

A Quantitative Structure-Activity Relationship and Molecular Modeling studies on a Series of Hydroxamate analogues Acting as HDACs inhibitors

¹Anubha Bajpai and ¹Neeraj Agarwal
Department of Biotechnology,
Meerut Institute of Engineering and
Technology,
Meerut-250005, India.
Email: anubha.mbi@gmail.com

²Dr. Amit Mishra
R&D Division,
DCM Shriram Industries Unit: Daurala Organics,
Meerut-250221, India.
Email: amitmishra1550@gmail.com

Abstract: The inhibition of HDACs is a rapidly growing and very promising area for cancer chemotherapy. In the QSAR study, we were observed that for the Hydroxamate analogues series which was HDAC inhibitory action, the correlation seems to be quite significant with indicator parameters and steric parameters such as Parachor, Index of refraction and Surface tension found to possess good predictive ability. A positive correlation found with Parachor, Index of refraction which increase activity with increases these parameters values which means presence of bulky groups increases the inhibitory property of molecules against HDACs while negative coefficient of surface tension suggests that the steric properties show great efficiency to inhibit HDACs activity. Values of $r^2_{cv} = 0.769$ and $r^2_{pred} = 0.765$ comes from series of Hydroxamate analogues which show good results from equation of series. On the basis of this equation comes from this series predict some higher activity compound. After finds higher activity compounds from multiple regression analysis (MRA) check their Lipinski values for having good lead compounds.

Keywords: Quantitative structure-activity relationship, Histone deacetylase

Introduction

The enzymes, histone deacetylases (HDACs) affect the acetylation status of histones and other important cellular proteins and have been recognized as potentially useful therapeutic targets for a broad range of human disorders. They mediate changes in nucleosome conformation and are important in the regulation of gene expression [1]. Further, HDACs are also involved in a cell cycle progression and differentiation and their deregulation is associated with several cancers [2]. Hydroxamates are a new class of anticancer agents reported to act by selective inhibition of the Histone deacetylase enzyme. HDACs has become a novel target for the discovery of drugs for the treatment of cancer and other diseases [3-8]. The number of HDAC enzyme subtypes has expanded considerably over the past few years, for the development of HDAC inhibitors with improved specificity. A number of natural inhibitors such as trichostatin A (TSA) [9], cyclic tetrapeptide trapoxin (TPX) [10], HC toxin [11]

and apicidin [12] have been reported. TSA has been identified as a potent and specific HDAC inhibitor. Synthetic inhibitors like sodium phenyl butyrate [13], sodium valproate [14], suberanilo hydroxamic acid (SAHA) [15], straight chain TSA and SAHA like analogues [16] and oxamflatin [17] have been reported. A QSAR can be derived correlating the differences in steric and electrostatic fields surrounding a set of molecules to the biological activity, this method can be used to develop a 3D pharmacophore model [18]. The multiple regression analysis and molecular modeling studies have been increasingly employed in rational drug designing process to understand the drug receptor interaction and to design new molecules with higher potency. Drug design has come to rely on the method of docking which seeks to mount molecules from a virtual database into the active sites of proteins and evaluate their binding affinities. This is becoming an increasingly useful way to predict potential inhibitors of proteins [19]. Lipinski's

1997 publication of the 'Rule of 5' (Ro5) was one of the most influential recent medicinal chemistry publications. Lipinski's publication of the Ro5 and subsequent work by other authors has made medicinal chemists more aware of the relationships between physical properties and ADME/T, it has hardly been a panacea. Pharmaceutical productivity continues to lag, and the industry is exploring new models to improve its output [20]. By the application of Quantitative Structure-Activity Relationship (QSAR) and MOLECULAR MODELING methodologies our aim is to predict more potent HDAC inhibitors.

Materials and Methods

We performed a simple multiple regression analysis (MRA) on the series of Hydroxamate analogues (1) as listed in Table 1 and reported for HDAC inhibitory activity by Juvale et al. [21]. The total 56 compounds of table 1 has been divided into two subsets: the training set comprising of 39 compounds and the test set

$$\begin{aligned} \text{pIC}_{50} &= 0.255 (\pm 0.149) \text{ PAR} + 17.770 (\pm 6.537) \text{ RI} - 0.797 (\pm 0.346) \text{ ST} - 1.795 (\pm 0.855) \\ \text{IP}_1 &+ 0.709 (\pm 0.334) \text{ IP}_2 - 0.198 (\pm 0.143) \text{ IP}_3 - 17.311 (8.680) \\ n &= 36, r = 0.925, r^2_{cv} = 0.769, r^2_{pred} = 0.765, s = 0.296 F = 28.506 (3.50) \end{aligned} \quad (1)$$

In Eq. (1), n is the number of data points, r is the correlation coefficient, r^2_{cv} is the square of the cross-validated correlation coefficient obtained from leave-one-out (LOO) jackknife procedure, s is the standard deviation, F is the Fischer ratio between the variances of calculated and observed activities, and the data within the parentheses with \pm sign are 95% confidence intervals. The figure within the parenthesis for F is the standard F-value at 99% level. In the activity term, IC_{50} refers to molar concentration of the compound leading to 50% inhibition of the enzyme. The values of these statistical parameters exhibit that the correlation obtained is quite significant. The internal validity of the correlation is judged by the value of its r^2_{cv} which is calculated as:

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

Where $y_{i,obsd}$ and $y_{i,pred}$ are the observed and predicted (from LOO) activity values of

comprising of 17 compounds, as was done by Juvale et al., two test compound taken as training set molecule i.e., 43 and 54 having significant difference in observed and calculated activity, who had performed CoMFA and CoMSIA Studies. The test set compounds in the table are marked with a superscript 'b' and are given in bold. For multiple regression analysis, a large number of physicochemical parameters were calculated using Chem-sketch (version 11.0) [22] tools. The physicochemical parameters that were found to be significant in MRA were simply Parachor, Index of refraction and Surface tension of the compounds. The values of these parameters for all the compounds are listed in Table 1.

Results and Discussion

When a multiple regression analysis was performed on the compounds of the training set, it revealed the following correlation.

compound i , respectively, and $y_{av,obsd}$ is the average of the observed activities of all compounds used in the correlation. The correlation is supposed to be valid if $r^2_{cv} > 0.60$. From this point of view, the correlation expressed by Eq. (1) seems to be quite valid. However, the predictive ability of any correlation equation is judged by predicting the activity of the compounds in the test set using it and calculating the value of r^2_{pred} , which is defined as:

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

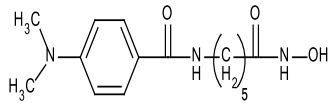
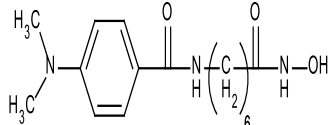
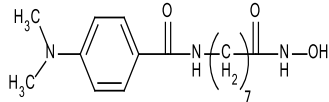
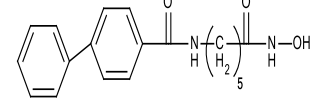
Where $y_{i,obsd}$ and $y_{i,pred}$ refer to the observed and predicted (from eq. obtained) activity of compound i in the test set and $y_{av,obsd}$ is same as in Eq.(2). A value of r^2_{pred} equal to 0.765, signifies a good predictive ability of the correlation. The activity values predicted from this equation for the test set compounds are given (in bold) in Table 1. A comparison shows

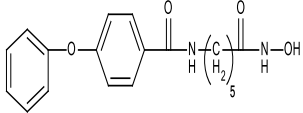
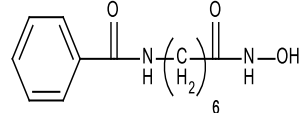
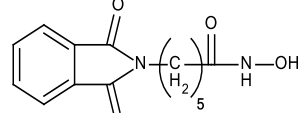
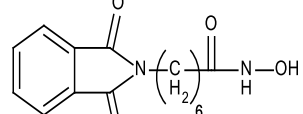
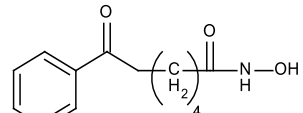
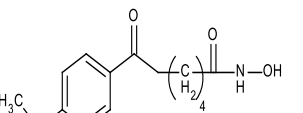
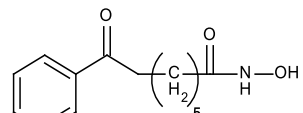
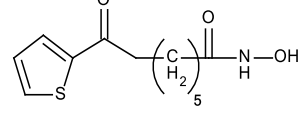
that these predicted values are in very good agreement with the corresponding observed ones. In the training set also, the calculated values are found to be in excellent agreement with the observed ones. All these observations can be better visualized in the graphs drawn between the predicted and observed activities (Fig. 1.). It is also to be noted that all the six parameters of the Eq.(1) are statistically quite significant in the correlation. Further, as shown in Table 2, these variables have no significant mutual correlation. Using Eq.(1), we have predicted the activity of some new prospective compounds with high potency (Table 3). The activities of these compounds are higher than any compound in the present series (Table 1).

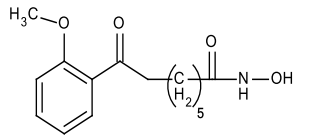
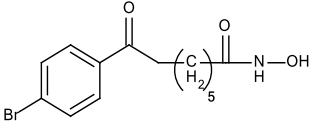
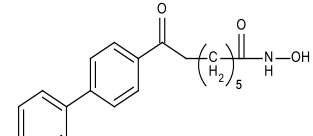
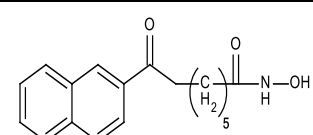
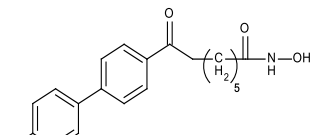
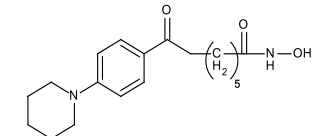
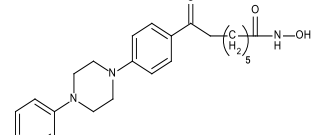
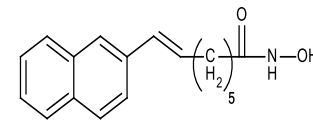
Conclusion

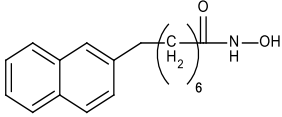
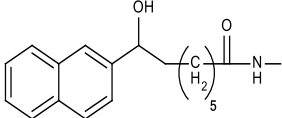
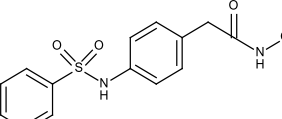
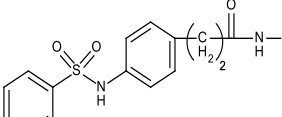
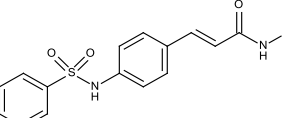
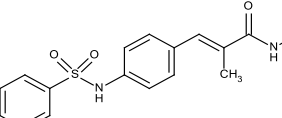
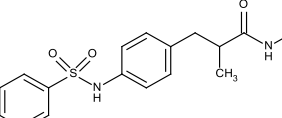
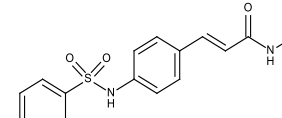
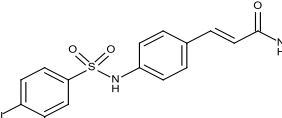
From this QSAR study, it is found that for this series which was HDAC inhibitory action, for the series of Hydroxamate analogues reported for HDAC inhibitory activity by Juvalé, the correlation seems to be quite significant with indicator parameters and steric parameters such as Parachor, Index of refraction and Surface tension found to possess good predictive ability. Indicator parameters IP_1 , IP_2 and IP_3 for presence of So2 group, single chain for 7 bonds and number of single bonds respectively. A positive coefficient of Parachor, Index of refraction and indicator parameter IP_2 indicates that has shown conducive effect respectively while negative coefficient of Surface tension and Indicator parameters IP_1 and IP_3 suggests that these steric properties show great efficiency to inhibit HDACs active.

Table 1: Hydroxamate Analogues and Their Observed and Predicted HDAC Inhibition Activities and Their Physicochemical Properties

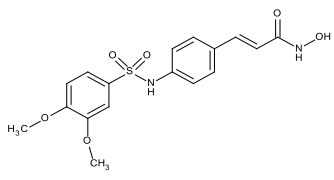
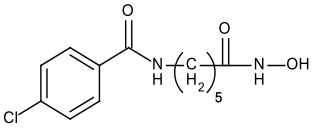
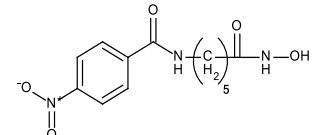
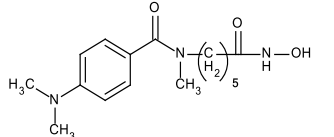
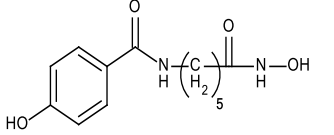
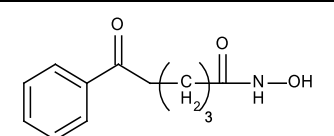
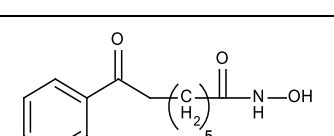
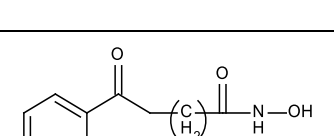
Sr. No.	Structure	PAR	RI	ST	IP_1	IP_2	IP_3	pIC_{50}	Cald.	LOO
1		6.711	1.563	4.890	0.000	0.000	6.000	7.300	7.089	7.055
2		7.109	1.558	4.780	0.000	1.000	7.000	8.000	7.701	7.684
3		7.506	1.553	4.690	0.000	1.000	8.000	7.540	7.587	7.599
4		7.424	1.574	4.870	0.000	0.000	6.000	7.280	7.483	7.527

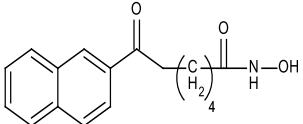
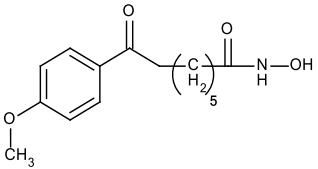
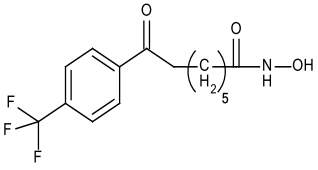
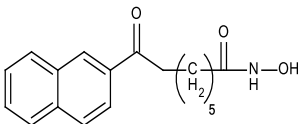
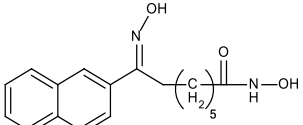
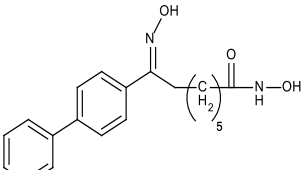
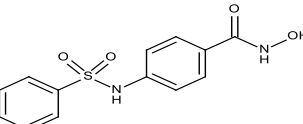
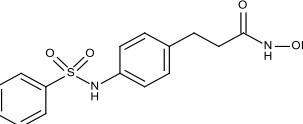
5		7.602	1.575	4.960	0.000	0.000	6.000	7.360	7.474	7.503
6		6.088	1.539	4.720	0.000	1.000	7.000	6.950	7.151	7.176
7		5.800	1.589	5.860	0.000	0.000	6.000	6.990	6.546	6.373
8		6.200	1.580	5.650	0.000	1.000	7.000	7.150	7.167	7.166
9		5.411	1.536	4.670	0.000	0.000	6.000	6.300	6.454	6.496
10		5.978	1.529	4.480	0.000	0.000	6.000	6.350	6.625	6.708
11		5.750	1.533	4.960	0.000	1.000	7.000	6.820	6.767	6.757
12		5.640	1.547	4.920	0.000	1.000	7.000	7.300	7.019	6.967

13		6.375	1.525	4.400	0.000	1.000	7.000	7.020	7.230	7.263
14		6.314	1.555	4.820	0.000	1.000	7.000	7.350	7.413	7.416
15		7.542	1.563	4.650	0.000	1.000	7.000	8.300	8.004	7.974
16		6.847	1.591	4.990	0.000	1.000	7.000	8.100	8.053	8.047
17		8.047	1.580	4.850	0.000	1.000	7.000	8.700	8.275	8.212
18		7.752	1.548	4.720	0.000	1.000	7.000	8.020	7.735	7.694
19		9.353	1.583	5.040	0.000	1.000	7.000	8.350	8.510	8.609
20		6.627	1.622	4.890	0.000	0.000	6.000	8.220	8.116	8.024

21		6.761	1.583	4.650	0.000	1.000	8.000	7.600	7.962	8.067
22		6.915	1.601	5.230	0.000	1.000	7.000	7.460	8.057	8.170
23		6.141	1.651	6.680	1.000	0.000	2.000	6.000	6.078	6.103
24		6.943	1.626	6.160	1.000	0.000	4.000	6.000	5.857	5.437
25		6.418	1.667	6.860	1.000	0.000	0.000	6.700	6.686	6.682
26 ^c		6.781	1.651	6.450	1.000	0.000	1.000	5.700	6.623	-
27 ^c		6.922	1.624	6.050	1.000	0.000	4.000	4.770	5.904	-
28		6.789	1.671	6.970	1.000	0.000	0.000	7.120	6.764	6.716
29		7.161	1.674	7.080	1.000	0.000	0.000	7.000	6.824	6.805

30		7.161	1.674	7.080	1.000	0.000	0.000	6.520	6.824	6.863
31		6.989	1.683	7.840	1.000	0.000	0.000	6.150	6.335	6.403
32		7.494	1.682	7.830	1.000	0.000	0.000	6.100	6.454	6.609
33		7.241	1.597	5.710	1.000	0.000	0.000	6.220	6.568	6.680
34		6.801	1.653	6.530	1.000	0.000	0.000	6.520	6.798	6.836
35		6.801	1.653	6.530	1.000	0.000	0.000	7.000	6.798	6.775
36		7.966	1.618	5.720	1.000	0.000	0.000	7.000	7.118	7.155
37 ^c		9.909	1.582	4.970	1.000	0.000	0.000	6.220	7.572	-
38		7.004	1.643	6.470	1.000	0.000	0.000	7.220	6.720	6.672

39		7.590	1.624	6.150	1.000	0.000	0.000	7.050	6.786	6.742
40 ^b		6.049	1.555	4.980	0.000	0.000	6.000	7.130	6.707	6.707
41 ^b		6.245	1.570	5.670	0.000	0.000	6.000	6.520	6.473	6.473
42 ^b		7.086	1.563	4.830	0.000	0.000	6.000	7.050	7.233	7.233
43 ^c		5.841	1.568	5.530	0.000	0.000	6.000	7.520	6.446	-
44 ^b		5.013	1.542	4.800	0.000	0.000	5.000	5.820	6.553	6.553
45 ^b		5.809	1.531	4.560	0.000	1.000	7.000	7.190	7.065	7.065
46 ^b		6.207	1.527	4.470	0.000	1.000	8.000	6.870	6.969	6.969

47 ^b		6.449	1.599	5.120	0.000	0.000	6.000	7.460	7.479	7.495
48 ^b		6.375	1.525	4.400	0.000	1.000	7.000	7.820	7.230	7.230
49 ^b		6.381	1.489	3.800	0.000	1.000	7.000	7.350	7.070	7.070
50 ^b		6.847	1.591	4.990	0.000	1.000	7.000	8.300	8.053	8.053
51 ^b		6.962	1.582	4.650	0.000	1.000	7.000	8.070	8.193	8.193
52 ^b		7.734	1.571	4.490	0.000	1.000	7.000	8.400	8.322	8.322
53 ^b		5.740	1.660	6.800	1.000	0.000	0.000	6.050	6.436	6.436
54 ^c		6.542	1.637	6.390	1.000	0.000	3.000	7.000	5.965	-

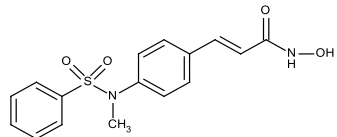
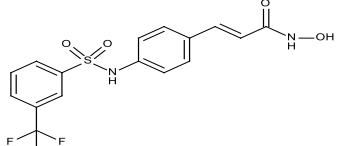
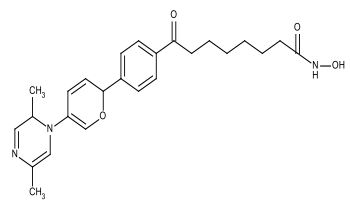
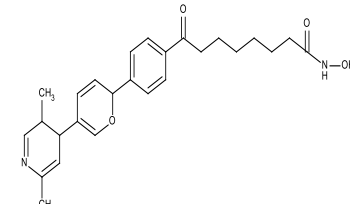
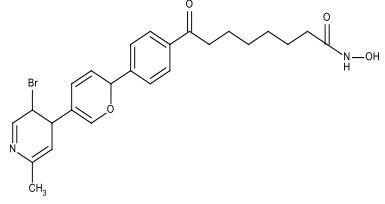
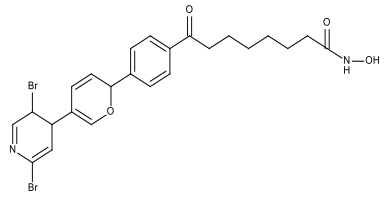
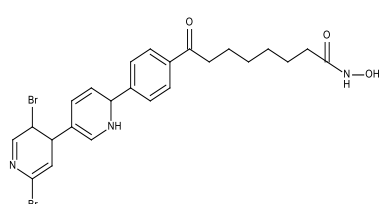
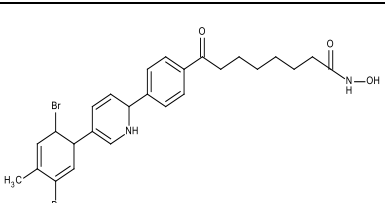
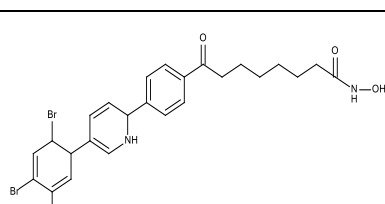
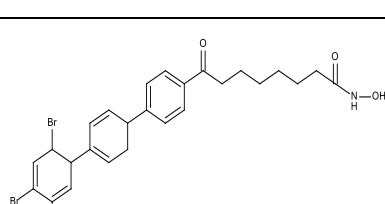
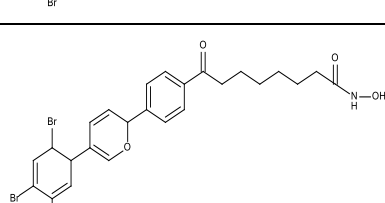
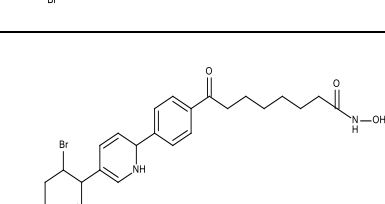
55 ^b		6.799	1.648	6.280	1.000	0.000	0.000	6.220	6.908	6.908
56 ^b		7.038	1.602	5.670	1.000	0.000	0.000	7.000	6.637	6.637

Table 2. Correlation Matrix Showing the Mutual Correlations among the Variables Used

	PAR	RI	ST	IP ₁	IP ₂	IP ₃
PAR	1.000	- 0.422	0.397	0.014	- 0.289	0.134
RI		1.000	- 0.770	- 0.108	0.197	- 0.072
ST			1.000	- 0.084	- 0.201	0.193
IP ₁				1.000	- 0.240	0.826
IP ₂					1.000	- 0.505
IP ₃						1.000

Table 3: Some Proposed Compounds belonging to the Series of Table 1 and Their Predicted Activity

Sr.No.	Structure	PAR	RI	ST	IP ₁	IP ₂	IP ₃	pIC ₅₀
1		9.500	1.591	4.540	0.000	1.000	7.000	9.088
2		9.544	1.584	4.360	0.000	1.000	7.000	9.118

3		9.669	1.608	4.730	0.000	1.000	7.000	9.282
4		9.793	1.632	5.120	0.000	1.000	7.000	9.429
5		9.799	1.646	5.240	0.000	1.000	7.000	9.584
6		10.669	1.613	5.230	0.000	1.000	7.000	9.227
7		10.797	1.642	5.810	0.000	1.000	7.000	9.313
8		10.915	1.641	5.830	0.000	1.000	7.000	9.309
9		10.718	1.642	5.940	0.000	1.000	7.000	9.189
10		10.810	1.599	5.050	0.000	1.000	7.000	9.158

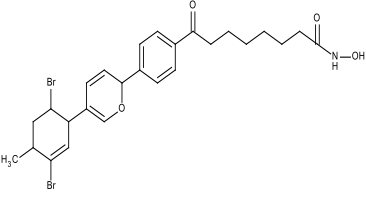
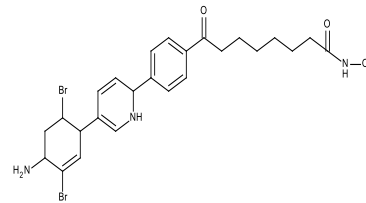
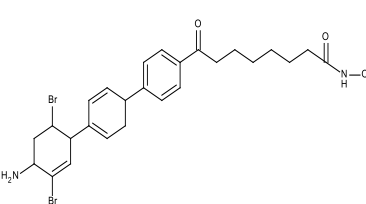
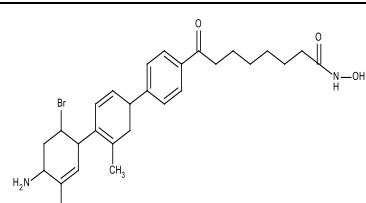
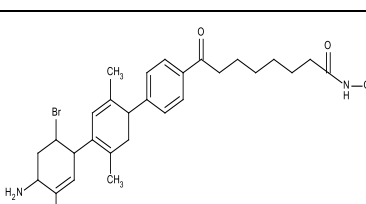
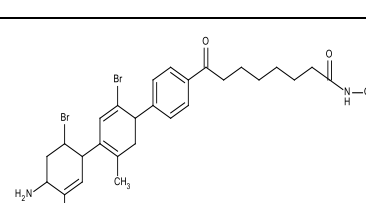
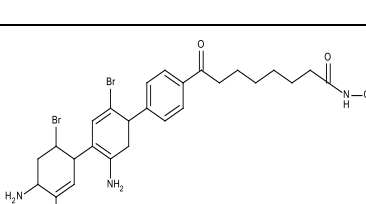
11		10.731	1.599	5.160	0.000	1.000	7.000	9.050
12		10.707	1.616	5.510	0.000	1.000	7.000	9.067
13		10.825	1.616	5.530	0.000	1.000	7.000	9.081
14		11.188	1.607	5.280	0.000	1.000	7.000	9.213
15		11.552	1.600	5.070	0.000	1.000	7.000	9.349
16		11.680	1.625	5.580	0.000	1.000	7.000	9.419
17		11.576	1.643	6.060	0.000	1.000	7.000	9.330

Table 4: Data related to Lipinski rules of predicted compounds and last two compound are FDA approved compounds

Sr.No.	MW	HD	HA	log P	pIC₅₀
1	437.531	2	6	2.52	9.088
2	436.543	2	5	3.25	9.118
3	501.413	2	5	3.45	9.282
4	566.282	2	5	4.34	9.429
5	565.297	3	5	4.02	9.584
6	578.336	3	4	4.75	9.227
7	643.205	3	4	4.91	9.313
8	642.217	2	3	6.03	9.309
9	644.190	2	4	5.23	9.189
10	580.352	3	4	4.92	9.158
11	581.337	2	4	5.24	9.050
12	581.340	4	5	3.3	9.067
13	580.352	3	4	4.42	9.081
14	594.378	3	4	4.66	9.213
15	608.405	3	4	4.91	9.349
16	673.275	3	4	4.99	9.419
17	674.263	4	5	3.28	9.330
SAHA	264.32	3	3	2	6.730
TSA	302.368	4	2	2.41	8.651

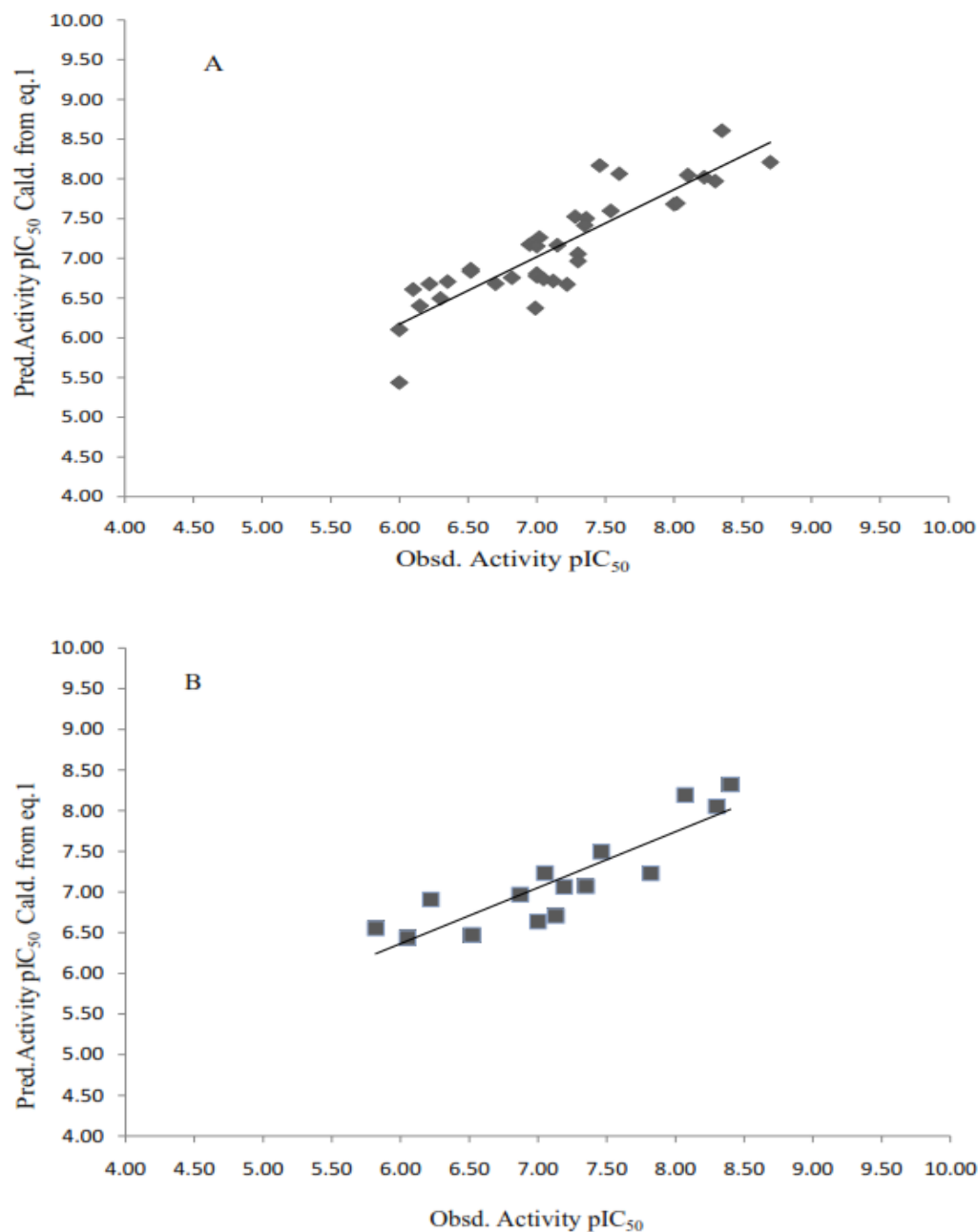


Fig.1. Training (A) and Test (B) set Graph between observed Vs predicted activity [Eq.1]

REFERENSES

1. Davie, J. R.; Chadee, D. N .Regulation and regulatory parameters of Histone modificatins, *J. Cell Biochem. (Suppl.)*, **1998**, *203*, 30-31.
2. Kourzarides, T. Histone acetylase and deacetylase in cell poriferlation, *Curr. Opin. Genet.*, **1999**, *9*, 40-48.
3. P. A. Marks, R. A. Rifkind, V. M. Richon, R. Breslow, T. Miller and W. K. Kelly, *Nat. Rev. Cancer*, **2001**, *1*, 194-202.
4. W. K. Kelly, O. A. O'Connor and P. A. Marks, *Expert Opin. Invest. Drugs*. **2002**, *11*, 1695-1713.
5. P. A. Marks, V. M. Richon, R. Breslow and R. A. Rifkind, *Curr. Opin. Oncol.*, **2001**, *13*, 477-483.

6. P. A. Marks, V. M. Richon, R. Breslow and R. A. Rifkind, *Clin. Cancer Res.*, **2001**, *7*, 759-760
7. C. J. Phiel, F. Zhang, E. Y. Huang, M. G. Guenther, M. A. Lazar and P. S. Klein, *J. Biol. Chem.*, **2001**, *76*, 36734-36741.
8. P. T. Meinke and P. Liberator, *Curr. Med. Chem.*, **2001**, *8*, 211-235.
9. M. Yoshida, M. Kijima, M. Akita and T. Beppu, *J. Biol. Chem.*, **1990**, *265*, 17174-17179.
10. M. Kijima, M. Yoshida, K. Suguta, S. Horinouchi and T. Beppu, *J. Biol. Chem.*, **1993**, *268*, 22429-22435.
11. R. E. Shute, B. Dunlap and D. H. Rich, *J. Med. Chem.*, **1987**, *30*, 71-78.
12. J. W. Han, S. H. Ahn, S. Y. Wang and G. U. Seo, *Cancer Res.*, **2000**, *60*, 6068-6074.
13. S. D. Gore and M. A. Carducci, *Expert Opin. Invest. Drugs*. **2000**, *9*, 2923-2934.
14. M. Gottlicher, S. Minucci, P. Zhu, O. H. Kramer, A. Schimpf, S. Giavara, J. P. Sleeman, F. Lo Coco, C. Nervi, P. G. Pelicci and T. Hainzel, *EMBO J.*, **2001**, *20*, 6969-6978.
15. L. M. Butler, D. B. Agus, H. I. Scher, B. Higgins, A. Rose, C. Cordon-Cardo, H. T. Thaler, R. A. Rifkind, P. Marks and V. M. Richon, *Cancer Res.*, **2000**, *60*, 5165-5170.
16. M. Jung, G. Brosch, D. Kölle, H. Scherf, C. Gerhäuser and P. Loidl, *J. Med. Chem.*, **1999**, *42*, 4669-4679.
17. S. W. Remiszewski, L. C. Sambucetti, P. Atadja, K. W. Bair, W. D. Cornell, M. A. Green, K. L. Howell, M. Jung, P. Kwon, N. Trogani and H. Walker, *J. Med. Chem.*, **2002**, *45*, 753-757.
18. S. H. Woo, S. Frechette, E. A. Khalil, G. Bouchain, A. Vaisburg, N. Bernstein, O. Moradei, S. Leit, M. Allan, M. Fournel, M.-C. Trachy-Bourget, Z. Li, J. M. Besterman and D. Delorme, *J. Med. Chem.*, **2002**, *45*, 2877-2885.
19. Y. B. Kim, K. H. Lee, K. Sugita, M. Yoshida and S. Horinouchi, *Oncogene*, **1999**, *18*, 2461-2470.
20. Ondej Skehota, Radka Svobodova Vaekova, Stanislav Geidl, Michal Kudera, David Sehnal, Janídek, Jaroslav Koa, Crina-Maria Ionescu, Janídek, Jaroslav Koca, **2011**. QSPR designer employ your own descriptors in the automated QSAR modeling process pp.37.
21. Greg Costakes, **2005**. Overview of current docking programs eJournal of basic biotechnology.
22. <http://www.acdlabs.com>
23. PDB database is available at the URL <http://www.pdb.org/>
24. <http://www.molegro.com>.
25. <http://www.chemaxon.com>