Journal of Current Chemical & Pharmaceutical Sciences

*J. Curr. Chem. Pharm. Sc.: 2(3), 2012, 161-178*
*ISSN 2277-2871*

# QSAR RATIONALE OF MATRIX METALLOPROTEINASE INHIBITION ACTIVITY IN A CLASS OF SELECTIVE α-SULFONE HYDROXAMATES

**P. SINGH***

Emeritus Fellow, University Grants Commission, Department of Chemistry, S.K. Government Post Graduate College, SIKAR – 332001 (Raj.) INDIA

(*Received : 29.05.2012; Accepted : 06.06.2012*)

## ABSTRACT

The matrix metalloproteinase (MMP) inhibition activity in a class of selective α-sulfone hydroxamates has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the MMP-2 and MMP-13 inhibition activities of these compounds. The descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of different atomic properties such as masses, polarizabilities, electronegativities and van der Waals volumes as the weighting components. Additionally, the structural information content, the topological charges, the number of total secondary and tertiary C ($sp^3$) atoms and the functionalities R--CR--X and R--CR--X have also shown their importance for inhibition of the MMP. PLS analysis has further corroborated the dominance of the CP-MLR identified descriptors.

**Key words**: Isonipecotamide α-tetrahydropyranyl and α-piperidine sulfone hydroxamates, MMP inhibitors, Combinatorial protocol in multiple linear regression (CP-MLR) analysis, Chemometric descriptors, QSAR.

## INTRODUCTION

Matrix metalloproteinases (MMPs), the zinc-dependent enzymes, are responsible for remodeling and degradation of all components of the extracellular matrix[1,2], but excessive activity of MMPs has been associated in cancer[3,4], arthritis[5] and cardiovascular disease[6–9]. MMP inhibitors (MMPIs) have therefore been identified as therapeutic agents to apprehend the development of such disease states[10-12]. The MMP family of enzymes incorporates at least 24 dissimilar mammalian isozymes, but MMP-13 has been recognized as a important target which is involved in cancer, osteoarthritis, and cardiovascular disease.

Treatment of patients with broad-spectrum MMPIs results into stiffening of the joints known as musculoskeletal syndrome[9] (MSS). MMP-1 has long been thought to be a cause whose inhibition plays a role in MSS. In addition, MMP-14 knockout mice suffer connective tissue disease due to inadequate collagen turnover[13] and impaired endochondral ossification[14] reminiscent of joint lesions in MSS. Efforts have therefore been made earlier on potential inhibitors of MMP-13 while sparing other MMPs to achieve joint safety[15-28].

In order to enhance MMP-13 selectivity, sparing MMP-1 and MMP-14, the P′ region of the MMPI's interacting with $S'_1$ pocket of the enzyme continues the research field of interest. Following the dual-sparing

Available online at www.sadgurupublications.com
*Author for correspondence; E-mail: psingh_sikar@rediffmail.com

hypothesis, Kolodziej et al.[29,30] have explored two series of hyroxamates as the novel selective inhibitors of MMP-13. The first series[29] consists of isonipecotamide α-tetrahydropyranyl and α-piperidine sulfone hydroxamates (Fig. 1A), while the second series[30] comprise of N-aryl piperazine α-sulfones and N-aryl piperazine α-tetrahydropyranyl sulfone hydroxamates (Fig. 1B). Interaction of the amide N-substituents (first series) or the distal phenyl group (second series) deep in the S′1 pocket was expected to affect isozyme selectivity across the MMP family. These studies have therefore afforded in learning about P′1 manipulations toward optimizing MMP-13 selectivity.

In both reported series, the structure-activity relationship (SAR) studies were, however, targeted at the alterations of substituents at different positions and provided no rationale to reduce the trial-and-error factors. Hence, the present study is aimed at to establish the quantitative relationships between experimental activities and 0D-2D chemometric descriptors which may address the molecular structures of the compounds. Such a 2D-QSAR may provide the rationale for drug-design and help to explore the possible mechanism of action at the molecular level. In a 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.

## EXPERIMENTAL

### Materials and methods

In the present work, the hyroxamates represented by general structures in Fig. 1(A) and 1(B) and their inhibition activities ($IC_{50}$s) towards MMP-2 and MMP-13 have been taken from the literature[29,30]. The activity, $IC_{50}$, represents the concentration of a compound to accomplish 50% inhibition of MMP-2 (gelatinase) and MMP-13 (collagen). The same is expressed as $-\log IC_{50}$, on a molar basis, and stand as the dependent variable for present study. The compounds along with their $-\log IC_{50}$ values for inhibition of MMP-2 and MMP-13 are documented in Table 1. MMP-2 is highly involved in the process of tumor invasion and metastasis and has been considered as a promising target for cancer therapy while inhibitors of MMP-13 can offer protection from the cartilage degradation associated with osteoarthritis.

For modeling purpose the compounds from both series, listed in Table 1, have been numbered consecutively and considered as one data-set so that it includes diverse structural features of the compounds and exhibits sufficient activity variation. This data-set was next divided into training- and test-sets to insure external validation of derived models. Nearly 25% of the total compounds have been selected for the test-set. The selection has been made through SYSTAT[31] using the single linkage hierarchical cluster procedure involving the Euclidean distances of the inhibition activity, -log $IC_{50}$ values. The compounds were then selected from the generated cluster tree in such a way to keep them at a maximum possible distance from each other. The statistical index, $r^2_{Test}$, representing the squared correlation coefficient between the observed and predicted activity values of compounds from test-set has been 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 0.500-0.600 are regarded as indicative predictive models. For simplicity, a test-set common both to MMP-2 and MMP-13 inhibition activities was preferred to impress upon similar structural features which are able to address two different activities.

### Molecular descriptors

The structures of the compounds (Table 1) under study have been drawn in 2D ChemDraw[32] using the standard procedure. These structures 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 series under investigation. Also, the compounds attain a common level of minimum energy; a condition suitable for interaction with receptor site(s). All these energy minimized structures of individual compounds have been ported to DRAGON software[33] for computing the descriptors corresponding to 0D-, 1D- and 2D-classes. Table 2 provides the definition and scope of these descriptor-classes in addressing the structural features which were employed in present work. The combinatorial protocol in multiple linear regression (CP-MLR) computational procedure[34] has been used in developing QSAR models. Prior to the application of the CP-MLR procedure, all those descriptors which are inter-correlated beyond 0.90 and showing a correlation of less than 0.1 with the biological endpoints (descriptor vs. activity, $r < 0.1$) were excluded. The remaining descriptors, able to address the biological activity of the compounds, have been retained in the descriptor's pool at the end of this initial stage.

**Model development**

The CP-MLR is a 'filter'-based variable selection procedure for model development in QSAR studies[34]. Its procedural aspects and implementation are discussed in some of our recent publications[35-40]. The thrust of this procedure is in its embedded 'filters'. They are briefly as follows: filter-1 seeds the variables by way of limiting inter-parameter correlations to predefined level (upper limit $\leq 0.79$); filter-2 controls the variables entry to a regression equation through t-values of coefficients (threshold value $\geq 2.0$); 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 $Q^2$ 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 ascertain any chance correlations associated with the models recognized in CP-MLR, each cross-validated model has been put to a randomization test[41,42] by repeated randomization of the activity profiles to discover 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.

**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[43,44]. 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 $\pm$ x $\times$ (standard deviations) 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 to 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.

## RESULTS AND DISCUSSION

A total number of 417 descriptors, belonging to 0D- to 2D-classes of DRAGON, have been computed for 56 compounds of Table 1 to quantify their MMP-2 and MMP-13 inhibition activities. A test-set has been selected through SYSTAT and the same was used for external validation of the models, derived from the training-set compounds. Fourteen compounds (S. Nos. **5**, **16**, **18**, **19**, **20**, **27**, **28**, **33**, **37**, **39**, **48**, **51**, **55** and **56**, Table 1) have been identified for the test-set while remaining compounds constitute the training-set for MMP-2 and MMP-13 activities. Next, the descriptors which were inter-correlated above 0.90 and exhibited correlation less than 0.1 with biological activities have been eliminated in the initial stage. The remaining 108 and 109 descriptors able to address, respectively, the MMP-2 and the MMP-13 inhibition activities of the compounds have been collated in the separate pools for CP-MLR analyses. A number of models in two-, three-, four-, five- and six-descriptors have been derived in succession. In doing so, filter-3 was in turn 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.

In order to quantify MMP-2 inhibition activity in terms of molecular descriptors, compounds **8**, **14** and **53** (Table 1), due to their uncertain activity values, have been eliminated from the data-set. The training-set was then employed to explore predictive models through CP-MLR. This resulted into 55 models in two-descriptors, 33 models in three-descriptors, 19 models in four-descriptors and 2 models in five-descriptors. However, models in five-descriptors only remained statistically more sensible and the same, in increasing level of significance, are given through Equations (1)-(2).

$$\begin{aligned} -\log IC_{50}\,(\text{MMP-2}) &= 5.586 - 1.341\,(0.226)\,\text{BELm7} - 1.197\,(0.275)\,\text{JGI4} + 1.359\,(0.338)\,\text{MATS4e} \\ &\quad + 1.405\,(0.309)\,\text{GATS8v} + 1.336\,(0.277)\,\text{C-028} \\ n = 39,\ r = 0.913,\ s &= 0.437,\ F = 33.980,\ Q^2_{LOO} = 0.765,\ Q^2_{L5O} = 0.738,\ r^2_{Test} = 0.602 \qquad \ldots(1) \end{aligned}$$

$$\begin{aligned} -\log IC_{50}\,(\text{MMP-2}) &= 5.701 + 2.697\,(0.552)\,\text{SEigp} - 1.973\,(0.227)\,\text{BELm7} - 1.392\,(0.344)\,\text{JGI3} \\ &\quad + 1.246\,(0.308)\,\text{GATS8v} + 1.628\,(0.272)\,\text{C-028} \\ n = 39,\ r = 0.916,\ s &= 0.430,\ F = 34.213,\ Q^2_{LOO} = 0.769,\ Q^2_{L5O} = 0.789,\ r^2_{Test} = 0.578 \qquad \ldots(2) \end{aligned}$$

The parameters 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 of regression coefficients. In all above equations, the F-values remained significant at 99% level [$F_{5,33}(0.01) = 3.630$]. The indices $Q^2_{LOO}$ and $Q^2_{L5O}$ (> 0.5) have accounted for internal robustness of these models while the index $r^2_{Test}$ greater than 0.5 specified that the selected test-set is accountable for external validation of above models. The descriptors, in all above models, have been scaled[45] between 0 and 1 so that they would have equal potential to influence the QSAR models and none of them dominate simply because it has larger or smaller pre-scaled values compare to the other descriptors. 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. In fact, a total number of 2 models, sharing 7 descriptors among them, have been obtained through CP-MLR. The shared 7 descriptors along with their class, brief description, average regression coefficients and total incidences are given in Table 3.

The descriptor SEigp, from the topological class, represents the eigenvalue sum from polarizability weighted distance matrix. The descriptor BELm7, from the BCUT class, is the lowest eigenvalue n. 7 of Burden matrix/ weighted by atomic masses. The descriptors, MATS4e and GATS8v are from the class of 2D autocorrelations. The MATS4e is the Moran autocorrelations of lag 4, weighted by atomic Sanderson

electronegativities while the descriptor GATS8v represents the Geary autocorrelation-lag 8/ weighted by atomic van der Waals volumes. The descriptors JGI3 and JGI4, from the Galvez class, are the mean topological charge indices of order 3 and order 4 respectively, the descriptor, C-028, from the atom-centred fragments class, corresponds to the functionality, R--CR--X.

The descriptors, BELm7 (Eqs. 1 and 2), JGI3 (Eq. 2) and JGI4 (Eq. 1) have negative impact on the MMP-2 inhibition activity. Therefore, the lower or more negative values of these descriptors are desirable to improve the activity of a compound. On the other hand, the descriptors, SEigp (Eq. 2), MATS4e (Eq. 1), GATS8v (Eqs. 1 and 2) and C-028 (Eq. 1) have shown positive influence on activity. The higher positive values of these descriptors would, therefore, enhance the activity.

Equations (1)-(2) emerged as significant predictive models and could estimate up to 83.90 percent of variance in observed activity of the compounds. These two models have been used to calculate the MMP-2 inhibition activity profiles of all the compounds and are included in Table 1 for the sake of comparison with observed ones. A close agreement between them has been observed. Moreover, the graphical display showing the variation of observed versus calculated activities is given in Fig. 2 to depict the goodness of fit for individual models.

The preliminary assessment of complete data-set for MMP-13 inhibition activity suggested that compound **24** and **32** (Table 1) could not fit into the trend followed by other compounds of the series. Both these analogues were, therefore, treated as the 'outliers'. However no appropriate reason is apparent, at present, for their unusual behavior. Considering 109 descriptors and the test-set, mentioned previously, the CP-MLR resulted into 5 models in three-descriptors, 6 models in four-descriptors, 7 models in five-descriptors and 2 models in six-descriptors from the training-set compounds. The six-descriptors models are given through Equations (3)-(4)

$$\begin{aligned} -\log IC_{50}\,(\text{MMP-13}) &= 9.236 - 0.605(0.175)\,\text{SIC3} - 0.561(0.167)\,\text{JGI4} - 1.214(0.174)\,\text{MATS8v} \\ &\quad -1.040(0.195)\,\text{nCs} - 0.663(0.203)\,\text{nCt} + 0.480(0.170)\,\text{C-026} \\ n = 40,\ r = 0.894,\ s &= 0.270,\ F = 21.987,\ Q^2_{LOO} = 0.710,\ Q^2_{L5O} = 0.720,\ r^2_{Test} = 0.742 \end{aligned} \quad \ldots(3)$$

$$\begin{aligned} -\log IC_{50}\,(\text{MMP-13}) &= 7.426 + 0.765(0.171)\,\text{HNar} - 0.732(0.164)\,\text{SIC3} + 0.534(0.168)\,\text{MATS4m} \\ &\quad +1.136(0.193)\,\text{MATS6m} - 1.051(0.173)\,\text{MATS8v} + 0.680(0.168)\,\text{C-026} \\ n = 40,\ r = 0.903,\ s &= 0.260,\ F = 24.230,\ Q^2_{LOO} = 0.739,\ Q^2_{L5O} = 0.705,\ r^2_{Test} = 0.616 \end{aligned} \quad \ldots(4)$$

In fact, the derived five-descriptor and six-descriptor models have shared a total number of 9 descriptors and the same are listed in Table 3.

The descriptors HNar and SIC3 (from the topological class) represent, the Narumi harmonic index and the structural information content (neighborhood symmetry of 3-order) respectively. The descriptor JGI4 (from the GLVZ class) is the mean topological charge index of order 4. The MATS4m, MATS6m and MATS8v (from the 2DAUTO class) are the representative of Moran autocorrelations, respectively, of lag 4 and lag 6/ weighted by atomic masses each and of lag 8/ weighted by atomic van der Waals volumes. The nCs and nCt (from the functional class) account, respectively for the number of total secondary and tertiary C ($sp^3$) atoms. Finally the descriptor, C-026 (from the atom-centred fragments class) accounts for functionality R--CX--R.

From Equations (3) and (4), it appeared that the descriptors, HNar, MATS4m, MATS6m, and C-026 have imparted positive impact on MMP-13 inhibition activity, suggesting that their higher values would be

advantageous in improving the activity. On the other hand, the descriptors, SIC3, JGI4, MATS8v, nCs and nCt have contributed negatively to the activity. Therefore, the lower or more negative values of these descriptors would be beneficial.

These two models (Eqs. 3 and 4) have accounted, respectively, for 79.92 and 81.54 percent of variances in the observed activity and have been used to calculate the MMP-13 inhibition activity profiles of all the compounds (Table 1). A close agreement between observed and calculated values has been observed. For identification of the goodness of fit, a plot between observed and calculated activities is given in Fig. 2.

The descriptors participated in individual models (Eqs. 1-4) have been found poorly intercorrelated. The intercorrelation matrices, amongst descriptors appeared in the individual models, have been avoided for the sake of brevity.

Further, the PLS analyses has also been carried out on 7 and 9 descriptors identified through CP-MLR for MMP-2 and MMP-13 inhibition activities respectively 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, two components have been found to be the optimum for each of these 7 and 9 descriptors and they explained, respectively, 84.82% and 81.36% of variance in the said activities. The PLS equations of two optimum components and MLR-like PLS coefficients of identified descriptors for MMP-2 and MMP-13 activities are given in Table 4. The calculated activity values of training- and test-set compounds remained in close agreement to that of the observed ones and are listed in Table 1. For comparison, the plot between observed and calculated activities (through PLS analysis) for the training- and test-set compounds is given in Fig. 2. Fig. 3 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity (Table 4). In the decreasing level of significance, 7 descriptors, being the part of Equations (1) and (2), have been arranged as BELm7, C-028, JGI4, GATS8v, MATS4e, SEigp and JGI3 for MMP-2 activity while 9 descriptors, shared Equations (3) and (4), have been arranged as MATS8v, SIC3, MATS6m, C-026, nCs, JGI4, MATS4m, nCt and HNar for MMP-13 activity. Similar conclusions have been observed from the PLS models for these two activities. The descriptors SEigp, MATS4e, GATS8v and C-028 have positive contribution to MMP-2 activity and the descriptors BELm7, JGI3 and JGI4 have negative contribution to it. Similarly, the descriptors HNar, MATS4m, MATS6m and C-026 have positively contributed to MMP-13 activity while the descriptors SIC3, JGI4, MATS8v, nCs and nCt have negatively contributed to it. The descriptors, in a given significant model, having positive contribution will augment the activity and their higher values are desirable to further improve it. On the other hand, the descriptors having negative contribution will diminish the activity. The lower or more negative values of such descriptors may, therefore, enhance the activity of a compound. It has also been observed that PLS model from the dataset devoid of 7 descriptors for MMP-2 activity and 9 descriptors for MMP-13 activity remained inferior in explaining the activity of the analogues.

The applicability domain (AD) has been analyzed for the models based on whole data set. It is characterized by the Williams plot of standardized residuals versus legerage ($h_i$) values. For this purpose, the most significant models for the MMP-2 and MMP-13 activities (Table 5) have been considered to calculate the standardized residuals and leverage values. The limits of normal values for the standardized residuals (response or Y outliers) were set as ± 2.5 × (standard deviation) while leverage threshold as h*. The graphical representations for the models, related to MMP-2 and MMP-13 activities, delineating the training-set and the test-set compounds is given in Fig. 4. As expected, compound **24** and **32** (Table 1) have been identified as the Y outliers through Equations (3) and (4). The standardized residuals of these two congeners were larger than the acceptable confines. Similarly compounds **45** (for MMP-2 activity) and **3**, **10** and **53** (for MMP-13 activity) have been predicted as the X outliers or the structurally influential analogues of the series as their leverage values, revealed by model Equations (2) and (3), remained greater than the threshold

estimate. For both the training- and test-set compounds, the suggested models for MMP-2 and MMP-13 activity match the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the AD of these two models and were evaluated correctly.

**Table 1: Observed and modeled MMP-2 and MMP-13 inhibition activities in a class of selective α-sulfone hydroxamates[a]**

| S. No. | X | R | -log$IC_{50}$ (M) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MMP-2 | | | | MMP-13 | | | |
| | | | Obsd.[b] | Cald. Eq. (1) | Cald. Eq. (2) | PLS | Obsd.[b] | Cald. Eq. (4) | Cald. Eq. (5) | PLS |
| **1** | O | Allyl(methyl)amino | 5.54 | 5.53 | 5.53 | 5.54 | 7.46 | 7.62 | 7.33 | 7.58 |
| **2** | O | Methyl(prop-2-ynyl) amino | 5.68 | 5.64 | 5.68 | 5.65 | 7.61 | 7.79 | 7.67 | 7.87 |
| **3** | N-*c*Pr | Benzyl(methyl) amino | 5.46 | 5.51 | 5.81 | 5.57 | 7.70 | 7.66 | 8.02 | 7.83 |
| **4** | O | 3,4-Dihydroisoquinolin-2(1*H*)-yl | 5.68 | 5.78 | 6.44 | 5.87 | 8.05 | 8.04 | 8.18 | 8.10 |
| **5**[c] | O | 6,7-Dimethoxy-3,4-dihydroisoquinolin-2(1*H*)-yl | 5.62 | 4.69 | 5.63 | 4.67 | 8.21 | 8.02 | 8.15 | 7.95 |
| **6** | O | 3,5-Dimethylpiperidin-1-yl | 5.14 | 5.55 | 5.45 | 5.42 | 8.36 | 7.94 | 7.84 | 7.94 |
| **7** | N-$CH_2CH_2$OMe | 3,5-dimethylpiperidin-1-yl | 5.77 | 5.51 | 5.62 | 5.40 | 7.30 | 7.43 | 7.34 | 7.39 |
| **8** | O | *cis*-2,6-Dimethylmorpholin-4-yl | < 5.00[d] | 5.14 | 5.08 | 5.04 | 7.74 | 7.81 | 7.66 | 7.76 |
| **9** | O | 4-Acetylpiperazin-1-yl | 5.60 | 5.10 | 5.12 | 5.13 | 7.30 | 7.86 | 7.65 | 7.78 |
| **10** | O | 4-Isopropylpiperazin-1-yl | 5.26 | 5.72 | 5.47 | 5.65 | 7.55 | 7.35 | 7.74 | 7.54 |
| **11** | O | 4-(2-Methoxyethyl) piperazin-1-yl | 5.15 | 5.59 | 5.38 | 5.59 | 7.35 | 7.18 | 7.14 | 7.14 |
| **12** | O | 4-Phenethylpiperazin-1-yl | 6.05 | 6.25 | 6.51 | 6.39 | 7.60 | 7.69 | 7.99 | 7.80 |
| **13** | O | 4-(2-Hydroxyethyl) piperazin-1-yl | 5.22 | 5.71 | 5.58 | 5.68 | 7.40 | 7.40 | 7.65 | 7.54 |
| **14** | O | 4-(2-(Dimethylamino) ethyl)piperazin-1-yl | < 5.00[d] | 5.30 | 5.39 | 5.35 | 7.05 | 7.15 | 7.14 | 7.10 |
| **15** | O | 4-(2-Fluorophenyl) piperazin-1-yl | 5.85 | 5.80 | 5.65 | 5.75 | 8.17 | 7.81 | 8.00 | 7.87 |
| **16**[c] | O | 4-(2-Methoxyphenyl) piperazin-1-yl | 5.35 | 5.67 | 5.73 | 5.72 | 7.74 | 7.60 | 7.54 | 7.49 |

Cont…

| S. No. | X | R | -logIC$_{50}$ (M) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MMP-2 | | | | MMP-13 | | | |
| | | | Obsd.[b] | Cald. Eq. (1) | Cald. Eq. (2) | PLS | Obsd.[b] | Cald. Eq. (4) | Cald. Eq. (5) | PLS |
| **17** | O | 4-(4-Fluorophenyl) piperazin-1-yl | 5.80 | 6.01 | 6.11 | 6.06 | 8.22 | 8.05 | 8.33 | 8.17 |
| **18**[c] | O | 4-(4-Acetylphenyl) piperazin-1-yl | 6.15 | 6.73 | 6.73 | 6.73 | 8.17 | 8.34 | 8.16 | 8.28 |
| **19**[c] | O | 4-(2,4-Dimethylphenyl) piperazin-1-yl | 5.25 | 6.17 | 6.28 | 6.11 | 7.91 | 7.83 | 7.92 | 7.88 |
| **20**[c] | O | 4-(Pyridin-2-yl)piperazin-1-yl | 5.44 | 5.54 | 5.54 | 5.73 | 7.97 | 7.74 | 7.68 | 7.70 |
| **21** | O | 4-(Pyrimidin-2-yl) piperazin-1-yl | 5.80 | 5.32 | 5.42 | 5.54 | 8.19 | 7.85 | 7.84 | 7.83 |
| **22** | O | 4-(Pyridin-4-yl) piperazin-1-yl | 5.76 | 6.63 | 6.50 | 6.68 | 7.52 | 7.94 | 7.98 | 7.96 |
| **23** | O | 4-(Pyrazin-2-yl)piperazin-1-yl | 6.05 | 5.54 | 5.41 | 5.68 | 7.57 | 7.75 | 7.75 | 7.75 |
| **24** | O | 4-(2,3-Dimethylphenyl) piperazin-1-yl | 5.22 | 5.21 | 5.46 | 5.20 | 8.40[e] | -- | -- | -- |
| **25** | N-*c*Pr | 4-(2,3-Dimethylphenyl) piperazin-1-yl | 5.20 | 5.32 | 5.80 | 5.32 | 7.56 | 7.53 | 7.38 | 7.34 |
| **26** | N-$CH_2CH_2$ OMe | 4-(2,3-Dimethylphenyl) piperazin-1-yl | 6.40 | 5.29 | 5.64 | 5.27 | 7.15 | 6.86 | 6.86 | 6.78 |
| **27**[c] | N-$CH_2CH_2$ OMe | Aniline amide | 6.40 | 5.86 | 6.46 | 5.97 | 8.05 | 7.72 | 7.51 | 7.55 |
| | | **Y** **$R_1$** | | | | | | | | |
| **28**[c] | O | N H | 7.63 | 8.61 | 8.76 | 8.74 | 8.77 | 8.70 | 8.56 | 8.60 |
| **29** | O | N 2-F | 7.68 | 7.72 | 7.83 | 7.71 | 8.57 | 8.40 | 8.47 | 8.42 |
| **30** | O | N 2-Me | 7.06 | 7.00 | 6.75 | 6.90 | 8.11 | 8.14 | 8.12 | 8.14 |
| **31** | O | N 2-Cl | 7.40 | 7.57 | 8.09 | 7.71 | 8.05 | 8.20 | 8.25 | 8.20 |
| **32** | O | N 2-OMe | 6.43 | 6.31 | 6.24 | 6.29 | 6.89[e] | -- | -- | -- |
| **33**[c] | O | N 3-OMe | 7.31 | 7.28 | 7.12 | 7.17 | 8.10 | 8.49 | 8.20 | 8.29 |
| **34** | O | N 3-$CF_3$ | 6.48 | 6.90 | 6.84 | 6.82 | 7.70 | 7.77 | 7.93 | 7.94 |
| **35** | O | N 4-OMe | 7.74 | 7.70 | 7.71 | 7.67 | 9.20 | 9.27 | 9.16 | 9.20 |
| **36** | O | N 4-Me | 7.46 | 7.66 | 7.52 | 7.61 | 8.72 | 8.73 | 8.37 | 8.54 |
| **37**[c] | O | N 2,4-diMe | 5.85 | 7.07 | 6.89 | 6.89 | 7.54 | 8.31 | 8.35 | 8.35 |
| **38** | N-*c*Pr | N H | 8.27 | 7.54 | 7.40 | 7.53 | 8.48 | 8.72 | 8.50 | 8.52 |
| **39**[c] | N-*c*Pr | N 4-$CF_3$ | 7.49 | 6.91 | 6.13 | 6.64 | 8.62 | 8.77 | 8.87 | 8.89 |

Cont…

| S. No. | X | R | | -logIC$_{50}$ (M) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | MMP-2 | | | | MMP-13 | | | |
| | | | | Obsd.[b] | Cald. Eq. (1) | Cald. Eq. (2) | PLS | Obsd.[b] | Cald. Eq. (4) | Cald. Eq. (5) | PLS |
| | | Y | $R_1$ | | | | | | | | |
| **40** | O | CH | H | 8.36 | 7.99 | 7.90 | 8.12 | 8.22 | 8.21 | 8.01 | 8.13 |
| **41** | O | CH | 2-OMe | 5.43 | 5.68 | 5.46 | 5.70 | 7.76 | 7.58 | 7.47 | 7.50 |
| **42** | O | CH | 4-Cl | 8.70 | 7.91 | 8.37 | 8.09 | 9.15 | 8.69 | 8.87 | 8.86 |
| **43** | O | CH | 2-Cl | 7.09 | 6.91 | 7.24 | 7.08 | 7.37 | 7.81 | 7.75 | 7.79 |
| **44** | O | CH | 2-Me | 6.15 | 6.50 | 5.95 | 6.39 | 7.49 | 7.84 | 7.80 | 7.85 |
| **45** | O | CH | 2-$CF_3$ | 6.05 | 6.29 | 5.80 | 6.00 | 7.54 | 7.46 | 7.53 | 7.57 |
| **46** | O | CH | 2-OEt | 5.51 | 6.08 | 5.79 | 6.10 | 7.46 | 7.58 | 7.25 | 7.36 |
| **47** | O | CH | 2-OH | 7.15 | 7.15 | 7.20 | 7.21 | 7.95 | 8.01 | 8.04 | 8.05 |
| **48**[c] | O | CH | 2-(4-F-$C_6H_4$) | 5.27 | 6.28 | 6.00 | 6.26 | 7.47 | 8.01 | 8.08 | 8.02 |
| **49** | O | CH | 2,3-(CH=CH) (naphthyl) | 6.30 | 6.42 | 6.07 | 6.34 | 7.94 | 7.44 | 7.70 | 7.56 |
| **50** | O | CH | 2-Me,4-OMe | 6.30 | 6.37 | 6.15 | 6.19 | 8.15 | 8.18 | 8.11 | 8.16 |
| **51**[c] | O | CH | 2,4-diOMe | 5.76 | 5.71 | 5.46 | 5.57 | 7.96 | 7.89 | 7.67 | 7.69 |
| **52** | O | CH | 2,5-diOMe | 6.10 | 5.48 | 5.83 | 5.46 | 7.94 | 8.10 | 7.88 | 7.89 |
| **53** | O | CH | 2-OMe,5-*i*Pr | < 5.00[d] | 5.95 | 6.13 | 5.89 | 7.15 | 7.35 | 7.28 | 7.24 |
| **54** | O | N | 4-$CF_3$ | 7.92 | 8.04 | 7.48 | 7.88 | 9.22 | 8.83 | 9.00 | 9.05 |
| **55**[c] | O | N | 4-Cl | 8.43 | 8.53 | 9.22 | 8.70 | 9.00 | 9.21 | 9.53 | 9.40 |
| **56**[c] | N-*c*Pr | N | 4-OMe | 8.28 | 7.48 | 7.73 | 7.46 | 9.38 | 9.14 | 8.96 | 8.96 |

[a]For general structures, Fig. 1(A) for compounds 1-27 and Fig. 1(B) for compounds 28-56,
[b]$IC_{50}$ represents the concentration of a compound required to bring out 50% inhibition of MMP-2/MMP-13,
[c]compounds in test-set, [d]eliminated compound due to uncertain activity, [e]" outlier" compound in present study

**Table 2: Descriptor classes used for the analysis of MMP-2 and MMP-13 inhibition activity of the compounds**

| Descriptor class (acronyms) | Definition and scope |
|---|---|
| Constitutional (CONST) | Dimensionless or 0D descriptors; independent from molecular connectivity and conformations. |
| Topological (TOPO) | 2D-descriptor from molecular graphs and independent conformations. |
| Molecular walk counts (MWC) | 2D-descriptors representing self-returning walk counts of different lengths. |
| Modified Burden eigenvalues (BCUT) | 2D-descriptors representing positive and negative eigen values of the adjacency matrix, weights the diagonal elements and atoms. |
| Galvez topological charge indices (GLVZ) | 2D-descriptors representing the first 10 eigen values of corrected adjacency matrix. |

Cont…

| Descriptor class (acronyms) | Definition and scope |
|---|---|
| 2D-autocorrelations (2DAUTO) | 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). |
| Functional groups (FUNC) | Molecular descriptors based on the counting of the chemical functional groups. |
| Atom-centred fragments (ACF) | Molecular descriptors based on the counting of 120 atom-centred fragments, as defined by Ghose-Crippen. |
| Empirical (EMP) | 1D-descriptors represent the counts of non-single bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule. |
| Properties (PROP) | 1D-descriptors representing molecular properties of a molecule. |

**Table 3: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the MMP-2 and MMP-13 inhibition activities**

| S. No. | Descriptor | Descriptor class | Physical meaning | Average regression coefficient (incidence) | |
|---|---|---|---|---|---|
| | | | | MMP-2 | MMP-13 |
| **1** | HNar | TOPO | Narumi harmonic index. | | 0.765 (1) |
| **2** | SIC3 | TOPO | Structural information content (neighborhood symmetry of 3-order). | | -0.669 (2) |
| **3** | SEigp | TOPO | Eigenvalue sum from polarizability weighted distance matrix. | 2.697 (1) | |
| **4** | BELm7 | BCUT | Lowest eigenvalue n. 7 of Burden matrix/ weighted by atomic masses. | -1.657 (2) | |
| **5** | JGI3 | GLVZ | Mean topological charge index of order 3. | -1.392 (1) | |
| **6** | JGI4 | GLVZ | Mean topological charge index of order 4. | -1.197 (1) | -0.561 (1) |
| **7** | MATS4m | 2DAUTO | Moran autocorrelation - lag 4/ weighted by atomic masses. | | 0.535 (1) |
| **8** | MATS4e | 2DAUTO | Moran autocorrelation - lag 4/ weighted by atomic Sanderson electronegativities. | 1.359 (1) | |
| **9** | MATS6m | 2DAUTO | Moran autocorrelation - lag 6/ weighted by atomic masses. | | 1.136 (1) |
| **10** | MATS8v | 2DAUTO | Moran autocorrelation - lag 8/ weighted by atomic van der Waals volumes. | | -1.133 (2) |
| **11** | GATS8v | 2DAUTO | Geary autocorrelation - lag 8/ weighted by atomic van der Waals volumes. | 1.326 (2) | |
| **12** | nCs | FUN | Number of total secondary C ($sp^3$). | | -1.040 (1) |

Cont…

| S. No. | Descriptor | Descriptor class | Physical meaning | Average regression coefficient (incidence) | |
|---|---|---|---|---|---|
| | | | | MMP-2 | MMP-13 |
| **13** | nCt | FUN | Number of total tertiary C ($sp^3$). | | -0.663 (1) |
| **14** | C-026 | ACF | Corresponds to R—CX—R. | | 0.580 (2) |
| **15** | C-028 | ACF | Corresponds to R—CR—X. | 1.482 (2) | |

[a]The descriptors have been identified from the models, emerged from CP-MLR protocol with a training-set of 39 and 40 compounds for MMP-2 and MMP-13 inhibition activities respectively. [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

**Table 4: PLS and MLR-like PLS models from the descriptors of five and six parameter CP-MLR models for MMP-2 and MMP-13 inhibition activities**

**A: PLS equation**

| PLS components | PLS coefficient (s.e.)[a] | |
|---|---|---|
| | MMP-2 | MMP-13 |
| Component-1 | -0.724 (0.052) | -0.214 (0.020) |
| Component-2 | -0.155 (0.053) | 0.211 (0.032) |
| Constant | 6.313 | 7.876 |

**B: MLR-like PLS equation**

| S. No. | MMP-2 | | | S. No. | MMP-13 | | |
|---|---|---|---|---|---|---|---|
| | Descriptor | MLR-like coefficient (f. c.)[b] | Order | | Descriptor | MLR-like coefficient (f. c.)[b] | Order |
| **1** | Seigp | 0.115 (0.062) | 6 | **1** | Hnar | 0.107 (0.050) | 9 |
| **2** | BELm7 | -0.502 (-0.270) | 1 | **2** | SIC3 | -0.287 (-0.134) | 2 |
| **3** | JGI3 | -0.075 (-0.040) | 7 | **3** | JGI4 | -0.173 (-0.081) | 6 |
| **4** | JGI4 | -0.320 (-0.172) | 3 | **4** | MATS4m | 0.162 (0.076) | 7 |
| **5** | MATS4e | 0.197 (0.106) | 5 | **5** | MATS6m | 0.267 (0.125) | 3 |
| **6** | GATS8v | 0.310 (0.166) | 4 | **6** | MATS8v | -0.538 (-0.251) | 1 |
| **7** | C-028 | 0.343 (0.184) | 2 | **7** | nCs | -0.216 (-0.101) | 5 |
| | Constant | 21.001 | | **8** | nCt | -0.140 (-0.065) | 8 |
| | | | | **9** | C-026 | 0.253 (0.118) | 4 |
| | | | | | Constant | 8.344 | |

Cont…

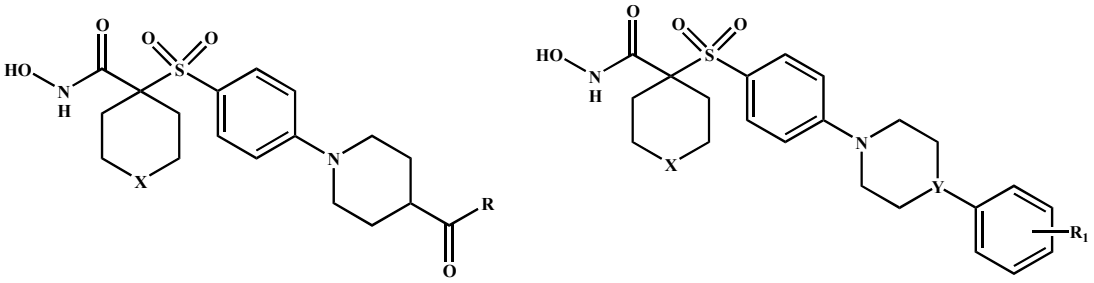
**C: PLS regression statistics**

| Symbol | Value | |
|---|---|---|
| | **MMP-2** | **MMP-13** |
| n | 39 | 40 |
| r | 0.921 | 0.902 |
| s | 0.339 | 0.246 |
| F | 100.555 | 80.879 |
| $Q^2_{LOO}$ | 0.824 | 0.785 |
| $Q^2_{L5O}$ | 0.816 | 0.785 |
| $r^2_{Test}$ | 0.588 | 0.600 |

[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

**Table 5: Models derived for the whole data set (n = 53) for the MMP-2 and (n = 56) for the MMP-13 inhibition activities**

| Model | r | s | F | $Q^2_{LOO}$ | $Q^2_{L5O}$ | Eq. |
|---|---|---|---|---|---|---|
| $-\log IC_{50}$ (MMP-2) = 5.551 - 1.302 (0.244) BELm7 - 0.957 (0.265) JGI4 + 1.127 (0.352) MATS4e + 1.561 (0.301) GATS8v + 1.170 (0.276) C-028 | 0.879 | 0.516 | 32.015 | 0.710 | 0.703 | 1(a) |
| $-\log IC_{50}$ (MMP-2) = 5.900 + 1.818 (0.570) SEigp - 1.772 (0.237) BELm7 - 1.395 (0.348) JGI3 + 1.407 (0.300) GATS8v + 1.408 (0.268) C-028 | 0.881 | 0.512 | 32.658 | 0.709 | 0.721 | 2(a) |
| $-\log IC_{50}$ (MMP-13) = 9.091 - 0.579 (0.193) SIC3 - 0.427 (0.173) JGI4 -1.104 (0.182) MATS8v - 0.980 (0.231) nCs-0.669 (0.255) nCt + 0.401 (0.167) C-026 | 0.824 | 0.347 | 17.314 | 0.593 | 0.567 | 3(a) |
| $-\log IC_{50}$ (MMP-13) = 7.582 + 0.790 (0.213) HNar -0.690 (0.192) SIC3 + 0.425 (0.195) MATS4m + 0.837 (0.216) MATS6m - 0.955 (0.190) MATS8v + 0.585 (0.176) C-026 | 0.819 | 0.352 | 16.615 | 0.568 | 0.566 | 4(a) |



(A) (B)

**Fig. 1 (A) and 1 (B): General structures for the compounds in Table 1**

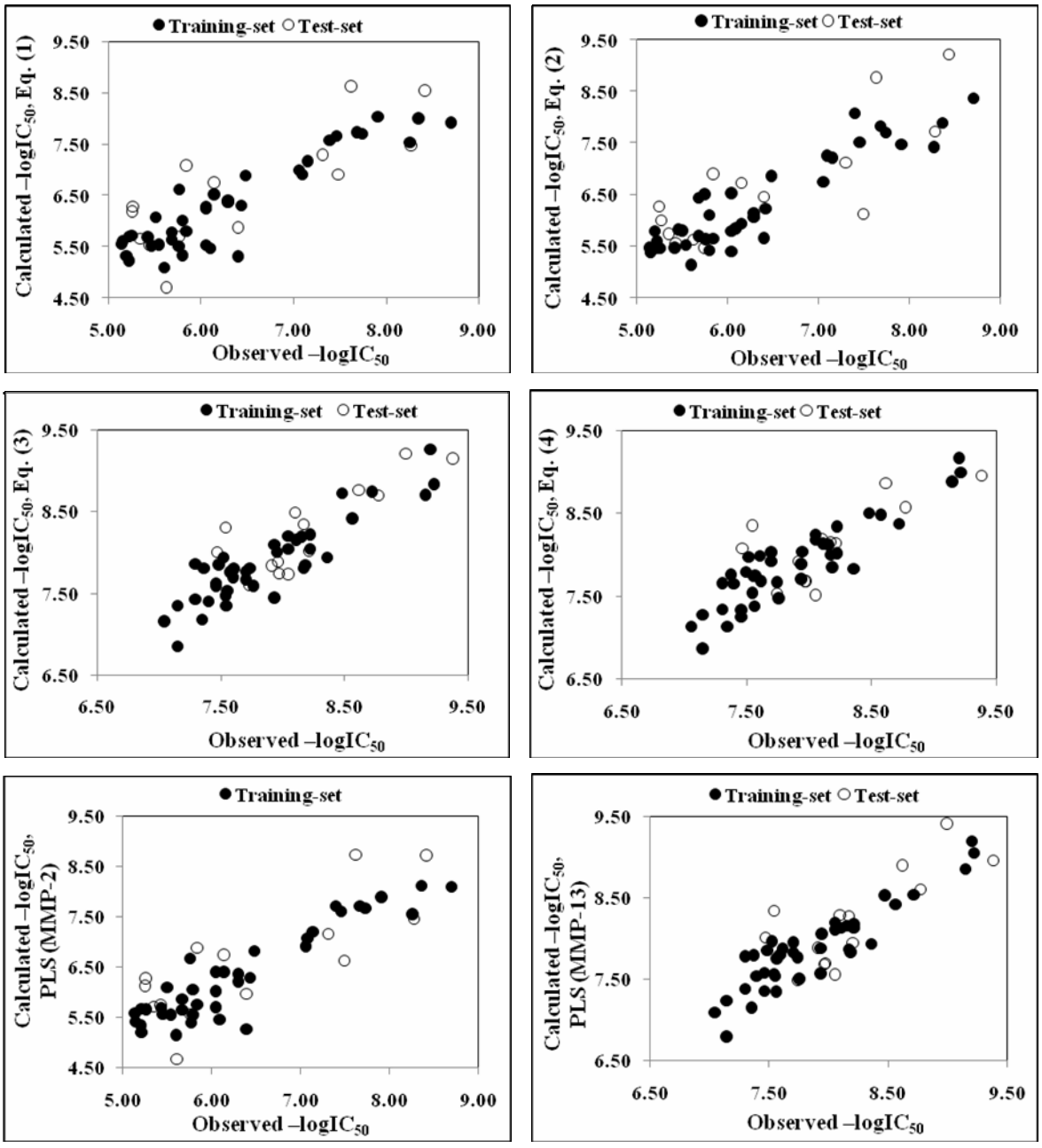



**Fig. 2: Plot of observed versus caculated -log$IC_{50}$ values relating to the inhibition of MMP-2 and MMP-13 for training-set and test-set compounds**

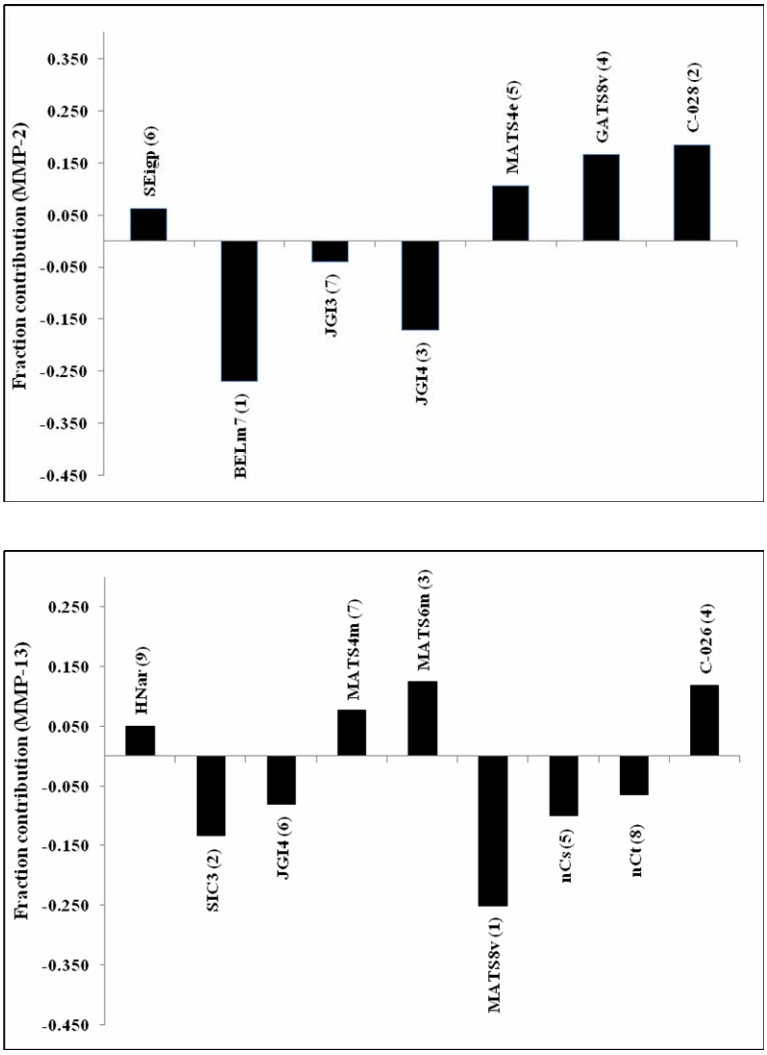
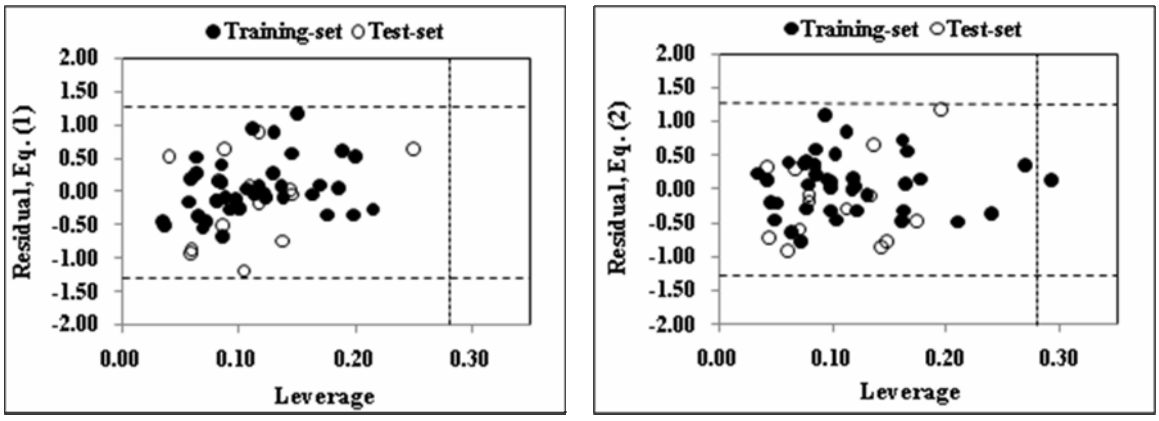



**Fig. 3: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 7 and 9 identified descriptors (Table 4) associated, resepectively, with MMP-2 and MMP-13 inhibition activities of the compounds**



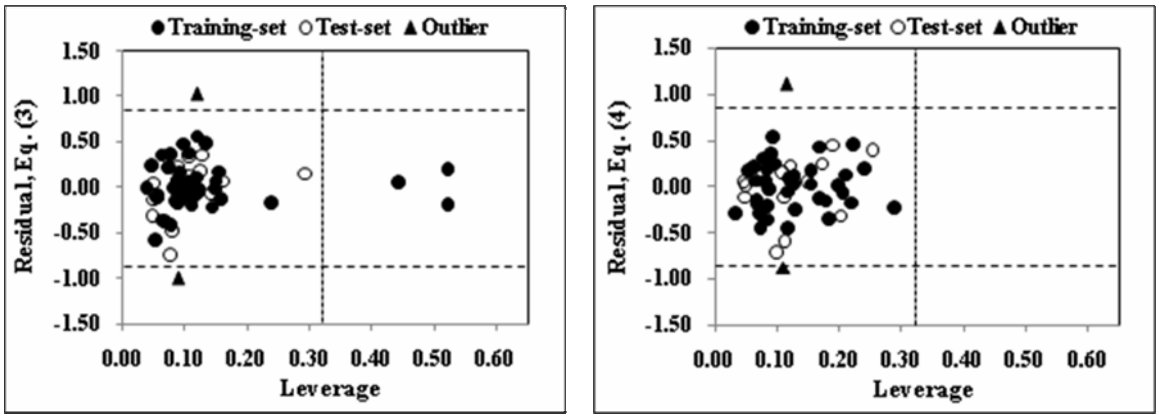



**Fig. 4: Williams plot for the MMP-2 and MMP-13 inhibition activities of training-set and test-set compounds (Table 1). The horizontal dotted line refers to the residual limit ± 2.5 × (standard deviation) and the vertical dotted line represents threshold leverage, h* (= 0.283 and 0.321 for inhibition of MMP-2 and MMP-13 respectively)**

## CONCLUSION

The MMP-2 and MMP-13 inhibition activity of selective α-sulfone hydroxamates have been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the inhibition activities of these congeners. For MMP-2 inhibition activity, the descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of the eigenvalue sum from polarizability weighted distance matrix (SEigp), the lowest eigenvalue n. 7 of Burden matrix/ weighted by atomic masses (BELm7), the Moran autocorrelation-lag 4/ weighted by atomic Sanderson electronegativities (MATS4e), the Geary autocorrelation-lag 8/ weighted by atomic van der Waals volumes (GATS8v), the mean topological charge indices of order 3 (JGI3) and order 4 (JGI4), and the functionality, R--CR--X (C-028). For MMP-13 activity, the Narumi harmonic index (HNar), the structural information content of neighborhood symmetry of 3-order (SIC3), the mean topological charge index of order 4 (JGI4), the Moran autocorrelations of lag 4 and 6/ weighted by atomic masses each (MATS4m and MATS6m) and of lag 8/ weighted by atomic van der Waals volumes (MATS8v), the number of total secondary C ($sp^3$) (nCs), the number of total tertiary C ($sp^3$) (nCt), and the functionality R--CX--R (C-026) remained the significant contributors. Such guidelines may be helpful in exploring more potential analogues of the series. The statistics emerged from the test-set 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. Except a few structurally influential analogues, all the compounds are within the applicability domain of the proposed models and were evaluated correctly.

## ACKNOWLEDGEMENT

Author is thankful to the University Grants Commission, New Delhi for financial support provided under the scheme of Emeritus Fellowship. He also expresses his gratitude to Institution for making available necessary facilities to complete this work.

## REFERENCES

1. C. E. Brinckerhoff and L. M. Matrisian, Nat. Rev. Mol. Cell Biol., **3**, 207 (2002).
2. H. Nagase and J. F. Jr., Woessner, J. Biol. Chem., **274**, 21491 (1999).

3. C. A. Corbitt, J. Lin, and M. L. Lindsey, Recent Pat. Anti-Cancer Drug Discov., **2**, 135 (2007).

4. J. F. Fisher and S. Mobashery, Cancer Metastasis Rev., **25**, 115 (2006).

5. J. Martel-Pelletier, D. J. Welsch and J. Pelletier, Best Pract. Res. Clin. Rheumatol., **15**, 805 (2001).

6. M. P. Hudson, P. W. Armstrong, W. Ruzyllo, J. Brum, L. Cusmano, P. Krzeski, R. Lyon, M. Quinones, P. Theroux, D. Sydlowski, H. E. Kim, M. J. Garcia, W. A. Jaber and W. D. Weaver, J. Am. Coll. Cardiol., **48**, 15 (2006).

7. R. J. MacFadyen, Curr. Opin. Pharmacol., **7**, 171 (2007).

8. M. L. Lindsey, Heart Fail. Rev., **9**, 7 (2004).

9. J. T. Peterson, Heart Fail. Rev., **9**, 63 (2004).

10. T. M. Doherty, K. Asotra, D. Pei, H. Uzui, D. J. Wilkin, P. K. Shah and T. B. Rajavashisth, Expert Opin. Ther. Pat., **12**, 665 (2002).

11. F. E. Jacobsen, J. A. Lewis and S. M. Cohen, Chem. Med. Chem., **2**, 152 (2007).

12. J. W. Skiles, N. C. Gonnella and A. Y. Jeng, Curr. Med. Chem., **11**, 2911 (2004).

13. K. Holmbeck, P. Bianco, J. Caterina, S. Yamada, M. Kromer, S. A. Kuznetsov, M. Mankani, P. G. Robey, A. R. Poole, I. Pidoux, J. M. Ward and H. Birkedal-Hansen, Cell (Cambridge, Mass.), **99**, 81 (1999).

14. Z. Zhou, S. S. Apte, R. Soininen, R. Cao, G. Y. Baaklini, R. W. Rauser, J. Wang, Y. Cao and K. Tryggvason, Proc. Natl. Acad. Sci. (U.S.A.), **97**, 4052 (2000).

15. Y. Hu, J. S. Xiang, M. J. DiGrandi, X. Du, M. Ipek, L. M. Laakso, J. Li, W. Li, T. S. Rush, J. Schmid, J. S. Skotnicki, S. Tam, J. R. Thomason, Q. Wang and J. I. Levin, Bioorg. Med. Chem., **13**, 6629 (2005).

16. J. Li, T. S. Rush, W. Li, D. DeVincentis, X. Du, Y. Hu, J. R. Thomason, J. S. Xiang, J. S. Skotnicki, S. Tam, K. M. Cunningham, P. S. Chockalingam, E. A. Morris and J. I. Levin, Bioorg. Med. Chem. Lett., **15**, 4961 (2005).

17. J. Wu, T. S. Rush, R. Hotchandani, X. Du, M. Geck, E. Collins, Z. Xu, J. Skotnicki, J. I. Levin and F. E. Lovering, Bioorg. Med. Chem. Lett., **15**, 4105 (2005).

18. K. D. Freeman-Cook, L. A. Reiter, M. C. Noe, A. S. Antipas, D. E. Danley, K. Datta, J. T. Downs, S. Eisenbeis, J. D. Eskra, D. J. Garmene, E. M. Greer, R. J. Griffiths, R. Guzman, J. R. Hardink, F. Janat, C. S. Jones, G. J. Martinelli, P. G. Mitchell, E. R. Laird, J. L. Liras, L. L. Lopresti-Morrow, J. Pandit, U. D. Reilly, D. Robertson, M. L. Vaughn-Bowser, L. A. Wolf-Gouviea and S. A. Yocum, Bioorg. Med. Chem. Lett., **17**, 6529 (2007).

19. D. P. Becker, T. E. Barta, L. Bedell, G. DeCrescenzo, J. Freskos, D. P. Getman, S. L. Hockerman, M. Li, P. Mehta, B. Mischke, G. E. Munie, C. Swearingen and C. I. Villamil, Bioorg. Med. Chem. Lett., **11**, 2719 (2001).

20. D. P. Becker, G. DeCrescenzo, J. Freskos, D. P. Getman, S. L. Hockerman, M. Li, P. Mehta, G. E. Munie and C. Swearingen, Bioorg. Med. Chem. Lett., **11**, 2723 (2001).

21. T. E. Barta, D. P. Becker, L. J. Bedell, G. A. De Crescenzo, J. J. McDonald, G. E. Munie, S. Rao, H. Shieh, R. Stegeman, A. M. Stevens and C. I. Villamil, Bioorg. Med. Chem. Lett., **10**, 2815 (2000).

22. T. E. Barta, D. P. Becker, L. J. Bedell, G. A. De Crescenzo, J. J. McDonald, P. Mehta, G. E. Munie and C. I. Villamil, Bioorg. Med. Chem. Lett., **11**, 2481 (2001).

23. D. P. Becker, C. I. Villamil, T. E. Barta, L. J. Bedell, T. L. Boehm, G. A. DeCrescenzo, J. N. Freskos, D. P. Getman, S. Hockerman, R. Heintz, S. C. Howard, M. H. Li, J. J. McDonald, C. P. Carron, C. L. Funckes-Shippy, P. Mehta, G. E. Munie and C. A. Swearingen, J. Med. Chem., **48**, 6713 (2005).

24. V. Aranapakam, J. M. Davis, G. T. Grosu, J. Baker, J. Ellingboe, A. Zask, J. I. Levin, V. P. Sandanayaka, M. Du, J. S. Skotnicki, J. F. DiJoseph, A. Sung, M. A. Sharr, L. M. Killar, T. Walter, G. Jin, R. Cowling, J. Tillett, W. Zhao, J. McDevitt and Z. B. Xu, J. Med. Chem., **46**, 2376 (2003).

25. V. Aranapakam, G. T. Grosu, J. M. Davis, B. Hu, J. Ellingboe, J. L. Baker, J. S. Skotnicki, A. Zask, J. F. DiJoseph, A. Sung, M. A. Sharr, L. M. Killar, T. Walter, G. Jin and R. Cowling, J. Med. Chem., **46**, 2361 (2003).

26. J. S. Condon, D. Joseph-McCarthy, J. I. Levin, H. Lombart, F. E. Lovering, L. Sun, W. Wang, W. Xu and Y. Zhang, Bioorg. Med. Chem. Lett., **17**, 34 (2007).

27. K. Park, A. Aplasca, M. T. Du, L. Sun, Y. Zhu, Y. Zhang and J. I. Levin, Bioorg. Med. Chem. Lett., **16**, 3927 (2006).

28. Y. Zhang, X. Fan, B. Xiang, D. Chakravarty, R. Scannevin, S. Burke, P. Karnachi, K. Rhodes and P. Jackson, Bioorg. Med. Chem. Lett., **16**, 3096 (2006).

29. S. A. Kolodziej, S. L. Hockerman, G. A. DeCrescenzo, J. J. McDonald, D. A. Mischke, G. E. Munie, T. R. Fletcher, N. Stehle, C. Swearingen and D. P. Becker, Bioorg. Med. Chem. Lett., **20**, 3561 (2010).

30. S. A. Kolodziej, S. L. Hockerman, T. L. Boehm, J. N. Carroll, G. A. DeCrescenzo, J. J. McDonald, D. A. Mischke, G. E. Munie, T. R. Fletcher, J. G. Rico, N. W. Stehle, C. Swearingen, P. Daniel and D. P. Becker, Bioorg. Med. Chem. Lett., **20**, 3557 (2010).

31. SYSTAT (version 7.0); SPSS Inc 444 North Michigan Avenue, Chicago IL, 60611.

32. Chemdraw Ultra 6.0 and Chem3D Ultra, Cambridge Soft Corporation, Cambridge, USA; http://www.cambridgesoft.com

33. Dragon Software (version 1.11-2001) by R. Todeschini V. Consonni and Milano, Italy, http//www.talete.mi.it/dragon.htm

34. Y. S. Prabhakar, QSAR Comb. Sci., **22**, 583 (2003).

35. S. Sharma, Y. S. Prabhakar, P. Singh and B. K. Sharma, Eur. J. Med. Chem., **43**, 2354 (2008).

36. S. Sharma, B. K. Sharma, S. K. Sharma, P. Singh and Y. S. Prabhakar, Eur. J. Med. Chem., **44**, 1377 (2009).

37. B. K. Sharma, P. Pilania, P. Singh and Y. S. Prabhakar, SAR QSAR Environ. Res., **21**, 169 (2010).

38. B. K. Sharma, P. Singh, K. Sarbhai and Y. S. Prabhakar, SAR QSAR Environ. Res., **21**, 369 (2010).

39. B. K. Sharma, P. Pilania, K. Sarbhai, P. Singh and Y. S. Prabhakar, Mol. Divers., **14**, 371 (2010).

40. B. K. Sharma, P. Singh, M. Shekhawat, K. Sarbhai and Y. S. Prabhakar, SAR QSAR Environ. Res., **22**, 365 (2011).

41. S. S. So and M. Karplus, J. Med. Chem., **40**, 4347 (1997).

42. Y. S. Prabhakar, V. R. Solomon, R. K. Rawal, M. K. Gupta and S. B. Katti, QSAR Comb. Sci., **23**, 234 (2004).

43. P. Gramatica, QSAR Comb. Sci., **26**, 694 (2007).

44. L. Eriksson, J. Jaworska, A. P. Worth, M. T. D. Cronin, R. M. McDowell and P. Gramatica, Env. Health Persp., **111**, 1361 (2003).

45. A. Golbraikh and A. Tropsha, J. Mol. Graph Model, **20**, 269 (2002).