Bioorganic & Medicinal Chemistry xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Robust design of some selective matrix metalloproteinase-2 inhibitors over matrix metalloproteinase-9 through in silico/ fragment-based lead identification and *de novo* lead modification: Syntheses and biological assays

Nilanjan Adhikari [a,†], Amit K. Halder [a,†], Sumana Mallick [b], Achintya Saha [c], Kishna D. Saha [b], Tarun Jha [a,*]

[a] *Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, PO Box 17020, Jadavpur University, Kolkata 700032, India*
[b] *Cancer Biology & Inflammatory Disorder Division, CSIR-Indian Institute of Chemical Biology, Kolkata 700032, India*
[c] *Department of Chemical Technology, University of Calcutta, 92, APC Ray Road, Kolkata 700009, India*

ARTICLE INFO

Article history:
Received 24 May 2016
Revised 11 July 2016
Accepted 12 July 2016
Available online xxxx

Keywords:
Matrix metalloproteinase 2
QSAR
Pharmacophore mapping
A549 cell line
Invasion assay

ABSTRACT

Broad range of selectivity possesses serious limitation for the development of matrix metalloproteinase-2 (MMP-2) inhibitors for clinical purposes. To develop potent and selective MMP-2 inhibitors, initially multiple molecular modeling techniques were adopted for robust design. Predictive and validated regression models (2D and 3D QSAR and ligand-based pharmacophore mapping studies) were utilized for estimating the potency whereas classification models (Bayesian and recursive partitioning analyses) were used for determining the selectivity of MMP-2 inhibitors over MMP-9. Bayesian model fingerprints were used to design selective lead molecule which was modified using structure-based *de novo* technique. A series of designed molecules were prepared and screened initially for inhibitions of MMP-2 and MMP-9, respectively, as these are designed followed by other MMPs to observe the broader selectivity. The best active MMP-2 inhibitor had IC_{50} value of 24 nM whereas the best selective inhibitor (IC_{50} = 51 nM) showed at least 4 times selectivity to MMP-2 against all tested MMPs. Active derivatives were non-cytotoxic against human lung carcinoma cell line—A549. At non-cytotoxic concentrations, these inhibitors reduced intracellular MMP-2 expression up to 78% and also exhibited satisfactory anti-migration and anti-invasive properties against A549 cells. Some of these active compounds may be used as adjuvant therapeutic agents in lung cancer after detailed study.

1. Introduction

Remodeling of extracellular matrix (ECM) is important in many physiological and pathological events. Matrix metalloproteinases (MMPs) are zinc-dependent endopeptidases which are involved in remodeling of ECM. The MMPs are implicated in numerous disease conditions such as cardiovascular disorders (atherosclerosis, restenosis, hypertension, heart failure, aortic aneurysm, etc), pulmonary disorders [chronic obstructive pulmonary diseases (COPD), bronchial asthma, pulmonary fibrosis, etc], rheumatic diseases (rheumatoid arthritis, lupus erythematosus, systemic sclerosis, etc) and diabetes mellitus as well as cancer. Owing to these broad ranges of applications, MMPs are still considered as potential targets for drug development.[1]

Although several MMP inhibitors (MMPIs) entered clinical trials, none of these has been established as an anticancer drug due to the adverse effects that mainly stem from the broad spectrum of MMP inhibition. Poor pharmacokinetic profiles are well-known obstruction for the discovery of MMPIs.[2] Most of the trialed compounds are hydroxamate derivatives that show poor aqueous solubility as

Abbreviations: CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity analysis; DS, discovery studio software; FT-IR, Fourier transform infrared spectroscopy; HDAC, histone deacetylase; HMBC, heteronuclear multiple bond correlation; MMP, matrix metalloproteinase; MM-GBSA, molecular mechanics combined with generalized Born and surface-area salvation; Mp, melting point; MS, mass spectroscopy; MLR, multiple linear regression; NMR, nuclear magnetic resonance; QPLD, quantum polarized ligand docking; QSAR, quantitative structure activity relationship; RMSD, root mean square deviation; ROC, receiver operating characteristic; RMSF, root mean square fluctuation.

* Corresponding author. Tel.: +91 33 2457 2495 (o), +91 33 2438 3814 (r), +91 9433187443 (m).
E-mail address: tjupharm@yahoo.com (T. Jha).
† Authors have equal contribution in this work.

http://dx.doi.org/10.1016/j.bmc.2016.07.023
0968-0896/

well as non-specificity towards MMPs and other zinc containing enzymes.[2,3] So far, 28 subtypes of MMPs have been identified and are classified into six subfamilies. One of the subfamilies, gelatinase is having two enzymes—MMP-2 (gelatinase A) and MMP-9 (gelatinase B).[4–6] The MMP-2 has been characterized as the most validated target for cancer whereas MMP-9 is reported as an anti-target in advanced state of the disease due to its anti-angiogenic and anti-tumorigenic functions.[7–12] It was earlier reported that MMP-9 is responsible for the formation of tumstatin that suppress angiogenesis via αVβ3 integrin.[13] The MMP-9 deficient mice were found to have increased invasiveness in neuroendocrine tumorigenesis.[14] Apart from the risks of angiogenesis and metastasis, increased hemorrhage and brain edema were also reported for MMP-9 suppression.[6] In addition, long time suppression of MMP-9 disrupts recovery in cardiac ischemic patients and may lead to cardiac failure.[15] Therefore, it is necessary to design MMPIs that would be active as well as selective against MMP-2 but non-selective towards other MMPs particularly MMP-9.

Computational chemistry gained significance in drug design and development in last few decades. In the current study, robust in silico fragment-based technique was adapted for the development as well as lead modification of the potent and the selective MMP-2 inhibitors. The 'fragments' are smaller and less complicated molecular residues that may be efficiently optimized into the lead compound series if the structural insight is obtained at the outset for the binding interaction between the fragment hit and the target protein of interest. This hypothesis is supported by the reports of clinical trials of drug molecules developed from weakly acting fragment.[16] As MMP-2 and MMP-9 enzymes belong to same MMP subgroup (gelatinase), these have highly homologous catalytic sites[12] for the enzyme actions. Therefore, initially in silico studies were performed considering the higher differences in affinities between MMP-2 and MMP-9. The designed compounds, obtained through the molecular modeling study were synthesized and their enzymatic activities were measured initially against these two MMPs followed by other MMPs to observe the broader selectivity. In addition, some of these active compounds were also tested against nuclear extract of histone deacetylase (HDAC) enzymes consists mainly of histone deacetylase 1 (HDAC1) and histone deacetylase 2 (HDAC2) to justify the selectivity of these compounds against other Zn dependent metalloenzymes. Some highly active designed MMP-2 inhibitors were further investigated for cytotoxicity, apoptosis, cellular expression, anti-migratory and anti-invasive properties against A549 cell line- a highly invasive as well as MMP-2 and -9 over-expressing human lung carcinoma cell line.[17] The initial work of these studies is reported here. The work may help to get some useful leads as adjuvant therapeutic agents in lung cancer.

2. Materials and methods

2.1. Molecular modeling study

Two different types of molecular modeling techniques were performed. The regression analyses (2D and 3D QSAR studies as well as pharmacophore mapping technique) were considered with an aim to understand the mechanistic activity of the designed molecules. The classification analyses (Bayesian modeling and recursive partitioning techniques), on the other hand, were performed to ensure the selectivity of the designed molecules towards MMP-2 comparing with that of MMP-9 which is of the same class.

2.1.1. Dataset

Two hundred fourteen (214) structurally diverse compounds were collected from the published work.[18–37] The biological activity value [IC_{50} (nM)] was converted to the negative logarithmic scale [$pIC_{50} = \log 10^9/IC_{50}$] and was used as the dependent variable in different regression analyses except the pharmacophore mapping method where IC_{50} value was used. Another one hundred fortynine (149) compounds with both MMP-2 and MMP-9 inhibitory activities were collected from the published work[38–43] of Johnson and Johnson Pharmaceutical Research and Development. Around 20% of the data (33 compounds) with the higher MMP-2 selectivity [i.e., the higher MMP-9 IC_{50} (nM)/MMP-2 IC_{50} (nM)] were denoted as the 'selective' and other compounds as the 'non-selective'. These compounds were used for the classification model development.

2.1.2. Division into training and test sets

For regression analyses, initially 25% of the dataset was selected as the test set by diversity based data splitting technique in Accelrys Discovery Studio 3.0 (DS)[44] following the method described earlier.[45] Fifty three (53) compounds were selected as the test set and remaining one hundred sixtyone (161) compounds were considered as the training set. To understand whether the current splitting maintains uniformity in both sets[46], these compounds were analyzed by the principal component analyses (PCA) method.[45] The PCA plot is provided in Supporting information (Fig. S1), which demonstrates that the test set compounds are uniformly distributed in three dimensional PCA plot. Therefore, similarity and uniformity of both of these sets are justified. The classification analyses data were divided randomly into one hundred twenty (120) [21 selective and 99 nonselective] training set and twenty nine (29) [12 selective and 17 non selective] test set compounds.

2.1.3. 2D QSAR model

Different molecular 2D and 3D descriptors were calculated for the dataset compounds through Dragon 2.1[47], DS[44] and Canvas[48] tools. The multiple linear regression (MLR) models were developed on the training set by the forward stepwise regression method following the procedure described earlier.[49] Statistical qualities of the training set MLR equations were justified by the square of correlation coefficient (R^2), adjusted R^2 (R^2_A), variance ratio (F), probability factor related to F-ratio (p) and standard error of estimate (s). To check the predictability of these 2D-QSAR models, leave-one-out (LOO) cross-validation method[50] was used as an internal validation tool. The LOO cross-validated regression coefficient (Q^2) was used for justifying internal predictabilities of the model along with parameters like $r^2_{m(LOO)}$ and $\Delta r^2_{m(LOO)}$.[51] The R^2_{Pred}[52] was used as an external validation parameter to verify the model predictability on the test set. Moreover, $r^2_{m(Test)}$ and $\Delta r^2_{m(Test)}$[53] were also used for justifying external predictabilities of the 2D QSAR model. To predict the overall predictability of models, $r^2_{m(Overall)}$ and $\Delta r^2_{m(Overall)}$[53] parameters were calculated. In order to verify the null hypothesis, Y-based randomization technique[54] was performed for the stepwise multiple linear regression (sMLR) model. Two parameters—R^2_P and cR^2_P[55] were taken into consideration as the validation parameters for the Y-based randomization test. For the MLR models, variance inflation factors (VIF) values[56,57] and applicability domain[58] were also determined to justify significance, robustness and reliability of these 2D-QSAR models.

2.1.4. Pharmacophore mapping method

Hypogen algorithm[59] was used to develop 3D-QSAR pharmacophore models using DS.[44] Thirty (30) structurally diverse compounds were selected as the training set from the current database with the help of *Find diverse molecules* tool of DS.[44] Remaining one hundred eighty four (184) compounds were used as the test set. The BEST conformation generation method with an energy threshold of 20 kcal/mol was used for conformation generation. All parameters were set to default value except uncertainty value and excluded volume. These were set to 1.5 and 8,

respectively, for getting optimum statistical results. The generated pharmacophore models were validated by three different techniques, i.e., (a) cost analysis, (b) Fischer's randomization test and (c) test set prediction method.[60] Configuration cost depends on the complexity of the hypothesis space being optimized. Pharmacophore hypotheses having a difference of 40–60 bits between the total cost and the null cost hypotheses indicates a 75–90% possibility for a true correlation in the data. Fischer's randomization test was performed to ensure that the generated models were not developed by chance. This method was computed with 95% confidence level. The third validation method, i.e., the test set prediction method proves that the pharmacophore models are well predictive over the test set. A test set containing one hundred twenty eight (128) compounds was chosen for the validation of pharmacophore models. The test set prediction was performed using the *Ligand pharmacophore mapping* protocol of DS[44] with BEST conformational search through rigid or flexible fitting options.

2.1.5. 3D QSAR analyses

Compounds fitted (aligned) with the best ligand-based pharmacophore model were used as the aligned structures for 3D QSAR analyses. Two 3D QSAR techniques, comparative molecular field analyses (CoMFA)[61] and comparative molecular similarity analyses (CoMSIA)[62] methods were performed using Sybyl[63] tool. The CoMFA method correlates between molecular properties (steric and electrostatic) and the biological activity in the form of contour maps. It is also possible to predict the biological activity of these compounds by using this method. For the CoMFA method, steric and electrostatic field energies were calculated using Lennard-Jones and Coulomb potentials, respectively. The partial least square (PLS) technique was used for the development of the model. The optimum number of components in the final PLS model was determined by the Q^2 value (*LOO* cross-validation technique). The optimization is based on '5% rule' according to which the number of components is increased only if it increases Q^2 value of 5% or more. Conventional correlation coefficient (R^2), standard error of estimate (SEE) and *F* ratio were considered as the validation parameters. Bootstrapping analysis for 20 runs was also performed to assess the robustness and statistical confidence of the derived models. External predictability of these models was determined by estimating $R^2_{\text{Pred.}}$[45]

2.1.6. Bayesian classification modeling

A Bayesian model[45] on the 'selective' and the 'non-selective' MMP-2 inhibitors was developed following the methodologies described previously.[45,64] The good (favorable) and bad (unfavorable) fingerprints for MMP-2 selectivity were obtained. The model predictability was checked on the test set compounds.

2.1.7. Recursive partitioning (RP) analyses

Create recursive partitioning model module of DS[44] was used to build RP classification model on the 'selective' and the 'non-selective' compounds. The model was developed by minimizing the Gini index to divide these compounds into branches.[45,64] The splitting process was continued until no significant nodes are found or when a minimum number of samples per node are reached. The values of 'Minimum number of samples per node' and 'Maximum tree depth' were set to 10 and 20, respectively. Leave-one out cross validation was used for estimating the error rates of tree depth.

2.1.8. *De novo* fragment based lead modification

In the current study, the best docked pose of the lead molecule was subjected to identify proper fragments for the higher interaction through *de novo link* protocol in DS.[44] This protocol uses Ludi algorithm[65] to identify suitable fragments and it also scores the fragments on the basis of the higher interactions. A site sphere with 13 Å radius was created from the ligand pose. The DS input fragment library was used as a source of fragments. The protocol was applied with all hydrogen atoms of the lead using default parameters including link point maximum alignment angle: 15, link/lipophilic/polar weight: 1, maximum fit attempts: 1000, maximum RMSD: 0.3, bond rotation: 2, density of lipophilic/polar site: 25, pre-selection factor: 2.

2.1.9. Quantum polarized ligand docking and MM-GBSA analyses

Since MMP-2 is a metalloenzyme with a catalytic transition state zinc metal, accurate and precise docking estimation requires quantum mechanics (QM) level treatment.[66] In the present study, quantum polarized ligand docking (QPLD) technique[67] was utilized to dock the designed inhibitors into the active site of MMP-2 enzyme. The NMR structure of MMP-2 enzyme (PDB ID: 1HOV) was obtained from the protein data bank (PDB).[68] The ligand atoms were treated at the QM level whereas the receptor including zinc metal was described using OPLS force field parameters. The partial charges and the protonation states of the receptor molecules were described using Epik module and the catalytic zinc atom was assigned with +2 atomic charge. Histidine residue coordinating with catalytic zinc atom was set with an epsilon protonation state whereas other two adjacent histidine residues were assigned with delta protonation states. The restrained minimization of the hydrogen atom of the protein molecule was done using the *protein preparation wizard* and OPLS2005 force field. A receptor grid box of size 15 Å × 15 Å × 15 Å was generated considering the bound ligand of the protein as the centre of the box. The initial docking of the prepared ligand was performed using the standard precision (SP) protocol that generated the preliminary ligand poses. The single point calculation of the ligand poses using the density functional theory (DFT) based Becke–Lee–Yang–Parr (BLYP)/6-31G* model was performed to calculate the polarizable charges of these inhibitors. These charges were assigned for inhibitors by electrostatic potential (ESP) fitting. Finally, these inhibitors with new partial charges were re-docked into the enzyme active site to obtain the most energetically favorable ligand poses. These selected poses were analyzed through molecular mechanics combined with generalized Born and surface-area salvation (MM-GBSA) approach[69–71] that calculates the free energy of binding of these ligands. The GBSA continuum model was employed for this analysis. The binding free energy of the internal dataset compounds was compared with their observed activities. Different contributors of MM-GBSA binding energy values (Coulomb energy, van der Waals energy, solvation free energy, covalent energy, etc) were also analyzed to understand structural requirements of these molecules.

2.1.10. Molecular dynamic simulation

The MD simulation of the protein–ligand complexes was carried out through Desmond/Maestro.[72–74] Briefly, OPLS2005 molecular mechanics forcefield parameters for the protein–ligand complex were set. The complex was soaked by an orthorhombic water box (extended 10 Å from the complex) of TIP3P water model. The system was neutralized with Na^+ ions applying periodic boundary conditions in each direction. The SHAKE algorithm was used to constrain the motion of covalent bonds using hydrogen atoms. Before MD simulations, a six step relaxation procedure was performed. The solvated system was initially minimized with the solute restrained followed by a minimization without the solute restrained with the help of hybrid method of steepest descent and limited memory Broyden–Flecher–Goldfarb–Shanno (LBFGS) algorithm with maximum 2000 iterations. The relaxed complex was subjected to a 12 ps simulation in the NVT ensemble (temperature 10 K) with non-hydrogen solute atoms restrained. The

similar simulation was carried out for subsequent NPT ensemble at 24 ps with a temperature of 300 K. The temperature and pressure were maintained by Berendsen thermostats and barostats, respectively. The system was simulated for 10 ns with a step of 2 fs NPT ensemble using a Nose–Hoover chain thermostat at 300 K and Martyna–Tobias–Klein barostat at 1.013 bar pressure. The atomic coordinate data and system energies were recorded every 10 ps with a trajectory of 4.8.[74]

2.2. Organic chemistry

The designed molecules were subjected to syntheses and characterizations followed by biological validation for exploring potent and selective MMP-2 inhibitors as well as for other biological activities.

2.2.1. General

The mass spectroscopy analysis was performed in a LC MS/MS instrument [The LC, Agilent coupled to electrospray ion (ESI) mass spectrometer controlled by Mass Hunter Quantitative Analysis software]. The ^{1}H nuclear magnetic resonance (NMR) spectra were recorded on a AC Bruker 300 mHz FT-NMR machine with TMS as the internal standard. The synthesized compounds were dissolved individually in dimethyl sulfoxide-d_6 (DMSO-d_6) solvent. Infrared spectroscopy (IR) analyses were performed in a Bruker alpha 11960095 FT-IR instrument. Splitting patterns are designated as s (singlet), d (doublet) and m (multiplet). The heteronuclear multiple bond correlation spectroscopy (HMBC) was performed with a Bruker Avance DPX-500 spectrometer using DMSO-d_6 solvent. Optical rotation of these compounds was observed by Perkin–Elmer Type 141 *polarimeter*. Melting point of synthesized compounds was measured on Mel-Temp Electrothermal apparatus and a capillary melting point apparatus. These were verified in CTRONICS-a digital melting point apparatus. Chemical reactions were monitored by the analytical thin layer chromatography (TLC) performed on silica gel G plates (TLC silica gel 60 F_{254}, Merck, Germany). The spots were located by keeping the TLC plates in iodine chamber.

2.2.1.1. Syntheses of substituted benzenesulfonyl chlorides (I1). The chemical synthesis was started with chlorosulfonylation of the substituted benzenes (S1) individually in dry chloroform by dropwise addition of chlorosulfonic acid over a period of 45–60 mins at 0–5 °C. After completion of chlorosulfonic acid addition, the reaction mixture was stirred for additional 45 min at room temperature and the mixture was poured over crushed ice. The mixture was extracted thrice with 50 ml chloroform which was dried overnight in presence of anhydrous sodium sulfate. After evaporation of the solvent chloroform, substituted benzenesulfonyl chlorides (**I1**) were obtained individually in solid forms which were used for next step without any purification.

2.2.1.2. Syntheses of 2*N*-substituted benzenesulfonyl L(+) glutamic acids (I2). L(+)-Glutamic acid was taken in a 250 ml conical flask and 2 N NaOH solution was added until the acid was solubilized and reaction mixture becomes alkaline (pH 8–9). In stirring condition, the mixture was heated at 60–70 °C and the substituted benzenesulfonyl chlorides (**I1**) was added individually in small portions. The 2 N NaOH solution was added from time-to-time maintaining the alkalinity of the mixture. The end-point of the reaction was determined by TLC. After completion of the reaction, the mixture was cooled to room temperature and was filtered to remove undissolved solids. The filtrate was acidified with 6 N HCl solution saturated with sodium chloride and extracted thrice with 50 ml portions of ethyl acetate. The substituted benzenesulfonyl-L(+) glutamic acids (**I2**) were obtained individually in solid forms after distillation of the solvent ethyl acetate. These products were purified separately by recrystallization with water.

2.2.1.3. Syntheses of 1-(substituted benzenesulfonyl)-5-oxopyrrolidine-2-carboxylic acids (I3). The 2*N*-substituted benzenesulfonyl L(+) glutamic acids (**I2**) were refluxed separately with acetyl chloride for 2–3 h. The completion of the reaction was estimated by TLC. The mixture was then cooled to room temperature and poured onto crushed ice with continuous stirring. The precipitated product was filtered and recrystallized with water.

2.2.1.4. Syntheses of 5-*N*-substituted 2-(substituted benzenesulfonyl) L(+) glutamines (A1–A26). The 1-(substituted benzenesulfonyl)-5-oxopyrrolidine-2-carboxylic acids (**I3**) were suspended in 20 ml water and to these, excess amines were added individually. The reaction mixture was allowed to stand for 14–15 h and was heated on a steam bath to remove excess unreacted amines. The mixture was cooled to room temperature and 6 N HCl solution was added in chilled condition. The precipitate was filtered off, washed with cold water and recrystallized with dilute ethanol (60%) to get the final products (**A1**–**A26**).

Chemical characterizations of the synthesized compounds (**A1**–**A26**) are presented in Supporting information (Text S1). Isomeric purities of the compounds were determined by HMBC analyses. The HMBC spectra of **A9** are shown in Supporting information (Fig. S2) as an example.

2.3. Biological assays

The synthesized designed molecules were screened for inhibition of MMPs and HDAC activities, cytotoxicity against lung carcinoma cell line (A549), flow cytometry apoptotic assay, cellular expression by immunofluorescence assay, wound healing migration and invasion assays.

2.3.1. Enzyme assays

Two kinds of enzyme assays were preformed namely matrix metalloproteinase (MMP) inhibition assay and histone deacetylase (HDAC) assay.

2.3.1.1. Matrix metalloproteinase inhibition assay. The MMPs inhibition assays[75] were carried out using MMP inhibitor profiling kits purchased from Enzo Life Science International Inc., USA following the manufacturer's protocol. The absorbance value at 410 nm was measured with the help of a microplate photometer (Thermo Scientific Multiscan FC, USA). An inhibitor—*N*-Isobutyl-*N*-(4-methoxyphenylsulfonyl) glycyl hydroxamic acid (NNGH) was included as the prototype control inhibitor.

2.3.1.2. Histone deacetylase inhibition assay. The inhibitory potential of the designed compounds towards HDAC was done by the HDAC colorimetric drug discovery kit (Enzo Life Science International Inc., USA).[75] The colorimetric substrate containing an acetylated lysine side chain was incubated at 37 °C with HeLa nuclear extract having high HDAC activity (especially HDACs 1 and 2) and a range of inhibitor concentrations. The substrate was sensitized by deacetylation and was reacted with the developer causing an increase in the absorption at 405 nm using a microplate photometer (Thermo Scientific Multiscan FC, USA). Trichostatin A was used as the positive control.

2.3.2. Other biological assays

2.3.2.1. Cytotoxicity assay. The lung carcinoma cell line A549 was collected from National Centre for Cell Science, Pune, India and was maintained as monolayer cultures in Dulbecco's Modified Eagle Medium (DMEM) (Gibco Life Technologies, USA)

supplemented with 10% heat-inactivated FBS (Gibco Life Technologies, USA) and Penicillin–Streptomycin (100 IU ml^{-1}–100 μg ml^{-1}). The growth inhibitory effect of compounds was evaluated by using the MTT assay following the earlier described method.[75]

2.3.2.2. Flow cytometry apoptosis assay. The apoptosis was assayed by using an annexin V-FITC apoptosis detection kit (Calbiochem, Germany) following the protocols described earlier.[75] Briefly, after treating the A549 cells with different concentrations of compounds, these were stained with propidium iodide (PI) and annexin V-FITC according to the manufacturer's instructions. The percentage of the live, apoptotic and necrotic cells were analyzed by BD LSR Fortessa cell analyzer (Becton Dickinson, USA). Data from 10^6 cells were analyzed for each sample.

2.3.2.3. Immunofluorescence assay. Anti-MMP-2 antibody (sc-10736) and CFL-tagged secondary antibody (sc-362252, IgG-CFL 488) were purchased from Santa Cruz Biotechnology Inc., USA. The A549 cell line (1×10^6 cells/cm^2) was seeded on cover slips and was subsequently treated with the test compounds (except the control). After 24 h, these cells were fixed with 4% paraformaldehyde in PBS for 15 min, permeabilized with 0.1% Triton X-100 in PBS for 15 min and exposed to the blocking solution (PBS containing 5% bovine serum albumin and 0.2% Tween 20) for 1 h. The fixed cells were incubated overnight with anti-MMP-2 primary antibody (1:250 dilutions), washed with PBS and treated with secondary antibody (1:250 dilutions) for 1 h. The culture was treated with 10 mg/ml of DAPI solution for 1 min for nuclear staining. Samples were mounted in Vectashield hard set media (Vector Laboratories, USA). Images were acquired by an Andor spinning disk confocal microscope. The corrected total cell fluorescence (CTCF) was calculated by ImageJ tool.[76] All integrated density of each cells (presented in these images) and at least six background areas were measured and CTCF value was determined as per the Eq. 1:

$$\mathrm{CTCF} = [\text{Integrated density} - (\text{area of selected cell} \times \text{mean fluorescence of background readings})] \quad (1)$$

2.3.2.4. Would healing migration assay. The A549 cells were seeded in 24 well plate at a density of 5×10^4 cells per well. After a certain confluence (60–75%), these were starved overnight. A scratch was made in the cell monolayer by 20–200 μl pipette tip. The detached and loosely bound cells were washed thrice with PBS. These remaining cells were treated with different concentrations of these compounds for 48 h and the rates of migration were then photographed.[75]

2.3.2.5. Invasion assay. The anti-invasiveness of these compounds was evaluated by a fluorimetric QCM ECMatrix cell invasion assay kit (Millipore, USA) following the methods previously described.[75] The A549 cells were starved overnight in FBS-free DMEM. These staved cells were resuspended in FBS-free medium with 5% BSA and were harvested (1×10^5 cells) in each well of the insert. The feeder tray (the lower portion of the well plate) was filled with 5% FBS. The synthesized compounds were added to the cell suspensions in triplicates. These cells were incubated for 24 h. Invasive cells are able to invade through a basement layer of matrix membrane solution and cross the pores of polycarbonate membranes to adhere at the bottom of the inserts. The adhered cells were dissociated from the bottom with cell detachment solution and the detached cells were stained by CyQuantGR dye. The fluorescence was measured with a multi-mode microplate reader at 480 nm excitation and 520 nm emission (SpectraMax, Molecular Devices, USA) spectra.

3. Results and discussion

In order to design potent and selective MMP-2 inhibitors, two different types of molecular modeling techniques were considered. First, regression analyses (three different methods) based models [2D and 3D quantitative structure activity relationship (QSAR) studies and pharmacophore mapping method] were developed and validated. These predictive and validated models were utilized to understand the structural requirements of MMP-2 inhibitors and to predict the activity of the designed molecules. The biological activity of these inhibitors [IC_{50} (nM)] was converted to the negative logarithmic scale [$pIC_{50} = \log 10^9/IC_{50}$] which was used as the dependent variable in different regression analyses (except pharmacophore mapping where the IC_{50} values were used). In the second step, the classification models were developed by two methods (Bayesian modeling and recursive partitioning analyses) to explore the important fragments responsible for selective MMP-2 inhibition. Bayesian model depicts useful fingerprint fragments responsible for the higher selectivity towards MMP-2 comparing with that of MMP-9. These fragments were joined to design the primary lead molecule.

3.1. Regression analyses

A dataset containing structurally diverse MMP-2 inhibitors were collected from the reported work of *Phizer Global Research and Development*[18–37] [Table S1, Supporting information] and were used for developing the regression models.

3.1.1. 2D-QSAR study

The best 2D QSAR model (Model 1) was developed through the stepwise multiple linear regression (sMLR) analysis and is depicted in the Eq. 2.

$$\begin{aligned} pIC_{50} = {} & 14.037(\pm 2.497) - 11.523(\pm 2.771)BIC5 \\ & + 4.185(\pm 0.592)HOMA - 0.056(\pm 0.008)RDF110u \\ & + 0.727(\pm 0.062)C - 026 + 1.012(\pm 0.116)Jurs\_RPCS \\ & - 0.095(\pm 0.023)STD(O,S) - 0.024(\pm 0.006)PEOE2 \\ & + 0.351(\pm 0.062)\#rtvFG \end{aligned} \quad (2)$$

$n_{Tr} = 161$; $R = 0.882$; $R^2 = 0.777$; $R^2_A = 0.766$; $F(8,152) = 66.397$; $p < 0.00001$; SEE = 0.750; $Q^2 = 0.750$; PRESS = 96.230; SDEP = 0.373; $r^2_{m(LOO)} = 0.735$; $\Delta r^2_{m(LOO)} = 0.160$; $n_{Ts} = 50$; $R^2_{Pred} = 0.690$; $r^2_{m(Test)} = 0.665$; $\Delta r^2_{m(Test)} = 0.181$; $r^2_{m(Overall)} = 0.713$; $\Delta r^2_{m(Overall)} = 0.158$; $k = 1.008$; $k' = 0.975$; $R^2_r = 0.210$; $cR^2_p = 0.755$.

Where n_{Tr} and n_{Ts} are the numbers of compounds used in the training and the test sets, respectively. Regarding the 2D-QSAR model validation, other statistical parameters and their meanings are shown in Table 1. The definitions related to the statistical validation parameters are shown in the Supporting information (Text S2).

Three compounds (Compd **R57**, **R75** and **R77**, Table S1, Supporting information) were removed as outliers from the test set as the standard residual values (observed-predicted) were found to be higher for these compounds. These compounds may act through different mechanism(s) of action. The model explains 76.60% and predicts 75.00% of variances of the experimental activity. The values of R^2_{Pred} of 0.690 and the $r^2_{m(Test)}$ of 0.665 show satisfactory external predictability of this model. The $r^2_{m(Overall)}$ value of 0.713 justifies the overall predictability of the model. The experimental vs predicted activity plot of Model 1 is provided in Figure 1.

The Y test randomization parameters R^2_r and cR^2_p indicate that the model is unique and not developed by chance. The correlation matrix among variables, variation inflation factor (VIF), *p*- and

Table 1
Statistical parameters and their meanings

Sl no.	Validation parameters	Meaning
1	n_{Tr}	Number of compounds in the training set
2	n_{Ts}	Number of compounds in the test set
3	R	Correlation coefficient of regression analysis
4	R^2	Squared correlation coefficient
5	R^2_A	Adjusted square correlation coefficient
6	F	Variance ratio between explained to the residual at specified degrees of freedom
7	p	Probability factor related to F-ratio
8	SEE	Standard error of estimate
9	Q^2	Leave-one-out cross-validated R^2
10	PRESS	Predicted residual sum of squares
11	SDEP	Standard deviation error of prediction
12	$r^2_{m(LOO)}$	Leave-one-out predicted modified R^2 for the training set compounds
13	$\Delta r^2_{m(LOO)}$	Difference of $r^2_{m(LOO)}$ using reverse axes for the training set compounds
14	R^2_{Pred}	Predicted R^2
15	$r^2_{m(Test)}$	Modified R^2 for the test set compounds
16	$\Delta r^2_{m(Test)}$	Difference of $r^2_{m(Test)}$ using reverse axes for the test set compounds
17	$r^2_{m(Overall)}$	Leave-one-out modified R^2 for the whole dataset
18	$\Delta r^2_{m(Overall)}$	Difference of $r^2_{m(Overall)}$ using reverse axes for the whole dataset
19	k	Σ(observed $\times$ predicted)activity/Σ(predicted activity)^2
20	k'	Σ(observed $\times$ predicted)activity/Σ(observed activity)^2
21	R^2_r	Squared randomized scrambled correlation coefficient
22	cR^2_p	Degree of variation between random and nonrandom model

t-values as well as leverage values of Model 1 are provided in the Supporting information (Tables S2, S3 and S4, respectively).

As per the Model 1, *HOMA*, *C-026*, *Jurs_RPCS*, *STD(O,S)* and *#rtvFG* variables have the positive influences on the activity whereas descriptors like *BIC5*, *RDF110u* and *PEOE2* are detrimental to the higher activity. The *HOMA*[77] is an aromaticity index parameter that stands for the harmonic oscillator model of aromaticity index. The positive correlation of *HOMA* indicates that the aromaticity may have a favorable contribution for the higher biological activity. The atom centre fragment index *C-026*[78] signifies the R-CX-R fragment [the number of substituent group attached to the phenyl ring excluding the carbon groups (e.g., methyl, carboxyl, etc)]. The presence of this fragment may increase the MMP-2 inhibitory potential. It also emphasizes that not only the aromatic ring but also its substitutions are important for the higher activity. The *Jurs_RPCS*[79] is a 3D descriptor that stands for the relative positive charged surface area and is positively correlated with the solvent accessible surface area (SASA) of the most positive atom as well as with the total positive charge of the molecule. Therefore, the positively charged atoms and SASA may be favorable for the higher MMP-2 inhibitory activity. Topological descriptor, *STD(O,S)*[80] signifies the sum of the topological distance between the sulfur and the oxygen atoms. The *#rtvFG*[81] descriptor is the number of reactive groups in these molecules. Three topological descriptors, *BIC5* (Bond information content neighborhood symmetry of 5th order), *RDF110u* (unweighted radial distribution function-11.0)[82] and *PEOE2* (partial equalization of orbital electronegativity of second order)[81] are found to be negatively correlated with the biological activity, indicating that the higher values of these descriptors are unfavorable for the activity.

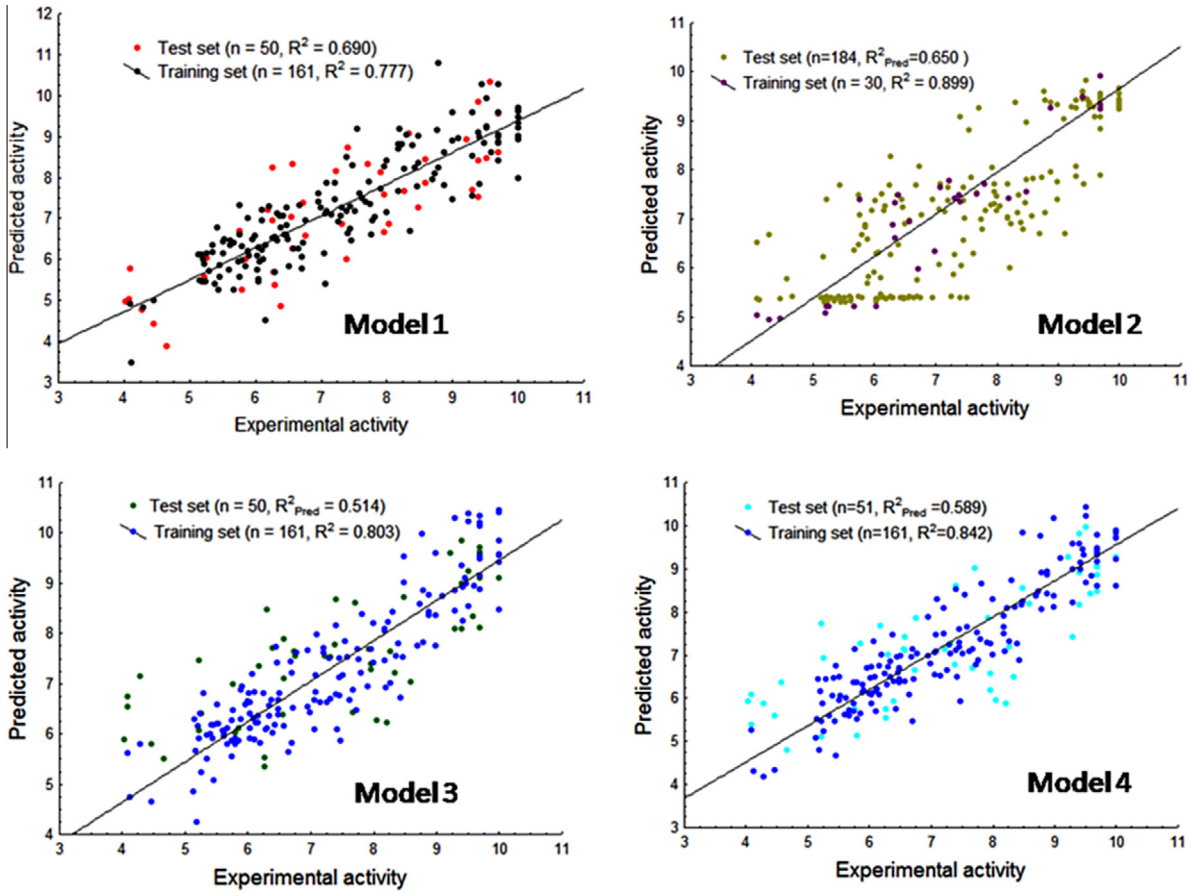


Figure 1. Observed vs predicted activity (pIC_{50}) plots of Models 1–4.

3.1.2. Ligand based pharmacophore modeling

30 structurally diverse training set molecules [Fig. S3, Supporting information] with a considerable biological activity variations were used for developing hypogen 3D QSAR pharmacophore models with the fixed, null and configuration costs of 108.71, 328.82 and 11.87, respectively.[45,60] The hypogen model helps to derive important pharmacophore features (such as hydrogen bond acceptor, hydrogen bond donor, hydrophobic, hydrophobic aromatic, ring aromatic, positive ionizable and negative ionizable groups) from a structurally diverse set of biologically active molecules (variation of activities at least four orders of magnitude). Moreover, this model differentiate the active molecules from the inactive ones by means of identifying important chemical and structural features required for the activity by means of generating top scoring predictive pharmacophores.[45,60] The higher value of the correlation coefficient along with the lower value of the total cost is characteristics of a good pharmacophore model. Furthermore, the value of the total cost should be close to the fixed cost whereas the difference between the total cost and the null cost should be high. The null cost is the cost of generating a hypothesis where the error cost is maximum. The total cost is the actual cost of hypothesis generation that is the sum of weight cost, error cost and configuration cost. The fixed cost is produced for an ideal hypothesis where the error cost is minimal. In addition, the root mean square deviation (RMSD) value should be less than 1.5 and the configuration cost should be less than 17. The RMSD value indicates the overall conformational difference between the final conformations of these structures with the minimized conformations. The cost values, correlation coefficients (R), root mean square deviations (RMSD) and the pharmacophore features of top ten hypotheses are listed in Table 2.

It is evident from Table 2 that the Hypo 1 produced the most reliable model in terms of statistical quality. The low RMSD (<1.50) indicates that the model is generated with a good alignment of these compounds. Hypotheses with the higher number of exclusion volumes (E) produced the better results implying that the steric hindrance plays a significant role for the ligand–receptor interactions. Interestingly, the zinc binding feature (Z) has appeared to be an important feature in the hypotheses, emphasizing its role in MMP-2 inhibition.

To estimate the predictability of the Hypo1 (Model 2) on the external dataset, the remaining 184 compounds were screened with this hypothesis. The test set showed a correlation coefficient (R^2) of 0.491 in the preliminary trial with the rigid fitting. Some compounds (Compd **R22**, **R20**, **R14**, **R42**, **R38**, **R64**, **R69**, **R68**, **R70**, **R71**, **R89**, **R85**, **R145**, **R146**, **R129**, **R133**, **R149**, **R127** and **R180**, Table S1, Supporting information) showed large variations between the observed and the estimated activities. It was also evident that some of these compounds could not map this pharmacophore because of rigidity. Therefore, the flexible mapping was allowed for these derivatives in Hypo 1. It was found that except two compounds (Compd **R64** and **R180**), the flexible fitting lowered the variations between the observed and the estimated activities. Therefore, it was presumed that these compounds require more restrictions in their mapping. These compounds were mapped with Hypo3 that contained four extra E features as compared to Hypo1. This mapping adjusted the predicted activity of these two compounds. Finally, 184 test set compounds produced a predictive correlation coefficient (R^2) of 0.650 in the final flexible fitting. The observed vs predicted activity plot of Model 2 is provided in Figure 1. The interfeature distance constrains of the Model 2 is illustrated in Figure 2A along with the mappings of the best active (Compd **R198**) of the training set (Fig. 2B) and the least active (Compd **R180**) of the dataset on this pharmacophore (Fig. 2C).

The highly active Compd **R198** is fitted with four features of the Model 2 whereas the least active Compd **R180** is mapped with only three features. The phenoxyphenyl moiety of the Compd **R198** is mapped with 'Ha' and 'A' features and the sulfonyl residue of this molecule is mapped with another A feature. The hydroxamate residues of both these compounds are fitted with the Z feature.

3.1.3. 3D QSAR studies

The alignments obtained in the best ligand based pharmacophore mapping (Hypo1) were used for developing comparative molecular field (CoMFA)[61] and similarity (CoMSIA)[62] analyses to understand 3D features required for these molecules. The results are shown in Table 3.

Both the CoMFA (Model 3) and the CoMSIA (Model 4) models were generated with 4 components. The CoMSIA model [Model 4] emerged as the superior to the CoMFA model [Model 3] as far as the statistical results are concerned. The CoMSIA model [Model 4] showed the importance of steric, electrostatic, hydrophobic and acceptor features. The presence of the donor feature deteriorated the statistical qualities and it was therefore, excluded to develop Model 4. Two compounds (Compd **R65** and **R68**, Table S1, Supplementary materials) in the CoMSIA model and three compounds (Compd **R65**, **R68** and **R77**, Table S1, Supporting information) in the CoMFA model were found to be outliers and were removed to develop the present models. The CoMFA and CoMSIA contour maps of the best active (Compd **R189**) and the least active (Compd **R180**) compounds are shown in Figure 3. The observed vs predicted activity plots of Models 3 and 4 are provided in Figure 1. Detail descriptions of model 3 and 4 are provided in Supporting information (Text S3).

Table 2
Statistical information of top 10 pharmacophore hypotheses

Hypo No.	Total Cost	Cost difference[a]	RMSD[b]	Error cost	Correlation	Features[b]
Hypo1	**134.10**	**194.72**	**1.251**	**118.94**	**0.948**	**AAHaHZ11E**
Hypo2	148.77	180.05	1.610	134.33	0.912	AAHaHZ14E
Hypo3	162.93	165.89	1.896	149.36	0.877	AAHaHZ15E
Hypo4	163.99	164.83	1.917	150.61	0.874	AAHaHZ10E
Hypo5	185.86	142.96	2.260	172.08	0.820	AAHaHZ7E
Hypo6	191.78	131.04	2.422	183.48	0.789	AAHaHZ5E
Hypo7	212.98	115.84	2.632	199.34	0.745	AAHaHZE
Hypo8	213.42	115.40	2.637	199.76	0.744	AAHaHZE
Hypo9	216.99	111.83	2.686	203.71	0.732	AAHaHZ3E
Hypo10	222.72	106.10	2.756	222.72	0.715	AAHaHZ

[a] Cost difference between the null and the total cost. The null cost, the fixed cost and the configuration cost are 328.82, 108.71, and 11.87, respectively.
[b] *Abbreviations used for features:* RMSD, root mean square deviation; A: hydrogen bond acceptor; H: hydrophobic; Ha: hydrophobic aromatic; Z: zinc bonding feature, E: exclusion volume.

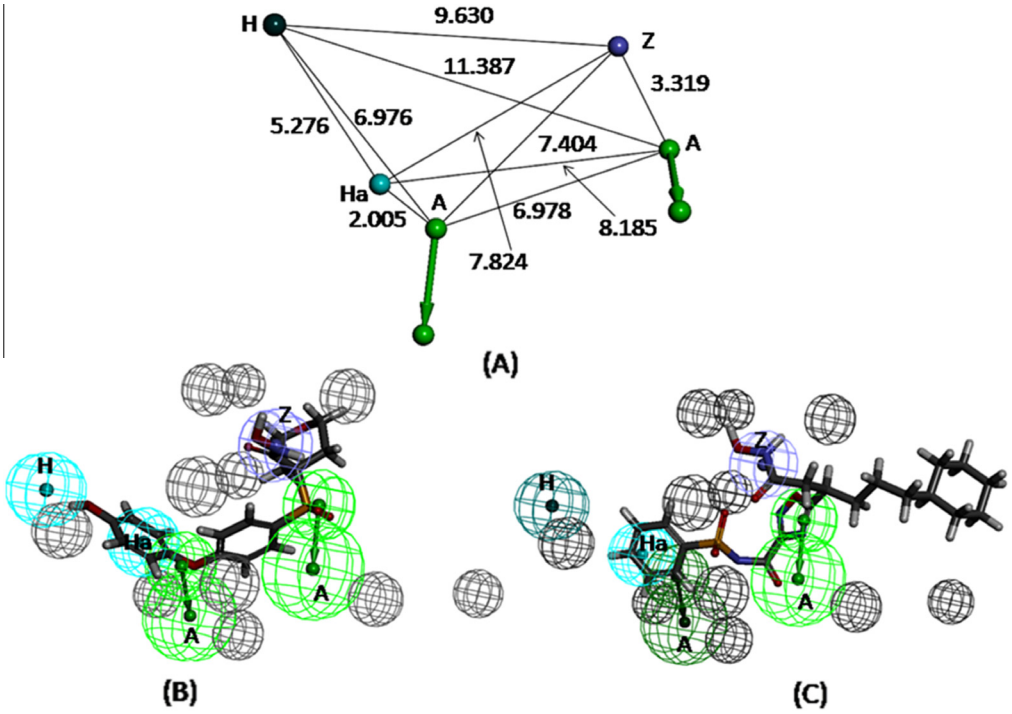


Figure 2. (A) Ligand-based pharmacophore model (Model 2) with its interfeature distance constraints (Å), (B) Mapping of the best active (Compd **R198**) of training set and (C) the least active compounds (Compd **R180**) of the dataset.

Table 3
Statistical results of 3D-QSAR models

PLS statistics	CoMFA	CoMSIA
Non-validated R^2	0.803 (n_{Tr} = 161)	0.842 (n_{Tr} = 161)
SEE	0.696	0.625
F	159.463	207.236
LOO-cross validated Q^2	0.512	0.589
10 fold cross validation	0.501	0.586
$R^2_{bootstrap}$	0.845 (±0.014)	0.885 (±0.008)
$Q^2_{Scrambled}$	0.353	0.324
CSDEP	1.133	1.148
dq^2/dr^2_{yy}	0.917	0.811
Optimal components	4	4
R^2_{Pred}	0.514 (n_{Ts} = 50)	0.599 (n_{Ts} = 51)
Field contributions		
Steric	0.580	0.207
Electrostatic	0.420	0.289
Hydrophobic	—	0.285
Acceptor	—	0.219

3.2. Classification analyses

Since the aim of the current work is not only limited to explore the structural requirements of active MMP-2 inhibitors but also to obtain selectivity towards MMP-2 comparing MMP-9 of the same class, two classification models were built on a dataset containing 149 MMP-2 and MMP-9 inhibitors.[38–43] The structures and activities of these derivatives are provided in Supporting information (Table S5). To build the classification models, 33 compounds having the higher selectivity towards MMP-2 (comparing with MMP-9) were assigned as the 'selective' whereas the remaining 116 compounds were assigned as the 'non-selective'. The whole dataset was randomly divided into 120 training set (21 selective and 99 non-selective) and 29 test set (12 selective and 17 non-selective) compounds. Fingerprint-based Bayesian model[83,84] was developed to understand these structural requirements of the compounds for having more MMP-2 selectivity. Recursive partitioning[64] model was also developed with the same training and test set combination to understand the structural and physicochemical requirements of these compounds for the higher MMP-2 selectivity.

3.2.1. Bayesian model development and validation

Eight descriptors [lipophilicity ($A\log P$), molecular weight (MW), number of hydrogen bond acceptor (nHBA), number of hydrogen bond donor (nHBD), number of ring (nR), number of aromatic ring (nArR) and molecular fractional polar surface area (MFPSA) and one fingerprint feature FCFP_6 (extended connectivity fingerprint of maximum diameter 4)] were used for developing the Bayesian model (Model 5). The definition of these descriptors used in bayesian model development is shown in Table 4.

The model was validated using leave-one-out (LOO) cross-validation method and the receiver operating characteristic curve (ROC)[85] was generated. The ROC curves of the training and the test sets are provided in Supporting information (Fig. S4). The model showed leave-one-out (LOO) cross-validation ROC of 0.853, 5-fold cross validation ROC of 0.851, respectively, and satisfactory enrichments data (Fig. S4, Supporting information). The 'best split' value obtained in the current model (Model 5) is −3.523 which demarcates the 'selective' compounds from the 'non-selective' inhibitors. Based on this value, a contingency table was obtained which helped to calculate the number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) compounds. The test set compounds were used to validate this model and it produced ROC score of 0.833 which indicated satisfactory predictability of the model. The good and bad fingerprints are shown in the Supporting information (Fig. S5A and B, respectively) and five each selected favorable and unfavorable fingerprints are listed in Figure 4.

Figure 4 shows that the biphenyl residues may be responsible for the selectivity towards MMP-2 enzyme comparing to that of MMP-9 whereas the phenoxy derivatives are mostly non-selective. The sulfonyl group emerged as an important fingerprint for assuring selectivity when it is attached to the long aliphatic chains but

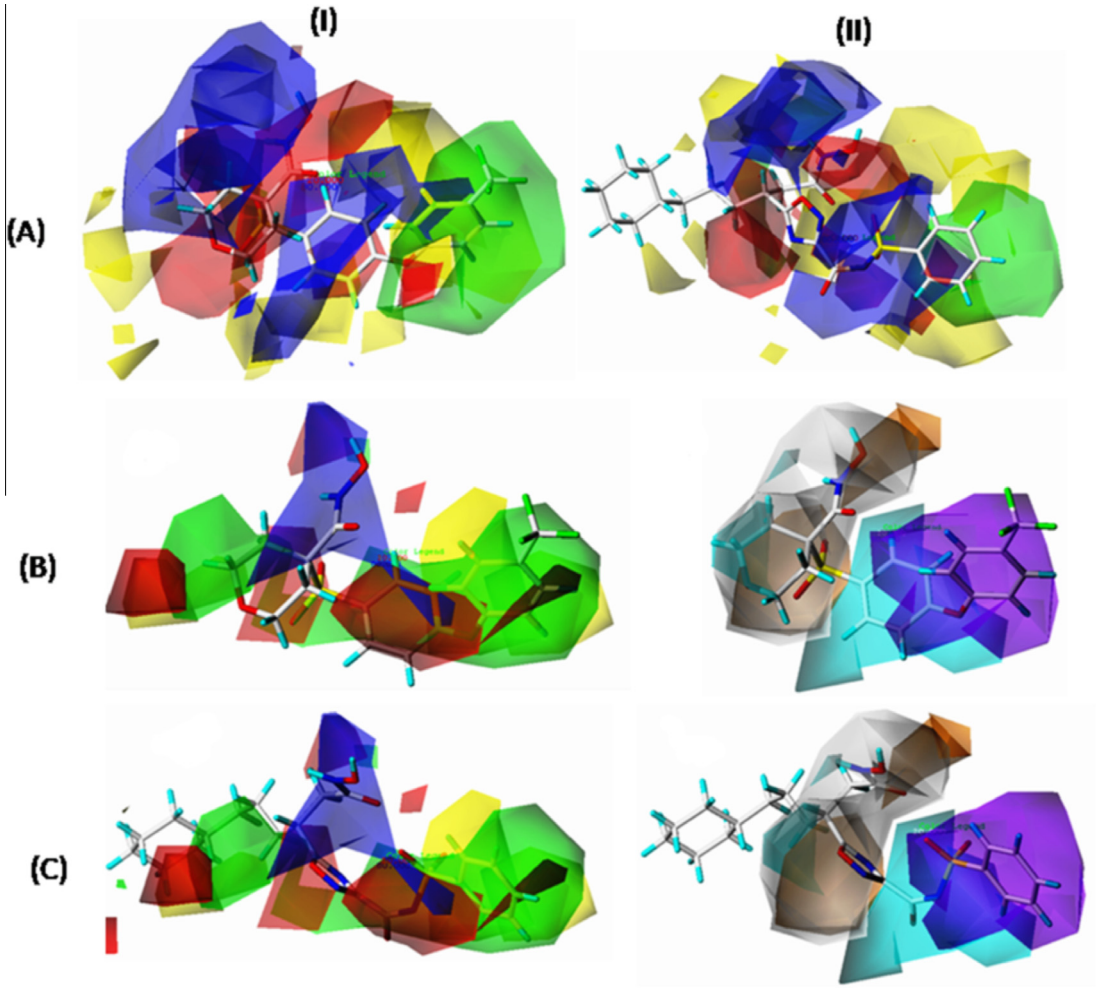


Figure 3. (A) CoMFA, (B) CoMSIA (steric, electrostatic) and (C) CoMSIA (acceptor and hydrophobic) contour maps of (I) the best active (Compd **R189**) and (II) the least active (Compd **R180**) compounds. Steric favorable: green and steric unfavorable: yellow; positive electrostatic: blue and negative electrostatic: red; hydrophobic favorable: purple and hydrophobic unfavorable: white; acceptor favorable: orange and acceptor unfavorable: cyan.

Table 4
Definition of descriptors used in Bayesian model development

Sl No.	Descriptors	Meaning
1	*A*log*P*	Lipophilicity
2	MW	Molecular weight
3	nHBA	Number of hydrogen bond acceptor groups
4	nHBD	Number of hydrogen bond donor groups
5	nR	Number of rings
6	nArR	Number of aromatic rings
7	MFPSA	Molecular fractional polar surface area
8	FCFP_6	Extended connectivity fingerprint of maximum diameter 4

the group is attached to tertiary nitrogen atom infers non-selectivity for MMP-2. The tertiary nitrogen atom was repeatedly found to be a bad fingerprint for MMP-2 selectivity. Among the zinc binding groups, the carboxylic acid group appeared as a good fingerprint for the higher selectivity whereas the hydroxamate group was found as a bad fingerprint.

3.2.2. Recursive partitioning (RP) analyses

The recursive partitioning (RP) on 121 training set compounds generated four classification trees (Tree 1–4). The Tree 1 was selected as the best model (Model 6) on the basis of the receiver operating characteristic (ROC) score (0.902) and 5-fold cross-validated ROC curve (0.816). Moreover, when the model was validated with 29 test set compounds, the ROC score was found as 0.779. The details statistical analysis of Model 4 is provided in Table 5. The schematic representation of the RP-tree 1 model is shown in Supporting information (Fig. S6).

The RP model shows the importance of descriptors like *ES_sum_ssO* (electrotopological state indices of two single bonded oxygen atom), *Jurs_RPCG* (relative positive charge)[79] and *CHI_V_2* (second order Kier & Hall valence-modified connectivity index). The model was primarily classified on the basis of *ES_sum_ssO* descriptor, the value of which should be less than 0.5 for the higher selectivity towards MMP-2. The Bayesian model (Model 5) highlighted the phenoxy containing fingerprints as unfavorable for the better selectivity and this information is consistent with RP prediction. The 2D QSAR model (model 1) described importance of Jurs descriptor, *Jurs_RPCS* that is related to the positively charged solvent accessible surface area (SASA). The descriptor *Jurs_RPCG* stands for the relative positive charge (charge of the most positive atom divided by the total positive charge). Therefore, it is inferred that the positive charge of these molecules are responsible for inferring the selectivity towards MMPs. The topological descriptor CHI_V_2 is the last classifier of the model. Definition of these descriptors used to develop the RP model was shown in Table 6.

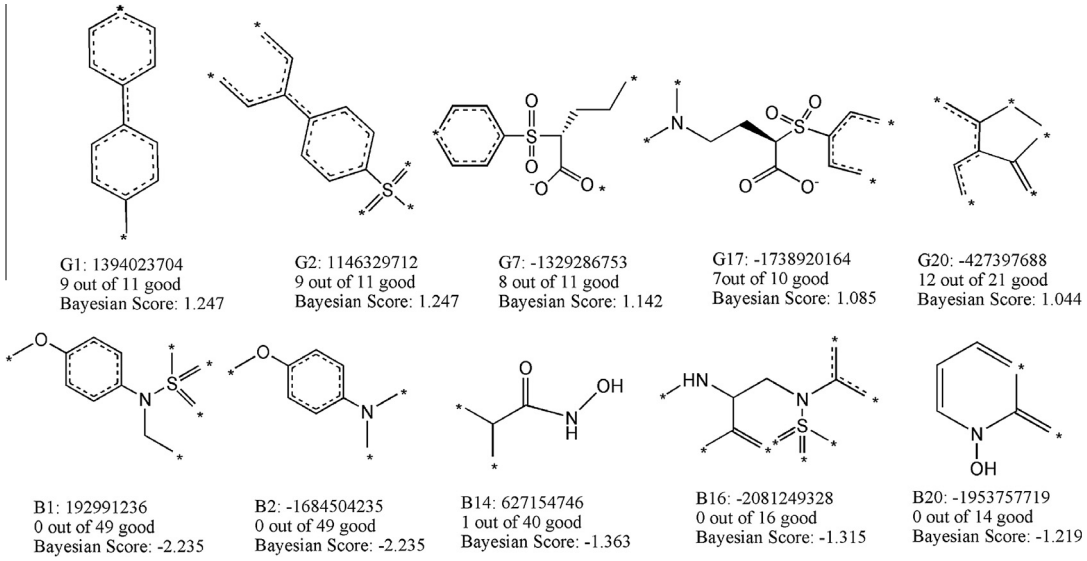


Figure 4. Top five good (G) and bad (B) molecular fingerprints identified by FCFP_6 descriptors.

Table 5
Results of recursive partitioning classification analyses

Set	*n*	Tree no.	TP	FN	TN	FP	Se	Sp	Acc	ROC	ROC_{CV}
Training	120	1	20	1	79	20	0.952	0.798	0.825	0.902	0.816
		2	21	0	70	29	1.000	0.707	0.758	0.894	0.810
		3	16	5	85	14	0.762	0.859	0.842	0.810	0.719
Test	29	1	8	4	14	3	0.750	0.823	0.759	0.779	–
		2	9	3	14	3	0.739	0.657	0.793	0.779	–
		3	8	4	14	3	0.750	0.823	0.759	0.745	–

TP: true positive, FN: false negative, FP: false positive, TN: true negative, Se: sensitivity, Sp: specificity, Acc: accuracy, ROC: receiver operating characteristic, ROC_{CV}: cross validated ROC curve

Table 6
Definition of these descriptors used to develop the RP model

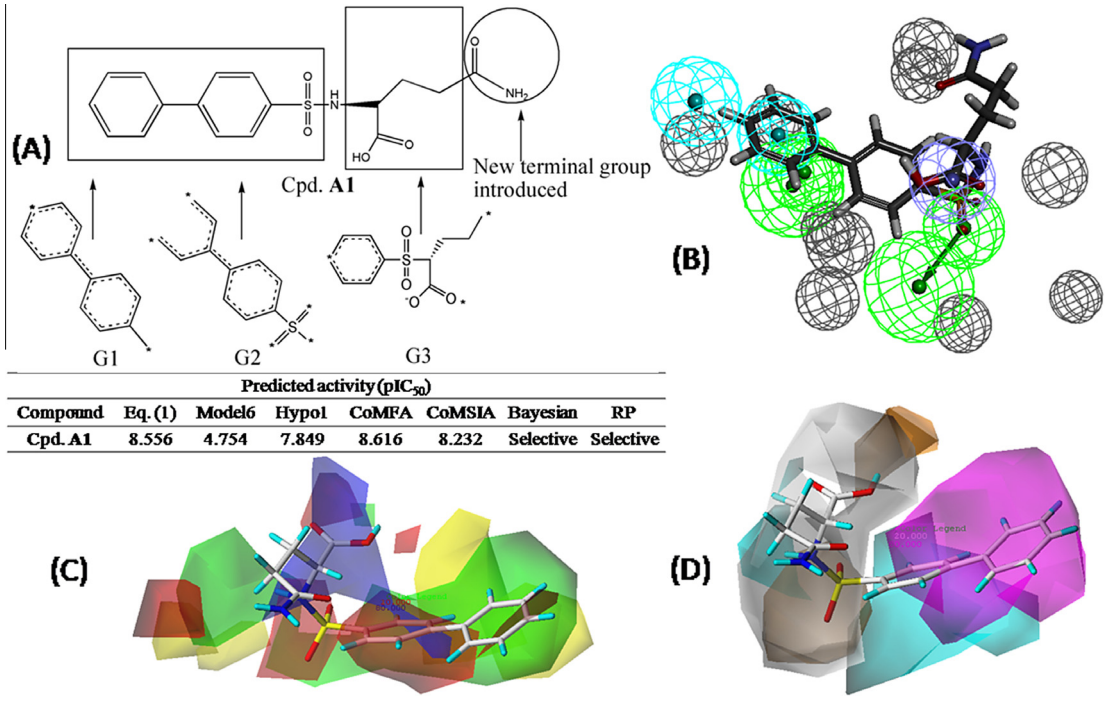
Sl no.	Descriptors	Meaning
1	*ES_sum_ssO*	Electrotopological state indices of two single bonded oxygen atoms
2	*Jurs_RPCG*	Relative positive charge (charge of the most positive atom/total positive charge)
3	*CHI_V_2*	Kier & Hall valence-modified connectivity index of second order

3.3. Design of selective MMP-2 inhibitors (over MMP-9) and *de novo* lead modification

The information obtained in the classification models were analyzed to design a novel compound that may show the higher selectivity towards MMP-2 comparing with that of MMP-9. The compound was designed from the Bayesian fingerprints. The fingerprints, like the biphenyl, the sulfonyl and the aliphatic chains were joined through *de novo* design to form the Compd **A1** [(2*S*)-5-amino-2-[(biphenyl-4-ylsulfonyl) amino]-5-oxopentanoic acid], depicted in Figure 5A. The **A1** was predicted as an 'active' MMP-2 inhibitor in all regression models (Models 1–4) and was also depicted as a 'selective' MMP-2 inhibitor in both of the classification models (Models 5 and 6).

To design the Compd **A1**, three topmost favorable fingerprints (G1–G3) of Bayesian model were taken into consideration so that the specificity to MMP-2 may be ensured. The sulfonamido linkage was included instead of the sulfonyl group since the **A1** showed the better estimated activity against different ligand-based models as compared to the respective sulfonyl derivative. In addition, the introduction of the sulfonamido moiety avoided the exact mimicking of molecules reported earlier in the classification analyses dataset. Moreover, a new terminal carboxamide group was introduced in the structure. Bayesian model predicted the tertiary amino group as a negative fingerprint for the higher selectivity but the roles of the secondary or the primary amines or amides were not revealed. The designed **A1** was mapped in different chemometric models (Models 2, 3 and 4). The current work reports only the (*S*)-isomer of **A1** and its modified derivatives. The (*R*)-isomer also showed similar activity patterns in the regression models. Biological activities (*R*)-isomers are under investigation and will be reported elsewhere.

The compound **A1** was docked at the active site of MMP-2 enzyme (PDB: 1HOV).[86] All poses were submitted for Molecular Mechanics with Generalized Born and Surface Area solvation (MM-GBSA) analyses for the binding energy calculations and the pose with the highest MM-GBSA binding energy of −57.81 kcal/-mol was considered as the best docked conformation. The best docked pose of **A1** along with its interactions is shown in Figure 6. The biphenyl moiety is found to form multiple hydrophobic and pi–pi interactions with catalytic amino acids like Leu83, His120, Thr143 and Tyr142. The catalytic zinc atom forms a metal interaction with the carboxylic acid group and also develops a pi–metal interaction with the biphenyl residue. The sulfonyl group interacts with Leu83 and Ala84 by hydrogen bonding. Additional hydrogen bonding interaction is observed between the carboxyl group of



Predicted activity (pIC_{50})

Compound	Eq. (1)	Model6	Hypo1	CoMFA	CoMSIA	Bayesian	RP
Cpd. A1	8.556	4.754	7.849	8.616	8.232	Selective	Selective

Figure 5. (A) Design of the lead compound (Compd **A1**) from Bayesian fingerprints (G1–G3) and its estimated activities in different ligand based models; (B) Mapping of compound **A1** in Hypo1 (Model 2), CoMSIA (Model 4) (C) electrostatic and steric and (D) acceptor and hydrophobic contour maps of **A1**.

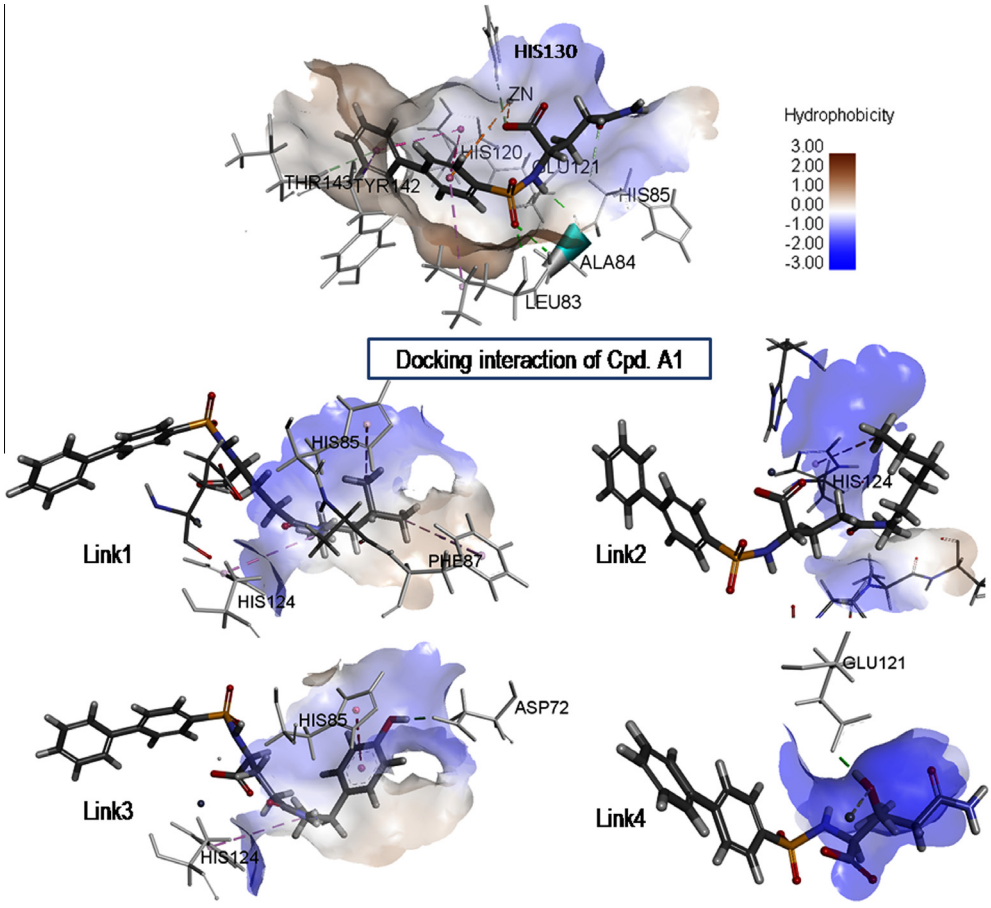


Figure 6. Docking interactions of compound **A1** and its lead modification using *de novo* links (Link1-4).

A1 and His85 residue. These interactions comply with the pharmacophore fitting of **A1**. Moreover, the hydrophobicity contour found in the receptor binding site is also consistent with the CoMSIA hydrophobic contours.

To modify the structure of the designed lead (**A1**), *de novo* fragment based lead modification technique was adopted. When the docked conformer of **A1** was subjected to *de novo* fragment-based search for better interactions, most of the generated links were found near the amide terminal and only a few links were observed at the biphenyl end and other positions. Four top *de novo* links are shown in Figure 6 (Links 1–4). Three of these links (Links 1–3) were found at the terminal amide residue are isobutyl (Link 1), *n*-pentyl (Link 2), 4-hydroxybenzyl (Link 3). These substitutions may give rise to additional interactions with the catalytic amino acids like Asp 72, His85, Phe87 and His124. The Link 4 was rejected since alkyl hydroxyl group was found to be detrimental to the higher MMP-2 selectivity as per Model 5. As the terminal amide group of **A1** was not involved in interaction, it was selected for different aliphatic and aromatic substitutions as suggested by *de novo* design. In order to validate these hypotheses, **A1** as well as 13 other derivatives and analogs of these compounds were synthesized and tested for MMP-2 inhibition. Moreover, these compounds were also screened against MMP-9 inhibition to explore the pattern of selectivity of these two enzymes of the same class.

3.4. Synthesis of the designed molecules

In order to understand the SAR of the designed compounds, **A1** and its 13 derivatives (**A2**–**A14**) were synthesized on the basis of prediction of *de novo* fragment based design to get a fruitful picture of the SAR. Syntheses and characterizations of the designed molecules are discussed in the experimental section. The general synthetic scheme is illustrated in Scheme 1.

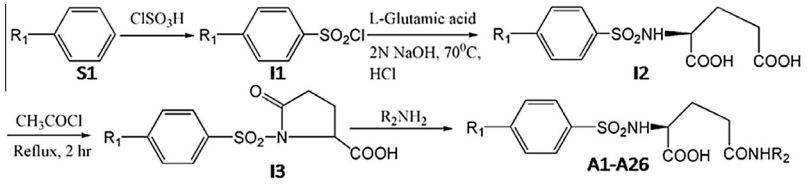
The structures of the designed leads are shown in Table 7.

3.5. Enzyme inhibition assay

Two types of enzyme inhibition assays were performed in the work:

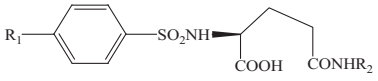
3.5.1. Matrix metalloproteinase (MMP) assay

MMPs are related to cancer migration, invasion and metastases.[7,75] The MMP-2 is a cancer target and MMP-9 is an anti-target.[7–12] We performed preliminary inhibition assays of both MMP-2 and MMP-9 and the results are shown in Table 7. The compound **A1** and its derivatives (**A2**–**A14**) showed variable activities against MMP-2 and -9 subtypes. All these compounds showed satisfactory isoform selectivity towards MMP-2 as compared to that of MMP-9. Though **A1** showed 8 folds selectivity to MMP-2, all of its derivatives show better activity as well as selectivity to this enzyme. The benzyl analog **A13** is found to be the most potent MMP-2 inhibitor (IC_{50} value of 24 nM) whereas the *n*-pentyl analog **A9** is the most selective as well as the most potent MMP-2 inhibitor. The selectivity of **A9** towards MMP-2 is more than 100 times. Interestingly, these two derivatives are directly predicted by *de novo* fragment-based design. Furthermore, the phenylethyl analog **A14** almost five times less potent than **A13** whereas the *n*-hexyl analog **A10** is twice less potent than **A9**. The difference in



Scheme 1. Syntheses of designed molecules (**A1**–**A26**).

Table 7
MMP-2 and -9 inhibitory activities of the designed molecules (Compd **A1**–**A14**)



Compd	R_1	R_2	IC_{50}[a] (nM)		Selectivity[b]	Lipinski rule	Veber rule
			MMP-2	MMP-9			
A1	C_6H_5	H	885	7432	8.40	Pass	Pass
A2	C_6H_5	CH_3	843	38449	45.61	Pass	Pass
A3	C_6H_5	C_2H_5	203	8710	42.78	Pass	Pass
A4	C_6H_5	*n*-C_3H_7	177	11280	63.72	Pass	Pass
A5	C_6H_5	*i*-C_3H_7	208	7407	35.61	Pass	Pass
A6	C_6H_5	*n*-C_4H_9	49	714.9	14.58	Pass	Fail
A7	C_6H_5	*i*-C_4H_9	98	8170	83.37	Pass	Pass
A8	C_6H_5	*t*-C_4H_9	186	8500	45.69	Pass	Pass
A9	C_6H_5	C_5H_{11}	41	5715	139.39	Pass	Fail
A10	C_6H_5	C_6H_{13}	103	5882	57.10	Pass	Fail
A11	C_6H_5	C_6H_{11}	30	636	21.23	Pass	Pass
A12	C_6H_5	C_6H_5	79	4282	54.20	Pass	Pass
A13	C_6H_5	$CH_2C_6H_5$	24	492	20.53	Pass	Pass
A14	C_6H_5	$CH_2CH_2C_6H_5$	123	7001	56.90	Pass	Fail
A15	4Br-C_6H_4-O	C_5H_{11}	472	5494	11.63	Pass	Fail
A16	4Br-C_6H_4-O	C_6H_{11}	672	ND	—	Pass	Pass
A17	4Br-C_6H_4-O	$CH_2C_6H_5$	419	5991	14.29	Pass	Fail
A18	C_6H_{11}	C_5H_{11}	>500	ND	—	Pass	Pass
A19	C_6H_{11}	$CH_2C_6H_5$	>500	ND	—	Pass	Fail

[a] Enzymatic data are mean values of minimum three independent experiments and standard deviations are with ±10%.
[b] MMP-9/MMP-2.

MMP-2 inhibitory activities of **A3** is highly significant since it demonstrates that the potency of the compound increases 4-times when the methyl group is replaced with the ethyl moiety. Moreover, these compounds are drug-like since most of these passed Lipinski[87] and Veber[88] rules. However, four compounds (**A6**, **A9**, **A10** and **A14**) failed Veber rule as total number of rotatable bonds are more than 10 for these derivatives. Therefore, further increase in the aliphatic chain in R_2 position [Table 7] is avoided. The best active compound **A13** passed both of these rules.

To understand the SAR systematically and validate the current in silico predictions regarding the activity and the selectivity of designed MMPIs, five other derivatives (**A15**–**A19**) were also designed and synthesized by the same method. Since the phenoxy moiety has shown to infer less selectivity to the MMPIs in Bayesian Model (Model 5, Fig. 4), three 4-bromophenoxy derivatives (**A15**–**A17**) were synthesized to check the experimental validation of the prediction. The aromaticity of **A1** phenyl ring (attached to sulfonamido residue) is found to be important in all regression models (Models 2–4) as well as docking interactions (Fig. 6). To validate these predictions, two cyclohexyl derivatives (**A18**–**19**) were also prepared. The other parts of the molecules were kept unchanged and the higher active R_2 side chains (*n*-pentyl, benzyl) were included in these derivatives. It is inferred that MMP-2 inhibitory activities of these phenoxy derivatives (**A15**–**A17**) are considerably lower than the biphenyl derivatives and their MMP-2 selectivity also suffered (Table 7). Similar results are also observed for the cyclohexyl derivatives (**A18**–**A19**). These indicate that predictions obtained from different classification and regression models (Models 1–6) are successfully validated experimentally.

In order to understand the selectivity of these designed molecules in other MMPs, some highly active compounds were finally tested against MMP-1, -8, -12 and -14. The enzyme inhibitory activities are depicted in Table 8.

Table 8
Experimental activities of higher active designed molecules in different MMPs

Compd	IC_{50}[b] (nM)					
	MMP-2[a]	MMP-9[a]	MMP-1	MMP-8	MMP-12	MMP-14
A4	177.70	11,280.00	>10,000	45.30	177.60	7783.00
A6	44.65	714.90	>10,000	59.60	71.02	4639.00
A7	98.00	8170.00	>10,000	226.00	246.60	8805.00
A9	41.00	5715.00	>10,000	59.00	52.30	1000.00
A10	103.90	5882.50	>10,000	81.40	96.10	807.00
A11	30.00	636.75	>10,000	64.80	73.20	2407.00
A12	79.80	4282.00	>10,000	56.20	295.80	1389.00
A13	24.00	492.60	>10,000	21.30	53.20	427.00
A14	123.70	7001.00	>10,000	104.30	296.60	6875.00

[a] The activities against MMP-2 and MMP-9 are included for comparison.
[b] Enzymatic data are mean values of minimum three independent experiments and standard deviations are within ±10%.

It is observed that all compounds are inactive against MMP-1. The non-selectivity to MMP-14 is also satisfactory for most of these derivatives. The highest active MMP-2 inhibitor **A13** is more than 20 times selective to MMP-2 compared to MMP-14. Some compounds like **A7**, **A11**, **A12**, **A13** and **A14** have at least twice selectivity to MMP-2 compared to MMP-14 but compounds like **A6**, **A9** and **A10** are either less selective or non-selective. The most significant result is found for MMP-8 activity since the binding affinity value of MMP-8 of **A13** is less than that of MMP-2. Only two analogs, **A7** and **A11** have shown more than twice selectivity to MMP-2 comparing MMP-8. Since **A13** is the most potent MMP-2 inhibitor and pi–pi interaction is observed with the benzyl portion (Fig. 6), some substituted benzyl derivatives were also synthesized by the same method and the activity against different MMPs were determined following the same procedure (Table 9).

The substituted benzyl derivatives (**A20**–**A26**) showed variable activities against MMP-2 enzyme. However, none of these had the higher MMP-2 inhibitory potential comparing with that of **A13**. The compounds **A23** (2-chlorobenzyl analog) has shown equal MMP-2 inhibitory activity to that of **A13**. However, **A23** is also non-specific towards MMP-8. Four derivatives (**A21**, **A23**, **A24** and **A25**) showed IC_{50} values of less than 100 nM against MMP-2. Interestingly, the 3,5-bis-trifluoromethylbenzyl analog, **A21** is found to be the most selective MMP-2 inhibitor of the current series. Other than being very much less selective towards MMP-9 (i.e., MMP-9/MMP-2 = 132), it showed at least five times selectivity to MMP-2 compared to MMP-8. Therefore, **A21** may be considered as the most potent as well as the selective MMP-2 inhibitor among the current series of compounds.

3.5.2. Histone deacetylase (HDAC) inhibition assay

As matrix metalloproteinase (MMP) inhibitors are zinc binders, the inhibitory activities of some important designed compounds (**A9**, **A11**, **A13**, **A21**, **A23**) were tested against HeLa nuclear extract of HDACs as another zinc-dependent protease. The catalytic center of these proteases is formed by narrower tunnel like pockets. None of these derivatives showed inhibition of HDACs upto 100 μM indicating that the designed compounds may be selective to MMPs only.

3.6. In silico physicochemical ADME studies

The designed MMP-2 inhibitors were studied for in silico ADME property prediction through QuikProp module.[81] Values of some important ADME parameters of these compounds are presented in Table 10. It is observed that in spite of the structural similarity, the ADME properties may vary and that in turn, may affect their cellular and biological responses. All these compounds showed the lower blood brain barrier penetration (QP log BB) and also have the lower predicted central nervous system (CNS) activity. The

Table 9
Activities of **A13** derivatives in different MMPs (common substructure is shown in Table 4)

Compd	R_1	R_2	IC_{50}[a] (nM)					
			MMP-2	MMP-9	MMP-1	MMP-8	MMP-12	MMP-14
A13[b]	C_6H_5	$CH_2C_6H_5$	24	492.6	>10,000	21.3	53.2	427
A20	C_6H_5	3,4-di$ClCH_2C_6H_3$	806	ND	ND	ND	ND	ND
A21	C_6H_5	3,5-di$CF_3CH_2C_6H_3$	51	6740	>10,000	228	462	6311
A22	C_6H_5	4-$ClCH_2C_6H_4$	263	2093	ND	ND	ND	ND
A23	C_6H_5	2-$ClCH_2C_6H_4$	31	880	>10,000	39	63	1614
A24	C_6H_5	4-$FCH_2C_6H_4$	76	1155	>10,000	47	84	5424
A25	C_6H_5	4-$OCH_3CH_2C_6H_4$	193	1824	ND	ND	ND	ND
A26	C_6H_5	4-$NO_2CH_2C_6H_4$	291	ND	ND	ND	ND	ND

[a] Enzymatic data are mean values of minimum three independent experiments and standard deviations are within ±10%.
[b] Included for comparison.

Table 10
The predicted ADME properties of designed derivatives

Descriptors	A6	A7	A9	A11	A12	A13	A21	A23	A24
CNS (−2 to +2)	−2	−2	−2	−2	−2	−2	−2	−2	−2
PSA (7.0–200)	128.31	127.42	134.04	117.00	124.87	128.36	128.70	131.00	127.35
QPlog*P* o/w (−2.0 to 6.5)	3.100	3.089	3.065	3.408	3.831	3.565	5.532	4.007	3.975
QPlog*S* (−6.5 to 0.5)	−4.914	−4.778	−4.383	−4.722	−5.866	−5.142	−7.984	−5.932	−5.800
QPlogBB (−3.0 to 1.2)	−2.355	−2.129	−2.477	−1.469	−2.064	−2.273	−1.854	−2.289	−2.109
QPPMDCK (<25 poor >500 good)	18.959	22.902	10.531	58.169	20.117	22.52	343.362	29.852	40.339
Metabolism (1–8)	2	2	2	2	3	3	5	3	3
% Hu Or Ab >80% high <25% poor	70.03	71.34	66.03	81.32	78.29	72.029	57.632	73.49	76.373
Violation of Rule of 5 (Max 4)	0	0	0	0	0	0	2	0	0
Violation of Rule of 3 (Max 3)	0	0	1	0	0	1	2	2	1

predicted *n*-octanol–water partition coefficient (QPlog*P* o/w), aqueous solubility (QPlog*S*) and polar surface area (PSA) are within the permissible limits for all compounds except **A21** that showed poor water solubility. However, the presence of the free carboxylic acid group in **A21** may help to form metallic salts of the free carboxylic acid to increase the water solubility. All compounds showed moderate to good human oral absorption. The permeability in MDCK cells (QPPMDCK) is low for most of these compounds except **A21** that is predicted with exceptionally high cellular permeability as compared to other derivatives. Apart from these, **A21** is predicted to have some unique ADME property since it showed the highest number of likely metabolism reactions and maximum violations in rule of 3 and 5.

3.7. Molecular docking and MM-GBSA binding energy analyses

To understand the SAR of the designed molecules, biphenyl compounds were docked at the binding site of the MMP-2 receptor (PDB: 1HOV).[86] All docked poses were subjected to MM-GBSA binding energy calculations. The (extra precision) XP_Gscores, docking scores and binding energy scores for these derivatives are depicted in Table 11. The energy components of MM-GBSA binding are provided in Supporting information (Table S6).

The MM-GBSA binding energy scores [ΔG_{Bind}] are mostly consistent with the experimental MMP-2 inhibitory activities of these biphenyl derivatives. The MMP-2 inhibitory activity increased almost three folds when the methyl substitution at R_2 position is replaced with the ethyl group. The difference in the ΔG_{Bind} is observed when the ethyl group was converted to the *n*-propyl residue. More branching in the side chains at this position is found to be detrimental since the experimental activity deteriorated when the *n*-propyl residue is replaced with the *i*-propyl and the *n*-butyl is replaced with the *i*-butyl or the *t*-butyl groups. Long straight aliphatic chain R_2-substituents favors the activity. These changes were properly reflected in the binding energy analyses. The MM-GBSA energy components reveal that mainly van der Waals and/or hydrophobic energy are responsible for these differences in affinities. Hydrophobic interaction is also found to be important for the higher activity of the long aliphatic chain derivatives like *n*-pentyl, *n*-hexyl, etc. Therefore, it may be inferred that the residual interaction predicted by *de novo* design (Fig. 6) was correct. The ΔG_{Bind} of the most active compound **A13** is −76.429 kcal/mol which is higher than the small aliphatic chain derivatives but lower than the higher aliphatic chain compounds. All substituted benzyl derivatives except **A13** and **A21** showed comparatively lower ΔG_{Bind} complying with the experimental results. Except these two (**A13** and **A21**), differences between solvated and non-solvated ΔG_{Bind} values increased implying the fact that the solvation free energy may play important role in binding affinities of these compounds. Moreover, improvements of Coulombic energies for these two compounds indicate involvements of more number of polar interactions. The higher Coulombic energies are mainly attributed to better π–π interactions with His85 residues as predicted in the *de novo* designing. Unexpectedly, the second best active **A11** showed extremely lower MM-GBSA binding energy owing to the lower lipophilic interaction.

Table 11
The extra-precision (XP) GLIDE docking scores and MM-GBSA binding energy values of designed derivatives

Compd	R_1	R_2	XP_GScore	XP_Dock Score	ΔG_{Bind}	ΔG_{Bind} (NS)	MMP-2 IC_{50}[a] (nM)
A1	C_6H_5	H	−7.201	−7.201	−55.623	−57.806	885
A2	C_6H_5	CH_3	−7.282	−7.282	−60.026	−66.420	843
A3	C_6H_5	C_2H_5	−6.934	−6.934	−63.998	−70.524	203
A4	C_6H_5	*n*-C_3H_7	−9.149	−9.149	−73.493	−74.900	177
A5	C_6H_5	*i*-C_3H_7	−9.073	−9.073	−70.630	−71.919	208
A6	C_6H_5	*n*-C_4H_9	−9.404	−9.404	−73.558	−79.481	44
A7	C_6H_5	*i*-C_4H_9	−6.162	−6.162	−54.568	−60.506	98
A8	C_6H_5	*t*-C_4H_9	−6.300	−6.300	−50.790	−56.310	186
A9	C_6H_5	C_5H_{11}	−9.338	−9.338	−82.228	−84.785	41
A10	C_6H_5	C_6H_{13}	−8.714	−8.714	−82.029	−89.772	103
A11	C_6H_5	C_6H_{11}	−5.535	−5.535	−52.814	−58.724	30
A12	C_6H_5	C_6H_5	−8.596	−8.596	−63.327	−73.627	79
A13	C_6H_5	$CH_2C_6H_4$	−5.256	−5.256	−73.137	−76.429	24
A14	C_6H_5	$CH_2CH_2C_6H_4$	−8.962	−8.956	−72.023	−80.556	123
A20	C_6H_5	3,4-diCl C_6H_3	−11.460	−11.46	−71.780	−83.056	806
A21	C_6H_5	3,5-diCF_3 C_6H_3	−12.230	−12.23	−75.020	−80.918	51
A22	C_6H_5	4-Cl C_6H_4	−11.457	−11.457	−70.658	−77.765	263
A23	C_6H_5	2-Cl C_6H_4	−9.905	−9.900	−75.749	−78.761	31
A24	C_6H_5	4-F C_6H_4	−11.897	−11.897	−67.520	−73.362	76
A25	C_6H_5	4-OCH_3 C_6H_4	−12.725	−12.725	−67.300	−80.729	193
A26	C_6H_5	4-NO_2 C_6H_4	−11.649	−11.649	−69.570	−78.538	291

[a] The activities against MMP-2 are included for comparison.

The docked conformations of **A13** and **A21** are shown in Figure 7 (A and B). Like **A1**, the biphenyl residues of both of these compounds strongly interacted with Tyr142, Thr143 and His 120 as well catalytic zinc residue. Similarly, the sulfonyl group formed hydrogen bond interactions with Leu83 and Ala84. The benzyl residue of **A13** interacted with His85 and Asp72 and these interactions are consistent with the findings of *de novo* design. The 3,5-bistrifluoromethyl residues of **A21** interacted with three different amino acid residues –His85, Phe87 and Ala88. Since other interactions are similar to **A13**, it may be assumed that these interactions with 3,5-bistrifluoromethyl residues may be responsible for the higher selectivity of **A21** towards MMP-2.

3.8. Molecular dynamics (MD) simulation studies

The most active compound **A13** and the most selective derivative **A21** were subjected to 10 ns MD simulation runs to understand whether the major interactions found in the docking were stable or not in the biological conditions. The percentage occurrences of interactions (more than 10%) are shown in Figure 7(C and D).

Figure 7 demonstrates that the carboxylic acid group of both of these compounds makes strong interaction with the catalytic zinc ion and the percentage of occurrences for this interaction ranges from 70 to 100%. Moreover, the sulfonyl moiety of both of these molecules interacts strongly with Leu83 and Ala84 residues. The amido group attached with the sulfonyl functionality of both of these compounds is found to interact with either water molecules or the catalytic amino acid residues like Ala84 and Gly81. The carboxamido residues showed similar hydrogen bond interactions. Overall, the number of interactions is higher for the compound **A13** that may justify its higher activity as compared to **A21**. The biphenyl aromatic moiety of **A13** interacts with His120 and Tyr 142 residues by π–π interactions and the percentage of occurrences are more than 30% implying strong interactions. In **A21**, however, only one aromatic residue of the biphenyl group interacts with these two amino acids (His120 and Tyr 142). The terminal benzyl moiety of both of these analogs is found to be solvent exposed and therefore, it may be justified that solvation free energy may be a determining factor for the higher binding energy for these synthesized compound as predicted by MM-GBSA binding energy analyses.

Some important MD simulation analyses results are presented in the Supporting information. These provide some interesting facts about the activity and the selectivity of this series of molecules. The average root mean square deviation (RMSD) of **A13** is found to be lower than **A21** (Supporting information, Fig. S7A and B). The **A13** showed some fluctuations in the ligand RMSD whereas **A21** is found with a steady RMSD. The histogram interaction plots show that **A13** generated more number of hydrophobic polar and water bridges interactions than **A21** (Supporting information, Fig. S7C and D). It may explain the lower RMSD and the

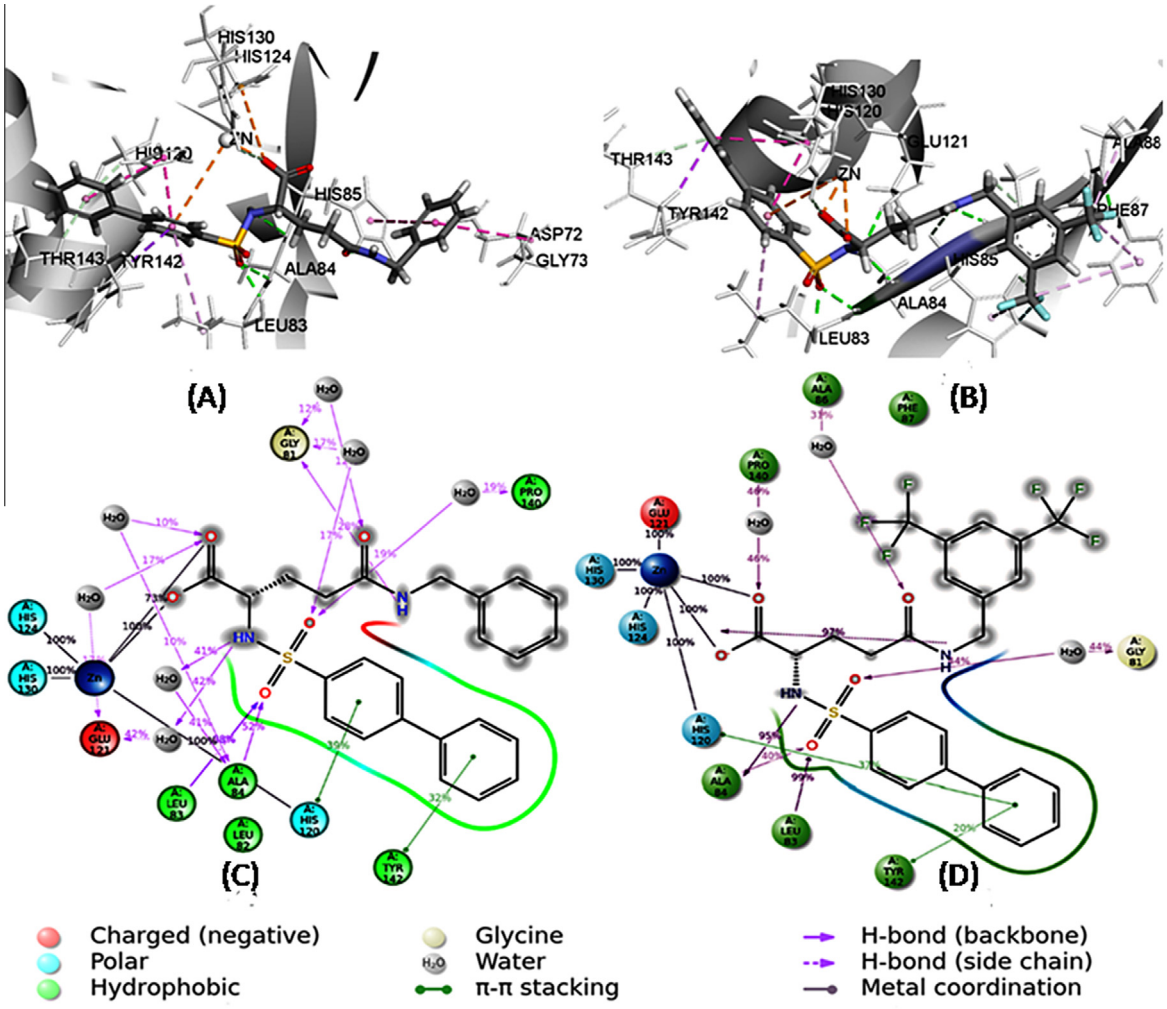


Figure 7. Docking interactions of (A) **A13** and (B) **A21**. The percentage occurrences of interactions observed in 10 ns MD simulation of (C) **A13** and (D) **A21**.

higher experimental activity of **A13**. Interactions with Tyr74, Le82, Tyr142 and Phe148 are observed for **A13** but these are either absent or less frequent in **A21**. Interestingly, **A21** showed more interactions with residues like His85, Ala86, Phe87, Ala88, Leu116 and Val117. To elaborate these interactions, the timeline representations of protein–ligand contacts of these two ligands are observed (Fig. S8, Supporting information). The plot reveals that these interactions of **A21** are not only more frequent than **A13** but also complementary to each other (i.e., in 10 ns MD run, the ligand is found to interact with some of these residues). Interestingly, these amino acid residues are part of S1′ pocket of MMP-2 and it has earlier been reported that more number of interactions with the hydrophobic pocket may ensure the higher selectivity towards MMP-2 against MMP-8.[38] The docking interaction of **A21** suggested that the 3,5-bis-trifluromethylbenzyl residue interacts with His85, Ala88 and Phe87. It may, therefore, be hypothesized that these interactions are responsible for non-selectivity of **A21** towards MMP-8. Two of the non-conservative residues between MMP-2 and MMP-8 (Thr143 of MMP-2 or Ala220 of MMP-8 and Thr145 of MMP-2 or Arg228 of MMP-8) have also been recognized for the differences in ligand affinities between MMP-2 and MMP-8. When the docking interactions predicted a possible π-donor interaction with Thr143 in both these ligands, no information of these interactions is deduced from the MD simulation. On the other hand, interaction with Thr145 residue is unlikely since the ligand poses are located away from this residue. The analyses of the ligand root mean square fluctuations (RMSFs) (Fig. S9, Supplementary material) further reveal that the benzyl residues of both ligands have the higher fluctuations (or variable interactions) whereas the rest of these molecules have the lower fluctuations (or constant interactions). However, the average RMSF of 3,5-bis trifluromethylbenzyl moiety of **A21** is slightly less than the benzyl moiety of **A13**. It may, therefore, be interpreted that the better stability in the terminal residues may be a determining factor for both the affinity and the selectivity of this series of molecules.

3.9. Cytotoxicity and flow cytometry apoptotic assays

The designed compounds having MMP-2 IC_{50} values < 105 nM were tested for cytotoxicity against A549 cell line by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. The cytotoxicity results are shown in the Supporting information (Table S7). All these compounds showed lower cytotoxicity against A549 cell line and most of these compounds showed cell viability 80–90% at concentration of 100 μM. Except **A21**, none of these inhibitors exhibited 50% or more cytotoxicity up to 300 μM (Table S7, Supporting information). This compound showed an IC_{50} value of 270 μM. The results of **A13** and **A21** treated A549 cells at four different concentrations between 0 and 300 μM are provided in the Supporting information (Fig. S10). None of these two compounds significantly reduce the cell viability compared with the untreated control. It was earlier reported that MMPIs may not exhibit cytotoxicity even at the higher concentrations.[89,90] To understand the mechanistic involvement of apoptosis in the antiproliferative activity, **A13** and **A21** were tested in annexin-V/propidium iodide (PI) flow cytometry assay (Fig. S10, Supporting information). Two different doses of 100 and 200 μM were chosen for this assay. At 100 μM concentration, both inhibitors showed minimal apoptotic as well as necrotic effects. However, the necrotic effects predominate in both compounds. At the higher concentrations (200 μM), necrotic effect increases approximately 2.5% whereas apoptotic effects remain unchanged. Two conclusions may be drawn from these studies—(I) these compounds are not cytotoxic below 100 μM and (II) concentrations higher that 200 μM may elicit undesirable necrotic effects in the cell lines. Therefore, a fixed noncytotoxic concentration of 50 μM was chosen for immunofluorescence as well as migration and invasion assays for the active compounds.

3.10. Immunofluorescence assay for cellular MMP-2 expression

To explore the potential of these synthesized molecules to inhibit MMPs in human cancer cell lines, immunofluorescence microscopy technique was adopted to find out cellular localizations of MMP-2 enzyme in the treated and untreated cell lines. The A549 cell line (1×10^6 cells) was cultured in chamber slides and treated with a fixed dose (50 μM) of three compounds (Compd **A9**, **A13** and **A21**) for 24 hours. The best active **A13** is chosen along with **A9** since the later is the most selective MMP-2 inhibitor with the aliphatic side chain. The **A21**, on the other hand, is the most selective MMP-2 inhibitor of this series. After 24 h, cells were washed with phosphate buffer saline (PBS) and fixed in chilled methanol and permeabilized in 0.01% triton X 100. Nonspecific bindings were blocked by bovine serum albumin (BSA) in phosphate buffer saline (PBS), followed by incubation with MMP-2 antibodies. Expression was detected with conjugated secondary antibody. For nucleus counterstaining, 4′-6-diamidino-2-phenylindole (DAPI) was used. The expressions of MMP-2 antibodies in untreated control and treated cells are presented in Figure 8A.

It is evident from Figure 8A that all treated cells showed the lower expression of MMP-2 comparing with the untreated cells. The mean fluorescence reading was recorded. The total intensity of fluorescence was measured by calculating the corrected total cell fluorescence (CTCF) of the control and samples are provided in Supplementary materials (Fig. S11). The cellular MMP-2 expression of **A13** treated cells is slightly lower than that of **A9**. The overall MMP-2 expression is reduced upto 60.50% in **A13** treated cells whereas the intensity is lowered upto 52.10% in **A9** treated cells. Notwithstanding the lowest observed binding affinity among these three derivatives, the most prominent inhibition is obtained for **A21**. This compound reduces MMP-2 expression upto 78.90% comparing with that of the control. The DAPI and phase contrast diagrams (Fig. 8A) depict that the designed compounds did not produce considerable DNA fragmentations and morphological feature changes at 50 μM concentrations.

3.11. Wound healing migration assay

To investigate anti-migratory properties of the designed compounds, migration assay was performed with A549 cell line which was treated with 50 μM of compounds **A9**, **A13** and **A21** for 48 h after a wound was made on the monolayer of cells. The observations are illustrated in Figure 8B. It demonstrates that the control cells extensively migrated into the denuded area. However, the cell migration was significantly reduced by the designed compounds. The inhibition was most prominent for the compounds **A13** and **A21** whereas **A9** also reduced the migration but overall effect was lower than that of these two compounds.

3.12. Invasion assay

Inhibitions of A549 cell invasion by the designed compounds were measured using fluorimetric QCM ECMatrix cell invasion assay (ECM 555, Millipore). Some higher active designed compounds (**A6**, **A9**, **A11**, **A13**, **A21** and **A23**) were investigated for their anti-invasive properties. The total numbers of migrated cells are graphically represented in Figure 8C. It was observed that all of these compounds were able to reduce the total number of migrated cells. The maximum inhibition was observed for the compound **A13** (55.8 ± 2.7%), followed by the compounds **A6** (52.0 ± 3.0%) and **A21** (51.0 ± 1.5%). The other three compounds showed less than 50% anti-invasive properties (**A9**: 37.0 ± 2.6%,

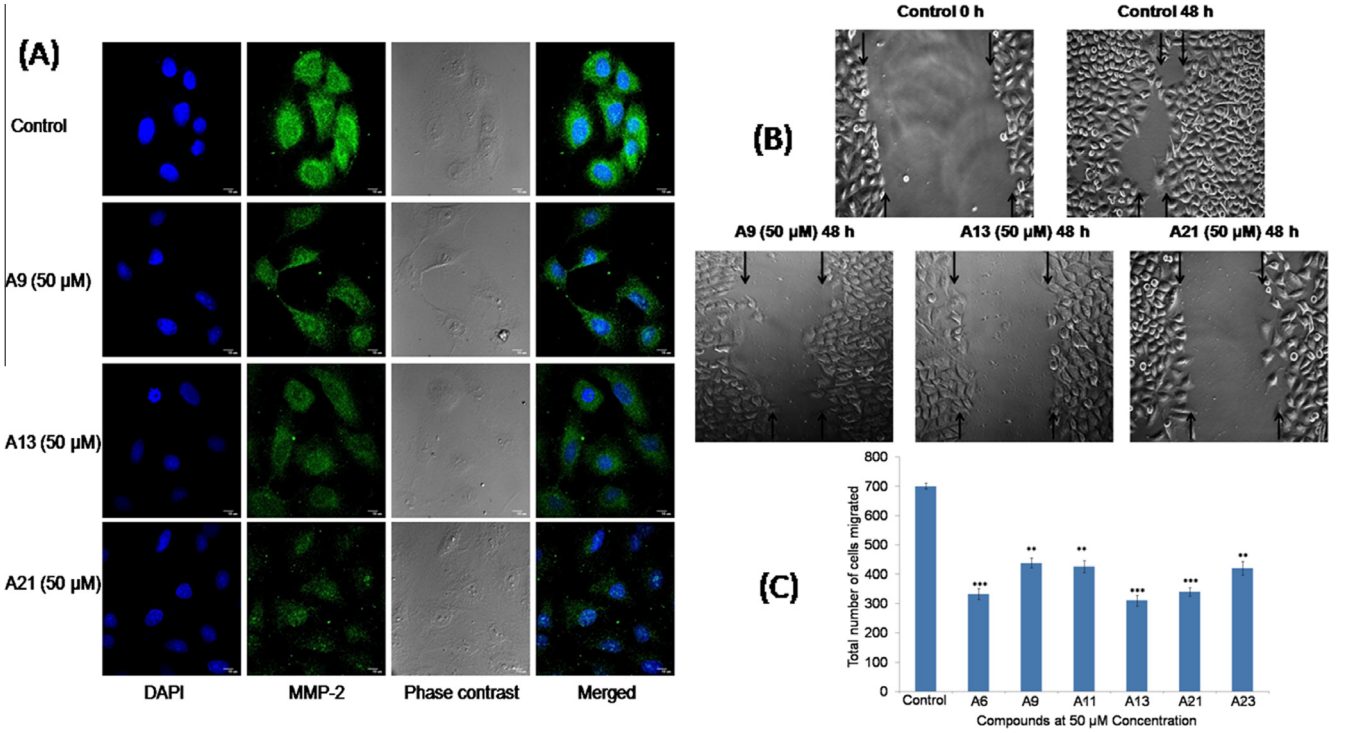


Figure 8. (A) Confocal microscopy images of cellular MMP-2 expressions in untreated, compounds **A9**, **A13** and **A21** treated A549 cell lines. 1×10^6 cells were seeded in glass coverslips and treated with fixed concentration (50 μM) of compounds (except the control) for 24 h. These cells were fixed with formaldehyde, permeabilized with Triton X-100 and blocked with PBS containing blocking solution. Cultures were incubated with primary antibody followed by secondary antibody. These were counter stained with DAPI. The changes of MMP-2 expression in untreated and treated cells were observed by confocal microscopy. (B) Effects of designed molecules (**A9**, **A13** and **A21**) on the migration in A549 cells. (C) The total number of cell migration in untreated and compounds (**A6**, **A9**, **A11**, **A13**, **A21** and **A23**) treated A549 cells. A549 cells were incubated with or without compounds for 24 h. The cells migrate to invade from the upper side of insert to the underside of the porous polycarbonate membrane. Migration abilities of A549 cells were measured using CyQuantGR dye. Values were expressed as means ± SD representing 3 independent experiments.

A11: 39.0 ± 3.1% and **A23**: 40.0 ± 2.4%). Therefore, it may be inferred that all these compounds have both anti-migratory and anti-invasive properties in A549 cell line.

4. Conclusion

In the present work, an initial attempt has been made to design some novel active as well as selective MMP-2 inhibitors over MMP-9 through the concepts of molecular modeling. The basic structural requirements of these molecules for potential MMP-2 inhibitory activity and selectivity (with respect to MMP-9) were analyzed by the regression and classification models. It was observed from Figure 5 that the outcomes of different regression models were similar as far as the predictability of the lead compound is concerned. Similarly, two classification models also depicted the same result for the lead compound. Application of more than one regression and classification tools consolidated the predictability for the lead compound by robust design. However, the scope of these models are not only limited to the design and predictability of the current work only. The information obtained from different modeling analyses may be utilized by other workers to understand structural requirements for potent and selective MMP-2 inhibitors. The 2D QSAR model provided some crucial structural and physicochemical descriptors for the higher potency. Pharmacophore mapping study demonstrated the essential pharmacophoric features responsible for the higher MMP-2 inhibitory activity. Similarly, 3D QSAR models showed the essential three dimensional requirements for the higher activity. The Bayesian classification method highlighted the favourable and the unfavourable fingerprints whereas the RP modeling suggested important structural and physicochemical descriptors for the higher MMP-2 selectivity (comparing with MMP-9). Considering the fact that these analyses are independently developed and validated in the current work, one or more models may be utilized for other investigators for the design of novel MMP-2 inhibitors.

During designing in the current work, the higher activity was preferred over the selectivity (i.e., the selectivity of a compound was discarded if it is found to be less potent than most of the higher active molecules). Initial enzyme assays were performed for both MMP-2 and MMP-9 followed by other MMPs to observe the broader selectivity. The results of these enzyme assays led to further synthesis of derivatives of the higher active compound **A13** and tested against different MMPs. The compound **A13** emerges as the most potent MMP-2 inhibitor whereas the compound **A21** shows the highest MMP-2 selectivity. The docking poses, binding energies, in silico ADME properties and molecular dynamics simulation results of important compounds were obtained and were analysed and described. It is revealed that most of these designed molecules are non-cytotoxic in nature even at high concentrations in A549 cell line. At non-cytotoxic concentrations, their cellular permeability is satisfactory to reduce the expression of intracellular MMP-2 enzyme. Moreover, these are also able to reduce the migration and invasion of lung carcinoma A549 cells though not showing cytotoxicity. Further studies of some of these active compounds may help to use them in lung cancer as adjuvant therapeutic agents.

Acknowledgments

Authors are thankful to the All India Council for Technical Education (AICTE) (Grant No. 8023/RID/RPS-53/2011-12), New Delhi and the Council of Scientific and Industrial Research (CSIR) (Grant

No. 02(0037)/11/EMR-II), New Delhi as well as University Grants Commission (UGC) (Grant No. 41-747/2012(SR)), New Delhi for providing financial support. NA is grateful to UGC, New Delhi for providing Rajiv Gandhi National Fellowship. AKH thanks CSIR, New Delhi for providing senior research fellowship. Authors are also thankful for grant from the University for Potential for Excellence-II (UPE-II) program of Jadavpur University by University Grants Commission (UGC), New Delhi. Authors are thankful to the authority of Jadavpur University, University of Calcutta and Indian Institute of Chemical Biology for providing facilities required for the work. We are thankful to Tapan Kumar Pal, Murari Mohan Pal and Pradipta Sarkar as well as Tapan Kumar Maity and Tanushree Sinha of Department of Pharmaceutical Technology, Jadavpur University for the help to conduct mass and IR spectroscopy, respectively.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2016.07.023.

References and notes

1. Vergova, V.; Pytliak, M.; Mechirova, V. In *Matrix Metalloproteanases*; Gupta, S. P., Ed., 1st ed.; Springer: Basel, 2012; pp 1–33. Chapter 4.
2. Devel, L.; Garcia, S.; Czarny, B.; Beau, F.; Lajeunesses, E.; Vera, L.; Georgiadia, D.; Stura, E.; Dive, V. *J. Biol. Chem.* **2010**, *285*, 35900.
3. Devel, L.; Beau, F.; Amoura, M.; Vera, L.; Cassar-Lajeunesse, E.; Garcia, S.; Czarny, B.; Stura, E. A.; Dive, V. *J. Biol. Chem.* **2012**, *287*, 26647.
4. Gialeli, C.; Theocharis, A. D.; Karamanos, N. K. *FEBS J.* **2011**, *278*, 16.
5. Kessenbrock, K.; Plaks, V.; Werb, Z. *Cell* **2011**, *141*, 52.
6. Defour, A.; Overall, C. M. *Trends Pharm. Sci.* **2013**, *34*, 233.
7. Fabre, B.; Filipiak, K.; Zapico, J. M.; Diaz, N.; Carbajo, R. J.; Schott, A. K.; Martinez-Alcazar, M. P.; Suarez, D.; Pineda-Lucena, A.; Ramos, A.; de Pascual-Teresa, B. *Org. Biomol. Chem.* **2013**, *11*, 6623.
8. Overall, C. M.; Kleifeld, O. *Nat. Rev. Cancer* **2006**, *6*, 227.
9. Serra, P.; Bruczko, M.; Zapico, J. M.; Puckpwska, A.; Garcia, M. A.; Martin-Santamaria, S.; Ramos, A.; de Pascual-Teresa, B. *Curr. Med. Chem.* **2012**, *19*, 1036.
10. Zapico, J. M.; Puckpwska, A.; Filipiak, K.; Coderch, C.; de Pascual-Teresa, B.; Ramos, A. *Org. Biomol. Chem.* **2015**, *13*, 142.
11. Dufour, A.; Overall, C. M. In *Extracellular Matrix: Pathobiology and Signaling*; Karamanos, N., Ed.; Walter De Gruyter: Gottingen, 2012; pp 377–400. Chapter 4.7.
12. Fabre, B.; Filipiak, K.; Diaz, N.; Zapico, J. M.; Suarez, D.; Ramos, A.; de Pascual-Teresa, B. *Chem. Biol. Chem.* **2014**, *15*, 399.
13. Hamano, Y.; Zeisberg, M.; Sugimoto, H.; Lively, J. C.; Maeshima, Y.; Yang, C.; Hynes, R. O.; Werb, Z.; Sudhakar, A.; Kalluri, R. *Cancer Cell* **2003**, *3*, 589.
14. Shchors, K.; Nozawa, H.; Xu, J.; Rostker, F.; Swigart-Brown, L.; Evan, G.; Hanahan, D. *Oncogene* **2013**, *32*, 502.
15. Cathcart, J.; Pulkoski-Gross, A.; Zucker, S.; Cao, J. In *Matrix Metalloproteinase Biology*; Sagi, I., Gaffney, J. P., Eds.; Wiley-Blackwell: New Jersey, 2015; pp 85–102. Chapter 5.
16. Hesterkamp, T.; Whittaker, M. *Curr. Opin. Chem. Biol.* **2008**, *12*, 260.
17. Li, S.; Chen, J.; Xu, H.; Long, J.; Xie, X.; Zhang, Y. *Plos One* **2014**, *9* e110670.
18. Fray, M. J.; Burslem, M. F.; Dickinson, R. P. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 567.
19. Fray, M. J.; Dickinson, R. P. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 571.
20. Robinson, R. P.; Laird, E. R.; Donahue, K. M.; Lopresti-Morrow, L. L.; Mitchell, P. G.; Reese, M. R.; Reeves, L. M.; Rouch, A. I.; Stam, E. J.; Yocum, S. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1211.
21. Reiter, L. A.; Mitchell, P. G.; Martinelli, G. J.; Lopresti-Morrow, L. L.; Yocum, S. A.; Eskra, J. D. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2331.
22. Reiter, L. A.; Robinson, R. P.; McClure, K. F.; Jones, C. S.; Reese, M. R.; Mitchell, P. G.; Otterness, I. G.; Bliven, M. L.; Liras, J.; Cortina, S. R.; Donahue, K. M.; Eskra, J.; Griffiths, R. J.; Lame, M. E.; Lopez-Anaya, A.; Martinelli, G. J.; McGahee, S. M.; Yocum, S. A.; Lopresti-Morrow, L. L.; Tobiassen, L. M.; Vaughn-Bowser, M. L. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3389.
23. Noe, M. C.; Snow, S. L.; Wolf-Gouveia, L. A.; Mitchell, P. G.; Lopresti-Morrow, L.; Reeves, L. M.; Yocum, S. A.; Liras, J. L.; Vaughn, M. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4727.
24. Blagg, J. A.; Noe, M. C.; Wolf-Gouveia, L. A.; Reiter, L. A.; Laird, E. R.; Chang, S. P.; Danley, D. E.; Downs, J. T.; Elliott, N. C.; Eskra, J. D.; Griffiths, R. J.; Hardink, J. R.; Haugeto, A. I.; Jones, C. S.; Liras, J. L.; Lopresti-Morrow, L. L.; Mitchell, P. G.; Pandit, J.; Robinson, R. P.; Subramanyam, C.; Vaughn-Bowser, M. L.; Yocum, S. A. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1807.
25. Noe, M. C.; Natarajan, V.; Snow, S. L.; Wolf-Gouveia, L. A.; Mitchell, P. G.; Lopresti-Morrow, L.; Reeves, L. M.; Yocum, S. A.; Otterness, I.; Bliven, M. A.; Carty, T. J.; Barberia, J. T.; Sweeney, F. J.; Liras, J. L.; Vaughn, M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3385.
26. Dublanchet, A.; Ducrot, P.; Andrianjara, C.; O'Gara, M.; Morales, R.; Compere, D.; Denis, A.; Blais, S.; Cluzeau, P.; Courte, K.; Hamon, J.; Moreau, F.; Prunet, M. L.; Tertre, A. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3787.
27. Whitlock, G. A.; Dack, K. N.; Dickinson, R. P.; Lewis, M. L. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6750.
28. Bailey, S.; Fish, P. V.; Billotte, S.; Bordner, J.; Greiling, D.; James, K.; McElroy, A.; Mills, J. E.; Reed, C.; Webster, R. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6562.
29. Kolodziej, S. A.; Hockerman, S. L.; DeCrescenzo, G. A.; McDonald, J. J.; Mischke, D. A.; Munie, G. E.; Fletcher, T. R.; Stehle, N.; Swearingen, C.; Becker, D. P. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3561.
30. Kolodziej, S. A.; Hockerman, S. L.; Boehm, T. L.; Carroll, J. N.; DeCrescenzo, G. A.; McDonald, J. J.; Mischke, D. A.; Munie, G. E.; Fletcher, T. R.; Rico, J. G.; Stehle, N. W.; Swearingen, C.; Becker, D. P. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3557.
31. Fobian, Y. M.; Freskos, J. N.; Barta, T. E.; Bedell, L. J.; Heintz, R.; Kassab, D. J.; Kiefer, J. R.; Mischke, B. V.; Molyneaux, J. M.; Mullins, P.; Munie, G. E.; Becker, D. P. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2823.
32. Barta, T. E.; Becker, D. P.; Bedell, L. J.; Easton, A. M.; Hockerman, S. L.; Kiefer, J.; Munie, G. E.; Mathis, K. J.; Li, M. H.; Rico, J. G.; Villamil, C. I.; Williams, J. M. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2820.
33. Wu, Y.; Li, J.; Wu, J.; Morgan, P.; Xu, X.; Rancati, F.; Vallese, S.; Raveglia, L.; Hotchandani, R.; Fuller, N.; Bard, J.; Cunningham, K.; Fish, S.; Krykbaev, R.; Tam, S.; Goldman, S. J.; Williams, C.; Mansour, T. S.; Saiah, E.; Sypek, J.; Li, W. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 138.
34. Robinson, R. P.; Laird, E. R.; Blake, J. F.; Bordner, J.; Donahue, K. M.; Lopresti-Morrow, L. L.; Mitchell, P. G.; Reese, M. R.; Reeves, L. M.; Stam, E. J.; Yocum, S. A. *J. Med. Chem.* **2000**, *43*, 2293.
35. Fray, M. J.; Dickinson, R. P.; Huggins, J. P.; Occleston, N. L. *J. Med. Chem.* **2000**, *46*, 3514.
36. Fish, P. V.; Allan, G. A.; Bailey, S.; Blagg, J.; Butt, R.; Collis, M. G.; Greiling, D.; James, K.; Kendall, J.; McElroy, A.; McCleverty, D.; Reed, C.; Webster, R.; Whitlock, G. A. *J. Med. Chem.* **2007**, *50*, 3442.
37. Becker, D. P.; Barta, T. E.; Bedell, L. J.; Boehm, T. L.; Bond, B. R.; Carroll, J.; Carron, C. P.; Decrescenzo, G. A.; Easton, A. M.; Freskos, J. N.; Funckes-Shippy, C. L.; Heron, M.; Hockerman, S.; Howard, C. P.; Kiefer, J. R.; Li, M. H.; Mathis, K. J.; McDonald, J. J.; Mehta, P. P.; Munie, G. E.; Sunyer, T.; Swearingen, C. A.; Villamil, C. I.; Welsch, D.; Williams, J. M.; Yu, Y.; Yao, J. *J. Med. Chem.* **2010**, *53*, 6653.
38. Wilson, L. J.; Wang, B.; Yang, S. M.; Scannevin, R. H.; Burke, S. L.; Karnachi, P.; Rhodes, K. J.; Murray, W. V. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 6485.
39. Yang, S. M.; Scannevin, R. H.; Wang, B.; Burke, S. L.; Huang, Z.; Karnachi, P.; Wilson, L. J.; Rhodes, K. J.; Lagu, B.; Murray, W. V. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1140.
40. Yang, S. M.; Scannevin, R. H.; Wang, B.; Burke, S. L.; Wilson, L. J.; Karnachi, P.; Rhodes, K. J.; Lagu, B.; Murray, W. V. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1135–1139.
41. Zhang, Y. M.; Fan, X.; Yang, S. M.; Scannevin, R. H.; Burke, S. L.; Rhodes, K. J.; Jackson, P. F. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 405.
42. Zhang, Y. M.; Fan, X.; Chakaravarty, D.; Xiang, B.; Scannevin, R. H.; Huang, Z.; Ma, J.; Burke, S. L.; Karnachi, P.; Rhodes, K. J.; Jackson, P. F. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 409.
43. Zhang, Y. M.; Fan, X.; Xiang, B.; Chakravarty, D.; Scannevin, R.; Burke, S.; Karnachi, P.; Rhodes, K. J.; Jackson, P. F. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3096.
44. Accelrys Inc. *Discovery Studio 3.0, San Diego, USA*, 2011.
45. Halder, A. K.; Saha, A.; Jha, T. *J. Pharm. Pharmacol.* **2013**, *65*, 1541.
46. Ojha, P. K.; Roy, K. *Eur. J. Med. Chem.* **2010**, *45*, 4645.
47. *DRAGON Web version 2.1*; Milano Chemometrics and QSAR Research group: Italy, 2000.
48. Canvas *version 1.8*; Schrödinger, LLC: New York, NY, 2013.
49. Adhikari, N.; Halder, A. K.; Saha, A.; Saha, K. D.; Jha, T. *Tox. In Vitro* **2015**, *29*, 1392.
50. Tetko, I. V.; Tanchuk, V. Y.; Villa, A. E. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1407.
51. Roy, P. P.; Paul, S.; Mitra, I.; Roy, K. *Molecules* **2009**, *14*, 1660.
52. Golbraikh, A.; Tropsha, A. *J. Mol. Graph. Model.* **2002**, *20*, 269.
53. Roy, P. P.; Roy, K. *QSAR Comb. Sci.* **2008**, *27*, 302.
54. Hemmatateenejad, B. *J. Chemometr.* **2004**, *18*, 475.
55. Ojha, P. K.; Roy, K. *Chemometr. Intell. Lab Sys.* **2011**, *109*, 146.
56. Jaiswal, M.; Khadikar, P. V.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3283.
57. Shapiro, S.; Guggenheim, B. *Quant. Struct. Act. Relat.* **1998**, *17*, 327.
58. Tropsha, A.; Gramatica, P.; Gombar, V. K. *QSAR Comb. Sci.* **2003**, *22*, 69.
59. Li, H.; Sutter, J.; Hoffmann, R. In *Pharmacophore Perception, Development, and use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 173–189. Chapter 10.
60. Debnath, A. K. *J. Med. Chem.* **2002**, *45*, 41.
61. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
62. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
63. SYBYL-X 2.0., Tripos Inc., St Louis, MO 63144, USA.
64. Halder, A. K.; Saha, A.; Jha, T. *Mol. Divers.* **2013**, *17*, 123.
65. Bohm, H. J. *J. Comput. Aided Mol. Des.* **1992**, *6*, 61.
66. Cho, A. E.; Rinaldo, D. *J. Comp. Chem.* **2009**, *30*, 2609.
67. Librando, V.; Pappalardo, M. *J. Mol. Graph. Model.* **2013**, *44*, 1.
68. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
69. *Prime Version 3.3*; Schrödinger, LLC: New York, 2013.
70. Das, D.; Koh, Y.; Tojo, Y.; Ghosh, A. K.; Mitsuya, H. *J. Chem. Inf. Model.* **2009**, *49*, 2851.

71. Lyne, P. D.; Lamb, M. L.; Saeh, J. C. *J. Med. Chem.* **2006**, *49*, 4805.
72. Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Scalable algorithms for molecular dynamics simulations on commodity clusters. In *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, Florida, November 11–17*, 2006.
73. Pradhan, D.; Priyadarshini, V.; Munikumar, M.; Swargam, S.; Umamaheswari, A.; Bitla, A. J. *Biomol. Str. Dyn.* **2014**, *32*, 171.
74. Halder, A. K.; Saha, A.; Saha, K. D.; Jha, T. J. *Biomol. Str. Dyn.* **2015**, *33*, 1756.
75. Halder, A. K.; Mallick, S.; Shikha, D.; Saha, A.; Saha, K. D.; Jha, T. *RSC Adv.* **2015**, *5*, 72373.
76. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat. Methods* **2012**, *9*, 671.
77. Viswanadhan, V. M.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. *J. Chem. Inf. Comp. Sci.* **1989**, *29*, 163.
78. Kruszewski, J.; Krygowsk, T. M. *Tetrahedron Lett.* **1972**, *36*, 3839.
79. Stanton, D. T.; Jurs, P. C. *Anal. Chem.* **1990**, *62*, 2323.
80. Ivanciuc, O.; Ivanciuc, T.; Balaban, A. T. *J. Chem. Inf. Comp. Sci.* **1998**, *38*, 395.
81. Schrödinger, L. L. C. *QikProp, version 3.5., New York, NY*, 2012.
82. Hemmer, M. C.; Steinhauer, V.; Gasteiger, J. *Vib. Spectrosc.* **1999**, *19*, 151.
83. Klon, A. E.; Lowrie, J. F.; Diller, D. J. *J. Chem. Inf. Model.* **2006**, *46*, 1945.
84. Prathipati, P.; Ma, N. L.; Keller, T. H. *J. Chem. Inf. Model.* **2008**, *48*, 2363.
85. Triballeau, N.; Acher, F.; Brabet, I.; Pin, I. J.; Bertrand, H. O. *J. Med. Chem.* **2005**, *48*, 2534.
86. Feng, Y.; Likos, J. J.; Zhu, L.; Woodward, H.; Munie, G.; McDonald, J. J.; Stevens, A. M.; Howard, C. P.; De Crescenzo, G. A.; Welsch, D.; Shieh, H. S.; Stallings, W. C. *Biochim. Biophys. Acta* **2002**, *1598*, 10.
87. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *46*, 3.
88. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615.
89. Wang, Y.; Xie, S.; Liu, C.; Wu, Y.; Liu, Y.; Cai, Y. *Nutr. Cancer* **2012**, *64*, 627.
90. Qu, X. J.; Yuan, Y. X.; Tian, Z. G.; Xu, W. F.; Chen, M. H.; Cui, S. X.; Guo, Q.; Gai, R.; Makuuchi, M.; Nakata, M.; Tang, W. *Int. J. Mol. Med.* **2006**, *18*, 609.