Chapter- 5

Comparative genome analysis of different strains of *Leishmania* species using available sequences

5.1. Abstract

Leishmaniasis, an intricate vector-borne malady caused by intracellular protozoan parasites of the *Leishmania* genus, poses a substantial public health challenge in tropical and subtropical regions globally. Subtractive genomics and structure-based methodologies were employed to unearth shared drug targets and inhibition of the target by potential ligand across five *Leishmania* species strains. The subtractive genomics approach unveiled Glutamate Dehydrogenase (GDH) as a promising drug target in combating *Leishmania* infections. The investigation meticulously followed established methodologies observed in analogous studies, orthologous groups, druggability tests, etc. Multiple sequence alignment uncovered conserved sequences in GDH along with phylogenetic tree analysis provided insights into the evolutionary origin and close relationships of GDH across *Leishmania* species, further reinforcing its potential as a therapeutic target. Conserved sequences in GDH along with its function in pathogenicity provided insights into the close relationships of GDH across *Leishmania* species. Employing a structure-based approach, our study delved into the molecular interactions between GDH and three ligands (Bithionol, GW5074, Hexachlorophene) through molecular docking and a 100ns MD simulation. Notably, GW5074 showcased a pronounced affinity for GDH, as evidenced by stable RMSD values, more compact conformation and GW5074 forming more H-bonds than bithionol. MMPBSA analysis affirmed the superior binding energy of the GW5074-GDH complex, underscoring its potential as a potent ligand for drug development. This comprehensive analysis of GW5074 as a promising candidate for impeding GDH activities in *Leishmania* species help in developing therapeutics against *Leishmania* infections.

This chapter considered Lig1, Lig2 and Lig3 to Bithionol, GW5074 and Hexachlorophene, respectively.

5.2. Introduction

Leishmaniasis is distinguished as a remarkably diverse and intricate vector-borne disease, marked by its complexity and the participation of obligatory intracellular protozoan parasites from the *Leishmania* genus [1]. This ailment presents a considerable public health dilemma,

impacting diverse regions globally, notably in tropical and subtropical zones [2]. Over 12 million individuals affected by *Leishmaniasis* worldwide, with an annual incidence of 0.9 to 1.6 million new cases. Furthermore, the number of fatalities resulting from *Leishmaniasis* ranges from 20,000 to 30,000, a considerable figure [3]. There are four distinct forms of leishmaniasis that impact the human body: cutaneous leishmaniasis, mucocutaneous leishmaniasis, VL, and post-kala-azar dermal leishmaniasis. As per the 2022 data provided by the World Health Organization (WHO), cutaneous leishmaniasis demonstrates significant endemicity worldwide, particularly impacting the nations in South America, Africa, the Mediterranean basin, the Middle East, and central Asia (WHO, 2023)[4]. Concurrently, VL cases are prevalent in various regions, including India, South Sudan, Sudan, Brazil, Ethiopia and Somalia. Moreover, in India, regions such as Bihar, Jharkhand, Uttar Pradesh and West Bengal are primarily afflicted by visceral leishmaniasis [5].

The life cycle of *Leishmania* involves two primary stages: the amastigote stage and the promastigote stage. The female sand fly carries the promastigote form, transmitting it through biting and then facilitating its entry into human macrophages. Once inside the macrophages, the parasite undergoes a transformation into the amastigote form and multiplies through asexual division [6]. In the human host, leishmaniasis predominantly affects internal organs, with a notable impact on the liver and spleen along with symptoms are fever, weight loss, anemia, etc. The severity of the infection can lead to detrimental consequences for these vital organs, potentially culminating in fatal outcomes [7].

The treatment of leishmaniasis primarily involves the use of chemical compounds such as miltefosine, paromomycin, pentavalent antimonial, and AmpB. These medications constitute the contemporary choices available for addressing leishmaniasis [8]. Concurrently, preventive measures aim to disrupt the pathogen's life cycle by implementing strategies such as avoiding sandfly bites, controlling sandfly habitats, promoting early diagnosis and treatment, and managing animal reservoir hosts. However, despite these efforts, several challenges persist in the fight against leishmaniasis [9]. One significant challenge is the emergence of resistant strains towards existing medicines, posing a threat to the effectiveness of current treatment approaches [10]. Issues related to the efficacy of available treatments, adverse effects, and the overall cost further contribute to the complexity of addressing leishmaniasis [11]. Thus, ongoing research and collaborative efforts are essential to overcome these challenges and enhance the effectiveness of both treatment and preventive strategies against this debilitating disease.

In this study, we employed an integrated methodology, combining subtractive genomics and structure-based approaches, to thoroughly examine and analyze the proteomic data of five distinct *Leishmania* species. The subtractive genomics technique, a well-established and widely utilized *in silico* method, entails the removal of host cell sequences and proteomes from those of the pathogens [12] and similar approach was also followed against human Bartonella, *Acinetobacter baumannii*, etc. [13, 14]. This method isolates pathogen-specific protein sequences absent in host cells. Key steps involve removing paralogous sequences, identifying pathogen-specific essential genes lacking homologs in the host proteome, and evaluating the druggability of these vital proteins [15]. Furthermore, a structure-based approach was integrated to observe the interactions between the selected target protein and potential inhibitory compounds against *Leishmania* species. Molecular docking methodology was involved in predicting and analyzing the spatial arrangement of the ligands within the binding site of the target protein, along with evaluating the strength of their interaction [16, 17]. Subsequent to the docking analysis, the MD simulation technique was employed a detailed scrutiny of the dynamic behavior and interactions among atoms and molecules over an extended period.

Numerous investigations have explored the activity of Glutamate dehydrogenase (GDH) across various pathogens, including *Plasmodium falciparum, Benjaminiella poitrasii, Aspergillus terreus*, among others [18-20], and have endeavored to inhibit this enzyme through diverse methods. GDH's crucial role in the conversion of glutamic acid to α-ketoglutarate within the tricarboxylic acid cycle underscored its significance in the cellular processes of *Leishmania* [21]. Unraveling the role of GDH in *Leishmania* species assumes significance in devising strategies to regulate the growth and development of these protozoa. Moreover, the exploration of GDH inhibitors holds promise as a potential therapeutic avenue for addressing *Leishmania* infections. To assess the potential inhibitory effect, ligands previously identified in the literature against GDH were subjected to molecular docking analyses to discern their binding affinities. Two compounds exhibiting favorable results in binding affinity with the target were identified for further investigation. Subsequently, MD simulations were conducted to explore the interactions between the selected ligands and GDH proteins of five *Leishmania* species, aiming to inhibit the pathogen [22].

Through this approach, a comprehensive understanding of the intricate interactions between drug targets and ligands was attained. The molecular docking provided insights into the initial binding modes and affinities of the ligands with the target protein, while MD

simulations allowed for the exploration of the dynamic behavior and stability of these protein-ligand complexes over time. These findings contribute to the groundwork for potential therapeutic strategies against *Leishmania* infections by elucidating the molecular interactions involved in inhibiting GDH, a crucial enzyme in the protozoan's metabolic processes.

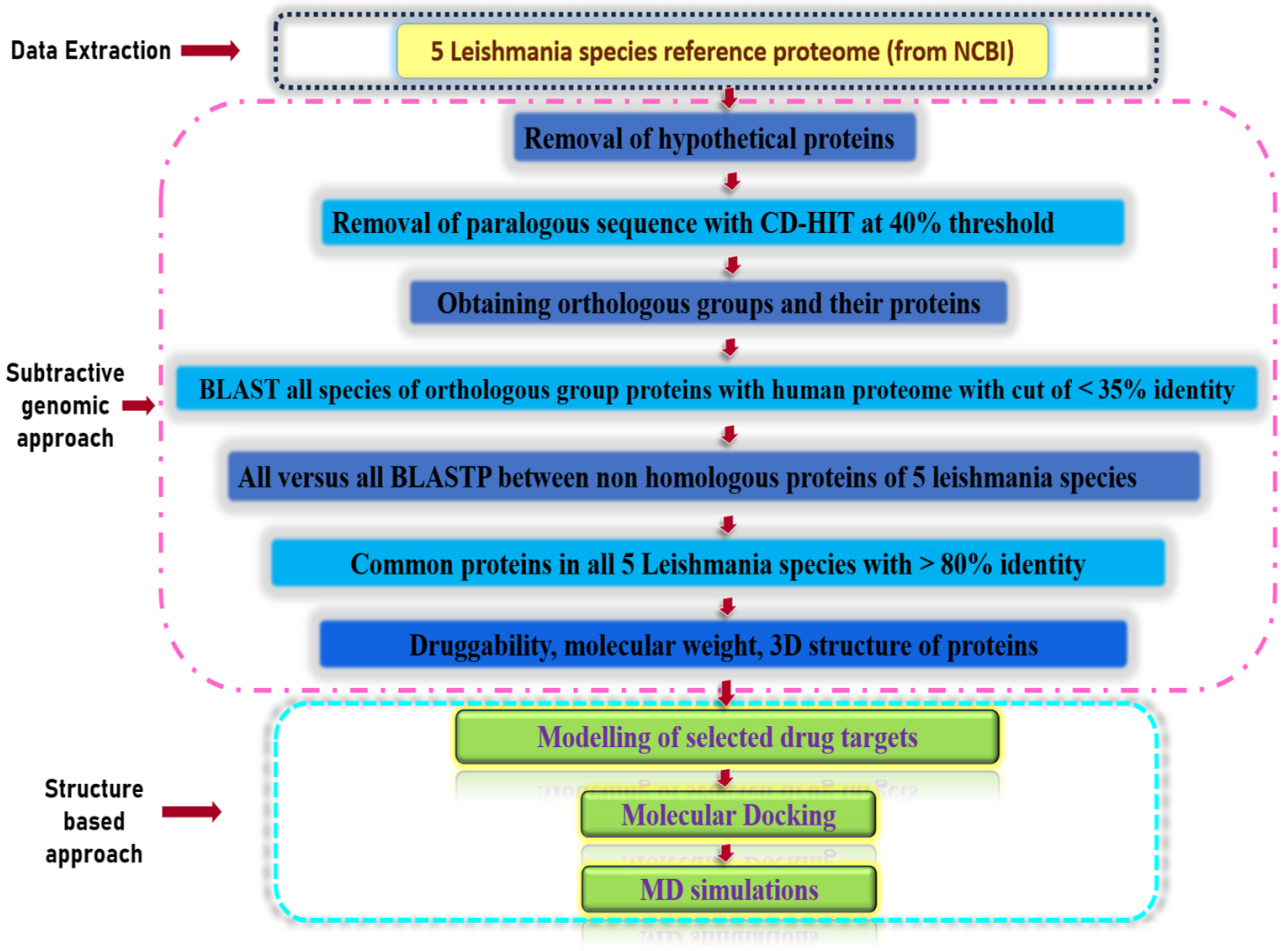

Figure 5.1: Illustration depicting the subtractive genomic and structure-based methodologies for identifying drug targets in *Leishmania* species.

5.3. Materials and Methods

A subtractive genomics approach was applied to analyze the entire proteome of five *Leishmania* species, aiming to identify a shared novel drug target to prevent different forms of Leishmaniasis. Subsequently, a structure-based approach was employed to inhibit the selected target. The results section outlines the comprehensive workflow for subtractive

genomic analysis and the structure-based approach. The workflow of the study is depicted in Figure 5.1.

5.3.1. Data collection and screening of proteins

The proteomes of five *Leishmania* species were retrieved from the National Centre for Biotechnology Information (NCBI) database as of July 2023 [23]. The *Leishmania* strains included *Leishmania braziliensis* MHOM/BR/75/M2904 (assembly ASM284v2), *Leishmania donovani* (assembly ASM22713v2), *Leishmania infantum* JPCM5 (assembly ASM287v2), *Leishmania mexicana* MHOM/GT/2001/U1103 (assembly ASM23466v4), and *Leishmania major* strain Friedlin (assembly ASM272v2). The proteome of Homo sapiens (assembly GRCh38.p14) was acquired for comparison. Hypothetical proteins, lacking defined functions, were prevalent in all proteomes and thus excluded from analysis to ensure the identification of potential drug targets with known functions. After removing hypothetical proteins, proteomes underwent refinement, eliminating paralogous proteins with over 40% sequence identity via CD-HIT software [24]. This procedure aimed to eliminate duplicates while retaining the non-paralogous protein sequences within the entire proteome [15].

5.3.2. Identification of orthologous groups and non-homologous sequences

The previously filtered proteomes of five *Leishmania* species, underwent analysis using the OrthoFinder software with default parameters [25]. By utilizing BLAST searches and the Markov Cluster Algorithm, we identified regions with homologous sequences. During this procedure, the software systematically generated orthologous groups by leveraging the identified protein sequences in them [26]. Proteins in humans that bear similarities to proteins of infectious agent could hinder the binding of potent ligands to the active sites of pathogen proteins. Consequently, we screened for functionally comparable homologous sequences between the *Leishmania* species and human proteomes. For further analysis, non-homologous sequences from the *Leishmania* proteome, exhibiting an identity of less than 35%, were selected. BLASTp processes were executed between the screened proteomes of the five *Leishmania* species and the human proteome, with a threshold expectation value cut-off (e-value) of 0.0001 [27]. Our focus is on the proteomes of five Leishmania species, aiming to find shared proteins among them. We conducted BLAST searches, extracting common proteins with an identity threshold exceeding 80%.

5.3.3. Druggability analysis and identification of novel drug targets

A protein target is deemed druggable when it possesses the capability to strongly bind with therapeutic ligands. To assess this, common non-homologous proteins from *Leishmania* species underwent evaluation through BLASTp against the FDA approved drug targets in the DrugBank database, using an E value of 10^{-5} [28]. This process aimed to determine the drug-target-like potential of the shortlisted proteins, ultimately identifying novel drug targets. Subsequently, proteins from the *Leishmania* proteome exhibiting over 50 percent identity with established targets were chosen.

5.3.4. Protein active site prediction and molecular docking studies

Since the chosen protein drug target lacks experimental 3D structures in the Protein Data Bank, we utilized the AlphaFold Protein Structure Database to acquire predicted structures for the selected protein across all five *Leishmania* species [29]. Subsequently, multiple sequence alignments (MSA) were conducted for the selected protein in each species, aiming to further scrutinize and explore the conservation of these proteins with the help of Jalview tool [30]. Additionally, phylogenetic tree analysis of the selected proteins was carried with the help of MEGA4 software to observe the evolutionary relationships among them [31].

To determine the active site of the chosen protein, the CASTp 3.0 software was employed to identify binding pockets having active sites across the entire protein, maintaining a consistent probe radius of 2.1 Å. The software assessed cavity volumes and surface areas using both solvent-accessible and molecular surface models, presenting the outcomes in descending order according to binding pocket volume and areas [32]. The identified active sites were chosen for the docking of ligands against their corresponding proteins. To gain insights into the ligand's binding mechanism within the protein specific area, molecular docking experiments were implemented. To achieve this, AutoDock 4.2 program was employed [33]. The active sites of the GDH protein in five *Leishmania* species were predicted using a computational tool. Meanwhile, the ligands (bithionol, GW5074, and hexachlorophene) were sourced from a previous study where these compounds were reported to inhibit *Plasmodium falciparum* GDHs in an *in vitro* experiment [18]. Consequently, these selected compounds were employed to assess their binding affinity with the GDH protein in the five *Leishmania* species.

5.3.5. Molecular Dynamic Simulations

MD simulations were employed to comprehend the behavior of proteins in specific scenarios, considering factors such as the presence or absence of chemical compounds [16]. The dynamics of complexes involving the GDH protein of *Leishmania braziliensis* and *Leishmania donovani* without ligands and selected ligands were investigated by MD. In this regard, the GROMACS 2020.4 package was employed for conducting all-atom simulations, and the GROMOS 54a7 force field was applied [11]. The topology for bithionol and GW5074, along with the derivation of force field parameters, was created using the Automated Topology Builder (version 3.0) online tool. Binding energy calculation approach is known as the MM/PBSA (Molecular Mechanics Poisson–Boltzmann Surface Area) method, employed for determining the interaction energy between proteins and ligands [34].

5.4. Results

The current investigation employed subtractive genomics and a structure-based approach to identify and characterize potential drug targets in *Leishmania* species. The study process of target identification and structure-based approach comprises three main segments of research. The initial phase involved extracting the proteomes of *Leishmania* species from the NCBI database. The second phase encompassed a subtractive analysis of the acquired proteomes. Through protein sequence-level analysis, we excluded non-paralogous proteins and identified orthologous groups. Along with it, proteins non-homologous to humans were removed, and at last the druggability of proteins were assessed. In the third segment, molecular docking and molecular dynamics simulation were performed between the selected protein target and ligands. The findings of the analysis are elucidated in the subsequent subdivisions.

5.4.1. Data collection of proteomes and protein sequence exclusion

Reference proteomic sequences from the strains of *Leishmania* species, including *Leishmania braziliensis* MHOM/BR/75/M2904, *Leishmania donovani*, *Leishmania infantum* JPCM5, *Leishmania mexicana* MHOM/GT/2001/U1103, and *Leishmania major* strain Friedlin, comprising 8160, 8014, 8150, 8147, and 8216 protein sequences, respectively, were retrieved from the NCBI database.

Initially, we removed hypothetical proteins without specified functions from the five *Leishmania* species. Next, CD-hit software was used to filter out paralogous sequences from

the proteomes. Using a 40% sequence identity cut-off, around 2400 non-paralogous proteins were identified collectively among the five species. [35]. This process resulted in the rejection of approximately 350-450 proteins for each proteome, which exhibited paralogous characteristics. The outcomes of this analysis are detailed in Table 5.1.

Table 5.1: Details of reference proteome of 5 *Leishmania* species

Species	Total Proteome (Mb)	Total number of proteins	After removal of hypothetical proteins	CD-HIT of 40%
***L. donovani* (BPK282A1)**	32.45	8014	2767	2404
***L. major* (Friedlin)ASM272v2**	32.79	8316	3043	2408
***L. infantum* (JPCM5)**	32.12	8150	2905	2412
***L. braziliensis* (MHOM/BR/75/M2904)**	33.50	8160	2940	2360
***L. mexicana* (MHOM/CT/2001/U1103)**	32.08	8147	2929	2377

5.4.2. Analysis of orthologous groups and non-homologous protein

The subsequent step involved leveraging the set of five proteomes to eliminate protein sequences within *Leishmania* species that did not constitute the core proteome. To achieve this, OrthoFinder software was employed, aiming to identify orthologous groups that encompass shared proteins present in all five proteomes [26]. This comprehensive analysis resulted in the initial identification of 2035 orthologous groups. Each *Leishmania* species—*Leishmania braziliensis*, *Leishmania donovani, Leishmania infantum, Leishmania mexicana and Leishmania major*— contributed to these groups with 2193, 2185, 2184, 2182, and 2194 proteins, respectively. This detailed approach ensured a thorough exploration of orthologous proteins shared among the *Leishmania* species.

Proteins that function similarly in both pathogenic and human systems have evolved as homologs across time and are more likely to trigger autoimmunity in humans. BLASTp was utilized to identify non-homologous protein sequences from the *Leishmania* proteome compared to the human proteome [13]. Employing a cut-off (threshold E-value) of 10^{-5}, this

analysis resulted in the exclusion of approximately 150-170 homologous proteins for each proteome. Subsequently, non-homologous sequences of *Leishmania braziliensis*, *Leishmania donovani, Leishmania infantum, Leishmania mexicana and Leishmania major*, with less than 35% similarity to the human host were 729, 733, 731, 728 and 729, respectively. Additionally, common proteins among the refined *Leishmania* species proteomes, showing over 80% identity after performing BLAST and consistent functions across all proteomes leads to 421 proteins per *Leishmania* species.

5.4.3. Druggability analysis and drug target identification

Next, the 421 non-homologous protein sequences from each *Leishmania* species underwent BLASTp against 5350 known target proteins from the DrugBank database. The aim was to detect sequence similarities with drug target proteins. Only proteins sharing over 50% sequence identity with FDA-approved therapeutic targets were retained for further analysis, while others were disregarded. [36]. Consequently, only five therapeutic targets exhibited more than 50% sequence identity among the 421 protein sequences within each proteome of the five *Leishmania* species. Conversely, excluded proteins from *Leishmania* did not show significant similarity to DrugBank-listed drug targets. Table 5.2 provides a detailed list of proteins with desirable identity to DrugBank targets.

Among the five potential drug targets, Glutamate dehydrogenase (GDH) emerged as a crucial, non-homologous, and innovative druggable target against *Leishmania* species. Its selection was influenced by its function as a precursor in the tricarboxylic acid (TCA) cycle. Although not previously identified as a drug target in *Leishmania* species, it's been reported in another protozoan, *Plasmodium falciparum*. [21]. Moreover, GDH proteins across *Leishmania braziliensis, Leishmania donovani, Leishmania infantum, Leishmania mexicana,* and *Leishmania major*, a multiple sequence alignment was conducted, revealing an identity of over 88 percent. [37, 38]. Additionally, a phylogenetic tree was constructed, indicating a close relationship between the GDH of *L. donovani* and *L. infantum*, while slight differences were observed in the phylogeny of GDH between *L. donovani* and *L. braziliensis*.

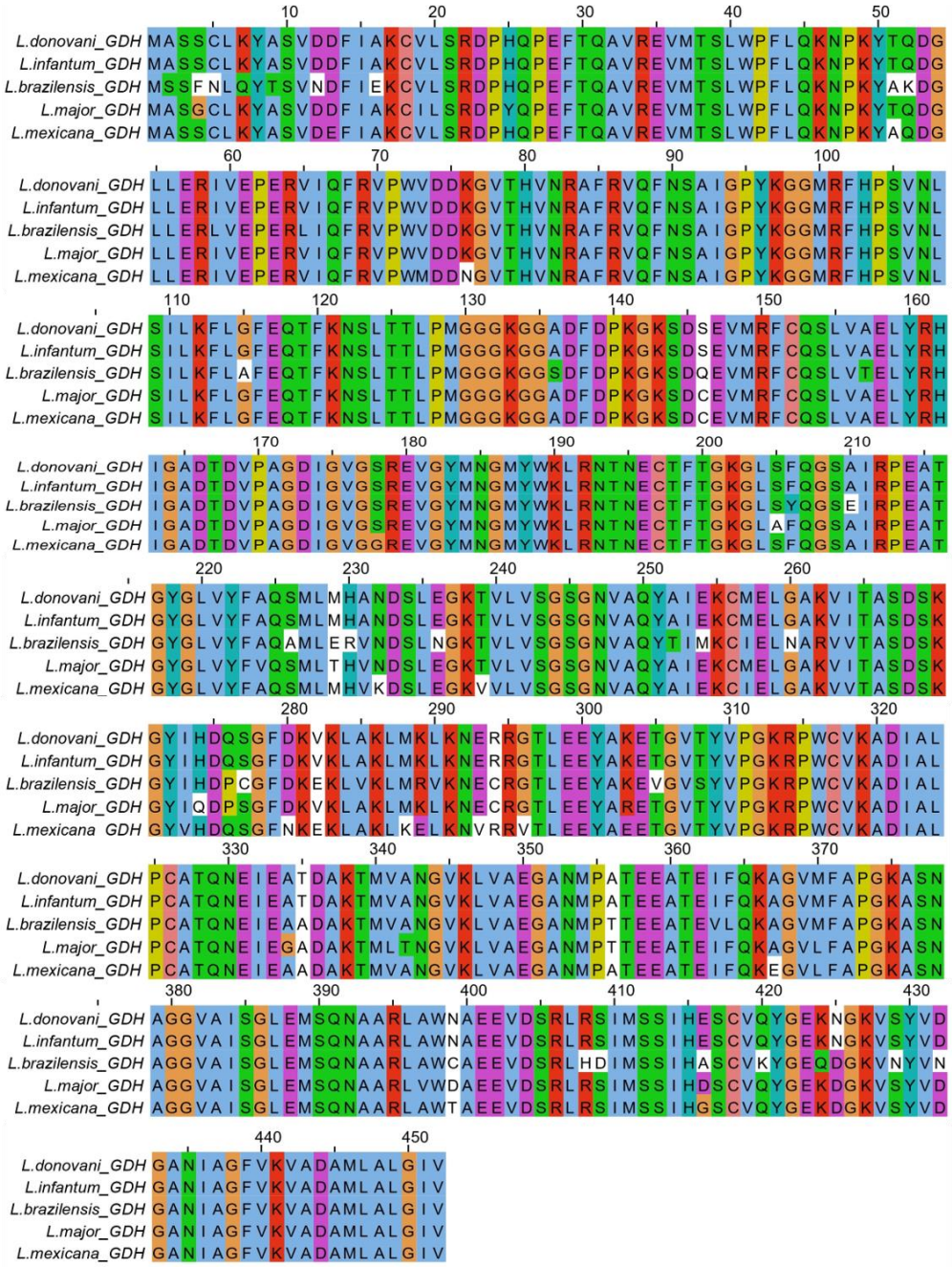

Figure 5.2: Multiple Sequence Alignment of GDH proteins

Table 5.2: Druggability test of proteins

Species	Q7K9G0	%	P39126	%	Q8T8E9	%	Q8ILF7	%	Q53ZE5	%
L. brazilensis	MST*	85.67	IDH*	57.75	UGE*	57.45	GDH*	54.66	DDH*	56.57
L. donovani	MST	95.67	IDH	58.47	UGE	57.99	GDH	55.15	DDH	54.92
L. infantum	MST	95.94	IDH	58.47	UGE	57.99	GDH	55.15	DDH	54.95
L. major	MST	100	IDH	58.23	UGE	58.24	GDH	54.66	DDH	53.35
L. mexicana	MST	95.96	IDH	58.71	UGE	57.54	GDH	54.66	DDH	54.95

* 3-mercaptopyruvate sulfurtransferase (**MST**), Isocitrate dehydrogenase (**IDH**), UDP-glucose 4 epimerase (**UGE**), Glutamate dehydrogenase (**GDH**) and Dihydroorotate dehydrogenase (**DDH**)

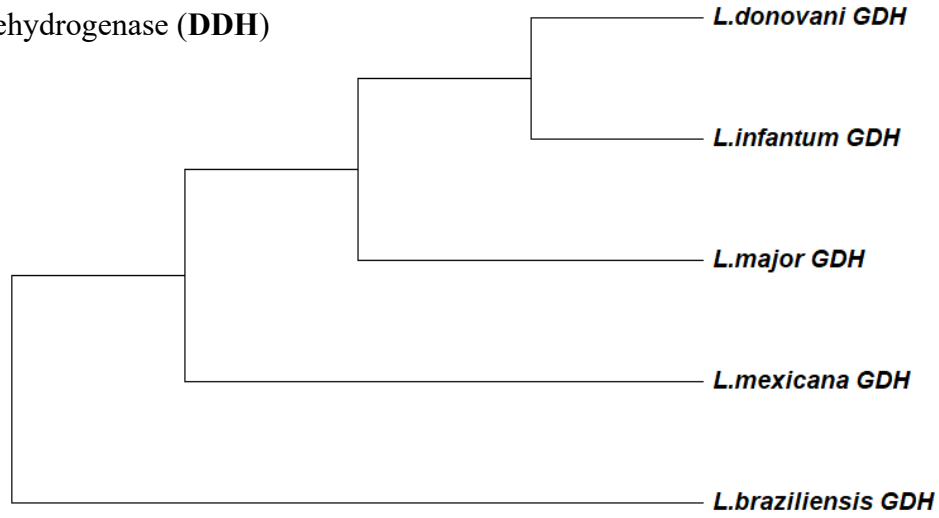

Figure 5.3. Phylogenetic tree of GDH proteins.

Table 5.3: List of selected protein targets.

Sl. No.	Protein targets	Proteins	Species
1	3-mercaptopyruvate sulfurtransferase	Q7K9G0	*Leishmania major*
2	Isocitrate dehydrogenase	P39I26	*Bacillus subtilus*
3	UDP-glucose 4 epimerase	Q8T8E9	*Trypanosona brucci*
4	Glutamate dehydrogenase	Q8ILF7	*Plasmodium falciparum*
5	Dihydroorotate dehydrogenase	Q53ZE5	*Lactococcus lactics*

5.4.4. Analysis of protein target

Alpha Fold method structures of GDH for *L. braziliensis, L. donovani, L. infantum, L. mexicana*, and *L. major* were retrieved from the UniProt database with the accession numbers A4HGZ4, A0A3S7X1Z4, A4I426, E9B0B3 and Q4Q7X1, respectively. The stereochemical quality of the structure generated by the AlphaFold method was evaluated using PROCHECK. Ramachandran plot analysis revealed that all five GDH proteins mainly had residues in the most favored regions, surpassing 90% [39]. The analysis confirmed the structural integrity of the selected proteins, validating both the overall protein structure and the residue-by-residue geometry. Figure 5.4 depicted the three-dimensional GDH protein structures, with detailed Ramachandran plot results in Table 5.4. In Table 5.5, list of GDH proteins with their IDs obtained from Alphafold method is mentioned whereas Table 5.6 provides information about percentage identity and RMSD of the GDH protein among five different *Leishmania* species.

Table 5.4: Ramachandran plot analysis of all five proteins

Sl. No.	GDH Protein Species	Favored regions (%)	Residues in favored regions	Residues in allowed regions
1	*L. braziliensis*	90.8	354	34
2	*L. donovani*	90.2	351	36
3	*L. infantum*	90.5	352	36

4	*L. mexicana*	91.5	355	32
5	*L. major*	90.7	350	35

Table 5.5: List of GDH proteins obtained from Alphafold method

Sl. No.	GDH Protein Species	Alphafold ID
1	*L. braziliensis*	AF-A4HGZ4-F1-v4
2	*L. donovani*	AF-A0A3S7X1Z4-F1-v4
3	*L. infantum*	AF-A4I426-F1-v4
4	*L. mexicana*	AF-E9B0B3-F1-v4
5	*L. major*	AF-Q4Q7X1-F1-v4

Table 5.6: Percentage identity and RMSD of the GDH protein among five different *Leishmania* species.

Sl. No.	Alphafold GDH Protein	Identity (%)	RMSD (Å)
1	*L. donovani* + *L. infantum*	100.00	1.13
2	*L. donovani* + *L. major*	95.13	1.14
3	*L. donovani* + *L. mexicana*	94.46	1.14
4	*L. donovani* + *L. braziliensis*	88.94	1.25

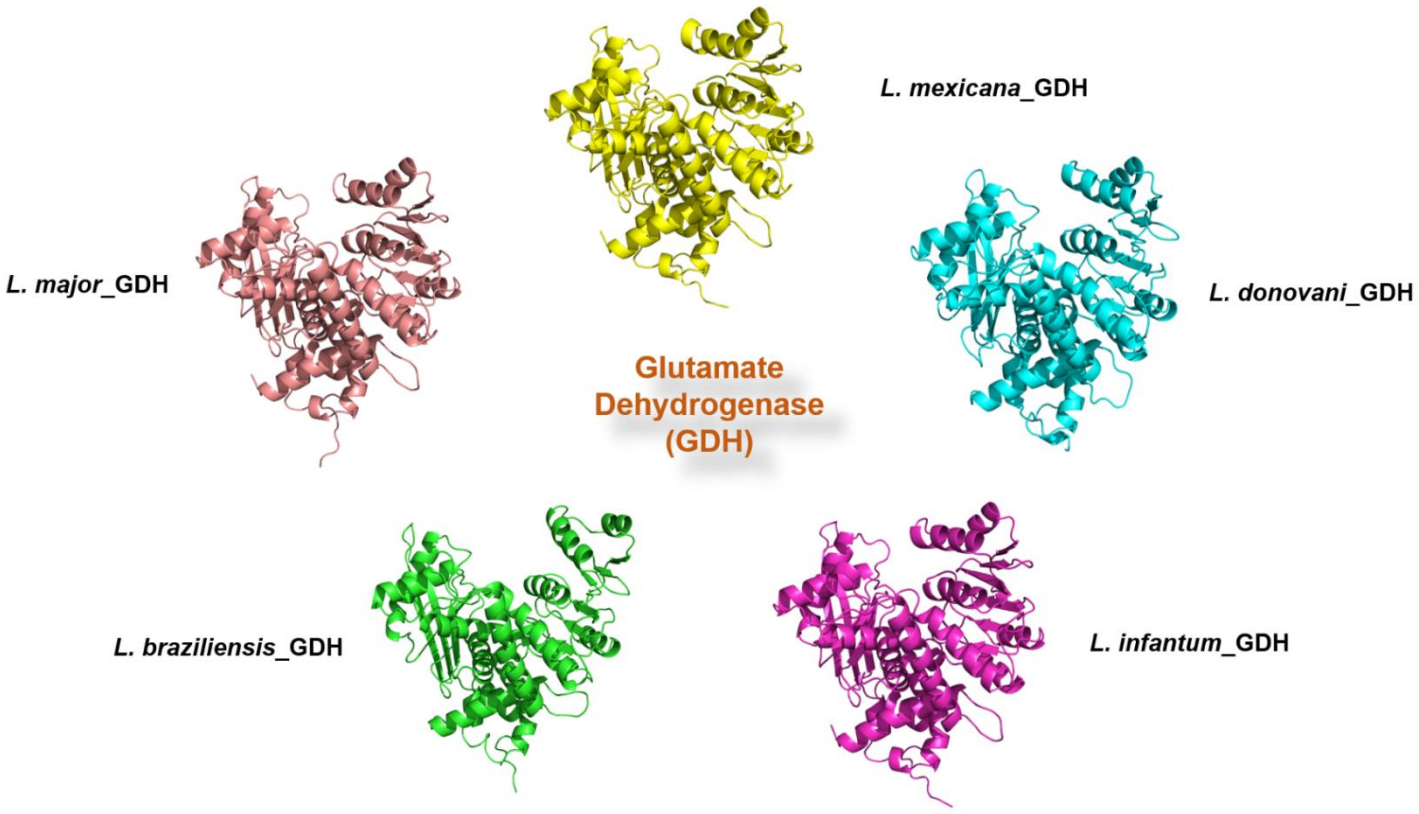

Figure 5.4: 3D structure of five GDH *Leishmania* species

5.4.5. Active site prediction of proteins and molecular docking analysis

In our investigation, we focused on selecting binding pockets with high area ($Å^2$) and volume ($Å^3$) in GDH for further analysis. The amino acids located on the active sites of GDH proteins in the five *Leishmania* species were prepared. Consequently, the pockets with elevated area and volume across GDH proteins in all five species were considered as potential binding sites. Table 5.7 provides information about ligand-binding active sites of the GDH protein based on CASTp predictions. These sites, integral in both structure and function, are specific regions on the protein surface where diverse compounds interact to produce desired effects. To assess the impact of ligands on GDH protein, we chose compounds (bithionol, GW5074, and hexachlorophene) previously reported to inhibit *Plasmodium falciparum* GDH protein in an *in vitro* experiment and the 2D structure of the ligands is shown in Figure 5.5. These compounds exhibited IC^{50} values of >20 µM, 18 µM, and 7.5 µM, respectively [18]. Molecular docking investigations employing AutoDock 4.2 were performed to anticipate binding affinities and interaction patterns between GDH proteins from five *Leishmania* species and chosen ligands. Table 5.8 displays the outcomes,

revealing that all GDH proteins displayed the least binding energy with bithionol and GW5074. Complexes with the lowest docking scores underwent additional analysis via MD simulations. Figures 5.6 and 5.7 depict the interactions of GDH with Bithionol and GW5074 structures.

Table 5.7: Ligand-binding active sites of the GDH protein based on CASTp predictions.

Sl. No.	Proteins	Area ($Å^2$)	Volume ($Å^3$)	Pocket amino acid residues
1	GDH *(L. braziliensis)*	733.09	1585.52	Gly98, Arg101, Leu114, Glu117, Lys121, Lys133, Pro140, Lys141, Asp145, Val148, Ala171, Gly172, Asp173, Ile174, Gly175, Val176, Gly177, Ser178, Arg179, Gly180, Gly201, Gly203, Leu204, Gln207, Glu210, Ile211, Arg212, Pro213, Glu214, Ala215, Thr216, Ser245, Gly246, Asn247, Val248, Gln250, Tyr251, Met254, Ser267, Asp268, Ser269, Met288, Arg289, Lys291, Asn292, Arg295, Gly296, Ala352, Asn353, Met354, Asn378, Gly381, Val382, Ile384, Ser385
2	GDH *(L. donovani)*	736.64	1413.96	Arg101, Phe102, His103, Pro104, Ile110, Phe113, Leu114, Asp139, Pro140, Lys143, Ser144, Asp145, Val148, Gly172, Asp173, Ile174, Gly175, Val176, Gly177, Ser178, Arg179, Glu180, Val181, Tyr183, Gly201, Lys202, Gly203, Leu204, Arg212, Pro213, Gly246, Asn247, Val248, Gly250, Tyr251, Met288, Lys291, Asn292, Arg295, Tyr301, Cys326, Ala327, Thr328, Gln329, Ala352, Asn353
3	GDH *(L. infantum)*	788.57	963.36	Gly99, Arg101, His103, Leu114, Gln118, Lys121, Lys133, Pro140, Lys141, Gly142, Lys143, Ala171, Gly172, Asp173, Ile174, Gly175, Val176, Gly177, Ser178, Arg179, Glu180, Arg212, Thr216, Gly244, Ser245, Gly246, Asn247, Val248, Asp268, Ser269, Lys270, Lys291, Asn292, Arg295, Gly296, Thr297, Cys326, Ala327, Thr328, Gln329, Gly351, Ala352, Asn353, Asn378, Gly381, Val382, Ser385
4	GDH *(L. mexicana)*	894.57	1191.34	Gly98, Gly99, Arg101, His103, Phe113, Leu114, Glu117, Gln118, Lys121, Lys133, Asp139, Pro140, Lys141, Gly142, Lys143, Ala171, Gly172, Asp173, Ile174, Gly175, Val176, Gly177, Gly178, Arg179, Glu180, Arg212, Thr216, Gly244, Ser245, Gly246, Asn247, Val248, Asp268, Ser269, Lys270, Lys291, Asn292, Arg295, Val296, Thr297, Cys326, Ala327,

				Thr328, Gln329, Gly351, Ala352, Asn353, Met354, Lys375, Asn378, Gly381, Val382, Ser385
5	GDH *(L. major)*	897.13	1211.26	Gly98, Gly99, Arg101, His103, Phe113, Leu114, Gln118, Lys121, Lys133, Asp139, Pro140, Lys141, Gly142, Lys143, Ala171, Gly172, Asp173, Ile174, Gly175, Val176, Gly177, Ser178, Arg179, Glu180, Arg212, Thr216, Gly244, Ser245, Gly246, Gly247, Val248, Asp268, Ser269, Lys270, Lys291, Asn292, Arg295, Gly296, Thr297, Cys326, Ala328, Gln329, Gly351, Ala352, Asn353, Asn378, Gly381, Val382, Ser385

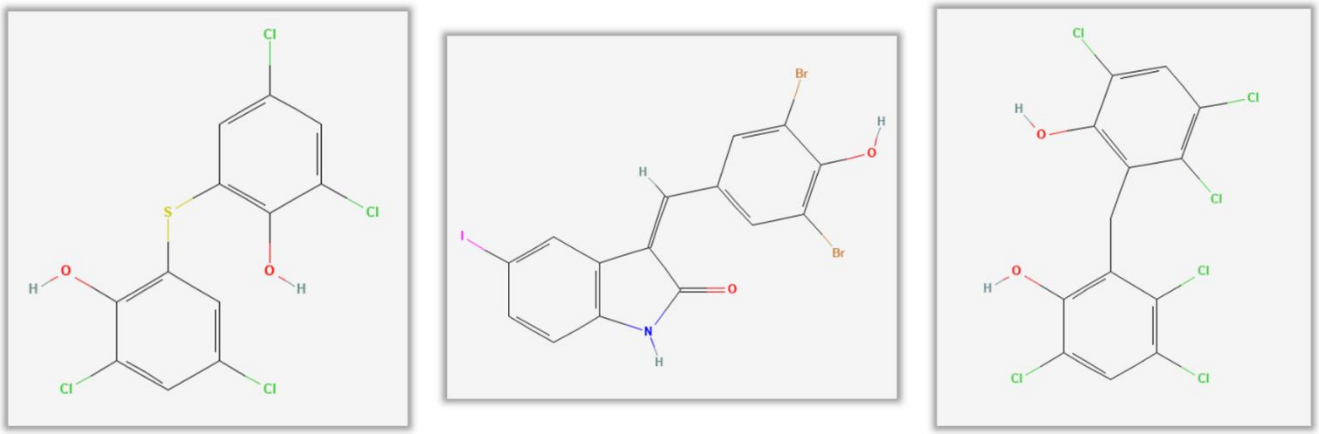

Figure 5.5: 2D structure of three ligands.

Table 5.8: Docking score between selected ligands and GDH of various *Leishmania* species

Ligand No.	Ligands	GDH (kcal/mol)				
		L. donovani	***L. braziliensis***	***L. infantum***	***L. major***	***L. mexicana***
1	Bithionol	-4.54	-4.34	-5.7	-4.36	-3.74
2	GW5074	-5.20	-5.08	-4.65	-4.78	-6.6
3	Hexachlorophene	-4.52	-4.11	-3.92	-3.68	-3.49

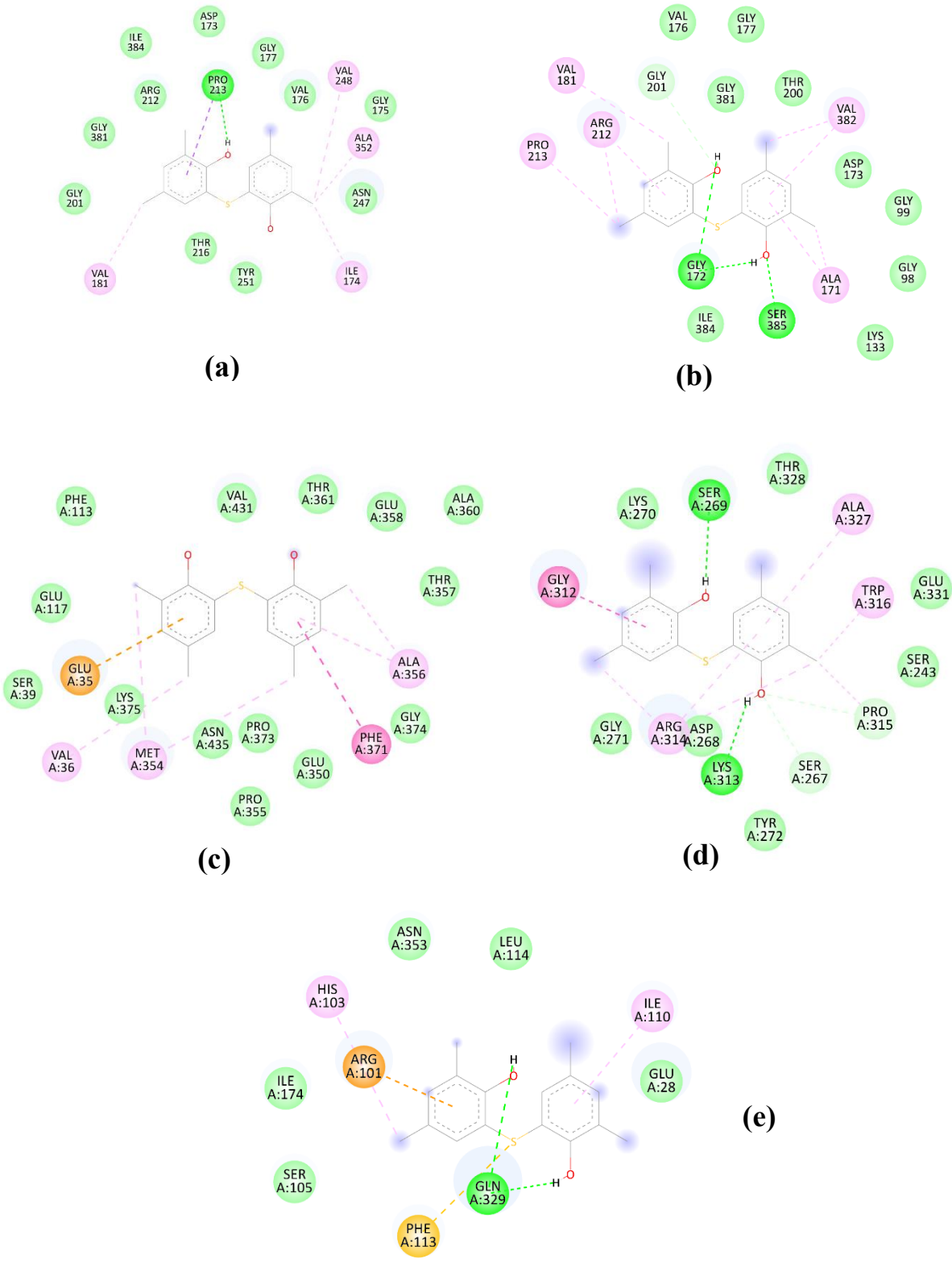

Figure 5.6: Docked structure of bithionol with GDH of *L. donovani* (a), *L. braziliensis* (b), *L. infantum* (c), *L. major* (d) and *L. mexicana* (e). The interactions present are van der Waals bonds (cyan), H-bonds (green), π-σ bonds (purple), and alkyl bonds (pink).

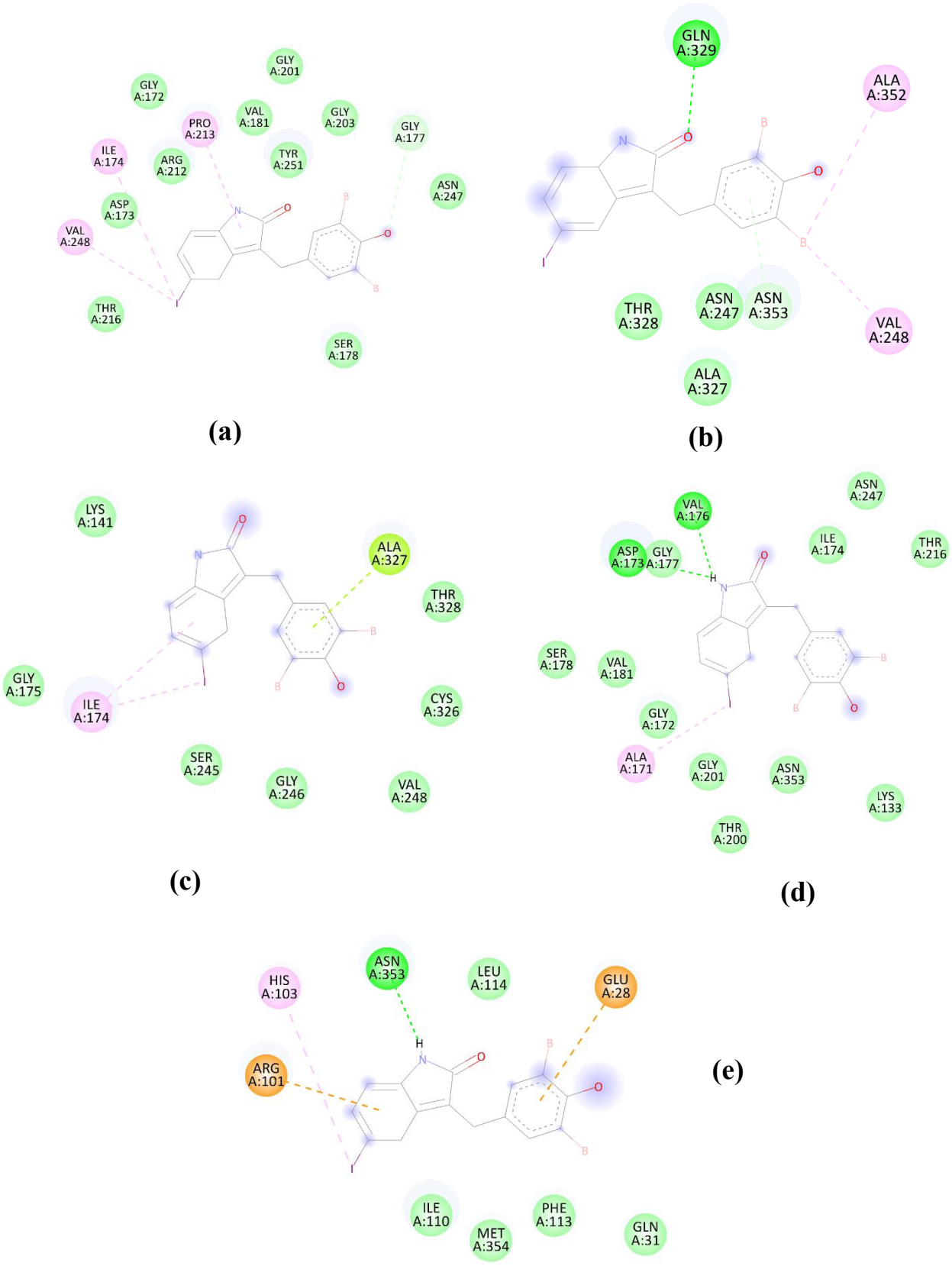

Figure 5.7: Docked structure of GW5074 with GDH of *L. donovani* (a), *L. braziliensis* (b), *L. infantum* (c), *L. major* (d) and *L. mexicana* (e). The interactions present are van der Waals bonds (cyan), H-bonds.

5.4.6. MD simulation analysis

MD simulation is a computational technique frequently used to determine the dynamics and interactions of diverse macromolecules. Additionally, it aids in exploring the behavior of small compounds binding to specific target proteins[34]. As GDH proteins of the five *Leishmania* species showed percentage identity above 88% among themselves providing the evidence that GDH can be considered as common target for the same. Moreover, on the basis of the phylogenetic tree analysis of GDH proteins of five *Leishmania* species helped us to select GDH of *L. donovani* and *L. infantum* which are more closely related whereas GDH of *L. donovani* and *L. braziliensis* are less closely related for MD simulations. Thus, observing their more closely and less closely related protein-ligand complexes behavior among themselves. The top two compounds (bithionol and GW5074) selected for MD simulation in complex with the GDH protein also possess potential antimalarial properties. Among them, only the complexes formed by bithionol and GW5074 with GDH of *L. donovani, L. braziliensis* and *L. infantum* were chosen for MD simulation along with control GDH proteins of the same for 100ns. Additionally, to scrutinize the dynamic characteristics of the system, various analyses, including root mean square deviations (RMSD), radius of gyration (Rg), and root mean square fluctuations (RMSF), Hydrogen bond analysis and MMPBSA analysis were conducted.

5.4.6.1. RMSD analysis

The RMSD plots for the GDH protein of *L. donovani*, *L. braziliensis* and *L. infantum* both in the presence and absence of bithionol and GW5074, depict the stability of the proteins during the simulations. In the control group, the *L. donovani* GDH exhibited higher fluctuations within the range of 0.3 to 0.4 nm, while the GDH of *L. braziliensis* showed fluctuations ranging from 0.4 to 0.5 nm. For the complexes, when *L. donovani* GDH bound with bithionol (Lig1), there was less fluctuation between 30-60 ns, but after 70-90 ns, comparative fluctuations were observed. In the case of GW5074 (Lig2), fluctuations were observed between 50-60 ns, followed by stabilization. For GDH of *L. braziliensis* with ligands, comparative fluctuations were observed between 35-65 ns, and then minor fluctuations persisted until the end of the simulation when GDH bound to bithionol. GW5074 showed fluctuations until 35 ns, followed by very minor fluctuations until the end of the simulation. In case of GDH of *L. infantum,* both ligands provide stability to the protein compared to the control. Moreover, GW5074 with GDH showed less fluctuations among the two complexes

of *L. infantum.* The 3D conformations of the complexes were extracted from the production run at 0 and 100 ns to visualize the orientation of both ligands within the binding site, maximizing their interaction with the GDH protein. RMSD analysis of GDH proteins are depicted in Figure 5.8.

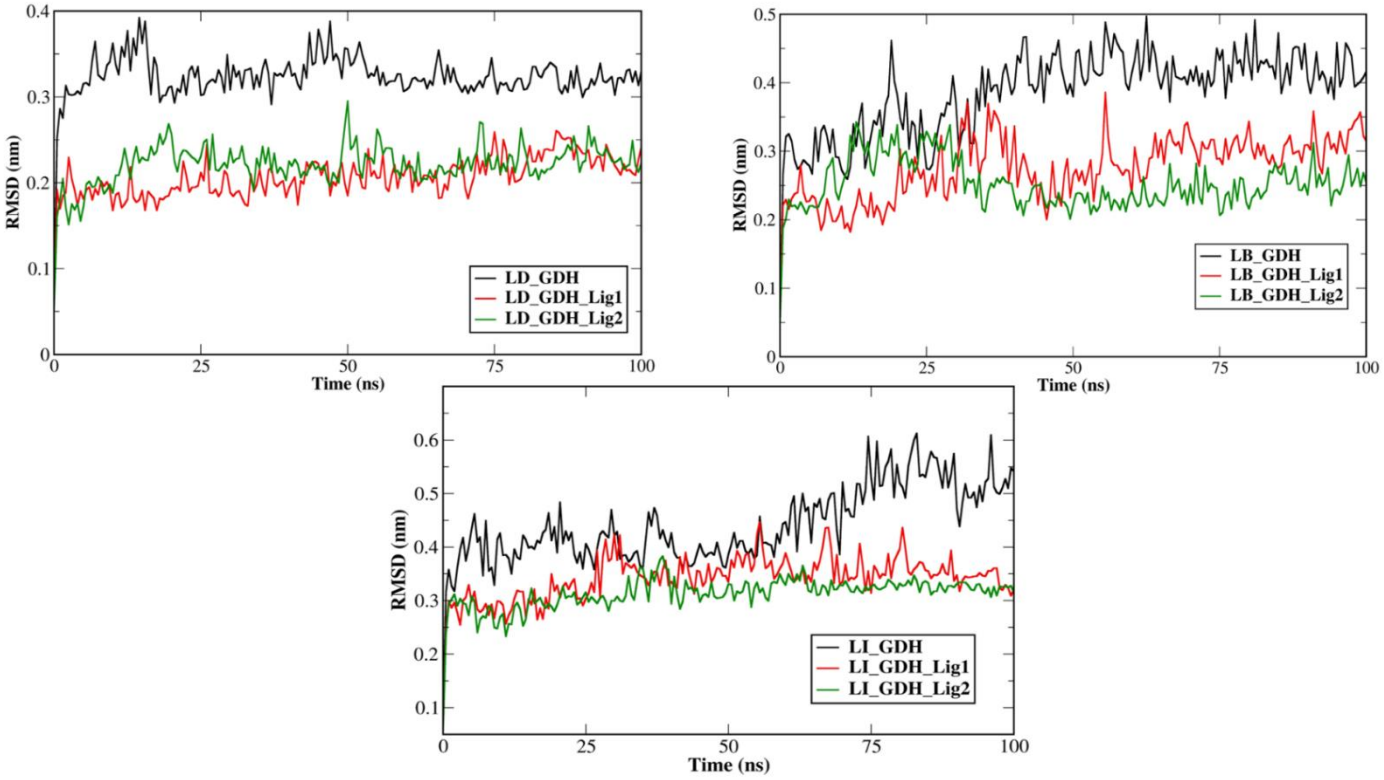

Figure 5.8: Comparison of average Root Mean Square Deviation (RMSD) of Glutamate dehydrogenase of *L. donovani, L. braziliensis* and *L. infantum.*

5.4.6.2. RMSF analysis

RMSF analysis was utilized to assess residue flexibility in GDH proteins under diverse conditions, including ligand-bound and ligand-free states. Comparable fluctuation patterns were observed in GDH systems of both *Leishmania* species between apo-GDH and GDH bound to various ligands. Minor average fluctuations were noted in ligand-bound cases for *L. donovani, L. braziliensis*, and *L. infantum*, with no significant variations compared to the control. Figure 5.9 illustrates RMSF analysis of GDH proteins.

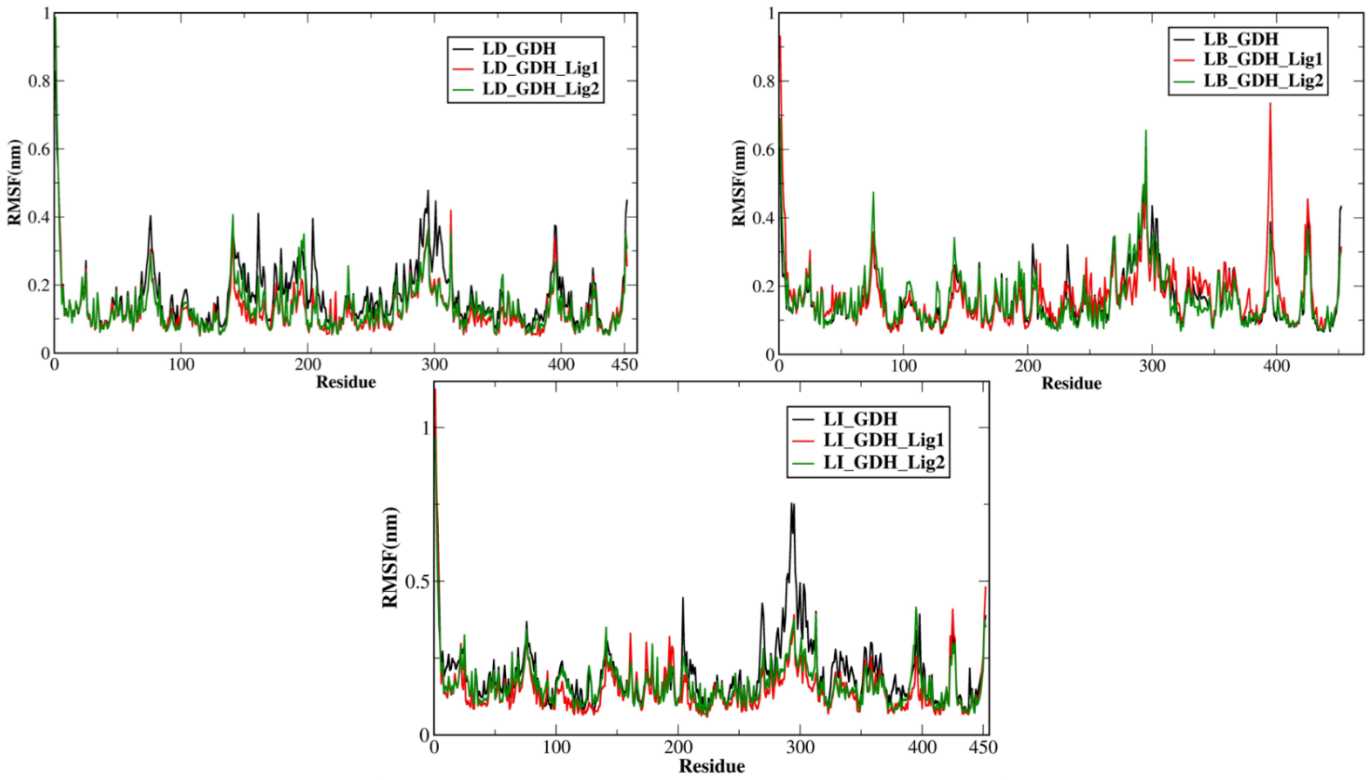

Figure 5.9: Comparison of average Root Mean Square Fluctuations (RMSF) of Glutamate dehydrogenase of *L. donovani, L. braziliensis* and *L. infantum.*

5.4.6.3. Rg analysis

The Rg analysis was employed to assess the compactness of the GDH proteins in the presence and absence of ligands during the simulation. In the case of the GDH system with *L. donovani*, the apo-GDH exhibited fluctuations up to 50 ns, followed by reduced fluctuations throughout the simulations. Contrarily, the Lig1 and Lig2 bound proteins showed lower fluctuations and greater compactness after 80 ns. Similarly, in the case of *L. braziliensis* and *L. infantum*, the GDH-apo demonstrated high fluctuations throughout the simulation. The GDH protein of *Leishmania* species bound to Lig1 exhibited substantial fluctuations, while Lig2 displayed stable behavior after 85 ns, indicating enhanced compactness of the structure due to the presence of ligands. Rg analysis of GDH proteins is depicted in Figure 5.10.

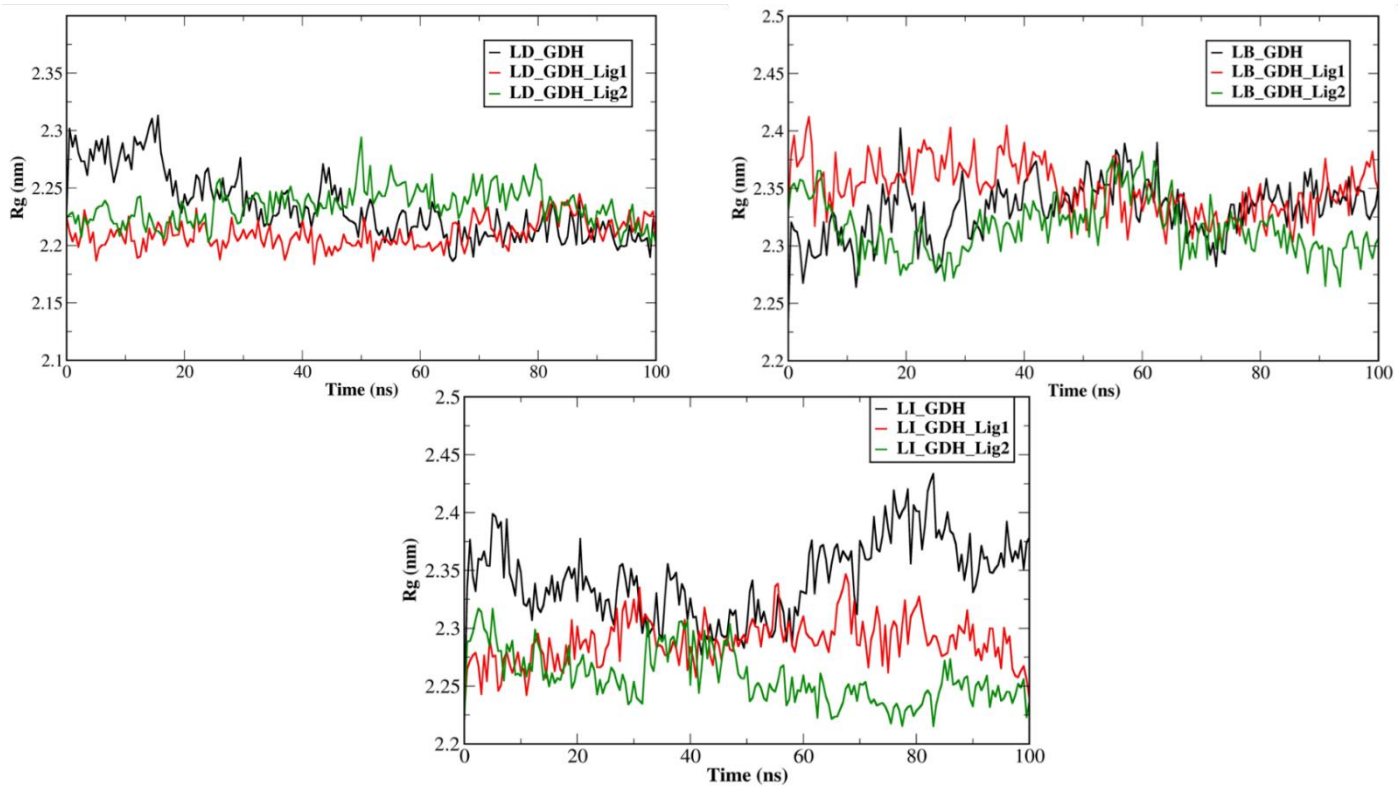

Figure 5.10: Comparison of average Radius of Gyration (Rg) of Glutamate dehydrogenase of *L. donovani, L. braziliensis* and *L. infantum.*

5.4.6.4. H-bond analysis

Throughout MD simulations, the calculation of hydrogen-bond interactions provides valuable insights into the quality of interactions between ligands and the GDH proteins, offering a measure of complex stability. The objective of this analysis is to assess interactions contributing to the stability of the complexes. The findings, depicted in the figure, show that when GDH of *L. donovani* binds to Lig1, it typically forms 1-2 hydrogen bonds, while Lig2 binding results in 2-3 hydrogen bonds. For *L. braziliensis*, GDH binding to Lig2 forms 1-3 hydrogen bonds, while Lig1 maintains a consistent single hydrogen bond. In *L. infantum*, Lig1 initially forms 2-3 hydrogen bonds during the simulation's first half, while Lig2 exhibits more hydrogen bonds in the latter half. Figure 5.11 illustrates the hydrogen bond analysis of GDH proteins with ligands.

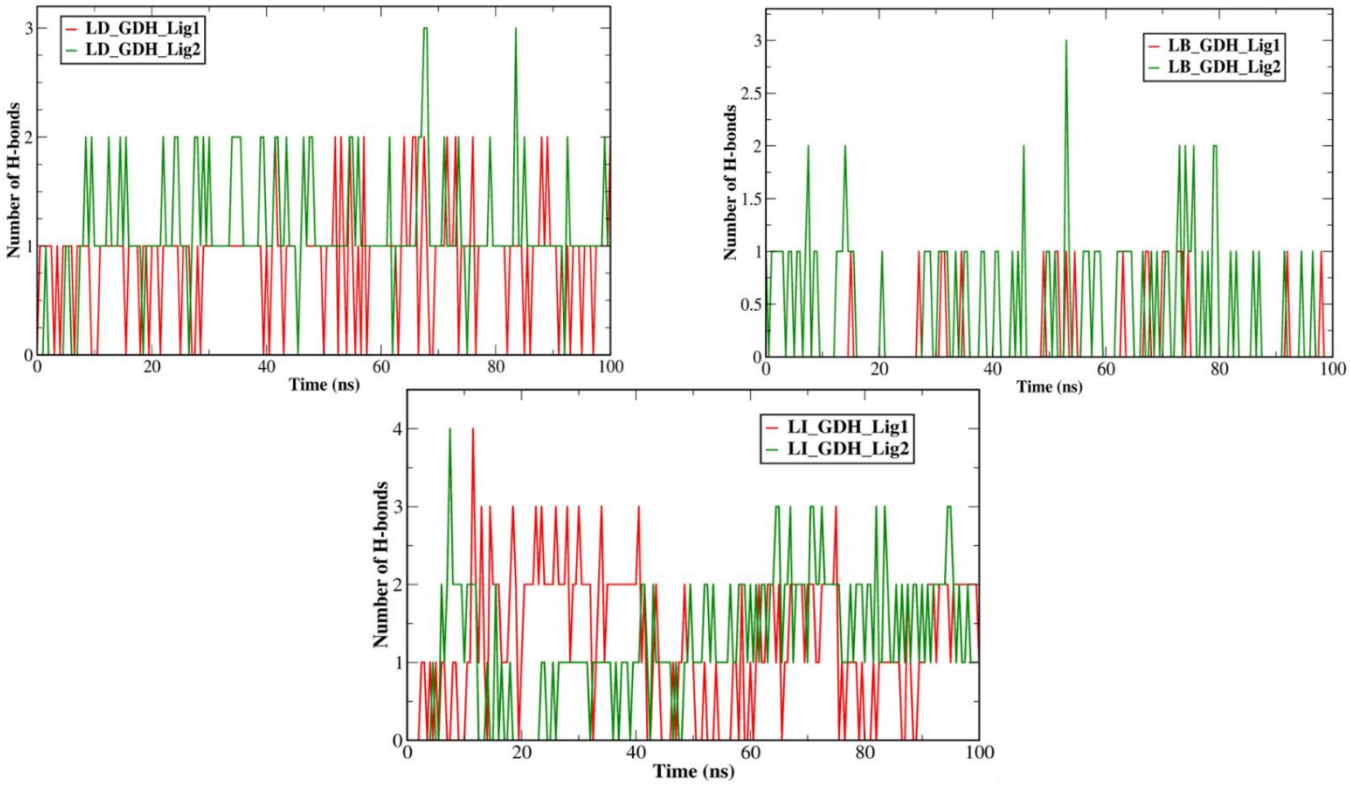

Figure 5.11: Comparison of average number of Hydrogen bonds of Glutamate dehydrogenase of *L. donovani, L. braziliensis* and *L. infantum.*

5.4.7. Interaction analysis

After simulations, we aimed to understand interactions between ligands and GDH proteins across five *Leishmania* species. Figures 5.12.1 and 5.12.2 depict two- and three-dimensional interaction profiles of Lig1 and Lig2 with GDH proteins from three species. Analysis revealed that Lig2 formed more significant interactions than Lig1. Table 5.9 summarizes simulation results, emphasizing predominant interactions like hydrophobic interactions, van der Waals forces, and hydrogen bonds between GDH proteins and ligands. These interactions underscored that Lig2 role in facilitating stable interactions with the active site residues of proteins.

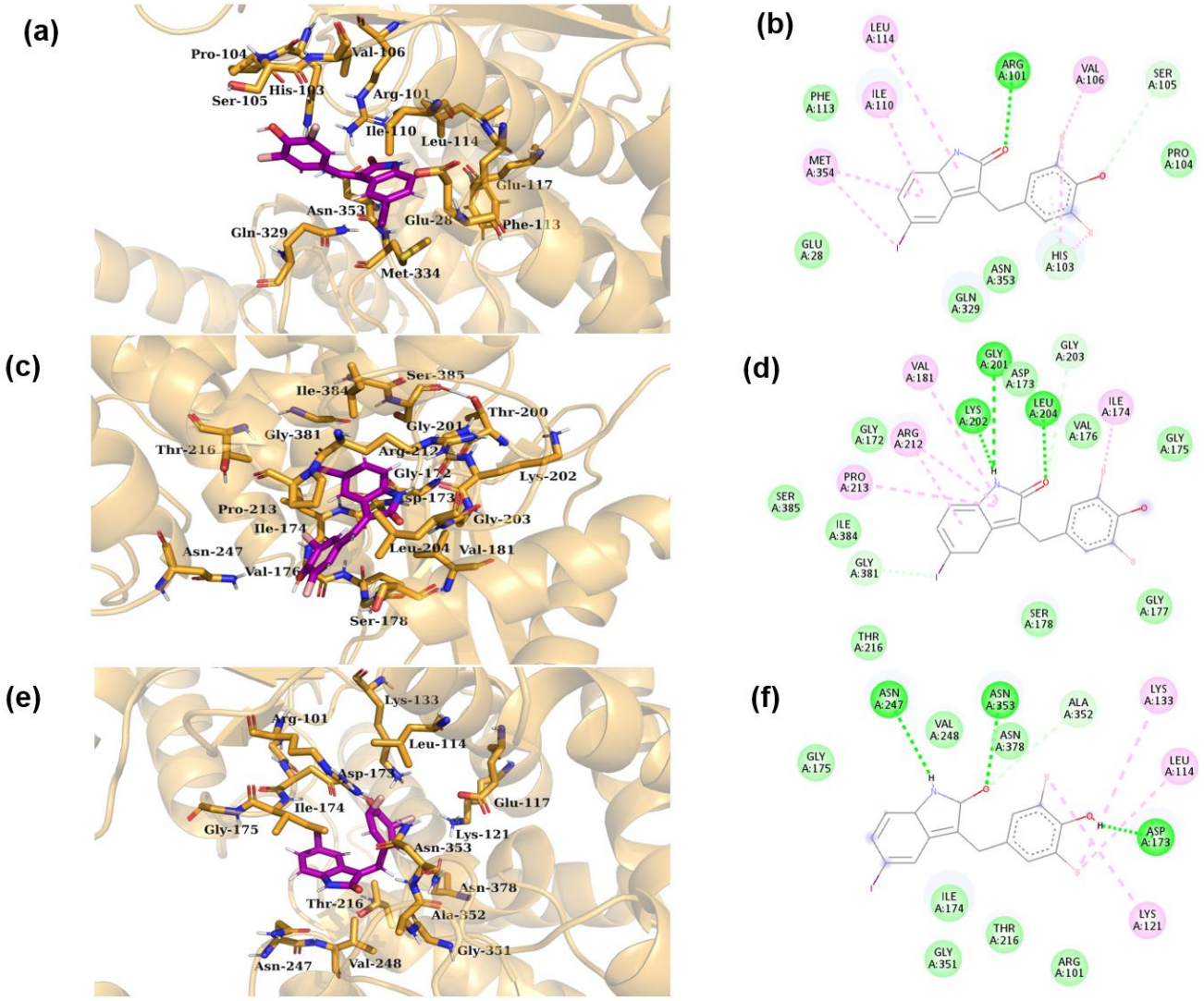

Figure 5.12.1: Interaction profile of GDH protein with Lig1. (a) GDH of *L. braziliensis* (3D visualization); (b) GDH of *L. braziliensis* (2D visualization); (c) GDH of *L. donovani* (3D visualization); (d) GDH of *L. donovani* (2D visualization); (e) GDH of *L. infantum* (3D visualization); (f) GDH of *L. infantum* (2D visualization).

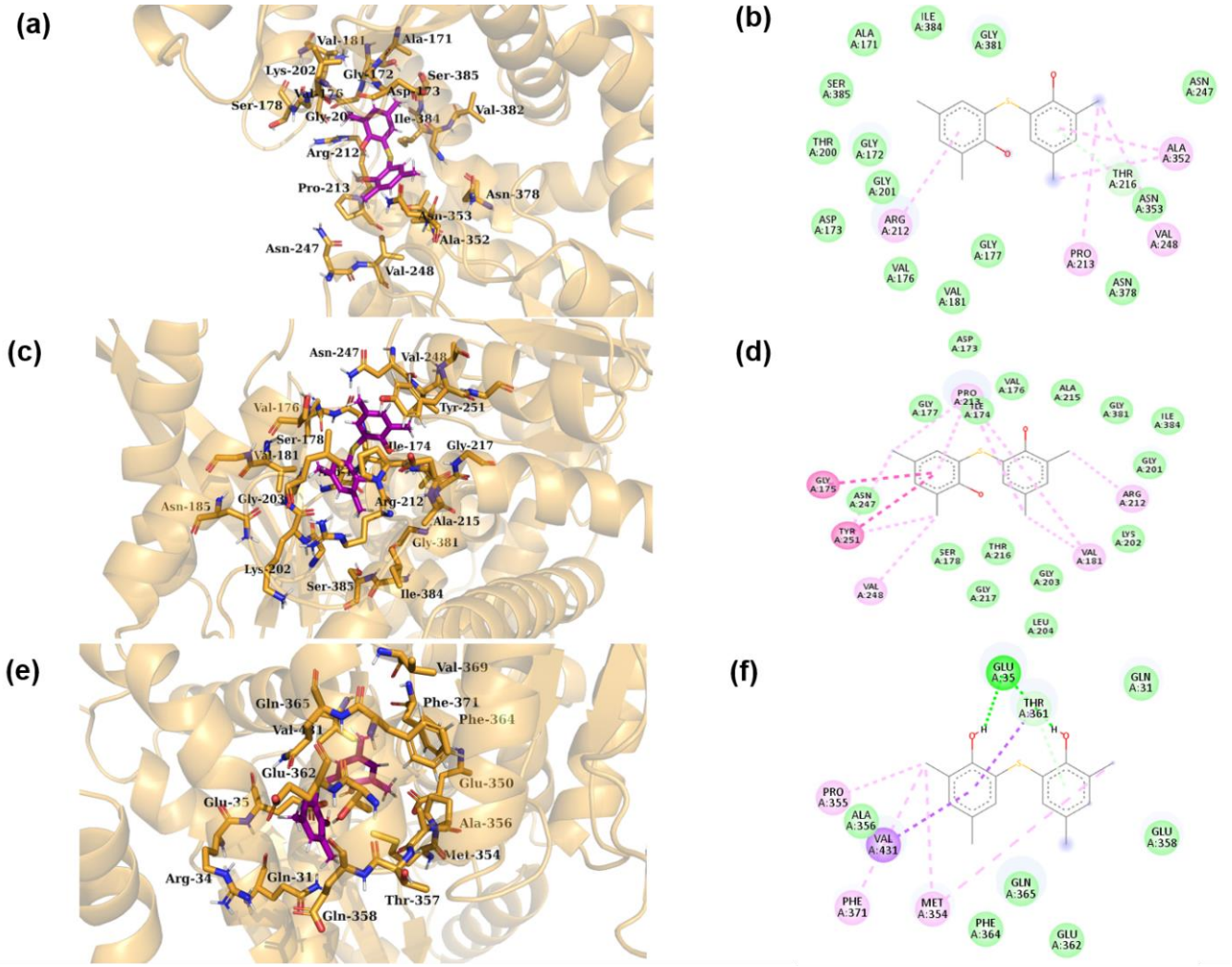

Figure 5.12.2: Interaction profile of GDH protein with Lig2. (a) GDH of *L. braziliensis* (3D visualization); (b) GDH of *L. braziliensis* (2D visualization); (c) GDH of *L. donovani* (3D visualization); (d) GDH of *L. donovani* (2D visualization); (e) GDH of *L. infantum* (3D visualization); (f) GDH of *L. infantum* (2D visualization).

Table 5.9: Numbers of interactions between GDH proteins and ligands during the course of the simulation

Protein–ligand complex	H-bonds	van der Waals bonds	Hydrophobic interactions
LB_GDH-Lig1	0	15	4
LB_GDH-Lig2	1	7	4
LD_GDH-Lig1	0	15	6
LD_GDH-Lig2	3	11	4
LI_GDH-Lig1	2	7	4
LI_GDH-Lig2	3	8	3

5.4.8. MMPBSA and binding free energy analysis

To elucidate the individual components contributing to the total binding energy between the compounds and GDH, an MM/PBSA analysis was conducted. The calculation involved the last 20 ns of the trajectory obtained from MD simulations [39]. Table 5.10 presents the distinct contributions from various components, including van der Waals energy (kJ/mol), electrostatic energy (kJ/mol), polar solvation energy (kJ/mol), and SASA energy (kJ/mol). For *L. braziliensis*, the binding energy of GDH-Lig1 and GDH-Lig2 is -84.85 and -101.82 kJ/mol, respectively. In the case of *L. donovani*, Lig1 with GDM exhibits a binding energy of -123.46 kJ/mol, while the GDH-Lig2 complex has a binding energy of -90.870 kJ/mol. When bound to Lig1, the GDH of *L. infantum* demonstrates a binding energy of -100.18 kJ/mol, while with Lig2, it exhibits a binding energy of -150 kJ/mol. High binding energy were observed between GDH of *L. braziliensis* and *L. infantum* when bind with Lig2 showed strong binding among them.

Table 5.9: MM/PBSA analysis of the bound complexes

Sl. No.	System	Binding energy (kJ/mol)	van der Waal energy (kJ/mol)	Electrostatic energy (kJ/mol)	Polar solvation Energy (kJ/mol)	SASA energy (kJ/mol)
1	LB_GDH-Lig1	-84.85 ± 13.09	-125.48 ± 9.92	-93.64 ± 17.18	148.69 ± 14.90	-14.42 ± 1.01
2	LB_GDH-Lig2	-101.82 ± 19.16	-96.30 ± 12.04	-78.86 ± 27.86	162.21 ± 22.12	-17.25 ± 0.99
3	LD_GDH-Lig1	-123.46 ± 14.52	-168.91 ± 9.04	-33.91 ± 11.11	96.95 ± 6.68	-17.59 ± 0.94
4	LD_GDH-Lig2	-90.87 ± 24.25	-152.92 ± 12.17	-140.48 ± 23.85	219.22 ± 9.53	-16.68 ± 1.00
5	LI_GDH-Lig1	-100.18 ± 21.96	-106.67 ± 11.41	-68.58 ± 27.07	90.14 ± 19.41	-15.07 ± 1.61
6	LI_GDH-Lig2	-150.06 ± 28.62	-63.56 ± 15.43	-488.91 ± 43.35	416.98 ± 26.56	-14.56 ± 0.90

5.5. Discussion

The integration of subtractive genomics and structure-based methodologies involved scrutinizing the proteomic data of *L. braziliensis*, *L. donovani*, *L. infantum*, *L. mexicana*, and *L. major*. The principal objective of this investigation was to identify novel treatment targets for various *Leishmania* species. To achieve this, a subtractive genomic strategy encompassing the entire proteome was employed, leveraging multiple online databases and

computational tools. Notably, GDH emerged as a novel drug target across the five *Leishmania* species examined in this study.

This approach aligns with methodologies observed in related studies, such as those investigating the identification of drug targets in *Micrococcus luteus* targeting Superoxide Dismutase (SOD) [35], *Acinetobacter baumannii* [14], and *Streptococcus pneumoniae* [36]. In these studies, protein sequences were identified as paralogous sequences based on a cut-off at 40% sequence identity. It's worth noting that the selection of this cut-off value is not universal and may vary between 40% and 80%, depending on various factors. The choice of the cut-off is influenced by considerations such as the dataset under examination, the number of sequences, the diversity of species involved, and the intended application [35]. The proteome sequences of *Leishmania* species were specifically retrieved from the National Center for Biotechnology Information (NCBI) database, underscoring the scientific rigor and reliance on a reputable source for genomic data in this study.

By employing BLASTp analyses with specific filtering parameters across five *Leishmania* species, we can deduce that the quantity of orthologous protein groups and non-homologous proteins is intricately tied to the diversity present within the protein sequences of the screened organism. The identification of non-homologous target proteins, particularly those distinct from the human proteome, holds the potential to unveil novel protein or enzyme targets within pathogens.

A refined investigation, involving the focused analysis of shared proteins and rigorous druggability assessments, is crucial in the application of filtering parameters for identifying potential drug targets. Through this process, we isolated five proteins through druggability tests, and from this set, one protein demonstrated commonality across all five *Leishmania* species. This meticulous approach aids in the identification of a novel essential target for *Leishmania* species, showcasing the potential for therapeutic interventions in the context of these pathogens.

The identification of the Glutamate Dehydrogenase (GDH) protein involved a comprehensive examination of its functions, sequence evaluation, phylogenetic analysis and structural considerations to yield substantive insights. The primary function of GDH in *Leishmania* species centers on facilitating the conversion of glutamic acid into α-ketoglutarate (α-KG), a crucial step leading to the tricarboxylic acid (TCA) cycle. This metabolic role underscores the significance of GDH in the cellular processes of *Leishmania*.

Subsequent to functional elucidation, Multiple Sequence Alignment (MSA) of GDH proteins across the five species revealed a high degree of identity among the protein sequences, indicative of their conserved nature. Phylogenetic tree analysis further demonstrated the evolutionary origin of GDH proteins and highlighted their close relationships across the Leishmania species. The conservation of both function and sequence strengthens the argument for considering GDH as a potential drug target.

To delve into the structural aspects, protein structures were obtained using the alpha fold method, and their validity was confirmed through Ramachandran plot analysis. This step ensured the reliability and accuracy of the predicted protein structures. Moreover, ligands with known antimalarial activity on GDH, namely Bithionol (Lig1), GW5074 (Lig2), and Hexachlorophene (Lig3), were collected and subjected to screening using the AutoDock tool. The detailed investigation of the binding site of GDH and the interactions with these ligands provided valuable insights. Notably, the lowest docking scores for Bithionol and GW5074 with GDH across the five species were crucial in selecting complexes for subsequent MD simulation analysis.

In the MD simulation analysis, Root Mean Square Deviation (RMSD) values were employed to assess the structural stability of the Glutamate Dehydrogenase (GDH) proteins of *L. donovani* and *L. braziliensis* in complex with GW5074 over the course of the simulation. The RMSD values indicated that both *L. donovani* and *L. braziliensis* and *L. infantum.* GDH proteins, when bound to GW5074, reached convergence within the simulation after 80 nanoseconds. This convergence suggests a stable and well-equilibrated system, providing evidence for a strong and sustained interaction between GW5074 and the GDH proteins of these *Leishmania* species. The observed stability supports the notion that GW5074 exhibits a high affinity for the GDH protein in these organisms. Furthermore, Root Mean Square Fluctuation (RMSF) analysis was conducted to examine the flexibility of the protein structure throughout the simulation, particularly in regions where ligands (Bithionol and GW5074) were bound. The results of RMSF analysis demonstrated that the binding of Bithionol and GW5074 to the GDH protein did not induce significant changes in the protein structure. The stable association observed in the RMSF analysis indicates that the ligands formed a persistent and well-tolerated interaction with the protein, suggesting that these ligands do not disrupt the overall stability of the GDH protein structure.

In the context of molecular dynamics simulations, the Radius of Gyration (Rg) values, representing the compactness of the system, exhibited consistent fluctuations within the range of 2.2 to 2.4 nanometers across all systems involving Glutamate Dehydrogenase (GDH) proteins of the two *Leishmania* species. Notably, among the six systems studied, those involving the binding of GW5074 to GDH proteins displayed less pronounced fluctuations, ultimately leading to a more compact conformation at the conclusion of the production run. The analysis of the total number of hydrogen bond contacts throughout the simulation duration revealed that GW5074 binding to GDH proteins formed a comparatively higher number of hydrogen bonds than Bithionol binding to the proteins. This observation underscores the significance of hydrogen bonds in stabilizing the interactions between ligands and the GDH protein. Examination of the interactions between GDH proteins and ligands indicated the presence of various contacts, including hydrogen bonds, van der Waals bonds, and hydrophobic interactions, which collectively contribute to their interaction. Notably, GW5074 exhibited a higher number of hydrogen bonds with the GDH protein compared to other types of interactions. Consequently, our analysis suggests that GW5074 has a propensity to occupy the binding pocket of the protein more prominently.

To gain further insights into the strength and stability of biomolecular interactions, Molecular Mechanics/Poisson-Boltzmann Surface Area (MMPBSA) analysis was conducted. This analysis provides the binding energy of complexes, offering a quantitative measure of the stability of ligand-protein interactions. The results indicated that GW5074, when bound to GDH proteins, exhibited superior binding energy compared to other ligand-bound complexes. This suggests that the binding of GW5074 is energetically more favorable, supporting its role as a potent ligand for the GDH proteins in the context of potential drug development.

Bibliography

[1] Sasidharan, S. and Saudagar, P. Leishmaniasis: where are we and where are we heading?. *Parasitology research*, 120(5):1541-1554, 2021.

[2] Croft, S. L. and Coombs, G. H. Leishmaniasis–current chemotherapy and recent advances in the search for novel drugs. *Trends in parasitology*, 19(11):502-508, 2003.

[3] Costa, C. H., Chang, K. P., Costa, D. L. and Cunha, F. V. M. From infection to death: An overview of the pathogenesis of visceral leishmaniasis. *Pathogens*, 12(7):969, 2023.

[4] Retrived on 20 October, 2023 from https://www.who.int/news-room/fact-sheets/detail/leishmaniasis.

[5] Bora, N. and Nath Jha, A. An integrative approach using systems biology, mutational analysis with molecular dynamics simulation to challenge the functionality of a target protein. *Chemical biology & drug design*, 93(6):1050-1060, 2019.

[6] Schlein, Y. Leishmania and sandflies: interactions in the life cycle and transmission. *Parasitology Today*, 9(7):255-258, 1993.

[7] Tashboltayevna, A. S., Mirzaali o‘g‘li, A. J. and Dilshod o‘g‘li, O. R. LEISHMANIOSIS DISEASE, ITS SYMPTOMS, PRIMARY CONSEQUENCES AND DISTRIBUTION. *Galaxy International Interdisciplinary Research Journal*, 10(12):836-838, 2022.

[8] Rajkhowa, S., Hazarika, Z. and Jha, A. N. Systems biology and bioinformatics approaches in leishmaniasis. In *Applications of Nanobiotechnology for Neglected Tropical Diseases* (pp. 509-548). Academic Press, 2021.

[9] Stockdale, L. and Newton, R. A review of preventative methods against human leishmaniasis infection. *PLoS neglected tropical diseases*, 7(6):e2278, 2013.

[10] Wijnant, G. J., Dumetz, F., Dirkx, L., Bulté, D., Cuypers, B., Van Bocxlaer, K. and Hendrickx, S. Tackling drug resistance and other causes of treatment failure in leishmaniasis. *Frontiers in Tropical Diseases*, 3:837460, 2022.

[11] Saha, D., Borah, N. J. and Jha, A. N. Molecular scaffold recognition of drug molecules against essential genes of Leishmania donovani using biocomputing approach. *South African Journal of Botany*, 162:52-63, 2023.

[12] Shahid, F., Shehroz, M., Zaheer, T. and Ali, A. Subtractive Genomics Approaches: Towards Anti-Bacterial Drug Discovery. *Front. Anti-Infect. Drug Discov*, 8:144-158, 2020.

[13] Khan, M. T., Mahmud, A., Iqbal, A., Hoque, S. F. and Hasan, M. Subtractive genomics approach towards the identification of novel therapeutic targets against human Bartonella bacilliformis. *Informatics in Medicine Unlocked*, 20:100385, 2020.

[14] Kaur, H., Kalia, M. and Taneja, N. Identification of novel non-homologous drug targets against Acinetobacter baumannii using subtractive genomics and comparative metabolic pathway analysis. *Microbial Pathogenesis*, 152:104608, 2021.

[15] Ashraf, B., Atiq, N., Khan, K., Wadood, A. and Uddin, R. Subtractive genomics profiling for potential drug targets identification against Moraxella catarrhalis. *Plos one*, 17(8): e0273252, 2022.

[16] Indari, O., Singh, A. K., Tiwari, D., Jha, H. C. and Jha, A. N. Deciphering antiviral efficacy of malaria box compounds against malaria exacerbating viral pathogens-Epstein Barr virus and SARS-CoV-2, an in silico study. *Medicine in Drug Discovery*, 16:100146, 2022.

[17] Rather, M. A., Saha, D., Bhuyan, S., Jha, A. N. and Mandal, M. Quorum quenching: a drug discovery approach against Pseudomonas aeruginosa. *Microbiological Research*, 264: 127173, 2022.

[18] Zocher, K., Fritz-Wolf, K., Kehr, S., Fischer, M., Rahlfs, S. and Becker, K. Biochemical and structural characterization of Plasmodium falciparum glutamate dehydrogenase 2. *Molecular and biochemical parasitology*, 183(1):52-62, 2012.

[19] Pathan, E. K., Kulkarni, A. M., Prasanna, N. V., Ramana, C. V. and Deshpande, M. V. NADP-dependent glutamate dehydrogenases in a dimorphic zygomycete Benjaminiella poitrasii: Purification, characterization and their evaluation as an antifungal drug target. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1864(11):129696, 2020.

[20] Godsora, B. K. J., Prakash, P., Punekar, N. S. and Bhaumik, P. Molecular insights into the inhibition of glutamate dehydrogenase by the dicarboxylic acid metabolites. *Proteins: Structure, Function, and Bioinformatics*, 90(3):810-823, 2022.

[21] Ferreira, C., Mesquita, I., Barbosa, A. M., Osório, N. S., Torrado, E., Beauparlant, C. J., ... and Silvestre, R. Glutamine supplementation improves the efficacy of miltefosine treatment for visceral leishmaniasis. *PLoS neglected tropical diseases*, 14(3):e0008125, 2020.

[22] Lyndem, S., Hazarika, U., Bhatta, A., Prakash, V., Jha, A. N. and Roy, A. S. In vitro interactions of esculin and esculetin with bovine hemoglobin alter its structure and

inhibit aggregation: insights from spectroscopic and computational studies. *New Journal of Chemistry*, 47(30):14447-14468, 2023.

[23] Pruitt, K. D., Tatusova, T. and Maglott, D. R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic acids research*, 35(suppl_1): D61-D65, 2007.

[24] Huang, Y., Niu, B., Gao, Y., Fu, L. and Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, 26(5):680-682, 2010.

[25] Emms, D. M.and Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20:1-14, 2019.

[26] Prado, L. C. D. S., Giacchetto Felice, A., Rodrigues, T. C. V., Tiwari, S., Andrade, B. S., Kato, R. B., ... and Soares, S. D. C. New putative therapeutic targets against Serratia marcescens using reverse vaccinology and subtractive genomics. *Journal of Biomolecular Structure and Dynamics*, 40(20):10106-10121, 2022.

[27] Bagewadi, Z. K., Aakanksha, U. K., Yaraguppi, D. A., Yunus Khan, T. M., Deshpande, S. H., Dammalli, M., ... and Hiremath, S. V. Molecular docking and simulation studies against nucleoside diphosphate kinase (NDK) of Pseudomonas aeruginosa with secondary metabolite identified by genome mining from paenibacillusehimensis. *Journal of Biomolecular Structure and Dynamics*, 41(22):12610-12619, 2023.

[28] Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., ... and Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research*, *46*(D1):D1074-D1082, 2018.

[29] Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., ... and Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, 50(D1):D439-D444, 2022.

[30] Procter, J. B., Carstairs, G. M., Soares, B., Mourão, K., Ofoegbu, T. C., Barton, D., ... and Barton, G. J. *Alignment of biological sequences with Jalview* (pp. 203-224). Springer US, 2021.

[31] Tamura, K., Dudley, J., Nei, M. and Kumar, S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular biology and evolution*, 24(8):1596-1599, 2007.

[32] Tian, W., Chen, C., Lei, X., Zhao, J. and Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids research*, 46(W1):W363-W367, 2018.

[33] Saha, D. and Nath Jha, A. Computational multi-target approach to target essential enzymes of Leishmania donovani using comparative molecular dynamic simulations and MMPBSA analysis. *Phytochemical Analysis*, 34(7):842-854, 2023.

[34] Quraishi, S., Saha, D., Kumari, K., Jha, A. N. and Roy, A. S. Non-covalent binding interaction of bioactive coumarin esculetin with calf thymus DNA and yeast transfer RNA: A detailed investigation to decipher the binding affinities, binding location, interacting forces and structural alterations at a molecular level. *International Journal of Biological Macromolecules*, 257:128568, 2024.

[35] Bagewadi, Z. K., Khan, T. Y., Gangadharappa, B., Kamalapurkar, A., Shamsudeen, S. M. and Yaraguppi, D. A. Molecular dynamics and simulation analysis against superoxide dismutase (SOD) target of Micrococcus luteus with secondary metabolites from Bacillus licheniformis recognized by genome mining approach. *Saudi Journal of Biological Sciences*, 30(9):103753, 2023.

[36] Khan, K., Jalal, K., Khan, A., Al-Harrasi, A. and Uddin, R. Comparative metabolic pathways analysis and subtractive genomics profiling to prioritize potential drug targets against streptococcus pneumoniae. *Frontiers in Microbiology*, 12:796363, 2022.

[37] Marcili, A., Sperança, M. A., da Costa, A. P., Madeira, M. D. F., Soares, H. S., de OCC Sanches, C., ... and Gennari, S. M. Phylogenetic relationships of Leishmania species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of Leishmania (L.) infantum chagasi in South America. *Infection, Genetics and Evolution*, 25: 44-51, 2014.

[38] Hollingsworth, S. A. and Karplus, P. A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins, 2010.

[39] Jakhmola, S., Hazarika, Z., Jha, A. N. and Jha, H. C. In silico analysis of antiviral phytochemicals efficacy against Epstein–Barr virus glycoprotein H. *Journal of Biomolecular Structure and Dynamics*, 40(12):5372-5385, 2022.