WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 8.084
Volume 9, Issue 15, 978-999. Research Article ISSN 2277– 7105

QSAR RATIONALES FOR THE MMP-13 INHIBITION ACTIVITY OF NON-ZINC BINDING QUINAZOLINE-2-CARBOXAMIDE DERIVATIVES

Jahan Afsar[a], Sharma Brij Kishore[a,*] and Yadav Rama Nand[b]

Department of Chemistry

[a]Government College, Bundi-323 001.

[b]Government R. R. Autonomous College, Alwar-301001, India.

Article Received on
23 Sept. 2020,

Revised on 13 Oct. 2020,
Accepted on 03 Nov. 2020

DOI: 10.20959/wjpr202015-19225

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

ABSTRACT

The MMP-13 inhibition activity of non-zinc binding quinazoline-2-carboxamide derivatives has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the inhibition activity of these congeners. The descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of positive and negative maximal electrotopological variations (MAXDP and MAXDN, respectively), path/walk 2- and 4-Randic shape index (PW2 and PW4, respectively), bond information content of neighborhood symmetry of 4-order (BIC4), Moran autocorrelations of lag-1 and -2/weighted by atomic van der Waals volumes (MATS1v and MATS2v, respectively) and of lag-8/ weighted by atomic Sanderson electronegativities (MATS8e). In addition to these 4th order Galvez topological charge index (GGI4), number of nitrogen atoms (nN), aromatic ethers functionality (nRORPh) and R-C(=X)-X/ R-C#X/X=C=X type structural fragments (C-040) have also shown prevalence to model the inhibition actions. From statistically validated models, it appeared that the descriptors MAXDP, BIC4, PW4, GGI4, nRORPh, MATS8e, MATS1v and C-040 make positive contribution to activity and their higher values are conducive in improving the MMP-13 inhibition activity of a compound. On the other hand, the descriptors nN, MAXDN, PW2 and MATS2v render detrimental effect to activity. Therefore, lower values of descriptors nN, MAXDN, PW2 and MATS2v would be advantageous. Such guidelines may be helpful in exploring more potential analogues of the series. The statistics emerged from

the test sets have validated the identified significant models. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested models have acceptable predictability. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

KEYWORDS: QSAR, MMP-13 inhibitors, Combinatorial protocol in multiple linear regression (CP-MLR) analysis, Chemometric descriptors, Quinazoline-2-carboxamide derivatives.

1. INTRODUCTION

The matrix metalloproteinases (MMPs) are involved in the progression of osteoarthritis (OA). Progressive cartilage degradation causes pain and reduces mobility in affected joints of a patient suffering from osteoarthritis. Nonsteroidal anti-inflammatory drugs (NSAIDs) or selective cylooxygenase-2 (COX-2) inhibitors, intra-articular injections of hyaluronic acid and surgical joint replacement are the current treatments which are limited to symptomatic relief. Coxibs (COX-2 inhibitors) have been withdrawal from the market due to cardiovascular side effects.[1] Thus there is a demand of such disease-modifying osteoarthritis drugs (DMOADs) which may alter disease progression with minimal side effects.

A vital role of MMP-13 (collagenase-3) in the destruction of articular cartilage during the advancement of OA has been suggested because it has shown a substrate specificity favoring degradation of cartilage type II collagen[2,3] which is the main structural component of the cartilage matrix with expressions at higher levels by OA chondrocytes than by normal chondrocytes.[3,4] MMP-13, the most efficient type II collagen-degrading MMP[5,6], has become an attractive therapeutic target because its inhibition reduces cartilage degradation associated with the progression of osteoarthritis in animal models.[7-9] However, broad-spectrum MMP inhibitors exhibit a dose-limiting toxicity leading to side effects such as skin rash and painful joint stiffening (musculoskeletal syndrome, MSS) and inflammation.[10-17] It was suggested that MSS is caused by the inhibition of normal extracellular matrix turnover due to inhibition of other MMPs rather than MMP-13.[18-23] At present, it is unclear which MMP isoforms may be involved[24-26] and to what extent they contribute to MSS. Thus, selective inhibition of MMP-13, devoid of MSS, may prove to be better therapeutic research area.

Most of the MMP inhibitors are representative of zinc binding inhibitors possessing a motif capable to achieve affinity through chelating the catalytic zinc ion[27,28] and substituents recognizing subsites of the target enzyme. Very few studies have been reported on the MMP-13 selective inhibitors devoid of a zinc chelating group.[29-32] The designing of new MMP inhibitors with different selectivity profiles is based on the observation that there is an interaction which involves hydrogen bonding between the inhibitor and the main chain amino acid residues flanking the catalytic site.

In view of this, a new series of potent and selective MMP-13 inhibitors involving unique binding mode at the active site and not interacting with the catalytic zinc, have recently been reported.[33] The general structure of these analogues is shown in Figure 1 and structural variations are given in Table 1.

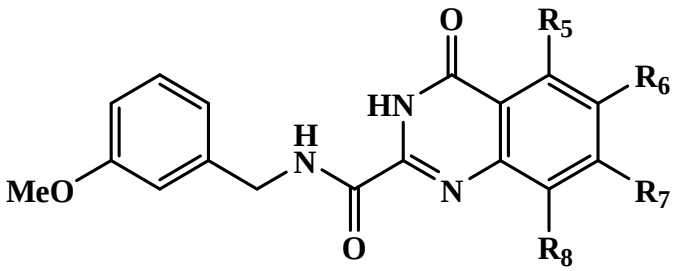
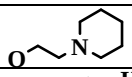
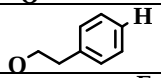
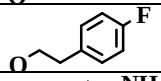
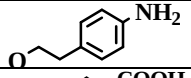
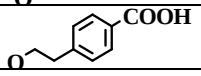
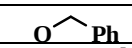
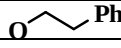
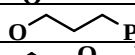
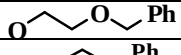
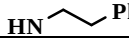
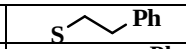
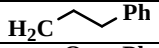
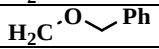
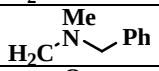
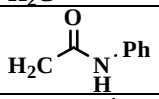
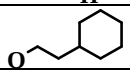


Figure 1: General structure of quinazoline-2-carboxamide derivatives.

Table 1: Structural variations and MMP-13 inhibition activities of quinazoline-2-carboxamides (See Figure 1 for general structure).

Cpd.	R_5	R_6	R_7/ R_8	Obs.[a] pIC_{50}		Cpd.	R_5	R_6	R_7/ R_8	Obs.[a] pIC_{50}
1	H	H	H/H	7.92		19		H	H/H	7.47
2	F	H	H/H	8.30		20		F	H/H	10.40
3	Me	H	H/H	7.54		21		F	H/H	9.82
4[b]	OMe	H	H/H	7.60		22[b]		F	H/H	10.52
5[b]	CN	H	H/H	8.07		23		F	H/H	11.41
6[b]	Ph	H	H/H	8.66		24	H	F	H/H	7.96
7	OPh	H	H/H	9.28		25	H	Me	H/H	7.59
8		H	H/H	9.51		26	H	CF_3	H/H	7.01
9[b]		H	H/H	9.16		27	H	OMe	H/H	8.40
10		H	H/H	9.28		28	H	OEt	H/H	8.62
11		H	H/H	8.80		29	H	OCF_3	H/H	7.57
12[b]		H	H/H	8.37		30[b]	H	OBn	H/H	6.96

13		H	H/H	8.01		31	H	SMe	H/H	8.74
14		H	H/H	8.54		32	H	SO_2Me	H/H	7.64
15		H	H/H	8.77		33	H	Ph	H/H	8.01
16		H	H/H	7.80		34	H	CN	H/H	8.20
17[b]		H	H/H	9.80		35	H	H	OMe/H	5.91
18		H	H/H	9.00		36	H	H	H/OMe	<5.00

[a]On molar basis, IC_{50} represents the concentration of a compound to bring out 50% inhibition of MMP-13, ref.[33]

These quinazoline-2-carboxamide compounds were obtained through optimization with the aid of co-crystal structural information of the complex of lead compound with MMP-13. These MMP-13 inhibitors are involved in interactions with both the distinct deep S1′ pocket and the adjacent side pocket.

In the present communication a 2D-quantitative SAR (2D-QSAR) has been conducted to provide the rationale for drug-design and to explore the possible mechanism of the action. In the congeneric series, where a relative study is being carried out, the 2D-descriptors may play important role in deriving the significant correlations with biological activities of the compounds. The novelty and importance of a 2D-QSAR study is due to its simplicity for the calculations of different descriptors and their interpretation (in physical sense) to explain the inhibition actions of compounds at molecular level.

2. MATERIAL AND METHODS

2.1 Data-set

For present work the quinazoline-2-carboxamide derivatives (Table 1), along with their in vitro inhibition activity of MMP-13, have been taken from the literature.[33] The inhibition activity, IC_{50}, represents the concentration of a compound to achieve 50% inhibition of MMP-13. The same is expressed as pIC_{50} on a molar basis and considered as the dependent variable for the present quantitative analysis. For modeling purpose, the complete data-set was divided into training- and test-sets. The training-set was used to derive statistical significant models while the test-set, consisting nearly 20% of total compounds, was employed to validate such models. The selection of test-set compounds was made through SYSTAT[34] using the single linkage hierarchical cluster procedure involving the Euclidean distances of the inhibition activity, pIC_{50} values. The test-set compounds were selected from

the generated cluster tree in such a way to keep them at a maximum possible distance from each other. In SYSTAT, by default, the normalized Euclidean distances are computed to join the objects of cluster. The normalized distances are root mean-squared distances. The single linkage uses distance between two closest members in clustering. It generates long clusters and provides scope to choose objects at intervals. Due to this reason, a single linkage clustering procedure was applied.

2.2 Molecular descriptors

The structures of the compounds (Table 1), under study, have been drawn in 2D ChemDraw[35] and were converted into 3D objects using the default conversion procedure implemented in the CS Chem3D Ultra. The generated 3D-structures of the compounds were subjected to energy minimization in the MOPAC module, using the AM1 procedure for closed shell systems, implemented in the CS Chem3D Ultra. This will ensure a well defined conformer relationship across the compounds of the study. All these energy minimized structures of respective compounds have been ported to DRAGON software[36] for computing the descriptors corresponding to 0D-, 1D-, and 2D-classes. The combinatorial protocol in multiple linear regression (CP-MLR)[37] analysis and partial least-squares (PLS)[38-40] procedures have been used in the present work for developing QSAR models. A brief description of the computational procedure is given below.

2.3 Model development

The CP-MLR is a 'filter'-based variable selection procedure for model development in QSAR studies. Its procedural aspects and implementation are discussed in some of our publications.[41-46] The thrust of this procedure is in its embedded four 'filters'. They are briefly as follows: filter-1 seeds the variables by way of limiting inter-parameter correlations to predefined level (upper limit ≤ 0.79); filter-2 controls the variables entry to a regression equation through t-values of coefficients (threshold value ≥ 2.0); filter-3 provides comparability of equations with different number of variables in terms of square root of adjusted multiple correlation coefficient of regression equation, r-bar; filter-4 estimates the consistency of the equation in terms of cross-validated r^2 or q^2 with leave-one-out (LOO) cross-validation as default option (threshold value $0.3 \leq q^2 \leq 1.0$). All these filters make the variable selection process efficient and lead to a unique solution. In order to collect the descriptors with higher information content and explanatory power, the threshold of filter-3 was successively incremented with increasing number of descriptors (per equation) by

considering the r-bar value of the preceding optimum model as the new threshold for next generation. Furthermore, in order to discover any chance correlations associated with the models recognized in CP-MLR, each cross-validated model has been put to a randomization test[47,48] by repeated randomization of the activity to ascertain the chance correlations, if any, associated with them. For this, every model has been subjected to 100 simulation runs with scrambled activity. The scrambled activity models with regression statistics better than or equal to that of the original activity model have been counted, to express the percent chance correlation of the model under scrutiny.

To support the findings, a partial least squares (PLS) analysis has been carried out on descriptors identified through CP-MLR. The study facilitates the development of a 'single window' structure-activity model and help to categorize the potentiality of identified descriptors in explaining the MMP-13 inhibition activity profiles of the compounds. It also gives an opportunity to make a comparison of the relative significance among the descriptors. The fraction contributions obtainable from the normalized regression coefficients of the descriptors allow this comparison within the modeled activity.

2.4 Applicability Domain

The utility of a QSAR model is based on its accurate prediction ability for new compounds. 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 is assessed by the leverage values for each compound.[49,50] The Williams plot (the plot of standardized residuals versus leverage values, h) can then be used for an immediate and simple graphical detection of both the response outliers (Y outliers) and structurally influential chemicals (X outliers) in the model. In this plot, the applicability domain is established inside a squared area within ± x (s.d.) and a leverage threshold h^*. The threshold h^* 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. Prediction must be considered unreliable for compounds with a high leverage value ($h > h^*$). 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

For the compounds in Table 1, a total number of 471 descriptors belonging to 0D- to 2D-classes of DRAGON have been computed and were subjected to CP-MLR analysis. The preliminary assessment of complete data-set suggested that the lone compound **30**, having a OBn group at R_6 remained as an 'outlier'. Similarly compound **36**, due to its uncertain activity value, could not fit into the trend of remaining compounds of the series. Both these compounds were, therefore, ignored in the subsequent analyses. The remaining 34 compounds were further divided into training-set and test-set. Seven compounds (nearly 20% of total population) have been selected for test-set through SYSTAT. The identified test-set was then used for external validation of models derived from remaining twenty seven compounds in the training-set. The squared correlation coefficient between the observed and predicted values of compounds from test-set, r^2_{Test}, was calculated to explain the fraction of explained variance in the test-set which is not part of regression/model derivation. It is a measure of goodness of the derived model equation. A high r^2_{Test} value is always good. But considering the stringency of test-set procedures, often r^2_{Test} values in the range of 0.5 to 0.6 are regarded as logical models. Following the strategy to explore only predictive models, CP-MLR resulted into 02 models in two descriptors, 10 models in three and four descriptors each and 05 models in five descriptors. However, the highest significants of them, in statistical sense, are given through Equations (1)-(11):

pIC_{50} = 7.545 + 3.060(0.440)MAXDP – 1.389(0.480)nNR2

n = 27, r = 0.825, s = 0.646, F = 25.709, q^2_{LOO} = 0.619, q^2_{L5O} = 0.595, r^2_{Test} = 0.825 (1)

pIC_{50} = 5.778 + 3.186(0.584)MATS1v + 1.990(0.337)nRORPh

n = 27, r = 0.791, s = 0.699, F = 20.180, q^2_{LOO} = 0.508, q^2_{L5O} = 0.532, r^2_{Test} = 0.786 (2)

pIC_{50} = 7.069 + 3.173(0.413)MAXDP + 1.065(0.487)MATS8e – 1.401(0.447)nNR2

n = 27, r = 0.858, s = 0.600, F = 21.437, q^2_{LOO} = 0.644, q^2_{L5O} = 0.635, r^2_{Test} = 0.803 (3)

pIC_{50} = 6.428 + 1.796(0.553)PW4 + 3.591(0.608)GGI4 – 1.588(0.461)MATS2v

n = 27, r = 0.857, s = 0.601, F = 21.291, q^2_{LOO} = 0.599, q^2_{L5O} = 0.600, r^2_{Test} = 0.663 (4)

pIC_{50} = 5.393 + 1.385(0.516)MAXDP + 2.352(0.590)MATS1v + 1.438(0.451)MATS8e + 1.545(0.370)nRORPh

n = 27, r = 0.892, s = 0.538, F = 21.610, q^2_{LOO} = 0.690, q^2_{L5O} = 0.632, r^2_{Test}= 0.705 (5)

pIC_{50} = 4.260 + 1.165(0.473)BIC4 + 3.163(0.470)MATS1v + 1.430(0.460)MATS8e + 2.367(0.280)nRORPh

n = 27, r = 0.888, s = 0.549, F = 20.547, q^2_{LOO} = 0.647, q^2_{L5O} = 0.675, r^2_{Test}= 0.641 (6)

pIC_{50} = 4.722 + 1.336(0.454)MAXDP + 1.118(0.408)BIC4 + 2.194(0.521)MATS1v + 1.364(0.397)MATS8e + 1.704(0.330)nRORPh

n = 27, r = 0.922, s = 0.473, F =23.892, q^2_{LOO} = 0.704, q^2_{L5O} = 0.703, r^2_{Test} = 0.681 (7)

pIC_{50} = 3.535 + 1.009(0.377)PW4 + 1.207(0.419)BIC4 + 2.992(0.420)MATS1v + 1.707(0.420)MATS8e + 2.310(0.249)nRORPh

n = 27, r = 0.917, s = 0.485, F =22.454, q^2_{LOO} = 0.713, q^2_{L5O} = 0.692, r^2_{Test} = 0.581 (8)

pIC_{50} = 3.890 + 1.636(0.501)PW4 + 1.192(0.455)BIC4 + 2.228(0.676)GGI4 + 1.979(0.525)MATS1v + 1.392(0.339)nRORPh

n = 27, r = 0.902, s = 0.527, F = 18.425, q^2_{LOO} = 0.654, q^2_{L5O} = 0.567, r^2_{Test} = 0.638 (9)

pIC_{50} = 5.076 –1.484(0.471)MAXDN + 1.365(0.464)BIC4 + 1.948(0.640)GGI4 + 2.064(0.523)MATS1v + 1.542(0.325)nRORPh

n = 27, r = 0.899, s = 0.533, F =17.880, q^2_{LOO} = 0.660, q^2_{L5O} = 0.647, r^2_{Test} = 0.833 (10)

pIC_{50} = 8.563 – 1.139(0.370)nN – 1.826(0.521)PW2 + 2.947(0.582)GGI4 – 1.412(0.429)MATS2v + 1.187(0.466)C-040

n = 27, r = 0.897, s = 0.540, F = 17.371, q^2_{LOO} = 0.669, q^2_{L5O} = 0.647, r^2_{Test}= 0.558 (11)

where n and F represent respectively the number of data points and the F-ratio between the variances of calculated and observed activities. The data within the parentheses are the standard errors associated with regression coefficients. In all above equations, the F-values remained significant at 99% level. The indices q^2_{LOO} and q^2_{L5O} (> 0.5) have accounted for their internal robustness. For all above models the r^2_{Test} values, obtained greater than 0.5, specified that the selected test-set is fully accountable for their external validation. The descriptors, in all above models, have been scaled between the intervals 0 to 1[51] to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors. In this way, the scaled descriptors would have equal potential to influence the QSAR models.

The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above 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.

Though Equations (1)-(11) emerged as significant predictive models but Equations (7)-(11) remained statistically more efficient. The later five models, involving five descriptors in each, could estimate up to 85.04 percent of variance in observed activity of the compounds. In fact, a total number of five such models, sharing 12 descriptors among them, have been obtained through CP-MLR and all of them have been documented through Equation (7)-(11). The shared 12 descriptors along with their brief description, average regression coefficients and total incidences are given in Table 2.

Table 2: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the MMP-13 inhibition activity.

S. No.	Descriptor	Descriptor class	Physical meaning	Average regression coefficient (incidence)
1	nN	Constitutional	Number of nitrogens	-1.139(1)
2	MAXDN	Topological	Maximal electrotopological negative variation	-1.484(1)
3	MAXDP	Topological	Maximal electrotopological positive variation	1.336(1)
4	PW2	Topological	Path/walk 2-Randic shape index	-1.826(1)
5	PW4	Topological	Path/walk 4-Randic shape index	1.323(2)
6	BIC4	Topological	Bond information content of neighborhood symmetry of 4-order	1.220(4)
7	GGI4	Galvez topological charge indices	Topological charge index of order 4	2.374(3)
8	MATS1v	2D autocorrelations	Moran autocorrelation of lag-1/ weighted by atomic van der Waals volumes	2.307(4)
9	MATS2v	2D autocorrelations	Moran autocorrelation of lag-2/ weighted by atomic van der Waals volumes	-1.412(1)
10	MATS8e	2D autocorrelations	Moran autocorrelation of lag-8/ weighted by atomic Sanderson electronegativities	1.406(2)
11	nRORPh	Functional	Number of ethers (aromatic)	1.737(4)
12	C-040	Atom-centered fragments	R-C(=X)-X/ R-C#X/X-=C=X	1.187(1)

[a]The descriptors are identified from the five parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.871, and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 27 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.

Besides listed descriptors in Table 2, the other identified descriptor nNR2 is from functional groups class and corresponds to number of tertiary aliphatic amines (Equation 1). The further discussion is, however, based on the highest significant Equations (7)-(11). The derived statistical parameters of these five models have shown that these models are significant. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 3 for the sake of comparison with observed ones. A close agreement between them has been observed.

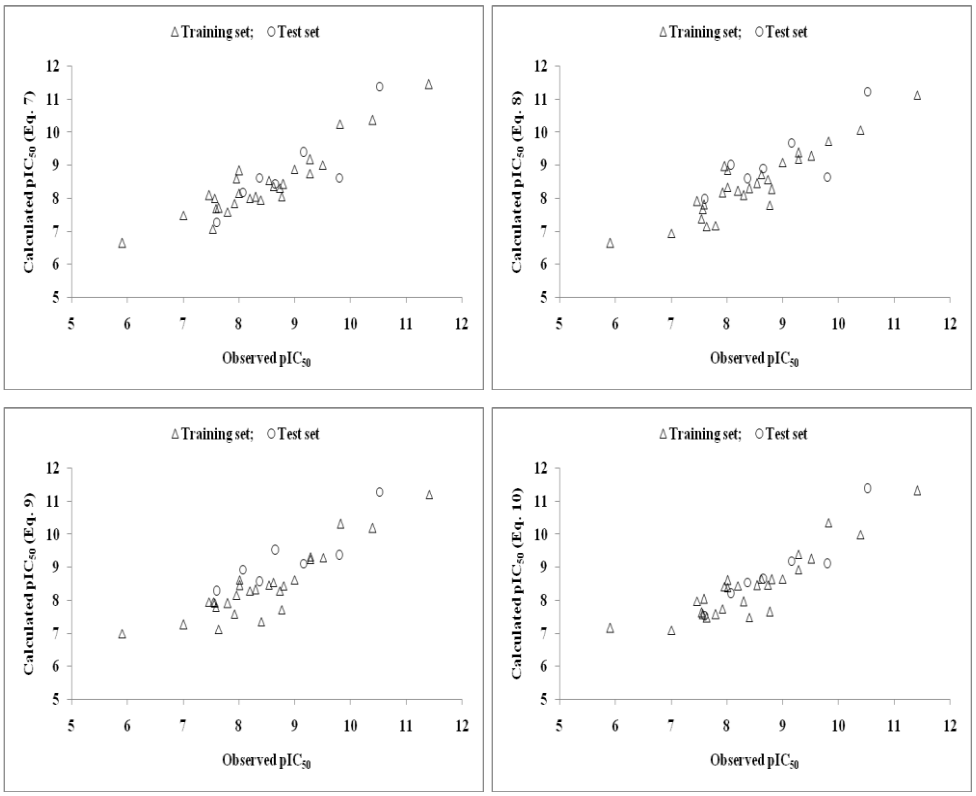
Table 3: Observed and calculated MMP-13 inhibition activity of quinazoline-2-carboxamide compounds.

Cpd.	$pIC_{50}(M)$[a]					
	Obsd.[b]	Calc. Eq. (7)	Calc. Eq. (8)	Calc. Eq. (9)	Calc. Eq. (10)	Calc. Eq.(11)	Calc. PLS
1	7.92	7.85	8.18	7.59	7.75	7.53	7.83
2	8.30	8.05	8.10	8.33	7.96	7.59	8.19
3	7.54	7.08	7.39	7.96	7.64	7.04	7.21
4[c]	7.60	7.28	7.97	8.28	7.54	7.49	7.68
5[c]	8.07	8.19	9.00	8.94	8.20	8.17	8.35
6[c]	8.66	8.44	8.91	9.52	8.64	8.89	8.86
7	9.28	8.76	9.19	9.24	8.94	9.12	8.92
8	9.51	9.01	9.29	9.29	9.27	8.91	8.83
9[c]	9.16	9.39	9.68	9.12	9.20	8.88	9.05
10	9.28	9.18	9.39	9.33	9.40	9.34	9.07
11	8.80	8.44	8.27	8.45	8.64	9.11	8.79
12[c]	8.37	8.63	8.59	8.56	8.55	7.83	7.80
13	8.01	8.85	8.86	8.62	8.61	8.88	8.79
14	8.54	8.56	8.46	8.48	8.46	8.96	8.69
15	8.77	8.06	7.81	7.72	7.66	8.50	8.56
16	7.80	7.60	7.18	7.91	7.58	7.48	7.57
17[c]	9.80	8.62	8.66	9.36	9.10	9.59	8.87
18	9.00	8.89	9.09	8.63	8.66	9.08	8.86
19	7.47	8.10	7.92	7.95	7.98	7.58	7.75
20	10.40	10.38	10.08	10.19	10.00	9.86	10.35
21	9.82	10.26	9.74	10.32	10.36	10.11	10.33
22[c]	10.52	11.40	11.24	11.28	11.38	8.82	9.54

23	11.41		11.46		11.14		11.22		11.32		11.20		11.34
24	7.96		8.60		8.98		8.17		8.42		8.14		8.69
25	7.59		7.70		7.81		7.80		8.04		7.59		7.76
26	7.01		7.49		6.95		7.28		7.11		7.51		7.02
27	8.40		7.95		8.31		7.37		7.49		7.74		8.07
28	8.62		8.37		8.73		8.55		8.66		8.66		8.67
29	7.57		8.01		7.67		7.92		7.58		7.86		7.53
30[d]	-[d]		-[d]		-[d]		-[d]		-[d]		-[d]		-[d]
31	8.74		8.32		8.58		8.29		8.47		8.14		8.28
32	7.64		7.71		7.15		7.14		7.48		7.52		7.36
33	8.01		8.16		8.34		8.46		8.40		8.41		8.18
34	8.20		8.01		8.22		8.29		8.45		8.41		8.03
35	5.91		6.66		6.65		7.00		7.17		7.25		6.82
36[e]	-[e]		-[e]		-[e]		-[e]		-[e]		-[e]		-[e]

[a]On molar basis; [b]Taken from reference [33]; [c]Compounds in test set; [d]Outlier compound not considered in study; [e]Compound with uncertain activity not part of data set.

Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 2 to ensure the goodness of fit for each of these five models.



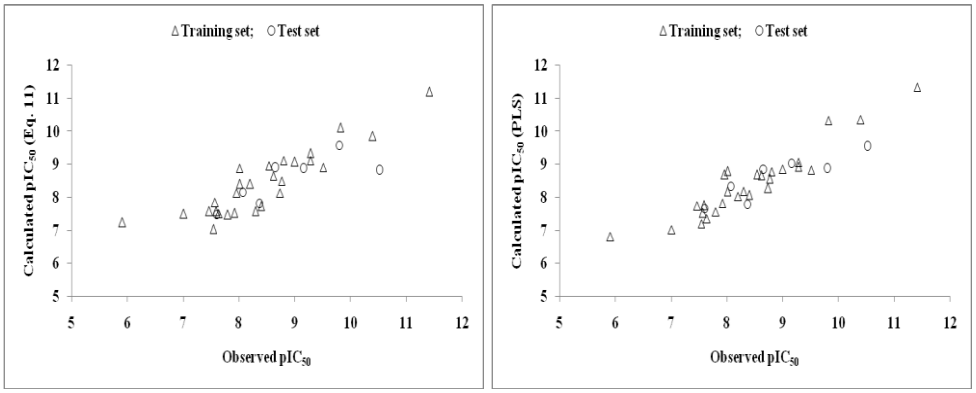


Figure 2. Plot of observed versus caculated pIC_{50} values for training- and test-set compounds.

The descriptors MAXDP, MAXDN, PW2, PW4 and BIC4 participated in these models are topological class descriptors. These descriptors represent, respectively, maximal electrotopological positive variation, maximal electrotopological negative variation, path/walk 2-Randic shape index, path/walk 4-Randic shape index and bond information content of neighborhood symmetry of 4-order. The positive sign of regression coefficient of descriptors MAXDP, PW4 and BIC4 suggest that a higher value of these descriptors is beneficiary to the MMP-13 inhibition activity whereas a lower value of descriptors MAXDN and PW2 will augment the inhibition activity.

The descriptors MATS1v (Moran autocorrelation of lag-1/weighted by atomic van der Waals volumes), MATS2v (Moran autocorrelation of lag-2/weighted by atomic van der Waals volumes) and MATS8e (Moran autocorrelation of lag-8/weighted by atomic Sanderson electronegativities) are 2D autocorrelation descriptors. It is evinced from the models mentioned above the descriptors MATS1v and MATS8e contributed positively and descriptor MATS2v negatively to the activity. Thus a higher value of descriptors MATS1v and MATS8e and a lower value of descriptor MATS2v will be supportive to enhance the inhibition activity.

From Equations (7)-(11), it appeared that the descriptors nRORPh, a functional group accounting descriptor representing number of aromatic ethers functionality in a structure; GGI4, Galvez topological charge index of order 4 and atom centered fragment accounting descriptor C-040 showing R-C(=X)-X/ R-C#X/X-=C=X type fragments in a molecular structure make positive contribution to activity and nN, number of nitrogen atoms present in molecular constitution shown negative correlation to the activity. In this way presence of

aromatic ethers functionality, R-C(=X)-X/ R-C#X/X=C=X type fragments in a molecular structure along with Galvez topological charge index of order 4 and absence of nitrogen atoms in a compound would be advantageous in improving the MMP-13 inhibition activity of a compound.

To corroborate the study further, a PLS analysis has also been carried out on 12 descriptors identified through CP-MLR and results are given in Table 4. For this purpose, the descriptors have been autoscaled (zero mean and unit s.d.) to give each one of them equal weight in the analysis. In the PLS cross-validation, three components have been found to be the optimum for these 12 descriptors and they explained 88.2% variance in the activity (r^2 = 0.882). The MLR-like PLS coefficients of these 12 descriptors are given in Table 4.

Table 4: PLS and MLR-like PLS models from the descriptors of five parameter CP-MLR models for MMP-13 inhibition activity.

A: PLS equation			
PLS components		PLS coefficient (s.e.)[a]	
Component-1		-0.619 (0.049)	
Component-2		-0.162 (0.048)	
Component-3		0.170 (0.073)	
Constant		8.425	
B: MLR-like PLS equation			
S. No.	Descriptor	MLR-like coefficient (f.c.)[b]	Order
1	nN 3	-0.271(-0.117)	3
2	MAXDN 12	-0.099(-0.042)	12
3	MAXDP 1	0.278(0.120)	1
4	PW2 10	-0.104(-0.045)	10
5	PW4 9	0.135(0.058)	9
6	BIC4 8	0.157(0.068)	8
7	GGI4 2	0.276(0.119)	2
8	MATS1v 11	0.102(0.044)	11
9	MATS2v 6	-0.207(-0.089)	6
10	MATS8e 5	0.238(0.103)	5
11	nRORPh 4	0.261(0.113)	4
12	C-040 7	0.178(0.077)	7
	Constant	6.383	
C: PLS regression statistics		Values	
n		27	
r		0.939	
s		0.402	
F		57.249	
q^2_{LOO}		0.842	
q^2_{L5O}		0.831	
r^2_{Test}		0.703	

[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; 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.

The calculated activity values of training- and test-set compounds are in close agreement to that of the observed ones and are listed in Table 2. For the sake of comparison, the plot between observed and calculated activities (through PLS analysis) for the training- and test-set compounds is given in Figure 2. Figure 3 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity (Table 4).

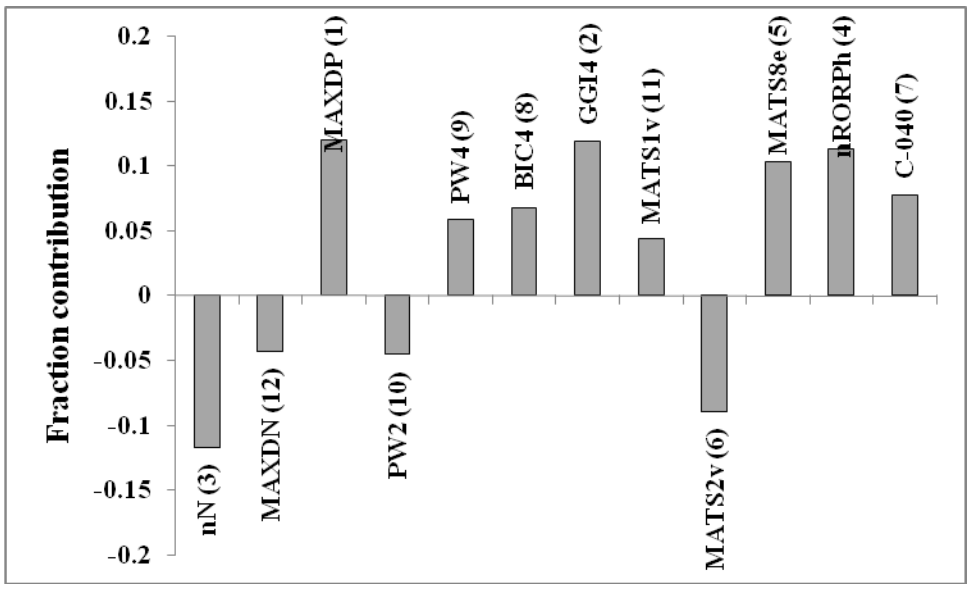


Figure 3: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 12 identified descriptors (Table 4) associated with MMP-13 inhibition activity of the compounds.

The PLS analysis in 12 identified descriptors revealed three components (Table 4) as optimum to explain the MMP-13 inhibition activity. The decreasing order of significance of all the participated descriptors is MAXDP, GGI4, nN, nRORPh, MATS8e, MATS2v, C-040, BIC4, PW4, PW2, MATS1v, MAXDN (Table 4, figure 3). All the descriptors are part of Equations discussed above and convey same inferences in PLS analysis. It is also observed that PLS model from the dataset devoid of 12 descriptors (Table 4) remained inferior in explaining the activity of the analogues.

3.2 Applicability domain

On analyzing the applicability domain (AD) in the Williams plot (Figure 3) of the model based on the whole data set (Table 5), no any compound has been identified as an obvious 'outlier' for the MMP-13 inhibitory activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units.

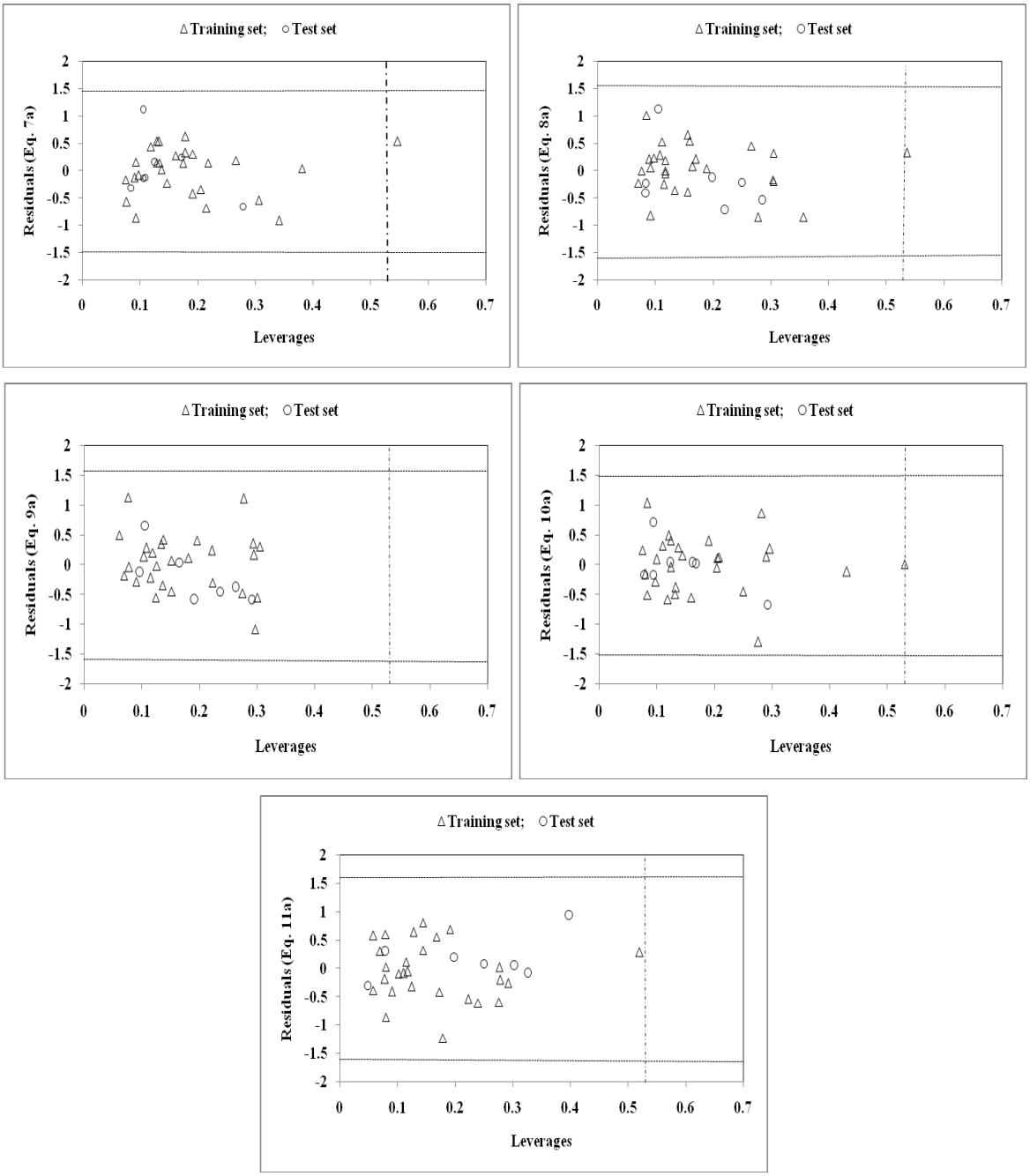


Figure 4. Williams plot for the training-set and test- set for inhibition activity of MMP-13 for the compounds in Table 1. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.529).

Table 5: Models derived for the whole data set (n = 34) for the MMP-13 inhibition activity in descriptors identified through CP-MLR.

Model	r	s	F	q^2_{LOO}	Eq.
pIC_{50} = 5.141 + 1.385(0.460)MAXDP + 1.020(0.402)BIC4 + 1.946(0.488)MATS1v + 1.101(0.395)MATS8e + 1.456(0.311)nRORPh	0.907	0.493	26.279	0.705	(7a)
pIC_{50} = 3.988 + 0.736(0.331)PW4 + 1.164(0.429)BIC4 + 2.900(0.367)MATS1v + 1.352(0.430)MATS8e + 2.145(0.230)nRORPh	0.895	0.523	22.736	0.689	(8a)
pIC_{50} = 4.241 + 1.114(0.379)PW4 + 1.295(0.429)BIC4 + 1.793(0.580)GGI4 + 2.030(0.438)MATS1v + 1.480(0.281)nRORPh	0.894	0.525	22.488	0.676	(9a)
pIC_{50} = 5.278 – 1.511(0.438)MAXDN + 1.337(0.412)BIC4 + 1.919(0.558)GGI4 + 1.867(0.432)MATS1v + 1.396(0.274)nRORPh	0.903	0.503	24.999	0.713	(10a)
pIC_{50} = 8.419 – 0.678(0.266)nN – 1.825(0.495)PW2 + 3.475(0.496)GGI4 – 1.512(0.360)MATS2v + 0.760(0.333)C-040	0.891	0.533	21.722	0.673	11(a)

One of the compound (**27**; Table 1) was found to have leverage (h) values greater than the threshold leverage (h*); suggesting it as chemically influential compound. 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.

4. CONCLUSION

The MMP-13 inhibition activity of quinazoline-2-carboxamide compounds has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the inhibition activity of these congeners. The descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of maximal electrotopological positive and negative variations (MAXDP and MAXDN, respectively), path/walk 2- and 4-Randic shape index (PW2 and PW4, respectively), bond information content of neighborhood symmetry of 4-order (BIC4), Moran autocorrelations of lag-1 and -2/weighted by atomic van der Waals volumes (MATS1v and MATS2v, respectively) and of lag-8/ weighted by atomic Sanderson electronegativities (MATS8e). In addition to these 4^{th} order Galvez topological charge index (GGI4), number of nitrogen atoms (nN), aromatic ethers functionality (nRORPh) and R-C(=X)-X/ R-C#X/X=C=X type structural fragments (C-040) have also shown prevalence to model the inhibitory activity.

From statistically validated models, it appeared that the descriptors MAXDP, BIC4, PW4, GGI4, nRORPh, MATS8e, MATS1v and C-040 make positive contribution to activity and their higher values are conducive in improving the MMP-13 inhibition activity of a compound. On the other hand, the descriptors nN, MAXDN, PW2 and MATS2v render detrimental effect to activity. Therefore, lower values of descriptors MAXDN, PW2 and MATS2v and the absence of nitrogen atoms in molecular structure (nN) would be advantageous. Such guidelines may be helpful in exploring more potential analogues of the series. The statistics emerged from the test sets have validated the identified significant models. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested models have acceptable predictability. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

Compliance with ethical standards

Acknowledgements

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

Disclosure of conflict of interest

The authors declare no conflict of interest.

REFERENCES

1. Fitzgerald GA. (2004). Coxibs and cardiovascular disease. New England Journal of Medicine, 351: 1709-1711.
2. Knauper V, Lopez-Otin C, Smith B, Knight G, and Murphy G. (1996). Biochemical characterization of human collagenase-3. Journal of Biological Chemistry, 271: 1544-1550.
3. Mitchell PG, Magna HA, Reeves LM, Lopresti-Morrow LL, Yocum SA, Rosner PJ, Geoghegan KF and Hambor JE. (1996). Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. Journal of Clinical Investigation, 97: 761-768.
4. Yocum SA, Lopresti-Morrow LL, Reeves LM and Mitchell PG. (1999). MMP-13 and MMP-1 expression in tissues of normal articular joints. Annals of New York Academy of Sciences, 878: 583-586.

5. Wernicke D, Seyfert C, Hinzmann B and Gromnica IE. (1996). Cloning of collagenase 3 from the synovial membrane and its expression in rheumatoid arthritis and osteoarthritis. Journal of Rheumatology, 23: 590-595.
6. Knäuper V, Will H, López-Otín C, Smith B, Atkinson SJ, Stanton H, Hembry RM and Murphy G. (1996). Cellular mechanisms for human procollagenase-3 (MMP-13) activation. Evidence that MT1-MMP (MMP-14) and gelatinase a (MMP-2) are able to generate active enzyme. Journal of Biological Chemistry, 271: 17124-17131.
7. Takaishi H, Kimura T, Dalal S, Okada Y and D'Armiento J. (2008). Joint diseases and matrix metalloproteinases: a role for MMP-13. Current Pharmaceutical Biotechnology, 9: 47-54.
8. Rowan AD, Litherland GJ, Hui W and Milner JM. (2008). Metalloproteases as potential therapeutic targets in arthritis treatment. Expert Opinion on Therapeutic Targets, 12: 1-18.
9. Sabatani M, Lesur C, Thomas M, Chomel A, Anract P, de Nanteuil G, Pastoureau P, Janusz MJ, Bendele AM, Brown KK, Taiwo YO, Hsieh L and Heitmeyer SA. (2005). Effect of inhibition of matrix metalloproteinases on cartilage loss in vitro and in a guinea Pig model of osteoarthritis. Arthritis and Rheumatology, 52: 171-180.
10. Drummond AH, Beckett P, Brown PD, Bone EA, Davidson AH, Galloway WA, Gearing AJ, Huxley P, Laber D, McCourt M, Whittaker M, Wood LM and Wright A. (1999). Preclinical and clinical studies of MMP inhibitors in cancer. Annals of New York Academy of Sciences, 878: 228-235.
11. Hutchinson JW, Tierney GM, Parsons SL and Davies TRC. (1998). Dupuytren's disease and frozen shoulder induced by treatment with a matrix metalloproteinase inhibitor. Journal of Bone and Joint Surgery British Volume, 80: 907-908.
12. Wojtowicz-Praga S, Torri J, Johnson M, Steen V, Marshall J, Ness E, Dickson R, Sale M, Rasmussen HS, Chiodo TA and Hawkins MJ. (1998). Phase I trial of Marimastat, a novel matrix metalloproteinase inhibitor, administered orally to patients with advanced lung cancer. Journal of Clinical Oncolology, 16: 2150-2156.
13. Griffioen AW. (2000). AG-3340 (Agouron Pharmaceuticals Inc). *IDrugs*, 3: 336-345.
14. Levitt NC, Eskens FALM, O'Byrne KJ, Propper DJ, Denis LJ, Owen SJ, Choi L, Foekens JA, Wilner S, Wood JM, Nakajima M, Talbot DC, Steward WP, Harris AL and Verweij J. (2001). Phase I and pharmacological study of the oral matrix metalloproteinase inhibitor, MMI270 (CGS27023A), in patients with advanced solid cancer. Clinical Cancer Research, 7: 1912-1922.

15. Skiles JW, Gonnella NC and Jeng AY. (2001). The design, structure, and therapeutic application of matrix metalloproteinase inhibitors. Current Medicinal Chemistry, 8: 425-474.
16. Clark IM and Parker AE. (2003). Metalloproteinases: their role in arthritis and potential as therapeutic targets. Expert Opinion on Therapeutic Targets, 7: 19-34.
17. Peterson JT. (2004). Matrix metalloproteinase inhibitor development and the remodeling of drug discovery. Heart Failure Reviews, 9: 63-79.
18. Stickens D, Behonick DJ, Ortega N, Heyer B, Hartenstein B, Yu Y, Fosang AJ, Schorpp-Kistner M, Angel P and Werb Z. (2004). Altered endochondral bone development in matrix metalloproteinase 13-deficient mice. Development, 131: 5883-5895.
19. Inada M, Wang Y, Byrne MH, Rahman MU, Miyaura C, López-Otín C and Krane SM. (2004). Critical roles for collagenase-3 (Mmp13) in development of growth plate cartilage and in endochondral ossification. Proceedings of the National Academy of Sciences USA, 101: 17192-17197.
20. Kennedy AM, Inada M, Krane SM, Christie PT, Harding B, López-Otín C, Sánchez LM, Pannett AJ, Dearlovem A, Hartley C, Byrne MH, Reed A, Nesbit MA, Whyte MP and Thakker RV. (2005). MMP13 mutation causes spondyloepimetaphyseal dysplasia, Missouri type (SEMD(MO)). Journal of Clinical Investigations, 115: 2832-2842.
21. Johnson AR, Pavlovsky AG, Ortwine DF, Prior F, Man CF, Bornemeier DA, Banotai CA, Mueller WT, McConnell P, Yan C, Baragi VM, Lesch C, Roark WH, Wilson M, Datta K, Guzman R, Han HK and Dyer RD. (2007). Discovery and characterization of a novel inhibitor of matrix metalloprotease-13 that reduces cartilage damage in vivo without joint fibroplasia side effects. Journal of Biological Chemistry, 282: 27781-27791.
22. Baragi VM, Becher G, Bendele AM, Biesinger R, Bluhm H, Boer J, Deng H, Dodd R, Essers M, Feuerstein T, Gallagher BM, Gege C, Hochgürtel M, Hofmann M, Jaworski A, Jin L, Kiely A, Korniski B, Kroth H, Nix D, Nolte B, Piecha D, Powers TS, Richter F, Schneider M, Steeneck C, Sucholeiki I, Taveras A, Timmermann A, Veldhuizen JV, Weik J, Wu X and Xia B. (2009). A new class of potent matrix metalloproteinase 13 inhibitors for potential treatment of osteoarthritis: Evidence of histologic and clinical efficacy without musculoskeletal toxicity in rat models. Arthritis and Rheumatology, 60: 2008-2018.
23. Reiter LA, Freeman-Cook KD, Jones CS, Martinelli GJ, Antipas AS, Berliner MA, Datta K, Downs JT, Eskra JD, Forman MD, Greer EM, Guzman R, Hardink JR, Janat F, Keene NF, Laird ER, Liras JL, Lopresti-Morrow LL, Mitchell PG, Pandit J, Robertson D,

Sperger D, Vaughn-Bowser ML, Waller DM and Yocum SA. (2006). Potent, selective pyrimidinetrione-based inhibitors of MMP-13. Bioorganic and Medicinal Chemistry Letters, 16: 5822-5826.

24. Reiter LA, Robinson RP, McClure KF, Jones CS, Reese MR, Mitchell PG, Otterness IG, Bliven ML, Liras J, Cortina SR, Donahue KM, Eskra JD, Griffiths RJ, Lame ME, Lopez-Anaya A, Martinelli GJ, McGahee SM, Yocum SA, Lopresti-Morrow LL, Tobiassen LM and Vaughn-Bowser ML. (2004). Pyran containing sulfonamide hydroxamic acids: potent MMP inhibitors that spare MMP-1. Bioorganic and Mediinal Chemistry Letters, 14: 3389-3395.
25. Hoekstra R, Eskens FA and Verweij J. (2001). Matrix metalloproteinase inhibitors: current developments and future perspectives. Oncologist, 6: 415-427.
26. Peterson JT. (2006). The importance of estimating the therapeutic index in the development of matrix metalloproteinase inhibitors. Cardiovascular Research, 69: 677-687.
27. Puerta DT and Cohen SM. (2004). A bioinorganic perspective on matrix metalloproteinase inhibition. Current Topics in Medicinal Chemistry, 4: 1551-1573.
28. Breuer E, Frant J and Reich R. (2005). Recent non-hydroxamate matrix metalloproteinase inhibitors. Expert Opinion on Therapeutic Patents, 15: 253-269.
29. Engel CK, Pirard B, Schimanski S, Kirsch R, Habermann J, Klinger O, Schlotte V, Weithmann KU and Wendt KU. (2005). Structural basis for the highly selective inhibition of MMP-13. Chemical Biology, 12: 181-189.
30. Hu Y, Xiang JS, DiGrandi MJ, Du X, Ipek M, Laakso LM, Li J, Li W, Rush TS, Schmid J, Skotnicki JS, Tam S, Thomason JR, Wang Q and Levin JI. (2005). Potent, selective, and orally bioavailable matrix metalloproteinase-13 inhibitors for the treatment of osteoarthritis. Bioorganic and Mediinal Chemistry, 13: 6629-6644.
31. Wu J, Rush TS III, Hotchandani R, Du X, Geck M, Collins E, Xu ZB, Skotnicki J, Levin JL and Lovering FE. (2005). Identification of potent and selective MMP-13 inhibitors. Bioorganic and Mediinal Chemistry Letters, 15: 4105-4109.
32. Reiter LA, Freeman-Cook KD, Jones CS, Martinelli GJ, Antipas AS, Berliner MA, Datta K, Downs JT, Eskra JD, Forman MD, Greer EM, Guzman R, Hardink JR, Janat F, Keene NF, Laird ER, Liras JL, Lopresti-Morrow LL, Mitchell PG, Pandit J, Robertson D, Sperger D, Vaughn-Bowser ML, Waller DM and Yocum SA. (2006). Potent, selective pyrimidinetrione-based inhibitors of MMP-13. Bioorganic and Medicinal Chemistry Letters, 16: 5822-5826.

33. Nara H, Sato K, Naito T, Mototani H, Oki H, Yamamoto Y, Kuno H, Santou T, Kanzaki N, Terauchi J, Uchikawa O and Kori M. (2014). Discovery of novel, highly potent, and selective quinazoline-2-carboxamide-based matrix metalloproteinase (MMP)-13 inhibitors without a zinc binding group using a structure-based design approach. Journal of Medicinal Chemistry, 57: 8886-8902.
34. SYSTAT, Version 7.0; SPSS Inc 444 North Michigan Avenue, Chicago IL, 60611.
35. Chemdraw ultra 6.0 and Chem3D ultra, Cambridge Soft Corporation, Cambridge, USA. http://www.cambridgesoft.com
36. Dragon software (version 1.11-2001) by Todeschini R.; Consonni V. Milano, Italy. http//www.talete.mi.it/dragon.htm
37. Prabhakar YS. (2003). A combinatorial approach to the variable selection in multiplelinear regression: analysis of Selwood et al. Data Set-a case study. QSAR and Combinatorial Science, 22: 583-595.
38. Wold S. (1978). Cross-validatory estimation of the number of components in factor and principal components models. Technometrics, 20: 397–405.
39. Kettaneh N, Berglund A and Wold S. (2005). PCA and PLS with very large data sets. Computational Statistics and Data Analysis, 48: 69-85.
40. Stahle L and Wold S. (1988). Multivariate data analysis and experimental design. In: Ellis GP and West WB. (Eds.), Biomedical research. Progress in medicinal chemistry. Elsevier Science Publishers, BV, Amsterdam, 25: 291–338.
41. Sharma S, Prabhakar YS, Singh P and Sharma BK. (2008). QSAR study about ATP-sensitive potassium channel activation of cromakalim analogues using CP-MLR approach. European Journal of Medicinal Chemistry, 43: 2354-2360.
42. Sharma S, Sharma BK, Sharma SK, Singh P and Prabhakar YS. (2009). Topological descriptors in modeling the agonistic activity of human A3 adenosine receptor ligands: The derivatives of 2-Chloro-N^6-substituted-4'-thioadenosine-5'-uronamide. European Journal of Medicinal Chemistry, 44: 1377-1382.
43. Sharma BK, Pilania P, Singh P and Prabhakar YS. (2010). Combinatorial protocol in multiple linear regression/partial least-squares directed rationale for the caspase-3 inhibition activity of isoquinoline-1,3,4-trione derivatives. SAR and QSAR in Environmental Research, 21: 169-185.
44. Sharma BK, Singh P, Sarbhai K and Prabhakar YS. (2010). A quantitative structure-activity relationship study on serotonin 5-HT6 receptor ligands: Indolyl and piperidinyl sulphonamides. SAR and QSAR in Environmental Research, 21: 369-388.

45. Sharma BK, Pilania P, Sarbhai K, Singh P and Prabhakar YS. (2010). Chemometric descriptors in modeling the carbonic anhydrase inhibition activity of sulfonamide andsulfamate derivatives. Molecular Diversity, 14: 371-384.
46. Sharma BK, Singh P, Shekhawat M, Sarbhai K and Prabhakar YS. (2011). Modelling of serotonin reuptake inhibitory and histamine H_3 antagonistic activity of piperazine and diazepane amides: QSAR rationales for co-optimization of the activity profiles., SAR and QSAR in Environmental Research, 22: 365-383.
47. So S-S and Karplus M. (1997). Three-dimensional quantitative structure activity relationships from molecular similarity matrices and genetic neural networks. 1. Method and validations. Journal of Medicinal Chemistry, 40: 4347-4359.
48. Prabhakar YS, Solomon VR, Rawal RK, Gupta MK and Katti SB. (2004). CP-MLR/PLS directed structure–activity modeling of the HIV-1 RT inhibitory activity of 2,3-diaryl-1,3-thiazolidin-4-ones. QSAR and Combinatorial Science, 23: 234-244.
49. Gramatica P. (2007). Principles of QSAR models validation: internal and external. QSAR and Combinatorial Science, 26: 694-701.
50. Eriksson L, Jaworska J, Worth AP, Cronin MTD, McDowell RM and Gramatica P. (2003). Methods for reliability and uncertainty assessment and for applicability evaluations of classification and regression-based QSARs. Environmental and Health Perspectives, 111: 1361-1375.
51. Golbraikh A and Tropsha A. (2002). Beware of q2!. Journal of Molecular Graphics and Modeling, 20: 269-276.