International Journal of Pharma and Bio Sciences V1(2)2010

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

PRADEEP K. NAIK*, SENEHA SANTOSHI AND ASHIMA BIRMANI

Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan 173215, Himachal Pradesh, India.

***Corresponding Author** pknaik73@rediffmail.com .

ABSTRACT

Pseudomonas syringae is a plant pathogen which infects a wide range of plant species, and is also the causative organism of bacterial brown spot of bean. The complete genome sequence of pathogen has provided all the determinants of host-pathogen interactions and possible drug targets, computational methods for selecting suitable candidates for further experimental analyses are currently limited. We have performed comparative analysis of genomes and metabolic pathways of the pathogen *P. syringae* and the host *phaseolus vulgaris* (kidney bean) and identified eight potent therapeutic targets. After critical evaluations of the targets we have considered murD and murE as the potent targets for virtual screening of lead molecules. Virtual screening was carried out using the high throughput virtual screening protocol and PubChem drug database. The interaction studies provided promising ligands for future experimental screening to inhibit the proliferation of *Pseudomonas syringae*.

KEY WORDS

Pseudomonas, database of Essential Genes (DEG), comparative microbial genomics, comparative metabolomics, homology, drug targets, virtual screening.

INTRODUCTION

The cost of research and development in the pharmaceutical industry has been rising steeply and steadily in the last decade, but the amount of time required bringing a new product to market remains around ten to fifteen years[1]. It necessitates investment in inexpensive technologies that shorten the length of time spent in drug discovery. The target identification stage is the first step in the drug discovery process [2] and as such can provide the foundation for years of dedicated research in the pharmaceutical industry. This stage is complicated by the fact that the identified drug target must satisfy a variety of criteria to permit progression to the next stage. Important factors in this context include homology between target and host[3], activity of the target in the diseased state[4,5] and the essentiality of the target to the pathogen's growth and survival[6,7,8]. In order to prevent host toxicity such homology must be low or nonexistent. The values of some of these selection criteria

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

can be found easily by querying publicly available bioinformatics resources, including metabolic pathway databases such as KEGG (Kyoto encyclopedia of genes and genomes)[9], protein classification sets such as COGs (clusters of orthologous groups) and databases of ‘‘druggable’’ (potentially useful as drug targets) proteins [4,10,11].

Availability of genome sequences of pathogens has provided a tremendous amount of information that can be useful in drug target identification [12,13]. The strategies for drug design and development are progressively shifting from the genetic approach to the genomic and metabolomic approach. Novel drug targets are required in order to design new defense against antibiotic resistance pathogens. Comparative genomics and metabolomics provide new opportunities for finding optimal targets among previously unexplored cellular functions based on an understanding of their related biological processes in bacterial pathogens and their hosts. In general, a target should provide adequate selectivity; yielding a drug which is specific or highly selective against the pathogen with respect to the host[14,15]. Moreover, the target should be essential for growth and viability of the pathogen at least under the condition of infection. A gene is deemed to be essential if the cell cannot tolerate its inactivation by mutation, and its status is confirmed using conditional lethal mutants. The microbial target for treatment should not have any well-conserved homolog in the host, in order to address cytotoxicity issues. This can help to avoid expensive dead-ends when a lead target is identified and investigated in great detail only to discover at a later stage that all its inhibitors are invariably toxic to the host. Genes that are conserved in different genomes often turn out to be essential [16,17,18,19].

A simple and efficient computational tool designated as "subtractive genomics" and “comparative metabolomics” can determine concordances of putative gene products showing sets of proteins conserved across one set of user-specified genomes, but are not present in another set of user-specified genomes [20,21,22]. The functions encoded by essential genes are considered to constitute the foundation of life of the organism, and are therefore likely to be common to all cells [23]. Identification and characterization of essential genes for the establishment and/or maintenance of infection may be the basis to elaborate novel and effective antimicrobials against bacteria, especially if these genes are conserved in various bacterial pathogens. With the availability of Database of Essential Genes (DEG)[24], it is now possible to predict the list of essential genes in a bacterial pathogen. Galperin and Koonin (2004) suggested searching for drug targets among previously characterized proteins that are specific and essential for a particular pathogen. Computational pathway analysis also facilitates the identification of enzymes that participate in several pathways. This approach relies on the fact that comparison of pathways present in host and in microbial pathogen reveal unique essential pathways present in the pathogen and further essential enzymes that are unique for the pathogen but absent in host thus suitable as the drug targets. Using metabolic pathway information as the starting point for the identification of potential targets has its advantages as each step in the pathway is validated as essential function for the survival of the bacterium.

In this study we have taken *Pseudomonas syringae* (B728a), whole genomes and applied subtractive genomics and comparative metabolomics

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

approaches in order to identify a list of potential therapeutic targets. *Pseudomonas syringae* is a rod shaped, Gram-negative bacterium with polar flagella. It is a member of the Pseudomonas genus, and based on 16S rRNA analysis. *P. syringae* has been placed in the *P. syringae* group[25]. It is a plant pathogen which can infect a wide range of plant species, and exists as over 50 different pathovars, *P. syringae* strain B728a is the cause of bacterial brown spot of bean[25]. It is a highly fit colonist of bean with predominant leaf surface localization[25,26]. It is also active in ice nucleation at warm subfreezing temperatures[27]. Brown spot is a recurring problem that can cause serious losses in snap beans. Halo blight has caused serious losses in both dry and snap beans. Moreover the bacterium is quite variable in nature which has made the development of resistant varieties difficult. These resources provide a basis for addressing the "complexities and conundrums" in drug discovery by computational methods.

Here, we demonstrate the unprecedented potential of the available complementary genome datasets, and the application of both subtractive genomics and comparative metabolomics approaches for the identification of essential genes that may be considered as candidates for antibacterial drug discovery, using *P. syringae* as an example. As a proof of concept, many of the genes identified by our approaches are also reported as essential by experimental methods. Furthermore, our approach successfully identified a number of promising protein targets for new antibiotic development.

MATERIALS AND METHODS

In this work we discuss the current state of the art for some of the bioinformatics approaches to identifying drug targets. It makes use of database of essential genes (DEG) (http://tubic.tju.edu.cn/deg/)[24] and the comparative genomic as well as comparative metabolomics approaches to compare with the pathogen bacteria versus host, including identifying new members of successful target classes and their functions, predicting disease relevant genes.

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

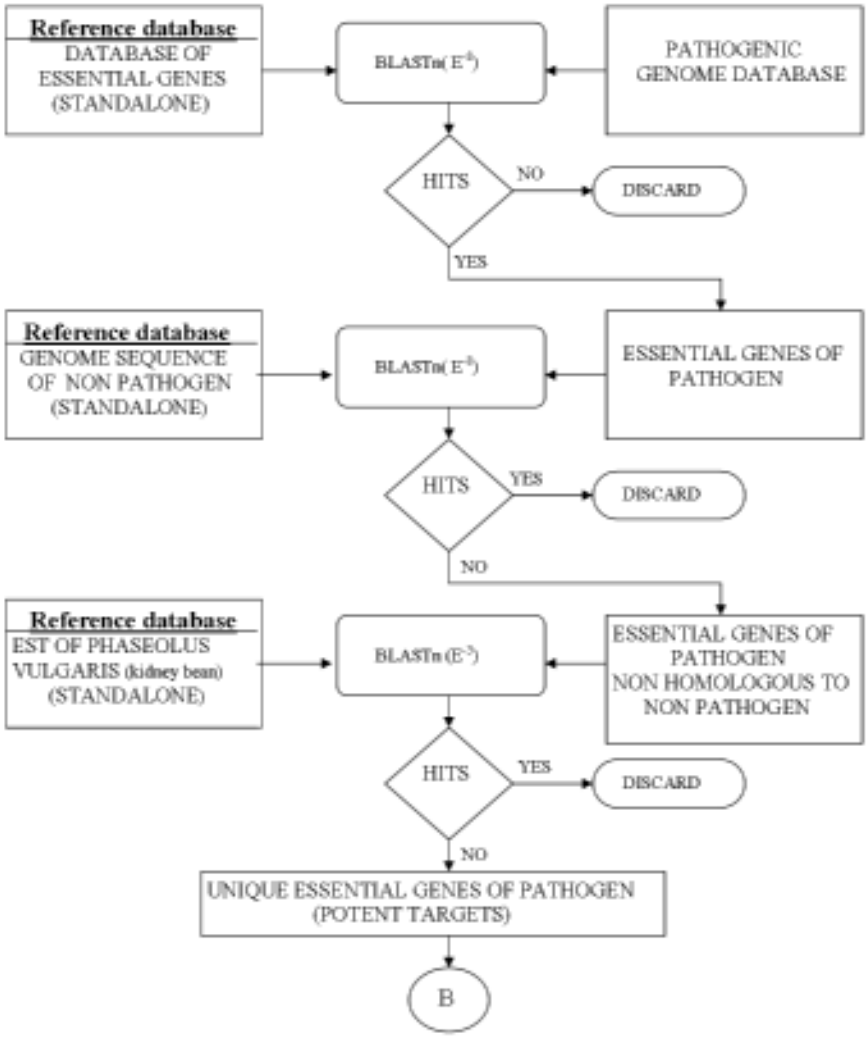


Figure 1. ***Flowchart for excluding essential genes/proteins of Pseudomonas syringae having homology to non-pathogen and host.***

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

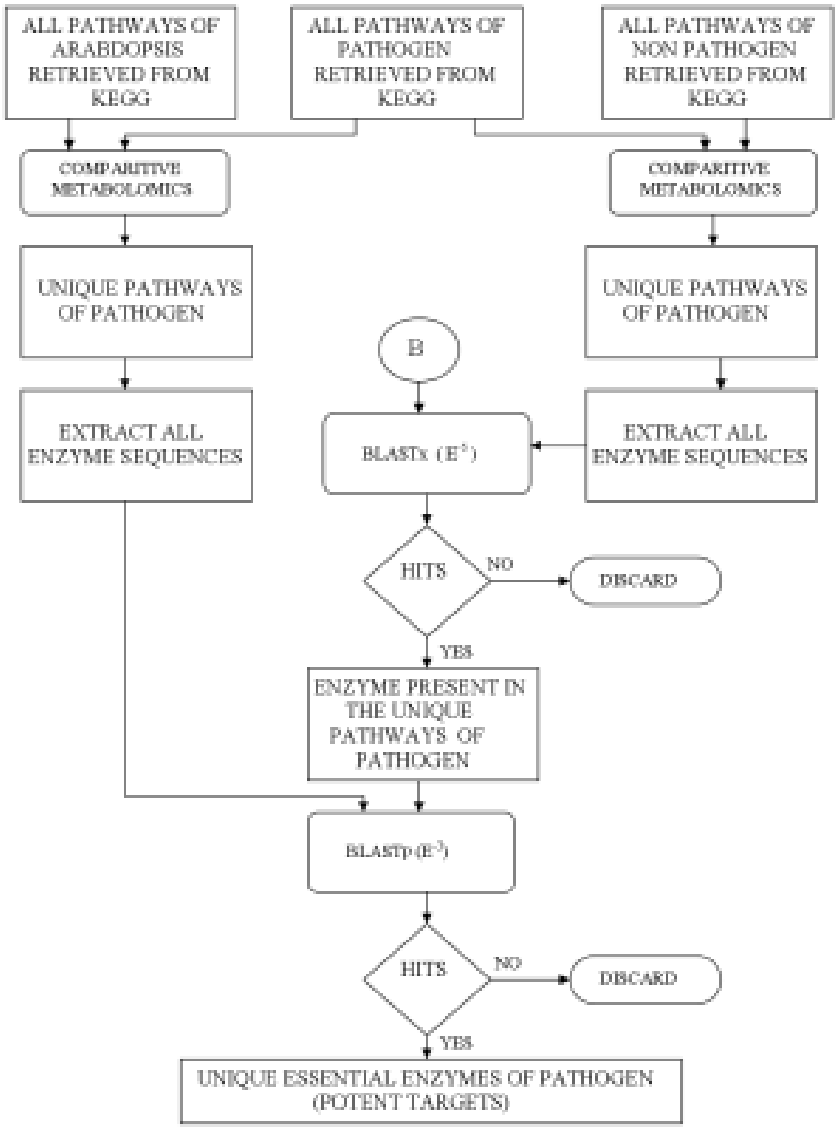


Figure 2. ***Flowchart for excluding metabolic pathway enzymes in Pseudomonas syringae having homology to non-pathogen and host.***

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

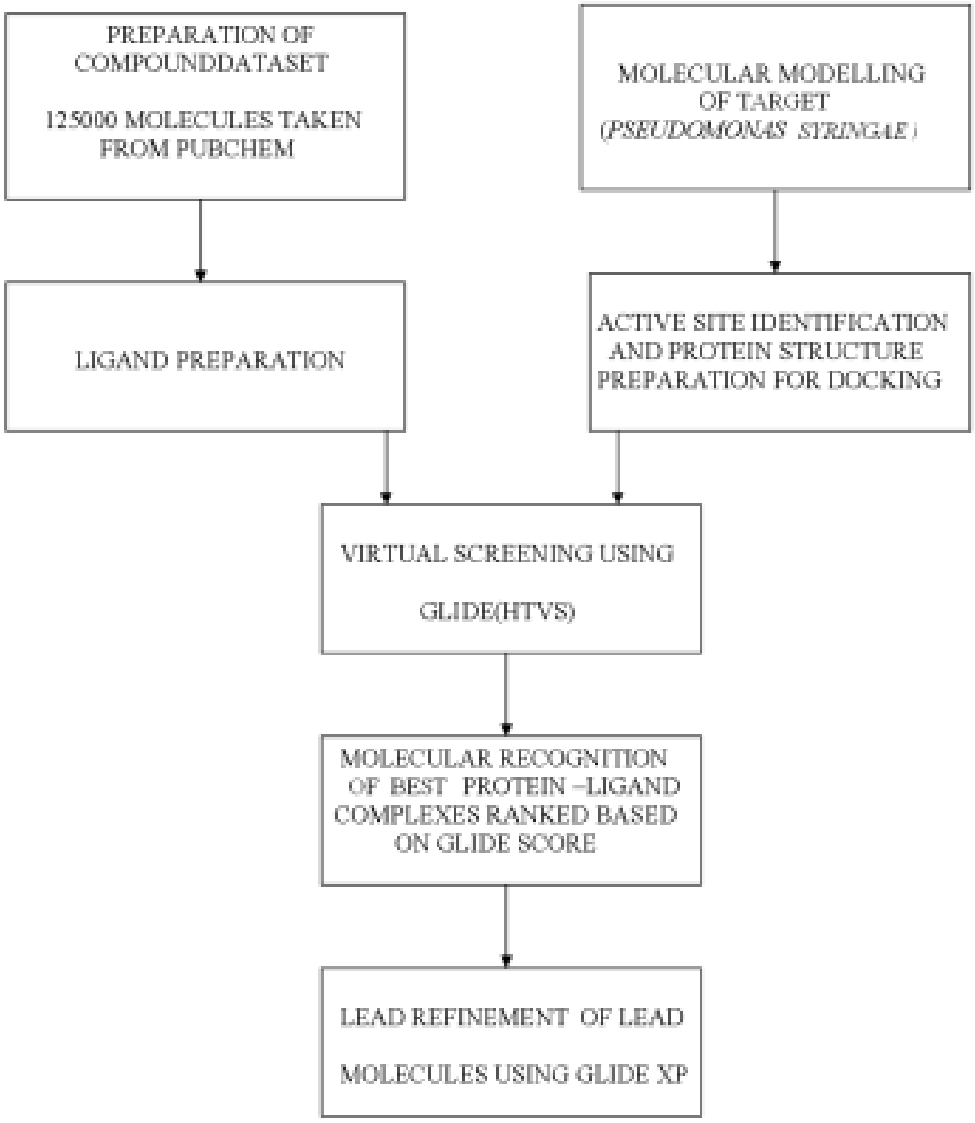


Figure 3. ***Workflow diagram followed in the in silico virtual screening for inhibitors of murD and murE of Pseudomonas syringae.***

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

Table 1.

Unique targets identified in Pseudomonas syringae (B728a) after three levels of subtraction based on comparative genomics

	Gene ID	Gene name	DEG ID	Score	Metabolic pathway
	CARBOHYDRATE METABOLISM				
1	Psyr_0517		DEG10040028	66	Glycolysis / Gluconeogenesis
2	Psyr_0517		DEG10040028	66	Pyruvate metabolism
3	Psyr_1120		DEG10040316	62	Pentose phosphate pathway
4	Psyr_2176		DEG10110051	70	Propanoate metabolism
5	Psyr_2988		DEG10100164	68	Pentose phosphate pathway
	ENERGY METABOLISM				
1	Psyr_1238		DEG10040392	70	Nitrogen metabolism
2	Psyr_3483		DEG10040067	70	Nitrogen metabolism
	LIPID METABOLISM				
1	Psyr_0500		DEG10100527	74	Fatty acid biosynthesis
2	Psyr_1349		DEG10040043	62	Biosynthesis of steroids
3	Psyr_1647	fabG	DEG10040181	72	Fatty acid biosynthesis
4	Psyr_1647	fabG	DEG10040181	72	Biosynthesis of unsaturated fatty acids
5	Psyr_1754		DEG10110097	62	Fatty acid biosynthesis
6	Psyr_2176		DEG10110051	70	Fatty acid metabolism
	NUCLEOTIDE METABOLISM				
1	Psyr_1269		DEG10130197	72	Purine metabolism
	AMINO ACID METABOLISM				
1	Psyr_0012	glyQ	DEG10040536	230	Glycine, serine and threonine metabolism
2	Psyr_0183	dapF	DEG10040559	64	Lysine biosynthesis
3	Psyr_0487		DEG10040446	66	Glutamate metabolism
4	Psyr_0517		DEG10040028	66	Alanine and aspartate metabolism
5	Psyr_1735	cysS	DEG10040100	74	Cysteine metabolism
6	Psyr_1985		DEG10100480	66	Valine, leucine and isoleucine biosynthesis
7	Psyr_2176		DEG10110051	70	Valine, leucine and isoleucine degradation
8	Psyr_4107	murE	DEG10110006	64	Lysine biosynthesis
9	Psyr_4893		DEG10100263	66	Histidine metabolism
10	Psyr_4104	murD	DEG10110006	52	D-Glutamine and D-glutamate metabolism
	METABOLISM OF OTHER AMINO ACID METABOLISM				
1	Psyr_0487		DEG10040446	66	Glutathione metabolism
2	Psyr_1120		DEG10040316	62	Glutathione metabolism
3	Psyr_2176		DEG10110051	70	beta-Alanine metabolism
4	Psyr_2988		DEG10100164	68	Glutathione metabolism
	GLYCAN BIOSYNTHESIS AND METABOLISM				
1	syr_4107	murE	DEG10110006	64	Peptidoglycan biosynthesis
2	Psyr_4104	murD	DEG10110006	52	Peptidoglycan biosynthesis
	METABOLISM OF COFACTORS AND VITAMINS				
1	Psyr_0387		DEG10040563	96	Ubiquinone biosynthesis
	TRANSLATION				
1	Psyr_0012	glyQ	DEG10040536	230	Aminoacyl-tRNA biosynthesis
2	Psyr_1735	cysS	DEG10040100	74	Aminoacyl-tRNA biosynthesis
3	Psyr_2165	rplT	DEG10010205	64	Ribosome
4	Psyr_4120	rplM	DEG10040487	86	Ribosome
5	Psyr_4544	rpsS	DEG10040508	58	Ribosome
6	syr_4639	rpsU	DEG10130100	60	Ribosome
	MEMBRANE TRANSPORT				
1	Psyr_0759		DEG10100446	60	ABC transporters
2	Psyr_3076		DEG10110057	62	ABC transporters

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

Identification of unique targets based on comparative genomics

Gene sequences for *P. syringae* and ESTs of *phaseolus vulgaris* (the host) were downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/genomes/). The DEG database[24] was downloaded from http://tubic.tju.edu.cn/deg/ and manually compiled to use as a stand-alone database for the BLAST program[28]. The *P. syringae* genes were purged at 60% using CD-HIT to exclude paralogs from further analysis. The standalone BLAST executables including BLASTn, BLASTp and BLASTx were downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/) and installed locally. The set of essential genes in *P. syringae* have been predicted based on homologous sequence search against DEG using BLASTn (E-value 10^{-8}). An interesting approach designated "differential genome display" has been proposed for the prediction of potential drug targets[20,21]. The genes that are present in the genome of a pathogenic bacterium, but absent in the genome of a closely related free-living bacterium (non-pathogenic), are therefore likely to be important for pathogenicity and may be considered candidate drug targets. The resultant set of essential genes of pathogenic bacteria was subjected to BLASTn (E-value, 10^{-8}) against complete genome of non pathogenic bacteria (*Pseudomonas putida*; KT2440). The subtracted essential genes of the pathogen were subjected to BLASTn (E-value, 10^{-3}) with ESTs of *phaseolus vulgaris* to identify pathogen genes non-homologous to the host. These identified genes were classified into different groups based on gene name, gene id and the pathway in which the gene is involved. These sets of genes were named as set B as shown in Table 1. This set of unique essential genes can be considered as potent targets. From this set of genes, no genes were selected for lead designing and virtual screening, as it was further carried forward to be compared and analyzed through comparative metabolomics approach. ***Further subtraction of the gene set obtained using comparative metabolomics approach.***

In this approach two step comparison of metabolic pathways were performed. First the entire metabolic pathways from the pathogen (*P. syringae*) were compared with the model plant, (Arabidiopsis) as host since the metabolic pathways from the natural host (*phaseolus vulgaris*) were not available. Secondly the entire metabolic pathways from the pathogen (*P. syringae*) were compared with the entire metabolic pathways of non-pathogen (*pseudomonas putida*; KT2440). Metabolic pathways for these organisms were downloaded from the database Kyoto Encyclopedia of Genes and Genomes (KEGG)[29]. Comparison of entire metabolic pathways between *Pseudomonas syringaea* (KEGG i.d. –psb) and Arabidopsis resulted in 24 unique pathways present only in *P. syringae* (Table 2) and could be considered as putative targets. Comparative study of metabolic pathways between pathogen and non-pathogen bacteria resulted in 2 unique pathways present only in the pathogen (*P. syringae*) and thus relevant to pathogenesity (Table 2). The enzymes from these pathways were subjected to BLASTx (E-value 10^{-5}) against the dataset of unique essential genes obtained

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

during comparative genomics approach (Table 1) to check whether the enzymes are essential to the pathogen or not for its survive. All these enzyme sequences were subjected to BLASTp (E-value 10^{-3}) against the data set obtained after comparison between *P. syringae* and Arabidopsis in the first step. The obtained unique essential enzymes (Table 3) are thus could be considered as potent therapeutic targets. Out of 8 unique enzymes only murD and murE were considered for virtual screening of lead molecules. The overall protocol used for further filtration of targets based on comparative metabolomics approach is represented in Figure 2

Table 2.
Unique metabolic pathways identified in Pseudomonassyringae (B728a) after comparative metabolomics.

Unique pathways of *Pseudomonas syringae* in comparison to *Pseudomonas putida*
Lipid Metabolism
C21-steroid hormone metabolism
Androgen and estrogen metabolism
Unique pathways of *Pseudomonas syringae* in comparison to *Arabidopsis thaliana*
Carbohydrate Metabolism
C5-Branched dibasic acid metabolism
Lipid Metabolism
C21-steroid hormone metabolism
Androgen and estrogen metabolism
Metabolism of Other Amino Acids
D-Glutamine and D-glutamate metabolism
D-Alanine metabolism
Biosynthesis of Polyketides and Nonribosomal Peptides
Polyketide sugar unit biosynthesis
Biosynthesis of Secondary Metabolites
Streptomycin biosynthesis
Novobiocin biosynthesis
Xenobiotics Biodegradation and Metabolism
Caprolactam degradation
Toluene and xylene degradation
2,4-Dichlorobenzoate degradation
1,2-Dichloroethane degradation
Ethylbenzene degradation
Fluorene degradation
Atrazine degradation
Bisphenol A degradation
Trinitrotoluene degradation
Geraniol degradation
Folding, Sorting and Degradation
Type II secretion system
Type III secretion system
Membrane Transport
Phosphotransferase system (PTS)
Signal Transduction

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

Two-component system
Cell Motility
Bacterial chemotaxis
Flagellar assembly

Table 3.

Eight potential targets identified in Pseudomonas syringae after the complete genomic and metabolomic comparisons

METABOLISM OF COFACTORS AND VITAMINS			
GENE ID	Name	Protein name	pathway
Psyr_0387		2-polyprenylphenol 6-hydroxylase	Ubiquinone biosynthesis
LIPID METABOLISM			
Psyr_2176		Acyl-CoA dehydrogenase, C-terminal:Acyl-CoA dehydrogenase, central region:Acyl-CoA dehydrogenase, N-terminal	Fatty acid metabolism
AMINO ACID METABOLISM			
Psyr_2177		Acyl-CoA dehydrogenase, C-terminal:Acyl-CoA dehydrogenase, central region:Acyl-CoA dehydrogenase, N-terminal	Valine, leucine and isoleucine degradation
Psyr_4107	murE	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	Lysine biosynthesis
CARBOHYDRATE METABOLISM			
Psyr_2178		Acyl-CoA dehydrogenase, C-terminal:Acyl-CoA dehydrogenase, central region:Acyl-CoA dehydrogenase, N-terminal	Propanoate metabolism
OTHER AMINO ACID METABOLISM			
Psyr_2178		Acyl-CoA dehydrogenase, C-terminal:Acyl-CoA dehydrogenase, central region:Acyl-CoA dehydrogenase, N-terminal	beta-Alanine metabolism
Psyr_3615		aminotransferase AlaT	
Psyr_4104	murD	UDP-N-acetylmuramoylalanyl-D-glutamate synthetase	D-glutamine&Dglutamate metabo metabolism
GLYCAN BIOSYNTHESIS AND METABOLISM			
Psyr_4107	murE	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	Peptidoglycan biosynthesis
Psyr_4104	murD	UDP-N-acetylmuramoylalanyl-D-glutamate synthetase	Peptidoglycan biosynthesis

Sequence analysis

Domain analysis (ProDom, Pfam) was carried out to check whether any domain present in the target protein could interfere in the action of the drug. ProDom, a database of protein domain families, is useful for analyzing the protein domain arrangements and helps to analyze homology relationship in modular proteins. The ProDom building procedure MKDOM2 is based on recursive PSI-BLAST searches. The source protein sequences are non-fragmentary, derived from SWISS-PROT and TrEMBL databases. Whereas, Pfam, a database of multiple alignments and HMM profiles of protein domains or conserved protein regions, is actually composed of two sets of families. Pfam-A family is based on curated multiple alignment whereas Pfam-B family is derived from ProDom, a comprehensive set of protein domain family. The screening of the

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

database in this study proved that the domains present in murD and murE of *pseudomonas syringae* are not ubiquitous in host.

BLASTp of murD protein sequence against PDB databank gave two significant hits:1EEH and 2UAG. These proteins were different forms of UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase from *E. coli* and had given a hit of 50% identity, 64% positive scores and 4% gaps with an expect value of 1.1. Again these sequences were aligned using clustalW (Thompson et al.1997), obtaining a score of 50. Similarly BLASTp of murE sequence with PDB databank gave two significant hits: 3UAG and 1E8C. These proteins were different forms of UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase from *E. coli* had given a hit of 41% identity, 62% positive scores and 0% gaps with an expect value of 0.07. Again these sequences were aligned using clustalW[30], obtaining a score of 48. This provided the broadly studied templates to model the target proteins.

Homology model construction

The homology models of the proteins: murD and murE were built using Prime (Prime version 1.5, Macromodel version 9.1, Schrodinger, LLC, New York, NY, 2005) accessible through the Maestro interface (Schrodinger, Inc.). All water molecules were removed and the bound ligand was kept. During the homology model building, Prime keeps the backbone rigid for the cases in which the backbone does not need to be reconstructed due to gaps in the alignment. The model was screened for unfavorable steric contacts and remodeled using a rotamer library database of Prime. Explicit hydrogens were added to the protein and the protein model subjected to energy minimization using the Macromodel (Prime version 1.5) force-field MMFFS. Energy minimization and relaxation of the loop regions was performed using 300 iterations in a simple minimization method. Again the steepest descent was carried out until the energy showed stability in the sequential repetition. Model evaluation was performed using PROCHECK v3.4.4 [31] producing plots that were analyzed for the overall and residue-by-residue geometry. Ramachandran Plot[32] provided by the program PROCHECK assured very good confidence for the predicted protein. There were only 0.3% residues in the disallowed region and 0.9% residues in generously allowed regions. Nevertheless, PROCHECK assured the reliability of the structure and the protein was subjected to VERIFY3D[33], available from NIH MBI Laboratory Servers.

Ligand binding site prediction

In silico prediction of binding sites of these proteins was done using SiteMap (Schrodinger Inc.). SiteMap treat entire proteins to locate binding sites whose size, functionality, and extent of solvent exposure meet user specifications. SiteScore, the scoring function used to assess a site's propensity for ligand binding, accurately ranks possible binding sites to eliminate those not likely to be pharmaceutically relevant. It identifies potential ligand binding sites by linking together "site points" that are suitably close to the protein surface and sufficiently well sheltered from the solvent. Given that similar terms dominate the site scoring function, this approach ensures that the search focuses on regions of the protein most likely to produce tight protein-ligand or protein-protein binding. Subsites are merged into larger sites when they are

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

sufficiently close and could be bridged in solvent-exposed regions by ligand atoms. SiteMap evaluates sites using a series of properties. The binding site with highest site score was taken for docking and virtual screening of the lead molecules.

Docking studies

A library of 1,25,000 compounds taken from pubchem drug database (http://zinc.docking.org/vendor0/index_nfs.shtml) were compiled together to form a standalone library. These compounds were available in 3D-MOL2 file. These molecules were imported into maestro and finally prepared using ligPrep. LigPrep is a utility of Schrodinger suit that combines tools for generating 3D structures from 1D and 2D representation, searching for tautomers and steric isomers and perform a geometry minimization of ligands. The Schrodinger Glide program version 4.0 has been used for docking. The best 10 poses and corresponding scores have been evaluated using Glide high throughput virtual screening (HTVS) (Glide HTVS) for each ligand from the virtual libarary. For each screened ligand, the pose with the lowest Glide HTVS score has been taken as the input for the Glide calculation in extra precision mode (Glide XP). To soften the potential for non-polar parts of the receptor, we scaled vander Waals radii of receptor atoms by 1.00 with partial atomic charge 0.25. The workflow used for the virtual screening is depicted in Figure 3.

RESULTS AND DISCUSSION

One major criterion of the modern drug discovery is the specificity and accuracy. The drug should never disturb any normal process of the host system. The objective of the work was to find and locate those essential genes of *P. syringae* that play important roles in the normal functioning of the bacterium within the host and to shortlist them in the view of drug targeting. Identification of essential and non-host homologs genes of *P. syringae* with subsequent screening of the proteome to find the corresponding protein products are likely to lead to development of drugs that specifically interact with the pathogen. Computational methods have been previously employed for the identification of probable drug targets by predicting only the essential genes in pathogen bacteria[34,14]. However, these works simply compare the bacterial genome sequences with host genome only to discard those essential genes which are homologous to host. Those genes whose products have sequence similarities with the host may lead to drug reactions with the host and thus led to toxic effects. Therefore, here we have excluded those essential genes having sequence similarities with the host and considered only unique essential genes which present only in the pathogen bacteria.

Comparative genomics of P. syringae

We have used two step processes of subtracting genome sequences of *P. syringae* to predict possible therapeutic targets from the pathogenic bacteria. First the genes of pathogenic bacteria were subjected to BLASTn (E-value 10^{-8}) against database of essential genes to find out the essential genes in the pathogenic bacteria. The predicted essential genes of pathogenic bacteria were compared with genome of non-pathogenic bacteria using BLASTn (E-value10^{-8}). It is revealed that out of 160 essential genes only 43 genes are uniquely present in *P. syringea*, which

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

could be related with pathogenicity. The essential genes obtained after first subtraction were then compared with ESTs of *Phaseolus vulgaris* using BLASTn (E-value 10^{-3}), 31 essential genes were uniquely identified from *P. syringea.* The predicted probable therapeutic targets based on comparative genomic approaches were taken further for comparative metabolomic.

Comparative metabolomic

The targeting of biosynthetic pathways has several advantages. Each step in the pathway is already well validated as an essential function for bacterial growth. Discarding target enzymes from the pathogen which share a similarity with the host proteins ensures that the targets have nothing in common with the host proteins, thereby, eliminating undesired host protein–drug interactions. Metabolic pathways belonging to the pathogen and the host are compared and pathways present in the pathogen but not in the host are considered as unique pathway. This results in further reduction in the number of unique therapeutic targets. Comparison of metabolic pathways of the pathogen (*P. syringae*) and the non pathogenic bacteria (*Pseudomonas putida* KT2440) resulted in 2 pathways (Table 2) which are unique to the pathogen. The enzyme sequences of these pathways were further checked with the resultant data set obtained from comparative genomics study to check if any matches. These are the essential unique enzymes present in the unique pathways of the pathogen. Similarly comparative study of metabolic pathways between the host (Arabidopsis) and *P. syringae* was carried out which resulted in 24 unique metabolic pathways present only in the pathogen (Table 2). The enzyme sequences of these 24 pathways were subjected to BLASTp (E-value 10^{-3}) against the dataset obtained after comparison of pathogenic and non-pathogenic bacteria. Only the enzymes which showed matches were considered as potent therapeutic targets for the pathogen. A total of 8 potent targets have been identified (Table 3). These 8 proteins underlie the important aspects of host-pathogen interaction and would therefore represent promising candidates for further study and characterization with an intention for lead design.

The targeted genes that have been selected further for lead design in this study include murD and murE. MurE encodes for UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase involved in peptidoglycan biosynthesis and lysine biosynthesis. Whereas murD encodes for UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase involved in D-Glutamine and D-glutamate metabolism and also in peptidoglycan biosynthesis.

One of the most attractive targets for new antibacterial compounds is the bacterial peptidoglycan biosynthetic pathway[35]. Peptidoglycan is an essential component of the bacterial cell wall which completely surrounds the cytoplasmic membrane and thereby provides a mechanical protection against the turgor pressure of the cytoplasm. Since the osmoprotective function is required in continuity throughout the cell cycle, peptidoglycan metabolism is intimately involved in cell division[36]. It is responsible for a defined cell shape and preserves cell integrity by compensating internal osmotic pressure. Any perturbation of the multi-step peptidoglycan biosynthesis may lead to cell lysis[35]. Peptidoglycan also provides a scaffold to anchor various surface polymers that interact with

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

host cells and the immune system [37,38]. As a result, the biosynthetic pathway involving UDP-*N*-acetylmuramoylpentapeptide (cytoplasmic peptidoglycan precursor) represents an attractive target for the development of new antibacterial agents. Peptidoglycan is formed as a linear chain of repeating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) units, interconnected by short peptide chains. Four ADP-forming ligases (MurC, MurD, MurE and MurF) catalyze the assembly of the peptide moiety by the successive additions of *L*-alanine, *D*-glutamate, *meso*-diaminopimelate (or *L*-lysine) and *D*-alanyl-*D*-alanine to UDP-Mur*N*Ac. These essential cytoplasmic enzymes are highly specific, thus making them attractive as targets for the development of new therapeutic agents against bacterial infections[39]. Out of the enzymes involved in peptidoglycan synthesis, MurD and MurE mediates the formation of a peptide bond between the γ-carboxylate of D-glutamate and the amino group of L-lysine. Presumably these structures render the exposed peptidoglycan resistant to the action of proteases, but they also imply that the active sites of the enzymes must have unusual structures in order to handle the somewhat uncommon substrates. These unusual active sites are targets to bind novel inhibitors that can have antimicrobial activity. Structural studies of enzymes involved in the pathway have already provided strategies for the rational design of novel inhibitors.

Virtual screening of lead molecule

The impact of microbial genomics on drug discovery has led to the identification of novel antibacterial drugs. Enzymes mediate the synthesis of many complex molecules from simpler ones in a series of chemical reactions. Targeting enzymes present in the pathogen but absent in the host ensures the elimination of pseudo drug targets in the pathways[17]. The computational genomics[14] as well as comparative metabolomics approaches stated herein, are likely to speed up drug discovery process by removing hindrances like dead ends or toxicity that are encountered in classical approaches. The strategy is also likely to locate critical pathways and stages in the pathogenecity of *P. syringae*. Sequence as well as structural analysis was carried out to avoid any interference in the normal functioning of the host system. Although, in general, sequence dissimilarity means possible dissimilar binding sites, however, exceptions where targets with significant sequence dissimilarity bind similar molecules do exist. Presumably, screening against such novel targets for functional inhibitors will result in discovery of novel therapeutic compounds active against bacteria, including the increased number of antibiotic resistant clinical strains[40].

Homology modeling protocol was employed to predict the 3D structure of the selected proteins viz. murD and murE. Three dimensional structure predictions by comparative modeling were done using Prime (Schrodinger Inc.). We used both PROCHECK and the VERIFY3D softwares to check the quality of the modeled protein. Ramachandran plot obtained from the program PROCHECK, which checks the stereochemical quality of a protein structures, producing a number of postscript plots, analyzing its overall and residue-by-residue geometry, assured the reliability of the modeled protein with 91.1% and 92.6% residues in most allowed region as well as 7.7% and 6.3% in additional allowed region respectively. There were only 0.36% and 0.41% residues in disallowed region

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

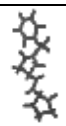
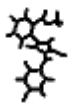
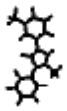
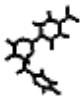
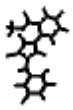
as well as 0.84% and 0.69% respectively in generously allowed region. The assessment with VERIFY3D, which derives a ‘‘3D-1D’’ profile based on the local environment of each residue, described by the statistical preferences for the area of the residue that is buried, the fraction of side-chain area that is covered by polar atoms (oxygen and nitrogen), and the local secondary structure, also substantiated the reliability of the three dimensional structure. The residues that deviated from the standard conformational angles of Ramachandran plot were found to be more than 8 A° distance to the active site residues, which suggested that those residues would interfere little with the binding of ligands in the active site region. The binding sites were predicted using Sitemap (Schrodinger Inc.). For each targeted protein the binding site with maximum site score was selected for docking and virtual screening of lead molecules using Glide (Schrodinger Inc.). The standard methodology accepted for virtual screening with the program Glide involves, applying the parameters of library screening followed by the standard default setting. Library screening was done to filter the non-docked compounds from database we used for virtual screening. After filtering non-docked compounds, remaining compounds were used for detailed docking. Docking procedure consisted of three interrelated components; (a) identification of binding site, (b) a search algorithm to effectively sample the search space (the set of possible ligand positions and conformations on the protein surface) and (c) a scoring function. For each ligand in the virtual library, the pose with the lowest Glide (HTVS) score was refined using Glide (XP) docking. The best ligands as shown in Table 4-5 respectively, chosen with Glide score proved their reliability in the Glide (XP).

Table 4.

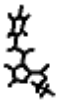
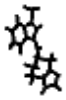
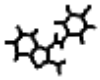
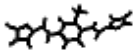
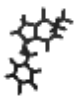
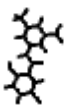
Database ID number, chemical structure, IUPAC names and Glide scores of top scored ligands docked with murD of Pseudomonas syringae using docking program Glide.

S. No.	Ligand structure	Ligand Name	ZINC ID	Glide (HTVS) score	Glide (XP) score
1		**methyl 4-methyl-3-[(4,5,6,7-tetrahydro-1-benzothien-3-ylcarbonyl)amino]benzoate**	ZINC00999929	-6.268	-9.361
2		methyl 2-[3-chloro-4-(3-chloro-2-methylanilino)-2,5-dioxopyrrol-1-yl]benzoate	ZINC01938126	-6.636	-9.349
3		3-[(5-amino-2-chlorophenyl)amino]-1-(2,4,6-trichlorophenyl)-1H-pyrazol-5-ol	ZINC01031540	-6.591	-9.172

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

4		N-cyclopentyl-5-oxo-1-(1-phenylethyl)pyrrolidine-3-carboxamide	ZINC00219356	-6.046	-9.169
5		(6R,8R,9R)-6-(2-fluorophenyl)-9-methyl-7-oxo-5,6,8,9,10, 11-hexahydrobenzo[c][1,5]benzodiazepine-8-carboxylate	ZINC00720035	-6.061	-9.138
6		8-(2-hydroxyethylamino)-1, 3-dimethyl-7-(naphthalen-1-ylmethyl)-9H-purin-7-ium-2,6-dione	ZINC02391940	-6.874	-9.093
7		(2R)-2-[(3,5-dioxo-2H-1,2,4-triazin-6-yl)amino]propanehydrazide	ZINC01317711	-6.861	-9.092
8		methyl 2-[3-chloro-4-(4-chloroanilino)-2,5-dioxopyrrol-1-yl]benzoate	ZINC01938091	-6.674	-8.954
9		5-methyl-4-phenyl-2-[3-(trifluoromethyl)phenyl]-1H-pyrazol-3-one	ZINC03133798	-6.889	-8.895
10		4-[[3-(3,5-dioxo-2H-1,2, 4-triazin-6-yl)propanoylhydrazinylidene]methyl]benzoate	ZINC01748535	-6.385	-8.819
11		[5-acetyloxy-2-[(4R)-4,6,6-trimethyl-2-sulfanylidene-1, 3-diazinan-4-yl]phenyl] acetate	ZINC00704941	-6.207	-8.814
12		2-amino-4-(3,4-difluorophenyl)-5-oxo-4,6-dihydropyrano[3, 2-c]quinoline-3-carbonitrile	ZINC00753445	-6.067	-8.752
13			ZINC00595805	-6.531	-8.739
14		methyl 2-[3-chloro-4-(3-chloroanilino)-2,5-dioxopyrrol-1-yl]benzoate	ZINC01938089	-6.501	-8.730

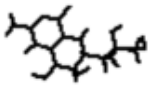
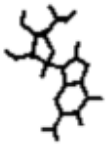
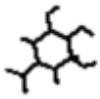
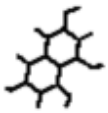
COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

15			ZINC02144749	-7.551	-8.727
16		N-(4-chlorophenyl)-2,8-dimethyl-4-quinolinamine	ZINC00112329	-6.200	-8.727
17		ethyl 4-[(3-chlorophenyl)carbamoyl]-1H-imidazole-5-carboxylate	ZINC00608429	-6.672	-8.695
18		(1-methylindol-3-yl)-(4,5,6,7-tetrahydro-3H-benzimidazol-5-yl)methanone	ZINC03873275	-6.921	-8.666
19		2-amino-N-(4-bromophenyl)-4,5,6,7-tetrahydro-1-benzothiophene-3-carboxamide	ZINC00666142	-6.114	-8.651
20		(6S,7S)-7-(3-hydroxyphenyl)-2-(3-hydroxypropyl)-5-methylidene-6,7-dihydro-4H-[1,2,4]triazolo[1,5-a]pyrimidine-6-carboxamide	ZINC01795351	-6.645	-8.642
21		5-bromo-N-[4-(furan-2-carbonylamino)-3-methylphenyl]furan-2-carboxamide	ZINC00884397	-6.048	-8.61
22		8-[(4-fluorophenyl)methylsulfanyl]-3,3-dimethyl-6,7,8,8a-tetrahydro-2H-pyrrolo[1,2-a]pyrazine-1,4-dione	ZINC03839604	-6.634	-8.59
23		N-[3,5-bis(trifluoromethyl)phenyl]-5-chloro-2-hydroxybenzamide	ZINC03986227	-6.214	-8.582
24		3H-benzimidazol-1-ium-5-yl(2,3-dihydroindol-1-yl)methanone	ZINC03332308	-6.742	-8.579
25		2-(2-bromo-4-formylphenoxy)-N-(2-methylphenyl)acetamide	ZINC00611426	-6.144	-8.577

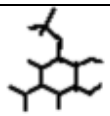
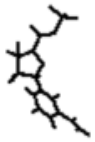
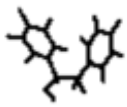
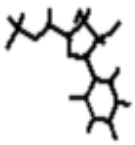
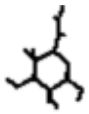
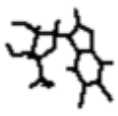
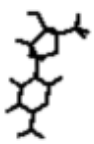
COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

Table 5.

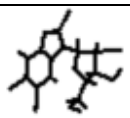
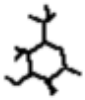
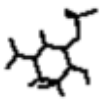
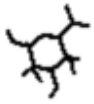
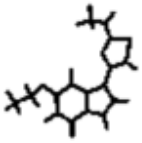
Database ID number, chemical structure, IUPAC names and Glide scores of top scored ligands docked with murE of Pseudomonas syringae using docking program Glide.

S. No.	Ligand structure	Ligand Name	ZINC ID	Glide (HTVS) score	Glide (XP) score
1		1,3,4-trihydroxy-5-oxocyclohexane-1-carboxylic acid	ZINC03869824	-6.266	-7.504
2		2-amino-7,8-dihydroxy-6-(1,2,3-trihydroxypropyl)-1,5,6, 7-tetrahydropteridin-4-one	ZINC03869200	-6.258	-7.377
3		1,3,4,5-tetrahydroxycyclohexanecarboxylic acid	ZINC03870192	-6.664	-7.304
4		2-amino-9-[(2S,3S,4R,5S)-3, 4-dihydroxy-5-(hydroxymethyl) oxolan-2-yl]-3H-purin-6-one	ZINC01532637	-6.134	-7.292
5		1,3,4-trihydroxy-5-oxocyclohexane-1-carboxylic acid	ZINC03869823	-6.968	-7.269
6		(2S,3S,4S,5R,6R)-3,4,5-trihydroxy-6-phosphonooxyoxane-2-carboxylic acid	ZINC04095578	-6.001	-7.227
7		3,4,5-trihydroxycyclohexene-1-carboxylic acid	ZINC03870234	-6.242	-6.995
8		naphthalene-1,3,6,8-tetrol	ZINC00901903	-6.471	-6.94

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

9		4,5-dihydroxy-3-phosphonooxycyclohexene-1-carboxylic acid	ZINC03870238	-6.355	-6.922
10		[5-(3-carbamoylpyridin-1-ium-1-yl)-3,4-dihydroxyoxolan-2-yl]methyl dihydrogen phosphate	ZINC03870107	-6.055	-6.902
11		1,2-bis(4-fluorophenyl)ethane-1,2-diol	ZINC03852703	-6.419	-6.779
12		[5-(2,4-dioxo-1H-pyrimidin-5-yl)-3,4-dihydroxyoxolan-2-yl]methyl dihydrogen phosphate	ZINC03870178	-6.471	-6.753
13		3,4,5,6-tetrahydroxyoxane-2-carboxamide	ZINC03872700	-6.499	-6.740
14		3,4,5,6-tetrahydroxyoxane-2-carboxamide	ZINC03872698	-6.375	-6.666
15		(2R,3R,4S,5R)-2-(5,6-dichlorobenzimidazol-1-yl)-5-(hydroxymethyl)oxolane-3,4-diol	ZINC03873418	-6.478	-6.623
16		4-amino-1-[(2R,3S,4S,5S)-5-(chloromethyl)-3,4-dihydroxyoxolan-2-yl]pyrimidin-2-one	ZINC03800994	-6.391	-6.553

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

17		(2R,3S,4R,5R)-2-(hydroxymethyl)-5-(2,5,6-trichlorobenzimidazol-1-yl)oxolane-3,4-diol	ZINC03797914	-6.0632	-6.48
18		3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-one	ZINC03869937	-6.01	-6.464
19		3,4-dihydroxy-5-oxocyclohexene-1-carboxylic acid	ZINC03869253	-6.273	-6.463
20		4,5-dihydroxy-3-phosphonooxycyclohexene-1-carboxylic acid	ZINC03870239	-6.446	-6.456
21		3,4,5,6-tetrahydroxyoxane-2-carboxylic acid	ZINC03869799	-6.588	-6.429
22		4-(5-ethoxy-1H-indol-3-yl)-N-methyl-1,3-thiazol-2-amine	ZINC02330112	-6.184	-6.418

Since computational screenings always demand experimental testing in order to confirm the accurate drug molecule(s), the proposed LEAD molecules need to be optimized in further studies. The significance of this work is in providing a relatively inexpensive approach to screen compounds that are likely to inhibit the action of murD and murE in *Pseudomonas syringae*.

CONCLUSION

The data presented here demonstrate that stepwise prioritization of genome open reading frames using simple biological criteria can be an effective way of rapidly reducing the number of genes of interest to an experimentally manageable number. This process is an efficient way for enriching potential target genes, and for identifying those that are critical for normal cell function. The generation of a comprehensive essential genes list will allow an accelerated genetic dissection of traits such as metabolic flexibility and inherent drug resistance that render *P. syringae* such a tenacious pathogen. Such a strategy will enable us to locate critical pathways and steps in pathogenesis; to target these steps by designing new drugs; and to inhibit the infectious agent of interest with new antimicrobial agents. We propose probable chemical compounds, which could be tested to devise drug

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

molecules to retard the hazardous proliferation of *P. syringae.* The scope of this work could be to use this data to do cost-effective experimental screening. The proposed potential chemical compounds could provide the prime lead for future experimental screening.

REFERENCES

1. F. Humer. Innovation in the Pharmaceutical Industry—Future Prospects, Available: http://www.roche.com/fbh_zvg05_e.pdf (2005).
2. G.C. Terstappen, and A. Reggiani. In silico research in drug discovery, Trends Pharmacol Sci. 22: 23–26 (2001).
3. C. Freiberg. Novel computational methods in anti-microbial target identification, Drug Discovery Today 6: S72–S80 (2001).
4. P. Sanseau. Impact of human genome sequencing for in silico target discovery, Drug Discov Today 6: 316–323 (2001).
5. S. Wang, T.B. Sim, Y.S. Kim, and Y.T. Chang. Tools for target identification and validation, Curr Opin Chem Biol 8: 371–377 (2004).
6. C. Freiberg, B. Wieland, F. Spaltmann, K. Ehlert, and H. Brotz. Identification of novel essential Escherichia coli genes conserved among pathogenic bacteria,J. Mol. Microbiol. Biotechnol 3: 483–489 (2001).
7. T. Roemer, B. Jiang, J. Davison, T. Ketela, and K. Veillette. Large-scale essential gene identification in Candida albicans and applications to antifungal drug discovery, Mol. Microbio. 50: 167–181 (2003).
8. C.M. Sassetti, D.H. Boyd, and E.J. Rubin. Genes required for mycobacterial growth defined by high density mutagenesis, Mol. Microbiol 48: 77–84 (2003).
9. J. Wixon, and D. Kell. The Kyoto encyclopedia of genes and genomes-KEGG, Yeast 17:48–55 (2000).
10. A.L. Hopkins and C.R. Groom. The druggable genome, Nat. Rev. Drug Discov 1: 727–730 (2002).
11. J.G. Robertson. Mechanistic basis of enzyme-targeted drugs, Biochemistry 44:8918-8923 (2005).
12. L. Miesel, J.Greene, and T.A. Black. Genetic strategies for antibacterial drug discovery, Nat. Rev. Genet 4: 442-456 (2003).
13. D.L. Wheeler, D.M. Church, R. Edgar, S. Federhen, W. Helmberg, T.L. Madden, J.U.Pontius, D.G. Schuler, L.M. Schriml, E. Sequeira, T.O. Suzek, T.A. Tatusova, and L. Wagner. Database resources of the National Center for Biotechnology Information: update, Nucleic Acids Res. 32: D35-D40 (2004).
14. K.R. Sakharkar, M.K. Sakharkar, and V.T.K. Chow. A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*, In Silico Bio. 4: 0028 (2004).
15. K.R. Sakharkar, M.K.Sakharkar, and V.T.K. Chow. Gene fusion in *Helicobacter pylori*: making the ends meet, Antonie Van Leeuwenhoek 89: 169-180 (2006).
16. M. Itaya, An estimation of minimal genome size required for life, FEBS Lett 362: 257-260 (1995).
17. E.V. Koonin, R.L. Tatusov, and M.Y. Galperin. Beyond complete genomes: from sequence to structure and function, Curr. Opin. Struct. Biol. 8: 355-363 (1998).

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

18. Kobayashi. Essential *Bacillus subtilis* genes, Proc. Natl. Acad. Sci USA 100: 4678-4683 (2003).
19. M.Y. Galperin, and E.V. Koonin. 'Conserved hypothetical' proteins: prioritization of targets for experimental study, Nucleic Acids Res 32: 5452-5463 (2004).
20. M. Huynen, Y. Diaz-Lazcoz, and P. Bork. Differential genome display, Trends Genet 13: 389-390 (1997).
21. M. Huynen, T. Dandekar, and P. Bork. Differential genome analysis applied to the species-specific features of *Helicobacter pylori*, FEBS Lett 426: 1-5 (1998).
22. R.E. Bruccoleri, T.J. Dougherty, and D.B. Davison. Concordance analysis of microbial genomes. Nucleic Acids Res, 26: 4482-4486 (1998).
23. A.R. Mushegian, and E.V. Koonin. A minimal gene set for cellular life derived by comparison of complete bacterial genomes, Proc. Natl. Acad. Sci. USA 93: 10268-10273 (1996).
24. R. Zhang, H.Y. Ou, and C.T. Zhang. DEG: a database of essential genes, Nucleic Acids Res 32: D271-D272 (2004).
25. G.A. Beattie and S.E. Lindow. Comparison of the behavior of epiphytic fitness mutants of Pseudomonas syringae under controlled and field conditions, Appl. Environ. Microbiol 60:3799-3808 (1994).
26. M. Wilson, S.S. Hirano, and S.E. Lindow. Location and survival of leaf-associated bacteria in relation to pathogenicity and potential for growth within the leaf, Appl. Environ. Microbiol 65: 1435-1443 (1999).
27. D. Gurian-Sherman and S.E. Lindow. Bacterial ice nucleation: significance and molecular basis. FASEB Journal 9: 1338-1343 (1993).
28. S.F. Altschul, L.M. Thomas, A.S. Alejandro, Z. Jinghui, Z. Zheng, M. Webb, and J.L. David. Gapped BLAST and PSI BLAST: a new generation of protein database search programs, Nucleic Acids Res 25 17: 3389–3402 (1997).
29. M. Kanehisa, S. Goto, S. Kawashima, and A. Nakaya. The KEGG databases at Genome Net. Nucleic Acids Res. 30 (1): 42–46 (2002).
30. J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res 24: 4876–4882 (1997).
31. R.A. Laskowski, M.W. MacArthur, D.S. Moss, and Thornton. J.M. *PROCHECK*: a program to check the stereochemical quality of protein structures, J Appl Crystallogr 26: 283-291 (1993).
32. G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations, J. Mol. Biol. 7: 95- 99 (1963).
33. D. Eisenberg, R. Luthy, and J.U. Bowie. VERIFY3D: assessment of protein models with three-dimensional profiles, Methods Enzymol 277: 396-404 (1997).
34. A. Dutta, S.K. Singh, P. Ghosh, R. Mukherjee, S. Mitter, and D. Bandyopadhyay. In silico identification of potential therapeutic targets in the human pathogen *Helicobacter pylori*. In Silico Biol 6: 0005 (2006).
35. J.A. Bertrand, A.E.F. Geneviève, M. Lydie, B. Didier, Jean van Heijenoort, and D. Otto.

COMPUTATIONAL PREDICTION OF POTENT THERAPEUTIC TARGETS OF *PSEUDOMONAS SYRINGAE* AND IN SILICO VIRTUAL SCREENING FOR NOVEL INHIBITORS

Crystal structure of UD{-N-acetylmuramoyl-L-alanine : D glutamate ligase from E. coli, The EMBO Journal 16: 3416–3425 (1997).

36. T. Den Blaauwen, M.A. de Pedro, M. Nguyen-Distèche, and J.A. Ayala. Morphogenesis of rod-shaped sacculi, FEMS Microbiol. Rev. 32: 321–344 (2008).
37. S. Dramsi, S.Magnet, S. Davison, and M. Arthur. Covalent attachment of proteins to peptidoglycan. FEMS Microbiol. Re 32: 307–320 (2008).
38. W. Vollmer. Structural variation in the glycan strands of bacterial peptidoglycan, FEMS Microbiol. Rev 32: 287–306 (2008).
39. E.A. Zoeiby, F. Sanschagrin, and R.C. Levesque. Structure and function of the Mur enzymes: development of novel inhibitors, Mol. Microbio. 47: 1-12 (2003).
40. J.A. Thanassi, S.L. Hartman-Neumann, T.J. Dougherty, B.A. Dougherty, and M.J. Pucci. Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumonia*, *Nucleic Acid Res* 30: 3152-3162 (2002).