SUMMARY AND CONCLUSIONS

Sir Alexander Ogston introduced the term “Staphylococcus” (from the Greek staphyle) to describe grape-like clusters of bacteria responsible for inflammation and suppuration (Ogston et al., 1882). *Staphylococci* are Gram-positive cocci (0.5 to 1.5 μM in diameter), non-motile, non-spore-forming, usually catalase positive and oxidase negative (Kloos et al., 1991, 1994; Entenza et al., 2013). *Staphylococci* are tolerant to high concentrations of sodium chloride (Wilkinson et al., 1997) and show resistance to high temperature. The genus *Staphylococcus* contains 32 species, 16 of which are found in humans. Only a few of them are pathogenic in the absence of predisposing host conditions such as immune suppression or the presence of a foreign body. In humans and animals, the most virulent one is *Staphylococcus aureus* (Waldvogel et al., 1995; Projan et al., 1997).

*Staphylococcus aureus* infection

*Staphylococcus aureus* has long been recognized as one of the most important human pathogens. It colonizes and infects both hospitalized and healthy people in the community (Waldvogel, 2000; Palavecino et al., 2014). It is responsible for a wide range of infections ranging from skin and soft tissue infections to bacteraemia, infections of the central nervous system, bone and joints, skeletal muscles, respiratory and urinary tracts, infections associated with intravascular devices and foreign bodies. Other life threatening infections caused by *S. aureus* include endocarditis, pneumonia, scalded skin syndrome, or bone and joint infections. Serious infections typically require hospitalization and treatment with intravenous antibiotics (Faria et al., 2009).

Multidrug resistance - driving the need for new drugs

*Staphylococcal* resistance against five groups of life saving antibiotics in the years 2003-2005 indicates urgent need of new chemical class antibiotics for therapy (Fatima et al., 2013). The rise in the levels of antibacterial drug resistance in human pathogens is most common phenomenon (Park et al., 2013). Resistance is defined as bacteria that are not inhabited by usually achievable systematic concentration of an agent with normal dosage schedule and /or fall in the minimum inhibitory concentration ranges. Drug resistance is of major concern for severely ill and hospitalized patients as therapeutic
efficacy of current drugs in practice is declining. First clear proof of resistance to penicillin was reported by an accidental observation in 1958 (Ley et al., 1958). Microorganisms developing resistance towards an antibacterial substance is an inherent mechanism (Piotr et al., 2013; Park et al., 2013). Considering the present need for discovery and development of novel antibiotics we are already too late (Bradley et al., 2013; Fatima et al., 2013).

**Target selection criteria**

The Aminoacyl-tRNA synthetases (AaRS) are essential proteins found in all living organisms. The enzymes catalyze the attachment of amino acids to transfer RNAs (tRNA) and thereby establish the rules of the genetic code by virtue of matching the nucleotide triplet of the tRNA anticodon with its cognate amino acid. They have emerged as leading targets for the development of new antibiotics (Zhou et al., 2013). Among AaRS Phenylalanine-tRNA synthetase (PheRS) was selected to develop the inhibitors because (1) it is structurally unique among the AaRS (it is a class II AaRS but is significantly different from other class II AaRS); (2) it is considerably different from human cytosolic and human mitochondrial AaRS; (3) it is essential and conserved across bacterial species; (4) previously described antibacterial PheRS inhibitors (Beyer et al., 2004; Yu et al., 2004; Jarvest et al., 2005) have demonstrated antibacterial activity in vitro and in vivo.

Phenylalanine-tRNA synthetase is a functional heterodimer, that assembles into an obligate (ab) 2 heterotetramer in all cases studied to date (Sanni et al., 1988; Yao et al., 2013) with the exception of mitochondrial PheRS, which is a monomer. The small (a) subunit of the enzyme carries the conserved class-II catalytic module where acyl transfer takes place. Our understanding of the structure, function and modulation of cytoplasmic PheRS has been greatly enhanced by a series of studies conducted using the PheRS from *Thermus thermophilus* (TTH) (Fishman et al., 2001).

Recent studies describe the discovery of PheRS small-molecule inhibitors with submicromolar potency. Unlike the previously mentioned compounds, the latter inhibitors were discovered by means of high-throughput screening and therefore are not conceptually based on prior knowledge of PheRS biochemistry.
Little literature is available with respect to overcoming antibiotic resistance in *Staphylococcus aureus* PheRS. Hence the present study focused on the following objectives.

**OBJECTIVES**

The present study focused on:

1). To study the Ligand based pharmacophore (LBPP) of PheRS inhibitors in *S. aureus*

2). To develop the 3D-QSAR lead molecule of PheRS inhibitors in *S. aureus*

3). To develop the Homology model and perform Molecular dynamics and simulation study.

4). To generate the Structure based pharmacophore model (SBP)

and

5) To study the virtual screening and antibacterial activity of novel molecules.

**Ligand based Pharmacophore modeling of Phenylalanyl-tRNA synthetase inhibitors:**

We have generated a 3D-QSAR pharmacophore model using Catalyst software (CATALYST, 2005) for diverse set of Phenylalanyl-tRNA synthetase (PheRS) inhibitors with an aim to obtain pharmacophore model that could provide a rational hypothetical picture of the primary chemical features responsible for activity. As crystal structure of *Staphylococcus aureus* PheRS is not available, ligand based drug discovery is useful to identify novel lead compounds with an improved PheRS inhibitory activity.

**Biological data:** Extensive literature collection was done with respect to PheRS protein information and ligands binding to the protein. A data set of 76 molecules having activities against PheRS was identified in the literature (Xiang et al., 2004; Richard et al., 2005). Dataset perfectly meets the requirements to use for pharmacophore generation. Twenty four compounds forming the training set was used to generate pharmacophore, where biological activity data span over 4 orders of magnitude. To validate the pharmacophore hypothesis, 52 compounds with available IC$_{50}$ values were used as a test set. The
compounds were built using Catalyst 2D–3D sketcher (CATALYST, 2005) and a family of representative conformations was generated for each compound using the best conformational analysis method with Poling algorithm (Smellie et al., 1995) and CHARMM force field parameters (Brooks et al., 1983).

The process of pharmacophore development has four stages: data preparation, pharmacophore generation, data analysis and model validation.

With the input of a full range of training set compounds from inactive to active, the HypoGen algorithm generates hypotheses with features common amongst active molecules and missing from the inactive molecules. This is accomplished in three steps, a constructive step, a subtractive step and an optimization step. Finally in the optimization step, the resultant hypotheses are then optimized using simulated annealing to further fine tune the model parameters, thereby improving model quality.

The first hypothesis (Hypo1) is the best pharmacophore hypothesis in this study, characterized by the highest cost difference 63.348, lowest root mean square error 0.920 and the best correlation coefficient 0.944. The fixed cost, pharmacophore cost and null cost are 100.679, 90.8878 and 174.417 respectively. The difference between null cost and fixed cost is large indicating that hypo1 has 75–90% probability of correlating the data. Configuration cost value must be around 17 and accordingly 18.82 was obtained. Hypo 1 is presented in Figure 1 consists of spatial arrangement of four chemical features: one hydrogen bond acceptor (HBA), one hydrophobic point (HY) and two ring aromatic (RA) features.

Among the 10 hypothesis, first hypothesis (Hypo1) is exhibiting the better statistics than the other and hence we considered Hypo1 for further analysis. The activities were estimated for all compounds based on the best ranking pharmacophore (Hypo1). Among 24 training set compounds, two moderately active compounds (++) were predicted as active (+), one inactive (+) compound was predicted as active as moderately active (++)

The discrepancy between the actual and estimated activity observed for the three compounds was only about 1 order of magnitude, which might be an artifact of the program that uses different number of degrees of freedom for these compounds to mismatch the pharmacophore model.
Interestingly, in the training set, all highly active compounds map all the features that are hydrophobic point (HY), hydrogen bond acceptor (HBA) and hydrogen bond donor (HBD), with a few exceptions in moderately active and inactive compounds one feature are missing. All the compounds in the training set map HBA and HY feature revealing that these two features should be mainly responsible for the activity. The most active compound 48 has a fitness score of 8.6 when mapped to Hypo1 whereas the least active 58 maps to a value of 4.77. Fit value indicates how well the features in the pharmacophore overlaps the chemical features in the molecule and thereby aid in understanding the chemical meaning of the hypothesis.

In compound 48, HBD feature corresponds to Nitrogen present in middle of thiophene and phenyl parts of structure, one HY feature corresponds to halogen group on thiophene ring. Two HBA features correspond to hydroxyl group which attached to alkyl chain and Oxygen group present on phenyl ring, whereas in compound 58, HBD feature does not map. For molecules with lesser activity (70, 77, 69 and 71), at least one feature is missing. Pharmacophore superimposed with twelve potent compounds in the training set is shown in figure 1 revealing that hydrophobic point, hydrogen bond donor and one hydrogen bond acceptor mainly responsible for the activity.

Figure 1: Pharmacophore superimposed with twelve potent compounds in the training set using Hypo1. Green: Hydrogen bond acceptor (HBA), Orange: Ring Aromatic (RA), Cyan: Hydrophobic (HY) and Pink: Hydrogen bond donor (HBD).
Ligand based computational approach was employed to identify molecular structural features required for an effective PheRS inhibitor, with an aim to discover drugs to prevent and cure *Staphylococcus aureus* bacterial infections. A highly predictive pharmacophore model was generated based on 24 training set compounds. The pharmacophore model generated can be used as: (1) Three-dimensional query in database searches for new potent molecules in a virtual screening system. (2) To design new molecules and forecasting their inhibitory activity for PheRS quantitatively before undertaking any synthesis.

### 3D-QSAR Studies

Since, its introduction in 1988, comparative molecular field analysis (CoMFA) (Cramer et al., 1988) has emerged as one of the most powerful tools in ligand based drug design strategies (Desiraju, 2002; Feng et al., 2013). The CoMFA methodology assumes that a suitable sampling of steric and electrostatic fields around a set of aligned molecules provides all the information necessary for understanding their biological properties (Bohm et al., 1999; Sanjeev Kumar Singh, 2006). In the present study, we apply two methodologies, comparative molecular field analysis (CoMFA) (Cramer et al., 1998) and comparative molecular similarity indices analysis (CoMSIA) (Klebe et al., 1998) based on a set of novel PheRS inhibitors to build the 3D-QSAR models, which can predict the biological activity values of new compounds prior to their synthesis.

The statistical data obtained from the standard CoMFA and CoMSIA model constructed are depicted in table 1.
Table 1: PLS statistics of CoMFA and CoMSIA

<table>
<thead>
<tr>
<th>Parameters</th>
<th>CoMFA</th>
<th>CoMSIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOO-(r^2)</td>
<td>0.828</td>
<td>0.754</td>
</tr>
<tr>
<td>No validation ((r^2))</td>
<td>0.986</td>
<td>0.991</td>
</tr>
<tr>
<td>F value</td>
<td>374.015</td>
<td>465.930</td>
</tr>
<tr>
<td>Standard error estimate</td>
<td>0.101</td>
<td>0.083</td>
</tr>
<tr>
<td>Steric contribution</td>
<td>39.6%</td>
<td>48%</td>
</tr>
<tr>
<td>Electrostatic contribution</td>
<td>60.4%</td>
<td>39.8%</td>
</tr>
<tr>
<td>Cross validation ((r^2))</td>
<td>0.80</td>
<td>0.70</td>
</tr>
</tbody>
</table>

To understand key requirements of PheRS inhibitors CoMFA and CoMSIA studies were carried out. The CoMFA model yielded satisfactory statistical data with the cross-validated \(q^2\) value of 0.88 and the non-cross-validated \(r^2\) value of 0.986. The cross-validated \(q^2\) value for CoMSIA model was 0.75 and the non-cross-validated \(r^2\) value was 0.99. The high leave-one-out (LOO) cross-validated correlation coefficient \(q^2\) reveals that the model is a useful tool for the prediction of 10 test set compounds that were not included in the training set of 38 compounds. The results not only lead to better understanding of structural requirements of PheRS inhibitors but also can help in the design of new potent inhibitors.

The \(q^2\) values of the CoMSIA model compared to the CoMFA model indicating stable analyses of high quality. Further the CoMSIA model comprise of valuable complementary information, as they offer comparative and additional explanation for ligands by introducing three auxiliary field types, the hydrophobic field, H-bond donor and acceptor fields. The CoMSIA statistical details are given in Table 1.
Interpretation of Contour Maps

**CoMFA:** The green and yellow colored contours of steric feature mapped on to the most active compound 25 were shown in the figure 2. The figure depicts the favorable (green contour) and non-favorable (yellow) regions around the 3 dimensional spaces of the compound 38. A large green contour mapped near to ortho-position of phenyl ring signifies the importance of bulky group substitution at this position. Further the substituted phenyl ring of second best compound 43 also overlapping well on to this contour.

![Figure 2: Electrostatic and steric fields CoMFA contour maps with best active compounds](image)

(a) Best active compound 25 (b) Compound 43. The green contours indicate sterically favored regions whereas the yellow contours denote sterically unfavorable regions. The blue contours identify regions that favor electropositive substituent and the red regions favor electronegative substituent.

The compounds 8, 17 and 18 having bulky substitution (OCH2Ph) are mapping well to this contour and showing higher activity than the compound 1 substituted with methyl group. On comparisons with compounds 26 and 46, the compound 46 is having larger group ((CH2)4NH2) than 26 ((CH2)2NH2) and showing one log order increase in activity. The large yellow contour appeared in the CoMFA, suggests the presence of bulky group decreases the activity. The substituted phenyl groups of the compounds 9-12 and 16 mapped well to the yellow contour and were showing lowest activity.
CoMSIA:

The steric and hydrophobic contours obtained from CoMSIA are similar to the contours found in CoMFA. These contours are unambiguous in distinguishing the important hydrophobic features necessary for higher activity. The electrostatic contour map is more significant in CoMSIA than CoMFA. A large blue colored electro-positive contour appeared near the ortho position of the phenyl ring of the compound 38 establishes the importance of electro-positive substitution at that position and more over two more electro-positive contours which appeared on the other side of the compound is essential for activity, specially for the compounds having substituted phenyl groups at R1 position of ethanolamine derivatives.

In this study, by using the alignment scheme generated on most active molecule, a highly predictive CoMFA and CoMSIA models were developed and were used to predict the pIC$_{50}$ activity of a set of Phenylalanyl tRNA synthetase inhibitors. The QSAR models gave good statistical results in terms of q$_2$ and r$_2$ values and have been validated using a test set, obtained from the hierarchical clustering. The CoMFA region focusing model provided the most significant correlation of steric and electrostatic fields with the biological activities. Totally, the CoMFA models provided better statistical results than CoMSIA, which shows the significance of steric and electrostatic fields in the selectivity and activity of these compounds. From these analyses, it is possible to predict the ligand activities of newly designed PheRS inhibitors and design better anti-bacterial inhibitors.

Homology modeling studies

A better understanding of a protein function can be learned from its 3D structure. The function of a protein is highly correlated with its structure. In this regard, a major effort is underway to develop strategies of high throughput structure determination of proteins. In the lack of an experimentally established crystal structure of a given protein, homology modeling is the best alternative to construct a reasonable three-dimensional (3D) model of the target. At present, homology modeling is the most accurate technique for 3D structure prediction of proteins (Bordoli et al., 2009).

To achieve the aim of the study, The 3D model of *S. aureus* PheRS α subunit is developed by homology modeling using the X-ray structures of *E. coli* PheRS and
Summary

*Thermus thermophilus* PheRS (PDB: 3PC0 and 1PYS) as templates. A recursive approach comprising sequence alignment and model building was used to build the multiple models, followed by the refinement of non-conserved regions. With an aim to investigate the stability of the generated model, the homology model was refined in the presence of lipid bilayer molecular dynamics simulation.

Percent similarity and identity of the PheRS with different proteins have been analyzed. Alignment of *Staphylococcus aureus* PheRS, *E. coli* PheRS (3PC0) and *Thermus thermophilus* PheRS (1PYS) reveals that both templates have more than 40% similarity and belong to same class of proteins i.e., Amino acyl tRNA synthetase. Hence homology models for *S. aureus* PheRS can be generated using these structures as templates. The final multiple sequence alignment was submitted to MODELLER 9.2 (Sali et al., 1993, 1995) for generating the homology model of the PheRS. The whole model was subjected to a stepwise energy refinement in Macromodel using the OPLS-2005 force field.

Known active compounds were sketched with their amide bonds in trans (R-Chirality) conformation. SiteMap program from Schrodinger was used to identify the potential binding pockets (Halgren et al., 2007) in the model. Docking study was carried out using Glide software to predict bioactive conformation and binding modes of active compounds.

**MD simulation of receptor-membrane complex**

The compound 47 protein complex was used for further study. The complex was inserted in a pre-equilibrated DPPC/TIP3P membrane system using the system builder tool within Desmond software (Bowers et al., 2006) as implemented in Maestro graphical interface (Schrodinger Inc., Portland, OR, USA). The longest axis of the Model as well as the membrane normal was aligned with Z-axis. Lowest minimum energy complex with stable statistics was selected to understand the binding mode of the active ligand.

**Ligand binding mode analysis**

Docking of best active antagonists was carried out to assess the validity of our generated model and examine its ability to guide the structure-based drug design. In the
proposed binding pose, the ligand NH of amine forms an ionic and hydrogen bond with Glu216 and forms key interaction with Met314. The ligand phenyl moiety occupies a relatively wide hydrophobic pocket (P1) formed by Val261, Phe254, Thr257, Phe312, Gly217 and Gly290. In addition to the essential interactions between PheRS antagonists and Met314 ligand has the cation–π interactions with Arg318. As a rule of thumb, we can conclude that binding to PheRS requires ionic interaction with Met314, Glu216. Hydrophobic complimentary with Phe254 and Cation – π interaction with Arg218 may improve the activity.

To our knowledge there are no reports on binding mode of the ligands to PheRS, to gain this knowledge, the initial PheRS homology model was subjected to a strategic ligand supported refinement protocol using known best active compound. This model was also validated using Procheck software and Ramachandran plot. The final model, in complex with active compound, was embedded in a membrane aqueous environment and then simulated for 20ns in an NPT ensemble. This protocol shows the stability of the generated model. The knowledge gained from the developed 3D model of PheRS can be useful in the identification and structure based design of novel antagonists that can represent promising anti Staphylococcus agents.

Structure Based Pharmacophore generation and Virtual screening

The ‘Receptor ligand pharmacophore generation’ module of DS 3.1 program package was used to generate pharmacophore models. Resulted Pharmacophore composed of five features viz., two hydrogen bond donors, two hydrophobic and one positive ionizable feature. Resulted structure based pharmacophore (SBP) model has good statistical validation with good ROC curve accuracy 0.88. SBP has good selectivity power to differentiate actives from inactive. Top ten active molecules have a good alignment with SBP (Figure 3). This pharmacophore can be further used to screen the library for new class of PheRS inhibitors.
Understanding the activity difference

It is very important to understand activity difference of the known molecules before Virtual screening. Molecules 3 and 9 have only polar interactions with Glu216. Inactive molecules do not show any interaction with Met314 backbone, which is key interaction for binding to PheRS. Molecules do not show crucial interactions with Arg318 and Val261. These interactions play a key role in gaining the activity.

Virtual screening

The optimal structure based pharmacophore model PharA was used to screen several chemical databases including the commercial databases like Asinex, NCI and SPECS databases. This consensus screening resulted in potential PheRS inhibitors with novel scaffolds. The number of maximum omitted features was set to n-1. Virtual screening against 680000 molecules resulted in 12000 molecules as hits.

Drug-likeness analysis

Drug-likeness properties are usually used for selecting compounds for invitro studies. Here, Lipinski’s rule of five was slightly modified according to the properties of
existing inhibitors. After drug-likeness screening 12000 molecules reduce to 2500 compounds.

**ADMET analysis**

ADMET (absorption, distribution, metabolism, elimination, toxicity) analysis is an important step in drug designing. Some properties including human intestinal absorption, aqueous solubility levels, BBB penetration levels, CYP2D6 inhibition and hepato toxicity of these 2500 compounds were analyzed using TopCat software. Molecules within the acceptable range of ADMET value were selected out as hits. Only 380 molecules could able to pass ADMET filters. These results indicate that although there have been some high active inhibitors of PheRS, their ADMET properties are not satisfactory. It may be troublesome to use them for further experiments in vivo.

**In-vitro enzymatic and antibacterial screening of novel molecules**

Total 380 hits passed ADMET filter were analyzed for the difference in the chemical space. Approximately 30 diverse molecules were identified in the hits. Final hits were tested for antibacterial activity against *Staphylococcus aureus*.

Biology testing details are as below.

*Test organism: Staphylococcus aureus NCIM 5345 (Wild)*

*Assay medium: Antibiotic assay medium (Hi-medium).*

*Culture medium: Nutrient agar medium.*

*Protocol: Standard antibiotic assay protocol from Indian pharmacopeia.*

Incubation time: 24 hrs after adding the molecule.

Standard compound: Ciprofloxacin.

**Biological assay results for compound 4299493:**

Most scored virtual screen hit 4299493 was tested for inhibitory activity against *Staphylococcus aureus*. MIC results show that tested compound has better clear zones of inhibition than the reference compound see figure 4. This gives the confidence on the ligand and structure based methods.
A = Standard compound  
B= Sample compound

<table>
<thead>
<tr>
<th>Structure</th>
<th>Name</th>
<th>SBPP_FitValue</th>
<th>Glide XP score</th>
<th>MappedFeatures</th>
<th>kDiverse</th>
<th>Predicted IC50 nM</th>
<th>Fit value</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image.png" alt="Chemical Structure" /></td>
<td>4299493</td>
<td>3.82</td>
<td>-12.09</td>
<td>HBD2, HYP85, HYP6, Pi7</td>
<td>Yes</td>
<td>10.4</td>
<td>6.53641</td>
</tr>
</tbody>
</table>

Figure 4: Virtual screening hit compound #4299493 inhibition zones against *Staphylococcus aureus* in petri dish MIC assay. Both standard compound and sample are tested at concentrations 5, 10, 20 and 30 µg/ml.

Total 15 compounds were selected from the top-ranking compounds sorted by consensus scoring and pharmacophore fit values. Among them, the compound # 1040757 attracted the attention due to that it possesses a novel scaffold and has the smallest molecular weight. Few top predicted virtual screening hit compounds 1040757, 20406024 and 4324899 has better inhibition zone than standard compound.
CONCLUSIONS

- In this study, 3D-QSAR pharmacophore was built to understand the common features of PheRS inhibitors. Hypo1, 3D-Pharmacophore model was developed based on 24 currently available PheRS inhibitors. Best model consists of four features namely two HBA, one HY and one HBD, with test set prediction correlation of 0.731.

- CoMFA and CoMSIA models were generated to understand the SAR changes in the known ligands. The CoMFA model provided the most significant correlation of steric and electrostatic fields with the biological activities. CoMSIA, explained the significance of steric electrostatic and hydrophobic fields in the selectivity and activity of these compounds.

- To understand the binding mode of the actives crystal structure of PheRS is not available; homology model was developed based on its homologues protein 3PCO. Quality of the homology model was assessed using Ramachandran plot analysis. ERRAT plot was generated to know the overall confidence in the modeled structure.

- A Molecular Dynamics (2ns) study was carried out to understand the stability of the generated homology model; this also gives the information on binding mode of the active ligand and movement of flexible amino acid side chains. Based on MD trajectory analysis and clustering, final refined homology model was picked to perform docking studies.

- Docking studies on top 10 active and 10 inactive molecules suggests the interaction with MET314 backbone is crucial for binding to PheRS. Polar interactions with GLU216 and Val 261. plays a critical role in the gaining the activity.

- PharA, the structure-based pharmacophore was developed using docked complex of the high active 47. PharA composed of Pharmacophore composed of five features: Two Hydrogen bond donors, two hydrophobic and one Positive ionizable feature. Pharmacophore validated using ROC curve and enrichment studies for predicting capability.

- Virtual screening was carried out using validated structure-based pharmacophore PharA, against 680,000 molecules database for new potential scaffolds against PheRS protein. Selected compounds from database screening based on fitness score. This resulted in 12000 VS hits.
• Drug-likeness and ADMET analysis: The selected 12000 compounds were further analyzed and refined using drug-like filters and ADMET analysis. At last, 380 hits were selected out.

• Molecular docking procedure was carried out for these hits to understand the interaction between these hits and the target protein.

• Invitro biological studies were carried to understand the activities of 15 novel scaffold molecules. Results show that, five VS hits shows inhibition activity against *Staphylococcus aureus*.

• Thus a humble attempt was made to study the molecular modeling strategies in overcoming antibiotic resistance in *Staphylococcus aureus*. However much work has to be done to derive precise conclusions.