Target identification & molecular docking of *Burkholderia pseudomallei*

Mitali B Ghormade  
Department of Bioinformatics  
Shri Shivaji Science College, Amravati, India.

**Abstract:**  
Target based or structural based drug designing is the rapidly growing area. The explosion of genomic, proteomic and structural information has provided hundreds of new targets and opportunity to find new drug lead components. Based on burgeoning structural data of *Burkholderia pseudomallei*. I focused on tomato disease, *Burkholderia pseudomallei* also known as *Pseudomonas pseudomallei* is a Gram- negative, bipolar, aerobic, motile rod-shaped bacterium. It infects human and animals and cause the disease Melioidosis, it also infect plants. In this study I have taken 592 protein sequences from NCBI (National Centre of Biotechnological Information) protein database for which protein structure is available at PDB Database. From this study we have selected 4 druggable targets & 4 virulence protein sequences using Tid tool out of detected virulence protein we have choose 4bko protein as target.

**1. Introduction**

Bacterial diseases include any type of illness caused by bacteria. *Burkholderia pseudomallei* (also known as *Pseudomonas pseudomallei*) are a Gram-negative, bipolar, aerobic, motile rod-shaped bacterium. It is a soil-dwelling bacterium endemic in tropical and subtropical regions worldwide, particularly in Thailand and northern Australia. It infects humans and animals and causes the disease Melioidosis. It is also capable of infecting plants. *B. pseudomallei* measures 2–5 μm in length and 0.4–0.8 μm in diameter and is capable of self-propulsion using flagella. The bacteria can grow in a number of artificial nutrient environments, especially betaine- and arginine-containing ones.

*In vitro*, optimal proliferation temperature is reported around 40 °C in neutral or slightly acidic environments (pH 6.8–7.0). The majority of strains are capable of fermentation of sugars without gas formation (most importantly, glucose and galactose; older cultures are reported to also metabolize maltose and starch). Bacteria produce both exo- and endotoxins. The role of the toxins identified in the process of 13 melioidosis symptom development has not been fully elucidated.

The Bacteria *Burkholderia pseudomallei* I have choose because it infects both plant as well as animals & causes serious disease that is Melioidosis. The aim was to access the deuggable targets for *Burkholderia pseudomallei* Tid tool. TiD is a standalone application, which relies on basic assumption that a protein must be essential for
pathogens survival and non-homologous with host to qualify as putative target. With an input bacterial proteome, TiD removes paralogous proteins, picks essential ones, and excludes proteins homologous with host organisms. The targets illustrate non-homology with at least 40 out of 84 gut microbes, considered safe for human. TiD classifies proposed targets as known, novel and virulent. Users can perform pathway analysis, choke point analysis, interactome analysis, subcellular localization and functional annotations through web servers cross-referenced with the application. Drug targets identified by TiD for *Listeria monocytogenes*, *Bacillus anthracis* and *Pseudomonas aeruginosa* have revealed significant overlaps with previous studies. TiD takes < 2 h to scan putative targets from a bacterial proteome with ~ 5000 proteins.

Molecular docking is one of the most frequently used methods in structure-based drug design, due to its ability to predict the binding-conformation of small molecule ligands to the appropriate target binding site. Characterization of the binding behavior plays an important role in rational design of drugs as well as to elucidate fundamental biochemical processes. Docking accuracy represents one measure to quantify the fitness of a docking program by rationalizing the ability to predict the right poses of a ligand with respect to that experimentally observed.

Molecular docking can demonstrate the feasibility of any biochemical reaction as it is carried out before experimental part of any investigation. There are some areas, where molecular docking has revolutionized the findings. Ligand binding is one of the major biochemical functions of proteins, and thus the identification of ligands and their binding sites is the starting point for the function identification.

**RESEARCH METHODOLOGY**

2. Data and Sources of Data

2.1. NCBI(National Center for Biotechnological Information)

The NCBI houses a series of databases relevant to biotechnology and biomedicine and is an important resource for bioinformatics tools and services. Major databases include GenBank for DNA sequences and PubMed, a bibliographic database for the biomedical literature. Other databases include the NCBI Epigenomics database. All these databases are available online through the Entrez search engine. After submitting the query of *Burkholderia pseudomallei* (protein) obtained, from which the first 592 sequences were selected which are about same base pairs.

2.2. PDB (Protein Data Bank)-

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that help students and researchers understand all aspects of biomedicine.
and agriculture, from protein synthesis to health and disease, the PDB ID of *Burkholderia pseudomallei* was downloaded.

2.3. Target identification tool (TiD)

TiD is a standalone application, which relies on basic assumption that a protein must be essential for pathogens survival and non-homologous with host to qualify as putative target. With an input bacterial proteome, TiD removes paralogous proteins, picks essential ones, and excludes proteins homologous with host organisms. The targets illustrate non-homology with at least 40 out of 84 gut microbes, considered safe for human. TiD classifies proposed targets as known, novel and virulent. Users can perform pathway analysis, choke point analysis, interactome analysis, subcellular localization and functional annotations through web servers cross-referenced with the application.

TiD is fast and efficient tool for detecting potential drug target from bacterial proteome. TiD takes < 2 h to scan putative targets from a bacterial proteome with ~ 5000 proteins. Hence, it is a useful tool for rational drug design.

<table>
<thead>
<tr>
<th>Home</th>
<th>Paralog Analysis</th>
<th>Essentiality Analysis</th>
<th>Non Homologous Analysis</th>
<th>Target Prioritization</th>
<th>Single Click Target Mining</th>
<th>Help</th>
</tr>
</thead>
</table>

TiD is fast and efficient tool for detecting potential drug targets from bacterial proteome using subtractive channel analysis. It offers automated target prioritization through gut flora non-homology, druggability and virulent factor analysis. A single click recommended workflow is included in the software to identify drug targets from whole proteome in less than two hours.

**Step1: Paralog Analysis**

- It performs CD-HIT at 60% identity to remove redundant paralogous sequences. After paralog analysis 96 sequences were filter.

**Step2: Essentiality Analysis**

It looks for pathogen specific essential genes in DEG, CEG and common from both DEG & CEG based on threshold E-values and bit scores, after performing essentiality analysis 43 essential sequences was obtained from paralogs analysis result.
Step3: Non-Homologous Analysis

It helps in identifying non-homolog proteins of pathogenic bacteria in different Hosts and Gut flora based on threshold E-values and bit scores, fig(3.7)After performing non homologous analysis for host (i.e. human) 12 sequences were obtained from essentiality analysis result. The targets illustrate non-homology with at least 40 out of 84 gut microbes, considered safe for human. After performing non homologous analysis for Gut Flora 5 sequences were obtained from essentiality analysis result.

Step 4: Target Prioritization

4.1. Druggability Analysis-

It finds homology with druggable target evaluates pathogenicity of target and provides links to online target prioritization servers, in this I performed druggability analysis I obtained 4 sequences from Gut Flora result.

4.2. Virulence Analysis-

After Druggability Analysis the next set was performed, that was Virulence Analysis & there were 3 virulent sequences obtained.

2.4. Basic Local Alignment Search Tool

BLAST for Basic Local Alignment Search Tool is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold. Fig (3.11)-As I obtained 3 virulence sequences against human, but as this bacteria also infected to tomato plant, considering this I perform blast against tomato proteome. All 4 protein id were identified as a target for human and tomato.

Pdb|3N3R|B chain, Pdb|2X3Y|H chain, Pdb|3GK0|H chain, Pdb|4BKO|A chain, Out of which 4BKO was selected for Docking

2.5. AutoDock

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure.

AutoDock consists of two main programs: AutoDock for docking of the ligand to a set of grids describing the target protein; AutoGrid for pre-calculating these grids. The complete interface of autodock tool- Autodock software has different step to complete the Docking process, the steps are such as Ligand, Flexible residues, Grid, Docking, Run and then Analyze. Also other tabs are present on it such as File, 3D Graphics, Edit, Select, Display, Color, Compute, Hydrogen bonds, Grid 3D.
3.4 Statistical tools

3.4.1. Tid tool

<table>
<thead>
<tr>
<th>Steps</th>
<th>Input Sequences</th>
<th>Filter Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paralog Analysis</td>
<td>592</td>
<td>96</td>
</tr>
<tr>
<td>Essentiality Analysis</td>
<td>96</td>
<td>43</td>
</tr>
<tr>
<td>Non-Homologous Analysis</td>
<td>43</td>
<td>12</td>
</tr>
<tr>
<td>Target Prioritization-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) Druggability Analysis</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>ii) Virulence Analysis</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

3.4.2. AutoDock Tool

![AutoDock Tool Interface](image)

Fig- Autodock tool interface
IV. RESULTS AND DISCUSSION

Fig. Result for target (4BKO) and known ligand and the binding energy is -4.93 on conformation 1.

Fig. Result for Target (4BKO) & screened ligand 1 and binding energy is -5.41 on conformation 1.
My first aim for Docking is to find the Protein-Ligand interaction and to provide an opportunity for the designing of active compounds. In order to recognize an accurate docking routine for carrying out molecular docking studies of ASMT protein as well as to identify the potent inhibitors against that protein, four widely used docking routines (AutoDock/Vina, GOLD) are compared in this work. Each docking routine returned top ten ranked docked poses for each ligand. AutoDock/Vina was found to be the best for carrying out blind docking and in generating poses that bind best deep inside the 5 Å of the binding pocket. Evaluation of docking accuracy of docking programs requires the programs run at approximately comparable speeds.

Currently, Molecular Docking is a standard computational tool that has been successfully employed in drug design and discovery studies. Some theoretical and computational challenges remain to be overcome; doing so would increase the predictive power and widen the applications of this important computational tool.

Computer-aided docking is an important tool for gaining understanding of the binding interactions between a ligand (small molecule) and its target receptor (enzyme) and has emerged as a reliable, cost-effective and time-saving technique for the discovery of lead compounds.

The virtual screening approach for docking small molecules into a known protein structure is a powerful tool for drug design and has become an integral part of the drug discovery process. The virtual screening result shows similar ligands for the known ligand. The application of AutoDock in virtual screening is constrained only by the chemical compounds features that can be calculated and the relation between these features and the target.
Computational tools like AutoDock offer the advantage of delivering new drug candidates more quickly and at a lower cost. The use of complementary experimental and informatics techniques increases the chance of success in many stages of the discovery process. AutoDock outputs a result which is the lowest energy conformation of the ligand it found during that run. This conformation is a combination of translation, quaternion and torsion angles and is characterized by intermolecular energy, internal energy and torsional energy. The first two of these combined give the ‘docking energy’ while the first and third give ‘binding energy’. AutoDock also breaks down the total energy into a vdW energy and an electrostatic energy for each atom. Once the calculation has finished it is possible to analyze the conformations obtained and their energy, the lowest energy obtained is -4.93 (binding energy is the sum of the intermolecular energy and the torsional free-energy penalty) and other factors such as docking energy (docking energy is the sum of the intermolecular energy and the ligand’s internal energy) ligand efficiency, inhibition constant, inhibition constant units, intermolecular energy, electrostatic energy, torsional energy, rseed1 and rseed2, unbound energy.

Acknowledgment

I would like to thank my beloved father Mr. Babarao G. Ghormade and inspiring mother Mrs. Jayamala B. Ghormade and brother Mr. Nikhil B. Ghormade being such a wonderful family. They all stood by me, encouraged me and held me whenever I fell and I am also thankful of my friend Gaurav.

References


