Prioritization of known drugs against panC of M. tuberculosis

8.1. Introduction

Pantothenate synthetase encoded by the panC (Rv3602c) gene, catalyzes the essential ATP-dependent condensation of D-pantoate and beta-alanine to form pantothenate in M. tuberculosis (Zheng and Blanchard, 2001). Pantothenate is known to be essential for the synthesis of CoA and other important molecules involved in fatty acid biosynthesis and degradation, intermediary metabolism, and other cellular processes (Smith, 2003). A comprehensive study based on transposon site hybridization identified panC as one the essential genes for optimal growth of the M. tuberculosis (Sassetti et al., 2003). Pantothenate synthetase (panC) along with panD have been reported to be essential for intracellular survival and pathogenesis of M. tuberculosis (Sambandamurthy et al., 2002). M. tuberculosis H37Rv strain with double deletion of genes panC along with panD, which encodes for aspartate-1-decarboxylase and collectively involved in pantothenate biosynthesis, was demonstrated to lose virulence; further, when Severe Combined Immune Deficiency (SCID) mice infected with the panC and panD mutant strain, they were observed to live for an average 250 days versus the wild-type M. tuberculosis H37Rv infected mice living only for 35 days; in addition, lesser bacterial burden, much smaller histopathologic changes in the lungs of mice infected with the mutant compared to mice infected with the wild-type bacteria were observed (Sambandamurthy et al., 2002). All these findings established the importance of the panC in the virulence of M. tuberculosis.

Several bioinformatics based studies identified panC as one of the important drug target for M. tuberculosis. For example, based on comparative analysis of
metabolic pathways and proteins of host, *Homo sapiens* and *M. tuberculosis*, (Anishetty *et al.*, 2005) reported *panC* as one of the potential drugs targets as it is uniquely present in *M. tuberculosis* alone. Later, a comprehensive bioinformatics-based target identification pipeline that incorporated multiple biologically important analysis including protein-protein interactome, a flux balance analysis of the reactome, experimentally derived phenotype essentiality data, sequence analyses, a structural assessment of targetability, gene expression data analysis, non-similarity to gut flora proteins and 'anti-targets' in the host, resulted in 451 high-confidence targets including *panC* from *M. tuberculosis* (Raman *et al.*, 2008). Further, based on computational analysis of metabolic pathways, protein sequences, choke-point analysis, and protein interaction networks, 18 proteins including *panC* were identified as most interactive metabolic proteins and hence as potential drug targets (Kushwaha and Shakya, 2010). All these evidences support that *panC* could be a promising target for *M. tuberculosis* for drug discovery. By comparing the protein sequence of *panC* with human proteome using BLASTP it was also confirmed that *panC* has no homologous protein in human (Altschul *et al.*, 1997). The Drugability score of the *panC* was found to be 0.867 (Radusky *et al.*, 2014) indicating that *panC* could be potentially be targeted by small molecules efficiently. By querying the TBDB database, it was found that the *panC* was also found to be conserved in 27 different closely related bacteria indicating its essentiality (Reddy *et al.*, 2009; Galagan *et al.*, 2010). Using TDR targets database it was found that *panC* was upregulated in dormant condition suggesting that *panC* helps the *M. tuberculosis* to maintain the dormancy in the host. The importance of the pantothenate synthetase for the *M. tuberculosis* have encouraged researchers to put tremendous efforts to understand the structural and molecular interactions aspects that resulted in crystallographic structures for this enzyme (The UniProt Consortium,
Furthermore, several studies have already attempted to identify inhibitors for \textit{panC} of \textit{M. tuberculosis} (White \textit{et al}., 2007; Kumar \textit{et al}., 2013; Devi \textit{et al}., 2014; Xu \textit{et al}., 2014; Devi \textit{et al}., 2015). All of these previous studies suggest that \textit{panC} could be used as a target protein for drug discovery against \textit{M. tuberculosis}. Therefore, \textit{panC} was selected as one of the potential drug targets in the present study and all of the 1554 known drugs were prioritized using virtual screening against \textit{panC} of \textit{M. tuberculosis}.

\textbf{8.2. Objective}

The objective of the present study is to prioritize 1554 FDA-approved small molecule drugs against \textit{panC} of \textit{M. tuberculosis}.

\textbf{8.3. Methods}

The three-dimensional structure of \textit{panC} (PDB ID: 1N2I) was downloaded from Protein Data Bank (Wang and Eisenberg, 2003). Among many structural entries available for the \textit{panC}, the entry 1N2I was chosen because this structure complex with the substrate molecule Pantoyl Adenylate (PAJ). The protein structure (1N2I) and drug-library were prepared for docking-based virtual screening using the protocol described previously. All of the 1554 FDA approved drugs were docked against the \textit{panC} using the protocol described previously in chapter 3.3. The \textit{panC}-specific method is provided in the Figure 8.1.
Figure 8.1: The workflow of virtual screening and prioritization of 1554 FDA approved drugs against panC of M. tuberculosis.
8.4. Results

The rigid body docking of 1554 known drugs against *panC* was employed for preliminary shortlisting of drug. As per the original procedure, these drugs ranked in top 10% by both Glide and ADV were selected for next round of screening. However, surprisingly, only eight drugs were found to be consistently ranked within top 10% (155) by both Glide and ADV. Therefore, these eight drugs were subjected for induced fit docking by both Glide and AutoDockVina were presented as prioritized drugs. All of these eight drugs were observed to have RMSD value lesser than 1 (Table 8.1).

Intermolecular interactions between docked drugs and the target protein *panC* was analysed. Based on the X-ray crystallographic structure of the *panC*, Wang *et al.*, (2003) (Wang and Eisenberg. 2003) reported GLN72, GLN164, GLY158, MET40, ASP161, VAL187, MET195, GLY46, LYS160, HIS47 to be involved in the intermolecular interactions with other ligands. Diosmin, which was known to treat venous disease (https://www.drugbank.ca/) (Glide rank: 1 and ADV rank: 2) was observed to make hydrogen bond interaction with nine amino acids namely, GLY158, HIS47, GLN164, PRO38, ARG132, ARG198, HIS135, ARG278, SER196 (Figure 8.2). The amino acids GLY158, HIS47, GLN164 were already reported to form the active site in the protein *panC* (Wang and Eisenberg. 2003). An anti-cancer drug, (https://www.drugbank.ca/) Lapatinib (Glide rank: 2 and ADV rank: 6) was found to make hydrogen bond interaction with three amino acids namely, ASP161, GLN163, HIS135 (Figure 8.3). Amino acid ASP161 is already reported in the active site residues of the protein *panC* (Wang and Eisenberg. 2003). Zafirlukast was found to be ranked as 3 by Glide while 5 by ADV and interacted with five the amino acids (GLN163, ARG198, ARG132, HIS135 and ASP78) of the *panC*. Candoxatril, used in the
treatment of chronic heart failure (https://www.drugbank.ca/) was found to be ranked as 4 and 8 by Glide and ADV respectively and interacted with GLN164, GLN163 and ARG132 of panC. Olmesartan, used for the treatment of high blood pressure (https://www.drugbank.ca/) was observed to interact with five amino acids in the panC, GLN164, ARG198, ARG132, HIS135, ASP78 of which GLN164 is a known active site residue (Wang and Eisenberg. 2003). Estradiol valerate/Dienogest, (Figure 8.4) used for the treatment of heavy menstrual bleeding in women (https://www.drugbank.ca/) was ranked 6th and 1st by Glide and ADV respectively, and interact with 2 of the amino acids, GLN164 and GLN163, of which the former is a known active site residue (Wang and Eisenberg. 2003).

The pantoyl adenylate (PAJ) which was also docked with panC was found to have Glide score of -15.2788 kcal/mol and ADV score of -9.7 kcal/mol. All of the eight prioritized drugs were calculated to have better binding energy by ADV. However, surprisingly only one of the eight drugs (diosmin) was calculated to have better Glide score than the PAJ. These observed discrepancies in calculation of binding affinity by Glide and ADV needs to be studied further.
Table 8.1: List of prioritized drugs against panC of *M. tuberculosis*

<table>
<thead>
<tr>
<th>S.No</th>
<th>DrugBank ID</th>
<th>Drug name</th>
<th>Primary use</th>
<th>Glide score (kcal/mol)</th>
<th>Glide Rank</th>
<th>AutoDock Vina Binding Affinity (kcal/mol)</th>
<th>AutoDock Vina Rank</th>
<th>RMSD (Å)</th>
<th>Amino acids of panC interacting with drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DB08995</td>
<td>Diosmin</td>
<td>Treatment of venous disease</td>
<td>-17.6212</td>
<td>1</td>
<td>-17</td>
<td>2</td>
<td>0.844</td>
<td>GLY158, HIS47, GLN164, PRO38, ARG132, ARG198, HIS135, ARG278, SER196</td>
</tr>
<tr>
<td>2</td>
<td>DB01259</td>
<td>Lapatinib</td>
<td>Anti-cancer drug</td>
<td>-11.233</td>
<td>2</td>
<td>-15.6</td>
<td>6</td>
<td>0.868</td>
<td>ASP161, GLN163, HIS135</td>
</tr>
<tr>
<td>3</td>
<td>DB00549</td>
<td>Zafirlukast</td>
<td>Treatment of asthma</td>
<td>-9.3915</td>
<td>3</td>
<td>-15.7</td>
<td>5</td>
<td>0.851</td>
<td>GLN163, ARG198, ARG132, HIS135, ASP78</td>
</tr>
<tr>
<td>4</td>
<td>DB00616</td>
<td>Candoxatril</td>
<td>Treatment of chronic heart failure</td>
<td>-7.9621</td>
<td>4</td>
<td>-13.7</td>
<td>8</td>
<td>0.875</td>
<td>GLN164, GLN163, ARG132</td>
</tr>
<tr>
<td>5</td>
<td>DB00275</td>
<td>Olmesartan</td>
<td>Treatment of high blood pressure</td>
<td>-7.77796</td>
<td>5</td>
<td>-13.8</td>
<td>7</td>
<td>0.86</td>
<td>GLN164, ARG198, ARG132, HIS135, ASP78</td>
</tr>
<tr>
<td>6</td>
<td>DB08866</td>
<td>Estradiol valerate/Dienog est</td>
<td>Treatment of heavy menstrual bleeding in women.</td>
<td>-7.63079</td>
<td>6</td>
<td>-19.4</td>
<td>1</td>
<td>0.854</td>
<td>GLN164, GLN163</td>
</tr>
<tr>
<td>7</td>
<td>DB08909</td>
<td>Glycerol Phenylbutyrate</td>
<td>Treatment of urea cycle disorders (UCDs)</td>
<td>-7.46079</td>
<td>7</td>
<td>-16</td>
<td>3</td>
<td>0.85</td>
<td>GLN163</td>
</tr>
<tr>
<td>8</td>
<td>DB00947</td>
<td>Fulvestrant</td>
<td>Treatment of hormone receptor-positive metastatic breast cancer</td>
<td>-6.63079</td>
<td>8</td>
<td>-16</td>
<td>4</td>
<td>0.84</td>
<td>HIS135</td>
</tr>
</tbody>
</table>

Amino acids in bold are from the active site of the panC. NA- Not Applicable
Figure 8.2: Intermolecular interactions between the protein panC and diosmin (ranked 1 and 2 by Glide and ADV respectively). (A) The interaction between the protein panC from Mtb (grey color ribbon form) and the drug diosmin (red color stick form). (B) The grey color solid ribbon form indicates the protein panC while the drug diosmin is in red color stick form. The active site residues of panC are in yellow color surface while (C) displays the interacting amino acids of panC with hydrogen bonds in dotted lines. (D) shows the interactions between the panC and the diosmin in two dimensional view.
Figure 8.3: Intermolecular interactions between the protein *panC* of *M. tuberculosis* and lapatinib (ranked 2 and 6 by Glide and ADV respectively). (A) The interaction between the protein *panC* from *Mtb* (grey color ribbon form) and the drug lapatinib (red color stick form). (B) The grey color solid ribbon form indicates the protein *panC* while the drug lapatinib is in red color stick form. The active site residues of *panC* are in yellow color surface. (C) shows the interacting amino acids of *panC* in the cyan color while the drug lapatinib is displayed in red color stick form with hydrogen bonds in pink color dotted lines. (D) displays the interactions in two dimensional view.
Figure 8.4: Intermolecular interactions between the protein \textit{panC} from \textit{M. tuberculosi}s and estradiol valerate/dienogest (ranked 6 and 1 by Glide and ADV respectively). (A) The interaction between the protein \textit{panC} from \textit{Mtb} (grey color ribbon form) and the drug estradiol valerate/dienogest (red color stick form). (B) The grey color solid ribbon form indicates the protein \textit{panC} while the drug estradiol valerate/dienogest is in red color stick form. The active site residues of \textit{panC} are in yellow color surface. (C) displays interacting amino acids of \textit{panC} in the cyan color with hydrogen bonds in pink color dotted lines while (D) shows the molecular interactions in two dimensional view.
8.5. Discussion

One of the eight prioritized drugs, zafirlukast was found already reported to have anti-mycobacterial activity (Pinault et al., 2013). The researchers reported zafirlukast to inhibit lsr2 and subsequently inhibit the growth of M. tuberculosis. Lsr2 is a nucleoid-associated protein, it influences the organization of chromatin and gene expression with a preference for AT-rich sequences, and bridging distant DNA segments; it also enable the genomic DNA to resist the oxidative damage thus it is essential for M. tuberculosis viability (Colangeli et al., 2009; Pinault et al., 2013 and http://www.uniprot.org/uniprot/P9WIP7). Since zafirlukast was found as one of prioritized drugs and also been reported previously to have anti-mycobacterial activity, other drugs prioritized along with zafirlukast gain importance for further evaluation towards repurposing against tuberculosis.

Glycerol phenylbutyrate was observed to be ranked 3 by ADV with -16 kcal/mol in the present study; Glycerol phenylbutyrate is a nitrogen-binding agent used for the chronic management of urea cycle disorders. Phenylbutyrate is a similar compound to that of glycerol phenylbutyrate that has been demonstrated to inhibit the growth of M. tuberculosis (Coussens et al., 2015) and induce intracellular killing of M. tuberculosis in human macrophages (Rekha et al., 2015). Chemically, glycerol phenylbutyrate is a triglyceride in which three molecules of phenylbutyrate are linked to a glycerol backbone (https://www.drugbank.ca/). Therefore, based on the previously reported anti-mycobacterial activity of phenylbutyrate (Coussens et al., 2015; Rekha et al., 2015) and strong binding of Glycerol phenylbutyrate against panC observed in the present study, it can be suggested that glycerol phenylbutyrate may have potential
biological activity against *M. tuberculosis* though it needs to be experimentally studied.

Since chemically similar molecule of glycerol phenyl butyrate that is phenylbutyrate is known already to have anti-mycobacterial activity, the glycerol phenyl butyrate can also be used as lead molecules to find new drugs with improved activity against *M. tuberculosis*. However, further *in vitro* studies are needed to establish the anti-mycobacterial activity of glycerol phenyl butyrate.

**8.6. Conclusion**

Pantothenate synthetase encoded by *panC* in *M. tuberculosis* is a promising target for drug discovery against tuberculosis, which has also been supported by previous studies. A total of eight drugs were prioritized against *panC* of *M. tuberculosis* from the list of 1554 known drugs consistently by employing four rounds of docking. Since one of these eight prioritized drugs, zafirlukast was already demonstrated to inhibit *M. tuberculosis* growth (Pinault *et al*., 2013), other seven prioritized drugs gain importance. Thus, these seven drugs can be taken up for further experimental studies towards repurposing against tuberculosis.