Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2012, 4(5):1776-1784
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X
CODEN (USA): PCHHAX

Pharmacophore mapping and 3D-QSAR analysis of *Staphylococcus aureus* Sortase a inhibitors

Heena Mehta, Sukhbir Lal Khokra*, Kanika Arora, Pawan Kaushik

Institute of Pharmaceutical Sciences, Kurukshetra University, Kurukshetra, India.

ABSTRACT

Staphylococcus aureus causes a variety of human infections, ranging from superficial abscesses to life threatening bacteremias. Staphylococcus aureus Sortase A inhibitors are widely used for the treatment of bacterial infections. 3D-QSAR analysis has been applied to a structurally diverse set of 34 compounds as Staphylococcus aureus Sortase A inhibitors, which are of special interest because of their role in bacterial infections. The present study has been focused on pharmacophore mapping study that can explore 3D features and configurations responsible for biologically activity of structurally diverse compounds. A four point pharmacophore (ADRR) with one hydrogen bond acceptor (A), one hydrogen bond donor (D) and two aromatic rings (R) as pharmacophore features was developed. The generated best pharmacophore hypothesis yielded a statistically significant 3D-QSAR model, with a correlation coefficient of R^2 = 0.929 for training set molecules. The model generated showed excellent prediction power, with Q^2 = 0.887 for an external set of 10 test set molecules. The geometry and features of pharmacophore are expected to be useful for the design of selective Staphylococcus aureus Sortase A inhibitors.

Keywords: Pharmacophore mapping, 3D-QSAR, *Staphlococcus aureus* Sortase A.

INTRODUCTION

Gram positive pathogenic bacteria display proteins on their surface that play important roles during infection. In Staphylococcus aureus, these surface proteins are anchored to the cell wall by two sortase enzymes, Sortase A (srtA) and Sortase B (srtB), they recognize specific surface protein sorting signals [1]. *Staphylococcus aureus* Sortase A is an enzyme naturally located in the cell wall of gram positive bacteria, where it catalyzes transpeptidation reactions without the need of ATP [2]. Srt A is a polypeptide of 206 amino acids [3], responsible for anchoring proteins containing a C-terminal tripartite sorting signal, which consists of a) an LPXTG pentapeptide (where X represents amino acid) followed by, b) a less well conserved hydrophobic domain and c) a basic charged tail. Both the hydrophobic domain and the charged tail help to retain the putative surface protein in the membrane prior to sortase-catalysed anchoring [4]. Srt A, a transpeptidase with an active site cysteine, cleaves surface proteins between the threonine (T) and the glycine (G) of the LPXTG motif [3] and catalyzes the formation of peptide bond between the carboxyl group of the threonine and the amine group of the cell wall precursor lipid II. The lipid II linked protein is then incorporated into the peptidoglycan of the cell wall via the transglycosylation and transpeptidation reactions of bacterial cell wall synthesis. Sortases represent an attractive target for new anti-infective agents, since they are widely distributed among a variety of bacterial pathogens [5].

QSAR studies are mathematical methodologies, statistically validated and mostly used to correlate experimental or calculated properties derived from chemical structures with biological activities. With the advent of 3D molecular space modeling, a pharmacophore hypothesis can visualize the potential interaction between the ligand and the receptor [6]. 3D pharmacophore model is a ligand-based approach that provides a unique tool for drug design. A 3D pharmacophore is a collection of chemical features in space that are required for a desired biological activity. These may include hydrophobic groups, charged / ionisable groups, hydrogen bond donors / acceptors and other features properly assembled in 3D space to reflect structural requirements [7]. The objective of the present study is to develop ligand based pharmacophore hypothesis and to derive 3D-QSAR model with Pharmacophore Alignment and Scoring Engine (PHASE) [8]. PHASE is a highly flexible system for common pharmacophore identification and assessment, 3D-QSAR model development, and 3D database creation and searching.

MATERIALS AND METHODS

Data sets

A set of 34 novel *Staphylococus aureus* Sortase A inhibitors (Tables 1 and 2) with available IC_{50} data were taken from literature for the development of ligand-based pharmacophore hypothesis and 3D-QSAR model [9, 10]. The IC_{50} values, i.e., the concentration (µM) of inhibitor that produces 50% inhibition to *Staphylococcus aureus* Sortase A, were converted to molar concentrations and then into pIC_{50} values as reported in Tables 1 and 2. In order to obtain a validated and predictive QSAR model, an available data set should be divided into the training and test sets. For the prediction statistics to be reliable, the test set must include at least five compounds [11]. The data set was divided into a training set of 24 molecules and a test set of 10 molecules. The training set molecules were selected in such a way that they contained information in terms of both their structural features and biological activity ranges. The most active molecules, moderately active and less active molecules were included, to spread out the range of activities [12]. In order to assess the predictive power of the model, a set of 10 compounds was arbitrarly set aside as the test set. The test compounds were selected in such a way that they truly represent the training set.

TRAINING SET & TEST SET BASIC STRUCTURE

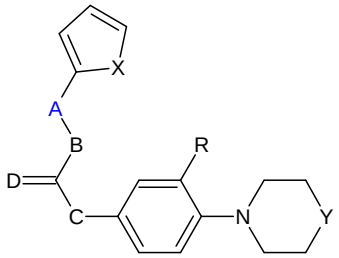


Table 1: Experimental inhibitory activity of training set compounds

S.No.	X	A-B	C	D	Y	R	IC_{50} (µM)	pIC_{50} observed	pIC_{50} predicted
1.	S	CH=CH	NH	O	O	COOH	75	4.125	4.00
2.	O	CH=CH	NH	O	O	$COOCH_3$	58	4.237	4.04
3.	O	CH=CH	NH	O	O	COOH	181	3.742	3.77
4.	O	CH_2-CH_2	NH	O	O	COOH	600	3.222	3.24
5.	O	CH_2-CH_2	NH	O	O	$COOCH_3$	600	3.222	3.42
6.	S	C≡C	NH	O	O	$COOCH_3$	165	3.783	3.83
7.	S	C≡C	NH	O	O	COOH	183	3.738	3.75
8.	S	CH=CH	NH	O	O	$COOCH_3$	61	4.215	4.22
9.	S	CH=CH	NH	O	O	COOH	154	3.812	3.82
10.	O	CH=CH	NMe	O	O	$COOCH_3$	514	3.289	3.28
11.	S	CH=CH	NMe	O	O	COOH	600	3.222	3.29
12.	O	CH=CH	NMe	O	O	COOH	600	3.222	3.21
13.	S	CH=CH	NH	HH	O	$COOCH_3$	249	3.604	3.56
14.	S	CH=CH	NH	HH	O	COOH	600	3.222	3.24
15.	O	CH=CH	NH	O	CH_2	COOH	463	3.334	3.62
16.	O	CH=CH	NH	O	O	CH_2OH	111	3.955	3.97
17.	S	CH=CH	NH	O	O	CHO	77	4.114	4.14
18.	O	CH=CH	NH	O	O	CHO	107	3.971	3.97

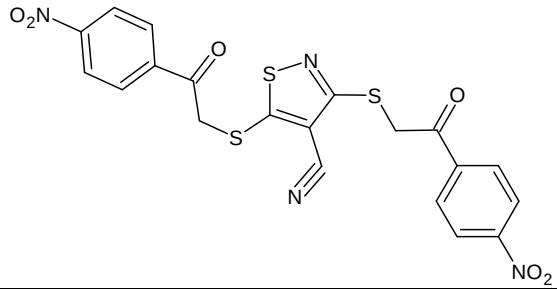
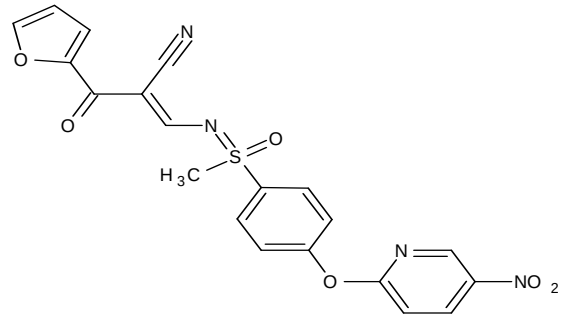
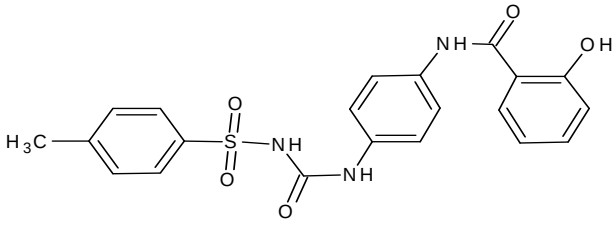
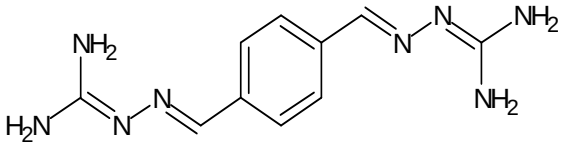
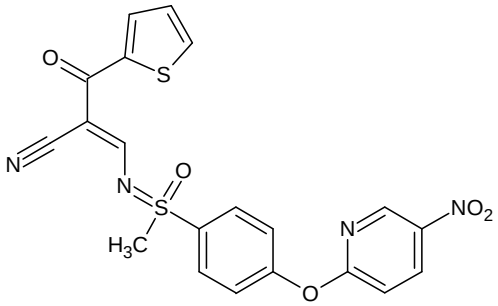
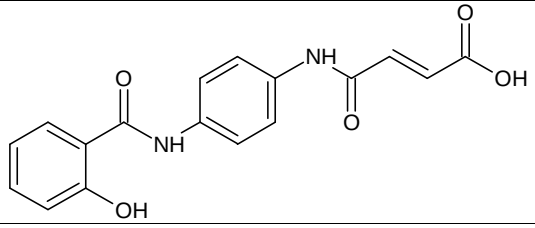
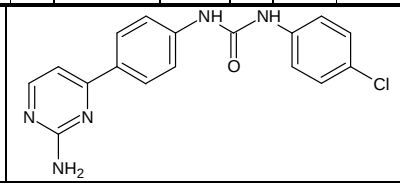
S.No	Structure	IC_{50} (µm)	pIC_{50} observed	pIC_{50} predicted
19.		97.3	4.012	4.04
20.		150	3.824	3.84
21.		226	3.646	3.67
22.		275	3.561	____
23.		400	3.398	3.35
24.		125	3.903	3.79

Table 2: Experimental inhibitory activity of test set compounds

S.No	X	A-B	C	D	Y	R	IC_{50} (µm)	pIC_{50} observed	pIC_{50} predicted
1.	S	CH=CH	NH	O	O	$COOCH_3$	71	4.149	4.11
2.	S	CH_2-CH_2	NH	O	O	COOH	600	3.222	3.63
3.	S	CH_2-CH_2	NH	O	O	$COOCH_3$	600	3.222	3.53
4.	S	CH=CH	NMe	O	O	$COOCH_3$	571	3.243	3.28
5.	S	CH=CH	NH	O	CH_2	$COOCH_3$	92	4.036	4.01
6.	O	CH=CH	NH	O	CH_2	$COOCH_3$	131	3.883	3.86
7.	S	CH=CH	NH	O	CH_2	COOH	181	3.742	3.79
8.	S	CH=CH	NH	O	O	CH_2OH	73	4.137	4.17
9.	S	CH=CH	NH	O	O	$CONH_2$	105	3.979	3.98
10.							400	3.398	3.59

Pharmacophore modeling

Pharmacophore modeling and 3D database searching are now recognized as integral components of lead discovery and lead optimization. Pharmacophore modeling was carried out using PHASE: a module of Schrödinger's software program 'MAESTRO' [13].

Generation of Common Pharmacophore Hypothesis

The chemical structures of all the compounds were drawn in maestro and geometrically refined using LigPrep module. Tautomers were generated using MacroModel method discarding current conformers. The conformations were generated by the Monte Carlo (MCMM) method as implemented in MacroModel version 9.6 using a maximum of 2,000 steps with a distance-dependent dielectric solvent model and an OPLS-2005 force field. All the conformers were subsequently minimized using truncated Newton conjugate gradient (TNCG) minimization up to 500 iterations. For each molecule, a set of conformers with a maximum energy difference of 30kcal/mol relative to the global energy minimum conformer was retained. The conformational searches were done for aqueous solution using the generalized born/solvent accessible surface (GB/SA) continuum solvation model [14].

The next step in developing a pharmacophore model is to use a set of pharmacophore features to create pharmacophore sites (site points) for all the ligands.

Common pharmacophoric features were then identified from a set of variants-a set of feature types that define a possible pharmacophore. Common pharmacophores are identified using a tree-based partitioning technique that groups together similar pharmacophores according to their inter site distances, i.e., the distances between pairs of sites in the pharmacophore.

In the next step, common pharmacophore hypothesis were examined using a scoring function to yield the best alignment of the active ligands using an overall maximum root mean square deviation (RMSD) value of 1.2 Å for distance tolerance. The quality of alignment was measured by a survival score, defined as:

$$S = W_{site}S_{site} + W_{vec}S_{vec} + W_{vol}S_{vol} + W_{sel}S_{sel} + W^{m}_{rew}$$

Where W_{Ds} are weights and S_{Ds} are scores, and S_{site} represents an alignment score, the RMSD in the site point position. S_{vec} represents the vector score, and averages the cosine of the angles formed by corresponding pairs of vector features in aligned structures. S_{vol} represents the volume score based on the overlap of the Vander Walls models of non-hydrogen atoms in each pair of structures. S_{sel} represents the selectivity score, and accounts for the fractions of molecules that are likely to match the hypothesis regardless of their activity towards a receptor. Weights are user adjustable. Wsite, Wvec, Wvol, and W_{rew} have a default value of 1.0 while W_{sel} has a default value of 0.0, so that a useful hypothesis is not missed. W^{m}_{rew} represents the reward weights, where *m* is the number of actives that match the hypothesis minus one. In the hypothesis generation, all default values were used [15].

Validation of QSAR model and Assessment of significant common pharmacophore hypothesis (CPH) using partial least square (PLS) analysis

Validation is a crucial aspect of pharmacophore design, particularly when the model is built for the purpose of predicting activities of compounds in external test series [16, 17]. External validation is considered to be a conclusive proof for judging predictability of a model. Our priority was to develop QSAR models that were statistically robust both internally as well as externally. The main target of any QSAR modelling is that the developed model should be robust enough to be capable of making accurate and reliable predictions of biological activities of new compounds. In the present case, the validation of generated common pharmacophore hypothesis was performed by correlating the observed and estimated activity for 24 molecules of the training set and 10 molecules of the test set. PLS analysis were carried out using PHASE with ligands in the training set using a grid spacing of 1 Å. A common pharmacophore hypothesis with the best predictivity and significant statistics was selected for molecular alignments and a QSAR study of the test set was carried out.

RESULTS

Pharmacophore generation and 3D-QSAR model

A total of 19 different variant hypothesis were generated upon completion of common pharmacophore identification process. A maximum of four features were allowed to develop hypothesis and a number of CPHs were reported common to all molecules based on hydrogen bond acceptor (A), hydrogen bond donor (D) and aromatic ring (R). There were 9 hypothesis based on the combination ADRR, 3 hypothesis based on AADR and 7 hypothesis based on AARR. We have selected those pharmacophore models whose survival scores ranked in the top. The top model was found to be associated with the four point hypothesis (Figure 1), which consist of one hydrogen bond acceptor (A), one hydrogen bond donor (D) and two aromatic rings (R). This is denoted as A_2D_3 R_7R_8. The pharmacophore hypothesis showing distance between pharmacophoric sites is depicted in Figure 1. Predicted and observed biological activity of training and test sets are shown in Table 1 and 2.

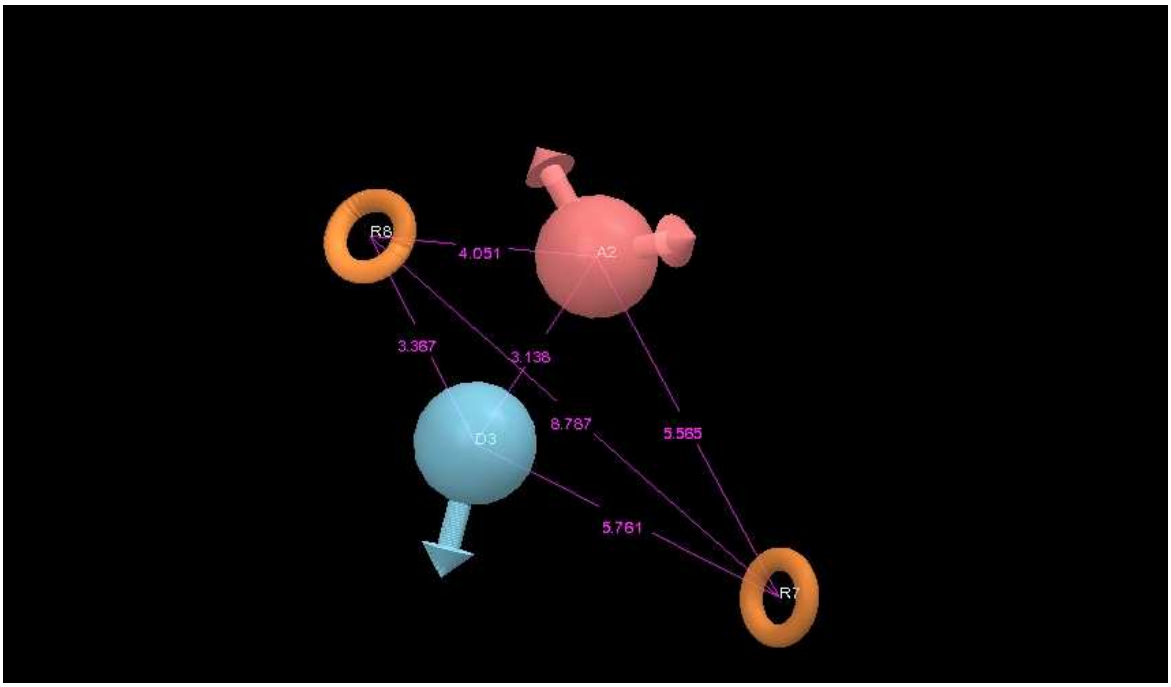


Figure 1: Pharmacophore hypothesis and distance between pharmacophoric sites. All distances are in Å unit.

The pharmacophore hypothesis yielded a 3D-QSAR model with good PLS statistics. The training set correlation is characterized by PLS factors (r^2 = 0.929, SD = 0.12, F = 123, P = 6.307e – 20). The test set correlation is characterized by PLS factors (Q^2_{ext} = 0.887, RMSE = 0.18, Pearson-*R* = 0.91). Results of PLS statistics of 3D-QSAR model is shown in Table 3. Graph of observed versus predicted biological activity of training and test sets are shown in Figures 2 and 3, respectively.

Table 3: Parameters of best four featured pharmacophore hypothesis

Training set	Test set
r^2 = 0.929	Q^2_{ext} = 0.887
SD = 0.12	RMSE = 0.18
F = 123, *P* = 6.307 - 20	Pearson-*R* = 0.91

SD = standard deviation of the regression, r^2 = correlation coefficient
P = significance level of variance ratio, F = variance ratio
Q^2_{ext} = for the predicted activities, RMSE = root-mean-square error
Pearson-R = correlation between the predicted and observed activity for the test set

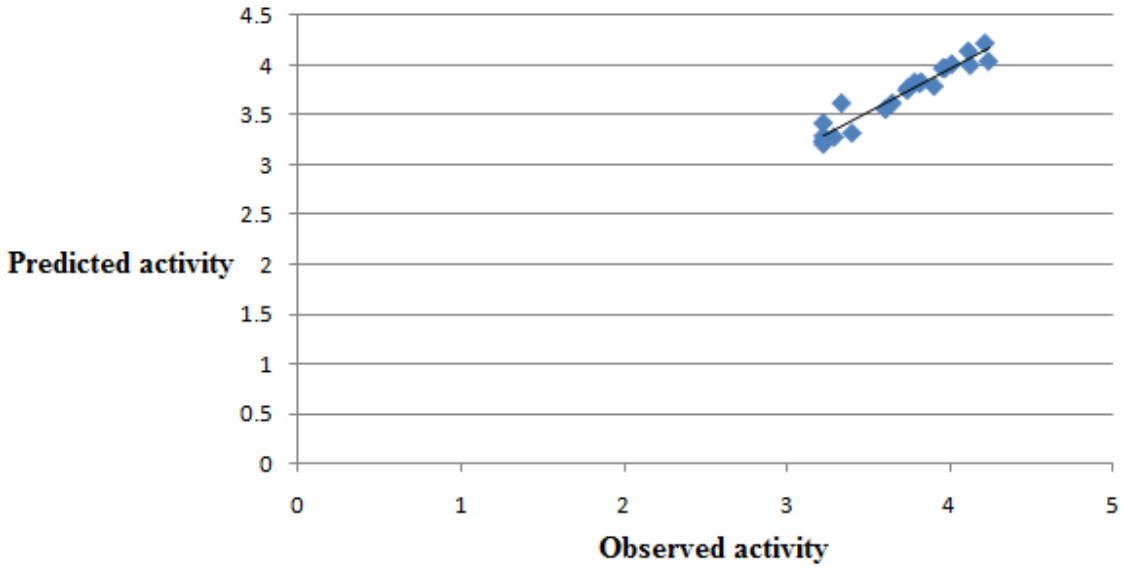


Figure 2: Graph of observed versus predicted biological activity of training set molecules

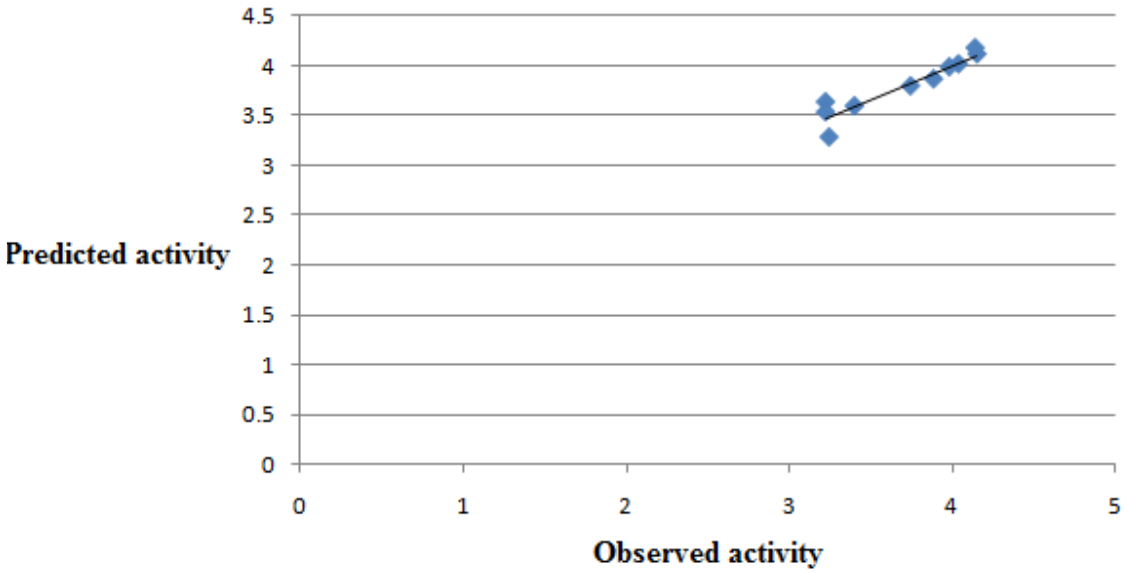


Figure 3: Graph of observed versus predicted biological activity of test set molecules

DISCUSSION

The purpose of pharmacophore modeling is to perform *in silico* screening searches in a 3 dimensional database of a virtual or real compound library to find diverse structures with desired binding activity and selectivity. In the present study, a series of *S. aureus* Sortase A inhibitors were considered for molecular modeling studies. The present study aimed to develop ligand based pharmacophore hypothesis and 3D-QSAR model which can provide information regarding structural modifications with which to design analogs with better activity prior to synthesis.

We have previously shown the results of PLS analysis applied to generated pharmacophore hypothesis. For a reliable model, the squared predictive correlation coefficient should be > 0.6 [12]. The results of this study reveal that model $A_2D_3R_7R_8$ can be used for the prediction of srt A inhibitory activity.
Additional insights into the inhibitory activity can be gained by visualizing the 3D-QSAR model.

3D-QSAR model of compound 9 (5-(3-(yhiophen-2-yl)acrylamido)-2-morpholino benzoic acid) of the training set illustrating the hydrogen bond acceptor feature is shown in Figure 4. The red cubes around the oxygen of OH group of benzoic acid and at position 2 of morpholine ring favours the srt A inhibitory activity and substitutions at these positions by groups having more hydrogen bonding acceptor property favours the srt A inhibitory activity. Green region around the S of thiophene ring, near the donor (D_3) and at position 5 and 6 of morpholine ring do not favour the srt A inhibitory activity. Substitutions at these positions by electron withdrawing groups such as NO_2, Cl, F, etc. will result in increase in srt A inhibitory activity.

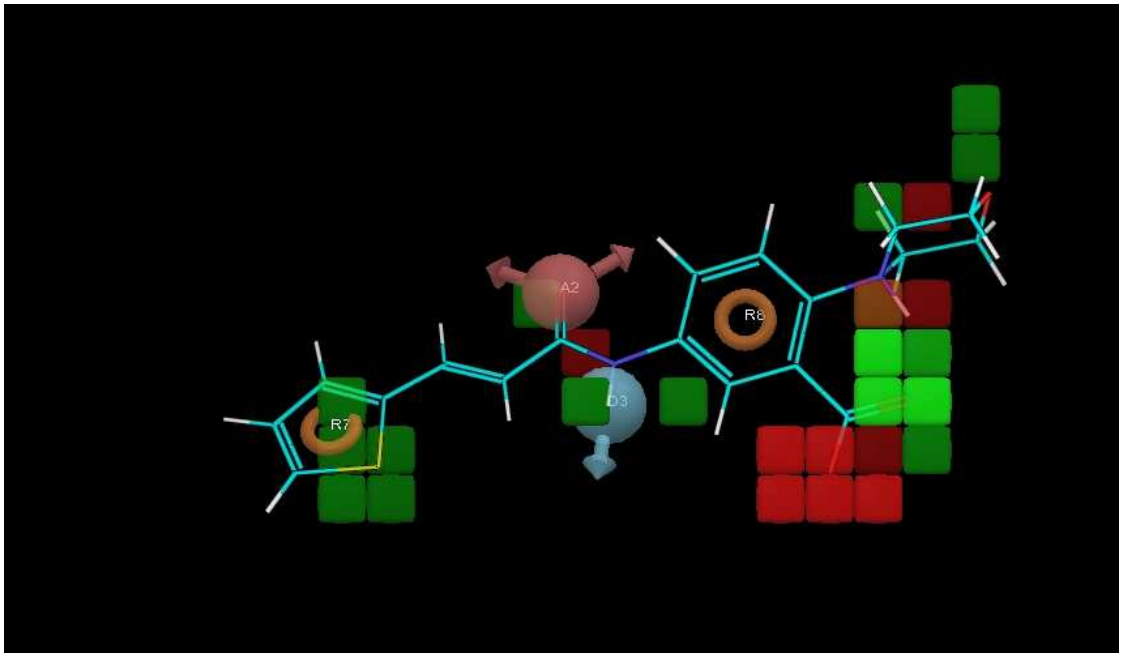


Figure 4: Pictorial representation of the cubes generated using the 3D-QSAR model based on molecule 9 (most active) of training set illustrating hydrogen bond acceptor prediction. *Red cubes* indicate favourable regions, while *green cubes* indicate unfavourable region for the activity.

The 3D-QSAR model based on molecule 9 of the training set using hydrophobicity feature is shown in Figure 5. Red region around oxygen of OH group and C=O group of benzoic acid, at position 3 of morpholine ring and around S of thiophene ring favours srt A inhibitory activity and substitutions at these positions by more hydrophobic groups will result in increase in srt A inhibitory activity. Green region at position 1, 3 and 4 of thiophene ring and position 2 of benzoic acid do not favour the srt A inhibitory activity and substitution at these position with hydrophobic group increase the srt A inhibitory activity.

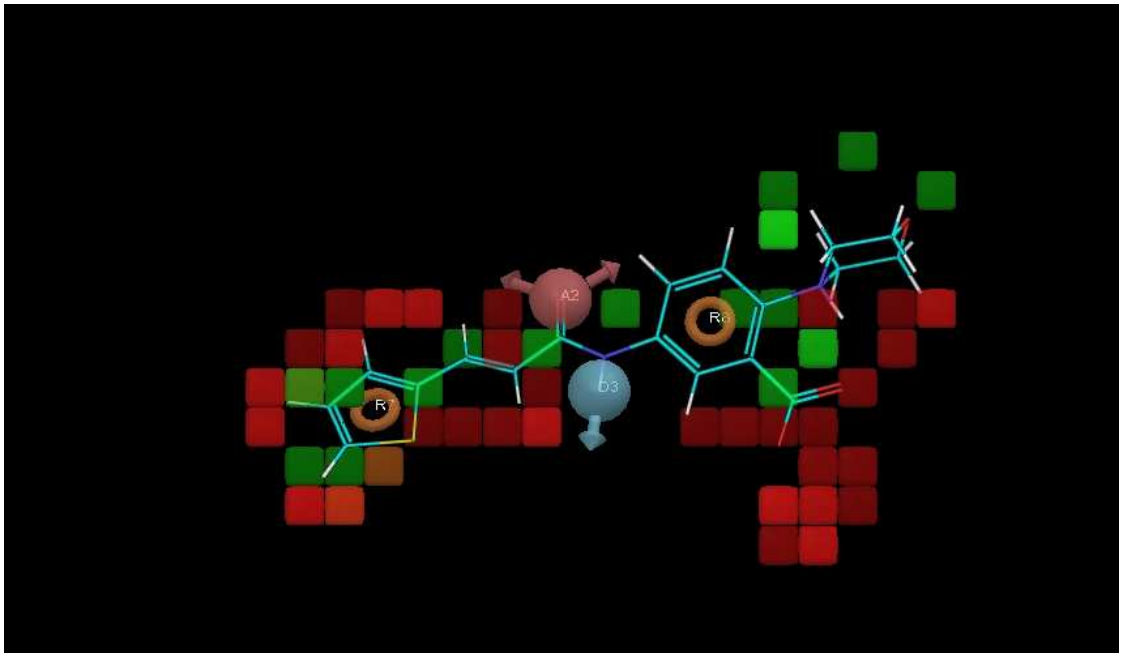


Figure 5: 3D-QSAR model based on molecule 9 (most active) of training set illustrating hydrophobic feature

The 3D-QSAR model based on molecule 9 of the training set using hydrogen bond donor feature is shown in Figure 6. Red region around oxygen of OH group of benzoic acid and at the H of –CONH group (near D_3 region) favours srt A inhibitory activity and substitutions at these positions by more hydrogen bond donor groups will result in increase in srt A inhibitory activity. Green region around the N of morpholine ring do not favour the srt A inhibitory activity and substitution at these position with electron donating groups increase the srt A inhibitory activity.

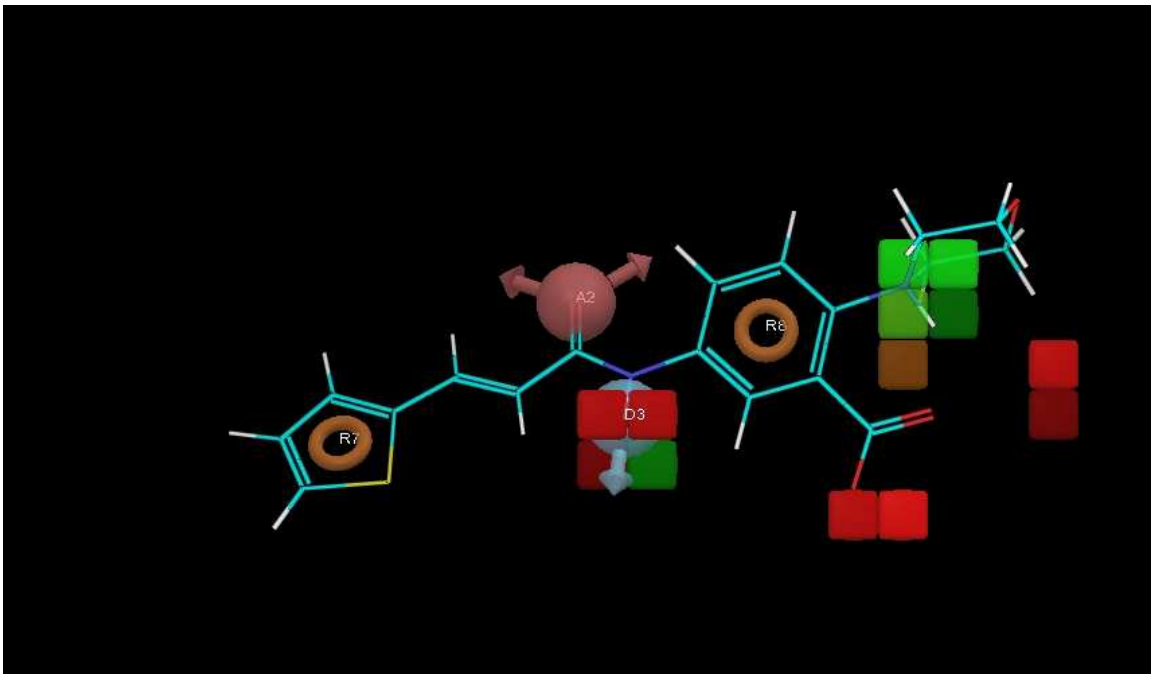


Figure 6: 3D-QSAR model based on molecule 9 (most active of highest fitness score 3) of training set illustrating hydrogen bond donor feature

CONCLUSION

In conclusion, a highly predictive pharmacophore based 3D-QSAR model was generated using a training set of 24 molecules which consists of four point pharmacophore hypothesis with one hydrogen bond acceptor (A), one hydrogen bond donor (D) and two aromatic rings (R). The developed 3D-QSAR model can provide insights into the structural requirement of novel *Staphylococcus aureus* Sortase A inhibitors.

The present study aimed to develop ligand based pharmacophore hypothesis and 3D-QSAR model which give detailed structural insights as well as highlights important binding features of novel *Staphylococcus aureus* Sortase A inhibitors, which can provide guidance for the rational design of novel srt A inhibitors.

Acknowledgement
This work is supported in part by grants to Institute of Pharmaceutical Sciences, Kurukshetra University, Kurukshetra.

REFERENCES

[1] IM Jonsson; SK Mazmanian; O Schneewind; T Bremell; A Tarkowski. *Microbes and Infection*, **2003**, 5, 775-780.
[2] D Zhulenkovs; A Leonchiks. *Environmental and Experimental Biology*, **2010**, 8, 97-101.
[3] Y Zong; TW Bice; HT That; O Schneewind; SVL Narayana. *The Journal of Biological Chemistry*, **2004**, *279*, 31383-31389.
[4] ML Bentley; H Gaweska; JM Kielec; DG McCafferty. *The Journal of Biological Chemistry*, **2007**, *282*, 6571-6581.
[5] MT Naik; N Suree; U Ilangovan; CK Liew; W Thieu; DO Campbell; JJ Clemens; ME Jung; RT Clubb. *The Journal of Biological Chemistry*, **2006**, *281*, 1817-1826.
[6] S Nagar; MA Islam; S Das; A Mukherjee; A Saha. *Mol Divers*, **2008**, *12*, 65-76.
[7] PP Roy; K Roy. *Molecular Simulation*, **2010**, *36*, 887-905.
[8] Phase, version: 9.1, Schrodinger**,** LLC, New York, USA.
[9] BC Chenna; JR King; BA Shinkre; AL Glover; AL Lucius; SE Velu. *Europeon Journal of Medicinal Chemistry*, **2010**, *45*, 3752-3761.
[10] BC Chenna; BA Shinkre; JR King; AL Lucius; SVL Narayana; SE Velu. *Bioorganic & Medicinal Chemistry*, **2008**, *18*, 380-385.
[11] A Golbraikh; A Tropsha. *J. Comput.-Aided Mol. Design*, **2002**, *16*, 357-369.
[12] A Golbraikh; M Shen; Z Xiao; YD Xiao; KH Lee; A Tropsha. *Journal of Computer-Aided Molecular Design*, **2003**, *17*, 241-253.
[13] Maestro, version: 9.1, Schrodinger, LLC, New York, USA.
[14] G Chang; WC Guida; WC Still. *J Am Chem Soc*, **1989**, *111*, 4379-4386.
[15] UA Shah; HS Deokar; SS Kadam; VM Kulkarni. *Mol Divers*, **2010**, *14*, 559-568.
[16] DB Boyd. In **:** *Reviews in computational chemistry*, **1990**, *4*, 355–371.
[17] A Golbraikh; A Tropsha. *J. Mol. Graphics Mod.*, **2002**, *20*, 269-276.