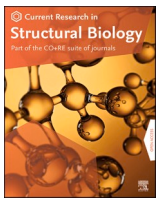
Current Research in Structural Biology 7 (2024) 100127

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology



Research Article

Computational identification of candidate inhibitors for Dihydrofolate reductase in *Acinetobacter baumannii*

Saurabh Kumar Bhati, Monika Jain, Jayaraman Muthukumaran, Amit Kumar Singh *

Department of Biotechnology, Sharda School of Engineering and Technology, Sharda University, Greater Noida, India

ARTICLE INFO

Handling Editor: Dr A Wlodawer

Keywords:
DHFR enzyme
Acinetobacter baumannii
Molecular dynamics simulation
Virtual high throughput screening
MM/PBSA

ABSTRACT

Acinetobacter baumannii is one of the emerging causes of hospital acquired infections and this bacterium, due to multi-drug resistant and Extensive Drug resistant has been able to develop resistance against the antimicrobial agents that are being used to eliminate it. *A.baumannii* has been the cause of death in immune compromised patients in hospitals. Hence it is the urgent need of time to find potential inhibitors for this bacterium to cease its virulence and affect its survival inside host organisms. The Dihydrofolate reductase enzyme, which is an important biocatalyst in the conversion of Dihydrofolate to Tetrahydrofolate, is an important drug target protein. In the present study high throughput screening is used to identify the inhibitors of this enzyme. The prioritized ligand molecular candidates identified through virtual screening for the substrate binding site of the predicted model are Z1447621107, Z2604448220 and Z1830442365. The Molecular Dynamics Simulation study suggests that potential inhibitor of the Dihydrofolate reductase enzyme would prevent bacteria from completing its life cycle, affecting its survival. Finally the complexes were analysed for binding free energy of the Dihydrofolate reductase enzyme complexes with the ligands.

1. Introduction

One of the main causes of nosocomial infections is the gram-negative bacterium *Acinetobacter baumannii*, which is a serious risk to immunocompromised patients who are admitted to intensive care units and who are on ventilator support. This bacterium has the potential to spread several diseases like pneumonia, meningitis, soft tissue infections, and urinary tract infections through blood. It is a multidrug-resistant (MDR) bacterium due to its high level of genetic flexibility and ability to produce resistant biofilm (Jha et al., 2022). The bacterium develops resistance towards the antimicrobial agents mainly by three mechanisms, it has certain enzymes causing the inactivation of antibiotics, second due to restriction of entry of antimicrobial agents to the target site within the cell and third overcoming the effect of antibiotics by altering the functional metabolic pathway with other one (Skariyachan et al., 2020). This bacterium has a propensity to thrive on any surface, both biotic and abiotic, and shows fastidious approach to its growth (Harding et al., 2018).

A.baumannii has been regarded as one of the biggest hazards to human health by the majority of organisations working in the field of disease control and management. It has been categorised as a serious health hazard by WHO (https://www.who.int/) (Skariyachan et al., 2020). The WHO's (https://www.who.int/) Global Priority List of Antibiotic-Resistant Bacteria, published in 2017, recommends urgent need of new antibiotics for the Carbapenem resistant *A. baumannii* and has been categorised as CRITICAL among the three categories as of CRITICAL, HIGH and MEDIUM. In Intensive Care Units, polymyxin is still a successful treatment for CRAB infections but there have also been reports of *A. baumannii* developing polymyxin resistance (Jiang et al., 2022). As per the 2022 report from NCDC-India (National Center for Disease Control) (https://ncdc.mohfw.gov.in/), data concerning 9531 isolates of *Acinetobacter* spp. Was made available through network sites, with 8840 of these isolates originating from distinct patients. Most of *Acinetobacter* spp. isolates deposited to NARS-Net (3659) were from blood cultures. Among these blood isolates, there was a 56% isolates showing resistance against imipenem. For all types of specimens containing *Acinetobacter* spp., the highest level of resistance was observed in the case of Ceftazidime. Colistin-resistant isolates were detected in two instances of urine, the first case of blood, and the second instance of pleural fluid. In the US and Europe, *A. baumannii* causes 2% of hospital acquired diseases, while in the Middle East, this percentage is almost twice as high.

One can conduct research on proteins and enzymes associated with

* Corresponding author.
E-mail address: amitk.singh@sharda.ac.in (A.K. Singh).

https://doi.org/10.1016/j.crstbi.2024.100127
Received 27 October 2023; Received in revised form 5 January 2024; Accepted 8 January 2024
Available online 17 January 2024

Abbreviations

ADME	Absorption, Distribution, Metabolism and Excretion
MM/PBSA	Molecular mechanics/Poisson–Boltzmann surface area
MDR	Multi Drug Resistant
XDR	Extensive Drug Resistant
vHTS	virtual high throughput screening
DHF	Dihydrofolate
THF	Tetrahydrofolate
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuations
Rg	Radius of Gyration
SASA	Solvent Accessible Surface Area
H-Bond	Intermolecular Hydrogen Bonding
PCA	Principal Component Analysis

virulence in order to discover potential inhibitors. The growth of bacteria can be inhibited by blocking essential enzymes with suitable competitive inhibitors.

Folic acid biosynthesis is one such metabolic pathway catalysed by various enzymes (Diaset al., 2018). Folate is a cofactor made up of glutamate, pterin, and para-amino benzoic acid. Folic acid is an important component as it is an important component in the production of amino acids like glycine, serine, and methionine and also in the biosynthesis of purines and thymidine (Korduset al., 2019). Animals fulfill their folic acid requirements by dietary supplements. Dihydrofolate reductase (DHFR) enzyme catalyses the conversion of DHF to THF by using $NADPH_2$ as a cofactor. The crucial fact is that both prokaryotic and eukaryotic organisms contain this enzyme including humans, but bacterial enzymes are non-homologous to humans. The substrate binding grooves are different in DHFR enzyme of *A. baumannii* and *Homo sapiens* (Franklin et al., 2015). Many bacteria have been targeted using this enzyme as an antimicrobial agent against human diseases (Dias et al., 2018). In some of the bacteria, such as *Coxiellaburnetti*, the structure of DHFR has already been established (Franklin et al., 2015). The X-ray crystal structure of the DHFR enzyme from *Streptococcus pneumonia* has been determined (Lee et al., 2010). The structural information about the DHFR enzyme from *Coxiellaburnetti*, as obtained from PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) (Laskowski et al., 1993), shows that DHFR enzyme comprises two sheets, one beta alpha unit, one beta-hairpin, five beta bulges, ten strands, four helices, two helix-helix interactions, sixteen beta turns, and three gamma turns. Thus DHFR enzyme is an important enzyme required for the survival of *A. baumannii*. In this current investigation, we delve into the DHFR enzyme as a drug target protein for combating this bacterium. We examine both the DHFR enzyme in its unbound state and when it is bound to ligands to gain insights into their interactions, binding affinities, and their influence on the structural dynamics of DHFR. This study aims to identify ligand molecules which tends to spontaneously interact with the active site of the DHFR enzyme, potentially serving as inhibitors and thereby impeding the bacterium's survival.

2. Methodology

2.1. Enzyme model prediction and analysis

The amino acid sequence with ID A0A0R4J8G5, of DHFR's enzyme was fetched from Uniprot datatbase (https://www.uniprot.org/). Protein BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was employed to ascertain similarities with human proteins (Johnson et al., 2008). The physicochemical characteristics of DHFR protein were computed and calculated utilising the Protparam server (https://web.expasy.org/protparam/). This polypeptide sequence was utilised in the prediction of the three-dimensional structure of DHFRenzyme using the Robetta server (https://robetta.bakerlab.org/) (Baek et al., 2021). Refinement of model was done by using the YASARA server (Krieger, Koraimann and Vriendet, 2002). The model was validated using the ProSAWeb server (https://prosa.services.came.sbg.ac.at/prosa.php) (Wiederstein and Sippl, 2007) and the Structural Analysis and Verification Server (SAVES 6.0) (https://saves.mbi.ucla.edu/).

2.2. Virtual screening of ligands and pharmacokinetics

The University of Texas in Austin hosts the Texas Advanced Center's portal (TACC) (drug.discovery.tacc.utexas.edu), where virtual screening was carried out. The screening process involved the utilization of the Enamine High Throughput Screening Core (HTSC) Library, comprising a total of 2,141,514 compounds. This library was employed to identify potential inhibitor candidates for the DHFR protein, employing the Auto Dock Vina software (Morris et al., 2009). The screening was limited to the DHFR's active site. Using literature analysis of the *Coxiellaburnetti,* DHFR protein template model (PDB ID-3TQ8), the substrate binding groove of the protein was reported, and the residues in the substrate binding site were analysed, identified and superimposition was used to validate the presence of residues in the target protein (Franklin et al., 2015). For Docking analysis, the following center coordinates: center_x = 12.5640 Å, center_y = 21.9998 Å, center_z = -14.5053 Å, and size dimensions: size_x = 25.0000 Å, size_y = 25.0000 Å, size_z = 25.0000 Å, a grid box encompassing predicted active site residues was produced. The exhaustiveness was set to 10. DataWarrior (https://openmolecules.org/datawarrior/) (Lopez, Naveja and Franco, 2019), was used to screen molecules against Lipinski's rule of five (Chen et al., 2020), toxicological properties, as well as other parameters like carcinogenic impact, mutagenesis ability, effects on reproductive health, and potential to induce irritation. These ligands were further subjected to ADME screening by using the SwissADME server (http://www.swissadme.ch/) (Daina et al., 2017). Intermolecular interactions were examined in the protein ligand complexes and the best three shortlisted ligands with the highest binding efficiency and stable interaction were chosen for further studies.

2.3. Molecular dynamics simulations

DHFR in unbound and ligand bound state was put to 100ns molecular dynamics simulation for which GROMACS 2021; Van Der Spoel et al., 2005) was utilised. This was done to explore the behaviour of the DHFR enzyme during interacting with shortlisted ligand molecules in a physiologically relevant environment. The topology of ligand molecules was created by using the PRODRG online server (https://link.springer.com/) (AW Schuttelkopfet al., 2004), while the structure of the enzyme was parameterized with the aid of the GROMOS 54A7 forcefield with using *gmx pdb2gmx* command (Silva et al., 2018). GROMOS 54A7 forcefield is well recognised for stabilising practically all sections of proteins, including α and βhelices (Huang, Lin and Gunstreren, 2011). To solvate the DHFR enzyme system, the SPC/E (Simple Point Charge/Extended) explicit water model (J Yin and Landau, 2011) was used. The best representation of bulk dynamics and structure comes from the SPC/E water model (Mark and Nilson, 2001). The SPC/E is a three-point interaction water model and is economical with respect to computational studies (J Yin and Landau, 2011). The SPC/E water model provides the highest viscosity and a good structure of liquid physiological model in comparison to experimental liquidation (Mark and Nilsson, 2001). The steepest descent method was employed to reduce the energy of DHFR and DHFR enzyme-ligand combinations, aiming for a tolerance level of 100 kJ/mol, while conducting a maximum of 50,000 iterations (Spoel et al., 2005). In conjunction with a reduction in restrictions, the steepest descent algorithm is capable of decreasing the energy levels of

both proteins and complexes of enzyme and ligand molecules, transitioning them from a highly distorted and energetically unfavourable state to a more stable, minimal-energy state across a wide range of conditions. For free DHFR and DHFR-ligand complexes, two equilibration phases were carried out. The initial phase, at constant volume and temperature equilibration (NVT), was carried out for 1 ns, and the second phase, at constant pressure and temperature equilibration (NPT), was carried out for 1ns since the system was already equilibrated for 1 ns? Following this, 100ns MD simulations with 2 fs time steps were performed and the resulting trajectories were considered for more studies. The simulation trajectories were examined by built-in GROMACS 2021 commands (McGibbonet al., 2015). Several structural characteristics and global dynamics studies, including RMSD, RMSF, Rg, SASA, H-bond and essential dynamics such as PCA, revealed the stability and dynamic behaviour of protein-ligand complexes. To graphically represent the various structural analysis, Qtgrace (https://sourceforge.net/projects/qtgrace/) was utilised.

2.4. MM/PBSA based binding free energy analysis

The free energy of DHFR enzyme complexes during the binding of ligand with enzyme binding groove, was assessed by *g_mmpbsa* tool (Kumari et al., 2014), which employs the MM/PBSA (Molecular mechanics/Poisson-Boltzmann surface area) methodology. To compute binding free energy, the following equation was employed:

$$\Delta G_{binding} = \Delta E_{ele} + \Delta E_{vdw} + \Delta G_{pol} + \Delta G_{np}$$

$\Delta G_{binding}$ signifies binding free energy of DHFR complexes in the above equation whereas, the ΔE_{vdw},ΔE_{ele},ΔG_{pol} and ΔG_{np}, represent the varying van der Waals energy, electrostatic energy, polar solvation energy, and non-polar solvation energy, respectively. The *g_mmpbsa* program's built-in SASA non-polar model was used to determine the contribution of non-polar solvation energy. The solute and solvent dielectric constants used to calculate various energy change factors were 2 and 80, respectively. At every 20 ps, 1000 frames from the final 20ns convergent MD trajectories from 80ns to 100ns were considered for computing binding free energy.

3. Results

3.1. Enzyme model prediction and analysis

Using Protein BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), the DHFR enzyme was found non-homologous to human proteins. ProtParam server (https://web.expasy.org/protparam/) gave the basic physiochemical properties as molecular weight of 18.9 kDa, pI of 5.88, and instability index score of 29.61 which is less than 40 making it a stable protein. The substrate binding site of the DHFR enzyme was identified with interacting residues such as Arg 50, Thr57, Leu60, Ile 106, Ile107, Gly 108, and Gly 109 (Franklin et al., 2015). The Robetta web server (https://robetta.bakerlab.org/)gave the predicted 3D model. This was followed by refinement by energy minimization of the 3D model by employing the YASARA server. The refined DHFR enzyme model was examined and analysed for its credibility as three dimensional model. ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) and SAVES 6.0 (https://saves.mbi.ucla.edu/) were used to validate the 3D model (Fig. 1.). Proteins are analysed by the SAVES 6.0 server using programmes like ERRAT (Colovos and Yeates, 1993), Verify 3D (Eisenberget al., 1997) PROCHECK (Laskowski et al., 1993). The output from the SAVES 6.0 server is as shown in Table 1. The Z score was −8.18, and −8.02 for pre and post energy minimization models of DHFR enzyme, respectively as reported by ProSA-web Server (https://prosa.services.came.sbg.ac.at/prosa.php)which also depicted the clashing interaction within the protein structure (see Fig. 2).

3.2. Virtual Screening of ligand and pharmacokinetics

The virtual screening through the TACC server yielded 1000 compounds with an estimated binding energy range from −13.2 kcal/mol to −10.5 kcal/mol, all targeting the active site of the DHFR enzyme. To refine the selection, 157 molecules were shortlisted out of the initial 1000 based on parameters detailed in Table 4 and Table 5 by DataWarrior tool (https://openmolecules.org/datawarrior/). Subsequently, an ADME analysis was conducted on the 157 compounds using the SwissADME server (http://www.swissadme.ch/). This resulted in a list of 09 molecules at the end. Finally 3 molecules (with the highest binding energy and good bonding with the active site of enzyme) (details given in Table 2) were selected based on their capability to cross the BBB, had adequate GI absorption, high solubility in hydrophilic environments,

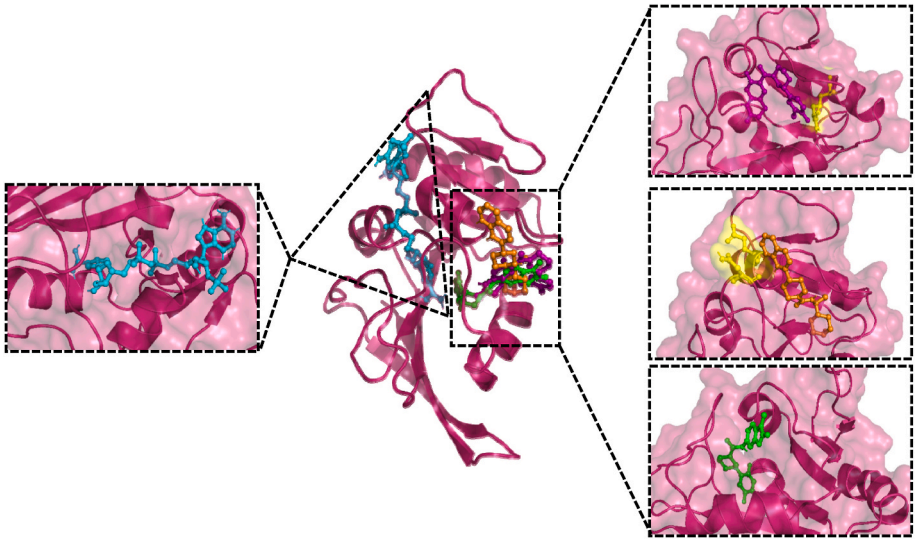


Fig. 1. Three dimensional representation of DHFR complex with Z1447621107 (Purple), Z2604448220 (Orange), Z1830442365 (Green). Zoomed cleft represent the active site of DHFR enzyme. Blue colour compound represent prosthetic group $NADPH_2$ which reduces Dihydrofolate to Tetrahydrofolate. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

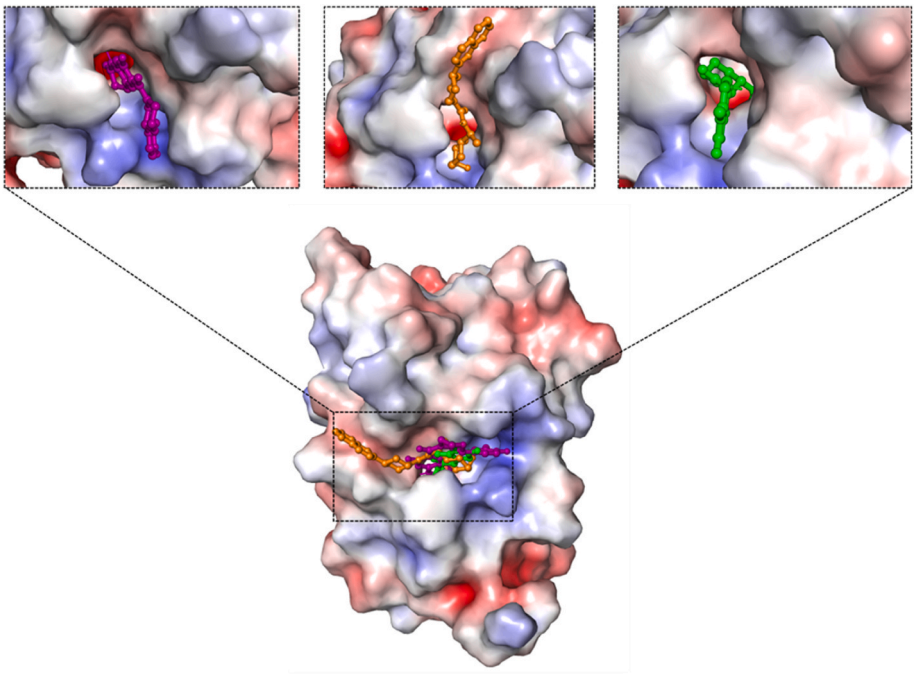


Fig. 2. Electrostatic surface potential maps of DHFR enzyme along with bound ligands in active site cleft. Z1447621107 (Purple), Z2604448220 (Orange), Z1830442365 (Green). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
DHFR enzyme model pre and post energy minimization analysis by Saves 6.0 server.

S. No	DHFR Model	ERRAT	VERIFY 3D	PROCHECK			
				Residues in most favored regions	Residues in additional permitted regions	Residues in generously permitted regions	Residues in disallowed regions
1	Initial model	89.44%	Pass (100%)	141 (93.4%)	9 (6.0%)	1 (0.7%)	0 (0.0%)
2	Refined model	92.54%	Pass (100%)	141 (93.4%)	10 (6.6%)	0 (0.0%)	0 (0.0%)

Table 2
Top three ligands with details of Chemical formula, AutoDock estimated binding affinity, Molecular weight; AutoDock estimated inhibition constant and 2-D Structure.

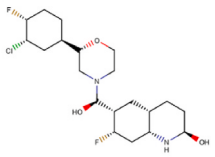
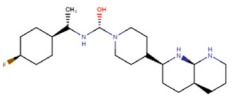
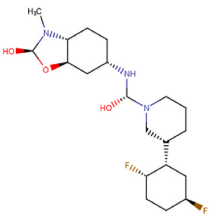
S. No.	Selected ligands	Molecular formula	Molecular weight (g/mol)	AutodockEstimated binding affinity (kcal/mol)	Autodock inhibition constant (nano molar)	2D structure
1	Z1447621107	C20H17ClF2N2O3	406.8	−8.1	1154.99	
2	Z2604448220	C22H29FN4O	384.5	−6.9	8753.00	
3	Z1830442365	C20H19F2N3O3	387.4	−8.4	695.99	

Table 3
Top three ligands with ADME profiles.

S.no	Ligands	GI	BBB Permeability	cLogP	TPSA (Å^2)	ESOL	CYP Inhibitor	Pains Alert
1	Z1447621107	High	Yes	2.53	64.96	Soluble	No	0
2	Z2604448220	High	Yes	2.81	59.56	Soluble	No	0
3	Z1830442365	High	Yes	1.99	68.2	Soluble	No	0

Table 4
Toxicological properties of top three ligands.

S.no	Ligand	Carcinogenicity	Mutagenicity	Reproducibility	Irritability
1	Z1447621107	None	None	None	None
2	Z2604448220	None	None	None	None
3	Z1830442365	None	None	None	None

Table 5
Lipinski's rule of five depicting physiochemical properties of selected three ligands.

S.no	Ligand	rBonds	HBA	HBD	DL
1	Z1447621107	3	7	3	2.06
2	Z2604448220	5	6	4	3.95
3	Z1830442365	4	8	3	3.06

cLogP: Consensus octanol-water partition coefficient, rBonds: Rotatable Bonds, HBD: Hydrogen Bond donor, HBA: Hydrogen Bond acceptor, DL: Drug likeness, GI: Gastrointestinal absorption, BBB: Blood Brain Barrier, CYP: Cytochrome, TPSA:Topological Surface Area.

with cLogP values less than 5 and TPSA values less than 100 Å^2, not-inhibiting CYP enzymes and exhibiting no PAINS alert, solubility as shown in Table 3. These were used to do further prioritization by molecular dynamics simulation.

The bonding interaction of the DHFR enzyme with its ligands shows the various bonds formed during interaction which stabilises the complex. The interactions are depicted in Figs. 2 and 3. The ligand Z1447621107 formed a hydrogen bond with Arg63, van der Waal bonds with residues like Phe37, Leu34, Leu26 and Arg58. It also interacts by forming Pi-Pi interactions with the residues like Pro61 (4.49, 6.12), Lys38 (5.62), Leu60 (6.40), Trp28 (6.69), and Met56 (4.98, 5.39). DHFR and ligand Z2604448220 interact by forming three Hydrogen bonds with residues Glu54, Asn24 (5.39,4.17) (two bonds were formed). The complex of DHFR enzyme and ligand Z2604448220 forms Pi-Pi interaction with residues Ala25 (4.72, 6.05), Leu60 (4.51, 5.28), Trp28 (6.67) and Lys38 (7.07) while van der Waal bonds are formed with residues Gly57, Leu26, Pro61, Met56, Ser55, Phe37 and Leu34. The DHFR and ligand Z1830442365 complex does not show any Hydrogen bond interaction but the ligand interact with good number of Pi-Pi interactions with residues Pro61 (5.43), Met56 (5.54), Trp28 (7.12), leu60 (5.48), Asp33 (5.28), Ile107 (4.60), Phe37 (4.83), Leu34 (6.06), Leu26 (3.38) and Pro27 (4.74). The complex of DHFR enzyme and ligand Z1830442365 also formed van der Waal bonds with residues Lys38, Val11, Ile30 and Pro27. Summarised details of bonds along with distance are shown as in Table 6.

3.3. Molecular dynamics simulation

Molecular dynamics simulations were employed to investigate the atomic alterations of entire macromolecules under physiological conditions over a specific time frame. This approach also enables the assessment of various aspects of protein-ligand interactions, such as their strength, stability, and interaction patterns. Additionally, it allows for the elucidation of conformational changes that macromolecules undergo when exposed to a hydrophilic biological environment. Several structural characteristics, including RMSD, RMSF, Rg, SASA, intermolecular hydrogen bonds, and principal component analysis (PCA), were scrutinized.

3.3.1. Root mean square deviation

RMSD, a crucial mechanism, can be employed to compute the structural stability of proteins and protein-ligand complexes over time. In unbound or native state DHFR enzyme, attains stability at around 3ns shows stable condition with minor fluctuations for the first 20 ns. Further, the deviation slightly increases but again remains stable till the end of simulation period with a peak around 56ns with 0.34 nm deviation. RMSD for DHFR enzyme and the ligand Z1447621107 complex is distorted the one with no stable state noticed. For the first 7ns the deviation increases followed by a gradual decrease till 15 ns. Then there is a continuous increase till 80ns after which an abrupt spike is seen with maximum deviation (0.49 nm) of the entire simulation period. Deviation, then slightly decreases till last point of tracjectory. The DHFR enzyme-Z2604448220 complex shows increase in deviation till 35 ns, and then stable deviation which are of low magnitude are seen with minor fluctuations till the end of the simulation, thus the complex is stable after 35 ns. For DHFR enzyme-Z1830442365 complex, RMSD value first increases till 25ns followed by an immediate decrease and again increases till 50 ns. Following 50 ns, the complex demonstrates stability up to the 90 ns mark, but during the last 10 ns, deviations exhibit random fluctuations (as depicted in Fig. 4a). The mean RMSD values for the unbound DHFR, DHFR-Z1447621107, DHFR-Z2604448220, and DHFR-Z1830442365 are 0.27 nm, 0.35 nm, 0.47 nm, and 0.27 nm, respectively.

3.3.2. Root mean square fluctuation

RMSF quantifies typical variations experienced by a protein molecule within its natural environment to comprehend the elastic nature of C-alpha atoms of amino acid units of both unbound proteins and protein-ligand complexes. DHFR enzyme in unbound state shows noticeable fluctuations in the residues located in the loop region. Residues such as Met1, Trp28, Glu 100, Glu156 shows peak in the graph, are all located in the loop region of enzyme among which Met1 show the maximum fluctuation of 0.56 nm. In DHFR enzyme-Z1447621107 complex, the residues such as Met1, Ile 72, Glu156, Lys169 shows noticeable fluctuations and all the residues are located in the loop regions of enzyme where Met1 again shows the maximum fluctuation of 0.19 nm in the bound and unbound DHFR enzyme. For DHFR enzyme-Z2604448220 complex fewer fluctuations are noticed and only Met1 and Lys169 (1.02 nm) show some noticeable motions as both the residues are found in the secondary loop regions of enzyme and Lys169 shows maximum fluctuation, whereas rest other residues shows fluctuations less than 0.40 nm. DHFR enzyme-Z1830442365 complex also mimic second complex (DHFR enzyme-Z2604448220) in showing fewer noticeable fluctuations among which Met1, Trp 74 are prominent and both the fluctuating residues are found within loop regions of enzyme, where Met1 again fluctuates to maximum RMSF of 0.67 nm as shown in Fig. 4b. The average RMSF values for unbound DHFR, DHFR- Z1447621107,

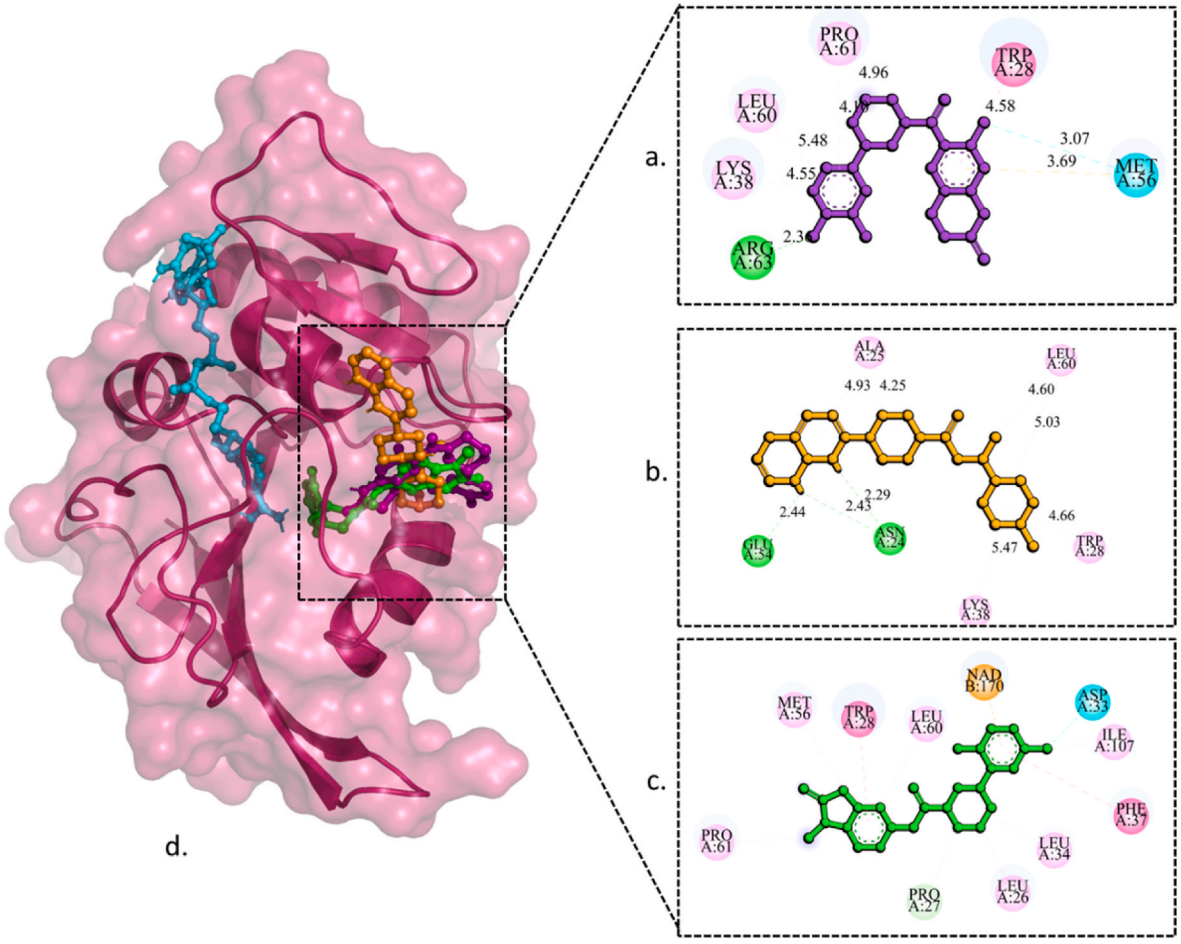


Fig. 3. (a) A 2-D depiction of the DHFR enzyme complex with the ligand Z1447621107 (in purple). **(b)** The engagement of active site residues of the DHFR enzyme with the ligand Z2604448220 (in orange). **(c)** The interaction between the active site residues of the DHFR enzyme and the ligand Z1830442365 (in green). **(d)** A three-dimensional representation showcasing the interactions of ligands with the active site of the DHFR enzyme. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 6
Different interaction between DHFR and Ligand Z1447621107, Ligand Z2604448220 and Ligand Z1830442365.

S.no	Ligands	HB	D (Å)	Pi-SR	D (Å)	vdWISR
1	Z1447621107	Arg63	5.88	Trp28	6.69	Phe37
				Leu60	6.40	Leu34
				Lys38	5.62	Arg58
				Met56	4.98, 5.39	Leu26
				Pro61	4.49, 6.12	
2	Z2604448220	Asn24	5.39, 4.17	Ala25	4.72, 6.05	Gly57
		Glu54	5.58	Leu60	4.51, 5.28	Leu26
				Trp28	6.67	Pro61
				Lys38	7.07	Met56
						Ser55
						Phe37
						Leu34
3	Z1830442365	–	–	Pro61	5.43	Lys38
				Met56	5.54	Val11
				Trp28	7.12	Ile30
				Leu60	5.48	Pro27
				Asp33	5.28	
				Ile107	4.60	
				Phe37	4.83	
				Leu34	6.06	
				Leu26	3.38	
				Pro27	4.74	

(HB: Hydrogen Bond, D: Distance, Pi-SR: Pi-Interaction, and vdWISR: van der Waals Interaction).

DHFR- Z2604448220, DHFR- Z1830442365 are 0.14 nm, 0.19 nm, 0.20 nm, 0.12 nm respectively.

3.3.3. *radius of gyration*

For understanding stable shape, density and folding behaviour of proteins and their interactions with ligands, it is crucial to examine a structural metric known as the radius of gyration (Rg). This particular parameter holds significant importance as it helps determine whether the binding of a ligand molecule contributes to the reinforcement of the protein's structure.The graph shows that unbound DHFR enzyme shows stable radius of gyration with occasional peaks among which one noticeable spike of 1.62 nm is obtained around 99 ns. The unbound enzyme remains stable till 55ns and then radius of gyration shows a slight spike. After that radius of gyration increases from 57ns to 80ns and for the last 20ns, the radius of compaction is unstable as it does not hold a uniform trajectory. For the complex of DHFR enzyme and ligand Z1447621107 the data analysis shows that for the first 40 ns, the complex is unstable where radius of gyration increases attaining a maximum value of 1.68 nm around 7ns and then decreases till 40 ns? After 40ns the complex shows stability till the end of simulation with small fluctuations and insignificant changes in the radius of gyration. The DHFR enzyme and Z2604448220 ligand shows stability with minor changes throughout entire simulation where around 30 ns a maximum peak of 1.65 nm is seen. The complex of DHFR enzyme and ligand Z1830442365 imitate the second complex of DHFR enzyme with ligand Z2604448220 as the radius of compaction seems stable throughout the simulation

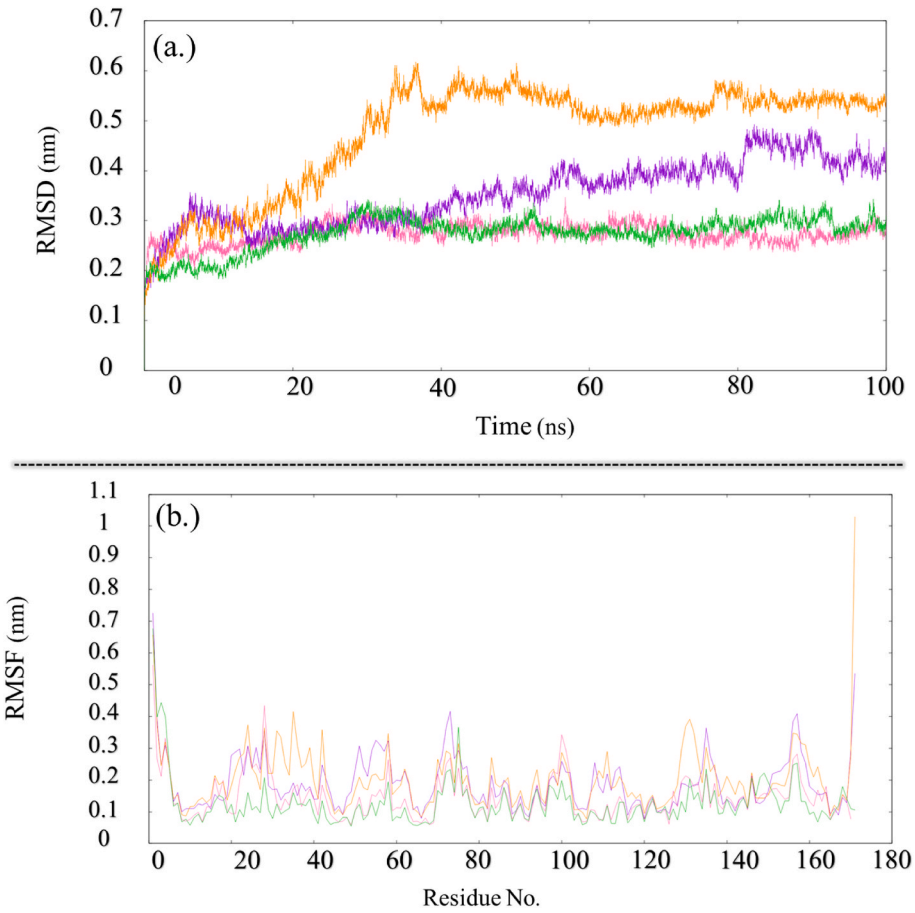


Fig. 4. MD simulation analysis **(a)** RMSD**(b)** RMSF. Pink (Free DHFR enzyme), Purple (DHFR-Z1447621107 complex), Orange (DHFR-Z2604448220 complex), Green (DHFR-Z1830442365 complex). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

same like DHFR and Z2604448220 complex, where a maximum radius of compaction of 1.60 nm is seen around 86 ns as shown in Fig. 5a. The average standard Rg values for unbound DHFR, DHFR- Z1447621107, DHFR- Z2604448220, DHFR- Z1830442365 are 1.58 nm, 1.59 nm, 1.60 nm, 1.57 nm respectively.

3.3.4. Solvent accessible surface area

SASA is used to describe a change in a protein molecule's solvent behaviour brought on by possible conformational changes upon binding to a ligand. Changes induced in the SASA profile reflect the conformational changes in enzyme induced due to interaction with the solvent environment and due to binding of ligand molecule to the active site in the biologically active system. The SASA profile for free DHFR enzyme is stable with minor fluctuations where the trajectory shows the maximum area accessible to solvent is 104.13 nm^2 during the 23ns and the area accessed by solvent is 96.51 nm^2. The DHFR enzyme with ligand Z1447621107 complex shows a standard SASA profile of 99.02 nm^2 which is slight more as compared to unbound DHFR enzyme and it suggests the conformational changes induced in the DHFR enzyme due to binding of ligand to its active site region. The trajectory starts with high SASA profile and there is a continuous decrease in the SASA profile towards end of simulation. The complex of DHFR enzyme with ligand Z2604448220shows the similar SASA profile to the DHFR enzyme and ligand Z1447621107 as the SASA values starts with high profile but continuously decreases towards the end and maximum SASA profile was seen 109.44 nm^2 around 7 ns. The complex shows a standard SASA profile of 97.91 nm^2 which is just slight more than unbound DHFR enzyme. For the DHFR enzyme and ligand Z1830442365 the SASA profile shows stable trajectory with slight fluctuations. The SASA profile for the complex is 94.99 nm^2 which is the least as compared to the unbound DHFR and its complex with the other two ligands as shown in Fig. 5b. The mean SASA values for unbound DHFR, DHFR-Z1447621107, DHFR- Z2604448220, DHFR- Z1830442365 are 96.51 nm^2, 99.02 nm^2, 97.91 nm^2, 94.99 nm^2 respectively.

3.3.5. Intermolecular hydrogen bonding

The hydrogen bonds formed during interaction of DHFR enzyme and ligands are very crucial for stability of the complex. Adequate number of H-bonds typically indicates a ligand's strong propensity for binding to the protein molecule. The DHFR enzyme complex with the ligand Z1447621107, for the first 40ns of simulation, shows 2 H-bonds continuously and at an instance of 40ns,3 H-bonds were also seen indicating the stable and appreciable binding. From 40ns to 80ns vague bonding was seen with one H-bond seen intermittently. During last 20ns of simulation, one H-bond was seen continuously seen but within 80ns–90 ns, but two to three H-bonds were also formed at some instances. The DHFR complex with ligand Z2604448220shows much stable H-bonding especially during 40ns–100ns period where twoH-bonds were continuously seen while three H-bonds were intermittently seen. In the initial 40 ns, a consistent presence of one hydrogen bond was noted, while the observation of two hydrogen bonds was somewhat indistinct. Surprisingly, at the 10 ns, there was an unexpected increase to five hydrogen bonds, deviating from the typical H-bonds in the subsequent trajectory. The DHFR enzyme and ligand Z1830442365complex shows one H-bond continuously during interaction and two H-bonds were seen vaguely while at some instances three H-bonds were also seen. Hence ligand Z1830442365 shows stable interaction with active site of DHFR enzyme during the entire simulation

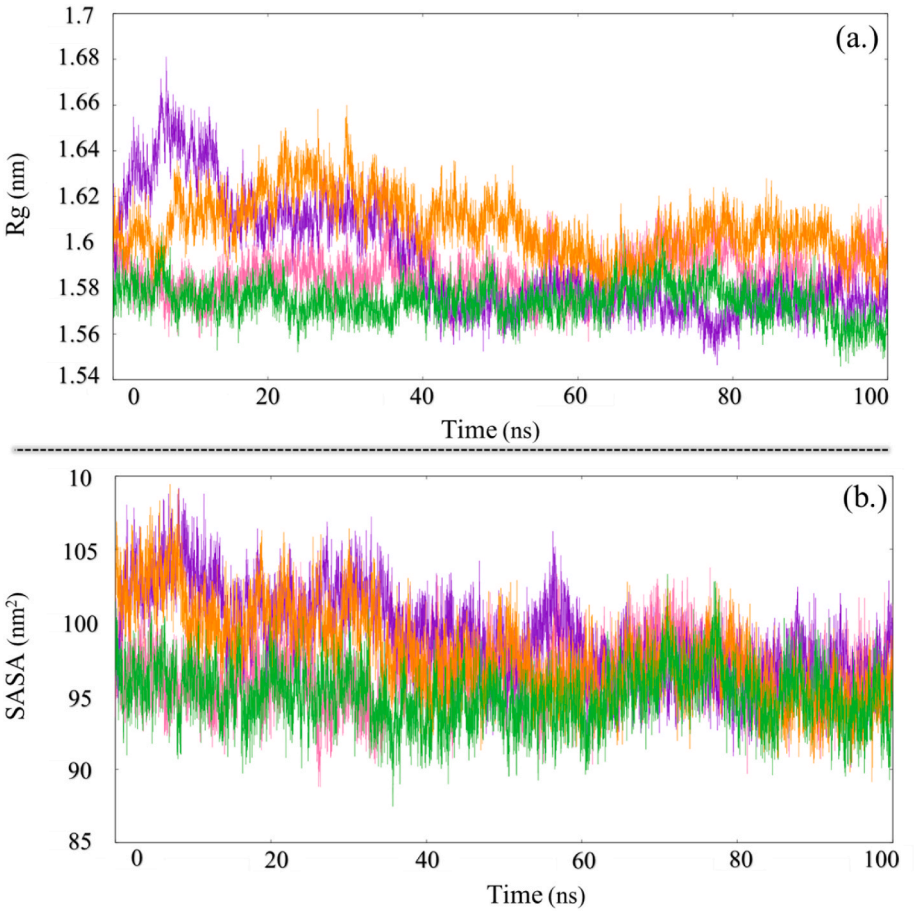


Fig. 5. MD simulation analysis showing **(a)** Radius of Gyration **(b)** Solvent accessible surface area. Pink colour (Free DHFR enzyme), Purple colour (DHFR-Z1447621107 complex), Orange (DHFR-Z2604448220 complex), Green (DHFR-Z1830442365 complex). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

as evident from the analysis as shown in Fig. 6a.

3.3.6. Principal component analysis (PCA)

Motion analysis is usually employed for evaluating the stability of enzyme and ligand complexes, as an enzyme's functionality typically arises from the collective atomic movements it undergoes. A method known as PCA uses the movement of the C-alpha atoms in the native DFHR enzyme to represent changes. Two eigenvectors (PC1 and PC2) with associated eigenvalues that are used to examine the key motions pertinent to the DHFR enzyme system are represented by the matrix of PCA analysis. Free DHFR enzyme, DHFR enzyme-Z1447621107 complex, DHFR enzyme-Z2604448220 complex, and DHFR enzyme-Z1830442365 complex have diagonalized covariance matrices with traces of 49.30 nm^2, 98.29 nm^2, 128.01$nm^2$and 48.64 nm^2 respectively. The distribution of variations in the free DHFR enzyme and DHFR enzyme-ligand complexes is shown in Fig. 6b. In the prospect of PCA analysis, it is evident that unbound DHFR enzyme has narrower PC1 and PC2 while DHFR enzyme-Z1447621107 and DHFR enzyme-Z2604448220 complexes have wider magnitude of eigenvectors (PC1 and PC2). DHFR enzyme and Z1830442365 complex shows the least magnitude in PC1 and PC2 which suggests that unbound DHFR enzyme and DHFR enzyme complex covers less conformational space as compared to complexes of DHFR enzyme with other two ligands. The narrow magnitude of PC1 and PC2 in analysis of DFHR enzyme and Z1830442365 complex suggest that stability is acquired by DHFR enzyme when ligand Z1830442365 interacts with it.

3.4. MM/PBSA calculations for interacting free energy

The g_mmpbsa tool, computed binding free energy of protein-ligand complexes by using input parameters of 1000 snapshots collected at intervals of 20ps between 80ns and 100 ns. DHFR enzyme and its complexes with ligands Z1447621107, Z2604448220, Z1830442365 shows overall binding energy as −8.56 kcal/mol, 0.52 kcal/mol, −20.84 kcal/mol respectively and are shown in Table 7. The graphical representation of binding energies of all ligand-DHFR complexes are depicted in Fig. 7.

4. Discussion

The sporadic effect of diseases occurring due infective activity of *A. baumannii* in hospital admitted patients and its increasing MDR and XDR propensity has an urgent need of the time to find potential antimicrobial agents against it. Our present study utilised a computational approach to find out potential inhibitors of DHFR enzyme. Virtual high throughput (vHTS) screening was used to identify three ligands to be potential inhibitors of the DHFR enzyme of *A. baumannii*. RMSD analysis shows that DHFR-Z1830442365 complex shows the mean deviations of 0.27 nm that are similar to unbound DHFR. While the other two complexes DHFR-Z1447621107, DHFR-Z2604448220 shows the higher deviation in the structural parameters, thus it can be suggested that higher stability is induced by the binding of the ligand Z1830442365. In spite that no hydrogen bond is formed with active site residues, the stability seems to be induced by the good number of Pi-Pi interactions and van-der Waal interactions which shows bonding all around ligand from every

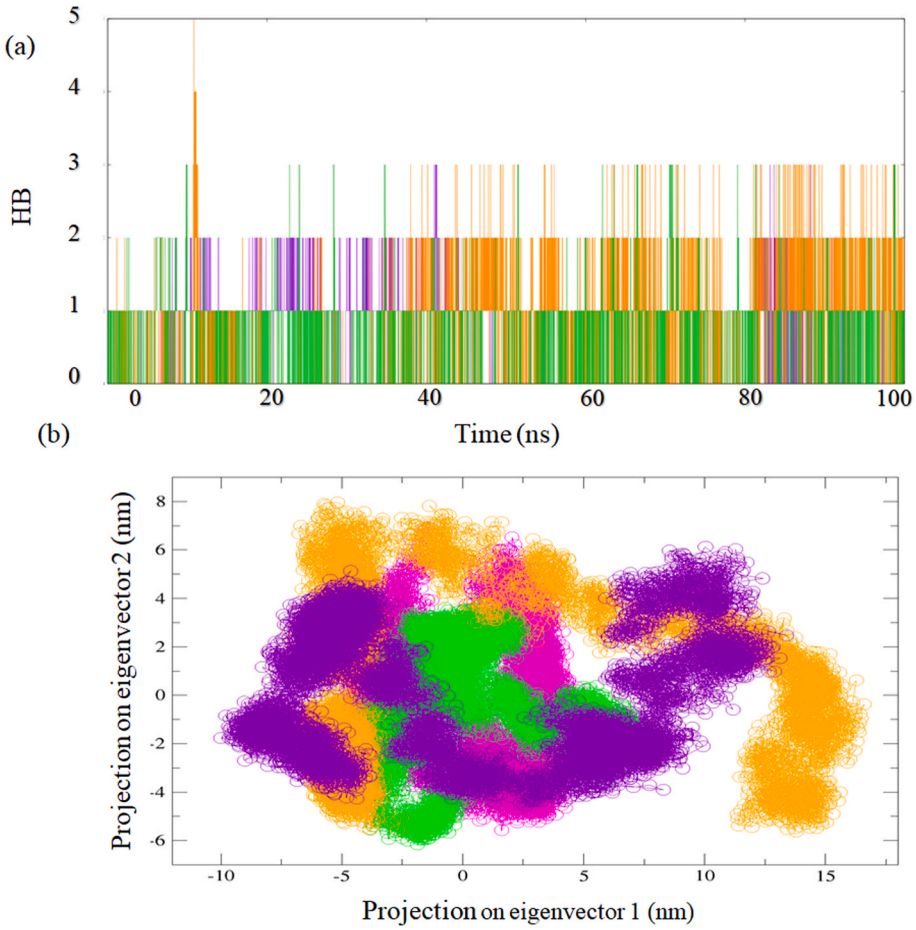


Fig. 6. MD simulation analysis (a) Intermolecular Hydrogen Bond (b) Principal Component Analysis. Pink colour (Free DHFR enzyme), Purple colour (DHFR-Z1447621107 complex), Orange (DHFR-Z2604448220 complex), Green (DHFR-Z1830442365 complex). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 7
Different energy contributions in the formation of DHFR complex with ligands based on MM/PBSA analysis. (Energies mentioned are in kcal/mol).

S.no	Complex	Van der Waals energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
1.	DHFR-Z1447621107	−19.48	−3.40	21.95	−2.31	−8.56
2.	DHFR-Z2604448220	−2.44	−2.57	4.77	−0.27	−0.52
3.	DHFR-Z1830442365	−32.02	−15.08	29.70	−3.44	−20.84

direction in the active site cleft. RMSF analysis shows that complex DHFR-Z1830442365 has even lower mean fluctuations than unbound DHFR enzyme while DHFR-Z1447621107 and DHFR-Z2604448220 complexes has higher mean fluctuations. The residues that show noticeable peaks in the trajectory are found to be located in loop region, especially Met1 which is common in depicting significant fluctuations in unbound DHFR enzyme and all other three complexes. According to Rg analysis, the DHFR enzyme and its complexes have roughly similar mean compaction. Nevertheless, when compared, the DHFR-Z1830442365 complex exhibits the least amount of compaction, while DHFR-Z1447621107 and DHFR-Z2604448220 exhibit somewhat higher mean compaction than unbound DHFR. This indicates that after contact, unbound DHFR and its complexes exhibit stability, but the ligand Z1830442365 has the highest stability. According to SASA analysis, the DHFR-Z1830442365 complex exposes the least amount of surface area to the solvent, but the DHFR-Z1447621107 and DHFR-Z2604448220 complexes have higher SASA profiles than unbound DHFR. This suggests that the DHFR-Z1830442365 complex had the highest level of stability during MD simulation when compared to the free DHFR enzyme, DHFR-Z1447621107 and DHFR-Z2604448220. According to H-Bond analysis, DHFR-Z1447621107 and DHFR-Z2604448220 show fluctuations in number of bonds with intermittent formation and breakage, but the ligand Z1830442365 shows continuous interaction through one H-Bond formation which shows stable and significant binding. According to the confirmation given to the DHFR enzyme and its complexes with ligands on the basis of pertinent motions, PCA analysis which includes diagonalized covariance matrix, demonstrates a stable nature. It also demonstrates that the DHFR complex with ligand Z1830442365 exhibits covariance matrices with least traces when compared to unbound DHFR, DHFR-Z1447621107, and DHFR-Z2604448220. The high trace of covariance matrices in the DHFR-Z1447621107 and DHFR-Z2604448220 complex indicate an unstable state for the DHFR enzyme in its interaction with the ligands. Hence, according to PCA analysis, the DHFR- Z1830442365 complex has the highest stable state. MM/PBSA analysis shows that the DHFR-Z1830442365 complex has the highest free binding energy among all the complexes which again points towards the stable state of DHFR enzyme induced after the Z1830442365 ligand binding. In the similar

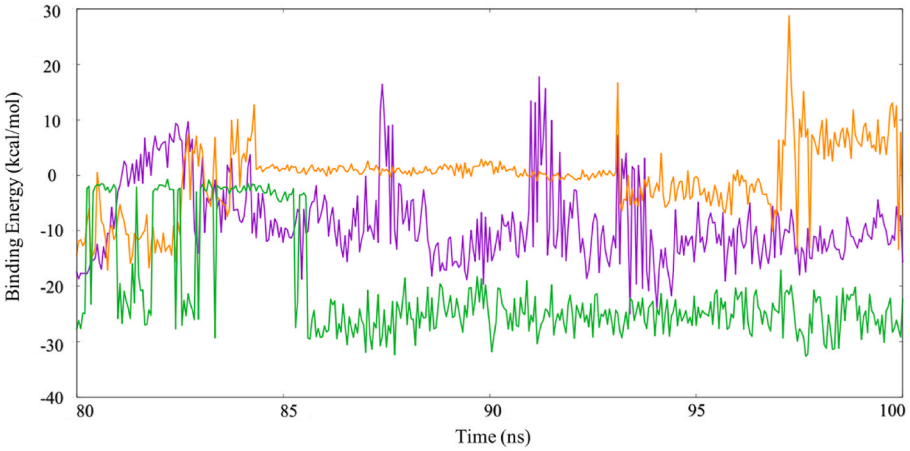


Fig. 7. Analysis of MM/PBSA binding energies. Pink colour (Free DHFR enzyme), Purple colour (DHFR-Z1447621107 complex), Orange (DHFR-Z2604448220 complex), Green (DHFR-Z1830442365 complex). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

study, conducted to find potential inhibitors of DHFR enzyme, the potential inhibitors (2, 4 dihydroxyphenyl and methanone) against *Mycobacterium tuberculosis,* possessed the binding free energy of −24.69 kal/mol and −23.58 kcal/mol (Sharma et al., 2020). The result for the binding free energy of the DHFR- Z1830442365 suggests the value of −20.84 kcal/mol, that is in comparison to energy of above mentioned inhibitors of DHFR enzyme from *Mycobacterium tuberculosis*.

5. Conclusion

The present study, which targets the DHFR enzyme from *A. baumannii* to find the lead molecule candidates which could bind to active site cleft and inhibit it. For this virtual high throughput screening (vHTS) was done and three ligands were selected which were fit to bind to DHFR enzyme's active site. The validation of the complex formed by ligands and DHFR enzyme was done using MD simulation studies. Various analyses of MD simulation studies predicted that ligand Z1830442365is the best potential inhibitor of DHFR enzyme as compared to other two candidates. There are certain limitations of *insilico* studies as they carry out studies based on virtual conditions and not the real physiological conditions. Therefore, additional experimental investigations are necessary to confirm the potential of ligand Z1830442365 as an inhibitor of the DHFR enzyme.

CRediT authorship contribution statement

Saurabh Kumar Bhati: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Monika Jain:** Formal analysis, Methodology, Software, Validation, Writing – review & editing. **Jayaraman Muthukumaran:** Investigation, Supervision, Validation, Software. **Amit Kumar Singh:** Conceptualization, Investigation, Supervision, Validation, Software.

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.

Data availability

Data will be made available on request.

Acknowledgements

Dr. Amit Kumar Singh thanks Indian Council of Medical Research (ICMR) and Indian National Science Academy (INSA), New Delhi, India. Mr. Saurabh Kumar Bhati thanks Council of Scientific and Industrial Research (CSIR) for providing Senior Research Fellowship. The authors thank Sharda University for support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crstbi.2024.100127.

References

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., et al., 2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science 373 (6557), 871–876.

Chen, X., Li, H., Tian, L., Li, Q., Luo, J., Zhang, Y., 2020. Analysis of the physicochemical properties of acaricides based on Lipinski's rule of five. J. Comput. Biol. 27 (9), 1397–1406.

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2 (9), 1511–1519.

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.

Franklin, M.C., Cheung, J., Rudolph, M.J., Burshteyn, F., Cassidy, M., Gary, E., et al., 2015. Structural genomics for drug design against the pathogen Coxiellaburnetii. Proteins: Struct., Funct., Bioinf. 83 (12), 2124–2136.

Harding, C.M., Hennon, S.W., Feldman, M.F., 2018. Uncovering the mechanisms of Acinetobacterbaumannii virulence. Nat. Rev. Microbiol. 16 (2), 91–102.

Jha, R.K., Khan, R.J., Singh, E., Kumar, A., Jain, M., Muthukumaran, J., Singh, A.K., 2022. An extensive computational study to identify potential inhibitors of Acyl-homoserine-lactone synthase from Acinetobacterbaumannii (strain AYE). J. Mol. Graph. Model. 114, 108168.

Jiang, Y., Ding, Y., Wei, Y., Jian, C., Liu, J., Zeng, Z., 2022. Carbapenem-resistant Acinetobacterbaumannii: a challenge in the intensive care unit. Front. Microbiol. 13, 1045206.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T.L., 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36 (Suppl. l_2), W5–W9.

Kumari, R., Kumar, R., Lynn, A., Open Source Drug Discovery Consortium, 2014. g_mmpbsa– A GROMACS tool for high-throughput MM-PBSA calculations. J. Chem. Inf. Model. 54 (7), 1951–1962.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26 (2), 283–291.

Mark, P., Nilsson, L., 2001. Molecular dynamics simulations of the ala-pro dipeptide in water: conformational dynamics of trans and cis isomers using different water models. J. Phys. Chem. B 105 (33), 8028–8035.

Sharma, K., Neshat, N., Sharma, S., Giri, N., Srivastava, A., Almalki, F., Akhter, M., 2020. Identification of novel selective Mtb-DHFR inhibitors as antitubercular agents through structure-based computational techniques. Arch. Pharmazie 353 (2), 1900287.

Silva, T.F., Vila-Viçosa, D., Reis, P.B., Victor, B.L., Diem, M., Oostenbrink, C., Machuqueiro, M., 2018. The impact of using single atomistic long-range cutoff schemes with the GROMOS 54A7 force field. J. Chem. Theor. Comput. 14 (11), 5823–5833.

Skariyachan, S., Muddebihalkar, A.G., Badrinath, V., Umashankar, B., Eram, D., Uttarkar, A., Niranjan, V., 2020. Natural epiestriol-16 act as potential lead molecule against prospective molecular targets of multidrug resistant Acinetobacter baumannii-Insight from in silico modelling and in vitro investigations. Infect. Genet. Evol. 82, 104314.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem. 26 (16), 1701–1718.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35 (Suppl. l_2), W407–W410.

Yin, J., Landau, D.P., 2011. Structural properties and thermodynamics of water clusters: a Wang–Landau study. J. Chem. Phys. 134 (7).