ISSN: 0974 - 7516 Volume 7 Issue 1

Trade Science Inc.

Organic CHEMISTRY

An Indian Journal

Full Paper

OCAIJ, 7(1), 2011 [7-16]

Pharmacophore based 3D-QSAR study of matrix metalloproteinase (MMPs) inhibitors

Divya Rastogi, Sachi Mittal, P.P.Singh*

Department of Chemistry, Bareilly College, Bareilly, U.P., (INDIA)

E-mail: dr_ppsingh@sify.com

Received: 23rd May, 2010 ; Accepted: 2nd June, 2010

ABSTRACT

Matrix metalloproteinase (MMPs also known as metzincins generally expressed in very low amounts and their transcription is tightly regulated either positively or negatively by cytokines and growth factors or tumor necrosis factor. MMPs dealing with various aspects of their physiologic and pathologic roles in biologic processes. Therefore, inhibitors of MMPs have several interests. In order to understand the possible interactions and necessary structural features we have derived common pharmacophore hypotheses (CPHs) among a series of hydroxamate derivatives using PHASE module of Schrödinger suit. The study implies that a hydrophobic centre along with four different hydrogen bond acceptor site with in a set of conditions was common to all the derivativs. This CPH has significantly contributed (q^2=0.92, r^2=0.94, F=55.2) to inhibitory activity. Overall, the Pharmacophore models have given an insight to the importance of hydrophobic and hydrogen bond acceptor groups which is helpful for further modification and designing of such inhibitors.

KEYWORDS

Drug design;
Pharmacophore;
MMPs;
3D-QSAR;
Matrix metalloproteinase.

INTRODUCTION

Matrix metalloproteinase (MMPs) belong to the family of zinc endopeptidases[1] collectively referred as metzincins. The metzincin super family is distinguished by a highly conserved motif containing three histidines that bind to zinc at the catalytic site and a conserved methionine that sits beneath the active site[2]. The metzincins are subdivided into four multigene families: seralysins, astacins, ADAMs/adamalysins, and MMPs. Although our knowledge of the metzincin biology is rapidly expanding, we still do not fully understand how these enzymes regulate biological functions. MMP family is comprised of more than 20 related zinc-dependent enzymes that share common functional domains. These enzymes are generally expressed in very low amounts and their transcription is tightly regulated either positively or negatively by cytokines and growth factors such as interleukins (IL-1, IL-4, IL-6), transforming growth factors (EGF, HGF, TGFß), or tumor necrosis factor alpha (TNFα)[3]. Some of these regulatory molecules can be proteolytically activated or inactivated by MMPs. The MMPs dealing with various aspects of their physiologic and pathologic roles in biologic processes. In cancer, special emphasis was initially placed on the degradation of type IV collagen, a major protein component of basement membranes by MMP2 and MMP9. Subsequently, it has been demonstrated that many non-

ECM proteins can be cleaved by selected MMPs. MMPs also play a role in pathological conditions involving untimely and accelerated turnover of ECM in rheumatoid arthritis, osteoarthritis, periodontitis, autoimmune blistering disorders of the skin, dermal photoaging, and chronic ulceration[4]. Both macromolecular inhibitors (natural TIMPs and monoclonal antibodies) and small molecules (synthetic and natural products) have been considered as potential therapies for diseases in which excess MMP activity has been implicated[5-7]. However, technical difficulties with the biotechnological production of macromolecular proteins and limited patient compliance because of parenteral administration have often been cited as limitations for their development. Nevertheless, monoclonal-antibody derivatives[8] are promising drugs to be used as therapeutics, especially if a high MMP specificity is required[9]. TIMPs that have affinities for MMPs in the Pico molar range seem ideal inhibitors but they lack good selectivity and possess other biological functions, which could lead to side effects[10]. Therefore, inhibition of MMPs have several interest and for such purpose the Hydroxamate-based MMPs inhibitors were largely been studied.

The requirement for a molecule to be an effective synthetic inhibitor of the MMP class of enzymes, is a functional group capable of chelating the active site Zn (II) ion (this group is known as zinc binding group, ZBG), and at least one functional group which provides an interaction with the enzyme backbone, and one more side chain which undergo effective van der Waals interactions with the enzyme sub sites. The in silico techniques provided a coast effective window. There are many techniques including QM based QSAR, graph theory based QSAR as well as 3D QSAR is our practice[11-16]. The former interaction were largely been studied with quantum chemical calculations, here we are focused to the later kind of interactions. The other possiable interaction were considered using pharmacophore modeling by PHASE module of Schrödinger package, which significantly help understand the responsible factors affecting the inhibition.

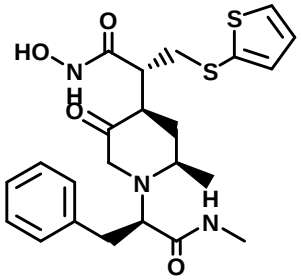
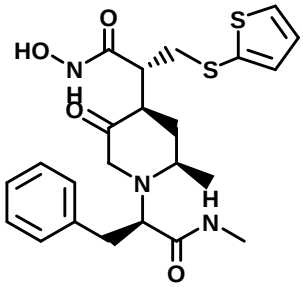
MATERIAL AND METHODS

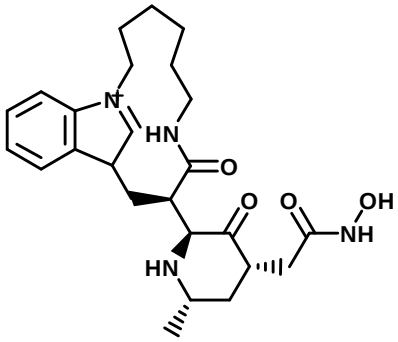
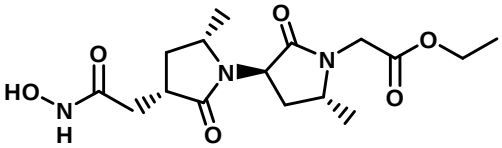
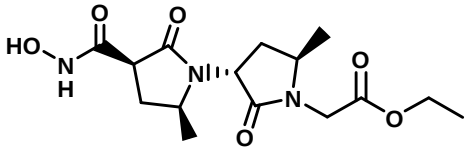
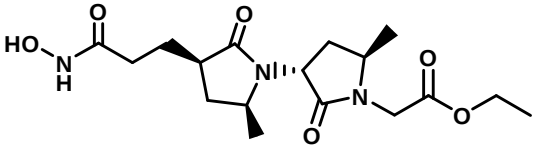
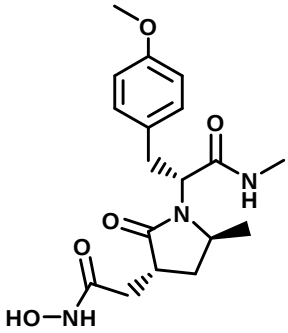
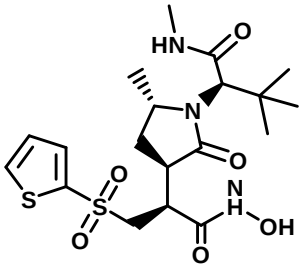
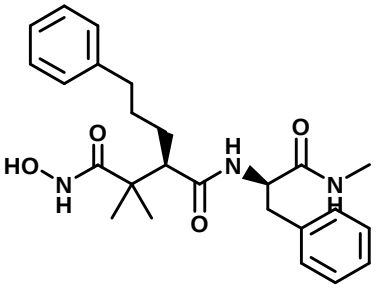
Twenty-two molecules have been taken from literature with their biological activities in terms of IC_{50} values. The IC_{50} values, (i.e., the concentration (μM) of inhibitor that produces 50% inhibition to MMPs) were converted into molar concentration then into pIC_{50} values as reported in TABLE 1.

Generation of the common pharmacophore hypothesis (CPHs)

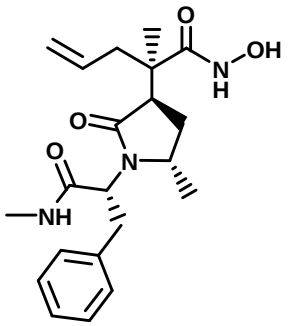
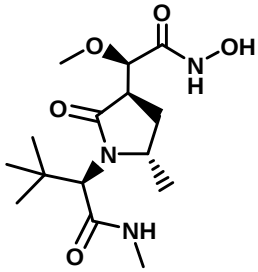
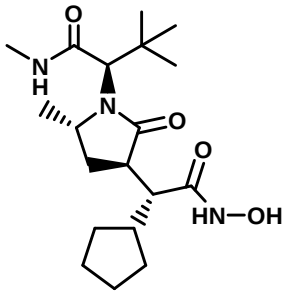
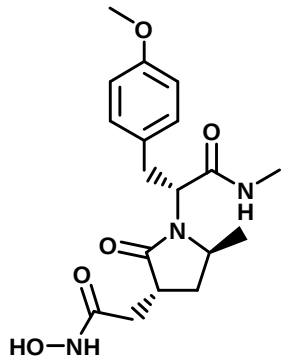
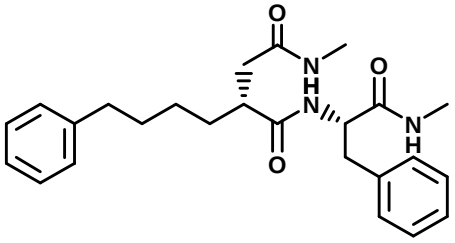
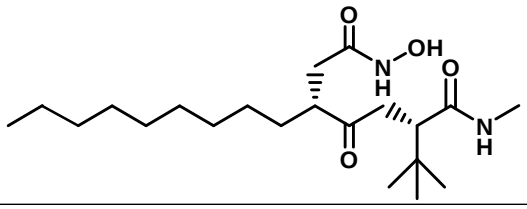
The common pharmacophore hypotheses were generated using PHASE[17]. Conformers generated using MCMM/LMOD with OPLS-2005 force field

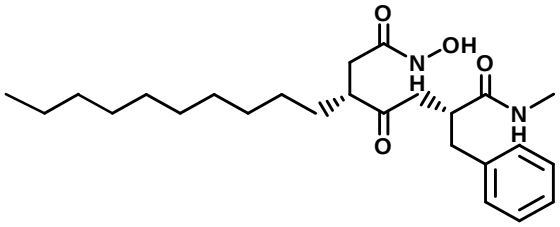
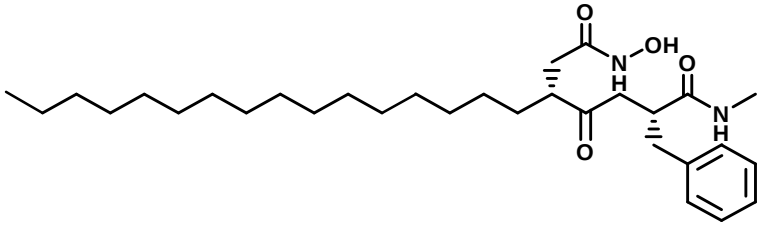
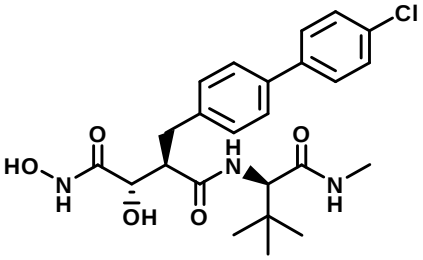
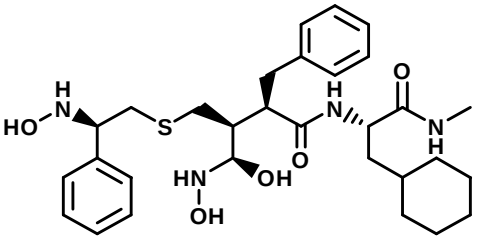
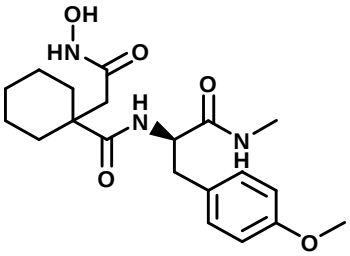
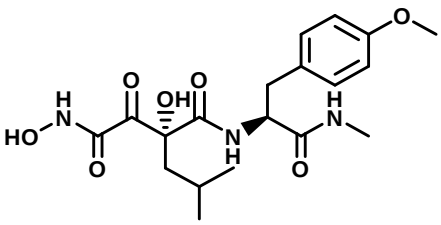
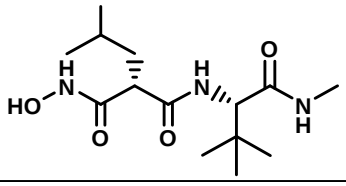
TABLE 1 : The molecular structure and observed activities of MMps.

No	Structure	IC50(nM)	pIC50 (M)	Conformer
1		10	8	138
2		20	7.7	139

No	Structure	IC50(nM)	pIC50 (M)	Conformer
3		2	8.7	1
4		3.1	8.51	54
5		6	8.22	23
6		4	8.4	104
7		5	8.3	61
8		5	8.3	118
9		3	8.52	221

Full Paper

No	Structure	IC50(nM)	pIC50 (M)	Conformer
10		5.4	8.27	91
11		11	7.96	30
12		3	8.52	33
13		1.1	8.96	64
14		13	7.89	277
15		10	8	279

No	Structure	IC50(nM)	pIC50 (M)	Conformer
16		6	8.22	344
17		24	7.62	363
18		2.5	8.6	169
19		25	7.6	431
20		15	7.82	106
21		10	8	238
22		5	8.3	40

The IC50 values were collected from literature and were converted into molar unit and pIC50 using formula pIC50=-logIC50. The conformers were generated using OPLS2005 force field with Monte Carlo multiple minima Low MOD (MCMM/LowMOD) technique.

and a set of conformers of each molecule, with maximum energy difference of 10 kcal/mol relative to global energy minima were retained. Pharmacophore features; hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P), and aromatic ring (R), were defined by three chemical structure patterns namely point, vector and group as SMARTS queries. These were assigned as one of three possible geometries, which define physical characteristics of the site. Point is known as the site located on a single atom in the SMARTS query. Vector is located on a single atom in the SMARTS query, and assigned directionality according to one or more vectors originating from the atom.

In addition, group is located at the centroid of a group of atoms in the SMARTS query. For aromatic rings, the site includes directionality, defined by a vector that is normal to the plane of the ring.

The final size of pharmacophore box was 1Å, which governs the tolerance on matching; the smaller the box size, the more closely pharmacophores must match. Any single pharmacophore in the group could ultimately become a CPH. The analyses indicate that maximum five sites can match to all 22 molecules. These CPHs examined using a scoring function to yield the best alignment of the active ligands and quality of alignment measured by a survival score, which defined as equation 1:

$$S = W_{site}S_{site} + W_{vec}S_{vec} + W_{vol}S_{vol} + W_{sel}S_{sel} + W^{m}_{rew} \quad (1)$$

where W_{Ds} are weights and S_{Ds} are scores, S_{Site} represents an alignment score, the root mean square deviation at the site point position. S_{vec} represents vector score, and averages cosine of the angles formed by corresponding pairs of vector features in aligned structures. S_{vol} represents volume score based on overlap of Van der Waals models of non-hydrogen atoms in each pair of structures. S_{sel} represents the selectivity score,

TABLE 2 : The details of pharmacophore hypotheses.

In	ID	Survival	Site	Vector	Volume	Selectivity	In	ID	Survival	Site	Vector	Volume	Selectivity
1	AAAAD.342	2.857	0.610	0.869	0.375	1.042	25	AAADH.418	2.675	0.500	0.850	0.321	1.253
2	AAAAD.294	2.857	0.620	0.866	0.374	1.042	26	AAAAD.304	2.668	0.470	0.887	0.309	1.033
3	AAAAD.309	2.820	0.570	0.901	0.347	1.057	27	AADDH.182	2.667	0.490	0.862	0.311	1.434
4	AAADD.77	2.811	0.520	0.914	0.380	1.185	28	AAADD.29	2.665	0.500	0.868	0.295	1.182
5	AAADH.1053	2.793	0.530	0.838	0.427	1.379	29	AAAAD.293	2.661	0.480	0.861	0.316	1.039
6	AAADD.76	2.771	0.520	0.904	0.350	1.192	30	AAAAD.341	2.659	0.480	0.864	0.312	1.039
7	AAAAD.297	2.750	0.540	0.875	0.339	1.026	31	AAAAH.471	2.645	0.450	0.865	0.327	1.416
8	AAADD.79	2.728	0.470	0.908	0.348	1.206	32	AADDH.2	2.638	0.470	0.853	0.310	1.374
9	AAADD.83	2.723	0.500	0.879	0.340	1.175	33	AADDH.60	2.638	0.480	0.846	0.309	1.374
10	AAADD.63	2.719	0.510	0.882	0.332	1.167	34	AAADH.107	2.627	0.490	0.882	0.250	1.277
11	AADDH.21	2.717	0.530	0.873	0.317	1.426	35	AADDH.176	2.626	0.460	0.870	0.295	1.436
12	AADDH.79	2.717	0.530	0.873	0.317	1.426	36	AADDH.191	2.624	0.450	0.853	0.322	1.451
13	AAAAD.308	2.710	0.530	0.873	0.308	1.036	37	AADDH.184	2.623	0.430	0.868	0.322	1.437
14	AAADD.35	2.710	0.520	0.874	0.312	1.201	38	AAAAH.475	2.615	0.440	0.840	0.332	1.411
15	AAAAH.474	2.706	0.490	0.863	0.348	1.408	39	AAAAH.476	2.608	0.460	0.850	0.295	1.415
16	AAADH.21	2.697	0.530	0.879	0.290	1.236	40	AADDH.22	2.600	0.420	0.848	0.336	1.431
17	AAADH.219	2.697	0.530	0.879	0.290	1.236	41	AADDH.80	2.600	0.420	0.848	0.336	1.431
18	AAADH.417	2.697	0.530	0.879	0.290	1.236	42	AAADD.30	2.597	0.470	0.839	0.283	1.183
19	AAADH.29	2.697	0.530	0.874	0.292	1.236	43	AAADD.67	2.592	0.390	0.853	0.346	1.199
20	AAADH.227	2.697	0.530	0.874	0.292	1.236	44	AAADD.45	2.592	0.390	0.853	0.346	1.199
21	AAADH.425	2.697	0.530	0.874	0.292	1.236	45	AAAAH.492	2.541	0.380	0.845	0.315	1.435
22	AAAAD.295	2.688	0.460	0.861	0.365	1.009	46	AAAAH.482	2.484	0.350	0.828	0.306	1.392
23	AAAAD.310	2.683	0.480	0.880	0.326	1.065	47	AAAAH.484	2.458	0.310	0.797	0.348	1.435
24	AAADH.30	2.675	0.500	0.851	0.321	1.253	48	AAAAH.491	2.399	0.390	0.739	0.268	1.400

and accounts for what fractions of molecules are likely to match the hypothesis regardless of their activity toward a receptor. Weights are user adjustable. W_{site}, W_{vec}, W_{vol}, and W_{rew} have a default value of 1.0 while W_{sel} has a default value of 0.0, so that a useful hypothesis are not missed out. W_{rew}^{m} represents the reward weights, where *m* is the number of actives that match the hypothesis minus one. In the hypothesis generation, all default values used.

Assessment of significant CPH using partial least square analysis (PLS)

The evaluation of generated CPHs was performed by correlating the observed and estimated activity for 22 molecules. PLS analyses were performed using grid spacing of 1Å. The CPH of best predictivity and significant statistics selected for QSAR study.

Quantum chemical descriptors

Partial least square (PLS) analysis and validation of QSAR models

To derive 3D-QSAR models, pharmacophore based descriptors were used as independent variables and the pIC_{50} values as the dependent variable. PLS method[18,19] was used to linearly correlate these descriptors to the activity. The cross validation analysis performed using the leave one out (LOO) method in which one compound removed from the data set and its activity predicted using the model derived from the rest of the dataset. The cross-validated correlation coefficient (q^2) that resulted in optimum number of components and lowest standard error of prediction considered for further analysis and calculated using equation 2:

$$q^2 = 1 - \frac{\sum_{y}(y_{pred} - y_{observed})^2}{\sum_{y}(y_{observed} - y_{mean})^2} \quad (2)$$

where, γ_{pred}, γ_{actual} and γ_{mean} are predicted, actual and mean values of the target property (pIC_{50}), respectively, and PRESS is the sum of predictive sum of squares. The non-cross-validated PLS analyses were performed with a column filter value of 2.0, to reduce analysis time with small effect on the q^2 values. To have robustness and statistical confidence of the derived models, bootstrapping analysis was used for 100 runs.

RESULTS AND DISCUSSION

The all 22 molecules were used to identify the common pharmacophore (CPH) with tree based partition algorithms. Maximum five features were allowed to de-

TABLE 3 : The regression summary of pharmacophore based best-fitted QSAR models

NO.	ID	n	SD	r^2	F	P	RMSE	q^2
1	AAAAH	1	0.26	0.52	20	0.00029	0.2181	-1.11
		2	0.18	0.8	34.2	1.09E-06	0.1224	0.333
		3	0.13	0.89	45.2	4.86E-08	0.0923	0.621
		4	0.11	0.94	55.2	8.56E-09	0.0433	0.917

TABLE 4 : The angular details of all the pharmacophoric features.

Entry	Site1	Site2	Site3	Angle
1	A4	A2	A5	45.2
2	A4	A2	A6	36.6
3	A4	A2	H12	41
4	A5	A2	A6	13.7
5	A5	A2	H12	32.8
6	A6	A2	H12	20.4
7	A2	A4	A5	98.4
8	A2	A4	A6	127.1
9	A2	A4	H12	115.2
10	A5	A4	A6	32.4
11	A5	A4	H12	49
12	A6	A4	H12	29.7
13	A2	A5	A4	36.4
14	A2	A5	A6	148.3
15	A2	A5	H12	100.1
16	A4	A5	A6	117.7
17	A4	A5	H12	83.9
18	A6	A5	H12	50.5
19	A2	A6	A4	16.3
20	A2	A6	A5	18
21	A2	A6	H12	46.7
22	A4	A6	A5	30
23	A4	A6	H12	48.6
24	A5	A6	H12	62.5
25	A2	H12	A4	23.8
26	A2	H12	A5	47.1
27	A2	H12	A6	112.9
28	A4	H12	A5	47
29	A4	H12	A6	101.6
30	A5	H12	A6	67

TABLE 5 : The distance details of all the pharmacophoric features.

Entry	Site1	Site2	Distance
1	A2	A4	3.494
2	A2	A5	5.83
3	A2	A6	9.928
4	A2	H12	7.841
5	A4	A5	4.185
6	A4	A6	7.42
7	A4	H12	5.686
8	A5	A6	4.484
9	A5	H12	4.318
10	A6	H12	3.758

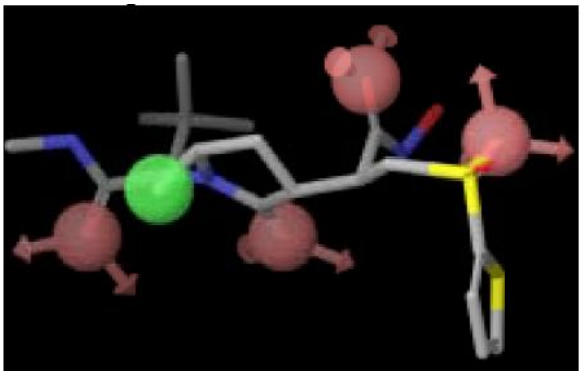


Figure 1 : Perception of CPHs (AAAAH) with compound 8

velop hypotheses and a number of CPHs were reported common in all 22 molecules using features like hydrogen bond acceptor (A), hydrogen bond donor (D), ring (R), hydrophobic (H), negative center (N) and positive centre (P). There were 10 hypotheses based on combination AAAAD, 8 hypotheses based on AAAAH, 10 hypotheses based on AAADD, 10 hypotheses based on AAADH and 10 hypotheses based on AADDH. The variation of biological activity was statistically correlated as a function of quantitative information in terms of different score derived from these common feature. PLS analysis was conducted using up to four factors with default grid spacing. The 192 different regression models based on 48 hypotheses were derived. The best-fitted model was based on features AAAAH. The regression summary is reported in TABLE 3. The specific angel and distances to these pharmacophore features were reported in TABLE 4 and 5 respectively. The figure 1 showed the perception of pharmacophore with template molecule-8.

The red ball showing hydrogen bond acceptor site while the green ball demonstrates the hydrophobic feature as pharmacophore. The figure 2 displayed the three dimensional molecular alignment based on best fitted CPHs. The same model was subsequently used for prediction of activities of molecules as reported in TABLE 6. The trend of predicted and observed activity with different PLS components for best fitted CPH as shown in figure 3(a-d).

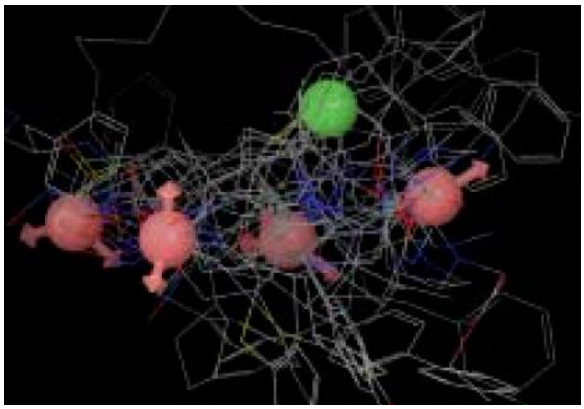
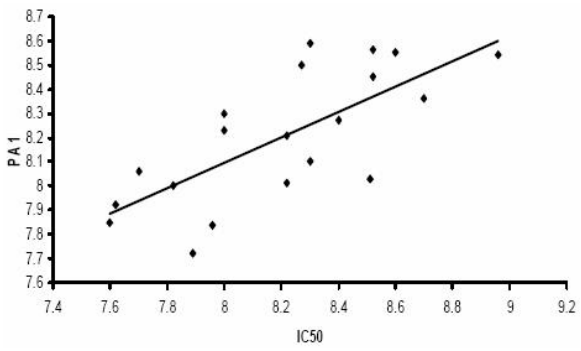


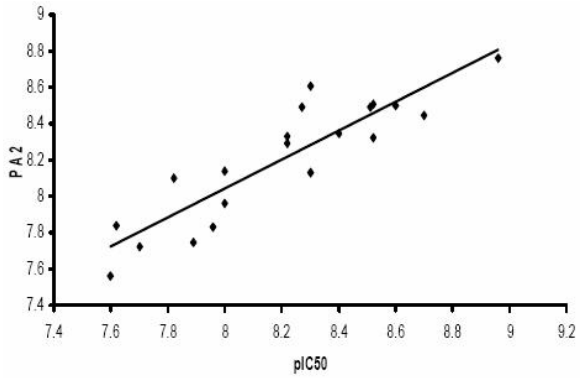
Figure 2 : Pharmacophore based molecular alignment used for ligand based 3D-QSAR

TABLE 6 : The observed and predicted activities by best-fitted CPHs based QSAR model.

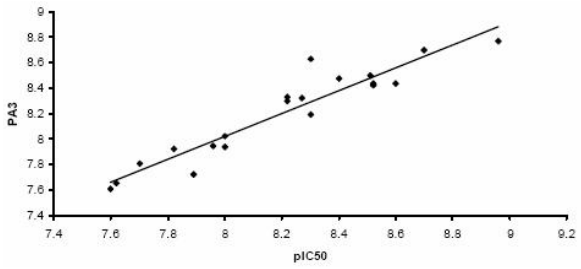
No.	pIC_{50}	PA1	PA2	PA3	PA4
1	8.00	8.23	8.14	8.02	8.04
2	7.70	8.06	7.72	7.81	7.72
3	8.70	8.36	8.45	8.70	8.76
4	8.51	8.03	8.49	8.50	8.50
5	8.22	8.01	8.29	8.30	8.28
6	8.40	8.27	8.35	8.48	8.45
7	8.30	8.59	8.61	8.63	8.54
8	8.30	8.10	8.13	8.19	8.14
9	8.52	8.56	8.51	8.42	8.57
10	8.27	8.50	8.49	8.32	8.26
11	7.96	7.84	7.83	7.95	7.95
12	8.52	8.45	8.32	8.44	8.36
13	8.96	8.54	8.76	8.77	8.81
14	7.89	7.72	7.75	7.72	7.85
15	8.00	8.30	7.96	7.94	7.92
16	8.22	8.21	8.33	8.33	8.29
17	7.62	7.92	7.84	7.65	7.66
18	8.60	8.55	8.50	8.44	8.54
19	7.60	7.85	7.56	7.61	7.69
20	7.82	8.00	8.10	7.92	7.79
21	8.00	8.11	7.90	7.97	8.02
22	8.30	8.01	8.16	8.17	8.24



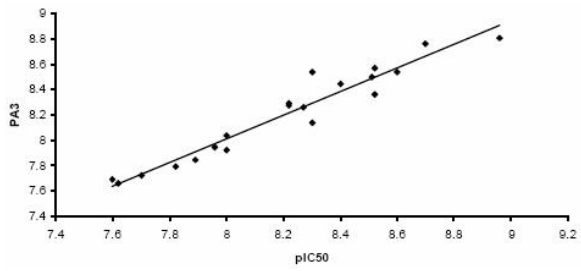
(A) Best model with one component



(B) Best model with two components



(C) Best model with three components



(D) Best model with four components

Figure 3 : Trend of observed and predicted activities using pharmacophore based fitted models (A-D corresponds to number of components 1-4 respectively)

This is clear from figure 3 (D) the best fitted CPH (AAAAH) derived model-4 based on 4 PLS components has a significant correlation with observed activity. Structural data of MMPs revealed that the S1'. Subsite is a deep pocket for several of the enzymes, including MMP2, MMP3, MMP8 and MMP9, but is occluded for MMP1 and MMP7. At the same time It is proven that in MMP1 inhibition, the hydroxamate oxyanion forms a strong, short hydrogen bond to the carboxylate oxygen of the catalytical Glu219 . Also contributing to the binding is another hydrogen bond between the hydroxamate-NH group and the carbonyl oxygen of Ala182. Thus, a set of interactions is known to be involved by hydrogen bonding.

CONCLUSION

In our study, we built a pharmacophore model applying ligand based pharmacophore generation approach, using PHASE. Different pharmacophore base QSAR models have been developed by using PLS analysis. The resulting best hypothesis is consisted of five features: four hydrogen bond acceptor and one-hydrophobic site. The alignment rule of best-fitted model ($q^2 = 0.80$, $r^2 = 0.90$ and $R = 0.91$) was used to develop ligand based QSAR model. The pharmacophore-based study indicates a possible hydrophobic and hydrogen bond acceptor interaction of ligand to MMPs. The study helpfully indicates a possible common structural feature, which might be involved in the inhibition of MMPs. The refinement of inhibitor designing against MMPs using X-ray structure based information is on the way and we expect to include in further communications.

REFERENCES

[1] J.L.Hu, P.E.Van den Steen, Q.X.A.Sang, G.Opdenakker; Nature Reviews Drug Discovery, **6(6)**, 480-498 **(2007)**.

[2] W.Stocker, W.Bode; Current Opinion in Structural Biology, **5(3)**, 383-390 **(1995)**.

[3] M.D.Sternlicht, Z.Werb; Annual Review of Cell and Developmental Biology, **17**, 463-516 **(2001)**.

[4] M.Egeblad, Z.Werb; Nature Reviews Cancer, **2(3)**, 161-174 **(2002)**.

[5] M.Whittaker, C.D.Floyd, P.Brown, A.J.H.Gearing; Chemical Reviews, **101(7)**, 2205-2205 **(2001)**.

[6] J.W.Skiles, N.C.Gonnella, A.Y.Jeng; Current Medicinal Chemistry, **8(4)**, 425-474 **(2001)**.

[7] M.Whittaker, C.D.Floyd, P.Brown, A.J.H.Gearing; Chemical Reviews, **99(9)**, 2735-2776 **(1999)**.

[8] J.L.Hu, P.E.Van den Steen, M.Houde, T.T.Ilenchuk, G.Opdenakker; Biochemical Pharmacology, **67(5)**, 1001-1009 **(2004)**.

[9] E.Martens, A.Leyssen, I.Van Aelst, P.Fiten, H.Piccard, J.Hu, F.J.Descamps, P.E.Van den Steen, P.Proost, J.Van Damme, G.M.Liuzzi, P.Riccio, E.Polverini, G.Opdenakker; Biochimica Et Biophysica Acta-General Subjects, **1770(2)**, 178-186 **(2007)**.

[10] R.Chirco, X.W.Liu, K.K.Jung, H.R.C.Kim; Cancer and Metastasis Reviews, **25(1)**, 99-113 **(2006)**.

[11] P.P.Singh, H.K.Srivastava, F.A.Pasha; Bioorganic & Medicinal Chemistry, **12(1)**, 171-177 **(2004)**.

[12] F.A.Pasha, H.K.Srivastava, A.Srivastava, P.P.Singh; Qsar & Combinatorial Science, **26(1)**, 69-84 **(2007)**.

[13] F.A.Pasha, H.K.Srivastava, P.P.Singh; International Journal of Quantum Chemistry, **104(1)**, 87-100 **(2005)**.

[14] F.A.Pasha, H.K.Srivastava, P.P.Singh; Bioorganic & Medicinal Chemistry, **13(24)**, 6823-6829 **(2005)**.

[15] V.K.Sahu, A.K.R.Khan, R.K.Singh, P.P.Singh; International Journal of Quantum Chemistry, **109(6)**, 1243-1254 **(2009)**.

[16] R.K.Singh, A.K.R.Khan, V.K.Sahu, P.P.Singh; International Journal of Quantum Chemistry, **109(2)**, 185-195 **(2009)**.

[17] Phase; Schrödinger, LLC, New York, NY, version 8.0, **(2007)**.

[18] P.Geladi, Y.L.Xie, A.Polissar, P.Hopke; Journal of Chemometrics, **2**, 231 **(1998)**.

[19] S.Wold, A.Ruhe, H.Wold, W.J.Dunn; Siam Journal on Scientific and Statistical Computing, **5(3)**, 735-743 **(1984)**.