



# Potent inhibitors precise to S1' loop of MMP-13, a crucial target for osteoarthritis



Sukesh Kalva, K. Saranyah, P. Rathi Suganya, M. Nisha, Lilly M. Saleena\*

Department of Bioinformatics, SRM University, SRM Nagar, Kattankulathur, Kancheepuram District, Chennai 603203, India

## ARTICLE INFO

### Article history:

Accepted 23 June 2013

Available online 4 July 2013

### Keywords:

MMP-13

Pharmacophore

Docking

QSAR

Zinc database

E-model energies

## ABSTRACT

Matrix metalloproteinase-13 (MMP-13) is the primary MMP involved in cartilage degradation through its particular ability to cleave type-II collagen. This protein is expressed by chondrocytes and synovial cells in human osteoarthritis and rheumatoid arthritis; hence, it is an attractive target for the treatment of arthritic diseases. Currently available inhibitors lack specificity for metalloproteinase because of a common Zn binding site in MMPs; thus, there is a need to identify selective MMP-13 inhibitors for osteoarthritis therapy. Because selectivity is the major concern, both ligand-based and protein-based pharmacophore methodologies were used to identify potent and selective MMP-13 inhibitors. Different hypotheses were validated, and the best hypothesis was used to screen Zinc (natural and chemical) databases to seek novel scaffolds as MMP-13 inhibitors. The identified hits were validated using different strategies, such as Glide Standard precision, extra precision, E-model energies and receiver operating curve (ROC). In addition, potent inhibitors were selected based on two criteria: a similar binding mode as that of the active site PB3 crystal ligand and crucial amino acid interactions that are catalytically important for the function of MMP-13. The candidate potent inhibitors ZINC 02535232, ZINC 08399795, ZINC 12419118 and ZINC 00624580 nearly reproduced the H-bond interactions formed in the crystal structure of 1XUC with reasonable RMSD values exhibiting a novel interaction pattern that was not previously observed in MMP-13 inhibitors. The identified potent hits with diverse chemical scaffolds may be useful in designing new MMP-13 inhibitors.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

Osteoarthritis (OA), which is characterized by the breakdown of joint cartilage, chronic joint pain, and inflammation, affects millions of people worldwide [1–3]. This disease is characterized by a progressive loss of articular cartilage that eventually leads to denudation of the joint surface. The cartilage loss observed in OA is a complex process that results in the degradation of various components of the cartilage matrix [4]. Matrix metalloproteinases (MMPs) play a key role in the degradation of cartilage extracellular matrix (ECM) components. In human osteoarthritis, many MMPs, including MMP-1, MMP-2, MMP-3, MMP-7, MMP-8, MMP-9, MMP-13 and MT1-MMP, are expressed in articular cartilage [5]. Among the various matrix metalloproteinases, MMP-13 is thought to be the most important for the degradation of collagen within cartilage due to its preferential digestion of type II collagen over types I and III [6]. MMP-13 is specifically expressed in the cartilage of human OA patients and is not present in normal adult cartilage

[7]. Data from MMP-13 transgenic animals indicate that MMP-13 induces joint abnormalities that are characteristic of human OA because MT1-MMP efficiently activates proMMP-2 and proMMP-13 in osteoarthritic cartilage [8]. The increased levels of MMP-13 result in an imbalance in the regulation of MMPs by tissue inhibitors of metalloproteinases (TIMPs), thus likely contributing to the diseased state [9].

Many peptidomimetic, hydroxamate, non-hydroxamate sulfonamide inhibitors of MMP-13 have been identified, but they failed to serve as a remedy because they cause parallel issues, such as a painful joint stiffening side effect, which is termed musculoskeletal syndrome, and lack selectivity. The low specificity of these inhibitors is because the active site is highly conserved among MMPs. This failure is mainly attributed to the complex biology of MMPs, which was unexplored during the early period of MMP inhibitor discovery [10,11].

The structure of MMP-13 consists of a prodomain, a catalytic domain, a hinge region, and a hemopexin domain [12]. MMP-13 has a conserved active site motif in the catalytic domain where a tris (histidine)-bound zinc (II) acts as a catalytic site for substrate hydrolysis. The active site motif in MMP-13 is highly conserved in other MMPs with the exception of the S1' loop region [13]. The

\* Corresponding author. Tel.: +91 44 27417813.

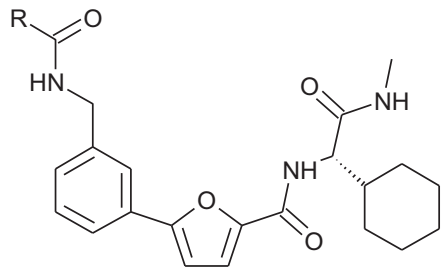
E-mail address: [saleena.m@ktr.srmuniv.ac.in](mailto:saleena.m@ktr.srmuniv.ac.in) (L.M. Saleena).

S1' loop region for MMP-13 is located at residues 244–255. The size and amino acid residues of the S1' loop in MMP-13 differ from other MMPs [14]. The long loop size of MMP-13 has been suggested to be a determining factor for the selective binding of inhibitors of MMP-13, while in other MMPs, the S1' loop is too shallow and narrow. The S1' pocket in MMP-1 is smaller than those of MMP-8 and MMP-13, which have a long and open channel that can accommodate large P1 groups [13]. Thus, targeting the S1' loop region helps in identifying MMP-13-specific inhibitors. To achieve this, a novel virtual screening procedure, including ligand- and protein-based pharmacophore modeling, molecular docking studies were performed. Candidate novel inhibitors were selected based on their amino acid interaction and binding mode with the specific residues Lys 249, Phe 252, and Met 253 in the S1' loop of MMP-13.

## 2. Materials and methods

### 2.1. Generation of pharmacophore models: ligand-based or common feature approaches

For pharmacophore modeling, a set of 32 human MMP-13 non-zinc binding inhibitors with activity values ranging from  $pIC_{50}$  5.481 to 8.699 nM were collected based on the literature (Fig. 1) [15]. All compounds used in this study were subjected to minimization with the Ligprep module of the Schrödinger software [16]. The conformations of the above structures were generated using MMFFs force fields with an implicit GB/SA solvent model. A maximum of 1000 conformers were generated per structure using a preprocess minimization of 1000 steps with the ConfGen algorithm.



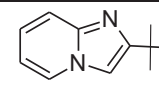

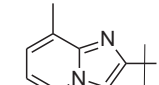
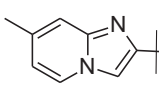







S.No	Compound	R	IC <sub>50</sub>	pIC <sub>50</sub>
1	3 <sup>a</sup>		100	7.00
2	5 <sup>a</sup>		3300	5.48
3	6 <sup>a</sup>		110	6.95
4	7 <sup>a</sup>		140	6.85
5	8		75	7.12
6	9 <sup>a</sup>		11	7.95
7	10		180	6.74
8	11		330	6.48
9	12 <sup>a</sup>		310	6.50
10	13		24	7.62
11	14		72	7.14

Fig. 1. Chemical structures and biological activity (IC<sub>50</sub> and pIC<sub>50</sub>) values for the MMP-13 inhibitors.

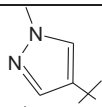
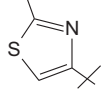


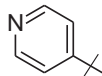
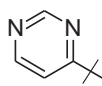

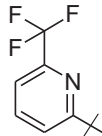

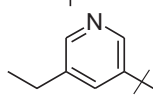
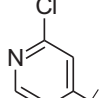
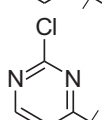
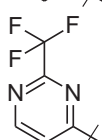
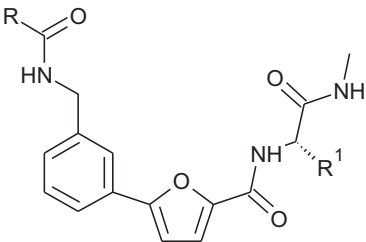
12	15		140	6.85
13	16 <sup>a</sup>		8	8.09
14	17 <sup>a</sup>		8	8.09
15	18 <sup>a</sup>		42	7.37
16	19 <sup>a</sup>		31	7.50
17	20 <sup>a</sup>		10	8.00
18	21 <sup>a</sup>		4	8.39
19	22		9	8.04
20	23		100	7.00
21	24		29	7.53
22	25		27	7.56
23	26		3	8.52
24	27		7	8.15

Fig. 1. (continued)

Quantitative pharmacophores were generated using a set of the above molecules based on the diversity of their chemical structures and biological activity against human MMP-13 inhibitors. During the search, hydrogen-bonding interactions were suppressed because conformations in which ligands bind to the receptor are needed in the model, not just conformations with internal hydrogen bonding.

The development of quantitative pharmacophore models was performed using the PHASE software [17]. Active and inactive  $pIC_{50}$  thresholds of 8.00 and 7.00, respectively, were applied to yield 12 actives and 9 inactives, and the remaining were moderately active ligands. A set of pharmacophore features were produced

using the create sites option of the Develop pharmacophore model, which creates a site point for each conformer of the above ligands. A default setting including acceptor (A), donor (D), hydrophobic (H), negative (N), positive (P), and aromatic ring (R) features was used to create pharmacophore sites. Using the sites defined in the previous step, pharmacophores common to the active set of ligands were generated. PHASE uses a tree-based partitioning algorithm that places pharmacophore configurations in multidimensional boxes and groups together according to their intersite distances. Common pharmacophores were identified from a set of variants, which is a feature-type set that defines potential pharmacophores. Common pharmacophore hypotheses for the active



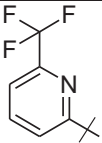
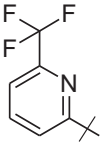
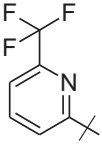
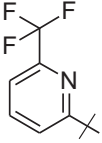
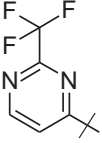
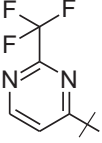
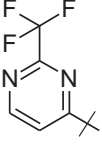
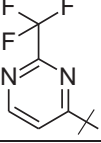
Compound	R	R <sup>1</sup>	IC <sub>50</sub> (nM)	pIC <sub>50</sub> (nM)
25	28		6	8.22
26	29 <sup>a</sup>		28	7.55
27	30 <sup>a</sup>		2	8.69
28	31 <sup>a</sup>		25	7.60
29	32		9	8.04
30	33		36	7.44
31	34		2	8.69
32	35 <sup>a</sup>		5	8.30

Fig. 1. (continued)

ligands were scored using the survival and post hoc scores. Survival scores were calculated by mapping the generated common pharmacophore models with highly active molecules by considering various geometric and heuristic factors, such as site, vector and volume scores. Higher survival scores resemble better mapping of pharmacophores with active molecules. In addition to the survival score, we have generated a post hoc score to validate the best pharmacophore hypothesis.

## 2.2. Generation of pharmacophore models: protein or structure-based approaches

The MMP-13 complex structure (PDB ID: 1XUC with a resolution of 1.70 Å) (<http://www.pdb.org>) was used for protein-based pharmacophore and docking studies. A protein-based pharmacophore model was generated using the Energy-based pharmacophore (E-pharmacophore) module of Schrodinger version 9.2. Using

excluded volumes in the receptor active site, crucial interactions between the ligand and amino acid residues were converted into pharmacophore features.

Energy-based pharmacophores are pharmacophoric features that are generated by using the protein–ligand energetic terms computed by the Glide XP scoring function [18]. The Glide scoring function involves ligand-based pharmacophores and protein-based docking to create energetically optimized protein-based pharmacophores. Energy-based pharmacophores were performed using the E-pharmacophore option in the Schrödinger software. Docked pose viewer file was given as input where the “Single ligand mode” and “receptor based excluded volumes” options were chosen to generate a hypothesis. Pharmacophoric sites were generated, scores were assigned and the pharmacophoric features were ranked based on the Glide XP scoring terms. The best desired hypotheses were then exported using the option “write hypothesis with selected features.”

### 2.3. Pharmacophore validation

The pharmacophore models selected were then validated by Güner–Henry (GH) scoring methods. This method quantifies hits by recalling the actives from inactives from a database containing a 2060 molecule dataset consisting of known actives and inactives. Of these molecules, 60 molecules are known inhibitors of MMP-13 with diverse activities, while the other 2000 are decoy molecules downloaded from the Schrodinger website. This database of known actives and inactives was screened with ligand- and protein-based pharmacophore models to calculate GH scores. GH scores range from 0, which indicates a null model, to 1, which indicates an ideal model.

The formula for analyzing the hits list by a ligand- and protein-based pharmacophore model-based database search was as follows [19,20]:

$$\%A = \frac{Ha}{A} \times 100$$

$$\%Y = \frac{Ha}{Ht} \times 100$$

$$E = \frac{Ha/Ht}{A/D}$$

$$GH = \left( \frac{Ha(3A + Ht)}{4HtA} \right) \left( 1 - \frac{Ht - Ha}{D - A} \right)$$

where %A is the percentage of known active compounds obtained from the database, *Ha* is the number of actives in the hits list (true positives), *A* is the number of active compounds in the database, %Y is the percentage of known actives in the hits list (recall), *Ht* is the number of hits retrieved, *D* is the number of compounds in the database, *E* is the enrichment of the concentration of actives by the model relative to random screening without a pharmacophoric approach, and GH is the Güner–Henry score.

In addition to common validation methods for ligand- and protein-based pharmacophore models, the common feature-based pharmacophore model was also validated using a 3D-QSAR model. PHASE provides options to build pharmacophore-based 3D-QSAR models for a set of ligands that are aligned to a selected hypothesis [17]. Pharmacophore-based QSAR models are only concerned with the sites in a molecule that can be matched to the hypothesis. Such sites are treated as a sphere with a user-adjustable radius, and categories are assigned according to the feature type (A, D, H, N, P, R, X, Y, and Z). PHASE QSAR models are created by applying partial least-squares (PLS) regression to the pool of binary-value independent variables. The PLS procedure results in the assignment of a

regression coefficient to each bit, which facilitates the identification of specific chemical features that tend to increase or decrease the estimated activity. PHASE creates a series of regression models, progressively incorporating more PLS factors with the maximum number of factors being no larger than 1/3 the number of training set molecules. All molecules were aligned with a common pharmacophore hypothesis (CPH) and validated using PLS analysis to generate a more reliable 3D-QSAR model.

### 2.4. Database search

The ligand- and protein-based pharmacophore models developed using PHASE and E-pharmacophore were used in a 3D structural query to retrieve compounds with novel structural scaffolds from a Zinc database (natural and chemical) containing 6 lakh compounds. All of the compounds used in this study were subjected to an option for conformational sampling using the ConGen algorithm to generate a maximum of 1000 conformations. The molecules associated with these conformations were provided as an input for the pharmacophore search using the Find matches to Hypothesis option from the Schrodinger website. The model was considered best if the database search retrieved all of the inhibitors with good fitness values. The fitness score is a measure of how well the pharmacophoric sites of a ligand match with the hypothesis and how well the conformation superimposes with the vector features (acceptors, donors, and aromatic rings). Compounds with fitness scores above 1.5 were considered good inhibitors.

### 2.5. Docking studies

Three dimensional coordinates from the X-ray crystal structure of MMP-13 in complex with an inhibitor (PDB code: 1XUC) were selected as a receptor model for docking simulations. Before performing the docking, protein preparation was performed using the protein preparation wizard of Maestro version 9.2. Some residues with missing side chain atoms were detected in the protein, these residues were modeled using prime modeler followed by refinement, and H atoms were added to the structure. Water molecules and co-crystallized sulfur atoms were deleted, with the exception of the ligand PB3. Finally, the hydrogen atoms were assigned, and the protein was minimized using the OPLS2005 force field. All of the selected inhibitors were docked into the active site of the receptor using Glide (Schrodinger Inc.) in the SP (standard precision) and XP (Extra precision) modes [21]. The binding region was defined by selecting the co-crystal ligand PB3 (n,n'-bis (3-methylbenzyl) pyrimidine-4,6-dicarboxamide), and a cubic box was centered on the centroid of the crystal ligand. The docking poses were selected based on the SP and XP Glide scores, the E-model energy and interactions formed between ligands and the S1' loop region amino acids of MMP-13 receptor. RMSD is one of the methods used to validate docking programs and scoring functions [22]. Programs that are able to return poses below a preselected RMSD value from a known conformation (1.5–2 Å) are considered to have performed successfully [23]. All of the docked images were captured using the PyMOL software (<http://www.pymol.org/>).

### 2.6. Validation of the docking program by ROC curves

To validate the glide docking program, the docking results were analyzed using a ROC (receiver operating characteristic) curve, which describes the ability of a docking method to avoid false positives and false negatives [24]. A total of 2000 compounds were used as a decoy set from the Schrodinger to validate the Glide docking program [25]. This program consists of drug-like molecules with an average molecular weight of 400 Da. These molecules were downloaded as a 3D file from the Schrodinger website. The candidate hits

**Table 1**

The best common pharmacophore hypotheses with active and inactive survival scores.

Hypothesis	Survival	Survival inactive	Post hoc score	Site	Vector	Volume
AADHR	3.84	1.43	3.84	0.94	0.99	0.91
DHRRR	3.76	1.40	3.76	0.75	0.99	0.89
AHRRR	3.63	1.24	3.63	0.76	0.99	0.87
AAAHH	3.61	1.15	3.61	0.76	0.99	0.85
AADHH	3.57	1.02	3.57	0.77	0.99	0.81

from the docking procedure were seeded into each of the decoy sets. The decoy set with candidate compounds was docked using the Glide program. After ranking the docked decoy sets by score, the enrichment was calculated, and enrichment plots or receiver operating characteristic curves were plotted. ROC curves plot the sensitivity (Se) of a given docking/scoring combination against the specificity (Sp), and areas under the curve can be calculated for comparison. The advantage of ROC curves over enrichment plots is that they are independent of the number of actives and also include sensitivity and specificity information. An ROC curve is a plot of true-positive versus false-positive rates for all compounds. The area under the ROC (AU-ROC) curve is the probability of active compounds being ranked earlier than decoy compounds. A theoretically perfect AUC value is 1.0.

In an ROC curve, a true positive rate (sensitivity) is plotted as a function of the false positive rate (100-specificity) for different cut-off points. Each point on an ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. A test with perfect discrimination (no overlap in the two distributions) has a ROC curve that passes through the upper left corner (100% sensitivity and 100% specificity). Therefore, the closer the ROC curve is to the upper left corner, the higher the overall accuracy of the test.

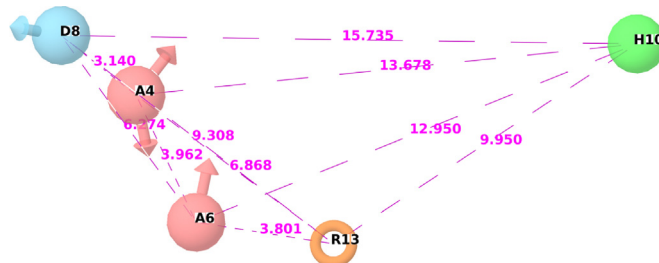
### 2.7. ADME property analysis

To examine the drug likeliness properties for identified hit compounds, we have used the ADME (absorption, distribution, metabolism, and excretion) filtration criteria and the Lipinski rule of five. We analyzed the molecular weight, H-bond acceptors, H-bond donors, QPlogS (aqueous solubility), QPlogPo/W (octanol/water partition coefficient) and QPlogBB (brain blood partition coefficient) for the searched compounds using Qikprop, which predicts physiochemical descriptors of for possible drug compounds. This program predicts properties necessary for the drug design of all of the searched compounds with detailed analysis of the Lipinski rule of five [26].

## 3. Results

### 3.1. Generation of pharmacophore models: ligand-based or common feature approach

The MMP-13 inhibitors with activity values used to generate a pharmacophore model are displayed in Fig. 1. Out of 32 compounds, 12 highly active, 11 moderately active and 9 inactive compounds were selected to generate a common pharmacophore hypothesis. Using a tree-based partition algorithm, PHASE generates 5 featured pharmacophores from a set of 28 different variant lists (AADH, AADDR, DHRRR, AARRR, ADRRR, AADRR, ADDRR, DDHRR, AAAHR, ADDHR, HHRRR, AADHR, AAAHH, AHRRR, ADHRR, AAARR, DDHRR, AAHRR, AHRRR, ADDHH, AAADR, AADHH, DHRRR, AAHRR, ADHRR, AADDH, AAADD, and DDDRR). All of these pharmacophores were mapped using an intersite distance of 2 Å in a terminal box size of 1 Å with 12 highly active molecules. Hypotheses emerging from this process were subsequently scored with respect to the 20 inactives



**Fig. 2.** AADHR, the best hypothesis produced by the PHASE module in Maestro 9.2 software. Pharmacophore features are colored in green, blue, brown and brick red contours represent the hydrophobic feature (H), H-bond donor (D), ring aromatic (R), and H-bond acceptor (A). The pink dotted lined indicates the distance between the pharmacophore features (Å). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of the article.)

**Table 2**

Distances between the different pharmacophoric sites of the AADHR hypothesis (ligand-based method).

Entry	Site1	Site2	Distance
AADHR	A4	A6	3.962
AADHR	A4	D8	3.14
AADHR	A4	H10	13.67
AADHR	A4	R13	6.86
AADHR	A6	D8	6.27
AADHR	A6	H10	12.95
AADHR	A6	R13	3.80
AADHR	D8	H10	15.73
AADHR	D8	R13	9.30
AADHR	H10	R13	9.95

using a weight of 1.0. A summary of five pharmacophore hypotheses generated with five features are listed in Table 1.

### 3.2. Pharmacophore validation (ligand-based model): survival and post hoc score

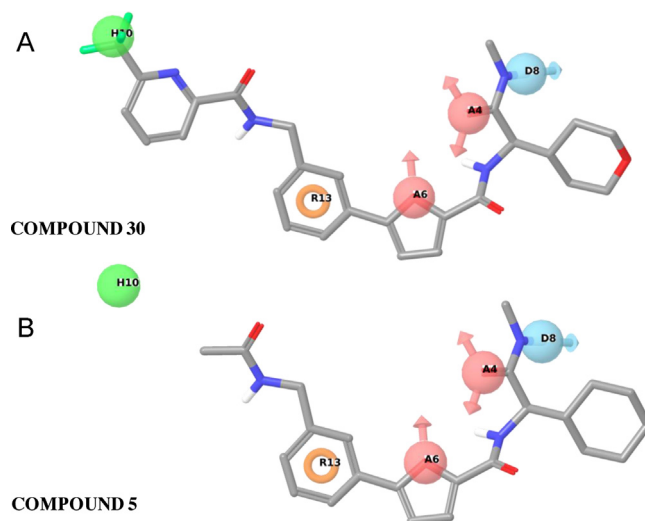
The generated hypotheses were characterized by highest survival and post hoc scores. The survival and post hoc scores for generated 5 different hypotheses ranged from 3.84 to 3.57 (Table 1). Among the 5 hypothesis we had chosen the best hypothesis i.e., AADHR had the highest survival score (3.84). The post hoc score is the result of rescoring, and it is a weighted combination of the vector, site, volume, and selectivity scores. The best hypothesis gave a 3.84 post hoc score, which indicated that the model is highly reliable and more likely to be unique to the actives. The pharmacophore features present in this hypothesis are two H-bond acceptors, one H-bond donor, one hydrophobic and one ring aromatic groups, which are shown in Fig. 2. The distances and angles between the different sites of AADHR are provided in Tables 2 and 3, respectively. The greater the fitness score, the greater the activity prediction of the compound. The fitness scores for all of the active ligands were observed on the best pharmacophore model AADHR. The best hypothesis was superimposed with a highly active compound (30), indicating that all of the features were perfectly mapped and had a fitness score of 3.00 (Fig. 3a), whereas the least active compound (1) was mapped with only 4 features and had a fitness score



**Table 3**  
Angles between different pharmacophoric sites of the AADHR hypothesis.

Entry	Site1	Site2	Site3	Angle
AADHR	A6	A4	D8	123.7
AADHR	A6	A4	H10	71.1
AADHR	A6	A4	R13	27.2
AADHR	D8	A4	H10	126.2
AADHR	D8	A4	R13	133.4
AADHR	H10	A4	R13	44
AADHR	A4	A6	D8	24.6
AADHR	A4	A6	H10	92.1
AADHR	A4	A6	R13	124.4
AADHR	D8	A6	H10	104.4
AADHR	D8	A6	R13	133.5
AADHR	H10	A6	R13	32.4
AADHR	A4	D8	A6	31.7
AADHR	A4	D8	H10	44.6
AADHR	A4	D8	R13	32.4
AADHR	A6	D8	H10	52.8
AADHR	A6	D8	R13	17.2
AADHR	H10	D8	R13	36.6
AADHR	A4	H10	A6	16.8
AADHR	A4	H10	D8	9.3
AADHR	A4	H10	R13	28.6
AADHR	A6	H10	D8	22.7
AADHR	A6	H10	R13	11.8
AADHR	D8	H10	R13	33.9
AADHR	A4	R13	A6	28.4
AADHR	A4	R13	D8	14.2
AADHR	A4	R13	H10	107.4
AADHR	A6	R13	D8	29.3
AADHR	A6	R13	H10	135.8
AADHR	D8	R13	H10	109.5

of 1.9 (Fig. 3b). The selected hypothesis was superimposed with the highly active compound, where the A4 and A6 acceptor features correctly mapped to the oxo group and the furan ring, and



**Fig. 3.** Pharmacophore mapping of the most active and least active compounds on the best hypothesis compound, (a) pharmacophore mapping of the highest active compound (**30**) on the best hypothesis compound from the test set and (b) pharmacophore mapping of the least active compound (**5**) on the best hypothesis compound from the training set.

another D8 donor feature was mapped to the amino group. The hydrophobic groups H10 feature mapped to the trifluoro methyl group and the R13 feature mapped to the benzyl group of the compound. The best hypothesis was validated with high and less active compounds, and results indicated that the generated model correctly mapped with the active compound with the best fitness score.

**Table 4**  
The actual and predicted activity values for the training and test set molecules for MMP-13 inhibitors.

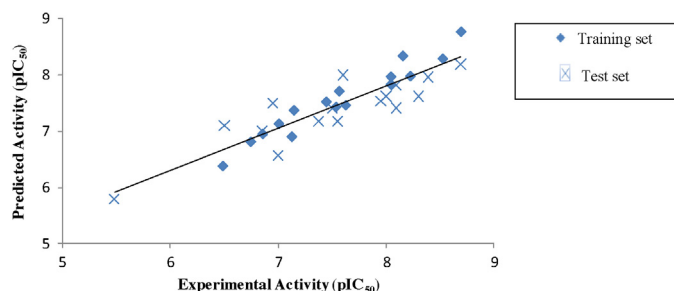
S. no	Compound name	Experimental IC <sub>50</sub> values (nM)	Experimental pIC <sub>50</sub> values (nM)	Predicted IC <sub>50</sub> values (nM)	Residual value
1	3 <sup>a</sup>	100	7.00	6.57	0.43
2	5 <sup>a</sup>	3300	5.48	5.79	−0.31
3	6 <sup>a</sup>	110	6.95	7.5	−0.55
4	7 <sup>a</sup>	140	6.85	7.01	−0.16
5	8	75	7.12	6.91	0.21
6	9 <sup>a</sup>	11	7.95	7.54	0.41
7	10	180	6.74	6.82	−0.08
8	11	330	6.48	6.39	0.09
9	12 <sup>a</sup>	310	6.50	7.1	−0.6
10	13	24	7.62	7.47	0.15
11	14	72	7.14	7.38	−0.24
12	15	140	6.85	6.96	−0.11
13	16 <sup>a</sup>	8	8.09	7.82	0.27
14	17 <sup>a</sup>	8	8.09	7.41	0.68
15	18 <sup>a</sup>	42	7.37	7.17	0.2
16	19 <sup>a</sup>	31	7.50	7.4	0.1
17	20 <sup>a</sup>	10	8.00	7.62	0.38
18	21 <sup>a</sup>	4	8.39	7.96	0.43
19	22	9	8.04	7.84	0.2
20	23	100	7.00	7.14	−0.14
21	24	29	7.53	7.44	0.09
22	25	27	7.56	7.72	−0.16
23	26	3	8.52	8.3	0.22
24	27	7	8.15	8.35	−0.2
25	28	6	8.22	7.99	0.23
26	29 <sup>a</sup>	28	7.55	7.17	0.38
27	30 <sup>a</sup>	2	8.69	8.19	0.5
28	31 <sup>a</sup>	25	7.60	8	−0.4
29	32	9	8.04	7.98	0.06
30	33	36	7.44	7.53	−0.09
31	34	2	8.69	8.78	−0.09
32	35 <sup>a</sup>	5	8.30	7.62	0.68

<sup>a</sup> Test set molecules.

**Table 5**  
Summary of PLS analysis results for the best common pharmacophore hypotheses (CPHs) with survival scores.

CPH	AADHR
Survival score	3.84
SD	0.20
$r^2$	0.94
$F$	50
$q^2$	0.75
RMSE	0.56
Pearson $R$	0.73

SD – standard deviation of the regression;  $r^2$  – correlation coefficient;  $F$  – Fischer test;  $q^2$  –  $q$ -squared cross validation test set; RMSE – root-mean-squared error; Pearson  $R$  – correlation coefficient between the actual and predicted activity for test set molecules.



**Fig. 4.** Relationship between the experimental and predicted MMP-13 inhibitory activity values of the training set and test set molecules using the ligand-based pharmacophore model (AADHR).

### 3.3. Pharmacophore validation (ligand-based model): 3D-QSAR model

The best hypothesis (AADHR) was validated with a pharmacophore-based 3D-QSAR model. During the QSAR model generation, molecules were divided into a training set consisting of 16 compounds and a test set consisting of 16 compounds. The selection of the training and test sets was performed manually so that low, moderate, and high MMP-13 inhibitory activities were all represented. The above hypothesis, i.e., AADHR, showed better statistical results for the  $r^2$  and  $q^2$  values (i.e., 0.94 and 0.75, respectively, with an  $F$  value = 50.0, SD (standard deviation) = 0.20, RMSE (root mean square error) = 0.56 and Pearson  $R$  = 0.73, which indicated that the generated pharmacophore-based 3D-QSAR model was robust enough to be capable of making accurate and reliable predictions for the biological activities of new compounds. The predictive ability of the model was assessed by predicting the activity of test set molecules. A squared predictive correlation coefficient of 0.75 was also observed between the experimental and predicted activity values of the test set molecules. The actual and predicted activity values for the training and test set molecules are provide in Table 4. Summary of the PLS analysis results for the best common pharmacophore hypotheses (CPHs) with survival scores are shown in Table 5. Fig. 4 shows the correlation between experimental and predicted MMP-13 inhibitory activity values for the training and test set molecules using the AADHR model.

### 3.4. Generation of pharmacophore model: protein or structure-based approach

A set of 13 crystal structures of MMP-13 in complex with diverse ligands is available in the PDB (Table 6). Out of 13 crystal ligands, 7 were bound to the S1' loop alone. To identify the specific ligands that bind MMP-13, analysis of the intermolecular interactions (both bonded and non-bonded) between those 7 crystal ligands with the S1' loop was performed. One of the co-crystal ligands (PB3) had

**Table 6**  
The set of PDB entries for MMP-13 in complex with diverse ligands.

S. no	PDBID	Zinc binding	S1' loop binding	Resolution
1	830C	Yes	Yes	1.6
2	1XUC <sup>a</sup>	–	Yes	1.7
3	2OW9 <sup>a</sup>	–	Yes	1.7
4	1XUD <sup>a</sup>	–	Yes	1.8
5	1XUR <sup>a</sup>	–	Yes	1.8
6	3ELM	Yes	–	1.9
7	3I7G <sup>a</sup>	–	Yes	1.9
8	1ZTQ	Yes	–	2
9	3I7I <sup>a</sup>	–	Yes	2.2
10	1YOU	Yes	–	2.3
11	2OZR <sup>a</sup>	–	Yes	2.3
12	2D1N	Yes	Yes	2.3
13	456C	Yes	–	2.4

<sup>a</sup> Crystal ligands bound alone in the S1' loop region.

**Table 7**  
The scores for pharmacophoric sites based on the energetic terms from glide XP docking.

Pharmacophore sites	Energy score (kcal/mol)
A3	–1.26
A4	–0.63
D5	–0.02
R9	–0.73
R10	–1.25

more interactions with the S1' loop and showed high specificity with MMP-13 (PDB ID: 1XUC) [27]. Because specificity is an important criterion for MMPs, a protein-based pharmacophore would be a more reliable approach for identifying specific inhibitors of the receptor. Here, the crystal structure of MMP-13 bound with the synthetic molecule PB3, an MMP-13 inhibitor, was used to derive a reliable protein-based pharmacophore model. This model was evaluated by cross checking the PB3 docking result to elucidate the essential features responsible for MMP-13 inhibitory activity.

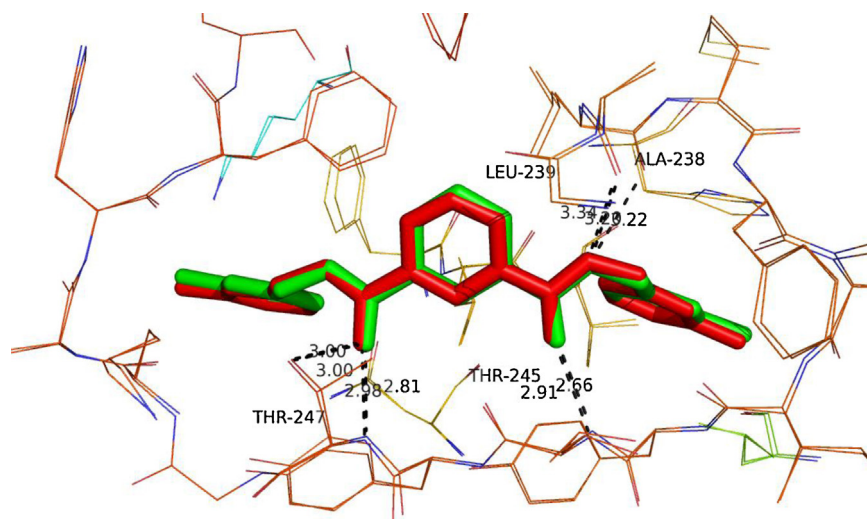
### 3.5. Redocking results for the crystal ligand PB3

Before generating the protein-based pharmacophore model, the crystal ligand PB3 was redocked using Glide XP and SP to validate the binding mode in the S1' loop of MMP-13. The best method for evaluating the accuracy of a docking procedure is to determine the closeness of the lowest energy poses predicted by the scoring function and Glide score with a crystal ligand. The binding poses and hydrogen bond interactions were very similar to PB3 with a RMSD value of 0.17 Å. Fig. 5 shows the alignment of redocked (red) and co crystal ligands (green). Similar to Glide XP, SP is successful in reproducing the native pose; this pose was used to generate the protein-based pharmacophore model using the E-pharmacophore module of maestro.

### 3.6. E-pharmacophore for PB3

Seven pharmacophoric sites were generated. By careful observation of ligand interactions in the S1' loop of MMP-13 and the pharmacophore; two sites in the pharmacophore were excluded. These two excluded sites do not share interactions with the S1' loop amino acid residues. Finally, 5 pharmacophoric sites that map the essential hydrogen- and hydrophobic-selective amino acid residues in the S1' loop of MMP-13 were selected. Glide XP energies from the atoms that comprise each pharmacophore site are summarized in Table 7. The features involved two hydrogen bond acceptors, two aromatic rings and one hydrogen bond donor (AARRD). The two hydrogen bond acceptor groups A3 and A4 shared H-bonds with the main chain of Thr 245 and Thr 247, while the HBD



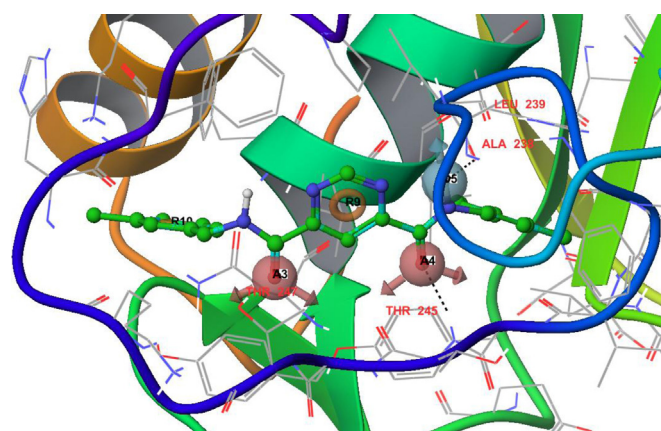


**Fig. 5.** Docking Program Validation: Alignment of the redocked (red) and co-crystal ligand (green) binding pose using the glide docking module. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of the article.)

D5 forms a H-bond with Ala 238 and Leu 239. The two ring aromatic groups R9 and R10 form hydrophobic interactions with Phe 252, Pro 255 and Tyr 246. It was observed from the pharmacophoric features scores that R9 and R10 are necessary for selectively blocking the MMP-13 enzyme. Similarly, the other groups (A3, A4 and D5) showed equal significance for maintaining the binding affinity to MMP-13. The pharmacophoric sites and their energy scores are listed in Table 7, the pharmacophoric features mapped with respect to redocked PB3 are shown in Fig. 6 and distances between the pharmacophoric sites are shown in Fig. 7. These features generated for the above protein–ligand complex will help identify specific novel inhibitors for the MMP-13 receptor.

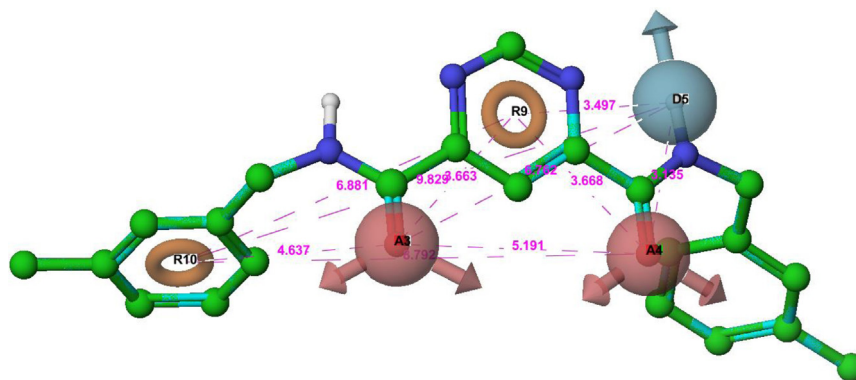
### 3.7. Validation of pharmacophore models

The quality of the ligand- and protein-based pharmacophore models was validated using the GH scoring method. The developed pharmacophore model should search for known active MMP-13 inhibitors to distinguish them from inactive molecules. The results were analyzed using the hit list (*Ht*), number of active percent of yields (%Y), percent ratio for actives in the hit list (%A), the enrichment factor (*E*), false negatives, false positives and a goodness of hit score. The ligand- and protein-based pharmacophore models were successful in retrieving 80 and 91% of the active compounds from



**Fig. 6.** Mapping of protein-based pharmacophoric sites with respect to the H-bond and hydrophobic interactions formed in the S1' loop.

the decoy set, respectively (Table 8). In addition, the calculated EF and GH scores for the ligand- and protein-based pharmacophore models were found to be 23.88, 0.731 and 25.86 and 0.802, respectively, indicating that quality of the pharmacophore models was acceptable. Hence, this study suggests that both pharmacophore



**Fig. 7.** Protein-based hypothesis using the E-pharmacophore module in the Maestro 9.2 software. Pharmacophore features are shown as brick red for the hydrogen bond acceptor (A), and the blue and brown contours represent the hydrogen bond donor (D) and ring aromatic (R), respectively. The distance between the pharmacophore features is shown in angstroms (Å). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of the article.)

**Table 8**  
Ligand- and protein-based pharmacophore model evaluation using the Güner–Henry scoring method.

	Ligand-based	Protein-based
Total number of molecules in the database ( <i>D</i> )	2060	2060
Total number of actives in the database ( <i>A</i> )	60	60
Total Hits ( <i>Ht</i> )	69	73
Active Hits ( <i>Ha</i> )	48	55
% yield of actives [ $(Ha/Ht) \times 100$ ]	69.56	75.34
% Ratio of actives [ $(Ha/A) \times 100$ ]	80.00	91.66
Enrichment factor ( <i>E</i> ) [ $(Ha \times D)/(Ht \times A)$ ]	23.88	25.86
False negatives ( <i>A – Ha</i> )	12	5
False positives ( <i>Ht – Ha</i> )	21	18
Goodness of hit score <sup>a</sup>	0.731	0.802

$$^a [(Ha(3At + Ht)/4HtA)] \times (1 - ((Ht - Ha)/(D - A))).$$

models are useful for discriminating known actives from inactives and suitable for retrieving active inhibitors of MMP-13.

### 3.8. Database screening using ligand- and protein-based pharmacophore models

The overall workflow for ligand-based and protein-based pharmacophore in screening the novel compounds are provided in Fig. 8. Using this two different pharmacophore hypothesis compounds were screened from chemical (300,000) and natural compound databases (300,000). The zinc chemical and natural compounds were filtered based on the fitness score i.e., >1.5, which gave 66 and 41 compounds for ligand-based, 53 and 112 compounds for protein-based pharmacophore. Then, the compounds were subjected to molecular docking studies using SP and XP docking studies. Docking results were used as a post-docking filter to select the compounds which interact with active site amino acids, and to predict the binding orientations of the hit

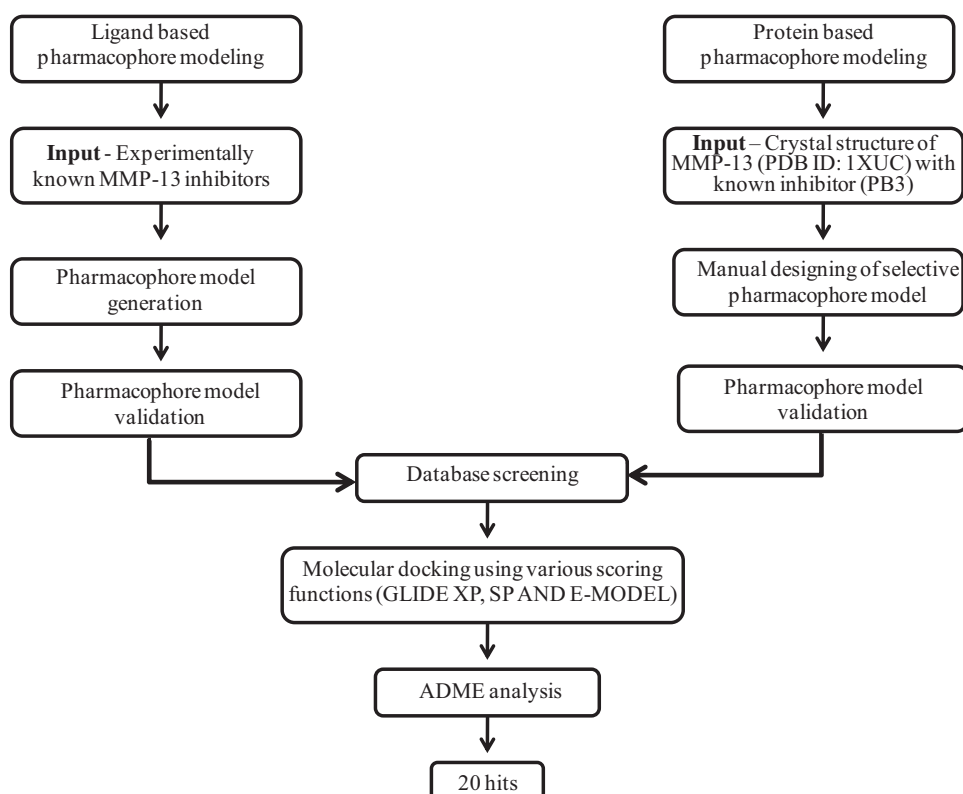
compounds. The compounds with the glide score (>7 kcal/mol), similar binding mode and H-bond and hydrophobic interactions for both ligand- and protein-based hits are listed in Tables 9 and 10. Finally 20 best compounds were shortlisted from ligand- and protein-based studies based on drug likeliness properties. All compounds had shown drug likeliness properties in acceptable range (Table 11).

### 3.9. Docking studies for ligand- and protein-based hits

#### 3.9.1. Binding mode of compounds with the co-crystallized protein 1XUC after docking using SP and XP

The X-ray structure of MMP-13 (PDB ID: 1XUC) was selected as a starting reference for molecular docking analysis. To study the molecular basis of the interaction and affinity of the binding of novel searched compounds, hits were docked into the active site of the MMP-13 receptor using SP and XP docking scoring functions. A comparison of docking poses for known and searched compounds with SP and XP suggested that these compounds showed a similar binding mode with the active site residues of MMP-13 (Ala 238, Leu 239, Thr 245, Thr 247, and Thr 247). Further stabilization of the binding mode was mediated by hydrophobic interaction with Leu 218, Leu 239, Phe 241, Tyr 244, Tyr 246, Lys 249, Phe 252 and Pro 255. The searched inhibitors deeply bind in the specificity loop i.e., the S1' loop, and do not come closer than 5.5 Å to the catalytic zinc ion, showing remarkable selectivity against MMP-13. In addition, most of these inhibitors form hydrophobic interactions with the highly selective amino acid residues Lys 249 and Phe 252, which are directly involved in selectively inhibiting MMP-13 and favor a protein ligand complex by tightly binding with the side chain residues Leu 218, Tyr 246, Phe 252, and Pro 255.

By comparing the different scoring functions SP and XP and the E-model energies, novel potent inhibitors were selected that adopt a similar binding mode and active site amino acid interaction



**Fig. 8.** Workflow for identifying novel inhibitors using ligand- and protein-based pharmacophore modeling.

**Table 9**

Glide docking XP and SP scores, energies, E-model and RMSD scores for ligand- and protein-based hits.

	S. no	Compound ID	Glide score		Glide energy		Glide E-model		RMSD		
			XP	SP	XP	SP	XP	SP	SP	XP	
		<b>Ligand-based hits</b>									
		<b>Crystal compound</b>									
	Natural	1	<b>ZINC 12419118</b>	−10.05	−9.69	−68.737	−65.43	−76.16	−74.35	0.17	0.17
		2	ZINC 67879717	−10.43	−9.82	−77.70	−53.99	−123.07	−72.92	0.54	0.51
		3	ZINC 04200924	−10.06	−8.77	−57.21	−52.04	−99.56	−68.90	0.61	0.67
		4	ZINC 03839736	−9.74	−7.92	−65.30	−47.69	−109.38	−51.77	0.63	0.63
		5	ZINC 03839736	−7.71	−7.21	−59.43	−46.89	−93.82	−61.47	0.51	0.48
	Chemical	6	ZINC 04222324	−7.36	−7.02	−62.40	−41.31	−96.49	−62.74	0.63	0.65
		7	<b>ZINC 00624580</b>	−9.34	−8.73	−79.21	−67.01	−135.71	−100.23	0.65	0.61
		8	ZINC 08437972	−8.75	−7.22	−76.29	−65.24	−124.38	−95.40	0.60	0.59
9		ZINC 01777238	−7.44	−7.11	−71.21	−62.20	−105.28	−99.21	0.73	0.72	
10		ZINC 01819795	−7.62	−7.48	−65.18	−59.78	−106.76	−81.38	0.62	0.60	
		ZINC 19797161	−7.42	−7.02	−57.32	−55.44	−86.22	−77.79	0.59	0.56	
		<b>Protein-based hits</b>									
		<b>Crystal compound</b>									
	Natural	11	<b>ZINC 02535232</b>	−10.05	−9.69	−68.77	−65.43	−76.16	−74.35	0.17	0.17
		12	ZINC 00161877	−9.58	−8.68	−51.00	−45.11	−83.97	−81.74	0.53	0.51
		13	ZINC 00161877	−8.39	−8.35	−43.22	−41.37	−61.56	−65.49	0.57	0.56
		14	ZINC 49449833	−8.11	−8.81	−47.76	−47.76	−76.39	−73.28	0.52	0.50
		15	ZINC 32928079	−7.40	−8.45	−50.63	−49.43	−57.25	−75.91	0.43	0.52
		16	ZINC 68303930	−7.23	−8.23	−45.03	−51.00	−78.07	−67.31	0.38	0.45
		17	ZINC 00032849	−7.52	−7.89	−49.38	−49.23	−75.78	−72.49	0.52	0.56
	Chemical	18	<b>ZINC 08399795</b>	−10.33	−9.80	−83.45	−74.23	−134.20	−92.87	0.56	0.32
19		ZINC 01920523	−9.69	−7.88	−71.94	−65.62	−131.82	−74.62	0.60	0.61	
20		ZINC 00633964	−7.66	−8.91	−74.50	−68.29	−133.22	−108.00	0.33	0.42	
		ZINC 08442293	−7.77	−8.80	−60.21	−61.12	−91.095	−90.44	0.51	0.61	

in the S1' loop of MMP-13. The novel compounds with docking scores, H-bonds, RMSDs and hydrophobic interactions are provided in Tables 3–6. Here, the binding modes of the 4 best identified compounds from zinc chemical and natural compound databases are described.

### 3.10. Validation of the docking program using an ROC plot

The docking studies were subsequently validated by using a dataset that includes 20 candidate compounds from the protein- and ligand-based studies and 2000 randomly sampled compounds.

**Table 10**

Glide XP and SP H-bond and hydrophobic interactions for ligand- and protein-based hits.

	COMPOUND ID	SP GLIDE DOCKING	XP GLIDE DOCKING	Hydrophobic
<b>Ligand-based hits</b>				
Natural	<b>ZINC 12419118</b>	Thr 245, Ile 243	Phe 241, Gly 248, Lys 249	Tyr 244, Tyr 246, Leu 239, Phe 252 Met 253, Pro 255,
	ZINC 67879717	Thr 245, Thr 245	Phe 241, Thr 247	Met 253, Pro 255, Tyr 244, Tyr 246
	ZINC 04200924	Thr 247, Thr 247, Lys 249	Thr 247, Thr 247, His 251	Phe 241, Phe 252
	ZINC 03839736	Thr 247, Thr 247, Lys 249	Thr 247, Thr 247, Gly 248, Lys 249	Phe 252, Met 253
	ZINC 04222324	Thr 247, Lys 249	Gly 237, Thr 247, Thr 247	Tyr 244, Tyr 246, Phe 252, Pro 255
Chemical	<b>ZINC 00624580</b>	Thr 247, Thr 247, Lys 249	Thr 245, Thr 245, Thr 247	Leu 218, Leu 239, Val 219, Tyr 246, Tyr 244 Phe 252,
	ZINC 08437972	Ala 238, Leu 239, Phe 241, Lys 249	Phe 241, Ala 238, Thr 245, Lys 249	Tyr 244, Tyr 246, Phe 252, Pro 255
	ZINC 01777238	Ala 238, Thr 245, Lys 249	Ala 238, Thr 245, Lys 249	Leu 239, Tyr 246, Phe 252, Pro 255,
	ZINC 01819795	Tyr 246	Thr 245, Thr 245, Thr 247	Leu 239, Tyr 246, Phe 252, Pro 255
	ZINC 19797161	Ala 238, Leu 239, Thr 245,	Ala 238, Leu 239, Thr 245,	Tyr 246, Phe 252, Pro 255
<b>Protein-based hits</b>				
Natural	<b>ZINC 02535232</b>	Ala 238, Phe 241, Ile 243, Thr 245, Thr 247, Thr 247, Gly 248	Ala 238, Phe 241, Thr 245, Thr 247, Thr 247	Leu 218, Tyr 246, Phe 252, pro 255.
	ZINC 00161877	Ala 238, Ile 243, Thr 245, Thr 247, Thr 247	Ala 238, Thr 245, Thr 247, Thr 247	Phe 217, Leu 218, Tyr 246, Phe 252.
	ZINC 49449833	Ala 238, Thr 245, Thr 247, Thr 247, Lys 249	Ala 238, Thr 245, Thr 247, Thr 247	Phe 217, Leu 218, Tyr 246, Phe 252, Pro 255.
	ZINC 32928079	Ala 238, Thr 245, Thr 247, Thr 247	Ala 238, Thr 245, Thr 247, Thr 247	Phe 217, Leu 218, Leu 239, Tyr 244, Phe 252, Pro 255.
	ZINC 68303930	Ala 238, Thr 245, Thr 247, Thr 247, Lys 249	Ala 238, Thr 245, Thr 247, Thr 247	Tyr 246, Phe 252, Pro 255
Chemical	ZINC 00032849	Ala 238, Thr 245, Thr 247,	Ala 238, Thr 245, Thr 247,	Phe 241, Tyr 244
	<b>ZINC 08399795</b>	Lys 140, Asn 215, Ala 238, Thr 245, Thr 247, Thr 247	Ala 238, Thr 245, Thr 247, thr 247	Leu 218, Pro 255, Phe 252, Tyr 246
	ZINC 01920523	Ala 238, Thr 245, Thr 247, Thr 247	Thr 245, Thr 247	Phe 217, Leu 218, Val 219, Phe 252, Tyr 246
	ZINC 00633964	Thr 245, Thr 247, Ala 238	Ala 238, Thr 245, Thr 247, Met 253	Leu 218, Val 219, Pro 255, Phe 252, Tyr 246
	ZINC 08442293	Ala 238, Thr 245	Ala 238, Thr 245, Thr 247, Met 253	Leu 218, Val 219 Pro 255, Phe 252, Tyr 244, Tyr 246

**Table 11**

Physiochemical descriptors calculated for the ligand- and structure-based hits using the Qikprop module of Maestro 9.2 software.

	S. no	Compound ID	Mol wt	Accept B	donorHB	QPlogS	QPlogPo/W	QPlogBB
<b>Ligand-based hits</b>								
Natural	1	<b>ZINC 12419118</b>	392.58	8.00	4.00	−2.28	1.20	−0.27
	2	ZINC 67879717	492.43	13.75	5.00	−3.14	0.10	−2.71
	3	ZINC 04200924	541.59	18.70	6.00	−2.43	−1.13	−2.38
	4	ZINC 03839736	578.67	10.00	4.20	−5.85	2.93	−2.16
	5	ZINC 04222324	430.47	8.95	2.25	−3.21	1.87	−0.97
Chemical	6	<b>ZINC 00624580</b>	572.71	13.15	5.00	−4.52	1.75	−2.41
	7	ZINC 08437972	455.40	6.50	2.00	−7.12	4.70	−0.58
	8	ZINC 01777238	448.61	7.50	3.00	−6.20	3.89	−1.42
	9	ZINC 01819795	468.62	9.00	4.00	−5.54	2.14	−2.42
	10	ZINC 01819795	411.47	6.50	3.00	−6.26	3.83	−2.23
		ZINC 19797161	432.48	6.50	1.00	−6.05	3.36	−1.16
<b>Protein-based hits</b>								
Natural	11	<b>ZINC 02535232</b>	232.24	5.25	3.00	−2.54	0.97	−1.26
	12	ZINC 00161877	239.27	4.50	2.00	−3.20	1.98	−1.02
	13	ZINC 49449833	230.22	6.50	4.00	−2.34	−0.10	−1.60
	14	ZINC 32928079	244.24	5.50	3.00	−2.72	1.01	−1.00
	15	ZINC 68303930	244.24	6.00	3.00	−1.33	0.01	−1.57
Chemical	16	ZINC 00032849	233.29	5.75	2.25	−2.38	0.25	−1.57
	17	<b>ZINC 08399795</b>	235.19	10.00	3	−3.075	−0.35	−2.84
	18	ZINC 01920523	371.3	8.00	2	−5.555	2.43	−2.76
	19	ZINC 00633964	410.4	7.00	2	−6.214	4.22	−1.07
	20	ZINC 08442293	464.3	7.00	2	−5.614	3.68	−0.80

Mol wt: 130.0–725.0, donorHB: 0.0–6.0, acceptHB: 2.0–20.0, QPlogS: −6.5 to −0.5, QPlogPo/W: −2.0 to −6.5, and QPlogBB: −3.0 to −1.2.

Enrichment was calculated based on how well the candidate or annotated compounds were retrieved from the decoy database. The enrichment was calculated using ROC plots, which provides a sensitivity and specificity report. The glide docking performance for discriminating known inhibitors from decoys is shown in Fig. 9. Eighty-five percent of the true positives were retrieved. The ROC curve has a gentle increase in the beginning, which means that number of true positives is sacrificed to reduce the amount of false positives when screening a large database of compounds. The area under curve (AUC) value for glide docking is 0.91, which indicates that the positives are ranked higher than the negatives. Fig. 9 shows the evaluation of glide docking using the ROC plot.

### 3.10.1. Binding mode prediction for identified hits

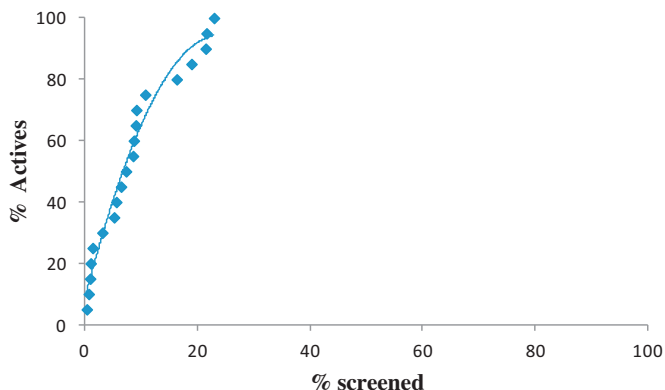
Fig. 10a shows the amino acid interactions of the ZINC 12419118–MMP-13 complex, where the oxygen group (O2 and O9) of ZINC 12419118 donates hydrogen bonds with amino acid residues Phe 241 and Lys 249. The oxygen (O11) donates an H-bond to the amino group of Thr 247. This complex is further stabilized by forming hydrophobic interactions with the S1' loop residues Tyr 244, Tyr 246, Leu 239, Phe 252 Met 253, Pro 255. In the ZINC 00624580–MMP-13 complex, the amide group residing beside the sulfur containing pyrazole ring donates a hydrogen bond

interaction with main chain amino acid residue Thr 245 (Fig. 10b). The pyrazole ring receives a hydrogen bond from main chain residue Thr 247. In this complex, groups of atoms forming the hydrogen bond interaction with amino acid residues were similar to that of the crystal ligand PB3 with a RMSD of 0.61 Å, and the inhibitor was further stabilized by hydrophobic interactions with the specific S1' loop residues Phe 252 and Tyr 244.

Fig. 10c shows the stable binding mode of ZINC 02535232 in the MMP-13 binding site. The binding mode of ZINC 02535232 differs from ZINC 08399795 in the way that side chains Lys 140 and Asn 215 are not involved in a hydrogen bond interaction with the inhibitor. In this MMP-13–ZINC 02535232 complex, terminal amino groups of the inhibitor donate a hydrogen bond interaction from the main chain residues Ala 238, Phe 241 and Ile 243. The amino group in pyrrole ring forms a hydrogen bond interaction with Thr 247 and Gly 248. The two oxygen atoms in the inhibitor form a stable hydrogen bond interaction with the amino and oxygen atoms of Thr 245 and Thr 247, respectively. This complex is further stabilized by hydrophobic interactions with the S1' loop specific amino acid residues Tyr 246, Phe 252, which is common to that of the ZINC 08399795–MMP-13 complex. The calculated binding mode of the most potent inhibitor ZINC 08399795 in the binding site of MMP-13 is shown in Fig. 10d. In this complex, the two terminal oxygen atoms of the inhibitor form hydrogen bond interactions with the side chain amino group consisting of Lys 140 and Asn 215. The presence of multiple hydrogen bonds in the protein ligand complex plays an anchor role for the binding of ZINC 08399795 to MMP-13. The oxygen atoms of the sulfonyl groups in the inhibitor form a hydrogen bond interaction with the main chain oxygen and amino groups of Thr 247 and Thr 245. The terminal amino group donates a hydrogen bond to the oxygen atom of Ala 238. This complex further establishes stable hydrophobic interactions in the binding site with Tyr 246 and Phe 252, which are specific amino acid residues in the MMP-13 S1' loop.

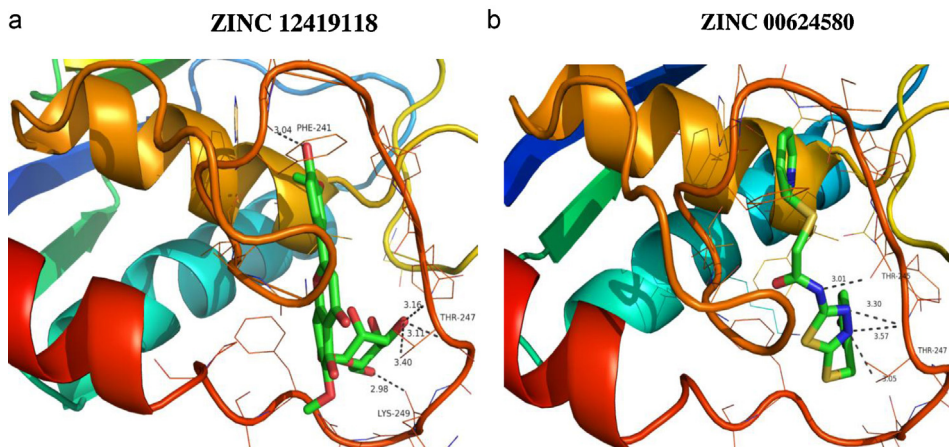
### 3.11. Discussion

MMP-13 is the leading MMP involved in the cleavage of type II collagen, a major component of cartilage. Abnormal activity for this enzyme results in osteoarthritis and rheumatoid

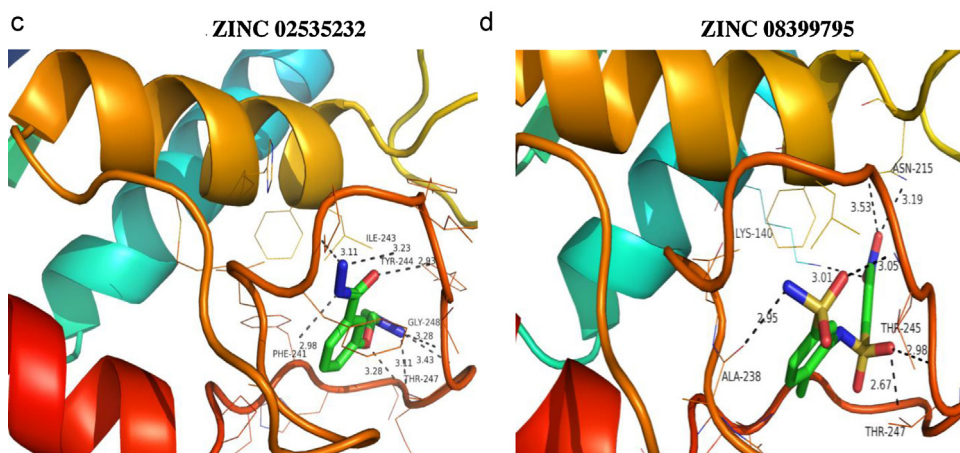
**Fig. 9.** Evaluation of glide docking using an ROC curve.



## Ligand-based hits



## Protein-based hits



**Fig. 10.** (a–d) Binding mode of the ligand- and protein-based hits in the S1' loop of the MMP-13 receptor.

arthritis. The development of potent and selective inhibitors for this enzyme is a difficult task because all of the MMP members are structurally similar. Structural alignment of the MMPs suggests that the S1' specificity loop (residues 244–55) is a structural determinant design a highly selective MMP-13 inhibitor. The MMP-13 crystallographic structure, which includes the *n,n'*-bis(3-methylbenzyl)pyrimidine-4,6-dicarboxamide inhibitor, has provided the opportunity to use virtual screening methods to identify novel inhibitory compounds that specifically dock in the S1' loop of MMP-13. To identify novel and selective inhibitors, combinations of ligand- and protein-based methodologies were applied to screen for the best molecules that interact with the S1' loop in the MMP-13 receptor. The purpose of ligand-based pharmacophore modeling is to elucidate the spatial arrangement of the structural features of various structurally diverse and potent inhibitors crucial for the biological recognition of MMP-13. In addition, protein-based pharmacophore modeling has the ability to predict ligand interactions by generating a map based on protein binding sites in a specific manner.

In ligand-based pharmacophore modeling, 5 common pharmacophore models were generated. The best model, AADHR, was selected based on its highest survival active and post hoc scores, which means that this model is highly reliable. In addition to the survival score analysis, the quality of the AADHR model was further validated by predicting the pharmacophore-based 3D-QSAR model. AADHR showed better statistical results for  $r^2$  and  $q^2$

values, which were 0.94 and 0.75, respectively. The results of this study reveal that the pharmacophore model AADHR accurately predicts MMP-13 inhibitory activity.

In a protein-based study, E-pharmacophore was generated using the predicted binding mode for the highly selective known inhibitor PB3, which showed good energy involvement in the portions reported to interact with the S1' pocket of MMP-13. Because the zinc atom in the catalytic domain is a common catalytic site for all MMPs, excluded volumes at the zinc atom were added to identify selective MMP-13 inhibitors. The generated E-pharmacophore model ranks the portions of the known ligand using glide XP energies. The generated model has two H bond acceptors, two ring aromatic groups and one H bond donor that maps the essential H bond and selective hydrophobic interactions formed in the S1' loop of MMP-13. Because the generated pharmacophore model maps the selective amino acid residues (Lys 249, His 251 and Phe 252) in the S1' loop of MMP-13, this model was considered to be robust enough to screen 3D databases to identify selective non zinc-binding inhibitors for MMP-13. The developed ligand- and protein-based pharmacophore models were validated using the GH scoring method, which is a metric for quantifying the precision of hits and the recall of actives mined from a database consisting of known actives and inactives. Ligand- and protein-based pharmacophore models were successful in retrieving 80 and 91% of the active compounds from the decoy sets, respectively. In addition, the calculated EF and GH scores for the ligand- and protein-based

pharmacophore models were found to be 23.88 and 0.731 and 25.86 and 0.802, respectively, indicating that quality of the pharmacophore models are acceptable.

We have used pharmacophores to screen natural and chemical Zinc databases. Hits above a 1.5 fitness score were used in docking studies with three different scoring combinations, including Glide SP and XP scores, E-model energies and Qikprop, to predict any properties making them likely to be useful drugs. To validate the docking results, we used RMSD values and an ROC plot. The best candidate 20 compounds were docked along with the decoys. Finally, ROC calculations were able to retrieve 85% of the compounds from the decoys. Our results found that potent inhibitors, including ZINC 02535232, ZINC 08399795, ZINC 12419118 and ZINC 00624580, nearly reproduced the H-bond interactions formed in the crystal structure of 1XUC with reasonable RMSD values, exhibiting a novel interaction pattern previously not observed for MMP-13 inhibitors. The above reported novel inhibitors showed <0.7 Å RMSD with respect to the known ligand and bound deeply in the specificity loop i.e., the S1' loop. As reported in previous studies, the novel inhibitors do not come closer than 5.5 Å to the catalytic zinc ion, which confirms their significant selectivity against MMP-13 [27]. The pharmacophore developed for this study was purely based on the inhibitor used in the above-mentioned reference, which demonstrated its potent activity against MMP-13 in vitro. Thus, it is believed that the screened ligands would have more specificity and activity for MMP-13. In addition, most of these inhibitors form hydrophobic interactions with highly selective amino acid residues, such as Lys 249 and Phe 252, which are directly involved in selectively inhibiting MMP-13 and favors the protein ligand complex by tightly binding with the side chain residues Leu 218, Tyr 246, Phe 252, and Pro 255.

#### 4. Conclusion

In this study, we used different virtual screening methods, such as ligand- and protein-based pharmacophore studies and docking methods, to discover novel MMP-13 inhibitors. Ligand- and protein-based pharmacophore models were validated using the GH scoring method. These pharmacophore models were used to screen zinc chemical and natural databases to obtain novel leads, which were further refined using docking studies. The identified hits were validated by comparing the crystal ligand binding pose, different scoring functions, E-model energies, and ROC curves. The potential compound binding modes were predicted, and the results indicated that ZINC 02535232, ZINC 08399795, ZINC 12419118 and ZINC 00624580 interacts with specific residues (Lys 249 and Phe 252) in the S1' loop of MMP-13. These residues are well known to be unique for their interaction with the MMP-13 receptor rather than other MMPs. The ADME analyses including physiochemical descriptors of the above hits were within the acceptable range defined for human use, thereby indicating their likelihood of being used as drugs. Therefore, these compounds may be used as new lead compounds for developing an effective osteoarthritis drug.

#### Acknowledgment

I would like to acknowledge S.R.M. University for their constant encouragement and support.

#### References

- [1] A. Litwic, M.H. Edwards, E.M. Dennison, C. Cooper, Epidemiology and burden of osteoarthritis, *British Medical Bulletin* (2013).
- [2] O. Yoshinari, P. Ann Marone, H. Moriyama, M. Bagchi, Y. Shiojima, Safety and toxicological evaluation of a novel, water-soluble undenatured type II collagen, *Toxicology Mechanisms and Methods* (2013).

- [3] D. Bhatia, T. Bejarano, M. Novo, Current interventions in the management of knee osteoarthritis, *Journal of Pharmacy and Bioallied Sciences* 5 (2013) 30.
- [4] H. Madry, F.P. Luyten, A. Facchini, Biological aspects of early osteoarthritis, *Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA* 20 (2012) 407–422.
- [5] G. Murphy, V. Knauper, S. Atkinson, G. Butler, W. English, M. Hutton, et al., Matrix metalloproteinases in arthritic disease, *Arthritis Research* 4 (2002) S39–S49.
- [6] R.C. Billingham, L. Dahlberg, M. Ionescu, A. Reiner, R. Bourne, C. Rorabeck, et al., Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage, *The Journal of Clinical Investigation* 99 (1997) 1534–1545.
- [7] X. Wang, P.A. Manner, A. Horner, L. Shum, R.S. Tuan, G.H. Nuckolls, Regulation of MMP-13 expression by RUNX2 and FGF2 in osteoarthritic cartilage, *Osteoarthritis and cartilage/OARS, Osteoarthritis Research Society* 12 (2004) 963–973.
- [8] V.M. Baragi, G. Becher, A.M. Bendele, R. Biesinger, H. Bluhm, J. Boer, et al., A new class of potent matrix metalloproteinase 13 inhibitors for potential treatment of osteoarthritis: evidence of histologic and clinical efficacy without musculoskeletal toxicity in rat models, *Arthritis and Rheumatism* 60 (2009) 2008–2018.
- [9] P. Yazdan-Ashoori, P. Liaw, L. Tolft, B. Webb, G. Kilmer, D.E. Carter, et al., Elevated plasma matrix metalloproteinases and their tissue inhibitors in patients with severe sepsis, *Journal of Critical Care* 26 (2011) 556–565.
- [10] C.A. Kontogiorgis, P. Papaioannou, D.J. Hadjipavlou-Litina, Matrix metalloproteinase inhibitors: a review on pharmacophore mapping and (Q)SARs results, *Current Medicinal Chemistry* 12 (2005) 339–355.
- [11] D. Georgiadis, A. Yiotakis, Specific targeting of metzincin family members with small-molecule inhibitors: progress toward a multifarious challenge, *Bioorganic & Medicinal Chemistry* 16 (2008) 8781–8794.
- [12] S. Das, M. Mandal, T. Chakraborti, A. Mandal, S. Chakraborti, Structure and evolutionary aspects of matrix metalloproteinases: a brief overview, *Molecular and Cellular Biochemistry* 253 (2003) 31–40.
- [13] L. Aureli, M. Gioia, I. Cerbara, S. Monaco, G.F. Fasciglione, S. Marini, et al., Structural bases for substrate and inhibitor recognition by matrix metalloproteinases, *Current Medicinal Chemistry* 15 (2008) 2192–2222.
- [14] S. Kalva, S. Vadivelan, R. Sanam, S.A. Jagarlapudi, L.M. Saleena, Lead identification and optimization of novel collagenase inhibitors: pharmacophore and structure based studies, *Bioinformation* 8 (2012) 301–308.
- [15] D.A. Gao, Z. Xiong, A. Heim-Riether, L. Amodeo, E.M. August, X. Cao, et al., SAR studies of non-zinc-chelating MMP-13 inhibitors: improving selectivity and metabolic stability, *Bioorganic & Medicinal Chemistry Letters* 20 (2010) 5039–5043.
- [16] A.C. Good, T.I. Oprea, Optimization of CAMD techniques 3. Virtual screening enrichment studies: a help or hindrance in tool selection, *Journal of Computer-aided Molecular Design* 22 (2008) 169–178.
- [17] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *Journal of Computer-aided Molecular Design* 20 (2006) 647–671.
- [18] D. Lokwani, R. Shah, S. Mokale, P. Shastry, D. Shinde, Development of energetic pharmacophore for the designing of 1,2,3,4-tetrahydropyrimidine derivatives as selective cyclooxygenase-2 inhibitors, *Journal of Computer-aided Molecular Design* 26 (2012) 267–277.
- [19] S. Thangapandian, S. John, S. Sakthiah, K.W. Lee, Ligand and structure based pharmacophore modeling to facilitate novel histone deacetylase 8 inhibitor design, *European Journal of Medicinal Chemistry* 45 (2010) 4409–4417.
- [20] S. John, S. Thangapandian, S. Sakthiah, K.W. Lee, Potent BACE-1 inhibitor design using pharmacophore modeling, in silico screening and molecular docking studies, *BMC Bioinformatics* 12 (Suppl. 1) (2011) S28.
- [21] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, et al., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes, *Journal of Medicinal Chemistry* 49 (2006) 6177–6196.
- [22] D. Yusuf, A.M. Davis, G.J. Kleywegt, S. Schmitt, An alternative method for the evaluation of docking performance: RSR vs RMSD, *Journal of Chemical Information and Modeling* 48 (2008) 1411–1422.
- [23] K.E. Hevener, W. Zhao, D.M. Ball, K. Babaoglu, J. Qi, S.W. White, et al., Validation of molecular docking programs for virtual screening against dihydropteroate synthase, *Journal of Chemical Information and Modeling* 49 (2009) 444–460.
- [24] W. Zhao, K.E. Hevener, S.W. White, R.E. Lee, J.M. Boyett, A statistical framework to evaluate virtual screening, *BMC Bioinformatics* 10 (2009) 225.
- [25] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, et al., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *Journal of Medicinal Chemistry* 47 (2004) 1739–1749.
- [26] G. Barreiro, C.R. Guimaraes, I. Tubert-Brohman, T.M. Lyons, J. Tirado-Rives, W.L. Jorgensen, Search for non-nucleoside inhibitors of HIV-1 reverse transcriptase using chemical similarity, molecular docking, and MM-GB/SA scoring, *Journal of Chemical Information and Modeling* 47 (2007) 2416–2428.
- [27] C.K. Engel, B. Pirard, S. Schimanski, R. Kirsch, J. Habermann, O. Klingler, et al., Structural basis for the highly selective inhibition of MMP-13, *Chemistry & Biology* 12 (2005) 181–189.