Docking, CoMFA and CoMSIA Studies of 5-Amino-1-aryl-1*H*-tetrazole Derivatives as Cholera Vulgaris

RANJITHREDDY PALREDDY[1], JAHEER MOHMED[1],
SHRAVAN KUMAR GUNDA[2] and CH. SARALA DEVI[*1]

[1]Department of Chemistry, Osmania University, Hyderabad, Telangana State, India
[2]Division of Bioinformatics, Osmania University, Hyderabad, Telangana State, India
prof.saraladevi.ch@gmail.com

Abstract: The 3D-QSAR studies with 5-amino-1-aryl-1*H*-tetrazole derivatives were conducted using CoMFA and CoMSIA. Internal and external validation techniques were investigated using leave-one-out, no-validation, cross-validation and bootstrapping. The CoMFA model predicted satisfactory correlation coefficient q^2 value of 0.806 and conventional correlation coefficient r^2 value of 0.990, while CoMSIA model predicted q^2 value of 0.675 and r^2 value of 0.976, inferring the important role of steric and electrostatic properties of candidate compounds. The models were graphically interpreted using contour plots which gave more accuracy into the structural requirements for increasing the activity of a molecule and proved a solid basis for future rational design of more active inhibitors for cholera vulgaris. The resulting CoMFA and CoMSIA contour maps were used to classify the structural features relevant to the biological activity in selected series of 5-amino-1-aryl-1*H*-tetrazole derivatives and further aided in designing seven novel molecules that showed higher inhibitory activity against cholera vulgaris which shed new light on effective therapeutic agents against these classes of enzymes. Molecular docking studies were also carried out with 5-amino- aryl-1*H*- tetrazole derivatives in the pocket of "Cholera toxin *B*-pentamer with decavalent ligand BMSC-0013".

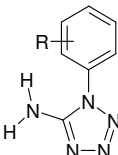**Keywords**: 5-Amino-1-aryl-1*H*-tetrazole derivatives, CoMFA, CoMSIA, Autodock

Introduction

Tetrazoles have a wide range of applications in medicinal chemistry. 5-Aminotetrazole and its derivatives possess a unique combination of properties, due to their reported anti-asthmatic, anti-allergic[1,2], anti-neoplastic[3], anti viral, anti inflammatory[4], anti cancer, anti bacterial, anti fungal, anti type 2 diabetes mellitus[5] and cognition disorder activities[6]. Tetrazole derivatives are potential compounds for the development of drugs for HIV or other immune diseases[7,8]. Tetrazoles found applications as explosives, rocket propellants[9] and as ligands in coordination chemistry[10-12]. Keeping in view the importance of title compounds, the 3D-QSAR technique is employed for drug designing. In general the 3D-QSAR techniques are valuable methods of ligand-based drug design by correlating physicochemical

properties from a set of related compounds to their known molecular property or molecular activity values. The present investigation reports the first application of 3D-QSAR to study 5-amino-1-aryl-1*H*-tetrazole as potent cholera vulgaris receptor agonists[15]. We studied twenty 5-amino-1-aryl-1*H*-tetrazole derivatives using CoMFA (comparative molecular field analysis)[16] and CoMSIA (comparative molecular similarity indices analysis)[17]. Models obtained from 3D-QSAR studies provide a strong basis for future rational design of more active and selective cholera vulgaris receptor agonists. Molecular docking studies were also performed with 5-amino-1-aryl-1*H*-tetrazole inhibitors by taking co-crystal structure from PDB (protein data bank).

Experimental

Twenty substituted compounds of 5-amino-1-aryl-1*H*-tetrazole were selected from the literature and their biological data is presented in Table 1. IC_{50} values of 5-amino-1-aryl-1*H*-tetrazole agonists were often converted to their negative logarithm (pIC_{50}) values. The pIC_{50} values of these compounds range from 2.08 to 4.01, providing a wide range and homogenous data set for 3D-QSAR study. The data set was validated by external test set by taking 3 compounds randomly and remaining 17 compounds were taken as the training set.

Table 1. Experimental, predicted and residual values of CoMFA, CoMSIA study

C.No	R	pIC_{50}	CoMFA		CoMSIA	
			Predicted	Residual	Predicted	Residual
1	H	2.65	2.64	0.01	2.89	-0.24
2	3-F	2.66	2.68	-0.02	3.01	-0.35
3	3-Cl	3.21	3.15	0.06	2.78	0.43
4	3-Me	3.09	2.97	0.12	2.66	0.43
5	**3-NO$_2$**	**2.79**	**3.17**	**-0.38**	**2.53**	**0.26**
6	3-OMe	2.92	2.97	-0.05	3.42	-0.50
7	4-F	2.49	2.62	-0.13	3.09	-0.60
8	4-Me	3.15	3.22	-0.07	3.06	0.09
9	4-NO$_2$	3.01	3.00	0.01	2.45	0.56
10	**4-OH**	**2.19**	**2.71**	**-0.52**	**2.45**	**-0.26**
11	4-OEt	3.31	3.29	0.02	3.45	-0.14
12	**3,4-Cl$_2$**	**3.84**	**3.29**	**0.55**	**3.71**	**0.13**
13	3-Cl-4-Me	3.76	3.45	0.31	3.65	0.11
14	3,4-(CH)$_4$	4.01	3.72	0.29	3.87	0.14
15	2-Cl	2.4	2.67	-0.27	2.87	-0.47
16	2-Me	2.46	2.29	0.17	2.74	-0.28
17	2-OMe	2.08	2.65	-0.57	2.51	-0.43
18	2,5-Cl$_2$	3.11	2.84	0.27	3.08	0.03
19	2,5-Me$_2$	2.98	3.10	-0.12	2.59	0.39
20	2,3-(CH)$_4$	3.07	3.09	-0.02	3.18	-0.11

Bold indicates test set C.No. 5,10 & 12

Structure building and   odeling   ion

All 3D structures of 5-amino-1-aryl-1*H*-tetrazole derivatives were sketched in SYBYL 6.7. Molecular   odeling package and energy minimization were performed on each 5-amino-1-aryl-1*H*-tetrazole derivative using Tripos force field with distance-dependent dielectric function and Powell conjugate gradient algorithm with a convergence criterion of 0.005 kcal/mol, all the molecules were minimized by adding Gasteiger-Huckel charges[18]. This resulted in 5-amino-1-aryl-1*H*-tetrazole molecules that are close to their native conformation and were used for subsequent studies.

Molecular alignment

Molecular alignment plays very important role in CoMFA and CoMSIA methods and their results are extremely sensitive to a number of factors like alignment rule and orientation rule of the aligned 20 compounds, probe atom type and lattice shifting step size[19]. The most active compound (compound 14) was used as the scaffold for aligning rest of the molecules by default SYBYL aligns database option. The structures of aligned molecules are shown in Figure 1.
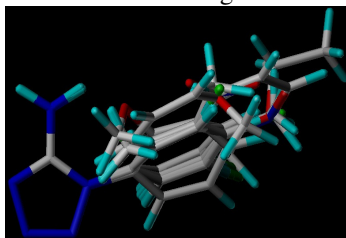

Figure 1. Molecular alignment of 5-amino-1-aryl-1*H*-tetrazole derivatives

3D-QSAR studies

For better comprehension of steric, electrostatic, hydrophobic, *H*-bond donor and *H*-bond acceptor field contributions for the set of molecules and to build predictive three dimensional quantitative structure activity relationship models, both CoMFA and CoMSIA studies were performed on the basis of molecular alignment as described earlier. CoMFA calculates steric and electrostatic properties whereas CoMSIA calculates hydrophobic, donor and acceptor properties along with steric, electrostatic properties. CoMFA and CoMSIA properties are calculated with respect to Lennard-Jones and Columbic potentials.

CoMFA studies

Steric and electrostatic properties were calculated using Tripos force field engine and the aligned molecules were placed in a 3D grid box so that the entire set was included. CoMFA descriptors were generated using sp^3 probe atom carrying +1 charge to generate steric and electrostatic fields. 30 kcal/mol cut-off was used for the analysis and standard options were used for the calculation of regression analysis[20]. Partial least square (PLS) was done by selecting leave-one-out (LOO) using 5 as the number of components and the column filtering was set as 2.0 kcal/mol.

CoMSIA studies

CoMSIA analysis was performed with the QSAR option in SYBYL. Five different properties were used in CoMSIA studies, which are steric, electrostatic, hydrophobic, donor and acceptor, based on which similarity indices between a probe atom and compound were calculated. The probe atom with 1Å radius, +1 charge and +1 for hydrophobicity was set at the intersections of the lattice.

Molecular docking studies

Molecular docking studies were performed using Autodock 4.2[21,22] to calculate possible binding modes, screening and validation for all 5-amino-1-aryl-1*H*-tetrazole derivatives. Default parameters were used in present study. All the molecules were docked into the binding site of a crystal structure of "Cholera toxin *B*-pentamer with decavalent ligand BMSC-0013" (PDB id: 1MD2) with 1.45Å resolution to provide the interaction between the receptor and the ligand. All the molecules were docked into the binding site using Autodock module, which uses genetic algorithm for generating protein-ligand interactions and molecules from the smaller fragments in the cavity of the receptor site. Grid was created, x, y, z (-13.201, 13.033, 7.125 for 1MD2) coordinates of Asn90 and Lys26 were selected.

Results and Discussion

CoMFA and CoMSIA

CoMFA and CoMSIA methods were applied to derive 3D-QSAR models for 5-amino-1-aryl-1H-tetrazole derivatives as cholera inhibitors. The statistical results of CoMFA and CoMSIA analysis are summarized in Table 2. Best predictions were obtained with CoMFA standard model involving $q^2 = 0.806$, $r^2 = 0.990$, SEE =0.057, $r^2cv = 0.810$ and F-value =227.939 with number of components as 5 and column filtering 2.0 kcal/mol. In CoMSIA, the standard model predictions obtained are $q^2 = 0.675$, $r^2 = 0.976$, SEE =0.102, $r^2cv = 0.716$ and F-value =71.766 with number of components as 6 and column filtering 1.0 kcal/mol selected for CoMSIA. CoMFA results show a higher predictive ability for 5-amino-1-aryl-1*H*-tetrazole derivatives against cholera vibrio on comparison with the CoMSIA results. Biological activities, predicted and residual values of both CoMFA and CoMSIA are shown in Table 1.

Table 2. Statistical Analysis of CoMFA and CoMSIA models

	CoMFA		CoMSIA	
q^2	0.806		0.675	
r^2	0.990		0.976	
SEE	0.057		0.102	
F-value	227.939		71.766	
CV	0.810		0.716	
Bootstrapping				
	Mean	Std.Dev	Mean	Std.Dev
SEE	0.039	0.028	0.088	0.087
r^2	0.995	0.003	0.982	0.018
Field contribution, *%*				
Steric	84.8		21.1	
Electrostatic	15.2		34.1	
Hydrophobic	-		34.2	
Donor	-		0.0	
Acceptor	-		10.6	

Table 2 shows the results of relative contributions for CoMFA and CoMSIA methods wherein 84.8% contribution was observed for steric and 15.0% field contribution was observed for electrostatic in CoMFA, while in CoMSIA, the observed contributions for steric, electrostatic, hydrophobic, donor and acceptor properties are 21.1%, 34.1%, 34.2%, 0.0% and 10.6% respectively. Electrostatic property is the main contributor in CoMSIA analysis. The affinities between experimental and calculated values of the training set and

the test set of CoMFA and CoMSIA models derived from non-cross-validated analysis are plotted in Figure 2(a) and 2(b). CoMSIA and the best CoMFA models were used to predict the inhibitory activities of the compounds in the test set.
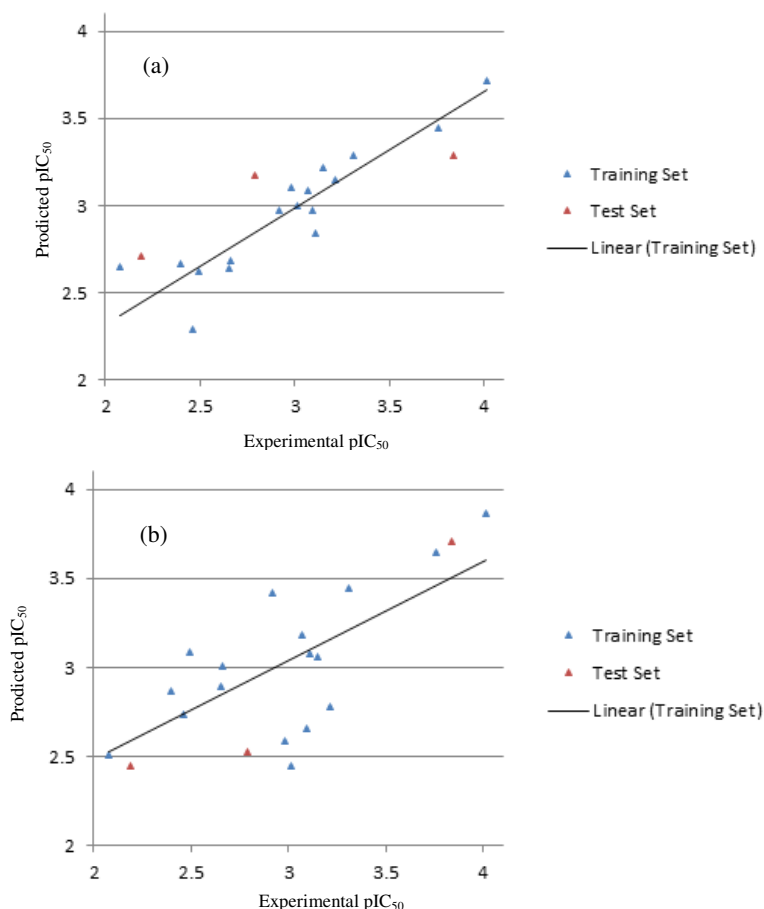

Figure 2. Experimental *versus* predicted activities of (a) CoMFA and (b) CoMSIA analysis

Contour analysis

The contour maps for CoMFA (steric and electrostatic) and CoMSIA (steric, electrostatic, donor, acceptor and hydrophobic) fields are based on PLS analysis, which are represented in 3D contour plots shown in Figures 3-5. The default parameters used in contour analysis by contribution are 80% favored region and 20% contribution for disfavored region.

Steric contour analysis

The steric field contour plots of CoMFA and CoMSIA for the most active compound are depicted in Figure 3(a) and (b), respectively. CoMFA and CoMSIA steric maps are shown in green and yellow regions, respectively. Green contour maps represent sterically favored regions where more bulky groups are expected to increase biological activity, whereas yellow contour maps represent less sterically favored regions where less bulky substituents are predicted to increase biological activity.

For the most active compound 14 the CoMFA steric maps showed small green contours and large yellow contour maps are present around phenyl group and in CoMSIA one medium green and large yellow contour maps are present around phenyl group. This indicates at yellow region less bulky groups are required to increase the activity while at green region more bulky groups are required to increase the biological activity.
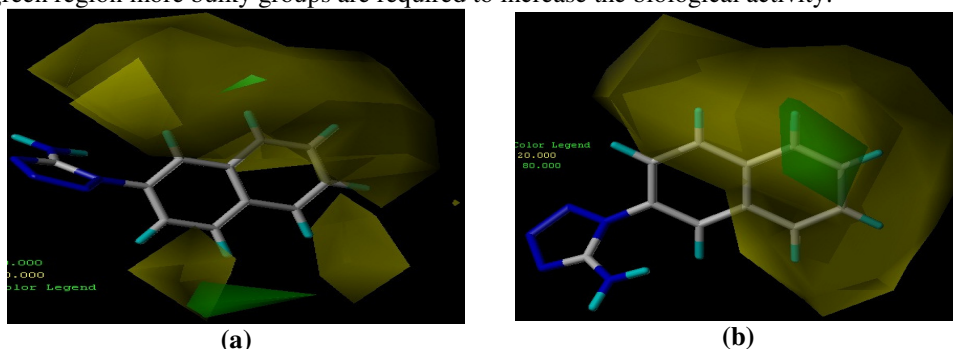

Figure 3. Steric field contour plots of (a) CoMFA and (b) CoMSIA

Electrostatic contour map

The contour plots of electrostatic fields of CoMFA and CoMSIA for the most active compound are depicted in Figure 4(a) and (b), respectively. Electrostatic contour maps are displayed in red and blue colored contours. The introduction of positively charged groups in blue contour map electrostatic regions and negatively charged groups in red contour maps enhance biological activity of the compound.

In CoMFA, for the most active compound 14, a medium sized blue contour, two medium-sized red contours and a small red contours are present above the benzene ring. Red contour maps refer to regions for the substitution of electron donating groups and blue contour maps for electron withdrawing groups which can enhance biological activity. In CoMSIA two very small red contours and a medium sized blue contour are present. The substituents at these sites with the above mentioned groups would enhance biological activity.
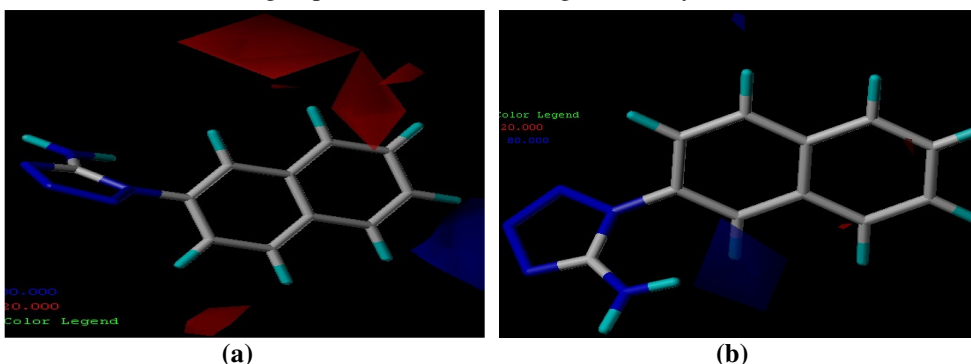

Figure 4. Contour plots of electrostatic fields of (a) CoMFA and (b) CoMSIA

Hydrophobic contour analysis

CoMSIA hydrophobic contour maps are represented by yellow and white colors. The contour plot of CoMSIA hydrophobic fields for the most active compound is depicted in Figure 5(a). This contour map showed two small yellow contours and big white contour. The white contour refers

to regions where hydrophilic substituents are favored while yellow contours refer to regions where hydrophobic substituents are favored to increase the biological activity.

Hydrogen bond donor and acceptor contour map

The contour maps of donor and acceptor fields explain the   pecial arrangement of the favorable and disfavourable H-bond interactions to donor or acceptor groups of the target protein. The contour plot of CoMSIA for the most active compound 14   depicted in Figure 5(b) showed only hydrogen bond acceptor fields with no donor contours present. While CoMSIA contour plots displayed hydrogen bond acceptor fields as red contours and donor group as pink contour maps. The hydrogen bonding at these sites is favorable for biological activity.
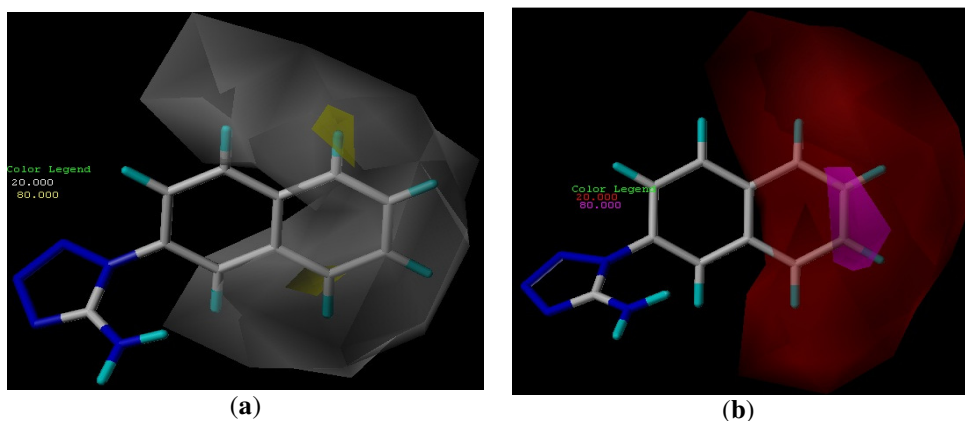

(**a**)   (**b**)

Figure 5. (a) Contour plot of CoMSIA hydrophobic fields (b) Contour plot of CoMSIA

Docking results

Molecular docking studies were performed using Autodock4.2. The results obtained by using molecular docking method disclosed the possible molecular interactions or orientation of 5-amino-1-aryl-1*H*-tetrazole derivatives in the binding pocket of PDB id 1MD2. All the molecules were docked against the protein 1MD2. The most active compound showed a docking energy of -4.48 kcal/mol for 1MD2 with Thr92. Least active compound shows interaction with Thr92 for 1MD2 protein. Figure 6(a) and (b) shows receptor ligand interactions for most and least active compounds for 1MD2.
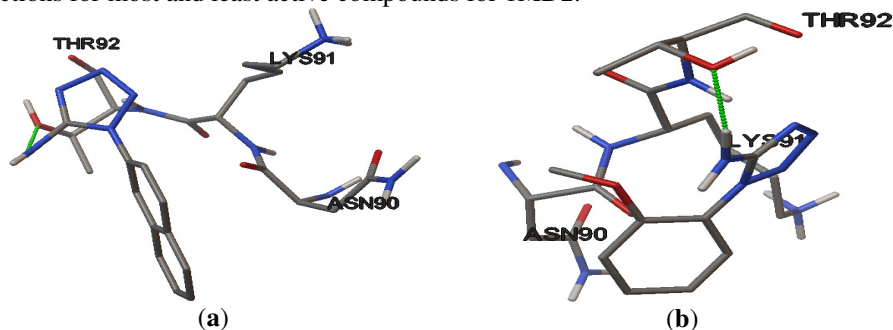

(**a**)   (**b**)

Figure 6. (a) Docking interaction of most active compound and (**b**) Docking interaction of least active compound for 1MD2

Designing of novel molecules

In our quest for finding better inhibitory molecules against cholera, we used the information obtained from the contour plots of the highest active molecule to design 7 new molecules that would act effectively against cholera. Table 3 shows the list of compounds designed that showed better inhibitory activity against cholera. The most active compound has pIC_{50} 4.01 and the predicted compounds showed higher pIC_{50} than experimental pIC_{50} values. The predicted pIC_{50} values for all the 7 molecules are given in Table 4.

Table 3. Structures of novel molecules
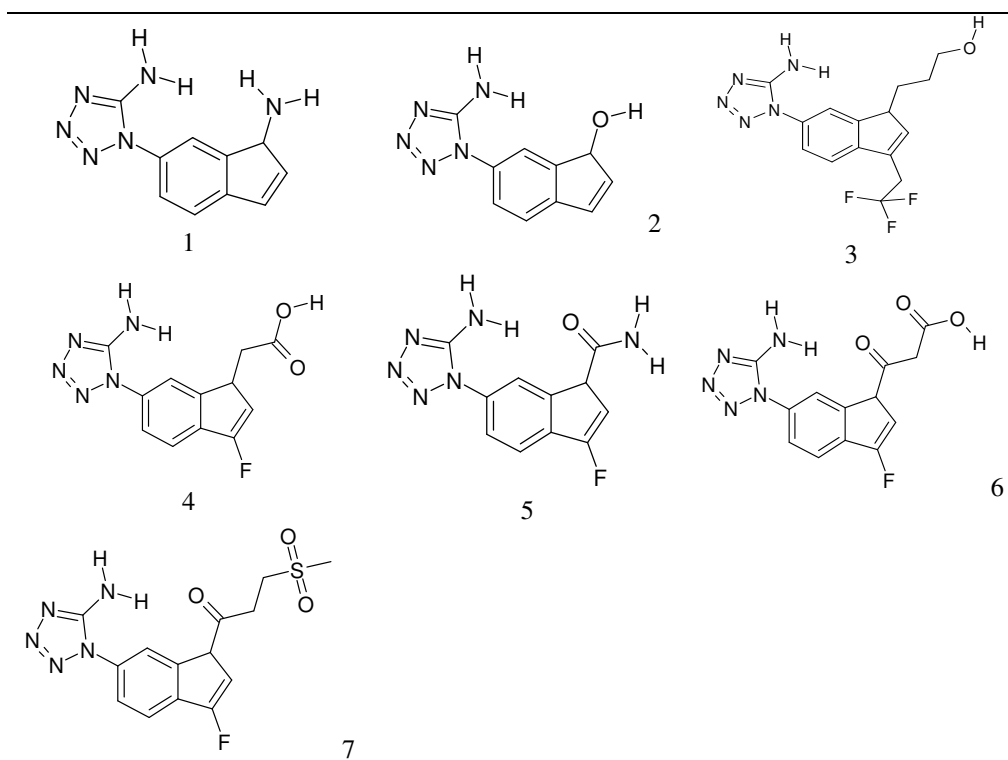
Among the seven novel molecules whose structures have been interpreted based on the pIC_{50} values, structure 3 and 4 showed pIC_{50} value of 4.24 and 4.23 respectively. These values which are relatively high among the lead molecules designed are ascribable to presence of corresponding electron donating group in red region and electron withdrawing group in the blue region of contour maps of precursor molecule (*c.f.* Figure 4a and 4b).

The presence of moderately donating 4-(methyl sulfonyl) 2-butanone in the red region and also the presence of F in the blue region, in the structure 7 correspond to medium pIC_{50} value of 4.09. Same pIC_{50} value of 4.07 in structure 2 and 5 is ascribable to nature of substituents in regions, one favoring electron withdrawing group and other electron donating group. In structure 2 electron donating OH group is present at red region and there is no substituent in the blue region. While in structure 5 the presence of contradicting electron withdrawing group (amide) in red region can cause decrease of biological activity but it is compensated by the presence of a withdrawing group at blue region.

In structure 1 which exhibits Pic_{50} value of 4.06 is justifiable due to presence of only NH_2 electron donating group at red region and absence of any substituent in blue region. Structure 6 has lowest Pic_{50} value of 4.03 indicating least biological activity which may be probably due to presence of contrary strong electron withdrawing 3-oxo butanoic acid group in the red region in spite of electron withdrawing moiety F in blue region. The decreasing trend of Pic_{50} values observed follows the order: structure 3> structure 4> structure 7> structure 2 ,5 > structure 1> structure 6 which is ascribable to presence or absence of electron donating and electron withdrawing groups with varying strengths present in their respective red or blue regions.

Table 4. Novel molecule predicted Pic_{50} values

| C.No. | Predicted Pic_{50} values of novel compounds | 1MD2 | |
		Docking energy of predicted molecules	Interacting amino acids
1	4.06	-4.14	Thr92
2	4.07	-4.11	Thr92
3	4.24	-4.15	Asn90, Thr92
4	4.23	-4.33	Thr92
5	4.07	-4.61	Asn90
6	4.03	-4.65	Asn90, Thr92
7	4.09	-5.09	Asn14, Asn90, Thr92

Conclusion

In the last decade, ligand-based and structure-based methods have become great tools in drug design, including lead discovery and optimization. It has also been shown that ligand-based and structure-based methods are now able to predict, with an acceptable degree of accuracy, the position of a substituent in the ligand. The molecular docking and 3D-QSAR studies together could be employed to design the more potent cholera vulgaris inhibitors. By the aid of molecular docking, the derived 3D-QSAR models are also able to indicate which interaction sites in the binding pockets strongly responsible for the variance in biological activities. In this study, calculated CoMFA and CoMSIA 3D-QSAR models for 5-amino-1-aryl-1*H*-tetrazole derivatives were expanded. Because of the nature of cholera vulgaris active site, which is covered by hydrophobic residues, shape contribution in our CoMFA and CoMSIA model is a dominant factor. The number of hydrophobic interactions and hydrogen bonding between chemical inhibitors and cholera vulgaris active site could increase the inhibitory potency. Based on this useful information seven new novel molecules are predicted with good interactions.

References

1. Ford R E, Knowles P, Lunt E, Marshall S M, Penrose A J, Ramsden C A, Summers A J H, Walker J L and Wright D, *Eur J Med Chem*., 1986, **29**, 538.
2. Peet N P, Baugh L E, Sunder S, Lewis J E, Matthews E H, Olberding E L and Shah D N, *J Med Chem*., 1986, **29(11)**, 2403-2409; DOI:10.1021/jm00161a045
3. Katritzky A R, Rogovoy B V and Kovalenko K V, *J Org Chem*., 2003, **68(12),** 4941-4943; DOI:10.1021/jo0266543
4. Radheshyam S, Abhilash Singh and Jayashree Nagarkar M, *Tetrahedron Lett.*, 2013, **54(1)**, 106-109; DOI:10.1016/j.tetlet.2012.10.116
5. Mitch CH and Quimby SJ, PCT Int Appl WO 9851312, 1998.

6.   Wood E, Crosby R M, Dickerson S, Frye, Griffin R and Hunter R, *Anticancer Drug Des.*, 2001, **16**, 1-6.

7.   Barnaby Mayab C H and Abell A D, *J Chem Soc*., 2002, 172-178; DOI:10.1039/B109128J

8.   Fredrik E K, Wistrand L G and Frejd T, *Tetrahedron*, 2003, **59(35),** 6759-6769; DOI:10.1016/S0040-4020(03)00818-4

9.   Zhao-Xu C and Heming X, *Int J Quantum Chem*., 2000, **79(6),** 350-357; DOI:10.1002/1097-461X(2000)79:6<350::AID-QUA3>3.0.CO;2-T

10.   Flippin L A, *Tetrahedron Lett*., 1991, **32**, 6857-6860; DOI:10.1016/0040-4039(91)80425-6

11.   Rhonnstad P, Wensbo D and Rhonnstad P, *Tetrahedron Lett.*, 2002, **43(17),** 3137-3139; DOI:10.1016/S0040-4039(02)00490-2

12.   Jursic B S and Leblanc B W, *J Heterocycl Chem*., 1998, **35(2),** 405-408; DOI:10.1002/jhet.5570350224

13.   John E O, Kirchmeier R L and Shreeve J M, *Inorg Chem.*, 1989, **28**, 4629-4633; DOI:10.1021/ic00325a018.

14.   Katritzky A R, Jain, Petrukhin R, Denisenko S and Schelenz T, *SAR QSAR Environl Res.*, 2001, **12(3),** 259-266; DOI:10.1080/10629360108032915

15.   Gunda S K, Anugolu R K, Sri Ramya R and Mahmood S, *Acta Pharm*, 2012, **62(3)**, 287-304; DOI:10.2478/v10007-012-0029-7

16.   Gunda S K, Sri Swathi M, Sharada D, Swarna A and Mahmood S, *Int J Pharm Sci Rev Res*., 2015, **30**, 28-34.

17.   Hanley S, Louis Rd St. SYBYL 6.7 is available from Tripos Associates Inc USA 1699; MO 631444.

18.   Clark III M C, Cramer R D and Bosch N V O, *J Comput Chem*., 1989, **10(8),** 982-1012; DOI:10.1002/jcc.540100804

19.   Cho S J and Tropsha A, *J Med Chem*., 1995, **38(7),** 1060-1066; DOI:10.1021/-jm00007a003

20.   Wold S, *Technometrics,* 1978, **20(4)**, 397-405; DOI:10.1080/00401706.1978.10489693

21.   Gohlke H, Hendlich M and Klebe G, *J Mol Biol*., 2000, **295(2),** 337-356; DOI:10.1006/jmbi.1999.3371

22.   Morris G M, Huey R, Lindstrom W, Sanner M F, Belew R K, Goodsell D S and Olson A J, *J Compu Chem.*, 2009, **16,** 2785-2791; DOI:10.1002/jcc.21256