

## 2D and 3D QSAR using kNN-MFA method of the novel 3, 4-dihydropyrimidin-2(1H)-one urea derivatives of N-aryl urea as an antifungal agents.

Ashwini H. Pagare\*, Rani S. Kankate, Anwar R. Shaikh.

Bhujbal Knowledge City, MET's Institute of Pharmacy, Adgaon, Nashik-422003, India  
Received 20 January 2015; received in revised form 20 March 2015; accepted 20 March 2015  
Available online 22 March 2015

### Abstract

Quantitative structure–activity relationship (QSAR) analysis for recently synthesized 3,4-dihydropyrimidin-2(1H)-one urea derivatives of N-aryl urea derivatives was studied for their antifungal activity. The statistically significant 2D-QSAR models ( $r^2 = 0.9759$ ;  $q^2 = 0.9314$ ; F test = 72.73;  $r^2$  se = 5.1472;  $q^2$  se = 8.6750;  $\text{pred}_r^2 = -0.03094$ ;  $\text{pred}_r^2$  se = 40.9144) were developed using molecular design suite (VLifeMDS 4.3.1). The study was performed with 23 compounds (data set) using sphere exclusion (SE) algorithm, random selection and manual selection methods used for the division of the data set into training and test set. Multiple linear regression [MLR] methodology with stepwise (SW) forward-backward variable selection method was used for building the QSAR models. The results of the 2D-QSAR models were further compared with 3D-QSAR models generated by kNN-MFA, (k-Nearest Neighbor Molecular Field Analysis) investigating the substitutional requirements for the favorable antifungal activity against *Candida albicans* and providing useful information in the characterization and differentiation of their binding sites. The results derived may be useful in further designing novel 3,4-dihydropyrimidin-2(1H)-one urea derivatives of N-aryl urea derivatives prior to synthesis.

**Keywords:** 3, 4-dihydropyrimidin-2-(1H) One urea derivatives, Antifungal agents, 2D QSAR, 3D QSAR, MLR, kNN-MFA.

### 1. Introduction

3, 4 – dihydropyrimidin - 2 (1H) - one urea derivatives are known to possess simple structure with wide variety of pharmacological activities such as anti-inflammatory, antifungal, antibacterial, calcium channels blockers, antioxidant, anticancer etc.[1-5]. Biological importance of heterocyclic derivatives of N-aryl ureas have been reported in the literature. Compounds 1– 23 were evaluated for in vitro anti-inflammatory, antibacterial and antifungal activity against various Gram-positive, Gram-negative bacteria and fungal strains using agar well diffusion method.

The compounds 3 and 4 bearing F, Cl group at 2-position are most potent followed by compounds 5, 6, 22 and 23 bearing  $\text{CF}_3$ ,  $\text{OCF}_3$  and  $\text{OCH}_3$  at 2- or 4-position. Thus, compounds bearing substituent such as  $\text{CF}_3$ ,  $\text{OCF}_3$  and  $\text{OCH}_3$  at 2- or 4-position of the terminal benzene ring of urea part found to have higher potency than the compounds bearing such a groups at 3- or 5 positions or at both. Explicitly, 2- or 4-position is the favourable site for high antibacterial activity. The high potency of 3, 4, 22, and 23 may be attributed to the presence of lipophilic or H-bond acceptor type group's placement such as F, Cl,  $\text{CF}_3$ ,  $\text{OCF}_3$  and  $\text{OCH}_3$  at 2- or 4-positions. This is further supported by the fact that the presence of nonpolar lipophilic groups such as isopropyl, n-butyl etc. at 4-position, compounds 20 and 21, respectively, has no

\*Corresponding author.

E-mail address: [ranipharmacy@gmail.com](mailto:ranipharmacy@gmail.com)

(Ashwini H. Pagare)

2230-7842 / © 2015 JCPR. All rights reserved.

major effect on the activity. Rest of the compounds bearing substituent such as F, Cl, CF<sub>3</sub>, OCF<sub>3</sub>, OCH<sub>3</sub> and OPh at position 3 or 4 or both showed moderate or no activity with respect to standard drug against the test strains. No activity was observed in case of compounds 16–19 up to concentration of 200 µg/mL against some bacteria and fungi. Position 2 and 4 of terminal benzene ring is favorable site for high activity. The compounds 3 and 4 found to be 2.5-fold more potent than the standard drug Miconazole, while 22 and 23 exhibited comparable antifungal activity. Similar to the antibacterial activity trend, nonpolar lipophilic groups such as isopropyl or n-butyl at 4-position, compounds 20 and 21, respectively, has no major effect on the antifungal activity also [6]. Traditional Computer-assisted Quantitative Structure–Activity Relationship (QSAR) studies pioneered by C. Hansch et al. 1962 [7] have been proved to be one of the useful approaches for accelerating the drug design process [8] which helps to correlate the bioactivity of compounds with structural descriptors [9]. To gain further insights into the structure–activity relationships of these derivatives and understand the mechanism of their substitutional specificity, we have performed 2D and 3D-QSAR on 3,4-dihydropyrimidin-2(1H)-one urea derivatives of N-aryl urea by using multiple linear regression methodology (MLR) and k-Nearest Neighbor Molecular Field Analysis (kNN MFA), respectively. The significance of the QSAR models was evaluated using cross-validation tests, randomization tests and external test set prediction. The robust 2D/3D-models may be useful in further designing new candidates as potential antifungal agents prior to synthesis.

## 2. Materials and Methods

### 2.1. Selection of molecules

Data set of 23 compounds 3,4-dihydropyrimidin-2(1H)-one urea derivatives (Table 1) collected from published literature [10] were taken for the present QSAR study. The antifungal activity data of compounds were converted into log p values (MIC values µg/mL) to get the linear relationship in equation.

Molecules were rationally divided into the training set and test set based on the suggestions given by Alexander Tropsha et al. [11].

### 2.2. Molecular modeling

All computational experiments were performed using on LENOVO computer having genuine Intel Pentium Dual Core Processor and Windows XP operating system using the software Molecular Design Suite (VlifeMDS 4.3.1). [12] Structures were drawn using the 2D draw application and converted to 3D structures and subjected to an energy minimization and geometry optimization using Merck Molecular Force Field, force field and charges followed by Austin Model-1 with 10000 as maximum number of cycles, 0.01 as convergence criteria (root mean square gradient) and 1.0 as constant (medium's dielectric constant which is 1 for in vacuo) in dielectric properties. The default values of 30.0 and 10.0 Kcal/mol were used for electrostatic and steric energy cut off.

### 2.3. 2D-QSAR analysis

#### 2.3.1. Calculation of descriptors

Numbers of descriptors were calculated after optimization or minimization of the energy of the data set molecules. Various types of physicochemical descriptors were calculated: Individual (Molecular weight, H-Acceptor count, H-Donor count, X log P, slog P, SMR, Polarisability, etc.), retention index (Chi), atomic valence connectivity index (ChiV), Path count, Chi chain, ChiV chain, Chain Path count, Cluster, Path cluster, Kappa, Element count (H, N, C, S count etc.), Distance based topological (DistTopo, Connectivity Index, Wiener Index, Balaban Index), Estate numbers (SsCH<sub>3</sub>count, SdCH<sub>2</sub>count, SssCH<sub>2</sub>count, StCH count, etc.), Estate contribution (SsCH<sub>3</sub>-index, SdCH<sub>2</sub>-index, SssCH<sub>2</sub>-index, StCH index), Information theory based (Ipc, Id etc.) and Polar surface area. More than 200 alignment independent descriptors were also calculated using the following attributes. A few examples are T\_2\_O\_7, T\_N\_N\_5, T\_2\_2\_6, T\_C\_O\_1, T\_O\_Cl\_5 etc. The invariable descriptors (the descriptors that are constant for all the molecules) were removed, as they do not contribute to QSAR.

**2.3.2. Generation of training and test sets:**

In order to evaluate the QSAR model, data set was divided into training and test set using sphere exclusion, random selection and manual selection method. Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive power of the model which is not included in model generation.

**Sphere Exclusion method:** In this method initially data set were divided into training and test set using sphere exclusion method. In this method dissimilarity value provides an idea to handle training and test set size. It needs to be adjusted by trial and error until a desired division of training and test set is achieved. Increase in dissimilarity value results in increase in number of molecules in the test set.

**Random Selection Method:** In order to construct and validate the QSAR models, both internally and externally, the data sets were divided into training [90%-60% (90%, 85%, 80%, 75%, 70%, 65% and 60%) of total data set] and test sets [10%-40% (10%, 15%, 20%, 30%, 35% and 40%) of total data set] in a random manner. 10 trials were run in each case.

**Manual data selection method:** Data set is divided manually into training and test sets on the basis of the result obtained in sphere exclusion method and random selection method.

**2.3.3. Generation of 2D-QSAR models:**

Two dimensional quantitative structure activity relationship (2D QSAR) studies by means of multiple linear regression (MLR) method was performed on a series of 3,4-dihydropyrimidin-2-(1H)One urea derivatives as antifungal agents using software QSAR pro (VLife Science). MLR is a method used for modeling linear relationship between a dependent variable Y (Activity) and independent variable X (2D/3D descriptors). MLR is based on least squares. The model is fit such that sum-of-squares of differences of observed and a predicted value is minimized. MLR estimates values of regression coefficients ( $r^2$ ) by applying least squares curve fitting method. The model creates a relationship in the form of

a straight line (linear) that best approximates all the individual data points. In regression analysis, conditional mean of dependant variable (Activity) Y depends on (descriptors) X. MLR analysis extends this idea to include more than one independent variable [13].

**2.4. 3D-QSAR analysis:****2.4.1. kNN-MFA**

kNN-MFA is novel methodology, unlike conventional QSAR regression methods, this methodology can handle nonlinear relationships of molecular field descriptors with biological activity, thus making it a more accurate predictor of biological activity. Conventional correlation methods try to generate linear relationship with the activity, whereas, kNN is inherently non-linear method and is better able to explain activity trends. The kNN-MFA technique is a conceptually simple approach to pattern recognition problems. In this method, an unknown pattern is classified according to the majority of the class memberships of its k nearest neighbors in the training set. The nearness is measured by an appropriate distance metric (e.g. a molecular similarity measure, calculated using field interactions of molecular structures). The standard kNN method is implemented simply as follows: (i) calculate distances between an unknown object (u) and all the objects in the training set; (ii) select k objects from the training set most similar to object u, according to the calculated distances, (iii) classify object u with the group to which a majority of the k objects belong. An optimal k value is selected by the optimization through the classification of a test set of samples or by the leave-one out cross-validation. The variables and optimal k values are chosen using different variable selection methods as described below.

**kNN-MFA with Simulated Annealing**

Simulated Annealing (SA) is another stochastic method for function optimization employed in QSAR. Simulated annealing (SA) is the simulation of a physical process, 'annealing', which involves heating the system to a high temperature and then gradually cooling it down to a preset temperature (e.g., room temperature). During this process, the system samples possible configurations distributed according to the Boltzmann

distribution so that at equilibrium, low energy states are the most populated.

#### **kNN-MFA with Stepwise (SW) Variable Selection**

This method employs a stepwise variable selection procedure combined with kNN to optimize the number of nearest neighbors (k) and the selection of variables from the original pool as described in simulated annealing.

#### **kNN-MFA with Genetic Algorithm**

Genetic algorithms (GA) first described by Holland mimic natural evolution by modeling a dynamic population of solutions. The members of the population, referred to as chromosomes, encode the selected features. The encoding usually takes form of bit strings with bits corresponding to selected features set and others cleared. Each chromosome leads to a model built using the encoded features. By using the training data, the error of the model is quantified and serves as a fitness function. During the course of evolution, the chromosomes are subjected to crossover and mutation. By allowing survival and reproduction of the fittest chromosomes, the algorithm effectively minimizes the error function in subsequent generations.

#### **2.4.2. Creation of interaction energies**

Methyl probe with charge 1 and energy cut-off for electrostatic 10 Kcal/mol and for steric 30 Kcal/mol, dielectric constant 1 and charge type Gasteiger-marsili were used to calculate steric and electrostatic fields. The fields were computed at each lattice intersection of a regularly spaced grid of 2.0 Å within defined three-dimensional region.

#### **2.4.3. Generation of training and test sets**

In order to evaluate the QSAR model, data set was divided into training and test set using sphere exclusion, random selection and Manual selection method. Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive power of the model which is not included in model generation.

## **Results and Discussion**

### **D-QSAR models**

Different sets of 2D-QSAR models were generated using the MLR analysis in conjunction with stepwise forward-backward variable selection method. Different training and test set were constructed using sphere exclusion, random and manual selection method. Training and test set were selected if they follow the unicolon statistics, i.e. maximum of the test is less than maximum of training set and minimum of the test set is greater than of training set, which is prerequisite for further QSAR analysis. This result shows that the test is interpolative i.e., derived from the min-max range of training set. The mean and standard deviation of the training and test set provides insight to the relative difference of mean and point density distribution of the two sets. The statistical significant 2D-QSAR models for column "log p activity distribution." The selection of the best model is based on the values of  $r^2$  (squared correlation coefficient),  $q^2$  (cross-validated correlation coefficient),  $pred\_r^2$  (predicted correlation coefficient for the external test set),  $F$  (Fisher ratio) reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the  $F$ -test indicate that the model is statistically significant.  $r^2_{se}$ ,  $q^2_{se}$  and  $pred\_r^2_{se}$  are the standard errors terms for  $r^2$ ,  $q^2$  and  $pred\_r^2$  respectively. The statistically significant 2D-QSAR model is shown as follows.

**Test set:** 4, 5, 10, 11,13,19,20.

-0.5495 (T\_T\_O\_5); + 0.5543(H-Acceptor / Saas Count) +0.1528 (T\_N\_F\_6); + 0.1434 (T\_T\_O\_6) -0.0979 (T\_2\_T\_3) + 22.3512

### **Statistics**

[Optimum Components= 4; n= 15; Degree of freedom= 9;  $r^2 = 0.9759$ ;  $q^2 = 0.9314$ ;  $F$  test= 72.73;  $r^2_{se} = 5.1472$ ;  $q^2_{se} = 8.6750$ ;  $pred\_r^2 = 80.36$ .

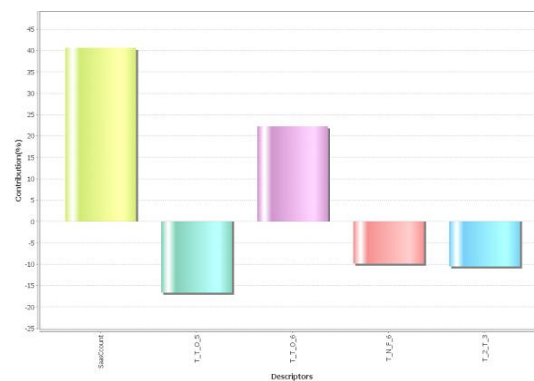
In the above QSAR equation, n is the number of molecules (Training set) used to derive the QSAR model,  $r^2$  is the squared correlation coefficient,  $q^2$  is the cross-validated correlation coefficient,  $pred\_r^2$  is the predicted correlation coefficient for the external test set,  $F$  is the Fisher ratio, reflects the ratio of the variance explained by the model and the variance due

to the error in the regression. High values of the F-test indicate that the model is statistically significant.  $r^2_{se}$ ,  $q^2_{se}$  and  $pred_r^2_{se}$  are the standard errors terms for  $r^2$ ,  $q^2$  and  $pred_r^2$  (smaller is better).

#### Interpretation of the Model:

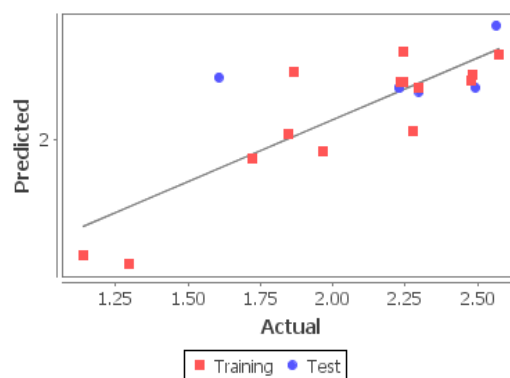
##### 2D-QSAR Model

From equation, 2D-QSAR model explains 97.56 % ( $r^2 = 0.9756$ ) of the total variance in the training set as well as it has internal ( $q^2$ ) and external ( $pred_r^2$ ) predictive ability of 80.36 %. The F test shows the statistical significance of 99.99 % of the model which means that probability of failure of the model is 1 in 10000. In addition, the randomization test shows confidence of 99.9999 (Alpha Rand Pred  $R^2 = 0.00000$ ) that the generated model is not random and hence chosen as the QSAR model. From QSAR model, negative coefficient value of T\_T\_O\_5 [count of number of double bonded atoms (i.e. any double bonded atom, T\_2) separated from carbon atom by 5 bonds], T\_T\_O\_6 [count of any bond separated from any atom by 6 bonds] on the biological activity indicated that lower values leads to good antifungal activity while higher value leads to reduced antifungal activity while positive coefficient value of H-Acceptor/Saas Count [number of hydrogen bond acceptor atoms], T\_T\_O\_5 [count of any atom (represented as T) separated from O atom by 5 bonds], T\_T\_O\_5 [count of number of double bonded atoms (i.e. any double bonded atom, T\_2) separated from carbon atom by 5 bonds], on the antifungal activity indicated that higher value leads to better antifungal activity whereas lower value leads to decrease antifungal activity. Contribution chart for model is represented in Figure 2 reveals that the descriptors H-Acceptor / Saas Count, T\_T\_O\_6, contributing 40.07 %, 25.26 % respectively. Three more descriptors T\_T\_O\_5 and T\_N\_F\_6 and T\_2\_T\_3 are contributing inversely 23.65 %, 9.79 %, and 2.56% respectively to biological activity.



**Fig. 2.** Contribution chart for 2D-QSAR model showing contribution of different descriptors.

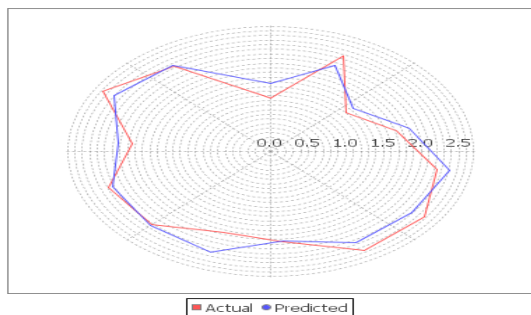
Data fitness plot for 2D-QSAR model is shown in Figure 3. The plot of observed vs predicted activity provides an idea about how well the model was trained and how well it predicts the activity of external test set.



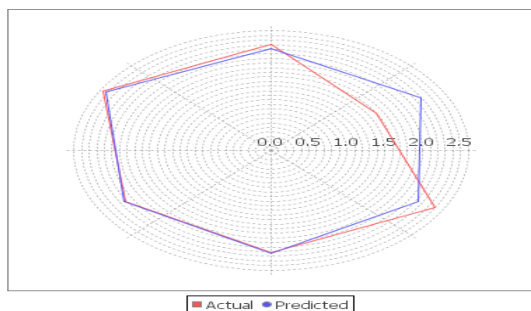
**Fig. 3.** Data fitness plot for 2D-QSAR model.

The graph of observed vs. predicted activity of training and test sets for 2D-QSAR model is shown in Figure 4. The 2D-QSAR model is able to predict the activity of training set quite well as well as external test set, providing confidence of model. Results of the observed and predicted antifungal activity are shown in Table 4.

**Training set**



**Test set**



**Fig. 4.** Graph between actual and predicted biological activity of test and training set for 2D-QSAR model.

**3.2. 3D-QSAR model**

kNN-MFA samples the steric and electrostatic fields surrounding a set of ligands and constructs 3D-QSAR models by correlating these 3D fields with the corresponding biological activities. The statistical significant results for 3D-QSAR models are given in Table 3.

The selection of the best model is based on the values of  $q^2$  (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).

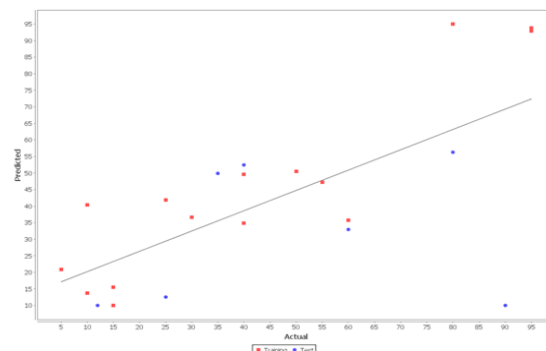
**3D-QSAR model: Test set:**  
4,5,10,11,13,19,20.  
E\_7952 (0.0702 0.0745); S\_246 (-0.1301 - 0.1222) S\_7841 (-0.0179 -0.0172)

**Statistics:**

[kNN= 2; n= 15; Degree of freedom= 12;  $q^2= 0.8032$ ;  $q^2\_se= 13.4025$ ;  $pred\_r^2= 0.5391$ ;  $pred\_r^2se= 37.0345$ .

The model 3D-QSAR explains values of k (2),  $q^2(0.8032)$ ,  $pred\_r^2(0.5391)$ ,  $q^2\_se (13.4025)$ , and  $pred\_r^2se (37.0345)$  prove that QSAR equation so obtained is statistically significant and shows the predictive power of the model

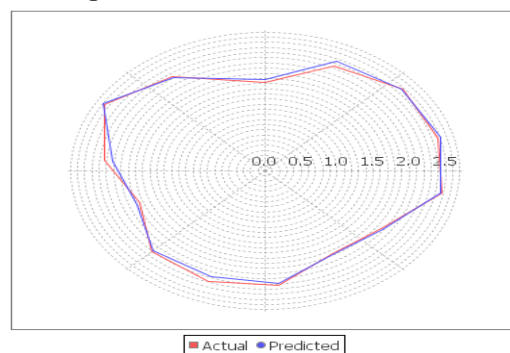
is 70.94% (internal validation). Table 4 represents the predicted antifungal activity by the model 3D-QSAR for training and test set. The data fitness plot for model 3D-QSAR is shown in Figure 5. The plot of observed vs. predicted activity provides an idea about how well the model was trained and how well it predicts the activity of the external test set.



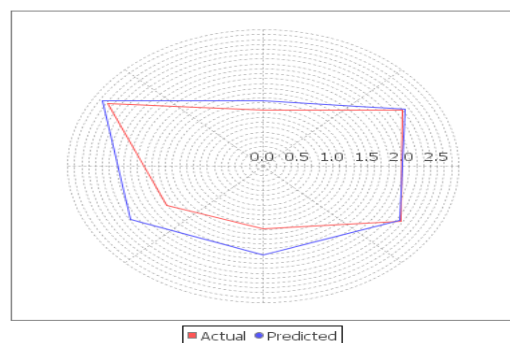
**Fig. 5.** Data fitness plot for model 3D-QSAR.

From figure 6, it can be seen that the model is able to predict the activity of the training set quite well as well as external test set, providing confidence of the model.

**Training set**



**Test set**

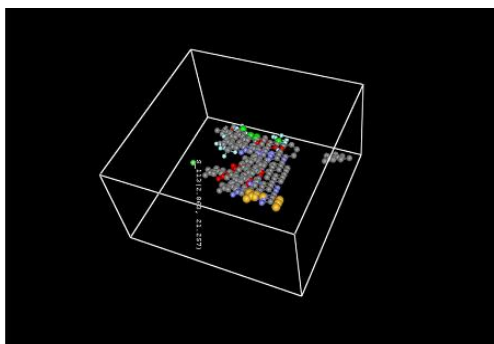


**Fig. 6.** Graph between actual and predicted biological activity of training and test set for Model-3D-QSAR.

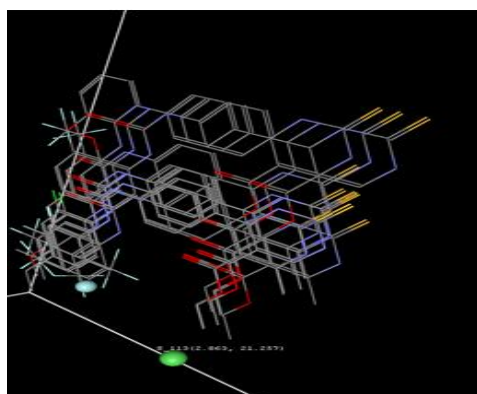
Result plot in which 3D-alignment of molecules with the important steric and electrostatic points contributing in the model 3D-QSAR with ranges of values shown in the parenthesis represented in Figure 7. It shows the relative position and ranges of the corresponding important steric and electrostatic fields in the model provides guideline for new molecule design as follows-

(a) Electrostatic field, E\_710 (0.0702, 0.0745) has positive range indicates that positive electrostatic potential is favorable for increase in the activity and hence less electronegative substituent group is preferred in that region.

(b) Steric field, S\_389 (-0.1301, -0.1222) has negative range indicates that negative steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.



**Fig. 7.** 3D-alignment of molecules (Ball and stick model) with the important steric and electrostatic points contributing model 3D-QSAR with ranges of values shown in parenthesis.



**Fig. 8.** 3D-alignment of molecules (Stick model) with the important steric and electrostatic points contributing model 3D-QSAR with ranges of values shown in parenthesis.

## Conclusion

Statistically significant 2D/3D-QSAR models were generated with the purpose of deriving structural requirements for the antifungal activities of some novel 3,4-dihydropyrimidin-2(1H)-one urea derivatives of N-aryl urea against candida albicans. The validation of 2D-QSAR models was done by the cross-validation test, randomization tests and external test set prediction. The best 2D-QSAR models indicate that the descriptors of H-Acceptor Count, T\_T\_O\_5, T\_T\_O\_6, and T\_2\_N\_6 influenced the antifungal activity. kNN-MFA investigated the substitutional requirements for the receptor-drug interaction and constructed the best 3D-QSAR models by PLSR method, providing useful information in characterization and differentiation of their binding sites. In conclusion, the information provided by the robust 2D/3D-QSAR models use for the design of new molecules and hence, this method is expected to provide a good alternative for the drug design.

## References

1. Inca S. Z. , Selma S. , Semra C., Kevser E., Bioorg Med Chem , 14 (2006) 8582.
2. Zamanova A.V., Kurbanova M.M., Rzaeva I.A., Farzaliev V. M., Allakhverdiev M.A., Russ J Appl Chem , 83 (2010) 293.
3. Hélio A.S., Carlindo B.O., Roberta B.A., Claudio M.P., Rodolpho C.B., Rodrigo C., Vanessa C.B.,Lucielli S., Cristina W. N. ,Eur J Med Chem , 41 (2006) 513.
4. Prashantha Kumar B.R., Gopu S., NasirBaig R.B., Srinivasan C., Eur J Med Chem , 14 (2009) 4192.
5. Sushilkumar S.B., &Devanand B.S., Bioorg Med Chem Lett, 14 (2004) 1733.
6. Rajesh H Tale, Atish H Rodge, Girish D Hatnapure , Ashish P Keche , Bioorg Med ChemLett , 21 (2011) 4648–4651.
7. Hansch C., Kurup A., Garg R., GaoH, Chem.-Bioinformatics and QSAR: A Review of QSAR Lacking Positive Hydrophobic Terms. Chem Rev, 101 (2001) 619–672
8. Lill M., Multi-dimensional QSAR in drug discovery, Drug Discovery Today, 12 (2007) 1013–1017.

9. Yang G., Huang X., *Curr Pharm Des*, 12 (2006) 4601–4611.
10. Tale R.H., Rodge A.H., Hatnapure G.D., Keche A.P., *Bioorg Med Chem Lett* 21 (2011) 4648–4651.
11. Golbraikh A., Shen M., Xiao Z., Xiao Y., Lee K., Tropsha A., *Journal of computer-aided molecular design*, 17, 2-4, (2003) 241-253.
12. VLife MDS 4.1.3, Molecular Design Suite, Vlife Sciences Technologies Pvt. Ltd, Pune, India, (2012) [www.vlifesciences.com](http://www.vlifesciences.com).
13. Dr. Kulkarni V.M., Dr. Bothra K.G., *Drug Design*, Nirali prakashan, (2007) 187-235.
14. [QSAR.en.wikipedia.org.in](http://QSAR.en.wikipedia.org.in)

**Table 1.** Substitutions in Structures and Biological activity of 3,4 dihydropyrimidin-2(1H)-one urea derivatives.

COMPOUNDS	-R Substitutions	ANTIFUNGAL ACTIVITY AGAINST CANDIDA ALBICANS (MIC values µg/mL)	Log P values
1	1-F	30	1.475
2	1-Cl	55	1.744
3	2-F	10	1.030
4	2-Cl	10	1.030
5	2-CF <sub>3</sub>	25	1.393
6	2-OCF <sub>3</sub>	15	1.172
7	2-OC <sub>6</sub> H <sub>5</sub>	40	1.604
8	2-F, 6-CH <sub>3</sub>	50	1.695
9	2-F, 6-CF <sub>3</sub>	55	1.746
10	2-Cl, 6-CH <sub>3</sub>	40	1.604
11	2-Cl, 6-CF <sub>3</sub>	60	1.775
12	2-Cl, 5-CF <sub>3</sub>	60	1.775
13	2-Cl, 4-CF <sub>3</sub>	35	1.541
14	2-Cl, 6-F	40	1.603
15	3-CF <sub>3</sub>	25	1.394
16	3-Cl, 4-F	80	1.905
17	3,5-F	95	1.975
18	3,4-CH <sub>3</sub>	95	1.978
19	4-F, 3-CH <sub>3</sub>	80	1.9095
20	4-isopropyl	90	1.956
21	4-butyl	5	0.692
22	4-CF <sub>3</sub>	15	1.174
23	4-OCH <sub>3</sub>	10	1.030



**Table 2.** Statistical evaluation of 2D QSAR models for “log p activity distribution”.

MODEL	r <sup>2</sup>	q <sup>2</sup>	r <sup>2</sup> se	q <sup>2</sup> se	Pred r <sub>-</sub> <sup>2</sup>	F test
2D QSAR	0.9759	0.9314	5.1472	8.6750	40.9144	72.73

**Table 3.** Statistical evaluation of 3D-QSAR model.

Model	k-Nearest Neighbor	n	Degree of freedom	q <sup>2</sup>	q <sup>2</sup> _se	Pred r <sup>2</sup>	Pred r <sup>2</sup> _se
3D-QSAR	2	15	12	0.8032	13.4025	0.5391	37.0345

**Table 4.** Actual and predicted activities for 23 compounds based on the best 2D/3D-QSAR models.

COMPOUNDS	ACTUAL VALUES (log P)	PREDICTED FOR 2D QSAR MODEL	PREDICTED FOR 3D QSAR MODEL
1	1.475	0.945	0.864
2	1.744	0.978	0.984
3	1.030	0.856	0.7778
4	1.030	0.876	0.943
5	1.393	0.908	0.982
6	1.172	0.943	0.880
7	1.604	0.934	0.876
8	1.695	0.967	0.916
9	1.746	0.977	0.9345
10	1.604	0.985	0.9823
11	1.775	0.980	0.972
12	1.775	0.981	0.8890
13	1.541	0.948	0.984
14	1.603	0.935	0.923
15	1.394	0.923	0.965
16	1.905	1.005	0.927
17	1.975	1.0704	0.988
18	1.978	1.678	0.908
19	1.9095	0.967	0.887
20	1.956	1.865	0.986
21	0.692	0.676	0.657
22	1.174	0.987	0.986
23	1.030	0.957	0.969

Source of Support: Nil. Conflict of Interest: None declared
---

\*\*\*\*\*