IJMPR 2022, 6(4), 89-101

International Journal of Modern Pharmaceutical Research
www.ijmpronline.com

ISSN: 2319-5878
IJMPR
Research Article

SJIF Impact Factor: 5.273

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY ON A-SPIROPIPERIDINE HYDROXAMATES AS MATRIX METALLOPROTEINASE-3 INHIBITORS

Jahan Afsar, Sharma Brij Kishore* and Meena Dinesh Kumar

Department of Chemistry, Government College, Bundi-323 001 (Rajasthan), India.

Received on: 15/03/2022
Revised on: 05/04/2022
Accepted on: 25/04/2022

***Corresponding Author**
Sharma Brij Kishore
Department of Chemistry, Government College, Bundi-323 001 (Rajasthan), India.

ABSTRACT

QSAR study has been carried out on the MMP-3 inhibitory activity of α-spiropiperidine hydroxamates in 0D- to 2D-Dragon descriptors. The derived QSAR models have revealed that the Balaban mean square distance index (descriptor MSD), path/walk 5 -Randic shape index (descriptor PW5), information content indices of 2nd order neighborhood symmetry (descriptors IC2 and SIC2) in addition to 5th order Galvez topological charge indices (descriptors GGI5 and JGI5) played a pivotal role in rationalization of MMP-3 inhibition activity of titled compounds. Atomic properties such as mass, volume and electronegativity in terms of atomic properties weighted descriptors MATS6m, MATS3e, GATS1p, GATS3p, BELm2, BELv7 and Sv are also predominant to explain MMP-3 inhibition actions of α-spiropiperidine hydroxamates. PLS analysis has also corroborated the dominance of CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

KEYWORDS: QSAR; MMP-3 inhibitory activity; Combinatorial protocol in multiple linear regression (CP-MLR) analysis; PLS analysis; Dragon descriptors; α-spiropiperidine hydroxamates.

1. INTRODUCTION

A risky complication of excess bleeding may result in the need for blood transfusion which is associated with substantial morbidity and mortality. Nearly 50% of deaths occurring within the first 24 hours of traumatic injury and up to 80% of intraoperative trauma mortality are due to hemorrhage.[1] The excessive blood loss after injury is prevented by coagulation and the blood clots after wound repair are removed by fibrinolysis, physiologically. Thrombosis or hemorrhage is outcome of the pthological alterations of this delicate balance. The key protein in fibrinolysis is plasmin and it results from plasminogen activation. The dysregulation of this system is hyperfibrinolysis, which plays an important role in major bleeding events.[2] To reduce blood loss and transfusion and in the treatment of heavy menstrual bleeding antifibrinolytic agents are widely used in major surgeries and trauma episodes.[3] The effects of antifibrinolytics are elicited by competitively reducing the binding of plasminogen to fibrin and inhibiting the activation of plasminogen to plasmin, an enzyme that degrades fibrin clots, fibrinogen and other plasma proteins.[4] Adverse side effects and potential risk of thrombotic complications has been associated with these compounds.[5-7] Due to the safety issues aprotinin, a polypeptidic, bovine-derived protease inhibitor, was withdrawal from the market although, it was an effective agent used to reduce bleeding during complex surgery.[8,9] The European Medicines Agency (EMA) has recently recommended that the suspension be lifted for a restricted range of indications.[10] Therefore new therapeutic strategies are needed to prevent major bleeding events.

The fibrinolytic and matrix metalloproteinase (MMP) systems cooperate in thrombus dissolution.[11] The action of MMPs either through direct fibrinogen targeting or by enhancing tissue plasminogen activator (tPA)-induced fibrinolysis may provide a new pharmacological target for fibrinolysis control.[12] The fibrinolytic activity and bleeding time markedly reduced in MMP-10 null mice and administration of active recombinant human MMP-10 reversed it.[12] The another fibrinolytic MMP, MMP-3, which is 82% homologous to MMP-10,12 has also been associated with intracranial hemorrhage.[13] In this way, the inhibition of MMP-10, MMP-3 and/or other MMPs may present a new opportunity for controlling bleeding. However, in human cancer trials the hydroxamate-based MMP inhibitors were of limited clinical success due mainly to lack of efficacy and to reports of side-effects such as musculoskeletal pain and inflammation and which occurred after 2-3 months of treatment.[14,15] and subsided within 1-3 weeks of

discontinuation.[15] but there is still hope for MMP inhibition as a therapeutic approach.[16]

The doses that required for acute use in antihemorrhagic treatment with MMP inhibitors are much lower than those for reducing metastasis.[17,18] and this indication only requires short-term MMP inhibition leading to a new therapeutic opportunity for MMP inhibitors, as antihemorrhagic agents. This hypothesis was recently demonstrated with a synthetic broad-spectrum MMP inhibitor GM6001 (Ilomastat) and a novel preclinical antihemorrhagic candidate, CM-352.[17] This novel preclinical candidate CM-352 showed remarkable efficacy and safety for the acute treatment of hemorrhage. MMP inhibitors (MMPi) have been explored as anticancer, anti-inflammatory and antiviral agents. To target the zinc catalytic center of these enzymes different structures have been designed with the majority of them containing a hydroxamic acid group acting as a chelating zinc binding group (ZBG).[19] α-sulfonyl-4-tetrahydropyranyl (THP) and α-sulfonyl-α-piperidinyl hydroxamates including SC-78080/SD-2590 and the α-THP 7 (SC-77774) have been explored as potent MMP-2, -9 and -13 inhibitors.[20,21] These compounds spare MMP-1. The α-sulfone α-THP series possesses better ADME and PK properties than the closely related β-sulfone α-THP core of the RS-130830.[22] The MMP inhibitor belonging to α-sulfone α-piperidinyl series showed clinically relevance in advanced rabbit osteoarthritis model.[23] A novel chemical series of compounds belonging to α-spiropiperidine hydroxamates has been synthesized by Orbe et al.[24] The aim of present communication is to establish the quantitative relationships between the reported activities and molecular descriptors unfolding the substitutional changes in titled compounds.

2. MATERIALS AND METHODS

2.1 Biological actions and theoretical molecular descriptors

The reported twenty seven α-spiropiperidine hydroxamates are considered as the data set for present study.[24] These derivatives were evaluated for their MMP-3 and MMP-10 inhibitory activities and were reported as IC_{50}. The MMP-3 and MMP-10 activities are very highly correlated to each other (n = 27, r = 0.988, s = 0.195, F = 1055.08). Thus only the MMP-3 activity has been taken into account in the present study. The reported MMP-3 activity on molar basis (as pIC_{50}) along with the structures of these analogues is shown in Table 1. In the dataset, the initial assessment of activity with all descriptors has suggested the compounds **14** and **19** as potential outliers. An outlier to a QSAR can indicate the limits of applicability of QSAR models. These outliers are not part of the data set. The data set was sub-divided into training set to develop models and test set to validate the models externally. The test set compounds which were selected using an in-house written randomization program, are also mentioned in Table 1.

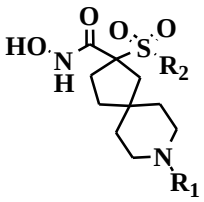
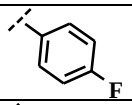
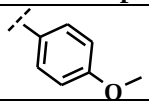
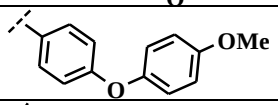
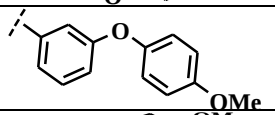
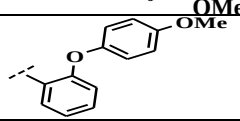
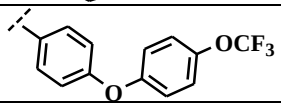
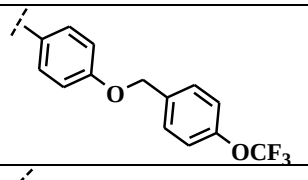
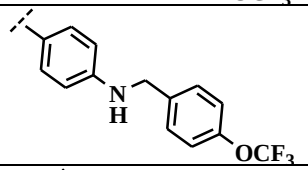
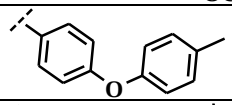
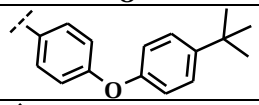
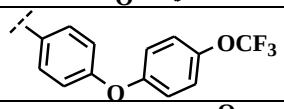
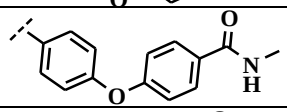
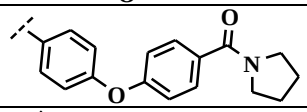
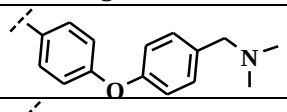
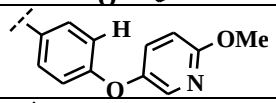
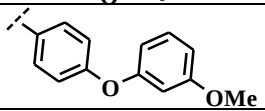
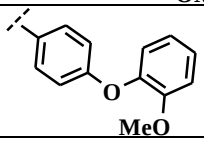
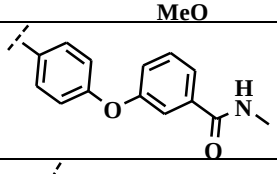
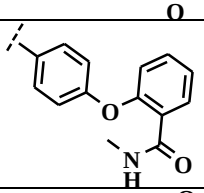
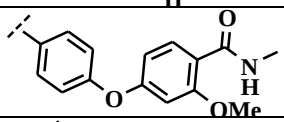
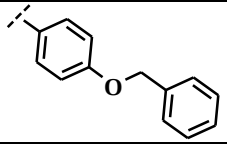
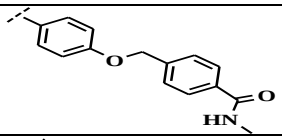
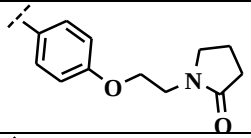
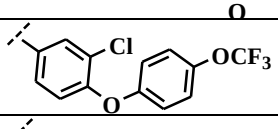
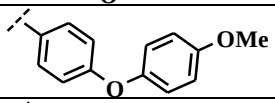
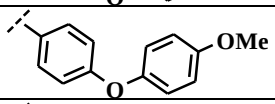
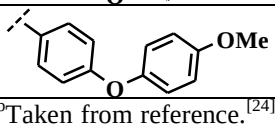


Table 1: Structures, observed and calculated MMP-3 inhibitory activities of α-spiropiperidine hydroxamates.

Cpd.	R_1	R_2		pIC_{50}[a]						
					Calculated					
			Obs.[b]	Eq. (2)	Eq. (3)	Eq. (4)	Eq. (5)	Eq. (6)	Eq. (7)	PLS
1[c]	H		4.72	6.11	6.05	5.94	5.69	5.63	6.00	6.52
2	H		7.02	6.52	6.28	5.94	6.85	6.66	6.56	6.92
3	H		8.05	7.96	8.04	8.36	7.86	8.41	8.24	8.29
4	H		6.75	6.96	6.95	7.01	6.97	6.56	6.49	6.79
5	H		4.70	4.88	4.90	4.97	4.82	5.10	5.08	4.60
6	H		8.40	7.91	7.85	8.04	8.38	8.23	8.28	8.43

7	H		7.70	7.77	7.90	7.59	7.90	8.07	7.94	7.66
8[c]	H		6.94	7.31	7.47	7.41	7.57	8.05	7.94	7.39
9	H		8.22	7.65	7.75	8.37	7.57	7.59	7.71	7.81
10	H		7.68	8.02	7.61	7.01	7.33	7.47	7.62	7.56
11	H		8.40	7.91	7.85	8.04	8.38	8.23	8.28	8.43
12	H		7.68	7.55	7.56	7.96	7.98	7.84	7.65	7.62
13[c]	H		7.72	7.26	7.23	8.10	7.78	7.55	7.58	7.39
14[d]	H		-[d]	-[d]	-[d]	-[d]	-[d]	-[d]	-[d]	-[d]
15	H		7.80	8.09	8.08	7.39	7.42	6.64	6.60	7.56
16[c]	H		8.15	7.39	7.51	7.83	7.17	7.46	7.27	7.34
17	H		6.04	6.02	6.03	6.18	7.27	6.88	6.67	6.52
18	H		7.19	7.36	7.27	7.48	7.77	6.67	6.55	7.00
19[d]	H		-[d]	-[d]	-[d]	-[d]	-[d]	-[d]	-[d]	-[d]
20	H		6.50	6.62	7.02	6.68	7.18	6.33	6.33	6.18
21	H		7.34	7.01	7.14	7.51	7.14	7.88	8.04	7.48

22	H		7.41	7.28	7.47	7.08	6.97	7.48	7.28	6.80
23	H		6.32	7.14	6.81	7.42	6.84	6.90	6.64	6.99
24	H		5.02	5.29	5.44	5.56	4.85	5.61	6.14	5.57
25	CH_3		8.00	8.85	8.96	8.13	7.61	8.38	8.47	8.57
26	$(CH_2)_2CF_3$		8.22	7.67	7.53	7.69	7.36	7.51	7.89	7.65
27[c]	$C(=O)CH_3$		8.30	7.79	7.58	7.74	7.64	7.05	7.40	7.62

[a]IC_{50} on molar basis; [b]Taken from reference.[24] [c]Compound included in test set; [d] "Outlier" compound not included in data set.

The structures of the all the compounds (listed in Table 1) were drawn in 2D ChemDraw.[25] and subjected to energy minimization in the MOPAC using the AM1 procedure for closed shell system after converting these into 3D modules. The energy minimization was carried out to attain a well defined conformer relationship among the congeners under study. The 0D- to 2D-molecular descriptors of titled compounds was computed using DRAGON software.[26] This software offers a large number of descriptors corresponding to ten different classes of 0D- to 2D-descriptor modules. The different descriptor classes include the constitutional, topological, molecular walk counts, BCUT descriptors, Galvez topological charge indices, 2D-autocorrelations, functional groups, atom-centered fragments, empirical descriptors and the properties describing descriptors. These descriptors offer characteristic structural information specific to the descriptor class. The definition and scope of these descriptor's classes is given in Table 2.

Table 2: Descriptor classes used for the modeling of MMP-3 inhibitory activity of α-spiropiperidine hydroxamates.

S. No.	Descriptor Class (Acronyms)[a]	Definition and Scope
1	Constitutional (**CONST**)	Dimensionless or 0D descriptors; independent from molecular connectivity and conformations
2	Topological (**TOPO**)	2D-descriptor from molecular graphs and independent conformations
3	Molecular walk counts (**MWC**)	2D-descriptors representing self-returning walk counts of different lengths
4	Modified Burden eigenvalues (**BCUT**)	2D-descriptors representing positive and negative eigenvalues of the adjacency matrix, weights of the diagonal elements and atoms
5	Galvez topological charge indices (**GALVEZ**)	2D-descriptors representing the first 10 eigenvalues of corrected adjacency matrix
6	2D-autocorrelatons (**2D-AUTO**)	Molecular descriptors calculated from the molecular graphs by summing the products of atom weights of the terminal atoms of all the paths of the considered path length (the lag)
7	Functional groups (**FUN**)	Molecular descriptors based on the counting of the chemical functional groups
8	Atom centered fragments (**ACF**)	Molecular descriptors based on the counting of 120 atom centered fragments, as defined by Ghose-Crippen
9	Empirical (**EMP**)	1D-descriptors represent the counts of nonsingle bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule
10	Properties (**PROP**)	1D-descriptors representing molecular properties of a molecule

[a]Reference.[26]

A total number of 467 descriptors, belonging to 0D- to 2D- modules, have been computed to obtain most appropriate models describing the biological activity. Prior to model development procedure, all those descriptors that are inter-correlated beyond 0.90 and showing a correlation of less than 0.1 with the biological endpoints (descriptor versus activity, $r < 0.1$) were excluded. This procedure has reduced the total descriptors from 467 to 79 as relevant ones to explain the biological actions of titled compounds.

2.2 Development and validation of model

The combinatorial protocol in multiple linear regression (CP-MLR).[27-31] and partial least squares (PLS).[32-34] procedures were used in the present work for developing QSAR models. The CP-MLR is a "filter"-based variable selection procedure, which employs a combinatorial strategy with MLR to result in selected subset regressions for the extraction of diverse structure–activity models, each having unique combination of descriptors from the generated dataset of the compounds under study. The embedded filters make the variable selection process efficient and lead to unique solution. Fear of "chance correlations" exists where large descriptor pools are used in multilinear QSAR/QSPR studies.[35,36] In view of this, to find out any chance correlations associated with the models recognized in CP-MLR, each cross-validated model has been subjected to randomization test.[37,38] by repeated randomization (100 simulation runs) of the biological responses. The datasets with randomized response vector have been reassessed by multiple regression analysis. The resulting regression equations, if any, with correlation coefficients better than or equal to the one corresponding to unscrambled response data were counted. This has been used as a measure to express the percent chance correlation of the model under scrutiny.

Validation of the derived model is necessary to test its prediction and generalization within the study domain. For each model, derived by involving n data points, a number of statistical parameters such as r (the multiple correlation coefficient), s (the standard deviation), F (the F ratio between the variances of calculated and observed activities), and Q^2_{LOO} (the cross-validated index from leave-one-out procedure) have been obtained to access its overall statistical significance. In case of internal validation, Q^2_{LOO} is used as a criterion of both robustness and predictive ability of the model. A value greater than 0.5 of Q^2 index suggests a statistically significant model. The predictive power of derived model is based on test set compounds. The model obtained from training set has a reliable predictive power if the value of the r^2_{Test} (the squared correlation coefficient between the observed and predicted values of compounds from test set) is greater than 0.5. Additional statistical parameters such as, the Akaike's information criterion, AIC.[39,40] the Kubinyi function, FIT.[41,42] and the Friedman's lack of fit, LOF.[43] have also been calculated to further validate the derived models. The AIC takes into account the statistical goodness of fit and the number of parameters that have to be estimated to achieve that degree of fit. The FIT, closely related to the F-value, proved to be a useful parameter for assessing the quality of the models. A model which is derived in k independent descriptors, its F-value will be more sensitive if k is small while it becomes less sensitive if k is large. The FIT, on the other hand, will be less sensitive if k is small whereas it becomes more sensitive if k is large. The model that produces the lowest AIC value and highest FIT value is considered potentially the most useful and the best. The LOF factor takes into account the number of terms used in the equation and is not biased, as are other indicators, toward large number of parameters.

2.3 Applicability domain

The usefulness of a model is based on its accurate prediction ability for new congeners. A model is valid only within its training domain and new compounds must be assessed as belonging to the domain before the model is applied. The applicability domain (AD) is evaluated by the leverage values for each compound.[44] A Williams plot (the plot of standardized residuals versus leverage values (h)) is constructed, which can be used for a simple graphical detection of both the response outliers (Y outliers) and structurally influential chemicals (X outliers) in the model. In this plot, the AD is established inside a squared area within $\pm x$ standard deviations and a leverage threshold h^*, which is generally fixed at $3(k + 1)/n$ (n is the number of training set compounds and k is the number of model parameters), whereas $x = 2$ or 3. If the compounds have a high leverage value ($h > h^*$), then the prediction is not trustworthy. On the other hand, when the leverage value of a compound is lower than the threshold value, the probability of accordance between predicted and observed values is as high as that for the training set compounds.

3. RESULTS AND DISCUSSION

3.1 QSAR results

In multi-descriptor class environment, exploring for best model equation(s) along the descriptor class provides an opportunity to unravel the phenomenon under investigation. In other words, the concepts embedded in the descriptor classes relate the biological actions revealed by the compounds. For the purpose of modeling study, 5 compounds have been included in the test set for the validation of the models derived from 20 training set compounds. A total number of 79 significant descriptors from 0D- to 2D- classes have been subjected to CP-MLR analysis with default "filters" set in it. Statistical models in three and four descriptors have been derived to achieve the best relationship correlating MMP-3 inhibitory activity. Only one model in three descriptors and six models in four descriptors, having $r^2_{Test} > 0.5$, were obtained through CP-MLR. The four parameter models have shared 13 descriptors among them. All these 13 descriptors along with their brief meaning, average regression coefficients, and total incidence are

listed in Table 3, which will serve as a measure of their estimate across these models.

Table 3: Identified descriptors[a] along with their class, physical meaning, average regression coefficient and incidence[b].

Descriptor class, average regression coefficient and (incidence)	
Constitutional descriptors (**CONST**):	Sv (sum of atomic van der Waals volumes scaled on Carbon atom), -4.505(1);
Topological descriptors (**TOPO**):	MSD (Balaban mean square distance index), 3.005(4); PW5 (path/walk 5 - Randic shape index), -1.390(1); IC2 (information content index, neighborhood symmetry of 2-order), -2.375(1); SIC2 (structural information content index, neighborhood symmetry of 2-order), -2.201(1)
Modified Burden eigenvalues (**BCUT**)	BELm2 (lowest eigenvalue n.2 of Burden matrix/weighted by atomic masses), 2.177(4); BELv7 (lowest eigenvalue n.7 of Burden matrix/weighted by atomic van der Waals volumes), -1.815(1)
Galvez topological charge indices (**GALVEZ**):	GGI5 (topological charge index of order 5), -2.558(2); JGI5 (mean topological charge index of order 5), -2.793(2)
2D autocorrelations (**2D-AUTO**):	MATS6m (Moran autocorrelation of lag-6/ weighted by atomic masses), 2.923(1); MATS3e (Moran autocorrelation of lag-3/ weighted by atomic Sanderson electronegativities), -2.400(3); GATS1p (Geary autocorrelation of lag-1/ weighted by atomic polarizabilities), 2.455(1)

[a]The descriptors are identified from the four parameter models for activity emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.813 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models.

The model in three descriptors is mentioned below

pIC_{50} = 6.297 + 3.660(0.677)MSD – 3.374(0.752)IC2 + 1.662 (0.662)C-025

n = 20, r = 0.845, s = 0.615, F = 13.409, Q^2_{LOO} = 0.515, Q^2_{L5O} = 0.529

r^2_{Test} = 0.517, FIT = 1.387, LOF = 0.618, AIC = 0.568 (1)

where n, r, s and F represent respectively the number of data points, the multiple correlation coefficient, the standard deviation and the F-ratio between the variances of calculated and observed activities. In above regression equations, the values given in the parentheses are the standard errors of the regression coefficients. The signs of the regression coefficients suggest the direction of influence of explanatory variables in the models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation.

The descriptors MSD and IC2 participated in above models are the topological descriptors representing Balaban mean square distance index and information content index of 2nd order neighborhood symmetry, respectively. The positive influence of descriptor MSD on the activity and negative influence of descriptor IC2 suggested that higher values of descriptor MSD and lower values of descriptor IC2 would be beneficiary to the activity. The other participated descriptor C-025 is an atom centered fragment class descriptor. The positive contribution shown by it suggested that structure having R--CR--R type fragment would be favorable to the activity.

This model in three descriptors has could account for nearly 71% variance in the observed activities. Considering the number of observations models upto four descriptors have been explored and all the six models having test set r^2 greater than 0.5 are presented below:

pIC_{50} = 9.187 – 1.390(0.438) PW5 – 3.081(0.523)JGI5 – 2.347(0.571)MATS3e + 3.313(0.499)GATS3p

n = 20, r = 0.919, s = 0.469, F = 20.436, Q^2_{LOO} = 0.625, Q^2_{L5O} = 0.716

r^2_{Test} = 0.638, FIT = 2.270, LOF = 0.458, AIC = 0.366 (2)

pIC_{50} = 7.509 + 1.521(0.574) MSD – 2.504(0.643)JGI5 – 2.372(0.609)MATS3e + 2.868(0.526)GATS3p

n = 20, r = 0.907, s = 0.500, F = 17.496, Q^2_{LOO} = 0.549, Q^2_{L5O} = 0.576

r^2_{Test} = 0.626, FIT = 1.944, LOF = 0.522, AIC = 0.417 (3)

pIC_{50} = 7.426 + 2.474(0.551) MSD + 1.546(0.553) BELm2 – 2.678(0.532)GGI5 – 2.480(0.650)MATS3e

n = 20, r = 0.889, s = 0.545, F = 14.138, Q^2_{LOO} = 0.502, Q^2_{L5O} = 0.716

r^2_{Test} = 0.735, FIT = 1.570, LOF = 0.620, AIC = 0.496 (4)

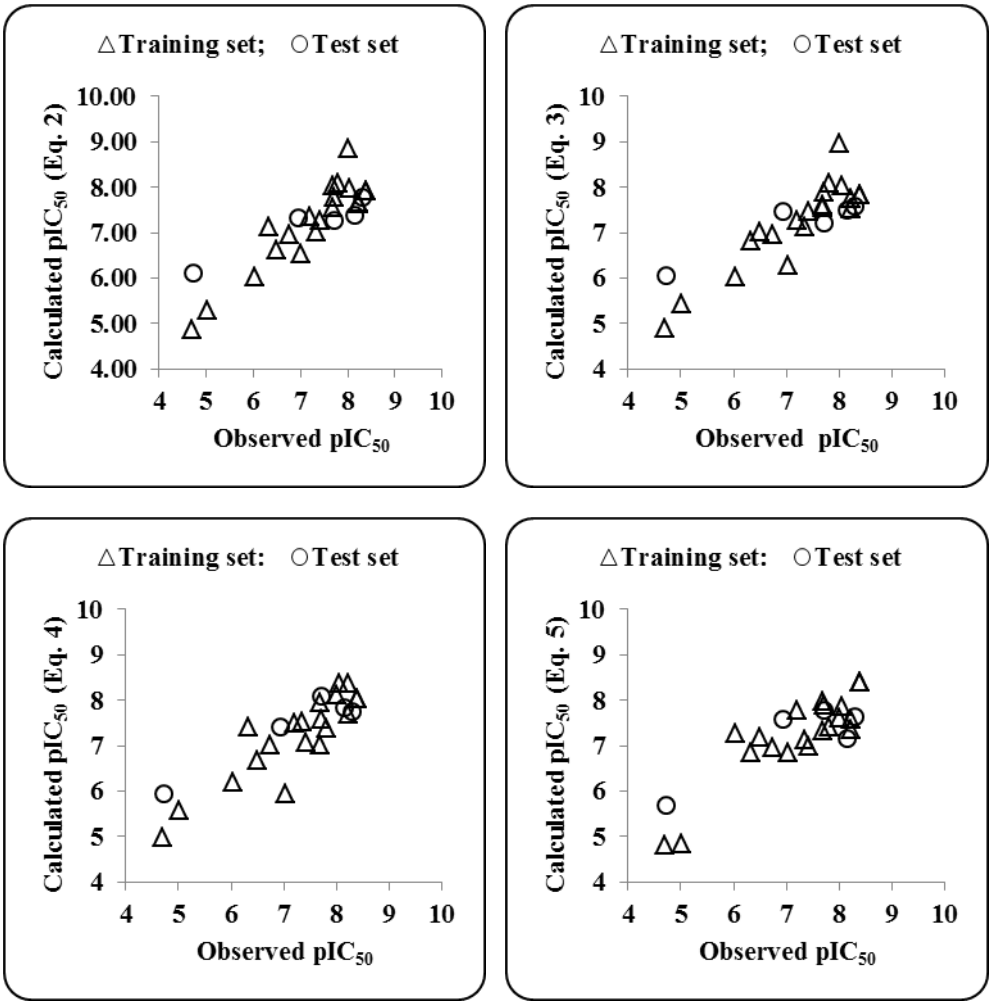
pIC_{50} = 2.851 + 3.638(0.655) BELm2 – 2.439(0.522) GGI5 + 2.922(0.563)MATS6m + 2.455(0.636)GATS1p

n = 20, r = 0.883, s = 0.559, F = 13.280, Q^2_{LOO} = 0.569, Q^2_{L5O} = 0.580
r^2_{Test} = 0.680, FIT = 1.475, LOF = 0.651, AIC = 0.521
(5)

pIC_{50} = 7.330 – 4.505(1.172)Sv + 4.225(0.691)MSD – 2.201(0.524) SIC2 + 2.092(0.679)BELm2
n = 20, r = 0.871, s = 0.583, F = 11.897, Q^2_{LOO} = 0.527, Q^2_{L5O} = 0.623
r^2_{Test} = 0.520, FIT = 1.321, LOF = 0.708, AIC = 0.567
(6)

pIC_{50} = 6.173 + 3.800(0.655)MSD -2.375(0.578)IC2 + 1.432(0.588) BELm2 – 1.815(0.863)BELv7
n = 20, r = 0.867, s = 0.592, F = 11.434, Q^2_{LOO} = 0.502, Q^2_{L5O} = 0.581
r^2_{Test} = 0.506, FIT = 1.270, LOF = 0.730, AIC = 0.584
(7)

These models have accounted for nearly 84% variance in the observed activities. The values greater than 0.5 of Q^2 index is in accordance to a reasonable robust QSAR model. The pIC_{50} values of training set compounds calculated using Eqs. (2) to (7) have been included in Table 1. The models (2) to (7) are validated with an external test set of 5 compounds listed in Table 1. The predictions of the test set compounds based on external validation are found to be satisfactory as reflected in the test set r^2 (r^2_{Test}) values and the same is reported in Table 1. The plot showing goodness of fit between observed and calculated activities for the training and test set compounds is given in Figure 1.



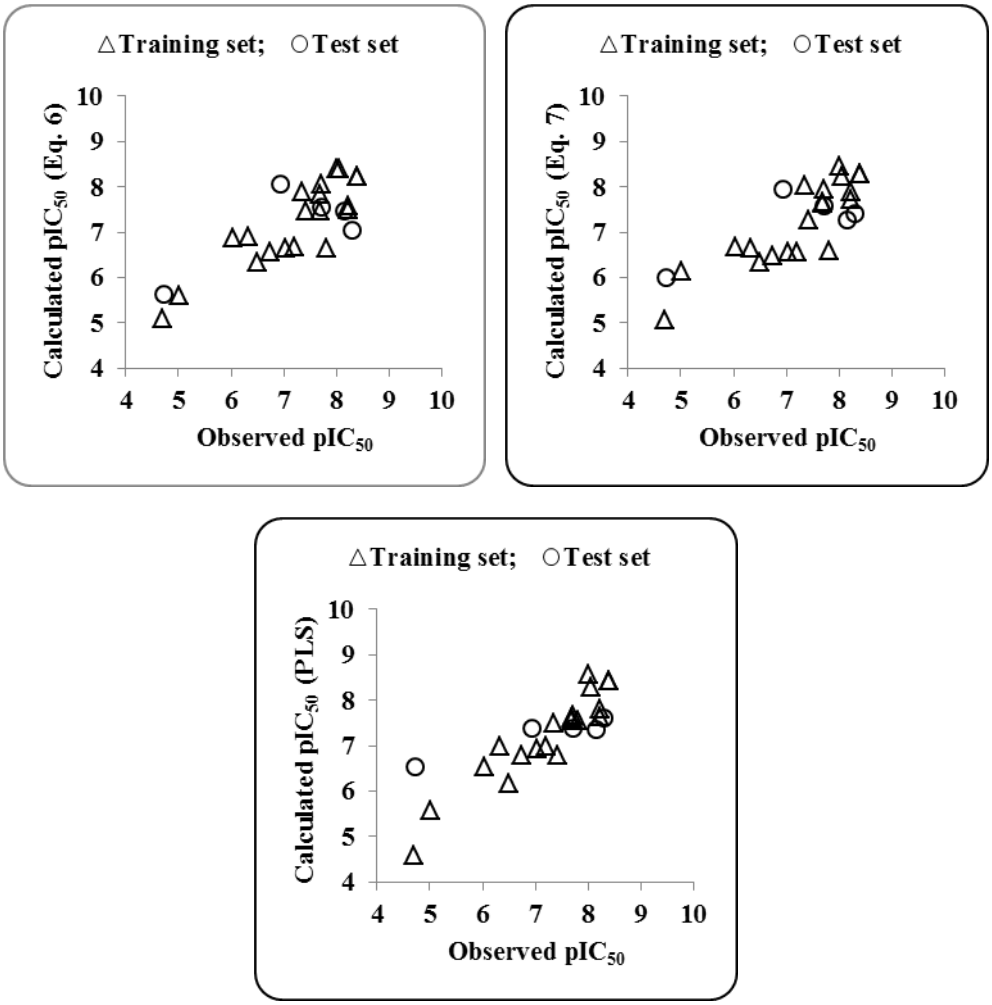


Figure 1: Plot of observed and calculated pIC_{50} values of training- and test-set compounds for MMP-3 inhibition.

It is evident from the signs of the regression coefficients that the newly appeared topological class descriptors PW5 and SIC2, and Galvez charge indices GGI5 and JGI5 contributed negatively to the activity. Thus lower values of path/walk 5 -Randic shape index (descriptor PW5) and structural information content index of neighborhood symmetry of 2-order (descriptor SIC2) and topological charge index of order 5 (GGI5) and mean topological charge index of order 5 (JGI5) would be beneficiary to the activity. The participated descriptors in above models MATS6m, MATS3e, GATS1p and GATS3p are 2D-autocorrelations. Except MATS3e all the descriptors have shown positive contribution to the activity advocating that a lower value of descriptor MATS3e (Moran autocorrelation of lag-3/ weighted by atomic Sanderson electronegativities) and a higher values of descriptors MATS6m (Moran autocorrelation of lag-6/ weighted by atomic masses), GATS1p and GATS3p (Geary autocorrelation of lag-1 and -3, respectively/ weighted by atomic polarizabilities) would be favorable to the activity. Similarly on the same grounds higher value of BCUT descriptor BELm2 (lowest eigenvalue n.2 of Burden matrix/weighted by atomic masses) and a lower values of descriptor BELv7 (lowest eigenvalue n.7 of Burden matrix/weighted by atomic van der Waals volumes) in addition to a lower value of constitutional class descriptor Sv representing sum of atomic van der Waals volumes scaled on Carbon atom would augment the activity of titled compounds.

A partial least square (PLS) analysis has been carried out on these 13 CP-MLR identified descriptors (Table 3) to facilitate the development of a "single window" structure–activity model. For the purpose of PLS, the descriptors have been autoscaled (zero mean and unit SD) to give each one of them equal weight in the analysis. In the PLS cross-validation, two components are found to be the optimum for these 13 descriptors and they explained 88.26% variance in the activity. The MLR-like PLS coefficients of these 13 descriptors are given in Table 4.

Table 4: PLS and MLR-like PLS models from the 13 descriptors of four parameter CP-MLR models for MMP-3 inhibitory activities.

A: PLS equation				
PLS components			PLS coefficient (s.e.)[a]	
Component-1			-0.564(0.055)	
Component-2			0.250(0.051)	
Constant			7.222	
B: MLR-like PLS equation				
S. No.	Descriptor	MLR-like coefficient[b]	(fraction contribution)[c]	Order
1	Sv	-0.004	-0.001	13
2	MSD	0.321	0.125	1
3	PW5	-0.157	-0.061	10
4	IC2	-0.146	-0.057	11
5	SIC2	-0.158	-0.062	9
6	BELm2	0.233	0.091	5
7	BELv7	-0.215	-0.084	6
8	GGI5	-0.247	-0.096	4
9	JGI5	-0.287	-0.112	3
10	MATS6m	0.207	0.081	7
11	MATS3e	-0.191	-0.075	8
12	GATS1p	0.089	0.035	12
13	GATS3p	0.310	0.121	2
Constant			7.833	
C: PLS regression statistics			Values	
n			20	
r			0.939	
s			0.383	
F			63.926	
FIT			5.327	
LOF			0.195	
AIC			0.198	
Q^2_{LOO}			0.833	
Q^2_{L5O}			0.835	
r^2_{Test}			0.503	

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values; [c]f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

For the sake of comparison, the plot showing goodness of fit between observed and calculated activities (through PLS analysis) for the training and test set compounds is also given in Figure 1. Figure 2 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity.

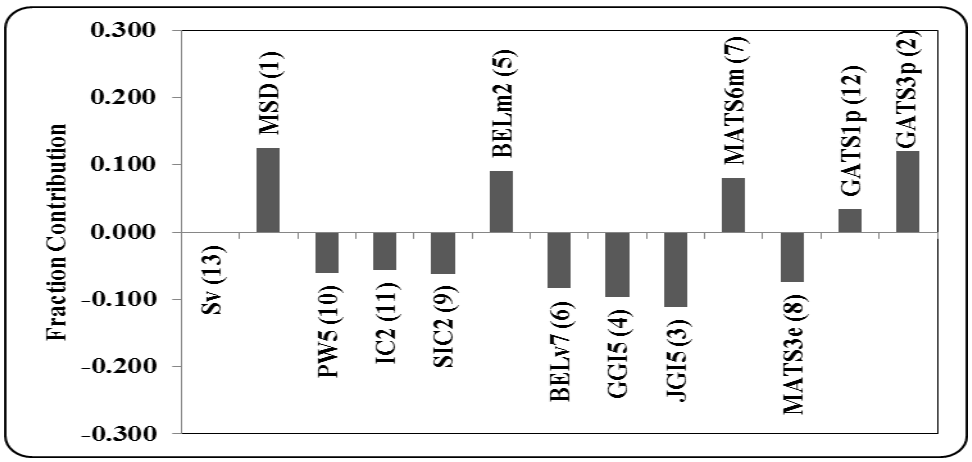


Figure 2: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 13 CP-MLR identified descriptors (Table 3) associated with MMP-3 inhibitory activity of α-spiropiperidine hydroxamates.

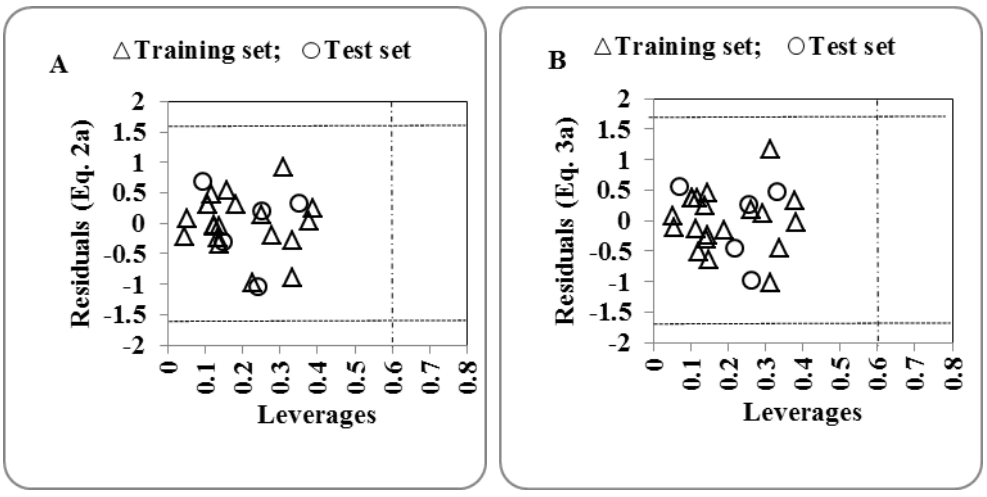
The PLS analysis has suggested MSD as the most determining descriptor for modeling the activity of the compounds (descriptor S. No. 2 in Table 4; Figure 2). The other descriptors in decreasing order of significance are GATS3p, JGI5, GGI5, BELm2, BELv7, MATS6m, MATS3e, SIC2, PW5, IC2, GATS1p and Sv. All these descriptors are part of Eqs. (1) to (7) and convey same inference in the PLS model as well. It is also observed that PLS model from the dataset devoid of CP-MLR identified 13 descriptors (Table 3) is inferior in explaining the activity of the analogues.

3.2 Applicability domain (AD)

On analyzing the model AD in the Williams plot, shown in Figure 3, of the model based on the whole dataset (Table 5), it has appeared that none of the compound was identified as an obvious outlier for the MMP-3 inhibitory activity if the limit of normal values for the *Y* outliers (response outliers) was set as 3 times of standard deviation units. An outlier to a QSAR is identified normally by having a large standard residual activity and can indicate the limits of applicability of QSAR models. None of the compounds listed in Table 1 were found to have leverage (h) values greater than the threshold leverage (h*=0.600). For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

Table 5: Models derived for the whole data set (n = 25) in descriptors identified through CP-MLR.

Model	r	s	F	Q^2_{LOO}	Eq.
pIC_{50} = 9.704 – 1.481(0.452)PW5 – 3.487(0.531)JGI5 – 2.696(0.535)MATS3e + 3.125(0.542)GATS3p	0.900	0.531	21.486	0.668	(2a)
pIC_{50} = 7.975 + 1.605(0.559)MSD – 2.824(0.649)JGI5 – 2.830(0.555)MATS3e + 2.638(0.553)GATS3p	0.891	0.554	19.348	0.625	(3a)
pIC_{50} = 7.388 + 2.488(0.520)MSD + 1.855(0.512)BELm2 – 2.612(0.512)GGI5 – 2.953(0.550)MATS3e	0.891	0.553	19.392	0.637	(4a)
pIC_{50} = 2.306 + 4.043(0.572)BELm2 – 2.387(0.534)GGI5 + 2.938(0.584)MATS6m + 2.894(0.616)GATS1p	0.873	0.595	16.099	0.573	(5a)
pIC_{50} = 7.359 – 3.710(1.167)Sv + 3.954(0.732)MSD – 2.644(0.546)SIC2 + 1.944(0.721)BELm2	0.851	0.641	13.174	0.548	(6a)
pIC_{50} = 5.678 + 3.804(0.688)MSD -2.857(0.601)IC2 + 1.665(0.608)BELm2 – 0.943(0.765)BELv7	0.848	0.646	12.866	0.519	(7a)



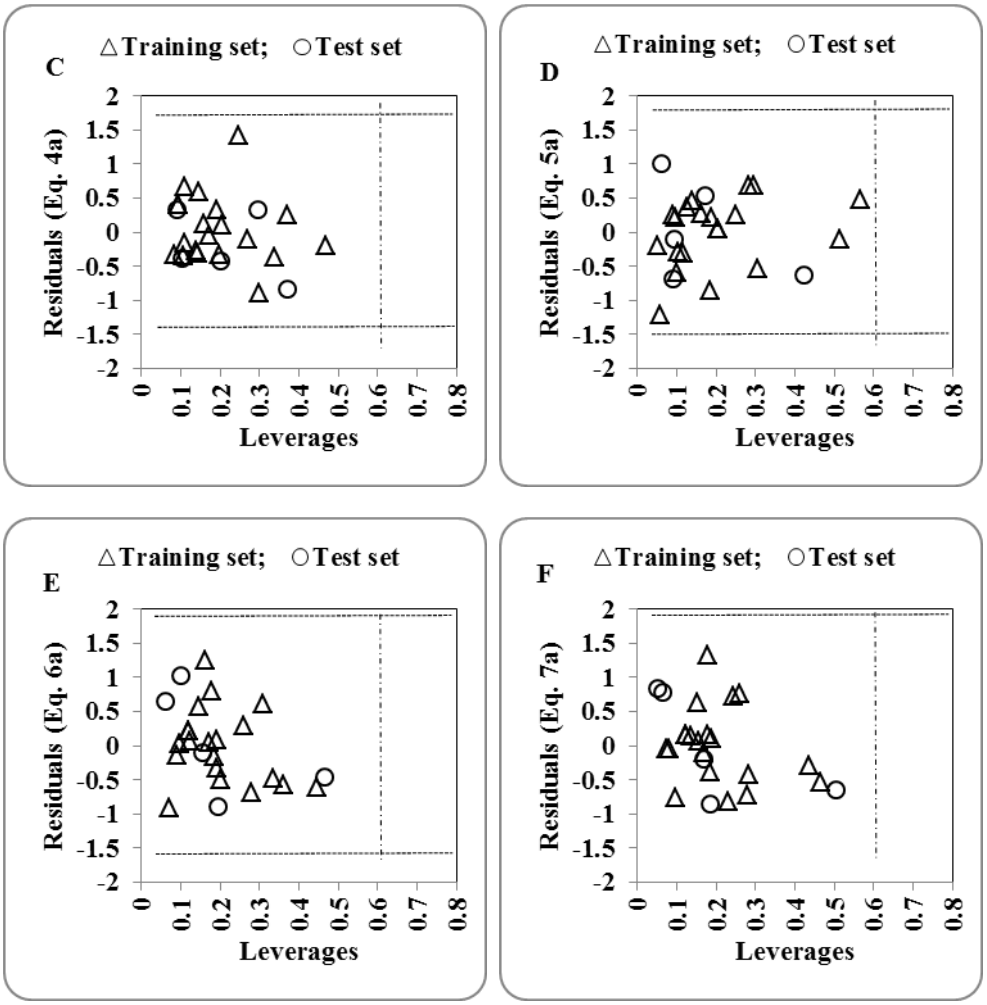


Figure 3: Williams plot for the training-set and test- set compounds for MMP-3 inhibitory activity. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.600).

CONCLUSIONS

QSAR study has been carried out on the MMP-3 inhibitory activity of α-spiropiperidine hydroxamates in 0D- to 2D-Dragon descriptors. The derived QSAR models have revealed that the Balaban mean square distance index (descriptor MSD), path/walk 5 -Randic shape index (descriptor PW5), information content indices of 2nd order neighborhood symmetry (descriptors IC2 and SIC2) in addition to 5th order Galvez topological charge indices (descriptors GGI5 and JGI5) played a pivotal role in rationalization of MMP-3 inhibition activity of titled compounds.

Atomic properties such as mass, volume and electronegativity in terms of atomic properties weighted descriptors MATS6m, MATS3e, GATS1p, GATS3p, BELm2, BELv7 and Sv are also predominant to explain MMP-3 inhibition actions of α-spiropiperidine hydroxamates. PLS analysis has also corroborated the dominance of CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

Compliance with ethical standards

Acknowledgements

Authors are thankful to their institution for providing necessary facilities to complete this study.

Disclosure of conflict of interest

The authors declare no conflict of interest.

REFERENCES

1. Kauvar DS, Lefering R and Wade CE. Impact of hemorrhage on trauma outcome: An overview of epidemiology, clinical presentations, and therapeutic considerations. Journal of Trauma, 2006; 60: S3-11.
2. McCormack PL. Tranexamic acid: a review of its use in the treatment of hyperfibrinolysis. Drugs, 2012; 72(5): 585-617.
3. Mannucci PM and Levi M. Prevention and treatment of major blood loss. New England Journal of Medicine, 2007; 356: 2301-2311.

4. Hochschwender SM and Laursen RA. The lysine binding sites of human plasminogen. Journal of Biological Chemistry, 1981; 256: 11172-11176.
5. McCormack PL. Tranexamic acid: a review of its use in the treatment of hyperfibrinolysis. Drugs, 2012; 2: 585-617.
6. Shakur H, Roberts I, Bautista R, Caballero J, Coats T, Dewan Y, El-Sayed H, Gogichaishvili T, Gupta S, Herrera J, Hunt B, Iribhogbe P, Izurieta M, Khamis H, Komolafe E, Marrero MA, Mejia-Mantilla J, Miranda J, Morales C, Olaomi O, Olldashi F, Perel P, Peto R, Ramana PV, Ravi RR and Yutthakasemsunt S. Effects of Tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. Lancet, 2010; 376: 23-32.
7. Ker K, Edwards P, Perel P, Shakur H and Roberts I. Effect of tranexamic acid on surgical bleeding: systematic review and cumulative meta-analysis. BMJ, 2012; 344: e3054.
8. Fergusson DA, Hebert PC, Mazer CD, Fremes S, MacAdams C, Murkin JM, Teoh K, Duke PC, Arellano R, Blajchman MA, Bussieres JS, Cote D, Karski J, Martineau R, Robblee JA, Rodger M, Wells G, Clinch J and Pretorius R (BART Investigators). A comparison of aprotinin and lysine analogues in high-risk cardiac surgery. New England Journal of Medicine, 2008; 358: 2319-2331.
9. Mangano DT, Tudor IC and Dietzel C. The risk associated with aprotinin in cardiac surgery. New England Journal of Medicine, 2006; 354: 353-365.
10. http://www.ema.europa.eu/ema/pages/news_and_events/news/2012/02/news_detail_001447.jsp.
11. Lijnen HR. Matrix metalloproteinases and cellular fibrinolytic activity. Biochemistry (Mosc), 2002; 67: 92-98.
12. Orbe J, Barrenetxe J, Rodriguez JA, Vivien D, Orset C, Parks WC, Birkland TP, Serrano R, Purroy A, Martinez de Lizarrondo S, Angles-Cano E and Paramo JA. Matrix metalloproteinase-10 effectively reduces infarct size in experimental stroke by enhancing fibrinolysis via a thrombin-activatable fibrinolysis inhibitor-mediated mechanism. Circulation, 2011; 124: 2909-2919.
13. Xue M, Fan Y, Liu S, Zygun DA, Demchuk A and Yong VW. Contributions of multiple proteases to neurotoxicity in a mouse model of intracerebral haemorrhage. Brain, 2009; 132: 26-36.
14. Coussens LM, Fingleton B and Matrisian LM. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. Science, 2002; 295: 2387-2392.
15. Krzeski P, Buckland-Wright C, Balint G, Cline GA, Stoner K, Lyon R, Beary J, Aronstein WS and Spector TD. Development of musculoskeletal toxicity without cflear benefit after administration of PG-116800, a matrix metalloproteinase inhibitor, to patients with knee osteoarthritis: a randomized, 12-month, double-blind, placebo controlled study. Arthritis Research and Therapy, 2007; 9: R109.
16. Vandenbroucke RE and Libert C. Is there new hope for therapeutic matrix metalloproteinase inhibition? Nature Reviews Drug Discovery, 2014; 13: 904-927.
17. Orbe J, Rodriguez JA, Sanchez-Arias JA, Salicio A, Belzunce M, Ugarte A, Chang HCY, Rabal O, Oyarzabal J and Paramo JA. Discovery and safety profiling of a potent pre-clinical candidate, (4-[4-[[(3R)-3-(hydroxycarbamoyl)-8-azaspiro[4.5]decan-3-yl] sulfonyl]phenoxy]-N-methyl-benzamide) (CM-352), for the prevention and treatment of hemorrhage. Journal of Medicinal Chemistry, 2015; 58(7): 2941-2957.
18. Lund LR, Romer J, Bugge TH, Nielsen BL, Frandsen TL, Degen JL, Stephens RW and Dano K. Functional overlap between two classes of matrix-degrading proteases in wound healing. EMBO Journal, 1999; 18: 4645-4656.
19. Yadav RK, Gupta SP, Sharma PK and Patil VM. Recent advances in studies on hydroxamates as matrix metalloproteinase inhibitors. Current Medicinal Chemistry, 2011; 18: 1704-1722.
20. Becker DP, Villamil CI, Barta TE, Bedell LJ, Boehm TL, DeCrescenzo GA, Freskos JN, Getman DP, Hockerman S, Heintz R, Howard SC, Li MH, McDonald JJ, Carron CP, Funckes-Shippy CL, Mehta PP, Munie GE and Swearingen CA. Synthesis and structure-activity relationships of β- and α-piperidine sulfone hydroxamic acid matrix metalloproteinase inhibitors with oral antitumor efficacy. Journal of Medicinal Chemistry, 2005; 48: 6713-6730.
21. Becker DP, Barta TE, Bedell LJ, Boehm TL, Bond BR, Carroll J, Carron CP, DeCrescenzo GA, Easton AM, Freskos JN, Funckes-Shippy CL, Heron M, Hockerman S, Howard CP, Kiefer JR, Li MH, Mathis KJ, McDonald JJ, Mehta PP, Munie GE, Sunyer T, Swearingen CA, Villamil CI, Welsch D, Williams JM, Yu Y and Yao J. Orally active MMP-1 Sparing α-tetrahydropyranyl and α-piperidinyl sulfone matrix metalloproteinase (MMP) inhibitors with efficacy in cancer, arthritis, and cardiovascular disease. Journal of Medicinal Chemistry, 2010; 53: 6653-6680.
22. Lollini L, Haller J, Eugui EM, Womble SW, Martin R and Campbell J. Disease modification by RS-130830, a collagenase-3 selective inhibitor, in experimental osteoarthritis (OA). Arthritis and Rheumatology, 1997; 40(Suppl.): S87-341.
23. Aranapakam V, Davis JM, Grosu GT, Baker J, Ellingboe J, Zask A, Levin JI, Sandanayaka VP, Du M, Skotnicki JS, DiJoseph JF, Sung A, Sharr MA, Killar LM, Walter T, Jin G, Cowling R, Tillett J, Zhao W, McDevitt J and Xu ZB. Synthesis and structure-activity relationship of N-substituted 4-arylsulfonylpiperidine-4- hydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. Journal of Medicinal Chemistry, 2003; 46: 2376-2396.

24. Orbe J, Sanchez-Arias JA, Rabal O, Rodriguez JA, Salicio A, Ugarte A, Belzunce M, Xu M, Wu W, Tan H, Ma H, Paramo JA and Oyarzabal J. Design, synthesis and biological evaluation of novel matrix metalloproteinase inhibitors as potent antihemorrhagic agents: From hit identification to an optimized lead. Journal of Medicinal Chemistry, 2015; 58(5): 2465-2488.
25. Chemdraw ultra 6.0 and Chem3D ultra, Cambridge Soft Corporation, Cambridge, USA. http://www.cambridgesoft.com.
26. Dragon software (version 1.11-2001) by Todeschini R.; Consonni V. Milano, Italy. http://www.talete.mi.it/dragon.htm.
27. Prabhakar YS. A combinatorial approach to the variable selection in multiple linear regression: Analysis of Selwood et al. Data Set-a case study. QSAR and Combinatorial Science, 2003; 22: 583-595.
28. Sharma S, Prabhakar YS, Singh P and Sharma BK. QSAR study about ATP- sensitive potassium channel activation of cromakalim analogues using CP-MLR approach. European Journal of Medicinal Chemistry, 2008; 43: 2354-2360.
29. Sharma S, Sharma BK, Sharma SK, Singh P and Prabhakar YS. Topological descriptors in modeling the agonistic activity of human A3 adenosine receptor ligands: The derivatives of 2-Chloro-N6-substituted-4'-thioadenosine-5'-uronamide. European Journal of Medicinal Chemistry, 2009; 44: 1377-1382.
30. Sharma BK, Singh P, Sarbhai K and Prabhakar YS. A quantitative structure-activity relationship study on serotonin 5-HT6 receptor ligands: Indolyl and piperidinyl sulphonamides. SAR and QSAR in Environmental Research, 2010; 21: 369-388.
31. Sharma BK, Pilania P, Sarbhai K, Singh P and Prabhakar YS. Chemometric descriptors in modeling the carbonic anhydrase inhibition activity of sulfonamide and sulfamate derivatives. Molecular Diversity, 2010; 14: 371-384.
32. Wold S. Cross-validatory estimation of the number of components in factor andprincipal components models. Technometrics, 1978; 20: 397-405.
33. Kettaneh N, Berglund A and Wold S. PCA and PLS with very large data sets. Computational Statistics and Data Analysis, 2005; 48: 69-85.
34. Stahle L and Wold S. Multivariate data analysis and experimental design, in Biomedical Research, Progress in Medicinal Chemistry, Ellis GP and West WB, eds., Elsevier Science Publishers, BV, The Netherlands, 1998; 25: 291-338.
35. Topliss JG and Edwards RP. Chance factors in studies of quantitative structure- activity relationships. Journal of Medicinal Chemistry, 1979; 22: 1238-1244.
36. Katritzky AR, Dobchev DA, Slavov S and Karelson M. Legitimate utilization of large descriptor pools for QSAR/QSPR models. Journal of Chemical Information and Modeling, 2008; 48: 2207-2213.
37. So S-S and Karplus M. Three-dimensional quantitative structure activity relationships from molecular similarity matrices and genetic neural networks. 1. Method and validations. Journal of Medicinal Chemistry, 1997; 40: 4347-4359.
38. Prabhakar YS, Solomon VR, Rawal RK, Gupta MK and Katti SB. CP-MLR/PLS directed structure–activity modeling of the HIV-1RT inhibitory activity of 2,3-diaryl-1,3- thiazolidin-4-ones. QSAR and Combinatorial Science, 2004; 23: 234-244.
39. Akaike H. Information theory and an extension of the minimum likelihood principle. In: Petrov BN and Csaki F. (Eds.). Second international symposium on information theory. Budapest: Akademiai Kiado, 1973; 267-281.
40. Akaike H. A new look at the statistical identification model. IEEE Transactions on Automatic Control, 1974; AC-19: 716-723.
41. Kubinyi H. Variable selection in QSAR studies. I. An evolutionary algorithm. Quantitative Structure-Activity Relationship, 1994; 13: 285-294.
42. Kubinyi H. Variable selection in QSAR studies. II. A highly efficient combination of systematic search and evolution. Quantitative Structure-Activity Relationship, 1994; 13: 393-401.
43. Friedman J. In Technical Report No. 102. Laboratory for Computational Statistics, Stanford University:Stanford, 1990.
44. Gramatica P. Principles of QSAR models validation: internal and external. QSA and Combinatorial Science, 2007; 26: 694-701.