WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 8.084
Volume 11, Issue 5, 1714-1736. Research Article ISSN 2277– 7105

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY ON THE MMP-13 INHIBITORY ACTIVITY OF FUSED PYRIMIDINE DERIVATIVES POSSESSING A NON-CARBOXYLATE ZINC-BINDING FUNCTION

Jahan Afsar, Meena Dinesh Kumar and Sharma Brij Kishore*

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

Article Received on
11 March 2022,

Revised on 01 April 2022,
Accepted on 22 April 2022

DOI: 10.20959/wjpr20225-23943

***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-13 inhibitory activity of fused pyrimidine derivatives possessing a non-carboxylate zinc-binding function in 0D- to 2D-Dragon descriptors. The derived QSAR models have revealed that the information content indices of 1st and 2nd order neighborhood symmetry (descriptors IC1 and IC2), Lovasz-Pelikan index (descriptor LP1), maximal electrotopological positive variation (descriptor MAXDP) and molecular walk count of order 09 (descriptor MWC09) played a pivotal role in rationalization of MMP-13 inhibition activity of titled compounds. Atomic properties such as mass and atomic Sanderson electronegativity in terms of atomic properties weighted descriptors MATS2m, GATS3e, GATS4e, GATS7e and Me, certain atom centred fragments such as R-C(=X)-X/R-C#X/X-=C=X (descriptor C-040) and aliphatic hydroxylamine functionality (descriptor nNHOH) are also predominant to explain MMP-13 inhibition actions of fused pyrimidines. 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-13 inhibitory activity; Combinatorial protocol in multiple linear regression (CP-MLR) analysis; PLS analysis; Dragon descriptors; Fused pyrimidines. Non-carboxylate zinc binding group.

1. INTRODUCTION

Osteoarthritis (OA), a chronic disease and at the top of arthralgia for aged patients[1-3], is characterized by articular cartilage destruction with aging and mechanical stress causes progressive degradation and a big health concern for aging population.[4] Environmental factors and multiple genetic factors are also supposed to be associated in the progression of OA.[5-8] It is the need of the hour to develop novel and safer disease modifying osteoarthritis drugs (DMOADs) which may reduce or reverse the cartilage destruction as there are limited existing OA treatments including symptomatic relief with NSAIDs, intra-articular injections of hyaluronic acid conjugates, or surgical joint replacement. Additionally, there is increased risk of cardiovascular side effects with the COX-2 inhibitors.[9] There may be a role of a family of zinc-dependent, calcium-containing endopeptidases, matrix metalloproteinases (MMPs), in OA as MMP inhibitors showed ability to prevent the destruction of cartilage in preclinical testing.[10] The dose-limiting toxicity such as skin rash and musculoskeletal side effects like joint stiffness, pain, inflammation, and tendinitis has stopped the progress of most clinical trials of broad spectrum MMP inhibitors. The postulate that these side effects are due to the inhibition of MMP-1, MMP-14, or sheddases such as TACE[11-13] has increased interest in more selective MMP-13 inhibitors. MMP-13 (collagenase-3) cleaves type II collagen efficiently[14,15] and has overexpression in OA cartilage[14,16] rendered MMP-13 inhibition as one of the most potential approach to cure the degradation of cartilage in arthritis.[17-32]

The highly selective non zinc binding MMP-13 inhibitors[33-36] having a pyrimidin-4-one-2-carboxamide core effectively bind to the primed regions (S1') of the catalytic active site[36,37] and the fused pyrimidine carboxamides have successful application to other selective MMP-13 inhibitors.[38-40] The zinc binding group (ZBG), the side chains that may interact with amino acids around the catalytic zinc ion and the pocket-occupying functionality referred to as the P1' group and that bound in the S1' pocket[41-47] are the major components of the MMP inhibitors. Most of the MMP inhibitors gain affinity by interaction with the catalytic zinc through a chelating moiety and by positioning hydrogen bonding groups near the catalytic zinc and this approach raises issue of selectivity as many of the important residues in the catalytic site are well-conserved among the MMP family whereas the conformational diversity is localized at the S1' loop region forming the bottom half of the S1' subsite. MMPs including MMP-2, MMP-3, MMP-8, MMP-9, MMP-12, MMP-13 and MMP-14 possess large S1' pockets where more diverse large P1' groups can be accommodated therefore

targeting the large S1' loop region may be helpful in the identification of MMP-13-specific inhibitors.[19-28,48-52]

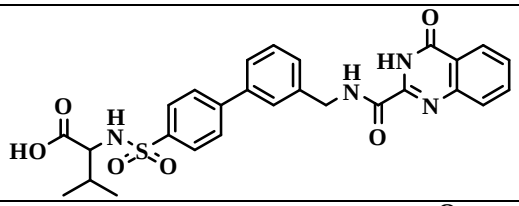
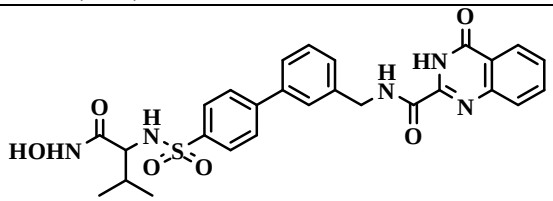
A drug used to treat OA has to reach the site of action through diffusion into the cartilage matrix. Efforts were made to incorporate non-carboxylate moieties or intra-articular therapy[53] as the ionization of carboxylic acid at physiological pH may lead to poor penetration through the negatively charged cartilage matrix.[54-56] A novel series of MMP-13 inhibitors, possessing a non-carboxylate zinc-binding function attached to the pyrimidin-4-one-2-carboxamide core has been synthesized by Nara *et al.*[57] 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.

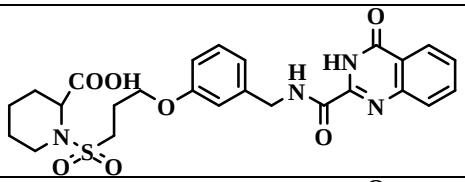
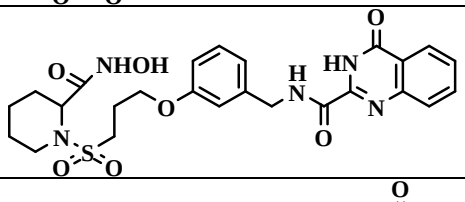
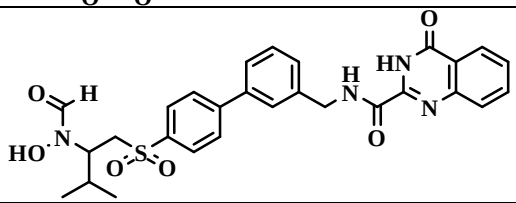
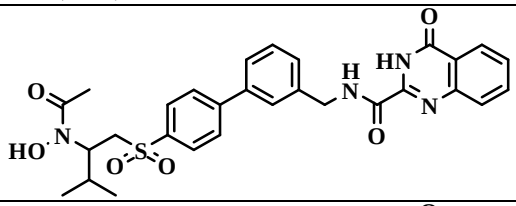
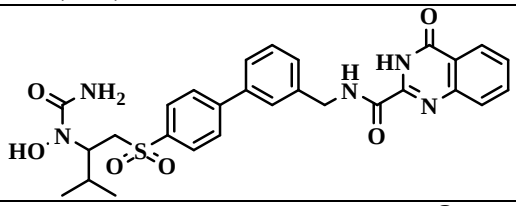
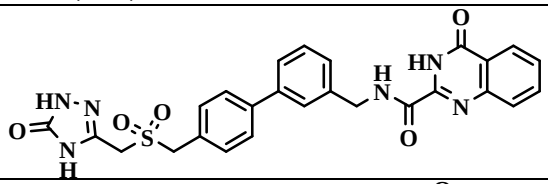
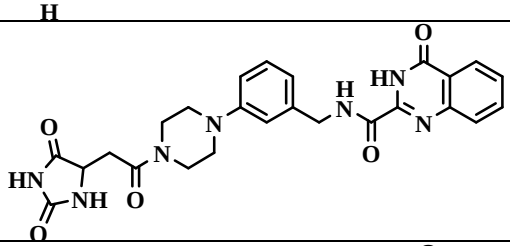
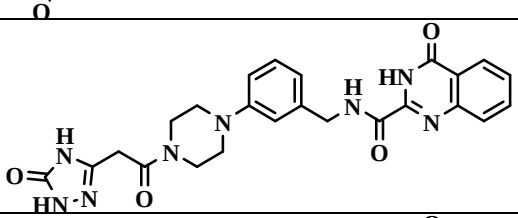
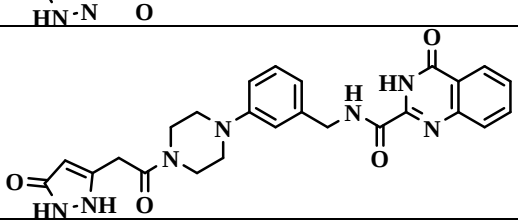
2. MATERIALS AND METHODS

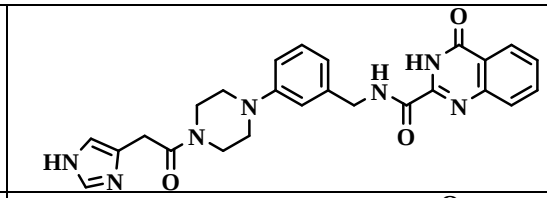
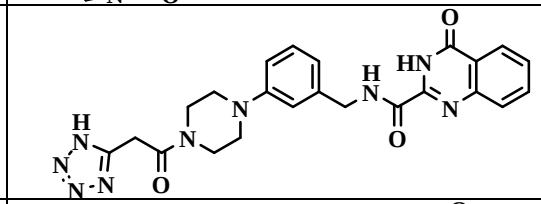
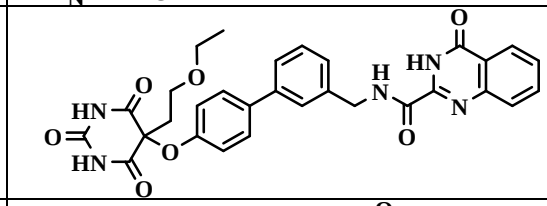
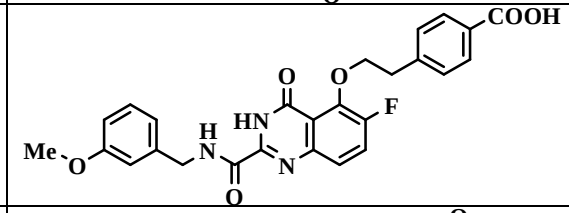
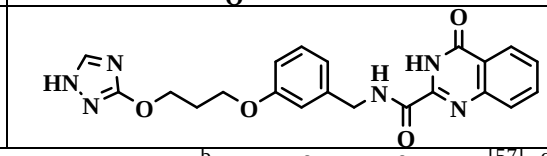
2.1 Biological actions and theoretical molecular descriptors

The reported seventeen pyrimidin-4-one-2-carboxamide derivatives possessing a non-carboxylate zinc-binding function are considered as the data set for present study.[57] These derivatives were evaluated for their MMP-13 inhibitory activities and were reported as IC_{50}. The reported MMP-13 activity on molar basis (as pIC_{50}) along with the structures of these analogues is shown in Table 1. 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-13 inhibitory activities of fused pyrimidine derivatives.

Cpd.	Structure	pIC_{50}[a]			
		Obsd.[b]	Calculated		
			Eq. (1)	Eq. (2)	PLS
1		9.77	9.56	10.12	9.95
2		11.46	11.21	10.72	10.95

3		9.62	9.67	9.79	9.62
4		10.28	10.53	10.32	10.30
5		10.29	10.04	10.61	10.13
6		9.52	9.51	9.18	9.55
7		9.82	9.75	10.35	10.14
8[c]		10.15	9.78	10.14	9.75
9		9.72	9.34	9.42	9.67
10[c]		9.96	9.66	9.96	9.65
11		9.53	10.01	9.63	9.48

12		8.82	9.31	9.07	8.62
13[c]		8.82	8.81	9.39	9.15
14		9.64	9.68	9.68	10.04
15		7.92	7.90	8.01	8.11
16		11.41	11.42	11.24	11.53
17		9.23	9.09	8.90	8.93

[a]IC_{50} on molar basis; [b]Taken from reference[57]; [c]Compound included in test set.

The structures of the all the compounds (listed in Table 1) were drawn in 2D ChemDraw[58] 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.[59] 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-13 inhibitory activity of fused pyrimidine derivatives.

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.[59]

A total number of 509 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 509 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)[60-64] and partial least squares (PLS)[65-67] 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.[68,69] 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[70,71] 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[72,73], the Kubinyi function, FIT[74,75] and the Friedman's lack of fit, LOF[76], 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.[77] 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 ±*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, 3 compounds have been included in the test set for the validation of the models derived from 14 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 one, two and three descriptors have been derived to achieve the best relationship correlating MMP-13 inhibitory activity. No any one model in one and two descriptors having $r^2_{Test} > 0.5$, were obtained through CP-MLR. The analysis resulted 13 such three parameter models which have shared 19 descriptors among them. All these 19 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**):	Me (mean atomic Sanderson electronegativity scaled on Carbon atom), 2.301(1); Ms (mean electrotopological state), 1.405(2); nO (number of Oxygen atpms), 1.301(2)
Topological descriptors (**TOPO**):	MAXDP (maximal electrotopological positive variation), 2.434(2); IC1 (information content index, neighborhood symmetry of 1-order), 1.072(1); IC2 (information content index, neighborhood symmetry of 2-order), 2.931(1); LP1 (Lovasz-Pelikan index (leading eigenvalue), 1.881(8);
Molecular Walk Counts (**MWC**):	MWC09 (molecular walk count of order 09), 1.671(1)
2D autocorrelations (**2D-AUTO**):	MATS2m (Moran autocorrelation of lag-2/ weighted by atomic masses), 1.424(1); MATS8m (Moran autocorrelation of lag-8/ weighted by atomic masses), -2.500(1); GATS3e (Geary autocorrelation of lag-3/ weighted by atomic Sanderson electronegativity), -1.735(1); GATS4e (Geary autocorrelation of lag-4/ weighted by atomic Sanderson electronegativity), -1.450(1); GATS7e (Geary autocorrelation of lag-7/ weighted by atomic Sanderson electronegativity), -2.118(11)
Atom centered fragments (**ACF**)	C-040 (R-C(=X)-X/R-C#X/X-=C=X), 1.579(1); O-056 (alcohol), 0.670(1)
Empirical (**EMP**)	Hy (hydrophilic factor), 1.591(3)
Functional groups (**FUN**)	nNHOH (number of aliphatic hydroxylamines), 1.370(1); nSO2N (number of sufonamides), 0.755(1)

[a]The descriptors are identified from the three parameter models for activity emerged from CP-MLR protocol with filter-1 as 0.30, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 14 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 representative highly significant models in three descriptors are presented below:

$pIC_{50} = 9.979 + 2.264(0.289)LP1 – 2.560(0.333)GATS3e + 1.370(0.229)nNHOH$

$n = 14, r = 0.958, s = 0.296, F = 37.819, Q^2_{LOO} = 0.854, Q^2_{L3O} = 0.768$

$r^2_{Test} = 0.797, FIT = 4.932, LOF = 0.192, AIC = 0.158$ (1)

$pIC_{50} = 9.647 + 2.931(0.449)IC2 – 1.735(0.388)GATS3e – 1.450(0.348)GATS4e$

$n = 14, r = 0.928, s = 0.386, F = 20.877, Q^2_{LOO} = 0.729, Q^2_{L3O} = 0.723$

$r^2_{Test} = 0.708, FIT = 2.723, LOF = 0.326, AIC = 0.268$ (2)

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 LP1 and IC2 participated in above models are the topological descriptors representing Lovasz-Pelikan index (leading eigenvalue) and information content index of 2^{nd} order neighborhood symmetry, respectively. The positive influence of descriptors LP1 and IC2 on the activity suggested that higher values of descriptor LP1 and IC2 would be beneficiary to the activity. The descriptor GATS3e and GATS4e are 2D-autocorrelations. Both of these descriptors showed negative contribution to the activity. Thus, lower values of atomic Sanderson electronegativities weighted Geary autocorrelations of lag-3 and -4 would be favorable to the activity. Additionally, positive sign of functional group class descriptor nNHOH representing number of aliphatic hydroxylamines advocated that more number of aliphatic hydroxylamine functionality in a molecular structure would be supportive to elevated activity.

These models have accounted for nearly 92% 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. (1) and (2) have been included in Table 1. The models (1) and (2) are validated with an external test set of 3 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.

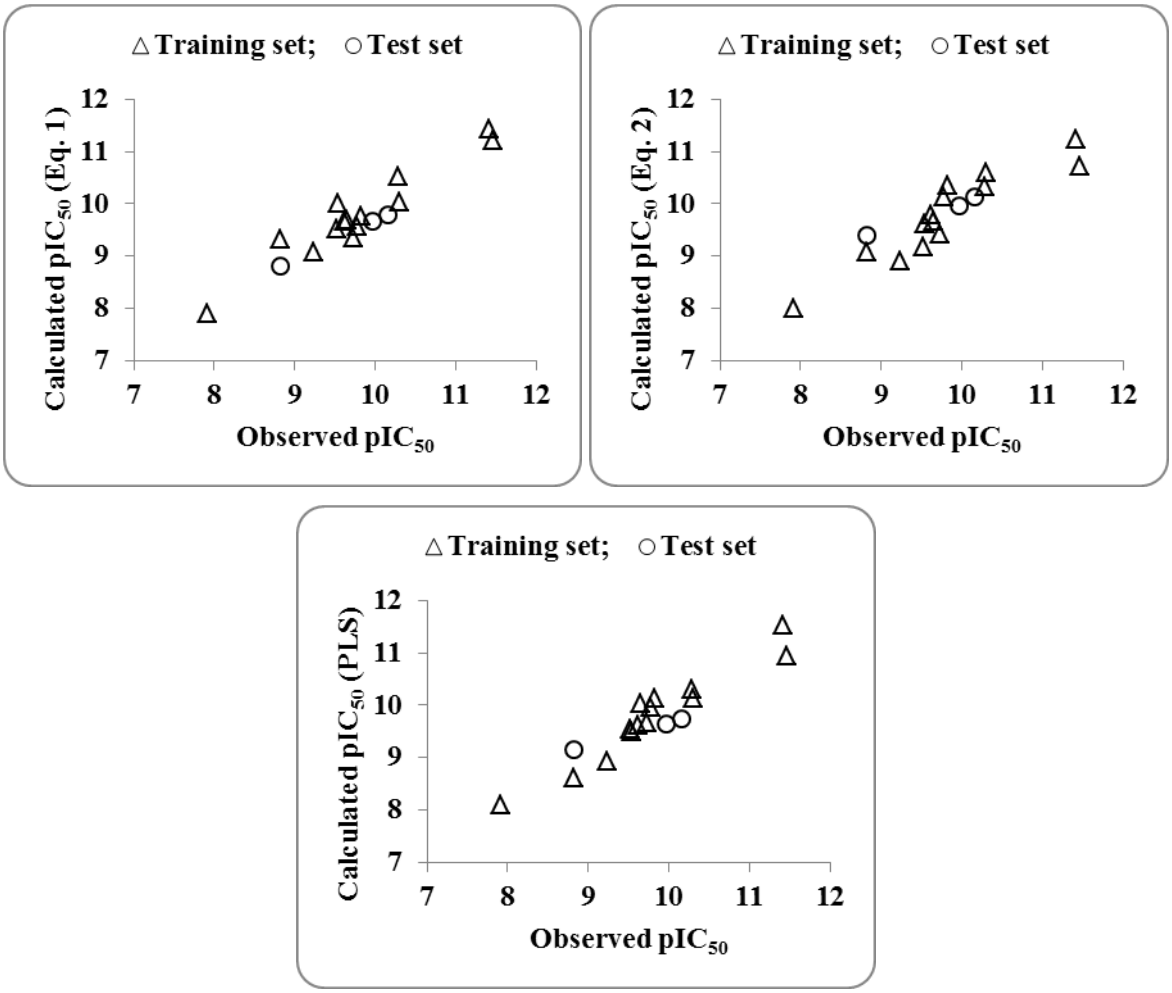


Figure 1: Plot of observed and calculated pIC_{50} values of training- and test-set compounds for MMP-13 inhibition.

A partial least square (PLS) analysis has been carried out on these 19 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 19 descriptors and they explained 92.73% variance in the activity. The MLR-like PLS coefficients of these 19 descriptors are given in Table 4.

Table 4: PLS and MLR-like PLS models from the 19 descriptors of three parameter CP-MLR models for MMP-13 inhibitory activities.

A: PLS equation									
PLS components						PLS coefficient (s.e.)[a]			
Component-1						0.356(0.031)			
Component-2						-0.177(0.043)			
Constant						9.787			
B: MLR-like PLS equation									
S. No.	**Descriptor**	**MLR-like coefficient[b]**	**(f.c.)[c]**	**Order**		**S. No.**	**Descriptor**	**MLR-like coefficient[b]**	**(f.c.)[c]**	**Order**
1	Me	0.149	0.073	4		11	MATS5e	-0.050	-0.025	16
2	Ms	0.105	0.052	11		12	GATS3e	-0.134	-0.066	6
3	nO	0.067	0.033	14		13	GATS4e	-0.168	-0.083	3
4	MAXDP	0.202	0.100	1		14	GATS7e	-0.174	-0.085	2
5	IC1	0.109	0.053	10		15	nNHOH	0.146	0.072	5
6	IC2	0.095	0.047	13		16	nSO2N	0.040	0.020	18
7	LP1	0.100	0.049	12		17	C-040	0.112	0.055	8
8	MWC09	0.112	0.055	9		18	O-056	0.050	0.025	17
9	MATS2m	0.127	0.063	7		19	Hy	0.065	0.032	15
10	MATS8m	-0.028	-0.014	19			Constant = 8.602			
C: PLS regression statistics						Values				
n						14				
r						0.963				
s						0.264				
F						72.229				
FIT						8.025				
LOF						0.107				
AIC						0.107				
Q^2_{LOO}						0.889				
Q^2_{L3O}						0.907				
r^2_{Test}						0.667				

[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.

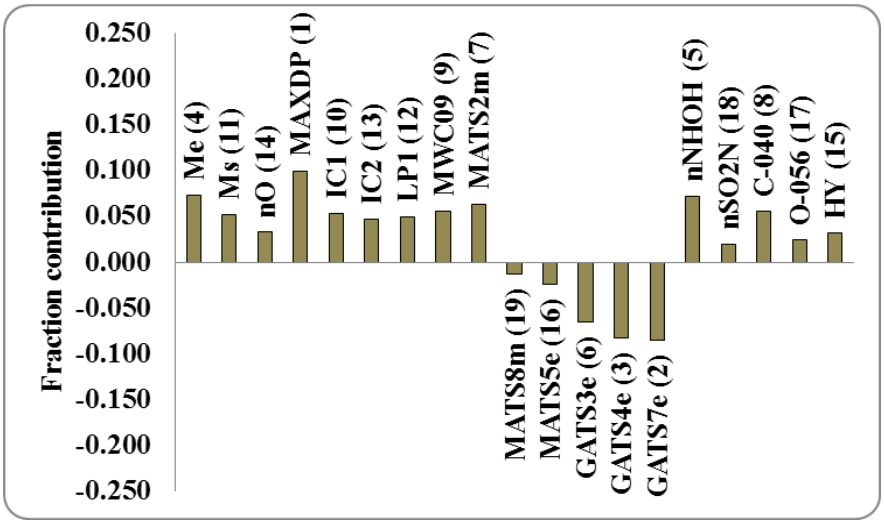


Figure 2: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 19 CP-MLR identified descriptors (Table 3) associated with MMP-13 inhibitory activity of fused pyrimidine derivatives.

The PLS analysis has suggested MAXDP as the most determining descriptor for modeling the activity of the compounds (descriptor S. No. 4 in Table 4; Figure 2). The other nine descriptors in decreasing order of significance are GATS7e, GATS4e, Me, nNHOH, GATS3e, MATS2m, C-040, MWC09 and IC1. Descriptor GATS3e, GATS4e and nNHOH are part of Eqs. (1) and (2) and convey same inference in the PLS model as well.

It is inferred from the PLS analysis that a higher values of descriptors MAXDP (maximal electrotopological positive variation), Me (mean atomic Sanderson electronegativity scaled on Carbon atom), MATS2m (Moran autocorrelation of lag-2/ weighted by atomic masses), MWC09 (molecular walk count of order 09) and IC1 (information content index, neighborhood symmetry of 1-order) in addition to a lower value of descriptor GATS7e (Geary autocorrelation of lag-7/ weighted by atomic Sanderson electronegativity) and presence of R-C(=X)-X/R-C#X/X-=C=X type atom centred fragment (descriptor C-040) in a molecular structure would be advantageous to the activity. It is also observed that PLS model from the dataset devoid of CP-MLR identified 19 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-13 inhibitory activity if the limit of normal values for the *Y*

outliers (response outliers) was set as 2 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*=0.706). 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 = 17) in descriptors identified through CP-MLR.

Model	r	s	F	Q^2_{LOO}	Eq.
pIC_{50} = 10.057 + 2.236(0.270)LP1 – 2.656(0.304)GATS3e + 1.316(0.216)nNHOH	0.954	0.285	44.607	0.858	(1a)
pIC_{50} = 9.572 + 3.032(0.422)IC2 – 1.770(0.362)GATS3e – 1.438(0.252)GATS4e	0.922	0.370	24.723	0.738	(2a)

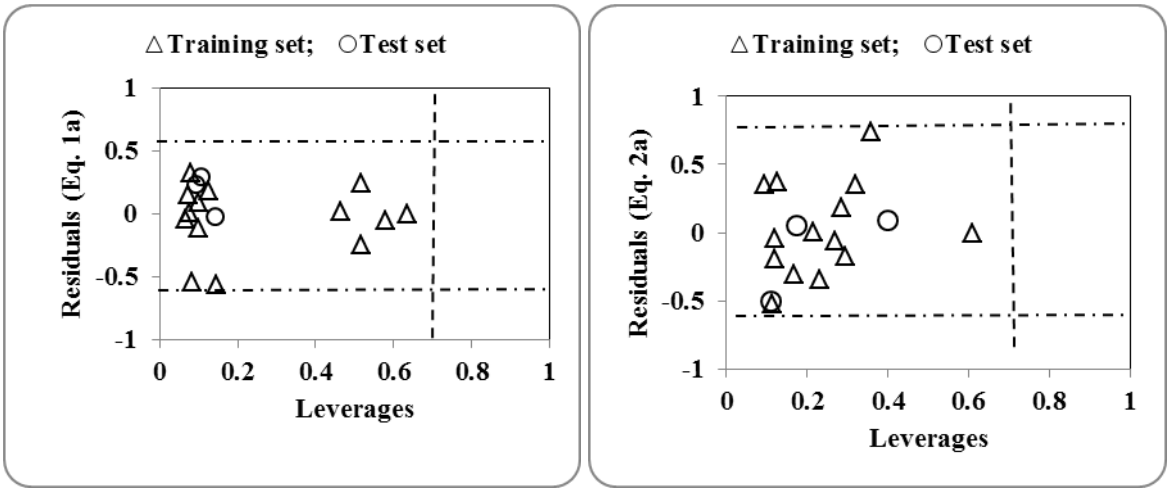


Figure 3: Williams plot for the training-set and test- set compounds for MMP-13 inhibitory activity. The horizontal dotted line refers to the residual limit (±2×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.706).

CONCLUSIONS

QSAR study has been carried out on the MMP-13 inhibitory activity of fused pyrimidine derivatives possessing a non-carboxylate zinc-binding function in 0D- to 2D-Dragon descriptors. The derived QSAR models have revealed that the information content indices of 1st and 2nd order neighborhood symmetry (descriptors IC1 and IC2), Lovasz-Pelikan index (descriptor LP1), maximal electrotopological positive variation (descriptor MAXDP) and molecular walk count of order 09 (descriptor MWC09) played a pivotal role in rationalization

of MMP-13 inhibition activity of titled compounds. Atomic properties such as mass and atomic Sanderson electronegativity in terms of atomic properties weighted descriptors MATS2m, GATS3e, GATS4e, GATS7e and Me, certain atom centred fragments such as R-C(=X)-X/R-C#X/X-=C=X (descriptor C-040) and aliphatic hydroxylamine functionality (descriptor nNHOH) are also predominant to explain MMP-13 inhibition actions of fused pyrimidines.

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. Badley EM and Tennant A. Disablement associated with rheumatic disorders in a British population: problems with activities of daily living and level of support. British Journal of Rheumatology, 1993; 32: 601-608.
2. Buckwalter JA and Martin JA. Osteoarthritis. Advanced Drug Delivery Reviews, 2006; 58: 150-167.
3. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, Liang MH, Kremers HM, Mayes MD, Merkel PA, Pillemer SR, Reveille JD, Stone JH. National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis and Rheumatology, 2008; 58: 15-25.
4. Miura H and Iwamoto Y. Bone, 2000; 14: 287.
5. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage, 2004; 12: S39-S44.

6. Ikegawa S. New gene associations in osteoarthritis: what do they provide, and where are we going? Current Opinion in Rheumatoloy, 2007; 19: 429-434.
7. Miyamoto Y, Shi D, Nakajima M, Ozaki K, Sudo A, Kotani A, Uchida A, Tanaka T, Fukui N, Tsunoda T, Takahashi A, Nakamura Y, Jiang Q, Ikegawa S. Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. Nature Genetics, 2008; 40: 994-998.
8. Mototani H, Iida A, Nakajima M, Furuichi T, Miyamoto Y, Tsunoda T, Sudo A, Kotani A, Uchida A, Ozaki K, Tanaka Y, Nakamura Y, Tanaka T, Notoya K, Ikegawa S. A functional SNP in EDG2 increases susceptibility to knee osteoarthritis in Japanese. Human Molecular Genetics, 2008; 15(17): 1790-1797.
9. FitzGerald GA. Coxibs and cardiovascular disease. New England Journal of Medicine, 2004; 351: 1709-11.
10. Sabatini M, Lesur C, Thomas M, Chomel A, Anract P, de Nanteuil G, Pastoureau P, Janusz MJ, Bendele AM, Brown KK, Taiwo YO, Hsieh L, Heitmeyer SA. Effect of inhibition of matrix metalloproteinases on cartilage loss in vitro and in a guinea pig model of osteoarthritis. Arthritis and Rheumatology, 2005; 52: 171-180.
11. Hoekstra R, Eskens FA, Verweij J. Matrix metalloproteinase inhibitors: current developments and future perspectives. Oncologist, 2001; 6: 415-427.
12. Peterson JT. The importance of estimating the therapeutic index in the development of matrix metalloproteinase inhibitors. Cardiovascular Research, 2006; 69: 677- 687.
13. 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 Vaughn-Bowser ML. Pyran containing sulfonamide hydroxamic acids: potent MMP inhibitors that spare MMP-1. Bioorganic and Medicinal Chemistry Letters, 2004; 14: 3389-3395.
14. Mitchell PG, Magna HA, Reeves LM, Lopresti-Morrow LL, Yocum SA, Rosner PJ, Geoghegan KF, Hambor JE. Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. Journal of Clinical Investigations, 1996; 97: 761-768.
15. Knäuper V, López-Otin C, Smith B, Knight G, Murphy G. Biochemical characterization of human collagenase-3. Journal of Biological Chemistry, 1996; 271: 1544-1550.

16. Yocum SA, Lopresti-Morrow LL, Reeves LM, Mitchell PG. MMP-13 and MMP-1 expression in tissues of normal articular joints. Annals of the New York Academy of Sciences, 1999; 878: 583-586.
17. Engel CK, Pirard B, Schimanski S, Kirsch R, Habermann J, Klingler O, Schlotte V, Weithmann KU, Wendt KU. Structural basis for the highly selective inhibition of MMP-13. Chemical Biology, 2005; 12: 181-189.
18. Li JJ, Nahra J, Johnson AR, Bunker A, O'Brien P, Yue WS, Ortwine DF, Man CF, Baragi V, Kilgore K, Dyer RD, Han HK. Quinazolinones and pyrido[3,4-d]pyrimidin-4-ones as orally active and specific matrix metalloproteinase-13 inhibitors for the treatment of osteoarthritis. Journal of Medicinal Chemistry, 2008; 28: 835-841.
19. Kim SH, Pudzianowski AT, Leavitt KJ, Barbosa J, McDonnell PA, Metzler WJ, Rankin BM, Liu R, Vaccaro W, Pitts W. Structure-based design of potent and selective inhibitors of collagenase-3 (MMP-13). Bioorganic and Medicinal Chemistry Letters, 2005; 15: 1101-1106.
20. Blagg JA, Noe MC, Wolf-Gouveia LA, Reiter LA, Laird ER, Chang SP, Danley DE, Downs JT, Elliott NC, Eskra JD, Griffiths RJ, Hardink JR, Haugeto AI, Jones CS, Liras JL, Lopresti-Morrow LL, Mitchell PG, Pandit J, Robinson RP, Subramanyam C, Vaughn-Bowser ML, Yocum SA. Potent pyrimidinetrione-based inhibitors of MMP-13 with enhanced selectivity over MMP-14. Bioorganic and Medicinal ChemistryLetters, 2005; 15: 1807-1810.
21. Aranapakam V, Grosu GT, Davis JM, Hu B, Ellingboe J, Baker JL, Skotnicki JS, Zask A, DiJoseph JF, Sung A, Sharr MA, Killar LM, Walter T, Jin G, Cowling R. Synthesis and structure-activity relationship of alpha-sulfonylhydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. Journal of medicinal Chemistry, 2003; 46: 2361-2375.
22. Aranapakam V, Davis JM, Grosu GT, Baker J, Ellingboe J, Zask A, Levin JI, Du M, Sandanayaka VP, Skotnicki JS, DiJoseph JF, Sung A, Sharr MA, Killar LM, Walter T, Jin G, Cowling R, Tillett J, Zhao W, McDevitt J, 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.
23. Wu J, Rush TS 3rd, Hotchandani R, Du X, Geck M, Collins E, Xu ZB, Skotnicki J, Levin JI, Lovering FE. Identification of potent and selective MMP-13 inhibitors. Bioorganic and Medicinal Chemistry Letters, 2005; 15: 4105-4109.

24. Li J, Rush TS 3rd, Li W, DeVincentis D, Du X, Hu Y, Thomason JR, Xiang JS, Tam S, Skotnicki JS, Cunningham KM, Chockalingam PS, Morris EA, Levin JI. Synthesis and SAR of highly selective MMP-13 inhibitors. Bioorganic and Medicinal Chemistry Letters, 2005; 15: 4961-4966.
25. 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, Yocum SA. Potent, selective pyrimidinetrione-based inhibitors of MMP-13. Bioorganic and Medical Chemistry Letters, 2006; 16: 5822-5826.
26. Freeman-Cook KD, Reiter LA, Noe MC, Antipas AS, Danley DE, Datta K, Downs JT, Eisenbeis S, Eskra JD, Garmene DJ, Greer EM, Griffiths RJ, Guzman R, Hardink JR, Janat F, Jones CS, Martinelli GJ, Mitchell PG, Laird ER, Liras JL, Lopresti-Morrow LL, Pandit J, Reilly UD, Robertson D, Vaughn-Bowser ML, Wolf-Gouviea LA, Yocum SA. Potent, selective spiropyrrolidine pyrimidinetrione inhibitors of MMP-13. Bioorganic and Medicinal Chemistry Letters, 2007; 17: 6529-6534.
27. 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, Levin JI. Potent, selective, and orally bioavailable matrix metalloproteinase-13 inhibitors for the treatment of osteoarthritis. Bioorganic and Medicinal Chemistry, 2005; 13: 6629-6644.
28. Chen JM, Nelson FC, Levin JI, Mobilio D, Moy FJ, Nilakantan R, Zask A, Powers R. Structure-based design of a novel, potent, and selective inhibitor for MMP-13 utilizing NMR spectroscopy and computer-aided molecular design. Journal of American Chemical Society, 2000; 122: 9648-9654.
29. Johnson AR, Pavlovsky AG, Ortwine DF, Prior F, Man CF, Bornemeier DA, Banotai Mueller WT, McConnell P, Yan C, Baragi V, Lesch C, Roark WH, Wilson M, Datta K, Guzman R, Han HK, Dyer RD. 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, 2007; 282: 27781-27791.
30. Vandenbroucke RE, Libert C. Is there new hope for therapeutic matrix metalloproteinase inhibition? Nature Reviews Drug Discovery, 2014; 13: 904-927.
31. Cathcart JM, Cao J. MMP Inhibitors: Past, present and future. Frontiers in Biosciences (Landmark Ed), 2015; 20: 1164-1178.

32. Wang K, Xu J, Hunter DJ, Ding C. Investigational drugs for the treatment of osteoarthritis. Expert Opinion on Investigational Drugs, 2015; 24: 1539-1556.
33. Nara H, Sato K, Naito T, Mototani H, Oki H, Yamamoto Y, Kuno H, Santou T, Kanzaki N, Terauchi J, Uchikawa O, Kori M. 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, 2014; 57: 8886-8902.
34. Nara H, Sato K, Naito T, Mototani H, Oki H, Yamamoto Y, Kuno H, Santou T, Kanzaki N, Terauchi J, Uchikawa O, Kori M. Thieno[2,3-d]pyrimidine-2-carboxamides bearing a carboxybenzene group at 5-position: highly potent, selective, and orally available MMP-13 inhibitors interacting with the S1″ binding site. Bioorganic and Medicinal Chemistry, 2014; 22: 5487-5505.
35. Nara H, Kaieda A, Sato K, Naito T, Mototani H, Oki H, Yamamoto Y, Kuno H, Santou T, Kanzaki N, Terauchi J, Uchikawa O, Kori, M. Discovery of novel, highly potent, and selective matrix metalloproteinase (MMP)-13 Inhibitors with a 1,2,4-Triazol- 3-yl Moiety as a zinc binding group using a structure-based design approach. Journal of Medicinal Chemistry, 2017; 60: 608-626.
36. Terauchi J, Kuno H, Nara H, Oki H and Sato K. (2005). Preparation of heterocyclic amides as MMP-13 inhibitors for treating osteoarthritis and rheumatoid arthritis. WO2005105760A1.
37. Nara H, Kaieda A, Sato K and Terauchi J. (2007). Preparation of heterocyclic amide compounds as matrix metalloproteinase inhibitors. WO2007049820A1.
38. Schnute ME, Ruminski PG, Hanau CE, Strohbach JW, Carroll JN and Sample K. (2008). Hetero bicyclic carboxamide derivatives, processes for preparing them, pharmaceutical compositions containing them, and their uses as inhibitors of matrix metalloproteinase enzymes. WO2008149191A1.
39. Pochetti G, Montanari R, Gege C, Chevrier C, Taveras AG and Mazza F. Extra binding region induced by non-zinc chelating inhibitors into the S1′ subsite of matrix metalloproteinase 8 (MMP-8). Journal of Medicinal Chemistry, 2009; 52: 1040-1049.
40. Gege C, Chevrier C, Schneider M, Bluhm H, Hochguertel M, Deng H, Gallagher BM, Sucholeiki I and Taveras AG. Preparation of heterotricyclic metalloprotease inhibitors. WO2008063667A1, 2008.
41. Breuer E, Frant J, Reich R. Recent non-hydroxamate matrix metalloproteinase inhibitors. Expert Opinion on Therapeutic Patents, 2005; 15: 253-269.

42. Rao BG. Recent developments in the design of specific Matrix Metalloproteinase inhibitors aided by structural and computational studies. Curr Pharm Des., 2005; 11(3): 295-322.
43. Skiles JW, Gonnella NC, Jeng AY. The design, structure, and therapeutic application of matrix metalloproteinase inhibitors. Curr Med Chem., 2001; 8(4): 425-74.
44. Puerta DT, Cohen SM. A bioinorganic perspective on matrix metalloproteinase inhibition. Current Topics in Medicinal Chemistry, 2004; 4: 1551-1573.
45. Whittaker M, Floyd CD, Brown P, Gearing AJ. Design and therapeutic application of matrix metalloproteinase inhibitors. Chemical Reviews, 1999; 99(9): 2735-76.
46. Skiles JW, Gonnella NC, Jeng AY. The design, structure, and clinical update of small molecular weight matrix metalloproteinase inhibitors. Current Medicinal Chemistry, 2004; 11(22): 2911-77.
47. Fischer T, Riedl R. Molecular Recognition of the Catalytic Zinc(II) Ion in MMP-13: Structure-Based Evolution of an Allosteric Inhibitor to Dual Binding Mode Inhibitors with Improved Lipophilic Ligand Efficiencies. Int J Mol Sci., 2016; 17(3): 314.
48. O'Brien PM, Ortwine DF, Pavlovsky AG, Picard JA, Sliskovic DR, Roth BD, Dyer RD, Johnson LL, Man CF, Hallak H. Structure-activity relationships and pharmacokinetic analysis for a series of potent, systemically available biphenylsulfonamide matrix metalloproteinase inhibitors. J Med Chem., 2000; 43(2): 156-66.
49. Natchus MG, Bookland RG, Laufersweiler MJ, Pikul S, Almstead NG, De B, Janusz MJ, Hsieh LC, Gu F, Pokross ME, Patel VS, Garver SM, Peng SX, Branch TM, King SL, Baker TR, Foltz DJ, Mieling GE. Development of new carboxylic acid-based MMP inhibitors derived from functionalized propargylglycines. Journal of Medicinal Chemistry, 2001; 44(7): 1060-71.
50. Pikul S, Ohler NE, Ciszewski G, Laufersweiler MC, Almstead NG, De B, Natchus MG, Hsieh LC, Janusz MJ, Peng SX, Branch TM, King SL, Taiwo YO, Mieling GE. Potent and selective carboxylic acid-based inhibitors of matrix metalloproteinases. Journal of Medicinal Chemistry, 2001; 44(16): 2499-502.
51. Natchus MG, Bookland RG, De B, Almstead NG, Pikul S, Janusz MJ, Heitmeyer SA, Hookfin EB, Hsieh LC, Dowty ME, Dietsch CR, Patel VS, Garver SM, Gu F, Pokross ME, Mieling GE, Baker TR, Foltz DJ, Peng SX, Bornes DM, Strojnowski MJ, Taiwo YO. Development of new hydroxamate matrix metalloproteinase inhibitors derived from functionalized 4-aminoprolines. Journal of Medicinal Chemistry, 2000; 43(26): 4948-63.

52. Uminski PG, Massa M, Strohbach J, Hanau CE, Schmidt M, Scholten JA, Fletcher TR, Hamper BC, Carroll JN, Shieh HS, Caspers N, Collins B, Grapperhaus M, Palmquist KE, Collins J, Baldus JE, Hitchcock J, Kleine HP, Rogers MD, McDonald J, Munie GE, Messing DM, Portolan S, Whiteley LO, Sunyer T, Schnute ME. (2016), Discovery of N-(4-Fluoro-3-methoxybenzyl)-6-(2-(((2S,5R)-5-(hydroxymethyl)-1,4-dioxan-2-yl)methyl)-2H-tetrazol-5-yl)-2-methylpyrimidine-4-carboxamide. A Highly Selective and Orally Bioavailable Matrix Metalloproteinase-13 Inhibitor for the Potential Treatment of Osteoarthritis. Journal of Medicinal Chemistry, 59(1):313-327.
53. Gege C, Bao B, Bluhm H, Boer J, Gallagher BM, Korniski B, Powers TS, Steeneck C, Taveras AG, Baragi VM. (2012). Discovery and evaluation of a non-Znchelating, Selective matrix metalloproteinase 13 (MMP-13) inhibitor for potential intra-articular treatment of osteoarthritis. Journal of Medicinal Chemistry, 55(2):709-16.
54. Janusz MJ, Hookfin EB, Brown KK, Hsieh LC, Heitmeyer SA, Taiwo YO, Natchus MG, Pikul S, Almstead NG, De B, Peng SX, Baker TR, Patel V. Comparison of the pharmacology of hydroxamate- and carboxylate-based matrix metalloproteinase inhibitors (MMPIs) for the treatment of osteoarthritis. Inflamm Res., 2006; 55(2): 60-65.
55. Kreiselmeier A, Ulmer W, Stöve J, Wieland HA, Gerwin N, Bartnik E, Schudok M, Heute S, Breitenfelder M, Scharf HP, Schwarz ML. Validation of a diffusion chamber as in vitro system for the analysis of compound diffusibility through cartilagetissue. Biomed Pharmacother., 2005; 59(7): 395-401.
56. Peng SX, VonBargen EC, Bornes DM, Pikul S. Permeability of articular cartilage to matrix metalloprotease inhibitors. Pharm Res., 1998; 15(9): 1414-1418.
57. Nara H, Sato K, Kaieda A, Oki H, Kuno H, Santou T, Kanzaki N, Terauchi J, Uchikawa O, Kori M. Design, synthesis, and biological activity of novel, potent, and highly selective fused pyrimidine-2-carboxamide-4-one-based matrix metalloproteinase (MMP)-13 zinc-binding inhibitors Bioorganic & Medicinal Chemistry, 2016; 24: 6149-6165.
58. Chemdraw ultra 6.0 and Chem3D ultra, Cambridge Soft Corporation, Cambridge, USA.
59. Dragon software (version 1.11-2001) by Todeschini R.; Consonni V. Milano, Italy.
60. Prabhakar, Y. S. 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. Sci., 2003; 22: 583-595.
61. Sharma, S.; Prabhakar, Y. S.; Singh, P.; Sharma, B. K. QSAR study about ATP-sensitive potassium channel activation of cromakalim analogues using CP-MLR approach, Eur J Med Chem., 2008; 43: 2354-2360.

62. Sharma, S.; Sharma, B. K.; Sharma, S. K.; Singh, P.; Prabhakar, Y. S. Topological Descriptors in modeling the agonistic activity of human A3 adenosine receptor ligands: The derivatives of 2-Chloro-N6-substituted-4'-thioadenosine-5'-uronamide, Eur. J. Med. Chem., 2009; 44: 1377-1382.
63. Sharma, B. K.; Pilania, P.; Singh, P.; Prabhakar, Y. S. 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 QSAR Environ. Res., 2010; 21: 169-185.
64. Sharma, B. K.; Singh, P.; Sarbhai, K.; Prabhakar, Y. S. A quantitative structure-activity relationship study on serotonin 5-HT6 receptor ligands: Indolyl and piperidinyl sulphonamides, SAR QSAR Environ. Res., 2010; 21: 369-388.
65. Wold S. Cross-validatory estimation of the number of components in factor and principal components models. Technometrics, 1978; 20: 397–405.
66. Kettaneh N, Berglund A, Wold S. PCA and PLS with very large data sets. Comput Stat Data Anal, 2005; 48: 69–85.
67. Stahle L, Wold S. Multivariate data analysis and experimental design. In: Ellis GP, West WB. Eds., Biomedical research. Progress in medicinal chemistry. Elsevier Science Publishers, BV, Amsterdam, 1988; 25: 291–338.
68. Topliss JG, Edwards RP. Chance factors in studies of quantitative structure–activity relationships. J Med Chem., 1979; 22: 1238–1244.
69. Katritzky AR, Dobchev DA, Slavov S, Karelson M. Legitimate utilization of large descriptor pools for QSAR/QSPR models. J Chem Inf Model., 2008; 48: 2207–2213.
70. So S-S, Karplus M. Three-dimensional quantitative structure–activity relationship from molecular similarity matrices and genetic neural networks. 1. Method and validation. J. Med Chem., 1997; 40: 4347–4359.
71. Prabhakar YS, Solomon VR, Rawal RK, Gupta MK, Katti SB. CP-MLR/PLS directed structure–activity modeling of the HIV-1 RT inhibitory activity of 2,3-diaryl-1,3-thiazolidin-4-ones. QSAR Comb Sci., 2004; 23: 234–244.
72. Akaike H. Information theory and an extension of the minimum likelihood principle. In: Petrov BN, Csaki F, editors. Second international symposium on information theory. Budapest: Akademiai Kiado, 1973; 267–281.
73. Akaike H. A new look at the statistical identification model. IEEE Trans Autom Control, 1974; AC-19: 716–723.

74. Kubinyi H. Variable selection in QSAR studies. I. An evolutionary algorithm. Quant Struct-Act Relat, 1994; 13: 285–294.
75. Kubinyi H. Variable selection in QSAR studies. II. A highly efficient combination of systematic search and evolution. Quant Struct-Act Relat, 1994; 13: 393–401.
76. Friedman J. In Technical Report No. 102. Laboratory for Computational Statistics, Stanford University: Stanford, 1990.
77. Gramatica, P. Principles of QSAR models validation: internal and external. QSAR Comb. Sci., 2007; 26: 694-701.