Abstract

Tuberculosis is a common and deadly infectious disease caused by Mycobacterium tuberculosis (Mtb), reemergence of TB as a public health threat has created a need to develop anti-mycobacterial agents. The shikimat pathway is an attractive target for anti-microbial agent’s development. Key enzymes are playing a major role in shikimat pathway metabolism. Among them, Mtb 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS) was proposed to be present by sequence homology. Accordingly, to pave the way for structural and functional efforts towards anti-mycobacterial agent’s development, here we describe the molecular modeling, dynamic, docking studies on Mtb- DAH7PS, and revels key differences that could be useful for development of anti-tuberculosis drugs.
INTRODUCTION:

Tuberculosis (TB), an infectious disease caused by Mycobacterium tuberculosis (MTB), it is a leading killer that has plagued mankind for centuries. This disease accounts for 7% of all deaths in developing countries and as much as 26% of avoidable adult deaths (Gomez JE and McKinney JD, 2004). One of the most effective first-line anti-TB drugs is isonicotinic acid hydrazide (INH), commonly known as isoniazid, it is a common therapeutics for MTB consist of a six-month regimen, using streptomycin/ethambutol in combination with INH, rifampicin, and pyrazinamide for 2 months followed by rifampicin and INH biweekly for 4 months[1]. During the past decade the increase in the prevalence of INH-resistant Mtb strains has become a severe setback to early therapeutic success and life-threatening complication. Among new TB cases, as much as 14.1% Mtb isolates are resistant to at least INH and rifampicin. The multidrug resistant (MDR) TB rate constitutes at least 5% and depends on the geographic sites[2]. Now a days, the incidence of MDR-TB infectious is constantly rising due to the fact that TB has become a leading opportunistic infection in AIDS patients. The treatment of MDR-TB patients thus requires the administration of second line drugs (amikacin, kanamycin, capreomycin, cycloserine, para-aminosalicylic acid, ethionamide, and fluoroquinolones). However, these drugs are more toxic and less efficient and have a longer regiment’s time. The cost of the treatment of MDR-TB is also 100 times higher than that of the basic six month short-course chemotherapy regiments[3]. Treatment for TB is more complicated and difficult with the appearance of extensively drug-resistant (EDR) TB strains [4,5] (resistant to at least INH and rifampicin and, in addition, to at least three of the six classes of second line anti-TB drugs: aminoglycosides, polypeptides, fluoroquinolones, thioamides, cycloserine, and p-aminosalicylic acid) since the year 2006 and the possible of extremely drug-resistant (XDR) TB strains 8 (resistant to all first- and second-line anti-TB drugs).

For this new emerging field, insilico drug design has been offered enormous benefit for the development of effective drugs against TB. In this context, we have chosen the enzymes involved in shikimate path[15] way of Mtb, which plays an essential structural role in the survival of bacterium. Inhibition of shikimate pathway would be lethal to
bacteria. Significantly this pathway doesn’t exist in mammals our enzyme is therefore represent potential therapeutic targets. we have employed insilico approaches to resolve and characterize the structure of this important enzyme by molecular modeling and simulation techniques. Global and local accuracy of the predicted model was assessed by various assessment programs. With the aim to built novel inhibitors for Mtb-DHAPS model, docking studies are done with series of E4P derived ligands. Results of ligand interactions have revealed specific residues in the binding domain of Mtb-DHAPS. This information could be exploited for future designing of more effective inhibitors for Mtb-DHAP enzyme.

Methodology

Retrieval of DAH7PS genome and proteome information from on line servers:

M. tuberculosis 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS) is an important enzyme in shikimate pathway which is responsible for the biosynthesis of the amino acids phenylalanine, tyrosine, and tryptophan of M. tuberculosis. This is novel target for development of anti-tuberculosis drugs as per published literature.

The first step in proposed project is retrieval of primary information like genetic map and protein coding region on chromosome 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase through NCBI (www.ncbi.nlm.nih.gov/).

Prediction of 3-D model for mtb-DAH7PS enzyme:

3-D structure model of DAH7PS enzyme will be predicted and generated by special restraining of amino acids using MODELLER, by selecting suitable template solved by experimental methods and stored at PDB database.

Structural Validation and Energy minimization:

The built 3-D model of DAH7PS enzyme will be refined by loop modeling and Energy minimization under specific force field. This model will be further used for analysis of molecular dynamics and simulations.
Identification of Inhibitor site and design of lead molecule or inhibitor:

The DAH7PS enzyme model will be analyzed with molecular surface, electrostatic potential surface for the identification of best cavity for docking process. Lead molecules will be designed and screened through optimization under AMBER force field.

Docking of lead compounds with DAH7PS

Designed and selected Lead molecules docked with designed DAH7PS model with suitable grid map and best interacting lead molecules with lowest binding energy are selected and represent for further pharmaceutical research purpose.

Results

3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase (DAHPS), the first enzyme of the aromatic biosynthetic pathway in microorganisms and plants, catalyzes the aldol-like condensation of phosphoenolpyruvate and D-erythrose-4-phosphate with the formation of 3-deoxy-D-arabino-heptulosonate-7-phosphate. In Escherichia coli, there are three isoforms of DAHPS, each specifically feedback-regulated by one of the three aromatic amino acid end products. Metabolic pathway database (KEGG) [6] is a source of metabolic pathway information. Metabolic pathway identification numbers of the host H. sapiens and the pathogen M. tuberculosis were extracted from the KEGG database. Pathways which do not appear in the host but present in the pathogen according to KEGG database annotation have been identified as pathways unique to M. tuberculosis as compared to the host H. sapien. Enzymes in the unique pathways metabolism were identified from the KEGG database. The corresponding protein sequences were retrieved from the KEGG database. They were subjected to a BLASTp [7] search against the non-redundant database with the e-value inclusion threshold set to 0.005. The search was restricted to proteins from H. sapiens through an option available in the BLAST program, which allows the user to select the organism to which the search should be restricted. In the current context, the objective is to find only those targets, which do not have detectable human homologues. Enzymes, which do not have hits below the e-value inclusion threshold of 0.005, were picked out as potential drug targets. These targets were checked against
Among the unique pathways in mycobacterium tuberculosis H37Rv shikimate pathway is one of the good targets for the development of new anti-tuberculosis drugs.

**Shikimate pathway**

The shikimic acid pathway is a seven step metabolic route used by bacteria, fungi, algae, parasites and plants for the biosynthesis of aromatic amino acids (phenylalanine, tyrosine, and tryptophan). This pathway is not found in animals; hence the products of this pathway represent essential amino acids that must be obtained from the animal's diet. The first enzyme involved is the shikimate kinase,
an enzyme that catalyzes the ATP-dependent phosphorylation of shikimate to form shikimate 3-phosphate.\textsuperscript{[1]} Shikimate 3-phosphate is then coupled with phosphoenolpyruvate to give 5-enolpyruvylshikimate-3-phosphate via the enzyme 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase. Due to its absence in humans and animals, Shikimate pathway has been an attractive target for development of antibiotics against a number of important microbial pathogens, such as Mycobacterium tuberculosis.

![Fig.1: KEGG database Mycobacterium tuberculosis Phenylalanine, Tyrosine and Tryptophan metabolic pathway](image)

3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase:

M. tuberculosis 3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase (DAHP) synthase is an enzyme that catalyzes the chemical reaction phosphoenolpyruvate + D-erythrose 4-phosphate + H$_2$O $\rightleftharpoons$ 3-deoxy-D-arabino-hept-2-ulosonate 7-phosphate + phosphate

The three substrates of this enzyme are phosphoenolpyruvate, D-erythrose 4-phosphate, and H$_2$O, whereas its two products are 3-deoxy-D-arabino-hept-2-ulosonate 7-phosphate and phosphate as shown in Figure 2. This step lies at the first step in shikimate pathway.

Available Online At www.ijprbs.com
Multiple sequence alignment

The PSI-BLAST program was used to search for homologues of DHAPS family proteins. The protein sequences of homologues DHAPS from different subspecies of Mycobacterium were aligned to identify the evolutionary conserve residues. Thereafter, very closely resembling sequences with better E score <0.005 and higher % similarity were used for finding conserved regions of Mtb-DHAPS with similar proteins from other species using Multiple sequence alignment (MSA; http://www.ebi.ac.uk/Tools/clustalw2/index.html) program ClustalW with routine default parameter. The results of multiple sequence alignment showed that DHAP 100% conserved with species of 3KGF (Fig.3) and almost 100% similarity with Mycobacterium tuberculosis strains.
Fig:3 Pairwise sequence alignment of DHAP with 3KGF, denotes the conserver regions, semi conserve regions and also less conserve regions in the sequences.

3D model building:

3D of DHAPS enzyme was built using Modeller 9v8 software by satisfaction of spatial restrains[14]. The program was carried out using standard parameter set and databases. Many runs of model building were carried out to obtain the most reasonable model and subsequently the best model (with the low RMS value of superposition using Swiss-pdb viewer was subjected to further analysis. To remove steric clashes arising from non-bonded interactions and to correct the bad geometry in RmlA model and to achieve a
good starting structure, refinement was done by energy minimization (EM). The 3-D model of Mtb-DHAPs (fig.5) is remarkably similar to those of crystal structures of 3kgf.

**Fig.5: Predicted 3D structure of Mtb-DHAPS Using Modeller 9v8.**

Energy minimization was performed to obtain the best accurate conformation of the developed 3D model of Mtb-DHAPS using GROMACS 3.2.1. During optimization, energy minimization of residue, bonds, angles, torsion, improper non-bonded lengths, electrostatic, constraint are minimized for best confirmation fig.

**Validation of homology model**

Energy minimized accurate structure obtained for Mtb-DHAPS was subjected for structural evaluation in PROCHECK [9]. The values obtained are within the values of determined homologs structures. The low overall RMS Z-score values for backbone superposition reflect the high accurate structural conservation of this complex through evolution, making it a good system for homology modeling. Ramachandran plot [RP] calculations computed with PROCHECK program by checking the detailed residue-by-residue stereo-chemical quality of a protein.
structure. According to PROCHECK program, RP of the shading represents the different regions of the plot, the darker the area the more favorable the \( \phi-\Psi \) combination. The RP depicted in fig.9a, reveal that 91.7\% of the amino acid residues in the favorable regions and 7.8\% of the amino acids are in additional allowed regions and 0.5\% residues in generously allowed and no in disallowed regions of the plot for the whole developed model.

Fig6. Ramachandran plots of Mtb-DHAPS: no amino acids in disallowed regions of the plot for the whole developed model (DHAPS).

**Preparation of ligands**

100 Substrate analogues were drawn in chemoffice and checked in molinspiration for drug likeliness whether it follows Lipinski rule or not. The lead analogues were further subjected to PRODRG server to prepare ligand .pdb files [17]. From the random ligands 5 analogues were selected as lead molecules for further analysis after binding interaction with DHAPS. The 5 ligands with their structure and Lipinski rule properties are listed below.
From the above the last best five are selected for further docking.

Docking studies:

Docking was performed with the program AutoDock4.0 [10] 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. When docking, DHAPS was kept rigid, but all the torsional bonds in ligand were set free to perform flexible docking. Polar hydrogens were added by using the Hydrogens module in AutoDock Tools (ADT) for DHAPS; after that, Kollman united atom partial charges were assigned [11]. Docking of Ligand to DHAPS was carried out using the empirical free energy function and the genetic algorithm, applying a standard protocol with an initial population of 100 randomly placed individuals, a maximum number of $2.0 \times 100$ energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1, where the average of the worst energy was calculated over a window of the previous 50 generations. Structures of ligand derivatives optimized by Prodrg. 50 independent docking runs were carried out for each ligand. Results were clustered according to the 1.0 Å root-mean-square deviation (RMSD) criterions. All torsion angles for each compound were considered flexible. The grid maps representing the proteins in the actual docking process were calculated with AutoGrid. The grids (one for each atom type in the ligand plus one for electrostatic interactions) were chosen to be sufficiently large to include not only the active site but also significant portions of the surrounding surface. The dimensions of the grids were thus $90 \times 90 \times 90$ Å, with a spacing of 0.375 Å between the grid points. AutoDock uses binding free energy evaluation to find the best binding mode. Energy items calculated by AutoDock include intermolecular energy (constituted by van der Waals energy, hydrogen bonding energy, desolvation energy, and electrostatic energy), internal energy, and torsional energy. The first two items build up docking energy; the first and the third item compose the binding energy. During all these interactions, the hydrogen bond between ligand and enzyme is the most important, because in most cases it can decide the binding strength and the location of ligand, whereas the hydrophobic
interaction of some certain groups can affect the inhibition specialty to a large extent[12,13,14,15,16]. The energy information is listed in Table.1, and the best interaction modes of ligand and DHAPS are shown in Fig 7, where only the amino acids located within 5 Å of the inhibitor are displayed with highest energy. For a better understanding on the docking conformation differences between ligand, a superposition is also given in fig 7. According to the difference between groups used to form hydrogen bonds Ligands orientation is close, in respect to their cluster orientation and hydrogen bond formation and far from the hydrophobic residues; all the other conformations are orientated similarly to another direction and immerged in a hydrophobic cavity. The binding modes of ligands are shown in fig 7. Out of these lead E4P derivatives, binding energies of E4P-3 and E4P-4 showed highest binding energies i.e -0.511 and -0.504 (Table.1) respectively with DHAPS of M.tb than that of other E4P substrate. The major amino acids involved with hydrogen bonding includes Tyr240,Tyr187,Gln239, Ala241,Ser188 as shown in below diagrams.

Table.1: Docking energies of DHAPS with e4p analogs.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Energy Total In kjoul</th>
<th>Energy Total In kcal</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.025469e+02</td>
<td>-0.483774959</td>
<td>-1.00</td>
</tr>
<tr>
<td>2</td>
<td>-2.017770e+02</td>
<td>-0.481936085</td>
<td>-1.00</td>
</tr>
<tr>
<td>3</td>
<td>-2.141813e+02</td>
<td>-0.511563246</td>
<td>-1.00</td>
</tr>
<tr>
<td>4</td>
<td>-2.113276e+02</td>
<td>-0.504747301</td>
<td>-1.00</td>
</tr>
<tr>
<td>5</td>
<td>-2.055432e+02</td>
<td>-0.490931499</td>
<td>-1.00</td>
</tr>
</tbody>
</table>

Fig.7.Binding modes of E4p-1 to E4P-5 to the DHAPS.
Fig. 7. Binding modes of E4p-1 to E4P-5 to the DHAPS.

Summary

Mtb- DHAPS is a vital enzyme chosen for development of new drugs against TB. This enzyme plays a key role in the survival of Mtb. In spite of the availability of a lot of drugs to treat TB, this enzyme has been targeted, due to the development of MDR-TB and XDR-TB in TB patients. There is urgency of development of new potent inhibitors against TB. In this work, we are developed new inhibitors based on substrate, which are binding more potentially to Mtb- DHAPS than that of available inhibitors and substrate. Based on the molecular interaction studies better drugs could be designed at a prime line level, lead molecule modification and designing, leads to identification of protein ligand interactions which do aid to rationalize effective Anti-TB drugs. For a virtual screening of protein and ligand molecules, Cheminformatics and Bioinformatics have provided tools for enhanced experimental research. This paves the way for researchers to further analyze the data obtained through computational studies. The systematic stepwise analysis conducted in the present work on in silico Screening of Effective inhibitors for Mtb- DHAPS using pharmacophore modeling yields useful information to design better drug formulations by pharmacists abiding to pharmaceutical norms. For Mtb- DHAPS, a systematic study was conducted to get an insight about Mtb- DHAPS enzyme and the corresponding inhibitors using in silico methods. To have a better understanding of
the enzyme, its molecular organization and other allied parameters, it is first necessary to build a computational 3-D model which requires the retrieval of amino acid sequences. The sequence of Mtb- DHAPS was retrieved from NCBI and this enzyme was identified from shikimate pathway. The 3-D model of Mtb- DHAPS was built using 3kgf (PDB id), the model was further refined and submitted to PROCHECK to evaluate reliability. PROCHECK analysis reports of Mtb- DHAPS have shown that 100% of amino acids are present in the most favored, additional and generously allowed regions, 0% in disallowed regions. The analysis reports of PROCHECK, revealed that Mtb- DHAPS model is of good quality and is reliable. The secondary structure conformations of Mtb- DHAPS model revealed that the protein is composed of 5 beta alpha beta units, 2 beta hairpins, 2 beta bulges, 12 strands, 35 beta turns, 5 gamma turns, 18 helix-helix interact and 18 helices. The built model was submitted to an online Protein Model Data Bank (PMDB) and can be viewed freely through online Google search with a PMDB id: PM0078265. As per the competitive binding of substrate (erythrose 4-phosphate(E4P)), we screened the E4P through AUTODOCK 4.0 tool which has provided information that substrate possesses the lowest docking energy and thus it has been selected for further analysis. The E4P molecule was further modified structurally and optimized through HYPERCHEM 7.5, CHEMOFFICE and PRODRG server. All the designed lead molecules were tested for Lipinski’s Rule-of-Five using MOLINSPIRATION server. The end results of Lipinski’s rule have enabled for identification of 5 ranked lead molecules(E4P-1 TO E4P-5) and were further docked on to Mtb-DHAPS with suitable grid map. Docking of the 5 lead molecules for best interactions with Mtb-DHAPS has given an insight that all the lead molecules have shown interactions with active site amino acids of Mtb-DHAPS. The analysis reports have shown that the conducted study can be further utilized in designing of better anti tuberculosis drugs. The obtained binding mode may be a guide to further mutational studies and structure-based search of potential DHAPS ligands.
Acknowledgements

Authors are thankful to UGC for financial support (F.No MRP-4246/12/UGC-SERO). Madhusudana and Srinivasaulu wish thanks to ICMR for providing financial assistance.

Reference:


17. Madhusudana P, Babajan B, Chaitanya Mulakayala, C.M.Anuradha, Chitta Suresh Kumar. Molecular modeling and docking studies of O-succinylbenzoate