

Identification of taurocyamine compound against histidinol-phosphate aminotransferase of *Mycobacterium tuberculosis* through bioinformatics and molecular docking approaches

Jiraporn Yongpisanphop*

Department of Agro-Industrial, Food and Environmental Technology, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

Received: October 17, 2025; Revised: January 17, 2026; Accepted: January 26, 2026

Abstract

Mycobacterium tuberculosis contaminates all types of environments, and it is a biohazard of tuberculosis disease. This concern has provided a strong motivation to search for a natural compound to inhibit the identified target enzyme against that pathogen, which is contaminated in the environment before spreading to other living things and the ecology. The procedure of finding and developing inhibitors begins with target identification. Therefore, the more accurate the identification, the more opportunity for the success of inhibition. Many academic databases and bioinformatics tools have been applied to complete this study via subtractive proteomics, biological networks, and molecular docking approaches. Many databases including Kyoto Encyclopedia of Genes and Genomes (KEGG), the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), the Database of Essential Genes (DEG), the virulence factor database (VFDB), The Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB), and The National Center for Biotechnology Information (NCBI) were used. Eventually, according to the subtractive proteomics results, a total of 30 cytoplasmic proteins were obtained. Finally, histidinol-phosphate aminotransferase was identified as a potential target via protein-protein interaction network analysis. Then molecular docking of 12 compounds obtained from the KEGG database was performed against this enzyme target. Only a taurocyamine compound was proposed successfully as the potential candidate based on its binding energy (-6.25 kcal/mol) compared to an original inhibitor (-7.10 kcal/mol). Moreover, the binding patterns, including hydrogen bonds and hydrophobic interactions between the taurocyamine and targeted enzyme were similar to those of the original inhibitor. The results reveal that the taurocyamine compound might be useful as an antibacterial agent in environmental remediation. However, experimental validation needs to be performed. Moreover, the added benefit from this study is that the *in silico* platform could be applied to identify the potent enzyme target and its inhibitor quickly, cheaply, and credibly.

Keywords: Biological network, Bioinhibitor, Comparative proteomics, Subtractive proteomics, Target identification

1. Introduction

The causative agent of tuberculosis disease, also called TB (tuberculosis), in humans is *Mycobacterium tuberculosis*. This bacterium is a large, non-motile, rod-shaped, aerobic, non-spore-forming, acid-fast bacterium, and has a slow generation period (Singh, 2014; Zhang *et al.*, 2022a). Unfortunately, TB has most likely reverted to being the most common infectious agent-related cause of death worldwide (World Health Organization, 2024). *M. tuberculosis* transmission occurs both with and without patient contact, causing the environment to become a source of *M. tuberculosis* infection (Velayati *et al.*, 2015). Moreover, *M. tuberculosis* contaminates many types of environments, including natural environments (e.g., air, soil, and rivers) and others (e.g., wastewater, raw sewage, sewage effluent, sanatorium sewage, inlet and outlet settling tanks, sewage from cattle farms, sewage sediment, dam sediment, and household dust) (Zhang *et al.*, 2022a). Moreover, this pathogen remains virulent in contaminated

environments, especially in soil (Velayati *et al.*, 2015). Therefore, it is necessary to inhibit *M. tuberculosis* to reduce the likelihood of epidemics and transmission to humans.

The greater the accuracy and speed of target identifying, the greater the rate of inhibitor discovery for pathogens contaminating the environment. Generally, in the past, traditional target identification has relied on experimental study, which is costly, time-consuming, and inaccurate (Zhang *et al.*, 2022b). To date, *in silico* strategies have gained prominence and the acceptance of researchers' attention because they enable the identification of possible targets (Melak and Gakkhar, 2015; Chowdhury *et al.*, 2021; Ashraf *et al.*, 2022; Khan *et al.*, 2022; Zhang *et al.*, 2022b; Duguma and Dinka, 2024). Additionally, computational techniques have facilitated the identification of effective targets for such a pathogen (Ashraf *et al.*, 2022).

Among the diversity of *in silico* strategies, comparative analysis is applied to find the unique enzymes of the pathogen by comparing the metabolic pathways of the

* Corresponding author. e-mail: jiraporn.y@sci.kmutnb.ac.th.

pathogens with the host (Zhang *et al.*, 2022b). Subtractive proteomics analysis is a notable approach to detect novel targets in the pathogen (Chowdhury *et al.*, 2021). It is a powerful concept via sequence subtraction between the host and pathogen proteomes (Ashraf *et al.*, 2022). Moreover, the protein-protein interaction network is also used to identify a hub-bottleneck node, which can be the potential target (Nithaya *et al.*, 2023). The biological network utilized for target identification operates under the assumption that the impact of particular nodes within the network might disseminate along its connections (Zhang *et al.*, 2022b). Generally, the interesting potential targets are based on the following characteristics: 1) an essential protein, 2) a non-homologous protein in the human host, 3) existing in unique metabolic pathways, 4) a cytoplasmic virulent enzyme, and 5) a hub-bottleneck node (Ashraf *et al.*, 2022; Khan *et al.*, 2022; Zhang *et al.*, 2022b; Nithaya *et al.*, 2023). If this enzyme loses function due to inhibition, the pathogen will die. Meanwhile, the two-dimensional (2D) similarity search combined with molecular docking performs well in compound screening (Szilágyi *et al.*, 2021). The 2D similarity search assumes that similar molecules are more likely to have similar biological properties than dissimilar ones (Szilágyi *et al.*, 2021). It is suitable for virtual screening due to its simplicity, ease of interpretation, and reasonably high recovery rate (Szilágyi *et al.*, 2021). The molecular docking is commonly used to calculate the binding affinity between a ligand and receptor in 3D space. Identifying the “best-fit” orientation of a ligand and a specific protein by reducing the system’s free energy helps to predict the molecule’s affinity and efficacy (Abuhamdah *et al.*, 2020).

The objectives in this study are 1) to identify and prioritize the potential target of *M. tuberculosis* H37Rv by a comparative, subtractive proteomics approach and biological network analysis, and 2) to propose a lead compound that could inhibit the potential target by molecular docking.

2. Material and Methods

2.1. Hardware and overall computational methodology

A laptop with an Intel (R) Core™ i7-10750H CPU at 2.60 GHz, 16 GB of RAM, and a 64-bit operating system was used in all steps of this study. Many databases and programs were used in this study (Fig. 1).

2.2. Comparative proteomics analysis

The metabolic pathways of *M. tuberculosis* H37Rv, which is the most studied strain (Melak and Gakkhar, 2015), and *Homo sapiens* were downloaded from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto, 2000). The organism codes, including “mtu” and “has”, were used for searching the metabolic pathways of *M. tuberculosis* and *Homo sapiens* (human), respectively. Those metabolic pathways were compared using the Venny 2.1 online tool (Oliveros, 2007). Subsequently, all proteins of each unique metabolic pathway were collected. The repeated proteins were discarded using the Cluster Database at High Identity with Tolerance (CD-HIT) program with cd-hit function at a 60% threshold (Li and Godzik, 2006). Finally, the non-redundant amino acid sequences in the unique metabolic pathways were obtained as dataset-1.

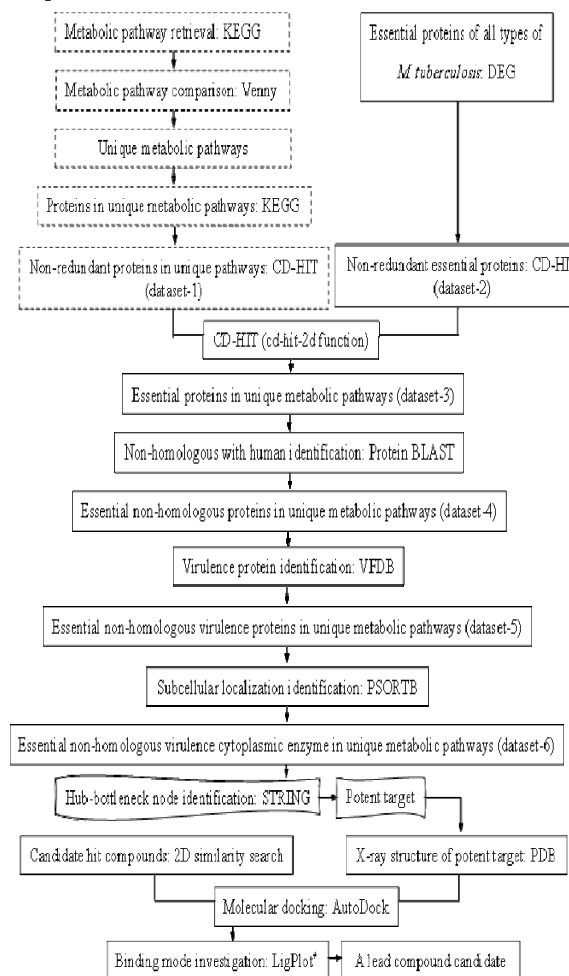


Figure 1. Methodology of the computational experiment. Dash, solid, and sketched boxes represent the comparative, subtractive proteomics, and biological analysis, and molecular docking, respectively.

#

2.3. Subtractive proteomics analysis

First, the essential genes of *M. tuberculosis* were downloaded from the Database of Essential Genes (DEG) (Luo *et al.*, 2021). A total of 614, 771, and 687 essential genes in *M. tuberculosis* H37Rv, H37Rv II, and H37Rv III, respectively, were downloaded (Sasseti *et al.*, 2003; Griffin *et al.*, 2011; Zhang *et al.*, 2012). The repeated sequences were then removed by the cd-hit function at a 60% threshold (Li and Godzik, 2006). After that, the non-redundant essential proteins were kept for the next step as dataset-2. Second, the cd-hit-2D function was utilized to compare dataset-1 and dataset-2 at a 100% threshold (Li and Godzik, 2006). This step was executed twice. The essential proteins in the unique metabolic pathways were then obtained as dataset-3. Third, it was subjected to the Basic Local Alignment Search Tool (BLAST) available on the NCBI database using BLASTp with the expectation threshold 10^{-5} against *Homo sapiens* (human) Protein BLAST (Altschul *et al.*, 1990). The amino acid sequences of the essential non-homologous proteins in the unique metabolic pathways were then obtained as dataset-4. Fourth, for identifying the proteins having the highest virulence factor, the shortlisted proteins of dataset-4 were subjected to BLASTp against the virulence factor database

(VFDB) (Liu *et al.*, 2019) by setting the E-value cut-off to 10^{-5} . The essential non-homologous virulent proteins were obtained as dataset-5. Fifth, PSORTb version 3.0.3 was employed to identify the subcellular location. Gram-positive with an outer membrane was chosen, and the return results are cytoplasm, cytoplasmic membrane, periplasm, outer membrane, extracellular space, and unknown (Yu *et al.*, 2010). These areas are critical because proteins are delivered to specific places once released (Ashraf *et al.*, 2022). Finally, the essential non-homologous virulence cytoplasmic enzymes in unique metabolic pathways were obtained as dataset-6.

2.4. Biological network analysis

The protein-protein interaction (PPI) was generated through the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) version 12.0 (Szklarczyk *et al.*, 2015). Twenty-one cytoplasmic proteins with scores higher than 9.97 of dataset-6 were used to generate the PPI with the *M. tuberculosis*.

2.5. 2D similarity search and molecular docking

The query ligand, 2-(N-morpholino) ethanesulfonic acid (MES), was structurally searched via the SuperNatural_3 (Gallo *et al.*, 2023), the Natural Product Activity and Species Source (NPASS) (Zhao *et al.*, 2023), and the SIMilar COMpound (SIMCOMP) (Hattori *et al.*, 2010). For SIMCOMP, the query ligand was subjected to a target database, including KNApSACk (Afendi *et al.*, 2012) and compounds. Each database used a structure search with the default parameters. The x-ray structure of the enzyme (PDB_ID: 4R8D) was used as a receptor (Nasir *et al.*, 2016), and was visualized by Discovery Studio Visualizer (Biovia, 2024). AutoDock 4.2.6 and AutoDockTools 1.5.7 (Morris *et al.*, 2009) was used for molecular docking. The optimal parameters were screened prior to redocking (data not shown). Redocking was performed so as to calculate the root mean square deviation (RMSD) value. The lower the RMSD value, the smaller the error deviation in the docking (Frimayanti *et al.*, 2021). If this value is less than 2 Å, it means that the method is credible (Shamsian *et al.*, 2024). The grid box size was 40 Å x 40 Å x 40 Å. The grid center was X: 12.762, Y: -4.456, and Z: -15.695. The number of individuals in the population, the maximum number of energy evaluations, and the hybrid Genetic Algorithm with Local Search (GA-LS) runs were set to 300, 2500000, and 50, respectively (Mishra *et al.*, 2023). The other parameters were set to default values. For candidate docking, all compounds having structural similarity with MES were docked in the receptor as well as the redocking technique. For post-docking analysis, enzyme-inhibitor interactions were analyzed by Ligplot⁺ tool (Laskowski and Swindells, 2011).

3. Results and Discussion

3.1. Proteins in unique metabolic pathways

The comparative analysis has now helped in generating a set of proteins existing in the unique metabolic pathways of *M. tuberculosis*. A total of 137 and 360 metabolic pathways were found in *M. tuberculosis* and humans, respectively. Consequently, 51 unique metabolic pathways of *M. tuberculosis* were identified, as shown in Fig. 2 and

Table 1. Because they are absent in humans (Khan *et al.*, 2022), they are important for *M. tuberculosis* in survival and growth. Most of them are involved in the compound biosynthesis and degradation, such as lysine and peptidoglycan biosynthesis, and also benzoate and flavonoid degradation. A total of 1113 proteins of unique pathways were obtained. The 540 redundant proteins were removed. A total of 573 proteins in unique metabolic pathways were identified. They participate in different cellular functions of the pathogen (Duguma and Dinka, 2024).

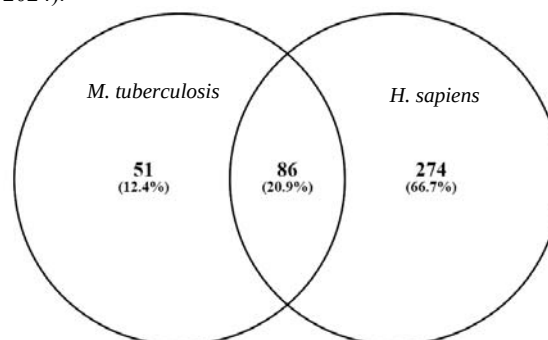


Figure 2. Metabolic pathways between pathogen and human using Venny diagram.

3.2. Subtractive proteomics analysis

Subtractive proteomics analysis has helped in identifying the set of proteins (essential non-homologous virulent enzymes in unique metabolic pathways) that could be a potential target, as shown in Table 2. A total of 1073 proteins were identified as essential proteins. They are required for the survival of *M. tuberculosis*. A total of 315 essential proteins were found in unique metabolic pathways. These proteins are indispensable for the survival and growth of the pathogen without changing the host's metabolic pathways (Melak and Gakkhar, 2015; Khan *et al.*, 2022). A total of 135 essential proteins in unique metabolic pathways showing sequence similarity with humans were removed so as to eliminate the risk of causing human toxicity (Melak and Gakkhar, 2015). Therefore, a total of 180 essential non-homologous proteins were selected. Non-homologous proteins are proteins of the pathogen that have no similar amino acid sequences in humans (Duguma and Dinka, 2024). Furthermore, a total of 50 proteins were identified as essential non-homologous and virulent proteins. Targeting the virulent protein is an interesting technique, because they determine the pathogenesis that enables them to cause infection (Liu *et al.*, 2019).

Among 50 essential non-homologous virulent proteins, thirty of them were found to be cytoplasmic proteins. They are more suitable for inhibitor targets, because the proteins found in the cell membrane are difficult to purify (Duguma and Dinka, 2024). Moreover, cytoplasmic proteins are frequently utilized as therapeutic targets (Khan and Uddin, 2022).

3.3. A hub-bottleneck protein identification

The refined lists of proteins were generated by the PPI network to prioritize potential targets (Fig. 3). The STRING results showed different nodes (proteins) and edges (interactions) of each protein. They also showed that the prioritized target may act as a hub node.

Table 1. List of proteins in unique metabolic pathways of *M. tuberculosis*.

No.	Metabolic pathways name	Pathway_ID	Protein number	No.	Metabolic pathways name	Pathway_ID	Protein number
1	Mycolic acid biosynthesis	mtu00074	36, 28	37	Caprolactam degradation	mtu00930	23, 1
2	Secondary bile acid biosynthesis	mtu00121	1, 1	38	Degradation of flavonoids	mtu00946	1, 0
3	Monobactam biosynthesis	mtu00261	7, 7	39	Steroid degradation	mtu00984	13, 13
4	Lysine biosynthesis	mtu00300	13, 9	40	“Biosynthesis of various plant secondary metabolites”	mtu00999	2, 1
5	Penicillin and cephalosporin biosynthesis	mtu00311	2, 2	41	“Biosynthesis of siderophore nonribosomal peptides”	mtu01053	9, 9
6	Carbapenem biosynthesis	mtu00332	2, 2	42	“Biosynthesis of secondary metabolites”	mtu01110	341, 236
7	Chlorocyclohexane and chlorobenzene degradation	mtu00361	5, 5	43	“Microbial metabolism in diverse environments”	mtu01120	219, 35
8	Benzoate degradation	mtu00362	40, 40	44	Degradation of aromatic compounds	mtu01220	12, 0
9	Fluorobenzoate degradation	mtu00364	2, 0	45	Nitrogen cycle	mtu01310	6, 0
10	Novobiocin biosynthesis	mtu00401	3, 3	46	Beta-Lactam resistance	mtu01501	7, 5
11	Cyanoamino acid metabolism	mtu00460	6, 6	47	Vancomycin resistance	mtu01502	6, 2
12	Streptomycin biosynthesis	mtu00521	11, 11	48	Cationic antimicrobial peptide (CAMP) resistance	mtu01503	2, 2
13	Polyketide sugar unit biosynthesis	mtu00523	7, 1	49	Two-component system	mtu02020	58, 36
14	Acarbose and validamycin biosynthesis	mtu00525	4, 0	50	Quorum sensing	mtu02024	24, 11
15	Lipopolysaccharide biosynthesis	mtu00540	3, 3	51	Bacterial secretion system	mtu03070	14, 5
16	O-Antigen nucleotide sugar biosynthesis	mtu00541	19, 13	Summation of protein			1113, 573
17	O-Antigen repeat unit biosynthesis	mtu00542	1, 1	Remark: Regular fonts represent all proteins. Bold fonts represent the unique protein.			
18	Exopolysaccharide biosynthesis	mtu00543	2, 1	Table 2. Stepwise applied in this study.			
19	Peptidoglycan biosynthesis	mtu00550	17, 15	Step used in this study			<i>M. tuberculosis</i>
20	Teichoic acid biosynthesis	mtu00552	2, 0	Essential proteins from DEG			2072
21	Lipoarabinomannan (LAM) biosynthesis	mtu00571	13, 13	Non-redundant essential proteins			1073
22	Arabinogalactan biosynthesis - Mycobacterium	mtu00572	14, 10	Proteins in unique metabolic pathways			573
23	Dioxin degradation	mtu00621	4, 0	Non-redundant essential proteins in unique metabolic pathways			315
24	Xylene degradation	mtu00622	4, 0	Non-redundant essential non-homologous proteins in unique pathways			180
25	Toluene degradation	mtu00623	2, 0	Virulence proteins			50
26	“Chloroalkane and chloroalkene degradation”	mtu00625	10, 7	Subcellular localization:			
27	“Naphthalene degradation”	mtu00626	4, 0	- Cytoplasmic			30
28	Aminobenzoate degradation	mtu00627	30, 9	- Cytoplasmic membrane			14
29	Styrene degradation	mtu00643	5, 0	- Periplasmic			-
30	“C5-Branched dibasic acid metabolism”	mtu00660	9, 9	- Outer membrane (Cell wall)			-
31	“Methane metabolism”	mtu00680	21, 18	- Extracellular			-
32	“Carbon fixation by Calvin cycle”	mtu00710	11, 0	- Unknown			5
33	“Other carbon fixation pathways”	mtu00720	37, 0	- Unknown (multiple sites)			1
34	Limonene degradation	mtu00903	5, 2	STRING is claimed to be a web resource database dedicated to experimentally known and predicted PPI, including both physical and functional interactions (Melak and Gakkhar, 2015; Khan <i>et al.</i> , 2022). This network composed of 21 nodes (protein) and 15 edges (interaction). The node hisC1 was a hub-bottleneck node, presenting			
35	Carotenoid biosynthesis	mtu00906	1, 1				
36	“Pinene, camphor and geraniol degradation”	mtu00907	23, 0				

high degree and betweenness (Yu *et al.*, 2007; Nithaya *et al.*, 2023). This node is very important to the cell (Yu *et al.*, 2007). Deletion of this node is lethal to the cell, leading to it being a potential target (Nithaya *et al.*, 2023). Furthermore, Raman *et al.* (2008) reported that hisC1 is a potential target of *M. tuberculosis*.

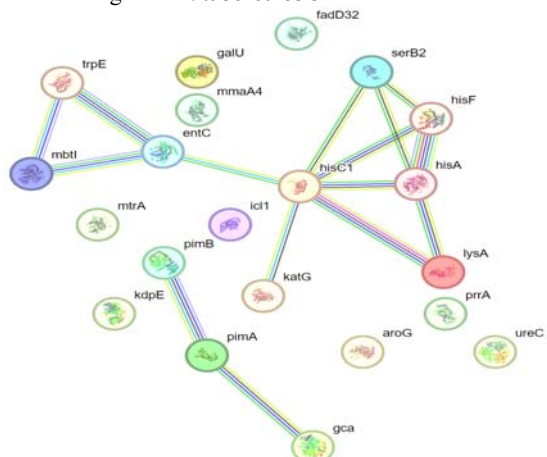


Figure 3. Protein-protein interaction network for 21 proteins by STRING.

3.4. Significance of the opted potential protein (hisC1)

The hisC1 is a histidinol-phosphate aminotransferase (EC: 2.6.1.9). It is classified as required for optimal growth (Lunardi *et al.*, 2013). It is the seventh enzyme in the L-histidine biosynthetic pathway, converting imidazole-acetol-P to L-histidinol-P (Kulis-Horn *et al.*, 2014). Thus, inhibition of this enzyme leads to the blockage of histidine biosynthesis, ultimately disrupting bacterial growth and survival, because histidine is an essential amino acid for cellular processes (Singh *et al.*, 2025). The histidine biosynthesis pathway comprises ten enzymatic steps that convert phosphoribosyl pyrophosphate (PRPP) and ATP to L-histidine, and it appears to be conserved in many living things (Kulis-Horn *et al.*, 2014). Importantly, this pathway is absent in mammals, making it a promising target (Lunardi *et al.*, 2013). In other words, this inhibition is a key strategy for developing novel antibiotics against pathogenic bacteria. Finally, the x-ray structure of this enzyme with PDB_ID: 4R8D has also supported its targetability and use for molecular docking (Fig. 4A and 4B).

A



B

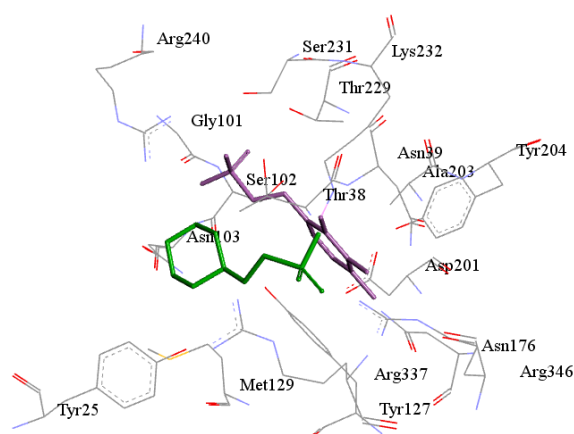


Figure 4. X-ray structure of hisC1 enzyme. (A) Sphere is the binding pocket around the ligand (MES: green) and cofactor (PLP: pink) within a radius of 4 Å. (B) A close-up view of the binding pocket, created with Discovery Studio.

3.5. Molecular docking of 2D similarity compounds

None of the compounds was obtained from NPASS and SuperNatural_3. A total of 12 compounds were obtained from KEGG, showing similarity with the original ligand (MES). Four compounds were obtained from KnapSack, whereas eight compounds were obtained from the compound target database. The similarity scores were in the range of 0.40-0.60. The RMSD value of redocking was 1.1035 Å (Fig. 5). Thus, this technique is valid and could be implemented in compound docking. The docked screening led to the identification of 12 compounds with binding energies ranging from -2.76 to -6.37 kcal/mol (Table 3). Only the top 5 compounds were selected with binding energies spanning between -5.97 and -6.37 kcal/mol, which was similar to those of redocking (-7.10 kcal/mol). The lower the value of the binding free energy, the stronger the attachment between the ligand and

receptor (Jain *et al.*, 2021). Moreover, the results also revealed that the taurocyamine compound could inhibit this enzyme as well as a specific inhibitor (MES) based on the estimated inhibition constant (K_i) value of 26.18 $\mu\text{mol/L}$ and 6.22 $\mu\text{mol/L}$, respectively. Increased inhibition arises from less dissociation, as evidenced by a lower K_i value (Pandey and Verma, 2020).

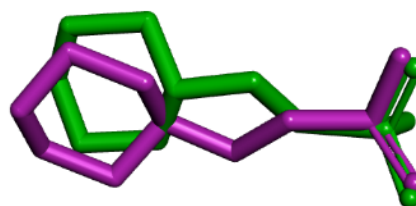


Figure 5. The superimposed co-crystallized MES (in pink) over redock (in green), created by Discovery Studio.

3.6. Post docking analysis

LigPlot⁺ version 2.3.1 was used for binding mode analysis (Laskowski and Swindells, 2011). For redocking complex, the results showed that Tyr25, Thr38, Asn39, Asn103, Tyr127, Met129, Asn176, Arg337, and Arg346 were common residues in the binding pocket (Fig. 6). It showed 6 hydrogen bonds with 5 residues, such as Asn39, Tyr127, Asn176, Arg337, and Arg346. It also displayed hydrophobic interaction with 4 residues, such as Tyr25, Thr38, Asn103, and Met129. These interactions determine the drug-receptor affinity (Sahu *et al.*, 2008). These results are similar to the results of Naris *et al.* (2016). Amino acids involved in binding between the MES ligand with histidinol-phosphate aminotransferase are Tyr25, Thr38, Asn39, Tyr127, Met129, Asn176, Pro260, Tyr261, Arg 337, and Arg346. Moreover, the residues involved in the active site region are Asn103, Tyr127, Met129, Lys232, Arg337, and Arg346 (Nasir *et al.*, 2016).

Table 3. Molecular docking of compounds posing the 2D structure similar to MES.

No.	Compound_ID	Binding energy (kcal/mol)	Ki value (μmol/L)
Lig_1	C00050208	-6.27	25.54
Lig_2	C00048188	-3.85	1500.00
Lig_3	C00000767	-5.42	107.32
Lig_4	C00001530	-5.94	43.88
Lig_5	C01616	-6.37	21.50
Lig_6	C05353	-6.10	33.81
Lig_7	C14452	-3.47	2850.00
Lig_8	C01959	-6.25	26.18
Lig_9	C22977	-5.03	204.14
Lig_10	C06771	-2.76	9410.00
Lig_11	C19283	-4.85	278.49
Lig_12	C00519	-3.95	1280.00
Lig_13	MES (redocking)	-7.10	6.22

Remark: Compounds with bold font were selected for binding pattern analysis.

For hydrogen bonds, the pattern is identical, while hydrophobic interactions are also similar with 3 residues, such as Tyr25, Thr38, and Met129. These results, therefore, also confirmed that the docking technique was valid. Among ligand candidates, only Lig_8 (2-(diaminomethylideneamino) ethanesulfonic acid: taurocyamine) could identically form hydrogen bonds with redocked results (Fig. 7), while the remaining compound could not form an interaction pattern similar to the redocked molecule.

Hydrogen bonds help in proper orientation of the compound, and they are primary contributors to the specificity of molecular recognition to make maximum interaction (Sahu *et al.*, 2008; de Freitas and Schapira, 2017). Moreover, Lig_8 could form hydrophobic interactions with Tyr25, Thr38, and Met129. Hydrophobic binding type is essential in ligand-protein interaction due to their predominant and major contribution (Sahu *et al.*, 2008). Therefore, the taurocyamine was selected as the best candidate.

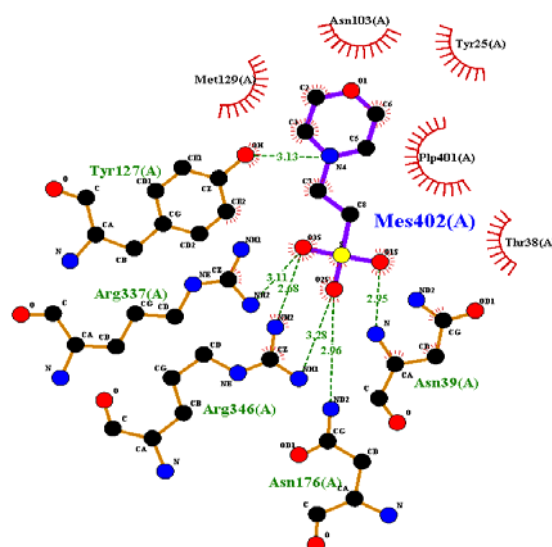


Figure 6. The binding interactions between enzyme and Lig_13. Hydrogen bonds are shown in green dotted lines, whereas hydrophobic interactions are shown in spikes. This figure was created by LigPlot⁺.

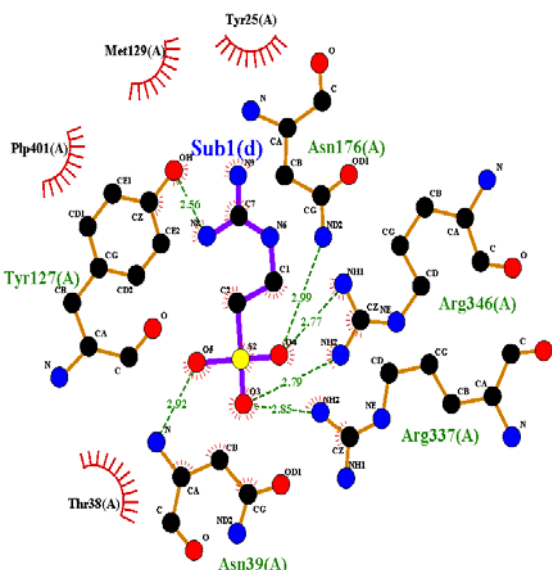


Figure 7. The binding interactions between enzyme and Lig_8. Hydrogen bonds are shown in green dotted lines, whereas hydrophobic interactions are shown in spikes. This figure was created by LigPlot⁺.

The taurocyamine has a molecular formula of $C_3H_9N_3O_3S$ with a molecular weight of 167.19 g/mol based on the KEGG database. It consists of a guanidine group attached to an ethanesulfonic acid molecule. Remarkably, the guanidine is a functional group of antibacterial agents due to hydrogen bond interaction with the target (Kim *et al.*, 2021). Based on the author's knowledge, the taurocyamine is not reported to be an antibacterial agent.

4. Conclusion

In this study, comparative and subtractive proteomics combined with many bioinformatics resources led to obtaining the histidinol-phosphate aminotransferase as a potent enzyme target. Then the taurocyamine compound was discovered as a potent inhibitor. These results may lead to stopping the spread of this pathogen that

contaminates the environment. However, the proposed inhibitor and target need to be studied further in depth and experimentally validated both *in vivo* and *in vitro*. Notably, applying the computational platform from this study may enable swift and economical environmental remediation in the future with the discovery of effective natural compounds against various other pathogens contaminating the environment before they become widespread.

Acknowledgements

The author would like to thank the owner of the academic databases and tools.

Funding

The author declares that no fund was received during the preparation of this manuscript.

Conflicts of interest

The author declares there are no conflicts of interest.

References

- Abuhamdah S, Abuhamdah R, Howes MJ, Uttley G and Chazot PL. 2020. A molecular docking study of *Aloysia citrodora* Palau. Leaf essential oil constituents towards human acetylcholinesterase: Implications for Alzheimer's disease. *Jordan J Biol Sci.*, **13**: 575-580.
- Afendi FM, Okada T, Yamazaki M, Hirai-Morita A, Nakamura Y, Nakamura K, Ikeda S, Takahashi H, Md. Altaf-UI-Amin Darusman LK, Saito K and Kanaya S. 2012. KNApSACk family databases: Integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.*, **53**: e1.
- Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.*, **215** (3): 403-410.
- Ashraf B, Atiq N, Khan K, Wadood A and Uddin R. 2022. Subtractive genomics profiling for potential drug targets identification against *Moraxella catarrhalis*. *PLoS One*, **17**(8): e0273252.
- Biovia, Dassault Systèmes, The Discovery Studio Visualizer, 2024, San Diego: Dassault Systèmes.
- Chowdhury UF, Saba AA, Sufi AS, Khan AM, Sharmin I, Sultana A and Islam MO. 2021. Subtractive proteomics approach to unravel the druggable proteins of the emerging pathogen *Waddlia chondrophila* and drug repositioning on its MurB protein. *Heliyon*, **7**(6): e07320.
- Duguma T and Dinka H. 2024. *In silico* identification and characterization of drug targets in *Streptococcus pneumoniae* ATCC 700669 (Serotype 23F) by subtractive genomics. *Biomed Res Int.*, **2024**: 5917667.
- de Freitas FR and Schapira M. 2017. A systematic analysis of atomic protein-ligand interactions in the PDB. *Med Chem Comm.*, **8**(10): 1970-1981.
- Frimayanti N, Nasution MR and Etavianti E. 2021. Molecular docking and molecular dynamic simulation of 1,5-benzothiazepine chalcone derivative compounds as potential inhibitors for Zika virus helicase. *Jurnal Riset Kimia.*, **12**(1): 44-52.
- Gallo K, Kemmler E and Goede A. 2023. SuperNatural 3.0-a database of natural products and natural product-based derivatives. *Nucleic Acids Res.*, **51**(D1): D654-D659.
- Griffin JE, Gawronski JD, DeJesus MA, Ioerger TR, Akerley BJ and Sasseti CM. 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.*, **7**(9): e1002251.
- Hattori M, Tanaka N, Kanehisa M and Goto S. 2010. SIMCOMP/SUBCOMP: Chemical structure search servers for network analyses. *Nucleic Acids Res.*, **38**: W652-W656.
- Jain AS, Sushma P, Dharmashekar C, Beelagi MS, Prasad SK, Shivamallu C, Prasad A, Syed A, Marraiki N and Prasad KS. 2021. *In silico* evaluation of flavonoids as effective antiviral agents on the spike glycoprotein of SARS-CoV-2. *Saudi J Biol Sci.*, **28**(1): 1040-1051.
- Kanehisa M and Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**(1): 27-30.
- Khan K and Uddin R. 2022. Integrated bioinformatics based subtractive genomics approach to decipher the therapeutic function of hypothetical proteins from *Salmonella typhi* XDR H-58 strain. *Biotechnol Lett.*, **44**: 279-298.
- Khan K, Jalal K, Khan A, Al-Harrasi A and Uddin R. 2022. Comparative metabolic pathways analysis and subtractive genomics profiling to prioritize potential drug targets against *Streptococcus pneumoniae*. *Front Microbiol.*, **12**: 796363.
- Kim SH, Semenya D and Castagnolo D. 2021. Antimicrobial drugs bearing guanidine moieties: A review. *Eur J Med Chem.*, **216**: 113293.
- Kulis-Horn RK, Persicke, M and Kalinowski J. 2014. Histidine biosynthesis, its regulation and biotechnological application in *Corynebacterium glutamicum*. *Microb Biotechnol.*, **7**(1): 5-25.
- Laskowski RA and Swindells MB. 2011. LigPlot⁺: Multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model.*, **51**: 2778-2786.
- Li W and Godzik A. 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**(13): 1658-1659.
- Liu B, Zheng D, Jin Q, Chen L and Yang J. 2019. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.*, **47**: D687-D692.
- Lunardi J, Nunes JE, Bizarro CV, Basso LA, Santos DS and Machado P. 2013. Targeting the histidine pathway in *Mycobacterium tuberculosis*. *Curr Top Med Chem.*, **13**(22): 2866-2884.
- Luo H, Lin Y, Liu T, Lai FL, Zhang CT, Gao F and Zhang R. 2021. DEG 15, an update of the Database of Essential Genes that includes built-in analysis tools. *Nucleic Acids Res.*, **49**(D1): D677-D686.
- Melak T and Gakkhar S. 2015. Maximum flow approach to prioritize potential drug targets of *Mycobacterium tuberculosis* H37Rv from protein-protein interaction network. *Clin Transl Med.*, **4**(1): 61.
- Mishra D, Mishra A, Rai SN, Singh SK, Vamanu E and Singh MP. 2023. In silico insight to identify potential inhibitors of BUB1B from mushroom bioactive compounds to prevent breast cancer metastasis. *Front Biosci (Landmark Ed)*, **28**(7): 151.
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS and Olson AJ. 2009. Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.*, **16**: 2785-2791.
- Nasir N, Anant A, Vyas R and Biswal BK. 2016. Crystal structures of *Mycobacterium tuberculosis* HspAT and ArAT reveal structural basis of their distinct substrate specificities. *Sci Rep.*, **6**: 18880.

- Nithaya C, Kiran M and Nagarajaram HA. 2023. Dissection of hubs and bottlenecks in a protein-protein interaction network. *Comput Biol Chem.*, **102**: 107802.
- Oliveros JC. 2007. Venny. An interactive tool for comparing lists with Venn's diagrams. <https://bioinfogp.cnb.csic.es/tools/venny/index.html>. (Oct 15, 2024).
- Pandey AK and Verma S. 2020. An in-silico evaluation of dietary components for structural inhibition of SARS-Cov-2 main protease, *J Biomol Struct Dyn.*, **40(1)**: 36-142.
- Raman K, Yeturu K and Chandra N. 2008. TargetTB: A target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol.*, **2**: 109.
- Sasseti CM, Boyd DH and Rubin EJ. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol.*, **48**: 77-84.
- Sahu VK, Khan AKR, Singh RK and Singh PP. 2008. Hydrophobic, polar and hydrogen bonding based drug-receptor interaction of tetrahydroimidazobenzodiazepinones. *Am J Immunol.*, **4(3)**: 33-42.
- Shamsian S, Sokouti B and Dastmalchi S. 2024. Benchmarking different docking protocols for predicting the binding poses of ligands complexed with cyclooxygenase enzymes and screening chemical libraries. *BioImpacts*, **14(2)**: 29955.
- Singh R. 2014. Determination of minimum inhibitory concentration of cycloserine in multidrug resistant *Mycobacterium tuberculosis* isolates. *Jordan J Biol Sci.*, **7(2)**: 139-145.
- Singh A, Tanwar M, Singh TP, Sharma S and Sharma P. 2025. Unveiling natural power: Morin and myricetin as potent inhibitors of histidinol-phosphate aminotransferase in drug-resistant *Acinetobacter baumannii*. *ACS Omega*, **10**: 7920-7936.
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ and von Mering C. 2015. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.*, **43(D1)**: D447-D452.
- Szilágyi K, Flachner B, Hajdú I, Szaszko M, Dobi K, Lőrincz Z, Cseh S and Dormán G. 2021. Rapid identification of potential drug candidates from multi-million compounds' repositories. Combination of 2D similarity search with 3D ligand/structure-based methods and *in vitro* screening. *Molecules*, **26**: 5593.
- Velayati AA, Farnia P, Mozafari M, Malekshahian D, Farahbod AM, Seif S, Rahideh S and Mirsaedi M. 2015. Identification and genotyping of *Mycobacterium tuberculosis* isolated from water and soil samples of a metropolitan city. *Chest*, **147(4)**: 1094-1102.
- World Health Organization. 2024. Global Tuberculosis Report 2024. <https://www.who.int/publications/i/item/9789240101531>. (Jan 16, 2025).
- Yu H, Kim PM, Sprecher E, Trifonov V and Gerstein M. 2007. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput Biol.*, **3(4)**: e59.
- Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ and Brinkman FSL. 2010. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*, **26(13)**: 1608-1615.
- Zhang YJ, Ierger TR, Huttenhower C, Long JE, Sasseti CM, Sacchettini, JC and Rubin EJ. 2012. Global assessment of genomic regions required for growth in *Mycobacterium tuberculosis*. *PLoS Pathog.*, **8(9)**: e1002946.
- Zhang H, Liu M, Fan W, Sun S and Fan X. 2022a. The impact of *Mycobacterium tuberculosis* complex in the environment on one health approach. *Front Public Health*, **10**: 994745.
- Zhang X, Wu F, Yang N, Zhan X, Liao J, Mai S and Huang Z. 2022b. *In silico* methods for identification of potential therapeutic targets. *Interdiscip Sci.*, **14**: 285-310.
- Zhao H, Yang Y, Wang S, Yang X, Zhou K, Xu C, Zhang X, Fan J, Hou D, Li X, Lin H, Tan Y, Wang S, Chu XY, Zhuoma D, Zhang F, Ju D, Zeng X and Chen YZ. 2023. NPASS database update 2023: Quantitative natural product activity and species source database for biomedical research. *Nucleic Acids Res.*, **51(D1)**: D621-D628.