Aspects of Molecular Medicine 5 (2025) 100084

Contents lists available at ScienceDirect

Aspects of Molecular Medicine

journal homepage: www.journals.elsevier.com/aspects-of-molecular-medicine

In silico identification of potential inhibitors for the universal stress G4LZI3 protein from *Schistosoma mansoni* using molecular docking and molecular dynamics simulation analyses

Lihle Mahamba, Mustafa Alhaji Isa, Abidemi Paul Kappo *

Molecular Biophysics and Structural Biology Group, Department of Biochemistry, Faculty of Science, University of Johannesburg, Auckland Park Kingsway Campus, Johannesburg, 2006, South Africa

ARTICLE INFO

Handling Editor: Prof A Angelo Azzi

Keywords:
Molecular docking
Modelling
MD simulation
Praziquantel
S. mansoni
Universal stress G4LZI3 protein

ABSTRACT

Human schistosomiasis is a debilitating, neglected tropical disease affecting millions worldwide. Control efforts primarily rely on health education, improved sanitation, snail host management, and mass drug administration with Praziquantel (PZQ). PZQ has some limitations, such as its lower effectiveness against immature parasites and the potential for developing resistance. This requires the urgent need for new treatment approaches. The universal stress protein G4LZI3 helps the *Schistosoma mansoni* parasite survive when it is under stress from its host. Because of this, it emerges as a promising target for developing new drugs. Despite its biological relevance, G4LZI3 has not been previously investigated as a druggable target, highlighting a significant research gap in schistosomiasis drug discovery. To find potential inhibitors of G4LZI3, we conducted a virtual screening using the RASPD^{+} tool, which led us to select 7889 ligands from the CoCoNut database. These ligands were filtered based on physicochemical properties (Lipinski's Rule of Five, Veber's Rule, Egan's Filter, and the Ghose filter), pharmacokinetics, and Pan-Assay Interference Structures (PAINS) criteria, followed by molecular docking. Fifteen compounds demonstrated strong binding affinities, with binding energies ranging from −10.6 to −8.50 kcal/mol, exceeding that of PZQ (−8.4 kcal/mol). From these, six compounds were selected for further analysis, including molecular dynamics (MD) simulation, solvent-accessible surface area (SASA), and molecular mechanics generalized Born surface area (MM-GBSA) calculations. MD simulation of 200 ns revealed that CNP0475438, CNP0415153, and CNP0353858 achieved significant stability and favourable interactions with G4LZI3. These findings show these compounds as promising candidates for *S. mansoni* inhibition, pending experimental validation. The results identify novel scaffolds with vigorous predicted activity and provide a rational starting point for experimental optimization and development of new antiparasitic therapies that address praziquantel resistance and efficacy limitations in endemic regions.

1. Introduction

Human schistosomiasis is a debilitating parasitic disease that remains one of the most significant neglected tropical diseases, affecting over 261 million individuals worldwide, with most cases reported in Africa, South America, and Asia (Pirzaman et al., 2024). It ranks second only to malaria in prevalence and is the second most neglected tropical disease in sub-Saharan Africa, following hookworm infection (Pirzaman et al., 2024; Gryseels et al., 2006). Trematode worms of the genus *Schistosoma mansoni*, *Schistosoma japonicum*, and *Schistosoma haematobium* are the primary species responsible for human infections causing the disease. *S. mansoni* causes intestinal schistosomiasis and accounts for a significant disease burden, particularly in sub-Saharan Africa, the Middle East, and South America. The lifecycle of *Schistosoma* involves a complex alternation between two hosts: a definitive human host, where the parasite undergoes sexual reproduction, and a freshwater snail, which serves as the intermediate host for asexual reproduction (Gryseels et al., 2006; Colley et al., 2014). Schistosomiasis pathology results from immune responses to eggs trapped in host tissues, where antigens released by the eggs trigger a granulomatous reaction involving T-cells,

This article is part of a special issue entitled: Science in Africa published in Aspects of Molecular Medicine.

* Corresponding author. Molecular Biophysics and Structural Biology (MBSB) Group C2 Lab Room 426, Department of Biochemistry University of Johannesburg Auckland Park Kingsway Campus Johannesburg, 2006, South Africa.
E-mail address: akappo@uj.ac.za (A.P. Kappo).

https://doi.org/10.1016/j.amolm.2025.100084
Received 4 March 2025; Received in revised form 13 April 2025; Accepted 23 April 2025
Available online 29 April 2025
2949-6888/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

eosinophils, and macrophages. Initially, this inflammatory response is reversible; however, chronic infection leads to collagen deposition, fibrosis, irreversible organ damage, and disability (Elbaz and Esmat, 2013; Masamba et al., 2016).

Efforts to control schistosomiasis have primarily relied on health education, improved sanitation, biological control measures targeting the snail hosts, and mass drug administration of praziquantel (PZQ), the cornerstone of treatment for all forms of the disease. However, PZQ presents several significant limitations, including its inability to target all developmental stages of the parasite, lack of efficacy against immature worms, and emerging reports of drug resistance in endemic regions despite its widespread use (Pirzaman et al., 2024; Colley et al., 2014). These limitations are further exacerbated by the expansion of the disease into previously non-endemic regions, which is driven by human activities such as dam construction, irrigation projects, and the effects of global warming, which alter ecosystems and facilitate transmission (Masamba et al., 2016). Some developing countries like Japan, Algeria, and Morocco have successfully eradicated schistosomiasis through well-organized control programs; however, the disease continues to pose a significant threat by causing public health challenges in many parts of the world. This widespread disease requires an urgent need for innovative treatments and alternative approaches to effectively manage and combat the disease (Masamba et al., 2016).

In the search for new therapeutic targets, universal stress proteins (USPs) have emerged as promising candidates. First identified in *Escherichia coli*, USPs have since been discovered in various organisms, including plants, archaea, bacteria, and metazoans. These proteins play crucial roles in cellular responses to stress conditions, such as oxidative damage, DNA damage, nutrient deprivation, elevated temperatures, and acidic environments (Masamba and Kappo, 2021). The USPs help survive under adverse conditions in parasites like *Schistosoma mansoni*, which enable the parasite to thrive within the host. One such protein is the G4LZI3 USP, which has been implicated in the resilience and adaptability of *S. mansoni* (Adenowo et al., 2021). The growing understanding of its functional significance in stress responses presents an opportunity to target this protein for therapeutic intervention. This study presses the need for novel therapeutic strategies by focusing on the G4LZI3 USP of S. mansoni through *in silico* approaches such as molecular docking and molecular dynamics simulations. This research aims to identify potential inhibitors of the G4LZI3 protein and propose candidates for further experimental validation. Using computational tools, this study offers a cost-effective and efficient pathway to drug discovery to advance the development of targeted therapies against schistosomiasis.

2. Materials and methods

2.1. Molecular modelling

The three-dimensional structure of the G4LZI3 USP from *S. mansoni* was predicted using AlphaFold 3.0 (Abramson et al., 2024), **accessed via Galaxy Australia. Galaxy's AlphaFold 3 implementation was run with the following default settings: use of AlphaFold v3.0 wt, model relaxation enabled, and generation of five ranked models with maximum sequence length capped at 1400 residues**. Five models were generated, and the relaxed model with the highest confidence score was selected for further validation. The stability and quality of the structure chosen were assessed through analyses of its stereochemistry and non-bonded interactions. The PROCHECK tool (Zhang et al., 2005) evaluated stereochemical features, providing detailed information about bond angles, torsion angles (Φ and ψ), dihedral angles, and atom-to-atom distances. Also, ERRAT (Colovos and Yeates, 1993) was used to analyze the statistical patterns of non-bonded interactions between different atom types. This analysis indicated the potential discrepancies by comparing the model's interactions to those in high-resolution experimental structures. The compatibility between the 3D structure and its amino acid sequence was assessed using Verify_3D (Lüthy et al., 1992). This tool evaluates the environment of each residue in the model by comparing it to known protein structures with similar residue environments, categorized as α-helices, β-sheets, loops, and polar and nonpolar regions. We also used Verify_3D to check if the predicted structure aligns with sequence-based expectations and meets established structural standards.

2.2. Virtual screening

We conducted a virtual screening of small molecules from the COCONUT (Collection of Open Natural Products) database to identify potential inhibitors of the G4LZI3 USP in *S. mansoni*. This protein is associated with the parasite's resilience and adaptability under stress conditions, which are critical for its survival and infectivity. Given the growing recognition of its role in stress response mechanisms, G4LZI3 USP represents a promising therapeutic target for anti-schistosomiasis interventions (Masamba et al., 2016). The COCONUT database, containing over 600,000 natural compounds, was used for this study using the RASPD^{+} tool (Holderbach et al., 2020). This tool efficiently predicts binding affinities by assessing the interaction between potential drug-like compounds and a target protein (He et al., 2012). For this screening, the 3D structure of G4LZI3 USP, including its ligand-binding site, was used as the input to model interactions with the database compounds. Through virtual screening, 7889 compounds were identified with binding free energies ranging from −12.01 to −9.35 kcal/mol, indicating potential strong interactions with G4LZI3 USP (Supplementary Material 1). These compounds were subsequently used for further study to determine the compounds with suitable drug-like properties.

2.3. Physicochemical analysis

The 7889 compounds identified through virtual screening were further analyzed to assess their physicochemical properties and determine their potential as Pharmaceutical Active Ingredients (PAIs). This analysis used established drug-likeness properties and computational tools to select compounds with favourable characteristics. Properties include Lipinski's Rule of Five, which focuses on molecular weight, the logarithm of the partition coefficient (logP), and the number of hydrogen bond donors (HBD) and acceptors (HBA) as predictors of oral bioavailability (Lipinski et al., 1997). Veber's Rule was then used to evaluate molecular flexibility and polarity by analyzing the number of rotatable bonds and the Topological Polar Surface Area (TPSA) (Veber et al., 2002). Egan's Pharmacia filter, which combines logP and TPSA, was employed to assess further bioavailability potential (Egan et al., 2000). Finally, the Ghose filter considered molecular weight, log P, molecular refractivity, and atom counts to purify the selection of drug-like candidates (Ghose et al., 1999). All these criteria were determined using SwissADME, a computational tool for drug-likeness and pharmacokinetic analyses (Daina et al., 2017). The compounds with desirable physicochemical properties were selected for further study.

2.4. Pan-assay interference structure (PAINS) and pharmacokinetics analyses

The compounds that exhibited favourable physicochemical properties were subjected to pharmacokinetics and the Pan-assay interference structure (PAINS) analysis, an analysis identifying possible toxicophores, which assessed the compounds' toxicity risks. Toxicophores refer to structural elements in molecules that can interact with biological processes. This potentially causes damage by interfering with DNA or proteins. This interference can lead to severe health conditions, which include carcinogenicity and hepatotoxicity (Baell and Holloway, 2010). In addition to PAINS analysis, the pharmacokinetic properties of the compounds, including absorption, distribution, metabolism, excretion,

and toxicity (collectively known as ADMET), were analyzed using SwissADME (Daina et al., 2017). The specific pharmacokinetic parameters assessed included gastrointestinal absorption (GIA), blood-brain barrier (BBB) penetration, the ability to act as a *P*-glycoprotein (Pgp) substrate, and inhibition potential for crucial cytochrome P450 enzymes (CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4). These properties are essential in determining how the compound is absorbed, distributed, and metabolized in the body, directly affecting its pharmacological activity and therapeutic efficacy. Also, the synthetic accessibility of each compound was evaluated to determine whether they could be feasibly synthesized in a laboratory setting. Following these analyses, only the compounds that did not trigger PAINS alerts exhibited desirable pharmacokinetic properties and demonstrated the ease of synthesis were selected for further experimental investigations.

2.5. Molecular docking analysis

The docking analysis was performed to determine the free binding energy of the G4LZI3 USP-ligand complex using AutoDock 4.2 (Morris et al., 1998). This technique evaluates how the ligands interact with the target protein (G4LZI3 USP) by providing their binding energy values. Praziquantel, a well-known drug for treating schistosomiasis, was considered a standard ligand and was docked to G4LZI3 USP for comparative analysis of binding energies between this standard and the selected ligands. The docking simulations were guided by a pre-defined active site corresponding to the UspA domain of G4LZI3, which spans residues 29–176. Key residues such as Met33, Arg43, Tyr60, Ser80, Asn92, Thr107, and Leu114 were selected based on domain annotation and structural predictions using AlphaFold. These residues represent a conserved and accessible cavity and were used as the target grid center during virtual screening and docking. The UspA domain is critical in stress response mechanisms and protein stabilization in *S. mansoni*. The binding affinities were computed using a Lamarckian genetic algorithm, an efficient method for optimizing the docking process. The G4LZI3 USP protein was protonated with polar hydrogen atoms while fixed Kollman charges were applied to the structure during the docking process. The PDBQT provided essential details about the G4LZI3 protein, which include partial charges, atom types, and torsional degrees of freedom. The ligands' side chains and torsional bonds were kept flexible, while the G4LZI3 USP structure was rigid. A grid box of 60 Å × 60 Å × 60 Å was set to encompass the protein's binding site, with a spacing of 0.375 Å. A total of 10 independent runs, a maximum of 27,000 generations, 2.5 million evaluations, and a population size of 150 were performed during the docking process. The free binding energy (ΔG_{bind}) was calculated by considering several energy components: van der Waals energy (ΔGvdw), electrostatic energy (ΔG_{elect}), hydrogen bond and desolvation energy (ΔG_{hbond}), total internal energy ($\Delta G_{conform}$), torsional free energy (ΔG_{tor}), and the energy of the unbound system (ΔG_{solv}). Finally, the protein-ligand complexes were analyzed to determine interactions such as hydrogen bonds, hydrophobic, and Van der Waals interactions.

2.6. Molecular dynamics (MD) simulation analysis

Molecular dynamics (MD) simulations were conducted to evaluate the stability of the G4LZI3-ligand complex using the Amber18 software suite (Case et al., 2023). The simulation began by preparing the system, where explicit hydrogen atoms were added to the protein-ligand complex using the Protonate 3D tool. Missing ligand parameters were generated with Antechamber. The topology and coordinate files for the complex were created using **tleap**, which also provided molecular graphics of the system's topology. The **GAFF force field** was applied to the ligand, while the protein was parameterized using the **ff12SB force field**. To simulate the system in a realistic environment, the complex was immersed in a TIP3P water model, neutralized with chloride ions, and placed within an octahedral box with a 10 Å buffer. Structural artifacts were eliminated through energy minimization in two stages. First, a minimization cycle of 10,000 steps of conjugate gradient and steepest descent methods was performed with a 544 kcal/mol/Å restraint applied to the protein-ligand complex. The restraint was removed, and another 5000 steps of steepest descent, followed by 5000 steps of conjugate gradient minimization, were executed to optimize the system further. The system was gradually heated from 0 K to 300 K using Langevin dynamics for temperature regulation, with a collision frequency of 1 ps. Heating was performed without pressure control. For production, the simulation was run at a constant temperature of 300 K and a pressure of 1 atm, maintained using the Berendsen barostat. The simulation used a time step of 2 fs and included SHAKE constraints to fix all hydrogen bond lengths. A 200-ns MD simulation was conducted to assess the dynamic stability of the complex. **This timescale was selected as it has been demonstrated in previous studies to be sufficient for capturing meaningful protein-ligand interaction patterns, conformational stability, and energetic convergence in medium-sized systems (**Isa, 2019; Genheden and Ryde, 2015**).**

2.7. Post-MD simulation analysis

The molecular dynamics (MD) simulation trajectories were analyzed using the CPPTRAJ tool within the Amber18 package (Case et al., 2023) to evaluate metrics essential for understanding the protein-ligand complex's stability and behaviour. **Root Mean Square Deviation (RMSD)** was calculated to determine the structural deviations from the initial configuration, which measures the system's stability throughout the 200 ns simulation. **Root Mean Square Fluctuation (RMSF)** was employed to explore dynamic behaviour further to assess individual residues' flexibility, showing regions with significant conformational changes. Also, the **Radius of Gyration (Rg)** was analyzed to determine the compactness of the complex based on its folding and structural organization. Solvent Accessible Surface Area (SASA) analysis, Clustering Analysis, and Comparative Analysis of Energy Fluctuations were performed to identify dominant motions and significant conformational changes by reducing the trajectory's dimensionality.

2.8. Molecular mechanics generalized born surface area (MMGBSA) analysis

The **MMGBSA** technique is widely recognized for estimating the free binding energy between a ligand and its target protein. This method combines principles from molecular dynamics (MD) simulations and thermodynamic calculations to balance computational efficiency and accuracy. MMGBSA has been increasingly adopted due to its reliability in predicting binding free energies compared to traditional empirical scoring functions (Genheden and Ryde, 2015). In this study, the MMGBSA method, as implemented in Amber18, was employed to compute the free binding energy of the G4LZI3 protein-ligand complexes. The analysis was based on MD simulation trajectories, which extracted 5000 snapshots at 100 ps intervals over the 150–200 ns simulation period. This method allowed for calculating an average binding energy that captures the dynamic nature of the protein-ligand interactions, which provide the strength and stability of the binding under simulated conditions.

3. Results and discussion

3.1. Homology modelling and model validation

The G4LZI3 USP from *S. mansoni* was successfully modelled using the AlphaFold 3.0 server (Fig. 1a). This method uses deep learning techniques to accurately predict protein structures, even for proteins with limited experimental data (Jumper et al., 2021). The model's reliability and quality were evaluated using standard structure validation tools such as PROCHECK, ERRAT, and Verify_3D. The PROCHECK analysis revealed that 92.4 % of residues were in the most favourite regions of the

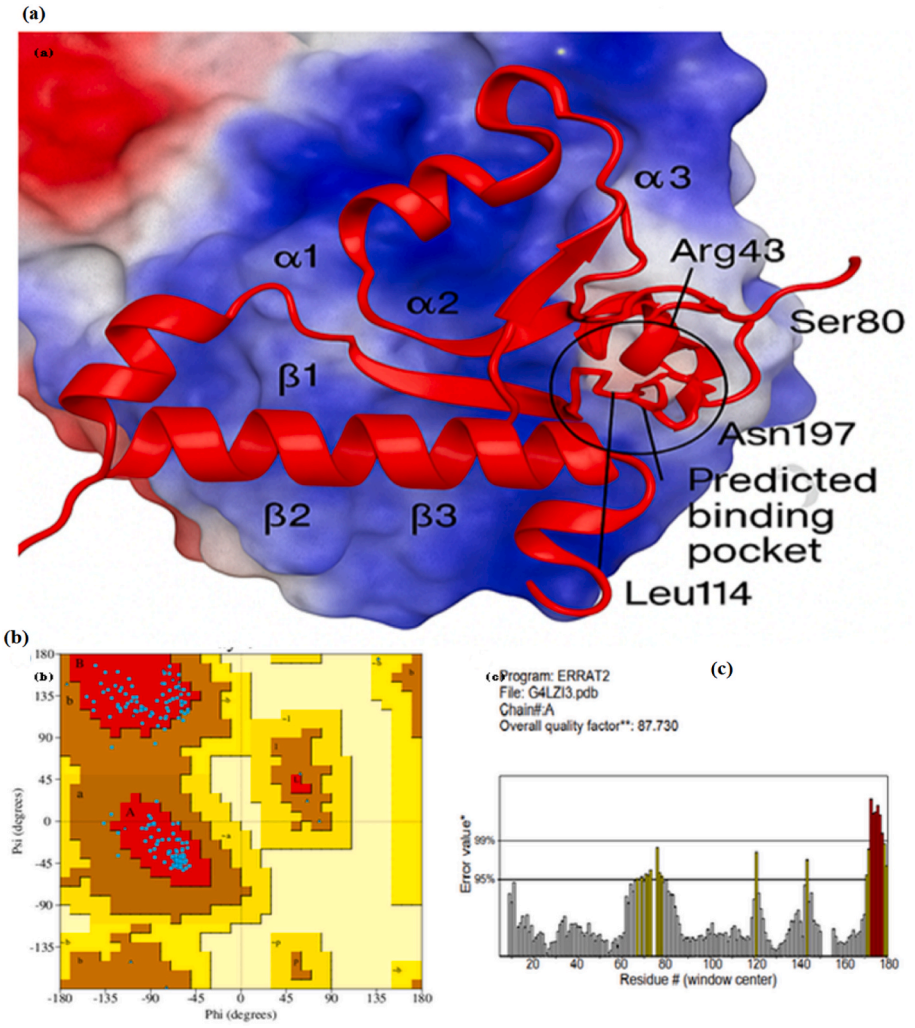

Fig. 1. Structural prediction and validation of *S. mansoni* Universal stress G4LZI3 protein. **(a)** Predicted three-dimensional structure of the G4LZI3 protein, generated using AlphaFold 3.0 via Galaxy Australia and visualized with PyMOL. The ribbon model is overlaid on the electrostatic surface potential map (red = negative charge; blue = positive charge) to highlight the physicochemical environment of the UspA domain. Key secondary structural elements [α-helices (α1–α3) and β-sheets (β1–β3)] as well as functionally relevant residues (Arg43, Tyr60, Ser80, Asn197, and Leu114) are annotated around the predicted ligand-binding pocket. The model achieved a predicted Template Modelling (pTM) score of 0.83, indicating high global structural accuracy (Abramson et al., 2024). **(b)** Ramachandran plot of the predicted structure showing the distribution of amino acid residues' backbone dihedral angles (Φ, Ψ). Most residues are located in the most favoured and additionally allowed regions, confirming acceptable stereochemistry. **(c)** ERRAT quality assessment plot displaying non-bonded atomic interaction reliability across the protein chain. The overall quality factor of 87.73 supports the model's suitability for downstream molecular docking and simulation workflows. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Ramachandran plot, with 7.6 % in additional allowed regions and 0 % in generously allowed or disallowed regions (Fig. 1b). These results indicate well-modelled backbone geometry, showing minimal steric clashes and energetically favourable torsional angles. The high percentage of residues in favoured regions is consistent with high-quality protein models, often correlating with accurate functional predictions. Further evaluation using ERRAT, which measures nonbonded atomic interactions, produced a quality factor of 87.73 % (Fig. 1c). This value exceeds the recommended threshold of 50 %, which indicates the structure's reliability for computational and experimental applications (Colovos and Yeates, 1993). Also, Verify_3D analysis revealed that 74.5 % of residues scored favourably in their environment, validating that the 3D structure aligns well with the protein's sequence-based features (Eisenberg et al., 1997). The G-factor analysis from PROCHECK assessed torsional angles, covalent geometry, and overall geometry, yielding values of 0.01, 0.36, and 0.10, respectively. These scores fall within acceptable ranges, affirming that the structural geometry is consistent with high-resolution experimental models. In addition, the predicted structure had a pTM (predicted Template Modelling) score of 0.83, as reported by AlphaFold, which suggests high global accuracy and justifies its use in downstream drug discovery applications. pTM scores above 0.7 typically reflect reliable backbone conformations suitable for virtual screening. Minor deviations in these parameters, while within tolerable limits, could be attributed to inherent challenges in predicting flexible regions or underrepresented motifs in the training dataset used by AlphaFold. The results collectively affirm the structural reliability of the G4LZI3 USP model, which makes it a strong candidate for further *in silico* and experimental investigations. The accurate prediction of its structure enables detailed studies into its enzymatic mechanisms, substrate binding, and interaction with potential inhibitors. This is particularly relevant for *S. mansoni*, as targeting USPs has emerged as a promising therapeutic target in parasitic diseases due to their role in proteostasis and immune evasion (Clague et al., 2019).

3.2. *Virtual screening, physicochemical and pharmacokinetics analyses*

The modelled structure of G4LZI3 USP was used for the virtual screening using the COCONUT database, which hosts over 600,000

natural compounds and provides a diverse library of bioactive molecules for potential therapeutic discovery. Virtual screening identified 7889 compounds with binding free energies ranging from −12.01 to −9.35 kcal/mol, which indicates strong interaction potential with the G4LZI3 USP (Supplementary Material 1). Binding free energy is a critical metric as it correlates with the strength and stability of the ligand and protein interaction (Isa et al., 2021). Lower binding free energy values suggest a high likelihood of inhibitory efficacy, particularly for a target as structurally conserved and functionally crucial as G4LZI3 (Isa et al., 2021). The screening process used the RASPD^{+} tool, known for its efficiency in evaluating large chemical libraries for potential hits. This step is particularly significant as it enables prioritization of compounds with a higher probability of biological relevance and efficacy. The compounds' drug-likeness and pharmacological potential were analyzed for their physicochemical properties against Lipinski's Rule of Five, Veber's Rule, Egan's Pharmacia filter, and the Ghose filter. These filters assess critical parameters, such as molecular weight, hydrogen bond donors and acceptors, lipophilicity (log P), and polar surface area (PSA), which collectively predict a compound's oral bioavailability and drug-likeness (Lipinski et al., 1997; Veber et al., 2002). Among the 7889 initial hits, only 2764 compounds met these criteria, which reflect their potential to be developed as orally bioavailable drugs (Supplementary Material 2). Excluding compounds that failed these filters and focused on those with a higher likelihood of successful pharmacokinetics, the remaining 2764 compounds were subjected to pharmacokinetics analysis and Pan-Assay Interference Structures (PAINS) evaluation.

Pharmacokinetics analysis encompasses absorption, distribution, metabolism, excretion, and toxicity (ADMET), which is crucial in predicting a compound's behaviour *in vivo*. Key parameters, such as gastrointestinal absorption, blood-brain barrier permeability, and cytochrome P450 interactions, were considered to identify the safety and efficacy of the candidates. PAINS analysis was employed to identify compounds that might produce false-positive results in bioassays due to their structural motifs interfering with biological targets non-specifically. This analysis, often termed "toxicophores," can lead to adverse effects, including carcinogenicity or hepatotoxicity (Baell and Holloway, 2010). This step provided compounds with minimal risk and was advanced for further studies. Ultimately, 40 compounds emerged as strong candidates for molecular docking analysis (Supplementary Material 3). The identified compounds hold significant promise as potential leads for developing innovative inhibitors targeting G4LZI3 USP, which is crucial for the survival of *S. mansoni*. These inhibitors could be vital in enhancing treatment options for schistosomiasis, a neglected tropical disease that impacts millions of people globally.

3.3. Molecular docking analysis

The molecular docking analysis of G4LZI3 USP provides an in-depth exploration of potential inhibitors that target the UspA domain, a critical area for the survival and adaptation of *S. mansoni*. This domain involves essential functions such as stress response and protein stabilization, which are crucial for the parasite's ability to withstand host immune attacks and environmental pressures (Masamba et al., 2016). Therefore, the disruption of these interactions by small molecules could significantly impair the survival mechanisms of the parasite, which suggests a promising direction for a therapeutic approach. Among the ligands investigated, 15 compounds were identified with binding energies ranging from −10.6 to −8.50 kcal/mol, which is less than the binding energy of praziquantel (−8.4 kcal/mol) (Fig. 2), the current standard treatment for schistosomiasis. These binding energies suggest stronger and potentially stable interactions with the G4LZI3 USP protein. These compounds possessed desirable physicochemical (Table 1), pharmacokinetics, and PAINS assay properties (Table 2). Also, a detailed residue-level analysis identified several amino acids playing significant roles in ligand binding (Table 3). The docking poses of the top binders were carefully examined and visualized using LigPlot^{+} (Laskowski and Swindells, 2011) and PyMOL to confirm the quality of protein-ligand interactions. Key hydrogen bonds and hydrophobic contacts were consistently formed with residues in the conserved UspA binding pocket, supporting the structural relevance of the predicted docking conformations. Glycine residues, such as Gly145 and Gly150, were found to contribute substantially via hydrogen bonding. Glycine's small size and flexibility facilitate interactions with various ligand functional groups, accommodating diverse chemical scaffolds (Table 3). CNP0475438 exhibits the most vital binding energy of −10.6 kcal/mol, formed hydrogen bonds with Gly150 and Ser159, and engaged in hydrophobic interactions with residues like Pro34 and Ile64, demonstrating the importance of combining polar and nonpolar interactions for strong binding (Fig. 3e).

Other high-affinity ligands, such as CNP0353858 (Fig. 3a) and CNP0415153 (Fig. 3i), Indicated the cooperative role of residues like Val160, Ile64, and Arg147, which contributed both hydrophobic contacts and hydrogen bonding to stabilize the ligand-protein complex. Polar residues, such as Ser159 and Thr155, often serve as key players in

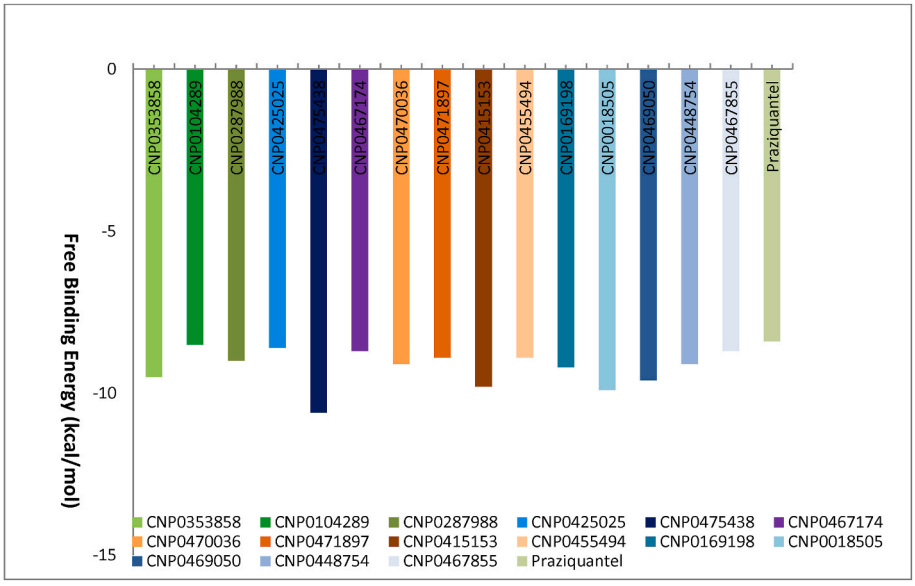

Fig. 2. Free binding energies of the selected ligands.

Table 1
Physicochemical analysis of the selected ligands with desirable properties.

S/No.	Compound ID	Molecular Weight	H Bond Donors	H Bond Acceptors	LogP	Rotatable Bonds	TPSA	Lipinski	Ghose	Veber	Egan
2	CNP0353858	416.47	2	7	2.145	6	110.13	TRUE	TRUE	TRUE	TRUE
3	CNP0104289	402.44	3	7	3.05	9	113.29	TRUE	TRUE	TRUE	TRUE
5	CNP0287988	422.61	2	4	2.97	6	78.87	TRUE	TRUE	TRUE	TRUE
7	CNP0425025	340.50	3	3	3.84	5	77.76	TRUE	TRUE	TRUE	TRUE
13	CNP0475438	372.42	4	4	3.56	6	115.06	TRUE	TRUE	TRUE	TRUE
20	CNP0467174	426.27	3	4	3.03	5	89.19	TRUE	TRUE	TRUE	TRUE
23	CNP0470036	397.46	3	5	2.13	6	106.26	TRUE	TRUE	TRUE	TRUE
26	CNP0471897	377.40	3	5	2.28	6	98.42	TRUE	TRUE	TRUE	TRUE
30	CNP0415153	411.80	2	7	2.42	3	118.19	TRUE	TRUE	TRUE	TRUE
31	CNP0455494	361.40	3	4	2.58	5	89.19	TRUE	TRUE	TRUE	TRUE
33	CNP0169198	372.38	4	5	1.43	8	124.96	TRUE	TRUE	TRUE	TRUE
34	CNP0018505	420.55	0	6	4.06	4	78.90	TRUE	TRUE	TRUE	TRUE
36	CNP0469050	375.38	2	5	3.20	6	91.10	TRUE	TRUE	TRUE	TRUE
37	CNP0448754	407.40	2	5	3.65	6	91.10	TRUE	TRUE	TRUE	TRUE
38	CNP0467855	423.40	2	6	3.35	7	100.33	TRUE	TRUE	TRUE	TRUE

Table 2
Pharmacokinetics analysis of the selected ligands with desirable properties.

S/ No.	Compound ID	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	PAINS	Synthetic Accessibility
2	CNP0353858	High	No	No	No	No	No	No	No	0	5.89
3	CNP0104289	High	No	No	No	No	No	No	No	0	3.58
5	CNP0287988	High	No	No	No	No	No	No	No	0	4.71
7	CNP0425025	High	No	No	No	No	No	No	No	0	4.17
13	CNP0475438	High	No	No	No	No	No	No	No	0	2.52
20	CNP0467174	High	No	No	No	No	No	No	No	0	3.02
23	CNP0470036	High	No	No	No	No	No	No	No	0	3.18
26	CNP0471897	High	No	No	No	No	No	No	No	0	2.89
30	CNP0415153	High	No	No	No	No	No	No	No	0	3.3
31	CNP0455494	High	No	No	No	No	No	No	No	0	2.94
33	CNP0169198	High	No	No	No	No	No	No	No	0	3.12
34	CNP0018505	High	No	No	No	No	No	No	No	0	5.11
36	CNP0469050	High	No	No	No	No	No	No	No	0	3.23
37	CNP0448754	High	No	No	No	No	No	No	No	0	3.33
38	CNP0467855	High	No	No	No	No	No	No	No	0	3.36

forming hydrogen bonds and act as either donors or acceptors. In contrast, hydrophobic residues like Pro34 and Ile149 contribute to a favourable binding environment through nonpolar interactions. This combination of interactions indicated the complex nature of ligand-protein binding, where a synergy of electrostatic forces, hydrophobic interactions, and hydrogen bonding collectively enhances binding affinity (Isa et al., 2021).

In the case of praziquantel's interaction with the G4LZI3 Ubiquitin-specific protease (USP), the engagement is less comprehensive, relying primarily on Gly148 for hydrogen bonding alongside Pro34 and Val160 for hydrophobic stabilization. This restricted interaction profile may explain the comparatively higher binding energy observed, suggesting potential avenues for designing new ligands that could influence additional binding opportunities within the UspA domain. Furthermore, the functionality of the ligands is crucial in shaping their interaction patterns. Hydroxyl and amino groups enabled strong hydrogen bonding with polar and charged residues, while aromatic and aliphatic groups facilitated hydrophobic contacts with residues such as Ile144 and Pro126. This shows the importance of chemical diversity in ligand design, with functional groups tailored to maximize interaction with critical residues in the binding pocket. Additionally, charged residues like Arg147 and Glu66 provided ionic interactions, further stabilizing some ligand-protein complexes (Table 3). The findings show the therapeutic potential of targeting the UspA domain with natural product-derived ligands. These compounds' strong binding affinities and diverse interaction patterns indicate their ability to disrupt critical protein functions in *S. mansoni*.

3.4. Molecular dynamics simulation analysis

From the molecular docking studies, 15 compounds were identified with binding energies ranging from −10.6 to −8.5 kcal/mol, all surpassing the binding energy of praziquantel (−8.4 kcal/mol), the standard drug. These values indicate that the identified compounds exhibit a stronger and more favourable binding affinity to the UspA domain of G4LZI3 USP. This is significant as binding energy is directly related to the strength of the interaction between the ligand and the target protein, which is a critical factor for drug efficacy (Isa, 2019). Out of these, six compounds with the most favourable binding energies were chosen for further evaluation through molecular dynamics (MD) simulations and praziquantel as a reference to assess their dynamic stability and interaction behaviour under physiological-like conditions.

The root mean square deviation (RMSD) values were used to monitor the overall stability of the protein-ligand complexes during the 200 ns simulation period. RMSD analysis showed that CNP0415153 had the lowest mean RMSD value of 3.17 Å, followed by CNP0469050 at 3.45 Å, indicating that these two ligands maintained a stable interaction with the UspA domain (Fig. 4). In contrast, praziquantel exhibited a higher mean RMSD of 4.70 Å, reflecting more significant structural fluctuations and comparatively weaker stability. RMSD maxima further indicated this trend, with CNP0415153 reaching a maximum of 4.72 Å, significantly lower than praziquantel's peak RMSD of 5.50 Å. This suggests CNP0415153 remains more tightly bound to the target protein throughout the simulation (Fig. 4).

Residue flexibility, as determined by root mean square fluctuation (RMSF), provided an understanding of the dynamic behaviour of individual amino acids within the protein-ligand complexes. For all the

Table 3
Free binding energies and various interactions between the G43LZI3 and the selected ligands.

S/ No.	Compound ID	Free Binding Energy (kcal/mol)	Hydrogen Bond	Distance (Å)	Hydrophobic Interaction	van der Waals Interactions
1	CNP0353858	−9.5	Ile64 Gly148 Gly148 Gly150 Gly150 Gly150 Val152	3.17 3.03 3.06 3.07 3.11 3.12 2.91	Pro34, Pro126, Arg147 and Val160	Val35, Asp36, His63, Glu66, Ile130, Ile144, Gly148, Ile149, Thr151, Thr155, Gly158 and Ser159
2	CNP0104289	−8.5	Gly145 Arg147 Ile149 Gly150 Val152	3.18 2.91 3.11 3.07 2.94	Pro34, Asp36, Ile144, Arg147 and Val160	Val35, Ser38, His40, Ser41, Ile64, Glu66, Pro126, Ile130, Gly148, Val152, Thr155, Gly158, Ser159 and Ser161
3	CNP0287988	−9.0	Pro7 Leu13	2.98 3.31	Phe87, Leu91, Leu94, Val95 and Ile120	Ser8, Thr9, Asp11, Gly12, Tyr74, Ala77, Ser78, Lys84, His121 and Val122
4	CNP0425025	−8.6	Gly145 Arg147 Gly150	2.98 2.93 2.88	Pro34, Ser41, Pro126, Ile144, Ile149 and Val160	Val35, Asp36, His63, Ile64, Ile130, Gly145, Gly148 and Ser161
5	CNP0475438	−10.6	Gly150 Ser159 Val160	3.23 3.13 2.80	Pro34, Ile64, Pro126 and Ile130	Val35, Ser41, Asp36, Glu66, Ile144, Arg147, Gly148, Ile149, Gly158 and Ser161
6	CNP0467174	−8.7	Thr155 Thr155 Ser159 Ser159 Val160	3.02 3.21 2.81 3.00 3.04	Pro34, Asp36, Ile64, Pro126, Ile130, Ile144 and Ile149	Val35, Val62, His63, Glu66, Arg147, Gly148, Gly150, Val152 and Gly158
7	CNP0470036	−9.1	Ile64 Gly148	3.10 3.31	Pro34, Asp36, Ile144, Ile149 and Val160	Val35,Ser41, Val62, His63, Val65, Glu66, Pro126, Ile130, Gly145, Arg147, Gly148, Gly150 and Ser161
8	CNP0471897	−8.9	His63 Gly148	2.91 3.09	Asp36 Arg147, Val152, Ser159 and Val160	Pro34, Val35, His63, Ile64, Glu66, Pro126, Ile130, Ile144, Gly148, Ile149, Gly150 and Thr155, Gly158 and Ser161
9	CNP0415153	−9.8	Gly148 Gly150 Ser159	3.10 2.82 3.13	Pro34, Asp36, Ile64, Glu66, Pro126, Gly148, Arg147, Val152 and Val160	Val35, Val62, His63, Ile130, Ile149, Gly150, Thr155, Leu157 and Gly158
10	CNP0455494	−8.9	His63 Gly148	2.97 3.08	Asp36, His63, Arg147, Val152, Ser159 and Val160	Pro34, Val35, Ile64, Glu66, Pro126, Ile130, Ile144, Gly148, Ile149, Gly150, Thr155, Gly158 and Ser161
11	CNP0169198	−9.2	Val62 Ile64 Ser159	2.83 3.20 2.81	Asp36, Ile64, Pro126, Arg147, Val152 and Val160	Pro34, Val35, Val62, Glu66, His63, Ile130, Gly148, Gly150, Thr155, Leu157 and Gly158
12	CNP0018505	−9.9	Gly148 Ser159 Ser159	3.34 3.12 3.21	Pro34, Ile64, Pro126 and Val160	Val35, Asp36, Glu66, Ile130, Ile144, Arg147, Gly148, Ile149, Gly150 and Gly158
13	CNP0469050	−9.6	Gly145 Arg147 Ser159 Val160	2.84 2.67 3.10 3.06	Asp36, Ile149 and Val152	Pro34, Val35, Ser38, His40, Ser41, Ile64, Pro126, Ile130, Ile144, Gly148, Gly150, Gly158 and Ser161
14	CNP0448754	−9.1	Gly145 Gly145 Ser161	2.76 3.26 2.87	Asp36, Arg147, Ile149, Val152 and Val160	Pro34, Ser41, His63, Ile64, Pro126, Ile130, Ile144, Gly148, Gly150, Gly158 and Ser159
15	CNP0467855	−8.7	Gly145 Gly145 Ser161	2.78 3.24 2.82	Asp36, His63, Ile149 and Val160	Pro34, Ser41, Ile64, Glu66, Pro126, Ile130, Ile144, Asn146, Arg147, Gly148, Gly150, Val152, Gly158, Ser159 and Ser161
16	Praziquantel	−8.4	Gly148	3.01	Pro34, Asp36, Ile64, Ile149 and Val160	Val35, Glu66, Pro126, Val62, His63, Glu66, Pro126, Ile130, Ile144, Arg147, Gly148, Gly150 and Gly158

ligands, most residues exhibited RMSF values below 6 Å, except for residues at the termini and loop regions, such as Met1-Thr4 and Lys184, which inherently display higher flexibility. Interestingly, Praziquantel showed elevated RMSF values for particular vital residues, particularly Lys184, compared to CNP0415153 and CNP0469050 (Fig. 5). This indicates that praziquantel induces a more dynamic environment in these regions, potentially destabilizing the overall interaction. In contrast, the lower RMSF values observed with CNP0415153 suggest a stronger anchoring of the ligand, contributing to a more stabilized protein-ligand interface (Fig. 5).

The radius of gyration (Rg) was analyzed to assess the compactness and folding behaviour of the protein-ligand complexes. Lower Rg values indicate a more compact and stable complex, a desirable characteristic in drug design. Among the ligands, CNP0415153 showed a mean Rg of 17.97 Å, identical to praziquantel, but its lower RMSD and RMSF values still pointed to superior stability (Fig. 6). Conversely, CNP0018505 exhibited the highest mean Rg of 18.35 Å, which suggests that it may partially unfold the protein structure, making it less suitable as a stable inhibitor. These findings align with the RMSD and RMSF analyses, which consistently identified CNP0415153 and CNP0469050 as the most stable candidates. The docking results, combined with the MD simulations, reveal that the stability of the protein-ligand complexes correlates well with their binding affinities. Compounds like CNP0415153 exhibited strong binding energies (−9.8 kcal/mol) and demonstrated improved simulation stability. This stability is likely due to favourable hydrophobic interactions and a well-defined hydrogen bonding network within the binding pocket. These interactions indicate that the ligand remains securely positioned within the active site, which reduces the likelihood of displacement or instability under dynamic conditions.

The structural superimposition of the initial and final complex structures obtained from the 200 ns molecular dynamics (MD) simulations provides a quantitative measure of structural deviations, shedding light on the stability of the protein-ligand complexes over time (Isa,

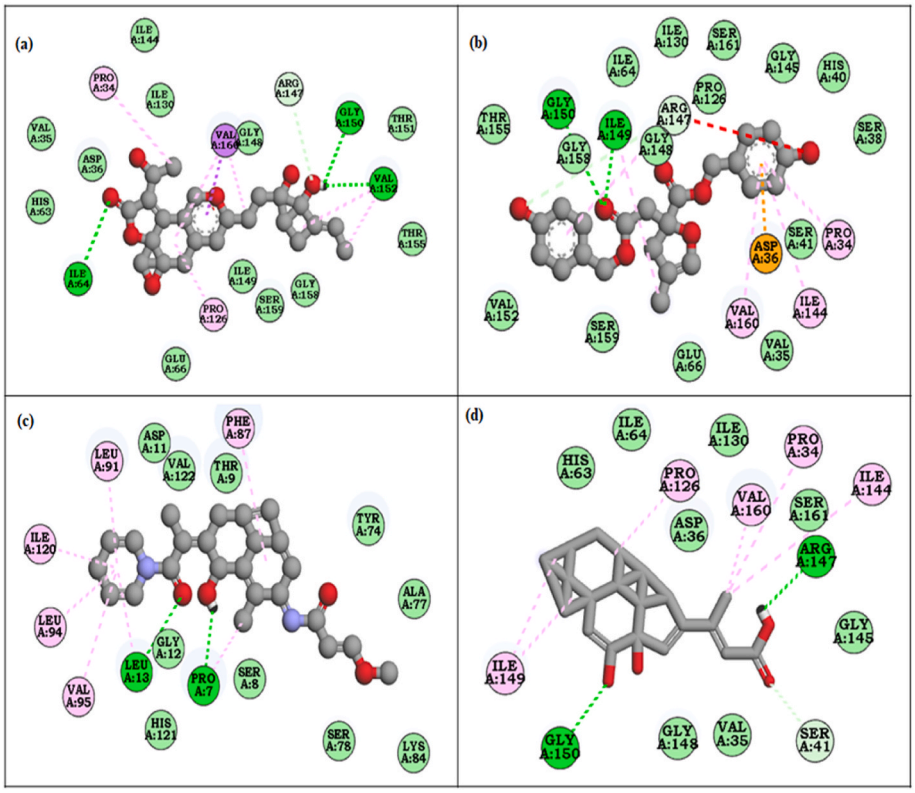

Fig. 3. Protein-ligand interaction profiles of *S. mansoni* Universal stress G4LZI3 protein with selected compounds. Panels (a) to (p) represent the docking interaction diagrams of the 15 top-ranked compounds and the reference drug Praziquantel with the Universal stress G4LZI3 protein. The panels correspond to: (a) CNP0353858, (b) CNP0104289, (c) CNP0287988, (d) CNP0425025, (e) CNP0475438, (f) CNP0467174, (g) CNP0470036, (h) CNP0471897, (i) CNP0415153, (j) CNP0455494, (k) CNP0169198, (l) CNP0018505, (m) CNP0469050, (n) CNP0448754, (o) CNP0467855, and (p) Praziquantel. The figure shows hydrogen bonds, hydrophobic contacts, and van der Waals interactions between the ligands and key residues within the UspA domain of G4LZI3. LigPlot+ (Laskowski and Swindells, 2011) was used to generate the 2D interaction diagrams.

2019). This analysis's root mean square deviation (RMSD) values reveal an important understanding when considered alongside the previous MD simulation results. Among the analyzed complexes, CNP0415153 demonstrated the lowest structural RMSD (1.746 Å), indicating minimal deviation from the initial structure (Fig. 7c). This result aligns well with the findings from the RMSD and RMSF analyses during MD simulations, where CNP0415153 exhibited the lowest mean RMSD (3.17 Å) and moderate residue flexibility. These observations strongly suggest that CNP0415153 forms a highly stable and consistent interaction with the UspA domain of G4LZI3 USP. The low structural RMSD reinforces its potential as a strong inhibitor, maintaining its conformational integrity while adapting optimally to the binding site. CNP0469050, with a structural RMSD of 1.868 Å, also showed minimal deviation, further supporting its stability as indicated by its low MD simulation mean RMSD (3.45 Å). Although CNP0469050 (Fig. 7f) was not as stable as CNP0415153 based on dynamic properties, its structural RMSD value confirms it as a promising ligand, retaining a strong and well-adapted binding pose.

The standard drug Praziquantel exhibited a structural RMSD of 1.902 Å, comparable to CNP0469050 and marginally higher than CNP0415153. This result suggests that while praziquantel maintains a relatively stable interaction, it undergoes slightly more significant conformational changes than the most stable ligands. This is consistent with its higher mean RMSD (4.70 Å) and increased residue flexibility (RMSF) in specific regions, particularly Lys184. These deviations may contribute to less efficient binding and stability under dynamic conditions. Other ligands showed moderate deviations from their initial structures, such as CNP0353858 (RMSD = 1.828 Å) (Fig. 7a) and CNP0475438 (RMSD = 1.930 Å) (Fig. 7b). Their structural RMSD values, combined with their dynamic RMSD results (mean RMSD of 3.55 Å and 5.30 Å, respectively), suggest that while they are relatively stable, they are less consistent than CNP0415153 or CNP0469050. The elevated RMSD values recorded for CNP0169198 (2.347 Å) (Fig. 7d) and CNP0018505 (2.074 Å) (Fig. 7e) suggest more pronounced structural deviations. This observation is consistent with their relatively higher dynamic RMSD and lower compactness in gyration (Rg) radius, potentially impacting their overall stability.

3.5. Solvent Accessible Surface Area (SASA) analysis

The results of the Solvent Accessible Surface Area (SASA) analysis reveal an understanding of the interactions and stability of the protein-ligand complexes formed between the USP protein and the various selected compounds, including Praziquantel. SASA provides an understanding of the extent of the surface area of the protein that is accessible to the solvent, and this can give important clues regarding the binding characteristics and structural stability of the complexes (Ausaf Ali et al., 2014). From the SASA values, the mean surface area of the G4LZI3 USP-CNP0475438 complex (11267.56 $Å^2$) is slightly higher than that of praziquantel (11001.67 $Å^2$), which indicates that CNP0475438 might have a greater exposure to solvent, possibly due to a more flexible interaction with the protein surface (Fig. 8). This could also suggest a potentially less stable interaction when compared to complexes with smaller SASA values. Similarly, the G4LZI3 USP-CNP0018505 complex shows a mean SASA of 11182.57 $Å^2$, indicating relatively high solvent accessibility. The G4LZI3 USP-CNP0353858 complex has a slightly lower mean value of 10996.77 $Å^2$ (Fig. 8), suggesting better protein-ligand packing and potentially more excellent stability. The

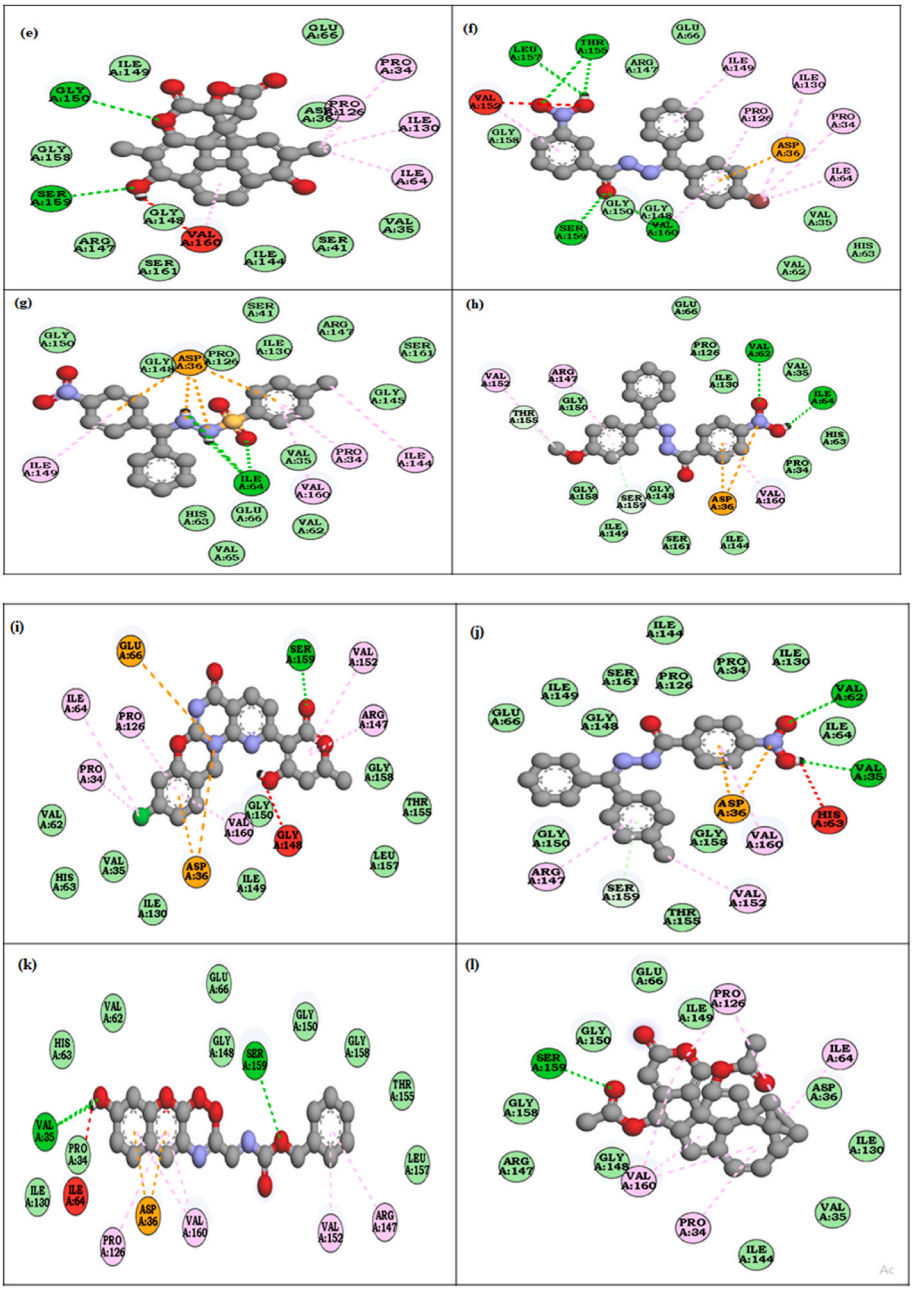

Fig. 3. (*continued*).

differences in SASA values are consistent with the findings from the Molecular Dynamics (MD) simulations. For instance, CNP0475438, with a higher SASA value, also exhibited higher RMSD (5.30 Å), which indicates a more significant fluctuation in the protein-ligand complex during the simulation. This could imply that the complex formed with this ligand is less stable compared to those with lower SASA values, such as CNP0353858 (mean SASA = 10996.77 $Å^2$, RMSD mean = 3.55 Å), which shows better structural stability during the simulation. The SASA results support the docking analysis where CNP0475438 showed a higher binding energy (−10.6 kcal/mol), suggesting a strong initial interaction. However, the MD simulation results indicated that this interaction might need to be more stable. On the other hand, compounds like CNP0353858, with lower SASA values and more favourable RMSD profiles, appear to maintain a more stable binding in the protein-ligand complex, as reflected by both the docking and MD simulation results.

3.6. Clustering Analysis

The clustering analysis of molecular dynamics (MD) simulations over 200 ns using 5000 trajectories provided a valuable understanding of the ligand-protein complexes' structural stability and dynamic behaviour (Daura et al., 1999). Across all complexes, the root mean square deviation (RMSD) ranged from 0 to 9 Å, with cluster 0 being the most frequently observed (Fig. 9). This dominance of cluster 0 indicates a preference for a central, stable conformation during the simulation, reflecting a consistent interaction between the ligands and the UspA domain of G4LZI3 USP. For the G4LZI3 USP ~ CNP0353858 complex, cluster 0 appeared 975 times, indicating moderate stability, as the significant presence of other clusters suggests occasional transitions between conformations (Fig. 9). Similarly, the G4LZI3 USP ~ CNP0169198 complex exhibited higher conformational variability, with cluster 1 dominating over cluster 0 (1755 vs. 875 occurrences) (Fig. 9).

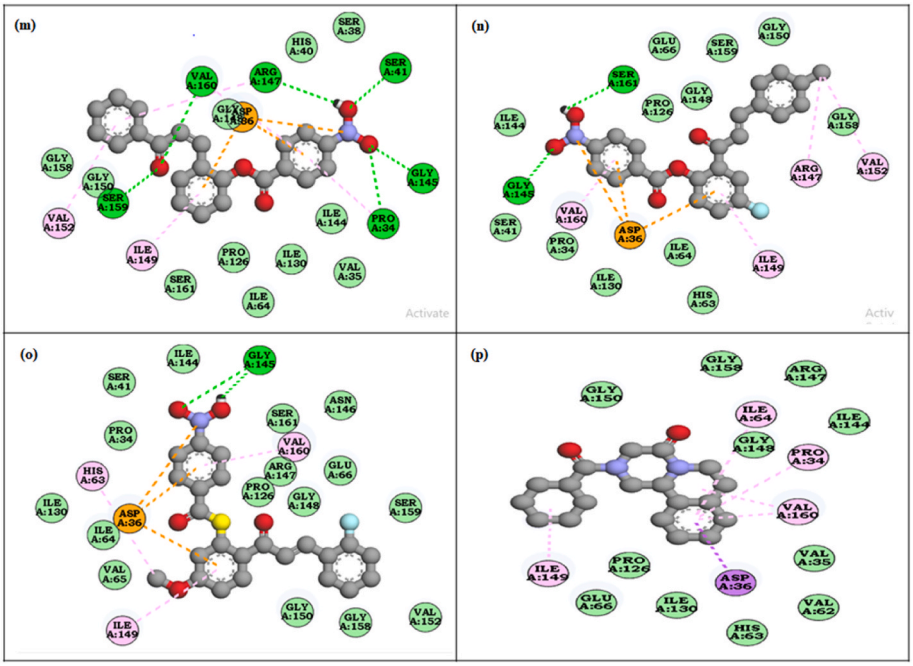

Fig. 3. (*continued*).

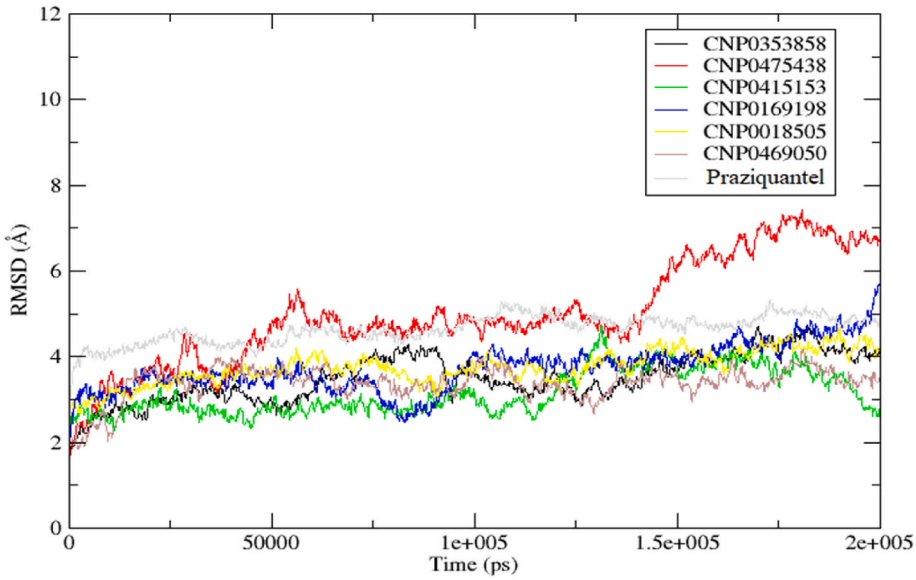

Fig. 4. Root Mean Square Deviation (RMSD) of the selected ligands against the G4LZI3 protein. The plot shows the backbone stability of protein-ligand complexes over 200 ns MD simulation, indicating dynamic behaviour and convergence trends for each ligand.

This suggests a more dynamic interaction that could affect how efficiently the binding occurs. On the other hand, the G4LZI3 USP ~ CNP0475438 complex had 2295 instances of cluster 0, indicating high stability and fewer changes between clusters, which means a strong and consistent binding interaction (Fig. 9).

Meanwhile, the complexes G4LZI3 USP ~ CNP0415153 and G4LZI3 USP ~ CNP0018505 also showed significant stability, recording 2180 and 2230 occurrences of cluster 0, respectively. This stability is associated with fewer transitions to other clusters, reinforcing the idea that these complexes also exhibit strong binding interactions. On the other hand, the G4LZI3 USP ~ CNP0469050 complex showed moderate stability, with 1720 occurrences of cluster 0, suggesting a slightly weaker binding compared to CNP0475438 and CNP0415153. Interestingly, the G4LZI3 USP ~ Praziquantel complex, which is used as a standard, showed 2100 occurrences in cluster 0, indicating good stability. This stability is notable, although it is slightly less pronounced than other complexes like CNP0475438. This observation supports praziquantel's recognized therapeutic effectiveness and is a reference point for comparing other ligands. The clustering results correspond well with the findings from docking and molecular dynamics (MD) simulations. Ligands frequently appearing in cluster 0, such as CNP0475438 and CNP0415153, displayed favourable binding energies during docking and maintained stable root-mean-square deviation (RMSD) profiles in the MD simulations. This alignment between the static docking results and the dynamic MD analyses suggests these ligands could be promising candidates for therapeutic use.

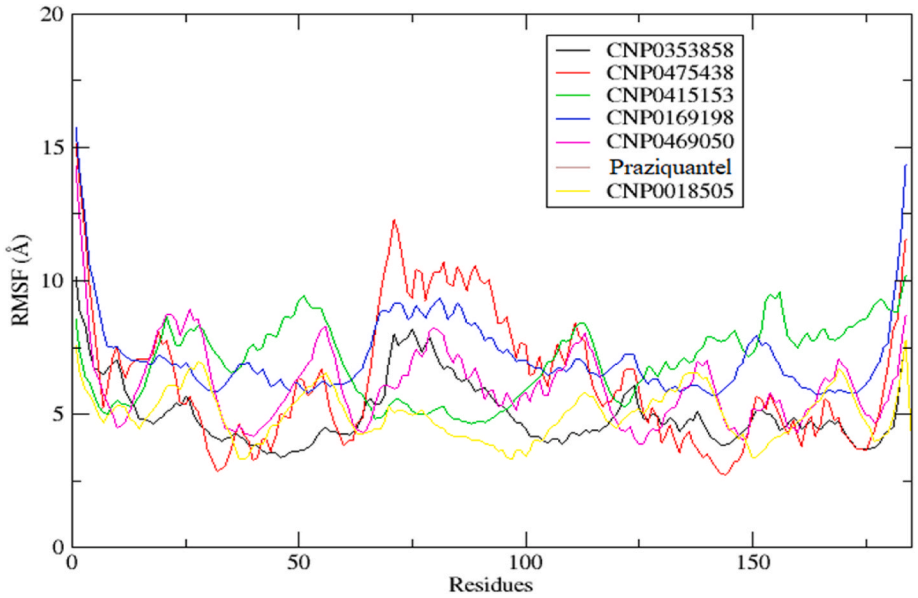

Fig. 5. Root Mean Square Fluctuation (RMSF) of the selected ligands against the G4LZI3 protein. This plot highlights the residue-level flexibility of the protein in complex with each ligand, identifying regions with higher mobility and structural perturbation during simulation.

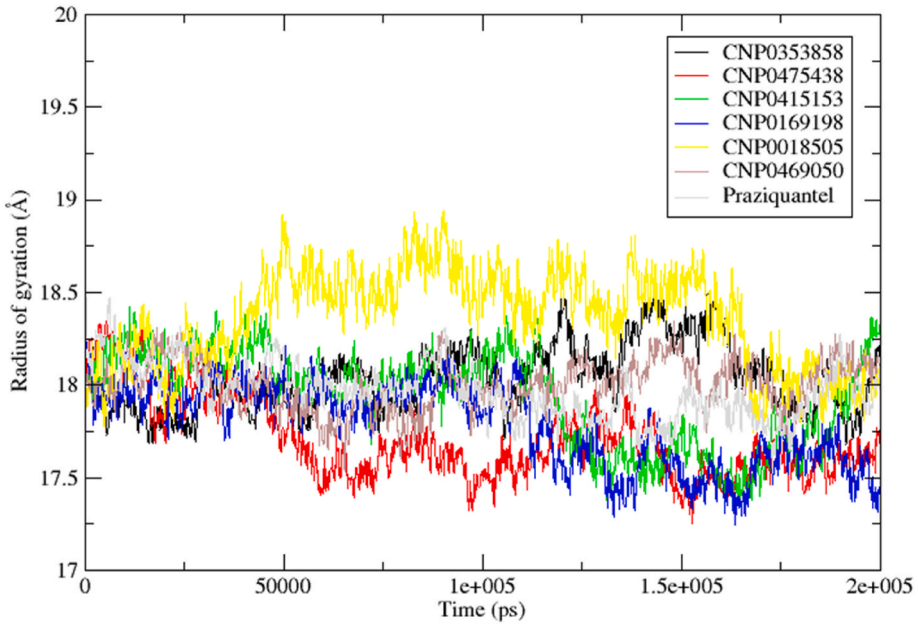

Fig. 6. The Radius of Gyration of the selected ligands against the G4LZI3 protein. The radius of gyration (Rg) reflects the compactness and folding stability of the protein-ligand complexes throughout the simulation period.

3.7. Comparative Analysis of Energy Fluctuations

The analysis of energy fluctuations over the 200 ns molecular dynamics (MD) simulation reveals a significant understanding of the stability and interactions of the G4LZI3 USP protein-ligand complexes (Fig. 10). By focusing on the total energy, van der Waals (vdW), and electrostatic (elec) energy components, we can evaluate the ligands' dynamic behaviour and binding efficiency, comparing them to the standard drug Praziquantel. The total energy, a measure of overall system stability (Hollingsworth and Dror, 2018), showed that CNP0353858 achieved the lowest minimum value (−12731.4895 kcal/mol), which indicates exceptional stability in the binding pocket. In contrast, praziquantel, with a minimum energy of −1759.6915 kcal/mol, exhibited considerable stability but was surpassed by CNP0353858 (Fig. 10). This superior stability for CNP0353858 shows its potential as a strong candidate for therapeutic development.

The vdW energy component, which reflects hydrophobic interactions, was most favourable for CNP0353858, ranging from −1411.667 to 8060.4469 kcal/mol (Fig. 10). These interactions are critical for ligand stabilization within the hydrophobic regions of the binding site. Praziquantel's vdW energy range (−1244.7774 to −1407.897 kcal/mol) demonstrated strong, albeit slightly less favourable, hydrophobic stabilization compared to CNP0353858. This suggests that CNP0353858 may exhibit better compatibility with the hydrophobic binding pocket. Electrostatic interactions, represented by the elec energy, are pivotal for ligand binding specificity (Hollingsworth and Dror, 2018). Praziquantel demonstrated consistently favourable electrostatic energy values (−12501.6662 to −13029.4356 kcal/mol),

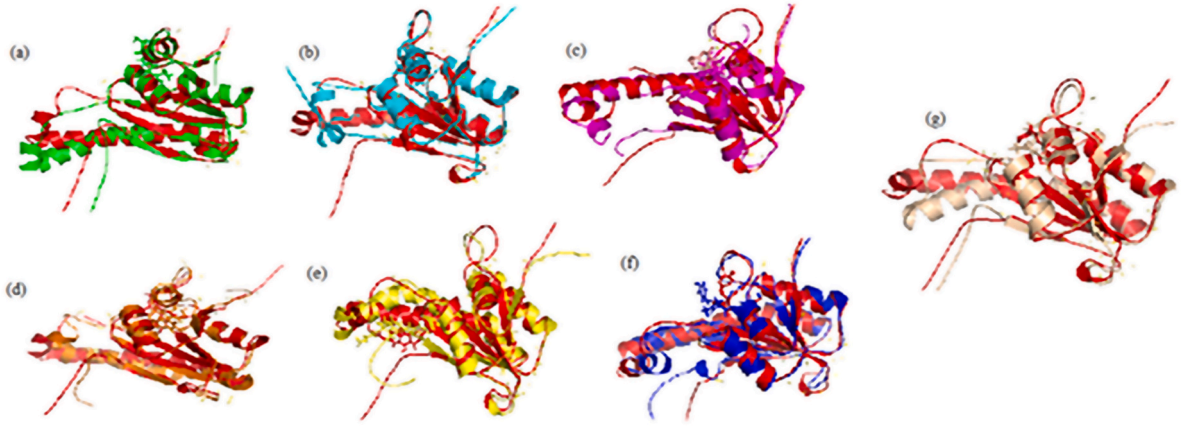

Fig. 7. Summarizes the structural comparison of initial and final conformations of protein-ligand complexes after 200 ns of MD simulation. (a) G4LZI3 USP-CNP0353858: Final structure in green, initial structure in red (RMSF = 1.828 Å). (b) G4LZI3 USP-CNP0475438: Final structure in cyan, initial structure in red (RMSF = 1.930 Å). (c) G4LZI3 USP-CNP0415153: Final structure in magenta, initial structure in red (RMSF = 1.746 Å). (d) G4LZI3 USP-CNP0169198: Final structure in orange, initial structure in red (RMSF = 2.347 Å). (e) G4LZI3 USP-CNP0018505: Final structure in yellow, initial structure in red (RMSF = 2.074 Å). (f) G4LZI3 USP-CNP0469050: Final structure in blue, initial structure in red (RMSF = 1.868 Å). (g) G4LZI3 USP-Praziquantel: Final structure in tint, initial structure in red (RMSF = 1.902 Å). This comparison illustrates the structural displacement and flexibility of the complexes after equilibration, supporting conclusions on binding stability. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

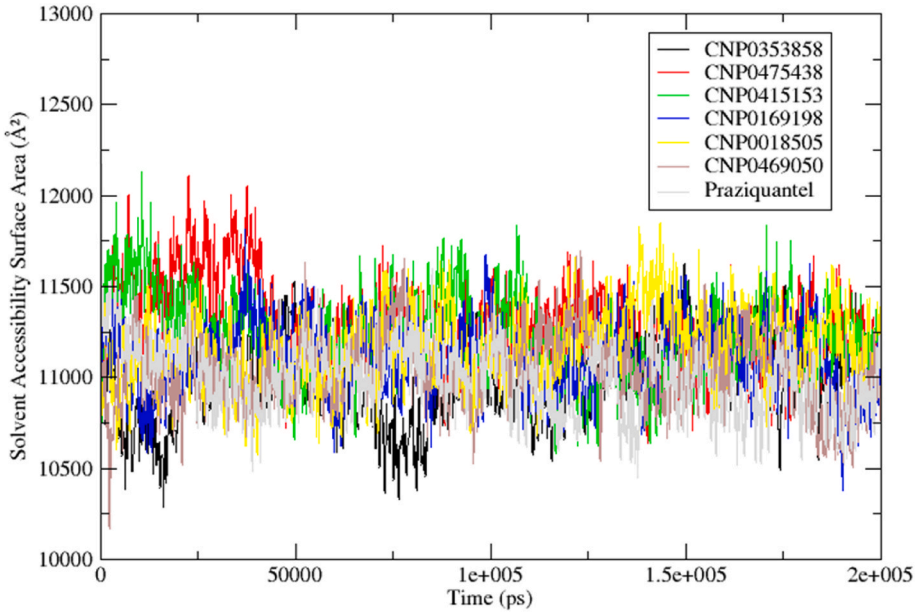

Fig. 8. Solvent Accessibility Surface Area (SASA) analysis of the selected ligands forming a complex with the G4LZI3 protein. SASA analysis indicates the degree of surface exposure and burial of residues upon ligand binding, providing insight into conformational changes and complex hydration.

which align with its established efficacy. However, CNP0353858 also exhibited strong electrostatic interactions, with a minimum energy of −12893.3456 kcal/mol, reinforcing its binding efficiency and stability. When these findings are contextualized with docking results, which provided initial indications of intense binding poses, the MD simulations improve our understanding of ligand dynamics and stability over time. The outstanding performance of CNP0353858 in terms of energy minima is consistent with its high stability observed during molecular dynamics simulations. On the other hand, ligands like CNP0415153, which showed higher minimum energy values of −1088.033 kcal/mol, demonstrated less stable interactions.

3.8. Free binding energy (MMGBSA) analysis

The MMGBSA analysis of binding free energies of MD simulation of 200 ns reveals a crucial understanding of the stability and interaction dynamics of the tested ligands with the G4LZI3 USP target (Genheden and Ryde, 2015). CNP0475438 emerged as the most stable among the evaluated ligands, with the lowest overall binding free energy of −48.51 kcal/mol (Table 4). This strong interaction is driven by significant van der Waals (ΔGvdw = −45.28 kcal/mol) and favourable solvation energy contributions (ΔGsolv = 50.26 kcal/mol). These findings indicate the hydrophobic solid and electrostatic interactions stabilizing the ligand-protein complex. In contrast, Praziquantel, the standard drug, demonstrated a slightly higher overall binding free energy (−41.41 kcal/mol), indicating lower stability.

Other ligands, such as CNP0415153 (−45.65 kcal/mol) and CNP0018505 (−44.06 kcal/mol), also exhibited superior binding energies relative to praziquantel, which further validated their potential as promising candidates (Table 4). Ligands like CNP0353858 (−43.40

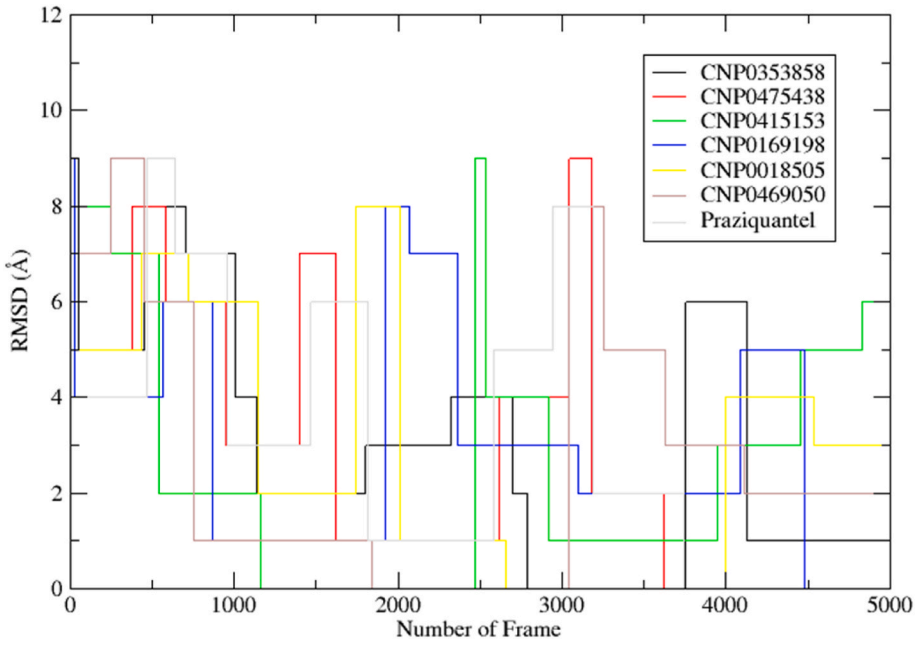

Fig. 9. Clustering Analysis between the selected ligands and the G4LZI3 protein after 200 ns MD simulation. Clustering reveals the most representative conformations and conformational diversity of protein-ligand complexes over the trajectory, aiding in understanding dominant binding modes.

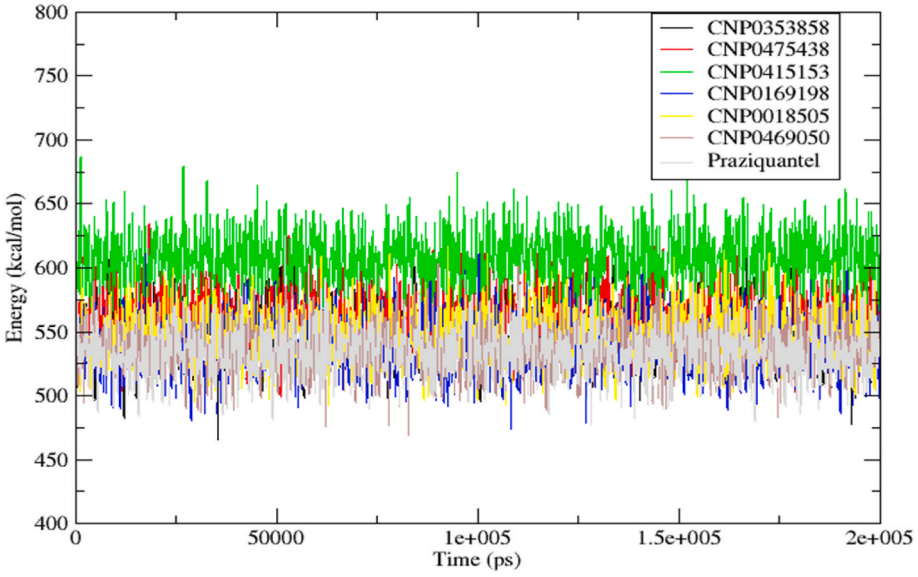

Fig. 10. Energy fluctuations of the system over time. The figure shows the total energy variation during MD simulation, confirming the thermodynamic stability of the protein-ligand complexes under the defined simulation conditions.

Table 4
Free Binding Energy (MMGBSA) analysis of the selected ligand.

S/No.	Compound	Δ G_{vdw} (kcal/mol)	Δ G_{ele} (kcal/mol)	Δ G_{polar} (kcal/mol)	Δ $G_{nonpolar}$ (kcal/mol)	Δ G_{gas} (kcal/mol)	Δ G_{solv} (kcal/mol)	Δ $G_{MM-GBSA}$ (kcal/mol)
1.	CNP0353858	−50.73 ± 0.37	−12.49 ± 0.68	24.57 ± 0.83	−4.83 ± 0.11	−57.30 ± 0.32	57.30 ± 0.11	−43.40 ± 0.17
2.	CNP0475438	−45.28 ± 0.27	−11.07 ± 0.17	20.23 ± 0.19	−5.97 ± 0.01	−56.68 ± 0.20	50.26 ± 0.17	−48.51 ± 0.09
3.	CNP0415153	−49.77 ± 0.41	−13.20 ± 0.38	19.89 ± 1.10	−4.95 ± 0.05	−52.01 ± 0.68	54.40 ± 1.05	−45.65 ± 0.33
4.	CNP0169198	−47.26 ± 0.26	−12.46 ± 1.17	23.96 ± 1.20	−4.53 ± 0.62	−59.72 ± 1.07	56.10 ± 0.16	−43.91 ± 1.18
5.	CNP0018505	−46.06 ± 0.23	−12.83 ± 0.40	26.75 ± 0.42	−5.89 ± 0.02	−59.89 ± 0.44	53.86 ± 0.41	−44.06 ± 0.31
6.	CNP0469050	−41.61 ± 0.19	−10.42 ± 0.76	22.02 ± 0.87	−5.22 ± 0.01	−56.04 ± 0.82	49.80 ± 0.87	−41.47 ± 0.24
7.	Praziquantel	−45.86 ± 0.86	−11.28 ± 0.51	25.28 ± 0.40	−4.84 ± 0.01	−55.14 ± 0.53	50.43 ± 0.41	−41.41 ± 0.27

kcal/mol) and CNP0169198 (−43.91 kcal/mol) showed comparable stability, while CNP0469050 (−41.47 kcal/mol) demonstrated performance like that of praziquantel. These results align well with previous docking and MD simulations, where CNP0475438 exhibited strong binding affinities and minimal RMSD fluctuations, which confirm its dynamic stability over 200 ns. This ligand's ability to maintain firm interaction profiles under dynamic conditions stresses its possibility for further drug development. Similarly, CNP0415153 and CNP0018505, with favourable MMGBSA scores, demonstrated stable binding conformations in MD simulations, correlating with their thermodynamic profiles. The robust van der Waals interactions observed for the top-performing ligands suggest their capacity for effective hydrophobic interactions within the protein's binding pocket. Also, the solvation energies indicate efficient desolvation upon ligand binding, a hallmark of favourable drug-like behaviour. When we compare these ligands with Praziquantel, CNP0475438 shows better binding stability and stronger interactions than the standard drug. This highlights the importance of using MMGBSA, docking, and molecular dynamics simulations to pinpoint and refine potential new therapies. For the accuracy and reproducibility of the reported quantitative findings, all docking simulations were performed in triplicate with consistent grid parameters and scoring functions using AutoDock 4.2. The best-ranked poses from independent runs were compared to confirm the convergence of docking energies and binding conformations. The MM-GBSA analysis used 5000 evenly spaced frames extracted from the final 50 ns of MD trajectories, reducing statistical bias and capturing the dynamic equilibrium. Energy values were reported as mean ± standard deviation for transparency in variation across the simulation window. Clustering analysis was also independently validated using RMSD thresholds to confirm dominant binding conformations.

4. Conclusion

Molecular modelling, docking, and molecular dynamics (MD) simulation were employed to evaluate the binding energies of selected compounds against G4LZI3 USP derived from *Schistosoma mansoni*. Docking analysis identified fifteen compounds with notable binding affinities ranging from −10.6 to −8.5 kcal/mol. From these, six compounds (CNP0353858: −9.5 kcal/mol, CNP0475438: −10.6 kcal/mol, CNP0415153: −9.8 kcal/mol, CNP0169198: −9.2 kcal/mol, CNP0018505: −9.9 kcal/mol, and CNP0469050: −8.4 kcal/mol) were selected for further analysis using MD simulation and MM-GBSA. CNP0475438, CNP0415153, and CNP0353858 emerged as the most promising candidates demonstrating superior binding affinity, stability, and favourable energetic profiles, outperforming the standard drug praziquantel and warranting further experimental validation. In addition to docking and MM-GBSA, key structural stability indicators such as RMSD, RMSF, SASA, and radius of gyration were analyzed across 200 ns simulations, confirming the stable and compact binding of the lead compounds. Furthermore, clustering analysis revealed consistent conformational behaviour, reinforcing the dynamic reliability of these interactions. This work introduces the *S. mansoni* Universal stress G4LZI3 protein as a novel drug target, filling a critical gap in developing schistosomiasis therapeutics. Focusing on stress-response proteins provides an alternative path to overcoming existing treatments' resistance and stage-specific limitations. These findings hold real-world relevance as they offer practical leads for drug development pipelines, particularly for regions facing rising treatment resistance and inadequate access to diverse anti-parasitic agents.

CRediT authorship contribution statement

Lihle Mahamba: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis. **Mustafa Alhaji Isa:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation. **Abidemi Paul Kappo:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization.

Ethical approval

Not Applicable.

Data availability

All data and materials are available in the supplementary materials.

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.amolm.2025.100084.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 1–3. https://doi.org/10.1038/s41586-024-07487-w.

Adenowo, A.F., Masamba, P., Qokoyi, N.K., Oyinloye, B.E., Kappo, A.P., 2021. Recombinant expression and biophysical characterization of a druggable *Schistosoma mansoni* universal stress G4LZI3 protein. Adv. Pharmaceut. Bull. 12 (2), 366. https://doi.org/10.34172/apb.2022.035.

Ausaf Ali, S., Imtaiyaz Hassan, M., Islam, A., Ahmad, F., 2014. A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. Curr. Protein Pept. Sci. 15 (5), 456–476. https://doi.org/10.2174/1389203715666140327114232.

Baell, J.B., Holloway, G.A., 2010. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J. Med. Chem. 53 (7), 2719–2740. https://doi.org/10.1021/jm901137j.

Case, D.A., Aktulga, H.M., Belfon, K., Cerutti, D.S., Cisneros, G.A., Cruzeiro, V.W.D., Merz Jr., K.M., 2023. AmberTools. J. Chem. Inf. Model. 63 (20), 6183–6191. https://doi.org/10.1021/acs.jcim.3c01153.

Clague, M.J., Urbé, S., Komander, D., 2019. Breaking the chains: deubiquitylating enzyme specificity begets function. Nat. Rev. Mol. Cell Biol. 20 (6), 338–352. https://doi.org/10.1038/s41580-019-0099-1.

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

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2 (9), 1511–1519. https://doi.org/10.1002/pro.5560020916.

Daina, A., Michielin, O., Zoete, V., 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci. Rep. 7 (1), 42717. https://doi.org/10.1038/srep42717.

Daura, X., Gademann, K., Jaun, B., Seebach, D., Van Gunsteren, W.F., Mark, A.E., 1999. Peptide folding: when simulation meets experiment. Angew. Chem. Int. Ed. 38 (1-2), 236–240. https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID.

Egan, W.J., Merz, K.M., Baldwin, J.J., 2000. Prediction of drug absorption using multivariate statistics. J. Med. Chem. 43 (21), 3867–3877. https://doi.org/10.1021/jm000292e.

Eisenberg, D., Lüthy, R., Bowie, J.U., 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277, 396–404. https://doi.org/10.1016/s0076-6879(97)77022-8.

Elbaz, T., Esmat, G., 2013. Hepatic and intestinal schistosomiasis. J. Adv. Res. 4 (5), 445–452. https://doi.org/10.1016/j.jare.2012.12.001.

Genheden, S., Ryde, U., 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expet Opin. Drug Discov. 10 (5), 449–461. https://doi.org/10.1517/17460441.2015.1032936.

Ghose, A.K., Viswanadhan, V.N., Wendoloski, J.J., 1999. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. J. Combin. Chem. 1 (1), 55–68. https://doi.org/10.1021/cc9800071.

Gryseels, B., Polman, K., Clerinx, J., Kestens, L., 2006. Human schistosomiasis. Lancet 368 (9541), 1106–1118. https://doi.org/10.1016/S0140-6736(06)69440-3.

He, G., Qiu, M., Li, R., Ouyang, L., Wu, F., Song, X., Li, C., Xiang, M., Yu, L., 2012. Multicomplex-based pharmacophore-guided 3D-QSAR studies of N-substituted 2′-(Aminoaryl)Benzothiazoles as aurora-A inhibitors. Chem. Biol. Drug Des. 79 (6), 960–971. https://doi.org/10.1111/j.1747-0285.2012.01366.x.

Holderbach, S., Adam, L., Jayaram, B., Wade, R.C., Mukherjee, G., 2020. RASPD+: fast protein-ligand binding free energy prediction using simplified physicochemical

features. Front. Mol. Biosci. 7, 601065. https://doi.org/10.3389/fmolb.2020.601065.
Hollingsworth, S.A., Dror, R.O., 2018. Molecular dynamics simulation for all. Neuron 99 (6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.
Isa, M.A., 2019. Homology modeling and molecular dynamic simulation of UDP-N-acetylmuramoyl-l-alanine-d-glutamate ligase (MurD) from Mycobacterium tuberculosis H37Rv using in silico approach. Comput. Biol. Chem. 78, 116–126. https://doi.org/10.1016/j.compbiolchem.2018.11.002.
Isa, M.A., Abubakar, M.B., Mohammed, M.M., Ibrahim, M.M., Gubio, F.A., 2021. Identification of potent inhibitors of ATP synthase subunit c (AtpE) from Mycobacterium tuberculosis using in silico approach. Heliyon 7 (12), e08482. https://doi.org/10.1016/j.heliyon.2021.e08482.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596 (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.
Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. J. Chem. Inf. Model. 51 (10), 2778–2786. https://doi.org/10.1021/ci200227u.
Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 23 (1–3), 3–25. https://doi.org/10.1016/s0169-409x(00)00129-0.
Lüthy, R., Bowie, J.U., Eisenberg, D., 1992. Assessment of protein models with three-dimensional profiles. Nature 356 (6364), 83–85. https://doi.org/10.1038/356083a0.
Masamba, P., Adenowo, A.F., Oyinloye, B.E., Kappo, A.P., 2016. Universal stress proteins as new targets for environmental and therapeutic interventions of Schistosomiasis. Int. J. Environ. Res. Publ. Health 13 (10), 972. https://doi.org/10.3390/ijerph13100972.
Masamba, P., Kappo, A.P., 2021. Parasite survival and disease persistence in cystic fibrosis, schistosomiasis and pathogenic bacterial diseases: a role for universal stress proteins? Int. J. Mol. Sci. 22 (19), 10878. https://doi.org/10.3390/ijms221910878.
Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A. J., 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 19 (14), 1639–1662. https://doi.org/10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.
Pirzaman, A.T., Sepidarkish, M., Alizadeh, F., Al-Obidy, S., Ebrahimi, P., Kianifard, N., et al., 2024. Prevalence of human Schistosoma mansoni infection in endemic regions (2010–2024): a systematic review and meta-analysis. eClinicalMedicine 77, 102855. https://doi.org/10.1016/j.eclinm.2024.102855.
Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., 2002. Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. 45 (12), 2615–2623. https://doi.org/10.1021/jm020017n.
Zhang, X., Zhang, S., Hao, F., Lai, X., Yu, H., Huang, Y., Wang, H., 2005. Expression, purification and properties of shikimate dehydrogenase from Mycobacterium tuberculosis. J. Biochem. Mol. Biol. 38 (5), 624–631. https://doi.org/10.5483/bmbrep.2005.38.5.624.