

STRUCTURE-BASED VIRTUAL SCREENING, MOLECULAR DOCKING, AND DYNAMICS SIMULATION APPROACHES FOR IDENTIFICATION OF INHIBITORS OF *HELICOBACTER PYLORI* MQNB

ARTHIKASREE ANANDAMURTHY¹, VASAVI GARISSETTI¹, ROSLIN ELSA VARUGHESE¹, GAYATHRI DASARARAJU*¹

Centre of Advanced Study in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai, Tamil Nadu, India.

*Corresponding author: Gayathri Dasararaju; Email: drdgayathri@gmail.com

Received: 02 December 2024, Revised and Accepted: 13 January 2025

ABSTRACT

Objectives: *Helicobacter pylori* is the major cause of duodenal ulcers and gastric cancer. Menaquinone is essential for the respiration and survival of *H. pylori*. Identification of compounds that have good binding affinity at the active site of MqnB will be a promising approach against *H. pylori*. This study aims to identify the natural lead compounds against MqnB enzymes in *H. pylori* through *in silico* methods.

Methods: As the three-dimensional structure of *H. pylori* MqnB is not available, we have modeled the three-dimensional structure followed by identification of the active site. The binding of the compounds from the COLleCtion of Open NatUral productS database was screened against MqnB, and dynamics simulation studies were carried out for the top three lead compounds.

Results: The predicted model of MqnB was subjected to a molecular dynamics (MD) simulation study to obtain the dynamic behavior of MqnB. The best representative model was validated, and further used for structure-based virtual screening. Based on the docking results, three lead compounds were chosen among the top hits and subjected to 200 ns MD simulation. The results highlight the dynamic nature of MqnB in complex with the lead compounds and favorable binding interactions were comparable with the substrate. All the compounds showed promising interactions with MqnB during MD simulations.

Conclusion: Identified lead compounds show good binding potential and also maintain interactions with amino acid residues at the active site of MqnB enzyme. These compounds could be further explored in the drug discovery process against *H. pylori*.

Keywords: Fufalosine, MqnB, *Helicobacter pylori*, Natural compounds, Molecular docking, Dynamics simulation.

© 2025 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2025v18i2.53566>. Journal homepage: <https://innovareacademics.in/journals/index.php/ajpcr>

INTRODUCTION

Menaquinone (MK) and ubiquinone are responsible for bacterial respiration [1]. MK biosynthesis happens through two distinct pathways in prokaryotes. The traditional MK pathway in certain bacteria employs six enzymes expressed as menFDHCEB genes to convert chorismate to 1,4-dihydroxy-2-naphthoyl-CoA thioesterase. Polyprenyltransferase (MenA) and methyltransferase (MenG) enzymes convert 1,4-dihydroxy-2-naphthoate to MK [2]. However, certain bacteria lack men homologs and have established an alternative pathway, the fufalosine pathway, which has been found in *Streptomyces coelicolor*, *Helicobacter pylori*, *Campylobacter jejuni*, and *Thermus thermophilus* [3]. In this alternative route, MqnABCD and unknown Mqn enzyme(s) catalyze the conversion of chorismate to MK.

MqnB, an enzyme in the fufalosine pathway, also known as fufalosine hydrolase utilizes fufalosine as a substrate and converts it into dehydropoxanthinyl fufalosine and hypoxanthine [4]. MqnB from *H. pylori* contains 180 amino acids and has a molecular weight of 20.02 kDa. Fufalosine is a nucleoside that was first discovered in 1999 in the fermentation broth of *Streptomyces*. It is a unique nucleoside derivative composed of an inosine core with a 3-carboxyphenyl methylene ketone group substituting the C-50 hydroxyl group [5]. MqnB has recently been found to be an essential enzyme in an alternative MK biosynthesis pathway. *H. pylori* which causes duodenal ulcers and stomach cancer requires MK for respiration and survival [6]. Identification of the potential lead compounds targeting the enzymes in the fufalosine pathway will be a promising approach towards the discovery and development of anti-*H. pylori* drugs. The structure of MqnB has not yet been determined in any organism, nor has its function been fully characterized.

In the current study, we have predicted the three-dimensional structural models of MqnB followed by structure-based high-throughput virtual screening (HTVS) of compounds from a natural compounds database. Based on the energy of the docked complexes, hydrogen bonds, and hydrophobic interactions at the active site of MqnB, two lead compounds were identified. Molecular dynamics (MD) simulation studies were carried out to explore the binding potential of the compounds and the stability of the active site interactions.

METHODS

Structure prediction and validation

The structure of MqnB was modeled computationally, as the structure was not yet determined experimentally. The primary sequence of MqnB was retrieved using the gene accession ID: A0A402DZF2 from the UniProt database [7]. The structure was then predicted using AlphaFold2 ColabFold v1.5.5. AlphaFold2, developed by DeepMind, an enhanced version of the original AlphaFold designed for predicting the protein structure [8]. It makes use of the MMseqs2 method for swift search against large databases to find homologous sequences that are related to the template sequence. AlphaFold2 uses a predicted local distance difference test (pLDDT) score, template modeling (TM) score, and predicted aligned error (PAE) for ranking the five predicted models to find the best reliable model. The most reliable predicted structure was validated using the Ramachandran plot which was generated using the PROCHECK available in SAVES v6.0 [9].

Active site prediction

SiteMap, a tool in the Schrödinger suite, was used to predict the binding site of MqnB [10]. It helps in the early stages of drug discovery in

offering insights into the druggability of the binding pockets of the enzymes. The predicted MqnB model was subjected to Schrödinger's Preparation Wizard which involves processes such as adding hydrogens and assigning bond orders. Site points that may be utilized for ligand interaction within proteins are produced by the OPLS2005 force field. Concave pockets on the protein surface where the ligand can bind are examined for its size and shape. Two scoring functions, Sitescore and Dscore, are assigned to the potential binding pockets identified by the program. Sitescore, which determines the quality of the binding pockets, ranges from 0 to 1.2. The score and the potential for a particular site to be druggable are directly related to each other. Dscore is more focused on the potential of the site to be druggable, helping to prioritize sites that are not only of good quality but also likely to be an effective target for drug discovery. It gives values that range from 0 to 1.2.

Structure-based virtual screening (SBVS)

Working database

Naturally derived small molecule compounds were downloaded from the COCONUT (COlleCtion of Open NatUral productS) database which contains over 400,000 compounds [11]. All the 2D ligand coordinates were downloaded in structured data file format and the LigPrep tool from the Schrödinger suite was used for minimizing all the ligands using the OPLS2005 forcefield. The COCONUT database is an extensively used and publicly accessible database for research on natural products and drug discovery. LigPrep helped in arriving at well-optimized, low-energy, three-dimensional structures for the ligands with correct protonation states and chirality.

Virtual screening workflow

To find the potential inhibitors for the enzyme, MqnB, we have employed a virtual screening technique using the HTVS module from Schrödinger. The HTVS module [10] utilized SBVS techniques, incorporating the three-dimensional structure of the MqnB receptor and its binding site into the screening process. It helped us to identify the hit molecules from the COCONUT database. HTVS workflow starts with protein preparation and grid generation using Schrödinger followed by Glide-HTVS docking, Glide-standard precision (SP), and Glide-extra precision (XP) docking. Glide-HTVS is mainly used for the screening of a large dataset within a short time period and 10% of the top-scored ligands are taken forward for Glide-SP which aids in reducing the false negatives by having a proper balance between speed and accuracy. 10% of the top-scored compounds from SP are directed to Glide-XP docking which uses a more sophisticated scoring function for a comparatively smaller dataset. The compounds which were ranked high based on their Glide energy were subjected to induced fit docking (IFD) in Schrödinger [12]. IFD incorporates both receptor and ligand flexibilities during receptor-ligand docking using both the Glide and Prime programs of Schrödinger [13]. Initially, a rigid receptor docking was carried out in SP mode with reduced van der Waals radii and an increased Coulomb-vdW cutoff. After prime side-chain minimization of the receptor and ligand for each pose, the structures within an energy limit were subjected to Glide re-docking using XP mode, and each pose was ranked based on an IFD scoring function. The best-docked pose for the ligands were analyzed based on their Glide energy and ligand interactions which were generated using LigPlot+ [14].

MD simulation

To understand the stability of the protein-ligand complexes after IFD, they were subjected to MD simulation studies in GROMACS [15,16]. Minimization, equilibration of the system, and the production MD run were performed in GROMACS using the CHARMM27 force field. The ligand parameter files were prepared using the SwissParam online tool [17]. A simple cubic box was generated and the protein was centered at 1.0 nm from the edge of the box. The system was solvated and neutralized by adding two Na⁺ ions by replacing the solvent molecules using the genion module. The neutralized system was energy minimized using the steepest descent method and a maximum force of 1000 KJ/mol/nm was used. This was followed by two steps

of equilibration, NVT, and NPT, with temperature coupling using a v-rescale thermostat (300 K), and pressure coupling using a Parrinello-Rahman barostat, respectively [18]. The system was then subjected to a 200 ns MD simulation run and the trajectories formed were analyzed using the gmx_rms, root mean square fluctuation (RMSF), gyrate, solvent accessible surface area (SASA), and hbond modules for a better understanding of the ligand's binding with the protein at its active site. The graphs were generated using xmgrace, and the structures were visualized in PyMOL [19].

QikProp analysis for lead compounds

To predict the absorption, distribution, metabolism, excretion, and toxicity characteristics of the lead molecules, we employed the QikProp module within the Schrödinger suite. It predicted key physicochemical descriptors such as molecular weight, polar surface area, and logP and properties relevant to pharmacokinetics such as solubility, permeability, and potential toxicity. In addition, QikProp provides a comparison range for each molecule's properties against those of 95% of known drugs. The identified lead compounds after IFD and MD simulation studies were subjected to QikProp analysis to ensure that the compounds have the proper pharmacokinetic properties.

RESULTS AND DISCUSSION

Structure prediction

To date, no crystal structure of MqnB has been reported. Therefore, to predict the three-dimensional structure, we have used AlphaFold2 which is an open-source software for protein structure prediction. The MqnB sequence was retrieved from UniProt (UniProt ID: A0A402DZF2) and the FASTA sequence was submitted to AlphaFold2 for modeling. Using the ColabFold module with default parameters, five structural models were generated. The per-residue confidence scores were assessed using the pLDDT and the PAE matrix (Fig. 1a). Shows the PAE matrix, whereas (Fig. 1b) illustrates the pLDDT plot, with regions having a high confidence score for structural accuracy.

Among the five models, the one with the highest pLDDT score of 93.9 and a TM score of 0.90 was selected as the best model. The selected model's structure and its corresponding Ramachandran plot are displayed in Fig. 2. The Ramachandran plot analysis shows that 96.3% of the residues fall within the most favored regions, 2.4% in additionally allowed regions and 1.2% in generously allowed regions, indicating the reliability of the predicted structure. We further conducted a 200 ns MD simulation, followed by cluster analysis using GROMACS to identify the best representative model. This selected model was then used for docking studies to evaluate the ligand binding potential at the active site of the MqnB enzyme. Fig. 3 shows the best representative model and Ramachandran plot. The Ramachandran plot of the best representative shows that 81.4% of the residues fall within the most favored regions, 14% in additionally allowed regions, and 1.8% in generously allowed regions.

Binding site prediction

The SiteMap module of the Schrödinger suite was used to predict binding sites for the MqnB model. The settings were chosen to generate four potential binding sites. Based on the SiteScore, site size, and Dscore, the top hit from the four sites was chosen. The chosen binding pocket has a SiteScore of 1.035, indicating a high affinity for ligand binding. The high Dscore value of 1.032 indicates that the binding pocket is highly druggable, meaning that compounds can bind at the site with high affinity and selectivity. Table 1 contains a detailed list of potential binding pockets. Based on the SiteScore and Dscore, site number 1 was chosen as the active site and the amino acid residues in the active site are Cys4, Ala5, Gly6, Arg7, Asn8, Glu9, Thr10, Leu11, Lys12, Ile17, Phe43, Ile44, Gly45, Ser46, Ala47, Gly48, Tyr50, Ser103, Ile106, His107, Glu124, Asn125, Met126, Glu127, Ser149, Asn150, Ala152, Gly153, Leu154, Ala156, His157, Phe160, His164, Val167, Lys168, Gln169, Leu171, and Glu172. Fig. 4 shows the active site residues.

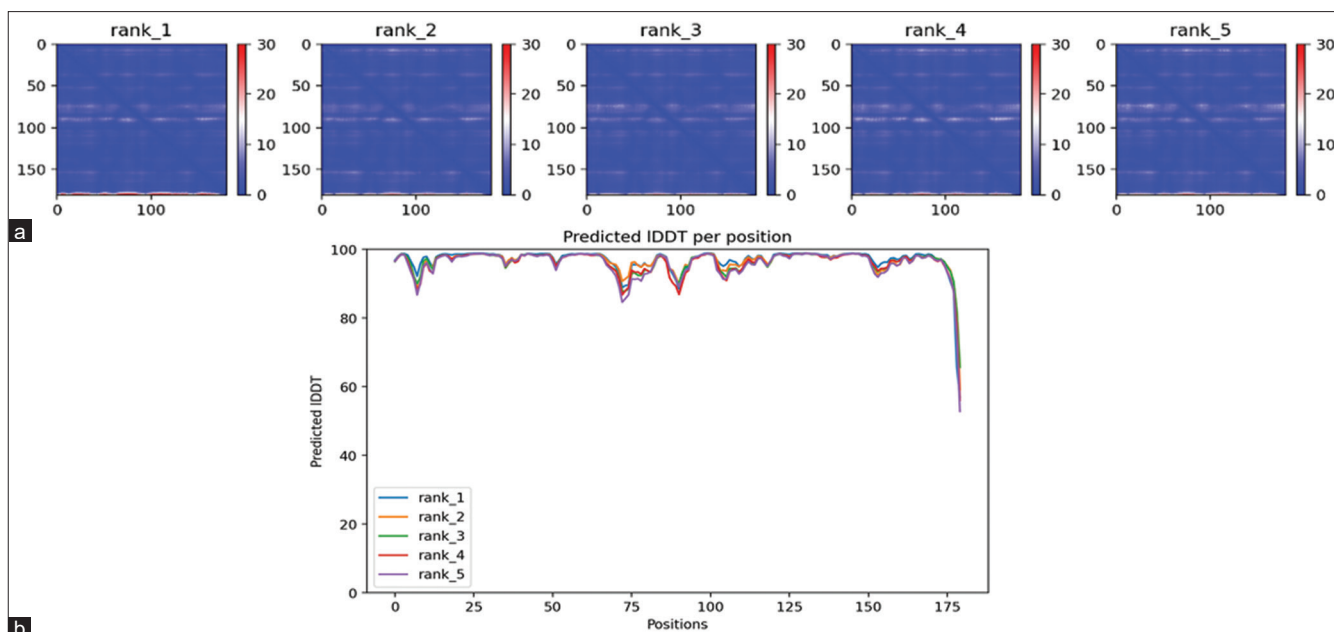


Fig. 1: AlphaFold2 results for the predicted models of MqnB (a) predicted aligned error matrix plots (b) predicted local distance difference test plot

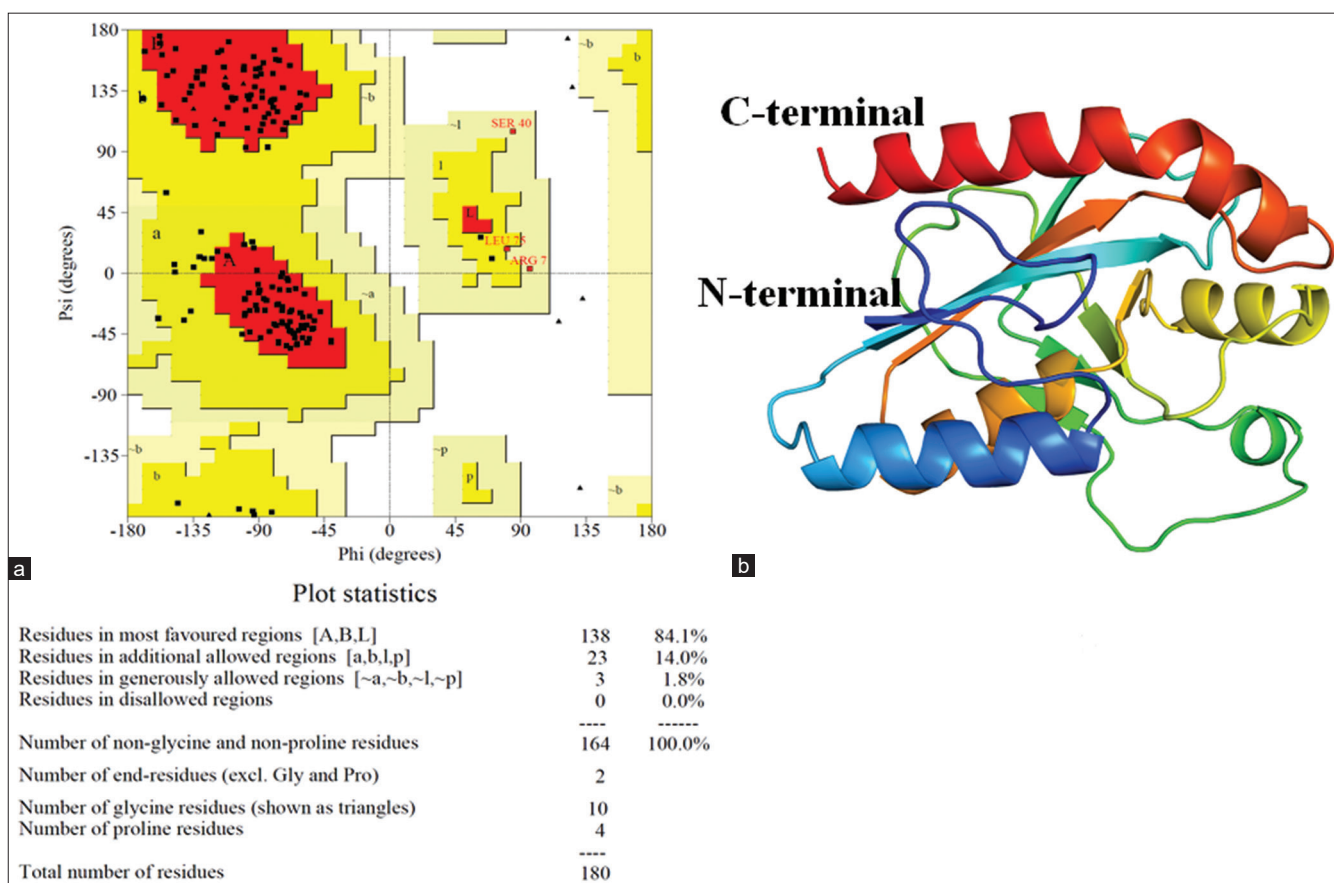


Fig. 2: (a) Ramachandran plot for the predicted model of MqnB (b) best predicted model from AlphaFold2

Molecular docking of lead compounds

A total of 406,760 compounds from the COCONUT database were docked into the predicted active site of MqnB using a SBVS approach in Schrödinger. The docking process followed a stepwise protocol. Initially,

compounds were docked using HTVS, followed by SP, and finally XP docking. Table 2 provides a detailed list of compounds screened during virtual screening. After XP docking, the top 20 compounds were shortlisted based on their Glide energy and docking scores, which were

then subjected to IFD. Table 3 presents the two-dimensional structures, molecular weights, and International Union of Pure and Applied Chemistry names of these shortlisted compounds. Based on the Glide energy, docking scores, and interactions at the active site, 10 compounds were selected after IFD, as they demonstrated stronger binding energies with the predicted binding pockets compared to other compounds. Table 4 summarizes the docking scores, Glide energies, hydrogen bonds, and hydrophobic interactions of the top 10 compounds (out of 20 compounds screened from IFD) with MqnB. The ligand interaction diagrams are shown in Fig. 5. Notably, compound 1 exhibited the lowest binding energy of -84.14 kcal/mol and formed hydrogen bonds with residues Arg7, Glu12, Asn125, Met126, and Asn150, along with hydrophobic interactions involving Asn8, Glu9, Leu11, Ser46, Ile106, His164, Lys168, and Leu171. Substrate (futasoline) shows binding energy of -68.26 kcal/mol and forms hydrogen bonds with Asn5, Gly6, Asn150, and Lys161 residues and forms hydrophobic interactions with Ala5, Gly6, Arg7, Ser46, Gly48, Ile106, Asn125, Met126, Ala 156, His157, and Phe160.

Based on the docking results, three compounds were shortlisted and the compounds showed strong interactions with active site residues (Fig. 6a-e). The binding affinities of these compounds were further confirmed through docking scores and protein-ligand interactions. Notably, compound 1 (lead1) showed the highest affinity with multiple hydrogen bond interactions involving key residues, suggesting a strong potential for binding at the active site.

MD simulation

MqnB-lead complexes were subjected to a 200 ns MD simulation to assess the stability of the protein-ligand interactions. Root mean square deviation (RMSD) analysis of the protein backbone indicated that all complexes stabilized after 100 ns, with minimal fluctuations (Fig. 7a). Similarly, the RMSF plot showed that the binding site residues exhibited limited flexibility, indicating stable binding interactions throughout the simulation (Fig. 7b). The SASA and radius of gyration graphs further confirmed that the complexes maintained their compact structure

and consistent solvent exposure (Fig. 7c-d). Hydrogen bond analysis revealed that the lead compounds formed stable hydrogen bonds throughout the simulation (Fig. 8a-d). The presence of stable hydrogen bonds over the course of the simulation indicates that the compound remains bound within the active site, supporting its potential as a high-affinity ligand.

Conformational analysis of MD snapshots

The conformational snapshots at initial, 100 ns, and 200 ns were analyzed to understand the ligand's behavior within the binding pocket over time (Fig. 9). All three lead compounds maintained stable interactions with the active site, with minor positional adjustments observed in the binding poses. The conformational analysis of MD snapshots shows that lead1 initially interacted with Arg7, Gly124, Asn125, Met126, and Asn150 and interacted with Glu9, Ser46, and Ser149 by 100 ns, and then with Asn8, Ser46, Ser149, and Lys161 at 200 ns.

Lead2 formed initial interactions with Gly6, Arg7, Asn8, Ser46, and Asn125, later made interactions with Ala5, Gly6, Glu9, and Ser46 at 100 ns, and Gly6, Arg7, Glu9, Ser46, and Asn125 at 200 ns. Lead3 made initial interactions with Gly6, Asn8, Ser46, Gly124, Asn125, Ser149, and Asn150, and then it showed key interactions with Asn8 at 100 ns and His164 at 200 ns. Substrate made initial interactions with Asn150, Asn8, Glu124, and Lys161, later made interactions with Asn150, Ser149, and Asn125 by 100 ns, and then with Asn150 and Ser149. These patterns indicate dynamic yet stable interactions, with residues such as Ser46, Asn125, Ser149, and His164 showing persistent interactions, suggesting their importance in maintaining ligand binding and stability. This consistency in binding interactions with active site residues highlights that these leads could be considered as a promising target for MqnB inhibition.

Hydrogen bond analysis in the best representative model

To further analyze the conformational change of the MqnB protein, clustering analysis was carried out over the last 100 ns simulation run

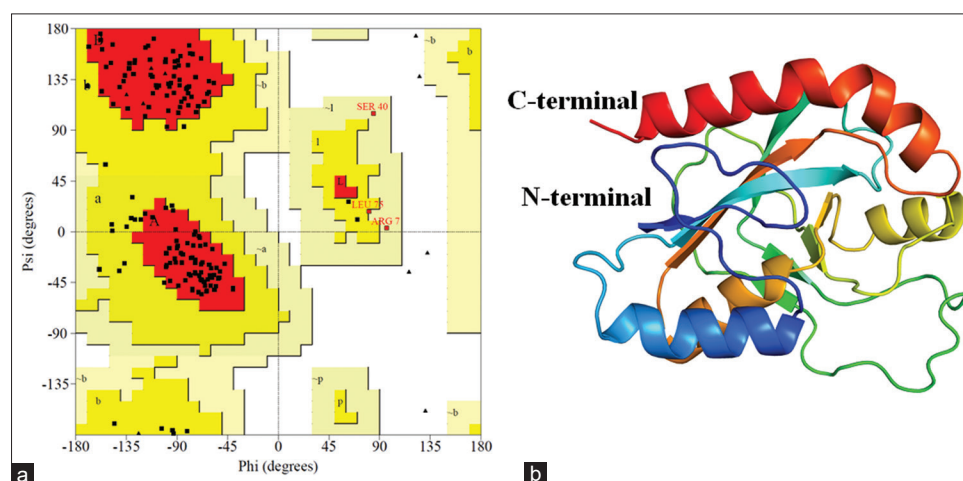


Fig. 3: (a) Ramachandran plot for the best representative model after 200 ns molecular dynamics run (b) the best representative model

Table 1: SiteMap results for the predicted binding sites of MqnB

Site no	Residues	SiteScore	Size (Å ³)	Dscore
1	Cys4, Ala5, Gly6, Arg7, Asn8, Glu9, Thr10, Leu11, Lys12, Ile17, Phe43, Ile44, Gly45, Ser46, Ala47, Gly48, Tyr50, Ser103, Ile106, His107, Glu124, Asn125, Met126, Glu127, Ser149, Asn150, Ala152, Gly153, Leu154, Ala156, His157, Phe160, His164, Val167, Lys168, Gln169, Leu171, Glu172	1.035	196	1.032
2	Leu21, Ile22, Ala25, Gln67, Val68, Glu69, Glu70, Lys72, His74, Asn76, Tyr78, Thr79, Pro80, Ser102, Ser103, Asn104, Tyr105, Phe129, Ser130, Ser133	0.988	95	1.022
3	Ser40, Glu61, Val63, Ile87, Thr89, Lys90, Glu91, Arg97, Lys141, Ala142, Lys143	0.737	32	0.673
4	Glu110, His151, Ala152, Gly153, Leu154, 155, Ala156, Glu159	0.507	24	0.426

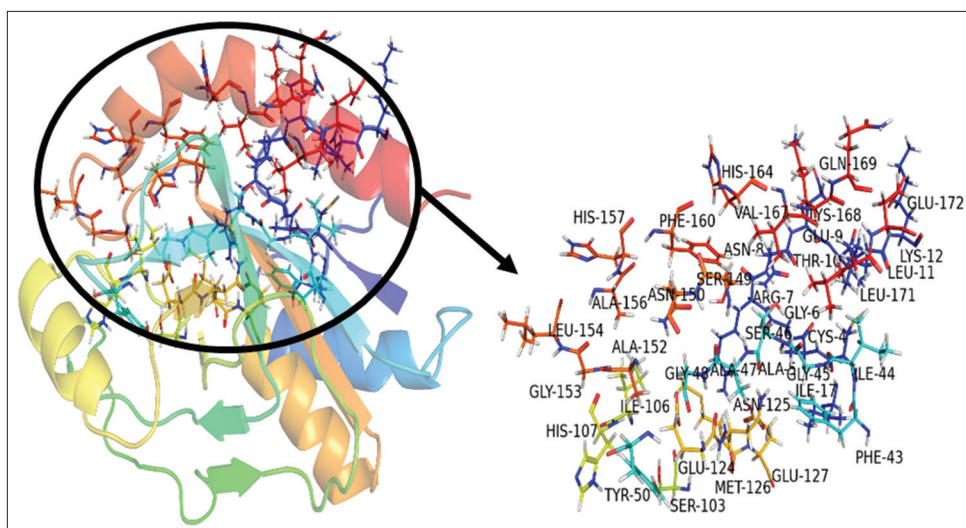


Fig. 4: MqnB best representative structure obtained after molecular dynamics analysis and the active site residues

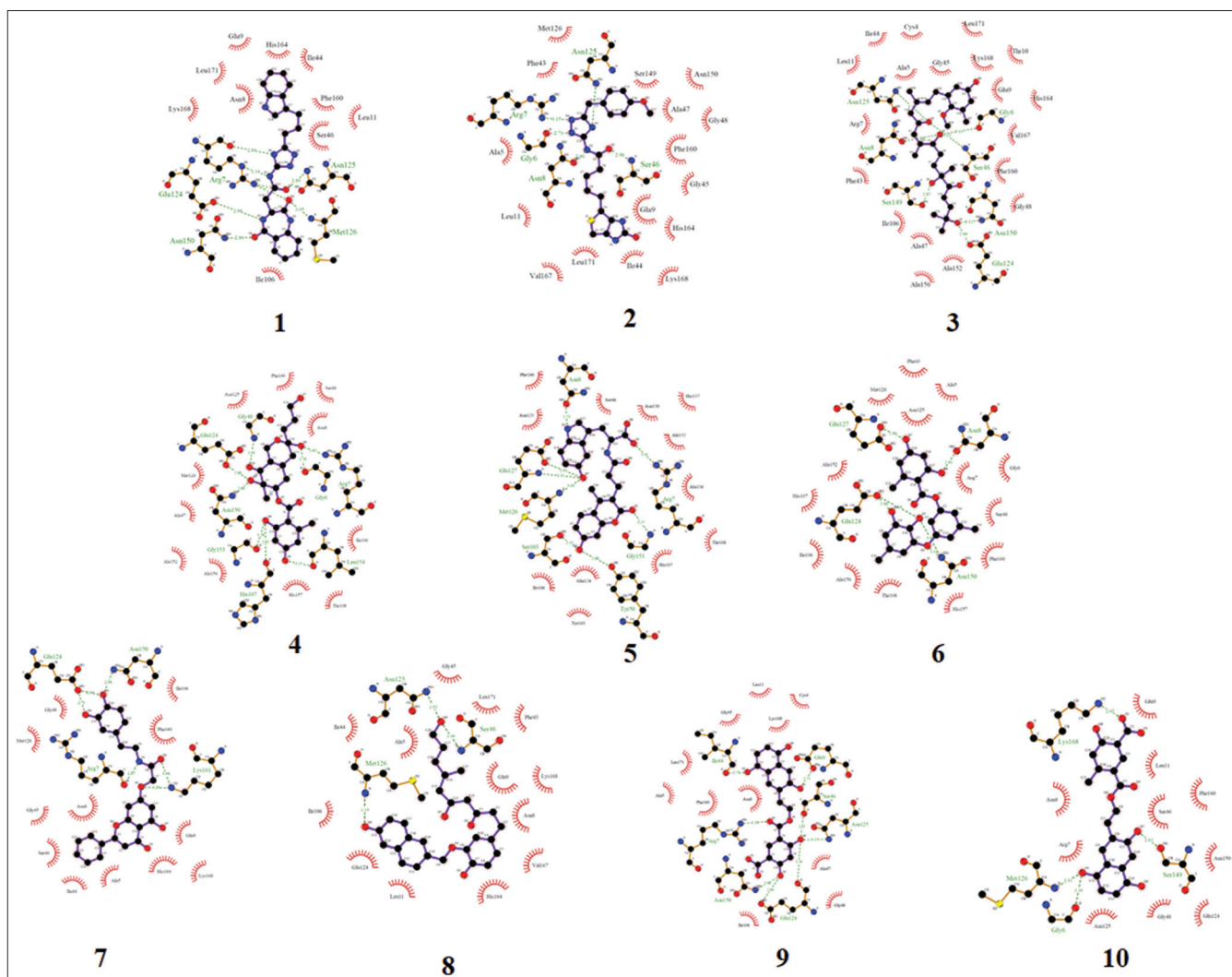


Fig. 5: Ligand interaction of top 10 compounds with MqnB

using GROMACS. Representative structures were extracted from the least RMSD frame in clustering analysis. Fig. 10a shows a superposed

image of the representative models and Fig. 10b shows the ligand interaction map for lead1, lead2, lead3, and substrate.

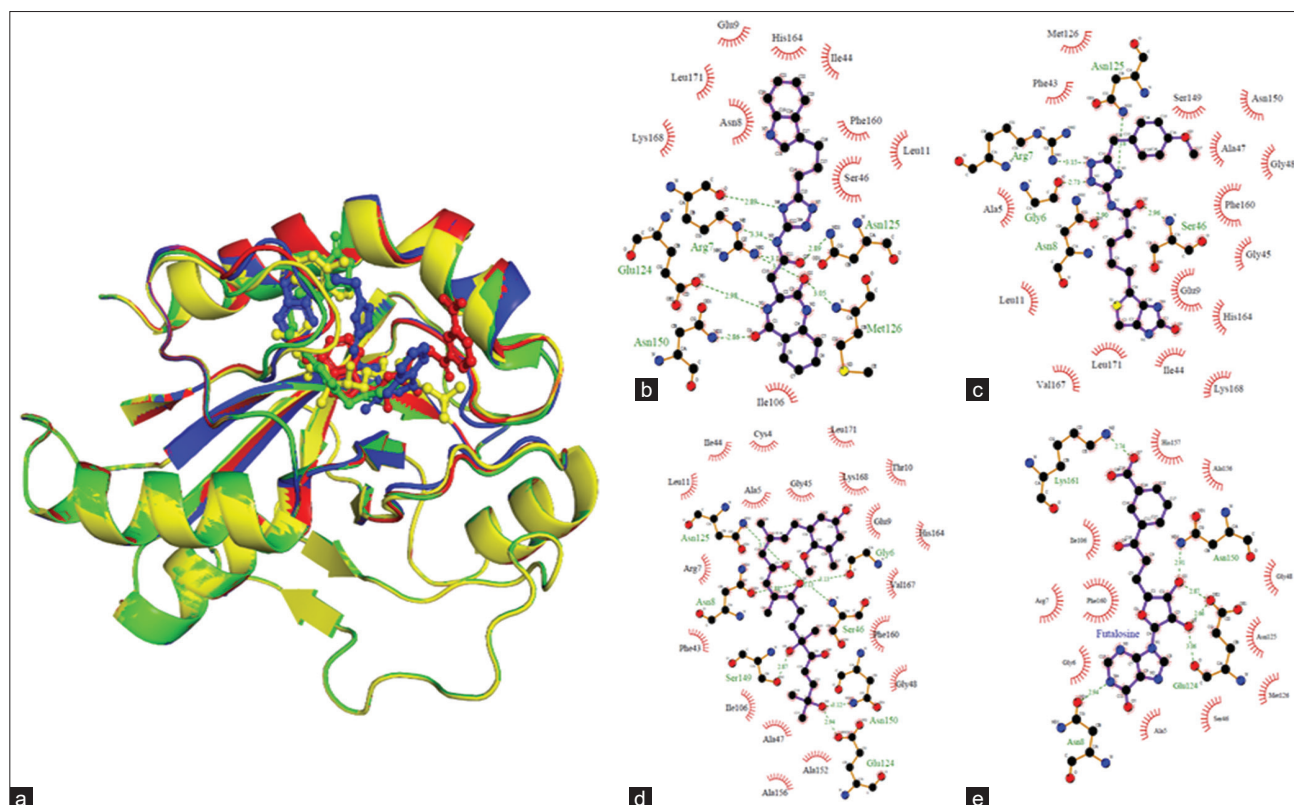


Fig. 6: (a) Superposition of the best-docked poses of MqnB in complex with lead1, lead2, lead3, and substrate. Ligplot representation showing active site interactions of MqnB complexed with (b) lead1 (c) lead2 (d) lead3, and (e) substrate

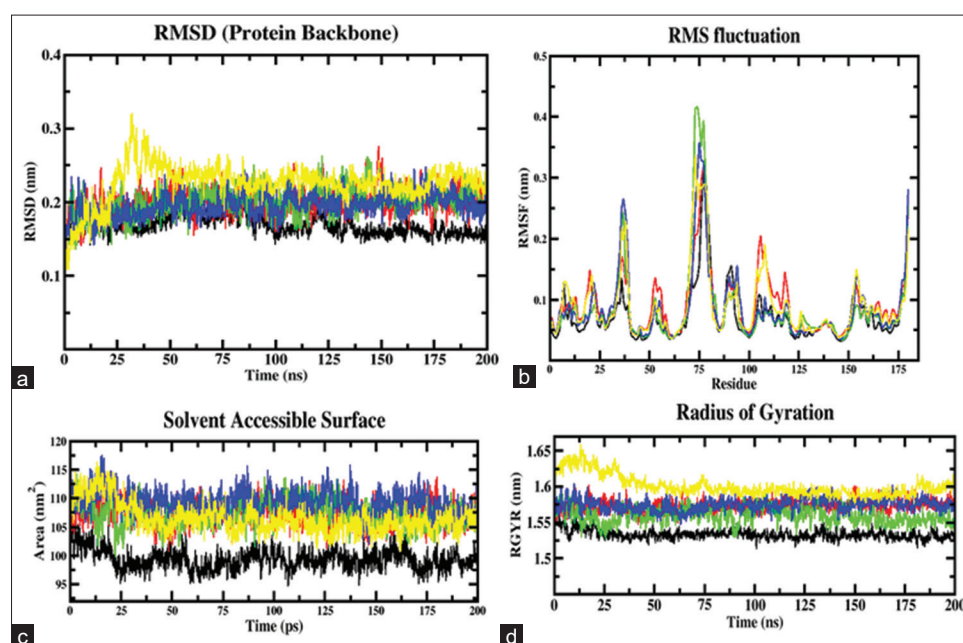


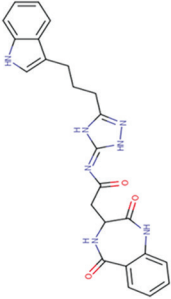
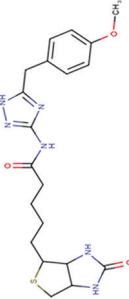
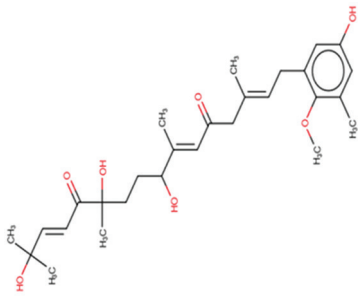
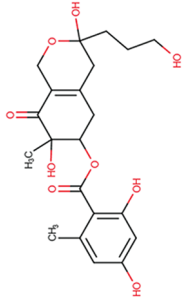
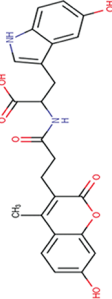
Fig. 7: Molecular dynamics trajectories of MqnB in complex with a substrate (red) lead1 (green), lead2 (blue), lead3 (yellow), and MqnB apo form (black) (A) protein backbone RMSD graph (B) RMSF graph (c) radius of gyration (d) SASA for the 200ns simulation run

Table 2: Virtual screening filters and shortlisted compounds

Filters	No. of compounds
Initial	4,06,760
After high throughput virtual screening docking	21,670
After standard precision docking	2162
After extra precision docking	200

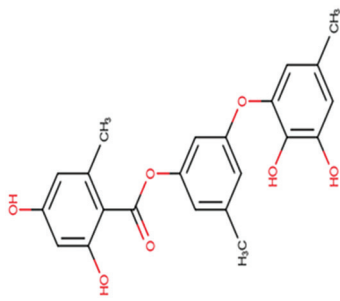
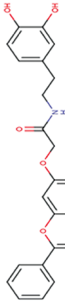
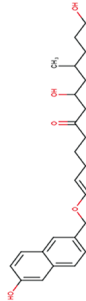
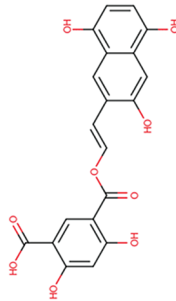
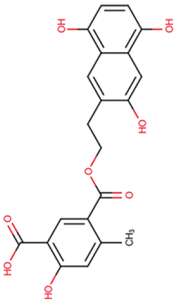
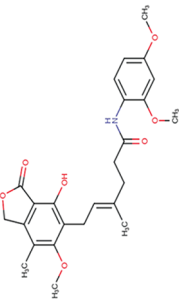
It was observed that the best representative structure for the MqnB-lead1 complex from the cluster showed interactions with residues Glu9, Ser46, and Ser149; lead2 showed interactions Ala5, Gly6, Glu9, Ser46, Asn125, and His164; lead3 showed interactions His164 and MqnB-substrate showed interactions with Arg7, Ser46, Ser149, and Asn150. The residues Ser46 and Ser149 are consistently observed in the conformational snapshots making hydrogen bond interactions at regular 100 ns interval frames, as well as in the clustering analysis.

Table 3. Compounds with COCONUT database ID, 2D structures, molecular weight, and IUPAC name

Compound no.	COCONUT database ID	Structure	Molecular weight	Compound IUPAC
1	CNP0223821		457.48	3-(2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl)-N-[5-[2-(1H-indol-3-yl) ethyl]-1H-1,2,4-triazol-3-yl] propanamide
2	CNP0130440		488.61	8,11,15-trihydroxy-1-(5-hydroxy-2-methoxy-3-methylphenyl)-3,7,11,15-tetramethylhexadeca-2,6,13-triene-5,12-dione
3	CNP0350743		430.52	N-[5-[(4-methoxyphenyl) methyl]-1H-1,2,4-triazol-3-yl]-5-(2-oxo-1,3,3a,4,6,6a-hexahydrothieno[3,4-d]imidazol-4-yl) pentanamide
4	CNP0128509		422.42	[3,7-dihydroxy-3-(3-hydroxypropyl)-7-methyl-8-oxo-1,4,5,6-tetrahydroisochromen-6-yl] 2,4-dihydroxy-6-methylbenzoate
5	CNP0387303		450.42	3-(5-hydroxy-1H-indol-3-yl)-2-[3-(7-hydroxy-4-methyl-2-oxochromen-3-yl) propanoylamino] propanoic acid

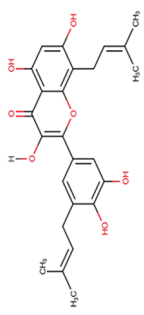
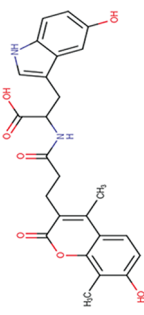
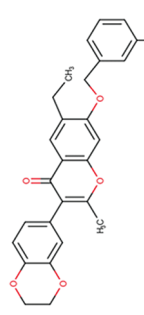

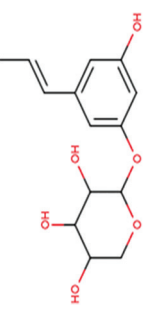
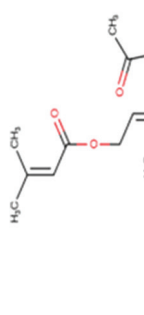
(Contd...)

Table 3. (Continued)

Compound no.	COCONUT database ID	Structure	Molecular weight	Compound IUPAC
6	CNP0354304		396.39	[3-(2,3-dihydroxy-5-methylphenoxy)-5-methylphenyl] 2,4-dihydroxy-6-methyl benzoate
7	CNP0141845		447.43	N-[2-(3,4-dihydroxyphenyl)ethyl]-2-(5-hydroxy-4-oxo-2-phenylchromen-7-yl)oxyacetamide
8	CNP0102189		466.56	5,10-dihydroxy-1-[4-hydroxy-3-[(6-hydroxynaphthalen-2-yl)methoxy]phenyl]-7-methyldecan-3-one
9	CNP0094298		398.32	2,4-dihydroxy-5-[2-(3,5,8-trihydroxynaphthalen-2-yl)ethenoxy]benzoic acid
10	CNP0103692		396.34	2-hydroxy-4-methyl-5-[2-(3,5,8-trihydroxynaphthalen-2-yl)ethenoxy]benzoic acid
11	CNP0398603		455.50	N-(2,4-dimethoxyphenyl)-6-(4-hydroxy-6-methoxy-7-methyl-3-oxo-1H-2-benzofuran-5-yl)-4-methylhex-4-enamide

(Contd...)

Table 3. (Continued)

Compound no.	COCONUT database ID	Structure	Molecular weight	Compound IUPAC
12	CNP0262248		438.47	2-[3,4-dihydroxy-5-(3-methylbut-2-enyl)phenyl]-3,5,7-trihydroxy-8-(3-methylbut-2-enyl)chromen-4-one
13	CNP0385398		464.46	2-[3-(7-hydroxy-4,8-dimethyl-2-oxochromen-3-yl)propanoylamino]-3-(5-hydroxy-1H-indol-3-yl)propanoic acid
14	CNP0372684		484.45	3-(2,3-dihydro-1,4-benzodioxin-6-yl)-6-ethyl-7-[(3-methoxyphenyl)methoxy]-2-methylchromen-4-one
15	CNP0277250		360.35	2-[3-hydroxy-5-[2-(4-hydroxyphenyl)ethenyl]phenoxy]oxane-3,4,5-triol
16	CNP0238215		450.56	2-[7-acetyl-2-(3,4-dihydroxy-4-methyl-pentyl)-5-(hydroxyl)methyl]-2-methyl-6-oxo-cyclohept-4-en-1-yl]vinyl 3-methylbut-2-enoate
17	CNP0244006		458.50	2-[2-[4-[5-[2-(2,5-dihydroxyphenyl)ethyl]-2-hydroxyphenoxy]phenyl]ethyl]benzene-1,4-diol

(Contd...)

Table 3. (Continued)

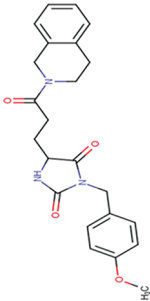
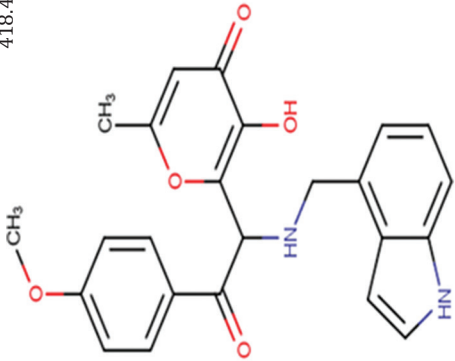
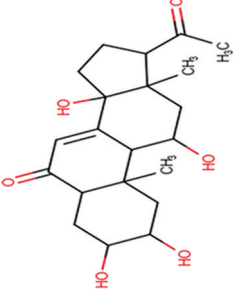
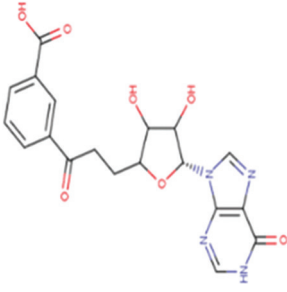
Compound no.	COCONUT database ID	Structure	Molecular weight	Compound IUPAC
18	CNP0403964		407.46	5-[3-(3,4-dihydro-1H-isoquinolin-2-yl)-3-oxopropyl]-3-[(4-methoxyphenyl) methyl] imidazolidine-2,4-dione
19	CNP0237765		418.44	3-hydroxy-2-[1-(1H-indol-4-yl)methylamino]-2-(4-methoxyphenyl)-6-methylpyran-4-one
20	CNP0404494		378.46	17-acetyl-2,3,11,14-tetrahydroxy-10,13-dimethyl-2,3,4,5,9,11,12,15,16,17-decahydro-1H-cyclopenta[a]phenanthren-6-one
Substrate	Futalosine		414.4	3-[3-[(2R,3S,4R,5R)-3,4-dihydroxy-5-(6-oxo-3H-purin-9-yl) oxolan-2-yl] propanoyl] benzoic acid

Table 4: Molecular docking results showing score, energy, hydrogen bond, and hydrophobic interactions

Compound no.	Docking score (kcal/mol)	Glide Energy (kcal/mol)	Hydrogen bond interactions	Hydrophobic interactions
1	-10.51	-84.14	Arg7, Glu12, Asn125, Met126, Asn150	Asn8, Glu9, Leu11, Ser46, Ile106, His164, Lys168, Leu171
2	-11.28	-79.79	Gly6, Asn8, Arg7, Ser46, Asn125	Ala5, Glu9, Leu11, Phe43, Ile44, Gly45, Ala47, Gly48, Met126, Ser149, Asn150, Phe160, His164, Val167, Lys168, Leu171
3	-9.04	-78.63	Gly6, Asn8, Ser46, Gly124, Asn125, Ser149, Asn150	Cys4, Ala5, Arg7, Glu9, Thr10, Leu11, Phe43, Ile44, Gly45, Ala47, Gly48, Ile106, Ala152, Ala156, Phe160, His164, Val167, Lys168, Leu171
4	-11.15	-72.87	Gly6, Arg7, Gly48, His107, Gly124, Asn150, Gly153, Leu154	Asn8, Glu9, Ser46, Ala47, Ile106, Thr108, Asn125, Met126, Ala152, Gly156, His157, Phe160
5	-11.52	-72.41	Arg7, Asn8, Tyr50, Ser103, Tyr105, Met126, Glu127	Ser46, Ile106, His107, Thr108, Gly124, Asn125, Asn150, Ala152, Gly156, His157, Phe160
6	-11.16	-72.31	Asn8, Gly124, Glu127, Asn150	Ala5, Gly6, Arg7, Phe43, Ser46, Ile106, His107, Thr108, Asn125, Met126, Ala152, Ala156, His157, Phe160
7	-10.83	-71.62	Arg7, Asn150, Lys161	Asn8, Glu9, Ile44, Gly45, Ser46, Gly48, Ile106, Met126, Phe160, His164, Lys168
8	-11.88	-70.36	Ser46, Asn125, Met126	Asn8, Glu9, Leu11, Phe43, Ile44, Gly45, Ile106, Glu124, His164, Val167, Lys168, Leu171
9	-6.139	-70.27	Arg7, Glu9, Ile44, Ser46, Gly124, Asn125, Asn150	Cys4, Ala5, Thr8, Leu11, Gly45, Ala47, Gly48, Ile106, Phe160, Lys168, Leu171
10	-10.16	-67.61	Gly6, Met126, Ser149, Lys168	Asn8, Glu9, Leu11, Ile44, Ser46, Gly48, Gly124, Asn125, Asn150, Phe160
Futalosine (substrate)	-11.36	-68.26	Asn5, Gly6, Asn150, Lys161	Ala5, Gly6, Arg7, Ser46, Gly48, Ile106, Asn125, Met126, Ala156, His157, Phe160

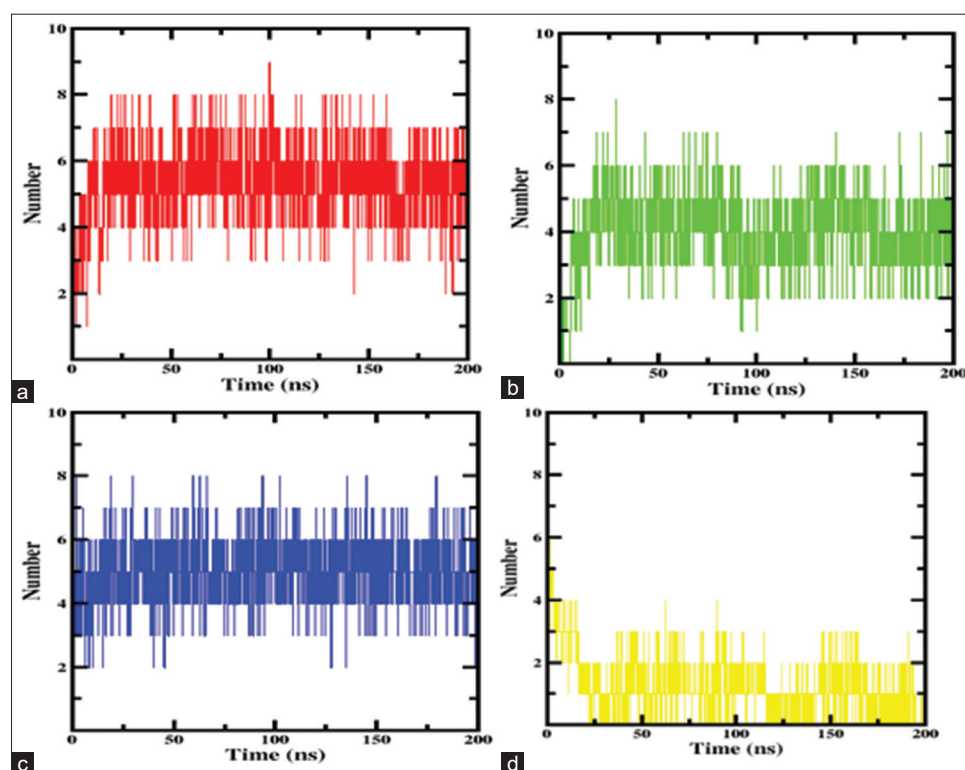


Fig. 8: Graphs show hydrogen bonds made by (a) substrate (b) lead1 (green) (c) lead2 (blue) and (d) lead3 (yellow)

These residues consistently played a role in ligand binding, highlighting their importance in maintaining the stability and affinity of the ligand-protein complexes throughout the simulation.

QikProp analysis

The QikProp results in Table 5 provide vital insights into the drug-like properties of the lead1, lead2, and lead3 compounds, as well as for the substrate. We have evaluated the compounds' potential as therapeutic

candidates by comparing their physicochemical features with those of the substrate.

The QikProp results indicate that the compounds have drug-like characteristics comparable to the substrate. Their molecular weights, SASA, hydrogen bonding properties, lipophilicity, partitioning behavior, and skin permeability are all within acceptable limits. Considering lipophilicity, both QPlogPoct and QPlogPw values are well within

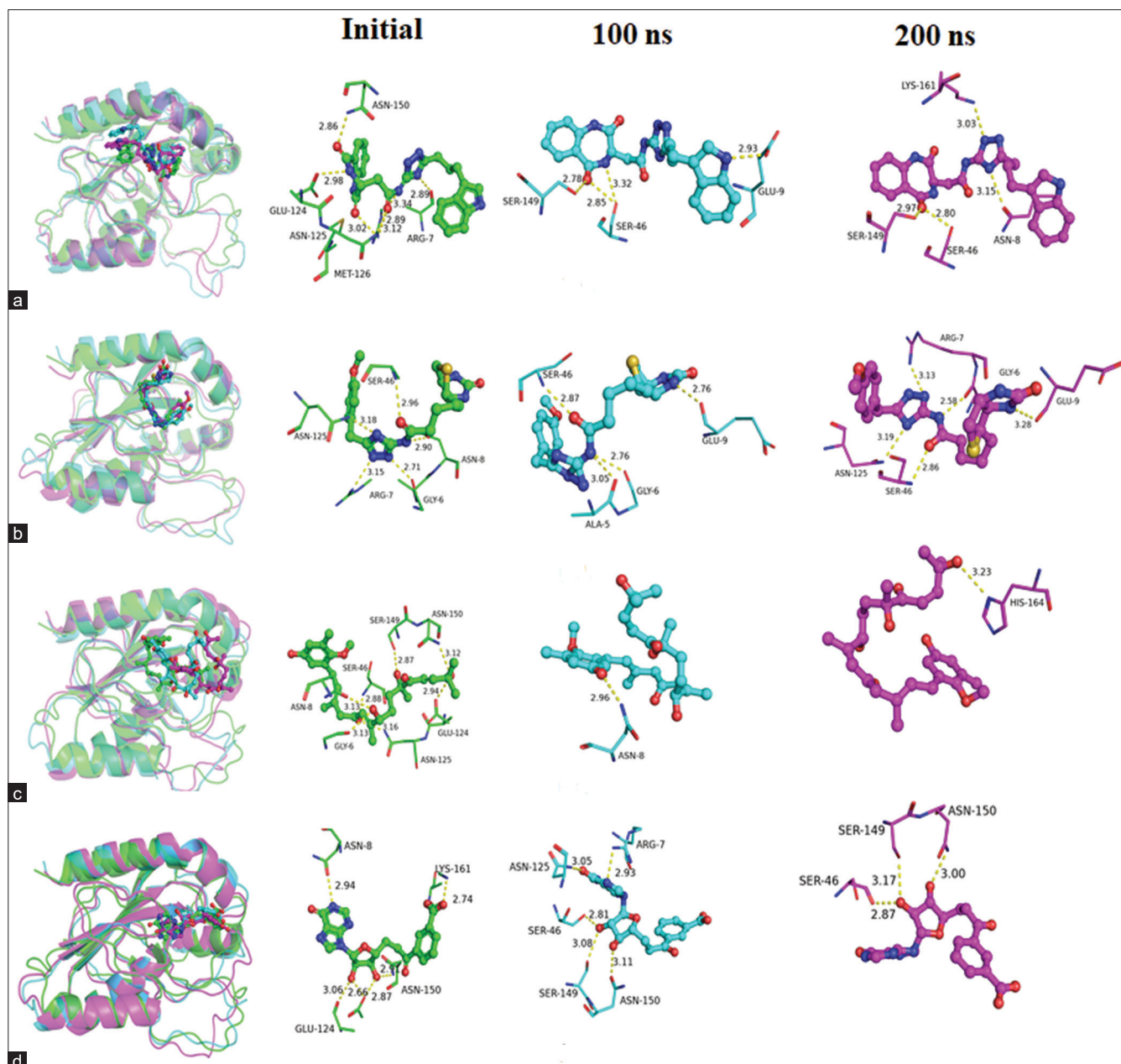


Fig. 9: Superposition of conformational snapshots of (a) lead1 (b) lead2 (c) lead3 (d) substrate at initial (green), 100ns (cyan), and 200ns (magenta)

Table 5: QikProp analysis of lead compounds

QikProp descriptors	Reference range	Lead1	Lead2	Lead3	Substrate
molMW	130–725	457.49	430.52	488.62	414.37
SASA	500–2000	826.18	766.48	772.43	569.41
FOSA	0–750	152.46	367.60	512.15	101.86
Donor HB	0.0–6.0	2	3	4	4
Accept HB	2.0–20.0	6	7.25	8.7	14.1
CNS	–2: inactive to+2: active	–2	–2	–2	–2
HOA	1: low 2: medium 3: high	1	2	1	1
QPlogPoct	8.0–35.0	23.75	23.51	25.64	26.96
QplogPw	4.0–45.0	13.36	13.64	14.02	22.47
QplogPo/w	–2.0–6.5	4.09	3.15	3.76	–0.90
QplogBB	–3.0–1.2	–3.25	–2.37	–2.15	–1.82
QPlogS	–6.5–0.5	–7.65	–6.08	–4.17	–1.89
QPlogKp	–8.0–1.0	–4.42	–4.20	–2.84	–5.98

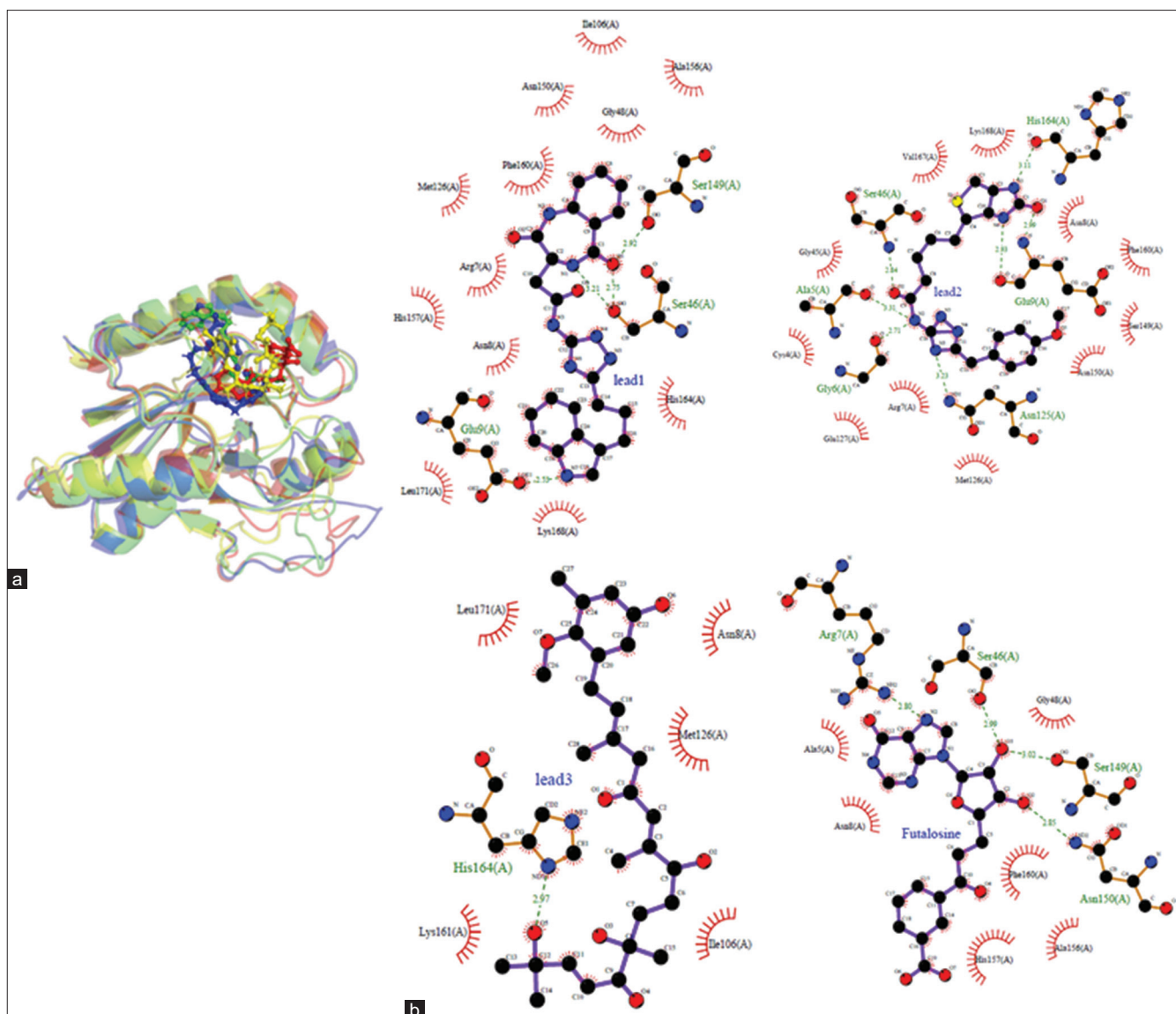


Fig. 10: (a) Superposition of the best representative model of MqnB after molecular dynamics where substrate in red, lead1 in green, lead2 in blue, and lead3 in yellow (b) ligand interactions with active site residues of MqnB

the reference ranges, showing that all molecules, including the substrate, have appropriate lipophilicity. The QPlogPo/w values for all compounds are within the reference range, indicating reasonable partitioning behavior. The aqueous solubility (QPlogS) of all substances falls within the reference range, suggesting that these molecules have high aqueous solubility, which is critical for drug development. The anticipated blood-brain barrier penetration (QPlogBB) values for lead2 and lead3 are within the reference range, indicating that they have a reasonable chance of crossing the blood-brain barrier. Finally, the skin permeability (QPlogKp) values for all compounds are within the reference range, indicating good skin permeability.

CONCLUSION

H. pylori uses the futasoline pathway for bacterial respiration and inhibiting this pathway can help control the growth of *H. pylori* infections, as humans and other gut microbiomes lack this pathway. To identify the novel potential inhibitors, the three-dimensional structure model of MqnB was predicted followed by active site prediction. SBVS was performed for compounds from the COCONUT database. Based on the results of virtual screening, the top 20 compounds were subjected to IFD. Based on a good docking score,

Glide energy, and hydrogen bonding, three lead compounds were chosen. The apo form, MqnB-substrate/lead complexes were subjected to 200 ns MD simulation using GROMACS. MD simulation results showed that the compounds showed minimal deviations in protein RMSD, SASA, and RGYR graphs. Consistent hydrogen bonds were maintained at regular intervals for all the complexes. The lead compounds also exhibited favorable physicochemical properties, suggesting their potential as drug candidates. Notably, our study demonstrated that the lead compounds made interactions with residues such as Ser46 and Ser149 to form hydrogen bonds in both the substrate and lead compounds. These results show that identified lead compounds interact with key residues with MqnB. The lead compounds demonstrate the potential as inhibitors, providing a strong foundation for further research and development of novel drugs against *H. pylori* infections.

ACKNOWLEDGMENT

AA thanks the RUSA 2.0 program for the fellowship. VG thanks the Department of Science and Technology, Government of India, for the INSPIRE fellowship (IF160596). Authors thank DST-FIST, Government of India for the computational facilities sanctioned to the department.

Authors are also thankful to the Schrödinger team for providing a trial version of the Schrödinger suite.

AUTHORS CONTRIBUTION

AA, VG, REV: Software, Validation, Formal analysis, Manuscript writing. GD: Conceptualization, Methodology, Software, Validation, Formal analysis, Manuscript writing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHOR'S FUNDING

The authors received no specific funding for this work.

REFERENCES

- Bentley R, Meganathan R. Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev.* 1982 Sep;46(3):241-80. doi: 10.1128/mr.46.3.241-280.1982, PMID 6127606
- Zhi XY, Yao JC, Tang SK, Huang Y, Li HW, Li WJ. The futasine pathway played an important role in menaquinone biosynthesis during early prokaryote evolution. *Genome Biol Evol.* 2014 Jan;6(1):149-60. doi: 10.1093/gbe/evu007, PMID 24398376
- Seto H, Jinnai Y, Hiratsuka T, Fukawa M, Furihata K, Itoh N, et al. Studies on a new biosynthetic pathway for menaquinone. *J Am Chem Soc.* 2008 Apr 30;130(17):5614-5. doi: 10.1021/ja710207s, PMID 18393499
- Joshi S, Fedoseyenko D, Mahanta N, Manion H, Naseem S, Dairi T, et al. Novel enzymology in futasine-dependent menaquinone biosynthesis. *Curr Opin Chem Biol.* 2018 Dec;47:134-41. doi: 10.1016/j.cbpa.2018.09.015, PMID 30447488
- Li X, Tanner ME. An efficient synthesis of futasine. *Tetrahedron Lett.* 2010 Oct;51(49):6463-5. doi: 10.1016/j.tetlet.2010.10.010
- Cover TL, Blaser MJ. *Helicobacter pylori* in health and disease. *Gastroenterology.* 2009 May;136(6):1863-73. doi: 10.1053/j.gastro.2009.01.073, PMID 19457415
- Coudert E, Gehant S, De Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. *Bioinformatics.* 2023 Jan 1;39(1):btac793. doi: 10.1093/bioinformatics/btac793, PMID 36484697
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. *Nat Methods.* 2022 Jun;19(6):679-82. doi: 10.1038/s41592-022-01488-1, PMID 35637307
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr.* 1993 Apr;26(2):283-91. doi: 10.1107/S0021889892009944
- Schrödinger LLC. Release 2022-3. New York: Schrödinger LLC.; 2022.
- Sorokina M, Merseburger P, Rajan K, Yirik MA, Steinbeck C. COCONUT online: Collection of open natural products database. *J Cheminform.* 2021 Jan 10;13(1):2. doi: 10.1186/s13321-020-00478-9, PMID 33423696
- Sachdeo R, Khanwelkar C, Shete A. *In silico* exploration of berberine as a potential wound healing agent via network pharmacology, molecular docking, and molecular dynamics simulation. *Int J Appl Pharm.* 2024;16(2):188-94. doi:10.22159/ijap.2024v16i2.49922
- Sherman W, Beard HS, Farid R. Use of an induced fit receptor structure in virtual screening. *Chem Biol Drug Des.* 2006 Jan;67(1):83-4. doi: 10.1111/j.1747-0285.2005.00327.x, PMID 16492153
- Laskowski RA, Swindells MB. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model.* 2011 Oct 24;51(10):2778-86. doi: 10.1021/ci200227u, PMID 21919503
- Abraham MJ, Murtola T, Schulz R, P'all S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1-2:19-25. doi.org/10.1016/J.SOFTX.2015.06.001
- Febrina E, Asnawi A, Abdulah R, Lestari K, Supratman U. Identification of flavonoids from *Acalypha indica* L. (*Euphorbiaceae*) as caspase-3 activators using molecular docking and molecular dynamics. *Int J Appl Pharm.* 2022;14(5):162-6. doi: 0.22159/ijap.2022.v14s5.34
- Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: A fast force field generation tool for small organic molecules. *J Comput Chem.* 2011 Aug;32(11):2359-68. doi: 10.1002/jcc.21816, PMID 21541964
- Bussi G, Zykova-Timan T, Parrinello M. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *J Chem Phys.* 2009 Feb 21;130(7):074101. doi: 10.1063/1.3073889, PMID 19239278
- Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.2r3pre. Schrödinger, LLC.; 2020.