

## QSAR STUDY OF SUBSTITUTED N-CONTAINING HETERO CYCLES AS AN ANTI-CANCER AGENTS

Saloni Fakroddin Mulani<sup>1\*</sup>, Dr. Sachin K. Bhosale<sup>2</sup>, Dr. Nirmala Shinde<sup>3</sup>, Mohd Ziyaul haque<sup>4</sup>

<sup>1\*,2,3</sup> Department of Pharmaceutical Chemistry, Savitribai Phule Pune University, S.M.B.T. College of Pharmacy, Dhamangaon, Nashik, Maharashtra, India.  
[salonishaikh3@gmail.com](mailto:salonishaikh3@gmail.com), [sachiniper@rediffmail.com](mailto:sachiniper@rediffmail.com),  
[nirmalampharm@rediffmail.com](mailto:nirmalampharm@rediffmail.com)

<sup>4</sup> Department of Pharmaceutical Chemistry, Kalsekar Technical Campus School of Pharmacy, Navi Mumbai, Maharashtra, India. [royalziya90@gmail.com](mailto:royalziya90@gmail.com)

*\*Corresponding Author:*  
[salonishaikh3@gmail.com](mailto:salonishaikh3@gmail.com)

### ABSTRACT:

2D and 3D QSAR analysis were performed on imidazo[4,5-b] pyridine and 4-heteroarylpyrimidine derivatives for their anti-cancer activities on MCF-7 and on HCT-116 cell line. The activity of the imidazo[4,5-b] pyridine and 4-heteroarylpyrimidine on MCF-7 and HCT-116 cell line were converted into  $-\log 1/C$ . The statistically significant 2D-QSAR models for MCF-7 are  $r^2 = 0.8507$  and  $q^2 = 0.7512$  and on HCT-116 giving  $r^2 = 0.8238$  and  $q^2 = 0.7267$ . 3D QSAR study was performed using k-nearest neighbor molecular field analysis (kNN-MFA) approach for electrostatic, steric and hydrophobic fields. Three different kNN-MFA 3D QSAR methods (SW-FB, SA, and GA) were used for the development of models and tested successfully for internal ( $q^2 = 0.8625$ ,  $q^2 = 0.8550$ ) and external (predictive  $r^2 = 0.8243$ , predictive  $r^2 = 0.7582$ ) validation criteria. Thus, 3D QSAR models showed that electrostatic effects dominantly determine the binding affinities. 2D QSAR studies revealed that  $T\_N\_O\_4$  descriptor was major contributing descriptor in case of MCF-7 and  $T\_N\_N\_3$  in case of HCT-116. The 3D QSAR was performed using kNN-MFA method. 3D QSAR results suggested the importance of some molecular characteristics, which should significantly affect the binding affinities of compounds. The results derived may be useful in further designing novel anti-cancer agents.

### KEYWORDS

QSAR, Descriptors, MCF-7, HCT-116, Imidazo(4,5-b) pyridine, 4-Heteroarylpyrimidine

### INTRODUCTION

#### 1.1 Cancer:

Ahead of World Cancer Day, the World Health Organization (WHO)'s cancer agency, the International Agency for Research on Cancer (IARC), released the [latest estimates](#) of the global burden of cancer. WHO also published survey results from 115 countries, showing a majority of countries do not adequately finance priority cancer and palliative care services, as part of universal health coverage (UHC). In 2022, there were an estimated 20 million new cancer cases and 9.7 million deaths. The estimated number of people who were alive within 5 years following a cancer diagnosis was 53.5 million. About 1 in 5 people develop cancer in

their lifetime, approximately 1 in 9 men and 1 in 12 women die from the disease. The global WHO survey on UHC and cancer shows that only 39% of participating countries covered the basics of cancer management as part of their financed core health services for all citizens, ‘health benefit packages’ (HBP). The new estimates available on IARC’s [Global Cancer Observatory](#) show that 10 types of cancer collectively comprised around two-thirds of new cases and deaths globally in 2022. Data covers 185 countries and 36 cancers. Lung cancer was the most commonly occurring cancer worldwide with 2.5 million new cases accounting for 12.4% of the total new cases. Female breast cancer ranked second (2.3 million cases, 11.6%), followed by colorectal cancer (1.9 million cases, 9.6%), prostate cancer (1.5 million cases, 7.3%), and stomach cancer (970 000 cases, 4.9%). For men, prostate and colorectal cancers were the second and third most commonly occurring cancers, while liver and colorectal cancers were the second and third most common causes of cancer death. For women, lung and colorectal cancer were second and third for both the number of new cases and of deaths. [Cervical cancer](#) was the eighth most commonly occurring cancer globally and the ninth leading cause of cancer death, accounting for 661 044 new cases and 348 186 deaths. It is the most common cancer in women in 25 countries, many of which are in sub-Saharan Africa. Even while recognizing varying incidence levels, cervical cancer can be eliminated as a public health problem, through the scale-up of the [WHO Cervical Cancer Elimination Initiative](#).<sup>1</sup>

### 1.2 HCT-116 Cell Line:

HCT-116 cell line is a widely studied human colorectal carcinoma cell line that was derived from a primary tumor of the colon. It was originally established in 1981 by J. Fogh et al. from the dukes’ type C colorectal carcinoma tissue of a 44-year-old female patient. The HCT-116 cell line is commonly used in cancer research due to its relevance to colorectal cancer, which is one of the leading causes of cancer-related deaths worldwide.<sup>2</sup>

### 1.3 MCF-7 Cell Line:

MCF-7 is another widely studied cell line used in cancer research. It is derived from the breast tissue of a 69-year-old woman with metastatic breast adenocarcinoma in 1970. MCF-7 cell line serves as a valuable tool for studying various aspects of breast cancer biology, including hormone receptor signaling, drug response, metastasis, and potential therapeutic interventions. Its widespread use has contributed significantly to our understanding of breast cancer and the development of novel treatment strategies.<sup>3</sup>

### 1.4 QSAR:

#### (a) 2D QSAR:

Affinity correlates with a structural pattern (e.g., chemical connectivity). 2D methodologies do not consider the 3D structure of a molecule directly. Instead, the molecule is represented by a set of molecular descriptors, numerical values characterizing various aspects of molecular structure. Together with the observed activity, a predictive model is built. It should be noted, that even though some descriptors are based on 3D coordinates, the method as a whole considers only the observed property and the descriptors, and hence is 2D in nature.<sup>4</sup>

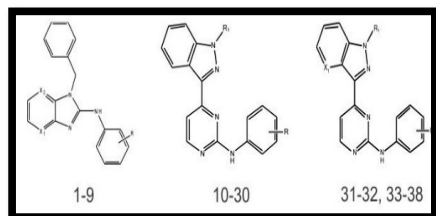
#### (b) 3D QSAR:

Affinity correlates with the three-dimensional structure and receptor modeling. QSAR is based on the interactions of the ligands with a 3D receptor surrogate. 3D QSAR should be used only where the conformation can be identified using the 3D structure of the target protein or when very rigid structures are analyzed.<sup>5</sup>

## EXPERIMENTAL SECTION

## 2.1. Chemistry

We have selected 38 molecule series from ‘Synthesis and biological evaluation of imidazo[4,5-b] pyridine and 4-heteroaryl-pyrimidine derivatives as anti-cancer agents’ having anticancer activity against Colon cancer cell line HCT-116 and Breast Cancer cell line MCF-7.<sup>6</sup> The chemical structures these derivatives are reported in Table 1.



NO.	Structure				Biological Activity			
	X <sub>1</sub>	X <sub>2</sub>	R <sub>1</sub>	R	IC <sub>50</sub> HCT-116)	pIC <sub>50</sub> HCT-116)	IC <sub>50</sub> MCF-7)	pIC <sub>50</sub> (MCF-7)
1	N	CH	-	H	64.74	1.8111	52.47	1.7199
2	N	CH	-	p-OH	72.62	1.8610	52.49	1.7200
3	N	CH	-	p-Cl	59.47	1.7742	48.03	1.6815
4	N	CH	-	m-Cl	36.33	1.8016	43.47	1.6381
5	N	CH	-	m-Br	77.66	1.8901	78.66	1.8957
6	N	CH	-	p-OMe	87.33	1.9411	59.08	1.7714
7	N	CH	-	m-NO <sub>2</sub>	41.71	1.6202	24.24	1.3845
8	N	CH	-	-NO <sub>2</sub> , p-NO <sub>2</sub>	40.36	1.6059	27.84	1.4446
9	CH	N	-	H	60.21	1.7796	51.81	1.7144
10	-	-	Me	p-OH	8.43	0.9258	8.91	0.9498
11	-	-	Me	m-OH	17.21	1.2357	28.16	1.4496
12	-	-	Me	p-Me	8.64	0.9365	6.10	0.7853
13	-	-	Me	m-Me	9.42	0.9740	9.15	0.9614
14	-	-	Me	p-Cl	8.73	0.9410	7.83	0.8937
15	-	-	Me	m-Cl	6.40	0.8061	5.97	0.7759
16	-	-	Me	p-NO <sub>2</sub>	8.22	0.9148	11.84	1.0733
17	-	-	Me	m-NO <sub>2</sub>	0.49	-0.3098	0.58	-0.2365
18	-	-	Me	p-SO <sub>2</sub> NH <sub>2</sub>	2.28	0.3579	0.94	-0.0268
19	-	-	Me	m-SO <sub>2</sub> NH <sub>2</sub>	1.98	0.2966	2.27	0.3560
20	-	-	Et	p-SO <sub>2</sub> NH <sub>2</sub>	2.05	0.3117	2.28	0.3579
21	-	-	Et	m-SO <sub>2</sub> NH <sub>2</sub>	9.01	0.9547	7.26	0.8609
22	-	-	Et	p-NO <sub>2</sub>	4.06	0.6085	23.96	1.3794
23	-	-	Et	m-NO <sub>2</sub>	2.82	0.4502	2.31	0.3636
24	-	-	Et	p-OH	5.29	0.7234	5.78	0.7619
25	-	-	Et	m-OH	36.59	1.5633	33.50	1.5280
26	-	-	n-Pr	p-OH	5.45	0.7363	3.34	0.5237
27	-	-	n-Pr	p-SO <sub>2</sub> NH <sub>2</sub>	6.74	0.8286	0.59	-0.2291
28	-	-	n-Pr	m-SO <sub>2</sub> NH <sub>2</sub>	7.95	0.9003	4.26	0.6294
29	-	-	-methylpyridine	p-SO <sub>2</sub> NH <sub>2</sub>	5.13	0.7101	2.63	0.4199
30	-	-	-methylpyridine	m-SO <sub>2</sub> NH <sub>2</sub>	8.28	0.9180	5.27	0.7218
31	CH	-	H	p-SO <sub>2</sub> NH <sub>2</sub>	0.59	-0.2291	1.87	0.2718

32	CH	-	H	m-SO <sub>2</sub> NH <sub>2</sub>	0.75	-0.1249	5.56	0.7450
33	N	-	H	m-SO <sub>2</sub> NH <sub>2</sub>	14.63	1.1652	6.93	0.8407
34	N	-	H	p-SO <sub>2</sub> NH <sub>2</sub>	8.04	0.9052	7.02	0.8463
35	N	-	Me	m-SO <sub>2</sub> NH <sub>2</sub>	24.84	1.3951	19.95	1.2999
36	N	-	Me	p-SO <sub>2</sub> NH <sub>2</sub>	8.84	0.9464	8.64	0.9365
37	N	-	Et	m-SO <sub>2</sub> NH <sub>2</sub>	65.14	1.8138	73.23	1.8646
38	N	-	Et	p-SO <sub>2</sub> NH <sub>2</sub>	6.34	0.8020	0.88	-0.0555

**Table 1.** Chemical structures and IC<sub>50</sub> values of Imidazo(4,5-b) pyridine and 4-Heteroarylpyrimidine derivatives.

## 2.2. QSAR Softwares and Descriptors

### 2.3.1. QSAR Softwares:

Softwares available for QSAR study are as follows:

1. VLife MDS 4.6
2. AutoQSAR
3. BuildQSAR
4. EasyQSAR

Among these VLife MDS 4.6 is used for designing 2D and 3D QSAR models.

### 2.3.2. Descriptors:

The descriptors are enlisted in Table 2.

Type	Descriptors
Hydrophobic Parameters	Partition coefficient : log P
	Hansch's substitution constant; $\pi$
	Hydrophobic fragmental constant; f, f'
	Distribution coefficient; log D
	Apparent log P
	Capacity factor in HPLC; log k', log k <sub>w</sub>
Electronic Parameters	Solubility parameter; log S
	Hammett constant; $\sigma$ , $\sigma^+$ , $\sigma^-$
	Taft's inductive (polar) constant; $\sigma^*$
	Swain and Lupton field parameter
	Ionization constant; pK <sub>a</sub> , $\Delta$ pK <sub>a</sub>
	Chemical shifts: IR, NMR
Steric Parameters	Taft's steric parameter; E <sub>s</sub>
	Molar volume; MV
	Van der waals radius
	Van der waals volume
	Molar refractivity; MR
	Parachor
Quantum chemical descriptors	Sterimol
	Atomic net charge; Q <sup>+</sup> , Q <sup>-</sup>
	Superdelocalizability
Spatial Descriptor	Energy of highest occupied molecular orbital; E <sub>HOMO</sub>
	Energy of lowest unoccupied molecular orbital; E <sub>LUMO</sub>
	Jurs descriptors, Shadow indices, Radius of Gyration, Principle moment of inertia

**Table 2.** Descriptors

## RESULTS AND DISCUSSION

### ▪ Step 1:

Convert the molecules 2D to 3D (mol2 file).

### ▪ Step 2:

Energy Minimization

### ▪ Step 3:

Generation of Descriptor file are shown in Fig 1.

VLifeMDS WorkSheet - 2D QSAR [E:/Saloni Ph D work Final/2D QSAR/HCT 116/HCT 2d descriptors.qsr]

File Edit Tools Graph QSAR Tools Regression Methods Advanced Methods Help

	log <sub>i</sub> IC50	Mol.Wt.	-AcceptorCou	+DonorCou	stableBondCc	XlogP	
1.mol2	1.811	300.363	1	1	4	4.004	4.2
10.mol2	0.926	317.35	2	2	5	3.428	3.4
11.mol2	1.236	317.35	2	2	5	3.428	3.4
12.mol2	0.936	315.378	1	1	5	4.083	4.0
13.mol2	0.974	315.378	1	1	5	4.083	4.0
14.mol2	0.941	335.796	1	1	4	4.421	4.4
15.mol2	0.806	335.796	1	1	4	4.421	4.4
16.mol2	0.915	346.348	4	1	5	3.674	3.6
17.mol2	-0.31	346.348	4	1	5	3.674	3.6
18.mol2	0.358	380.43	5	2	6	1.348	2.8
19.mol2	0.297	380.43	5	2	6	1.348	2.8
2.mol2	1.861	316.362	2	2	5	3.63	3.9
20.mol2	0.312	394.457	5	2	7	1.761	3.4
21.mol2	0.955	394.457	5	2	7	1.761	3.4
22.mol2	0.609	360.375	4	1	6	4.087	4.2
23.mol2	0.45	360.375	4	1	6	4.087	4.2
24.mol2	0.723	331.377	2	2	6	3.841	4.0

row 1, col 1 Rows : 38 Cols : 225

**Figure 1.** Descriptors for HCT-116 Cell Line and MCF-7 Cell Line

- **Step 4:**  
Data Selection- **Random Selection** method is used. In Random selection molecules are divided in 70:30 ratio in training and test set.
- **Step 5:**  
Selection of Regression Method. Selection of Regression Method: -
  1. Multiple regression.
  2. Principal component regression; coupled with stepwise variable selection.
- **Step 6:**  
Data Analysis.

Non-cross validated $r^2$	$\geq 0.6$
Cross validation $q^2$ (CV)	$\geq 0.5$
Coefficient of determination for external test set (pred $r^2$ )	$\geq 0.6$
F-value	$F_{\text{tab}}$
Standard error of estimate (pred $r^2$ se)	Min

**Table 4. Standard Values for 2D QSAR Parameter**

## 2D QSAR ON HCT-116

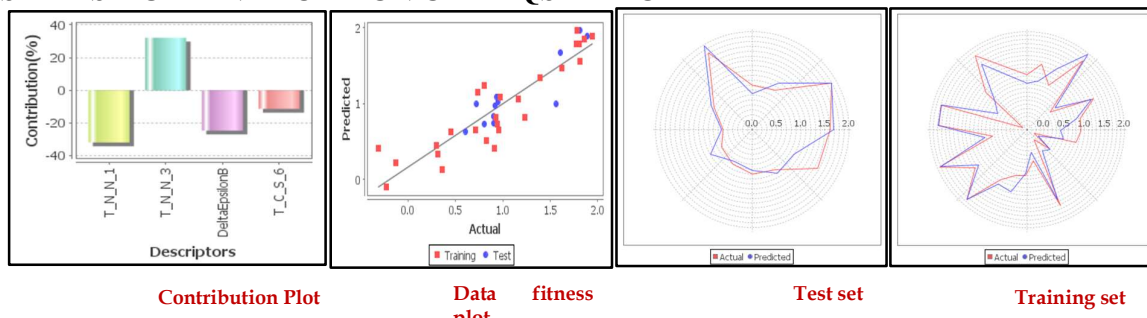
i. Models: -

No.	r <sup>2</sup>	q <sup>2</sup>	r <sup>2</sup> se	q <sup>2</sup> se	F test	Pred r <sup>2</sup>	Pred r <sup>2</sup> se	n	DOF
1	0.8238	0.7267	0.3056	0.3807	24.5500	0.8073	0.2129	26	21
2	0.8124	0.7017	0.2914	0.3675	22.7409	0.6449	0.3661	26	21

3	0.8042	0.7213	0.3056	0.3645	21.5605	0.7189	0.3099	26	21
---	--------	--------	--------	--------	---------	--------	--------	----	----

**Table 5. FINAL MODEL FOR 2D QSAR on HCT-116 CELL LINE**

**ii.QSAR parameters: -**  
**STATISTICAL EVALUATION OF 2D-QSAR MODEL-1**



**INTERPRETATION OF THE MODEL -1**

➤ **Model-**  
Training Set Size = 26  
Test Set Size = 12 (1,12,15,22,14,25,28,30,34,36,5,8)

➤  **$pIC_{50}$  -**  
-0.8983 (T\_N\_N\_1) +0.5222 (T\_N\_N\_3)- 19.9569(DeltaEpsilonB)-0.3197 (T\_C\_S\_6)3.3264

➤ **STATISTICS:**

- **Negative coefficient value of T\_N\_N\_1** on the biological activity indicated that **lower value** leads to better inhibitory activity.
- **Positive coefficient value of T\_N\_N\_3** on the biological activity indicated that **higher value** leads to better inhibitory activity.
- **Negative coefficient value of DeltaEpsilonB** on the biological activity indicated that **lower value** leads to better inhibitory activity.
- **Negative coefficient value of T\_C\_S\_6** on the biological activity indicated that **lower value** leads to better inhibitory activity.

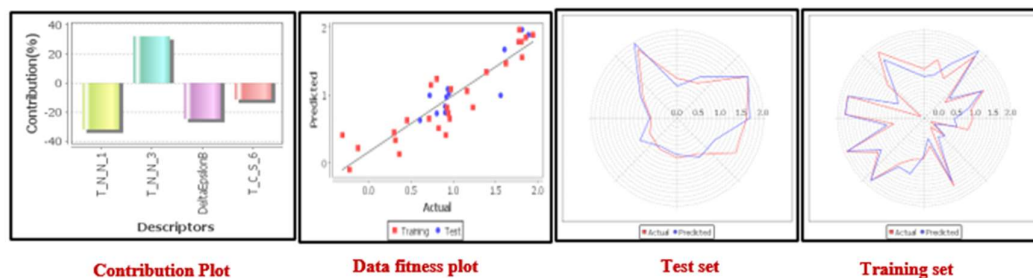
**2D QSAR ON MCF-7**

**i.Models: -**

No.	r <sup>2</sup>	q <sup>2</sup>	r <sup>2</sup> <sub>se</sub>	q <sup>2</sup> <sub>se</sub>	F test	Pred r <sup>2</sup>	Pred r <sup>2</sup> <sub>se</sub>	n	DOF
1	0.8507	0.7512	0.2364	0.3052	22.7893	0.7091	0.4075	26	20
2	0.8040	0.6618	0.2753	0.3616	16.4078	0.7737	0.3510	26	20
3	0.8268	0.7235	0.2612	0.3301	19.0984	0.6150	0.4480	26	20

**Table 6. FINAL MODEL FOR 2D QSAR on MCF-7 CELL LINE**

ii.QSAR parameters: -  
STATISTICAL EVALUATION OF 2D-QSAR MODEL-1



INTERPRETATION OF THE MODEL -1

➤ Model-

Training Set Size = 26 Test Set Size = 12 (1,24,27,3,32,34,36,37,38,6,8,9)

➤  $pIC_{50}$  -

-0.1666 (T\_N\_N\_6)-1.5334 (SaaaCE-index) +1.0326(T\_N\_O\_4)-0.7646 (T\_C\_S\_6) -0.2669 (T\_2\_O\_6).

➤ STATISTICS:

- **Negative coefficient value of T\_N\_N\_6** on the biological activity indicated that **lower value leads** to better inhibitory activity.
- **Negative coefficient value of SaaaCE-index** on the biological activity indicated that **lower value leads** to better inhibitory activity.
- **Positive coefficient value of T\_N\_O\_4** on the biological activity indicated that **higher value leads** to better inhibitory activity.
- **Negative coefficient value of T\_C\_S\_6** on the biological activity indicated that **lower value leads** to better inhibitory activity.
- **Negative coefficient value of T\_2\_O\_6** on the biological activity indicated that **lower value leads** to better inhibitory activity.
- **Negative coefficient value of T\_C\_S\_6** on the biological activity indicated that **lower value leads** to better inhibitory activity.

3D QSAR on HCT-116

k nearest neighbor	$\geq 0.6$
Cross validation $q^2$ (CV)	$\geq 0.5$
Degree of Freedom	+ve
Coefficient of determination for external test set ( $\text{pred } r^2$ )	$\geq 0.6$
F-value	High
Standard error of estimate ( $\text{pred } r^2 \text{ se}$ )	Min

Table 7. Standard Values for 3D QSAR Parameters

i.Models: -

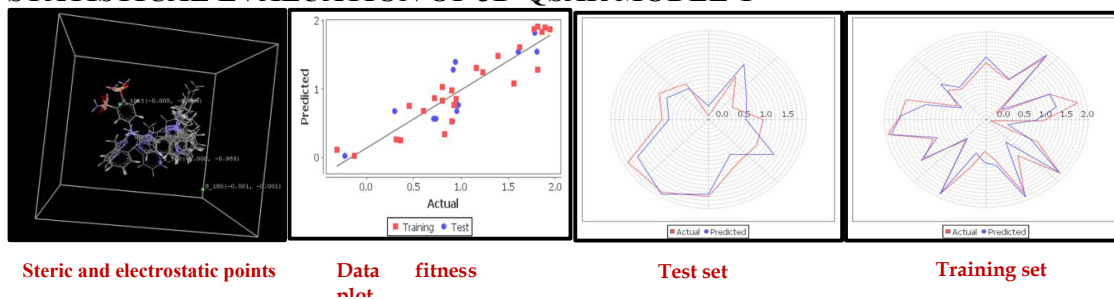
No.	kNN	DOF	$q^2$	$q^2 \text{ se}$	$\text{pred } r^2$	$\text{pred } r^2 \text{ se}$
1	2	22	0.8550	0.2380	0.7582	0.2918
2	2	23	0.8508	0.2331	0.7530	0.3230
3	2	23	0.8150	0.2606	0.7944	0.2921

Table 8. FINAL MODEL FOR 3D QSAR on HCT-116 CELL LINE



## ii.QSAR parameters: -

### STATISTICAL EVALUATION OF 3D-QSAR MODEL-1



### INTERPRETATION OF THE MODEL -1

➤ **Model-**  
 Training Set Size = 26  
 Test Set Size = 12 (13,14,16,19,21,26,29,30,31,4,8,9)

➤  **$pIC_{50}$  -**  
 E\_500+S\_1013+S\_180

#### ➤ STATISTICS:

Electrostatic field, E\_500 (-1.0000 -0.9550) Negative electrostatic potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.  
 Steric field, S\_1013 (-0.0050 -0.0050) Negative Steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.  
 Steric field, S\_180 (-0.0010 -0.0010) Negative Steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

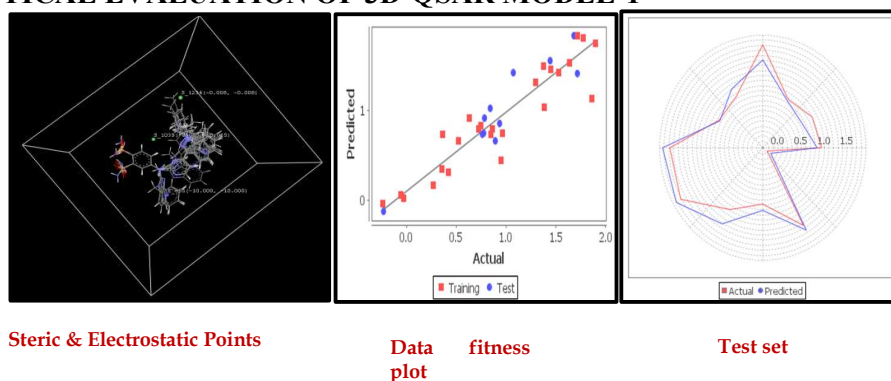
## 3D QSAR on MCF-7

### i.Models:-

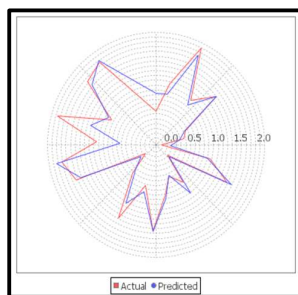
No.	kNN	DOF	q2	q2_se	pred_r2	pred_r2se
1	2	22	0.8625	0.2353	0.8943	0.1853
2	2	23	0.8001	0.2827	0.8259	0.2385
3	2	23	0.8207	0.2435	0.6329	0.4430

Table 9. FINAL MODEL FOR 3D QSAR on MCF-7 CELL LINE

### STATISTICAL EVALUATION OF 3D-QSAR MODEL-1







Training set

### INTERPRETATION OF THE MODEL -1

➤ **Model-**  
 Training Set Size = 26  
 Test Set Size = 12 (12,14,15,16,2,24,27,3,33,36,8,9)

➤  **$pIC_{50}$**  -  
 $E_{485} (-10.0000 -10.0000) + S_{1039} (-0.0253 -0.0247) + S_{1234} (-0.0076 -0.0076)$

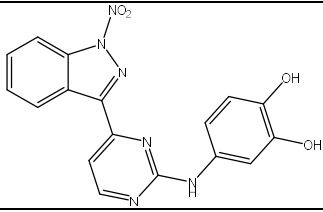
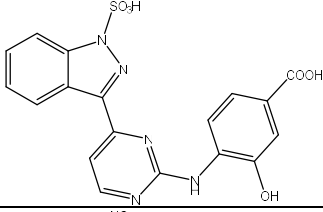
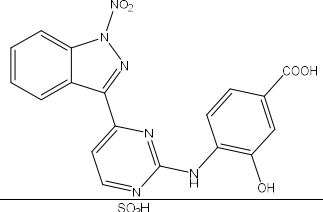
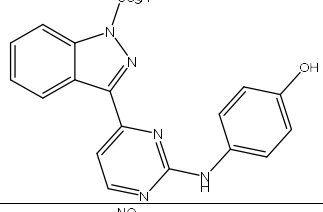
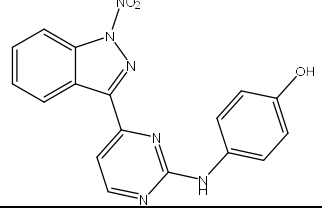
**STATISTICS:** Electrostatic field,  $E_{485} (-10.0000 -10.0000)$  Negative electrostatic potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

Steric field,  $S_{1039} (-0.0253 -0.0247)$  Negative Steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

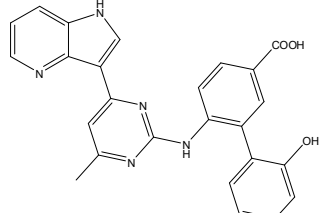
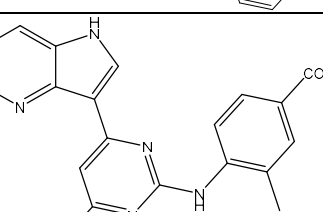
Steric field,  $S_{1234} (-0.0076 -0.0076)$  Negative Steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region. **NEWLY DESIGNED MOLECULES**

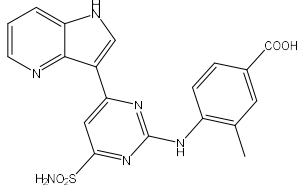
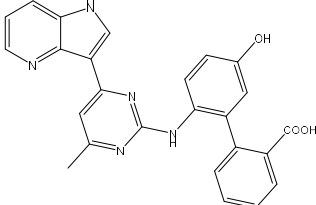
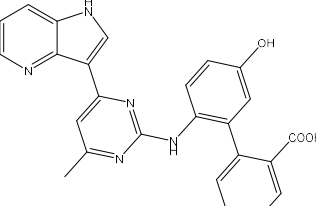
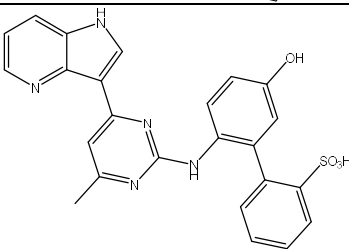
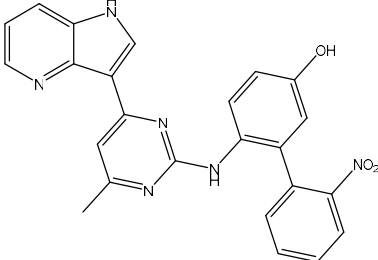
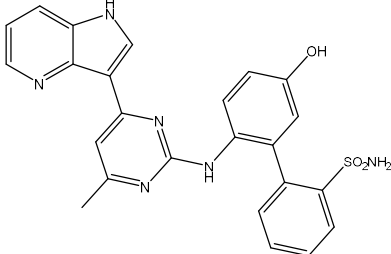
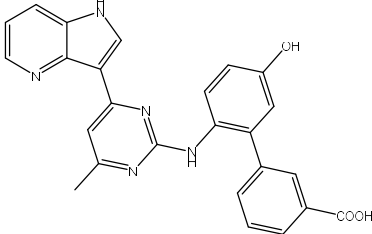
#### 1. HCT-116 Cell Line

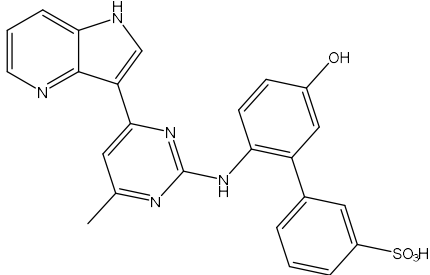
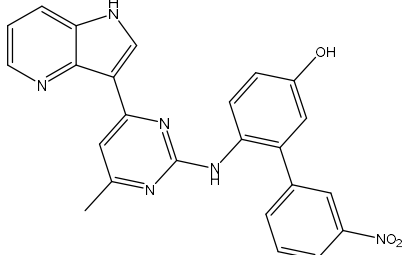
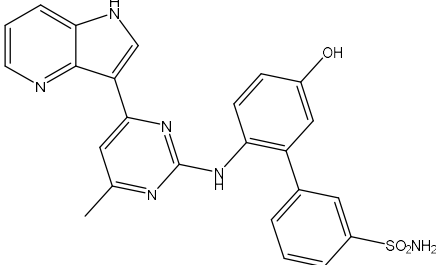
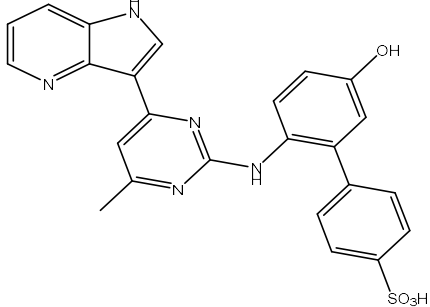
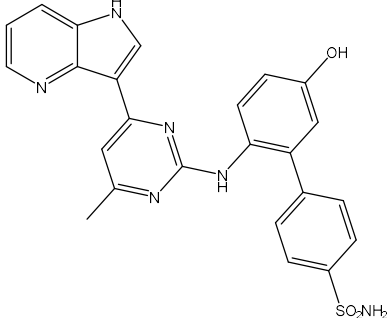
NO.	STURCTURES	Predicated $IC_{50}$	
		Log	Anti-log
1		-0.9884	0.1026
2		-1.1530	0.0703
3		-0.7982	0.1591

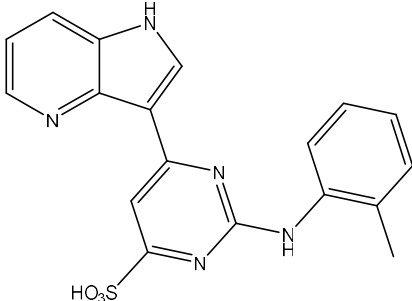
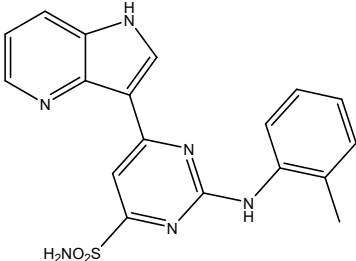
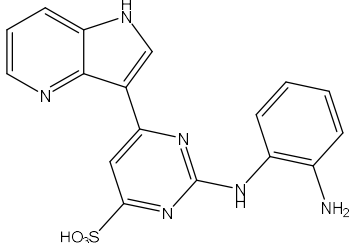
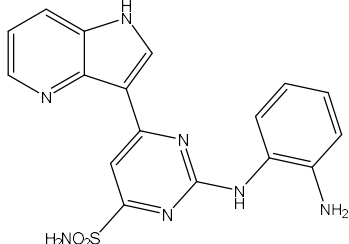
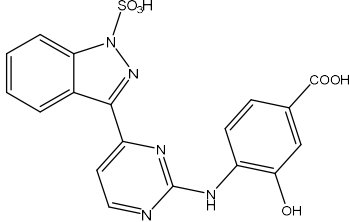
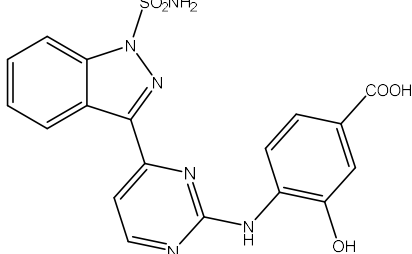
4		-0.95	0.1122
5		-0.9885	0.1026
6		-1.153	0.0703
7		-0.707	0.1961
8		-0.8557	0.1394

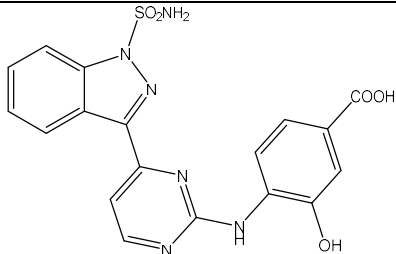
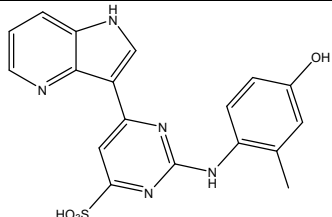
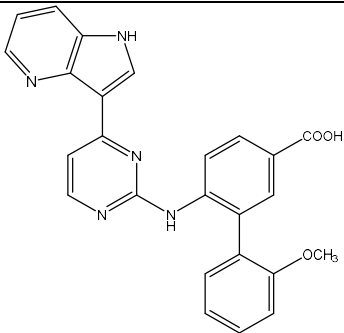
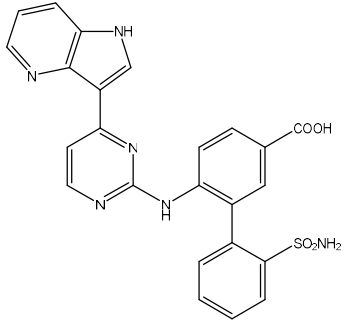
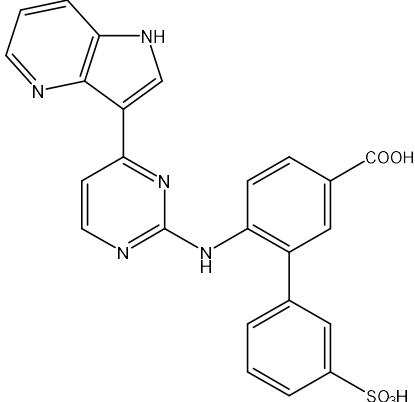
## 2. MCF-7 Cell Line

NO.	STURCTURES	Predicated IC <sub>50</sub>	
		Log	Anti-log
1		-0.5904	0.2567
2		-2.8343	0.0014

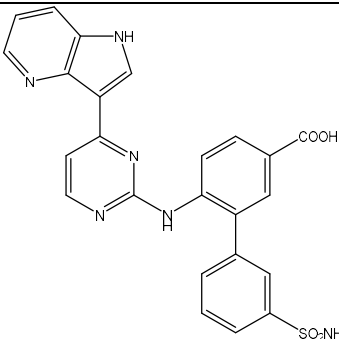
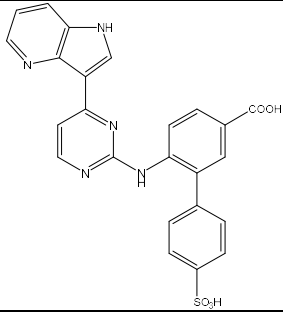
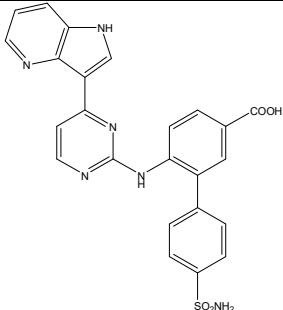
3		-2.1440	0.0071
4		-0.8573	0.1388
5		-0.8573	0.1388
6		-2.8466	0.0014
7		-0.8298	0.1479
8		-2.5188	0.003
9		-0.6162	0.2419

10		-2.6157	0.0024
11		-0.7593	0.174
12		-2.1156	0.0076
13		-1.8438	0.0143
14		-1.6061	0.0247

15		-3.0156	0.0007
16		-2.3253	0.0047
17		-2.9471	0.0011
18		-2.2568	0.0055
19		-1.3178	0.0480
20		-0.9980	0.1004

21		-1.2747	0.0531
22		-3.1908	0.0006
23		-0.6321	0.2337
24		-3.1839	0.0006
25		-2.7830	0.00164



26		-2.283	0.0052
27		-2.0111	0.0097
28		-1.7734	0.0168

## CONCLUSION

Statistically significant 2D/3D-QSAR models were generated with the purpose of deriving structural requirements for the inhibitory activities of some imidazo pyridine derivatives as an anti-cancer agent. 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  $T\_N\_N\_1$ ,  $T\_N\_N\_3$ , DeltaEpsilonB and  $T\_C\_S\_6$  (HCT-116) and  $T\_N\_N\_6$ , SaaaCE-index,  $T\_N\_O\_4$ ,  $T\_C\_S\_6$ , and  $T\_2\_O\_6$  (MCF-7) and Mominertia X Samosthydrophilic Distance,  $T\_N\_N\_3$ ,  $T\_C\_S\_6$  HCT-116 influenced the inhibition activity. The information provided by the robust 2D/3D- QSAR models use for the design of new molecules. Hence, this method is expected to provide a good alternative for the drug design.

## ACKNOWLEDGEMENTS

The authors express their deep thanks to SMBT College of Pharmacy and VLife Sciences for providing facilities to conduct this work.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this paper.

## REFERENCES

1. <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services>.
2. Elemam NM, Al-Jaderi Z, Hachim MY, Maghazachi AA. HCT-116 colorectal cancer cells secrete chemokines which induce chemoattraction and intracellular calcium mobilization in NK92 cells. *Cancer Immunol Immunother*. 2019 Jun;68(6):883-895. doi: 10.1007/s00262-019-02319-7. Epub 2019 Mar 7. PMID: 30847498.
3. Comşa Ş, Cimpean AM, Raica M. The story of MCF-7 breast cancer cell line: 40 years of experience in research. *Anticancer research*. 2015 Jun 1;35(6):3147-54.
4. Ramalakshmi N, Manimegalai P, Bhandare RR, Kumar SA, Shaik AB. 2D-Quantitative structure activity relationship (QSAR) modeling, docking studies, synthesis and in-vitro evaluation of 1, 3, 4-thiadiazole tethered coumarin derivatives as antiproliferative agents. *Journal of Saudi Chemical Society*. 2021 Jul 1;25(7):101279.
5. Rasulev B. Recent Developments in 3D QSAR and Molecular Docking Studies of Organic and Nanostructures. *Handbook of Computational Chemistry*. 2016 Dec 9:2133–61. doi: 10.1007/978-3-319-27282-5\_54. PMCID: PMC7123761.
6. Lukasik, P. M., Elabar, S., Lam, F., Shao, H., Liu, X., Abbas, A. Y., & Wang, S. (2012). Synthesis and biological evaluation of imidazo[4,5-b]pyridine and 4-heteroaryl-pyrimidine derivatives as anti-cancer agents. *European Journal of Medicinal Chemistry*, 57, 311–322.
7. QSARpro, VLife Sciences Technologies, Pvt. Ltd. Pune, India. Available from: <http://www.vlifesciences.com> .2023.
8. Kulkarni VM and Bothara KG, “Drug Design”; 3rd ed. Pune: Nirali prakashan;2008.
9. K. Khanna, R.M. Weier, K.T. Lentz, L. Swenton, D.C. Lankin, Facile, regioselective syntheses of N-alkylated 2,3-diaminopyridines and imidazo[4,5-b] pyridines, *J. Org. Chem.* 60 (1995) 960e965.
10. Akhtar, J.; Khan, A.A.; Ali, Z.; Haider, R.; Yar, M.S. Structure-activity relationship (SAR) study and design strategies of nitrogen-containing heterocyclic moieties for their anticancer activities. *Eur. J. Med. Chem.* 2017, 125, 143–189.
11. Zhang, L.; Peng, X.M.; Damu, G.L.V.; Geng, R.X.; Zhou, C.H. Comprehensive review in current developments of imidazole-based medicinal chemistry. *Med. Res. Rev.* 2014, 34, 340–437.
12. P.Lan,N. Chen, W.M. Chen, Molecular modeling studies on imidazo[4,5-b] pyridine derivatives as Aurora A kinase inhibitors using 3D-QSAR and docking approaches, *Eur. J. Med. Chem.* 46 (2011) 77e94. [j.ejmech.2010.10.017](https://doi.org/10.1016/j.ejmech.2010.10.017).
13. A.M. Sajith, K.K.A. Khader, N. Joshi, M.N. Reddy, M. Syed Ali Padusha, H.P. Nagaswarupa, M. Nibin Joy, Y.D. Bodke, R.P. Karuvalam, R. Banerjee, A. Muralidharan, P. Rajendra, Design, synthesis and structure-activity relationship (SAR) studies of imidazo[4,5-b]pyridine derived purine isosteres and their potential as cytotoxic agents, *Eur. J. Med. Chem.* 89 (2015) 21e31, <https://doi.org/10.1016/j.ejmech.2014.10.037>.