

https://dspace.org/

Unravelling selectivity discrepancies of protoporphyrin binding to glutathione transferase A comparative analysis of molecular dynamic simulated versus implicit solventminimized protein models

Neo Padi; Sadhna Mathura; Ikechukwu Achilonu

https://hdl.handle.net/10539/45806
WOS

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

Journal of Molecular Graphics and Modelling 136 (2025) 108971

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/jmgm

Unravelling selectivity discrepancies of protoporphyrin binding to glutathione transferase: A comparative analysis of molecular dynamic simulated versus implicit solvent-minimized protein models

Neo Padi [a], Sadhna Mathura [b], Ikechukwu Achilonu [a,*]

[a] Protein Structure-Function and Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg, 2050, South Africa
[b] Molecular Sciences Institute, School of Chemistry, University of the Witwatersrand, P.O. Wits, Johannesburg, 2050, South Africa

ARTICLE INFO

Keywords:
Schistosomiasis
Glutathione transferase
Porphyrin
Antihelminth
Inhibition
Molecular docking
Molecular dynamic simulation

ABSTRACT

Schistosomiasis is a neglected tropical disease caused by parasitic trematodes, which are an ongoing global health and veterinary concern owing to their acquired drug resistance pressure to available treatment. There is a need for a new generation of effective anthelmintics for preventive and therapeutic purposes. Natural products, such as porphyrins, have been reported to inhibit the main detoxification enzymes in these parasites, called glutathione transferases, which help them evade immune response and drug therapy, thus making them good drug targets. Computational modelling was used to screen potential inhibitors out of 461 protoporphyrin IX-like compounds, including a potent known inhibitor called bromosulfophthalein. However, unlike traditional docking, where the stable energy-minimized structure is used, a short molecular dynamic simulation step was added to yield the most averaged protein structure conformation as the starting point of high throughput virtual screening. Here, it was shown that the starting point is crucial as the results suggested different lead compounds; for the 26-kDa *japonicum* GST, the top-scoring compounds were CID: 122690402 for the minimized structure and CID: 137797052 for the MD-simulated structure. Similarly, for the 28-kDa *haematobium* GST, the lead compounds were CID: 70415734 and CID: 69301914, respectively. These results highlight the importance of incorporating protein dynamics into structure-based drug design and provide valuable insights into the development of porphyrin-based therapeutics against schistosomiasis and other helminthic infections.

1. Introduction

Neglected tropical diseases continue to plague the public health and veterinary systems in impoverished regions worldwide. Among these, schistosomiasis, caused by parasitic trematodes, is a formidable adversary, afflicting over 200 million individuals globally [1,2]. Its stronghold spans sub-Saharan Africa, Asia, and the Middle East, with a staggering impact on human health [3]. Schistosomiasis manifests in two primary forms – urogenital and intestinal, each associated with different species of parasitic worms. The accumulation of eggs produced by these parasites triggers complex immune responses, leading to a spectrum of symptoms, from the initial acute Katayama fever to the development of chronic lesions and fibrosis, potentially culminating into cancer in later stages [4].

For over four decades, mass administration of praziquantel, an effective treatment targeting adult worms, has been a mainstay in combating this disease [5]. Nevertheless, the battle against schistosomiasis faces multifaceted challenges that persist. Inadequate healthcare infrastructure contributes to schistosomiasis transmission, while increased rates of travel and migration foster genetic hybrid vigour among schistosome species [6]. Moreover, the emergence of drug-resistant strains poses a forthcoming threat to public health and medicine, potentially leading to increased treatment expenses, prolonged hospitalizations, and elevated mortality rates [7]. This underscores the urgent need to develop a new generation of anthelminthic drugs to combat these resilient strains.

Schistosomes are hemiparasites that rely on the host's haemoglobin for sustenance and reproduction, inadvertently producing toxic heme as a by-product. These parasites have evolved mechanisms for heme degradation and elimination, with the enzyme glutathione transferase

* Corresponding author.
E-mail address: Ikechukwu.Achilonu@wits.ac.za (I. Achilonu).

https://doi.org/10.1016/j.jmgm.2025.108971
Received 12 August 2024; Received in revised form 21 January 2025; Accepted 29 January 2025
Available online 31 January 2025
1093-3263/© 2025 Published by Elsevier Inc.

(GST) playing a crucial role [8]. GST exists in two isoforms, 26 and 28 kDa, compensating for the absence of phase I detoxification systems, such as cytochrome P-450, in these parasites [9]. These enzymes detoxify electrophilic compounds by binding them to glutathione, rendering them hydrophilic and facilitating excretion, as evidenced by the high abundance of glutathione adducts in schistosomes. Inhibiting GSTs could disrupt detoxification processes, leading to the accumulation of harmful substances such as hydrogen peroxide and ferriprotoporphyrin IX, ultimately resulting in parasite death. A promising GST inhibitor is bromosulfophthalein (BSP), a dye used in liver function tests [5,10,11]. This compound has served as a benchmark for developing more potent porphyrin-based inhibitors. There has been increasing interest in natural products, that exhibit biocompatibility and non-toxicity, with particular emphasis on porphyrins due to their abundance in haemoglobin the primary nutrient source for schistosomes. Moreover, porphyrins share structural similarities with toxic heme and have demonstrated inhibitory effects on GSTs [12].

This study investigates the binding selectivity of protoporphyrin to GSTs by comparing explicit solvent-equilibrated and implicit solvent-minimized protein models, addressing discrepancies in ligand-binding dynamics, stability, and selectivity. While implicit models offer computational efficiency by approximating solvation effects using Generalized Born (GB) or Poisson-Boltzmann (PB) equations, they often fail to capture critical solvent-solute dynamics such as hydrogen bonding. In contrast, explicit solvent models derived from molecular dynamic (MD) simulations more accurately represent a physiologically relevant framework by explicitly modelling individual solvent molecules and replicating detailed protein-ligand interactions under aqueous conditions. Fig. 1 illustrates the transition from the energy-minimized to equilibrated MD-simulated models, a key step in addressing challenges in protein structure preparation for high-throughput virtual screening (HTVS) in identifying viable drug candidates. Using MD simulations and binding free energy calculations, this study evaluates how solvent representation influences the structural and energetic determinants of protoporphyrin binding. The findings not only enhance our understanding of GST functionality and ligand specificity but also demonstrate that MD-simulated models significantly improve HTVS accuracy by providing a more physiologically relevant framework. This comparative approach bridges methodological gaps in computational biochemistry and advances precision in modelling protein-ligand interactions, offering robust strategies for targeting GST-mediated resistance mechanisms in parasitic diseases. Ultimately, this work contributes to rational drug design and the discovery of effective therapeutics by aligning computational predictions with experimental outcomes.

2. Materials and methods

2.1. Ligand and receptor preparations

The ligand of interest, protoporphyrin IX, was submitted to PubChem for an analogue search. This search yielded 3D structures of 913 similar compounds, which were extracted in Structured Data File (SDF) format for ligand preparation using Maestro v13. The preparation workflow involved energy minimization employing the OPLS 2005 force field, with heavy atom restraints set to achieve convergence at a cut-off of 0.3 Å root mean square deviation (RMSD). Possible protonation states of the molecules were generated at pH 7.0 ± 2, and the pKa of these states were predicted using the Epik module. Additionally, the ligands underwent desalting and the generation of tautomeric states (~3 per ligand). For molecules with multiple chiral centres, certain centres were restricted, while others were varied to produce chemically plausible structures in low-energy states. The finalized structures were compiled into a compressed Maestro file.

In terms of the protein structures, the atomic coordinates of Sj26GST (PDB 6RWD) and Sh28GST (PDB 1OE8) were extracted from the RCSB-PDB database and introduced to the protein preparation wizard module implemented in Maestro v12. During preparation, ligands, solvent molecules, heterogeneous groups and water molecules located 5 Å away from the protein structures were removed. Bond orders were assigned, and hydrogen atoms were added to the Apo crystal structures using default PROPKA parameters. The processed structures were then subjected to energy minimization (Minimz model) for refinement. To ensure structural integrity, the stereochemistry of the proteins was validated, confirming that no significant perturbations were introduced during preparation. Following this, the protein structures underwent a 200 ns molecular dynamic (MD) simulation to produce the average equilibrated models (MDS).

2.2. High-throughput virtual screening

The 461 molecules, previously generated utilizing the LigPrep tool and corresponding grid files for the 26-kDa Minimized, 26-kDa MD-simulated, 28-kDa Minimized and 28-kDa MD-simulated GST models

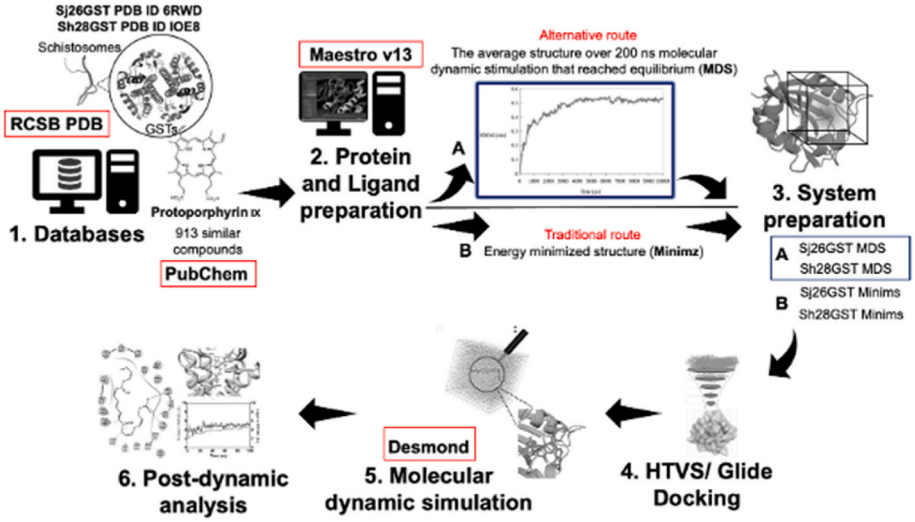


Fig. 1. Schematic representation of the computational workflow for identifying porphyrin-based antihelminthic compounds targeting schistosomiasis. The workflow incorporates an additional molecular dynamic (MD) simulation step following protein preparation to yield equilibrated structures. This approach enhances structural accuracy compared to traditional energy-minimized structures used in computational modelling.

in Maestro underwent a subsequent phase of analysis through a High-Throughput Virtual Screening (HTVS) module. The workflow commenced with Glide HTVS, wherein chains A and B were employed for blind ligand docking despite known characterized binding sites to accommodate novel allosteric sites. This process exclusively applied the QikProp filter, assessing pharmacokinetic and physicochemical (ADMET) properties. Prime MM/GBSA was also employed to predict the free-binding energy for ranking purposes. The implemented robust methodology resulted in identifying potential hits. Their assessment was conducted based on Glide docking scores and free binding energy, providing a quantitative measure of ligand-receptor binding affinity.

2.3. Induced-fit ligand docking

The highest-ranked molecules identified in the HTVS stage of Maestro were isolated and subjected to induced-fit docking (IFD) employing a mixed molecular docking and dynamic protocol. The standard IFD protocol was briefly applied to the centroid encompassing the entire protein dimer. Furthermore, SiteMap version 2,6 Schrödinger was used to validate potential binding sites and analyze alterations in their properties between the average MD-simulated model and the energy-minimized model. The druggability score (Dscore) was a focal metric correlating with Sitescore but possessing a distinct coefficient. Dscore gauges the druggability of a site by considering factors such as pocket volume, enclosure, and an unrestricted level of hydrophobicity, making it more discriminatory than Sitescore.

The docking simulations were conducted using the OPLS 2005 force field in an implicit solvent model. Various constraints were introduced, encompassing H-bonds and metal ions in the initial and re-docking stages. Additionally, constraints included sampling ring conformations with a 2.5 kcal/mol energy barrier and applying a non-polar conformation penalty on amide bonds. The scaling factor for both the receptor and ligand was set to 0.5, allowing a maximum of 20 poses for the ligand. Residues within 5 Å of the docked ligand underwent the Prime Refinement algorithm. The resulting refined protein-ligand complexes were ranked using Prime Energy. Receptor structures exhibiting minimal energy above 30 kcal/mol were advanced to the final Glide docking and scoring round phase. Each ligand was re-docked into every refined low-energy receptor structure in this stage, employing Glide XP default settings in the subsequent docking step.

2.4. Molecular dynamic simulation

The control Apo systems of Sj26GST and Sh28GST and complexes featuring top-scoring ligand poses of the various models underwent molecular dynamic simulation using Desmond, integrated into Maestro v13. The simulation encompassed two phases: system-building and production. The system builder module incorporated the TIP3P explicit solvent model with the OPLS 2005 force field in the system building phase. Precisely, the model was placed in an orthorhombic water box, encapsulating all protein and ligand atoms, with a 10 Å distance to the box face and angles $\alpha=\beta=\gamma = 90^{\circ}$. The box volume was minimized, and counter-ions from 0.15 M NaCl were added at least 20 Å from each atom to neutralize the system. Following solvation and ionization, the production phase consisted of eight stages employing default parameters of NVT and NTP statistical ensembles (constant atomic number, volume, temperature and pressure) to relax the system's geometry. The respective protein-ligand complexes were simulated at a constant temperature of 300 K and pressure of 1 atm for 500 ns.

Post-dynamic simulation analysis involved the simulations interaction diagram module, assessing RMSD and root-mean-square fluctuation (RMSF) of the protein backbone (Cα) and residues, radius of gyration (Rg) for target compactness, ligand properties (RMSD and solvent accessible surface area (SASA) and the distance between catalytic residues Tyr 7/9 and Arg 18/21 for Sj26GST and Sh28GST, respectively, both in the presence and absence of ligand These analyses collectively contributed to the determination of binding site affinity and stability.

2.5. Computational hardware

The computational tools used for molecular docking and simulation include two high-performance gaming computing units, namely the Windows OS PC with an AMD RYZEN Threadripper 1950X Processor, Asus Rog Strix X399-E Gaming Ryzen AMD with 40 MB of Cache Memory, 4 TB drive 6 GB/s 3D NAND internal SSD with 560 MB/s read, 520 MB/s write, a 4.0 GHz Precision Boost (up to 4.2 GHz) X399 chipset, Quad-Channel DDR4, motherboard, 64 GB (16 GB x 4) 3200 MHz DDR4 desktop gaming RAM and MSI GeForce RTX 2080 Ti 11 GB GDDR6 4352 CUDA core graphics card for Schrödinger Maestro v13. In contrast, the Ubuntu OS PC containing an AMD Threadripper 3990 × 4.3 GHz GT 1030 2 GB PRO high-performance workstation (288 MB Cache, 64x Cores, 4.3 GHz Turbo) with an MSI TRX40 PRO 10 G AMD Ryzen Threadripper motherboard, a GeForce RTX 2070 8 GB GDDR6 graphics card a 3200 MHz 64 GB gaming RAM, 1 TB M.2 SSD with up to 3.5 GB/s speed, and a 4 TB HDD was used for Schrödinger Desmond version, respectively.

3. Results

3.1. Protein conformation analysis

In pursuit of potential antischistosomal drugs, in silico investigations were conducted using Schrodinger software to focus on the interaction between *Schistosoma* GSTs and porphyrins. Before the virtual screening, the protein structures were prepared on Maestro, and it was observed that the protein conformation of the energy-minimized derived from a static crystal snapshot differed from the average conformation obtained from 200 ns molecular dynamics simulation that reached equilibration. Although the protein remained the same in both instances, the Sh28GST Minimized and MD-simulated models exhibited a slightly greater RMSD by 0.06 than the Sj26GST models with a value of 1.47 (Fig. 2).

3.2. Virtual screening

Only 50.49 % of the protoporphyrin IX-like molecules sourced from PubChem met the predefined criteria for drug-like properties and underwent subsequent virtual screening. A non-bias induced-fit docking approach was employed, mapping the entire protein rather than focusing on known binding sites. Notably, most porphyrin rings preferred binding at the dimer interface, presenting various poses with some projections towards the active hydrophobic and glutathione sites (Fig. 3). However, it is noteworthy that Sj26GST MD-simulated displayed an exception, as a substantial portion of the ligands predominantly occupied the active site.

3.3. Binding pocket Identification and druggability

The selective binding of protoporphyrin ligands to specific sites contingent upon the protein conformation was further elucidated by examining the binding sites' cavity alterations (Table 1). Consistent with the characteristics of GSTs, multiple binding sites were evident, with Dscores below 0.7 denoting difficulty, those above 0.8 deemed druggable, and values in between considered intermediate; values exceeding one were categorised as highly druggable, as described previously [13].

In the context of Sj26GST MD-simulated, a significant improvement is observed in the Dscore, elevating it to higher levels of druggability for both the active site and dimer interface. This contrasts with the cavities that were found to be difficult in the Minimized model. On the other hand, the Sh28GST MD-simulated model reveals a compromised dimer interface cavity compared to the Minimized model despite maintaining a comparable active site. Notably, some regions of the protein had cavities that had difficult druggability. Overall, the MD-simulated models

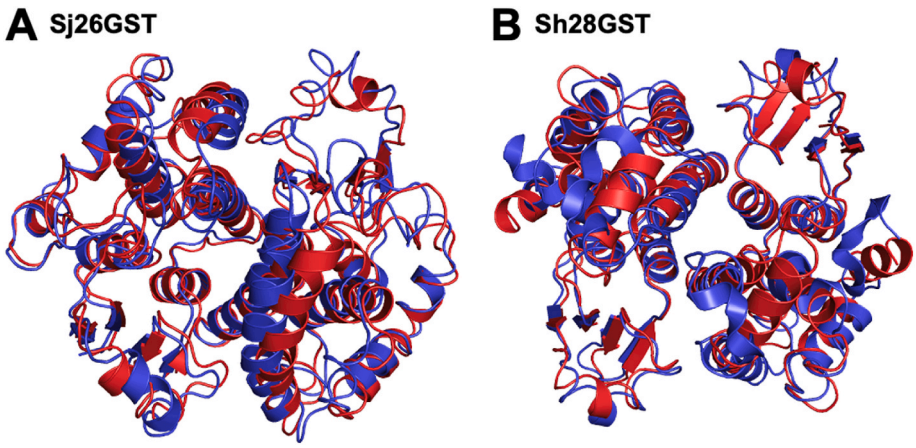


Fig. 2. Superimposed ribbon structures of energy-minimized protein structure (red) and average molecular dynamic simulated structure (blue) of **A**. *Schistosoma japonicum* 26-kDa glutathione transferase (Sj26GST) and **B**. *Schistosoma haematobium* 28-kDa glutathione transferase (Sh28GST). The structures were aligned using the Align tool in PyMol, whereby the RMSD for Sj26GST (PDB 6RWD) was 1.47 Å; then for Sh28GST (PDB 1OE8), it was 1.53 Å. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

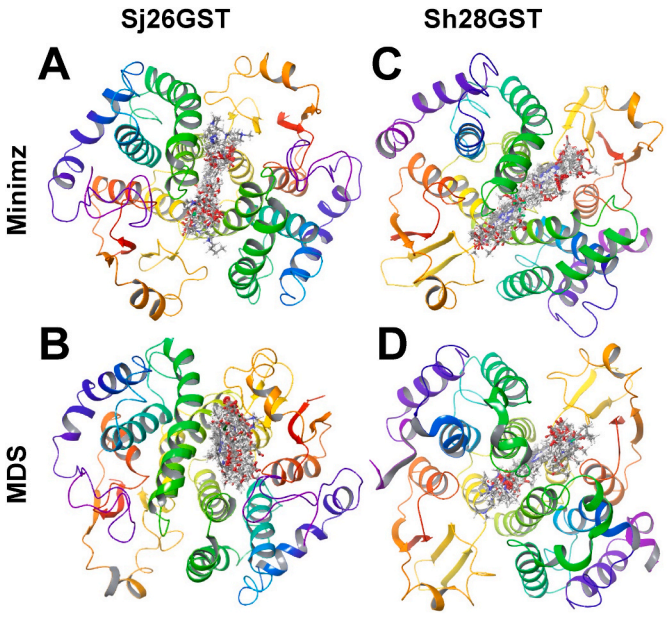


Fig. 3. The preferred binding site for protoporphyrin IX-like compounds in *Schistosoma* GSTs. Sj26GST Minimized (**A**), Sh28GST Minimized (**C**) and Sh28GST MD-simulated (**D**) tend to lean towards the L-site at the dimer interface and some ligands projections towards the H-site, whereas Sj26GST MD-simulated (**B**) on the other hand is towards the H-site.

exhibit more druggable cavities than the Minimized models (Table S1).

3.4. Binding energies

Nevertheless, the focus was on the most promising docked ligands and their respective properties and associated energies (Table 2). The lowest free binding energy (Δ*G*) (kcal/mol), a critical indicator of high affinity and stability [14], exhibited a range from −50 to −10.34 kcal/mol, following the order of CID: 137797052 (Sj26GST MD-Simulated), 70415734 (Sh28GST Minimized), 69301914 (Sh28GST MD-Simulated) and 122690402 (Sj26GST Minimized) (Table S2, Table S3 and Fig. S2). Remarkably, distinct protein conformations exhibited varying top ligands, even though they shared the same protein structure. Notably, these ligands surpassed the benchmark ligand, BSP, that was included in the drug library of HTVS, which is documented to possess a Glide XP score average of −7.20 kcal/mol [15] concerning *Schistosoma* GSTs as the target.

Table 1
SiteMap proposed binding cavities on different models of Sj26GST and Sh28GST in Maestro, whereby the pocket volume, enclosure, and hydrophobicity contributed to the druggability score (D-score).

Model	Cavity Rank	Dscore	
		Minimized	MD-Simulated
Sj26GST	1	1.07	1.059
	2	0.939	1.011
	3	0.555	0.962
	4	0.502	0.72
	5	0.548	0.69
Sh28GST	1	0.964	0.98
	2	0.751	0.624
	3	0.712	0.737
	4	0.545	0.416
	5	0.525	0.739

Table 2
Top docked protoporphyrins ligands for *Schistosoma* GST isoforms 26 and 28 kDa at varying conformations, energy-minimized vs the averaged equilibrated molecular dynamic simulation showing their free binding energy (Δ*G*bind) (kcal/mol) calculated by Molecular Mechanics/Generalized Born Surface Area (MM-GBSA.

System	PubChem CID	MM-GBSA ΔG_{bind} (kcal/mol)
Sj26GST MD-simulated	137797052	−50.08
Sj26GST Minimalized	122690402	−10.34
Sh28GST MD-simulated	69301914	−12.41
Sh28GSTMinimized	70415734	−32.11

3.5. Conformational stability

The top-ranking ligand complexes, encompassing various models and Apo models, underwent further investigation through Desmond simulations to gain insights into the contributions of molecular dynamics towards protein-ligand interactions and stability (Figure S8 and Figure S9). In the case of Sj26GST, the MD-simulated model demonstrated superior stability, achieving early convergence at 60 ns with a Cα RMSD of ±2 Å and minimal data distribution. In contrast, the Minimized model exhibited convergence at ± 3 Å, while the Apo structure was dynamic and failed to reach convergence, displaying the highest data

distribution skewed to the left (Fig. 4A and Fig. S4A). Analysis of Cα RMSF unveiled significant alterations in several residues, surpassing 3 Å across all three systems (Fig. 4B). The Minimized model exhibited comparable side-chain fluctuations to the Apo state, with exceptions noted in the C-terminal of chain A for the MD-simulated model and the dimer interface, chain B H-site and its C-terminus for the Apo state. The MD-simulated model demonstrated the least overall fluctuations, although significant outliers were observed in the active site of chain B. All the data is skewed to the right, emphasizing increased fluctuations in the specified regions (Figure S4 B). Regarding compactness (Fig. 4C), the Apo and MD-simulated were comparable at ±21.5 Å, while the Minimized model was less compact and less stable, converging at 22 Å after 180 ns (Fig. S4C).

For Sh28GST, the trend indicated convergence of the Apo and Minimized models at 220 ns at approximately 3 Å with negligible variation (Fig. 4D). In contrast, the MD-simulated achieved convergence later at 340 ns, displaying the highest data distribution skewed to the right (Fig. S4D). Several residues in the chain A C-terminus, chain B dimer interface, and active site were significantly impacted with RMSF above 3 Å. The MD-simulated model, however, showed the least fluctuations with the Apo leading the chain A C-terminus fluctuation, while the Minimized model led the chain B dimer interface (Fig. 4E and Fig. S4E). The compactness (Fig. 4F) remained relatively unchanged and comparable at 22 Å. However, the Apo slightly deviated to 21.5 Å at 360 ns, indicating increased compactness and stability (Fig. S4F).

Overall, the MD-simulated model in Sj26GST exhibits the least fluctuations in Cα RMSD, signifying greater stability than the Minimized and Apo models. Furthermore, Sj26GST systems converge earlier than the Sh28GST models, where the Apo, MD-simulated and Minimized models have a comparable Cα RMSD that is slightly higher. This observed trend aligns with existing literature indicating that Sj26GST is more stable than Sh28GST [16]. The Cα RMSF analysis elucidates distinctions in residue side chain dynamics. Sj26GST exhibits more residues with more pronounced fluctuations than Sh28GST, which lacks significant changes in chain A. However, the Cα RMSF profiles of Sh28GST models are generally comparable, with variations in the quantity or location of outliers observed in ascending order: Minimized, Apo, and MD-simulated. Furthermore, compactness is comparable upon convergence, despite Sh28GST models achieving convergence earlier than Sj26GST models, indicating differences in the temporal dynamics.

3.6. Ligand properties

Regarding ligand properties, ligands bound to MD-simulated models demonstrated minimal RMSD at ± 3 Å even though the Sh28GST systems only converged at 360 ns, signifying greater stability compared to Minimized models in both proteins (Fig. 5A, Figure S5A, Fig. 5C and Figure S5C). The ligands maintained their positions in the respective binding pockets. In Sj26GST, the area accessible to the solvent was comparable in the MD-simulated and Minimized models at ± 200 Å2.

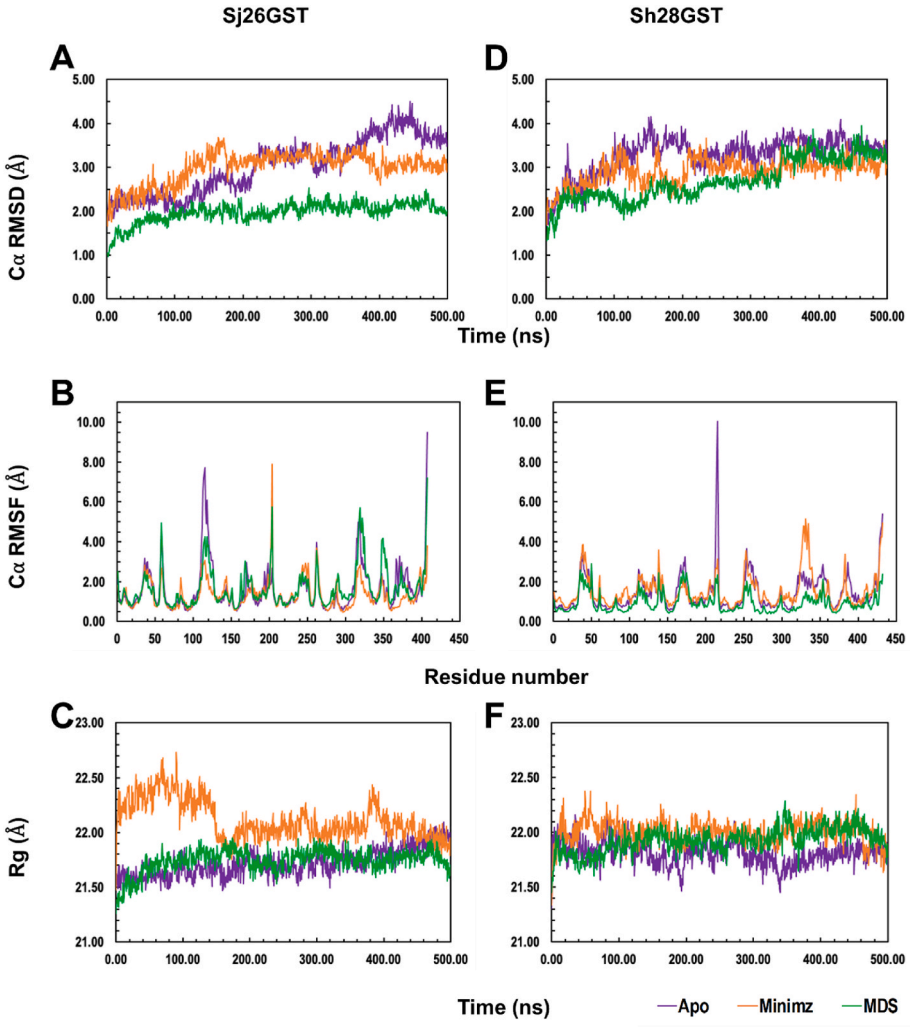


Fig. 4. Trajectory of 500 ns simulation showing the root mean square deviation (RMSD) of **A/D** Cα atoms and **B/E** root mean square fluctuation per residue side chain as well as **C/F** radius of gyration (Rg) for the Apo-proteins, Sj26GST and Sh28GST, respectively and ligand-bound complex of the various models.

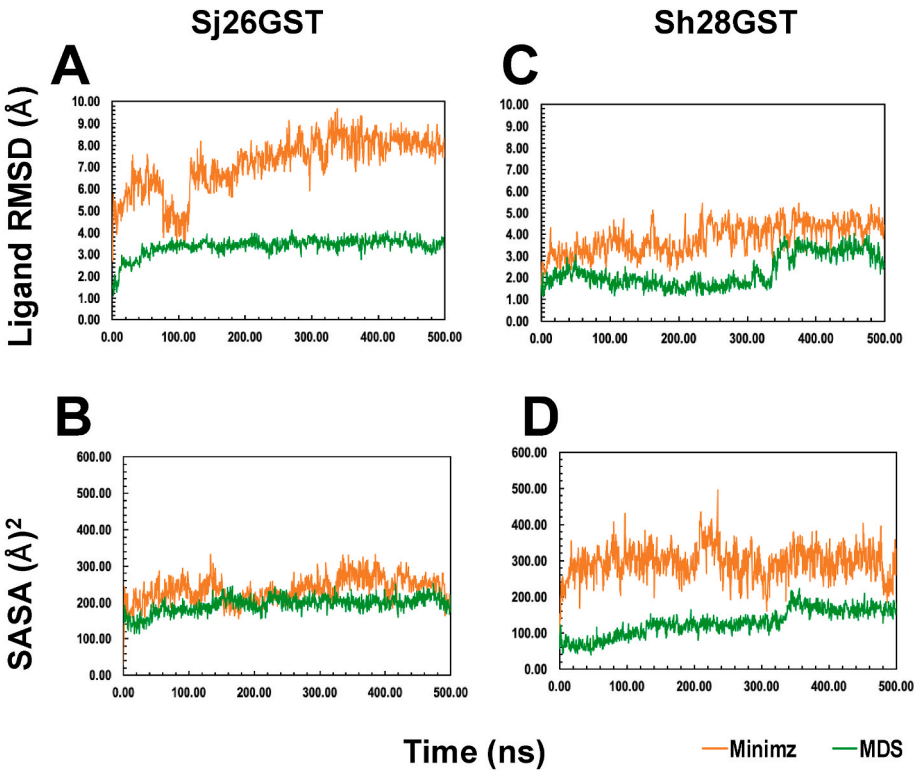


Fig. 5. The root mean square deviation of the ligand with respect to the receptor **(A and C)** and corresponding solvent-accessible surface area (**B and D**). A single protein structure in different models chose different ligands that have variable stability and solvent exposure.

However, Sh28GST MD-simulated exhibited only half of the exposed area compared to the ±300 Å2 of Minimized, increasing buried regions and decreasing flexibility/conformational entropy (Fig. 5B, Figure S5B, Fig. 5D and Figure S5D). The observed trend was also supported by the HTVS study which illustrated different ligand conformations stacking on top of each other without significant shifts, especially in MD-simulated models as shown by the whole at the centre which is not visible in Minimized systems (Fig. 6). In addition, SwissADME was used to substantiate the toxicity of the drugs (Fig. S3) and target predictions were all below 50 % and insignificant.

3.7. Residual level protein-ligand interaction

This section focuses on dissecting the ligand-protein interactions of the four models. The 500 ns simulation results unravel the nuanced dynamics of these interactions. In the Sj26GST Minimized scenario, the ligand interactions exhibit a milieu of positively charged contacts interspersed with subtle polar and hydrophobic groups and a minor fraction of negatively charged contacts (Fig. 7A). Notably, there is a significant Pi-cation interaction between the hydrophobic Tyr 103 and the amine moiety of the ligand ring. Water bridges are formed with the negatively charged Asp 100, and dominant hydrogen bonds are established via the positively charged Arg 107 and polar Gln 66 (Table 3 and Fig. S6A). Additionally, crucial residues such as polar Ser 67, mainly involved in a water bridge, and Tyr 110, which contributes to hydrogen bonds are identified (Fig. 8A).

In the case of Sj26GST MD-simulated, the top ligand showcases fewer positively charged groups, exhibiting well-distributed polar and hydrophobic contacts and minimal positively charged contacts (Fig. 7B). Stabilization is achieved through water bridges with specific residues on one side of chain B (Tyr 6, Lys 10, Leu 12, Val 13, Leu 54, Arg 102 and Leu 202), along with Asn 53 in chain A. Polar contacts are formed with Asn 53, Ser 67, Ser 106 and Gln 66 in chain B, as well as Lys 124 in chain A, through H-bonds. A unique salt bridge is observed between ligand amine and ionic Asp 100, accompanied by a hydrophobic contact with Phe 51 in chain B (Table 3 and Fig. S6B). Furthermore, highlighted interactions include a water bridge with Asp 120 in chain A and a 50 % probability between H-bonds and water bridge in the case of Ser 123 and Gln 203 (Fig. 8B).

Turning to Sh28GST, in the Minimized scenario, the ligand predominantly engages in polar contacts with minor positive and negative groups. Potential metal coordination with iron is highlighted (Fig. 7C). Unique amino acids contributing to these interactions included the positively charged Arg 16 in hydrophobic Pi-cation with the centre of a pyridine ring, Asp 104 forming a salt bridge in chain B and H-bond in chain A, and Glu 103, forming a salt bridge with amine and possibly coordinating with iron in chain A, with an additional H-bond in chain B (Table 3 and Fig. S6C). Additional ligands highlighted include Leu 72, Tyr 109, Leu 113 and Lys 131 involved in hydrophobic interactions, as well as His 107 and His 110 participating in Pi cation. Lastly, Gly 100 and Glu 106 contribute to water bridges (Fig. 8C).

Conversely, in the Sh28GST MD-simulated system, the ligand is characterised by negatively charged contacts, with a minor contribution of hydrophobic groups, hinting at potential metal coordination with iron (Fig. 7D). Detailed interactions included amine contacts with Asp 104 via ionic interactions and possible metal coordination with iron. His 107 engages in Pi-cation interactions, while residues like Arg 52, Lys 131, Glu 103, and Glu 106 form in H-bonds (Table 3 and Fig. S6D). Moreover, Gln 101 and Gly 70 participated in water bridges, and His 110 in hydrophobic interactions (Fig. 8D).

Overall, Fig. 8 provides additional insights into the various interactions between the respective top ligand and protein conformation, correlating with Fig. 7 and Table 3 interaction probabilities with a 30 % frequency threshold over the 500 ns simulation, thereby narrowing down other possibilities to showcase dominant bonds (Fig. S7).

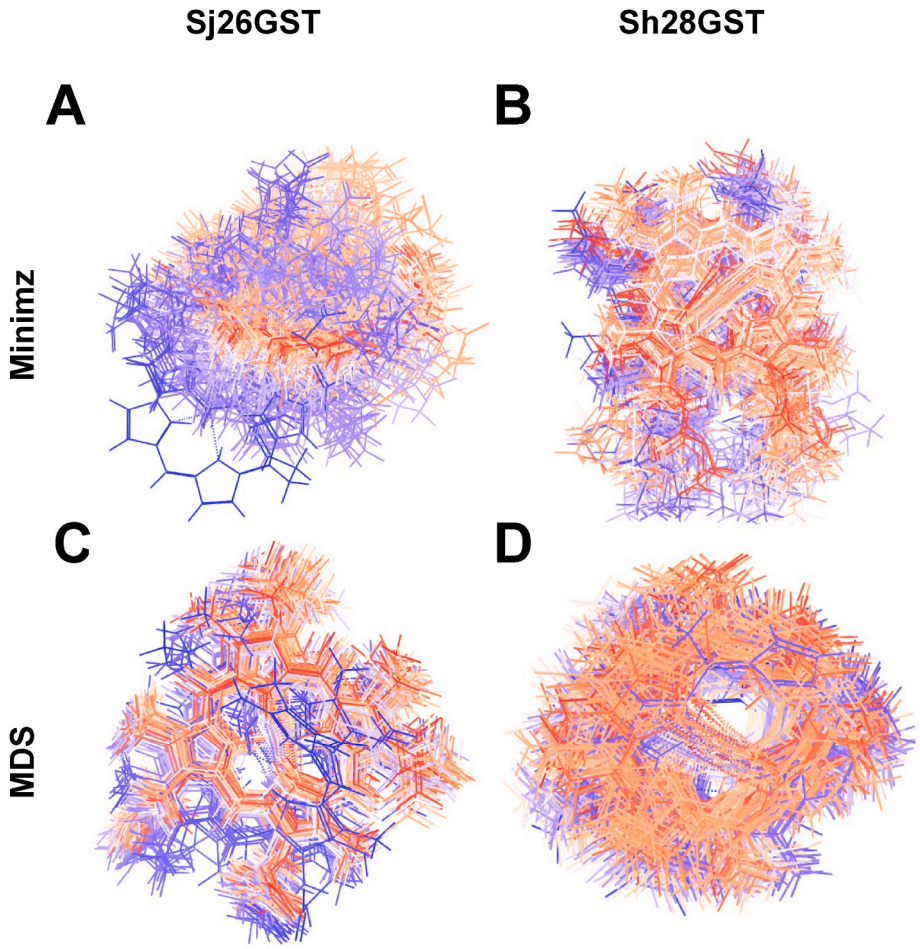


Fig. 6. The superimposed images of the top ligands' poses over the 500 ns simulation in the respective binding pocket. The Minimized systems, **A** CID: 122690402 and **B** CID: 70415734 show greater fluctuations than the MD-simulated ligands, **C** CID: 137797052 and **D** CID: 69301914.

3.8. Conformational shifts in catalytic residues

The catalytic Tyr is pivotal for stabilizing GSH thiolate anion facilitating nucleophilic attack, and its conformation holds significant importance, particularly concerning the phenyl ring's ability to adopt either an open or closed form. Conformational changes were monitored based on the ring structure distance between Tyr and the nearby conserved Arg, which is also involved in GSH stabilization (Fig. 9). For the Sj26GST system, both chains in the Minimized model closely resembled chain A in both Apo and MD-simulated models, all within a range of ±8 Å (Fig. 9A). Notably, in the Apo system, chain B reached convergence after 220 ns but at a slightly higher distance of ±10 Å. In the MD-simulated system, chain B exhibited an even higher distance, reaching ±12 Å (Fig. 9C).

In the case of Sh28GST, the revealed distance in chain B of the Apo structure displayed a slightly broader data distribution, converging to ±4 Å after 200 ns, and with chain A at ± 3 Å-overall smaller compared to Sj26GST. However, in the Minimized model, chain B demonstrated a greater distance of ±8 Å upon 60 ns convergence, while chain A remained at ± 3.5 Å (Fig. 9B). In the MD-simulated model, chain A exhibited the highest fluctuation in distance at ± 6 Å, while chain B remained at ± 3.5 Å (Fig. 9D).

4. Discussion

4.1. Therapeutic potential of Schistosoma GSTs

Schistosoma GSTs emerge as promising drug candidates, given their pivotal role in the detoxification system. They actively suppress the formation of toxic heme by binding porphyrins, facilitating their transport and metabolism. This interaction provides a molecular mechanism for regulating the porphyrin pathway. The substantial expression throughout the parasite's life cycle [17] enhances its suitability for therapeutic targeting. However, the similarity between cytosolic GSTs expressed by helminths and mammals presents a challenge. Basic alpha (α) GST mirrors Sh28GST, while neutral mu (μ) GST resembles Sj26GST, both sharing a common catalytic Tyr [18], which is also observed in acidic pi (π) GST that dominates the liver [19]. These GSTs exhibit Arg and Lys arrangements stabilizing GSH binding [18] (Fig. S1). Despite these similarities, distinct structural differences unique to helminth GSTs provide avenues for selective targeting [20]. Notably, *Schistosoma* GSTs exhibit a unique interaction with model inhibitors and anthelmintic compounds, such as BSP [10,11,15] and PZQ [21], respectively. These interactions occur at the dimer interface, demonstrating discrimination against human GSTs.

4.2. Conformational dynamics in structure-based drug design

In structure-based drug design, it is crucial to meticulously identify and characterize binding sites. While drugs often target specific protein conformations to modulate their activity, the understanding of protein dynamics through techniques like MD simulation unravels the physicochemical basis of their structures and biological function, thus aiding in designing targeted therapeutic interventions that have increased success at fewer costs of time and money [22] over the limitations of rigid crystal structures [23].

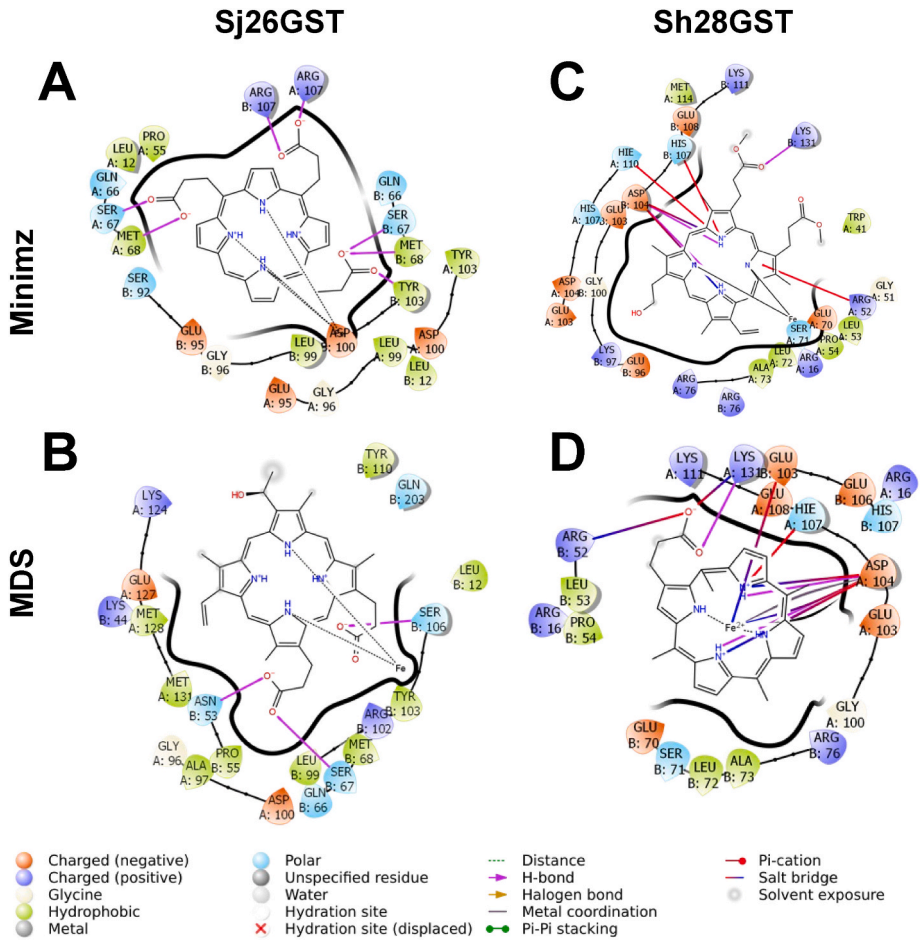


Fig. 7. 2D interaction plot of top poses of Sj26GST Minimized (**A**) and MD-simulated (**C**), respectively in complex with ligands. Similarly, the same information was captured for Sh28GST Minimized (**B**) and MD-simulated (**D**). The residues are within 4 Å from the ligand. The image was generated using the Maestro 2D interaction diagram implemented in Schrödinger Maestro v13.

Table 3
% Occupancy of the bond interaction between the top ligands and different models of *Schistosoma* GSTs. The colours represented are blue for salt bridge, grey for contacts made with the assistance of water in between, purple for H-bonds and red for hydrophobic Pi cations. Overall there is less engagement in the energy minimized models.

Sj26GST			Sh28GST		
Residue	Minimized	MD-simulated	Residue	Minimized	MD-simulated
Tyr 6	0	B: 72	Arg 16	A: 34	0
Lys 10	0	B: 73	Arg 52	0	B: 71, 98
Leu 12	0	B: 69	Glu 103[a]	A: 98	B: 93
Val 13	0	B: 41		B: 44	0
Asn 53	0	B: 37	Asp 104	B: 41, 52	B: 44, 50
Leu 54	0	B: 73	Glu 106[a]	0	B: 100
Gln 66	A: 31	B: 84	His 107	0	B: 30
Ser 67	0	B: 91, 98	Lys 131	0	A: 47, 98
Asp 100	A: 38, 39	B: 30, 85			
Arg 102	0	B: 64, 81			
Tyr 103	A: 60, 35 B: 61	0			
Ser 106	0	B: 73			
Arg 107	A: 77, 77, 79, 81	0			
Lys 124	0	A: 35			
Leu 202	0	B: 34, 50			

[a] Amino acids that interacted with nitrogen atoms bound to the metal ion centre.

In computational modelling studies, particularly in HTVS and other docking investigations, the conventional approach involves using a protein model prepared through a force-field applied energy minimization protocol. This protocol serves to assign H-bonding potential to the protein model and adjust the spatial orientation of side chains to alleviate clashes, fill in missing loops, and implement other structural corrections. Consequently, this process readies the protein model for downstream in silico modelling, including MD simulation. When such a minimized protein model undergoes MD simulation with an explicit solvent model, especially using the NPT (isothermal-isobaric ensemble) or canonical ensemble, the model evolves from its initial state and eventually converges after numerous simulation steps. However, the resultant converged or averaged model extracted from the simulation may exhibit variances compared to the initial minimized model, sometimes with RMSD reaching as high as 2.0 Å. This discrepancy suggests potential conformational changes that could influence ligand selectivity or affect the accuracy of docking or pharmacophore modelling, particularly concerning changes within the binding site cavities. This raises the fundamental question: which model should be favoured for in silico studies, including docking, HTVS, pharmacophore modelling, and others?

4.3. *Effect of protein conformation on ligand binding*

This study evaluates the optimal conformation for a class of protoporphyrins, comparing the energy-minimized model to the average accessed confirmation within 200 ns that reached MD simulation

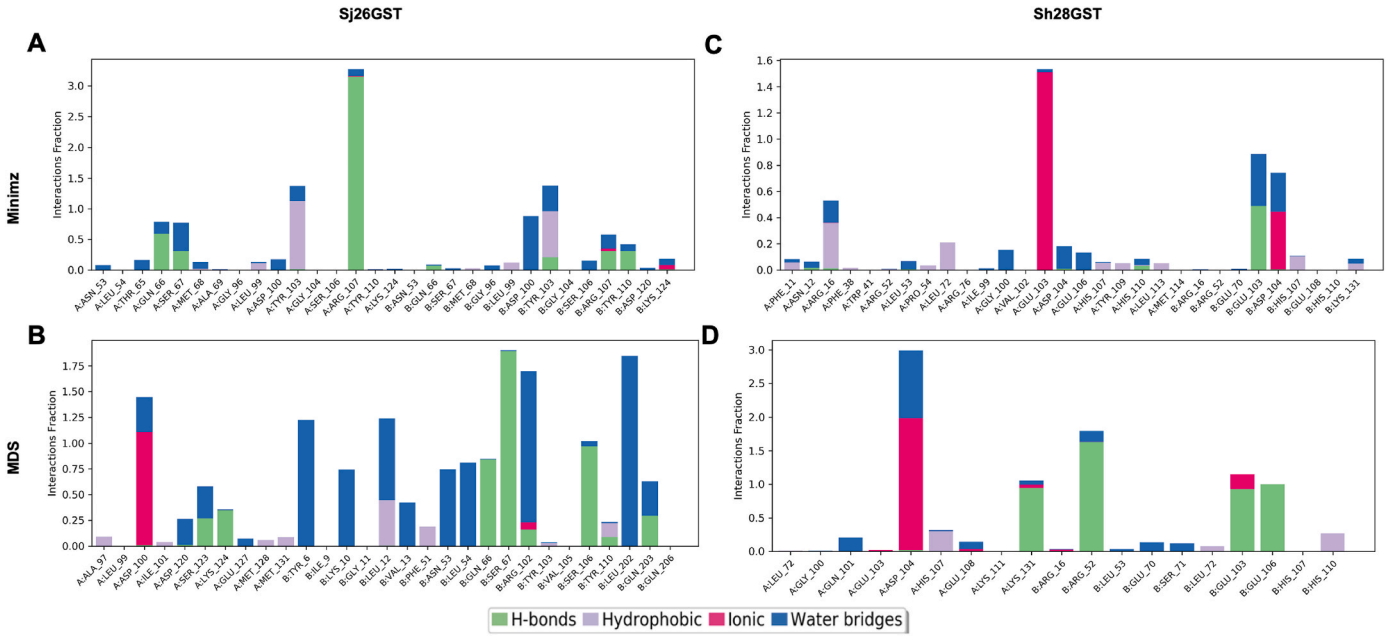


Fig. 8. The Stacked bar charts of the types of interactions the residues' side chains are involved in with the top ligands for Sj26GST Minimized **(A)** and MD-simulated **(B)** as well as for Sh28GST Minimized **(C)** and MD-simulated **(D).** The H-bonds are categorised into backbone acceptor, backbone donor, side-chain acceptor, and side-chain donor, whereas hydrophobic interactions include p-cation, p-p*, and other non-specific interactions. The stacked bar charts are normalised throughout the 500 ns trajectory, and the fraction of interaction can be interpreted to be a percentage time which could be more than 100 %.

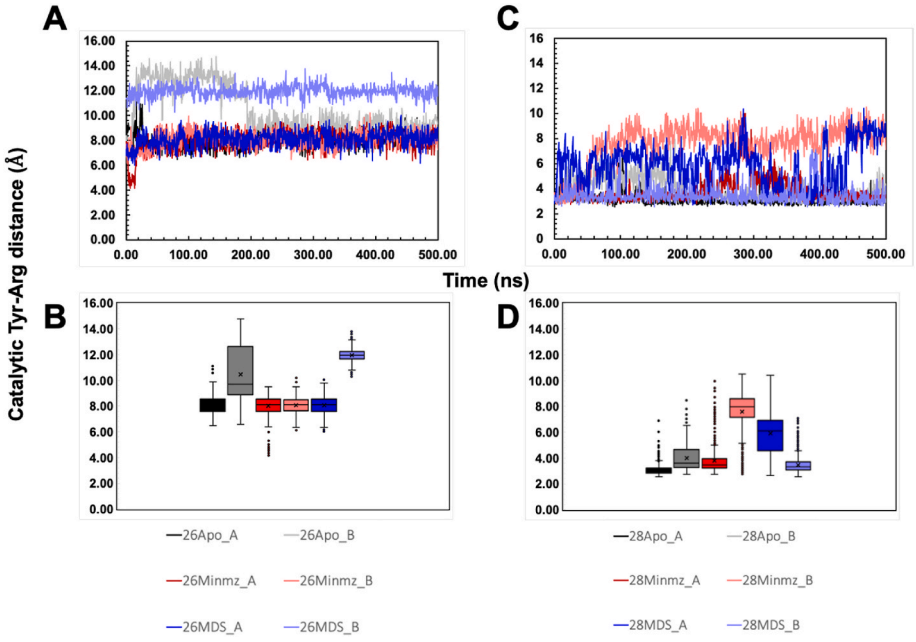


Fig. 9. The Atomic distance between key catalytic residues, namely Tyr10-Arg21 in Sh28GST and Tyr6-Arg17 in Sj26GST as a function of 500 ns simulation time.

convergence. The analysis specifically assesses two isoforms of *Schistosoma* GST. It is evident that proteins are flexible and access different models with relatively significant structural differences due to deviations in the atomic position of some residues as supported by the superimposition RMSD greater than one between the energy-minimized structures and the MD-simulated structures in both isoforms. The RMSD difference is comparable to the difference between isoforms Sj26GST and Sh28GST reported in the literature [15].

These different protein models of the same protein amongst *Schistosoma* GSTs were used to screen protoporphyrin compounds virtually. Only 50 % of the compounds from PubChem met the drug-like criteria and were used for induced fit docking. Interestingly, the porphyrins tend to bind to the dimer interface, but Sj26GST MD-simulated was the exception where they bound to the catalytic H-site. The trend displayed was further investigated by looking at the binding site cavities' character.

4.4. *Binding sites Variability and druggability in protein conformations*

For Sj26GST, the MD-simulated model enhances the druggability of the active site, contrasting with the Minimized model favouring only the dimer interface. Conversely, Sh28GST shows a different trend, with the

MD-simulated model decreasing druggability at the L-site compared to the Minimized model. The three binding sites; G, H and L, are collectively affected by conformational changes to accommodate a narrow range of ligands by fluctuating cavity volumes and solvent exposure, which moderately correlate with the Dscore with some exceptions. Larger, well-defined hydrophobic pockets are more likely to bind drug-like ligands [24] affecting the protein's function positively or negatively. The porphyrin ligands were biased to the H-site in Sj26GST MD-simulated as opposed to the dimer interface in Sj26GST Minimized and Sh28GST models. Moreover, the cavity that was ranked two with improved Dscore in Sj26GST MD-simulated to highly druggable correlated with the H-site; hence it was favoured, despite the dimer interface that was ranked one in all models, being closest to the latter ideal drug-binding pocket.

A notable finding emerged after a detailed examination of the top ligands within different systems, indicating that the same protein, at distinct conformations, exhibits a preference for different ligands. Specifically, in the case of Sj26GST, the MD-simulated model preferred 137797052, showcasing heightened spontaneity, while the Minimized model favoured 122690402. Conversely, for Sh28GST, the Minimized model bound to 70415734 with the greatest spontaneity, whereas the MD-simulated model preferred 693031914 with less pronounced spontaneity. It is essential to highlight that the Sj26GST MD-simulated model exhibited greater stabilization than the Minimized relative to the Apo state. This stabilization allowed some flexibility in sidechains without compromising the overall structural integrity, maintaining a comparable level of compactness and keeping the protein well-folded. In contrast, the Minimized model in Sj26GST experienced structural expansion and backbone variations at the expense of stability. Looking at Sh28GST, both models demonstrate similar backbone fluctuations and compactness. However, the MD-simulated model displayed the highest overall side-chain flexibility, contributing less energy towards a spontaneous reaction.

4.5. Key Residues and solvent contributions to protein-ligand binding

Analysis of ligand properties further revealed that ligands are confined within binding pockets, with MD-simulated models exhibiting the least fluctuations in both proteins. Furthermore, the SASA in Sj26GST systems were comparable. In contrast, the Sh28GST Minimized model displayed more exposure to solvent correlating with the highest molecular weight, thus leveraging energy from water molecules—a characteristic aligning with the native GST dynamic structure to accommodate a broad range of ligands, resulting in a more spontaneous reaction in the complex compared to the MD-simulated model, which had the lowest ligand efficiency in terms of surface area and lipophilic character.

Overall, MD-simulated models were characterized by attractive weak van der Waals forces and attractive strong electrostatic energy. The key distinguishing between the Sj26GST and Sh28GST isoforms was favourable solvation, rendering Sj26GST MD-simulated spontaneous and Sh28GST MD-simulated not. Conversely, minimized models encountered considerable steric hindrance, contributing to instability.

This discussion section focuses on unravelling the roles of residues and essential water molecules, often overlooked in conventional docking studies, through MD simulation. These water molecules play a crucial role in protein conformation, thereby influencing ligand binding affinity and effectiveness in modulating biological activity. The study underscores the substantial contribution of hydration and solvent effects to entropic energy within the thermodynamics framework [25].

The Sj26GST MD-simulated model revealed the incorporation of fourteen water molecules, resulting in a spontaneous shift in the top ligand preference towards the H-site. This contrasts with the Sj26GST Minimized model, which favoured the dimer interface. Notably, key residues such as Asp 100 and Ser 67 exhibited different orientations in the Sj26GST MD-simulated model, contributing to a salt bridge and H-bond, respectively, unlike the Minimized model, which relied solely on water bridges. Additionally, Gln 66 maintained the H-bond contacts, but there were chain-specific changes, with chain B involved in the MD-simulated model and chain A involved in the Minimized model. The arrangement of Arg 102 involved a water bridge in the MD-simulated model, while Arg 107 formed a salt bridge in the Minimized model. Comparison to human counterparts has highlighted alterations in the catalytic Tyr 6 forming a water bridge in the MD-simulated model while, in the Minimized model, Tyr 108 is typically involved in the structural integrity of the human GSTs L-site made a hydrophobic contact. Lastly, Arg 16, which is involved in GSH stabilization along with the C-terminal Asp 104, which forms a salt bridge, was implicated in contacting the ligand.

For Sh28GST, water molecules played a significant role in forming H-bonds, leading to a change in the properties of the dimer interface and favouring a ligand with a more negative charge in the MD-simulated system, in contrast to the polar Minimized model lacking water molecules. Noteworthy residues, including Asp 104, facilitated ionic contacts and the possibility of metal coordination with iron in chain B of the Minimized model and chain A of MD-simulated model. Asp 104 in Minimized also had a chance to participate in a water bridge in chain A, and Glu 106 in chain B, forming H-bonds in the MD-simulated system. In the Minimized model, the latter residue in chain A favoured a water bridge. Relatively to humans' GSTs, His 107 plays a role in the deprotonation of glutathione, facilitating its nucleophilic attack and the conformational changes inflicted on it.

In all systems, unless specified, residues contacted the oxygen groups from the propanoic acid side chains of the ligands. In essence, the ability of the same residue in different chains to partake in various interactions underscores the dynamic nature of proteins and ligand interactions.

4.6. Conformational influence on catalytic Mechanism and cooperative dimerization

The investigation into the Tyr-Arg distance in both scenarios reveals a consistent pattern wherein Apo protein structures naturally exhibit a small distance in chain A, while a more considerable distance is evident in chain B, attributed to the open and activating closed conformation of the catalytic Tyr phenyl ring towards Arg, respectively. However, this inherent structural feature undergoes modulation upon conformational changes, influencing ligand binding and further substantiating the observed trend.

In the Sj26GST system, this phenomenon is accentuated, correlating with the highest free energy observed, thereby reinforcing the characteristic cooperative dimerization [26] Nevertheless, the Minimized model compromised the close conformation in chain B by rendering it similar to Apo chain A, thereby failing to distinguish between the chains.

Regarding Sh28GST, in the Apo state, a comparable distance between chain A and chain B is observed in the open conformation. Similarly, the spontaneously favoured structure followed the closed conformation in chain A and open conformation in chain B, as exhibited in the Minimized model, with the opposite being true in the compromised structural integrity of the MD-simulated model.

Overall, chain B demonstrates a propensity to withstand the extended open conformation to collaborate with the Arg guanidino group in stabilizing the thiol group in GSH stabilization and forming a salt bridge with its Gln tail for activating conjugation. The dynamics of GSTs encompass adaptable regions prone to unpredictable structural changes, leading to their existence in various models. This insight into the Tyr-Arg distance dynamics provides a nuanced understanding of the interplay between ligand binding, protein integrity, and cooperative dimerization in GSTs with potential implications for enzymatic activity.

5. Conclusion

Undoubtedly, proteins exhibit a diverse array of conformations, with

ligands having the ability to stabilize specific configurations within this dynamic ensemble, thereby imparting distinct biological effects. This comprehensive investigation unveils the intricate dynamics of *Schistosoma* GSTs and their potential as promising therapeutic targets for combating schistosomiasis, utilizing protoporphyrin compounds to disrupt the detoxification of toxic heme. The incorporation of MD-simulated imparts crucial insights into the nuanced interplay of conformational changes, stability, and ligand interactions, all pivotal aspects for the strategic design of targeted therapeutic interventions. It is imperative to recognize that the efficacy of drugs extends beyond mere structural compatibility at a singular point; rather, it is intricately tied to the dynamic interplay of protein conformations. This consideration is paramount when embarking on the quest for lead drugs against diseases through virtual screening methodologies. The findings from this study underscore the significance of understanding and incorporating protein dynamics into the drug discovery process, emphasizing the necessity for a holistic approach to unravel the complexities of biological systems.

CRediT authorship contribution statement

Neo Padi: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Sadhna Mathura:** Writing – review & editing, Visualization, Supervision. **Ikechukwu Achilonu:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of ai and ai-assisted technologies in the Writing process

During the preparation of this, work the author(s) used a combination of Grammarly Inc. and ChatGPT 4o in order to improve the readability of the work. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

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.

Acknowledgements

TThis work was supported by the National Research Foundation (NRF) of South Africa through a CPRR research grant awarded to I.A. (Grant number CPRR23042899244). N.P. acknowledges the Department of Water and Sanitation for providing a PhD bursary.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmgm.2025.108971.

Data availability

The data generated and analyzed in this study primarily consists of theoretical data derived from computational modeling studies. Given the nature of the research, the data is not empirical but rather the result of simulations and predictive modeling. This dataset will be made available upon reasonable request from the corresponding author, subject to applicable institutional and ethical guidelines.

References

[1] P.J. Hotez, A. Fenwick, Schistosomiasis in Africa: an Emerging Tragedy in Our New Global Health Decade, Public Library of Science, San Francisco, USA, 2009, p. e485.
[2] W.H. Organization, World Health Statistics 2020, 2020.
[3] D.G. Colley, A.L. Bustinduy, W.E. Secor, C.H. King, Human schistosomiasis, Lancet 383 (9936) (2014) 2253–2264.
[4] C.H. King, M. Dangerfield-Cha, The unacknowledged impact of chronic schistosomiasis, Chron. Illness 4 (1) (2008) 65–79.
[5] C.R. Caffrey, Chemotherapy of schistosomiasis: present and future, Curr. Opin. Chem. Biol. 11 (4) (2007) 433–439.
[6] R. Bergquist, J. Utzinger, J. Keiser, Controlling schistosomiasis with praziquantel: how much longer without a viable alternative? Infect. Dis.f Poverty 6 (2017) 1–10.
[7] R.K. Prichard, M.-G. Basanez, B.A. Boatin, et al., A research agenda for helminth diseases of humans: intervention for control and elimination, PLoS Neglected Trop. Dis. 6 (4) (2012) e1549.
[8] D.P. McManus, A. Loukas, Current status of vaccines for schistosomiasis, Clin. Microbiol. Rev. 21 (1) (2008) 225–242.
[9] A. Torres-Rivera, A. Landa, Glutathione transferases from parasites: a biochemical view, Acta Trop. 105 (2) (2008) 99–112.
[10] 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 (12) (2021) 1844.
[11] K. Pooe, R. Worth, E.A. Iwuchukwu, H.W. Dirr, I. Achilonu, An empirical and theoretical description of Schistosoma japonicum glutathione transferase inhibition by bromosulfophthalein and indanyloxyacetic acid 94, J. Mol. Struct. 1223 (2021) 128892.
[12] I. Hamza, H.A. Dailey, One ring to rule them all: trafficking of heme and heme synthesis intermediates in the metazoans, Biochim. Biophys. Acta Mol. Cell Res. 1823 (9) (2012) 1617–1632.
[13] T.A. Halgren, Identifying and characterizing binding sites and assessing druggability, J. Chem. Inf. Model. 49 (2) (2009) 377–389.
[14] M.K. Gilson, J.A. Given, B.L. Bush, J.A. McCammon, The statistical-thermodynamic basis for computation of binding affinities: a critical review, Biophys. J. 72 (3) (1997) 1047–1069.
[15] A. Valli, I. Achilonu, Comparative structural analysis of the human and Schistosoma glutathione transferase dimer interface using selective binding of bromosulfophthalein, Proteins: Struct., Funct., Bioinf. 90 (8) (2022) 1561–1569.
[16] K. Pooe, M. Thulo, H. Makumbe, et al., Biophysical description of Bromosulfophthalein interaction with the 28-kDa glutathione transferase from Schistosoma japonicum, Mol. Biochem. Parasitol. 252 (2022) 111524.
[17] B.K. Kubata, M. Duszenko, K.S. Martin, Y. Urade, Molecular basis for prostaglandin production in hosts and parasites, Trends Parasitol. 23 (7) (2007) 325–331.
[18] B. Wu, D. Dong, Human cytosolic glutathione transferases: structure, function, and drug discovery, Trends Pharmacol. Sci. 33 (12) (2012) 656–668.
[19] P.M. Brophy, D.I. Pritchard, Parasitic Helminth Glutathione S-Transferases: an Update on Their Potential as Targets for Immuno-And Chemotherapy, 1994.
[20] D. Sheehan, G. Meade, V.M. Foley, C.A. Dowd, Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily, Biochem. J. 360 (1) (2001) 1–16.
[21] M.A. McTigue, D.R. Williams, J.A. Tainer, Crystal Structures of a Schistosomal Drug and Vaccine Target: Glutathione S-Transferase fromSchistosoma Japonicaand its Complex with the Leading Antischistomal Drug Praziquantel, Elsevier, 1995, pp. 21–27.
[22] A. Hospital, J.R. Goñi, M. Orozco, J.L. Gelpí, Molecular dynamics simulations: advances and applications, Adv. Appl. Bioinform. Chem. (2015) 37–47.
[23] R. Macarron, Critical review of the role of HTS in drug discovery, Drug Discov. Today 11 (7–8) (2006) 277–279.
[24] A.C. Cheng, R.G. Coleman, K.T. Smyth, et al., Structure-based maximal affinity model predicts small-molecule druggability, Nat. Biotechnol. 25 (1) (2007) 71–75.
[25] T. Beuming, Y. Che, R. Abel, B. Kim, V. Shanmugasundaram, W. Sherman, Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization, Proteins: Struct., Funct., Bioinf. 80 (3) (2012) 871–883.
[26] D.W. Nebert, V. Vasiliou, Analysis of the glutathione S-transferase (GST) gene family, Hum. Genom. 1 (2004) 1–5.