International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

2D AND 3D-QSAR STUDY OF ARYL-4,5-DIHYDRO-1H-PYRAZOLE-1-CARBOXAMIDE DERIVATIVES FOR HUMAN ISOZYME CARBONIC ANHYDRASE XII

TORASKAR MP[1*] AND SALVI SD[2]

1: Associate Professor, Department of Pharmaceutical Chemistry, Bharati Vidyapeeth's College of Pharmacy, Navi Mumbai, Maharashtra, India, 400614

2: Department of Pharmaceutical Chemistry, Bharati Vidyapeeth's College of Pharmacy, Navi Mumbai, Maharashtra, India, 400614

*Corresponding Author: E Mail: Dr. Mrunmayee P. Toraskar: mrunmayeetoraskar@gmail.com

ABSTRACT

Carbonic Anhydrase (CA) enzyme catalyse the inter-conversion of carbon dioxide and water to bicarbonate and proton, thus it is involved in crucial physiological processes in the body while the overexpression of this enzyme leads to various disorders. Among various human CA isozymes, CA XII isoform is mostly expressed in cancer cells. Because QSAR analysis is one of the most frequently used modelling technique, here 2D and 3D-QSAR studies were carried out on previously synthesized series of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1H-pyrazole-1-carboxamide derivatives as carbonic anhydrase XII inhibitors using software VLife MDS version 4.6. From the various developed 2D and 3D-QSAR models, the best model with good predictive capability of r^2 and pred r^2 and correlative coefficient (q^2) were selected. In 3D-QSAR technique, models were developed by using various methods of k-nearest neighbour molecular field analysis (kNN-MFA). 2D-QSAR model depicts contribution of descriptors like T_N_Cl_7, chi6chain and T_2_F_5 for inhibition of CA XII isozyme activity. The results of 3D-QSAR models provides a direct view of factors expressed in terms of molecular field (electrostatic and steric) influencing the activity of CA XII isozyme on the derivatives. The information provided by the 2D and 3D-QSAR

studies may lead to better understanding of the structural requirement for carbonic anhydrase XII inhibitors.

Keywords: 2D-QSAR, 3D-QSAR, hCA XII
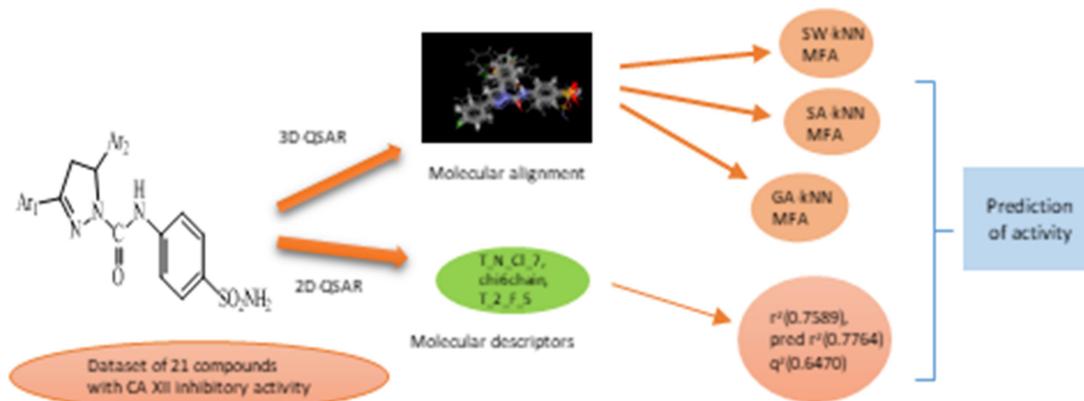
INTRODUCTION

The carbonic anhydrase (CA) belongs to the class of metalloenzymes, that catalyse the reversible interconversion between carbon dioxide and bicarbonate and participate in ion transport and pH control [1]. It is an essential enzyme in bacteria, archaea, and eukaryotes. In humans, CA-catalysed reactions are involved in various important biological processes such as respiration, pH balance, CO_2 homeostasis, electrolyte secretion in a variety of tissues/organs, gluconeogenesis, lipogenesis, ureagenesis, bone resorption, calcification, tumorigenicity, and many other physiological or pathological processes [2]. A total of 16 CA isozymes have been described so far in mammals from which many are well-established therapeutic target to treat wide variety of disorders. These isozymes are classified based on their cellular and sub-cellular location, such as cytosolic (CA I, CA II, CA III, CA VII, and CA XIII), membrane-associated (CA IV, CA IX, CA XII, and CA XIV), mitochondrial (CA VA and CA VB), and one is secreted in saliva and milk (CA VI). The membrane-associated CA IX and XII isoforms have been discovered in association with certain human cancers and also being expressed in a limited number of normal cells [3]. The environment near the solid tumor is dynamic which is mainly characterised by acidic pH and low levels of oxygen (i.e. hypoxia). Cancer cells acquire different adaptive feature for survival purpose which are probably responsible for the development of invasive and metastatic phenotypes. As there is need for high levels of ATP synthesis to the cancer cells, as a response to regional hypoxia, these cells obtained their need by switching the metabolism from aerobic respiration to fermentative (anaerobic), glycolysis. The tumor cell pH is regulated by CA IX and CA XII by maintaining an intracellular slightly

alkaline pH (7.2–7.4) while acidifying the extracellular pH (which arrives at 6.0–6.5), and this is exploited by the cancer cells for survival and proliferation. Overexpression of CA IX and CA XII is observed in solid tumor in hypoxic condition, which maintain the acidic tumor environment by catalysing the reversible hydration of tumor cell-generated CO_2 into a bicarbonate anion (HCO_3 −) and a proton (H+) and then trap H+ extracellularly to lower the pH [4]. Thus overexpression of both these isozymes contributes to the increased in acidification of extracellular hypoxic environment in contrast with the normal tissue. In breast tumors and in several cancer cell lines expression of both membrane-associated CA is induced by hypoxia [3].

Sulfonamides was the first organic inhibitor of CA isozymes. The primary sulfonamide group binds to the Zn2+ ion in the CA isozyme active site as an anion (SO2NH−) blocking the catalysis. Sulphonamide derivatives such as acetazolamide, methazolamide, ethoxzolamide, dichlorophenamide, dorzolamide and brinzolamide have been clinically used for decades as inhibitors of the zinc enzyme carbonic anhydrase. Aromatic and heterocyclic sulfonamides have been extensively investigated as CA inhibitors to treat various disorders. A thorough literature survey has shown that the presence of urea as a linker in sulfonamide-containing compounds is beneficial to generating effective CA inhibitory activity. In the research work, novel 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives were designed and synthesized on the basis of sulfonyl semicarbazides and iminoureido CA inhibitors previously synthesized in our laboratory [4, 5]. In vitro inhibition studies were carried out for all newly synthesized compounds against CA I, II, VII, IX, and XII. The synthesized derivatives were thoroughly characterized by ¹H-NMR, ¹³C-NMR, IR, and MS, and their purity was checked by HPLC [3].

Recently the advancement of computational chemistry led to new challenges of drug discovery. Molecular modelling study is an approach in drug designing that is used to minimize a library of high number of random molecules into smaller list of potentially effective inhibitors. QSAR is one of the tool of molecular modelling that is used for drug designing. QSAR methodologies save resources and minimize the process of development of new molecules or drug. It presents an attempt to correlate specific biological activities or physical properties of a series of compound with the measured or computed molecular properties of the compounds in terms of descriptors. QSAR studies provides better understanding of the interaction mechanisms between chemical compounds and biological targets [6]. Based on this advantages of QSAR studies, here in the present work QSAR studies was performed on the previously synthesized derivatives

which were evaluated as carbonic anhydrase inhibitors [4].

To the best of our knowledge, till date such QSAR study has not been reported on 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives as carbonic anhydrase inhibitors as discuss in this article. In the present work, 2D and 3D-QSAR studies were performed on a series of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives as carbonic anhydrase inhibitors using software VLife MDS version 4.6. The results obtained may contribute to further designing novel sulphonamide derivatives as carbonic anhydrase inhibitors.

MATERIALS AND METHODS

All molecular modelling studies i.e. 2D-QSAR and 3D-QSAR were performed using the molecular modelling software package VLife Molecular Design Suite (VLife MDS) version 4.6.

DATASET FOR ANALYSIS

A dataset of 21 compounds obtained from the literature consisted of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives as carbonic anhydrase inhibitors [4]. For QSAR analysis the in-vitro biological activity of compounds against the carbonic anhydrase XII [Ki(nM)] was converted into negative logarithm (-pKi). **Table 1** shows the structure of 21 compounds along with their biological activity values and their -log.
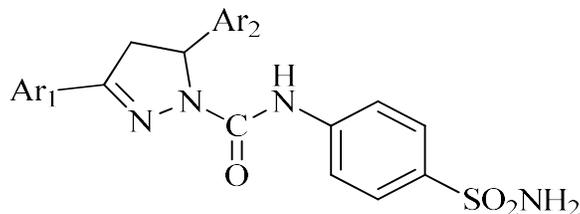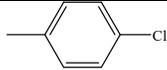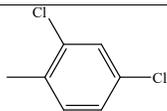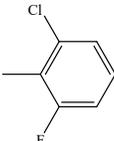

Table 1: Structure of dataset along with their biological activities and –log values

Compounds	Ar$_1$	Ar$_2$	CA XII (K$_i$)	pK$_i$
3A	C$_6$H$_5$	C$_6$H$_5$	7.71	8.11
3B	C$_6$H$_5$	(4-Cl-C$_6$H$_4$)	1.01	8.99
3D	C$_6$H$_5$	(2,4-diCl-C$_6$H$_3$)	0.84	9.07
3E*	C$_6$H$_5$	(4-F-C$_6$H$_4$)	6.32	8.19
3G*	C$_6$H$_5$	(2-Cl-6-F-C$_6$H$_3$)	1.05	8.97

Code	R1	R2	Value 1	Value 2
3H*	C₆H₅	(2-methylthiophene structure)	2.49	8.60
3J	C₆H₅	(9-methylanthracene structure)	6.82	8.16
3L	(4-bromophenyl structure)	C₆H₅	6.67	8.17
3M	(4-bromophenyl structure)	(4-chlorophenyl structure)	0.89	9.05
3N*	(4-bromophenyl structure)	(4-fluorophenyl structure)	2.38	8.62
S1	(4-chlorophenyl structure)	C₆H₅	9.16	9.00
S2	(4-chlorophenyl structure)	(4-methylphenyl structure, CH₃)	9.19	9.08
S3	(4-chlorophenyl structure)	(4-chlorophenyl structure, Cl)	1.15	8.99
S4	(4-chlorophenyl structure)	(4-fluorophenyl structure, F, Cl)	1.14	8.03
S5	(4-chlorophenyl structure)	(2-chlorophenyl structure, Cl)	1.03	8.03
S6*	(4-chlorophenyl structure)	(2,4-dichlorophenyl structure, Cl, Cl)	1.07	8.93
S8	(4-chlorophenyl structure)	(2-methylthiophene structure, S)	0.62	8.94
S9	(4-chlorophenyl structure)	(Cl, methyl, F substituted phenyl structure)	6.83	8.98
S13*	(4-chlorophenyl structure)	(2-hydroxyphenyl structure, HO)	0.99	8.97
S14	(4-chlorophenyl structure)	(4-methoxyphenyl structure, OCH₃)	0.82	9.20
S15	(4-chlorophenyl structure)	(2,3-dimethoxyphenyl structure, OCH₃, OCH₃)	1.01	8.16

Note: Compound with * are used in the test set

LIGAND PREPARATION

The 2D chemical structures were drawn using Chemdraw and converted to 3D using 3D ChemDBS module by VLife MDS version 4.6 software package. All the 3D structures were optimized using Merck molecular force field (MMFF) with distance dependent dielectric constant value 1 and the root-mean-square gradient (RMSgra) was set to at 0.01 kcal/mol Å, along with 10,000 number of cycles.

SELECTION OF TRAINING AND TEST SET

Dataset of 21 molecules was divided into training and test set to get reliable QSAR model. The sphere exclusion method was adopted for division of training and test sets. The dataset of 21 molecules was splitted into 15 training and 6 test set. Same procedure was used to build 3D-QSAR model. To assess the similarity of the distribution pattern of the molecules in the generated sets, statistical parameters (with respect to the biological activity) i.e. mean, maximum, minimum and standard deviation were calculated for the training and test sets. The min and max values in both train and test set should be compared in a way such that

i.   The max of the test should be less than or equal to max of train set

ii. The min of the test should be greater than or equal to min of train set

The mean and standard deviation of the train and test set provides insight to the relative difference of mean and point density distribution (along mean) of the two sets [7]. The statistical parameters were obtained such as mean, maximum, minimum and standard deviation are reported in **Table 2**.

Table 2: Unicolumn statistics of the training and test sets for 2D QSAR model

Dataset	Average	Max	Min	SD	Sum
2D-QSAR					
Training	7.9393	8.2800	6.6300	0.4707	111.010
Test	7.8625	8.2800	7.1700	0.4808	31.45
3DQSAR					
Training	8.6593	9.2000	8.0300	0.4703	129.8900
Test	8.7250	9.0000	8.1900	0.3189	52.3500

2D molecular descriptor calculation

The energy-minimized geometry was used for the calculation of the various 2D molecular descriptors. The 2D molecular descriptors such as molecular weight (M.W), hydrogen bond acceptor (HBA) count, hydrogen bond donor (HBD) count, rotatable bonds, slogp, XlogP, smr, chi2, SaaO count, negative potential surface area, SdsCHE-index, Individual, Chi, ChiV, Path count, Chi Chain, Chi Chain, Chain path count, Cluster, Path cluster, Kappa, Element count, Estate number, Estate contribution, Polar surface area, Alignment independent topological descriptors such as T_T_O_2, T_T_T_O, T_T_N_6, T_T_T_1 type were calculated [8]. Descriptors with constant values were removed, as they have no contribution to the model, leading to the

6

generation of 336 descriptors for QSAR analysis.

Molecular alignment of 3D-QSAR

The molecules in the dataset were align with the help of template-based alignment method of VLifeMDS using the common structures of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives. The alignment of all the molecules on the template is shown in **Figure 1**.
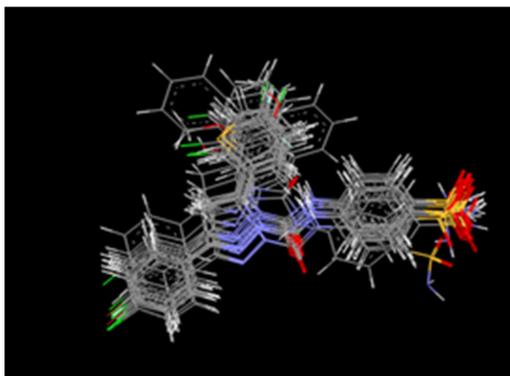

Figure 1: Common template used for alignment of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives

3D molecular descriptor calculation

Similar to many 3D QSAR methods, k-nearest neighbour molecular field analysis (kNN-MFA) requires suitable alignment of set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic energies are computed at the lattice points of the grid using methyl probe of charge +1. These interaction energy values at the grid points are considered for relationship generation using kNN method and utilized as descriptors to decide nearness between molecules. The term descriptor is utilized in the following discussion to indicate field values at the lattice points.

RESULT AND DISCUSSION

A series of 21 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives was selected for QSAR studies. [4] The dataset of 21 molecules were divided into 15 training and 6 test set using sphere exclusion method. Selection of molecules in the training set and test is an important feature of any QSAR model. Therefore, the care was taken in such a way that biological activities of all compounds in test lie within the maximum and minimum value range of biological activities of training set of compounds. The unicolumn statistical parameters values were found to be in good agreement and are given in **Table 2**. This observation showed that the test set was interpolative and derived within the minimum–maximum range of the training set. Statistically significant models were selected, considering the term selection criterion as r^2, q^2 and pred_r^2. The models were validated by both internal and external validation

procedures. Best 2D- and 3D-QSAR models are chosen for discussion.

2D-QSAR

The various 2D-QSAR models were developed using partial least square regression (PLSR) analysis in conjunction with stepwise forward-backward variable selection method. The best regression equation obtained is as follow **(Table 4)**:

chi6chain

The descriptor selected for this 2D-QSAR model are summarized in **Table 5**. The above equation of 2D-QSAR model explains 75% (r^2 =0.7589) of the total variance in the training set. It shows cross-validated squared correlation coefficient (q^2=0.6470) of 64 % and a predictivity for the external test set (pred r^2 =0.8220) of about 82%. The F-test = 21.5431 shows the statistical significance of 99.99% of the model which means that probability of failure of the model is 1 in 10,000. In addition, the randomization test shows confidence of 99.9% that the generated model is not random and hence it is chosen as the QSAR model. The fitness plot of observed versus predicted activity provides an idea about how well the model was trained and how well it predicts the activity of the external test set. From the fitness plot i.e **Figure 2** it can be seen that the model is able to predict the activity of training set quite well (all points are close to regression line) as well as external test set up to 69% (all points are close to regression line) providing confidence in

predictive ability of the model. The predicted activities & residuals of both training & test set compounds by 2D-QSAR model are shown in **Table 4**. The correlation matrix is shown in **Table 6** which shows good correlation of selected parameters with biological activity.

Interpretation of 2D-QSAR model

From the 2D-QSAR equation, negative coefficient value of T_2_F_5[count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from Fluorine atom by 5 bonds] and chi6chain(a retention index for six membered rings) on the carbonic anhydrase activity indicated that lower values leads to good carbonic anhydrase activity while higher value leads to reduce in the activity whereas positive coefficient value of T_N_Cl_7[count of number of Nitrogen atoms (single double or triple bonded) separated from any Chlorine atom (single or double bonded) by 7 bond] on the carbonic anhydrase activity indicated that higher value leads to better carbonic anhydrase activity while lower value leads to decrease the activity. The descriptor T_2_F_5 contributes negatively to the equation that means the molecules in the dataset containing fluorine atom has less carbonic anhydrase activity as compare to other molecules. The negative coefficient of chi6chain descriptor helps in understanding the effect of retention index of six membered ring in the molecules to the activity. The results for descriptor T_N_Cl_7 are in close agreement with the experimental observation

of some derivatives as they have substituent i.e. Cl group, respectively, in the aromatic ring (Ar1 and Ar2) that produce good carbonic anhydrase activity **[8]**.

3D-QSAR

3D-QSAR studies of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives with reported activities against carbonic anhydrase XII were performed using various k-nearest neighbour molecular field analysis methods such as stepwise variable forward backward analysis (SW-kNN-MFA), simulated annealing (SA-kNN-MFA) and genetic alogarithm (GA-kNN-MFA) method **[11]**. Several statistically s reliable and accurate. The LOO cross-validated analysis of the best model gave rise to a q^2 value of 0.944 suggesting that the model is a useful tool for predicting inhibitory activity against carbonic anhydrase enzyme (XII). A data set of six compounds was selected as the test set from the original data of 21 compounds for the validation experiments. The predictive ability of this forward stepwise variable selection kNN-MFA model was evaluated by predicting the biological activities of the test set compounds. Residual values obtained by subtraction of predicted activities from biological activities were found near to zero. Therefore, it was concluded that the resultant QSAR model by SW-kNN-MFA analysis has good predictive ability. The value of pred_r^2 was obtained for the test set and gave better results, with a value of 0.94, which

significant 3D-QSAR models were generated, of which the corresponding best model is reported herein.

3D-QSAR SW-kNN-MFA model

For inhibitory activity against carbonic anhydrase, selected the template-based 3D-QSAR model with 15 training set compounds was found to be statistically most significant, especially with respect to the internal predictive ability (q^2 = 0.944) of the model. As the cross-validated correlation coefficient (q^2) is used as a measure of reliability of prediction, the correlation coefficient suggests that   our   model   i

means 94% predictive power for the external test set. Thus, our model displays good predictivity in regular cross-validation that is mention in **Tabl**

e 3. The actual and predicted activities by 3D-QSAR model & residuals of both training & test set compounds are given in **Table 4**. The fitness plot of observed versus predicted activity of both training & test set compounds helped in cross-validation of SW-kNN-MFA 3D-QSAR model and are depicted in **Figure 3**.

Interpretation of 3D-QSAR SW-kNN-MFA model

The 3D-QSAR model selection criteria are the value of q^2, the internal predictive ability of the model, and that of pred_r^2, the ability of the model to predict the activity of external test set. The stereo view of molecular rectangular grid field of 3D-QSAR model that

is depicted in **Figure 4** indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules. The green-colored ball specifies the position of the steric descriptors and the blue-colored ball specifies the position of the electrostatic descriptors. The points generated in SW-kNN-MFA 3D-QSAR model are S_234(-0.1606, -0.1049), S_323(30.0, 30.0), E_488(2.1124, 2.2905) and E_544(-10.00, -10.00) that is steric and electrostatic interaction fields at lattice points234, 323, 488 and 544, respectively. From 3D-QSAR model, it is observed that electrostatic descriptor like E_488 with positive values is near pyrazoline ring. This indicates that positive electrostatic potential is favourable for increase in the activity and hence a less electronegative substituent group is preferred in that region. The steric effect i.e. S_234, as shown with green ball around the pyrazoline ring, implies that less bulky substituent is preferred in Ar1 position of pyrazoline ring and produces good activity. The positive value of S_323 shows that C6H4Cl, C6H4F, C6H3ClF, etc., at Ar2 position of pyrazoline ring is important for activity and it can be replaced with large bulky ring substituent. E_544 electrostatic descriptor with negative coefficient is near from the Ar2 position of the pyrazoline ring. This indicates that negative electrostatic potential is favourable for increase in the activity and hence more electronegative substituent group is preferred in that region **[10]**.

3D-QSAR SA-kNN-MFA model

The 3D-QSAR SA-kNN-MFA model is selected on the basis of statistical coefficients like q^2 (0.71) and pred r^2 (0.64). The fitness plot of observed versus predicted activity of both training & test set compounds helped in cross-validation of 3D-QSAR SA-kNN-MFA model and are depicted in **Figure 5**. The observed, predicted values, and residual values for activities as carbonic anhydrase inhibitors of both training and test set molecules are given in **Table 2**.

Interpretation of 3D-QSAR SA-kNN-MFA model

For 3D-QSAR SA-kNN-MFA model as shown in **Figure 6** it is observed that the electrostatic descriptor E_761 with negative value (-0.1161, 0.0259) represented with the blue ball is near sulphonamide moiety which indicates that more electronegative substituent is preferred in that region. The negative value of this descriptor is due to sulphonamide group.

3D-QSAR GA-kNN-MFA model

The statistically best significant model obtained by the 3D-QSAR GA-kNN-MFA method shows the values of q^2 and pred r^2 as 0.81 and 0.64 from which it can be said that the model is accurate. The fitness plot of predicted vs. observed activity of both training and test sets is shown in **Figure 7**. The predicted activities of the compounds in both test and training sets are shown in **Table 4**.

Interpretation of 3D-QSAR GA-kNN-MFA model

Three-dimensional data points were generated that contribute to GA-kNN-MFA model and are shown in **Figure 8**. S_446 (-0.1632, -0.1234) with positive descriptor indicates that less steric substituents are preferred near the pyrazoline ring. It is observed that electric field with positive coefficient E_185 (0.7539 1.1223) is near Ar1 position of the structure and positive range indicates that positive electrostatic potential is favourable for increase in the activity and hence a less electronegative substituent group is preferred in that region. The electrostatic descriptors E_564 and E_872 with negative and positive value are far away from the structure so it may hardly have any effect on the substituent nature.

Table 3: Statistical results of 2D-QSAR and 3D-QSAR model

Sr. no	Statistical parameter	Results			
		2D-QSAR	3D-QSAR		
			SW-kNN	SA-kNN	GA-kNN
1	r^2	0.7589	-	-	-
2	r^2se	0.2176	-	-	-
3	q^2	0.6470	0.944	0.7119	0.8119
4	q^2se	0.3581	0.5109	0.3286	0.5286
5	pred_r^2	0.8220	0.9453	0.6422	0.5422
6	pred_r^2se	0.7441	0.5611	0.5702	0.6806
7	F-test	21.5431	-	-	-
8	N	21	-	-	-
9	Nearest neighbour	-	2	2	2
10	Degree of freedom	-	10	13	13
11	Zscore_q^2	4.62576	-	-	-
12	best_rand_q^2	0.49388	-	-	-
13	alpha_rand_q2	0.00005	-	-	-

Table 4: Comparative observed and predicted activities of 2D and 3D-QSAR models

Comp	pKi	2D QSAR		3D SW-kNN		3D SA-kNN		3D GA-kNN	
	Observed activity	Pred.	Res.	Pred.	Res.	Pred.	Res.	Pred.	Res.
3A	8.11	8.29	-0.18	8.86	0.13	8.705	0.285	8.78	0.21
3B	8.99	8.875	0.115	8.86	0.13	8.87	0.12	8.87	0.12
3D	9.07	8.89	0.18	8.86	0.21	8.85	0.22	8.78	0.29
3E*	8.19	8.45	-0.26	8.48	-0.29	8.213	-0.023	8.45	-0.26
3G*	8.97	8.85	0.12	8.59	0.38	8.5698	0.272	8.88	0.09
3H*	8.60	8.57	0.03	8.84	-0.24	8.688	-0.08	8.69	-0.09
3J	8.16	8.06	0.1	8.34	-0.69	8.45	-0.29	8.2	-0.04
3L	8.17	8.42	-0.25	8.54	-0.37	8.475	-0.30	8.39	-0.22
3M	9.05	8.85	0.2	8.69	0.36	8.86	0.19	8.60	0.45
3N*	8.62	8.84	-0.22	8.74	-0.12	8.329	0.291	8.78	-0.16
S1	9.00	8.59	0.41	8.50	0.5	8.89	0.11	8.51	0.49
S2	9.08	8.80	0.28	8.966	0.114	8.91	0.17	8.86	0.22
S3	8.99	8.85	0.14	8.80	0.19	8.81	0.18	8.69	0.3
S4	8.03	8.23	-0.2	8.36	-0.33	8.1	-0.07	8.30	-0.27
S5	8.03	8.32	-0.29	8.05	-0.02	8.22	-0.19	8.45	-0.42
S6*	8.93	9.18	-0.25	8.84	0.09	8.65	0.28	8.93	0.15
S8	8.94	8.89	0.05	8.58	0.36	8.605	0.335	8.75	0.19
S9	8.98	8.63	0.35	8.55	0.43	8.685	0.295	8.79	0.19
S13*	8.97	8.85	0.12	8.55	0.42	8.69	0.28	8.96	0.01
S14	9.20	8.85	0.35	8.64	0.56	8.625	0.575	8.76	0.44
S15	8.16	8.23	-0.07	8.55	-0.39	8.72	-0.56	8.13	0.03

$$pKi = 9.2429 + 0.3373\ (\pm0.0223)\ T\_N\_Cl\_7 - 0.2467\ (\pm0.0477)\ T\_2\_F\_5 - 2.8971\ (\pm1.1665)$$
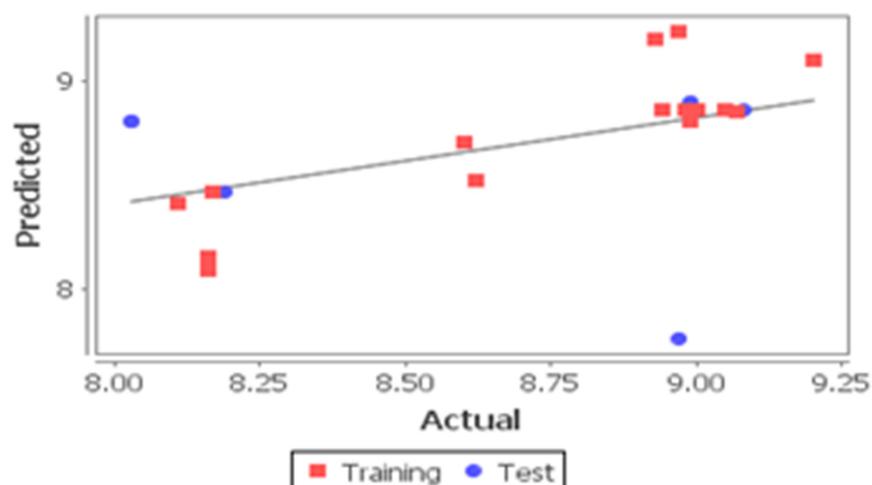
Figure 2: Fitness plot of observed activity versus predicted activity for training set & test set compounds according to 2D-QSAR model

Table 5: Descriptors used in 2D-QSAR model with values

Comp code	2D-QSAR descriptors		
	T_N_Cl_7	chi6chain	T_2_F_5
3A	0	0.287	0
3B	1	0.254	0
3D	2	0.253	0
3E*	0	0.25	1
3G*	1	0.245	0
3H*	0	0.185	2
3J	0	0.397	0
3L	0	0.269	0
3M	1	0.25	0
3N*	0	0.23	1
S1	1	0.269	0
S2	1	0.123	0
S3	2	0.25	0
S4	1	0.25	1
S5	2	0.25	0
S6*	2	0.235	0
S8	1	0.167	0
S9	1	0.235	3
S13*	1	0.32	0
S14	1	0.25	0
S15	1	0.25	0

Table 6: Correlation matrix for physico-chemical descriptors of 2D-QSAR model

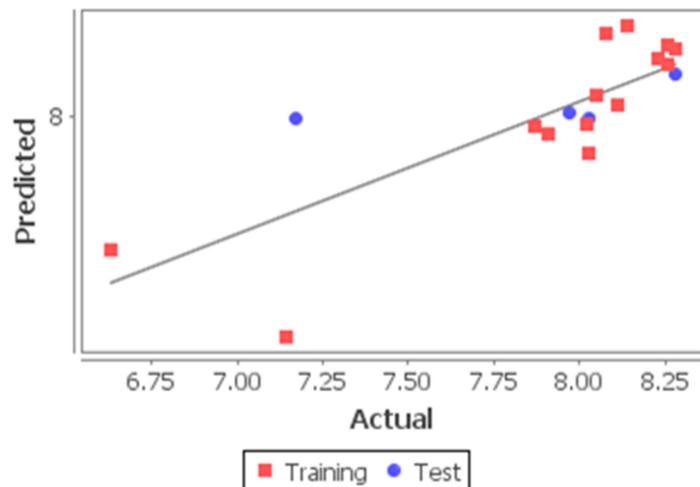| Parameters	T_N_Cl_7	chi6chain	T_2_F_5
T_N_Cl_7	1		
chi6chain	-0.17828	1	
T_2_F_5	-0.28727	-0.21377	1

Figure 3: Fitness plot of observed activity versus predicted activity for training set & test set compounds according to 3D-QSAR SW-kNN-MFA model
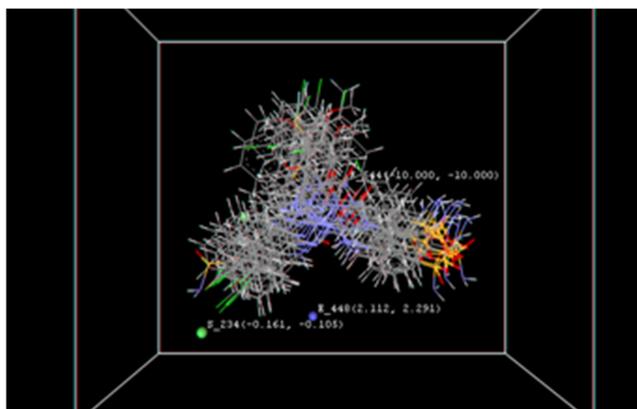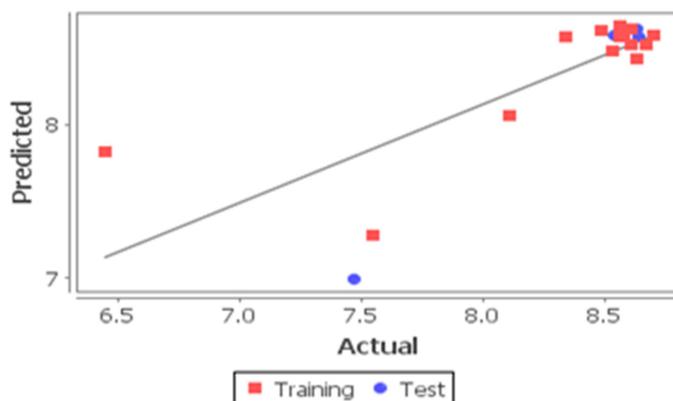

Figure 4: Field point for 3D-QSAR SW-kNN-MFA model

Figure 5: Fitness plot of observed activity versus predicted activity for training set & test set compounds according to 3D-QSAR SA-kNN-MFA model
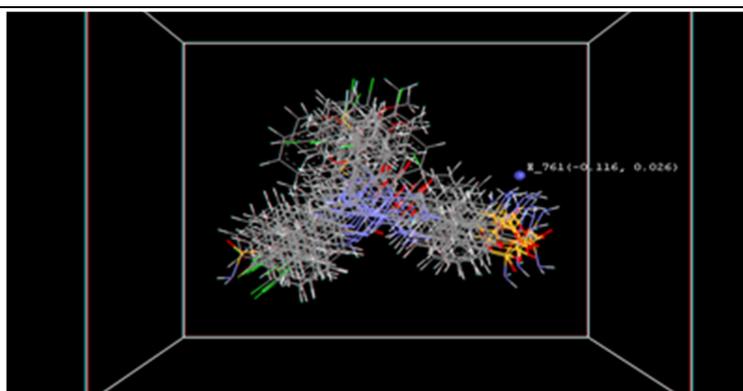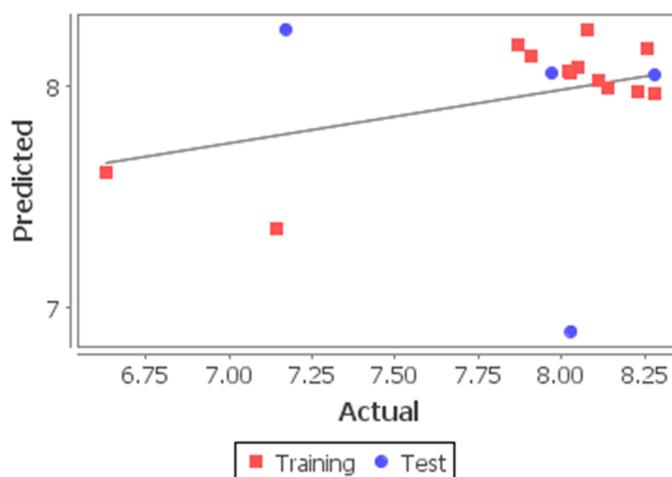
Figure 6: Field point for 3D-QSAR SA-kNN-MFA

Figure 7: Fitness plot of observed activity versus predicted activity for training set & test set compounds according to 3D-QSAR GA-kNN-MFA model
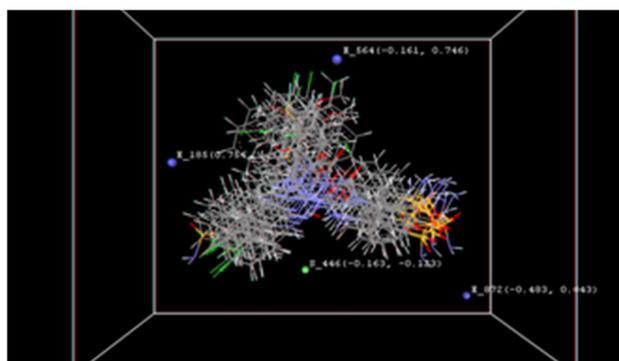

Figure 8: Field point for 3D-QSAR GA-kNN-MFA

CONCLUSION

In the present work, 2D and 3D-QSAR study was performed on the previously synthesized series of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-carboxamide derivatives as carbonic anhydrase inhibitors using software VLife MDS version 4.6. The results of 2D and 3D-QSAR studies provided the importance of substitution (electrostatic and steric pattern) required around the 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1*H*-pyrazole-1-

carboxamide derivatives. The 2D-QSAR model indicates that decrease in the value of the descriptor such as T_2_F_5 and chi6chain and increase in the value of the descriptor T_N_Cl_7 leads to increase in potency of molecules. The 3D-QSAR study was done using different kNN-MFA methods such as stepwise variable forward backward analysis, simulated annealing and genetic alogarithm. The results of 3D-QSAR investigated the importance of electrostatic and steric substituents of 3-phenyl-5-aryl-N-(4-sulfamoylphenyl)-4,5-dihydro-1H-pyrazole-1-carboxamide derivatives against carbonic anhydrase activity. In conclusion, the information provided by the 2D and 3D-QSAR model will guide us to design and synthesize new compounds with an increased biological activity in comparison to the reported compounds.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCE

[1] Supuran CT. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. Nature reviews Drug discovery. 2008 Feb;7(2):168-81.

[2] Smith KS, Jakubzick C, Whittam TS, Ferry JG. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. Proceedings of the National Academy of Sciences. 1999 Dec 21;96(26):15184-9.

[3] Pichake J, Kharkar PS, Ceruso M, Supuran CT, Toraskar MP. Carbonic anhydrase inhibitors: design, synthesis, and biological evaluation of novel sulfonyl semicarbazide derivatives. ACS medicinal chemistry letters. 2014 Jul 10;5(7):793-6.

[4] Hargunani P, Tadge N, Ceruso M, Leitans J, Kazaks A, Tars K, Gratteri P, Supuran CT, Nocentini A, Toraskar MP. Aryl-4, 5-dihydro-1H-pyrazole-1-carboxamide Derivatives Bearing a Sulfonamide Moiety Show Single-digit Nanomolar-to-Subnanomolar Inhibition Constants against the Tumor-associated Human Carbonic Anhydrases IX and XII. International journal of molecular sciences. 2020 Jan;21(7):2621.

[5] Singasane N, Kharkar PS, Ceruso M, Supuran CT, Toraskar MP. Inhibition of carbonic anhydrase isoforms I, II, IX and XII with Schiff's bases incorporating iminoureido moieties. Journal of enzyme inhibition and medicinal chemistry. 2015 Nov 2;30(6):901-7.

[6] Ajmani S, Jadhav K, Kulkarni SA. Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. Journal of chemical information and modeling. 2006 Jan 23;46(1):24-31.

[7] Bhadoriya KS, Sharma MC, Jain SV, Raut GS, Rananaware JR. Three-dimensional

quantitative structure–activity relationship (3D-QSAR) analysis and molecular docking-based combined in silico rational approach to design potent and novel TRPV1 antagonists. Medicinal Chemistry Research. 2013 May;22(5):2312-27.

[8] Thombare ST, Gonsalves SI, Shaikh AR. 2D and 3D QSAR using kNN-MFA method of pyrazolyl-thiazolinone derivatives as potential EGFR and HER-2 kinase inhibitors. Journal of Computational Methods in Molecular Design. 2012 Dec 2;2(3):107-21.

[9] Bhadoriya KS, Sharma MC, Sharma S, Jain SV, Avchar MH. An approach to design potent anti-Alzheimer's agents by 3D-QSAR studies on fused 5, 6-bicyclic heterocycles as γ-secretase modulators using kNN–MFA methodology. Arabian Journal of Chemistry. 2014 Dec 1;7(6):924-35.

[10] Sharma MC, Kohli DV. Insight into the structural requirement of substituted quinazolinone biphenyl acylsulfonamides derivatives as Angiotensin II AT1 receptor antagonist: 2D and 3D QSAR approach. Journal of Saudi Chemical Society. 2014 Jan 1;18(1):35-45.

[11] Sharma MC. A Structure–Activity Relationship Study of Imidazole-5-Carboxylic Acid Derivatives as Angiotensin II Receptor Antagonists Combining 2D and 3D QSAR Methods.

Interdisciplinary sciences: computational life sciences. 2016 Mar 1;8(1):1-0.