Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2016, 8(10):25-34

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Insilico Identification of Novel Inhibitors for Shikimate Kinase from *Mycobacterium Tuberculosis* using Natural and Synthetic Ligands

Mustafa Alhaji Isa* and Balwant Kishan Malik

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

---

ABSTRACT

Tuberculosis (TB) remains as a causes of death due to Mycobacterium tuberculosis. Tuberculosis (TB) is a major global health problem. It causes ill-health among millions of people each year and ranks alongside with human immunodeficiency virus (HIV) as a leading cause of death worldwide. Thus, the objective of this study was to determined novel inhibitors of Shikimate kinase from Mycobacterium tuberculosis (MTB) with aim of providing novel therapeutic candidates for treatment of tuberculosis in human. To understand the binding mode analysis and structural features, Shikimate kinase was modelled through homology modeling using Modeller 9.16. The modelled structures was validated using Ramachandran plot which revealed that 97.3% (142/146amino acid) of the entire residues lay in the most favourable regions [A, B, L], 2.7% (4/146amino acid) were in Additional allowed regions [a,b,l,p], while the none of the amino acid residues were found in both Generously allowed regions [~a,~b,~l,~p] and Disallowed regions[XX]. Validation with Pros A server shows a Z-Score of -8.44 which clearly indicates good quality model with high structural integrity, resembling structure determined through Nuclear Magnetic Resonance. The result of docking studies (synthetic ligands) among the several ligands used shows that, only thirteen (13) found to have promising activity against Shikimate kinase with a minimum free binding energy ranges from -11.52 and -9.00kcal/mol, while six natural ligands possess high activities with minimum free binding energy ranges from -11.75 and -8.49 kcal/mol. Therefore, these ligands were recommended for future drugs for the treatment of tuberculosis in human cause by both multidrug-resistant tuberculosis (MDR-TB) strain and extensive drug resistant strain

Keywords: Shikimate Kinase, MTB, Modeling, Docking, Natural and Synthetic Ligands

---

INTRODUCTION

Tuberculosis (TB) is considered as one of the most devastating global public health threat of the 21st century. It is infectious disease that responsible for second cause of death, after human immunodefiency virus (HIV). Resistance cases due to multidrug-resistant tuberculosis (MDR-TB) is a global threat posing a serious challenge to tuberculosis control programme. Since advent of HIV/TB co-infection many people died of tuberculosis [1-3]. TB is one of the most fatal diseases that can spread widely through coughing, sneezing or talking from person with active form of the disease. The organism mainly attack lungs, but it may also spread and affect other organs of the body such as central nervous system, bone, lymph node, joints and genitourinary. With all significant advances, still TB remains a cause of death in developing countries and it is also a leading known infectious disease worldwide. World Health Organization (WHO) reported that 1.3 million HIV negative patients and 38000 HIV positive patients died of tuberculosis, while 250, 000 patients reported to developed multi-drug resistant TB In 2009 [1]. In 2010, the prevalent of TB increase with 8.8 million new cases, while 1.45 million deaths were reported and mostly from developing countries. In 2012 about 450, 000 people had multidrug-resistant tuberculosis (MDR-TB) as compared to 62,000 cases in 2011. Of this total, 9.6% cases were found to have extensively drug-resistant tuberculosis (XDR-TB) [4-5]. The pandemic effect of multi-drug resistance TB to atleast two frontline drugs (isoniazid and rifampin), is a global threat. In view of this condition, WHO declared TB as a global emergency [6], and organized a surveillance project to assessed the anti-tuberculosis drug resistance trend all over the world, and it was reported that almost all countries have multi-drug resistance TB (MDR-TB), out of which more than fifty countries have extensive drug resistance TB (XDR-TB). Due to some lacuna attached to standard six month treatment for TB, MDR-TB case become ineffective, costly and time consuming. This have created much scientific interest in developing new anti-mycobacterial

agents to both treat *Mycobacterium tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance. Therefore, an alternative and less expensive drugs development is very important to completely eradicate the multidrug resistant cases of the tuberculosis. Because of this, it is essential to have breakthrough in treatment of tuberculosis; hence there is urgent need for computational drug design for combating such disease. Shikimate pathway plays indispensable role in synthesis of aromatic compounds and many other essential nutrients in plants and microorganisms (bacteria, fungi and algae) It involved series of reaction that produced secondary metabolites such as amino acids, foliates, ubiquinone and naphthoquinones and inability of mammal to undergo such pathway making it specific target for drug development. It is biosynthetic pathway that starts with reaction between erythrose 4-phosphate and phosphoenolpyruvate to yield chorismate which can be achieved in seven steps [7]. The final products of shikimate pathway in *Mycobacterium tuberculosis* are the precursor for the synthesis of protein, mycobactin and ubiquinones. Therefore, shikimate pathway was considered as a best target for drug development against *Mycobacterium tuberculosis.* One of the important enzymes in this pathway is shikimate kinase that catalyses the fifth of the seven steps involved in the pathway. This enzyme is responsible for catalysing phosphoryltransfer from ATP to shikimate to form shikimate-3-phosphate and ADP, thus, inhibition of this enzyme is a promising target for development of anti-tubercular drugs [8-9]. Therefore, this study was aimed to identify novel inhibitors of Shikimate kinase using both natural and synthetic ligands.

METHODOLOGY

Sequence identification

The sequence of *Shikimate kinase* was identified from National Centre for Biotechnology Information **(**NCBI Reference Sequence) with the Accession Number of NP_217055.1. The protein was subjected to protein blast (Blastp) against the Protein Data Bank (PDB) to detect similarity and perform alignment with protein of known three dimensional structures in other to choose templates for Homology modeling. After alignment six structures with high similarity and greater statistical significance (Less e-value) were selected for comparative modeling to build the 3D structure of query sequence (Table 1). The pairwise **s**equence alignment between query sequence and the sequence with high similarity from PBD was performed using T-Coffee and Boxshade to improve the sensitivity of the search and to find the conserved regions.

Determination of secondary structure

The secondary structure of the query sequence was analysed using Hierarchical Neural Network (HNN) (https://nspa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html) [10] to ascertained different regular local structures such as Alpha helix, 310 helix, Pi helix, Beta bridge, Extended strand, Beta turn, Bend region and Random coil that are essential intermediate in predicting the tertiary structure of protein.

Homology modeling

Homology modeling was used to determine the 3D structure of the query sequence, using Modeller 9.16 [11]. The 2IYQ (Shikimate kinase in complex with shikimate and ADP) was used as a template to build the model due to it high-resolution crystal structures. The program builds the model protein structure through satisfaction of spatial restraint using standard parameters and databases. The alignments of a query sequence with known related structures were used as an input. The output of the modelled protein contained three dimensional structure of Shikimate kinase including all non-hydrogen main-chain and side-chain atoms.

Validation of model

The model was validated using PROCHECK [12] which checks the stereochemical quality of a protein structure by analysing residue-by-residue geometry and overall structure geometry, PROSA [13] used for interactive web-based applications for the display of scores and energy plots that highlight potential problems spotted in protein structures.

Active site identification and preparation of ligands

The active site of Shikimate kinase was determined using 3DLigandSite-Ligand binding site prediction Server (http://www.sbg.bio.ic.ac.uk/3dligandsite/**)** [14]. Two classes of ligands were used in this study, synthetic ligands from Zinc database and natural ligands derived from phytochemical component of plants obtained from PubChem database of NCBI.

Molecular docking

Molecular docking of Shikimate kinase with different ligands was performed using AutoDock 4.0 [15], which combines a rapid energy evaluation through precalculated grids of affinity potentials with a variety of search algorithms to find suitable binding positions for a ligand on a given protein. It uses a scoring function based on

the AMBER force field, and estimates the free energy of binding of a ligand to its target. Novel hybrid global-local evolutionary algorithms were used to search the phase space of the ligand-macromolecule system. The query protein was kept rigid, but all the torsional bonds in ligand were set free to perform flexible docking. Polar hydrogens were added, after that, Kollman united atom partial charges were assigned [16]. All torsion angles for each compound were considered flexible. The grid maps representing the proteins in the actual docking process were determined using AutoGrid. The dimensions of the grids were set as 60×60× 60 Å, with a spacing of 0.375 Å between the grid points.

RESULTS AND DISCUSSION

Sequence identification

The sequence of *Shikimate kinase* (176aa) was obtained from NCBI (Accession number: NP_217055.1), protein-protein blast was used to compares the amino acid sequence against a protein sequence from PDB. The blastp results page shows around 100 "Hits", or sequences showing different range of similarity arranged in descending order. Six proteins (PDB Code: 1I44, 2GA8, 2IYQ, 3R20, 4XRP and 4XRU) were selected on the basis of high similarity and greater statistical significance (Less e-value) homology modeling. Although, 2IYQ was further selected on the basis of the high-resolution crystal structures (Table 1) and used as template for both pair sequence alignment and modeling of *Shikimate kinase*. The result of the pair sequence alignment revealed high identity and similarity of 95.7% (176/184aa) and with gap of 4.3% (8/184aa). This high similarity between the two proteins is clear indication that the proteins possessed similar three dimensional structure, since the 3D structure of proteins are more conserved than their amino acid sequence. The secondary structure of *Shikimate kinase* was determined using Hierarchical Neural Network (HNN) Secondary Structure Prediction Method (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html (Combet *et al.*, 2000) and the result shows that Alpha helix (Hh) was found to be dominant with 77/176aa (43.75%), followed by Random coil (Cc) 76/176aa (43.18%) and Extended strand (Ee) had the least 13/176aa (13.07%) while other secondary structure element were not found in the sequence (Figure 2).

Table 1: **Templates Selected for Homology Modeling**

S/No.	PDB Code	Sequence Identity %	E-Value	Query Coverage %	Crystallographic Resolution Å
1	1L4UA	100	9.00E-123	100	1.9
2	2IYQA	100	1.00E-122	100	1.8
3	1KAGA	38	2.00E-25	94	2
4	4Y0AA	40	1.00E-24	85	1.9
5	3TRFA	36	5.00E-23	89	2.6
6	1SHKA	35	1.00E-22	96	1.9

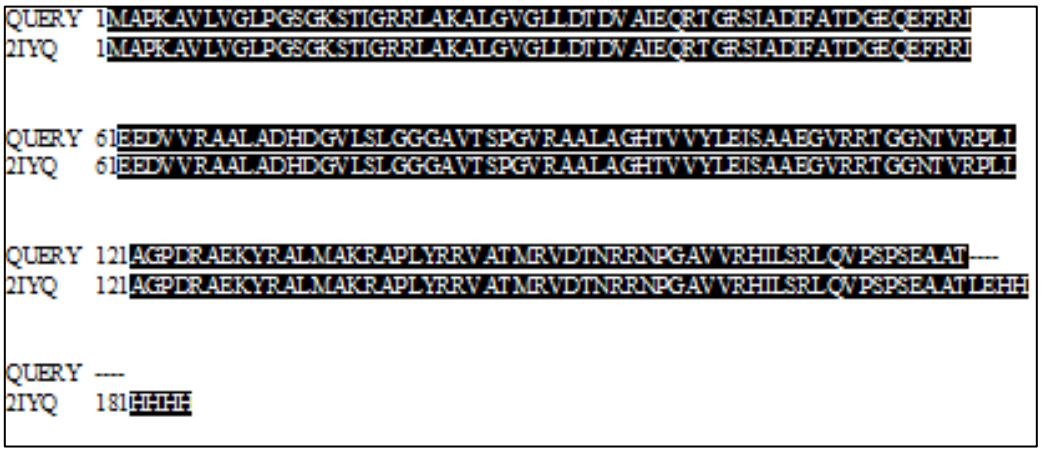


Figure 1: Pair sequence alignment using T-Coffee and Boxshade, between *Shikimate kinase* and crystal structures of *Shikimate kinase* in complex with shikimate and ADP from *Mycobacterium tuberculosis* but express in *E. coli* (PDB ID: 2IYQ)

Hierarchical Neural Network result for Shikimate Kinase

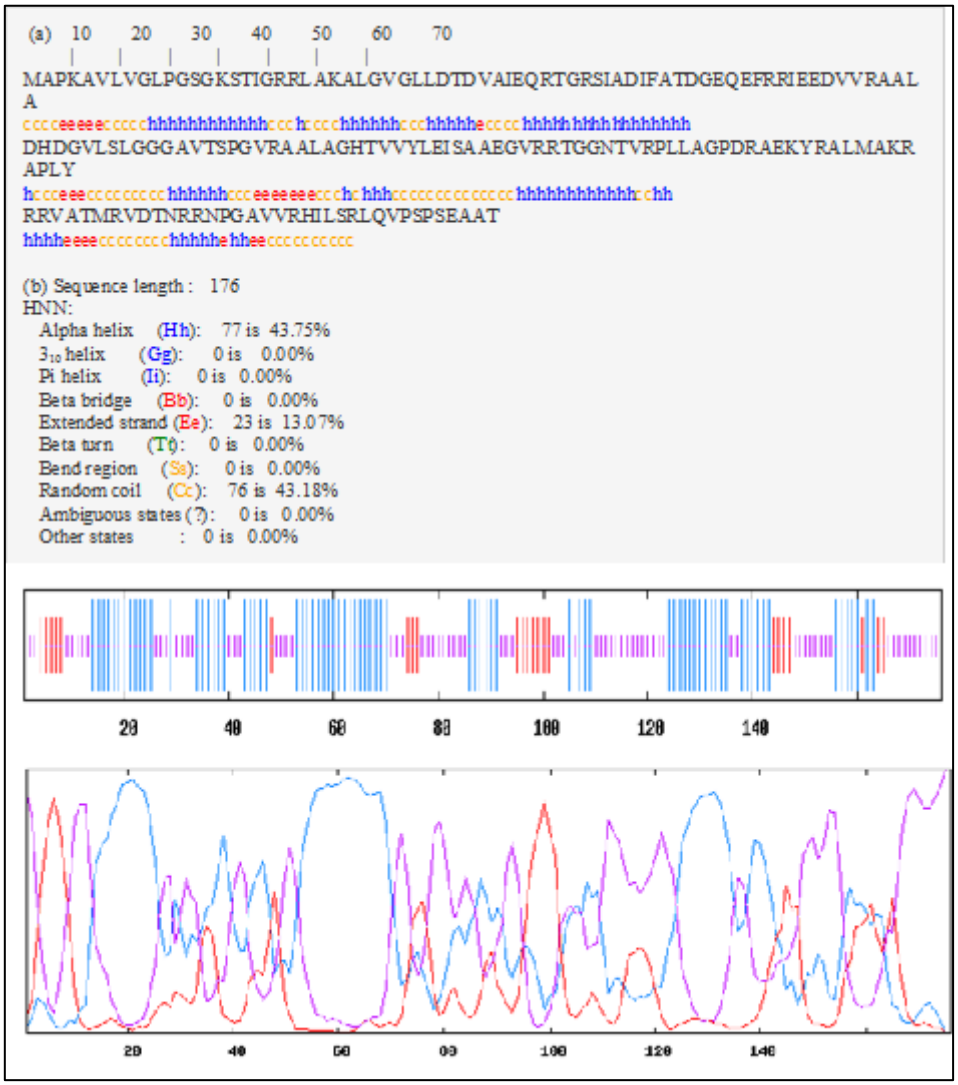


Figure 2: Secondary structure prediction of Shikimate kinase was performed using HNN Secondary Structure Prediction Method online server showing position of possible loop region (a) Residue wise prediction of secondary structure of Shikimate kinase (b) Tabular representation of secondary structural element (c) Graphical representation of secondary structural element

Homology Modeling and Validation of Model

The modelled structure of Shikimate kinase was build using crystal structure of (2IYQ) chain A, due to the high identity and similarity of 95.7%. This makes 2IYQ a good template for modeling. The coordinates of the template protein (2IYQ) such as structurally conserved regions (SCRs), structurally variable regions (SVRs), N and C–termini were assigned to the *Shikimate kinase* sequence based on the satisfaction of spatial restraints. All main chain, side chains of the modelled protein were set by rotamers (Figure 3a). The structural superimposition of Cα trace of template and Shikimate kinase was shown in Figure 3b. The RMSD of Cα trace between 2IYQ and Shikimate kinase calculated using the Pymol was shown in Figure (3b), and reveals that RMSD of Cα trace between template and the model built was 0.097 Å. This value clearly indicates that the structure was highly reliable. Hence, this model was used for the identification of active site and for docking of the inhibitors with Shikimate kinase. The final structure with least DOPE energy was further subjected to validation using PROCHECK (Ramachandran) to determine the stereochemical quality of a protein structure by analysing residue-by-residue geometry and overall structure geometry. The result of the verification with Ramachandran plot revealed that 97.3% (142/146 amino acids) of the entire residues lay in the most favourable regions [A, B, L], 2.7% (4/146 amino acids) were in Additional allowed regions [a,b,l,p], while the none of the amino acid residues were found in both Generously allowed regions [~a,~b,~l,~p] and Disallowed regions[XX] (Figure 4), which indicated good and acceptable quality model with high structural integrity. This is due to fact that a good quality model is expected to have over 90% in the most favoured regions based on the analysis of 118 structures of resolution of at least 2.0 Angstroms and R factor no greater than 20.0. On the other hand G-factors provide a measure of how unusual a property is. Values below -0.5 is considered as unusual while -1.0 is regarded as

highly unusual, although the overall average of G-factors in this study is -0.1 which is above the acceptable threshold of -0.5 (Table 2). ProSA server tool was used to check 3D models of predicted Shikimate kinase structures for potential errors [13]. The overall quality score of Shikimate kinase structure was displayed in a plot that shows the scores fall within the region of Nuclear Magnetic Resonance (NMR) protein chains available in the Protein Data Bank (PDB). The *Z*-Score of -8.44 was indication of good quality model with high structural integrity (Within the range of experimentally determined structure, -15 to 10) (Figure 5).

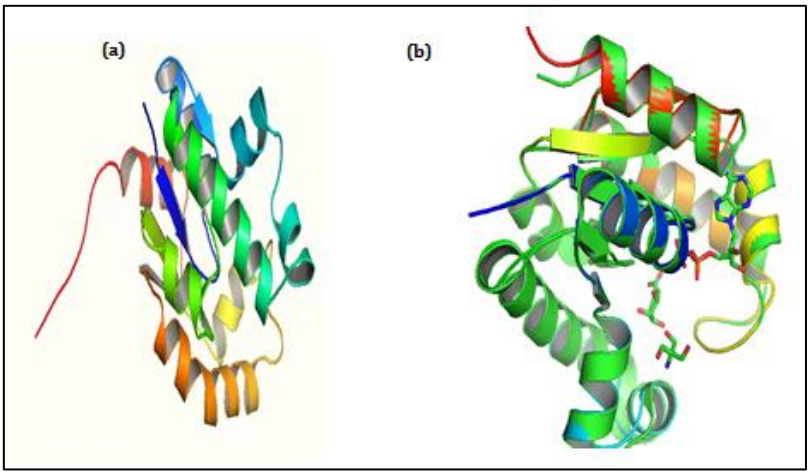


Figure 3: (a) Predicted three Dimensional (3D) structure of *Shikimate kinase* with least DOPE value using Modeller 9v16 (b) Structural superimposition of Cα trace of template and Shikimate kinase (RMS = 0.097 (142 to 142 atoms))

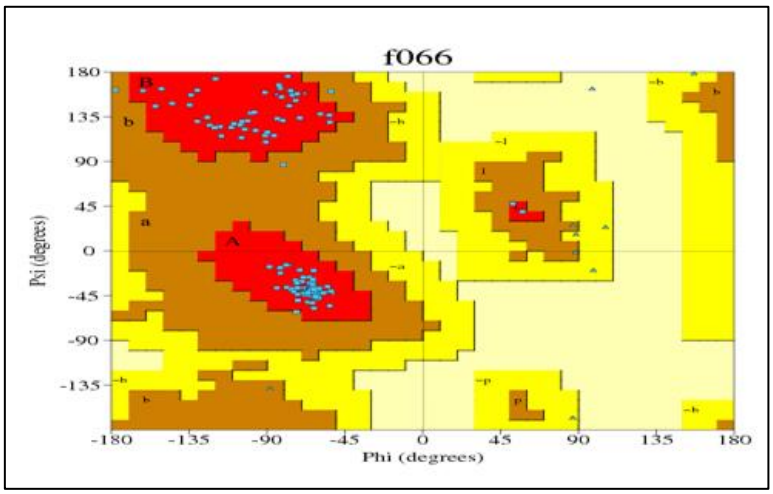


Figure 4: Ramachandran map showing distribution of residues for shikimate kinase

Active site prediction and Molecular docking analysis

The built Shikimate kinase model has been submitted to 3DLigandSite-Ligand binding site prediction Server (http://www.sbg.bio.ic.ac.uk/3dligandsite/**)**, which resulted in prediction of active sites present in the model protein. The residues found to be functional active sites of the protein includes; LEU10, PRO11, GLY12, SER14, LYS15, SER16, THR17, ILE18, ARG110, ARG117, AGR130, THR150, ARG153, ASN154, PRO155 and VAL158 (Figure 6). Also molecular docking analysis of protein-ligand complex was performed using AutoDock4.0 for both synthetic ligand (Zinc Database) and natural ligands (PubChem Database). The result of docking studies (synthetic ligands) among the several ligands used shows that only thirteen (13) found to have promising activity against Shikimate kinase with a minimum free binding energy ranges from -11.52 to -9.00 kcal/mol (Table 3), while on the other hand six natural ligands possess high activities with minimum free binding energy ranges from -11.75 to -8.49 kcal/mol (Table 4).

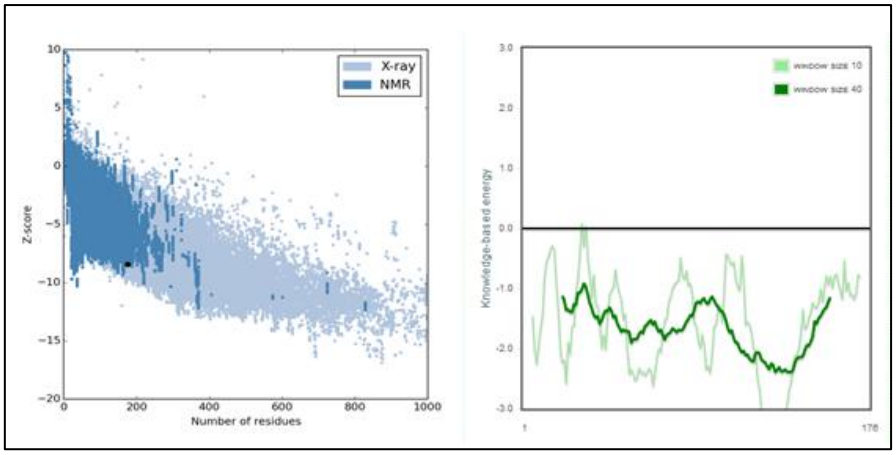


Figure 5: Overall quality of predicted model of Shikimate kinase using ProsA server

Among the ligands Digoxigenin (Natural ligand) which formed a complex hydrogen bond with LEU15, SER16 and ARG117 showed the lowest binding energy of -11.75kcal/mol and inhibitory constant of 2.43nM (nanomolar) (Table 4), while among the synthetic ligands compound with ZINC02838601 (Zinc Database ID) formed a strong hydrogen bond with ARG130 and exhibited a minimum free energy of -11.52kcal/mol and inhibitory constants of 3.59nM (nanomolar) (Table 3). These natural and synthetic ligands can be used for the treatment of tuberculosis and can also be included in treatment regime of MDR-TB after they undergo *in vitro* and *in vivo* evaluations of their bioactivity.

Table 2: Ramachandran Plot statistics

Residues	**Number of Residues**	**% of Residues**
Most favoured regions [A,B,L]	142	97.3
Additional allowed regions [a,b,l,p]	4	2.7
Generously allowed regions [~a,~b,~l,~p]	0	0
Disallowed regions [XX]	0	0
Non-glycine and non-proline residues	142	100
End-residues (excl. Gly and Pro)	2	
Glycine residues	19	
Proline residues	9	
Total number of residues	176	
G-Factors		
Parameter	Score	Average Score
Dihedral angles:-		
Phi-psi distribution	0.2	
Chi1-chi2 distribution	0.22	
Chi1 only	0.33	
Chi3 & chi4	0.54	
Omega	-0.18	
		0.2
Main-chain covalent forces:-		
Main-chain bond lengths	-0.07	
Main-chain bond angles	-0.04	
		-0.05
OVERALL AVERAGE		0.1

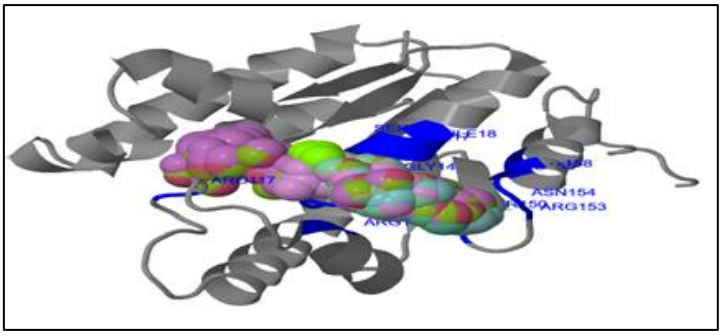


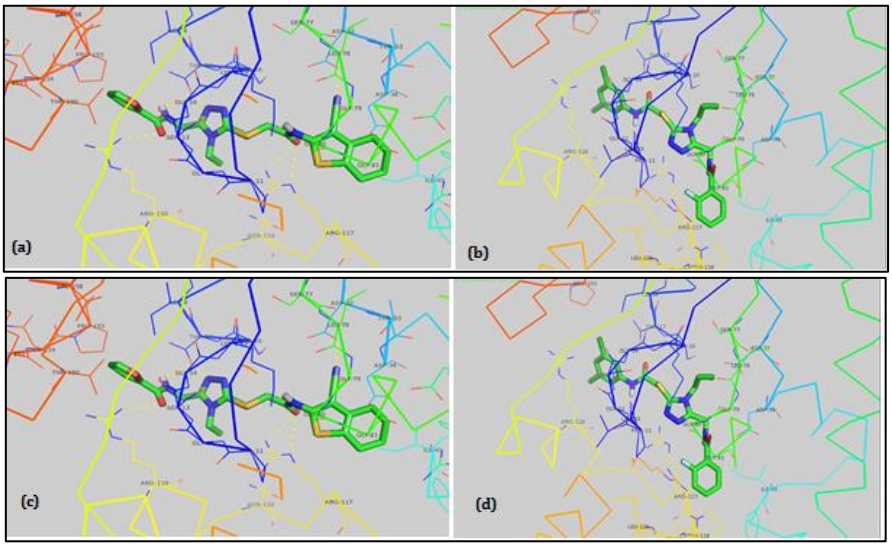
Figure 6: Functional active site prediction using 3DLigandSite-Ligand binding site prediction Server

Table 3: Thirteen (13) best synthetic ligands along with their Docking score

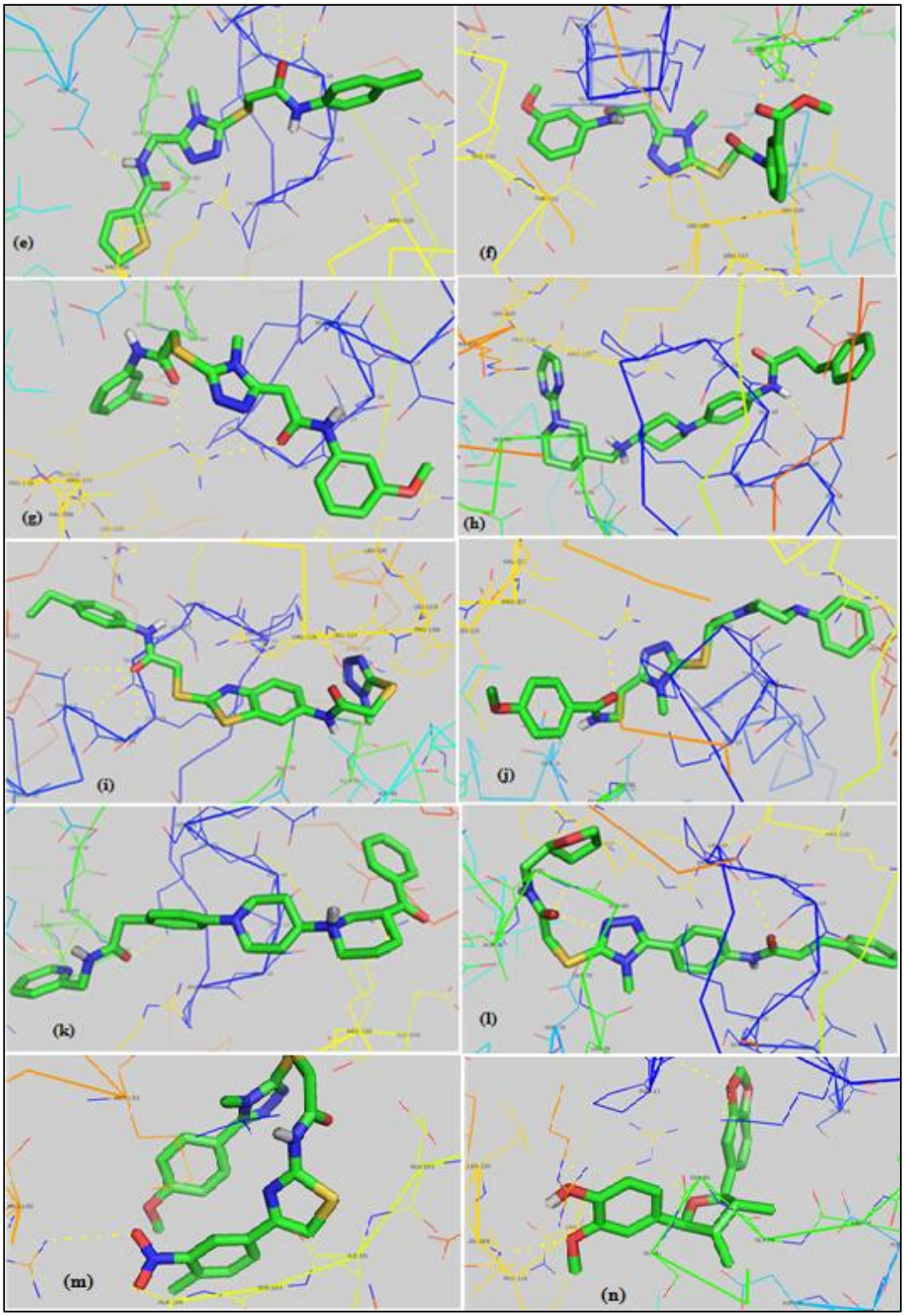
S/No.	Zinc Code	Docking Score/Minimum Free Energy of Binding (kcal/mol)	Inhibition Constant, Ki (nM (nanomolar)	Number of Hydrogen Bonding
1	ZINC02838601	-11.52	3.59	1
2	ZINC8442189	-10.41	23.31	5
3	ZINC11790367	-9.88	56.91	2
4	ZINC633842	-9.67	82.27	3
5	ZINC8442094	-9.51	107.71	3
6	ZINC02843658	-9.2	181.95	3
7	ZINC8442075	-9.17	190.08	6
8	ZINC634006	-9.14	198.07	4
9	ZINC633992	-9.1	212.08	4
10	ZINC16192643	-9.08	220.92	5
11	ZINC09191993	-9.02	243.55	3
12	ZINC633978	-9.01	250.21	7
13	ZINC11881196	-9	251.74	3

Table 4: Six (6) best natural ligands along with their Docking score

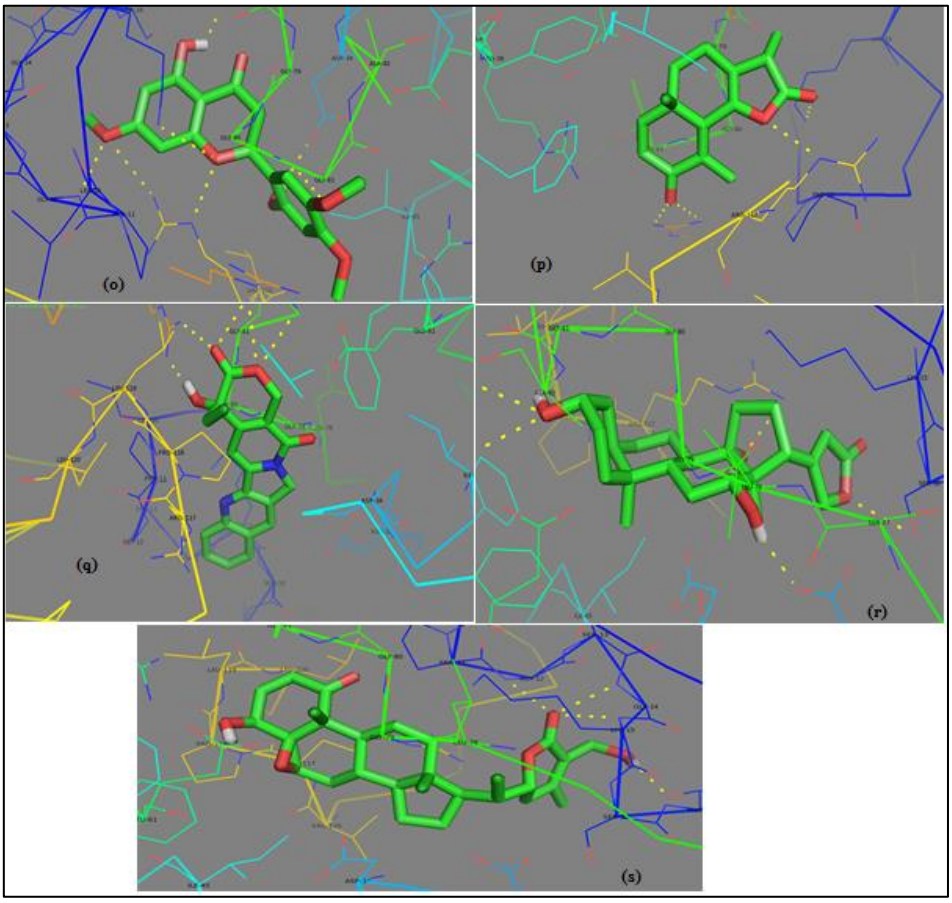
S/No.	Compound Names	PubChem ID	Docking Score/Minimum Free Energy of Binding (kcal/mol)	Inhibition Constant, Ki (nM (nanomolar)	Number of Hydrogen Bonding
1	Digoxigenin	PubChem15478	-11.75	2.43	5
2	Withaferin A	PubChem265237	-9.35	141.08	4
3	Santonin	PubChem221071	-9.1	213.83	4
4	Camptothecin	PubChem24360	-8.76	380.35	5
5	Lignans machilin F	PubChem13844301	-8.56	528.63	5
6	Hispidone	PubChem9997719	-8.49	597.94	7



(a) ZINC8442075; (b) ZINC633842; (c) ZINC8442094; (d) ZINC634006



(e) ZINC633992; (f) ZINC633978; (g) ZINC8442189; (h) ZINC11790367; (i) ZINC02843658; (j) ZINC09191993; (k) ZINC11881196; (l) ZINC16192643; (m) ZINC02838601; (n) Lignans machilin F



(o) Hispidone; (p) Santonin; (q) Camptothecin; (r) Digoxigenin; (s) Withaferin A

Figure 7: (a-m) Docking complex between Shikimate kinase and synthetic ligands from Zinc database (n-s) Docking complex between Shikimate kinase and natural ligands from PubChem database (NCBI)

CONCLUSION

Shikimate kinase is a potential drug target for the inhibition of Shikimate pathway which in return interfere with the growth of *Mycobacterium tuberculosis* and hence for the treatment of TB. To understand the binding mode analysis and structural features of the Shikimate kinase was modelled through homology modeling using Modeller 9.16. The modelled structures was validated using Ramachandran plot which revealed that 97.3% (142/146aa) of the entire residues lay in the most favourable regions [A, B, L], 2.7% (4/146aa) were in Additional allowed regions [a,b,l,p], while the none of the amino acid residues were found in both Generously allowed regions [~a,~b,~l,~p] and Disallowed regions[XX]. Also validation with Pros A server shows a *Z*-Score of -8.44 which clearly indicates good quality model with high structural integrity. The result of molecular docking analysis of protein-ligand complex was performed using both synthetic ligand (Zinc Database) and natural ligands (PubChem Database) revealed that ligands showed effective activity against the enzymes which indicates promising target for developing anti-tubercular drugs. Therefore, these ligands were recommended for future drugs for the treatment of *tuberculosis* in human cause by both multidrug-resistant tuberculosis (MDR-TB) strain and extensive drug resistant strain.

REFERENCES

[1] D Jothieswari; K Bhaskar Reddy, *I J Innov Drug Discov*, **2015**, 5(1), 40-45.
[2] HS Bello; MA Isa; M Bintube; Z Hassan, *Basic Res J Med Clin Sci*, **2014**, 3(12), 127-130.
[3] UK Muhammad; M Bello; SB Manga; MA Isa, *Int J Res*, **2014**, 1(11), 504-511.
[4] World Health Organization (WHO). Global tuberculosis repor. WHO/ HTM/TB/2013.11. **2014.**
[5] S Gordon; S Johayra; CG Douglas; IC Angela, *Perspect Med Chem*, **2015**, 7: 9-20.
[6] A Kanakaraju; PY Rajendra, *World J Pharmacy Pharm Sci*, **2015**, 4(9), 620-633.

[7] S Sinha; P Rajasulochana; BPB Ramesh; P Krishnamoorthy, *Res J Pharm Biol Chem Sci*. **2013**, 4(3): 715-720.
[8] LA Basso; LH Pereira Da Silva; AG Fett-Neto; WF De Azevedo Jr; IS Moreira; MS Palma; JB Calixto, SA Filho; RR Dos Santos; MBP Soares; DS Santos, *Mem Inst Oswaldo Cruz*. **2005**, 100:475–506.
[9] R Bentley, *Biochem Mol Bio*. **1990**, 25:307–84.
[10] C Combet; C Blanchet; C Geourjon; G Deléage, *Net Protein Seq Anal. TIBS*. **2000**, 25(3),147-150.
[11] A Webb; A Sali, *Curr Protocols Bioinforma*, **2014,** 5.6.1-5.6.32.
[12] AR Laskowski; MW MacArthur; DS Moss; JM Thornton, *J. Appl. Cryst*. **1993**, 26, 283–291.
[13] M Wiederstein; MJ Sippl, *Nucleic Acids Res*, **2007**, 35, W407-W410.
[14] MN Wass; LA Kelley; MJ Sternber, *NAR* 38 Suppl. **2010**, W469-73
[15] GM Morris; DS Goodsell; RS Halliday; R Huey; WE Hart; RK Belew; AJ Olson, *J Comput Chem*, **1998**, **19**: 1639-1662
[16] C LaMotta; S Sartini; L Mugnaini; F Simorini; S Taliani; S Salerno, *J Med Chem*. **2007**, 50, 4917–4924.