid: Insights into biophysical and structural studies. 240**,** 111319.

ANGELUCCI, F., BAIOCCO, P., BRUNORI, M., GOURLAY, L., MOREA, V. & BELLELLI, A. J. S. 2005. Insights into the catalytic mechanism of glutathione S-transferase: the lesson from Schistosoma haematobium. 13**,** 1241-1246.

ANGORA, E. K., ALLIENNE, J.-F., REY, O., MENAN, H., TOURÉ, A. O., COULIBALY, J. T., RASO, G., YAVO, W., N'GORAN, E. K. & UTZINGER, J. J. P. 2020. High prevalence of Schistosoma haematobium× Schistosoma bovis hybrids in schoolchildren in Côte d'Ivoire. 147**,** 287-294.

ARIYANAYAGAM, M. R., FAIRLAMB, A. H. J. M. & PARASITOLOGY, B. 2001. Ovothiol and trypanothione as antioxidants in trypanosomatids. 115**,** 189-198.

ARMSTRONG, R. N. J. A. I. E. & BIOLOGY, R. A. O. M. 1993. Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme. 1-44.

BABES, R.M., SELESCU, T., DOMOCOS, D., BABES, A. J. T. & PHARMACOLOGY, A. 2017. The anthelminthic drug praziquantel is a selective agonist of the sensory transient receptor potential melastatin type 8 channel. 336**,** 55-65.
BANK, P.D., 1971. Protein data bank. Nature New Biol, *233*(223), 10-1038.
BARYCKI, J. J., COLMAN, R. F. J. A. O. B. & BIOPHYSICS 1997. Identification of the Nonsubstrate Steroid Binding Site of Rat Liver Glutathione-S-Transferase, Isozyme 1-1, by the Steroid Affinity Label, 3β-(Iodoacetoxy) dehydroisoandrosterone. 345**,** 16-31.
BECKER, K., RAHLFS, S., NICKEL, C. & SCHIRMER, R. H. 2003. Glutathione–functions and metabolism in the malarial parasite Plasmodium falciparum.
BECKER, K., TILLEY, L., VENNERSTROM, J. L., ROBERTS, D., ROGERSON, S. & GINSBURG, H. J. I. J. F. P. 2004. Oxidative stress in malaria parasite-infected erythrocytes: host–parasite interactions. 34**,** 163-189.
BERAN, G. J. J. C. R. 2016. Modeling polymorphic molecular crystals with electronic structure theory. 116**,** 5567-5613.
BERMAN, H. M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T. N., WEISSIG, H., SHINDYALOV, I. N. & BOURNE, P. E. J. N. A. R. 2000. The protein data bank. 28**,** 235-242.
BOARD, P. G. & MENON, D. J. B. E. B. A.-G. S. 2013. Glutathione transferases, regulators of cellular metabolism and physiology. 1830**,** 3267-3288.
CAPPER, J. L. J. O. H. O. 2023. The impact of controlling diseases of significant global importance on greenhouse gas emissions from livestock production. 5**,** 17.
CARMENA-BARGUEÑO, M., MARTÍNEZ-CORTÉS, C., BANEGAS-LUNA, A. J. & PÉREZ-SÁNCHEZ, H. 2023. TOLEDO: Accelerated Maestro GUI molecular dynamics simulations.
CHARLIER, J., RINALDI, L., MUSELLA, V., PLOEGER, H. W., CHARTIER, C., VINEER, H. R., HINNEY, B., VON SAMSON-HIMMELSTJERNA, G., BĂCESCU, B. & MICKIEWICZ, M. J. P. V. M. 2020. Initial assessment of the economic burden of major parasitic helminth infections to the ruminant livestock industry in Europe. 182**,** 105103.
CHARLIER, J., VELDE, F. V., VAN DER VOORT, M., VAN MEENSEL, J., LAUWERS, L., CAUBERGHE, V., VERCRUYSSE, J. & CLAEREBOUT, E. J. V. P. 2015. ECONOHEALTH: Placing helminth infections of livestock in an economic and social context. 212**,** 62-67.
CHARLTON, M. J. J. O. T. C. S., PERKIN TRANSACTIONS 2 1998. Theoretical investigations of conformational aspects of polymorphism. Part 1: o-acetamidobenzamide. 763-772.
CHAUDHARY, P., SHARMA, R., SHARMA, A., VATSYAYAN, R., YADAV, S., SINGHAL, S. S., RAUNIYAR, N., PROKAI, L., AWASTHI, S. & AWASTHI, Y. C. J. B. 2010. Mechanisms of 4-hydroxy-2-nonenal induced pro-and anti-apoptotic signaling. 49**,** 6263-6275.
CROOM, E. J. P. I. M. B. & SCIENCE, T. 2012. Metabolism of xenobiotics of human environments. 112**,** 31-88.
DE LA TORRE-ESCUDERO, E., PÉREZ-SÁNCHEZ, R., MANZANO-ROMÁN, R., OLEAGA, A. J. M. & PARASITOLOGY, B. 2017. Schistosoma bovis-host interplay: Proteomics for knowing and acting. 215**,** 30-39.
DEMLEW, B. A. & TESSMA, A. K. J. S. A. R. J. A. M. S. 2020. Review on bovine schistosomiasis and its associated risk factors. 2**,** 44-55.
DOENHOFF, M. J., CIOLI, D. & UTZINGER, J. J. C. O. I. I. D. 2008. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. 21**,** 659-667.
DU, X., LI, Y., XIA, Y.-L., AI, S.-M., LIANG, J., SANG, P., JI, X.-L. & LIU, S.-Q. J. I. J. O. M. S. 2016. Insights into protein–ligand interactions: mechanisms, models, and methods. 17**,** 144.
DURRANT, J. D. & MCCAMMON, J. A. J. B. B. 2011. Molecular dynamics simulations and drug discovery. 9**,** 1-9.
ESCOBAR, L. S., JARA, W. H., NIZAM, Q. N. & PLAVŠIĆ, B. 2018. The perspective of the world organisation for animal health. *Advances in Agricultural Animal Welfare.* Elsevier.
FINDLAY, V. J., TOWNSEND, D. M. & TEW, K. D. J. C. D. R. 2006. Glutathione and Glutathione S-Transferases in drug resistance. 213-221.
FLOHÉ, L. J. B. A. 2012. The trypanothione system and the opportunities it offers to create drugs for the neglected kinetoplast diseases. 30**,** 294-301.
FRANCO, R., CIDLOWSKI, J. J. C. D. & DIFFERENTIATION 2009. Apoptosis and glutathione: beyond an antioxidant. 16**,** 1303-1314.
FROVA, C. J. B. E. 2006. Glutathione transferases in the genomics era: new insights and perspectives. 23**,** 149-169.
GAUR, R. K. J. A. C. S. F. S. B. C. 2021. Frequency distribution of space groups in soluble and membrane proteins and their complexes. 77**,** 187-191.

GINSBURG, H. & GOLENSER, J. J. R. R. 2003. Glutathione is involved in the antimalarial action of chloroquine and its modulation affects drug sensitivity of human and murine species of Plasmodium. 8**,** 276-279.
GREENBERG, R. M. J. F. M. 2007. Molecular target of the antischistosomal drug praziquantel. 2**,** 265-268.
GREVELDING, C. G., LANGNER, S. & DISSOUS, C. J. T. I. P. 2018. Kinases: molecular stage directors for schistosome development and differentiation. 34**,** 246-260.
HILLER, N., FRITZ-WOLF, K., DEPONTE, M., WENDE, W., ZIMMERMANN, H. & BECKER, K. J. P. S. 2006. Plasmodium falciparum glutathione S-transferase—Structural and mechanistic studies on ligand binding and enzyme inhibition. 15**,** 281-289.
HOEKSTRA, P. T., VAN DAM, G. J. & VAN LIESHOUT, L. J. F. I. T. D. 2021. Context-specific procedures for the diagnosis of human schistosomiasis–a mini review. 2**,** 722438.
HOLLINGSWORTH, S. A. & DROR, R. O. J. N. 2018. Molecular dynamics simulation for all. 99**,** 1129-1143.
HOSPITAL, A., GOÑI, J. R., OROZCO, M., GELPÍ, J. L. J. A., BIOINFORMATICS, A. I. & CHEMISTRY 2015. Molecular dynamics simulations: advances and applications. 37-47.
HUANG, H.-H., RIGOUIN, C. & L. WILLIAMS, D. J. C. P. D. 2012. The redox biology of schistosome parasites and applications for drug development. 18**,** 3595-3611.
ILARI, A., SAVINO, C. J. B. D., SEQUENCE ANALYSIS & EVOLUTION 2008. Protein structure determination by x-ray crystallography. 63-87.
ISLAM, M. K., ALIM, M. A., MIYOSHI, T., HATTA, T., YAMAJI, K., MATSUMOTO, Y., FUJISAKI, K., TSUJI, N. J. B. & COMMUNICATIONS, B. R. 2011. Longistatin, a novel plasminogen activator from vector ticks, is resistant to plasminogen activator inhibitor-1. 413**,** 599-604.
JOHNSON, K. A., ANGELUCCI, F., BELLELLI, A., HERVÉ, M., FONTAINE, J., TSERNOGLOU, D., CAPRON, A., TROTTEIN, F. & BRUNORI, M. J. B. 2003. Crystal structure of the 28 kDa glutathione S-transferase from Schistosoma haematobium. 42**,** 10084-10094.
JONSSON, N., MACLEOD, M., HAYWARD, A., MCNEILLY, T., FERGUSON, K. & SKUCE, P. J. P. V. M. 2022. Liver fluke in beef cattle–Impact on production efficiency and associated greenhouse gas emissions estimated using causal inference methods. 200**,** 105579.
JORTZIK, E. & BECKER, K. J. I. J. O. M. M. 2012. Thioredoxin and glutathione systems in Plasmodium falciparum. 302**,** 187-194.
KARPLUS, M. & MCCAMMON, J. A. J. N. S. B. 2002. Molecular dynamics simulations of biomolecules. 9**,** 646-652.
KEHR, S., STURM, N., RAHLFS, S., PRZYBORSKI, J. M. & BECKER, K. J. P. P. 2010. Compartmentation of redox metabolism in malaria parasites. 6**,** e1001242.
KING, C. H. & DANGERFIELD-CHA, M. J. C. I. 2008. The unacknowledged impact of chronic schistosomiasis. 4**,** 65-79.
KING, C. H., DICKMAN, K. & TISCH, D. J. J. T. L. 2005. Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. 365**,** 1561-1569.
KUHN, B., GERBER, P., SCHULZ-GASCH, T. & STAHL, M. J. J. O. M. C. 2005. Validation and use of the MM-PBSA approach for drug discovery. 48**,** 4040-4048.
LAMBERTON, P. H., KABATEREINE, N. B., OGUTTU, D. W., FENWICK, A. & WEBSTER, J. P. J. P. N. T. D. 2014. Sensitivity and specificity of multiple Kato-Katz thick smears and a circulating cathodic antigen test for Schistosoma mansoni diagnosis pre-and post-repeated-praziquantel treatment. 8**,** e3139.
LÉGER, E., BORLASE, A., FALL, C. B., DIOUF, N. D., DIOP, S. D., YASENEV, L., CATALANO, S., THIAM, C. T., NDIAYE, A. & EMERY, A. J. T. L. P. H. 2020. Prevalence and distribution of schistosomiasis in human, livestock, and snail populations in northern Senegal: a One Health epidemiological study of a multi-host system. 4**,** e330-e342.
LEGER, E. & WEBSTER, J. P. J. P. 2017. Hybridizations within the genus Schistosoma: implications for evolution, epidemiology and control. 144**,** 65-80.
MACKERELL JR, A. D., BASHFORD, D., BELLOTT, M., DUNBRACK JR, R. L., EVANSECK, J. D., FIELD, M. J., FISCHER, S., GAO, J., GUO, H. & HA, S. J. T. J. O. P. C. B. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. 102**,** 3586-3616.
MAKUMBE, H. H., PANDIAN, R., VALLI, A., SAYED, Y. & ACHILONU, I. J. J. O. M. S. 2024. Biophysical characterization, crystallization, and solution of the first crystal structure of the 28 kDa-Schistosoma bovis glutathione transferase. 1298**,** 136979.
MANNERVIK, B., BOARD, P. G., HAYES, J. D., LISTOWSKY, I. & PEARSON, W. R. J. M. I. E. 2005. Nomenclature for mammalian soluble glutathione transferases. 401**,** 1-8.

MASSOVA, I., KOLLMAN, P. A. J. P. I. D. D. & DESIGN 2000. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. 18**,** 113-135.
MAVEYRAUD, L. & MOUREY, L. J. M. 2020. Protein X-ray crystallography and drug discovery. 25**,** 1030.
MCMANUS, D., DUNNE, D., SACKO, M., UTZINGER, J., VENNERVALD, B. & ZHOU, X. 2018. Schistosomiasis Nat Rev Dis Primers 24: 13.
MCPHERSON, A. & GAVIRA, J. A. J. A. C. S. F. S. B. C. 2014. Introduction to protein crystallization. 70**,** 2-20.
MCTIGUE, M. A., WILLIAMS, D. R. & TAINER, J. A. J. J. O. M. B. 1995. Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase fromschistosoma japonicaand its complex with the leading antischistomal drug praziquantel. Elsevier.
MEISTER, A. & ANDERSON, M. E. J. A. R. O. B. 1983. Glutathione. 52**,** 711-760.
MFEKA, M., ONISURU, O., PANDIAN, R., SAYED, Y., KHOZA, T. & ACHILONU, I. J. R. I. C. 2025. Crystal enigma: Understanding diverse protein conformational dynamics, ligand selectivity and interaction in multi-space group crystals using computational modelling. 102288.
MOLLA, G., TINTAGU, T., YASIN, A., ALEMU, B., ASSEN, A. A. & TADESSE, K. J. P. O. 2022. Bovine schistosomiasis in some selected areas of South wollo and oromia zones of Amhara region, north-east Ethiopia. 17**,** e0259787.
MÜLLER, S. & KAPPES, B. J. T. I. P. 2007. Vitamin and cofactor biosynthesis pathways in Plasmodium and other apicomplexan parasites. 23**,** 112-121.
MÜLLER, S., LIEBAU, E., WALTER, R. D. & KRAUTH-SIEGEL, R. L. J. T. I. P. 2003. Thiol-based redox metabolism of protozoan parasites. 19**,** 320-328.
MÜLLER, S. J. M. M. 2004. Redox and antioxidant systems of the malaria parasite Plasmodium falciparum. 53**,** 1291-1305.
MUSTAFA, S., ALHARBI, L. M., ABDELRAHEEM, M. Z., MOBASHAR, M., QAMAR, W., A. AL-DOAISS, A. & ABBAS, R. Z. J. B. T. E. R. 2024. Role of Silver Nanoparticles for the Control of Anthelmintic Resistance in Small and Large Ruminants. 1-20.
NANGIA, A. J. A. O. C. R. 2008. Conformational polymorphism in organic crystals. 41**,** 595-604.
NESPOLO, M., AROYO, M. I. & SOUVIGNIER, B. J. J. O. A. C. 2018. Crystallographic shelves: space-group hierarchy explained. 51**,** 1481-1491.
NIU, X., HU, T., HONG, Y., LI, X. & SHEN, Y. J. J. O. T. M. 2022. The role of praziquantel in the prevention and treatment of fibrosis associated with schistosomiasis: a review. 2022**,** 1413711.
NOGUEIRA, R. A., LIRA, M. G. S., LICÁ, I. C. L., FRAZÃO, G. C. C. G., DOS SANTOS, V. A. F., MENDES FILHO, A. C. C., RODRIGUES, J. G. M., MIRANDA, G. S., CARVALHO, R. C., NASCIMENTO, F. R. F. J. M. & PARASITOLOGY, B. 2022. Praziquantel: An update on the mechanism of its action against schistosomiasis and new therapeutic perspectives. 252**,** 111531.
OAKLEY, A. J. J. C. O. I. S. B. 2005. Glutathione transferases: new functions. 15**,** 716-723.
OLVEDA, D. U., LI, Y., OLVEDA, R. M., LAM, A. K., CHAU, T. N., HARN, D. A., WILLIAMS, G. M., GRAY, D. J., ROSS, A. G. J. T. M. & SURGERY 2013. Bilharzia: pathology, diagnosis, management and control. 1.
ORGANIZATION, W. H. 2013. Schistosomiasis: progress report 2001-2011, strategic plan 2012-2020. *Schistosomiasis: progress report 2001-2011, strategic plan 2012-2020.*
ORGANIZATION, W. H. 2020. Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030. *Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030.*
ORGANIZATION, W. H. 2022. *WHO guideline on control and elimination of human schistosomiasis*, World Health Organization.
OTT, K. H. & MEYER, B. J. J. O. C. C. 1996. Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. 17**,** 1068-1084.
OUTLOOK, O. J. O. 2021. Oecd-fao agricultural outlook 2021–2030. 2030.
PAYNE, L., CHIODINI, P. & BUSTINDUY, A. J. C. M. N. 2023. An update on schistosomiasis. 45**,** 125-132.
PEARSON, M. S., TEDLA, B. A., BECKER, L., NAKAJIMA, R., JASINSKAS, A., MDULUZA, T., MUTAPI, F., OEUVRAY, C., GRECO, B. & SOTILLO, J. J. F. I. I. 2021. Immunomics-guided antigen discovery for praziquantel-induced vaccination in urogenital human schistosomiasis. 12**,** 663041.

PENNANCE, T., AME, S. M., AMOUR, A. K., SULEIMAN, K. R., CABLE, J., WEBSTER, B. L. J. C. R. I. P. & DISEASES, V.-B. 2021. The detection of Schistosoma bovis in livestock on Pemba Island, Zanzibar: A preliminary study. 1**,** 100056.
PEREIRA, T. A., DE MELO TRINDADE, G. V., SANTOS, E. T., PEREIRA, F. E. & DE SOUZA, M. M. J. I. J. F. P. 2021. Praziquantel pharmacotherapy reduces systemic osteopontin levels and liver collagen content in murine schistosomiasis mansoni. 51**,** 437-440.
PÉREZ-SÁNCHEZ, R., RAMAJO-HERNÁNDEZ, A., RAMAJO-MARTÍN, V. & OLEAGA, A. J. P. 2006. Proteomic analysis of the tegument and excretory-secretory products of adult Schistosoma bovis worms. 6**,** S226-S236.
PEREZ, A., LANKAS, F., LUQUE, F. J. & OROZCO, M. J. N. A. R. 2008. Towards a molecular dynamics consensus view of B-DNA flexibility. 36**,** 2379-2394.
POOE, K., WORTH, R., IWUCHUKWU, E. A., DIRR, H. W. & ACHILONU, I. J. J. O. M. S. 2021. An empirical and theoretical description of Schistosoma japonicum glutathione transferase inhibition by bromosulfophthalein and indanyloxyacetic acid 94. 1223**,** 128892.
PROCHÁZKOVÁ, D., BOUŠOVÁ, I. & WILHELMOVÁ, N. J. F. 2011. Antioxidant and prooxidant properties of flavonoids. 82**,** 513-523.
PYZER-KNAPP, E. O., SUH, C., GÓMEZ-BOMBARELLI, R., AGUILERA-IPARRAGUIRRE, J. & ASPURU-GUZIK, A. J. A. R. O. M. R. 2015. What is high-throughput virtual screening? A perspective from organic materials discovery. 45**,** 195-216.
RAHBARI, M., DIEDERICH, K., BECKER, K., KRAUTH-SIEGEL, R. L. & JORTZIK, E. J. B. C. 2015. Detection of thiol-based redox switch processes in parasites–facts and future. 396**,** 445-463.
RUEDA, M., FERRER-COSTA, C., MEYER, T., PÉREZ, A., CAMPS, J., HOSPITAL, A., GELPÍ, J. L. & OROZCO, M. J. P. O. T. N. A. O. S. 2007. A consensus view of protein dynamics. 104**,** 796-801.
SÁNCHEZ-LINARES, I., PÉREZ-SÁNCHEZ, H., CECILIA, J. M. & GARCÍA, J. M. J. B. B. 2012. High-throughput parallel blind virtual screening using BINDSURF. 13**,** 1-14.
SCOTT, J. C. & MCMANUS, D. P. J. P. I. 2000. Molecular cloning and enzymatic expression of the 28-kDa glutathione S-transferase of Schistosoma japonicum: evidence for sequence variation but lack of consistent vaccine efficacy in the murine host. 49**,** 289-300.
SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. J. B. J. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. 360**,** 1-16.
SIES, H. J. F. R. B. & MEDICINE 1999. Glutathione and its role in cellular functions. 27**,** 916-921.
STEINMANN, P., KEISER, J., BOS, R., TANNER, M. & UTZINGER, J. J. T. L. I. D. 2006. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. 6**,** 411-425.
STRYDOM, T., LAVAN, R. P., TORRES, S. & HEANEY, K. J. A. 2023. The economic impact of parasitism from nematodes, trematodes and ticks on beef cattle production. 13**,** 1599.
TANG, C.-L., ZHOU, H.-H., ZHU, Y.-W., HUANG, J. & WANG, G.-B. J. A. T. 2019. Glutathione S-transferase influences the fecundity of Schistosoma japonicum. 191**,** 8-12.
The Editors of Encyclopaedia Britannica (2017, December 8). triclinic system. Encyclopedia Britannica. https://www.britannica.com/science/triclinic-system
THOMAS, C. M. & TIMSON, D. J. J. C. M. C. 2020. The mechanism of action of praziquantel: can new drugs exploit similar mechanisms? 27**,** 676-696.
TORRES-RIVERA, A. & LANDA, A. J. A. T. 2008. Glutathione transferases from parasites: a biochemical view. 105**,** 99-112.
TRIPATHI, N. M. & BANDYOPADHYAY, A. J. E. J. O. M. C. 2022. High throughput virtual screening (HTVS) of peptide library: Technological advancement in ligand discovery. 243**,** 114766.
TSUJI, N. J. P. I. 2020. Schistosomiasis and hookworm infection in humans: Disease burden, pathobiology and anthelmintic vaccines. 75**,** 102051.
USE, F.P., 2019. Available online: https://www. fao. org/faostat/en/# data. *GN (accessed on 17 October 2021)*.
UTZINGER, J., BECKER, S., VAN LIESHOUT, L., VAN DAM, G., KNOPP, S. J. C. M. & INFECTION 2015. New diagnostic tools in schistosomiasis. 21**,** 529-542.
UTZINGER, J., RASO, G., BROOKER, S., DE SAVIGNY, D., TANNER, M., ØRNBJERG, N., SINGER, B. & N'GORAN, E. J. P. 2009. Schistosomiasis and neglected tropical diseases: towards integrated and sustainable control and a word of caution. 136**,** 1859-1874.
VALE, N., GOUVEIA, M. J., RINALDI, G., BRINDLEY, P. J., GÄRTNER, F., CORREIA DA COSTA, J. M. J. A. A. & CHEMOTHERAPY 2017. Praziquantel for schistosomiasis: single-drug metabolism revisited, mode of action, and resistance. 61**,** 10.1128/aac. 02582-16.

VAN DER WERF, M. J., DE VLAS, S. J., BROOKER, S., LOOMAN, C. W., NAGELKERKE, N. J., HABBEMA, J. D. F. & ENGELS, D. J. A. T. 2003. Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. 86**,** 125-139.
VERCRUYSSE, J., CHARLIER, J., VAN DIJK, J., MORGAN, E. R., GEARY, T., VON SAMSON-HIMMELSTJERNA, G. & CLAEREBOUT, E. J. P. 2018. Control of helminth ruminant infections by 2030. 145**,** 1655-1664.
WANG, W., WANG, L. & LIANG, Y.-S. J. P. R. 2012. Susceptibility or resistance of praziquantel in human schistosomiasis: a review. 111**,** 1871-1877.
WEATHERS, E. A., PAULAITIS, M. E., WOOLF, T. B., HOH, J. H. J. P. S., FUNCTION, & BIOINFORMATICS 2007. Insights into protein structure and function from disorder–complexity space. 66**,** 16-28.
WEERAKOON, K. G., GOBERT, G. N., CAI, P. & MCMANUS, D. P. J. C. M. R. 2015. Advances in the diagnosis of human schistosomiasis. 28**,** 939-967.
WIEGAND, H., BOESCH-SAADATMANDI, C., REGOS, I., TREUTTER, D., WOLFFRAM, S., RIMBACH, G. J. N. & CANCER 2009. Effects of quercetin and catechin on hepatic glutathione-S transferase (GST), NAD (P) H quinone oxidoreductase 1 (NQO1), and antioxidant enzyme activity levels in rats. 61**,** 717-722.
WOLMARANS, C. T., DE KOCK, K. N. J. S. A. J. O. E. & INFECTION 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. 24**,** 13-17.
WORMALD, M. R., PETRESCU, A. J., PAO, Y.-L., GLITHERO, A., ELLIOTT, T. & DWEK, R. A. J. C. R. 2002. Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. 102**,** 371-386.
WU, W., WANG, W. & HUANG, Y.-X. J. P. R. 2011. New insight into praziquantel against various developmental stages of schistosomes. 109**,** 1501-1507.
WUKOVITZ, S. W. & YEATES, T. O. J. N. S. B. 1995. Why protein crystals favour some space-groups over others. 2**,** 1062-1067.
XIAO, S.-H., JJ, Z., YU, S. C. X., CHONG, S., PARASITOLOGY, B. Z. Z. C. J. O. & DISEASES, P. 2008. Impact of host factors on the schistosome-killing process induced by praziquantel. 26**,** 217-225.
YANG, T., WU, J. C., YAN, C., WANG, Y., LUO, R., GONZALES, M. B., DALBY, K. N., REN, P. J. P. S., FUNCTION, & BIOINFORMATICS 2011. Virtual screening using molecular simulations. 79**,** 1940-1951.
YASSIN, Z., ORTIZ-SALMERÓN, E., GARCíA-MAROTO, F., BARÓN, C., GARCíA-FUENTES, L. J. B. E. B. A.-P. & PROTEOMICS 2004. Implications of the ligandin binding site on the binding of non-substrate ligands to Schistosoma japonicum-glutathione transferase. 1698**,** 227-237.
YEATES, T. O. & KENT, S. B. J. A. R. O. B. 2012. Racemic protein crystallography. 41**,** 41-61.
ZHANG, B., LI, H., YU, K. & JIN, Z. J. C. T. O. H. P. C. 2022. Molecular docking-based computational platform for high-throughput virtual screening. 1-12.
ZHANG, X., WONG, S. E. & LIGHTSTONE, F. C. 2014. Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized Born surface area rescoring to improve enrichment. ACS Publications.
ZHOU, T. & CAFLISCH, A. J. C. 2010. High-throughput virtual screening using quantum mechanical probes: discovery of selective kinase inhibitors. 5**,** 1007-1014.

Chapter 2

Crystallisation of *Schistosoma Bovis* 28 kDa Glutathione Transferase

Abstract

Glutathione S-transferases (GSTs) are potential therapeutic drug and vaccine targets against helminthic diseases and schistosomiasis. These essential enzymes prevent oxidative stress by inhibiting toxin build-up in the worm during animal host infection. *Schistosoma* infections affect farmers economically as the infection increases susceptibility to other infections, often leading to death. To improve upon computational methods of protein selection in molecular simulations, it is important to investigate whether different protein space groups bring about the selection of different ligands due to varying protein symmetries and packing within the system. In this study, the recombinant expression, purification, crystallization, X-ray diffraction and structure resolution of Sb28GST from *Schistosoma bovis* is described. The recombinant Sb28GST was purified by Immobilized metal affinity chromatography (IMAC). Diffraction quality crystals of Sb28GST were obtained by the hanging-drop vapor diffusion method with the presence of a high concentration of ammonium sulfate (2.1 M) in the reservoir solution. X-ray diffraction data were collected using a Bruker D8 Venture advance X-ray diffractometer with a Bio PHOTON III area detector. The Sb28GST crystal was diffracted up to 2.4 Å resolution. The crystals belonged to the orthorhombic space group C 2 2 2_1 with one whole biological molecule (dimer form) occupying the asymmetric unit with the unit-cell parameters are a = 73.796 Å, b = 77.57 Å, c = 77.185 Å and β=90° which is different to the monoclinic Sb28GST (8ALS) published.

2.1. Introduction

Schistosomiasis is a devastating neglected, tropical, and poverty-associated disease caused by trematode parasites from the *Schistosoma* genus (Utzinger et al., 2009). The disease is responsible for the debilitating health conditions of over 240 million infected people, with as many as 732 million more exposed to possible infection. The zoonotic nature of *Schistosoma* extends its burden of disease to domestic and wild animals alike from contaminated water sources (Adenowo et al., 2015). The prevalence of animal schistosomiasis particularly on cattle severely impacts global attempts to sustainably supply the growing demands of cattle meat and food products (Strydom et al., 2023). A schistosomiasis infection causes anaemia, emaciation, haemorrhagic enteritis and death in some cases. This drastically impacts milk yields, feed efficacy, carcass weight and composition while servery crippling food availability during crop failure (Charlier et al., 2020). This in turn leads to major financial setbacks for local farmers and greater economic losses on return of investments for cattle producers (Strydom et al., 2023).

Bovine schistosomiasis is caused by *Schistosoma bovis* whose life cycle is maintained in the bovine host and freshwater snail intermediate hosts. However, *Schistosoma bovis's* close immunological and phylogenetic relation to *Schistosoma haematobium* is of great concern owing to the increasing number of reported inter-species mating occurrences (de la Torre-Escudero et al., 2017). The growing prevalence of *Schistosoma bovis–haematobium* hybrids is assumed to alter both disease transmittance and distribution to both animal and bovine hosts alike. The threat that these hybrids pose is amplified by reports of increased virulence and increased intermediate host target, making the need for *schistosoma* intervention of vital importance (Angora et al., 2020).

The current treatment available against schistosomiasis is Praziquantel which is unfortunately exclusively efficacious to adult egg producing *Schistosomes* and fails to target each stage of the parasite's life cycle (Olveda et al., 2013). Much like many helminths parasites, *S. bovis'* possess the detoxification enzyme GST as their sole defence mechanism in the absence of vital phase I detoxification enzymes. Glutathione transferases possess ligandin capabilities that allow them to catalyse the nucleophilic conjugation of reduced glutathione to a variety of electrophiles, making them less harmful, more soluble compounds (Angelucci et al., 2005). Their

overexpression during chemotherapeutic interventions coupled with their presence in all stages of the parasite's life cycle, make them a promising target for anti-schistosomal drug development (Pérez-Sánchez et al., 2006, Sheehan et al., 2001).

Determining the selection of drugs to screen as possible inhibitors against specific enzymes has become an easy task. This is so due to the growing popularity of using high power atomic level computers that can simulate relevant biomolecule interactions (Durrant and McCammon, 2011). These computational simulations can predict the movement of every atom on a protein over a period using Newtonian physics to approximate atomic motion. The computational simulations are used in combination with X-ray crystallography as models build the simulation system around (Hollingsworth and Dror, 2018, Hospital et al., 2015). The use of X-ray crystallography is a powerful, reliable technique used to provide very detailed structural information to glean molecular mechanisms that are fundamental for the functioning of biomolecules. This technique however, is bottlenecked at the crystallisation step as the growth of protein crystals is a nuanced rate limiting step in structure determination (McPherson and Gavira, 2014). The stages of crystallization are nucleation where appropriate conditions allow for proteins in solution to form thermodynamically stable aggregates, crystal growth which is governed by principles of diffusion aiding in the ordered assembly of the nuclei and cessation of growth occurring once protein molecules are depleted from solution (Krauss et al., 2013). In some occurrence, proteins can grow in two or more different crystal states in a phenomenon called polymorphism. While this occurrence was once referred to as “Nemesis of crystal engineering”, it is now considered to be of grave significance in understanding crystal packing and the mechanisms of crystallization (Desiraju, 2010). Proteins that exhibit crystal polymorphism occupy different space groups as they are packed differently with varying symmetries within the unit cell they are packed in (Chruszcz et al., 2008). Whether this has any influence over the interaction with other molecules introduced into this system and thus affecting ligand recognition in computational simulations such as high throughput virtual screening is not known. In this work we aim to answer this question by obtaining the Sb28GST in a different space. Here, we describe the recombinant expression, purification, crystallization, and structure resolution of Sb28GST in an orthorhombic.

2.2. Materials and Methods

2.2.1. Materials

All chemicals and reagents used were analytical grade. Vector synthesis was carried out by GenScript (NJ, USA). The *Escherichia coli* (*E. coli*) T7 Express chemically competent cells were purchased from New England Biolabs (USA). Tryptone, yeast extract, NaCl, ampicillin, N,N,N',N'- tetramethylethylenediamine (TEMED), bisacrylamide, acrylamide, sodium dodecyl sulfate (SDS), β-mercarptoethanol, glycine, isopropyl-β-D-thiogalactoside (IPTG), reduced glutathione (GSH), 1-chloro-2,4-dinitrobenzene (CDNB), sodium phosphate, glacial acetic acid and coomassie Blue-G250 dye were all purchased from Sigma Aldrich (MO, USA). Thermo Scientific™ PageRuler™ Plus Prestained Protein Ladder and Snakeskin™ dialysis tubing (10K MWCO, 22 mm) were purchased from Thermo Fisher Scientific (MA, USA). The crystal buffers used for screening, reagents and other apparatus were purchased from Hampton research (CA, USA). The 15-well EasyXtal® plates were purchased from Qiagen (VVV, NL).

2.2.2. Methods

2.2.2.1. Recombinant Expression of Sb28GST

The recombinant expression of a target gene housed in a pET11a (+) vector, shown in Figure 2.1, is typically controlled by the *T7* promoter. However, a repressor protein encoded by the *lacI* gene typically present in the genomic and plasmid DNA disables both the T7 and *lac* promoters necessary for the expression of T7 RNA polymerase. This is circumvented by deactivating the repressor protein by way of culturing T7 Express *E. coli* cells in the presence of IPTG. As a result, the RNA polymerase freely expresses the T7 RNA polymerase which binds to the T7 promoter to induce the expression of the target protein (Smith and Johnson, 1988).

Uniport KB database was used to obtain the nucleotide sequence of Sb28GST (accession: M87800), following the addition of an N-terminal hexahistidine tag, the sequence was cloned into the multiple cloning site of the pET11a (+) vector between the restriction enzymes Bamh1 and Nde1 with ampicillin resistance (Genscript, USA). The *E. coli* T7 Express cells with the described vector were inoculated into a 2× YT broth supplemented with 100 µg/mL ampicillin. The cells were incubated for 16 h at

30°C and 230 revolutions per minute (rpm). The overnight cultures were then diluted in 1:100 ratio into fresh 2× YT broth supplemented with 100 µg/mL ampicillin and incubated once more (37°C at 200 rpm). Once the photographic density of the cells reached 0.4 – 0.6 at 600 nm, representing the mid-log phase, the cells were cold shocked in an ice bath for 30 min. Thereafter, the cells were induced with IPTG with a final concentration of 0.5 mM and further incubated for 6 h (30°C at 200 rpm). The cell contents were harvested through centrifugation at 7 500 × *g*, 4 °C for 10 min. The supernatant containing the media broth was discarded and the cell pellet was stored at -20 °C.

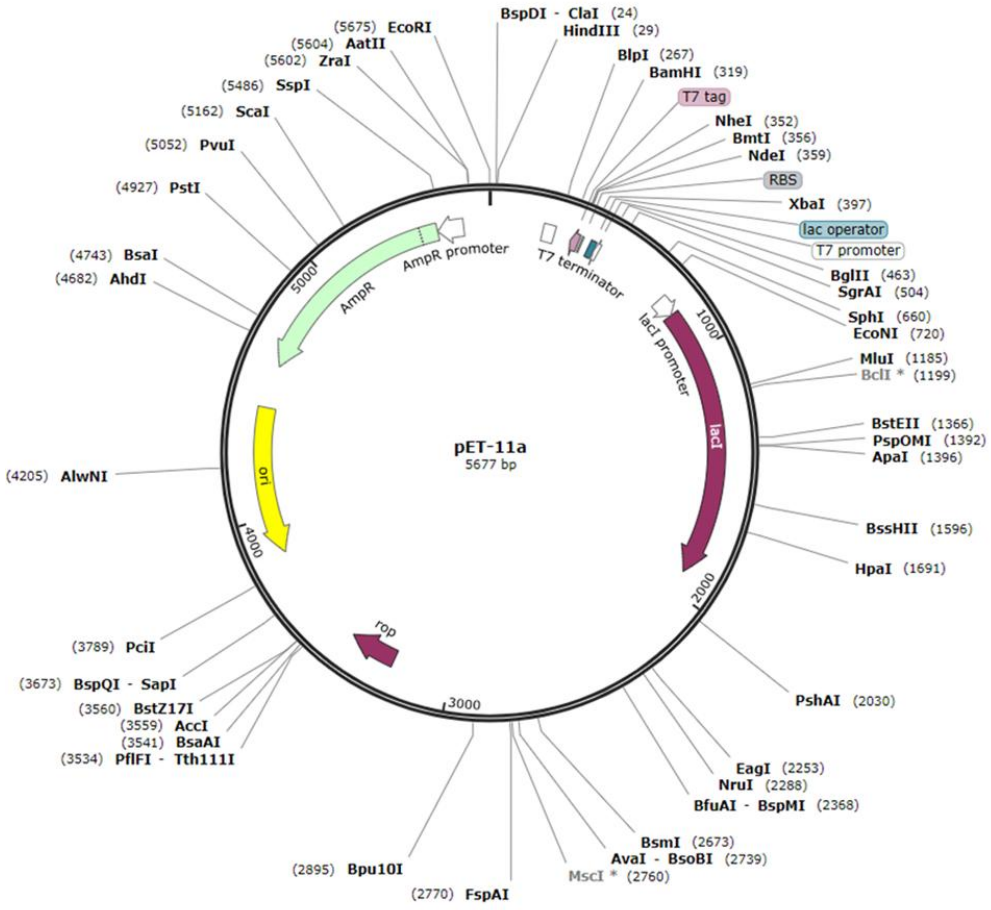


Figure 2. 1: Vector map of pET-11a

2.2.2.2. Purification of Sb28GST Protein

Obtaining a pure protein requires its isolation from numerous bacterial cell components expressed alongside the protein of interest. Appropriate fractionation techniques need to be utilized to successfully isolate the target protein. An effective isolation technique used for the purification of biomolecules is affinity chromatography.

The governing principles of affinity chromatography rely on the exploitation of biomolecules to bind to specific immobilized ligands to which they have a high affinity for. This allows the miscellaneous bacterial proteins to be eluted from the chromatography system. Thereafter, the bound protein is then eluted alone using an elution buffer that disrupts the interactions holding the protein bound to the ligand. This disruption can be achieved by the addition of a competitive ligand or a buffer that introduces a change in the pH or ionic strength (Labrou et al., 2014).

To isolate Sb28GST, the harvested cells were thawed at room temperature. The cell pellet was resuspended in equilibration-wash buffer containing phosphate-buffered saline (PBS), 25 mM imidazole, 0.02% NaN_3, pH 7.2 and mixed using an end over end rotor for 1 h. The cells were then lysed on ice through sonication using an ultrasonic processor (in 30 s intervals × 10) at 70 amps. The cells were then centrifuged at 18 000 × *g* for 30 min at 4 °C. The resulting supernatant containing the soluble cell fraction was filtered through a 0.45 µM cellulose acetate membrane filter to avoid clogging. The Sb28GST protein was the purified using immobilized metal affinity chromatography (IMAC). The filtered lysate was passed through an equilibrated HisTrap™ Fast Flow column (Sigma Aldrich). The bacterial cell components were washed using the equilibration-wash buffer. Thereafter, the desired protein was eluted using an elution buffer containing PBS, 500 mM imidazole, 0.02% NaN_3, pH 7.2. The purity and homogeneity of the purified protein were determined by Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE). The eluents were subsequently pooled and dialyzed against a PBS pH 7.4 buffer to remove the imidazole contents from the pure Sb28GST solution.

2.2.2.3. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

A popular and efficient tool used for the analysis of proteins is SDS-PAGE. This technique allows for the characterization of proteins based on their migration in the presence of an electric field. The SDS used in this technique is an anionic detergent that causes the disruption of both the secondary and non-disulfide linked tertiary protein structures while additionally masking their intrinsic charge. This in turn creates a uniform charge: mass ratio among the proteins treated with SDS to allow the migratory distance travelled to be governed solely by their molecular weights (Walker and Wilson, 2010).

The purity and homogeneity of the Sb28GST was assessed using 12.5% SDS-PAGE as described by (Laemmli, 1970). Samples to be analysed were prepared from the cell lysate, unbound protein flow through, and the imidazole induced affinity exchanged eluents. These samples were mixed in a 1:2 ratio with a reducing treatment buffer (125 mM Tris HCl, 4 % (w/v) SDS, 20 % (v/v) glycerol, 10 % (v/v) 2-mercaptoethanol) and subsequently boiled for 5 min to achieve denaturation. Thereafter, the samples were cooled on ice prior to loading into the gel. The gels were connected to a BIO-RAD Powerpac™ followed by protein separation at 40 V, 18 mA per gel in tank buffer (250 mM Tris-HCl, 192 mM glycine, 0.1 (w/v) SDS, pH 8.3). The gels were then stained overnight with staining solution (0.125 % (w/v) Coomassie blue R-250, 50 % (v/v) methanol, 10 % (v/v) acetic acid. They were then destained overnight with destaining solution (50 % (v/v) methanol, 10 % (v/v) acetic acid) until background was clear to capture the resulting gel.

2.2.2.4. Protein Concentration Determination

Proteins possess certain molecular features that allow them to absorb ultraviolet (UV) light. This ability is exploited to aid in determining the concentration of proteins using UV spectroscopy. Specifically, the peptide bonds, tryptophan, tyrosine and phenylalanine aromatic amino acids and to a lesser extent, disulfide bonds are the molecular components that allow proteins to absorb light at 280 nm (Walker and Wilson, 2010).

Following a 1: 50 dilution of the purified protein into PBS, the concentration and purity of Sb28GST was determined spectrophotometrically using the Jasco V-630 UV-VIS spectrophotometer (Jasco Inc., Tokyo, Japan). To determine protein purity, the diluted protein was scanned through a UV absorption spectrum monitored between 240 and 350 nm. For protein concentration, the diluted protein was additionally serial diluted in 1:1 ratios with a final dilution factor of 0.00125. The concentration of the protein was determined using the Beer-Lamberts Law (Swinehart, 1962):

$$A = \varepsilon_{\lambda} c l \qquad \textbf{Equation 1}$$

where A represents the absorbance in arbitrary units (AU), ε_{λ} represents the molar extinction coefficient ($M^{-1}.cm^{-1}$) at a particular wavelength ($\lambda = 280\ nm$), c represents the molecule of interest's concentration (M) and l represents the pathlength of the light through the solution (cm). The ε_{280} of Sb28GST was calculated in accordance to

Perkins (1986), combining the sum of the number of absorbing chromophores and their respective extinction coefficients demonstrated in equation 2:

$$\varepsilon_{280}(Sb28GST) = 5550\sum Trp + 1490\sum Tyr + 150\sum Cys \qquad \textbf{Equation 2}$$

$$\varepsilon_{280}(Sb28GST) = 5550(2) + 1490(8) + 150(1)$$

$$\varepsilon_{280}(Sb28GST) = 23045\ M^{-1}cm^{-1} \text{ per monomer}$$

2.2.2.5. Specific Activity Assay

Monitoring enzyme catalysed reactions is typically done by observing the concentration of the enzymatic substrate or product over a period of time. This exploits the unique spectroscopic properties of chromophoric compounds during reactions where structural changes can occur (Sheehan, 2009). The unique chromophoric products can be monitored at specific absorbance spectra using visible absorbance spectroscopy. A standard GSH-CDNB conjugation assay is monitored in a similar fashion, where GSTs catalyse the conjugation of GSH with the substrate CDNB to produce 1-(S-glutathionyl)-2, 4-dinitrobenzene shown in Figure 2.2. The concentration changes of this chromophore which absorbs light at 340 nm can then be monitored using the Beer-Lambert Law to determine the activity of the GST enzyme, allowing real-time monitoring of the enzymatic reaction through spectrophotometric analysis (Habig et al., 1974). The rate of increase in absorbance at 340 nm is directly proportional to GST activity, thus providing a quantitative measure of enzyme function. CDNB is particularly suited for this assay because it is a broadly reactive substrate across various GST isoforms, readily available, and chemically stable under standard assay conditions. Given its reliability, sensitivity, and adaptability, the GSH-CDNB conjugation assay remains a fundamental tool in biochemical, toxicological, and pharmacological research for studying GST-mediated detoxification mechanisms.

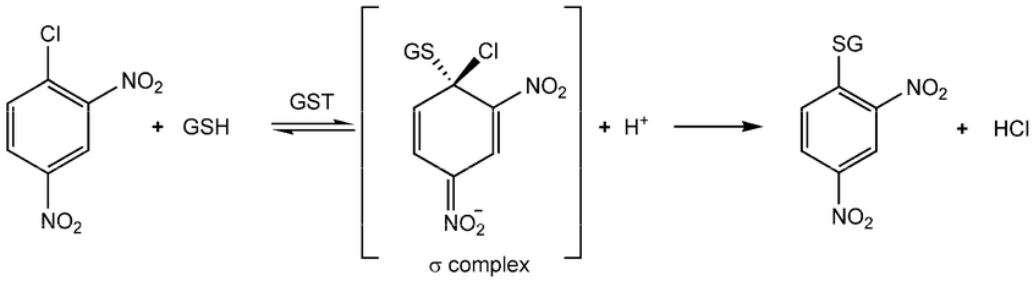


Figure 2. 2: Glutathione conjugation to 1-chloro-2,4-dinitrobenzene in a GST catalysed reaction

Purified recombinant Sb28GST was dialysed into 100 mM sodium phosphate, pH 6.5. Thereafter, the catalytic activity of Sb28GST was spectrophotometrically monitored using the CDNB-GSH assay as previously described (Zhan et al., 2010). Varying concentrations of Sb28GST (0 - 500 nM) were prepared in the reaction buffer, which contained 100 mM sodium phosphate, 1 mM EDTA, 5 mM DTT, 0.02% NaN_3, 5 mM GSH, 1 mM CDNB (3% (v/v) ethanol), pH 6.5. The formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene (ε_{340} = 9600 $M^{-1} \cdot cm^{-1}$) was monitored using a Jasco V–630 UV/Vis spectrophotometer (Jasco, Japan). Triplicates of the reactions were carried out for 60 s. The specific activity was obtained by plotting linear progress curves corrected for non-enzymatic reaction rates.

2.2.2.6. Protein Crystallography and Structure Determination

To elucidate three-dimensional structure of biomolecules, X-ray crystallography is popularly used to accurately determine structural details at atomic resolution. However, protein crystals are required to obtain this structure. The growth of protein crystals is a formidable task requiring specific supersaturation conditions that promote the slow precipitation of proteins in a thermodynamically stable state. The result is the periodic assembly of protein molecules in a three-dimensional crystal lattice held together by non-covalent bonds (Krauss et al., 2013).

A tried-and-true method used to achieve supersaturation is vapour diffusion. This method uses the diffusion of water between small protein droplets containing buffer with precipitants and the reservoir with the same buffer components. Sealing the reservoir creates a vacuum allowing the protein droplet to equilibrate over the reservoir either in a sitting or hanging position. This creates an environment where water diffuses from the droplet into the reservoir, leading to the slow increase of protein

concentration, thus causing supersaturation and consequently, crystal growth (Benvenuti and Mangani, 2007).

Once protein crystals are successfully grown X-ray diffraction is employed where incident X-ray beams are emitted into a single mounted crystal. The resulting diffraction pattern is the reciprocal of the protein crystal lattice. The intensities of the diffracted reflections are calculated to determine the distribution of electron density which is related to the atomic positions within a unit cell. This data is indexed, integrated, scaled and merged to build a model into the electron density. The model is finally refined and validated through sophisticated algorithms to obtain the crystal structure of the protein of interest (Ilari et al., 2008).

2.2.2.6.1. Crystallization of Sb28GST

Recombinant Sb28GST was concentrated to 15 mg/ml in PBS at 7.5 pH prior to crystallization trials. The crystallization experiments were performed at 298 K using hanging and sitting-drop vapor-diffusion methods. Initially, the crystallization conditions were screened using commercially available kits such as Index™ Screens I and II (Hampton Research, USA) through a sitting vapour diffusion method on a 96 well microplate using the Oryx8™ protein crystallisation robot (Doughlas Instruments, East Garston, UK). Additionally, the previously reported condition used by (Johnson et al., 2003) was added into the screening process. Apo Sb28GST crystals were produced through the hanging drop vapor-diffusion method using the Qiagen EasyXtal™ 15-well plates. Diffraction quality crystals were obtained using 2 µL drop of pure protein solution and 2 µL drop of the reservoir buffer consisting of 2.1 M ammonium sulfate, 100 mM Tris (pH 7.5), and 5 mM β-mercaptoethanol. Long needle-shaped crystals of apo Sb28GST grew within 2 - 4 days at 20°C.

2.2.2.6.2. X-ray Diffraction, Data Collection and Structure Solution

A single Sb28GST crystal was mounted in a 0.1 - 0.2 mm nylon loop and soaked in a cryoprotectant paratone solution (Parabar 10312; Hampton Research). This was flash frozen in liquid nitrogen at 100 K to deter radiation damage from the X-ray beam. The X-ray diffraction data for Sb28GST was collected using a rotating anode X-ray source (Cu Kα; 1.5418Å) and the pixel Bruker D8 Venture Bio PHOTON III area detector diffractometer with a crystal-to-detector distance of 41 mm. The unit cell and full data was collected using the PROTEUM^4 software suite (Evans, 2006). The data was

integrated with SAINT and SADABS was used to reduce and scale the data. SCALEPACK and AIMLESS/POINTLESS, which are built into the PROTEUM[4] suite, were used to process the diffraction data (Evans, 2011). The structure was determined by way of molecular replacement using the PHASER program built into PHENIX software (McCoy et al., 2007). The search model used for molecular replacement was 1OE7. Model building was performed with the WinCoot (Emsley et al., 2010). Water molecules were added during refinement using the "Flat Bulk Solvent Model" procedure on Coot. The protein structure was further refined using PHENIX, followed by structural validation using PROCHECK (Afonine et al., 2012, Laskowski et al., 1993). To generate the images, PyMOL was used. The resulting structure was deposited into the PDB under the code 8BHZ. To interpret the flexibility and disorder of different regions of the resolved Sb28GST molecule, normalised B-factors which are statistically adjusted B-factor values (with respect to the standard deviation around the mean value) were calculated using the following equation:

$$Bnorm = \frac{(B-\bar{B})}{\sigma B} \qquad \textbf{Equation 3}$$

The protein files were submitted into UCSF Chimera in PDB format to compute the normalised B-factors.

2.3. Results

2.3.1. Expression and Purification

To obtain a pure Sb28GST protein for downstream studies, the *E. coli* T7 Express cells transformed with the target protein vector were expressed at 30°C for 6 h, induced with 0.5 mM IPTG. The cellular fraction was harvested from the 2 × YT media upon recombinant expression yielding a 5.2 g/L wet cell pellet. The Sb28GST protein was present in the soluble fraction of the harvested *E. coli* contents and subsequently filtered to avoid clogging the purification column. The Sb28GST protein was then purified from the homogenised soluble fraction of harvested *E. coli* using Ni^{2+} affinity chromatography. The lysed cells were injected into the purification column allowing the hexa-histidine tagged protein to bind to Ni^{2+} while the unbound proteins were eluted in the flow-through as shown in Figure 2.3A. The protein was eluted through single step elution using 500 mM imidazole.

The purity of the eluted Sb28GST protein was analysed using SDS-PAGE. As seen in Figure 2.3B, the target protein successfully bound to the column while the unbound proteins flowed through the column. There was no loss of the target protein during this step owing to its absence in the flow-through. The protein eluted due to the addition of imidazole was seen to be pure, resulting in the successful purification of a 24 kDa Sb28GST protein.

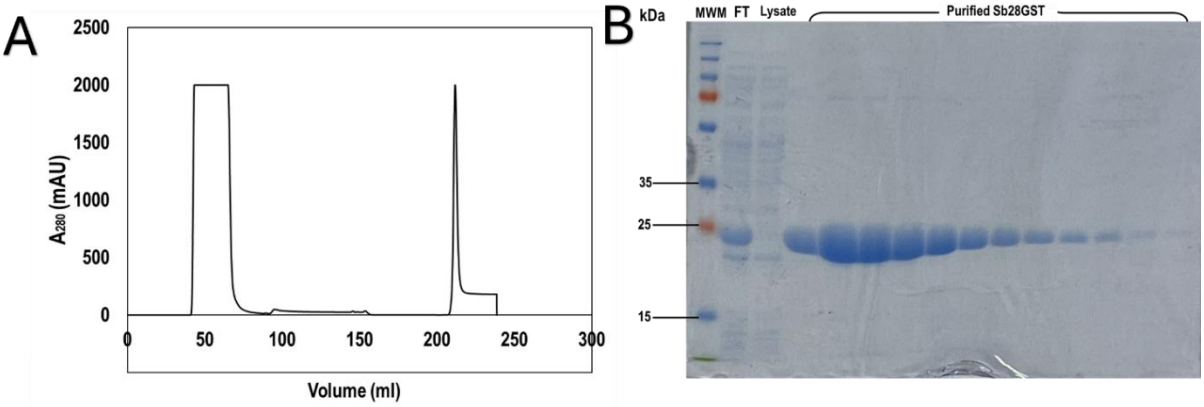


Figure 2. 3: Expression and purification profile of Sb28GST. A: The elution profile of Sb28GST purified using a Ni2+ affinity column eluted with 500 mM imidazole. B: A 12.5 % (w/v) polyacrylamide gel with the molecular weight marker (MWM), sonicated cell lysate, the flow through (FT) and eluents collected from the 500 mM imidazole elution. The monomer size of Sb28GST was 24 kDa.

2.3.2. Protein Concentration Determination

The purity and concentration of eluted Sb28GST was verified spectroscopically by way of UV-Vis absorbance. Analysis of Sb28GST purity across a UV absorption spectrum of 240 to 340 nm showed a peak at ~280 nm. The absorbance signal observed thereafter dropped approaching 0 AU at 340 nm as depicted in Figure 2.4A. There were no additional peaks observed at 320 and 340 nm, indicating the absence of protein aggregates and other UV absorbing impurities or unwanted dust particles. The protein concentration of Sb28GST was determined to be 18.37 mg/mL using a linear regression slope shown in Figure 2.4B.

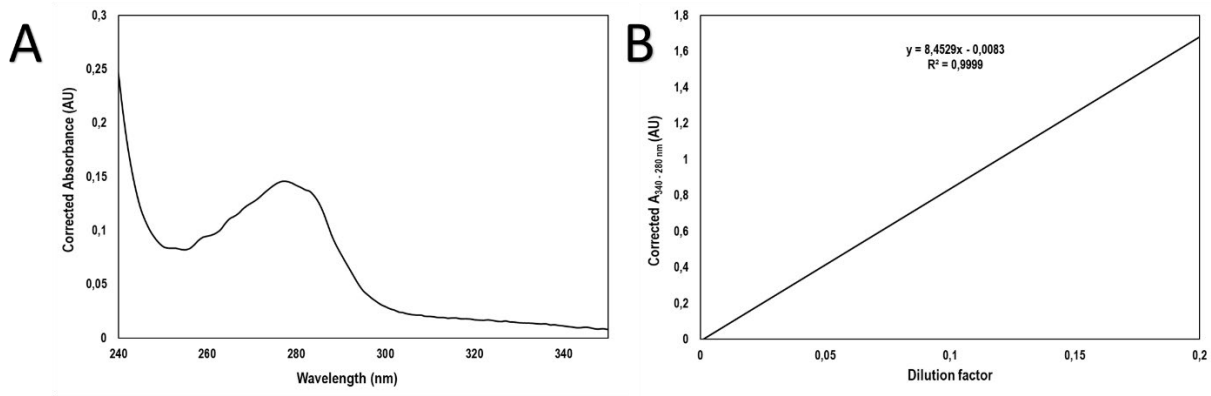


Figure 2. 4: Determination of Sb28GST purity and concentration. A: UV Scan of IMAC purified Sb28GST within a range of 240 - 340 nm. B: Concentration determination of purified Sb28GST. The purified protein was initially diluted (1:50) followed by the scan being taken across a 240-340 nm range to determine any impurities. The highest peak observed was at 277 nm and no other peaks were seen thereafter. The slope of linear regression for Sb28GST was $\boldsymbol{y = 8.4529x - 0.0083}$ resulting in a Sb28GST concentration of 18.37 mg/mL.

2.3.3. Specific Activity Assay

To ensure that the recombinant Sb28GST with which crystallisation studies would be performed is catalytically active, the specific activity of the protein was determined using the CDNB-GSH conjugation assay. As shown in Figure 2.5, the specific activity of Sb28GST was determined to be 13.3150 μmol.min^{-1}.mg^{-1}, confirming that the enzyme was not inactive.

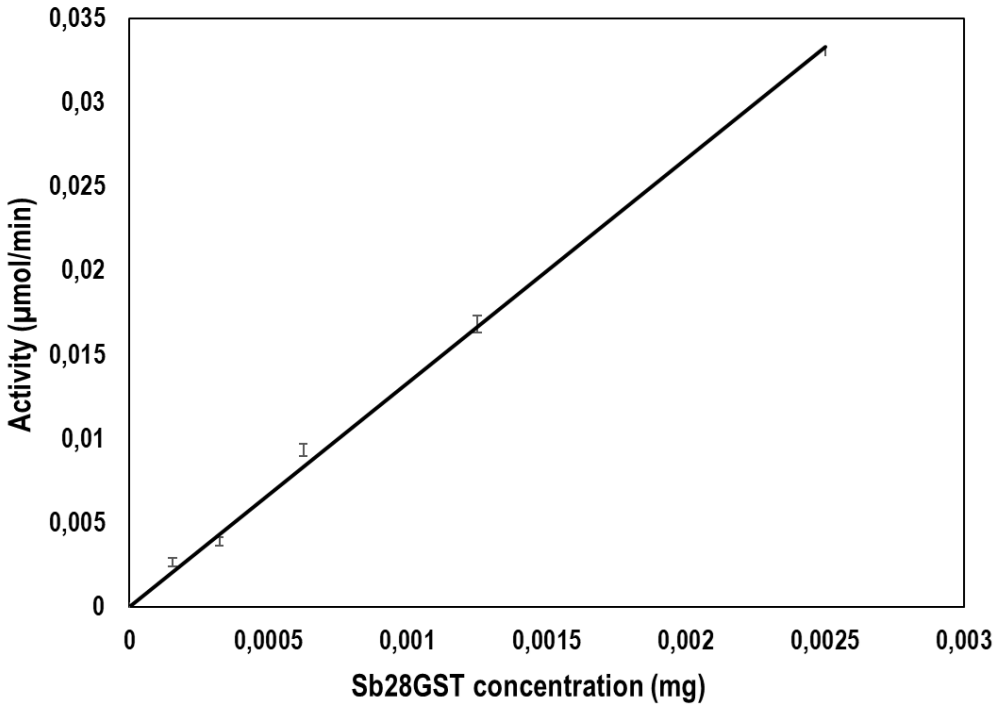


Figure 2. 5: Specific activity profile of Sb28GST. The GSH-CDNB conjugation assay was performed at 340 nm by monitoring the formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene catalysed by Sb28GST. Varying concentrations of Sb28GST (0 – 500 nM) were used to determine the specific activity of the enzyme. All assays were performed in triplicate, with the data corrected for non-enzymatic rates.

2.3.4. X-ray Crystallography

At the time of this study there was one Sb28GST crystal structure deposited into PDB with the code 8ALS belonging to a monoclinic P 1 2_1 1 space group. To determine whether Sb28GST could be crystallized in a different crystal space group, various crystal buffers and conditions were tested. Crystals of varying morphologies grew within 2 - 4 days however, the crystals obtained were not diffraction quality and were often small and brittle if not cloudy as seen in Figure 2.6.

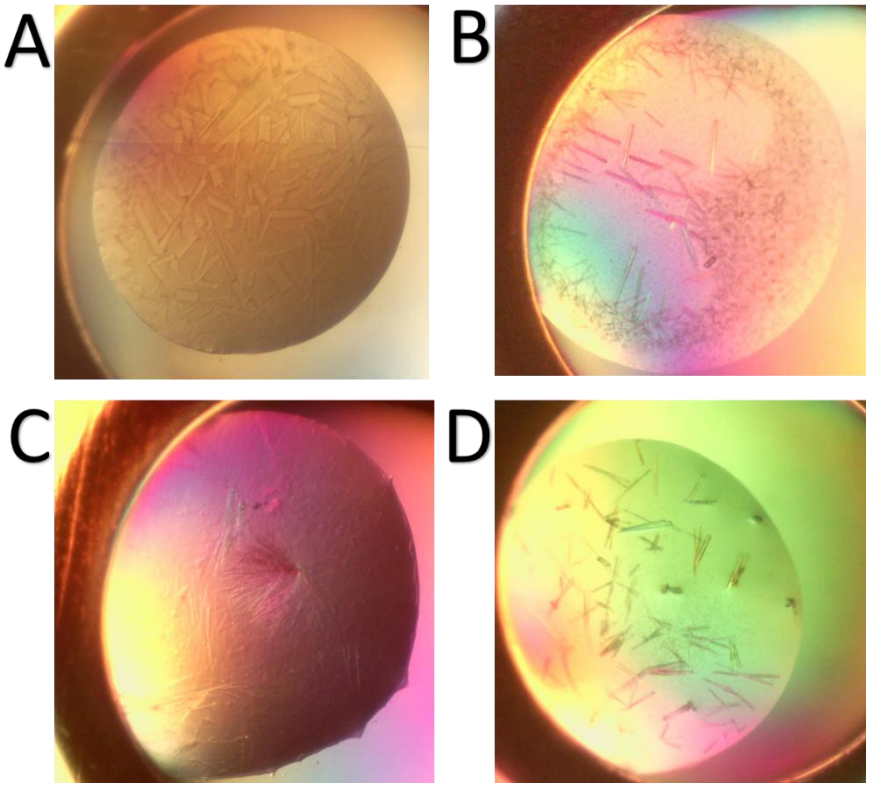


Figure 2. 6: Crystal trials of Sb28GST grown at 20°C. The crystal trial conditions that allowed 10 mg/mL of Sb28GST to precipitate in a 1:1 ratio with the following reservoir conditions A: 0.1 M Citric acid pH 3.5, 2.0 M Ammonium sulfate. B: 0.1 M Sodium acetate trihydrate pH 4.5, 25% w/v Polyethylene glycol 3,35. C: 27. 2.4 M Sodium malonate pH 7.0. D: 0.1 M Potassium thiocyanate, 30% w/v Polyethylene glycol monomethyl ether 2,000.

The condition that allowed satisfactory Sb28GST crystals to grow consisted of 2.1 M ammonium sulfate, 100 mM Tris (pH 7.5), and 5 mM β-mercaptoethanol. The protein crystals grew within two days by hanging drop vapour diffusion and had a needle-like morphology as shown in Figure 2.7 A. The crystals were harvested, mounted on a nylon loop and coated in a paratone solution prior to being flash frozen with liquid nitrogen at 100 K. The measures were taken to protect the Sb28GST crystal from radiation damage. A rotating anode X-ray source utilizing Cu Kα radiation was used to diffract the protein crystal. The data was subsequently collected using a Bruker D8 Venture Bio PHOTON III area detector diffractometer. The raw diffraction data was reduced to a consistent dataset and scaled across multiple frames using the SADABS program, integrated into the PROTEUM[4] software suite. This was carried out with the multi scan absorption correction algorithm, which effectively corrected for absorption effects presented as noise in the measured intensities for each reflection, corrupted by the crystal and paratone cryoprotectant. The multi scan algorithm successfully

aided in the improvement of data consistency by systematically minimising additional errors in the raw data which were caused by radiation decay and detector geometry. Data processing showed that the Sb28GST crystal unit cell had dimensions of $\alpha = \beta = \gamma = 90^{\circ}$, which is characteristic of orthorhombic crystal systems. The geometrical characteristics of the Sb28GST molecules adopted a C 2 2 2_1 space group. The C 2 2 2_1 space group crystals were thinner and more clustered than the previously described 8ALS crystals (Figure 2.7 B). The structure of Sb28GST was determined by molecular replacement using the coordinate file of 1OE7. The data collection and refinement statistics comparing the two crystals were compiled in Table 2.1. The Sb28GST crystal diffracted to a resolution of 2.4 Å with the structure refined to a final R_{factor} of 19.8%. The solvent content and Matthew's coefficient (V_M) were 46.55% and 2.33 $Å^3$ Da^{-1} respectively. The 3D structure was visualized on PyMol as shown in Figure 2.8. The structural alignment and superimposition of the two structures give a root mean square deviation (RMSD) values of 1.664 Å.

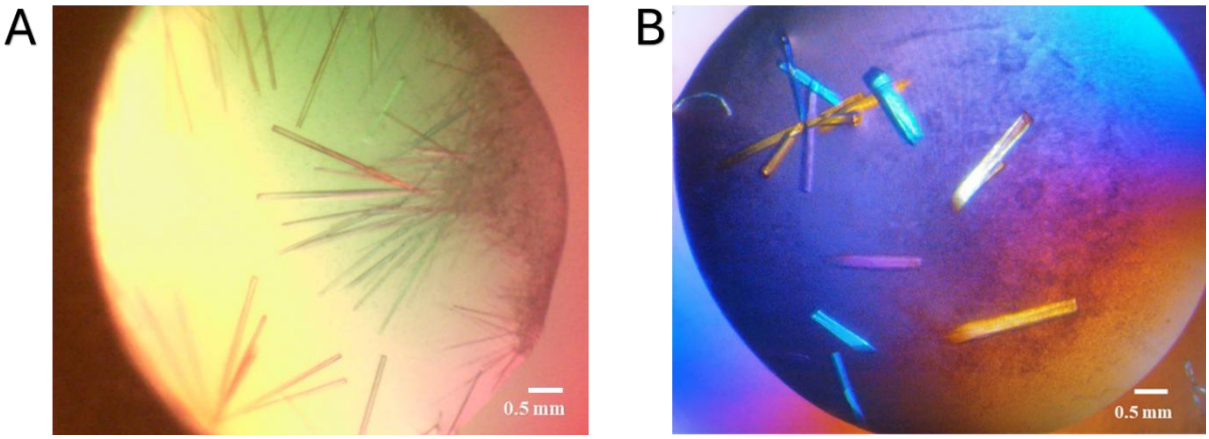


Figure 2. 7: Sb28GST crystals grown. The crystals were grown by hanging drop method at 20˚C. A: Needle like crystals grown in the presence of PBS deposited as 8BHZ. B: Thick needle like crystals grown in the presence of 25 mM NaH_2PO_4 buffer deposited at 8ALS.

Table 2. 1: Data collection and processing statistics

	Orthorhombic	**Monoclinic**
Light source	Cu Kα	Cu Kα
Wavelength (Å)	1.5418	1.5418
Detector	PIXEL Bruker PHOTON III	Bruker Venture Bio PHOTON III area detector
Temperature (K)	100	100
Space group	C 2 2 2_1	P 1 2_1 1
Unit-cell parameters (Å)		
*a, b, c (*Å)	73.796, 77.57, 77.185, 90	
β (°)		54.133, 77.086, 54.015, 93.15
Resolution range (Å)	24.4- 2.4 (2.49-2.4)*	24.68 -2.30 (2.38-2.30)*
No. of observed reflections	8849	19582
No. of unique reflections	8867	19612
Completeness (%)	99.2 (93.8)*	98.8 (89.9)*
I/σ(I)	8.6 (2.7)*	11.4 (3.4)*
R_{merge}¥	0.093 (0.329)*	0.132 (0.600)*
Multiplicity	3.8 (3.2)*	8.1 (6.1)*
Structure refinement		
Reflection used	8849	19.582
Resolution range	24.4- 2.4	24.68–2.30
Final overall Rfactor (%)	19.88	21.68
Rework (%) / ‡Rfree (%)	19.55/ 25.82	21.52 / 24.48
Number of protein atoms	1635	3274
Average B factor value (Å^2)	25	24.0
RMSD in bond length (Å)	0.0013	0.002
RMSD in bond angles (◦)	0.4277	0.506
Ramachandran statistics		
Favoured; Allowed; Outliers (%)	97.00, 3.00, 0.00	96.00, 4.00, 0.00
Matthew's coefficient V_M (Å^3 Da^{-1})	2.33	2.36
Solvent content (%)	46.55	47.98

*Values in parentheses are for the highest resolution shell.

¥$R_{merge} = \sum_{hkl}\sum_i |\, I_i\,(hkl) - \langle I\,(hkl)\rangle \,| / \sum_{hkl}\sum_i I_i\,(hkl)$, where $I_i\,(hkl)$ is the intensity of the observed reflection and $\langle I\,(hkl)\rangle$ is the mean intensity of the reflection

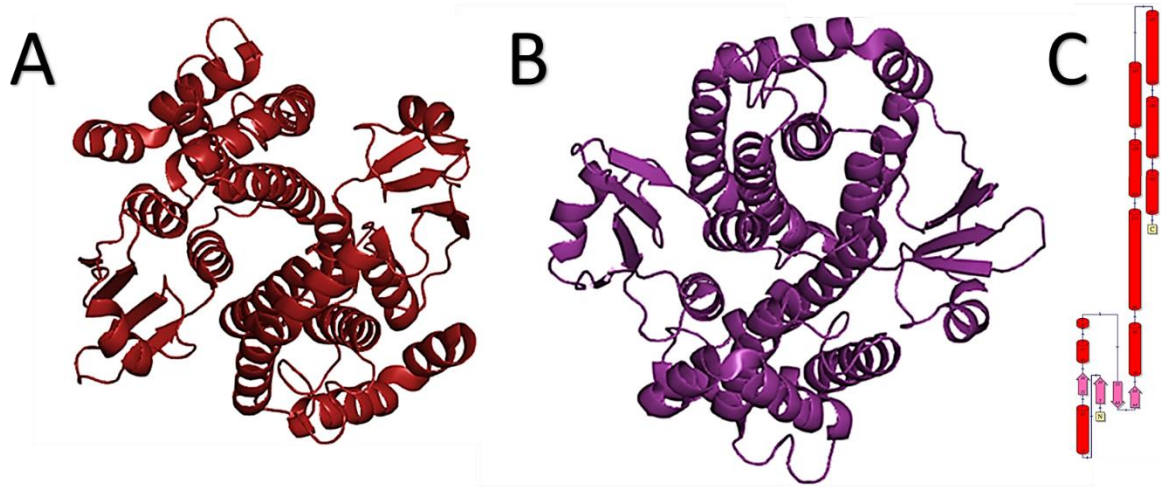


Figure 2. 8: Dimer ribbon representation of Sb28GST with the protein topology. A: The Sb28GST dimer of 8BHZ. B: The Sb28GST dimer of 8ALS. C: Topology of Sb28ST showing the secondary structure characteristics. The ribbon diagrams were generated using PyMol molecular graphics program.

2.4. Discussion

Crystal polymorphism is a regularly occurring phenomenon that can be induced in many ways (Desiraju, 2010). One of which being the manipulation of the crystallisation conditions by way of pH, temperature and precipitating agents' concentration (Krauss et al., 2013). To investigate this hypothesis, pure protein is required to grow high quality crystals. To this end, Sb28GST was recombinantly expressed under the conditions previously reported by (Makumbe et al., 2024), using 0.5 mM IPTG, for a 6 h induction at 30 °C. Prior to induction, the *E. coli* T7 Express cells were cold shocked in an ice bath. This aids in the reduction of metabolic stress on the host cells in order to reduce the risk of aggregation and protein misfolding, yielding amounts of soluble protein as seen in Figure 2.3 and 2.4 (Qing et al., 2004).

The Sb28GST protein was then successfully purified using a nickel affinity column. The absence of the target protein in the flow-through shown in Figure 2.3 indicating the successful binding of N-terminal hexa-histidine tagged Sb28GST to the Ni^{2+} chelator resin column. This was achieved by the selective coordinate covalent binding between Ni^{2+} and the imidazole ring of the hexa-histidine tag (Block et al., 2009). The purified protein was confirmed to be 24 kDa in size which corresponds to current literature (Makumbe et al., 2024). However, a cross-comparison with the net isotopic mass of the amino acids shows that the theoretical molecular mass of the protein is

23.4 kDa. The discrepancy is owed to the N-terminal addition of a hexa-histidine tag which accounts for approximately 0.8 kDa (Sørensen and Mortensen, 2005). The enzyme was additionally catalytically active, with the specific activity recorded being comparable to other Sb28GST enzymes recorded (Makumbe et al., 2024).

While the precipitation of proteins into high quality crystals is a gruelling endeavour, the conditions that lead to satisfactory crystal growth are achieved by a very narrow range of conditions. This is corroborated by the fact that living systems are operational almost exclusively on very particular aqueous chemistry and interactions, whose deviations or perturbations are rarely tolerated (McPherson and Gavira, 2014). As such, only four crystal conditions allowed for the precipitation of Sb28GST from the Index™ Screens I and II condition kit (Hampton Research, USA). The resulting crystals were not sufficient to proceed to X-ray diffraction due to their fragility. This is because protein crystals are loosely packed when precipitating into their crystalline state. Their assembly is made up of 40 – 60 % solvent-filled holes and channels, which make them particularly fragile and vulnerable to radiation damage and disintegration when undergoing X-ray diffraction (Ilari et al., 2008). However, alteration of the condition previously described by (Johnson et al., 2003, Krauss et al., 2013, Labrou et al., 2014, Laemmli, 1970, Laskowski et al., 1993, Makumbe et al., 2024, McCoy et al., 2007, McPherson and Gavira, 2014, Olveda et al., 2013, Pérez-Sánchez et al., 2006) lead to the precipitation of Sb28GST into diffraction-quality crystals.

It has been suggested that changes in crystallization conditions such as pH or the ionic strength of the crystal environment can impact crystal packing, resulting in changes of the crystal unit cell parameters (Jacobson et al., 2002). Comparison of the conditions that brought about 8ALS to those of 8BHZ revealed small but impactful differences that induced the formation of Sb28GST in a different space group. While 8ALS was suspended in 25 mM NaH_2PO_4, 8BHZ was suspended in PBS (Majeed et al., 2003). Despite both solvents possessing a phosphate based buffering system, the addition of KCl and NaCl in PBS largely influences the solubility of Sb28GST. The utilization of both monobasic and dibasic phosphate salts modifies the surface charges and ionic strength of the protein leading to reduced electrostatic variability, unique crystal contact and an altered arrangement of the crystal lattice (Vaney et al., 2001). Additionally, 5.3 mg/mL was used to obtain 8ALS crystals, while 15 mg/mL of Sb28GST was used to obtain 8BHZ. This difference in concentration directly impacts

the nucleation rate and consequently, the packing order and compactness of the crystal lattice as seen by 8BHZ forming an orthorhombic system (Table 2.1) (Krauss et al., 2013). Lastly, the ratios used to grow the Sb28GST crystals were 1:2 and 1:1 for 8ALS and 8BHZ respectively. The lower protein: reservoir buffer dilution allows for slower protein nucleation which allows Sb28GST to adopt a more thermodynamically stable packing arrangement (Rupp, 2009). The culmination of the buffer contents, protein concentration and protein: reservoir buffer ratios factored in the formation of Sb28GST in an orthorhombic crystal space group.

Obtaining Sb28GST in an orthorhombic C 2 2 2_1 space group will allow for the further exploration of the conformational landscape of polymorphic crystals in computational studies. This will be essential in determining whether space group polymorphism affects ligand selectivity across crystal variants in Sb28GSt drug design investigations.

2.5. References

ADENOWO, A. F., OYINLOYE, B. E., OGUNYINKA, B. I. & KAPPO, A. P. J. B. J. O. I. D. 2015. Impact of human schistosomiasis in sub-Saharan Africa. 19**,** 196-205.

AFONINE, P. V., GROSSE-KUNSTLEVE, R. W., ECHOLS, N., HEADD, J. J., MORIARTY, N. W., MUSTYAKIMOV, M., TERWILLIGER, T. C., URZHUMTSEV, A., ZWART, P. H. & ADAMS, P. D. J. A. C. S. D. B. C. 2012. Towards automated crystallographic structure refinement with phenix. refine. 68**,** 352-367.

ANGELUCCI, F., BAIOCCO, P., BRUNORI, M., GOURLAY, L., MOREA, V. & BELLELLI, A. J. S. 2005. Insights into the catalytic mechanism of glutathione S-transferase: the lesson from Schistosoma haematobium. 13**,** 1241-1246.

ANGORA, E. K., ALLIENNE, J.-F., REY, O., MENAN, H., TOURÉ, A. O., COULIBALY, J. T., RASO, G., YAVO, W., N'GORAN, E. K. & UTZINGER, J. J. P. 2020. High prevalence of Schistosoma haematobium× Schistosoma bovis hybrids in schoolchildren in Côte d'Ivoire. 147**,** 287-294.

BENVENUTI, M. & MANGANI, S. J. N. P. 2007. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. 2**,** 1633-1651.

BLOCK, H., MAERTENS, B., SPRIESTERSBACH, A., BRINKER, N., KUBICEK, J., FABIS, R., LABAHN, J. & SCHÄFER, F. J. M. I. E. 2009. Immobilized-metal affinity chromatography (IMAC): a review. 463**,** 439-473.

CHARLIER, J., RINALDI, L., MUSELLA, V., PLOEGER, H. W., CHARTIER, C., VINEER, H. R., HINNEY, B., VON SAMSON-HIMMELSTJERNA, G., BĂCESCU, B. & MICKIEWICZ, M. J. P. V. M. 2020. Initial assessment of the economic burden of major parasitic helminth infections to the ruminant livestock industry in Europe. 182**,** 105103.

CHRUSZCZ, M., POTRZEBOWSKI, W., ZIMMERMAN, M. D., GRABOWSKI, M., ZHENG, H., LASOTA, P. & MINOR, W. J. P. S. 2008. Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter? 17**,** 623-632.

CROOM, E. J. P. I. M. B. & SCIENCE, T. 2012. Metabolism of xenobiotics of human environments. 112**,** 31-88.

DE LA TORRE-ESCUDERO, E., PÉREZ-SÁNCHEZ, R., MANZANO-ROMÁN, R., OLEAGA, A. J. M. & PARASITOLOGY, B. 2017. Schistosoma bovis-host interplay: Proteomics for knowing and acting. 215**,** 30-39.

DESIRAJU, G. R. J. J. O. C. S. 2010. Crystal engineering: A brief overview. 122**,** 667-675.

DURRANT, J. D. & MCCAMMON, J. A. J. B. B. 2011. Molecular dynamics simulations and drug discovery. 9**,** 1-9.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. J. A. C. S. D. B. C. 2010. Features and development of Coot. 66**,** 486-501.

EVANS, P. J. A. C. S. D. B. C. 2006. Scaling and assessment of data quality. 62**,** 72-82.

EVANS, P. R. J. A. C. S. D. B. C. 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. 67**,** 282-292.
HABIG, W. H., PABST, M. J. & JAKOBY, W. B. J. J. O. B. C. 1974. Glutathione S-transferases: the first enzymatic step in mercapturic acid formation. 249**,** 7130-7139.
HOLLINGSWORTH, S. A. & DROR, R. O. J. N. 2018. Molecular dynamics simulation for all. 99**,** 1129-1143.
HOSPITAL, A., GOÑI, J. R., OROZCO, M., GELPÍ, J. L. J. A., BIOINFORMATICS, A. I. & CHEMISTRY 2015. Molecular dynamics simulations: advances and applications. 37-47.
ILARI, A., SAVINO, C. J. B. D., SEQUENCE ANALYSIS & EVOLUTION 2008. Protein structure determination by x-ray crystallography. 63-87.
JACOBSON, M. P., FRIESNER, R. A., XIANG, Z. & HONIG, B. J. J. O. M. B. 2002. On the role of the crystal environment in determining protein side-chain conformations. 320**,** 597-608.
JOHNSON, K. A., ANGELUCCI, F., BELLELLI, A., HERVÉ, M., FONTAINE, J., TSERNOGLOU, D., CAPRON, A., TROTTEIN, F. & BRUNORI, M. J. B. 2003. Crystal structure of the 28 kDa glutathione S-transferase from Schistosoma haematobium. 42**,** 10084-10094.
KRAUSS, I. R., MERLINO, A., VERGARA, A. & SICA, F. J. I. J. O. M. S. 2013. An overview of biological macromolecule crystallization. 14**,** 11643-11691.
LABROU, N. E. J. P. D. P. D., DEVELOPMENT, HIGH, A. O. & METHODS, L.-R. 2014. Protein purification: an overview. 3-10.
LAEMMLI, U. K. J. N. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. 227**,** 680-685.
LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D. S. & THORNTON, J. M. J. J. O. A. C. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. 26**,** 283-291.
MAJEED, S., OFEK, G., BELACHEW, A., HUANG, C.-C., ZHOU, T. & KWONG, P. D. J. S. 2003. Enhancing protein crystallization through precipitant synergy. 11**,** 1061-1070.
MAKUMBE, H. H., PANDIAN, R., VALLI, A., SAYED, Y. & ACHILONU, I. J. J. O. M. S. 2024. Biophysical characterization, crystallization, and solution of the first crystal structure of the 28 kDa-Schistosoma bovis glutathione transferase. 1298**,** 136979.
MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. J. J. O. A. C. 2007. Phaser crystallographic software. 40**,** 658-674.
MCPHERSON, A. & GAVIRA, J. A. J. A. C. S. F. S. B. C. 2014. Introduction to protein crystallization. 70**,** 2-20.
OLVEDA, D. U., LI, Y., OLVEDA, R. M., LAM, A. K., CHAU, T. N., HARN, D. A., WILLIAMS, G. M., GRAY, D. J., ROSS, A. G. J. T. M. & SURGERY 2013. Bilharzia: pathology, diagnosis, management and control. 1.
PÉREZ-SÁNCHEZ, R., RAMAJO-HERNÁNDEZ, A., RAMAJO-MARTÍN, V. & OLEAGA, A. J. P. 2006. Proteomic analysis of the tegument and excretory-secretory products of adult Schistosoma bovis worms. 6**,** S226-S236.
PERKINS, S. J. J. E. J. O. B. 1986. Protein volumes and hydration effects: the calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. 157**,** 169-180.
QING, G., MA, L.-C., KHORCHID, A., SWAPNA, G., MAL, T. K., TAKAYAMA, M. M., XIA, B., PHADTARE, S., KE, H. & ACTON, T. J. N. B. 2004. Cold-shock induced high-yield protein production in Escherichia coli. 22**,** 877-882.
RUPP, B. 2009. *Biomolecular crystallography: principles, practice, and application to structural biology*, Garland Science.
SHEEHAN, D. 2009. *Physical biochemistry: principles and applications*, John Wiley & Sons.
SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. J. B. J. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. 360**,** 1-16.
SMITH, D. B. & JOHNSON, K. S. J. G. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. 67**,** 31-40.
SØRENSEN, H. P. & MORTENSEN, K. K. J. M. C. F. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. 4**,** 1-8.
STRYDOM, T., LAVAN, R. P., TORRES, S. & HEANEY, K. J. A. 2023. The economic impact of parasitism from nematodes, trematodes and ticks on beef cattle production. 13**,** 1599.
SWINEHART, D. F. J. J. O. C. E. 1962. The beer-lambert law. 39**,** 333.
UTZINGER, J., RASO, G., BROOKER, S., DE SAVIGNY, D., TANNER, M., ØRNBJERG, N., SINGER, B. & N'GORAN, E. J. P. 2009. Schistosomiasis and neglected tropical diseases: towards integrated and sustainable control and a word of caution. 136**,** 1859-1874.

VANEY, M., BROUTIN, I., RETAILLEAU, P., DOUANGAMATH, A., LAFONT, S., HAMIAUX, C., PRANGE, T., DUCRUIX, A. & RIES-KAUTT, M. J. A. C. S. D. B. C. 2001. Structural effects of monovalent anions on polymorphic lysozyme crystals. 57**,** 929-940.

WALKER, J. M. & WILSON, K. 2010. *Principles and Techniques of biochemistry and molecular biology*, Cambridge university press.

ZHAN, B., PERALLY, S., BROPHY, P. M., XUE, J., GOUD, G., LIU, S., DEUMIC, V., DE OLIVEIRA, L. M., BETHONY, J., BOTTAZZI, M. E. J. I. & IMMUNITY 2010. Molecular cloning, biochemical characterization, and partial protective immunity of the heme-binding glutathione S-transferases from the human hookworm Necator americanus. 78**,** 1552-1563.

Chapter 3

Assessing Ligand Selectivity of *Schistosoma Bovis* Glutathione Transferase Through Computational Modelling in Different Crystal Space Groups

Abstract

Empirical structural methods have been instrumental in drug design, but they often fall short of capturing protein dynamics. To address this limitation, computer-aided drug design (CADD) has become essential. Given that *Schistosoma bovis* 28 kDa glutathione transferase (Sb28GST) is a promising anti-*Schistosome* drug target, a careful selection of crystal structure space groups is used for CADD. This study aimed to determine whether the conformational landscape provided by crystal polymorphs affected the selectivity of ligands across crystal variants. The Sb28GST crystal variants, namely 8BHZ and 8ALS underwent molecular dynamic simulation (MDS) to determine whether these crystal polymorphs undertook identical trajectories in a solvated simulation. Thereafter, high-throughput virtual screening (HTVS) studies were explored using a curated library of flavonoid compounds. The HTVS was performed on monoclinic 8ALS and orthorhombic 8BHZ to identify potential ligands to use as lead compounds. The HTVS showed a diverse selection of hit compounds, with apigenin 7-O-(2G-rhamnosyl)gentiobioside (apigenin) being the common ligand. However, quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside (quercetin) showed the highest affinity to 8ALS with a glide score of -15.66 kcal/mol. Thereafter, 500 ns MDS of the 8ALS-Sb28GST and 8BHZ-Sb28GST apo systems in complex with apigenin and quercetin showed two distinct trajectories which reveal significant differences in their dynamic behaviour and with varying interactions. This highlights the need for a collective assessment of protein polymorphs to comprehensively understand protein dynamics when used for the purposes of drug design.

3.1. Introduction

Computer-aided drug design (CADD) is fast gaining popularity with its inclusion as a rational approach to the drug design and development pipeline (Zhang et al., 2022). The computational aspect of the pipeline is initiated with the examination and extrapolation of the molecular structure of target proteins. This is typically extrapolated from X-ray crystallography, which is vital for understanding the structural conformation of proteins (Maveyraud and Mourey, 2020). However, understanding the overall dynamic nature of proteins is equally as important in drug design. While this information can be obtained using various experimental techniques, the cost to benefit ratio is largely disproportional when investigating multiple ligands as lead therapeutic compounds (Sánchez-Linares et al., 2012). Additionally, the dynamic capabilities of proteins and their complexes are typically reflected as an ensemble of average motions when examined experimentally, failing to showcase the individual protein molecule motions to qualify ligand selection (Kukol, 2008). The computational approaches typically included in CADD are high throughput virtual screening (HTVS) and molecular dynamic (MD) simulations.

High throughput virtual screening digitalises the screening of various compounds through the simulation of ligands and substrates interacting with all available binding sites of the protein of interest. This technique aims to identify the potential positions for ligand binding of screened compounds while simultaneously calculating the affinity between various ligands with the protein target (Zhang et al., 2014). This is determined by sampling all possible positions, orientations and conformations of the screened compounds relative to the binding sites of the target protein, providing atomistic details of interactions. The binding potential is scored using the most energetically favourable binding poses, which are then ranked in order of most probable (Karplus and McCammon, 2002, Tripathi and Bandyopadhyay, 2022). This has been useful for narrowing down lead compounds for subsequent experimental assays (Hospital et al., 2015).

Molecular dynamic simulations are then used to compute further inferences from the predicted top binding ligands. This method mimics the individual motions of complex molecular systems in their physiological environments at an atomic level (Durrant and McCammon, 2011). This is achieved through the calculation of forces with the energy of the participating particles combined to compute a MD system. The particle motions

are then calculated using the molecular mechanics force fields to determine the potential energies and forces of a system. These energies and forces are solved using Newtonian equations of motion to predict the trajectories of the atoms and compounds within the simulated system (Hospital et al., 2015, MacKerell Jr et al., 1998, Ott and Meyer, 1996). The use of MD simulations is advantageous due to the in-depth analysis of both the physical environment and the fluctuations of the participating particles over a specified time period to fully capture the dynamic evolution of a system (Hollingsworth and Dror, 2018).

Researchers often assume target proteins crystallized in different space groups behave similarly in *in-silico* modelling when predicting a lead compound. However, empirical testing suggests varying symmetries and arrangements could affect results. This has led to inaccurate predictions of lead compound performance in empirical testing (Jacobson et al., 2002). The impact of polymorphic protein space groups on conformational dynamics and ligand selectivity is underexplored. Crystal polymorphs' varying symmetries affect solvent content and distribution, potentially leading to one-dimensional interpretations in molecular dynamic simulations (Chruszcz et al., 2008). This study aims to identify novel flavonoid inhibitors targeting *Schistosoma bovis* 28 kDa GST and assess how polymorphic protein crystals affect ligand selectivity. This research showed that the pooling data from various crystal polymorphs could provide deeper insights into protein conformational states and diverse molecular interactions, offering a more comprehensive analysis than studying a single crystal space group.

3.2. Materials and Methods

3.2.1. Materials

All computational studies were performed using two high-power computing units. The bulk of the computational experiments were carried out using a Windows OS desktop housing the Schrödinger Maestro v13.0 software. This computer was equipped with an AMD RYZEN Threadripper 1950X Processor with 16 cores/ 32 threads and an Asus Rog Strix X399-E Gaming Ryzen motherboard all with a 4.0 GHz Precision Boost X399 chipset. The desktop had 4 TB of internal SSD storage with 6 GB/s 3D NAND technology speeds, 64 GB (16GB x 4) Quad-channel DDR4 desktop RAM, and an 11 GB GDDR6 MSI GeForce RTX 2080 Ti graphics card (GPU). The workstation housing the Schrödinger Desmond software was solely used for molecular dynamic

simulations. This computer had an Ubuntu OS desktop equipped with an AMD Threadripper 3990X processor with 64 cores/ 128 threads and a MSI TRX40 PRO 10 G motherboard all with a 4.3 GHz Turbo boost clock and TRX40 chipset. The desktop had a 4 TB HDD and 1 TB of internal M.2 SSD storage with up to 3.5 GB/s speeds, 64GB Quad-channel DDR4 desktop RAM and 8GB GDDR6 GeForce RTX 2070 GPU.

3.2.2. Methods

3.2.2.1. Protein Preparation

The monoclinic and orthorhombic crystal structures of Sb28GST, namely 8ALS and 8BHZ respectively were retrieved from the protein data bank. The crystal structure co-ordinates were submitted into the OPLS_2005 Protein Preparation wizard module implemented in Maestro. The ions, ligand and water molecules were removed from the structure and subjected to energy minimization. Thereafter, hydrogen atoms and disulfide bonds were added to the crystal structures using the default parameters. Het states were then generated at pH 7.0 ±2 using the Epik module.

To build the systems for MD simulation, the System Builder module implemented in Maestro was used. The two Sb28GST systems were built using the OPLS_2005 force field, solvating the system utilizing the TIP3P explicit solvent model. For the boundary conditions an orthorhombic box shape was used, encapsulating the protein at the centre of the box consisting of a 10 Å a = b = c distance with 90° $\alpha = \beta = \gamma$ angles. The box volume containing the Sb28GST systems was minimized. Counter ions (Na^+ or Cl^-) were added in accordance with the molecular systems' overall charge, thus conditioning the system physiologically through the addition of 0.15 M NaCl.

3.2.2.2. Molecular Dynamic Simulation of Apo Crystal Systems

The now solvated, ionized molecular systems were then subjected to MD simulation. The production phase of the simulation consisted of eight stages whose simulation parameters were specified for each phase. The simulation employed the NPT ensemble class using 1 bar pressure and 300 K temperature implemented in all runs. Stages 1 - 7 were for equilibration, where short simulations occurred. This was followed by stage 8 where the final long-range simulation occurred. The stages progressed as follows: Stage 1- The parameters of the solvated systems were

detected. Stage 2 - Brownian Dynamics were used to carry out a 100 ps simulation using NVT conditions at 10 K, restraining the solute heavy atoms. Stage 3 - NVT conditions were used for a 12 ps simulation now restraining heavy atoms. Stages 4, 6 and 7 had short simulation steps lasting 12, 12 and 24 ps respectively. The same conditions were used as stage 3 however stage 7 had no restraints placed on heavy atoms. Stage 8 had the long range 500 ns simulation at a constant temperature of 300 K.

3.2.2.3. Post-Dynamic Analysis

The post-dynamic analyses of the simulation trajectories were carried out on Maestro. This allowed for a comparative analysis of the conformational dynamics of the simulated molecular systems. The simulation interaction diagram module was used on Maestro to calculate the Root Mean Square Deviation (RMSD) and the Root Mean Square Fluctuation (RMSF) of the protein backbone alpha carbon atoms (Cα). The radius of gyration (Zhang et al., 2014) was determined using the Simulation Events Analysis tool on Maestro. The liganded protein simulations were analysed further to determine the ligand RMSD and the ligand properties which include the molecular surface area (MolSA), solvent accessible surface area (SASA) and intramolecular hydrogen bonds (IntraHB).

3.2.2.4. Receptor Grid Generation

To prepare the MD simulated 8BHZ and 8ALS systems for HTVS, the receptor grid was created. The receptor-grid generation module implemented on Maestro was used. The grid box was centroid of protein active site residues from both chains of the protein dimer. The size of the box allowed for the docking of ligands with a 36 Å length.

3.2.2.5. Ligand Library Preparation

A library of 1433 synthesizable flavonoid compounds were used to screen for potential hits. The Structure data files of the compounds were obtained from PubChem database. Maestro was used to prepare the compounds using the LigPrep module. The compounds were protonated, with ions removed, and the geometry optimized. In Maestro the ligands were subjected to energy minimization using OPLS 2005 force field to optimize the bond length, angle, and dihedral.

3.2.2.6. HTVS With MD Simulated Apo Crystal Systems

The virtual screening workflow module implemented in Maestro was used to perform HTVS. The flavonoid library was selected as the source of the ligands. The QikProp filter was exclusively applied whilst also excluding compounds with reactive functional groups. In preparation for the screening event, the Epik module was used to generate possible protonation states at a target pH of 7.0 ± 2.0. Additionally, Epik state penalties were used for docking thus factoring in the energetic cost of each state to avoid false positive hits. Thereafter, high energy ionization /tautomer states were removed. The stereochemical information of the ligands were obtained from the specified 2D stereo properties included in the default settings. The compounds were docked with Glide HTVS, the successful hits that could bind to the 8BHZ and 8ALS binding site were further screened using standard precision (SP) mode. Those outputs from that screening were then further filtered using extra precision (XP) mode. The docking results were processed with Prime MM/GBSA to predict the free-binding energy for ranking purposes. The binding energies was used to calculate the compounds binding affinity for the 8BHZ and 8ALS systems using the below equation:

$$\Delta G° = RT \ln Kd \qquad \textbf{Equation 1}$$

where $\Delta G°$ represents changes in the docking energy, R represents the Boltzmann gas constant ($R = 1.987 cal/mol/K$), *T* represents the temperature and Kd represents the binding affinity.

3.2.2.7. MD Simulation of Top Scoring and Common HTVS Outputs

Two compounds were chosen for further MD simulation; the top scoring compound with the lowest free binding energy and a common compound found in both 8BHZ and 8ALS SP mode screening outputs. Apigenin 7-O-(2G-Rhamnosyl) Gentiobioside herein named apigenin and quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside herein named quercetin were complexed with 8BHZ and 8ALS and subsequently prepared and subjected to 500 ns MD simulations as described previously section in 3.3.2.2.

3.3. Results

3.3.1. MD Simulations and Post-Dynamic Analyses of *Sb*28GST Systems

A 500 ns MDS was done to compare the trajectory of the two crystal polymorphs of Sb28GST. A RMSD evolution which depicts the dynamic behaviour and stability of the protein backbone throughout a simulation period, revealed that the structural conformation throughout the simulation of the polymorphs was not identical, it however followed a similar trend (Figure 3.1 A). A gradual increase in the first 100 ns of the simulation was followed by fluctuations which equilibrated at 400 ns at an average value of 3.5 Å for 8BHZ and 3.1 Å for 8ALS was observed. The RMSF shown in Figure 3.1 B indicated changes along the protein chain over the 500 ns simulation. The RMSF profiles showed similar fluctuation patterns throughout the simulation. The most notable distinctions were the marginally higher peaks observed from 8BHZ *N*-terminal amino acids, which are Lys115 and Ser204 on 8BHZ. The Rg trajectory observed from 8BHZ and 8ALS underwent minimal fluctuations throughout the 500 ns simulation, indicating a sustained degree of structural compactness (Figure 3.1 C). The lower Rg observed in the 8ALS model suggests a marginally more stable system than 8BHZ.

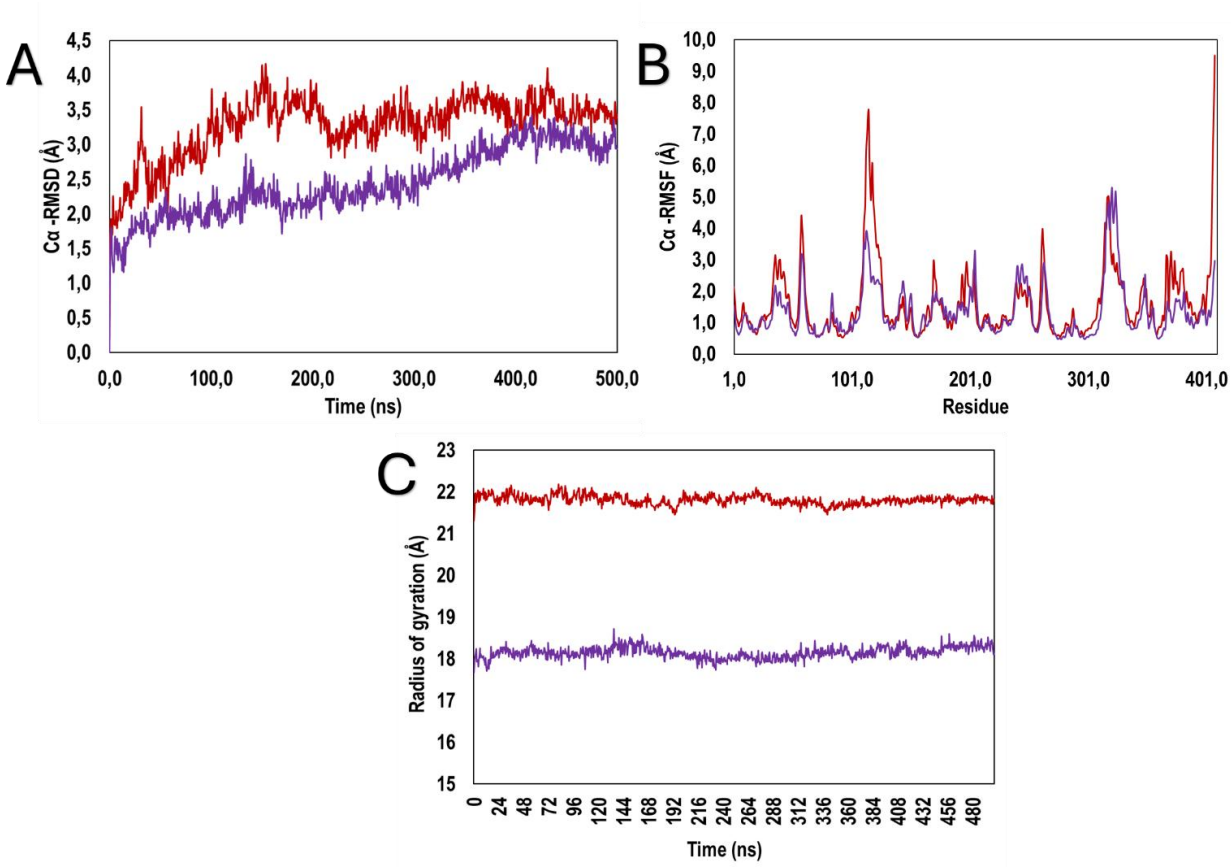


Figure 3. 1: Trajectory analysis of 500 ns molecular dynamic simulations of Sb28GST systems. The red and purple profiles represented 8BHZ and 8ALS respectively. A: The RMSD of the Cα atoms. B: The RMSF of the Cα atoms. C: Radius of gyration of the Cα atom.

3.3.2. Cavity Analysis of Sb28GST

Upon observation of the Sb28GST systems' staggered trajectories, data was collected to probe into the possible binding sites available within the two systems. There are typically four main classes to rank cavity druggability. Sites with D scores > 1.0 are considered very druggable, D scores between 0.75 – 1.0 are considered druggable, scores between 0.5 – 0.75 are considered moderately druggable, and scores below 0.5 are deemed difficult (Halgren and modeling, 2009). The frames from the MDS trajectories of the two Sb28GST systems showed that both structures possessed five potential binding cavities, as shown in Table 3.1. However, only three were putative due to their D score ranking. The druggability levels of these sites are relatively comparable, with the total volume being recorded showing similarity.

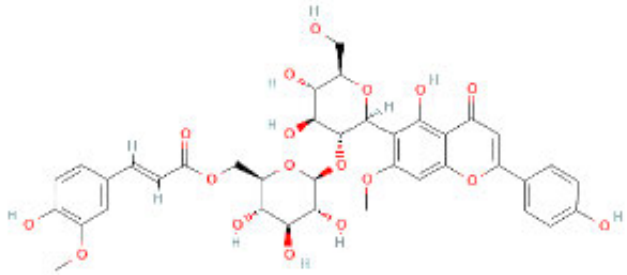
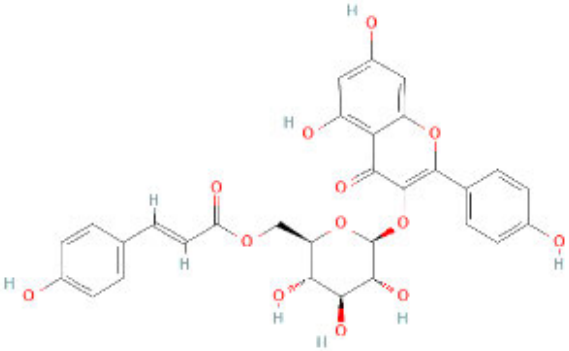
Table 3. 1: Cavity analysis of MD simulated trajectories of 8BHZ and 8ALS

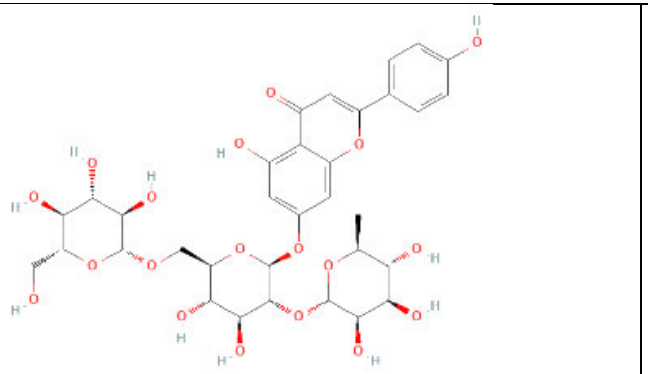
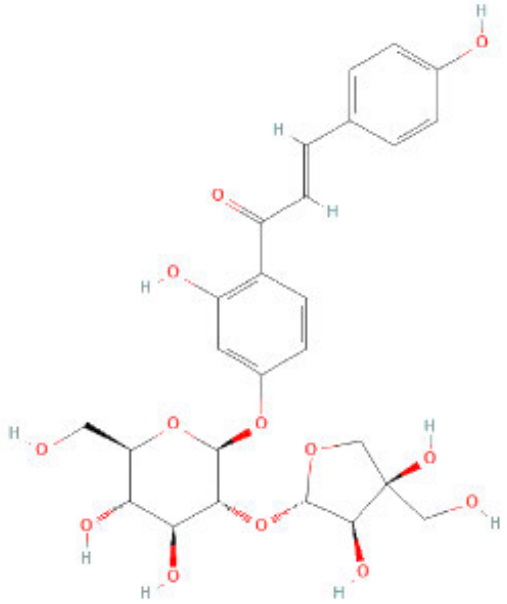
Cavity Rank	8BHZ		8ALS	
	D Score	Volume (Å^3)	D Score	Volume(Å^3)
1	0.9803	1161.7410	1.0586	1261.5540
2	0.6237	101.5280	0.9100	199.9690
3	0.7368	159.8380	0.7017	102.2140
4	0.4158	85.4070	0.6196	151.9490
5	0.7393	233.5830	0.5186	65.8560

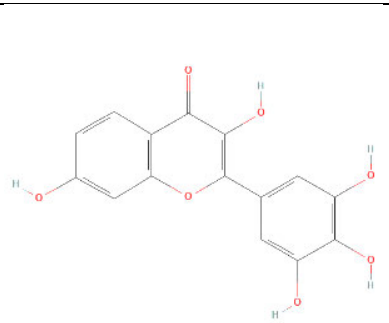
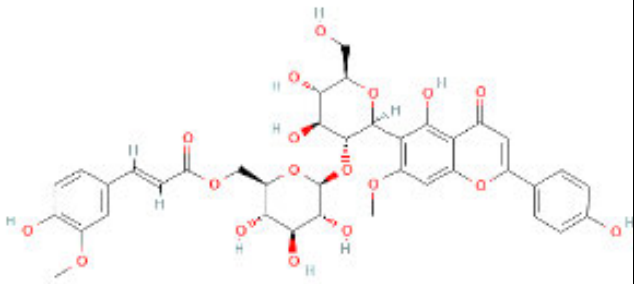
3.3.3. High throughput Virtual Screening of Flavonoid Library for Sb28GST.

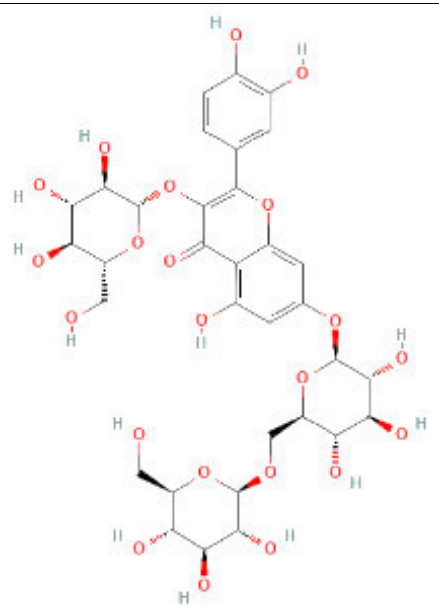
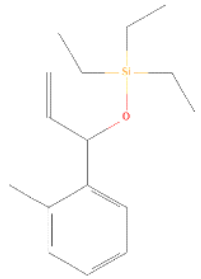
Given that the two Sb28GST MDS trajectories followed different courses, coupled with their assorted binding cavities, a comparative HTVS study was performed to show whether the hit ligands would be identical. This was simultaneously done to determine the best drug candidate for schistosomiasis treatment. The screened top-hit ligands showed compounds with the highest affinity for *Sb*28GST, ranked using their free binding energy (ΔG). The compounds with the lowest ΔG indicated a high affinity for the protein and high protein-ligand complex stability. The free binding energy range exhibited from the HTVS was -15.6645 to -3.591 kcal/mol, indicating that the binding occurred exergonically. Different top ligands from the HTVS were selected for the Sb28GST polymorph, with Table 3.2 showing the ligand 6'''-Feruloylspinosin being selected for 8BHZ, while Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside was selected for 8ALS. Interestingly, despite the two systems belonging to the same *Sb*28GST protein, only one ligand was common within the list of top 5 hits, which was Apigenin 7-O-(2G-Rhamnosyl)Gentiobioside. Additionally, the ligands selected surpassed bromosulfophthalein (BSP), which has been considered the benchmark ligand in previous GST inhibition studies (Padi et al., 2021, Onisuru et al., 2024). Although BSP passed the initial virtual screening, receiving a glide score of -5.93093 kcal/mol, it was subsequently filtered out when subjected to further refinement using SP mode and XP mode.

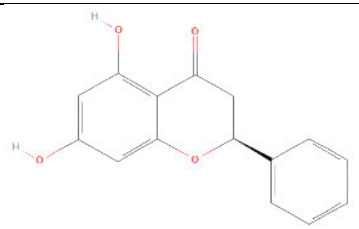
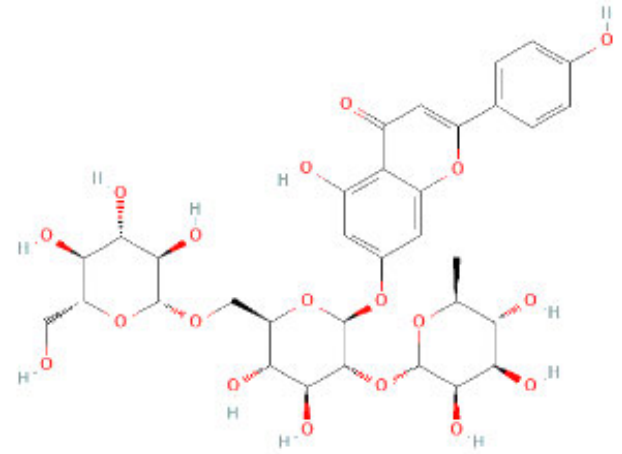
Table 3. 2: The top-ranking flavonoid ligands for Sb28GST selected from high-throughput virtual screening 1433 flavonoid compounds.

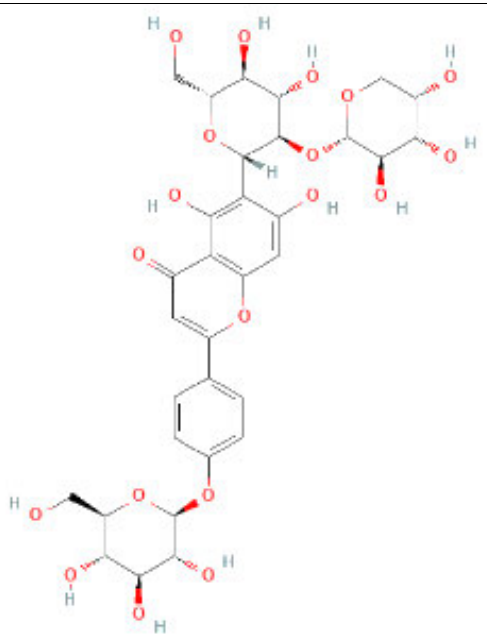
	HTVS Method	Ligands	PubChem ID	2D Structure	ΔGbind (kcal/mol)
8BHZ	SP	6'''-Feruloylspinosin	21597353		-8.22965
		Tiliroside	5320686		-7.65273

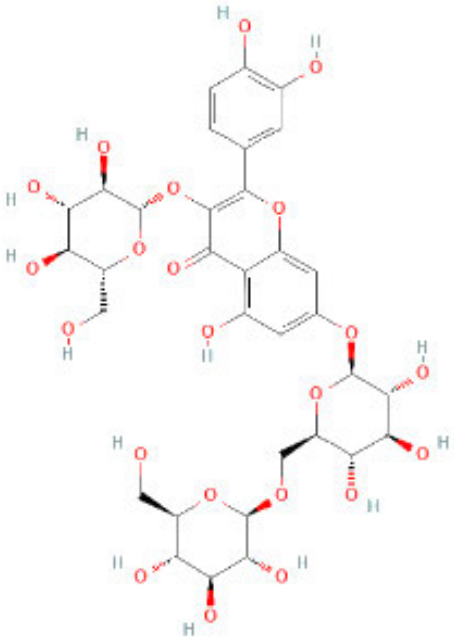
		Apigenin 7-O-(2G-Rhamnosyl) Gentiobioside	10795088		-7.38069
		Licuraside	14282455		-7.26056

		Robinetin	5281692		-6.70999
	XP	6'''-Feruloylspinosin	21597353		-11.2551

8ALS		Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside	131637041		-10.5487
		2-[1-(Triethylsiloxy)-2-propenyl]toluene	57831727		-9.76189

		Pinocembrin	68071		-9.6947
		Apigenin 7-O-(2G-Rhamnosyl) Gentiobioside	10795088		-9.32839

		Vaccarin	71307582		-9.07955

	XP	Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside	131637041		-15.6645

3.3.4. MD Simulations and Post-Dynamic Analyses of Ligand Complexed Sb28GST Systems

Another MDS was performed to test the dynamic behaviour and stability of the interactions between the ligands and the Sb28GST systems. This allowed the protein-ligand complex to be observed in a cell-like environment enclosed in a water box with Na^+ and Cl^- ions as neutralizers. Apigenin (the common ligand produced in the polymorph's HTVS) and Quercetin (the hit with the highest ∆G) were selected for further investigation. The Cα RMSD showed the conformation of the protein backbone throughout the 500 ns simulation with reference to the first frame. The Cα RMSD of Sb28GST-quercetin was very stable in both systems, showing minor fluctuations with an RMSD of 4.24 and 3.80 Å for 8BHZ and 8ALS respectively (Figure 3.2 B). The Sb28GST-apigenin complex however, was less comparable between the two Sb28GST systems but had lower RMSDs in comparison to the reference frame with RMSD values of 3.08 and 2.54 Å for 8HBZ and 8ALS respectively (Figure 3.2 A). The RMSF of the Sb28GST-quercetin and Sb28GST-apigenin complexes showed more homogeneity (Figure 3.2 C and D) with most of the peaks observed corresponding with the amino acids that interact with the ligands, an indication of the stability of the active site. However, additional peaks were observed for the liganded systems compared with the apo systems. In the apigenin complex system, the *N*-terminal amino acids 150, 228 and 352 as well as the amino acids 148, 178 and 260 in complex with quercetin, exhibited higher conformational changes. The Rg was examined to determine the compactness of the protein during the simulation. The Rg profiles of 8BHZ and 8ALS systems shown in Figure 3.2 E were relatively comparable, indicating that polymorphic Sb28GST systems shared similar compactness. However, the overall Rg for 8BHZ-apigenin is 21.75 (± 0.12) and 21.81 (±0.12) for 8ALS-apigenin. On the other hand, 8BHZ-quercetin and 8ALS-quercetin's Rg were 21.66 (± 0.11) and 21.97 (± 0.14) respectively, an indication of moderate conformational changes during the 500 ns MDS (Figure 3.2 F). Similarly, the ligand RMSD trajectory showed the stability of the ligand in its protein binding site, with Figure 3.3 A showing how apigenin exhibited significant fluctuations before reaching equilibrium when bound to 8BHZ and was less dynamic in the 8ALS binding pocket. This showed that while the ligand diffused away from the initial binding pocket (between 0–150 ns), it maintained a stable degree of motion

at 150 – 430 ns. On the other hand, similar activity was observed with quercetin complexed 8ALS, as depicted in Figure 3.3 B. Quercetin drastically diffused away from the initial binding pocket of 8ALS, followed by a stable degree of motion maintained from 200 – 500 ns while behaving completely stable in 8BHZ.

Additional molecular properties of apigenin and quercetin such as MolSA, SASA and IntraHB were studied to determine the stability of the ligands in their Sb28GST binding pockets, shown in Figure 3.4. The SASA profiles showed the surface area of the ligand that is accessible by the solvent molecules. The SASA of 8BHZ in Figure 3.4 A showed a minor increase in the first 100 ns of the simulation with apigenin, which subsequently equilibrated with an average SASA of 251.7 $Å^2$. The SASA of 8BHZ in contrast, was stable for the first 280 ns, fluctuating around a constant average of 199.4 $Å^2$ followed by a steep increase. For quercetin however, the SASA was maintained at a steady average of 211.5 $Å^2$ with 8BHZ while the SASA only reached equilibrium after 200 ns with 8ALS at 332.9 $Å^2$. The analysis of the intramolecular hydrogen bond (H-bond) formation throughout the MDS between 8BHZ and apigenin depicted in Figure 3.4 C, showed up to four H-bonds formed with two remaining intact at the end of the simulation. However, the H-bonds formed with 8ALS and apigenin were maintained at one, with two forming near the end of the simulation. In Figure 3.4 D, we see up to four H-bonds formed between both 8BHZ and 8ALS with quercetin. A majority of the two H-bonds were maintained throughout the simulation. The MolSA is equivalent to the Van Der Waals surface area probed at a radius of 1.4 Å. The MolSA of the protein-ligand complexes were relatively comparable, with the averages being 585.7 $Å^2$ for the 8BHZ-apigenin complex, 587.3 $Å^2$ for 8BHZ-quercetin and $Å^2$ 580.0 for 8ALS-quercetin. The lowest MolSA value obtained was 564.4 $Å^2$ obtained from 8ALS-apigenin.

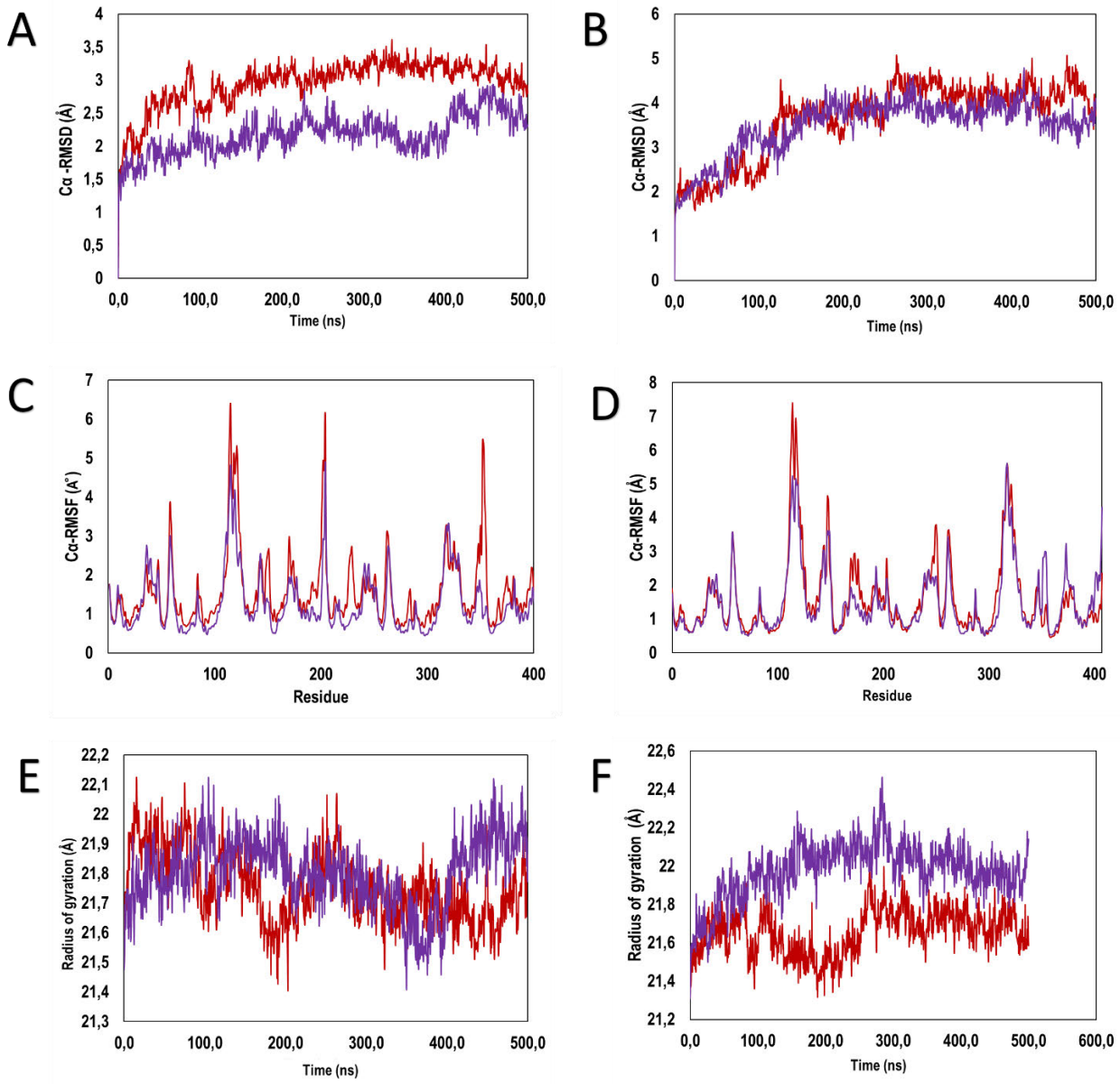


Figure 3. 2: Trajectory analysis of the 500 ns MDS of Sb28GST systems complexed apigenin and quercetin. The red and purple profiles represented 8BHZ and 8ALS respectively with apigenin interactions depicted on the left panels and quercetin interactions depicted on the right panels. A: The RMSD of the Cα atoms of Sb28GST systems with apigenin. B: The RMSD of the Cα atoms of Sb28GST systems with quercetin. C: The RMSF of the Cα atoms of Sb28GST systems with apigenin. D: The RMSF of the Cα atoms of Sb28GST systems with quercetin. E: The Rg of the Cα atoms of *Sb*28GST with apigenin. F: The Rg of the Cα atoms of Sb28GST with quercetin.

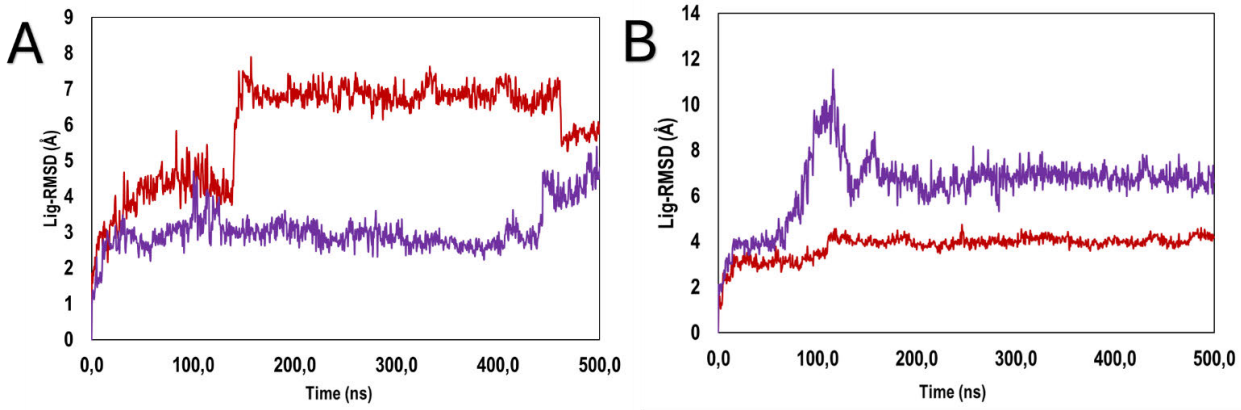


Figure 3. 3: Root mean square deviation of the apigenin and quercetin with respect to Sb28GST residues. The red and purple profiles represented 8BHZ and 8ALS residues respectively. A: The ligand RMSD of apigenin with Sb28GST receptors. B: The ligand RMSD of quercetin with Sb28GST receptors.

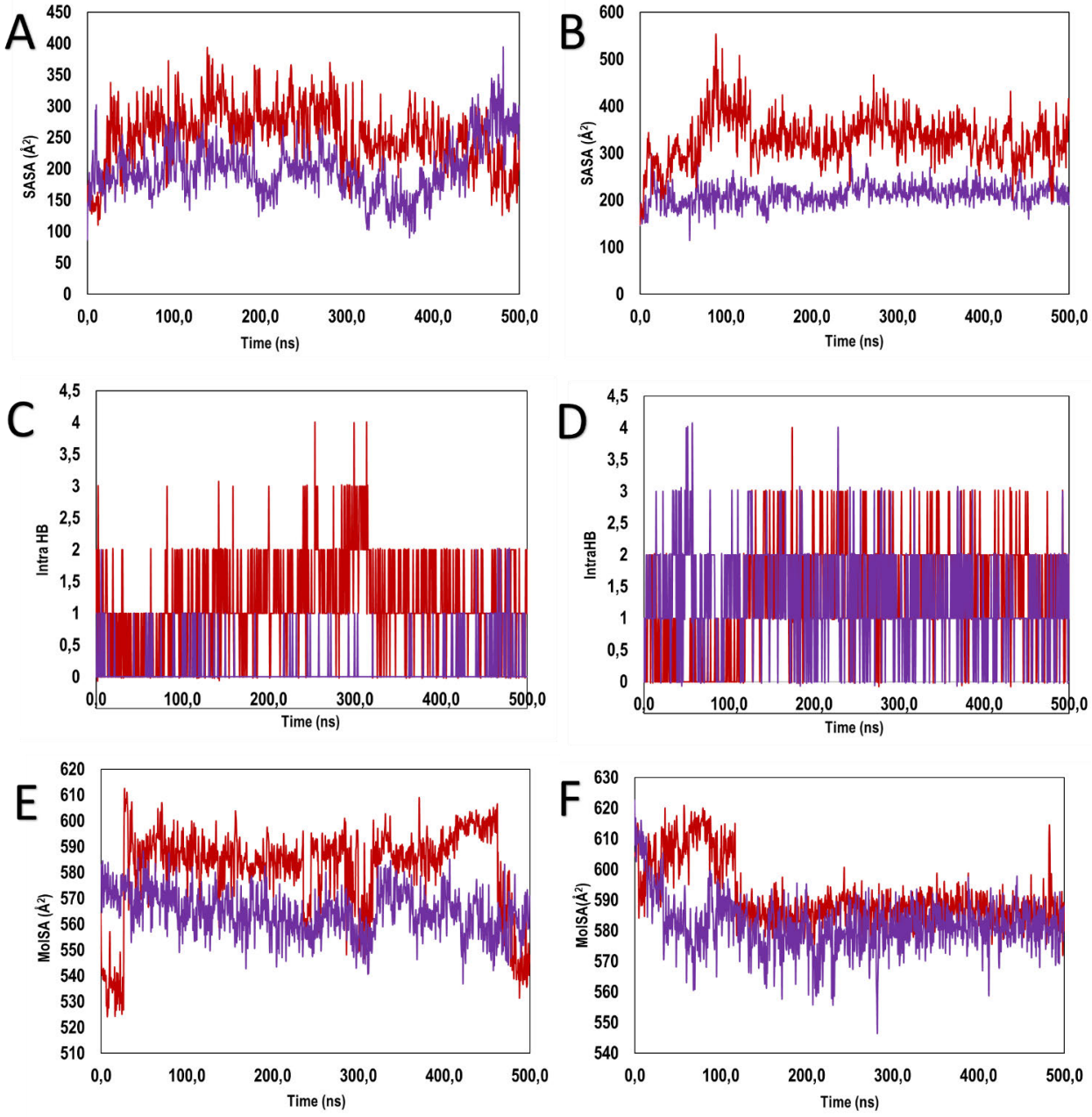


Figure 3. 4: Analysis revealing the ligand properties of apigenin and quercetin during the 500 ns MDS with Sb28GST systems. The red and purple profiles represented 8BHZ and 8ALS respectively with A: The solvent accessible surface area available to apigenin. B: The solvent accessible surface area available to quercetin. C: The intramolecular hydrogen bonds within apigenin. D: The intramolecular hydrogen bonds within quercetin. E: The molecular surface area of apigenin. F: The molecular surface area of quercetin.

Figure 3.5 showed the Sb28GST interacting amino acid and the trajectory cluster analysis, revealing that apigenin and quercetin were stabilized by hydrogen bonds, water bridges, hydrophobic and ionic interactions. In the 8BHZ-apigenin complex, water bridges played a larger role compared to 8ALS-apigenin. Figure 3.5 A showed that through 50% of the 500 ns simulation trajectory, apigenin formed H-bond interactions with B:Lys120 of 8BHZ. B:Arg14 (H-bond, 31%) and B:His107 (pi-pi stacking, 36%) interactions were also formed. B:His169 and B:Ser71 formed a water-mediated H-bond with the apigenin hydroxyl group (33%) and the carbonyl group (31%). Apigenin also formed an intramolecular H-bond between its hydroxyl and carbonyl group through 64% of the simulation. This bond was not formed in the 8ALS-apigenin simulation. However, the interactions observed were more complex and lasted longer than those seen in the 8BHZ simulation, as shown in Figure 3.5 B. In the 8ALS GST A:Glu70 and A:Leu72 formed H-bonds with the apigenin hydroxyl group through 85% and 32% of the simulation respectively. A:Glu103 formed bonds with two apigenin hydroxyl groups (H-bond, 82%), A:Ser72 formed a bond with a carbonyl group (H-bond, 80%) and A:Glu106 (H-bond 52%). B:Glu103 formed a water-mediated H-bond through 76% of the simulation and a pi-cation formation through 75% and 50% of the simulation.

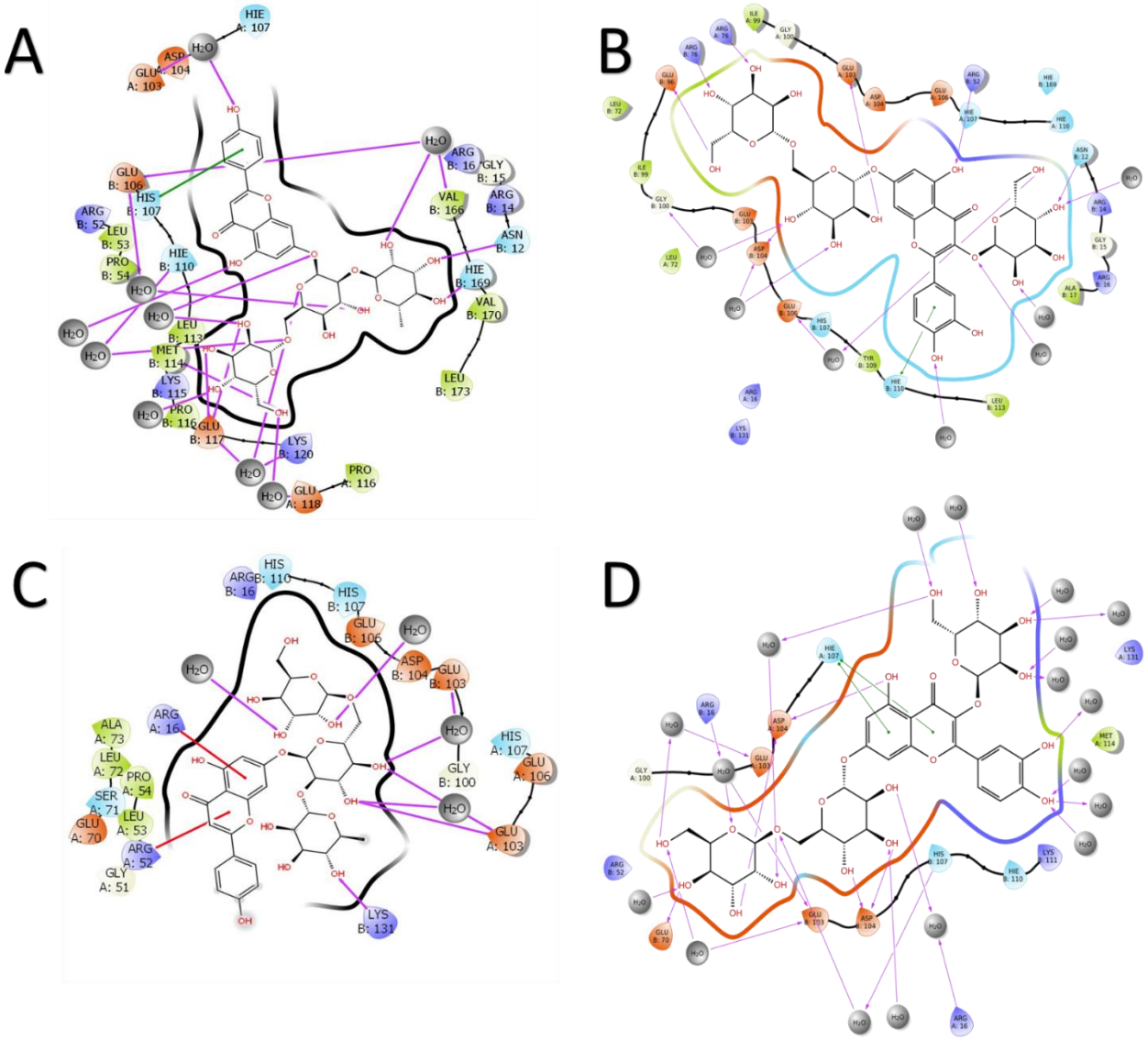


Figure 3. 5: 2D interaction plot of the trajectory frame based on the RMSD of most dominant snapshots throughout a 500-ns MDS. The protein: ligand complexes were depicted as follows; A: 8BHZ-Apigenin, B: 8BHZ-Quercetin, C: 8ALS-Apigenin, and D: 8ALS-Quercetin. Amino acid residues within 4 Å of the ligand are colour-coded: positively charged (violet), negatively charged (orange), non-polar (green), and polar (blue). Hydrogen bonds (purple lines) and salt bridges (red lines) are shown, with solvent exposure shaded grey. Images were created using Maestro v13.0.

Compared to apigenin, significantly more interactions occurred between the *Sb*28GST systems and quercetin. The top interactions shown in Figure 3.6 A are A:Glu103 (H-bond, 99%), A:Asp102 (H-bond, 94%), B:Glu106 (H-bond, 84%), B:Glu107 (H-bond, 77%). A:Glu106 also formed a mediated H-bond with the quercetin hydroxyl group through 82%

of the simulation. A pi-pi stack interactions were formed by A:His107 (56%) and B:His107 (56 and 34%), and lastly pi-cation interactions were formed with quercetin by B:Arg16 (33%). Again, pi-pi stacking interactions were also formed between 8ALS and quercetin by A:His107 through 70% of the simulation shown in Figure 3.6 B. Quercetin also formed intramolecular hydrogen bonds between its hydroxyl and carbonyl group through 64% of the simulation. Other interactions include A:Asp104 (H-bond, 70%), A:Glu70 (H-bond, 67%) and A:Glu103 (H-bond, 57%). Water-mediated hydrogen bonds were formed by B:Glu103 (64 and 38%), B:Arg16 (59%) and B:Asp104 (58 and 35%).

Between the two polymorphic systems, there was only one interaction shared with apigenin, although the contributing amino acid was from different subunits 8BHZ-B:Ser71, which was a water-mediated H-bond (31%), and 8ALS-A:Ser71, which was H-bonded to the apigenin carbonyl group. With quercetin, the two interactions shared between 8BHZ and 8ALS were B:His107, forming a pi-pi stack of 57% and 76%, respectively. The other interaction shared between 8BHZ and 8ALS was A:Glu103 with a hydroxyl group occurring throughout 99% and 57% of the simulation respectively.

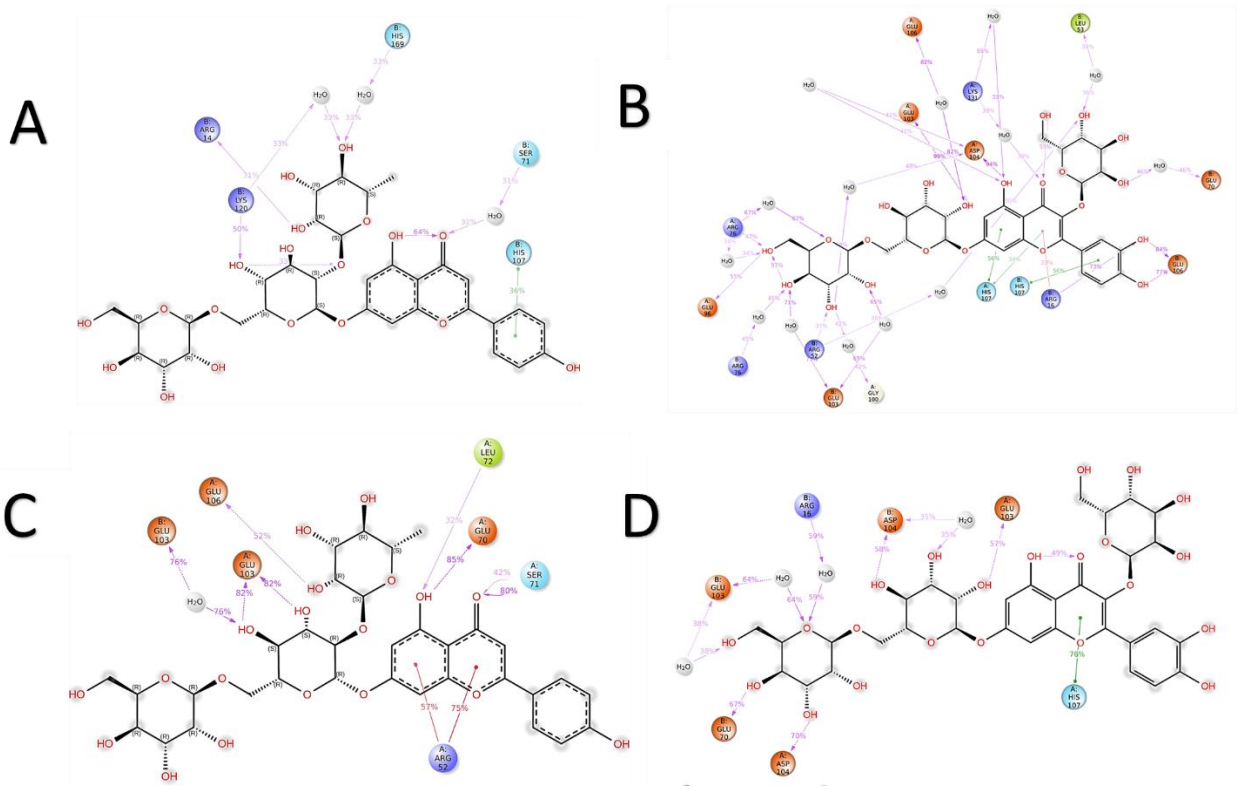


Figure 3. 6: Analysis of the 2D summary contact of ligand atoms interacting with *Sb*28GST amino acids. The protein: ligand complexes were depicted as follows; A: 8BHZ-Apigenin, B: 8BHZ-Quercetin, C: 8ALS-Apigenin, and D: 8ALS-Quercetin. Representation of positively charged amino acids are violet, negatively charged are orange, non-polar are green, and polar are blue. The percentages indicate ligand-*Sb*28GST interactions over 500 ns.

3.4. Discussion

To address the question of whether different crystal space groups could influence ligand selectivity in CADD we employed computational methods to examine the behaviour of two Sb28GST space group systems in solution through molecular dynamic simulation, namely 8BHZ (orthorhombic) and 8ALS (monoclinic). The simulations conveyed details of individual particle motions as a function of time, offering a way to quantify the degree of conformational changes in protein backbones and side chains (Hollingsworth and Dror, 2018).

Simulating the apo Sb28GST systems showed that the diverging trajectories of the protein backbone showed that different space group proteins possess unique conformational states relative to one another, which manifest through dissimilar dynamic properties. The protein side chains were comparably and relatively stable in solution with slight fluctuations observed with 8BHZ. These greater RMSF values indicated increased flexibility throughout the simulation. The lower RMSF values, however, reflected the stability of the 8ALS side chain residues. The radius of gyration showed us the degree of compactness of the protein systems. Although the degree of compactness remained relatively similar between the protein systems, 8ALS appeared to remain more compact throughout the simulation compared to 8BHZ. This highlights the divergence of polymorphic systems' dynamic properties from the anticipated identical molecular disposition.

The cavity detection of the Sb28GST systems revealed essential information for virtual screening, drug discovery and receptor binding analyses. This method allowed us to consider the flexibility of the protein and the potential cavity volume to quantitatively estimate the druggability of a particular site (Ricci et al., 2022). Once again, the cavities observed were not identical, although they were similar in volume. This could be due to variations in crystal packing forces inducing different side chain rearrangements (Chruszcz et al., 2008). The ranging orientations adopted by Sb28GST chains in orthorhombic and monoclinic space group systems thus induced regions of the protein surfaces to be involved in assorted interchain interactions, effectively resulting in slightly varied cavity arrangements. Since the D score is governed by the number of site points, enclosure score and hydrophilic score, inconsistencies in these parameters would not be surprising as the different side chain conformations could ultimately give

rise to different binding site possibilities due to the crystal packing environment (Jacobson et al., 2002, Vidler et al., 2012). There were three confirmed binding sites which were deemed favourable across the two models. This aligns with current literature which corroborates the existence of three known GST binding sites (Yassin et al., 2004). In a single monomer of the dimeric protein there exists a G-site which is highly covered and specific for GSH. Adjacent to that is in contrast a highly promiscuous H-site which accommodates a variety of electrophiles. Within the dimer interface of the GST monomers exists an additional L-site where substrate binding within the H-site is modulated (Sheehan et al., 2001).

The observations made from the diverging MD trajectories and cavity environments lead us to further investigate the selection of ligands predicted to be lead compounds in schistosomiasis treatment. A library of plant derived polyphenols were used to identify anti-*Schistosome* drugs due to the ease of procurement and the adeptness flavonoids possess to modulate drug metabolising enzymes (Boušová et al., 2012). It was evident that the Sb28GST systems exhibit an affinity for a wide variety of ligands. Analysis of the top five binding flavonoids for the orthorhombic and monoclinic Sb28GST systems revealed only one common ligand. This ligand was Apigenin 7-O-(2G-Rhamnosyl) Gentiobioside of the flavone subclass. Interestingly, this compound exhibited different binding energies across the two systems. The top binding ligand with the lowest binding energy was Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D Gentiobioside of the flavonol subclass. These flavonoids held therapeutic potential due to their polyphenolic structure and the specific arrangements of their hydroxyl groups on the A, B, and C rings available to facilitate stable binding within the protein active site. Additionally, the presence of the gentiobioside and rhamnosyl glycosides primed the flavonoids for enhanced interactions through increased water solubility and bioavailability (Boušová et al., 2012).

Although the screening showed apigenin and quercetin to have favourable binding energies towards the Sb28GST systems, analysis of the dynamic behaviours of the protein systems in complex with the ligands was performed. The Sb28GST side chains were relatively stable in the presence of both drugs as their presence led to interactions with near identical amino acid residues. Distinctions were observed with contrasting degrees of fluctuations between the two systems, showing that the

residues were impacted in varying degrees, with differing flexibility ranges while maintaining a notable level of stability.

The Rgs of the simulated systems in the presence of the ligands were relatively similar with an insignificant difference between them however, the apo 8ALS Rg was noticeably smaller. This suggests that the introduction of a ligand into a less packed monoclinic system allowed the protein structure to expand, significantly losing its compactness (Lobanov et al., 2008). This pattern was not seen with 8BHZ possibly because the orthorhombic unit cells do not allow for the same level of expansion. This suggests that monoclinical structures thus have room to expand and lose their compactness to accommodate the introduction of ligands into a system.

Assessing the ligand properties shows that both drugs had similar dynamic motions throughout the simulation with minor differences. Apigenin was not as stable in the Sb28GST active site as quercetin. The binding energy and the lig-RMSD suggested that quercetin diffused towards the active site and remained for the duration of the simulation. Apigenin however appeared to have diffused slowly towards the protein active site, then proceeded to drift away. This is further corroborated by the observation that the polar surface area of apigenin was not as consistent as quercetin, implying that apigenin did not have consistent and stable interactions with the solvent molecules as it failed to maintain a balanced distribution of polar contacts with the protein throughout the simulation. The ligands both seem to orient themselves within the active site prior to equilibration at certain positions as evidenced by the SASA. The number of intermolecular hydrogen bonds formed are important for ligand binding and heavily considered when designing drugs. They are essential for providing stability to the ligand-protein complex (da Fonseca et al., 2024). In both systems, there was a considerable network of hydrogen bonds formed throughout the simulation, with apigenin forming more bonds overall. The significantly higher number of hydrogen bonds formed within the monoclinical system showed that the monoclinical environment allowed for higher stability networks with the ligand compared to the orthorhombic environment.

Although the Sb28GST systems made strong interactions with both apigenin and quercetin, the amino acids utilized by the different space group systems to interact with the drugs vary in identity and intensity of the interaction. The increased presence

of water bridges observed in 8BHZ interactions could be as a result of spatial constraints contributed by fewer voids and channels of the tightly packed orthorhombic protein system (Shah et al., 2024). The restricted conformational flexibility can also result in fewer but more stable interactions aided by the abundant water molecules (Zhou and Caflisch, 2010). Limited access to the binding site would also result in increased reliance on more indirect interactions such as water bridges, making water molecules very crucial to mediate interactions as seen with 8BHZ (Chruszcz et al., 2008, Jacobson et al., 2002). While this is advantageous, the tight packing in environments such as those in the orthorhombic 8BHZ can limit the amount of water that can interact with the protein ligand complex as a result of fewer solvent channels (Chruszcz et al., 2008). On the other hand, the less regularly packed monoclinical 8ALS environment allows for more conformational stability, increased accessibility to the protein active site, larger solvent channels and increased overall hydration. These differences allow more complex, varied interactions which last longer allowing ligands to form more direct interactions with 8ALS amino acids. These differences allow direct hydrogen bonding and give way for more complex interactions due to the freedom ligands have to adopt multiple conformations, allowing optimal binding for a wide variety of interactions (Chruszcz et al., 2008).

In conclusion, adopting computational methods to identify lead compounds for schistosomiasis management is an effective time saving approach to drug discovery. Our study however, showed that it is crucial to be aware that the relationship between a protein's crystal structure relative to its dynamic behaviour in solution. The irregularities obtained from ligand screening, binding affinities and molecular interactions from two different space group models of the same protein highlighted the value in collectively analysing multi-space group crystals. This would allow a more comprehensive understanding of the possible conformational states a protein can assume and the resulting interactions formed because of it.

3.5. References

BOUŠOVÁ, I., HÁJEK, J., DRŠATA, J. & SKÁLOVÁ, L. J. X. 2012. Naturally occurring flavonoids as inhibitors of purified cytosolic glutathione S-transferase. 42**,** 872-879.

CHRUSZCZ, M., POTRZEBOWSKI, W., ZIMMERMAN, M. D., GRABOWSKI, M., ZHENG, H., LASOTA, P. & MINOR, W. J. P. S. 2008. Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter? 17**,** 623-632.

DA FONSECA, A. M., CALUACO, B. J., MADUREIRA, J. M. C., CABONGO, S. Q., GAIETA, E. M., DJATA, F., COLARES, R. P., NETO, M. M., FERNANDES, C. F. C. & MARINHO, G. S. J. M. B. 2024. Screening of potential inhibitors targeting the main protease structure of SARS-CoV-2

via molecular docking, and approach with molecular dynamics, RMSD, RMSF, H-bond, SASA and MMGBSA. 66**,** 1919-1933.
DURRANT, J. D. & MCCAMMON, J. A. J. B. B. 2011. Molecular dynamics simulations and drug discovery. 9**,** 1-9.
HALGREN, T. A. J. J. O. C. I. & MODELING 2009. Identifying and characterizing binding sites and assessing druggability. 49**,** 377-389.
HOLLINGSWORTH, S. A. & DROR, R. O. J. N. 2018. Molecular dynamics simulation for all. 99**,** 1129-1143.
HOSPITAL, A., GOÑI, J. R., OROZCO, M., GELPÍ, J. L. J. A., BIOINFORMATICS, A. I. & CHEMISTRY 2015. Molecular dynamics simulations: advances and applications. 37-47.
JACOBSON, M. P., FRIESNER, R. A., XIANG, Z. & HONIG, B. J. J. O. M. B. 2002. On the role of the crystal environment in determining protein side-chain conformations. 320**,** 597-608.
KARPLUS, M. & MCCAMMON, J. A. J. N. S. B. 2002. Molecular dynamics simulations of biomolecules. 9**,** 646-652.
KUKOL, A. 2008. *Molecular modeling of proteins*, Springer.
LOBANOV, M. Y., BOGATYREVA, N. & GALZITSKAYA, O. J. M. B. 2008. Radius of gyration as an indicator of protein structure compactness. 42**,** 623-628.
MACKERELL JR, A. D., BASHFORD, D., BELLOTT, M., DUNBRACK JR, R. L., EVANSECK, J. D., FIELD, M. J., FISCHER, S., GAO, J., GUO, H. & HA, S. J. T. J. O. P. C. B. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. 102**,** 3586-3616.
MAVEYRAUD, L. & MOUREY, L. J. M. 2020. Protein X-ray crystallography and drug discovery. 25**,** 1030.
ONISURU, O., ACHILONU, I. J. J. O. B. S. & DYNAMICS 2024. Describing the ligandin properties of Plasmodium falciparum and vivax glutathione transferase towards bromosulfophthalein from empirical and computational modelling viewpoints. 1-16.
OTT, K. H. & MEYER, B. J. J. O. C. C. 1996. Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. 17**,** 1068-1084.
PADI, N., AKUMADU, B. O., FAERCH, O., ALOKE, C., MEYER, V. & ACHILONU, I. J. B. 2021. Engineering a pseudo-26-kDa schistosoma glutathione transferase from bovis/haematobium for structure, kinetics, and ligandin studies. 11**,** 1844.
RICCI, F., GITTO, R., PITASI, G. & DE LUCA, L. J. B. 2022. In silico insights towards the identification of SARS-CoV-2 NSP13 helicase druggable pockets. 12**,** 482.
SÁNCHEZ-LINARES, I., PÉREZ-SÁNCHEZ, H., CECILIA, J. M. & GARCÍA, J. M. J. B. B. 2012. High-throughput parallel blind virtual screening using BINDSURF. 13**,** 1-14.
SHAH, A., CHOUDHARY, A., JAIN, M., PERUMAL, S., PATEL, V., PARMAR, G. & PATEL, A. J. B. 2024. Discovery of novel anticancer flavonoids as potential HDAC2 inhibitors: virtual screening approach based on molecular docking, DFT and molecular dynamics simulations studies. 14**,** 83.
SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. J. B. J. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. 360**,** 1-16.
TRIPATHI, N. M. & BANDYOPADHYAY, A. J. E. J. O. M. C. 2022. High throughput virtual screening (HTVS) of peptide library: Technological advancement in ligand discovery. 243**,** 114766.
VIDLER, L. R., BROWN, N., KNAPP, S. & HOELDER, S. J. J. O. M. C. 2012. Druggability analysis and structural classification of bromodomain acetyl-lysine binding sites. 55**,** 7346-7359.
YASSIN, Z., ORTIZ-SALMERÓN, E., GARCíA-MAROTO, F., BARÓN, C., GARCíA-FUENTES, L. J. B. E. B. A.-P. & PROTEOMICS 2004. Implications of the ligandin binding site on the binding of non-substrate ligands to Schistosoma japonicum-glutathione transferase. 1698**,** 227-237.
ZHANG, B., LI, H., YU, K. & JIN, Z. J. C. T. O. H. P. C. 2022. Molecular docking-based computational platform for high-throughput virtual screening. 1-12.
ZHANG, X., WONG, S. E. & LIGHTSTONE, F. C. 2014. Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized Born surface area rescoring to improve enrichment. ACS Publications.
ZHOU, T. & CAFLISCH, A. J. C. 2010. High-throughput virtual screening using quantum mechanical probes: discovery of selective kinase inhibitors. 5**,** 1007-1014.

Chapter 4

Characterization of The Ligandin Properties of *Schistosoma Bovis* 28 Kda Glutathione Transferase from Computationally Selected Flavonoid Compounds

Abstract

Schistosoma bovis 28 kDa glutathione transferase (Sb28GST) is a vital detoxification enzyme that aids in the parasite's control tactics employed by the host and through chemical intervention. This makes Sb28GST a promising therapeutic drug target for the devastating disease schistosomiasis. Herewith the computationally determined lead compounds apigenin 7-O-(2G-rhamnosyl)gentiobioside and quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside referred to as apigenin and quercetin, respectively, are validated as inhibitors against the parasitic enzyme. Upon the successful expression and purification of Sb28GST the inhibitory potency of apigenin and quercetin was determined using a reduced glutathione (GSH) - 1-chloro-2,4-dinitrobenzene (CDNB) conjugation assay. This activity assay confirmed that the lead compounds significantly reduced the specific enzyme activity of Sb28GST with both ligands having half-maximal inhibitory concentration (IC_{50}) of approximately 0.12 mM. Extrinsic fluorescence studies and thermal shift assay indicated that these compounds bind to the hydrophobic H-site and allosteric L-site at the dimer interface and have minimal effect on the protein's thermal stability. The quantitative thermodynamic parameters of apigenin and quercetin were examined using isothermal titration calorimetry. This analysis showed the interactions between Sb28GST and the compounds apigenin and quercetin were spontaneous and entropically driven (ΔG° apigenin = -46.76 kJ/mol and ΔG° quercetin = -26.17 kJ/mol), binding to Sb28GST with both compounds bound naturally to Sb28GST without the input of external energy. Apigenin bound endothermically to Sb28GST (ΔH° = 18.29 kJ/mol), while quercetin bound exothermically to Sb28GST (ΔH° = -17.69 kJ/mol), suggesting varying binding interactions.

4.1. Introduction

The schistosomiasis burden of disease is of major concern both to humans and animals. This disease is caused by trematodes of the *Schistosoma* parasite which are responsible for symptoms of anaemia, haemorrhagic enteritis and death in some cases (Molla et al., 2022). *Schistosoma bovis* most significantly impacts ruminants, causing devastating veterinary problems which influence animal growth, immune system and productivity (de la Torre-Escudero et al., 2017). Interferences of this nature translate to significant economic losses globally while crippling food security pathways in poor rural communities that rely on agricultural activities for survival thus perpetuating an endless cycle of poverty (Charlier et al., 2020, Strydom et al., 2023). The zoonotic potential that has arisen from *S. bovis-S. haematobium* hybrids amplifies the health burden caused by schistosomiasis. These *Schistosoma* hybrids have evolved to expand their host spectrum and exhibit increased virulence, resulting in increased reports of outbreaks of clinical disease in both cattle and humans (Angora et al., 2020, Léger et al., 2020). With insufficient control measures in health, education and sanitation to decrease schistosomiasis morbidity, disease mitigation relies heavily on chemotherapeutic intervention. However, the major anti-schistosomal drug used, named Praziquantel (PZQ), falls short of effective schistosomal control due to its predisposition to exclusively control egg producing adult *Schistosome* worms (Demlew and Tessma, 2020, Olveda et al., 2013). This shortfall highlights the urgent need for alternative chemotherapy interventions.

A promising therapeutic target for schistosomiasis intervention is the detoxification enzyme glutathione transferase (GST). The significance of GSTs is amplified due to their existence as *Schistosomes*' sole defence against host immune responses and chemotherapeutic disruptions, as they lack phase I detoxification enzymes (Huang et al., 2012, Pérez-Sánchez et al., 2006). The role of GSTs in *Schistosomes*, much like in eukaryotic cells is crucial for maintaining vital cell metabolic processes such as repairing oxidised macromolecules and biosynthesizing physiologically important metabolites (Müller, 2015). However, the primary role of GSTs is the detoxification of a wide range of electrophilic substrates through their conjugation to reduced glutathione (GSH) for subsequent excretion (Sheehan et al., 2001). The inhibition of this enzyme would render the parasite unable to evade parasite control interventions, both by the host immune system and external chemical aid, while simultaneously

disrupting cellular metabolic processes (Torres-Rivera and Landa, 2008). To strategically target this enzyme for anti-schistosomal intervention, observing its binding characteristics against lead compounds can empirically verify GST inhibitors pre-selected computationally. Glutathione transferases are dimeric in nature, possessing three binding sites (Sheehan et al., 2001). There exists a hydrophobic substrate binding site called the H-site and a GSH binding site called the G-site (Torres-Rivera and Landa, 2008). Adjoining the GST monomers at the dimer interface is a third non-substrate binding site within the enzyme's core which is typically implicated in drug binding, called the L-site (Yassin et al., 2004). The ligandin functionality that GSTs possess due to their promiscuous affinity to a wide range of xenobiotic substrates has been shown to result in the enzymes' inhibition. This allows for the exploitation of the enzymes' ligandin functionality as a noteworthy pathway to investigate lead compound efficacy that inhibits the enzymes' catalytic functionality (Kolobe et al., 2004, McTigue et al., 1995).

A rational approach to drug design systematically integrates computational and empirical studies to advance drug molecules through the drug development pipeline. In this study, the computationally determined lead compounds apigenin 7-O-(2G-Rhamnosyl)gentiobioside and quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside hence forth referred to as apigenin and quercetin respectively, were validated *in vivo* using spectroscopic, fluorescent and calorimetry techniques. Through monitoring the ability and rate at which Sb28GST can catalyse the conjugation of GSH to the substrate 1-chloro-2,4-dinitrobenzene (CDNB), the inhibitory capabilities of the lead compounds can be determined (Habig et al., 1974). To further characterize the drug compounds' binding characteristics within the protein, changes within the hydrophobic regions of the protein are probed using the fluorescent dye anilino-1-naphthalene sulphonate (Hawe et al., 2008). The displacement of this dye both competitively and non-competitively by hydrophobic ligands typically speaks to the specificity of the protein's binding sites (Kinsley et al., 2008). Fluorescent dyes such as SYPRO Orange are further exploited to monitor hydrophobic regions exposed during thermal denaturation events. Such techniques thus allow for the determination of protein melting points as well as the effect of drug compounds on the thermal stability of proteins (Huynh and Partch, 2015). Lastly, to gain insight into the binding affinities of the flavonoids for Sb28GST, the thermodynamic characteristics of the

drugs are typically investigated using calorimetry techniques (Du et al., 2016). This is the first reported study of synthesizable flavonoids investigated as therapeutic agents against Sb28GST.

4.2. Materials and Methods

4.2.1. Materials

All chemicals and reagents used were analytical grade. Vector synthesis was carried out by GenScript (NJ, USA). The *Escherichia coli* (*E. coli*) T7 Express chemically competent cells were purchased from New England Biolabs (USA). Tryptone, yeast extract, NaCl, ampicillin, *N,N,N′,N′*- tetramethylethylenediamine (TEMED), bisacrylamide, acrylamide, sodium dodecyl sulfate (SDS), β-mercarptoethanol, glycine, isopropyl-β-D-thiogalactoside (IPTG), glacial acetic acid and coomassie Blue-G250 dye, reduced glutathione (GSH), 1-chloro-2,4-dinitrobenzene (CDNB), anilino-1-naphthalene sulphonate, Dithiothreitol (DTT), Tris (2-carboxyethyl) phosphine (TCEP), sodium phosphate, Tris-HCl, ethylenediaminetetraacetic acid (EDTA), sodium azide (NaN_3) and (phosphate buffered saline (PBS) reagents were all purchased from Sigma Aldrich (MO, USA). Thermo Scientific™ PageRuler™ Plus Prestained Protein Ladder, Snakeskin™ dialysis tubing (10K MWCO, 22 mm) and SYPRO Orange dye were purchased from Thermo Fisher Scientific (MA, USA). Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside and Apigenin 7-O-(2G-Rhamnosyl)gentiobioside were purchased from ChemFaces (WUH, China)

4.2.2. Methods

4.2.2.1. Sb28GST Enzymatic Specific Activity and Inhibitory Assay

After the recombinant expression and purification of Sb28GST as shown in sections 2.2.2.1 and 2.2.2.2 the enzymatic specific activity of the Sb28GST and the inhibitory strength of the ligands quercetin and apigenin were tested using a reduced GSH-CDNB activity assay, where the catalytic activity of Sb28GST was spectrophotometrically monitored at varying enzyme concentrations as previously described (Zhan et al., 2010). The formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene (ε_{340} = 9600 $M^{-1}\cdot cm^{-1}$) using a Jasco V–630 UV/Vis spectrophotometer (Jasco, Japan) was monitored. The reaction buffer contained 100 mM sodium phosphate, 1 mM EDTA, 5 mM DTT, 0.02% NaN_3, 5 mM GSH, 1 mM CDNB (3% (v/v) ethanol), pH 6.5. Varying concentrations of Sb28GST (0 − 50 nM)

were incubated for 30 minutes in the absence or presence of 50 µM apigenin and quercetin. Triplicates of the reactions were carried out for 60 s. The specific activity was obtained by plotting linear progress curves corrected for non-enzymatic reaction rates.

The assay was additionally used to determine the half-maximal inhibitory [IC_{50}] of apigenin and quercetin. Varying concentrations of apigenin and quercetin, ranging from 0 – 1 mM were used with 50 nM Sb28GST. Experimental controls were conducted simultaneously in the absence of the protein. The data was then fitted using a four-parameter regression model equation 1:

$$y = \frac{bottom + (top - bottom)}{1 + 10^{((\log IC_{50} - x) \times hillslope)}} \qquad \textbf{Equation 1}$$

Where y is the enzyme activity, and the bottom and top are the minimum and maximum y - values for the curve asymptotes respectively. $\log IC_{50}$ is the concentration at which 50% of the enzyme activity was inhibited (nM), x the logarithm of inhibitor concentration and hillslope is the slope factor.

4.2.2.2. Extrinsic Fluorescence Spectroscopy Analysis of Sb28GST with Ligands

The use of fluorescence spectroscopy has aided in the structural characterization of proteins. The application of fluorescence in protein characterization exploits the principles of electromagnetic radiation interacting with matter, inducing energy transitions by way of electronic rearrangements (Pérez-Juste and Nieto Faza, 2015). When electrons of chromophores are excited to higher energy states upon light absorption, their relaxation back to an energy minimised state is accompanied with the emission of light representative of the energy lost by electrons (Sheehan, 2009). Extrinsic fluorescent dyes are used in combination with fluorescence spectroscopy due to their sensitivity to surrounding environments, allowing them to be effective probes used to determine their location on proteins. The extrinsic fluorophore ANS is most popularly used to probe the hydrophobic surfaces of proteins and consequently, protein active sites. While ANS has a low quantum yield in solution, when it binds non-covalently to a protein in a non-polar environment, its fluorescence exhibits both a hypsochromic blue shift and increased fluorescence intensity as shown in Figure 4.1 (Stryer, 1965, Uversky et al., 1998). Therefore, the amphiphilic nature of ANS allows

for its utilisation to monitor its competitive and non-competitive displacement by hydrophobic ligands bound to GST to probe ligand-protein interactions.

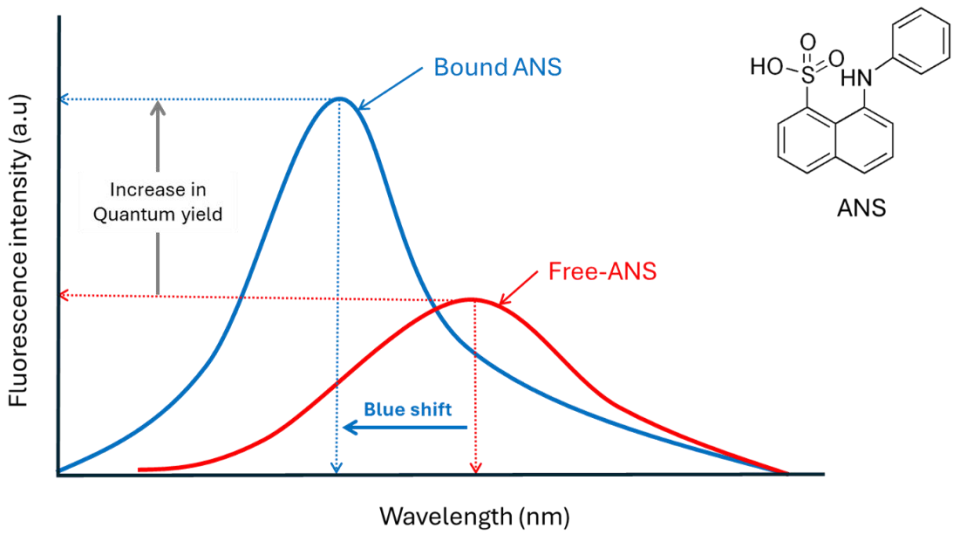


Figure 4. 1: ANS spectral shift and quantum yield changes anticipated for protein binding

To detect changes in protein conformation and probe hydrophobic clefts formed through Sb28GST interactions with apigenin and quercetin, a Jasco FP-6300 spectrofluorometer was used to measure the spectra using the Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan). The samples were prepared in 50 mM Tris-HCl pH 7.2, 1 mM EDTA and 1 mM DTT with 50 µM apigenin or quercetin. The samples contained 10 µM Sb28GST incubated with 50 µM of either GSH, CDNB, or PZQ co-substrates and additionally incubated with 200 µM of freshly prepared ANS. The samples were excited at 395 nm, and the emission spectra were collected at 400 - 650 nm with an excitation bandwidth of 5 nm, emission of 2.5 nm, data pitch of 1 nm, 1 cm pathlength and a scanning speed of 200 nm/min. The samples were collected in triplicate and were subsequently averaged and buffer corrected to determine protein–ligand interactions.

4.2.2.3. Thermal Shift Assay of Sb28GST with Ligands

To probe the thermal stability of Sb28GST in the presence of lead compounds, a fluorescence based thermal shift assay (TSA) was used. This assay monitors the fluorescent dye SYPRO Orange during protein unfolding to determine the melting temperature (*Tm*) of proteins in the presence and absence of ligands. This method, depicted in Figure 4.2, involves the systematic increase of temperature to a protein solution resulting in its unfolding. The denaturation that ensues leads to the exposure

of hydrophobic regions within the protein, typically buried within the protein's core (Huynh and Partch, 2015). The accessibility of these regions allows for SYPRO Orange to bind through hydrophobic interactions resulting in a significant increase in the quantum yield of the hydrophobic dye. As the protein continues to unfold and eventually aggregate due to intolerably high temperatures, the dye completely dissociates from the protein resulting in the decrease of the SYPRO Orange fluorescence emission (Kolobe et al., 2004). Thereafter, the melting temperature of the protein is determined using the signal curves obtained from the denaturation event. The *Tm* is finally calculated using the first derivative of the fluorescence emission with respect to temperature.

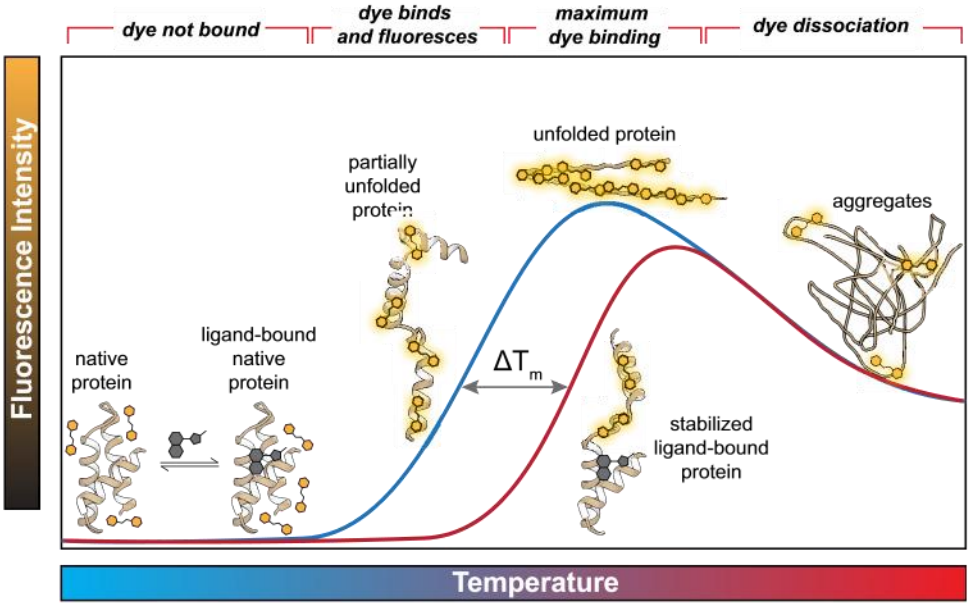


Figure 4. 2: Thermal unfolding profile of native and liganded protein denaturing in the presence of a fluorescent dye (Samuel et al., 2021)

The TSA was performed using a CFX96 Touch-Real-Time PCR detection system (Bio-Rad Laboratories, Inc., Hercules, CA, USA). This was monitored using a highly sensitive fluorescent dye called SYPRO Orange, whose quantum yield increases when hydrophobic patches are exposed during protein denaturation. Cocktails of 20 µM Sb28GST samples prepared in PBS containing 1 mM DTT and 1 mM EDTA supplemented with 0.5 mM GSH and 1 x SYPRO Orange dye were loaded into a 96-well qPCR plate. Additional samples containing apigenin [IC_{50}] and quercetin [IC_{50}] were also prepared and loaded into the qPCR plate. The melting curves were tracked from 10 ˚C to 95 ˚C with 0.5 ˚C increments at 10 s intervals. The relative fluorescence

unit (RFU) vs temperature (˚C) was recorded, and the derivative of the plot (-d(RFU)/dT) was used to determine the protein's *Tm*.

4.2.2.4. Isothermal Titration Calorimetry of Sb28GST with Ligands

To investigate the thermodynamic parameters associated with biomolecular interactions, biocalorimetry techniques are most suitably used. This isothermal titration calorimetry (ITC) technique offers a label free method to quantify the binding affinity (K_d), enthalpy (ΔH) and entropy (ΔS) occurring during protein–ligand binding events through the measurement of heat change (Freyer and Lewis, 2008). Figure 4.3 shows how the change in heat occurs as small aliquots of a ligand are titrated into a sample cell containing the protein of interest. The heat released or absorbed during these binding events is recorded against a reference cell to be plotted as an isotherm which reflects the change in heat against the molar ratio (Du et al., 2016). The isotherm is used to gain insight into how the hydrophobic or electrostatic interactions facilitate biomolecular interactions with respect to the structural and thermodynamic functioning of interacting biomolecules (Dutta et al., 2015).

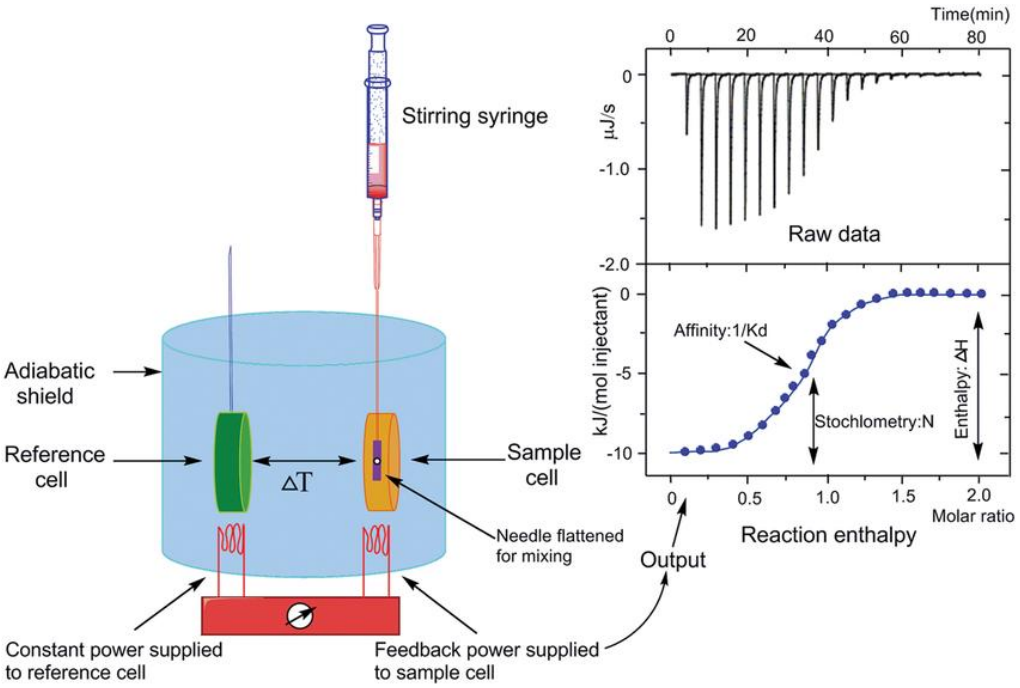


Figure 4. 3: Basic principle of isothermal titration calorimetry (Song et al., 2015)

The thermodynamics of interactions between Sb28GST with apigenin and quercetin were studied using the NanoITC instrument (TA Instruments, Delaware, USA). Freshly purified Sb28GST was dialysed in PBS containing 1 mM TCEP-HCl, 2 mM EDTA, 0.02 % (w/v) NaN3, pH 6.5. To avoid a buffer mismatch, the final dialysate was used to prepare 25 mM of both apigenin and quercetin as working stock and 25 µM of Sb28GST. The protein, ligand and buffer were degassed for 20 min at 0.3-0.5 atm. The reference cell contained Milli-Q® water whilst the sample cell contained a protein sample. The ITC experiment was performed at 20 ˚C. The ITC syringe volume which held the ligands, was 150 µL and the stir rate used was 250 rpm. A total of 19 injections were used with the initial injection volume being 2 µL, followed by subsequent 5 µL injections added every 300s. NanoAnalyze™ Software was used to analyze the data, allowing the integrated heats per an injection (in microjoules) to be corrected for the heats of dilution. The Data was then fitted to a multiple site binding model and visualized using the NanoAnalyze™ Software.

4.3. Results

4.3.1. Specific Activity and Inhibition Studies of Sb28GST

Figure 4.4 A shows the specific activity of Sb28GST with and without ligands apigenin and quercetin. The specific activity of Sb28GST was 19.9474 µmol.min^{-1}.mg^{-1}. In the presence of inhibitors, however the specific activity was drastically reduced. In the presence of apigenin the specific activity of Sb28GST was 2.5848 µmol.min^{-1}.mg^{-1}, while in the presence of quercetin, the specific activity was reduced to 0.0056 µmol.min^{-1}.mg^{-1}. In Figure 4.4 B, the IC_{50} of apigenin and quercetin were determined to be 0,130 and 0,120 mM respectively.

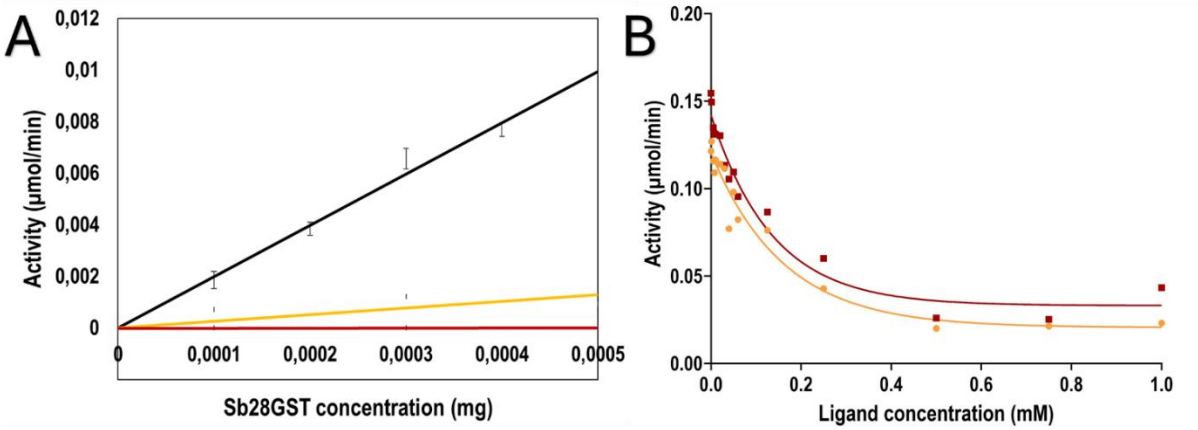


Figure 4. 4: Specific activity analysis and inhibition studies of Sb28GST with apigenin and quercetin. The GSH-CDNB conjugation assay was performed at 340 nm in the presence of apigenin (orange) and quercetin (Olveda et al.) by monitoring the formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene catalysed by Sb28GST. A: The specific activity of Sb28GST through varying concentrations of Sb28GST from 0 – 50 nM in the presence and absence of 50 µM apigenin and quercetin. B: Inhibition studies of Sb28GST to determine the IC_{50} of the ligands apigenin and quercetin. The ligand concentrations varied from 0 – 1 mM with 50 nM Sb28GST. All assays were performed in triplicate and the data were corrected for non-enzymatic rates.

4.3.2. Extrinsic Fluorescence Analyses of Sb28GST with Ligands

ANS fluorescence spectroscopy was performed to determine possible binding interactions between Sb28GST and the ligands apigenin and quercetin. When ANS binds to hydrophobic patches on proteins, a hypsochromic shift is typically observed in addition to increased fluorescent yield compared to free ANS. Figure 4.5 A shows the baseline fluorescence intensities of ANS incubated with Sb28GST and apigenin or quercetin, from which other spectra would be compared. The fluorescence of ANS-Sb28GST increased significantly compared to free ANS spectra. This was coupled with a blue shift in ANS emission from 520 nm to 492 nm in the presence of the protein. Upon the addition of GSH in Figure 4.5 B, the ANS - Sb28GST amplification was maintained with no notable fluorescence emission changes. The addition of CDNB and PZQ as co-substrates to the base samples led to a significant decrease in fluorescence intensity shown in Figure 4.5 C and D. Two anomalies were observed due to the presence of either quercetin or apigenin. In Figure 4.5 D the addition of quercetin showed a substantial hypsochromic shift to the ANS fluorescent intensity of Sb28GST with PZQ, while the opposite was observed to the ANS fluorescent intensity of Sb28GST with CDNB, shown in Figure 4.5 C.

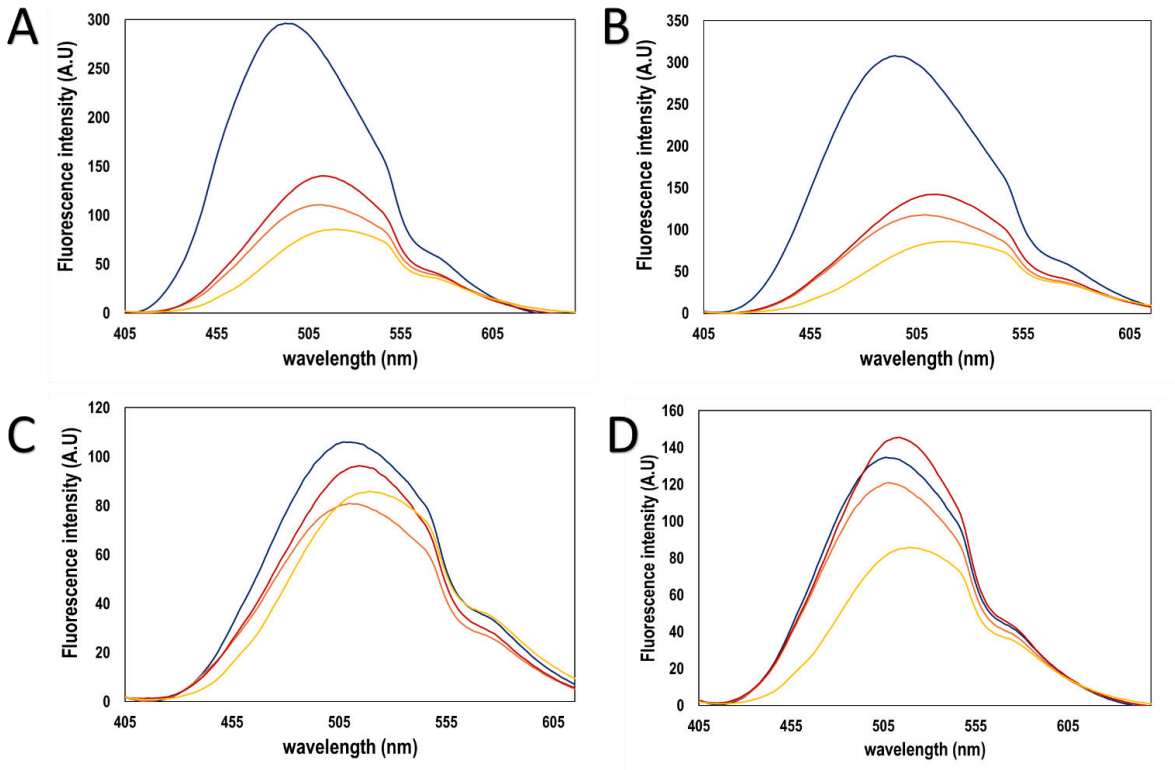


Figure 4. 5: Extrinsic ANS Fluorescence of Sb28GST. The emission spectra of 10 µM Sb28GST were collected at 400 - 640 nm upon excitation at 395 nm. The concentration of ANS used was 50 µM. Combinations of the Sb28GST and ANS (blue), apigenin, Sb28GST and ANS (orange), quercetin, Sb28GST and ANS (Olveda et al.), and free ANS (yellow) were recorded alongside varying co-substrates. A: No co-substrate. B: Glutathione. C: 1-chloro-2,4-dinitrobenzene. D: Praziquantel.

4.3.3. Thermal Stability of Sb28GST and Its Complexes

The fluorescent dye SYPRO orange has an affinity for hydrophobic patches, which typically result in increased quantum yield. This is seen upon protein denaturation where buried hydrophobic regions become exposed, allowing the dye to occupy more hydrophobic regions. Upon saturation, the dye becomes quenched indicated by decreasing fluorescence. The melting temperature of the Sb28GST was derived from the proteins' melting curve in the presence and absence of apigenin and quercetin. Figure 4.6 showed the thermal denaturation profile of Sb28GST in the presence and absence of apigenin and quercetin. The protein was stable at room temperature and began unfolding with increasing temperature, reaching a single transition peak at 70 ˚C. In the presence of apigenin, the T_m of the *Sb*28GST was reduced by 0.5 ˚C. Similarly, the T_m of Sb28GST in the presence of quercetin was reduced to 62 ˚C from 63.5 ˚C.

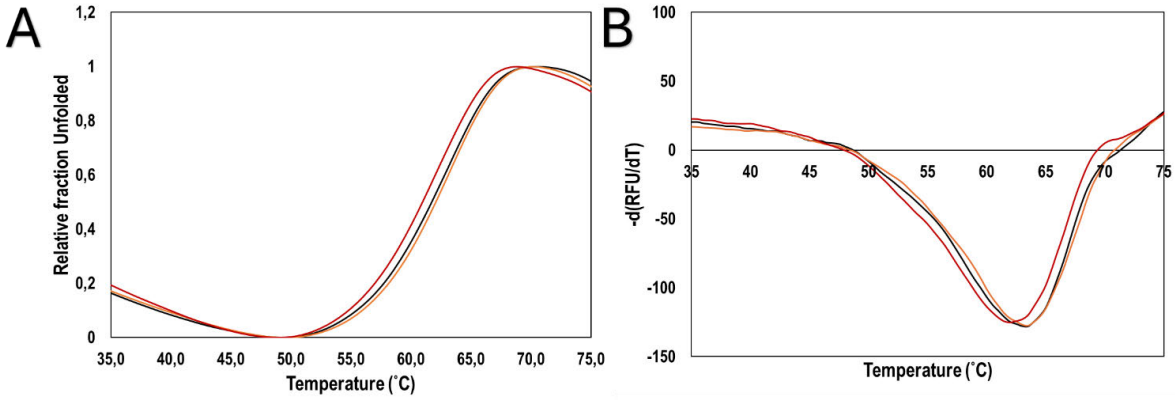


Figure 4. 6: Thermal stability profile of Sb28GST in the presence and absence of apigenin and quercetin. The readings of Sb28GST (black), Sb28GST - apigenin (orange) and Sb28GST - quercetin (Olveda et al.) were recorded in triplicates and averaged to generate individual thermal melt curves corrected with buffer blanks. A: The standardized RFU vs. temperature (˚C). B: The derived -d(RFU)/dT vs temperature (˚C).

4.3.4. ITC of Sb28GST with Apigenin and Quercetin

Isothermal titration calorimetry was used to determine the thermodynamic parameters of apigenin and quercetin binding to Sb28GST. The thermodynamic parameters of apigenin and quercetin binding to Sb28GST are summarised in Table 4.1. The thermograms and corresponding fitted data depicted in Figure 4.7 A show that the binding of apigenin to Sb28GST occurs endothermically as seen by the positive change in enthalpy (ΔH). Additionally, with Gibbs free energy (ΔG) of -46.75 and a TΔS˚ of 65.04 kJ/mol, this interaction appears to be a spontaneous reaction that is entropically driven. The binding of quercetin to Sb28GST is carried out in an exothermic reaction, as evidenced by the negative ΔH˚ and downward facing peaks in Figure 4.7 B. The negative ΔG˚ and positive TΔS˚ further show that this exothermic reaction occurs spontaneously and is entropically driven.

Table 4. 1: Thermodynamic parameters of Sb28GST interacting with either apigenin or quercetin

System	K_d (μM)	ΔH˚ (kJ/mol)	TΔS˚ (kJ/mol)	ΔG˚ (kJ/mol)
Sb28GST- apigenin	4.69×10^{-3}	18.29	65.04	-46.75
Sb28GST- quercetin	21.69	-17.69	8.48	-26.17

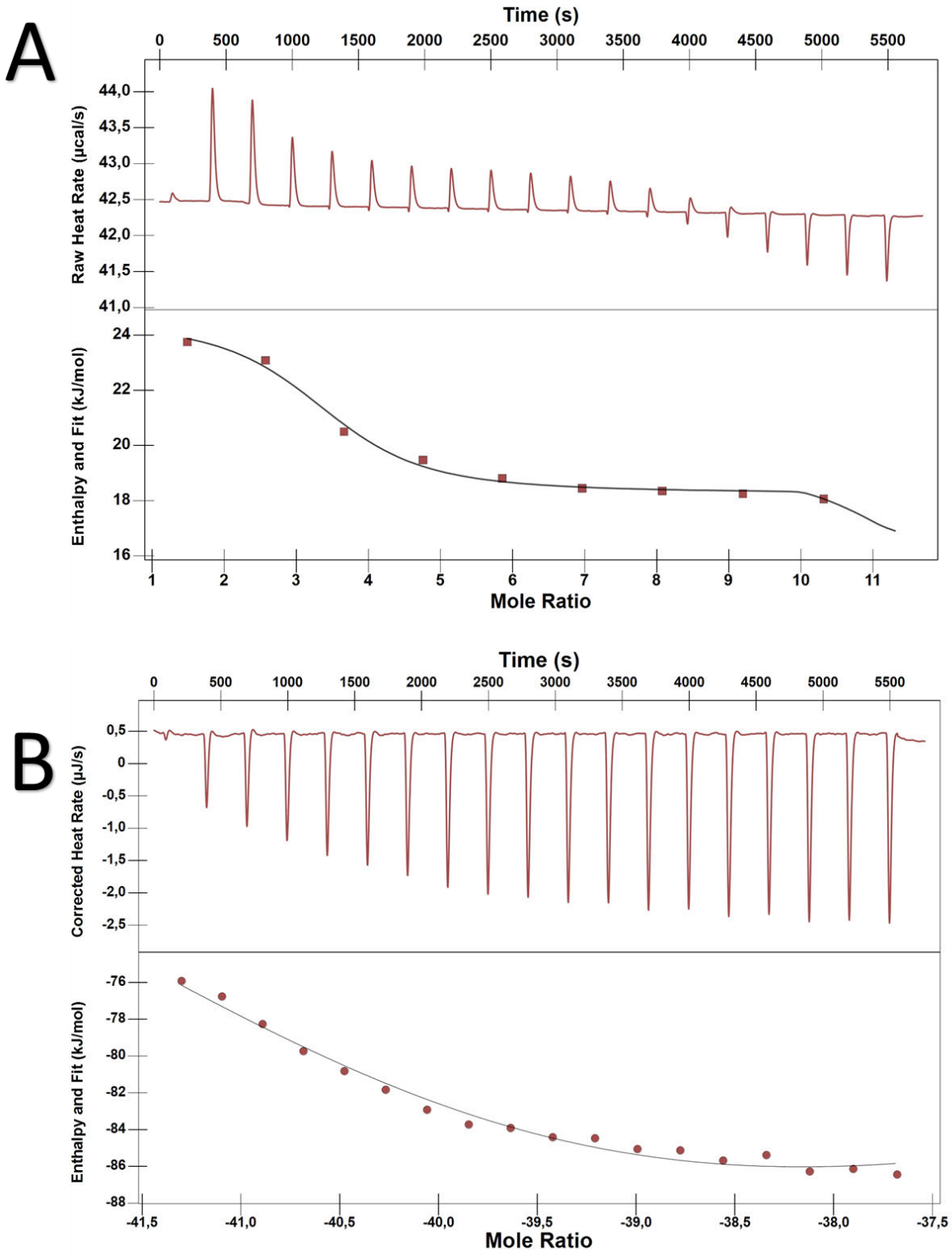


Figure 4. 7: The thermograms of Sb28GST interacting with either apigenin or quercetin. Using ITC, the interactions between A: 5 mM apigenin with Sb28GST and B: 7 mM quercetin with Sb28GST were monitored by titrating the ligands into 25 µM Sb28GST. NanoAnalyze software using the multiple sites model to fit the data to visualize the raw titration data depicted in the upper panel and fitted data in the lower panel.

4.4. Discussion

While the previous computational studies identified apigenin and quercetin as the lead compounds to execute flavonoid mediated Sb28GST inhibition, empirical validation is an important procedural part of the drug design and development pipeline.

Verification studies were done to empirically substantiate the ligands inhibitory potency of the lead compounds apigenin and quercetin. The catalytic activity of Sb28GST was monitored through the conjugation of CDNB to GSH and was found to have a specific activity of 19.94 $\mu mol.min^{-1}.mg^{-1}$ which is comparable to previously reported Sb28GST studies. In the presence of apigenin and quercetin the specific activity of Sb28GST was practically abolished, with the resulting activity reduced to 12.95% and 0.03% respectively. The potency at which the inhibition was achieved, determined through IC_{50} showed that quercetin is a more potent inhibitor compared to apigenin. As this was the first study to empirically examine Sb28GST inhibition, there is no research to date, available to compare the inhibitory potency of other potential Sb28GST inhibitors.

To determine the potential binding sites of these confirmed inhibitors, spectroscopic techniques were employed. Changes in the spectral properties of ANS by way of fluorescence intensity and/emission wavelength, in the presence of hydrophobic regions were used as probes to identify protein binding pockets (Hawe et al., 2008). The addition of ligands and co-substrates whose binding sites are known were incorporated to analyse the fluorescence displacement of ANS to determine the specific hydrophobic pockets that apigenin and quercetin bind to (Kinsley et al., 2008, Sluis-Cremer et al., 1996). The combination of Sb28GST with ANS saw a significant increase in the fluorescence intensity as typically expected, as the dye could occupy all accessible hydrophobic regions of the protein. The presence of apigenin and quercetin saw the reduction of the ANS fluorescence intensity, indicating the displacement of ANS within some hydrophobic pockets without identifying their specific binding sites. The lack of fluorescent changes observed upon the addition of GSH from the base profiles showed that ANS could not outcompete GSH for the highly specific G-site (Padi et al., 2021). The decrease in ANS fluorescence upon the addition of CDNB which binds at the GST H-site, showed its ability to displace ANS from that specific H-site hydrophobic pocket (Akumadu et al., 2020). The additional decrease in ANS fluorescence of GST in the presence of CDNB and quercetin indicated that the

inhibitor had a higher affinity for the H-site effectively displacing both ANS and CDNB. This trend was similarly observed with the displacement of PZQ and ANS from the GST L-site by apigenin, indicating its preferred binding site (Padi et al., 2021). An anomaly was observed upon the addition of CDNB and apigenin which induced a fluorescent decrease lower than that of free ANS. This indicated that the combination of both the inhibitor and CDNB completely displaced ANS from all available hydrophobic regions, effectively quenching ANS fluorescence. The second anomaly observed showed the ANS fluorescence of GST in the presence of PZQ and quercetin being greater than that of PZQ. This indicated that quercetin was able to bind to hydrophobic regions of the protein that did not directly overlap with the PZQ binding pocket, thus partially displacing ANS in hydrophobic pockets distinct from those occupied by PZQ. Both these anomalies indicate that the inhibitors might occupy both the H-site and L-site, corroborated by GST the amino acids the inhibitors were reported to interact with.

A quantitative approach was adopted to investigate the thermal stability of Sb28GST in the presence of apigenin and quercetin. The thermal stability of Sb28GST was slightly reduced by the addition of the flavonoid ligands. Although the reduction was not significant, it was however anticipated as previous research had shown that dietary flavonoids were sensitive to thermal conditions (Gao et al., 2022). While this is not entirely favourable, several investigations have proven that thermal exposure also promotes the flavonoids' bioactivity and bioavailability (Chaaban et al., 2017). While decreasing the thermal stability of Sb28GST might be seen as an advantage, as it could promote the denaturation of the protein, the extent to which the protein is destabilized is not significant enough to denature the protein. This does show however, that the ligands can be explored further as drug candidates.

The additional usage of heat analysis was employed to determine thermodynamic parameters of apigenin and quercetin binding to Sb28GST using ITC. This calorimetric technique allowed the determination of quantitative thermodynamic parameters through close analysis of heat changes and exchanges during protein-ligand interactions (Du et al., 2016). Both ligands exhibited spontaneous entropically driven binding to Sb28GST, showing that both flavonoids bound naturally to Sb28GST without the input of external energy. However, while apigenin bound endothermically to Sb28GST, quercetin bound exothermically. These different thermodynamic

signatures and varying heat exchanges highlight differences in molecular interactions occurring during binding. The endothermic reaction observed with apigenin suggests non-covalent interactions such as hydrophobic effects which compensate for the unfavourable enthalpy changes (Liu et al., 2012). On the other hand, the exothermic binding that occurred between quercetin and Sb28GST suggests favourable interactions such as van de Waals forces and hydrogen bonds formed (Perozzo et al., 2004). Interestingly, apigenin possessed the lowest dissociation constant, showing that even though quercetin was predicted to have a higher affinity, its appearance in only one of the Sb28GST computational systems showed that analysis leads to critical information being missed.

In conclusion, adopting computational methods to identify lead compounds for schistosomiasis management is an effective time saving approach implemented in drug discovery. Our study however, showed that it is crucial to be aware that the relationship between a protein's crystal structure and its dynamic behaviour in solution is relatively ambiguous. Diverging crystal structures of one Sb28GST protein could be obtained in monoclinic and orthorhombic space groups. The empirical validation of these drug candidates confirmed that apigenin and quercetin could significantly reduce Sb28GST activity. Extrinsic fluorescence studies showed that the inhibitors could bind both the hydrophobic H-site and the allosteric L-site at the dimer interface, with minimal impact on the protein’s thermal stability.

4.5. References

AKUMADU, B. O., PANDIAN, R., OLFSEN, J., WORTH, R., THULO, M., MENTOR, T., FANUCCHI, S., SAYED, Y., DIRR, H. W., ACHILONU, I. J. M. & PARASITOLOGY, B. 2020. Molecular basis of inhibition of Schistosoma japonicum glutathione transferase by ellagic acid: Insights into biophysical and structural studies. 240**,** 111319.

ANGORA, E. K., ALLIENNE, J.-F., REY, O., MENAN, H., TOURÉ, A. O., COULIBALY, J. T., RASO, G., YAVO, W., N'GORAN, E. K. & UTZINGER, J. J. P. 2020. High prevalence of Schistosoma haematobium× Schistosoma bovis hybrids in schoolchildren in Côte d'Ivoire. 147**,** 287-294.

CHAABAN, H., IOANNOU, I., CHEBIL, L., SLIMANE, M., GÉRARDIN, C., PARIS, C., CHARBONNEL, C., CHEKIR, L., GHOUL, M. J. J. O. F. P. & PRESERVATION 2017. Effect of heat processing on thermal stability and antioxidant activity of six flavonoids. 41**,** e13203.

CHARLIER, J., RINALDI, L., MUSELLA, V., PLOEGER, H. W., CHARTIER, C., VINEER, H. R., HINNEY, B., VON SAMSON-HIMMELSTJERNA, G., BĂCESCU, B. & MICKIEWICZ, M. J. P. V. M. 2020. Initial assessment of the economic burden of major parasitic helminth infections to the ruminant livestock industry in Europe. 182**,** 105103.

DE LA TORRE-ESCUDERO, E., PÉREZ-SÁNCHEZ, R., MANZANO-ROMÁN, R., OLEAGA, A. J. M. & PARASITOLOGY, B. 2017. Schistosoma bovis-host interplay: Proteomics for knowing and acting. 215**,** 30-39.

DEMLEW, B. A. & TESSMA, A. K. J. S. A. R. J. A. M. S. 2020. Review on bovine schistosomiasis and its associated risk factors. 2**,** 44-55.

DU, X., LI, Y., XIA, Y.-L., AI, S.-M., LIANG, J., SANG, P., JI, X.-L. & LIU, S.-Q. J. I. J. O. M. S. 2016. Insights into protein–ligand interactions: mechanisms, models, and methods. 17**,** 144.

DUTTA, A. K., RÖSGEN, J. & RAJARATHNAM, K. 2015. Using isothermal titration calorimetry to determine thermodynamic parameters of protein–glycosaminoglycan interactions. *Glycosaminoglycans.* Springer.
FREYER, M. W. & LEWIS, E. A. J. M. I. C. B. 2008. Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. 84**,** 79-113.
GAO, Y., XIA, W., SHAO, P., WU, W., CHEN, H., FANG, X., MU, H., XIAO, J. & GAO, H. J. C. O. I. F. S. 2022. Impact of thermal processing on dietary flavonoids. 48**,** 100915.
HABIG, W. H., PABST, M. J. & JAKOBY, W. B. J. J. O. B. C. 1974. Glutathione S-transferases: the first enzymatic step in mercapturic acid formation. 249**,** 7130-7139.
HAWE, A., SUTTER, M. & JISKOOT, W. J. P. R. 2008. Extrinsic fluorescent dyes as tools for protein characterization. 25**,** 1487-1499.
HUANG, H.-H., RIGOUIN, C. & L WILLIAMS, D. J. C. P. D. 2012. The redox biology of schistosome parasites and applications for drug development. 18**,** 3595-3611.
HUYNH, K. & PARTCH, C. L. J. C. P. I. P. S. 2015. Analysis of protein stability and ligand interactions by thermal shift assay. 79**,** 28.9. 1-28.9. 14.
KINSLEY, N., SAYED, Y., MOSEBI, S., ARMSTRONG, R. N. & DIRR, H. W. J. B. C. 2008. Characterization of the binding of 8-anilinonaphthalene sulfonate to rat class Mu GST M1-1. 137**,** 100-104.
KOLOBE, D., SAYED, Y. & DIRR, H. W. J. B. J. 2004. Characterization of bromosulphophthalein binding to human glutathione S-transferase A1-1: thermodynamics and inhibition kinetics. 382**,** 703-709.
LÉGER, E., BORLASE, A., FALL, C. B., DIOUF, N. D., DIOP, S. D., YASENEV, L., CATALANO, S., THIAM, C. T., NDIAYE, A. & EMERY, A. J. T. L. P. H. 2020. Prevalence and distribution of schistosomiasis in human, livestock, and snail populations in northern Senegal: a One Health epidemiological study of a multi-host system. 4**,** e330-e342.
LIU, S.-Q., TAN, D.-Y., ZHANG, K.-Q., JI, X.-L., TAO, Y. & FU, Y.-X. 2012. *Protein folding, binding and energy landscape: A synthesis*, INTECH Open Access Publisher Retrieved from.
MCTIGUE, M. A., WILLIAMS, D. R. & TAINER, J. A. J. J. O. M. B. 1995. Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase fromschistosoma japonicaand its complex with the leading antischistomal drug praziquantel. Elsevier.
MOLLA, G., TINTAGU, T., YASIN, A., ALEMU, B., ASSEN, A. A. & TADESSE, K. J. P. O. 2022. Bovine schistosomiasis in some selected areas of South wollo and oromia zones of Amhara region, north-east Ethiopia. 17**,** e0259787.
MÜLLER, S. J. M. 2015. Role and regulation of glutathione metabolism in Plasmodium falciparum. 20**,** 10511-10534.
OLVEDA, D. U., LI, Y., OLVEDA, R. M., LAM, A. K., CHAU, T. N., HARN, D. A., WILLIAMS, G. M., GRAY, D. J., ROSS, A. G. J. T. M. & SURGERY 2013. Bilharzia: pathology, diagnosis, management and control. 1.
PADI, N., AKUMADU, B. O., FAERCH, O., ALOKE, C., MEYER, V. & ACHILONU, I. J. B. 2021. Engineering a pseudo-26-kDa schistosoma glutathione transferase from bovis/haematobium for structure, kinetics, and ligandin studies. 11**,** 1844.
PÉREZ-JUSTE, I. & NIETO FAZA, O. J. S. E. I. O. C. T. S. F. T. R. T. 2015. Interaction of radiation with matter. 1-26.
PÉREZ-SÁNCHEZ, R., RAMAJO-HERNÁNDEZ, A., RAMAJO-MARTÍN, V. & OLEAGA, A. J. P. 2006. Proteomic analysis of the tegument and excretory-secretory products of adult Schistosoma bovis worms. 6**,** S226-S236.
PEROZZO, R., FOLKERS, G., SCAPOZZA, L. J. J. O. R. & TRANSDUCTION, S. 2004. Thermodynamics of protein–ligand interactions: history, presence, and future aspects. 24**,** 1-52.
SAMUEL, E. L., HOLMES, S. L. & YOUNG, D. W. J. J. O. C. 2021. Processing binding data using an open-source workflow. 13**,** 1-11.
SHEEHAN, D. 2009. *Physical biochemistry: principles and applications*, John Wiley & Sons.
SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. J. B. J. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. 360**,** 1-16.
SLUIS-CREMER, N., NAIDOO, N. N., KAPLAN, W. H., MANOHARAN, T. H., FAHL, W. E. & DIRR, H. W. J. E. J. O. B. 1996. Determination of a binding site for a non-substrate ligand in mammalian cytosolic glutathione S-transferases by means of fluorescence-resonance energy transfer. 241**,** 484-488.

SONG, C., ZHANG, S. & HUANG, H. J. F. I. M. 2015. Choosing a suitable method for the identification of replication origins in microbial genomes. 6**,** 1049.
STRYDOM, T., LAVAN, R. P., TORRES, S. & HEANEY, K. J. A. 2023. The economic impact of parasitism from nematodes, trematodes and ticks on beef cattle production. 13**,** 1599.
STRYER, L. J. J. O. M. B. 1965. The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: A fluorescent probe of non-polar binding sites. 13**,** 482-495.
TORRES-RIVERA, A. & LANDA, A. J. A. T. 2008. Glutathione transferases from parasites: a biochemical view. 105**,** 99-112.
UVERSKY, V. N., WINTER, S., LÖBER, G. J. B. E. B. A.-P. S. & ENZYMOLOGY, M. 1998. Self-association of 8-anilino-1-naphthalene-sulfonate molecules: spectroscopic characterization and application to the investigation of protein folding. 1388**,** 133-142.
YASSIN, Z., ORTIZ-SALMERÓN, E., GARCíA-MAROTO, F., BARÓN, C., GARCíA-FUENTES, L. J. B. E. B. A.-P. & PROTEOMICS 2004. Implications of the ligandin binding site on the binding of non-substrate ligands to Schistosoma japonicum-glutathione transferase. 1698**,** 227-237.
ZHAN, B., PERALLY, S., BROPHY, P. M., XUE, J., GOUD, G., LIU, S., DEUMIC, V., DE OLIVEIRA, L. M., BETHONY, J., BOTTAZZI, M. E. J. I. & IMMUNITY 2010. Molecular cloning, biochemical characterization, and partial protective immunity of the heme-binding glutathione S-transferases from the human hookworm Necator americanus. 78**,** 1552-1563.

Chapter 5

Overall Discussion, Summary and Conclusion

5.1. Overall Discussion

Efficacious therapy intervention strategies are urgently needed to achieve the World Health Organisation's (WHO) goal of "a world free of schistosomiasis" by 2030, after missing the previously set 2025 goal (Organization, 2013, Organization, 2020, Organization, 2022). With the WHO's revised roadmap to achieve this goal by adopting a One Health approach, greater emphasis has now been placed on animal health and environmental links (Organization, 2020). Schistosomiasis, caused by trematodes of the parasite *Schistosoma* spp has devastating medical and veterinary impacts, affecting over a quarter of a billion people globally and over 165 million livestock (de la Torre-Escudero et al., 2017, Moné et al., 1999). This prevalence severely threatens the profitability of beef agriculture and cripples protein food security together with the livelihood of cattle farmers and industry investors (Strydom et al., 2023). However, neglect of the zoonotic abilities of *Schistosoma* spp has allowed many interspecies hybridization cases to arise, expanding health risks through evolved disease transmissions and host spectrum expansion (Angora et al., 2020, Boon et al., 2018, Mogaji et al., 2023).

At present, Praziquantel (PZQ) is the only available anti-*Schistosome* vehicle for the treatment of schistosomiasis (Léger et al., 2020). However, the drug's efficacy is limited to the host antibody response caused by adult egg-producing *Schistosomes*, excluding immature *Schistosomes*. For that reason, PZQ administration typically requires repeat treatment, increasing the risk of drug resistance (Olveda et al., 2013). Additionally, the availability of veterinary formulations of PZQ is rather scant for many regions in Sub-Saharan Africa (Gower et al., 2017).

In this study, increasingly popular CADD methods were used as a rational approach to schistosomiasis intervention. We utilised polymorphic space group crystals to determine their impact on ligand selectivity and, consequently, computational drug discovery. We secondly sought to discover novel flavonoid inhibitors to Sb28GST that were identified by structure based HTVS and MDS for schistosomiasis treatment. Our

in-silico studies showed that crystal packing and, ultimately crystal space groups impact proteins' dynamic potential and, consequently, ligand selectivity.

This study successfully integrated structural biology, computational drug design, and biophysical assays to identify potential inhibitors of *Schistosoma bovis* 28 kDa glutathione S-transferase (Sb28GST), a critical enzyme in parasite detoxification and survival. The overexpression and crystallization of Sb28GST in a monoclinic alternative space group allowed for detailed structural characterisation, revealing key features of the enzyme's active and allosteric sites. High-throughput virtual screening (HTVS) coupled with molecular dynamics (MD) simulations facilitated the identification of flavonoid compounds exhibiting strong binding affinities. These computational predictions were validated through GST enzymatic assays, where selected ligands demonstrated inhibitory effects. The findings support the feasibility of Sb28GST as a therapeutic target and highlight the utility of combining crystallographic and computational approaches in antiparasitic drug discovery (Angelucci et al., 2005, Makumbe et al., 2024, Müller, 2015).

Importantly, this study addressed a significant gap in structure-based drug design by evaluating the influence of protein crystal polymorphism on ligand binding predictions. Comparative MD simulations of Sb28GST models from different space groups revealed that crystal packing variations can alter protein conformational dynamics and affect ligand binding modes and affinities. This underscores the need for considering multiple crystal forms when performing CADD to avoid oversimplification and potential misidentification of lead compounds (Beran, 2016, Yeates and Kent, 2012). Our findings align with previous reports that polymorphic crystal forms can induce differential protein flexibility, solvent channel architecture, and intermolecular interactions, all of which can impact druggability assessments (Du et al., 2016, Hospital et al., 2015).

The discovery of lead compounds capable of inhibiting Sb28GST enzymatic activity also reinforces the therapeutic relevance of targeting parasite-specific redox metabolism. Given the absence of phase I detoxification enzymes in *Schistosomes*, GSTs serve as the primary defence mechanism against host-derived oxidative stress and xenobiotic threats (Akumadu et al., 2020, Torres-Rivera and Landa, 2008). The flavonoid inhibitors identified in this study not only disrupted GST catalytic function but

may also impair the parasite's ligandin activity, potentially enhancing the susceptibility of *S. bovis* to host immune responses and environmental stressors (Grevelding et al., 2018). Moreover, the structural differences between host and parasite GSTs observed in this work support the possibility of designing selective inhibitors that minimize host toxicity (Sheehan et al., 2001).

Overall, this research advances the understanding of *Schistosome* GST structure-function relationships, provides a robust framework for incorporating crystallographic polymorphism into drug discovery workflows, and identifies promising molecular scaffolds for further development. Future studies should expand empirical validation of lead compounds in parasite culture and *in vivo* models, assess resistance potential, and explore structure-activity relationships to optimize pharmacological properties. By integrating empirical and computational strategies, this work contributes to addressing the urgent need for novel, effective anti-schistosomal therapeutics, especially in the context of emerging praziquantel resistance (Vale et al., 2017).

5.2. Conclusion

This thesis demonstrates the successful application of an integrated structure-based drug discovery approach to target *Schistosoma bovis* 28 kDa glutathione S-transferase (Sb28GST), a key enzyme in parasite redox metabolism. By combining protein crystallography, high-throughput virtual screening, molecular docking, and molecular dynamics simulations with enzymatic inhibition assays, the study identified flavonoid-based lead compounds exhibiting promising inhibitory potential. Importantly, it also highlighted how protein crystal polymorphism can influence conformational dynamics and ligand binding predictions, a factor often overlooked in computer-aided drug design (CADD). The findings validate Sb28GST as a viable therapeutic target and provide insights into the structural basis for selective inhibition. This research contributes to the broader goal of developing novel anti-schistosomal agents, addressing the increasing challenge of praziquantel resistance, and enhancing the methodological framework for antiparasitic drug discovery.

While this study has provided valuable insights, certain limitations should be acknowledged. First, although molecular dynamics simulations provided detailed predictions of ligand binding and stability, these computational models cannot fully capture the complexity of the parasite's intracellular environment. Second, the *in vitro*

inhibition assays, while robust, were limited to biochemical validation and did not include cellular or whole-organism assays to assess compound efficacy in live parasites. Additionally, only a single GST isoform from *S. bovis* was investigated, and the potential cross-reactivity or selectivity of the identified inhibitors against other *Schistosome* species or host GSTs was not assessed. Finally, although the influence of crystal polymorphism was explored, the structural models used still represent approximations and may not fully reflect the dynamic conformational landscape of Sb28GST under physiological conditions.

5.3. Future Directions

Building on these promising results, future research should focus on extending biochemical and biophysical validation of the identified flavonoid inhibitors through kinetic studies and mechanistic analyses. Evaluating the efficacy of lead compounds in *in vitro* parasite cultures and *in vivo* infection models will be essential to confirm antiparasitic activity and assess safety profiles. Structure-activity relationship studies should be undertaken to optimize compound potency, specificity, and pharmacokinetic properties. Expanding the investigation to include additional GST isoforms from other *Schistosome* species could identify broad-spectrum inhibitors. Moreover, incorporating more advanced molecular dynamics simulations that account for solvent effects, membrane interactions, and parasite-specific intracellular conditions may improve predictive accuracy. Ultimately, translating these findings toward clinical candidate development will require interdisciplinary collaboration and engagement with public health stakeholders in endemic regions.

5.4. References

AKUMADU, B. O., PANDIAN, R., OLFSEN, J., WORTH, R., THULO, M., MENTOR, T., FANUCCHI, S., SAYED, Y., DIRR, H. W., ACHILONU, I. J. M. & PARASITOLOGY, B. 2020. Molecular basis of inhibition of Schistosoma japonicum glutathione transferase by ellagic acid: Insights into biophysical and structural studies. 240**,** 111319.

ANGELUCCI, F., BAIOCCO, P., BRUNORI, M., GOURLAY, L., MOREA, V. & BELLELLI, A. J. S. 2005. Insights into the catalytic mechanism of glutathione S-transferase: the lesson from Schistosoma haematobium. 13**,** 1241-1246.

ANGORA, E. K., ALLIENNE, J.-F., REY, O., MENAN, H., TOURÉ, A. O., COULIBALY, J. T., RASO, G., YAVO, W., N'GORAN, E. K. & UTZINGER, J. J. P. 2020. High prevalence of Schistosoma haematobium× Schistosoma bovis hybrids in schoolchildren in Côte d'Ivoire. 147**,** 287-294.

BERAN, G. J. J. C. R. 2016. Modeling polymorphic molecular crystals with electronic structure theory. 116**,** 5567-5613.

BOON, N. A., VAN DEN BROECK, F., FAYE, D., VOLCKAERT, F. A., MBOUP, S., POLMAN, K. & HUYSE, T. J. P. 2018. Barcoding hybrids: heterogeneous distribution of Schistosoma haematobium× Schistosoma bovis hybrids across the Senegal River Basin. 145**,** 634-645.

DE LA TORRE-ESCUDERO, E., PÉREZ-SÁNCHEZ, R., MANZANO-ROMÁN, R., OLEAGA, A. J. M. & PARASITOLOGY, B. 2017. Schistosoma bovis-host interplay: Proteomics for knowing and acting. 215**,** 30-39.

DU, X., LI, Y., XIA, Y.-L., AI, S.-M., LIANG, J., SANG, P., JI, X.-L. & LIU, S.-Q. J. I. J. O. M. S. 2016. Insights into protein–ligand interactions: mechanisms, models, and methods. 17**,** 144.

GOWER, C. M., VINCE, L., WEBSTER, J. P. J. T. O. T. R. S. O. T. M. & HYGIENE 2017. Should we be treating animal schistosomiasis in Africa? The need for a One Health economic evaluation of schistosomiasis control in people and their livestock. 111**,** 244-247.

GREVELDING, C. G., LANGNER, S. & DISSOUS, C. J. T. I. P. 2018. Kinases: molecular stage directors for schistosome development and differentiation. 34**,** 246-260.

HOSPITAL, A., GOÑI, J. R., OROZCO, M., GELPÍ, J. L. J. A., BIOINFORMATICS, A. I. & CHEMISTRY 2015. Molecular dynamics simulations: advances and applications. 37-47.

LÉGER, E., BORLASE, A., FALL, C. B., DIOUF, N. D., DIOP, S. D., YASENEV, L., CATALANO, S., THIAM, C. T., NDIAYE, A. & EMERY, A. J. T. L. P. H. 2020. Prevalence and distribution of schistosomiasis in human, livestock, and snail populations in northern Senegal: a One Health epidemiological study of a multi-host system. 4**,** e330-e342.

MAKUMBE, H. H., PANDIAN, R., VALLI, A., SAYED, Y. & ACHILONU, I. J. J. O. M. S. 2024. Biophysical characterization, crystallization, and solution of the first crystal structure of the 28 kDa-Schistosoma bovis glutathione transferase. 1298**,** 136979.

MOGAJI, H. O., OMITOLA, O. O., BAYEGUN, A. A., EKPO, U. F. & TAYLOR-ROBINSON, A. W. J. Z. D. 2023. Livestock reservoir hosts: an obscured threat to control of human schistosomiasis in Nigeria. 3**,** 52-67.

MONÉ, H., MOUAHID, G. & MORAND, S. J. A. I. P. 1999. The distribution of Schistosoma bovis Sonsino, 1876 in relation to intermediate host mollusc–parasite relationships. 44**,** 99-138.

MÜLLER, S. J. M. 2015. Role and regulation of glutathione metabolism in Plasmodium falciparum. 20**,** 10511-10534.

OLVEDA, D. U., LI, Y., OLVEDA, R. M., LAM, A. K., CHAU, T. N., HARN, D. A., WILLIAMS, G. M., GRAY, D. J., ROSS, A. G. J. T. M. & SURGERY 2013. Bilharzia: pathology, diagnosis, management and control. 1.

ORGANIZATION, W. H. 2013. Schistosomiasis: progress report 2001-2011, strategic plan 2012-2020. *Schistosomiasis: progress report 2001-2011, strategic plan 2012-2020.*

ORGANIZATION, W. H. 2020. Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030. *Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030.*

ORGANIZATION, W. H. 2022. *WHO guideline on control and elimination of human schistosomiasis*, World Health Organization.

SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. J. B. J. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. 360**,** 1-16.

STRYDOM, T., LAVAN, R. P., TORRES, S. & HEANEY, K. J. A. 2023. The economic impact of parasitism from nematodes, trematodes and ticks on beef cattle production. 13**,** 1599.

TORRES-RIVERA, A. & LANDA, A. J. A. T. 2008. Glutathione transferases from parasites: a biochemical view. 105**,** 99-112.

VALE, N., GOUVEIA, M. J., RINALDI, G., BRINDLEY, P. J., GÄRTNER, F., CORREIA DA COSTA, J. M. J. A. A. & CHEMOTHERAPY 2017. Praziquantel for schistosomiasis: single-drug metabolism revisited, mode of action, and resistance. 61**,** 10.1128/aac. 02582-16.

YEATES, T. O. & KENT, S. B. J. A. R. O. B. 2012. Racemic protein crystallography. 41**,** 41-61.

5.5. Appendix I: Published Articles

Results in Chemistry 16 (2025) 102288

Contents lists available at ScienceDirect

Results in Chemistry

journal homepage: www.sciencedirect.com/journal/results-in-chemistry

Crystal enigma: Understanding diverse protein conformational dynamics, ligand selectivity and interaction in multi-space group crystals using computational modelling

Mbalenhle Mfeka [a,b], Olalekan Onisuru [b], Ramesh Pandian [b], Yasien Sayed [b], Thandeka Khoza [a], Ikechukwu Achilonu [b,*]

[a] Department of Biochemistry, School of Life Sciences, College of Agriculture, Engineering & Science, University of KwaZulu-Natal, Pietermaritzburg, 3209, South Africa
[b] Protein Structure-Function and Research Laboratory, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg 2000, South Africa

ARTICLE INFO

Keywords:
Schistosoma bovis
Polymorphs
High-throughput virtual screening
Molecular dynamics simulations
Empirical ligand

ABSTRACT

Empirical structural methods have been instrumental in drug design, but they often fall short of capturing protein dynamics. To address this limitation, computer-aided drug design (CADD) is essential. Given that *Schistosoma bovis* 28 kDa-Glutathione *S*-Transferase (*Sb*28GST) is a promising anti-schistosome drug target, a careful selection of crystal structure space groups used for CADD is emphasized in this study. *Sb*28GST was successfully over-expressed and purified to grow protein crystals with a resolution of 2.4 Å in an orthorhombic space group system.

Results in Chemistry 16 (2025) 102288

Contents lists available at ScienceDirect

Results in Chemistry

journal homepage: www.sciencedirect.com/journal/results-in-chemistry

Crystal enigma: Understanding diverse protein conformational dynamics, ligand selectivity and interaction in multi-space group crystals using computational modelling

Mbalenhle Mfeka [a,b], Olalekan Onisuru [b], Ramesh Pandian [b], Yasien Sayed [b], Thandeka Khoza [a], Ikechukwu Achilonu [b,*]

[a] Department of Biochemistry, School of Life Sciences, College of Agriculture, Engineering & Science, University of KwaZulu-Natal, Pietermaritzburg, 3209, South Africa
[b] Protein Structure-Function and Research Laboratory, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg 2000, South Africa

ARTICLE INFO

Keywords:
Schistosoma bovis
Polymorphs
High-throughput virtual screening
Molecular dynamics simulations
Empirical, ligand
Binding

ABSTRACT

Empirical structural methods have been instrumental in drug design, but they often fall short of capturing protein dynamics. To address this limitation, computer-aided drug design (CADD) is essential. Given that *Schistosoma bovis* 28 kDa-Glutathione *S*-Transferase (*Sb*28GST) is a promising anti-schistosome drug target, a careful selection of crystal structure space groups used for CADD is emphasized in this study. *Sb*28GST was successfully over-expressed and purified to grow protein crystals with a resolution of 2.4 Å in an orthorhombic space group system. High-throughput virtual screening (HTVS) of a flavonoid compound was performed on monoclinic 8ALS and orthorhombic 8BHZ Sb28GST to identify potential ligands. The results showed a diverse selection of hit compounds with apigenin 7-O-(2G-rhamnosyl)gentiobioside (apigenin) being the common ligand. However, quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside (quercetin) showed the highest affinity to 8ALS with a Glide gscore of −15.66 kcal/mol. Also, 500-ns molecular dynamics simulations (MDS) of the 8ALS-*Sb*28GST and 8BHZ-*Sb*28GST apo systems, as well as their corresponding apigenin and quercetin complexes, show two distinct trajectories which reveal significant differences in their dynamic behaviour and also showed varying interactions. This highlights the need for collective assessment of protein polymorphs to comprehensively understand protein dynamics. The inhibitory potency of apigenin and quercetin via a GSH-CDNB conjugation assay confirmed that they significantly reduced the specific enzyme activity of *Sb*28GST with both ligands having half-maximal inhibitory concentration (IC_{50}) of 0.13 mM. Extrinsic fluorescence studies and thermal shift assay indicated that these compounds bind to the hydrophobic H-site and allosteric L-site at the dimer interface and have a minimal effect on the protein's thermal stability.

1. Introduction

Schistosomiasis, caused by zoonotic trematodes from *Schistosoma* species, is the second most neglected tropical disease with a high socioeconomic impact after malaria. *Schistosoma bovis* severely affects ruminants, leading to high morbidity, mortality, reduced productivity, growth, and reproductive performance, crippling farmers' economic output and increasing livestock vulnerability [1,2]. Schistosomiasis intervention in animals is neglected, with praziquantel (PZQ) being the only treatment since the 1980s. However, PZQ only targets humans, leaving all other zoonotic reservoirs untreated and found ineffective against immature schistosome larvae [3]. An additional emerging threat schistosomiasis infections pose to global health is the increasing hybridization of *Schistosoma haematobium* and *Schistosoma bovis*. These hybrids, which are found in humans and ruminants (with infected children and snail hosts shedding the *S. haematobium* - *S. bovis* hybrids), complicate disease control efforts [4]. This evolutionary zoonotic adaptation will threaten control interventions, alter disease dynamics, and transmission if allowed to continue. Hence, these highlight the urgent need for *S. bovis* control to reduce the economic loss suffered by

* Corresponding author at: Protein Structure-Function and Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg 2000, South Africa.
E-mail address: Ikechukwu.Achilonu@wits.ac.za (I. Achilonu).

https://doi.org/10.1016/j.rechem.2025.102288
Received 12 December 2024; Accepted 21 April 2025
Available online 30 April 2025

farmers and prevent the spread of hybrid schistosomes from human vectors.

Schistosoma lacks cytochrome P450 and relies on glutathione transferases (GSTs) to conjugate glutathione to xenobiotics. These GSTs are the parasite's primary defence, aiding in toxin storage, transport, and protection against oxidative damage. They also prevent the harmful accumulation of hydrophobic molecules at hydrophilic sites within cells [5–7]. Gene knockout experiments revealed that silencing *Schistosoma* 28 kDa GST impairs egg growth due to redox imbalance and reduces prolificacy [8]. This is because GSTs bind xenobiotics with high affinity, and inhibiting this detoxification exposes the parasite to immune attack [7,9]. Their ligandin function, which inhibits catalysis, makes them promising drug targets for schistosomiasis treatment [6,10].

The empirical discovery of novel lead compounds is typically slow and costly, but computer-aided drug design (CADD) has accelerated and reduced costs in drug development [11,12]. CADD uses high-power atomic-level computers and Newtonian physics to simulate protein atom movements over time [13]. This approach reveals molecular functions and intermolecular dynamic interactions between biomolecules, offering insights beyond static crystallized protein snapshots. CADD has become an essential tool in the modern drug discovery pipeline, improving accuracy and efficiency in understanding biomolecular dynamics [14]. Molecular dynamics (MD) simulations are crucial in drug discovery, providing a dynamic perspective on biomolecular interactions beyond static docking models. High throughput virtual screening (HTVS) and molecular dynamic (MD) simulations were the CADD tools utilised in this research. High throughput virtual screening digitalises the screening of various compounds, identifying interactions with all available binding sites with the protein of interest [15]. While molecular docking predicts the binding orientation of small molecules to their protein targets, MD simulations refine these interactions by accounting for conformational changes, stability, and environmental effects. They help evaluate binding affinities, predict long-term stability, and assess the impact of point mutations on drug-target interactions. This allows for calculation of the affinity between various ligands with the protein target by sampling all possible positions, orientations and conformations of the screened compounds relative to the binding sites of the target protein [16,17]. The binding potential is scored using the most energetically favourable binding poses, which are then ranked in order of most probable [18]. Molecular dynamic simulations are then used to compute further inferences from the predicted top-binding ligands [19]. This method mimics the individual motions of complex molecular systems in their physiological environments at an atomic level, providing insights into the dynamic behaviour of biomolecular complexes [14]. Therefore, combining molecular docking with MD simulations enhances drug design by validating binding poses and optimizing potential inhibitors for therapeutic applications.

Computational simulations of biomolecules often rely on X-ray crystallography models, but caution is needed when selecting 3D protein structures [14]. Researchers often assume target proteins crystallized in different space groups behave similarly in *in-silico* modelling when predicting a lead compound. However, empirical testing suggests varying symmetries and arrangements could affect results. This has led to inaccurate predictions of lead compound performance in empirical testing [20]. The impact of polymorphic protein space groups on conformational dynamics and ligand selectivity is underexplored. Crystal polymorphs' varying symmetries affect solvent content and distribution, potentially leading to one-dimensional interpretations in molecular dynamics simulations [21]. This study aims to identify novel flavonoid inhibitors targeting *Schistosoma bovis* 28 kDa GST and assess how polymorphic protein crystals affect ligand selectivity. This is because pooling data from various crystal polymorphs will provide deeper insights into protein conformational states and diverse molecular interactions, offering a more comprehensive analysis than studying a single crystal space group.

2. Materials and methods

2.1. Materials

All chemicals and reagents used were of analytical grade. Vector synthesis was carried out by GenScript (NJ, USA). While crystal screening buffers were purchased from Hampton Research (CA, USA), other chemicals such as ampicillin, Isopropyl-β-D-thiogalactopyronoside (IPTG), Coomassie Blue-G250 dye, reduced glutathione (GSH), 1-chloro-2,4-dinitrobenzene (CDNB), anilino-1-naphthalene sulphonate (ANS) and dimethyl sulfoxide (DMSO) were all purchased from Sigma Aldrich (MO, USA). SYPRO Orange dye and PageRuler™ Plus Prestained Protein Ladder were purchased from Thermo Fisher Scientific (MA, USA). *Escherichia coli* (*E. coli*) T7 Express chemically competent cells were purchased from New England Biolabs (Massachusetts, USA). The 15-well EasyXtal® plates were purchased from Qiagen (Hilden, Germany). Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside and Apigenin 7-O-(2G-Rhamnosyl)gentiobioside were purchased from ChemFaces (WUH, China).

2.1.1. SbGST recombinant expression and purification

For the recombinant expression of *Sb*28GST, *E. coli* T7 expression cells were used following the protocol previously discussed in [22]. In this study, 0.5 mM IPTG was used to induce successful overexpression of *Sb*28GST, after which cells were harvested through centrifugation at 7500 x*g*, 4 °C for 10 min. The resulting pellet was resuspended in phosphate-buffered saline (PBS), 25 mM imidazole, 0.02 % NaN_3, pH 7.2 and lysed through sonification. The *Sb*28GST protein was subsequently purified using immobilized Ni^{2+} affinity chromatography, and soluble *Sb*28GST protein was eluted using PBS, 500 mM imidazole, 0.02 % NaN_3, pH 7.2. The purity and homogeneity of the purified protein were determined by Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE).

2.1.2. Crystallization of Sb28GST

Soluble and pure *Sb*28GST protein (15 mg/mL in PBS, pH 7.5) was first screened for crystallization conditions using Index Screens I and II (Hampton Research) through a sitting drop vapour diffusion method on a 96-well microplate. Apo *Sb*28GST was subsequently crystallized within 48–72 h using hanging drop methods. The diffraction quality crystals (long needle-shaped) were obtained with 1:1 of protein and crystallization buffer, i.e. 2 μL drop of protein solution and 2 μL drop of the crystallization/reservoir buffer (2.1 M ammonium sulfate, 100 mM Tris pH 7.5, and 5 mM β-mercaptoethanol). Long needle-shaped crystals of apo *Sb*28GST grew within a few days.

2.1.3. Data collection and Sb28GST crystal structure solution

A single *Sb*28GST crystal was mounted in a nylon loop and soaked in a cryoprotectant paratone (Parabar 10,312; Hampton Research) solution. This was flash-frozen in liquid nitrogen at 100 K. The X-ray diffraction data for *Sb*28GST was collected using a rotating anode X-ray source (Cu Kα; 1.5418 Å) and the pixel Bruker D8 Venture Bio PHOTON III area detector with a crystal-to-detector distance of 41 mm. The unit cell and complete data were collected using the PROTEUM4 software suite [23]. The data was integrated with SAINT and reduced and scaled using SADABS. SCALEPACK and AIMLESS/POINTLESS, which are built into the PROTEUM4 suite, were used to process the diffraction data [24,25]. The structure was determined by molecular replacement using the PHASER program built into PHENIX software [26]. The search model used for molecular replacement was 1OE7. Model building was performed with the WinCoot [27]. Water molecules were added during refinement using the "Flat Bulk Solvent Model" procedure on Coot. The protein structure was further refined using PHENIX followed by structural validation using PROCHECK [28,29], with the final structure deposited into protein data bank (PDB) as 8BHZ. PyMOL was used to generate the images [30].

2.2. Computational modelling

2.2.1. Computer hardware

All computational studies were performed using two high-power computing units. A Windows OS desktop housing the Schrödinger Maestro v13.0 software (used for MDS) and equipped with an AMD RYZEN Threadripper 1950× Processor, 16 cores/ 32 threads, an Asus Rog Strix X399-E Gaming Ryzen motherboard and a 4.0 GHz Precision Boost X399 chipset was used for the HTVS. Similarly, a desktop with 4 TB internal SSD storage and 6 GB/s 3D NAND technology speeds, 64 GB (16GB x 4) Quad-channel DDR4 desktop RAM, and an 11 GB GDDR6 MSI GeForce RTX 2080 Ti graphics card (GPU) was used for the MDS. Also, an Ubuntu OS desktop computer equipped with an AMD Threadripper 3990× processor with 64 cores/ 128 threads and an MSI TRX40 PRO 10 G motherboard, all with a 4.3 GHz Turbo boost clock and TRX40 chipset and another desktop with 4 TB HDD and 1 TB of internal M.2 SSD storage with up to 3.5 GB/s speeds, 64GB Quad-channel DDR4 desktop RAM and 8GB GDDR6 GeForce RTX 2070 GPU were used for the post-dynamic analysis.

2.2.2. Protein and ligand library preparation

The monoclinic (8ALS) [31] and orthorhombic (8BHZ) crystal structures of *Sb*28GST (which are crystal structures deposited into PDB from our research group) were retrieved from PDB. Whole structures were submitted into the OPLS_2005 protein preparation wizard module in Maestro. The structures were energy minimized, after which, using default parameters, hydrogen atoms and disulfide bonds were added. A System Builder module implemented in Maestro was used to build the systems for the MDS. Both *Sb*28GST systems were built using the OPLS_2005 force field, solvating the system utilizing the TIP3P explicit solvent model. For the boundary conditions, an orthorhombic box shape was used, encapsulating the protein at the centre of the box consisting of a 10 Å a = b = c distance with 90° $a = b = y$ angles. The box volume containing *Sb*28GST systems was minimized. Counter ions (Na^+ or Cl) were added per the molecular systems' overall charge, thus conditioning the system physiologically by adding 0.15 M NaCl. Similarly, a 1433 library of easily synthesizable flavonoid compounds was used to screen for potential hits. The structure data files of the compounds were obtained from PubChem database. The LigPrep module in Maestro was used to prepare the compounds. The compounds were protonated, ions were removed, and geometry optimized. The ligands were subjected to energy minimization using the OPLS 2005 force field to optimize the bond length, angle, and dihedrals.

2.2.3. Molecular dynamics simulations and post-dynamic analyses of Apo Sb28GST systems

The solvated and ionized molecular systems were subjected to MDS. The production phase of the simulation consists of eight stages, with simulation parameters specified for each phase. The simulation employed the NPT ensemble class with the 1 bar pressure and temperature of 300 K implemented in all runs. Stages 1–7 were for equilibration, where short simulations occurred. This was followed by stage 8, where the final long-range 500 ns simulation occurred at a constant temperature of 300 K. For a comparative analysis of the conformational dynamics of the simulated molecular systems, post-dynamic analyses of the simulation trajectories were carried out using the simulation interaction diagram module on Maestro. This calculates the Root-Mean-Square-Deviation (RMSD) and the Root-Mean-Square-Fluctuation (RMSF) of the protein backbone alpha carbon atoms (Cα). Similarly, the radius of gyration (Rg) was determined using the Simulation Events Analysis tool on Maestro.

2.2.4. High-throughput virtual screening

The receptor grids were first created to prepare the MD-simulated 8BHZ and 8ALS systems for high throughput virtual screening (HTVS) using the receptor-grid generation module implemented on Maestro. The grid box was a centroid of protein active site residues from both chains of the protein dimer with a 36 Å length size of the box allowed for the docking of ligands. The virtual screening workflow module implemented in Maestro was used to perform HTVS. The QikProp filter was exclusively applied while excluding compounds with reactive functional groups. For the preparation, Epik was used to generate possible states at a target pH of 7.0 ± 2.0 while also using Epik state penalties for docking. High energy ionization/tautomer states were removed. The stereochemical information of the ligands was obtained from the specified 2D stereo properties included in the default settings. The compounds were docked with Glide HTVS, and the successful hits that could bind to the 8BHZ and 8ALS binding sites were further screened using standard precision (SP) mode. Those screening outputs were further filtered using extra precision (XP) mode. The docking results were processed with Prime MM/GBSA to predict the free-binding energy for ranking purposes. The binding energies were used to calculate the compound's binding affinity for the 8BHZ and 8ALS systems using the below equation:

$$\Delta G = \quad RT\ln Ka \quad (1)$$

where G represents changes in the docking energy, R represents the Boltzmann gas constant ($R = 1.987$ cal/mol/K), T represents the temperature, and Ka represents the binding affinity.

Two compounds corresponding to the top-scoring compound with the lowest free binding energy and a common compound found in both 8BHZ and 8ALS SP mode screening outputs were then chosen for further MDS. The complexes of 8BHZ- apigenin 7-O-(2G-Rhamnosyl)gentiobioside herein named apigenin, 8BHZ- quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside herein named quercetin, 8ALS-apigenin and 8ALS-quercetin were prepared and subjected to a 500 ns MDS as described previously. The liganded protein simulations were further analysed for the ligand RMSD and the ligand properties, which include the molecular surface area (MolSA), solvent accessible surface area (SASA) and intramolecular hydrogen bonds (IntraHB).

2.3. Sb28GST enzymatic specific activity and inhibitory assay

The enzymatic specific activity of the two *Sb*28GST proteins and the inhibitory strength of the ligands (quercetin and apigenin) were tested using a reduced GSH-CDNB activity assay, where the catalytic activity of *Sb*28GST was spectrophotometrically monitored at varying enzyme concentrations as previously described [32]. The formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene (ε_{340} = 9600 $M^{-1}\cdot cm^{-1}$) using a Jasco V–630 UV/Vis spectrophotometer (Jasco, Japan) was monitored. The reaction buffer contained 100 mM sodium phosphate, 1 mM EDTA, 5 mM DTT, 0.02 % NaN_3, 5 mM GSH, 1 mM CDNB (3 % (v/v) ethanol), pH 6.5. Varying concentrations of *Sb*28GST (0–50 nM) were incubated for 30 min in the absence or presence of 50 μM apigenin and quercetin. Technical triplicates of the reactions were carried out for 60 s. The specific activity was obtained by plotting linear progress curves corrected for non-enzymatic reaction rates. The assay was also used to determine the half-maximal inhibitory (IC_{50}) of apigenin and quercetin. Varying concentrations of apigenin and quercetin were used with 50 nM *Sb*28GST. Experimental controls were conducted simultaneously in the absence of the protein. The data was then fitted using a four-parameter regression model (Eq. 2)

$$Y = bottom + \frac{top \quad bottom}{1 + 10^{((\log IC_{50} - x)\times hillslope)}} \quad (2)$$

where y is the enzyme activity, and the bottom and top are the minimum and maximum y - values for the curve asymptotes respectively. Log IC_{50} is the concentration at which 50 % of the enzyme activity is inhibited (nM), x is the logarithm of inhibitor concentration and Hillslope is the slope factor.

2.4. Fluorescence spectroscopy analysis of Sb28GST and with ligands

Extrinsic ANS fluorescence was employed to detect changes in protein conformation and probe hydrophobic clefts formed through *Sb*28GST interactions with apigenin and quercetin. A Jasco FP-6300 spectrofluorometer was used to measure the spectra using the Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan). The samples were prepared in 50 mM Tris-HCl pH 7.2, 1 mM EDTA and 1 mM DTT with 50 μM apigenin or quercetin. The samples contained 10 μM *Sb*28GST incubated with 50 μM of either co-substates GSH, CDNB, or PZQ and additionally incubated with 200 μM of freshly prepared ANS. The samples were excited at 395 nm, and the emission spectra were collected at 400–650 nm with an excitation bandwidth of 5 nm, emission of 2.5 nm, data pitch of 1 nm, 1 cm pathlength and a scanning speed of 200 nm/min. The samples were collected in triplicate and were subsequently averaged and buffer-corrected.

2.5. Thermal shift assay of Sb28GST and with ligands

To determine the thermal stability of *Sb*28GST in the presence of apigenin and quercetin, a thermal shift assay (TSA) was performed using a CFX96 Touch-Real-Time PCR detection system (Bio-Rad Laboratories, Inc., Hercules, CA, USA). This was monitored using a highly sensitive fluorescent dye called SYPRO Orange, whose quantum yield increases when hydrophobic patches are exposed during protein denaturation. Cocktails of 20 μM *Sb*28GST samples prepared in PBS containing 1 mM DTT and 1 mM EDTA supplemented with 0.5 mM GSH and 1 x SYPRO orange dye were loaded into a 96-well qPCR plate. Additional samples containing apigenin [IC_{50}] and quercetin [IC_{50}] were also prepared and loaded into the qPCR plate. The melting curves were tracked from 10 °C to 95 °C with 0.5 °C increments at 10 s intervals. The relative fluorescence unit (RFU) vs temperature (°C) was recorded, and the derivative of the plot (d(RFU)/dT) was used to determine the protein's melting temperature (*Tm*).

2.6. Isothermal titration calorimetry (ITC) of Sb28GST with ligands

The thermodynamics of interactions between *Sb*28GST with apigenin and quercetin were studied using the NanoITC instrument (TA Instruments, Delaware USA). Freshly purified *Sb*28GST was dialysed in PBS containing 1 mM TCEP HCl, 2 mM EDTA, 0.02 % (*w/v*) NaN_3, pH 6.5. The final dialysate was used to prepare 25 mM of apigenin and quercetin as working stock and 25 μM of *Sb*28GST to avoid a buffer mismatch. The protein, ligand and buffer were degassed for 20 min at 0.3–0.5 atm. The reference cell contained Milli-Q® water and the sample cell contained the protein sample. The ITC experiment was performed at 20 °C. The ITC syringe volume which held the ligands was 150 μL and the stir-rate used was 250 rpm. A total of 19 injections were used, with the initial injection volume being 2 μL, followed by subsequent 5 μL injections added every 300 s. NanoAnalyze™ Software was used to analyze the data, allowing the integrated heats per an injection (in microjoules) to be corrected for the heats of dilution. The data was then fitted to a multiple site binding model and visualized using the NanoAnalyze™ Software.

3. Results

3.1. SbGST's recombinant expression in E. coli and purification

Protein fractions containing soluble *Sb*28GST were harvested upon successful recombinant expression using a 0.5 mM IPTG concentration. The lysed cells were injected into the column after being centrifugally separated (15,000 x g, 10 mins), purified through a one-step elution method and using Ni^{2+} affinity chromatography to obtain a pure protein for downstream studies. While the result shows very little protein was lost to the flow through, as shown in Fig. 1, SDS-PAGE confirms the purity of the protein eluted and size to be 24 kDa.

3.2. Crystallization, data collection and structure solution of Sb28GST

Different crystallization buffers and conditions were tested before finally using 2.1 M ammonium sulfate, 100 mM Tris pH 7.5, and 5 mM *β*-mercaptoethanol to grow *Sb*28GST crystals. The protein crystals (with a needle-like morphology as shown in Fig. 2A), which grew within 48 h were diffracted to a resolution of 2.4 Å and had the structure refined with a final R_{factor} of 19.8 %. The solvent content and Matthew's coefficient (V_M) were 46.55 % and 2.33 $Å^3$ Da ^{1} respectively. Upon X-ray diffraction structural data collection and structure solution, it was found to belong to the C 2 2 2_1 space group, different from already deposited 8ALS (Fig. 2B) in PDB. The crystals were thinner and more clustered than the previously deposited 8ALS crystals. The structure was determined by molecular replacement method using the coordinate file of 8ALS. The data collection and refinement statistics comparing the two crystals were compiled in Supplementary Table 1. The 3D structure was visualized on PyMol and compared with the previously deposited, as shown in Fig. 3, with the superimposed structure showing a 1.664 Å RMSD to one another.

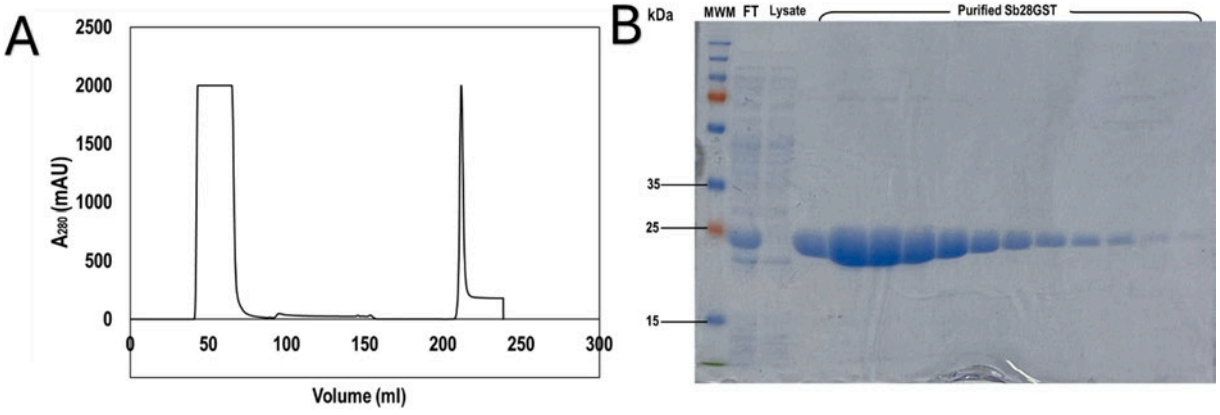


Fig. 1. The purification, purity and homogeneity of *Sb*28GST. (A) Shows the elution profile of *Sb*28GST purified using a Ni^{2+} affinity column eluted with 500 mM imidazole and (B) showing a 12.5 % (*w/v*) SDS-PAGE gel with the molecular weight marker (MWM), sonicated cell lysate, the flow through (FT) and *Sb*28GST eluents collected from the 500 mM imidazole elution.

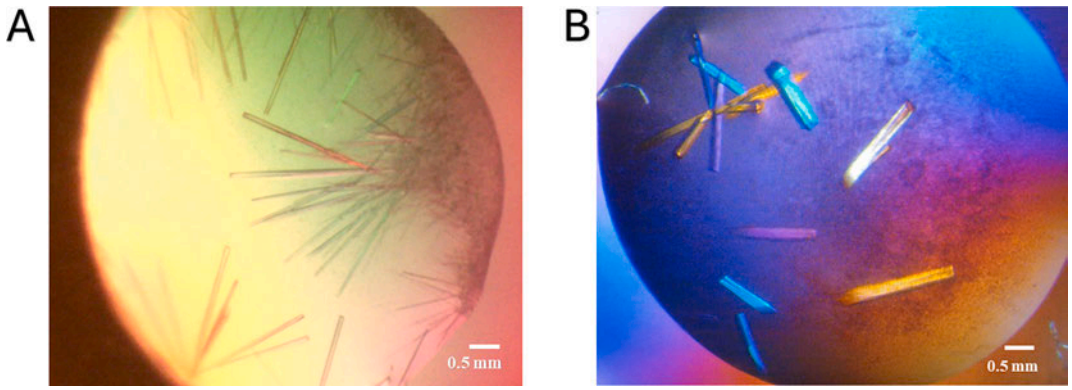


Fig. 2. Crystallization of *Sb*28GST by hanging drop method at 20 °C with (A) showing needle-like crystals grown as orthorhombic crystal system and deposited in the PDB as 8BHZ, and (B) thick needle-like crystals which belong to the monoclinic crystal system and deposited as 8ALS in the PDB.

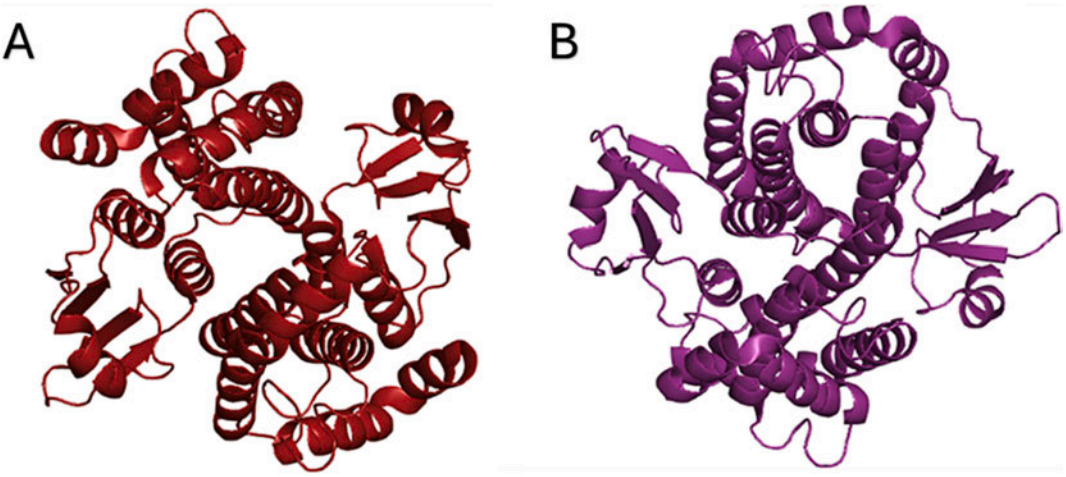


Fig. 3. The ribbon representation generated using the PyMol molecular graphics program and showing (A) the orthorhombic crystal and dimeric *Sb*28GST deposited in the PDB as 8BHZ and (B) the monoclinic crystal and dimeric *Sb*28GST deposited in the PDB as 8ALS.

3.3. Computational modelling results

3.3.1. Molecular dynamics simulations and post-dynamic analyses of apo orthorhombic and monoclinic Sb28GST systems before HTVS

A 500 ns MDS was done to compare the trajectory of the two crystal polymorphs of *Sb*28GST. A RMSD evolution which depicts the dynamic behaviour and stability of the protein backbone throughout a simulation period revealed that the structural conformation throughout the simulation of the polymorphs was not identical. However, followed a similar trend (Fig. 4A). A gradual increase in the first 100 ns of the simulation was followed by fluctuations which equilibrated at 400 ns at an average value of 3.5 Å for 8BHZ and 3.1 Å for 8ALS was observed. The RMSF shown in Fig. 4B indicated changes along the protein chain over the 500 ns simulation. The RMSF profiles showed similar fluctuation patterns throughout the simulation. The most notable distinctions were the marginally higher peaks observed from 8BHZ *N*-terminal amino acids, which are Lys115 and Ser204 on 8BHZ. The Rg trajectory observed from 8BHZ and 8ALS underwent minimal fluctuations throughout the 500 ns simulation, indicating a sustained degree of structural compactness (Fig. 4C). The lower Rg observed in the 8ALS model suggests a marginally more stable system than 8BHZ.

3.3.2. High-throughput virtual screening of flavonoid library for Sb28GST

Given that the two *Sb*28GST MDS trajectories are different and coupled with their assorted binding cavities, an HTVS (which is a further comparative study) was performed to show whether the hit ligands would be identical despite the different MDS trajectories. This was simultaneously done to determine the best drug candidate for schistosomiasis treatment. The screened top-hit ligands show compounds with the highest affinity for *Sb*28GST. This is because compounds with the lowest free binding energy (ΔG) indicate a high affinity for the protein and high protein-ligand complex stability [33]. The range exhibited from the HTVS was 15.6645 to 3.591 kcal/mol, indicating that the binding occurred exergonically. Different top ligands from the HTVS were selected for the *Sb*28GST polymorph, with Supplementary Table 3 showing the ligand 6″-Feruloylspinosin being selected for 8BHZ while Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D-Gentiobioside was selected for 8ALS. Interestingly, despite the two systems belonging to the same *Sb*28GST protein, only one ligand was common within the list of top 5 hits, which was apigenin 7-O-(2G-Rhamnosyl)gentiobioside. Additionally, the ligands selected surpassed bromosulfophthalein (BSP), which we considered our benchmark ligand, given its potency as elucidated in our previous GST inhibition studies [34,35]. Although BSP passed the initial virtual screening, receiving a Glide gscore of 5.93093 kcal/mol, it was, however, filtered out when subjected to further screening using SP and XP modes. Supplementary Table 2 details the assorted binding sites/cavities available within the two systems.

3.3.3. Molecular dynamics simulations and post-dynamic analyses of apo orthorhombic and monoclinic Sb28GST systems after HTVS

Another MDS was performed to test the dynamic behaviour and stability of the interactions between the ligands and the *Sb*28GST

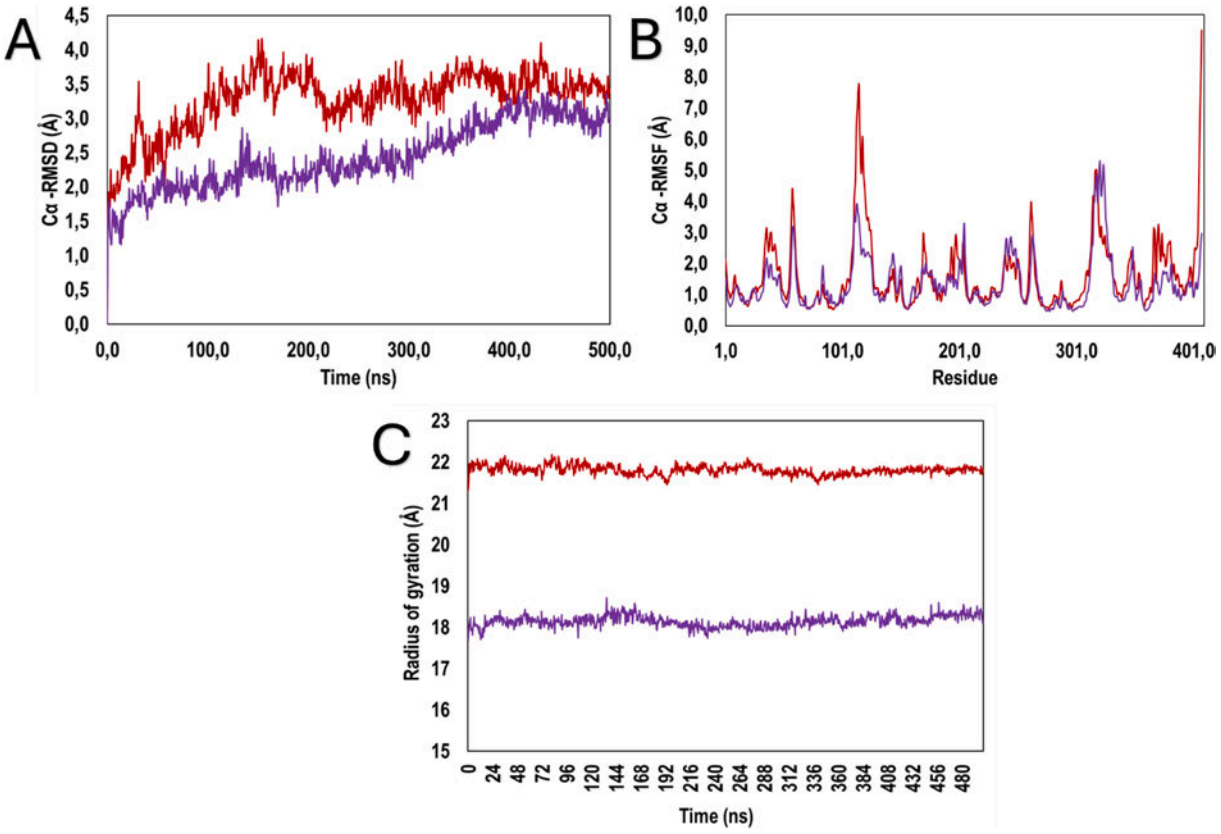


Fig. 4. Trajectory analysis of 500 ns molecular dynamic simulations of *Sb*28GST systems. The **red** and **purple** profiles represented 8BHZ and 8ALS respectively. (A) The RMSD of the Cα atoms. (B) The RMSF of the Cα atoms. (C) Radius of gyration of the Cα atom. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

systems. This allowed the protein-ligand complex to be observed in a cell-like environment enclosed in a water box with Na^+ and Cl ions as neutralizers. Apigenin (the common ligand produced in the polymorph's HTVS) and quercetin (the hit with the highest ΔG) were selected for further investigation. The Cα RMSD showed the conformation of the protein backbone throughout the 500 ns simulation with reference to the first frame. The *Sb*28GST:apigenin complex was less comparable between the two *Sb*28GST systems but had lower RMSD values in comparison to the reference frame with RMSD values of 3.08 and 2.54 Å for 8HBZ and 8ALS respectively (Fig. 5A). The Cα RMSD of *Sb*28GST: quercetin was very stable in both systems, showing very little fluctuation with an RMSD of 4.24 and 3.80 Å for 8BHZ and 8ALS respectively (Fig. 5B). The RMSF of the *Sb*28GST:apigenin and *Sb*28GST:quercetin complexes however showed more homogeneity (Fig. 5 C and D respectively) with most of the peaks observed corresponding with the amino acids that interact with the ligands, an indication of the stability of the active site. However, additional peaks were observed for the liganded systems compared with the apo systems. In the apigenin complex system, the *N*-terminal amino acids 150, 228 and 352 as well as the amino acids 148, 178 and 260 in complex with quercetin, exhibited higher conformational changes. The Rg was examined to determine the compactness of the protein during the simulation. The Rg profiles of 8BHZ and 8ALS systems shown in Fig. 5E were relatively comparable, indicating that polymorphic *Sb*28GST systems shared similar compactness. However, the overall Rg for 8BHZ:apigenin is 21.75 (± 0.12) and 21.81 (±0.12) for 8ALS:apigenin. On the other hand, 8BHZ:quercetin and 8ALS-quercetin's Rg were 21.66 (± 0.11) and 21.97 (± 0.14) respectively, an indication of moderate conformational changes during the 500 ns MDS (Fig. 5F). Similarly, the ligand RMSD trajectory showed the stability of the ligand in its protein binding site with Fig. 6A showing how apigenin exhibited significant fluctuations before reaching equilibrium when bound to 8BHZ and less dynamic in the 8ALS binding pocket. This shows that while the ligand diffused away from the initial binding pocket (between 0 and 150 ns), it however maintained a stable degree of motion at 150–430 ns. On the other hand, similar activity was observed with quercetin complexed 8ALS, as depicted in Fig. 6B. Quercetin drastically diffused away from the initial binding pocket of 8ALS, followed by a stable degree of motion maintained from 200 to 500 ns while behaving completely stable in 8BHZ.

Additional molecular properties of apigenin and quercetin such as MolSA, SASA and IntraHB were studied to determine the stability of the ligands in their *Sb*28GST binding pockets. The results are presented under Supplementary Fig. 1.

Similarly, Fig. 7 shows the interacting amino acid and the trajectory cluster analysis, revealing that apigenin and quercetin were stabilized by hydrogen bonds, water bridges, and hydrophobic and ionic interactions. The report presents a 2D interaction analysis based on the root-mean-square-deviation (RMSD) of the most dominant snapshots over a 500 ns molecular dynamics simulations (MDS). In 8BHZ:apigenin (Fig. 7A), water bridges played a more significant role than in 8 ALS:apigenin (Fig. 7C). On the other hand, water bridges were more pronounced in 8BHZ:quercetin (Fig. 7D) than 8 ALS:quercetin (Fig. 7B).

Compared to apigenin, significantly more interactions occurred between the *Sb*28GST systems and quercetin with the analysis of their atoms' 2D contact summary interacting with *Sb*28GST amino acid presented and reported in detail under Supplementary Fig. 2.

3.4. Specific activity and inhibition studies of Sb28GST

Fig. 8A shows the specific activity of *Sb*28GST with and without ligands apigenin and quercetin. The specific activity of *Sb*28GST was 19.95 ± 0.5 μmol.min ^{1}.mg ^{1}. In the presence of inhibitors, however

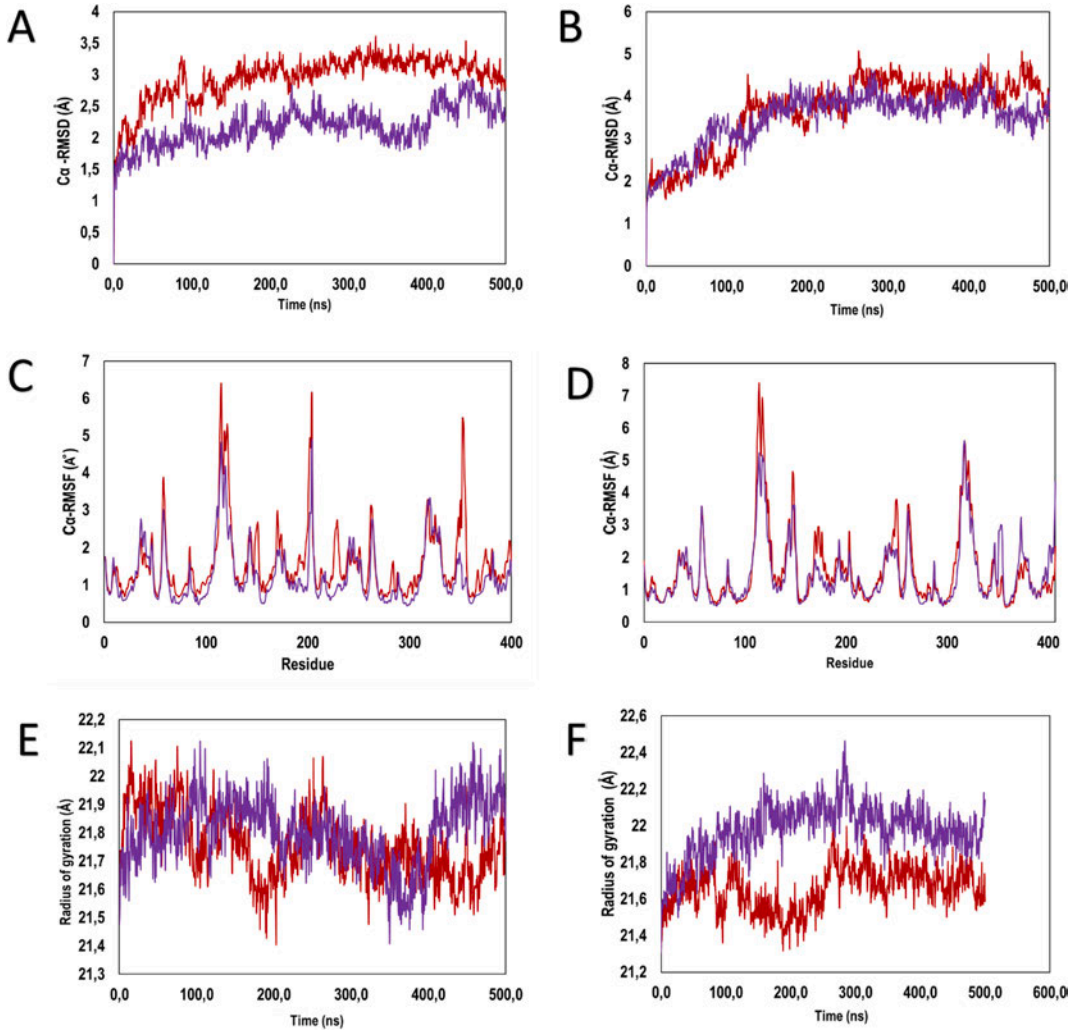


Fig. 5. Trajectory analysis of the 500 ns MDS of *Sb*28GST systems complexed apigenin and quercetin. The **red** and **purple** profiles represented 8BHZ and 8ALS respectively with apigenin interactions depicted on the left panels and quercetin interactions depicted on the right panels. (A) The RMSD of the Cα atoms of *Sb*28GST systems with apigenin. (B) The RMSD of the Cα atoms of *Sb*28GST systems with quercetin. (C) The RMSF of the Cα atoms of *Sb*28GST systems with apigenin. (D) The RMSF of the Cα atoms of *Sb*28GST systems with quercetin. (E) The Rg of the Cα atoms of *Sb*28GST with apigenin. (F) The Rg of the Cα atoms of *Sb*28GST with quercetin. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

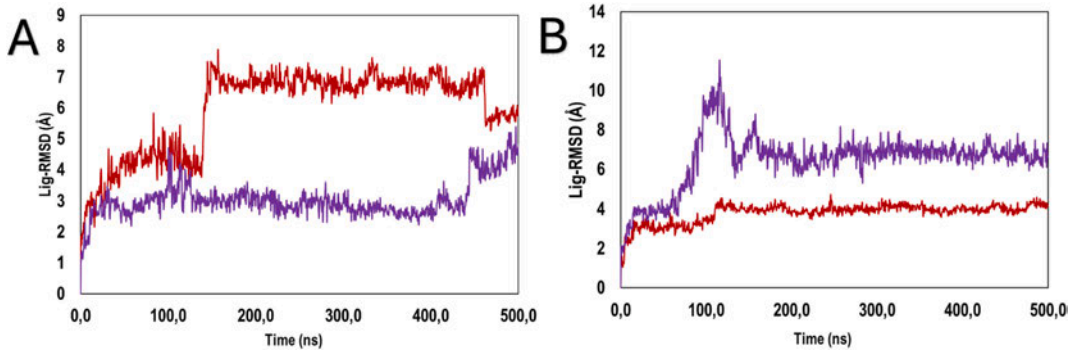


Fig. 6. Root mean square deviation of the apigenin and quercetin with respect to *Sb*28GST residues. The **red** and **purple** profiles represented 8BHZ and 8ALS residues respectively. (A) The ligand RMSD of apigenin with *Sb*28GST receptors. (B) The ligand RMSD of quercetin with *Sb*28GST receptors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the specific activity was drastically reduced. In the presence of apigenin the specific activity of *Sb*28GST was 2.59 ± 0.5 μmol.min ^{1}.mg ^{1}, while in the presence of quercetin, the specific activity was reduced to 0.01 μmol.min ^{1}.mg ^{1}. In Fig. 8B, the IC_{50} of apigenin and quercetin were determined to be 0.13 and 0,12 mM respectively.

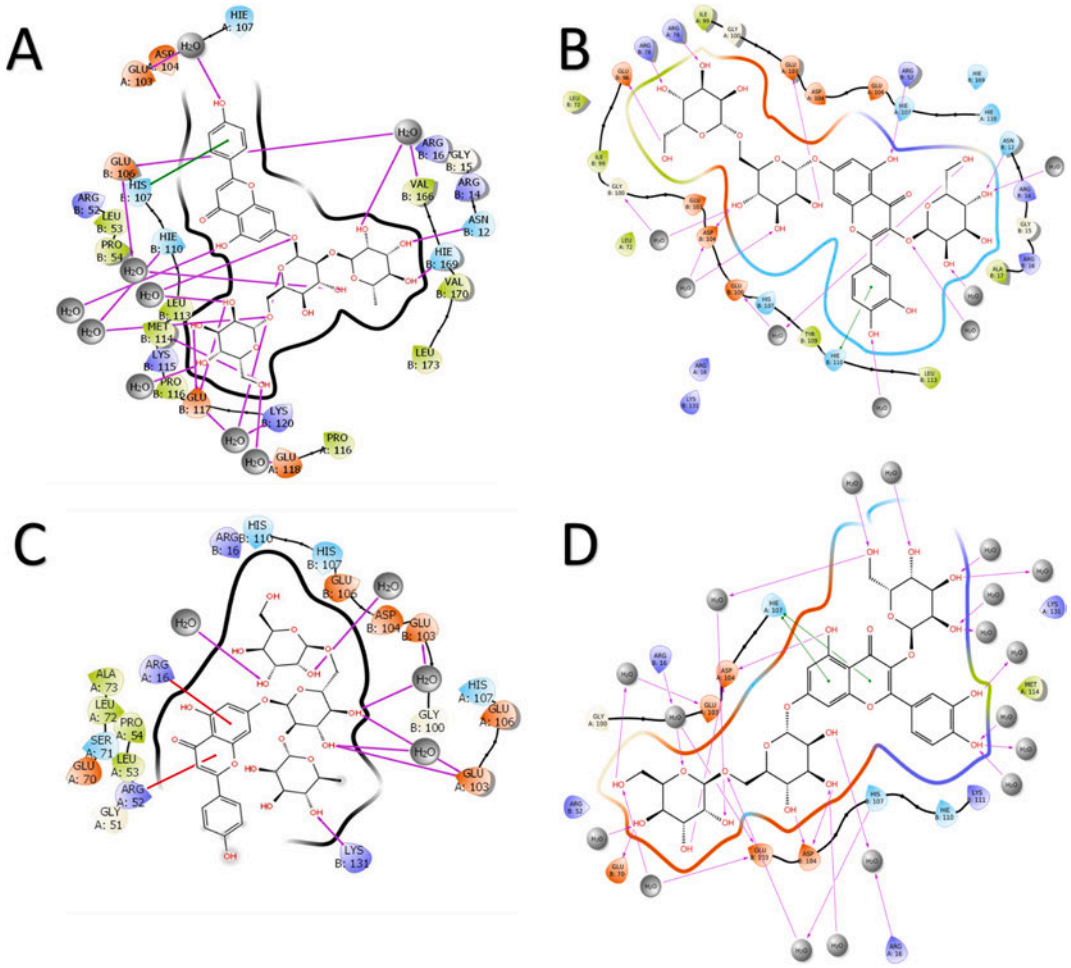


Fig. 7. 2D interaction plot of the trajectory frame based on the RMSD of most dominant snapshots throughout a 500 ns MDS showing complexes of (A) 8BHZ: Apigenin, (B) 8BHZ:Quercetin, (C) 8ALS:Apigenin, and (D) 8ALS:Quercetin. Amino acid residues within 4 Å of the ligand are colour-coded: positively charged (**violet**), negatively charged (**orange**), non-polar (**green**), and polar (**blue**). Hydrogen bonds (**purple**) and salt bridges (**red**) are shown, with solvent exposure shaded **grey**. Images were created using Maestro v13.0. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

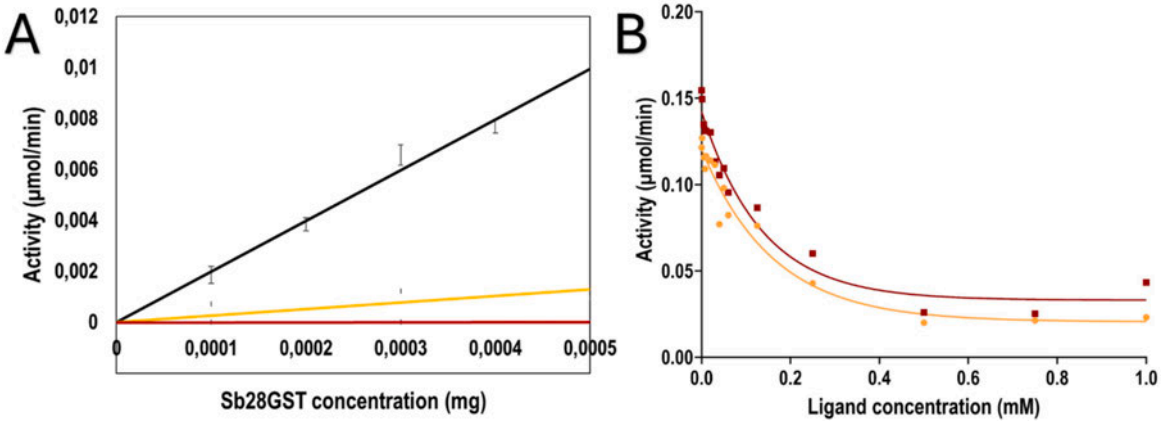


Fig. 8. Specific activity analysis and inhibition studies of *Sb*28GST with apigenin and quercetin**.** The GSH-CDNB conjugation assay was performed at 340 nm in the presence of apigenin (**orange**) and quercetin (**red**) by monitoring the formation of the chromophoric product 1-(S-glutathionyl)-2,4-dinitrobenzene catalysed by *Sb*28GST. (A) The specific activity of *Sb*28GST through varying concentrations of *Sb*28GST from 0 to 50 nM in the presence and absence of 50 µM apigenin and quercetin. (B) Inhibition studies of *Sb*28GST to determine the IC_{50} of the ligands apigenin and quercetin. The ligand concentrations varied from 0 to 1 mM with 50 nM *Sb*28GST. All assays were performed in triplicates, and the data were corrected for non-enzymatic rates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Extrinsic fluorescence analyses of Sb28GST and its complexes

ANS fluorescence spectroscopy was performed to determine possible binding interactions between *Sb*28GST and the ligands. When ANS binds to hydrophobic patches on proteins, a hypsochromic shift is typically observed in addition to increased fluorescent yield compared to free ANS. Fig. 9A shows the baseline fluorescence intensities of ANS incubated with *Sb*28GST and apigenin or quercetin, from which other spectra will be compared. The fluorescence of ANS:*Sb*28GST increased significantly compared to free ANS spectra. This was coupled with a blue shift in ANS emission from 520 nm to 492 nm in the presence of the protein. Upon the addition of GSH in Fig. 9B, the ANS-*Sb*28GST amplification was maintained with no notable fluorescence emission changes. The addition of CDNB and PZQ as co-substrates to the base samples led to a significant decrease in fluorescence intensity shown in Fig. 9C and D. Two anomalies were observed due to the presence of either quercetin or apigenin. In Fig. 9D, the addition of quercetin showed a substantial hypsochromic shift to the ANS fluorescent intensity of *Sb*28GST with PZQ, while the opposite was observed to the ANS fluorescent intensity of *Sb*28GST with CDNB, shown in Fig. 9C.

3.6. Thermal stability of Sb28GST and its complexes

The fluorescent dye SYPRO orange has an affinity for hydrophobic patches, which typically result in increased quantum yield. This is seen upon protein denaturation where buried hydrophobic regions become exposed allowing the dye to occupy more hydrophobic regions. Upon saturation the dye becomes quenched indicated by decreasing fluorescence. The melting temperature of the *Sb*28GST was derived from the proteins' melting curve in the presence and absence of apigenin and quercetin. Fig. 10A and B show the thermal denaturation profile of *Sb*28GST in the presence and absence of apigenin and quercetin respectively. The protein was stable at room temperature and began unfolding with increasing temperature, reaching a single transition peak at 70 °C. In the presence of apigenin, the T_m of the *Sb*28GST was reduced by 0.5 °C. Similarly, the T_m of *Sb*28GST in the presence of quercetin was reduced to 62 °C from 63.5 °C.

3.7. Isothermal titration calorimetry analyses of Sb28GST with apigenin and quercetin

Isothermal titration calorimetry was used to determine the thermodynamic parameters of apigenin and quercetin binding to *Sb*28GST. The thermodynamic parameters of apigenin and quercetin binding to *Sb*28GST are summarized in Table 1. The thermograms and corresponding fitted data depicted in Fig. 11A show that the binding of apigenin to *Sb*28GST occurs endothermically as seen by the positive change in enthalpy (ΔH). Additionally, with Gibbs free energy change (ΔG) of 46.75 kJ/mol and a TΔS° of 65.04 kJ/mol. This interaction appears to be an entropically driven spontaneous reaction. Quercetin binding to *Sb*28GST is carried out in an exothermic reaction as evidenced by the negative ΔH° and downward-facing peaks in Fig. 11B. The negative ΔG° and positive TΔS° further show that this exothermic reaction occurs spontaneously and is entropically driven.

4. Discussion

This study aims to evaluate the impact of polymorphic crystal space groups on ligand selectivity and identify novel flavonoid inhibitors that specifically target *Sb*28GST. Since *Sb*28GST has been recognized as a promising anti-schistosome drug target, carefully selecting crystal structure space groups for computer-aided drug design (CADD) is crucial in understanding how protein conformational dynamics influence inhibitor binding and efficacy. To the best of our knowledge, no

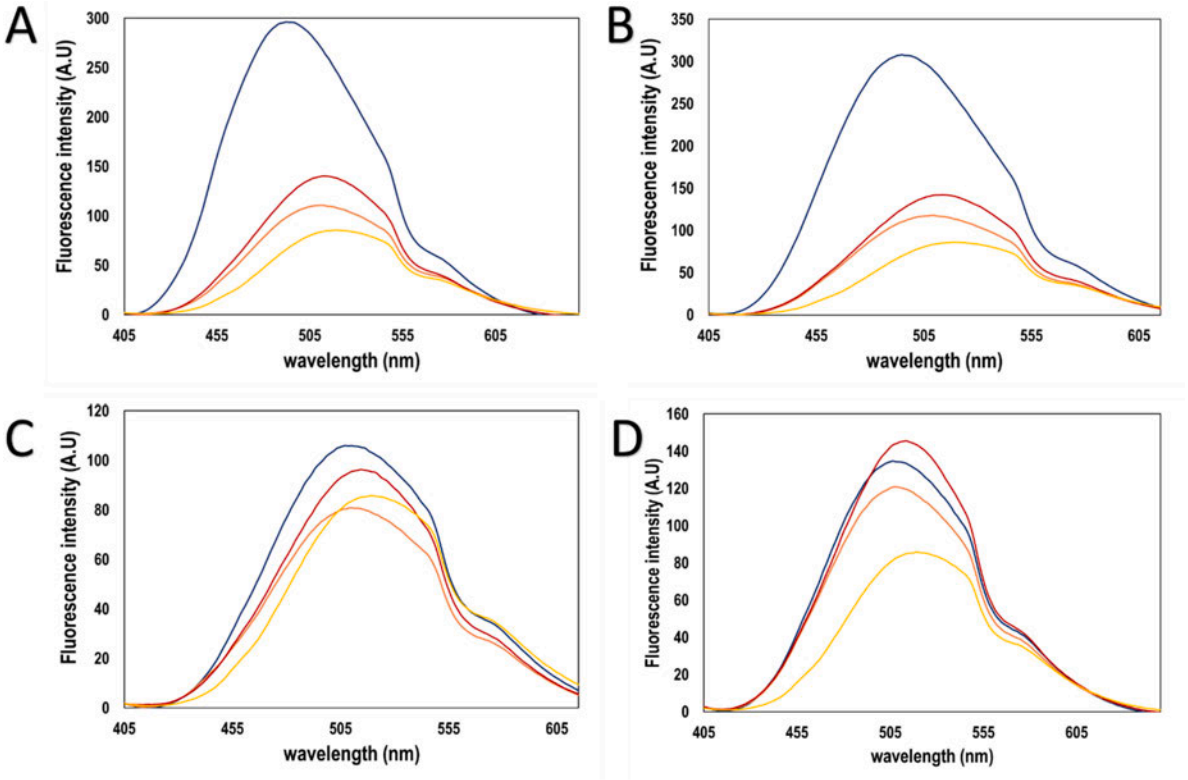


Fig. 9. Extrinsic ANS Fluorescence of *Sb*28GST. The emission spectra of 10 μM Sb28GST were collected at 400–640 nm upon excitation at 395 nm. The concentration of ANS used was 50 μM. Combinations of the *Sb*28GST and ANS (**blue**), apigenin, *Sb*28GST and ANS (**orange**), quercetin, *Sb*28GST and ANS (**red**), and free ANS (**yellow**) were recorded alongside varying co-substrates. (A) No co-substrate. (B) Glutathione. (C) 1-chloro-2,4-dinitrobenzene. (D) Praziquantel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

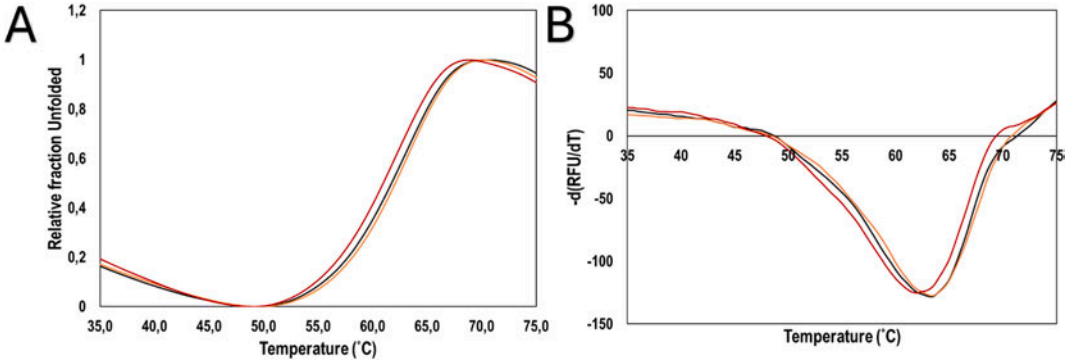


Fig. 10. Thermal stability profile of *Sb*28GST in the presence and absence of apigenin and quercetin. The readings of *Sb*28GST (**black**), *Sb*28GST-Apigenin (**orange**) and *Sb*28GST-Quercetin (**red**) were recorded in triplicate averaged to generate individual thermal melt curves corrected with buffer blanks. (A) The standardized RFU against temperature (°C), (B) The derived -d(RFU)/dT against temperature (°C). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Thermodynamic parameters of *Sb*28GST interacting with either apigenin or quercetin.

	K_d (μM)	ΔH° (kJ/mol)	TΔS° (kJ/mol)	ΔG° (kJ/mol)
*Sb*28GST: apigenin	4.69 × 10^{3}	18.29	65.04	46.75
*Sb*28GST: quercetin	21.69	17.69	8.48	26.17

polymorphic crystal space groups and ligand selectivity study has been considered in identifying novel flavonoid inhibitors that specifically target *Sb*28GST using an integration of in silico and empirical study approaches.

The *Sb*28GST protein was successfully and recombinantly expressed as a soluble protein, purified to homogeneity using immobilized Ni^{2+} affinity chromatography, and confirmed the molecular mass using SDS-PAGE to be 24 kDa, corresponding to the theoretical size. The purified protein was subsequently and successfully crystallized to 2.4 Å resolution. The crystal was found to belong to the C 2 2 2_1 space group and orthorhombic unit cell (8BHZ), different from the pre-existing monoclinic unit cell (8ALS) and P 1 2_1 1 space group, previously deposited into PDB from our research group. This polymorphic development (allowing us to explore downstream hypotheses) is attributable to changes in crystallization conditions, such as pH or the ionic strength of the crystal environment, which can impact crystal packing, resulting in changes in the crystal unit cell parameters [20].

Given that simulation conveys details of individual particle motions as a function of time, offering a way to quantify the degree of conformational changes in protein backbones and side chains, the behaviour of two *Sb*28GST space group systems in solution through MDS was evaluated. The resulting divergent trajectories (RMSD, RMSF, and Rg) of the protein backbone showed that these polymorphic space groups possess unique conformational states, manifesting through dissimilar dynamic properties but with relatively similar compactness degrees. However, the protein side chains were comparably and relatively stable in solution, with slight fluctuations observed with 8BHZ. These results highlight the divergence of polymorphic systems' dynamic properties from the anticipated identical molecular disposition. Similarly, considering the flexibility of the polymorphic proteins and their potential cavity volume to quantitatively estimate the druggability of their respective sites, cavity detection of the *Sb*28GST systems was carried out to reveal essential information for virtual screening, drug discovery and receptor binding analyses. The cavities observed were not identical but were similar in volume. This could be due to variations in crystal packing forces inducing different side chain rearrangements [21]. The ranging

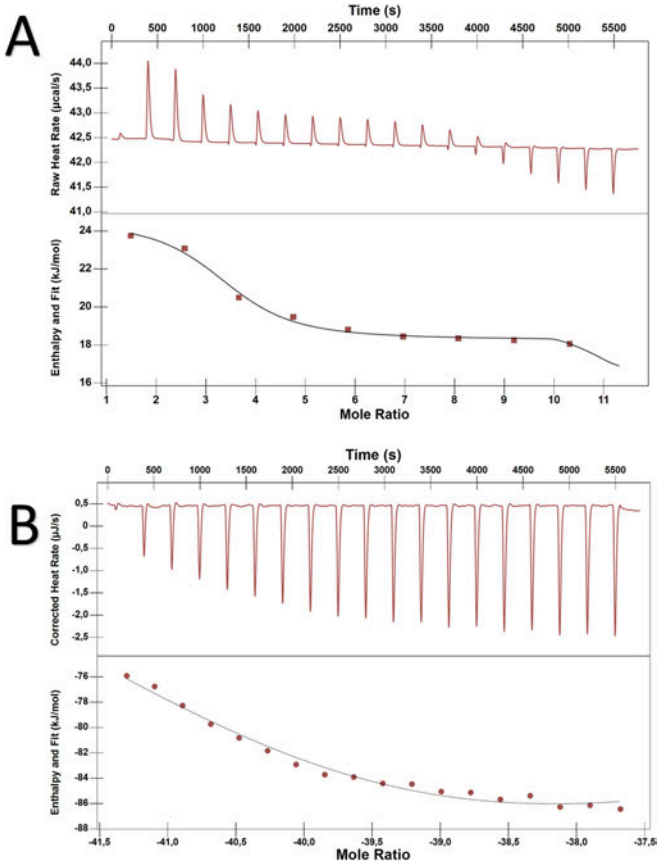


Fig. 11. Thermograms of *Sb*28GST interacting with either apigenin or quercetin. Using ITC, the interactions between (A) 5 mM apigenin with *Sb*28GST and (B) 7 mM quercetin with *Sb*28GST were monitored by titrating the ligands into 25 μM *Sb*28GST. NanoAnalyze software uses the multiple sites model to fit the data to visualize the raw titration data depicted in the upper panel and fitted data in the lower panel.

orientations adopted by *Sb*28GST chains in orthorhombic and monoclinic space groups thus induce regions of the protein surfaces to be involved in assorted interchain interactions, effectively resulting in slightly varied cavity arrangements. Three confirmed binding sites are deemed favourable across the two models. This aligns with current literature that corroborates three known GST binding sites [10].

A library of plant-derived and synthesizable polyphenols (flavonoids) was screened (via HTVS) to identify potential anti-schistosome

drugs due to their ease of procurement and adeptness in modulating drug-metabolizing enzymes [36,37]. The HTVS analysis of the top five binding flavonoids for the orthorhombic and monoclinic *Sb*28GST systems revealed only one common ligand (apigenin) among the top five hits. This is an indication that the polymorphic systems have different affinities for the ligands with different binding energies across the two systems. Quercetin-3-O-Beta-D-Glucose-7-O-Beta-D Gentiobioside, a flavanol, showed the lowest binding energy. Its polyphenolic structure, hydroxyl group arrangement, and glycosides enhance water solubility, bioavailability, and stable binding within the protein active site, offering therapeutic potential [36,37].

The MDS and post-dynamic analyses of the systems gave more insights. Apigenin and quercetin exhibited favourable binding energies to *Sb*28GST, and dynamic analysis showed stable side chains with similar amino acid interactions. However, varying degrees of fluctuations between the systems indicated differences in residue flexibility while maintaining overall stability in both complexes. The Rg values of ligand-bound systems were similar, but apo 8ALS showed a smaller Rg, indicating that ligand introduction into the less packed monoclinic system caused protein expansion and loss of compactness [38]. This expansion was not observed in the orthorhombic 8BHZ system, likely due to its structural constraints, which limit flexibility and expansion. Both apigenin and quercetin behaved similarly in simulations, but quercetin was more stable in the *Sb*28GST active site. Quercetin diffused quickly and remained throughout the simulation, while apigenin moved slowly towards the active site and then drifted away. Apigenin's inconsistent PSA suggested unstable solvent interactions and unbalanced polar contacts with the protein. Both ligands formed significant networks of hydrogen bonds, crucial for stability, with apigenin forming more bonds overall [39]. However, the monoclinic system facilitated more stable hydrogen bond networks than the orthorhombic system, likely due to its structural flexibility. This indicates that the monoclinic environment promotes greater ligand stability and stronger interactions, making it more favourable for ligand binding. Similarly, the *Sb*28GST systems exhibited strong interactions with both apigenin and quercetin, but the amino acids involved varied across space groups in identity and interaction intensity. In 8BHZ, fewer voids and channels in the orthorhombic structure resulted in restricted conformational flexibility, leading to more stable but fewer interactions, heavily reliant on water bridges. Limited solvent access increased dependence on water bridges, making them crucial mediators [20,21]. In contrast, the less compact monoclinic 8ALS structure allowed greater conformational stability, larger solvent channels, and more complex interactions, including direct hydrogen bonding and multiple ligand conformations, enhancing ligand binding.

Empirically, the catalytic performance of *Sb*28GST and the inhibitory ability of apigenin and quercetin as lead inhibitors of *Sb*28GST were tested. The specific activity of *Sb*28GST (19.94 μmol.min ^{1}.mg ^{1}) dropped to 12.95 % with apigenin and 0.03 % with quercetin. This result and that of the IC_{50} highlight quercetin's superior inhibitory potency. This study is the first to empirically evaluate *Sb*28GST inhibition, with no existing research for comparison with other potential inhibitors. Spectroscopically, the potential binding sites of apigenin and quercetin on polymorphic *Sb*28GST were carried out by analyzing ANS fluorescence spectra and intensities in hydrophobic regions. The fluorescence intensity of ANS increased when bound to accessible hydrophobic regions of the proteins, indicating that the dye occupied the accessible hydrophobic regions of the protein. However, upon the addition of apigenin and quercetin, a decrease in ANS fluorescence was observed, indicating the displacement of ANS from certain hydrophobic pockets. No fluorescence changes were seen with GSH, suggesting that ANS could not compete for the G-site. The decrease in ANS fluorescence upon adding CDNB, known to bind at GST's H-site [40], showed its ability to displace ANS from that specific H-site hydrophobic pocket. On the other hand, there is an additional decrease in ANS fluorescence of GST in the presence of CDNB and quercetin, indicating that the inhibitor had a higher affinity for the H-site. Thus effectively displacing both ANS and CDNB. Similarly, apigenin displaced ANS and PZQ from the protein's L-site, their known binding site [34], showing preference for that binding site. Two anomalies were observed: a significant decrease in ANS fluorescence with both CDNB and apigenin and a fluorescence increase when quercetin and PZQ were present. These results suggest that apigenin and quercetin might bind to both the H-site and L-site, corroborating the interactions with the amino acids reported in the inhibitors' dynamic studies.

The TSA revealed that apigenin and quercetin slightly reduced the thermal stability of *Sb*28GST, although the reduction was minimal. This is unexpected, as previous studies have shown that dietary flavonoids are sensitive to heat, and at the same time, thermal exposure can also enhance their bioactivity and bioavailability [41,42]. While the ligands' presence destabilization was not significant enough to denature the protein, it suggests potential for further exploration as drug candidates. Isothermal titration calorimetry (ITC) analyses revealed that both apigenin and quercetin bind entropically and spontaneously to *Sb*28GST, with apigenin binding endothermically and quercetin exothermically, reflecting distinct molecular interactions. Apigenin's binding suggests non-covalent interactions, like hydrophobic effects, compensating for unfavourable enthalpy changes. In contrast, quercetin's binding indicates favourable interactions, such as van der Waals forces and hydrogen bonds [43,44]. These results corroborate the MDS findings. Interestingly, despite quercetin's predicted higher affinity, apigenin had the lowest dissociation constant, showing stronger binding. This difference underscores a limitation within the study, highlighting how computational analyses might miss critical information about ligand behaviour across different systems. True interaction events occur over micro-millisecond time frames, the simulation of which is not readily available without costly computing power or cloud-based solutions [45]. While machine learning advances to bridge this gap, enhanced sampling methods could be explored in the interim, such as metadynamics or accelerated molecular dynamics [46,47].

5. Conclusion

While computational methods can effectively identify lead compounds for schistosomiasis management, understanding the relationship between a protein's crystal structure and its dynamic behaviour in solution is crucial. Our study revealed that *Sb*28GST protein structures obtained from monoclinic and orthorhombic space groups led to distinct molecular conformations due to differences in packing environments. This affected ligand selectivity, binding affinities and interactions. Molecular dynamics simulations of apigenin and quercetin with the polymorphic *Sb*28GST models revealed varied stabilizing interactions, underscoring the importance of analyzing multiple space group crystals to comprehensively understand protein conformational states. Empirical validation confirmed that both ligands identified significantly reduced *Sb*28GST activity. Additionally, fluorescence studies showed that apigenin and quercetin could bind to both the H-site and L-site at the dimer interface with minimal impact on the protein's thermal stability. These findings highlight the importance of combining computational and empirical approaches in drug discovery.

CRediT authorship contribution statement

Mbalenhle Mfeka: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Olalekan Onisuru:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Ramesh Pandian:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Yasien Sayed:** Writing – review & editing, Project administration, Funding acquisition. **Thandeka Khoza:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Ikechukwu Achilonu:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data

curation, Conceptualization.

Funding

This work received funding support from the National Research Foundation (NRF)- South Africa. The funding was awarded to I.A. as a CPRR research grant (Grant number CPRR23042899244). The author also acknowledges the Centre For High-Performance Computing (CHPC), Cape Town for access to the Schrodinger Molecular Modelling Suite.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rechem.2025.102288.

Data availability

Data will be made available on request.

References

[1] Y. Yang, M. Zhou, M. Hu, Y. Cui, Q. Zhong, L. Liang, F. Huang, UGT1A1* 6 and UGT1A1* 28 polymorphisms are correlated with irinotecan-induced toxicity: a meta-analysis, Asia-Pacific, J. Clin. Oncol. 14 (2018) e479–e489.
[2] M. Merrifield, P.J. Hotez, C.M. Beaumier, P. Gillespie, U. Strych, T. Hayward, M. E. Bottazzi, Advancing a vaccine to prevent human schistosomiasis, Vaccine 34 (2016) 2988–2991, https://doi.org/10.1016/j.vaccine.2016.03.079.
[3] E. Léger, A. Borlase, C.B. Fall, N.D. Diouf, S.D. Diop, L. Yasenev, S. Catalano, C. T. Thiam, A. Ndiaye, A. Emery, A. Morrell, M. Rabone, M. Ndao, B. Faye, D. Rollinson, J.W. Rudge, M. Sène, J.P. Webster, Prevalence and distribution of schistosomiasis in human, livestock, and snail populations in northern Senegal: a one health epidemiological study of a multi-host system, Lancet Planet. Heal. 4 (2020) e330–e342, https://doi.org/10.1016/S2542-5196(20)30129-7.
[4] H. You, P. Cai, B.M. Tebeje, Y. Li, D.P. McManus, Schistosome vaccines for domestic animals, Trop. Med. Infect. Dis. 3 (2018) 68, https://doi.org/10.3390/tropicalmed3020068.
[5] A.M. Martínez-Ibeas, C. González-Lanza, M.Y. Manga-González, Proteomic analysis of the tegument and excretory-secretory products of *Dicrocoelium dendriticum* (Digenea) adult worms, Exp. Parasitol. 133 (2013) 411–420, https://doi.org/10.1016/j.exppara.2013.01.010.
[6] H. Huang, The redox biology of Schistosome parasites and applications for drug development, Curr. Pharm. Des. 18 (2012) 3595–3611, https://doi.org/10.2174/138161212801327220.
[7] S. Müller, Role and regulation of glutathione metabolism in plasmodium falciparum, Molecules 20 (2015) 10511–10534, https://doi.org/10.3390/molecules200610511.
[8] N. Padi, B.O. Akumadu, O. Faerch, C. Aloke, V. Meyer, I. Achilonu, Engineering a pseudo-26-kda schistosoma glutathione transferase from bovis/haematobium for structure, kinetics, and ligandin studies, Biomolecules 11 (2021) 1844, https://doi.org/10.3390/biom11121844.
[9] A. Viana da Costa, S. Gaubert, S. Lafitte, J. Fontaine, A. Capron, J.M. Grzych, Egg-hatching inhibition in mice immunized with recombinant Schistosoma bovis 28 kDa glutathione S-transferase, Parasite Immunol. 21 (1999) 341–350, https://doi.org/10.1046/j.1365-3024.1999.00232.x.
[10] Z. Yassin, E. Ortiz-Salmerón, F. García-Maroto, C. Barón, L. García-Fuentes, Implications of the ligandin binding site on the binding of non-substrate ligands to Schistosoma japonicum-glutathione transferase, Biochim. Biophys. Acta - proteins, Proteomics 1698 (2004) 227–237, https://doi.org/10.1016/j.bbapap.2003.12.004.
[11] R.C. Godwin, R. Melvin, F.R. Salsbury, Molecular dynamics simulations and computer-aided drug discovery, methods, Pharmacol. Toxicol. 9 (2015) 1–30, https://doi.org/10.1007/7653_2015_41.
[12] I. Sánchez-Linares, H. Pérez-Sánchez, J.M. Cecilia, J.M. García, High-throughput parallel blind virtual screening using BINDSURF, BMC Bioinfo. 13 (2012) 1–14, https://doi.org/10.1186/1471-2105-13-S14-S13.
[13] M. Yang, Molecular dynamics simulations: chemical advances and applications, J. Phys. Conf. Ser. 2608 (2023) 37–47, https://doi.org/10.1088/1742-6596/2608/1/012044.
[14] S.A. Hollingsworth, R.O. Dror, Molecular dynamics simulation for all, Neuron 99 (2018) 1129–1143, https://doi.org/10.1016/j.neuron.2018.08.011.
[15] X. Zhang, S.E. Wong, F.C. Lightstone, Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized born surface area rescoring to improve enrichment, J. Chem. Inf. Model. 54 (2014) 324–337, https://doi.org/10.1021/ci4005145.
[16] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct. Biol. 9 (2002) 646–652, https://doi.org/10.1038/nsb0902-646.
[17] N.M. Tripathi, A. Bandyopadhyay, High throughput virtual screening (HTVS) of peptide library: technological advancement in ligand discovery, Eur. J. Med. Chem. 243 (2022) 114766, https://doi.org/10.1016/j.ejmech.2022.114766.
[18] F.I. Khan, F. Hassan, H. Ali, D. Lai, Mechanism of pH-induced conformational changes in MurE ligase obtained from *Salmonella enterica* serovar Typhi, J. Biomol. Struct. Dyn. 39 (2021) 1898–1905, https://doi.org/10.1080/07391102.2020.1739560.
[19] J.D. Durrant, J.A. McCammon, Molecular dynamics simulations and drug discovery, BMC Biol. 9 (2011) 1–9.
[20] M.P. Jacobson, R.A. Friesner, Z. Xiang, B. Honig, On the role of the crystal environment in determining protein side-chain conformations, J. Mol. Biol. 320 (2002) 597–608, https://doi.org/10.1016/S0022-2836(02)00470-9.
[21] M. Chruszcz, W. Potrzebowski, M.D. Zimmerman, M. Grabowski, H. Zheng, P. Lasota, W. Minor, Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter? Protein Sci. 17 (2008) 623–632, https://doi.org/10.1110/ps.073360508.
[22] K.A. Johnson, F. Angelucci, A. Bellelli, M. Hervé, J. Fontaine, D. Tsernoglou, A. Capron, F. Trottein, M. Brunori, Crystal structure of the 28 kDa glutathione S-transferase from Schistosoma haematobium, Biochemistry 42 (2003) 10084–10094, https://doi.org/10.1021/bi034449r.
[23] P. Evans, Scaling and assessment of data quality, Acta Crystallogr. Sect. D Biol. Crystallogr. 62 (2006) 72–82, https://doi.org/10.1107/S0907444905036693.
[24] P.R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics, Acta Crystallogr. Sect. D Biol. Crystallogr. 67 (2011) 282–292, https://doi.org/10.1107/S090744491003982X.
[25] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. 69 (2013) 1204–1214, https://doi.org/10.1107/S0907444913000061.
[26] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R. J. Read, Phaser crystallographic software, J. Appl. Cryst. 40 (2007) 658–674, https://doi.org/10.1107/S0021889807021206.
[27] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. Sect. D Biol. Crystallogr. 60 (2004) 2126–2132, https://doi.org/10.1107/S0907444904019158.
[28] P.V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, T.C. Terwilliger, A. Urzhumtsev, P.H. Zwart, P.D. Adams, Towards automated crystallographic structure refinement with phenix. Refine, Acta Crystallogr, Sect. D Biol. Crystallogr. 68 (2012) 352–367.
[29] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291, https://doi.org/10.1107/s0021889892009944.
[30] W.L. DeLano, Pymol: An open-source molecular graphics tool, {CCP4} Newsl, Protein Crystallogr. 40 (2002) 1–8. http://www.ccp4.ac.uk/newsletters/newsletter40/11_pymol.pdf.
[31] H.H. Makumbe, R. Pandian, A. Valli, Y. Sayed, I. Achilonu, Biophysical characterization, crystallization, and solution of the first crystal structure of the 28 kDa-Schistosoma bovis glutathione transferase, J. Mol. Struct. 1298 (2024) 136979, https://doi.org/10.1016/j.molstruc.2023.136979.
[32] B. Zhan, S. Perally, P.M. Brophy, J. Xue, G. Goud, S. Liu, V. Deumic, L.M. De Oliveira, J. Bethony, M.E. Bottazzi, D. Jiang, P. Gillespie, S.H. Xiao, R. Gupta, A. Loukas, N. Ranjit, S. Lustigman, Y. Oksov, P. Hotez, Molecular cloning, biochemical characterization, and partial protective immunity of the heme-binding glutathione S-transferases from the human hookworm *Necator americanus*, Infect. Immun. 78 (2010) 1552–1563, https://doi.org/10.1128/IAI.00848-09.
[33] X. Du, Y. Li, Y.L. Xia, S.M. Ai, J. Liang, P. Sang, X.L. Ji, S.Q. Liu, Insights into protein–ligand interactions: mechanisms, models, and methods, Int. J. Mol. Sci. 17 (2016) 144, https://doi.org/10.3390/ijms17020144.
[34] N. Padi, B.O. Akumadu, O. Faerch, C. Aloke, V. Meyer, I. Achilonu, Engineering a pseudo-26-kda schistosoma glutathione transferase from bovis/haematobium for structure, kinetics, and ligandin studies, Biomolecules 11 (2021), https://doi.org/10.3390/biom11121844.
[35] O. Onisuru, I. Achilonu, Describing the ligandin properties of plasmodium falciparum and vivax glutathione transferase towards bromosulfophthalein from empirical and computational modelling viewpoints, J. Biomol. Struct. Dyn. (2024) 1–16, https://doi.org/10.1080/07391102.2024.2329291.
[36] I. Boušová, J. Hájek, J. Dršata, L. Skálová, Naturally occurring flavonoids as inhibitors of purified cytosolic glutathione S-transferase, Xenobiotica 42 (2012) 872–879, https://doi.org/10.3109/00498254.2012.670737.
[37] I. Boušová, L. Skálová, Inhibition and induction of glutathione S-transferases by flavonoids: possible pharmacological and toxicological consequences, Drug Metab. Rev. 44 (2012) 267–286, https://doi.org/10.3109/03602532.2012.713969.
[38] L. Miu, N.S. Bogatyreva, O.V. Galzitskaia, Radius of gyration is indicator of compactness of protein structure, Mol. Biol. (Mosk) 42 (2008) 701–706.
[39] A.M. da Fonseca, B.J. Caluaco, J.M.C. Madureira, S.Q. Cabongo, E.M. Gaieta, F. Djata, R.P. Colares, M.M. Neto, C.F.C. Fernandes, G.S. Marinho, H.S. dos Santos, E.S. Marinho, Screening of potential inhibitors targeting the Main protease structure of SARS-CoV-2 via molecular docking, and approach with molecular dynamics, RMSD, RMSF, H-bond, SASA and MMGBSA, Mol. Biotechnol. 66 (2024) 1919–1933, https://doi.org/10.1007/s12033-023-00831-x.

[40] B.O. Akumadu, R. Pandian, J. Olfsen, R. Worth, M. Thulo, T. Mentor, S. Fanucchi, Y. Sayed, H.W. Dirr, I. Achilonu, Molecular basis of inhibition of Schistosoma japonicum glutathione transferase by ellagic acid: insights into biophysical and structural studies, Mol. Biochem. Parasitol. 240 (2020) 111319, https://doi.org/10.1016/j.molbiopara.2020.111319.

[41] H. Chaaban, I. Ioannou, L. Chebil, M. Slimane, C. Gérardin, C. Paris, C. Charbonnel, L. Chekir, M. Ghoul, Effect of heat processing on thermal stability and antioxidant activity of six flavonoids, J. Food Process. Preserv. 41 (2017) e13203, https://doi.org/10.1111/jfpp.13203.

[42] S. Lin, J. Xiao, Impact of thermal processing on dietary flavonoids, Adv. Food Nutr. Res. 108 (2024) 1–34, https://doi.org/10.1016/bs.afnr.2023.10.002.

[43] S.-Q. Liu, X.-L. Ji, Y. Tao, D.-Y. Tan, K.-Q. Zhang, Y.-X. Fu, Protein folding, binding and energy landscape: a synthesis, INTECH Open Access (2012), https://doi.org/10.5772/30440. Publisher Retrieved from.

[44] R. Perozzo, G. Folkers, L. Scapozza, Thermodynamics of protein-ligand interactions: history, presence, and future aspects, J. Recept. Signal Transduct. 24 (2004) 1–52, https://doi.org/10.1081/RRS-120037896.

[45] L. Zhao, H.L. Ciallella, L.M. Aleksunes, H. Zhu, Advancing computer-aided drug discovery (CADD) by big data and data-driven machine learning modeling, Drug Discov. Today 25 (2020) 1624–1638, https://doi.org/10.1016/j.drudis.2020.07.005.

[46] J. Wang, P.R. Arantes, A. Bhattarai, R.V. Hsu, S. Pawnikar, Y. Ming M. Huang, G. Palermo, Y. Miao, Gaussian accelerated molecular dynamics: principles and applications, Wiley Interdiscip. Rev. Comput. Mol. Sci. 11 (2021) e1521, https://doi.org/10.1002/wcms.1521.

[47] K. Zuo, A. Kranjc, R. Capelli, G. Rossetti, R. Nechushtai, P. Carloni, Metadynamics simulations of ligands binding to protein surfaces: a novel tool for rational drug design, Phys. Chem. Chem. Phys. 25 (2023) 13819–13824, https://doi.org/10.1039/d3cp01388j.