Innovare Academic Sciences

International Journal of Pharmacy and Pharmaceutical Sciences
ISSN- 0975-1491 Vol 6, Issue 9, 2014

Original Article

3D-QSAR AND MOLECULAR DOCKING STUDIES ON 1, 2, 4 TRIAZOLES AS MetAP2 INHIBITORS

STEPHEN PHILIP, BHARATHKUMAR INTURI, BHAVYA K, GURUBASAVARAJ V PUJAR, MADHUSUDAN PUROHIT*

Department of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS University, Mysore 570015, India.
Email: mnpurohit04@yahoo.com

Received: 13 Jul 2014 Revised and Accepted: 14 Aug 2014

ABSTRACT

Objective: Angiogenesis inhibitors are a novel class of promising therapeutic agents for treating cancer and other human diseases. Biological transformations and pathways that play a role in angiogenesis are, therefore, particularly attractive targets as potential methods for inhibiting solid tumors. MetAP2 is of particular interest because the enzyme plays a key role in angiogenesis, the growth of new blood vessels, which is necessary for the progression of diseases including solid tumor cancers and rheumatoid arthritis. In this paper we report the quantitative structure activity relationship and docking studies of 1, 2, 4 triazole derivatives for designing novel MetAP2 inhibitors.

Methods: Tripos Sybyl X 2.1 program was used to conduct docking based CoMFA, CoMSIA and Topomer CoMFA QSAR modeling for a dataset of 77 triazoles.

Results: The CoMFA, CoMSIA and Topomer CoMFA models demonstrated good statistical results with cross-validated coefficient (q^2) of 0.703, 0.704, 0.746 and correlation coefficient (r^2) of 0.894, 0.889, 0.886 respectively and these models have been externally validated.

Conclusion: Based on the statistical results obtained from the above model, the CoMFA, CoMSIA and Topomer CoMFA model can be utilized to design new molecules having 1, 2, 4 triazoles as common core with significant MetAP2 inhibitory activity.

Keywords: QSAR, CoMFA, CoMSIA, 1,2,4-triazole, MetAP2.

INTRODUCTION

Cancer is one of the most formidable afflictions in the world [1] Cancer known medically as malignant neoplasm is a broad group of diseases characterized by abnormal cell growth. In cancer, cells divide and grow hysterically, forming malignant tumors, and invading surrounding parts of the body. The cancer may also metastasize to more distant parts of the body through the lymphatic system or blood stream. All tumors are not cancerous; benign tumors do not invade neighboring tissues and do not spread throughout the body. There are more than two hundred different known cancers that affect humans [2-3]. Recent statistics indicates that, cancer accounts for about 23% of total deaths in the USA and the second most common cause of death after cardio vascular diseases. Therefore, there is an increasing need for new therapies, especially those based on current knowledge of cancer molecular biology and pathology [4].

Angiogenesis is a key process in the progression of a number of diseases such as diabetic retinopathy, rheumatoid arthritis, and cancer [5-6]. Cancerous tumors do not metastasize without the formation of new blood vessels. Biological transformations and pathways that play a role in angiogenesis are, therefore, particularly attractive targets as potential methods for inhibiting solid tumor growth/ metastasis. Methionine amino peptidase type II, in humans is encoded by the MetAP2 gene [7-8]. MetAP2, a member of the dimetallohydrolase family, is a cytosolic metallo-enzyme that catalyzes the hydrolytic removal of N-terminal methionine residues from nascent proteins [9-11]. MetAP2 is found in all organisms and is especially important because of its critical role in tissue repair and protein degradation [9]. It is expressed at higher concentrations in tumor as compared to normal cells. Available reports also suggested that MetAP2 plays an important role in growth of different type of tumors. MetAP2 is also the target of two groups of anti-angiogenic natural products (ovalicin and fumagillin) and their analogs which irreversibly inhibit MetAP2 through covalent modification of an epoxide [12-15]. MetAP2 is of particular interest because the enzyme plays a key role in angiogenesis, the development of new blood vessels, which is essential for the progression of diseases including solid tumor cancers and rheumatoid arthritis [16]. Due to the critical role of these enzymes for angiogenesis, MetAP2 has been one of the major targets in the anticancer drug development area. With an improved understanding of the genes and pathways accountable for cancer instigation and progression, cancer drug development has undergone a paradigm change in the recent years, from mainly cytotoxic agent based therapy to therapy aimed at molecular and genetic targets.

Quantitative structure-activity relationship (QSAR) enables the investigators to establish a reliable quantitative structure-activity relationship and structure-property relationships to derive an in silco QSAR model to predict the activity of novel molecules prior to their synthesis. The process of QSAR model development can be divided into three stages namely; data preparation, model development, and validation, representing a standard practice of any QSAR modeling. 3D-QSAR methodologies have been successfully used to generate models of various chemotherapeutic agents [17-19].

We have carried out 3D-QSAR studies employing comparative molecular field analysis (CoMFA) [20], comparative molecular similarity indices analysis (CoMSIA) [21] and Topomer CoMFA [22] techniques in order to study and deduce a correlation between structure and biological activity of 1,2,4 triazoles as MetAP2 inhibitors. Partial least square (PLS) based statistical analysis was carried out on aligned molecules to identify the correlation. The contour maps generated help in explaining the observed variation in activity. Validity of contour map analysis has been carried out using molecular docking studies. Important features observed in the developed model have been used to design new molecules, which showed higher predictivity and binding affinity in terms of different binding scores and selectivity in terms of interaction within the MetAP2 active site [23].

Computational analysis

Data set

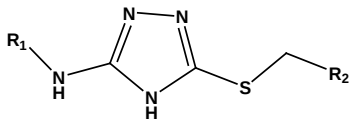
The *in vitro* biological activity reported as $K_{i,app}$ for the inhibition of MetAP2 by 1,2,4 triazoles were used for the current study. The *in vitro* assay involved inhibition of MetAP2 by an XTT assay. These compounds are shown in **Table 1**. All the molecules were obtained from previously reported literature [24] The biological activity obtained as $K_{i,app}$ values were converted into $pK_{i,app}$ ($-\log K_{i,app}$) and

used as dependent variable in the CoMFA, CoMSIA and Topomer CoMFA calculations. The *in vitro* $K_{i,app}$ values employed in this work were measured under the same experimental conditions [25], a fundamental requirement for QSAR studies. As a rule of thumb, a spread in affinity of at least three logarithmic units is considered necessary for developing a statistically significant 3D-QSAR model [26]. Here the biological data is spread over a range of nearly five logarithmic units from 5.76 to 10.39. The molecular structure of 1,2,4 triazole derivatives were sketched using ChemSketch.

The generation of consistent statistical model depends on the proper selection of both training and test sets in terms of structural diversity and property values distribution.

From the data 75% of the compounds were selected as members of the training set for model construction and the other 25% compounds as test set for external validation in the ratio of 3:1. The compounds of test and training sets were selected based on unity fingerprints and dissimilarity, which is available in SYBYL X2.1.

Table 1: Structures and experimental biological activity of compounds against MetAP2



Cmpd Code	R1	R2	$K_{i,app}$ (nm)	Cmpd Code	R1	R2	$K_{i,app}$ (nm)
6	Ph	Ph	0.5	52	4-OMe-Ph	Cyclohexyl	3.1
12	Ph	thiophen-2-yl	0.75	53	4-OMe-Ph	-CH=C(CH_3)$_2$	0.16
13	2-Me-Ph	Ph	0.04	54	4-OMe-Ph	5-methyl-3-isoxazol-3-yl	1.6
14	2-Me-Ph	2-Me-Ph	2.5	55	3,4-OMe-Ph	Ph	1.7
15	2-Me-Ph	2-MeO-Ph	61	56	3,4-OMe-Ph	2-Me-Ph	18
16	2-Me-Ph	2-F-Ph	0.23	58	3,4-OMe-Ph	2-F-Ph	1.3
17	2-Me-Ph	3,4-difluoro-Ph	2.4	59	3,4-OMe-Ph	3,4-difluoro-Ph	4.0
18	2-Me-Ph	4-pyridyl	3.7	60	3,4-OMe-Ph	2-pyridyl	24
19	2-Me-Ph	Cyclohexyl	3.3	61	3,4-OMe-Ph	4-pyridyl	9.6
20	2-Me-Ph	-CH=C(CH_3)$_2$	0.15	62	3,4-OMe-Ph	Cyclohexyl	8.3
21	2-Me-Ph	2-methyl-thiazo-4-yl	3.8	63	3,4-OMe-Ph	-CH=C(CH_3)$_2$	0.9
22	2-Me-Ph	5-methyl-3-isoxazol-3-yl	3.6	64	3,4-OMe-Ph	2-methyl-thiazo-4-yl	21
23	4-Me-Ph	Ph	0.07	65	3,4-OMe-Ph	5-methyl-3-isoxazol-3-yl	7.0
24	4-Me-Ph	2-Me-Ph	6.5	66	4-CO2Me-Ph	Ph	0.7

Table 1: (Continued)

Cmpd Code	R1	R2	$K_{i,app}$ (nm)	Cmpd Code	R1	R2	$K_{i,app}$ (nm)
26	4-Me-Ph	2-F-Ph	0.65	67	4-CO2Me-Ph	2-Me-Ph	135
27	4-Me-Ph	3,4-difluoro-Ph	3.3	68	4-CO2Me-Ph	2-MeO-Ph	10
28	4-Me-Ph	2-pyridyl	7.8	69	4-CO2Me-Ph	2-F-Ph	0.41
29	4-Me-Ph	4-pyridyl	5.7	70	4-CO2Me-Ph	3,4-difluoro-Ph	1.3
30	4-Me-Ph	Cyclohexyl	100	71	4-CO2Me-Ph	2-pyridyl	0.9
31	4-Me-Ph	-CH=C(CH_3)$_2$	1.2	72	4-CO2Me-Ph	4-pyridyl	0.9
31	4-Me-Ph	2-methyl-thiazo-4-yl	7.8	73	4-CO2Me-Ph	Cyclohexyl	2.7
33	4-Me-Ph	5-methyl-3-isoxazol-3-yl	3.5	74	4-CO2Me-Ph	-CH=C(CH_3)$_2$	0.5
34	4-Me-Ph	thiophen-2-yl	5.2	75	4-CO2Me-Ph	2-methyl-thiazo-4-yl	3.8
35	4-Cl-Ph	Ph	0.43	76	4-CO2Me-Ph	5-methyl-3-isoxazol-3-yl	0.8
36	4-Cl-Ph	2-Me-Ph	23	77	2-(Ph)-Ph	Ph	740
38	4-Cl-Ph	2-F-Ph	0.61	80	2-(Ph)-Ph	2-F-Ph	350
39	4-Cl-Ph	3,4-difluoro-Ph	1.5	83	2-(Ph)-Ph	Cyclohexyl	650
40	4-Cl-Ph	2-pyridyl	4.0	84	2-(Ph)-Ph	-CH=C(CH_3)$_2$	480
41	4-Cl-Ph	4-pyridyl	2.5	85	2-(Ph)-Ph	5-methyl-3-isoxazol-3-yl	1700
42	4-Cl-Ph	Cyclohexyl	5.2	86	3-pyridyl	Ph	0.04
43	4-Cl-Ph	CH=C(CH_3)$_2$	0.52	87	3-pyridyl	2-Me-Ph	1.7
44	4-Cl-Ph	2-methyl-thiazo-4-yl	4.8	88	3-pyridyl	2-F-Ph	0.3

Table 1: (Continued)

Cmpd Code	R1	R2	$K_{i,app}$ (nm)	Cmpd Code	R1	R2	$K_{i,app}$ (nm)
45	4-Cl-Ph	5-methyl-3-isoxazol-3-yl	1.8	89	3-pyridyl	3,4-difluoro-Ph	0.5
46	4-OMe-Ph	Ph	0.05	90	3-pyridyl	2-pyridyl	5.7
47	4-OMe-Ph	2-MeO-Ph	34	91	3-pyridyl	4-pyridyl	3.0
48	4-OMe-Ph	2-F-Ph	0.13	92	3-pyridyl	Cyclohexyl	6.5
49	4-OMe-Ph	3,4-difluoro-Ph	2.1	93	3-pyridyl	2-methyl-thiazo-4-yl	9.1
50	4-OMe-Ph	2-pyridyl	18	94	3-pyridyl	5-methyl-3-isoxazol-3-yl	1.5
51	4-OMe-Ph	4-pyridyl	6.1				

Molecular Docking

To determine the most probable binding confirmations of the whole data set, SYBYL-X 2.1 Surflex-Dock has been used.

The protein structure of human MetAP2 was obtained from protein data bank (PDB Code: 2OAZ). This structure is determined at 1.90 Å

resolution. We performed the minimization of MetAP2 using 2OAZ as template using the prepare protein module in SYBYL-X 2.1. Surflex-Dock. The ligand substructure was extracted and all unnecessary water molecules were removed. Explicit hydrogen's were added to the protein and protein model was charged with Gasteiger-Marsili and ligand using Gasteiger-Huckel charges. Energy minimization and relaxation of the loop region was performed using 1000 iterations with AMBER7 FF99 as force field. All the molecules were docked into the active site of MetAP2. The docking confirmation of most potent compound (**86)** was used for aligning all the molecules in the 3D-QSAR study.

Validation of docking results

Generally, the performance of the docking software is judged from the resemblance of docked pose to the corresponding co-crystallized ligand[27]. The available X-ray crystal structure of 2OAZ in complex with co-crystalized ligand was under taken to validate the docking reliability. Eventually all the docked solutions reside almost at the same coordinating position as the co-crystallized ligand. The selected docked pose of co-crystallized ligand remains the same at the binding site. This enables us to ensure that the docked possess of compounds resemble that of the compound 86 in the active site of MetAP2 (**Fig 1**).

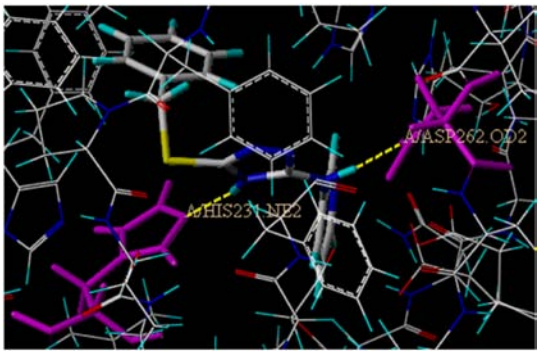


Fig. 1: Docking interaction of compound 86 with HIS 231 and ASP 262 in the MetAP2 active site (PDB Code: 2OAZ).

Alignment

Since there is a critical requirement of structure alignment in CoMFA analysis to generate a 3D QSAR model, the alignment rules remains to be a crucial process in 3D QSAR analysis. Because of the structural similarity of compounds in our dataset, all of the compounds in both training set and test set were assumed to interact with MetAP2 through the same binding motifs.

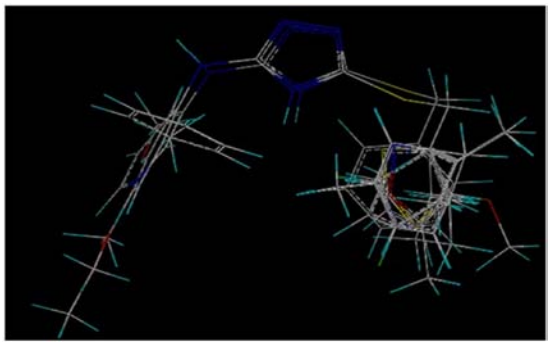


Fig. 2: 3D-QSAR structure alignment and superposition of 77 compounds using compound 86 as the template.

The docked conformation of most active compound **86** was regarded as a structural template for molecular superimposition. The molecules in their respective lowest energy conformation were superimposed on the template using the rigid body fit option in SYBYL-X 2.1. The alignment of the compounds is shown in **Fig 2**.

CoMFA field calculation

The alignment training set of molecules were positioned inside grid boxes and grid spacing value of 2 Å in all Cartesian directions. The grid box dimensions were determined automatically in such a way that the region boundaries were extended beyond 4 Å in each direction for the co-ordinates of each molecule.

The steric (vdW interaction) and electrostatic (Columbic terms) fields were calculated at each intersection using standard Tripos force field. A distance dependent dielectric constant of 1.00 was used. A sp^3 hybridized carbon atom with +1 charge served as probe atom to calculate steric and electrostatic fields. The cut off value for both steric and electrostatic interaction was set to 30 kcal/mol.

CoMSIA field calculation

All five physicochemical descriptors (electrostatic, steric, hydrophobic and hydrogen bond donor and acceptor) were evaluated at each lattice intersection of a regularly spaced grid of 2.0 Å. A probe atom with radius of 1 Å, +1 charge, hydrophobicity +1.0, H-bond donor and acceptor properties of +1.0 was used to evaluate the CoMSIA physicochemical descriptors. For attenuation factor (α) controlling the steepness of Gaussian function, the standard value of 0.3 was accepted.

The steric indices are related to the third power of the atomic radii, electrostatic descriptors are derived from atomic partial charge, hydrophobic fields are derived from atom based parameters, and H-bond donor and acceptor indices are obtained by a rule based method based on experimental results.

Topomer CoMFA

Topomer CoMFA is an alignment independent 3D-QSAR method that combines the topomer search method (a fragment alignment approach) with the conventional CoMFA method. A 3D-QSAR model was generated by splitting the molecules into fragments, topomerically aligning each fragment, and calculating steric and electrostatic field descriptor values for the topomerically aligned fragments to create a CoMFA table with the field descriptor values. Besides the common core of the ligand, we split side functional groups into two R-groups that refer to the R1 and R2 groups.

To evaluate the predictive ability of the Topomer CoMFA model, structure optimization of the test set was carried out as described previously for the training set. The $pK_{i,app}$ values of the test set were predicted on the basis of the constructed model.

Internal validation

Partial Least Square (PLS) regression [28-30] was used to analyse the training set by correlating the variation in the $pK_{i,app}$ values with variations in their CoMFA,CoMSIA and Topomer CoMFA interaction fields. The cross-validation analysis was performed using leave-one-out (LOO) method wherein one compound is removed from the data set and its activity is predicted using the model derived from the rest of the data set. The cross validated r^2 that resulted in optimum number of components and lowest standard error of prediction was taken. To speed up the analysis and reduce noise, a minimum column filtering value (r) of 2.00 kcal/mol was used for the cross-validation. Final analysis (non-cross-validation) was performed to calculate conventional r^2_{ncv} using the optimum number of components obtained from the LOO cross-validation analysis.

External Validation of the 3D-QSAR models

The predictive abilities of the CoMFA, CoMSIA and Topomer CoMFA models were determined from a test set of compounds not included in the model generation. The experimental and predicted $pK_{i,app}$ values based on the selected CoMFA and CoMSIA and topomer CoMFA models for the test set and training set of compounds are listed in **Table 2**.

Table 2: The experimental p $K_{i, app}$, predicted p $K_{i, app}$ and their residuals for the training and test set molecules using CoMFA, CoMSIA and topomer CoMFA.

Cmpd Code	Actual pIC_{50}	CoMFA		CoMSIA		Topomer CoMFA	
		Predicted pIC_{50}	Residual	Predicted pIC_{50}	Residual	Predicted pIC_{50}	Residual
Training Set							
6.	9.301	9.6495	0.3485	9.7179	0.4169	9.7046	0.4036
13	10.397	9.9483	-0.4496	9.8407	-0.5572	9.9064	-0.4915
14	8.6021	8.1868	-0.4153	8.1996	-0.4025	8.1767	-0.4254
15	7.2147	7.3478	0.1331	7.222	0.0073	7.2837	0.069
16	9.6383	9.6078	-0.0305	9.5797	-0.0586	9.6175	-0.0208
17	8.6198	8.9371	0.3173	8.9578	0.338	8.9713	0.3515
18	8.4318	8.7441	0.3123	8.7749	0.3431	8.8022	0.3704
19	8.4815	8.2516	-0.2299	8.1978	-0.2837	8.209	-0.2725
20	9.8239	9.7373	-0.0866	9.6767	-0.1472	9.7226	-0.1013
21	8.4202	8.4139	-0.0063	8.3903	-0.0299	8.3455	-0.0747
22	8.4437	8.8313	0.3876	8.8176	0.3739	8.8142	0.3705
23	10.154	9.476	-0.6789	9.4275	-0.7274	9.4771	-0.6778
26	9.2007	9.1372	-0.0635	9.1672	-0.0335	9.1882	-0.0125
27	8.4815	8.5299	0.0484	8.5542	0.0727	8.542	0.0605
28	8.1079	7.7828	-0.3251	7.7692	-0.3387	7.7658	-0.3421
29	8.2441	8.3057	0.0616	8.3717	0.1276	8.373	0.1289
30	7	7.7929	0.7929	7.7915	0.7915	7.7798	0.7798
31	8.9208	9.2476	0.3268	9.2707	0.3499	9.2933	0.3725
32	8.1079	7.8917	-0.2162	7.9791	-0.1288	7.9162	-0.1917
33	8.4559	8.3642	-0.0917	8.4163	-0.0396	8.385	-0.0709
34	8.284	8.2602	-0.0238	8.3674	0.0834	8.2729	-0.0111
35	9.3665	9.7021	0.3356	9.5925	0.226	9.6775	0.311
36	7.6383	7.9296	0.2913	7.9487	0.3104	7.9478	0.3095
38	9.2147	9.398	0.1833	9.3386	0.1239	9.3886	0.1739
39	8.8239	8.6802	-0.1437	8.7006	-0.1233	8.7424	-0.0815
41	8.6021	8.5473	-0.0548	8.5403	-0.0618	8.5734	-0.0287
42	8.284	8.0226	-0.2614	7.9658	-0.3182	7.9801	-0.3039
44	8.3188	8.1468	-0.172	8.1456	-0.1732	8.1166	-0.2022
45	8.7447	8.5652	-0.1795	8.575	-0.1697	8.5853	-0.1594
46	10.301	9.9717	-0.3293	9.9277	-0.3733	9.9522	-0.3488
47	7.4685	7.3463	-0.1222	7.3001	-0.1684	7.3296	-0.1389
48	9.8861	9.6321	-0.254	9.6659	-0.2202	9.6633	-0.2228
49	8.6778	8.9474	0.2696	9.0398	0.362	9.0171	0.3393
50	7.7447	8.2824	0.5377	8.2875	0.5428	8.2409	0.4962
52	8.5086	8.2869	-0.2217	8.3004	-0.2082	8.2549	-0.2537
53	9.7959	9.7489	-0.047	9.7666	-0.0293	9.7684	-0.0275
54	8.7959	8.8323	0.0364	8.907	0.1111	8.8601	0.0642
55	8.7696	9.2728	0.5032	9.172	0.4024	9.2392	0.4696
56	7.7447	7.4658	-0.2789	7.5148	-0.2299	7.5095	-0.2352
59	8.3979	8.2545	-0.1434	8.2823	-0.1156	8.3041	-0.0938
60	7.6198	7.5341	-0.0857	7.5082	-0.1116	7.5278	-0.092
61	8.0177	8.0539	0.0362	8.1028	0.0851	8.1351	0.1174
63	9.0458	9.0087	-0.0371	8.9934	-0.0524	9.0554	0.0096
65	8.1549	8.0968	-0.0581	8.137	-0.0179	8.1471	-0.0078
66	9.1549	9.5046	0.3497	9.3947	0.2398	9.3758	0.2209
67	6.8697	7.7475	0.8778	7.7475	0.8778	7.6461	0.7764
70	8.8861	8.5509	-0.3352	8.5276	-0.3585	8.4407	-0.4454
72	9.0458	8.3528	-0.693	8.3414	-0.7044	8.2717	-0.7741
74	9.301	9.3127	0.0117	9.2419	-0.0591	9.192	-0.109
77	6.1308	6.5054	0.3746	6.5017	0.3709	6.5203	0.3895
84	6.3188	6.1707	-0.1481	6.3327	0.0139	6.3365	0.0177
85	5.7696	5.5644	-0.2052	5.4621	-0.3075	5.4281	-0.3415
86	10.397	10.313	-0.0849	10.4296	0.0317	9.9751	-0.4228
87	8.769	8.21	-0.5596	8.1654	-0.6042	8.2454	-0.5242
88	9.522	9.5072	-0.0157	9.5615	0.0386	9.6861	0.1632
90	8.244	8.2674	0.0233	8.1964	-0.0477	8.2637	0.0196
91	8.522	8.6671	0.1442	8.7474	0.2245	8.8709	0.348
93	8.041	8.3858	0.3448	8.3775	0.3365	8.4141	0.3731
Test Set							
12	9.1249	8.4451	-0.6798	8.6582	-0.4667	8.5003	-0.6246
24	8.1871	7.7128	-0.4743	7.7901	-0.397	7.7475	-0.4396
40	8.3979	7.9951	-0.4028	7.9425	-0.4554	7.9661	-0.4318
43	9.284	9.4726	0.1886	9.4331	0.1491	9.4937	0.2097
51	8.2147	8.7674	0.5527	8.8612	0.6465	8.8481	0.6334
58	8.8861	8.9292	0.0431	8.9042	0.0181	8.9503	0.0642
62	8.0809	7.5503	-0.5306	7.5212	-0.5597	7.5418	-0.5391
64	7.6778	7.6808	0.003	7.7121	0.0343	7.6783	0.0005

68	8	6.9486	-1.0514	6.7918	-1.2082	6.7532	-1.2468
69	9.3872	9.167	-0.2202	9.1339	-0.2533	9.0869	-0.3003
71	9.0458	7.8401	-1.2057	7.7601	-1.2857	7.6644	-1.3814
73	8.5686	7.8341	-0.7345	7.7546	-0.814	7.6784	-0.8902
75	8.4202	7.9524	-0.4678	7.9432	-0.477	7.8149	-0.6053
76	9.0969	8.4017	-0.6952	8.376	-0.7209	8.2837	-0.8132
80	6.4559	6.1459	-0.31	6.2384	-0.2175	6.2314	-0.2245
83	6.1871	4.7418	-1.4453	4.8427	-1.3444	4.8229	-1.3642
89	9.301	8.9122	-0.3888	8.9453	-0.3557	9.0399	-0.2611
92	8.1871	8.2432	0.0561	8.1769	-0.0102	8.2777	0.0906
94	8.8239	8.762	-0.0619	8.8049	-0.019	8.8829	0.059

For a QSAR model, internal validation of leave-one-out cross-validated q^2 is commonly used. But Golbraikh and Tropsha [31-32] reported that the high value of q^2 was necessary and important but not the sufficient condition for a model to have a high predictive power. To estimate the predictive ability of the QSAR model, squared correlation coefficient values between the observed and predicted values of the test set compounds with intercept (r^2) and without intercept (r^2_0) were calculated. According to Golbraikh and Tropsha, models are considered acceptable if they satisfy all of the following conditions:

$q^2>0.5$

$r^2>0.6$

$[(r^2-r_0^2)/r^2]<0.1$

$$0{:}85 \leq k \leq 1{:}15$$

Where k is the slope of regression lines through the origin, k is obtained according to the following formula:

$$k = \frac{\sum y y_i}{\sum y_i^2}$$

Where

y= Experimental activity of Test set of Compounds

y_i=Predicted Activity of Test set of compounds

R^2_0 is calculated as follows

$$R_0^2 = 1 - \frac{\sum (y_i - y_i^{ro})^2}{\sum (y_i - \bar{y}_i)^2}$$

Where

$$y_i^{ro} = k y_i$$

According to Roy and Roy [30, 33-34] for a model with good external predictability, a difference between r^2 and r^2_0 values needs to be studied. An additional statistic for external validation r^2_m was introduced by the following equation.

$$r_m^2 = 1 - \sqrt{|r^2 - r_0^2|}$$

A value of r^2_m greater than 0.5 may be taken as an indicator of good external predictability.

RESULTS AND DISCUSSION

CoMFA Statistical Results

For the CoMFA model, the highest cross-validated q^2 was obtained by using the combination of steric and electrostatic fields. The best CoMFA model gave $q^2$and r^2 values of 0.703 and 0.816 respectively. The standard error for prediction was 0.358 with twelve components. The corresponding field contributions are 66.1&33.9 for steric and electrostatic respectively. Therefore, the steric field had greater influence than the electrostatic field on MetAP2 inhibitory activity. CoMFA analysis results are also summarized in **Table 2** and **3**. **Fig 3A** shows the relationship between the predicted and the experimental $pK_{i,\,app}$ values for the CoMFA model. The cross-validation results suggest that a reliable CoMFA model was successfully constructed.

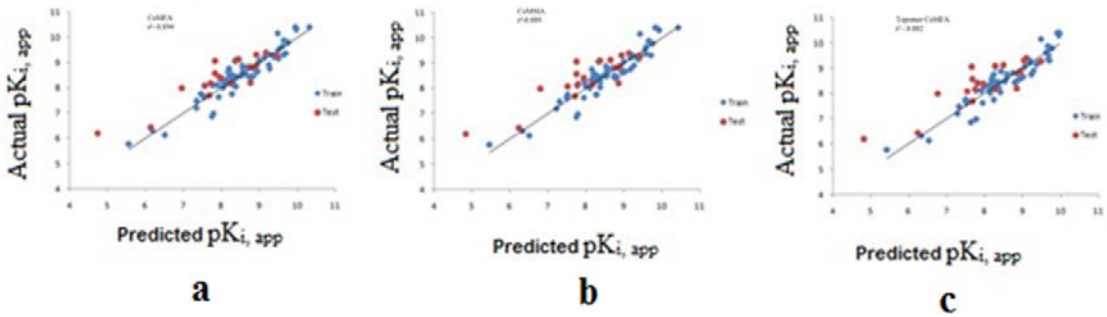


Fig. 3: Calculated $pK_{i,\,app}$ versus experimental p $K_{i,\,app}$ values for the 57 training set molecules and 19 test set molecules obtained by PLS analysis using CoMFA (a), CoMSIA (b) & Topomer CoMFA (c) models.

CoMFA contour Maps

CoMFA steric and electrostatic fields are shown in **Fig 4A** & **4B**, respectively. The green contour characterizes the regions where bulky substituents would increase the MetAP2 inhibitory activity; whereas yellow contour indicates the regions where steric bulk would decrease inhibitory action. The contribution of the green and yellow contours was maintained as the default value of 80 and 20, respectively. The blue contour depicts the favorable sites for electropositive groups while the red contour favors the electronegative groups.

CoMFA contour map analysis provided significant information about the steric and electrostatic favorable and unfavorable regions. CoMFA steric contours **(Fig 4A)** showed a green contour around the carbon adjacent to sulphur and C2 of the phenyl ring (**compound 86**). The most potent compounds of the series **13** & **86** possess no bulky group on second position of the phenyl group. Small steric

bulky groups are favorable at second and fourth position of R2 group. The phenyl group at R2 position is the most potent having no steric bulky groups. In case of compounds containing methyl group at second position of R2 (**14, 24, 36, 56, 67, 87**), the alignment of the steric group is towards the sterically favorable region, but reducing its activity as compared to compounds with no substitution (**86**). In compounds with methoxy group at ortho position of R2 (**15, 47, 68**) the activity has further decreased as compared to methyl group because the methoxy group is completely oriented in the sterically unfavorable region.

Table 3: The optimal results of the CoMFA, CoMSIA and Topomer CoMFA analyses for the MetAP2 Inhibitors

Parameters	MetAP2 Inhibitors				
	CoMFA	CoMSIA	Topomer CoMFA		
q^2	0.703	0.704	0.746		
r^2	0.894	0.889	0.886		
SEE	0.358	0.367	.36		
F	30.1	22.1	47.6		
N. C	12	12	8		
Donor	-	1.7			
Acceptor	-	20.1			
Hydrophobic	-	37.7			
Steric	66.1	13.1		R1	R2
			Steric+ve (green)	0.141	1.547
			Steric-ve (yellow)	-0.282	-0.886
Electrostatic	33.9	27.4	Electro+ve (blue)	0.316	0.757
			Electro-ve (red)	-0.007	-0.426

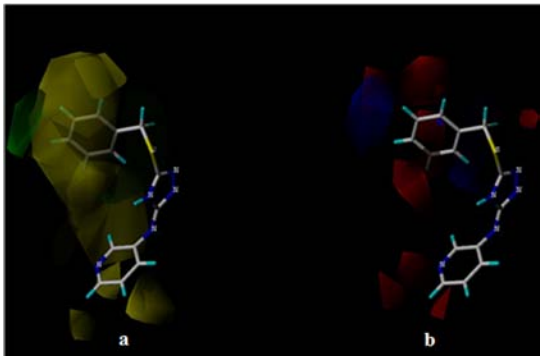


Fig. 4: CoMFA contour maps of compound 86. (a) CoMFA steric contours: Green contours indicate regions where steric interaction is favored. Yellow contours are areas where the steric interaction is disfavored. (b) CoMFA electrostatic contours: the blue region represents the area where an electropositive group is favorable and the red region refers to the area where an electronegative group is favorable.

In case of electrostatic contours (**Fig 4B**), the red contour occupies second position of R2 substitution of phenyl ring and blue at third position. In compound **16**, the fluoro group at second position is favorable because fluorine is an electronegative atom and aligns in the red contour. In case of 3,4 difluoro derivative (**17**) activity has decreased because the third fluoro group aligns in the electropositive region. Compound **21** containing thiazole ring, nitrogen of thiazole ring aligns in the electropositive region, while sulphur aligns in the electronegative region. This leads to decrease in activity compared to compound **22** which has an oxazole ring at R2 position. Oxygen and nitrogen aligns in the red contour enhancing activity over compound **21**. These results in case of electrostatic contours are observed in all sets of compounds. From these results we can conclude that in case of aromatic R2 groups, an electronegative group at ortho position is favorable for enhanced activity. In case of five member ring, electropositive group at third and fourth position, and electronegative group at second position is favorable.

CoMSIA statistical results

For the CoMSIA model, the highest cross-validated q^2 was obtained by using the combination of steric, electrostatic, hydrophobic and H-bond acceptor and H-bond donor fields. The CoMSIA statistical results are q^2-0.704, r^2-0.889, SEE-0.367, with twelve components. The corresponding CoMSIA field contributions are 1.7%, 20.1%, 37.7%, 13.1%, and 27.4%, respectively. CoMSIA analysis results are also summarized in **Table 2** and **3**. **Fig 3B** shows the relationship between the predicted and the experimental $pk_{i,\ app}$ values for the CoMSIA model. From the cross-validation results suggesting that a reliable CoMSIA model was successfully constructed.

CoMSIA Contour Maps

The CoMSIA contour maps derived using steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor fields are represented in **Fig 5**. The CoMSIA steric and electrostatic contours in **Fig 5D** and **5E** are similar to those of CoMFA.

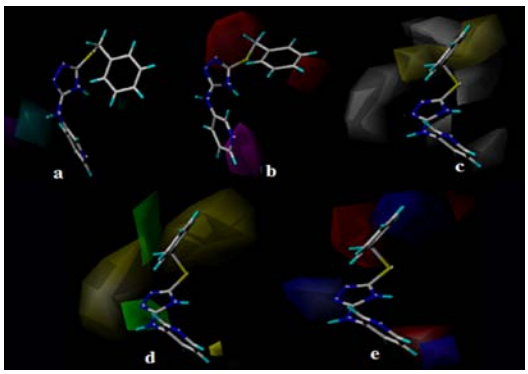


Fig. 5: CoMSIA contour maps based on compound 86: (a) CoMSIA hydrogen bond donor contours - cyan and purple contours represent favorable and unfavorable hydrogen bond donor regions, respectively (b) CoMSIA hydrogen bond acceptor contours - magenta and red contours represent favorable and unfavorable hydrogen bond acceptor regions, respectively (c) CoMSIA hydrophobic contour maps- yellow contours represent regions where hydrophobic groups increase activity, while grey contours highlight regions that would favor hydrophilic groups. (d) CoMSIA steric contour maps - green contour favors steric or bulky group and yellow contour denotes disfavored region. (e) CoMSIA electrostatic contour maps - blue contour indicates electropositive charge and red contour electronegative charge.

The hydrogen bond donor map of CoMSIA model is displayed in **Fig 5A**. In CoMSIA hydrogen bond donor field favored regions are

represented by cyan contours and unfavorable regions are represented by purple contours respectively. A big cyan contour was observed adjacent to the aniline nitrogen and a small contour above the NH of the triazole ring. The triazole NH group is essential for the formation of hydrogen bonding with HIS231of the MetAP2 active site. The anilino-NH group is essential for hydrogen bonding interaction with ASP262 for MetPA2 inhibitory activity (**Fig 1**).

In hydrogen bond acceptor contours (**Fig 5B**), magenta contour favors the regions of hydrogen bond acceptor substituents on ligands and red contour represent areas where such substituents and compounds may be unfavored. In compound **86**, a magenta contour is observed near the nitrogen on the pyridine ring. Hydrogen bond acceptors at this region greatly enhance the MetAP2 inhibitory action. The pyridine ring at R1 substitution of compounds **86 - 94** is showing greater activity than all other compounds in the series for the corresponding R2 substitution. Hydrogen bond acceptors are also favored at third position of R2 phenyl group substitutions. In compound **27**, having fluoro group at third and fourth position of the phenyl ring, the 4-fluoro group aligns in the red contour which results in decrease in activity as compared to compound **26**. These observations conclude that hydrogen bond acceptors at third position of R1 and R2 phenyl substitution greatly enhance the activity.

Hydrophobic contour maps are shown in **Fig 5C**. The yellow and grey regions indicate the areas where hydrophobic and hydrophilic properties are preferred respectively. A big yellow contour overlapped the R2 substituted phenyl ring. Compound **86** with highest inhibitory action was found to have a hydrophobic substitution at this region. In compounds **77-85**, the R1 substitution is a biphenyl ring which is highly hydrophobic in nature. However this region is covered by grey contours which are unfavorable for hydrophobic substituents. The hydrophobic nature of these compounds led to a large decrease in activity.

Topomer CoMFA

Statistical parameters obtained were LOO cross-validated q^2- 0.746 and r^2- 0.886 with eight optimum components. Standard error was found to be 0.0.36 for r^2 and 0.53 for q^2 having y-intercept of 8.75. Other statistical parameters for the topomer generated model are given in **Table 3**, and contours for each field are shown in **Fig 6**. The relationship between the predicted and the experimental $pK_{i,\ app}$ values are shown in scattered plot **Fig 3C**.

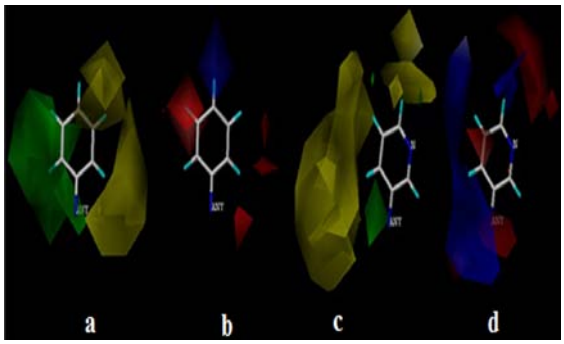


Fig. 6: Topomer CoMFA contour maps of compound 86. (R1 a & R2 c) CoMFA steric contours: Green contours indicate regions where steric interaction is favored. Yellow contours are areas where the steric interaction is disfavored. (R1 b & R2 d) CoMFA electrostatic contours: the blue region represents the area where an electropositive group is favorable and the red region refers to the area where an electronegative group is favorable.

The topomer CoMFA 3D contour maps around R1 and R2 were generated by plotting the coefficients from the model. The maps are shown using compound 86 as a reference structure. In the steric contour map (**Fig 6A**), the green contour at C1, C2 and C3 of R2 substitution indicate that a bulky substituent would be favorable, and the yellow contours denote where bulky substituents would not be tolerated. In the electrostatic contours (**Fig 6B**) of R2 substitution, the blue contours located at the C-4 site indicate that electropositive groups would be favorable, and the red contours indicate that electronegative groups would be favorable which suggested that bulky groups with electronegative potential would be favorable for activity. Regarding the contours of the R1 (**Fig 6C**) group, green contours were located near the C-1 site, and yellow contours were located at C2, C3 and C4 position. The red contours were located at C1 and C3. Blue contours were located at C1, C2 and C4 position. This demonstrated that a moderately bulky group with electronegative potential at R1 substitution would improve the MetAP2 inhibitory activity. The model can be used to search for R groups in large databases to identify potential MetAP2 inhibitors.

External validation results

Further, external validation was carried out for 3D-QSAR models as per Golbraikh and Roy method. CoMFA and CoMSIA models are satisfied the external validation parameters. External validation results for the CoMFA, CoMSIA and Topomer CoMFA models., the valid r^2_m (≥0.5) values of 0.912, 0.95 and0.941 as well as high slope of regression lines through the origin [k (0.85≤ k ≥1.15)] values of 1.0454, 1.044 and 1.04658 respectively. The calculated $[(r^2-r_0^2)/r^2]$ <0.1] values of -0.0086, 0.001985 and 0.00380 were also obtained respectively. These results suggest the generated best CoMFA, CoMSIA and Topomer CoMFA models have good predictive abilities. These models provide the tool to guide the design and synthesis of novel and more potent 1,2,4 triazoles as MetAP2 inhibitors. The external validation results are summarized in **Table 4.**

Table 4: Results of validation

Validation Parameter	CoMFA	CoMSIA	Topomer CoMFA
q^2	0.703	0.704	0.746
r^2	0.894	0.889	0.882
$[(r^2-r_0^2)/r^2]<0.1$	0.0086	0.001985	.00380
$0.85 \leq k \geq 1.15$	1.0454	1.044	1.04658
$r^2_m > 0.5$	0.912	0.95	0.941

Design of new molecules

Ligand-based method such as 3D QSAR is widely used not only because it is not very computationally intensive but also because it can lead to the rapid generation of QSARs from which the biological activity of newly designed compounds can be predicted. In contrast, an accurate prediction of activity of untested compounds based on the computation of binding free energies is both complicated and lengthy. The CoMFA contour maps are a clear indicator for intuitionistic medicinal chemist for predicting novel molecules with enhanced MetAP2 inhibitory activity. The key findings are depicted in **Fig 7**. The fourth triazole NH group and the adjacent NH group form hydrogen bonds in the MetAP2 active site. Electronegative groups are favored at C3 position of R1 and small bulky groups and electronegative groups beneficial at C2and C3 position of R2 respectively.

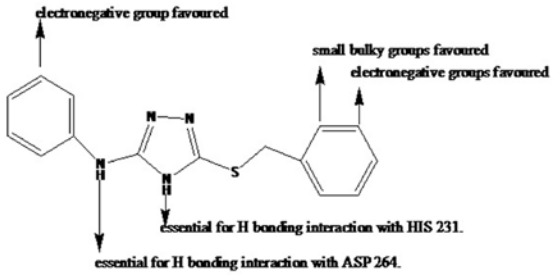


Fig. 7: Key findings of the 3D-QSAR study.

CONCLUSION

Present studies indicate the importance of ligand orientation and selection of training set molecules in the development of statistically significant QSAR models. All models demonstrated good statistical results which emphasizes the importance of steric, electrostatic, hydrophobic, H-bond donor and acceptor fields. Furthermore, the contour maps along with the docking results offered enough information to understand the structure–activity relationship and identified structural features influencing the inhibitory activity. Overall, the correlation of the results obtained from QSAR and the 3D topology of the protein structure suggested the robustness of the QSAR models. The most active compound from the series (**86**) was subjected to docking studies. The results showed specific interaction of the fourth triazole NH with HIS 231 and the anilino NH with Asp 262. Topomer CoMFA 3D-QSAR model with good internal and external prediction capability was established for a training set of 58 MetAP2 inhibitors, and a test set of 19 molecules were employed to validate the external predictive ability of the models. The Topomer search model developed can be used to screen R-groups various molecule databases, to identify new molecules with higher bioactivity. The present work provides references to drug design for MetAP2 inhibitors. The structural requirements identified in the present study can be utilized strategically in the design of novel, potent and selective MetAP2 inhibitors.

ACKNOWLEDGEMENT

Authors are thankful to The Principal, JSS College of Pharmacy, Mysore, India for providing necessary facilities. Authors are beholden to DBT for the financial assistance (Ref. No: BT/PR5594/MED/29/540/20l2).

CONFLICT OF INTEREST

Authors declare no conflict of interest

REFERENCES

1. Bruce NA. Endogenous Mutagens and the Causes of Aging and Cancer. Mutat Res 1991;250:3-16.
2. "How many different types of cancer are there? Cancer Research UK: Cancer Help UK". Retrieved 11 May 2012.
3. Anand P, Kunnumakkara AB, Kunnumakara AB, Sundaram C, Harikumar KB, Tharakan ST, *et al.* Cancer is a preventable disease that requires major lifestyle changes. Pharm Res 2008;25:2097–116.
4. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
5. Risau W. Mechanisms of angiogenesis. Nature 1997;386:671-4.
6. Yang G, Kirkpatrick RB, Ho T, Zhang G, Liang P, Johanson KO, *et al*. Steadystate kinetic characterization of substrates and metal-ion specificities of the full-length and N-terminally truncated recombinant human methionine aminopeptidases (type 2). Bio chem 2001;40:10645-54.
7. Arfin SM, Kendall RL, Hall L, Weaver LH, Stewart AE, Matthews BW, *et al*. Eukaryotic methionylaminopeptidases: two classes of cobalt-dependent enzymes. Proc Natl Acad Sci USA 1995;92:7714–8.
8. Li X, Chang YH. Evidence that the human homologue of a rat initiation factor-2 associated protein (p67) is a methionine aminopeptidase. Bio Chem Biophys Res Commun 1996;227:152–9.
9. Bennett B Holz. EPR Studies on the Mono-and Dicobalt (II)-Substituted Forms of the Aminopeptidase from Aeromonasproteolytica. Insight into the Catalytic Mechanism of Dinuclear Hydrolases. J Am Chem Soc 1997;119:1923–33.
10. Johansson FB, Bond AD, Nielsen UG, Moubaraki B, Murray KS, Berry KJ, *et al*. Dicobalt II-II, II-III, and III-III complexes as spectroscopic models for dicobalt enzyme active sites. Inorg Chem 2008;47:5079–92.
11. Larrabee JA, Leung CH, Moore RL, Thamrong-nawasawat T, Wessler BS. Magnetic circular dichroism and cobalt (II) binding equilibrium studies of Escherichia coli methionyl aminopeptidase. J Am Chem Soc 2004;126:12316–24.
12. Taunton J. "How to starve a tumor". Chem Biol 1997;4:493–6.
13. Sin N, Meng L, Wang MQ, Wen JJ, Bornmann WG, Crews CM. The anti-angiogenic agent fumagillin covalently binds and inhibits the methionine aminopeptidase, MetAP-2. Proc Natl Acad Sci USA 1997;94:6099–103.
14. Griffith EC, Su Z, Turk BE, Chen S, Chang YH, Wu Z, *et al*. Methionine aminopeptidase (type 2) is the common target for angiogenesis inhibitors AGM-1470 and ovalicin. Chem Biol 1997;4:461–71.
15. Lowther WT, McMillen DA, Orville AM, Matthews BW. The anti-angiogenic agent fumagillin covalently modifies a conserved active-site histidine in the Escherichia coli methionine aminopeptidase. Proc Natl Acad Sci USA 1998;95:12153–57.
16. Folkman J. Angiogenesis in cancer, vascular, rheumatoid and other disease. Nat Med 1995;1:27–31.
17. Yashwant S. Recent advancements of triazoles as anticancer agents. IJPCR 2010;2:95-7.
18. Verner E, Katz BA, Spencer JR, Allen D, Hataye J, Hruzewicz W, *et al*. Development of serine protease inhibitors displaying a multicentered short (<2.3 A) hydrogen bond binding mode: Inhibitors of Urokinase-type plasminogen activator and factor Xa. J Med Chem 2001;44:2753.
19. Mackman RL, Katz BA, Breitenbucher JG, Hui HC, Loung C, Liu L, *et al*. Exploiting subsite S1 of trypsin-like serine proteases for selectivity: potent and selective inhibitors of urokinase-type plasminogen activator. J Med Chem 2001;44:3856.
20. Barber CG, Dickinson RP, Horne VA. Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 1:2 pyridinylguanidines. Bioorg Med Chem Lett 2002;12:181.
21. Barber CG, Dickinson RP. Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 2:(3-substituted-5-halo-2-pyridinyl) guanidines. Bioorg Med Chem Lett 2002;12:185.
22. Cramer RD, Jilek RJ, Guessregen S, Clark SJ, Wendt B, Clark RD. Lead hopping. Validation of topomer similarity as a superior predictor of similar biological activities. J Med Chem 2004;47:6777–91.
23. Bhongade BA, Gadad AK. 3D-QSAR CoMFA/CoMSIA studies on Urokinase plasminogen activator (uPA) inhibitors: a strategic design in novel anticancer agents. Bioorg Med Chem 2004;12:2797–2805.
24. Joseph P M Jr, Paul WF, Glenn AH, Robert BK, Cheryl AJ, Randall KJ, *et al*. Highly Potent Inhibitors of Methionine Aminopeptidase-2 Based on a 1, 2, 4-Triazole Pharmacophore. J Med Chem 2007;50:3777-85.
25. Carvalho LL, Maltarollo VG, Lima EF, Weber KC, Honorio KM, Silva ABF. Molecular Features Related to HIV Integrase Inhibition Obtained from Structure-and Ligand-Based Approaches. PLOS one 2014;9:1-9.
26. Bostrom J, Bohm M, Gundertofte K, Klebe G. J Chem Inf Comput Sci 2003;43:1020–1027.
27. Zaheer UH, Khan W, Zia SR, Iqbal S. Structure-based 3D-QSAR models and dynamics analysis of novel N-benzyl pyridinone as p38α MAP kinase inhibitors for anticytokine activity. J Mol Graph Model 2012;36:48-61.
28. Cramer RD, Bunce JD, Patterson DE, Frank IE. Crossvalidation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. Quant Str Act Rel 1988;7:18-25.
29. Wold S, Ruhe A, Wold H, Dunn WJ. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM J Sci Stat Comput 1984;5:735–43.
30. Roy PP, Roy K. On Some Aspects of Variable Selection for Partial Least Squares Regression Models. Quant Str Act 2008;27:302-13.
31. Golbraikh A, Tropsha A. Beware of q^2!. J Mol Graph Model 2002;20:269-76.
32. Mohan KK, Bharathkumar I, Pujar GV, Purohit MN, Vijaykumar G S. Design, synthesis and 3D-QSAR studies of new diphenylamine containing 1,2,4-triazoles as potential antitubercular agents. E J Med Chem 2014;82:516-29.
33. Roy PP, Roy K. On two novel parameters for validation of predictive QSAR models. Mole 2009;14:1660-701.
34. Roy PP, Roy K. Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. Eur J Med Chem 2009;44:2913-22.