IJPSR (2021), Volume 12, Issue 6 (Research Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 18 June 2020; received in revised form, 07 November 2020; accepted, 04 May 2021; published 01 June 2021

2D QSAR ANALYSIS OF DIPEPTIDE NITRILE BASED CATHEPSIN S INHIBITORS

Sneha Kushwaha * and Sarvesh Kumar Paliwal

Department of Pharmacy, Banasthali University, Banasthali, Tonk - 304022, Rajasthan, India.

Keywords:

QSAR, Multiple Linear Regression, TSAR, Partial Least Square, dipeptide nitrile

Correspondence to Author:
Sneha Kushwaha

Assistant Professor,
Faculty of Pharmaceutical Sciences,
Rama University, NH-91, GT Road,
Mandhana, Kanpur - 209217, Uttar Pradesh, India.

E-mail: sneha.kush09@gmail.com

ABSTRACT: Cathepsin S enzyme has been considered as an evolving target for the development of novel therapeutic agents for the treatment of numerous autoimmune disorders and other inflammatory diseases. Using TSAR 3.3 2D QSAR has been performed on a series of dipeptide nitrile nucleus. Attempts have been made to derive and comprehend a correlation between biological activity and the generated descriptors. The study was carried out using 37 compounds by division into training and test set by a random selection method. A final QSAR model was generated from a set of 28 compounds with the Leave-out one row (LOO) method of cross-validation to estimate the model's predictive ability. The most significant model with n = 28, r = 0.969, r^2 = 0.939, r^2cv = 0.801, s value = 0.35, f value = 89.07 was developed using MLR analysis. For PLS, the fraction of variance explained = 0.928 was observed. A comparable PLS model with r^2 = 0.928 and Neural model with r^2 = 0.962 indicated good internal predictability of the model. External test set validation provided r^2 values of 0.721 and 0.821 for MLR and PLS analysis, respectively. QSAR model indicated the importance of Steric [Verloop B1 (Subs. 4)], Geometrical [Inertia moment 1 length (Subs. 4), topological [kier Chi V0 (atoms) index (Subs. 2)], and [Kier Chi 4 (path) index (Subs.4)] descriptors for the activity of Cathepsin S inhibitors. This study will be effective in the design of novel and more potent Cathepsin S inhibitors.

INTRODUCTION: The term "Cathepsin" was derived from the Greek word "Kathepsin" which means "digesting" [1, 2]. The human genome consists of a total of 11 human cysteine cathepsins [3]. They are cathepsins L, V, S, K, and F (endopeptidases), cathepsins X, B, C, and H (exopeptidases), and cathepsins O, and W of unknown category [1, 4, 5, 6]. Cathepsin S (gene symbol: CTSS), non-glycosylated cysteine proteinases belong to clan C1 (Papain family) [7, 8]. These are found intracellularly in the endolysosomal vesicles [1, 3, 9].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.12(6).3391-02

This article can be accessed online on www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.12(6).3391-02

These are majorly found in dendritic cells, macrophages, spleen, lymph nodes, monocytes, and/or thymic cortical epithelial cells [10, 11]. The enzyme has an integral role in antigen processing and presentation [12, 13, 14]. Their exclusive dispersal pattern specifies its profound contribution to the immune response [1].

All cysteine proteases are composed of three units- a signal peptide (10-20 amino acids long), a propeptide (variable lengths), and a catalytic domain (214-260 amino acids long) [15]. Signal peptides are responsible for the translocation into the endoplasmic reticulum during mRNA translation. Propeptides act as a skeleton to stimulate the folding of the catalytic domain. It acts as a chaperone to carry the proenzyme to the lysosomal compartment. It acts as a high-affinity reversible inhibitor to block the premature activation of the

catalytic domain. The catalytic domain represents the mature, proteolytically active enzyme and it's extremely preserved active site consists of Cysteine, Histidine, and Asparagine residues [1].

The structure of Cathepsin S was first disclosed by McGrath *et al*., 1998. It is a single chain mono-meric protein (217 amino acids) with a molecular mass of 30kDa. consisting of two domains. The left domain comprises of residues 12-111, and 208-211 with helices ranging from 25-40, 50-56, and 68-78. The right domain is based on a six-stranded β-barrel motif, residues 1-11, and 112-207, with a small helix coiling through residues 119-127, additional helix from residues 139-143. The active site lies in between the two domains and contains the residues Cys25 and His159 [16, 17, 18].

The optimal activity of cathepsins requires acidic pH. Cys25 as a catalyst forms thiolate ± imidazolium ion -pair with Hiatidine -159 at very low pka values (~2.5-3.5) [17, 19]. The thiolate ion acts as a nucleophile for the attack of the carbonyl carbon of the sessile peptide bond, which results in the release of the amine product. The acyl-enzyme, after reacting with water, releases the carboxyl product regenerating the free enzyme [4, 18].

Cysteine cathepsin S have a significant role in the growth and progression of various inflammation-associated diseases such as cancer [15, 20, 21, 22, 23, 24, 25], arthritis [18, 26], periodontitis [27], psoriasis [18, 28], various lung diseases [29, 30, 31, 32, 33, 34, 35], cardiovascular disease in patients with chronic kidney disease [36, 37, 38, 39, 40], bone [41], Sjogren's Syndrome [42, 43] and immune disorders [44]. Cathepsin S inhibitors also acts as immunomodulator [45]. Subsequently, there is a need to progress research efforts focused on cathepsin S use in diagnostics and as therapeutic targets in diseases [46, 47]. Dipeptidyl nitrile inhibitor of cathepsin S has proved emerging target for abrogation of tumour [25].

Rheumatoid arthritis (RA) is an autoimmune inflammatory disease of unknown etiology affecting all synovium joints. The genes encoding the major histocompatibility (MHC) molecule are clustered on a small segment of chromosome 6 in humans. MHC Complex or Human Leukocyte Antigen Complex (HLA) molecule plays a central role in the pathology of RA [48]. Antigen-presenting cells (APCs) engulfs the antigen. Peroxides inside the APCs break down the antigen into small portions [49]. The molecular mechanism commences with the MHC II αβ heterodimers synthesis in the endoplasmic reticulum and association of a protein, namely, the invariant chain (Li) in the peptide-binding cleft. The αβLi complex gets relocated to the lysosome, where a portion of the Li gets sliced by cathepsin S, leaving a short fragment- CLIP in the active site. It prevents any premature binding of antigenic peptides [50, 51, 52, 53, 54, 55]. Another protein HLA-DM assists in the release of CLIP from the MHC protein, which provides the binding site for the peptide fragments. The complex is transferred to the cell surface after binding to the MHC II molecule [50, 56]. This complex is exposed to T-cells (CD4 cells *i.e.*, T-helper cell). The T-cell receptor (TCR) recognizes and binds to and causes APCs to secrete cytokines like IL-1, IFN-α, IFN-γ, TNF, and other factors. These, in turn, activate lymphocytes and other immune cells to respond to the antigens, thus causing inflammation [18, 34, 48].

Quantitative Structure-Activity Relationship (QSAR) technique has been used in the modeling of biological activity and calculating ADME/ Toxicity properties [57]. A QSAR model correlates the structure/chemical characteristics of the molecule with their biological activities through mathematical equations. This relation is useful in designing more potent compounds. Biological activities can be predicted for new entities [58]. A QSAR study is significant in enzyme inhibition studies and the identification of the important active sites in the receptor. Thus, a QSAR study is emerging as an important tool in drug design [59, 60].

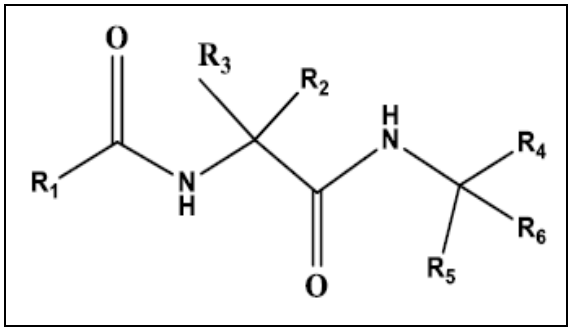
In the present paper, 2D QSAR analysis has been used because of its simplicity and fewer errors. It is more advantageous than 3D QSAR analysis as it does not involve any conformational search or structural alignment [61, 62]. Since structural descriptors encode all the chemical information [63], 2D methodology has been considered superior over 3D QSAR [61, 64, 65].

MATERIALS AND METHODS: The project was completed at Banasthali Vidyapith University, Jaipur, Rajasthan by Sneha Kushwaha, as a part of her M.Pharma project from July 2013 to June 2014. QSAR model has been developed using 37 congeneric molecules using Multiple Linear

Regression (MLR), Partial Least Squares (PLS), and Artificial neural network (ANN) [66].

Generation of 3-Dimensional Chemical Structures and their Optimization: All the chemical structures of dipeptide nitrile derivatives stated in the literature [67] and reported in **Table 1** were sketched using CHEM DRAW ULTRA 12.0 software. As compounds 5, 7, and 8 had uncertain IC_{50} values, and 27 was a racemic mixture, all four were excluded from the series.

TABLE 1: STRUCTURE OF CATHEPSINS INHIBITORS USED FOR QSAR ANALYSIS



Compd	R1	R2	R3	R4	R5	R6
6	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2CH_2Ph
9	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	H
10	morpholin-4-yl	CH_2(*i*-Pr)	H	CCH	H	H
12	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_3
13	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	*n*-Pr
14	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	*n*-Bu
15	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	*i*-Pr
16	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	*t*-Bu
17	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	Ph
18	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2Ph
19	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2(3,4-diCl)Ph
20	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	CH_2CH_2Ph	H
21	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	CH_3	CH_3
22	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2OCH_2Ph
23	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2OCH_2(*o*-Cl)Ph
24	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2OCH_2(*m*-Cl)Ph
25	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2OCH_2(*p*-Cl)Ph
26	morpholin-4-yl	H	CH_2(*i*-Pr)	CN	H	CH_2CH_2Ph
28	morpholin-4-yl	CH_3	H	CN	H	CH_2CH_2Ph
29	morpholin-4-yl	*n*-Bu	H	CN	H	CH_2CH_2Ph
30	morpholin-4-yl	*i*-Pr	H	CN	H	CH_2CH_2Ph
31	morpholin-4-yl	*c*-Hex	H	CN	H	CH_2CH_2Ph
32	morpholin-4-yl	CH_2Ph	H	CN	H	CH_2CH_2Ph
33	morpholin-4-yl	CH_2(*t*-Bu)	H	CN	H	CH_2CH_2Ph
34	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	H	CH_2CH_2Ph
35	morpholin-4-yl	CH_2(*t*-Bu)	H	CN	H	CH_2OCH_2Ph
36	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2Ph
37	morpholin-4-yl	CH_2(*i*-Pr)	H	CN	H	CH_2OCH_2(*o*-Me)Ph
38	morpholin-4-yl	CH_2(*t*-Bu)	H	CN	H	CH_2OCH_2(*o*-Me)Ph
39	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2(*o*-Me)Ph
40	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2(*o*-Cl)Ph
41	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	H	Ph
42	morpholin-4-yl	CH_2(*c*-Hex)	H	CN	CH_3	CH_3
43	pyridin-4-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2Ph
44	furan-2-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2Ph
45	thein-2-yl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2Ph
46	pyrazinyl	CH_2(*c*-Hex)	H	CN	H	CH_2OCH_2Ph

The chemical structures were then imported on the TSAR worksheet (version 3.3, Accelrys Inc., Oxford, England). The series consisted of six main substituents which were defined by the "define substituent" option in the TSAR worksheet. Molecular properties and receptor-ligand interaction depend on the connectivity of the atoms in a molecule and its 3D-Structure. The three-dimensional structure defines the physical, chemical, and biological properties of the molecule.

The molecules and their substituents are converted from 2D to 3D by using CORINA. The partial atomic charge of the molecule was calculated by using the "charge-2-derive charges" option, which is essential for several structural manipulations. By using COSMIC, 3D optimization of the structures was done. Low-energy conformation gets generated for each input structure by default [68]. COSMIC parameters include various parameters like valence terms (bond potentials, bond angle potentials, and torsional potentials) and non-bonded terms (electrostatic interaction and van der Waals interaction) [69]. Summing of all these parameters calculated total molecular energies using COSMIC. The force-field applied by COSMIC for energy calculations approves that only the additional energetically genuine confirmation is considered [70].

Molecular Descriptors Calculation: Molecular descriptor generates a link between chemical structure and biological activity. Descriptors map the chemical structure into a set of binary/ numerical values illustrating numerous molecular properties essential for explaining molecular property/ activity [71]. Descriptors are classified into different properties such as electronic, geometric, hydrophobic, and topological [72].

The activity data of 37 compounds have been imported into the TSAR worksheet after the experimental IC_{50} values have been converted to log ($1/IC_{50}$) [73]. This is done to obtain higher values for more effective analogs [70]. A total of 280 molecular descriptors, including Molecular attributes, Molecular Indices-Topological, Connectivity, Shape Indices, Atom Counts, and VAMP were generated using TSAR 3.3 software. These descriptors help in generating a good QSAR model [74, 75].

Data Reduction: Large data sets may increase the risk of overfitting; thus data must be minimized to reduce the risk of chance correlation. Descriptors with constant values were eliminated. A pair-wise correlation reduction method has been used to reduce data. The retained descriptors had a higher correlation with the biological activity and the least intercorrelation ($r^2 > 0.5$) [76]. Forward and backward elimination methods were used for the inclusion or rejection of descriptors. This was done based on t-values, the descriptors with poorer t-values were rejected [70]. After data reduction, four independent molecular descriptors- Verloop B1 (Subst.4), Inertia moment 1 length (Subst.4), Kier Chi V0 (atoms) index (Subst.2), and Kier Chi 4 (path) index (Subst.4) were left with high correlation with the dependent variable i.e., the Biological activity.

Data Set Preparation: The structures of the series were randomly divided into a training set consisting of 28 compounds and a test set with 9 compounds. The training set produced linear models relating to the structures and the biological activity. The molecules of the test set checked the predictive power of the developed model [70].

Model Development and Validation: Models can be linear or non-linear. Linear models are the backbone of QSAR methodology. They include Multiple Linear Regression Analysis (MLR) and Partial Least Square Analysis (PLS). The non-Linear Model includes an Artificial Neural Network (ANN). MLR has been carried out to produce the leading QSAR model.

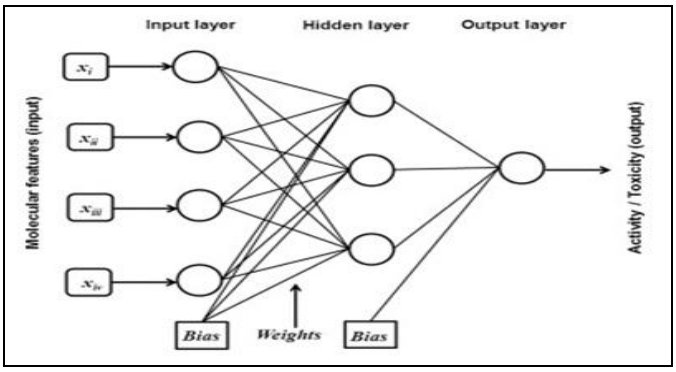
Several MLR models were created using the lasting descriptors as independent variables and biological activity data as dependent variables. These were used to compute the relationship between the variables. The models were generated in the form of a regression equation that described the activity data and was further used to predict the activity of new compounds. Positive values of the regression coefficient state that the given descriptor is positively correlated to the biological activity i.e., increase in the value of descriptor lead to enhancement in the activity value and vice-versa.

Statistical significance of the regression equations is tested based on the regression coefficient (r^2), Fischer's statistic (F), and the standard error of estimate [70].

The generated model is validated both internally and externally. Cross-validation analysis -leave one out (LOO) method was used for internal validation. External validation was done by using the model developed by the training set. Activities of the test set molecules were predicted by this method [77].

PLS is a multivariate analysis based on the principal component analysis. It gives the maximum correlation between the principal components (independent variables) and the dependent variable (activity) through linear equation [78]. PLS analysis

was performed on the same training set compounds to check the robustness and predictive ability of the models generated by MLR. The model generated during PLS was also validated using LOO method [79].



Artificial neural networks (ANNs) are computer-based mathematical models developed to have functions analogous to idealized simple biological nervous systems. They consist of layers of processing elements (neurodes), considered analogous to the nerve cells (neurons), and interconnected to form a network [80].

Dipeptide nitrile inhibitors also possessed a suitable pharmacokinetic (ADME) profile which is reported in the table.

Assessment of Druggability: To understand the pharmacodynamics and pharmacokinetics of a chemical entity, the knowledge of absorption, distribution, metabolism, and excretion is significant. For this, the violation of Lipinski's rule of five has been checked. According to this rule, H-bond donors should be less than five, H-bond acceptors should be less than ten, clog P (calculated log P) should be less than five, and molecular weight should be less than 500 Da. for excellent oral absorption of a compound. This calculation was done with the aid of "ADME check" option in the TSAR worksheet. **Table 2** shows the values of the calculated parameters for Lipinski's rule of five [70].

TABLE 2: VALUES OF THE CALCULATED PARAMETERS FOR LIPINSKI'S RULE OF FIVE

Comp. Name	ADME (Molecular weight)	ADME(H-Bond Acceptors)	ADME(H-Bond Donors)	ADME (Log P)	ADME Voilations
6	386.55	4	2	2.077	0
9	282.39	4	2	-0.542	0
10	281.4	3	2	0.023	0
12	296.42	4	2	-0.004	0
13	324.48	4	2	0.860	0
14	338.51	4	2	1.256	0
15	324.48	4	2	0.866	0
16	338.51	4	2	1.372	0
17	358.49	4	2	1.429	0
18	372.52	4	2	1.681	0
19	441.4	4	2	2.717	0
20	386.55	4	2	2.077	0
21	310.45	4	2	0.197	0
22	402.55	5	2	1.264	0
23	436.99	5	2	1.782	0
24	436.99	5	2	1.782	0
25	436.99	5	2	1.782	0
26	386.55	4	2	2.077	0
28	344.46	4	2	0.882	0
29	386.55	4	2	2.143	0
30	372.52	4	2	1.753	0
31	412.59	4	2	2.438	0
32	420.56	4	2	2.567	0
33	400.58	4	2	2.510	0
34	426.62	4	2	2.762	0
35	416.58	5	2	1.697	0
36	442.62	5	2	1.949	0
37	416.58	5	2	1.731	0
38	430.61	5	2	2.165	0
39	456.65	5	2	2.416	0
40	477.06	5	2	2.467	0

41	398.56	4	2	2.114	0
42	350.52	4	2	0.882	0
43	434.59	5	2	3.136	0
44	423.56	5	2	2.931	0
45	439.62	4	2	3.274	0
46	435.58	6	2	2.688	0

RESULTS: MLR was performed with 28 compounds in the training set and 9 compounds in the test set. None of the outliers were removed. The statistical values of the regression analysis are listed in **Table 3.**

TABLE 3: STATISTICAL VALUES OBTAINED BEFORE DATA REDUCTION AND AFTER PERFORMING MLR ANALYSIS

S. no.	Statistical tests	Values before data reduction	Values after MLR
1	s value	0.191	0.35
2	f value	147.614	89.07
3	Regression coefficient, r	0.991	0.969
4	r^2	0.982	0.939
5	Cross validation, r^2 (cv)	0.0381	0.801
6	Residual sum of squares	0.955	2.842
7	Predictive sum of squares	53.112	9.325

The value of r^2 (0.939) indicates that the MLR equation accounts for 93.9% variance in the biological activity depicting a quite reasonable fit. The cross-validation regression coefficient is greater than 0.6. The difference between r^2 (0.939) and r^2cv (0.801) is comparatively small which indicates the good internal predictive ability of the model.

Fischer statistic (f) is the measure of the probability of no chance correlation. The value of the F-test (89.07) has been found significant. The standard error (s=0.34) is significantly low for the regression to be significant. It measures the quality of the fit of the model.

Equation 1: Original Equation (By MLR Method)

$$Y = - 2.580 * X1 – 0.001 * X2 + 0.776 * X3 + 0.255 * X4 – 0.697$$

Equation 2: Standardized Equation (By MLR Method)

$$Y= - 0.560 * S1 - 0.528 * S2 + 0.845 * S3 + 0.235 * S4 – 1.915$$

Where X1 is Verloop B1 (Subst. 4), X2 is Inertia Moment 1 Length (Subst. 4), X3 is Kier ChiV0 (atoms) index (Subst. 2), X4 is Kier Chi4 (path) index (subst.4) and Y is the biological activity.

Table 4 represents the Correlation matrix showing a correlation between the biological activity and the molecular descriptors left after data reduction, and **Table 5** represents jacknife se, covariance se, and t-value for the molecular descriptors.

TABLE 4: CORRELATION MATRIX SHOWING CORRELATION BETWEEN THE BIOLOGICAL ACTIVITY AND THE MOLECULAR DESCRIPTORS LEFT AFTER DATA REDUCTION

	X1: Verloop B1 (Subst. 4)	X2: Inertia Moment 1 Length (Subst. 4)	X3: Kier ChiV0 (atoms) Index (Subst. 2)	X4: Kier Chi4 (path) index (Subst. 4)	Log ($1/IC_{50}$) Values
X1: Verloop B1 (Subst. 4)	1	-0.048578	-0.092444	-0.17934	-0.49709
X2: Inertia Moment 1 Length (Subst. 4)	-0.048578	1	-0.08263	-0.15681	-0.46068
X3: Kier ChiV0 (atoms) Index (Subst. 2)	-0.092444	-0.08263	1	0.12889	0.73655
X4: Kier Chi4 (path) index (Subst. 4)	-0.17934	-0.15681	0.12889	1	0.39979
Log ($1/IC_{50}$) Values	-0.49709	-0.46068	0.73655	0.39979	1

TABLE 5: JACKNIFE SE, COVARIANCE SE, AND T-VALUE FOR THE MOLECULAR DESCRIPTORS

Molecular Descriptors	Abbreviation	Jacknife SE	Covariance SE	t-value
Verloop B1 (Subst. 4)	X1	0.98321	0.31855	-8.0999
Inertia Moment 1 Length (Subst. 4)	X2	0.00097454	2.0466e-005	-7.6593
Kier ChiV0 (atoms) index (Subst. 2)	X3	0.0707	0.062936	12.322
Kier Chi4 (path) index (subst.4)	X4	0.082506	0.076224	3.3457
Constant	C	1.7283		

MLR analysis provided acceptable results with r^2 = 0.852 (Training set) and 0.721 (Test set) proposing good external validation.

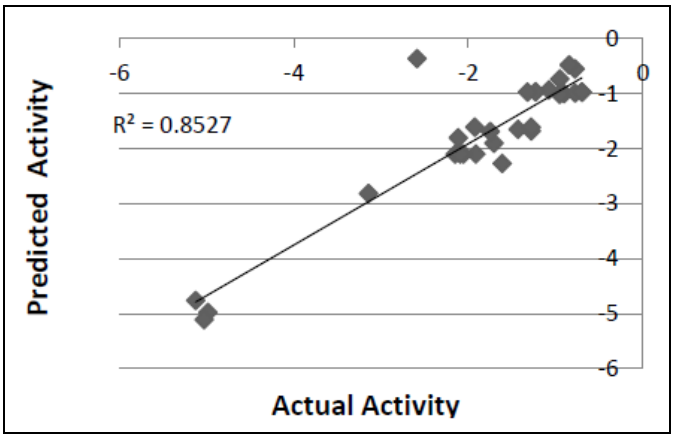


FIG. 1: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TRAINING SET COMPOUNDS DERIVED FROM MLR ANALYSIS

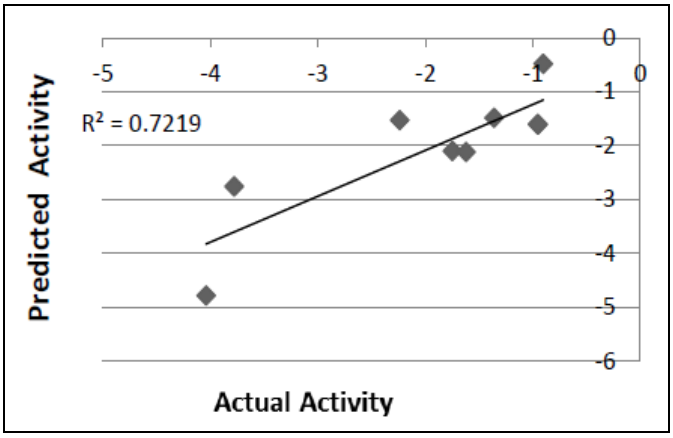


FIG. 2: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TEST SET COMPOUNDS DERIVED FROM MLR ANALYSIS

To confirm the liability of the generated model, the PLS analysis was performed using the same data set.

TABLE 6: STATISTICAL TEST SET VALUES OF THE MODEL DEVELOPED BY PLS ANALYSIS

Statistical significance	Fraction of Variance explained, r^2	r^2cv	Residual sum of squares	Predictive sum of squares
0.99643	0.9276	0.9135	60.28	63.96

PLS showed perfect results with r^2 = 0.928 (Training set) and 0.821 (Test set) which suggested good external prediction. This signifies a 92.8 % variance (greater than 0.6) in the biological activity. A small difference between r^2 and r^2cv predicts the good internal predictive ability of the developed model [81, 82].

Equation 3: Represents the PLS Equation (Dimension 1)

Y = - 2.386 * X1 – 0.0001* X2 +0.704* X3 + 0.453 * X4 – 0.892

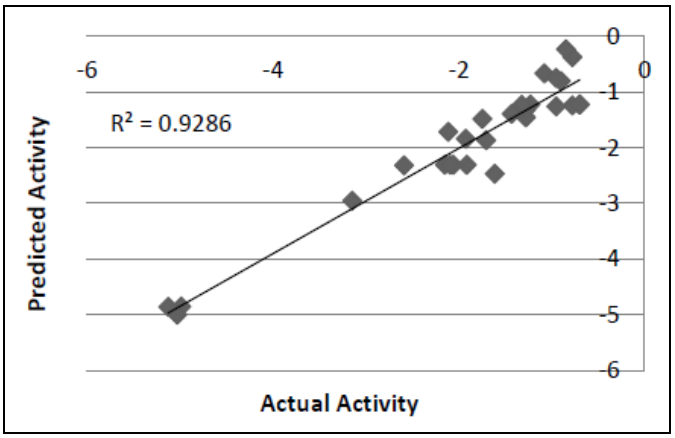


FIG. 3: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TRAINING SET COMPOUNDS DERIVED FROM PLS ANALYSIS

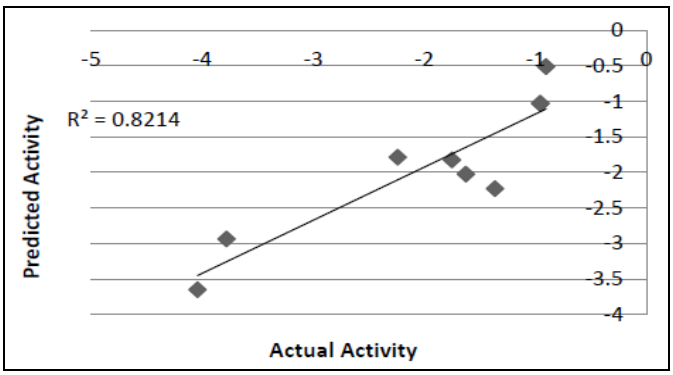


FIG. 4: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TEST SET COMPOUNDS DERIVED FROM PLS ANALYSIS

Further validation was done through ANN.

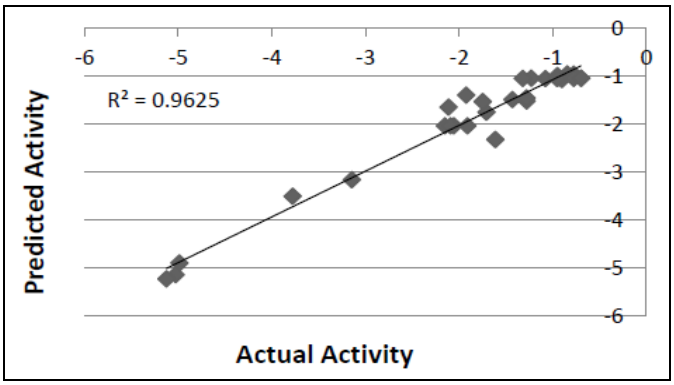


FIG. 5: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TRAINING SET COMPOUNDS DERIVED FROM ANN ANALYSIS

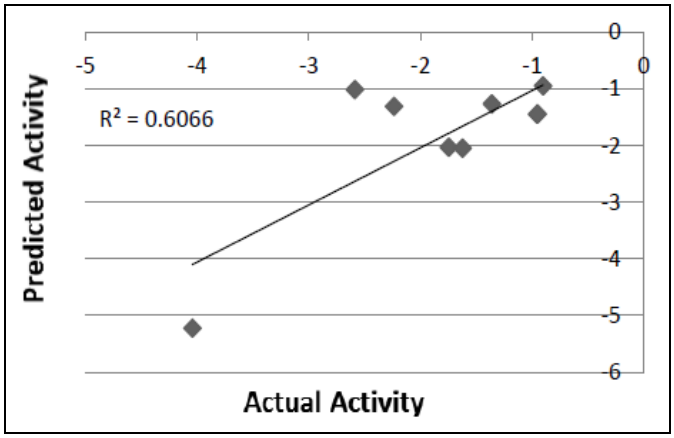


FIG. 6: ACTUAL *VS.* PREDICTED ACTIVITY PLOT FOR THE TEST SET COMPOUNDS DERIVED FROM ANN ANALYSIS

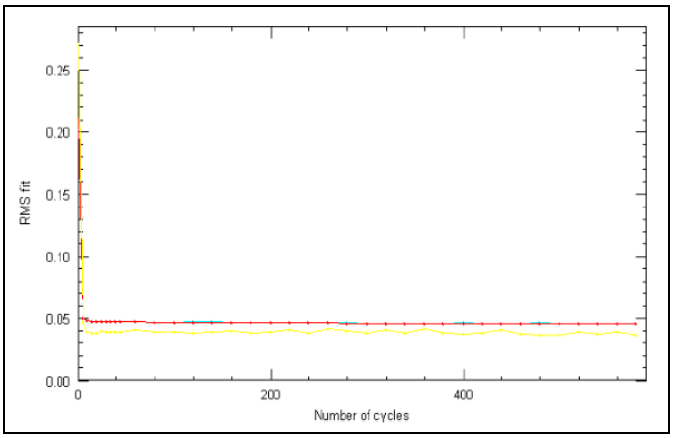


FIG. 7: TYPICAL TRAINING AND VALIDATION ERROR CURVE

Best RMS fit was found to be 0.0458 at 398 cycles. Net configuration was 4-1-1 and test RMS fit was 0.0389.

Verloop B1 (Subst. 4), Inertia Moment 1 Length (Subst. 4), Kier ChiV0 (atoms) index (Subst. 2), and Kier Chi4 (path) index (subst.4) were the inputs and negative log IC50 values were the output for ANN model.

The experimentally determined log ($1/IC_{50}$) values and predicted log ($1/IC_{50}$) for the compounds of training and test set are listed in **Table 7** and **8,** respectively.

TABLE 7: ACTUAL AND PREDICTED VALUES FOR THE TRAINING SET COMPOUNDS OBTAINED FROM MLR, PLS AND FFNN ANALYSIS OF TRAINING SET

S. no.	Comp. Name	Actual Activity (log $1/IC_{50}$)	Predicted Activity		
			MLR	PLS	FFNN
1	9	-2.093	-2.099	-2.310	-2.034
2	10	-4.986	-4.977	-4.854	-4.890
3	12	-2.152	-2.097	-2.308	-2.032
4	13	-1.913	-2.098	-2.308	-2.032
5	15	-3.146	-2.817	-2.952	-3.153
6	16	-5.029	-5.105	-4.999	-5.133
7	17	-1.707	-1.901	-1.864	-1.747
8	18	-2.113	-1.805	-1.716	-1.644
9	19	-1.748	-1.691	-1.485	-1.526
10	20	-2.587	-0.364	-2.319	-1.015
11	21	-2.056	-2.097	-2.308	-2.032
12	22	-1.278	-1.677	-1.455	-1.513
13	25	-1.431	-1.647	-1.393	-1.484
14	28	-5.127	-4.762	-4.861	-5.222
15	29	-1.612	-2.267	-2.467	-2.317
16	31	-1.924	-1.61	-1.840	-1.391
17	34	-0.698	-0.969	-1.226	-1.041
18	35	-0.903	-0.999	-0.805	-1.069
19	36	-0.778	-0.549	-0.373	-0.958
20	37	-1.278	-1.608	-1.315	-1.451
21	38	-1.079	-0.930	-0.665	-1.049
22	39	-0.845	-0.479	-0.233	-0.948
23	41	-0.954	-0.733	-0.747	-0.988
24	42	-1.230	-0.969	-1.225	-1.041
25	43	-0.778	-0.991	-1.245	-1.044
26	44	-1.322	-0.972∫∫	-1.228	-1.041
27	45	-0.698	-0.971	-1.227	-1.041
28	46	-0.954	-1.011	-1.263	-1.048

TABLE 8: ACTUAL AND PREDICTED VALUES FOR THE TEST SET COMPOUNDS OBTAINED FROM MLR, PLS AND FFNN ANALYSIS OF TEST SET

S. no.	Comp. Name	Actual Activity	Predicted Activity		
			MLR	PLS	FFNN
1	6	-1.623	-2.116	-2.026	-2.050
2	14	-1.748	-2.098	-1.824	-2.033
3	23	-0.954	-1.607	-1.030	-1.450
4	24	-0.954	-1.600	-1.029	-1.444
5	26	-4.041	-4.780	-3.654	-5.223
6	30	-3.778	-2.755	-2.937	-3.498
7	32	-2.238	-1.524	-1.788	-1.313
8	33	-1.361	-1.481	-2.228	-1.269
9	40	-0.903	-0.479	-0.511	-0.948

DISCUSSION: The first descriptor is the Verloop B1 (Subst. 4). It is negatively correlated with biological activity. Thus, a decrease in Verloop B1 value would increase biological activity.

The second descriptor is Inertia moment 1 length (Subst. 4). It is a geometrical descriptor that characterizes the mass distribution in a molecule and the susceptibility of a molecule to different rotational transitions. It is negatively correlated with biological activity. Thus, the substituent which increases the mass will decrease the biological activity. Hence, complex and bulky groups must be avoided to have a molecule with increased activity, as well as, its beneficial effects.

The third and fourth descriptors are Kier ChiV0 (atoms) index (Subst.2) and Kier Chi4 (path) index (Subst.4). These are well-known topological indices. They explain the atom's identity, bonding environment, and the number of hydrogen bonds. Precisely they explain the molecular connectivity of a molecule. As they are positively correlated, the presence of such groups is beneficial for biological activity.

CONCLUSION: QSAR study was successfully performed on a series of Dipeptide Nitrile analogs. Significant statistical values of MLR, PLS, and FFNN indicated the robustness of the model. The value of r^2 of 0.852, 0.928, and 0.962 for MLR, PLS, and FFNN (training set) respectively, indicated the soundness of the model. Value of r^2 of 0.721, 0.821, and 0.606 for MLR, PLS and FFNN (test set) respectively indicated better results. According to the classical QSAR models presented in the present work, the four molecular descriptors give predictive information about the overall behavior of the molecules and are considered to be the important contributors to their biological properties.

ACKNOWLEDGEMENT: The authors thank the vice-chancellor for fulfilling all the necessary requirements and Banasthali Vidyapith for providing all the computational resources.

CONFLICTS OF INTEREST: The author(s) confirm that this article's content has no conflict of interest.

REFERENCES:

1. Bromme D and Wilson S: Extracellular matrix degradation. Springer-Verlag Berlin Heidelberg, Canada, Vol. II, 2011: 23-51.
2. Reiser J, Adair B and Reinheckel T: Specialized roles for cysteine cathepsins in health and disease. The Journal of Clinical Investigation 2010; 120: 3421-31.
3. Mohamed MM and Sloane BF: Cysteine cathepsins: multifunctional enzymes in cancer. Nature Reviews Cancer 2006; 6: 764-75.
4. Guha S and Padh H: Cathepsins: Fundamental Effectors of Endolysosomal Proteolysis. Indian Journal of Biochemistry and Biophysics 2008; 45: 75-90.
5. Li YY, Fang J and Ao GZ: Cathepsin B and L inhibitors: a patent review. Expert Opinion on Therapeutic Patents 2016; 27: 643-56.
6. Sena BF, Figueiredo JL and Aikawa E: Cathepsin S As an inhibitor of Cardiovascular inflammation andCalcification in Chronic Kidney Disease. Frontiers in Cardiovascular Medicine 2018; 4: 1-7.
7. Chang WSW, Wu HR, Yeh CT, Wu CW and Chang JY: Lysosomal Cysteine Proteinase Cathepsin S as a Potential Target for Anti-cancer Therapy. Journal of Cancer Molecules 2007; 3: 5-14.
8. Pauly TA, Sulea T, Ammirati M, Sivaraman J, Danley DE, Griffor MC, Kamath AV, Wang IK, Laird ER, Seddon AP, Menard R, Cygler M and Rath VL: Specificity Determinants of Human Cathepsin S Revealed by Crystal Structures of Complexes. Biochemistry 2003; 42: 3203-13.
9. Hsing LC and Rudensky AY: The lysosomal cysteine proteases in the MHC class II antigen presentation. Immunological Reviews 2005; 207: 229-41.
10. Shi GP, Chapman HA, Webb AC, Foster KE, Knoll JH, Lemere CA and Munger JS: Human Cathepsin S: Chromosomal Localization, Gene Structure and Tissue Distribution. Journal of Biological Chemistry 1994; 269: 11530-6.
11. Nakagawa TY and Rudensky AY: The role of lysosomal proteinases in MHC class-II mediated antigen processing

and presentation. Immunological Reviews 1999; 172: 121-9.
12. Riese RJ and Chapman HA: Cathepsins and compartmentalization in antigen presentation. Current opinion in immunology 2000; 12: 107-13.
13. Shi GP, Bryant R, Riese R, Verhelst S, Driessen C, Li Z, Bromme D, Ploegh HL and Chapman HA: Role of cathepsin F in invariant chain processing and major histocompatibility complex class II peptide loading by macrophages. The journal of experimental Medicine 2000; 191: 1177-86.
14. Tolosa E, Li W, Yasuda Y, Wienhold W, Denzin LK, Lautwein A, Driessen C, Schnorrer P, Weber E, Stevanovic S, Kurek R, Melms A and Bromme D: Cathepsin V involved in the degradation of invariant chain in human thymus and is overexpessed in myasthenia gravis. The journal of clinical investigation 2003; 112: 517-26.
15. Huang CC, Lee CC, Lin HH and Chang JY: Cathepsin S attenuates endosomal EGFR signalling: A mechanical rationale for the combination of cathepsin S and EGFR tyrosine kinase inhibitors. Scientific Reports 2016; 6: 1-12.
16. McGrath ME, Palmer JT, Bromme D and Somoza JR: Crystal structure of human cathepsin S. Protein Science 1998; 7: 1294-1302.
17. Vasiljeva O, Reinheckel T, Peters C, Turk D, Turk V and Turk B: Emerging Roles of Cysteine Cathepsins in Disease and their Potential as Drug Targets. Current Pharmaceutical Design 2007; 13: 387-403.
18. Löser R and Pietzsch J: Cysteine cathepsins: their role in tumor progression and recent trends in the development of imaging probes. Frontiers in Chemistry 2015; 3: 1-36.
19. Brocklehurst K: A sound basis for pH-dependent kinetic studies on enzymes. Protein Engineering 1994; 7: 291-9.
20. Huang CC, Lee CC, Lin HH, Chen MC, Lin CC and Chang JY: Autophagy-Regulated ROS from Xanthine Oxidase Acts as an Early Effector for Triggering Late Mitochondria-Dependent Apoptosis in Cathepsin S-Targeted Tumor Cells. PLoS ONE 2015; 10: 1-20
21. Hsieh MJ, Lin CW, Chen MK, Chien SY, Lo YS, Chuang YC, His YT, Lin CC, Chen JC and Yang SF: Inhibition of cathepsin S confers sensitivity to methyl protodioscin in oral cancer cells *via* activation of p38 MAPK/JNK signaling pathways. Scientific Reports 2017; 7: 1-11.
22. Zhang L, Wang H, XU J: Cathepsin S as a cancer target. Neoplasma 2015; 62: 16-28.
23. Da Costa AC, Santa Cruz F, Mattos Jr LAR, Aquino MAR, Martins CR, Ferraz AAB and Figueiredo JL: Cathepsin S as a target in gastric cancer. Molecular and clinical oncology 2020; 12: 99-103.
24. Liu WL, Liu D, Cheng K, Liu YJ, Xing S, Chi P, Liu XH, Xue N, Lai Y, Guo L and Zhang G: Evaluating the diagnostic and prognostic value of circulating cathepsin S in gastric cancer. Oncotarget 2016; 7: 28124-38.
25. Wilkinson RDA, Young A, Burden RE, Williams R and Scott CJ: A bioavailable cathepsin S nitrile inhibitor. Molecular Cancer 2016; 15: 2-11.
26. Weitoft T, Larsson A, Manivel VA, Lysholm J, Knight A and Ronnelid J: Cathepsin S and cathepsin L in serum and synovial fluid in rheumatoid arthritis with and without autoantibodies arthritis. Rheumatology 2015; 54: 1923-8.
27. Memmert S, Damanaki A, Nogueira AVB, Eick S, Nokhbehsaim M, Papadopoulou AK, Till A, Rath B, Jepsen S, Götz W, Piperi C, Basdra EK, Cirelli JA, Jäger A and Deschner J: Role of Cathepsin S in Periodontal Inflammation and Infection. Mediators of Inflammation 2017; 1-11.
28. Ainscougha JS, Macleoda T, McGonagleb D, Brakefieldа R, Barond JM, Alaseb A, Wittmannb M and Stacey M: Cathepsin S is the major activator of the psoriasisassociated proinflammatory cytokine IL-36γ. Panas Plus 2017; 114: 1-10
29. Andrault PM, Samsonov SA, Weber G, Coquet L, Nazmi K, Bolscher JGM, Lalmanach AC, Jouenne T, Brömme D, Pisabarro MT, Lalmanach G, and Lecaille F: Antimicrobial Peptide LL-37 is Both a Substrate of Cathepsins S and K and a Selective Inhibitor of Cathepsin L. Biochemistry 2015; 54: 2785-98.
30. Brown R, Nath S, Lora A, Samaha G, Elgamal Z, Kaiser R, Taggart C, Weldon S and Geraghty P: Cathepsin S: investigating an old player in lung disease pathogenesis, comorbidities, and potential therapeutics. Respiratory Research 2020; 21: 1-17.
31. Wartenberg M, Andrault PM, Saidi A, Bigot P, Nadal-Desbarats L, Lecaille F and Lalmanach G: Oxidation of cathepsin S by major chemicals of cigarette smoke. Free Radical Biology and Medicine 2020; 150: 53-65.
32. Andrault PM, Schamberger AC, Chazeirat T, Sizaret D, Renault J, Staab-Weijnitz CA, Hennen E, Petit-Courty A, Wartenberg M, Saidi A, Baranek T, Guyetant S, Courty Y, Eickelberg O, Lalmanach G and Lecaille F: Cigarette smoke induces overexpression of active human cathepsin S in lungs from current smokers with or without COPD. American Journal of physiology: Lung Cellular and Molecular Physiology 2019; 317: L625-L638.
33. Nakajima T, Nakamura H, Owen CA, Yoshida S, Tsuduki K, Chubachi S, Shirahata T, Mashimo S, Nakamura M, Takahashi S, Minematsu N, Tateno H, Fujishima S, Asano K, Celli BR and Betsuyaku T: Plasma Cathepsin S and Cathepsin S/Cystatin C Ratios Are Potential Biomarkers for COPD. Disease Markers 2016; 1-10.
34. Zhou PP, Zhang WY, Li ZF, Chen YR, Kang XC and Jiang YX: Association between SNPs in the promoter region in cathepsin S and risk of asthma in Chinese Han population. European Review for Medical and Pharmacological Sciences 2016; 20: 2070-6.
35. Small DM, Brown RR, Doherty DF, Abladey A, Zhou-Suckow Z, Delaney RJ, Kerrigan L, Dougan CM, Borensztajn KS, Holsinger L, Booth R, Scott CJ, López-Campos G, Elborn JS, Mall MA, Weldon S and Taggart CC: Targeting of cathepsin S reduces cystic fibrosis-like lung disease. European Respiratory Journal 2018; 53: 1-11.
36. Sena BF, Figueiredo JL and Aikawa E: Cathepsin S As an inhibitor of Cardiovascular inflammation and Calcification in Chronic Kidney Disease. Frontiers in Cardiovascular Medicine 2018; 4: 1-7.
37. Figueiredo JL, Aikawa M, Zheng C, Aaron J, Lax L, Libby P, Filho JLL, Gruener S, Fingerle J, Haap W, Hartmann G and Aikawa E: Selective Cathepsin S Inhibition Attenuates Atherosclerosis in Apolipoprotein E-Deficient Mice with Chronic Renal Disease. The American Journal of Pathology 2015; 185: 1155-66.
38. Ahmad S and Siddiqi MI: Insights from molecular modeling into the selective inhibition of cathepsin S by its inhibitor. Journal of Molecular Modeling 2017; 23: 3255-6.
39. Figueiredo JL, Aikawa M, Zheng C, Aaron J, Lax L, Libby P, Lima Filho JL, Gruener S, Fingerle J, Haap W, Hartmann G and Aikawa E: Selective Cathepsin S Inhibition Attenuates Atherosclerosis in Apolipoprotein E-Deficient Mice with Chronic Renal Disease. The American Journal of Pathology 2015; 185, 1156-66.

40. Wu H, Du Qiuna, Dai Q, Ge J and Cheng X: Cysteine protease Cathepsins in Atherosclerotic Cardiovascular Diseases. Journal of Atherosclerosis and Thrombosis 2017; 24: 1-13.
41. Repnik U, Starr AE, Overall CM and Turk B: Cysteine Cathepsins Activate ELR Chemokines and Inactivate Non-ELR Chemokines. The journal of biological chemistry 2015; 290: 13800-11.
42. Hargreaves P, Daoudlarian D, Theron M, Kolb FA, Young MM, Bernhard Reis B, Andre Tiaden A, Bannert B, Kyburz D and Manigold T: Differential effects of specific cathepsin S inhibition in biocompartments from patients with primary Sjögren syndrome. Arthritis Research & Therapy 2019; 21: 1-11.
43. Edman MC, Janga SR, Meng Z, Bechtold M, Chen AF, Kim C, Naman L, Sarma A, Teekappanavar N, Kim AY, Madriga S, Singh S, Ortiz E, Christianakis S, Arkfeld DG, Mack WJ, Heur M, Stohl W, and Hamm-Alvarez SF: Increased Cathepsin S activity associated with decreased protease inhibitory capacity contributes to altered tear proteins in Sjögren's Syndrome patients. Scientific Reports 2018; 8: 1-12.
44. Vizovišek M, Fonović M, Turk B: Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond. Matrix Biology 2019; 75-76: 149-59.
45. Gupta S, Singh RK, Dastidar S and Ray A: Cysteine cathepsin S as an immunomodulatory target: present and future trends. Expert Opinion on Therapeutic Targets 2008; 12: 291-9.
46. Vizovišek M, Vidak E, Javoršek U, Mikhaylov G, Bratovš A and Turk B: Cysteine cathepsins as therapeutic targets in inflammatory diseases. Expert Opinion on Therapeutic Targets 2020; 24: 573-88.
47. Wilkinson RDA, Williams R, Scott CJ and Burden RE: Cathepsin S: therapeutic, diagnostic, and prognostic potential. Biological Chemistry 2015; 396: 867-82.
48. Kumar V, Abbas AK, Fausto N and Aster JC: Pathologic Basis of Disease. Saunders Elsevier, Eighth edition 2008.
49. Riese R, Villadangos JA, Shi GP, Ploegh HL, Chapman HA, Bryant RA, Deussing J, Driessen C, Lennon-Dumenil AM, Roth W, Saftig P and Peters C: Proteases involved in MHC class II antigen presentation. Immunological reviews 1999; 172: 109-20.
50. Riese RJ, Mitchell RN, Villadangos JA, Shi GP, Palmer JT, Karp ER, De Sanctis GT, Ploegh HL and Chapman HA: Cathepsin S Activity Regulates Antigen Presentation and Immunity. The Journal of clinical investigation 1998; 101: 2351-63.
51. Riese R, Villadangos JA, Shi GP, Ploegh HL, Chapman HA, Dranoff G, Small C, Gu L and Haley KJ: Cathepsin S Required for Normal Class II Peptide Loading and Germinal Center Development. Immu 1999; 10: 197-206.
52. Nakagawa TY, Brissette WH, Lira PD, Griffiths RJ, Petrushova N, Stock J, McNeish JD, Eastman SE, Howard ED, Clarke SR, Rosloneic EF, Elliott EA and Rudensky AY. Impaired Invariant Chain Degradation and Antigen Presentation and Diminished Collagen-Induced Arthritis in Cathepsin S Null Mice. Immunity 1999; 10: 207-17.
53. Lutzner N and Kalbacher H: Quantifying Cathepsin S Activity in Antigen Presenting Cells Using a Novel specific Substrate. Journal of Biological Chemistry 2008; 283: 36185-94.
54. Costantino CM, Ploegh HL and Hafler DA: Cathepsin S Regulates Class II MHC Processing in Human CD4 + HLA-DR+ T Cells. Journal of Immunology 2009; 183: 945-52.
55. Mcgrath ME: The Lysosomal Cysteine Proteases. Annual review of Biophysics and biomolecular structure 1999; 28: 181-204.
56. Kumar P and Mina U: Life Sciences, Fundamentals and Practices. Pathfinder Publication, Third Edition 2013.
57. Hansch C and Fujita T: p-δ-π analysis, a method for the correlation of biological activity and chemical structure. Journal of American Chemical Society 1964; 86: 1616-26.
58. Jamloki A, Karthikeyan C, Sharma SK, Moorthy NSHN and Trivedi P: QSAR studies on some GSK-3α Inhibitory 6-aryl-pyrazolo(3,4-b)pyridines. Asian Journal of Biochemistry 2006; 1: 236-43.
59. Thomas G: Medicinal Chemistry: An Introduction. John Wiley & Sons, Second Edition 2008.
60. Gupta SP and Kumaran S: Quantitative structure-activity relationship studies on benzodiazepine hydroxamic acid inhibitors of matrix metalloproteinases and tumour necrosis factor-α converting enzyme. Asian Journal of Biochemistry 2006; 1: 47-56.
61. Hoffman B, Cho S.J, Zheng W, Wyrick S, Nichols DE, Mailman RB and Tropsha A: Quantitative structure-activity relationship modeling of dopamine D1 antagonists using comparative molecular field analysis, genetic algorithms-partial least-squares and K nearest neighbour methods. Journal of Medicinal Chemistry 1999; 42: 3217-26.
62. Shen M, LeTiran A, Xiao Y, Golbraikh A, Kohn H and Tropsha A: Quantitative structure-activity-relationship analysis of functionalized amino acid anticonvulsant agents using k nearest neighbour and simulated annealing PLS methods. J Med Chem 2002; 45: 2811-23.
63. Rogers D and Hopfinger AJ: Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. Journal of Chemical Information and Computer Science 1994; 34: 854-66.
64. Golbraikh A, Bonchev D and Tropsha A: Novel chirality descriptors derieved from molecular topology. Journal of Chemical Information and Computer Science 2001; 41: 147-58.
65. Zheng W and Tropsha A: A novel variable selection quantitative structure-property relationship approach based on the k-nearest-neighbour principle. Journal of Chemical Information and Computer Science 2000; 40: 185-94.
66. Sen S and Sarker K: 2D QSAR study of some inhibitors of ehrlich ascites carcinoma. International Journal of Pharmaceutical Sciences Review and Research 2012; 15: 102-7.
67. Ward YD, Thomson DS, Frye LL, Cywin CL, Morwick T, Emmanuel MJ, Zindell R, McNeil D, Bekkali Y, Girardot M, Hrapchak M, DeTuri M, Crane K, White D, Pav S, Wang Y, Hao M, Grygon CA, Labadia ME, Freeman DM, Davidson W, Hopkins JL, Brown ML, and Spero DM: Design and Synthesis of Dipeptide Nitriles as Reversible and Potent Cathepsin S Inhibitors. Journal of Medicinal Chemistry 2002; 45: 5471-82.
68. Sadowski J and Gasteiger J: From atoms and bonds to three-dimensional atomic coordinates: Automatic model builders. Chemical Reviews 1993; 93: 2567-81.
69. Vinter JG, Davis A and Saunders MR: Strategic approaches to drug design. An integrated software framework for molecular modelling. Journal of Computer Aided Molecular Design 1987; 1: 31-51.
70. Paliwal SK, Pandey A and Paliwal S: Quantitative Structure Activity Relationship Analysis of N-(mercapto-alkanoyal)-[(acylthio)alkanoyal] glycine derivatives as

ACE inhibitors. American Journal of Drug Discovery and Development 2011; 1: 85-104.
71. Berhanu WM, Pillai GG, Oliferenko AA and Katritzky AR: Quantitative Structure–Activity/Property Relationships: The Ubiquitous Links between Cause and Effect 2012; 77: 507-17.
72. Datar PA: 2D-QSAR Study of indolylpyrimidines derivative as antibacterial against *Pseudomonas aeruginosa* and *Staphylococcus aureus*: A Comparative Approach. Journal of Computational Medicine 2014; 1-9.
73. Suh JT, Skiles JW, Williams BE, Youssefyeh RD, Jones H, Loev B, Neiss ES, Schwab A, Mann WS, Khandwala A, Wolf PS and Weinryb I: Angiotensin converting enzyme inhibitors. New orally active antihypertensive (mercaptoalkanoyl)- and [(acylthio) alkanoyl]glycine derivatives. Journal of Medicinal Chemistry 1985; 28: 57-66.
74. Klocker J, Wailzer B, Buchbauer G and Wolschann P: Bayesian neural networks for aroma classification. Journal of Chemical Information and Computer Science 2002; 42: 1443-9.
75. Kovatcheva A, Buchbauer G, Golbraikh A and Wolschann P: QSAR modelling of alpha campholenic derivatives with sadalwood odor. Journal of Chemical Information and Computer Science 2003; 43: 259-66.
76. Patil RB and Sawant SD: 2D QSAR Analysis on B-Ring Trifluoromethylated chromenone analogues as anticancer agents. International Journal of Advances in Pharmacy, Biology and Chemistry 2012; 1: 72-80.
77. Sharma MC and Sharma S: 2D QSAR study of 7-Methyljuglone Derivatives: An Approach to Design Anti-Tubercular Agents. Journal of Pharmacology and Toxicology 2011; 6: 493-504.
78. Kubinyi H: Evolutionary variable selection in regression and PLS analysis. J of Chemometrics 1996; 10: 119-33.
79. Vengurlekar S, Sharma R and Trivedi P: Two- and three-dimensional QSAR studies on benzyl amide-ketoacid inhibitors of HIV integrase and their reduced analogues. Medicinal Chemistry Research 2010; 19: 1106-20.
80. Winkler DA: Neural networks as robust tools in drug lead discovery and development. Chem Plus Chem 2004; 27: 139-67.
81. Paliwal S, Narayan A and Paliwal S: Quantitative structure activity relationship analysis of dicationic diphenylisoxazole as potent anti-trypanosomal agents. QSAR and Combinatorial Science 2009; 28: 1367-75.
82. Paliwal SK, Pal M and Siddiqui AA: Quantitative structure activity relationship analysis of angiotensin II AT1 receptor antagonists. Medicinal Chemistry Research 2010; 19: 475-89.

How to cite this article:
Kushwaha S and Paliwal SK: 2D QSAR analysis of dipeptide nitrile based cathepsin S inhibitors. Int J Pharm Sci & Res 2021; 12(6): 3391-02. doi: 10.13040/IJPSR.0975-8232.12(6).3391-02.