



REVIEW ARTICLE

QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP AND DOCKING STUDIES ON A SERIES OF PURINE-BASED HYDROXAMIC ACID DERIVATIVES AS HDAC1 INHIBITORS WITH ANTITUMOR ACTIVITIES

*Dr. Renu Kumari

Assistant Professor, Chemistry Department, D.B.S.P.G. College, Govind Nagar, Kanpur, India

ARTICLE INFO

Article History:

Received 20th January, 2025
Received in revised form
19th February, 2025
Accepted 26th March, 2025
Published online 26th April, 2025

Key words:

QSAR, molecular refractivity, Regression analysis, hydrophobicity, HDAC1 inhibitors with antitumor activities, inhibitory activity (pIC₅₀).

*Corresponding author:
Renu Kumari

Copyright©2025, Renu Kumari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Dr. Renu Kumari. 2025. Quantitative Structure Activity Relationship and Docking Studies on a Series of Purine-Based Hydroxamic Acid Derivatives as HDAC1 inhibitors with Antitumor activities". International Journal of Current Research, 17, (04), 32303-32321.

ABSTRACT

A series of 37 compounds of Purine-Based Hydroxamic Acid Derivatives were taken from the literature with inhibitory activity pIC₅₀. In the present study, we try to model some drug molecules acting as HDAC1 inhibitors with antitumor activities. For QSAR studies, out of 30 compounds, compounds (75%) were selected for the training set by random selection, using Statistica Data miner software, for the generation of the model, and the remaining seven compounds (25%) were used for the test set. The multiple linear regression method was applied to select the descriptors. The two parametric model were found to be the best which gave the variance of more than 90% (R²=0.902). Our results have indicated that the negative values of the descriptors ATSC6i and C-008 refer to a decrease in centered broto Moreau autocorrelation of lag 6 weighted by ionization potential, hydrophobicity and molecular refractivity will enhance the activity of the molecule. We also predicted some new compounds, as reported in Table 6, where each compound has a higher activity value than any compound in the existing series (Table 1).

INTRODUCTION

Molecular modeling encompasses a range of computational techniques used to represent and simulate the structures and behaviors of molecules. This interdisciplinary field combines chemistry, biology, physics, and computer science principles to understand and predict molecular properties and interactions, facilitating advances in drug discovery, material science, and biochemistry. The Key Techniques in Molecular Modeling comprises Quantum Mechanics (QM), Molecular Mechanics, molecular dynamics, and docking studies. Density Functional Theory (DFT) and Hartree-Fock (HF) are the widely used Quantum Mechanics methods to calculate electronic structures and properties of molecules at an atomic level. Molecular Mechanics employs force fields to model molecular systems, treating atoms as balls and bonds as springs. This approach is computationally less demanding than QM and is used to simulate larger biomolecular systems like proteins and nucleic acids. Popular force fields include AMBER, CHARMM, and OPLS. Molecular Dynamics simulations involve solving Newton's equations of motion for a system of particles to study the time-dependent behavior of molecules. MD provides insights into conformational changes, stability, and interactions over time, which are crucial for understanding protein folding and drug-receptor interaction. Another technique of molecular

modeling is docking. Molecular docking predicts the preferred orientation of a molecule (ligand) when bound to a target protein, facilitating drug design by identifying potential binding sites and affinities. Tools like AutoDock and Schrödinger's Glide are widely used in this context. Homology Modeling is another molecular modeling technique used when the 3D structure of a protein is unknown, and homology modeling builds a model based on the known structure of a related homologous protein. This technique is essential for studying proteins that are difficult to crystallize.

Applications in Drug Discovery: Molecular modeling is pivotal in drug discovery, allowing researchers to efficiently design and optimize drug candidates. By mimicking the interaction between drug molecules and their biological targets, molecular modeling helps predict efficacy, selectivity, and potential side effects, significantly reducing the time and cost associated with experimental procedures. Virtual Screening: High-throughput virtual screening of large compound libraries against target proteins helps identify promising drug candidates.

- **Lead Optimization:** Iterative cycles of modeling and experimental validation refine the chemical structure of

lead compounds to improve their pharmacological properties.

- **ADME/Tox Predictions:** Computational models predict the Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADME/Tox) profiles of compounds, guiding the selection of candidates with favorable drug-like properties (Leach, 2001; Jensen, 2017; Jorgensen, 2004; Karplus, 2005; Ekins, 2005; Ferreira *et al.*, 2015).

Challenges and Future Directions: While molecular modeling offers powerful tools, it faces challenges like the accuracy of force fields, the need for extensive computational resources, and the complexity of biological systems. Developments in machine learning and the integration of multi-scale modeling approaches promise to enhance the accuracy and efficiency of molecular simulations. Continued growth in hardware, such as quantum computing, and improved algorithms will further expand the capabilities of molecular modeling in scientific research and pharmaceutical development.

A class of chemicals known as histone deacetylase 1 (HDAC1) inhibitors is being studied closely in biomedical research because of its probable therapeutic uses, especially in treating cancer. The removal of acetyl groups from histone proteins by histone deacetylases (HDACs) enzymes results in chromatin condensation and transcriptional inhibition. HDAC1, a class I HDAC family member, plays a significant role in controlling differentiation, cell cycle progression, and gene expression. Since HDAC1's abnormal activity has been linked to several cancers, it is a prospective target for anticancer treatments. Histones that have been acetylated accumulated due to HDAC1 inhibitors' suppression of HDAC1's deacetylase activity. Tumor suppressor genes and other previously silenced genes can now be transcriptionally activated due to the more relaxed chromatin structure caused by this hyperacetylation of histones. Reactivating these genes can cause cell cycle arrest, decrease the growth of cancer cells, and encourage apoptosis.

Hydroxamates, benzamides, cyclic peptides, and aliphatic acids are among the various kinds of HDAC1 inhibitors. The chemical structures and ways in which they interact with the HDAC1 enzyme vary between the classes of inhibitors. To efficiently block HDAC1's enzymatic activity, hydroxamate-based inhibitors, including vorinostat (SAHA), bind to the zinc ion in the enzyme's active site. A wide range of possibilities for therapeutic development are provided by benzamide derivatives, such as entinostat, which interact with the enzyme through various molecular interactions (Bolden *et al.*, 2006; Ververis *et al.*, 2013; West, 2014; Bantscheff *et al.*, 2011; Haberland, 2009; Li, 2016; Khan, 2012). In the present study, we try to model some drug molecules acting as HDAC1 inhibitors with antitumor activities.

METHODOLOGY

Quantitative Structure-Activity Relationship (QSAR) and molecular docking are powerful drug discovery and design computational tools. Recent advances in these areas have been reported. These studies highlight the potential of QSAR and molecular docking in developing novel therapeutics (Basheerulla Shaik, 2017; Basheerulla Shaik, 2017; Izhar Ahmad, 2016; Neelu Singh; Basheerulla Shaik, 2016; Shweta Sharma, 2016). A series of 37 compounds of Purine-Based

Hydroxamic Acid Derivatives were taken from the literature (Yong Chen, 2016). All the compounds are listed in Table 1, along with their physicochemical properties and inhibition activity. For QSAR studies, out of 30 compounds, compounds (75%) were selected for the training set by random selection, using Statistica Data miner software, for the generation of the model, and the remaining seven compounds (25%) were used for the test set to evaluate the predictability of the developed model. ACD/Chem Sketch software has been used to draw all the chemical structures of the compounds listed in Table 1. 4888 descriptors, including 2D 3D, were calculated using DRAGON software. Among all the calculated physicochemical and topological descriptors, only three descriptors, as listed in Table 1, were found to be correlated with the activity. In this Table, test set compounds are marked with a superscript 'b,' and the compounds marked with the superscript 'c' acted as outliers and thus were removed in the model development. The most significant structural descriptors that were found to govern the activity of the compounds are listed in Table 2.

RESULTS AND DISCUSSION

A multiple linear regression (MLR) analysis was performed using NCSS software on the training set compounds to establish a correlation between observed activity and various calculated descriptors of the compounds. The most significant correlation achieved was as shown by Eq. (1).

$$pIC_{50} = -0.8459(\pm 0.3425)ATSC6i - 1.8363(\pm 0.2834)C-008 + 10.7061 \quad (1)$$

$n = 30, r^2 = 0.902, r^2_{cv} = 0.885, r^2_{pred} = 0.872, s = 0.311, F = 124.235$

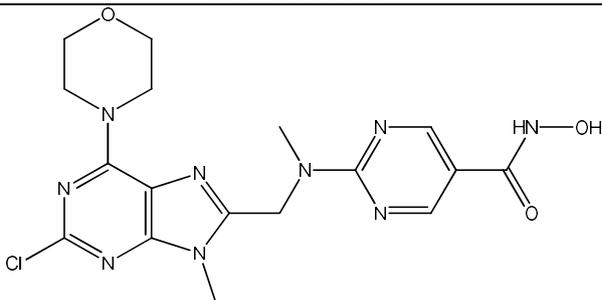
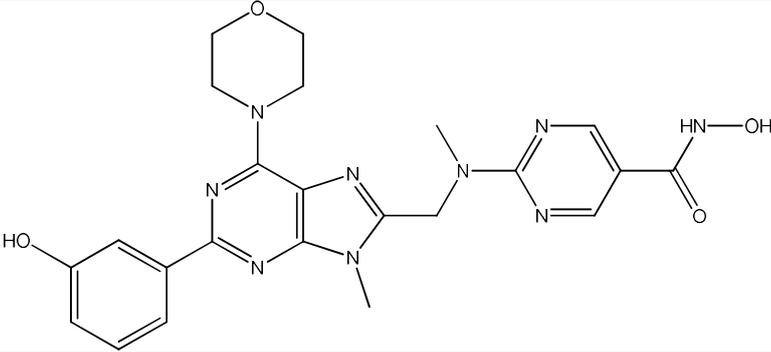
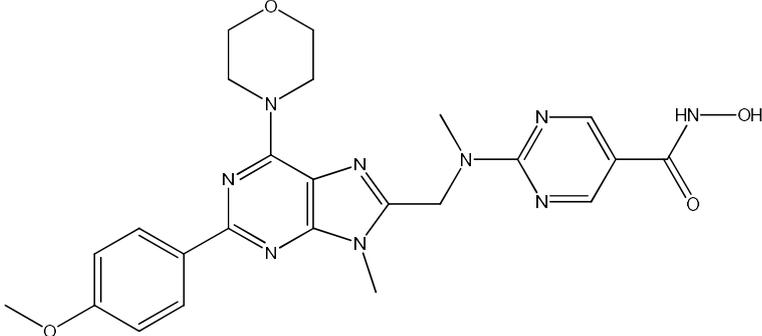
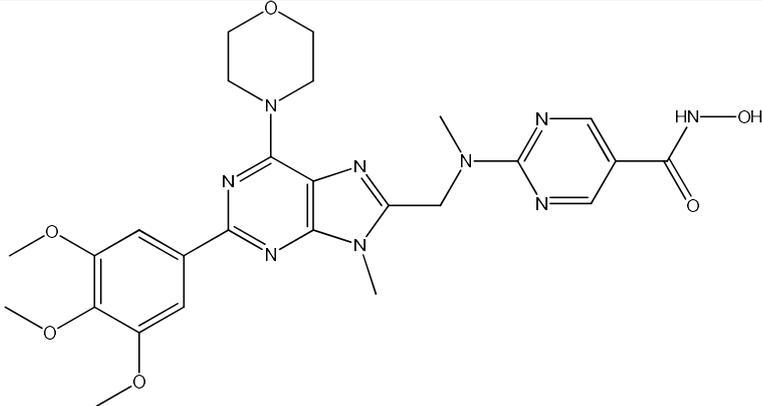
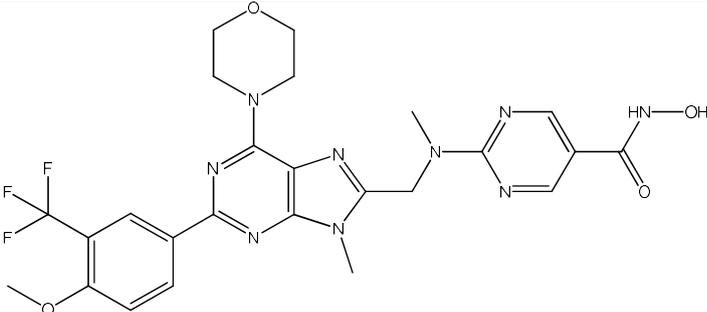
In Eq. (1), n denotes the number of data points used in the correlation, r^2 is the square of the correlation coefficient, r^2_{cv} is the square of cross-validated correlation coefficient obtained by the leave-one-out (LOO) jackknife procedure, and r^2_{pred} is the square of correlation coefficient obtained for test set compounds to judge the external validity of the correlation. Values of r^2_{cv} and r^2_{pred} are calculated according to Eqs. (2) and (3), respectively, where $y_{i,obsd}$ in Eq. (2) refers to the observed activity of compound i in the training set and that in Eq.(3) to the compound i in the test set. Similarly, $y_{i,pred}$ in Eq.(2) refers to the predicted activity of compound i in the training set obtained in leave-one-out jackknife procedure and that in Eq.(3) to that predicted for the test set compounds by the model obtained in the training set. However, $y_{av,obsd}$ in the equations refers to the average activity of the training set compounds. Now, eq. 1 indicates that the negative values of the descriptors ATSC6i and C-008 refer to a decrease in centered broto Moreau autocorrelation of lag 6 weighted by ionization potential, hydrophobicity and molecular refractivity will enhance the activity of the molecule.

$$r^2_{cv} = 1 - (\sum_i (y_{i,obsd} - y_{i,pred})^2 / \sum_i (y_{i,obsd} - y_{av,obsd})^2) \quad (2)$$

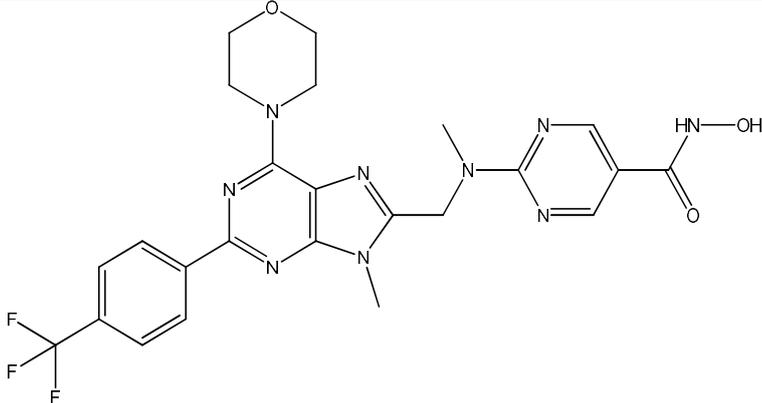
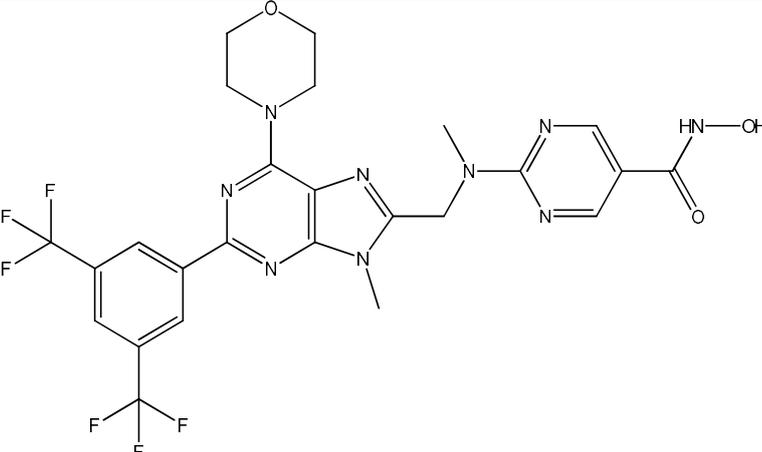
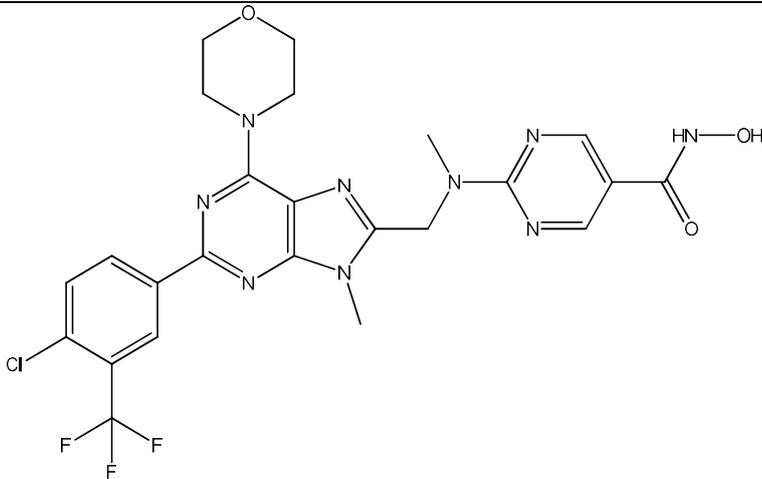
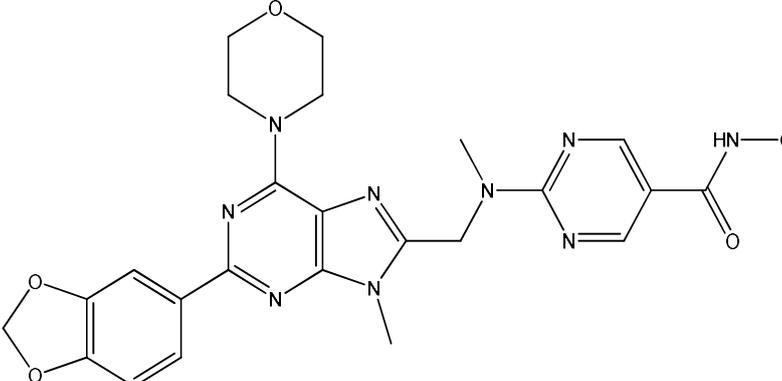
$$r^2_{pred} = 1 - (\sum_i (y_{i,obsd} - y_{i,pred})^2 / \sum_i (y_{i,obsd} - y_{av,obsd})^2) \quad (3)$$

The correlation is supposed to be valid and have a good internal predictive ability if $r^2_{cv} > 0.60$. Similarly, the external predictive ability of the model is supposed to be good if its $r^2_{pred} > 0.5$. From both parameters, the correlation expressed by Eq. (1) is found to be quite valid. Among the remaining two statistical parameters, s and F , s is the standard deviation, and F is the Fischer ratio between the variances of the calculated and observed activities. The figures within the parentheses

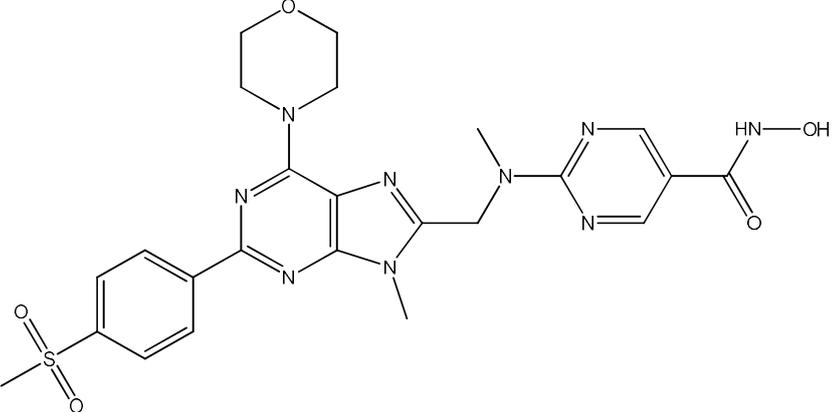
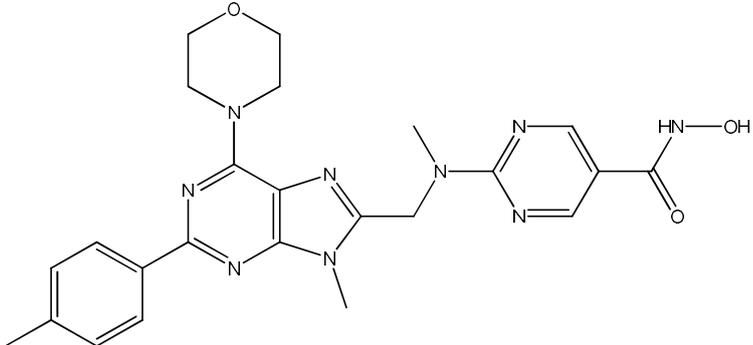
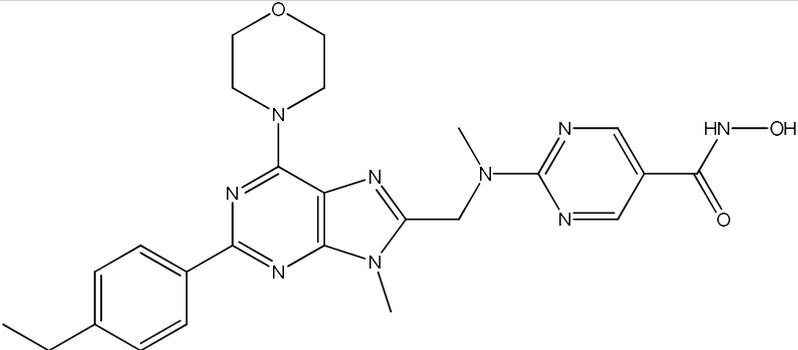
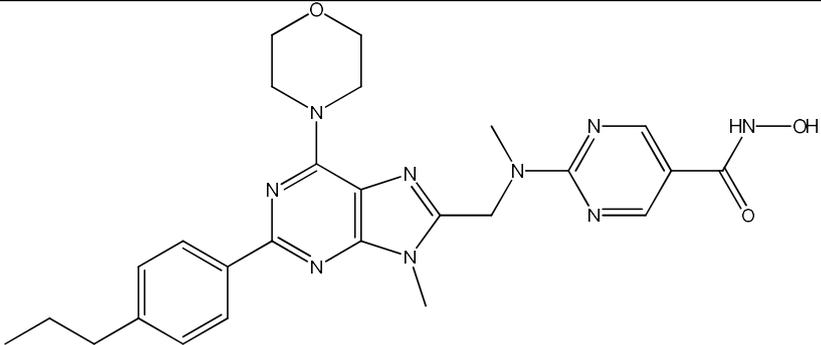
Table 1. Molecules used in the present study

S. No.	Molecular Structure	IC ₅₀ (nm)	pIC ₅₀
1.		0.45	9.35
2.		0.55	9.26
3.		0.45	9.35
4.		1.02	8.99
5.		1.59	8.80

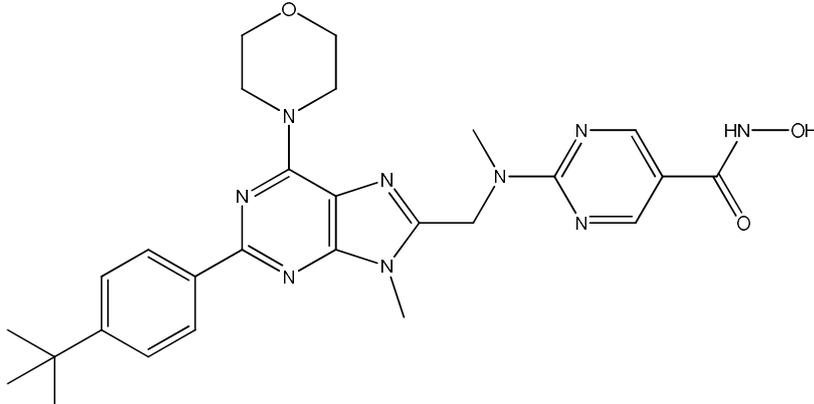
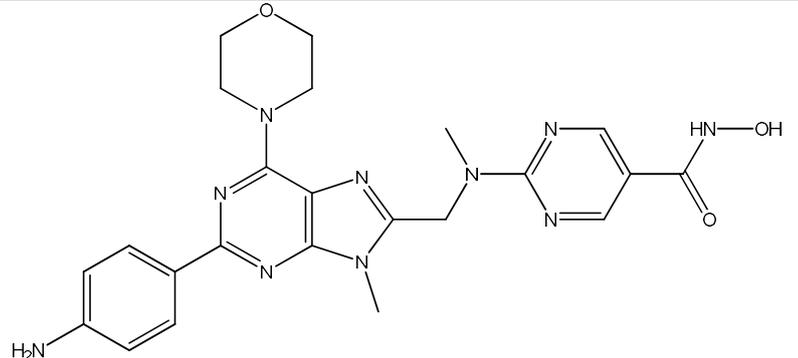
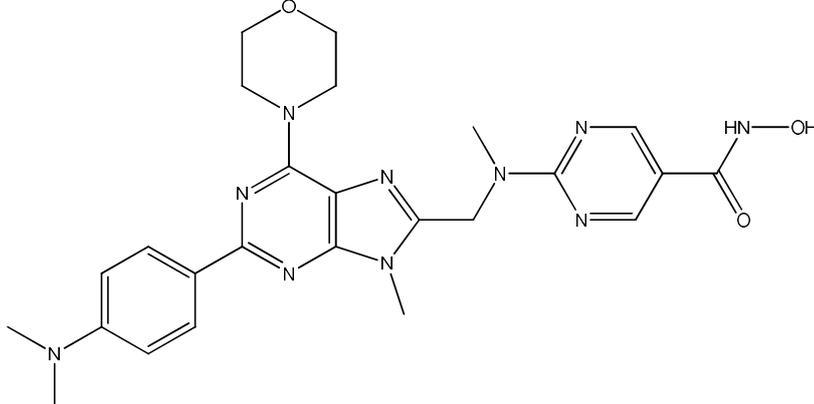
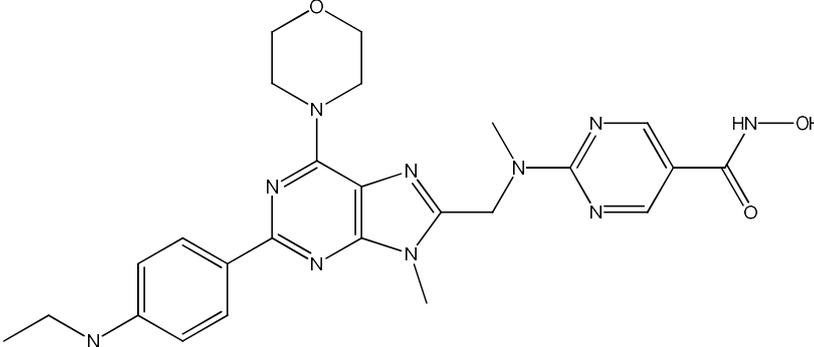
Continue.....

6.		5.10	8.29
7.		30.10	7.52
8.		1.01	9.00
9.		0.97	9.01

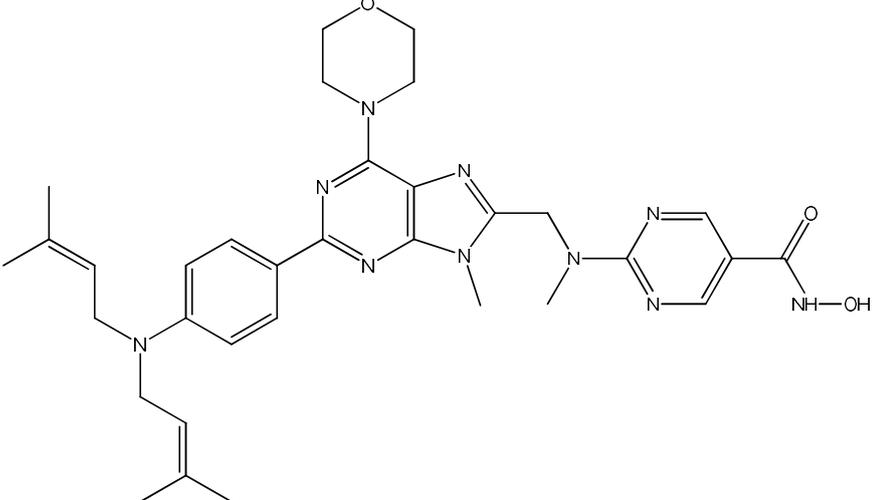
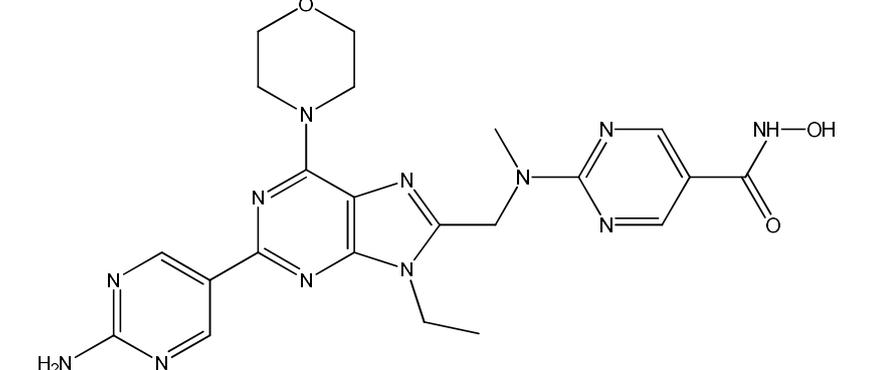
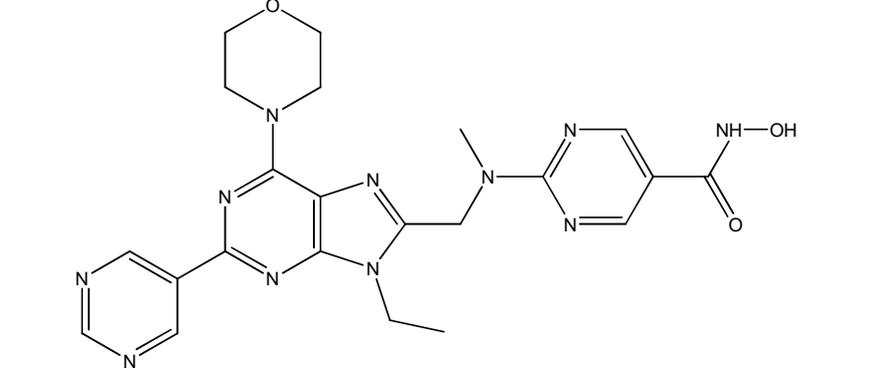
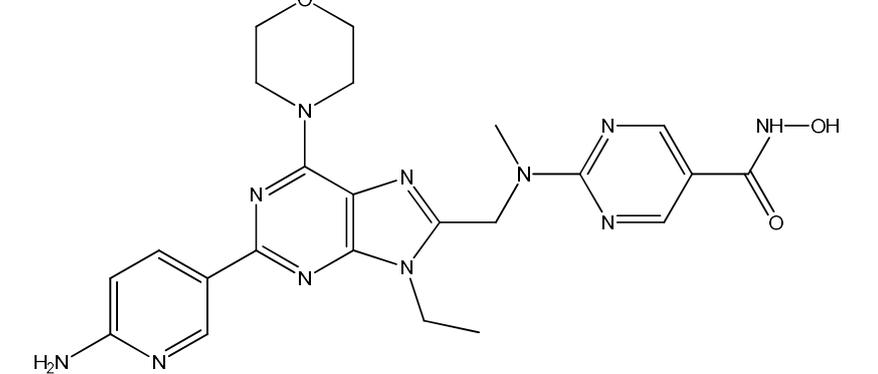
Continue.....

10.		0.85	9.07
11.		1.00	9.00
12.		1.98	8.70
13.		6.15	8.21

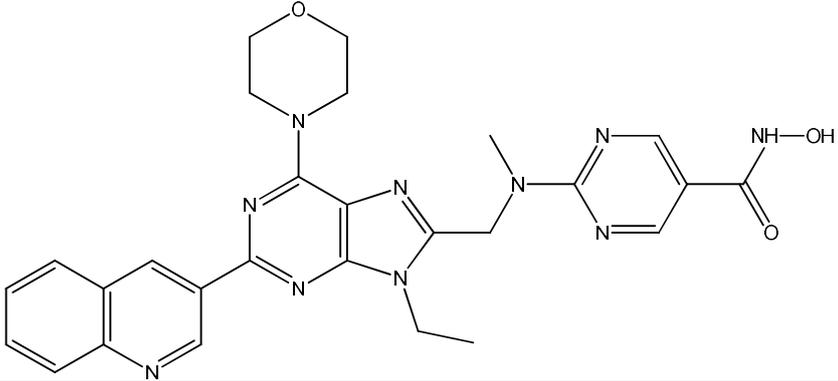
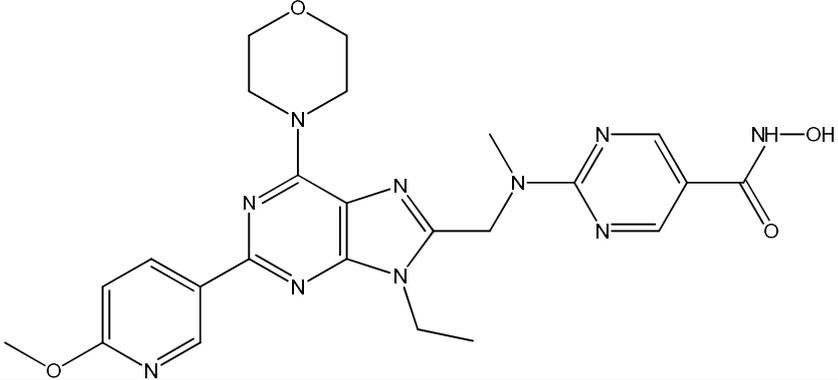
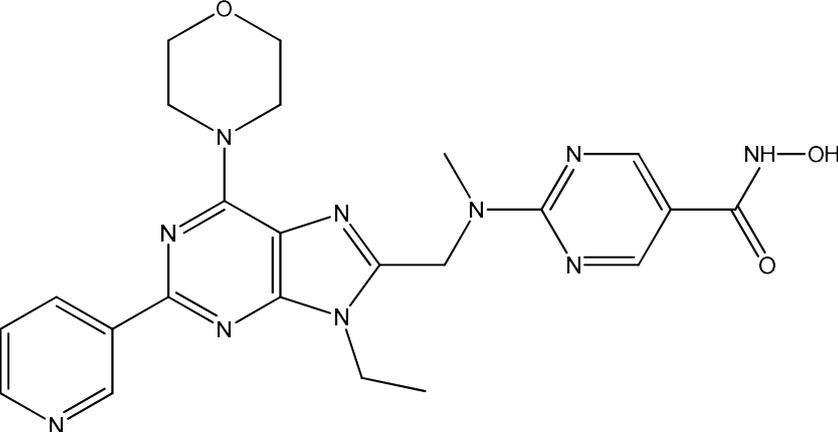
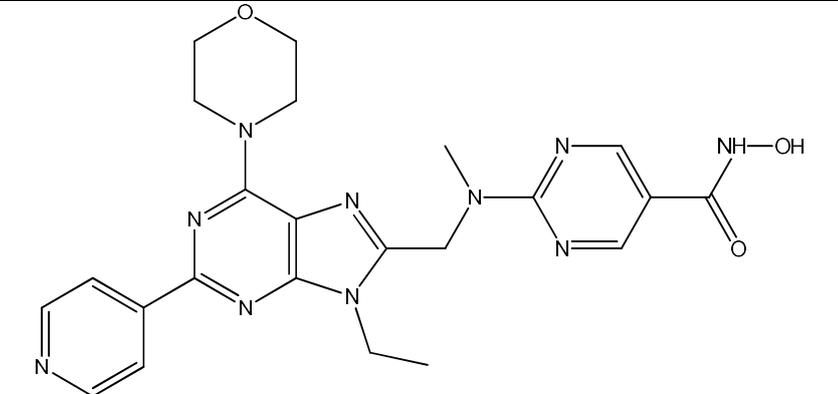
Continue.....

14.		10.01	8.00
15.		0.65	9.19
16.		0.81	9.09
17.		1.10	8.96

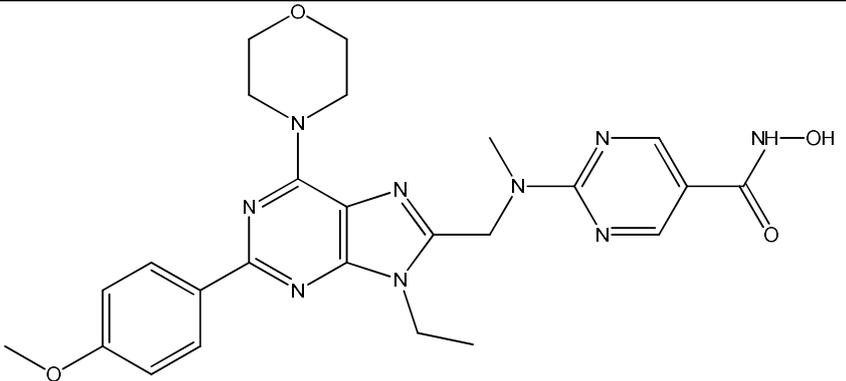
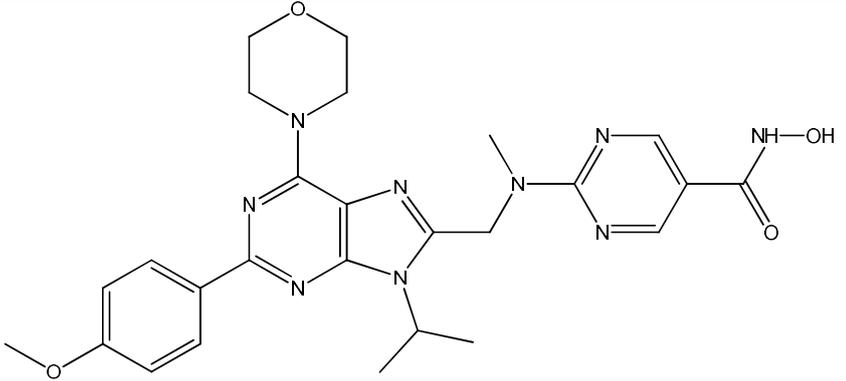
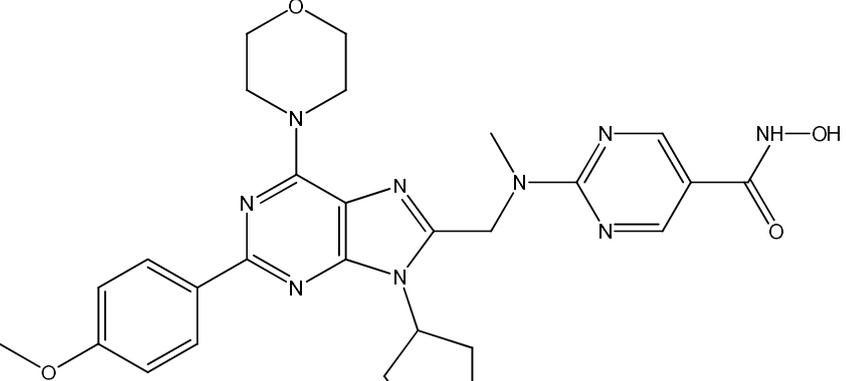
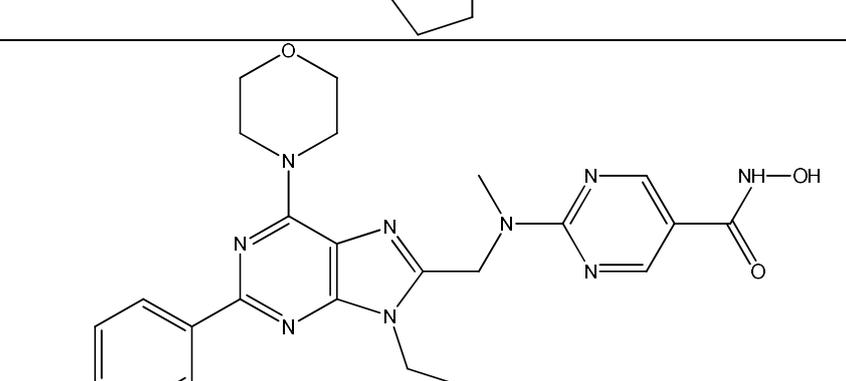
Continue.....

18.		10.00	8.00
19.		1.04	8.98
20.		1.14	8.94
21.		1.11	8.95

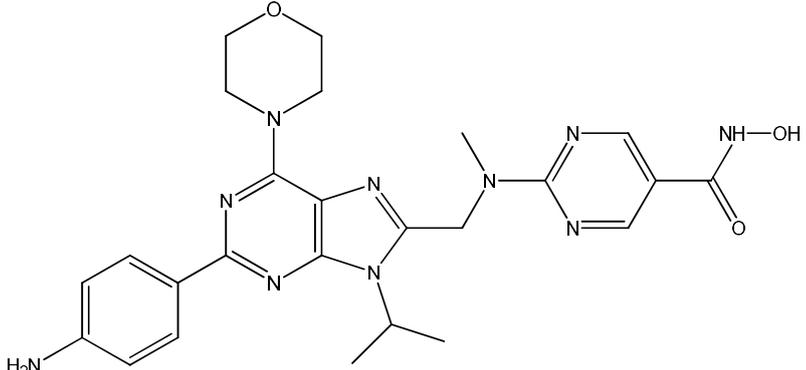
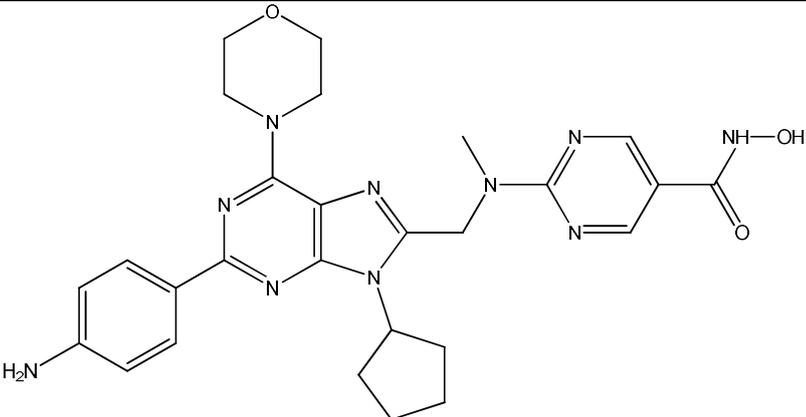
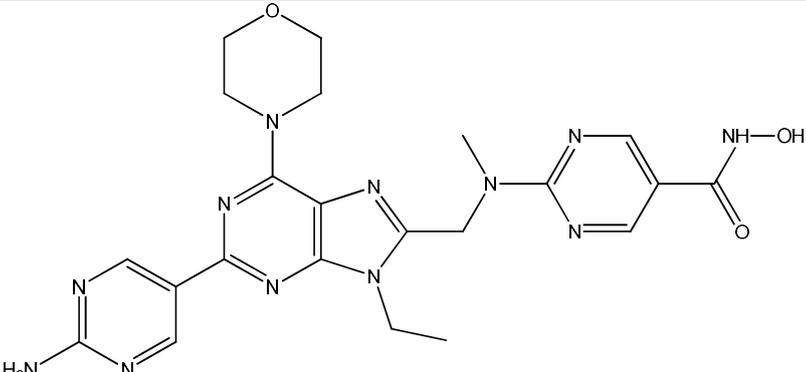
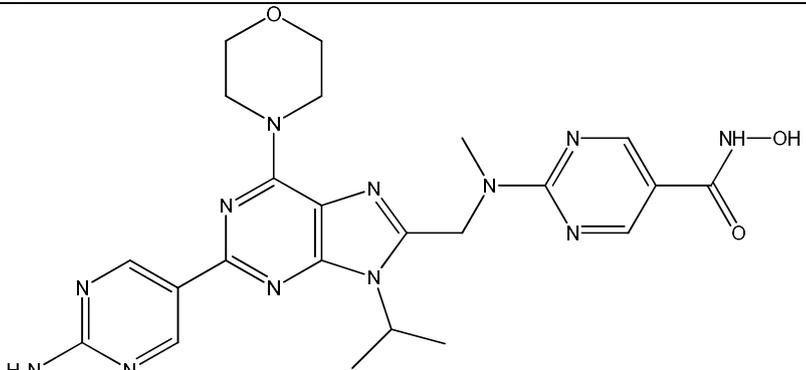
Continue.....

22.		1.60	8.80
23.		1.00	9.00
24.		0.49	9.31
25.		0.55	9.26

Continue.....

26.		1.00	9.00
27.		110	6.96
28.		125	6.90
29.		1.37	8.86

Continue.....

30.		165	6.78
31.		190	6.72
32.		2.62	8.58
33.		240	6.62

Continue.....

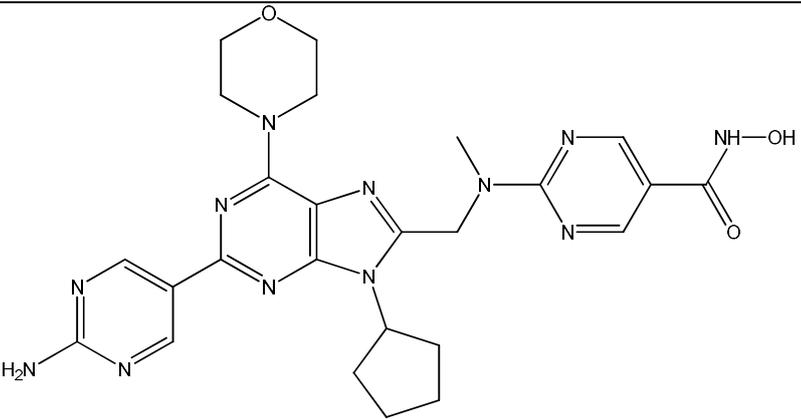
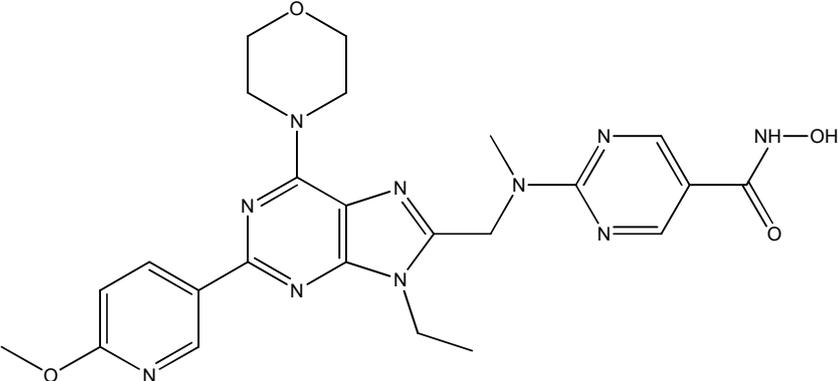
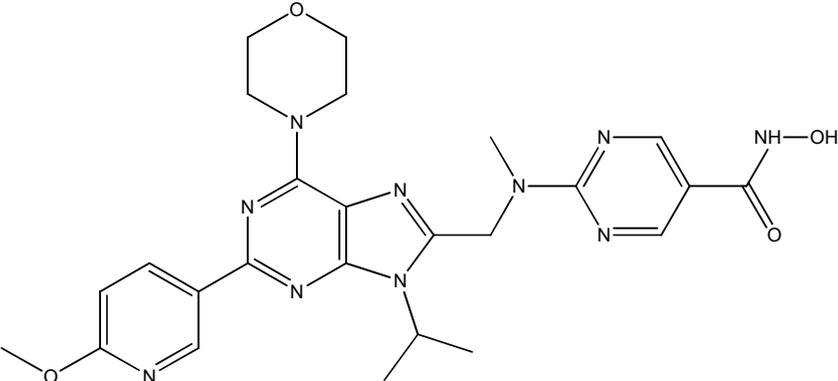
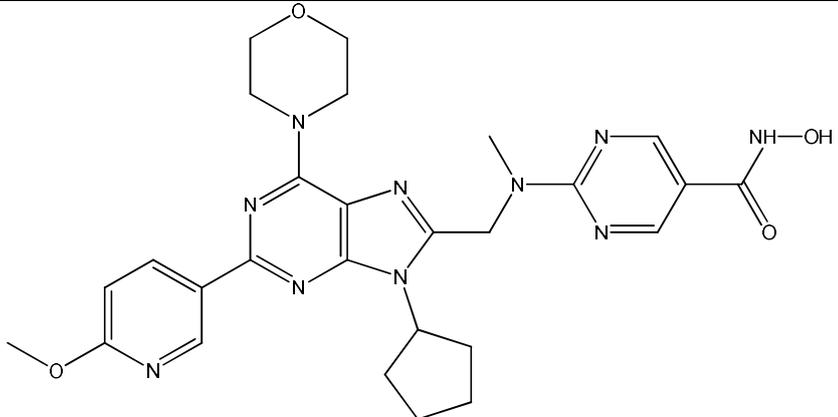
34.		300	6.52
35.		1.60	8.80
36.		62	7.21
37.		67	7.17

Table 2: Molecular Descriptors Used in The Present Study

Compd. No.	SpMax B(e)	SpDiam EA(bo)	GATS4m	R1v+	SpMax A	RDF025s	SpDiam AEA(bo)	MATS4v
1.	4.063	7.678	0.852	0.083	2.515	86.824	5.713	-0.045
2.	4.074	7.689	1.010	0.056	2.531	94.007	5.725	-0.013
3.	4.074	7.689	1.061	0.056	2.531	88.888	5.725	-0.077
4.	4.079	7.691	1.002	0.054	2.537	97.564	5.727	-0.085
5.	4.078	7.690	0.989	0.056	2.537	100.298	5.728	-0.053
6.	4.075	7.689	1.061	0.056	2.532	107.503	5.725	-0.044
7.	4.080	7.690	1.053	0.056	2.540	122.334	5.730	0.031
8.	4.078	7.690	0.990	0.056	2.536	94.128	5.728	-0.039
9.	4.077	7.690	1.055	0.056	2.535	84.396	5.726	-0.061
10.	4.079	7.693	1.006	0.055	2.532	144.783	5.803	-0.088
11.	4.074	7.689	1.022	0.056	2.531	92.234	5.725	-0.034
12.	4.074	7.689	1.046	0.055	2.531	97.078	5.725	-0.074
13.	4.074	7.689	1.002	0.055	2.531	101.384	5.725	-0.020
14.	4.074	7.689	0.948	0.054	2.532	112.998	5.725	0.026
15.	4.074	7.689	1.031	0.056	2.531	92.640	5.725	-0.031
16.	4.074	7.689	1.021	0.057	2.531	91.646	5.725	-0.044
17.	4.074	7.689	1.026	0.057	2.531	100.942	5.725	-0.038
18.	4.074	7.689	0.945	0.039	2.532	129.619	5.725	0.024
19.	4.073	7.696	1.019	0.051	2.538	104.329	5.734	-0.044
20.	4.072	7.696	0.987	0.058	2.537	99.366	5.734	-0.001
21.	4.072	7.696	1.014	0.056	2.538	100.462	5.734	-0.042
22.	4.080	7.698	1.020	0.051	2.541	101.818	5.735	-0.033
23.	4.072	7.696	1.041	0.055	2.538	96.419	5.734	-0.048
24.	4.071	7.696	0.996	0.060	2.537	92.718	5.734	-0.018
25.	4.071	7.696	1.003	0.056	2.537	93.636	5.734	-0.042
26.	4.072	7.696	1.035	0.056	2.538	100.592	5.734	-0.075
27.	4.080	7.707	0.960	0.038	2.549	113.015	5.754	-0.009
28.	4.080	7.708	0.926	0.037	2.560	115.806	5.758	0.076
29.	4.072	7.696	1.005	0.056	2.538	99.933	5.734	-0.033
30.	4.080	7.707	0.932	0.041	2.549	112.412	5.753	0.030
31.	4.080	7.708	0.898	0.040	2.560	115.132	5.758	0.113
32.	4.073	7.696	1.019	0.051	2.538	104.272	5.734	-0.044
33.	4.081	7.707	0.939	0.047	2.549	115.307	5.753	0.027
34.	4.081	7.708	0.902	0.046	2.560	116.689	5.758	0.119
35.	4.072	7.696	1.041	0.055	2.538	96.450	5.734	-0.048
36.	4.080	7.707	0.963	0.039	2.549	132.141	5.754	0.019
37.	4.081	7.708	0.927	0.039	2.560	111.568	5.758	0.105

Table 2. Molecular Descriptors Used in The Present Study

Compd. No.	GGI6	C-008	VE1sign X	Eig14 EA(ri)	Eig07 AEA(bo)	RDF120s	GGI9	Eig05 EA(ri)
1.	0.771	0.000	0.030	0.224	3.090	0.007	0.220	1.985
2.	1.066	0.000	0.029	0.946	3.328	5.205	0.327	2.404
3.	1.124	0.000	0.030	0.946	3.348	56.121	0.428	2.404
4.	1.164	0.000	0.048	0.946	3.355	85.910	0.515	2.429
5.	1.226	0.000	0.234	1.026	3.586	66.050	0.641	2.513
6.	1.287	0.000	0.023	1.000	3.396	108.845	0.630	2.531
7.	1.249	0.000	0.438	1.130	3.855	112.368	0.733	2.595
8.	1.226	0.000	0.053	1.023	3.541	56.710	0.605	2.510
9.	1.092	0.000	0.030	1.008	3.354	39.808	0.381	2.427
10.	1.287	0.000	0.023	1.000	3.441	53.079	0.630	2.538
11.	1.124	0.000	0.031	0.946	3.342	21.429	0.393	2.404
12.	1.124	0.000	0.030	0.946	3.348	35.391	0.428	2.404
13.	1.164	0.000	0.031	1.000	3.350	42.177	0.438	2.404
14.	1.287	0.000	0.023	1.000	3.396	68.688	0.630	2.544
15.	1.124	0.000	0.031	0.946	3.342	29.594	0.393	2.404
16.	1.205	0.000	0.028	1.000	3.362	49.549	0.519	2.439
17.	1.164	0.000	0.031	0.989	3.350	50.875	0.438	2.404
18.	1.571	0.000	0.328	1.161	3.641	133.911	0.619	2.511
19.	1.164	0.000	0.051	0.916	3.346	57.627	0.459	2.413
20.	1.085	0.000	0.049	0.916	3.229	41.139	0.355	2.413
21.	1.164	0.000	0.051	0.946	3.346	26.176	0.459	2.413
22.	1.103	0.000	0.047	1.027	3.365	59.243	0.442	2.470
23.	1.164	0.000	0.049	0.946	3.352	36.160	0.475	2.413
24.	1.085	0.000	0.049	0.946	3.229	11.372	0.355	2.413
25.	1.085	0.000	0.049	0.946	3.229	21.138	0.355	2.413
26.	1.164	0.000	0.049	0.946	3.352	51.139	0.475	2.413
27.	1.448	1.000	0.074	1.000	3.372	60.402	0.542	2.439
28.	1.404	1.000	0.044	1.000	3.451	50.236	0.510	2.464
29.	1.164	0.000	0.051	0.946	3.346	27.601	0.459	2.413
30.	1.448	1.000	0.076	1.000	3.367	52.207	0.526	2.439
31.	1.404	1.000	0.033	1.000	3.450	33.428	0.494	2.464
32.	1.164	0.000	0.051	0.916	3.346	56.498	0.459	2.413
33.	1.448	1.000	0.076	0.946	3.367	65.158	0.526	2.439
34.	1.404	1.000	0.033	0.960	3.450	44.077	0.494	2.464
35.	1.164	0.000	0.049	0.946	3.352	39.425	0.475	2.413
36.	1.448	1.000	0.074	0.956	3.372	9.245	0.542	2.439
37.	1.404	1.000	0.044	0.960	3.451	36.273	0.510	2.464

Table 2. Molecular Descriptors Used in The Present Study

Compd. No.	GGI1	DISPe	Eig10 AEA(ri)	ATSC6i
1.	6.000	0.153	2.045	1.256
2.	7.000	0.146	2.377	1.961
3.	7.000	0.179	2.395	2.012
4.	8.000	0.117	2.543	2.373
5.	9.500	0.293	2.543	2.500
6.	9.000	0.358	2.526	2.061
7.	12.000	0.563	2.720	3.748
8.	9.500	0.406	2.538	2.252
9.	7.000	0.129	2.480	2.015
10.	9.000	0.156	2.528	2.092
11.	7.000	0.246	2.378	1.994
12.	7.000	0.275	2.407	2.061
13.	7.000	0.331	2.434	2.165
14.	9.000	0.338	2.529	2.193
15.	7.000	0.185	2.374	1.966
16.	7.500	0.256	2.491	2.096
17.	7.000	0.269	2.424	2.134
18.	9.500	0.358	2.739	2.920
19.	7.000	0.157	2.391	2.042
20.	6.000	0.169	2.346	2.026
21.	7.000	0.149	2.393	2.090
22.	7.000	0.266	2.561	2.302
23.	7.000	0.142	2.407	2.139
24.	6.000	0.164	2.347	2.071
25.	6.000	0.155	2.346	2.071
26.	7.000	0.157	2.410	2.180
27.	7.500	0.207	2.543	2.347
28.	7.000	0.165	2.543	2.601
29.	7.000	0.179	2.395	2.133
30.	7.500	0.221	2.540	2.300
31.	7.000	0.179	2.540	2.554
32.	7.000	0.157	2.391	2.042
33.	7.500	0.126	2.539	2.210
34.	7.000	0.094	2.539	2.469
35.	7.000	0.142	2.407	2.139
36.	7.500	0.110	2.542	2.306
37.	7.000	0.108	2.543	2.563

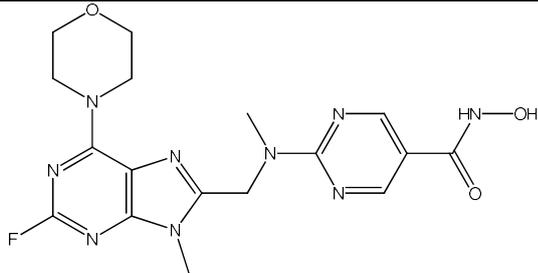
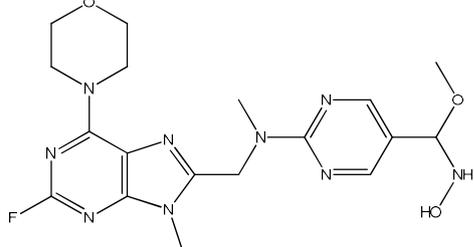
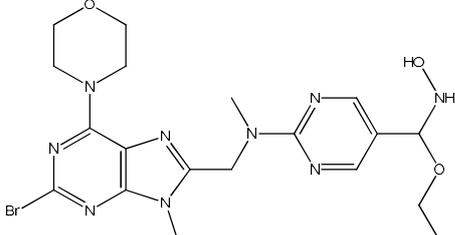
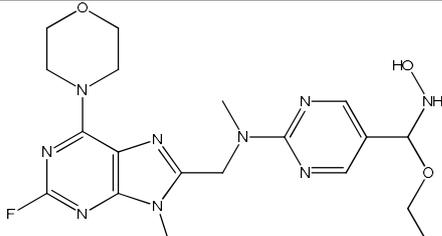
Table 3. Regression Parameters and Quality of Correlation

Model No.	Variables	$A_i = (I---2)$	C	R^2	R^2_{adj}	RMSE _{tr}	MAE _{tr}	S	F
1.	SpMax B(e)	-214.8028(±65.6184)	883.8133	0.616	0.603	0.583	0.461	0.604	44.964
2.	SpDiam EA(bo)	-104.0704(±31.6859)	809.2659	0.618	0.604	0.582	0.473	0.603	45.264
3.	GATS4m	16.4897(±4.8338)	-8.0239	0.636	0.623	0.569	0.436	0.589	48.829
4.	R1v+	110.7587(±30.3134)	2.6474	0.667	0.655	0.544	0.401	0.563	56.017
5.	SpMax A	-81.2872(±22.2178)	214.8044	0.667	0.655	0.543	0.456	0.562	56.166
6.	RDF025s	-0.0770(±0.0204)	16.3317	0.681	0.670	0.532	0.408	0.551	59.743
7.	SpDiam AEA(bo)	-65.2540(±16.9514)	382.6184	0.690	0.678	0.525	0.419	0.543	62.178
8.	MATS4v	-14.2415(±3.6766)	8.1895	0.692	0.681	0.523	0.393	0.541	62.957
9.	GGI6	-5.9285(±1.3411)	15.6557	0.746	0.736	0.475	0.354	0.492	81.999
10.	C-008	-2.0026(±0.3770)	8.8126	0.809	0.802	0.412	0.300	0.426	118.423
11.	VEI _{sign_X} C-008	-2.9472(±1.3616) -2.0764(±0.2943)	9.0469	0.890	0.881	0.313	0.231	0.330	108.672
12.	Eig14 EA(ri) C-008	-4.9136(±2.2107) -1.9923(±0.2890)	13.6212	0.892	0.884	0.310	0.256	0.326	111.479
13.	Eig07_AEA(bo) C-008	-2.2518(±0.9745) -1.9431(±0.2852)	16.4437	0.896	0.888	0.304	0.238	0.321	115.876
14.	RDF120s C-008	-0.0109(±0.0046) -2.0357(±0.2813)	9.3798	0.898	0.890	0.301	0.227	0.317	118.772
15.	GGI9 C-008	-3.1308(±1.3214) -1.8915(±0.2847)	10.3131	0.898	0.891	0.301	0.242	0.317	118.892
16.	Eig05_EA(ri) C-008	-6.0285(±2.4985) -1.9315(±0.2799)	23.5305	0.900	0.892	0.298	0.233	0.314	121.185
17.	GGI1 C-008	-0.2323(±0.0958) -2.0841(±0.2798)	10.5703	0.900	0.893	0.298	0.238	0.314	121.794
18.	ATSC6i C-008	-0.8459(±0.3425) -1.8363(±0.2834)	10.7061	0.902	0.895	0.295	0.224	0.311	124.235
19.	Eig10_AEA(ri) C-008	-3.0650(±1.2327) -1.7454(±0.2932)	16.3437	0.903	0.895	0.294	0.237	0.310	125.150
20.	DISPe C-008	-3.0150(±1.1898) -2.2330(±0.2866)	9.5168	0.905	0.897	0.291	0.243	0.307	127.789

Table 4. Internal and External Validation Parameters for the Obtained Models

No.	Variables	R ² _{CV}	RMSE _{ext}	MAE _{ext}	R ² _{Pred}	average \square^2_{\square}	$\Delta \square^2_{\square}$
1.	SpMax B(e)	0.5610	0.9569	0.7198	0.3399	0.1400	0.3905
2.	RDF025s	0.5681	1.5620	0.9625	0.2380	-0.0261	0.4763
3.	SpDiam EA(bo)	0.5710	0.4934	0.3743	0.6609	0.5239	0.2728
4.	GATS4m	0.5807	1.3798	0.9297	0.0167	-0.0043	0.0177
5.	R1v+	0.6198	0.9911	0.5638	0.5850	0.1750	0.5314
6.	SpMax_A	0.6256	0.5200	0.4177	0.6203	0.4858	0.1252
7.	GGI6	0.6517	0.7775	0.4949	0.6001	0.2567	0.4511
8.	SpDiam AEA(bo)	0.6528	1.9627	0.9759	0.0198	-0.0092	0.0502
9.	MATS4v	0.6595	0.4389	0.3885	0.6114	0.3189	0.3987
10.	C-008	0.7881	0.3514	0.3071	0.8031	0.7185	0.0974
11.	VEIsign X C-008	0.8631	0.3493	0.2518	0.7985	0.6843	0.1855
12.	Eig14 EA(ri) C-008	0.8633	1.2284	0.6728	0.5211	0.0668	0.6271
13.	GGI9 C-008	0.8673	0.3516	0.2501	0.8749	0.6383	0.1753
14.	Eig07 AEA(bo) C-008	0.8702	0.2750	0.2243	0.8688	0.7686	0.1291
15.	Eig05 EA(ri) C-008	0.8707	0.9052	0.5227	0.6039	0.1885	0.5249
16.	RDF120s C-008	0.8730	0.1505	0.1234	0.9514	0.9280	0.0290
17.	GGI1 C-008	0.8731	0.3040	0.2323	0.8882	0.7533	0.1222
18.	DISPe C-008	0.8743	0.1905	0.1693	0.9373	0.8696	0.0644
19.	Eig10 AEA(ri) C-008	0.8786	0.3980	0.3403	0.8358	0.5861	0.2155
20.	ATSC6i C-008	0.8855	0.3032	0.2086	0.8721	0.7371	0.1356

Table 5. Some Proposed Compounds Belonging to the Series of Table 1 and their Predicted Activity

Compd. No	Molecule	ATSC6i	C-008	Pred. pIC ₅₀
1.		1.51	0	9.43
2.		1.59	0	9.36
3.		1.57	0	9.37
4.		1.71	0	9.26

Continue

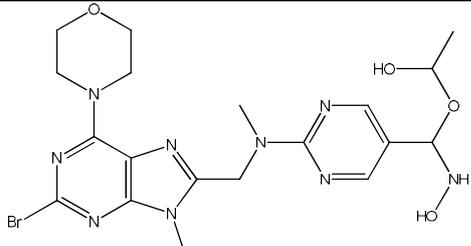
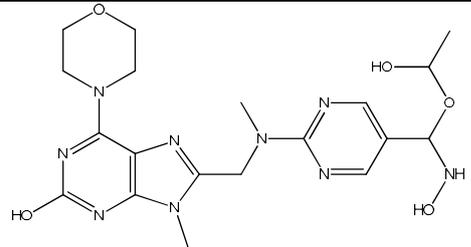
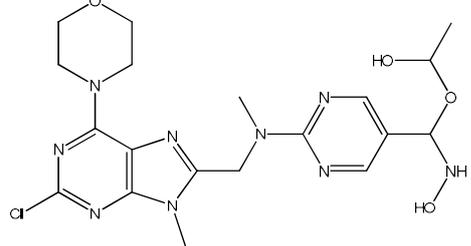
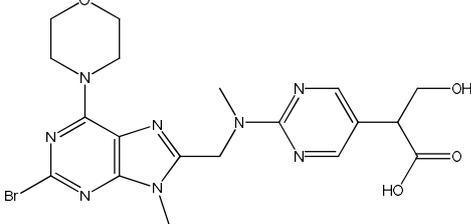
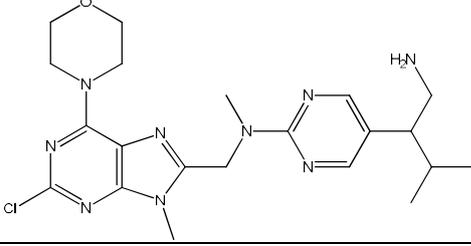
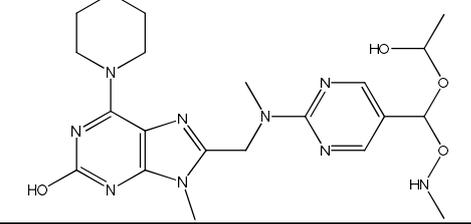
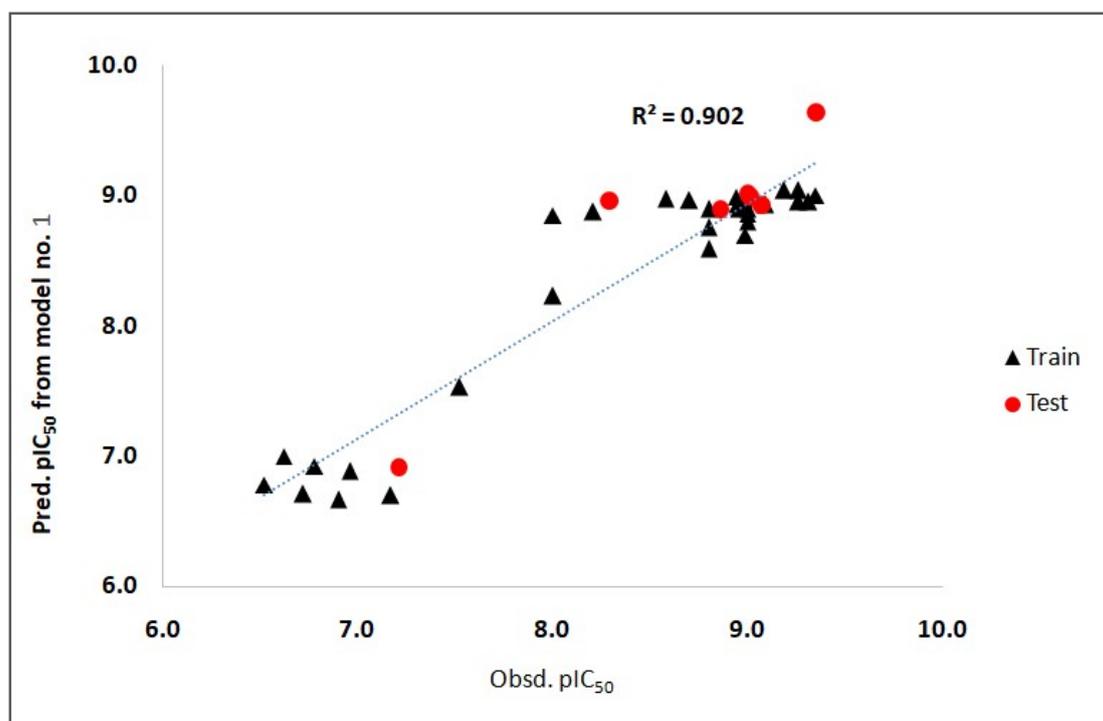
5.		1.59	0	9.36
6.		1.55	0	9.40
7.		1.48	0	9.45
8.		1.46	0	9.47
9.		1.58	0	9.37
10.		1.73	0	9.24

Table 6. Observed and Calculated pIC₅₀ Values of Two-Variable Model

S. No.	Status	pIC ₅₀		ApIC ₅₀ Model	Pred. LOO Model
		Obsd.	Cald. by Model		
1	Prediction	9.350	9.644	0.294	-
2	Training	9.260	9.047	-0.213	9.032
3	Training	9.350	9.004	-0.346	8.982
4	Training	8.990	8.698	-0.292	8.684
5	Training	8.800	8.591	-0.209	8.577
6	Prediction	8.290	8.963	0.673	-
7	Training	7.520	7.536	0.016	7.572
8	Training	9.000	8.801	-0.199	8.792
9	Prediction	9.010	9.002	-0.008	-
10	Prediction	9.070	8.936	-0.134	-
11	Prediction	9.000	9.019	0.019	-
12	Training	8.700	8.963	0.263	8.978
13	Training	8.210	8.875	0.665	8.906
14	Training	8.000	8.851	0.851	8.890
15	Training	9.190	9.043	-0.147	9.033
16	Training	9.090	8.933	-0.157	8.925
17	Training	8.960	8.901	-0.059	8.898
18	Training	8.000	8.236	0.236	8.287
19	Training	8.980	8.979	-0.001	8.979
20	Training	8.940	8.992	0.052	8.996
21	Training	8.950	8.938	-0.012	8.937
22	Training	8.800	8.759	-0.041	8.757
23	Training	9.000	8.897	-0.103	8.892
24	Training	9.310	8.954	-0.356	8.935
25	Training	9.260	8.955	-0.306	8.938
26	Training	9.000	8.862	-0.138	8.856
27	Training	6.960	6.884	-0.076	6.871
28	Training	6.900	6.669	-0.231	6.628
29	Prediction	8.860	8.902	0.042	-
30	Training	6.780	6.924	0.144	6.949
31	Training	6.720	6.709	-0.011	6.707
32	Training	8.580	8.979	0.399	9.002
33	Training	6.620	7.000	0.380	7.071
34	Training	6.520	6.781	0.261	6.825
35	Training	8.800	8.897	0.097	8.902
36	Prediction	7.210	6.919	-0.291	-
37	Training	7.170	6.702	-0.468	6.621

Figure 1. Correlation Between Observed and Calculated pIC₅₀ Using eq. 1

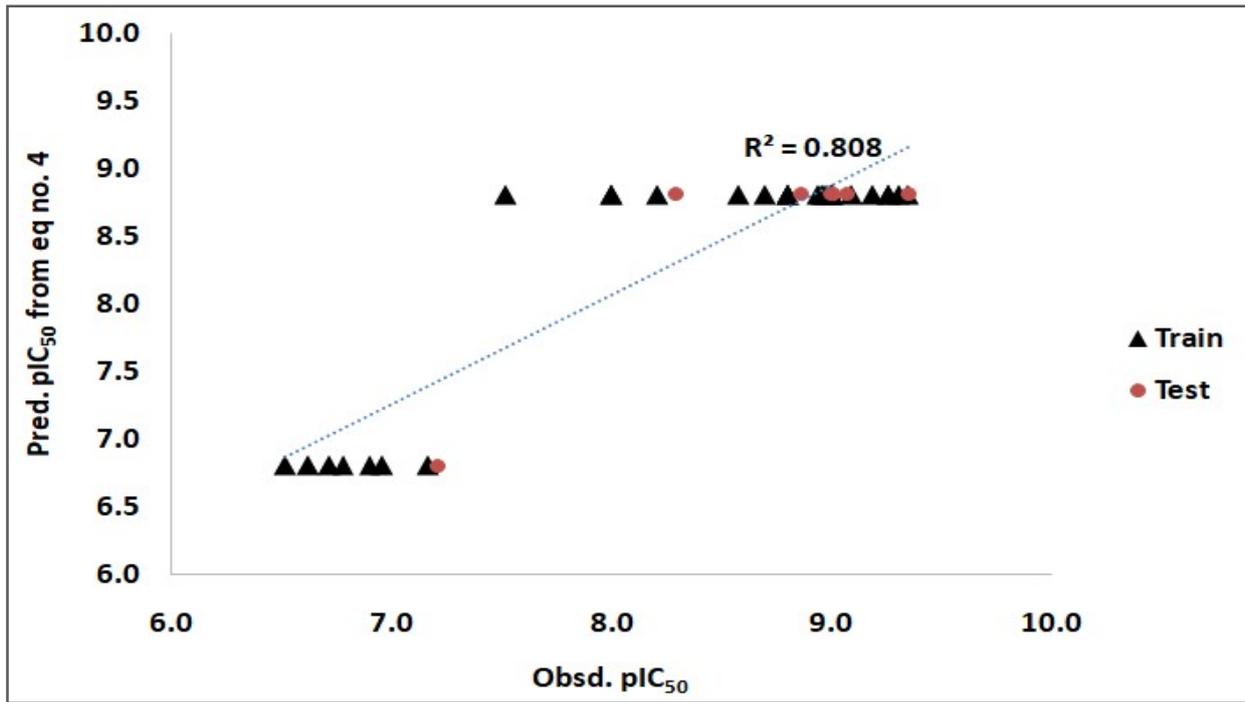


Figure 2. Correlation Between Observed and Calculated pIC₅₀ Using eq 4

Table 7. Observed and Calculated pIC₅₀ Values of one Variable Model

S. No.	Status	pIC ₅₀		ΔpIC ₅₀ Model	Pred. LOO Model
		Obsd.	Cald. by Model		
1	Prediction	9.350	8.813	-0.537	-
2	Training	9.260	8.813	-0.447	8.792
3	Training	9.350	8.813	-0.537	8.788
4	Training	8.990	8.813	-0.177	8.805
5	Training	8.800	8.813	0.013	8.813
6	Prediction	8.290	8.813	0.523	-
7	Training	7.520	8.813	1.293	8.871
8	Training	9.000	8.813	-0.187	8.804
9	Prediction	9.010	8.813	-0.197	-
10	Prediction	9.070	8.813	-0.257	-
11	Prediction	9.000	8.813	-0.187	-
12	Training	8.700	8.813	0.113	8.818
13	Training	8.210	8.813	0.603	8.840
14	Training	8.000	8.813	0.813	8.850
15	Training	9.190	8.813	-0.377	8.796
16	Training	9.090	8.813	-0.277	8.800
17	Training	8.960	8.813	-0.147	8.806
18	Training	8.000	8.813	0.813	8.850
19	Training	8.980	8.813	-0.167	8.805
20	Training	8.940	8.813	-0.127	8.807
21	Training	8.950	8.813	-0.137	8.806
22	Training	8.800	8.813	0.013	8.813
23	Training	9.000	8.813	-0.187	8.804
24	Training	9.310	8.813	-0.497	8.790
25	Training	9.260	8.813	-0.447	8.792
26	Training	9.000	8.813	-0.187	8.804
27	Training	6.960	6.810	-0.150	6.785
28	Training	6.900	6.810	-0.090	6.795
29	Prediction	8.860	8.813	-0.047	-
30	Training	6.780	6.810	0.030	6.815
31	Training	6.720	6.810	0.090	6.825
32	Training	6.580	6.813	0.233	6.823
33	Training	6.620	6.810	0.190	6.842
34	Training	6.520	6.810	0.290	6.858
35	Training	8.800	8.813	0.013	8.813
36	Prediction	7.210	6.810	-0.400	-
37	Training	7.170	6.810	-0.360	6.750

with \pm sign refer to the 95% confidence intervals. The F -value given in parentheses refers to the standard F -value at the 99% level. A higher value of F than this indicates a good correlation. Thus, all descriptors used in this correlation are found to be quite significant and if we remove them one by one, the significance of the correlation is appreciably dropped (Eqs. 4).

$$pIC_{50} = -2.0026(\pm 0.3770)C-008 + 8.8126$$

$$n = 30, r^2 = 0.809, r_{cv}^2 = 0.788, r_{pred}^2 = 0.803, s = 0.492, F = 118.42 \quad (4)$$

Thus, from the above results, Eq. (1) has a significant correlation between the inhibitory activity values and the structural descriptors of the compounds. Although the correlation does not have any mechanistic aspects, it has a good predictive ability. A graph is drawn between the predicted and observed activities for both the training and test sets, showing that the model has a good predictive ability. Figure 1 shows that almost all the points, except a few, lie near the straight line. Thus, using Eq.1, we predicted some new compounds, as reported in Table 6, where each compound has a higher activity value than any compound in the existing series (Table 1).

Docking Analysis: Molecular docking analysis was performed on the predicted compounds (Table 5) using LeadIT FlexX software to get the binding mode of these compounds. The ability of a molecule to interact with an enzyme decides its potency. For the study of molecular docking, the crystal structure of the related enzyme is vital, which can now be retrieved from the RCSB protein data bank. We selected the enzyme with PDB entry code 4BKX (<http://www.pdb.org>). All predicted compounds listed in Table 5 were docked in this enzyme, but there is no significant docking results were obtained.

CONCLUSION

Topological as well as physicochemical parameters are capable of modeling the antitumor activity. On the basis of best model following conclusion can be drawn. The negative values of the descriptors ATSC6i and C-008 refer to a decrease in centered broto Moreau autocorrelation of lag 6 weighted by ionization potential, hydrophobicity and molecular refractivity will enhance the activity of the molecule.

ACKNOWLEDGEMENT

Authors are thankful to the referees for their valuable suggestions making the paper in this form. I am also thankful to Dr. Basheerullah Seikh, NITTTR, Bhopal for their generous help and encouragement for presenting this work. I am also thankful to CSJM University, Kanpur, for giving us a financial support as CV Raman Minor research Project.

REFERENCES

Leach, A. R. (2001). *Molecular Modelling: Principles and Applications* (2nd ed.). Prentice Hall.

Jensen, F. (2017). *Introduction to Computational Chemistry* (3rd ed.). Wiley.

Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: current applications and future

challenges. *Drug Discovery Today*, 14(13-14), 676-683. <https://doi.org/10.1016/j.drudis.2009.04.006>

Jorgensen, W. L. (2004). The many roles of computation in drug discovery. *Science*, 303(5665), 1813-1818. <https://doi.org/10.1126/science.1096361>

Karplus, M., & Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences*, 102(19), 6679-6685. <https://doi.org/10.1073/pnas.0408930102>

Ekins, S., & Honeycutt, J. D. (2005). Applications of computational molecular modeling in structure-based drug discovery. *Current Pharmaceutical Design*, 11(3), 277-295. <https://doi.org/10.2174/1381612053381643>

Ferreira, L. G., dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, 20(7), 13384-13421. <https://doi.org/10.3390/molecules200713384>

Bolden, J. E., Peart, M. J., & Johnstone, R. W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nature Reviews Drug Discovery*, 5(9), 769-784. <https://doi.org/10.1038/nrd2133>

Ververis, K., Hiong, A., Karagiannis, T. C., & Licciardi, P. V. (2013). Histone deacetylase inhibitors (HDACIs): multitargeted anticancer agents. *Biologics: Targets & Therapy*, 7, 47-60. <https://doi.org/10.2147/BTT.S29965>

West, A. C., & Johnstone, R. W. (2014). New and emerging HDAC inhibitors for cancer treatment. *Journal of Clinical Investigation*, 124(1), 30-39. <https://doi.org/10.1172/JCI69738>

Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A. M., ... & Drewes, G. (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnology*, 29(3), 255-265. <https://doi.org/10.1038/nbt.1759>

Haberland, M., Montgomery, R. L., & Olson, E. N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature Reviews Genetics*, 10(1), 32-42. <https://doi.org/10.1038/nrg2485>

Li, Y., Seto, E. (2016). HDACs and HDAC Inhibitors in Cancer Development and Therapy. *Cold Spring Harbor Perspectives in Medicine*, 6(10), a026831. <https://doi.org/10.1101/cshperspect.a026831>

Khan, O., & La Thangue, N. B. (2012). HDAC inhibitors in cancer biology: emerging mechanisms and clinical applications. *Immunology and Cell Biology*, 90(1), 85-94. <https://doi.org/10.1038/icb.2011.100>

Basheerulla Shaik, Omar Deeb, Vijay Kumar Agrawal, Satya P. Gupta, QSAR and Molecular Docking Studies on a Series of Cinnamic Acid Analogues as Epidermal Growth Factor Receptor (EGFR) Inhibitors. *Letters in Drug Design & Discovery*, 14(1), 83-95, 2017. DOI: 10.2174/1570180813999160721160833

Basheerulla Shaik, Vijay Agrawal, Satya P. Gupta and Urvana Menon. Quantitative Structure-Activity Relationship and Docking Studies on a Series of Oxadiazole and Triazole Substituted Naphthyridines as HIV-1 Integrase Inhibitors. *Letters in Drug Design and Discover*, 14(1), 10 - 27, 2017. DOI: 10.2174/1570180813666160610090525

Izhar Ahmad, Basheerulla Shaik, Neelu Singh, Vijay K. Agrawal, Anita K, and Satya P Gupta. Quantitative Structure-Activity Relationship and Molecular Modeling Studies on a Series of H⁺-K⁺-ATPase Inhibitors. *Journal*

- of Applied Biopharmaceutics and Pharmacokinetics, 4, 20-39, 2016.
- Neelu Singh, Basheerulla Shaik, Neeraj Agrawal, Anita K, Vijay K. Agrawal and Satya P Gupta. QSAR and Molecular Modeling Studies on a Series of Indole-based Pyridone Analogues as HCV NS5B Polymerase Inhibitors. *Letters in Drug Design & Discovery*, 13(8), 757-770. DOI: 10.2174/1570180813666160815122359
- Basheerulla Shaik, Tripti Kaushal, Vijay K. Agrawal, Quantitative Structure Activity Relationship Studies on a Series of 4-Pyridones as Antimalarial Agents. *Journal of Indian Chemical Society*, 93, 871-876, 2016.
- Shweta Sharma, Basheerulla Shaik, Satya P Gupta, Vijay K. Agrawal, Quantitative Structure Activity Relationship and Docking Studies on a Series of Substituted Acyl(thio)ureas and Thiadiazolo(2,3- α)pyrimidine Derivatives as Potent Inhibitors of Influenza Virus Neuraminidase. *Journal of Applied Biopharmaceutics and Pharmacokinetics*, 4, 1-12, 2016.
- Yong Chen, Xiaoyan Wang, Wei Xiang, Lin He, Minghai Tang, Fang Wang, Taijin Wang, Zhuang Yang, Yuyao Yi, Hairong Wang, Ting Niu, Li Zheng, Lei Lei, Xiaobin Li, Hang Song, and Lijuan Chen. Development of Purine-Based Hydroxamic Acid Derivatives: Potent Histone Deacetylase Inhibitors with Marked in Vitro and in Vivo Antitumor Activities. *Journal of Medicinal Chemistry* 2016 59 (11), 5488-5504. DOI: 10.1021/acs.jmedchem.6b00579
