"Development of ability anticancer sellers the usage of pharmacophore optimization with the aid of molecular modeling studies"

(Akash R.Pagare),(Hrutuja R.Wagh),(Ravina A.Shaha),(Monicaben C. Ramani)

AISSMS College of Pharmacy, Near R.T.O, Kennedy Road, Pune-411001, India.

Abstract:

At some stage in the method of cellular replication, microtubules and key cytoskeletal filaments play an important function. At some stage in cellular mitosis, the microtubules extend from the cellular centrosome to form mitotic spindle and connect to the kinetochore of chromosomes. Then, the kinetochores are collected around the equatorial plate. Microtubules are in dynamic equilibrium with tubulin dimers. Disruption of the dynamic equilibrium will lead to cellular cycle arrest or cell apoptosis. Given their great roles in the boom and feature of cells, microtubules or tubulins are some of the maximum essential molecular goals for cancer chemotherapeutic retailers. A number of small molecules were observed to bind tubulins, interfering with microtubule polymerization or depolymerization, and then result in mobile cycle arrest, main to cellular death.2d and 3D (QSAR) research were executed for correlating chemical composition of Cinnamic Acyl Sulfonamide analogues. Generation of established qsar version is crucial to make sure that the version have suited predictive energy. New chemical entities (NCE's) were designed the usage of consequences of 2D & 3D QSAR research. Binding affinities of designed NCE's have been studied on protein tubulin polymerization enzyme (PDB code: 1sa0) the usage of docking research and their absorption, distribution, metabolism and excretion (ADME) homes were also predicted to make sure drug like pharmacokinetic profile. Generating r^2-0.735, q^2-0.6549, pred. r^2-0.8874 for 2D QSAR and q^2-0.6025, pred.r^2-0.8262 for 3D QSAR.

Keywords: Molecular Modelling, 2D QSAR, 3D QSAR, Design NCEs, Docking, ADME.

1. Introduction:

The inducement of cancer is complex involving many molecular mechanisms. Loss of control in cell replication is one of the most important mechanisms that cause cancer. During the process of cell replication, microtubules and key cytoskeletal filaments involved in numerous cellular functions, play an important role. During cell mitosis, the microtubules extend from the cell centrosome to form mitotic spindle and fasten to the kinetochore of chromosomes. Then, the kinetochores are gathered around the equatorial plate[1]. Microtubules are in dynamic equilibrium with tubulin dimers. Disruption of the dynamic equilibrium will lead to cell cycle arrest or cell apoptosis[2,3]. Given their significant roles in the growth and function of cells, microtubules or tubulins are among the most important molecular targets for cancer chemotherapeutic agents. A number of small molecules were found to bind tubulins, interfering with microtubule polymerization or depolymerization, and then induce cell cycle arrest, leading to cell death [4-7]. Colchicine is the first drug that is well known to bind tubulin, and its binding site on tubulin has been characterized recently [8, 9]. Several

other new compounds have also been synthesized as antitubulin agents, such as E7010, CA-4, HMN-214, cyanocombretastatins, Metronidazole acid acyl sulfonamide. They bound tubulins through the same site as colchicines[10-14].

Computational methods provide powerful tool to design a new highly active molecules. Quantitative Structure Activity Relationship (QSAR) studies helped us to correlate the physicochemical properties of chemical compounds with biological activities, which could shed light on the key structural components that are important for biological activity. One of the important roles of QSAR methodology is to predict the biological activity of in-silico designed New Chemical Entities (NCEs) [15]. To increase the predictive power of QSAR methodology various methods were applied for model validation. Among them most popular method for model validation is leave-one-out cross-validated r^2 (LOO q^2). But it has been reported that the widely accepted Leave-One-Out (LOO) cross-validated r^2 (q^2) is an inadequate characteristic to assess the predictive ability of the models [16]. Thus to evaluate the predictive ability of a QSAR model; its validation must be done using an external test set of compounds (i.e. those, which were not included in the training set) with known activities. It includes the methodologies in constructing main components of QSAR model, namely the methods for selection of informative descriptors, validating the model for anticancer activity prediction. QSAR model was developed on a series of compounds containing cinnamic acyl sulfonamide pharmacophore to identify key structural fragments required around pharmacophore for anti-proliferative activity. In the present studies 2D QSAR and 3D QSAR studies were carried out using predictive QSAR modeling method [17,18]. The New Chemical Entities (NCEs) were designed using the results of best 2D and 3D QSAR model, and Combinatorial Library was generated using Lipinski's screen as filter. Finally, molecular docking studies were carried out to understand interactions between Tubulin Polymerization target and designed NCEs in better way. Hydrophobic and hydrogen bond interactions were used to find out active binding sites of Tubulin Polymerization protein in the docked complex. Rational drug design should take both pharmacokinetic and metabolic information into consideration, and the information should be incorporated with molecular, biochemical and pharmacological data to provide well rounded drug design. Hence, Prediction of ADME properties was though worthwhile as screen to sort out those compounds that follow Lipinski's rule to ensure drug like pharmacokinetic profile of the designed NCEs[19].

2. Material & Methods:

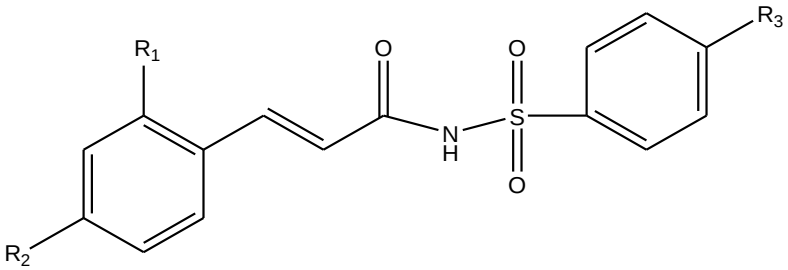
2.1. Biological dataset:

For present study, 40 molecules of Cinnamic Acyl Sulfonamide series reported for the inhibition of Tubulin Polymerization **(Table 1)** was chosen to optimize substitution requirement around selected Pharmacophore by using QSAR and Molecular Modeling studies like Docking & ADME Prediction [20].Six molecules did not fit well into either training or test set reported for anticancer activity; hence were dropped from present studies. Selected series of compound were evaluated for their ability to inhibit of tubulin polymerization using tubulin assembly assay. Biological activity was expressed in terms of IC_{50} which was converted to pIC_{50} using formula ($pIC_{50} = \log 1/IC_{50}$).

2.2. Software:

All QSAR studies were performed using V-Life Molecular Design Suite Software, version 3.5 & 4.2 [21]. Molecules were optimized by Merck Molecular Force Field (MMFF) energy minimization method.

Table1. Selected series of compounds containing Cinnamic Acyl Sulfonamide Pharmacophore:



Sr. No.	Comp. No.	R_1	R_2	R_3	IC_{50}	pIC_{50} (log 1/ IC_{50})
1	9a	-H	-H	-H	37	-1.5682
2	9b	-H	-H	$-CH_3$	152	-2.18184
3	9c	-H	-H	-F	4.5	-0.65321
4	9d	-H	-H	-Cl	14.8	-1.17026
5	9e	-H	-H	-Br	63	-1.79934
6	*10a	-H	-F	-H	4.5	-0.65321
7	10b	-H	-F	$-CH_3$	40	-1.60206
8	10c	-H	-F	-F	2.4	-0.38021
9	*10d	-H	-F	-Cl	4.5	-0.65321
10	10e	-H	-F	-Br	14.7	-1.16732
11	11a	-H-	-Cl	-H	15.3	-1.18469
12	11b	-H	-Cl	$-CH_3$	30	-1.47712
13	11c	-H	-Cl	-F	8.9	-0.94939
14	11d	-H	-Cl	-Cl	9.3	-0.96848
15	11e	-H	-Cl	-Br	10.8	-1.03342
16	12a	-H	-Br	-H	74	-1.86923
17	12b	-H	-Br	$-CH_3$	137	-2.13672
18	12c	-H	-Br	-F	3.8	-0.57978
19	12d	-H	-Br	-Cl	55	-1.74036
20	12e	-H	-Br	-Br	53	-1.72428
21	13a	-H	$-OCH_3$	-H	160	-2.20412
22	13b	-H	$-OCH_3$	$-CH_3$	146	-2.16435
23	13c	-H	$-OCH_3$	-F	45	-1.65321
24	13d	-H	$-OCH_3$	-Cl	52	-1.716

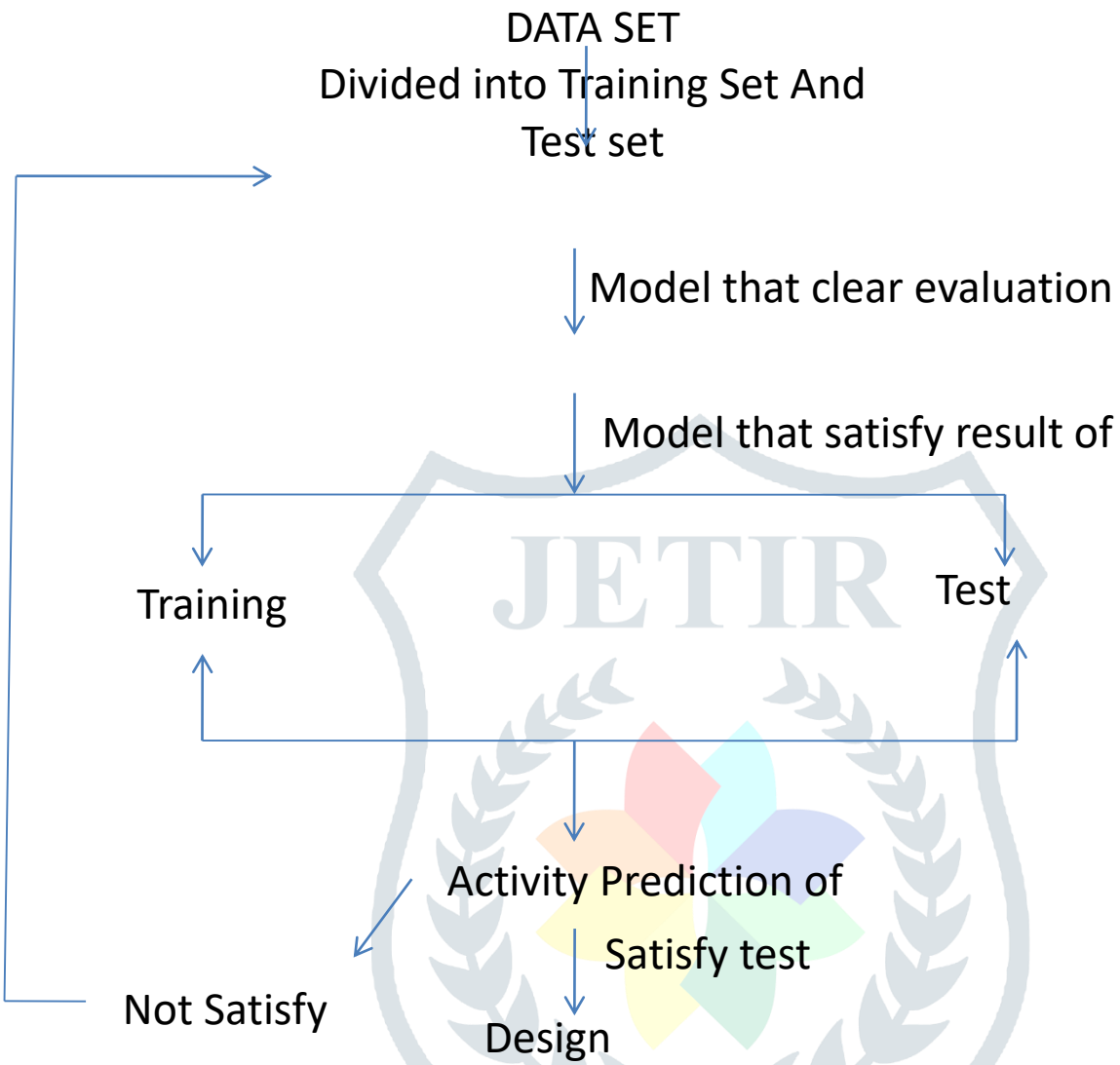
25	*13e	-H	$-OCH_3$	-Br	53	-1.72428
26	14a	-H	$-N(CH_3)_2$	-H	26	-1.41497
27	*14b	-H	$-N(CH_3)_2$	$-CH_3$	37	-1.5682
28	14c	-H	$-N(CH_3)_2$	-F	10	-1
29	14d	-H	$-N(CH_3)_2$	-Cl	14	-1.14613
30	14e	-H	$-N(CH_3)_2$	-Br	34	-1.53148
31	15a	-H	$-NO_2$	-H	40	-1.69897
32	15b	-H	$-NO_2$	$-CH_3$	151	-2.17898
33	15c	-H	$-NO_2$	-F	6.3	-0.79934
34	15d	-H	$-NO_2$	-Cl	134	-2.12711
35	15e	-H	$-NO_2$	-Br	163	-2.21219
36	*16a	$-NO_2$	-H	-H	74	-1.86923
37	16b	$-NO_2$	-H	$-CH_3$	102	-2.0086
38	16c	$-NO_2$	-H	-F	69	-1.83885
39	16d	$-NO_2$	-H	-Cl	78	-1.8921
40	*16e	$-NO_2$	-H	-Br	134	-2.12711

2.3. 2D QSAR Studies:

2.3.1 Experimental design for 2D QSAR:

Dataset of 34 molecules was divided into training and test set in an attempt to ensure robustness of the model and increase predictive ability of QSAR model. The experimental design for our study is shown in **Figure1**. The data set was divided in training and test sets and the training set which followed all model evaluation parameters were subjected to randomization test. Two models were selected which satisfy the results of randomization test which were named as Training set-A and Training set-B. These model were subjected to external validation by splitting test set into Test set-a for training set-A and Test set- b for training set-B respectively. If model generated by training set satisfy all parameters for prediction of test set, then only the concerned model was used to predict activity of test set molecules. If model does not satisfy test set prediction parameters, the whole process is repeated from first step i.e. selection of training set. Only those models which satisfy the test set were selected for design of NCEs. We have ensured that selected training and test sets also satisfy the following criteria: (i) Representative points of the test set must be close to those of the training set; (ii) Representative points of the training set must be close to representative points of the test set; (iii) Training set must have chemical and biological diversity; (iv) Training and test set have uniform representation of molecules; uni-column statistics were performed.

Figure 1. Selection of molecules in training and test sets:



2.3.2. Uni-column statistics:

The comparative statistical parameters of training and test sets created by manual data selection are reported in **Table 2**. Standard deviations of Training set A as well as test set A were found to be nearly close to each other. This showed that even though the selected molecules in training/test set are different, still the distribution pattern with respect to the biological activities of the molecules in both the selection methods is quite similar.

Table 2. Uni-Column statistics for training set and test set:

Column Name	Average	Max	Min	Std.dev.	Sum
Training	-1.5851	-o.5798	-2.2122	0.5087	-45.9685
Test	-1.1608	-0.3802	-1.5315	0.4841	-5.8038

2.4. Descriptor selection:

Various 2D descriptors (a total of 229) like structural, hydrophobic, steric, electrostatic, topological descriptors etc, were calculated using V life MDS software. The pre-processing of the independent variables (i.e., descriptors) was carried out by removing invariables (one which do not vary for majority of compounds) after doing so only 192 descriptors remained in the worksheet. It has been reported that there is high probability of chance correlation between the observed and predictive activity when number of independent variable (descriptors) are comparable or more than the compounds in dataset for any QSAR analysis [22, 23]. Thus, reduction in no. of descriptor is very important step to avoid the occurrences of chance correlation and by chance incorporation of irrelevant descriptors in final QSAR model. As there are some reported descriptor selection method like-forward, backward, simulated annealing etc, but one must understand and remember that each method have its own limitations. Hence, we have applied some methods for variable (descriptor) reduction which may improve performance as well as improve predictability of QSAR model.

2.4.1. Correlation matrix:

In the present study we have considered the correlation between descriptor with activity as well as their inter-correlation i.e. descriptor-descriptor correlation[24]. We have considered only those descriptors which show either direct or indirect correlation with activity by more than 0.35 and shows inter correlation less than 0.8 generated for the selected series of compounds is shown in **Table 3.**

Table 3. Correlation Matrix:

Descriptor	**T_C_O_6**	**T_2_F_6**	**SsCLE-index**	**T_C_N_6**
T_C_O_6	1	-	-	-
T_2_F_6	O.53	1	-	-
SsCLE-index	-0.184	0.968	1	-
T_C_N_6	-0.182	0.037	-0.18	1

2.5. QSAR model generation:

QSAR models were generated by MLR (multiple linear regressions) method by selecting descriptors manually. MLR is usually used to fit the regression model (Equation 1), which models a response variable, y, as a linear combination of the X-variables, with the coefficients b. The deviations between the data (y) and the model (Xb) are called residuals, and are denoted by:

$y = Xb + e$ ---------- [1]

2.6. 3D QSAR by SA-KNN method:

3D-QSAR studies were performed by KNN–MFA using SA variable selection method[25]. KNN-MFA method requires suitable alignment of given set of molecules after optimization; alignment was carried out by template based alignment method. Molecular alignment was used to visualize the structural diversity in the given set of molecules. It was followed by generation of common rectangular grid around the molecules,

steric and electrostatic interaction energies were computed at the lattice points of the grid. Resulting set of aligned molecules was then used to build 3D QSAR models.

2.7. Model validation:

2.7.1. Internal validation:

Internal validation was carried out using leave-one-out (q^2, LOO) method [15]. To calculate q^2, each molecule in the training set was sequentially removed, the model was refitted using same descriptors, and the biological activity of the removed molecules were predicted using the refit model. This attempt was made to calculate robustness of QSAR model. All cross-validation studies were performed by considering the fact that a value of $q^2 > 0.5$.

2.7.2 External validation:

External validation of generated models was carried out by predicting the activity of test set of compounds. The pred_r^2 value is calculated as follows

$$Pred\_r^2 = 1 - \Sigma(yi-\hat{y}i)^2 / \Sigma(yi-\hat{y}mean)^2 \quad \text{---------- (2)}$$

Where yi, ŷi are the actual and predicted activity of the *i*th molecule in the test set, respectively, and y mean is the average activity of all molecules in the training set [16]

2.7.3 Randomization test:

This is a most popular tool used by researchers to prevent from chance correlation. In this method, keeping the x-variable intact, a repeated permutation of response variable is done. After each permutation r^2 and q^2 is recorded. If in each case the r^2 and q^2 gives very low value than original data, then we can say with some confidence that original QSAR model is real and not generated by chance. In our study we have calculate Z-score to check significance of the model. Following formula was used for the same[26].

$$Z\ score = (^{q2org}\_^{q2a})\ q^2std \text{ - - - - - - - - - [3]}$$

Where ^{q2org} is the ^{q2} value calculated for the actual data set, ^{q2} a is the average ^{q2}, and ^{q2} std is the standard deviation of ^{q2}, calculated for various iterations using different randomized datasets.

2.8. Model Evaluation:

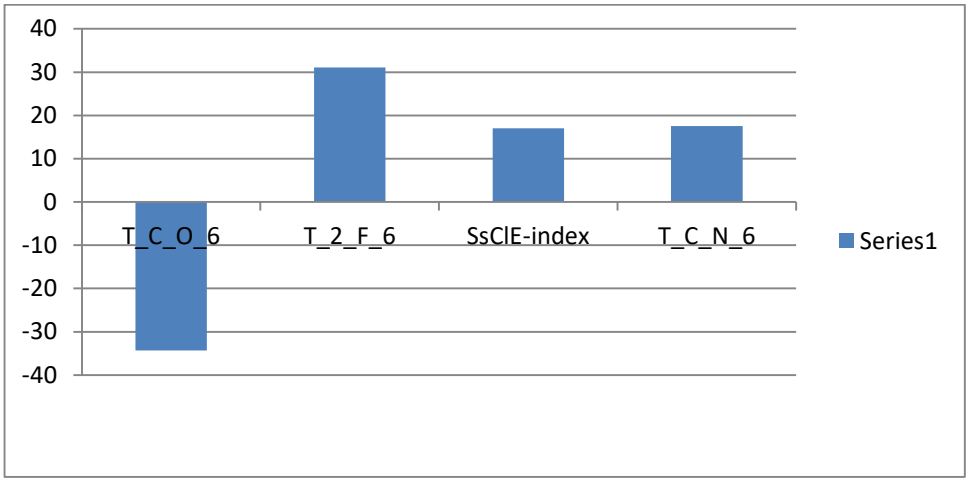
The developed QSAR models were evaluated using the following statistical measures: number of observations (molecules); k, number of variables; r^2, co-efficient of determination; ^{q2}, cross validated r^2 (by leave-one-out); pred_ r^2, r^2 for external test set; Z score, Z score calculated by the randomization test; best_ran_ ^{q2}, highest ^{q2} value in the randomization test; best_ran_ r^2, highest r^2 value in the randomization test; α, statistical significance parameter obtained by the randomization test; SEE, standard error of estimate of the model; SECV, standard error of cross validation; and SEP, standard error of external test set prediction.

3. Result and Discussion:

3.1. 2D QSAR models:

We have got four meaningful descriptor viz. T_C_O_6, T_2_F_6, SsClE-index, T_C_N_6, r^2= 0.7355, ^{q2} = 0.5965, F test = 16.6839, pred_ r^2 = 0.8874 shown in **Figure 2.**

Figure 2: Contribution plot of selected descriptor.



Contribution plot of selected descriptors:

By looking at the results we came to know that this descriptor alone satisfies all evaluation parameters. The above mentioned descriptors show highest correlation with activity (as shown in correlation matrix) and also show proper distribution of data points. On the basis of the statistical parameters viz. r^2> 0.7, cross-validated r^2 .i.e. q^2> 0.5 and parameter to assess external validation i.e. pred_r^2> 0.8; the generated regression equation of model was used for further studies. Following regression equation was used to design NCEs.

pIC50= -0.2773(T_C_O_6) + 0.3825T_2_F_6 + 0.426 SsClE-index + 0.3151T_C_N_6.

The plot of actual versus predicted activity for model considering only there three descriptors for training set-A is shown in

Figure 3. Shows Plot of Actual versus predicted Activity:

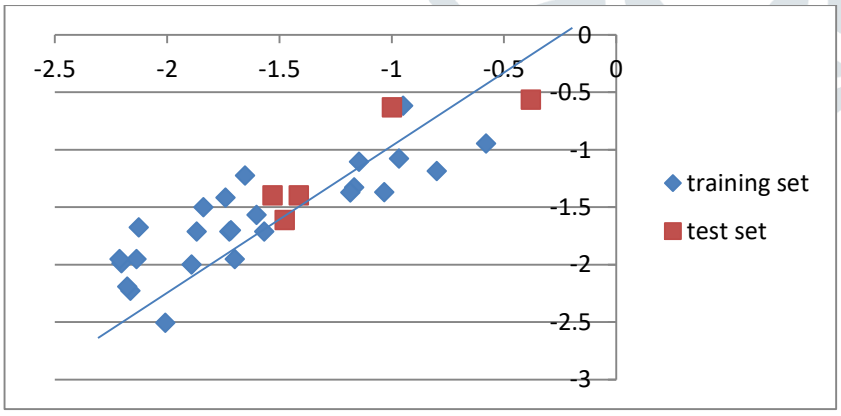


Table 4. Statistical parameters of developed 2D QSAR model:

Statistical parameter	MLR Model 1	MLR Model 2
r^2	0.7355	0.8165
q^2	0.6549	0.5965
F test	16.683	21.3793
r^2se	0.2826	0.2562
q^2se	0.2691	0.2856
Pred_r^2	0.8874	0.6322
Pred_r^2se	0.2274	0.2350
Best rand r^2	0.40189	0.44281
Best rand q^2	0.15695	0.20728
Z Score_ran_ r^2	6.90599	6.30521
Z Score_ran_ q^2	5.96042	4.04202
α_ran_ r^2	0.0000	0.0000
α_ran_ q^2	0.0000	0.0000
Descriptors	T_C_O_6 T_2_F_6 SsClE-index T_C_N_6	T_2_F_6 T_C_O_6 SsClE-index T_2_C_1

3.1.1 Accuracy of Model (Residual Value Calculation):

The accuracy of Model-1 has well because residual value of that model was found to be nearer to zero and we can take Model-1 for the pharmacophore optimization.

The residual value of Model-1 shows in **Table 5.**

Table 5. Residual value calculation:

Sr. No.	Compd.	Biological Activity Training set	Predicted Activity	Residual Value
TRAINING SET				
1	9a	-1.5682	-1.7097	0.1415
2	9b	-2.1818	-1.9492	-0.2326
3	9c	-0.6532	-0.9448	0.2916
4	9d	-1.1703	-1.417	0.2467
5	9e	-1.7993	-1.7097	-0.0896
6	10b	-1.6021	-1.5667	-0.0354
7	10e	-1.1673	-1.3273	0.16
8	11a	-1.1847	-1.3694	0.1847
9	11c	-0.9494	-0.6163	-0.3331

10	11d	-0.9685	-1.0752	0.1067
11	11e	-1.0334	-1.3685	0.3351
12	12a	-1.8692	-1.7097	-0.1595
13	12b	-2.1367	-1.9492	-0.1875
14	12c	-0.5798	-0.9448	0.365
15	12d	-1.7404	-1.416	-0.3244
16	12e	-1.7243	-1.7097	-0.0307
17	13a	-2.2041	-1.987	-0.2171
18	13b	-2.1644	-2.2264	0.062
19	13c	-1.6532	-1.2221	-0.4311
20	13d	-1.716	-1.6987	-0.0173
21	14d	-1.1461	-1.1046	-0.0415
22	15a	-1.699	-1.9493	0.2502
23	15b	-2.179	-2.1886	0.0096
24	15c	-0.7993	-1.1842	0.3427
25	15d	-2.1271	-1.6757	-0.4514
26	15e	-2.2122	-1.9492	-0.263
27	16b	-2.0086	-2.5037	0.4951
28	16c	-1.8389	-1.4993	-0.3396
29	16d	-1.8921	-1.9967	0.1046
TEST SET				
1	10c	-0.3802	-0.5623	0.1821
2	11b	-1.4771	-1.6093	0.1322
3	14a	-1.4149	-1.3946	-0.0203
4	14c	-1	-0.6096	-0.3904
5	14e	-1.5314	-1.3946	-0.1368

3.2. Interpretation of 2D QSAR:

The present QSAR model reveals that Baumann's alignment independent descriptor has major contribution in explaining variation in activity. The definition for the descriptors that contributing significantly for the QSAR models are given below, the value given in parenthesis are percentile contribution of descriptor for the activity[27].

Descriptors T_X_Y_Z can be defined as total count of fragments formed with atom types X and Y separated by topological distance of Z bonds.

T_C_O_6 ; This is the count of number of Carbon atoms (single double or triple bonded) separated from any Oxygen atom (single or double bonded) by Six bond distance in a molecule. This descriptor shows negative contribution in the activity. This signifies That presence of Oxygen atom which leads to decreases Anticancer Activity.

T_2_F_6 ; This is the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from Fluorine atom by Six bonds distance in a molecule. This descriptor shows positive contribution in the activity. The presence of fluorine atom increases the activity.

SsClE-index ; Electro-topological state indices for number of –Cl group connected with one single bond. This descriptor shows positive contribution in the activity.

T_C_N_6 ; This is the count of number of Carbon atoms (single double or triple bonded) separated from any Nitrogen atom (single or double bond Six bond distance in a molecule. This signifies that increase in number such groups increases anticancer activity.

From the above observation it can be concluded that the descriptor in the model shows the positive contribution like T_2_F_6 shows that substitution of fluorine at R2 and R3 position increases the tubulin polymerization activity. Where as the T_C_N_6 descriptor also shows the positive contribution leads to increase the activity. The presence of nitrogen at R_2 position shows the anticancer activity. Other descriptor like T_C_O_6 which shows decreases the biological acivity. The SsClE-index descriptor shows that presence of chlorine on the ring which increases the activity.

3.3. 3D QSAR Studies:

The SA-KNN-MFA method is used for 3D QSAR studies, the error obtain by this method was low (q^2 _se = 0.2578, pred_r^2se = 0.2428). By using SA-KNN-MFA method the value of q^2, pred_r^2 and K of model 1 were found to be 0.6025, 0.8262 and 3 respectively. The points which contributed to SA KNN-MFA model 1 are display in **Figure 4.** It gives the idea about hydrophobic, steric, electronic parameter generated by SA-KNN-MFA method with their positive or negative contribution. The model generated by this method showed good internal and external predictivity. The results obtained are show in

Figure 4: Data points generated using SA-kNN–MFA method (3D-QSAR):

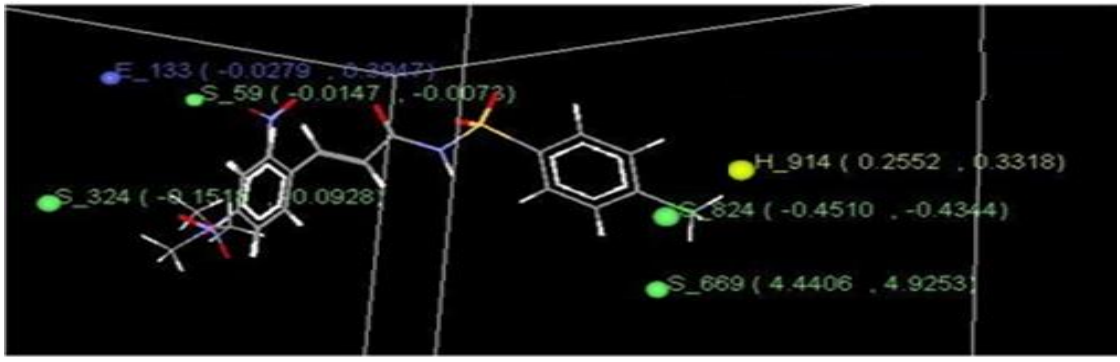


Table 6. The statistical results of 3D QSAR generated by SA- KNN-MFA Method:

Statistical Parameter	SA-kNN MFA Model 1
q^2	0.6025
q^2se	0.2578
$Pred\_r^2$	0.8262
$Pred\_r^2se$	0.2428
N	27
K nearest neighbor	3
Contributing descriptors	E_133(-0.0279, 0.3942) S_324(-0.1518, -0.0928) S_824(-0.4510, -0.4344) S_669(4.4406, 4.9253) H_914(0.2552, -0.3318) S_59 (-0.0147, -0.0073)

3.4. Interpretation of 3D QSAR model:

The electrostatic, steric, and hydrophobic requirement around Cinnamic Acyl Sulfonamide pharmacophore was optimized by 3D QSAR. The 3D data point was generated that contribute to SA-KNN-MFA 3D QSAR model, are shown in **Figure 4.**

The generated data point helped for the design of potent NCEs, range was based on the variation of the field values at the chosen point using the most active molecules and its nearest neighbor set. Negative and positive values in electrostatic descriptor indicated the requirement of negative and positive electrostatic potential respectively for enhancing the biological activity of Cinnamic Acyl Sulfonamide derivative. Point generated in SA-KNN-MFA 3D QSAR model are E_133(-0.0279, 0.3942) , S_324(-0.1518, -0.0928), S_824(-0.4510, -0.4344), S_669(4.4406, 4.9253), S_59 (-0.0147, -0.0073), H_914(0.2552, -0.3318), i.e electronic, sreric and hydrophobic interaction at lattice point 133,324,824,669,59,914 respectively. The nagative electronic value indicate the less electronic group are required to increase the activity. The sterically positive values like S_669(4.4406, 4.9253) and sterically negative values like S_324(-0.1518, -0.0928), S_824(-0.4510, -0.4344), S_59 (-0.0147, -0.0073) indicate the less steric and bulky group are required to enhancing the activity. Where as the hydrophobic group like H_914(0.2552, -0.3318) are required for the activity.

Figure 4. Data points generated using SA-KNN–MFA method (3D-QSAR):

3.5. Design of new chemical entities (NCEs) containing cinnamic acyl sulfonamide pharmacophore:

The information obtained from 2D and 3D QSAR studies had helped a lot in optimizing cinnamic acyl sulfonamide pharmacophore and for design of NCEs containing for potent anti-cancer activity. Substitution

pattern around cinnamic acyl sulfonamide pharmacophore showed in **Figure 5** was used for the design of NCEs using CombiLib tool of V Life MDS software. **Figure 6** shows common template used for design of NCE's. Designed compounds were subjected to Lipinski's screen[28] to ensure drug like pharmacokinetic profile of the designed compounds in order to improve their bioavailability. The following parameters were used as Lipinski's filters (Values in parenthesis indicate ideal requirements).

1. Number of Hydrogen Bond Acceptor (A) (<10)

2. Number of Hydrogen Bond donor (D) (<5)

3. Number of Rotatable Bond (R) (<10)

4. X log P (X) (<5)

5. Molecular weight (W) (<500 g/mol)

6. Polar surface area (S) is (<140 Å)

Figure 5. Pharmacophoric requirements around Cinnamic Acyl Sulfonamide.

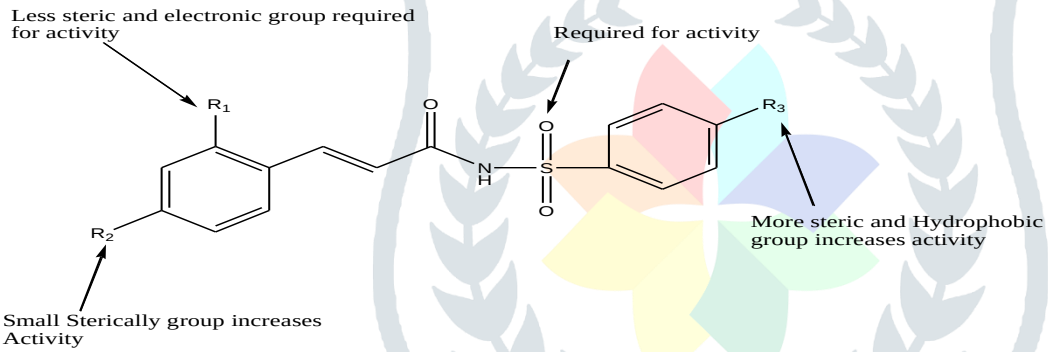
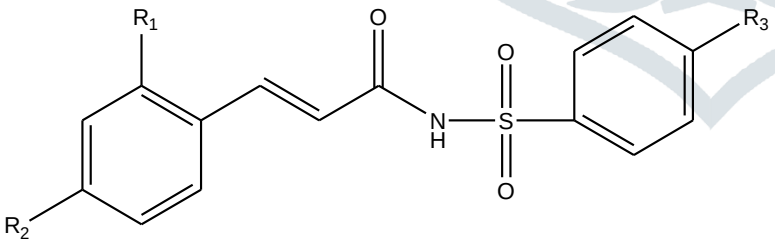


Figure.6. Common template used to design NCEs:



The designed NCE's along with predicted activity obtained by MLR equation generated by 2D-QSAR models shows in Table 7.

Table7. Structures of designed NCE's along with predicted activity:

Sr. No	Comp. No.	R_1	R_2	R_3	Predicted activity	Lipinski Screen	Lipinski Score
1	N-1	-H	-Br	$-C_6H_5$	0.812	ADRXWS	6
2	N-2	-H	-Cl	$-CH_3$	0.735	ADRXWS	6
3	N-3	-H	-F	$-C_6H_5$	0.721	ADRXWS	6
4	N-4	-H	-Cl	—C≡CH	0.702	ADRXWS	6
5	N-5	-H	-OH	$-CH_3$	0.667	ADRXWS	6
6	N-6	-H	-F	-Cyclopropane	0.648	ADRXWS	6
7	N-7	-H	-Br	$-CH_2-CH_3$	0.612	ADRXWS	6
8	N-8	$-CH_3$	-Br	—C≡CH	0.492	ADRXWS	6
9	N-9	-H	-Cl	$-C_6H_5$	0.419	ADRXWS	6
10	N-10	-OH	-OH	-F	0.353	ADRXWS	6
11	N-11	$-CF_3$	-Br	$-CH=CH_2$	0.273	ADRXWS	6
12	N-12	$-CF_3$	-H	$-CH=CH-CH_3$	0.239	ADRXWS	6
13	N-13	$-CF_3$	-OH	—C≡CH	0.116	ADRXWS	6
14	N-14	$-CF_3$	-Br	$-C_6H_5$	0.105	ADRXWS	6
15	N-15	-H	$-SO_2$	-Cyclopropane	-0.153	ADRXWS	6
16	N-16	-H	$-SO_2H$	$-CH_3$	-0.248	ADRXWS	6

3.6 Molecular Docking Studies:

The molecular docking tool, GLIDE (Schrodinger, LLC, New York, NY) was used for ligand docking studies in to the receptor tubuline polymerase enzyme binding pocket. The crystal structures of enzyme and co-crystalized ligand were obtained from protein data bank. (PDB Code: 1SA0). The goal of ligand–protein docking is to predict the predominant binding mode(s) of a ligand with a protein of known three-dimensional structure. Successful docking methods search high-dimensional spaces effectively and use a scoring function of dockings. Docking can be used to perform virtual screening on large libraries of compounds, rank the results, and propose structural hypotheses of how the ligands inhibit the target, which is invaluable in lead optimization[29].

3.6.1 Methodology:

The Glide (Schrodinger, LLC, New York, NY) software[30] was used to dock potential inhibitors (Ligand) in the binding pocket of the Protein (Tubuline) structure. Glide is most commonly used and validated software designed to assist in high-throughput screening of potential ligands based on binding mode and affinity for a given receptor molecule. One can compare ligand scores with those of other test ligands, or compare ligand geometries with those of a reference ligand. Glide approximates a complete systematic search of the

conformational, orientation, and positional space of the docked ligand. Glide (Grid-based Ligand Docking with Energetics) is one of the most accurate docking tools available for ligand-protein, protein-protein binding studies. Glide was found to produce least number of inaccurate poses and 85% of Glides binding models had an RMSD of 1.4 A^0 or less from native co-crystallized structures.

3.6.2 Receptor Preparation and Selection:

Docking studies were carried out using Tubuline complexe with DAMA colchicines (N-deacetyl-N-(2-mercaptoacetyl) colchicines). It was solved by X-ray diffraction techniques with a resolution of 3.58 Å. We retrieved it from the Gen-bank (code 1SA0). The quality of the results obtained from Glide depends critically on the quality of the starting structures. These starting structures must include all hydrogen's, have correct charge states near the binding site, and be reasonably free of major steric clashes. A typical PDB protein complex structure, as downloaded from the Research Collaboratory for Structural Bioinformatics (RCSB) web site (http://www.rcsb.org), has no hydrogen's and may have residues in unusual charge states. Therefore comprehensive protein preparation to ensure chemical correctness and optimization of protein structure was done in order to achieve best results. The typical structure file from the PDB is not suitable for immediate use in molecular modeling calculations as the PDB structure file consists only of heavy atoms and may include a co-crystallized ligand, water molecules, metal ions, and cofactors. Schrödinger has therefore assembled a set of tools to prepare proteins in a form that is suitable for modeling calculations. All structures were prepared for docking using the 'protein preparation wizard' in Maestro wizard 8.5[31]. In the refinement component, a restrained impact minimization of the co-crystallized complex was performed.

3.6.3 Receptor Grid Generation:

For receptors that adopt more than one conformation on binding, it is necessary to prepare grids for each conformation to ensure that possible actives are not missed. Grid files represent physical properties of a volume of the receptor (specifically the active site) that are searched when attempting to dock a ligand. Also the shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. Grids were generated by Receptor Grid Generation panel which define receptor structure by excluding any co-crystallized ligand that may be present, determine the position and size of the active site as it will be represented by receptor grids, and set up Glide constraints. Grids were defined by centering them on the ligand in the crystal structure using the default box size.

3.6.4 Ligand Preparation:

Ligand preparation was carried out using LigPrep panel in the software. The use of LigPrep produces a single low-energy 3D structure with correct chiralities for each successfully processed input structure. All the structures in format were imported in the project file and subjected to ligand preparation using OPLS 2005 force field using default setting[32].

Possible ionization states for each structure at the pH 7.0 ± 2.0 were generated using the ionizer option and only one low energy ring conformer per ligand was allowed to generate. Low energy stereo isomers 32

per ligand were allowed to generate to identify additional chiral atoms in the structures and generate additional structures with the same molecular formula but different chiral properties.

3.6.5 Docking and Scoring Function:

The ligands were docked with the active site using the 'Extra precision' Glide algorithm. Glide uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor. Final scoring of docked ligand is carried out on the energy-minimized poses Glide Score scoring function. Glide Score is based on Chem Score, but includes a steric-clash term and adds buried polar terms devised by Schrodinger to penalize electrostatic mismatches.

G Score = 0.065*vdW + 0.130*Coul + Lipo + Hbond + Metal + BuryP + RotB +

Site ---------- (3.3)

Where, vdW: - Vander Waal energy;

Cool: - Coulomb energy;

Lipo: - Lipophilic contact term;

H Bond: - Hydrogen-bonding term;

Metal: - Metal-binding term;

Bury P: - Penalty for buried polar groups;

Rot B: - Penalty for freezing rotatable bonds;

Site: - Polar interactions at the active site; and

The Coefficients of vdW and Coul are: - a = 0.065, b = 0.130.

The all 3D QSAR studies were carried out by using partial least square statistical method in combination with stepwise forward variable selection method.

3.6.6. Colchicine binding site on Tubuline:

In the tubulin–DAMA Colchicine crystal structure the A and C rings interacts with the β subunit and the B ring side chain interacts with the α subunit. The colchicine site in the tubulin DAMA colchicine crystal structure lies within the intermediate domain of the b subunit, surrounded by strands S8 and S9, loopT7 and helices H7 and H8. Besides β-subunit, colchicine also interacts with the loop T5 of the adjacent α-subunit. This structure supports the observation that tubulin heterodimer is stabilized upon colchicine binding[33]. The TMP moieties of DAMA-colchicine occupy space and are buried in the β -tubulin structure near residue Cys β 241. The dimensions of the colchicine site are ~ 10 Å× ~10 Å× ~4.5 Å. The colchicine site is located mostly in the β -subunit and is bordered in β -tubulin by helix 7, which contains Cys β 241, and helix 8. Although not as extensively as β -tubulin, α -tubulin also forms crucial interactions at the colchicine site, notably the loop connecting sheet 5 and helix 5. The latter contains Thr α 177 and Val α 179, both of which appear to form hydrogen bonds to the colchicines site inhibitors[34].

3.6.7. Results and Discussion:

All the designed compounds that confirmed top expected interest and accompanied Lipinski's rule were docked into Tubuline (PDB code 1sa0) for studying the binding mode of designed compounds and further screening to sort out the nice compound having right binding affinity which became in comparison with binding mode of Tubuline polymerization inhibitors i.e. Colchicines, consequences of which can be depicted in desk 8. The reliability of the docking results turned into first checked by means of evaluating the quality docking poses obtained for the co-crystallized inhibitor with its certain conformation. This become carried out with the aid of casting off co-crystallized ligand from their lively site and subjecting once more to docking into the binding pocket in the conformation discovered in the crystal structure. As a end result, a root suggest square deviation (RMSD) of 0.7 Å changed into found suggesting that the docking process may be depended on to predict the binding mode of our compounds. The correct prediction of protein-ligand interaction geometries is crucial for the success of digital- screening methods in shape-primarily based drug layout. The docking consequences were evaluated based on drift score (G- score), hydrogen bonds (H-bond) and vander waals (vdw) interplay among ligands and receptor.

Table 8. Results of Molecular Docking Studies:

Sr. No	Title	G-score	E-Model	H-Bond	Good VDW	Bad VDW	Ugly VDW
1	N-3	-8.46	-57.6	1	187	4	0
2	N-5	-7.84	-55.2	1	211	7	0
3	Std.	-7.22	-54.2	1	213	6	0
4	N-7	-6.96	-49.5	1	169	2	0
5	N-9	-6.84	-43.7	1	174	5	0
6	N-10	-6.27	-39.6	1	189	5	0
7	N-16	-6.01	-40.3	1	164	6	0

G score of compounds N-3 and N-5 was found to be -8.46 and -7.84 respectively which was comparable with G-score of standards colchicine (G score: -7.22) indicated that designed compounds have good binding affinity for binding to tubuline. The best poses obtained by docking results are reported in **Fig. 7 & 8,** where main interaction between ligands and receptors can be observed. All designed compounds adopt a very similar conformation at tubuline binding pocket.

Figure 7. Docking pose of colchicines at colchicines binding site:

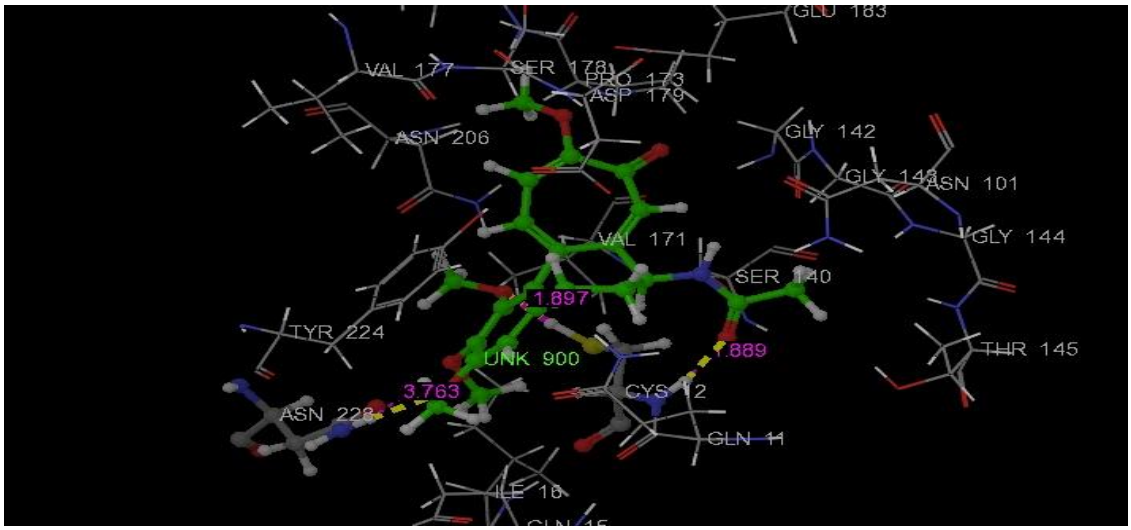
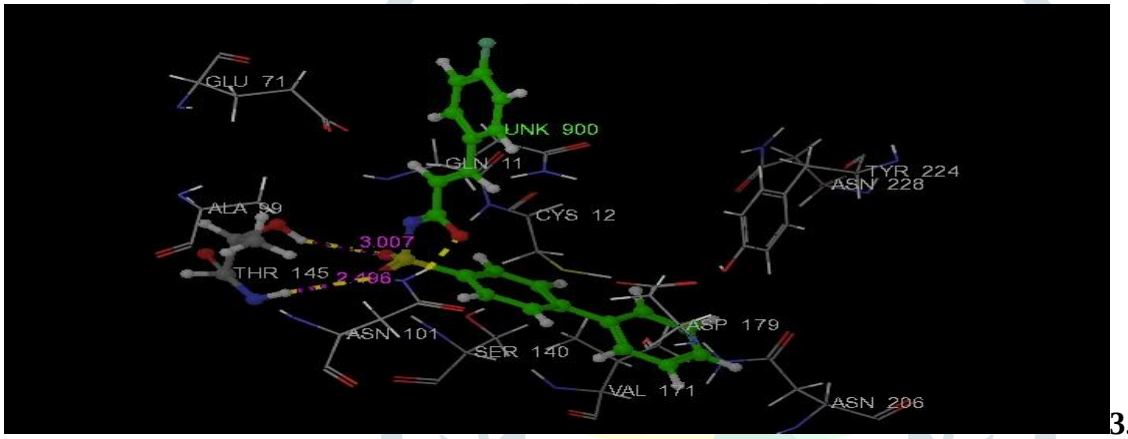


Figure 8. Docking pose of compounds N-3 at colchicines binding site:



3.7.

ADME Predictions:

All designed compounds were filtered by predicting their Absorption, Distribution, Metabolism, and Excretion (ADME) properties by means of Qikprop 2.2 Tool of, Schrodinger. It was used as last screening tool to design cinnamic acyl sulfonamide derivatives for anticancer activity. In addition to predicting molecular properties, Qikprop provides ideal ranges of these properties for comparing a particular property with those of 95% of known drugs. Number properties of designed analogues are predicted by Quikpro tool but here we have reported significant descriptors which are required for predicting drug like properties of the molecule. These properties are:

1) Rule of five: It includes Molecular Weight (MW) (150-650), Predicted octanol/water partition coefficient. (QP log Po/w < 5), estimated number of hydrogen bond donor (donor HB ≤ 5), estimated number of hydrogen bond acceptor (accpt HB≤ 10). Compounds that satisfy these rules are considered to have drug-like pharmacokinetic profile.

2. Brain/blood partition coefficient (CNS) (-2 to 2)

3. Percent Human Oral absorption (>80% is high, <25%is poor).

4. Number of possible metabolites (should range from 1-8)

The parameters illustrated in **Table 9** Qikprop analysis show significant results. CNS parameter is related with absorption of entity through Blood brain barrier, standard limit for CNS is -2 to +2, where -2 shows inactive CNS penetration and +2 shows active CNS penetration. All the designed entities show satisfactory results within the range. Percent Oral Absorption parameter is related with extent of oral absorption of drug, indicating suitable route of administration, if it is going to be formulated. Here all entities shows more than 80% oral absorption, it is considered to be highly absorbed. The result of prediction of ADME properties shows in **Table 9**.

Table 9. Prediction of ADME properties:

Sr. No	Comp. Code	Mol. Wt	Donor HB	Accept HB	QPlog Po/w	Percent Human Oral Absorption	CNS	No.Of possible metabolites
1	N-3	381.42	1	5.45	2.977	100	0	2
2	N-5	317.35	0	9.65	3.883	100	-1	3
3	N-7	394.28	0	6.95	3.731	100	0	4
4	N-9	397.87	2	6.2	2.358	87.307	-2	3
5	N-10	337.32	1	8.15	3.218	100	-1	3
6	N-16	365.42	2	8.9	2.582	91.579	-2	4

4. Conclusion:

On this look at, Goal became targeted on improvement of the capability by-product of Cinnamic Acyl Sulfonamide for anticancer interest by using pharmacophore optimization the use of QSAR research. In this we've got completed 2D QSAR studies the usage of a couple of linear regression (MLR) technique and 3D QSAR studies using simulated annealing okay-nearest neighbour molecular subject evaluation approach (SA-KNN) which offers meaningful descriptor having their contribution in hobby. On the idea of 2D & 3D QSAR research we have optimized the pharmacophore. The optimized molecule were then subjected to the lipinski screening and sixteen designed compounds suggests the coolest Lipinski's rule. ADRXWS. All the designed compounds that show exact predicted activity and follow Lipinski's rule have been Docked and filtered by way of predicting their absorption, distribution, metabolism, and excretion (ADME) homes of designed compounds and further screening to type out the pleasant six compound having appropriate predicted activity.

5. Acknowledgement:

Authors are thankful to Dr. A. R. Madgulkar, Principal of our institute for continuous motivation, support and for providing the necessary infrastructure to carry out this work.

5. References:

1. Frace, A.; Loge, C.; Gallet, S.; Lebegue, N.; Carato, P.; Chavatte, P.; Berthelot, P.;Lesieur, D. J. Enz. Inh. Med.Chem.2004,19, 541.
2. Fadeyi, O. O.; Adamson, S. T.; Myles, E. L.; Okoro, C. O. Bioorg. Med. Chem. Lett. 2008, 18, 4172.
3. Valiron, O.; Caudron, N.; Job, D. Cell. Mol. Life Sci. 2001,58, 2069.
4. Hamel, E. Med. Chem. Rev. 1996, 16, 207.
5. Kim, D. Y.; Kim, K. H.; Kim, N. D.; Lee, K. Y,; Han, C. K.; Yoon, J. H.; Moon, S. K.; Lee, S. S.; Seong, B. L. J. Med. Chem. 2006, 49,5664.
6. Chinigo, G. M.; Paige, M.; Grindrod, S.; Hamel, E.; Dakshamamurthy, S.; Chruszcz, M.; Minor, W.; Brown, M. L. J. Med. Chem. 2008, 51, 4620.
7. Romagnoli, R.; Baraldi, P. G.; Carrion, M. D.; Cruz-Lopez, O.; Cara, C. L.; Basso, G.; Viola, G.; Khedr, M.; Balzarini, J.; Mahboobi, S.; Sellmer, A.; Brancale, A.; Hamel, E. J. Med. Chem. 2009, 52,5551.
8. Ravelli, R. B.; Gigant, B.; Curmi, P. A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Nature 2004, 428, 198.
9. Chang, J. Y.; Hsieh, H. P.; Chang, C. Y.; Hsu, K. S.; Chiang, Y. F.; Chen, C. M.; Kuo, C. C.; Liou, J. P. J. Med. Chem. 2006,49, 6656.
10. Park, E. J.; Park, H. R.; Lee, J. S.; Kim, J. Planta Med. 1998, 64, 464.
11. Claude, A. C.; Jean, C. L.; Patric, T.; Christelle, P.; Gerard, H.; Albert, J. C.; Jean, L. D. Anticancer Res. 2001,21, 3949.
12. Kumar, S. K.; Erin, H.; Catherine, P.; Halluru, G.; Davidson, N. E.; Khan, S. R. J. Med. Chem. 2003, 46, 2813.
13. Pouria H. Jalily, John A. Hadfield, Nicholas Hirst, Steven B. Rossington, Bioorganic & Medicinal Chemistry Letters 22 (2012) 6731–6734.
14. Yin Luo, Yao Li, Ke-Ming Qiu, Xiang Lu, Jie Fu, Hai-Liang Zhu, Bioorganic & Medicinal Chemistry 19 (2011) 6069–6076.
15. Dudek A.Z., *Combinatorial Chemistry*, (**2006**) 213-228.
16. Golbraikh, A. Tropsha, A. J. Mol. Graphics Model. 20, (**2002**) 269–276.
17. Eriksson, L.; Jaworska, J. Worth, A.P.; Cronin, M.T.et al,Environ. Health Perspect., 111 (10),.(**2003**).1361-1375.
18. S. R. Johnson, J. Chem. Inf. Model., 48, (**2008**) 25–26.
19. Alexander Tropsha J. Mol. inf , 29, (**2010**) 476-488.
20. Yin Luo, Ke-Ming Qiu, Xiang Lu, Kai Liu, Jie Fu, Hai-Liang Zhu, Bioorganic &

Medicinal Chemistry 19 (2011) 4730–4738.

21. VLifeMDS; Molecular Design Suite version 3.5,(**2004**) V-life Sciences Technologies Pvt. Ltd., Pune, India. (www.vlifesciences.com).
22. John G. Topliss, Robert P. Edwards, J. Med. Chem. 1979, 22, 1238-1244.
23. Weifan Zheng, Alexander Tropsha, J. Chem. Inf. Comput. Sci. 2000, 40, 185-194.
24. Andrew R. Leach, An Introduction to Chemoinformatics, Springer press, **(2007),** pp125-170.
25. Subhash Ajmani, Kamalakar Jadhav,Sudhir Kulkarni; J. Chem. Inf. Model. **(2006)**, 46, 24-31.
26. W. Zheng, A. Tropsha ,J. Chem. Inf. Comput. Sci., 40, (**2000**) 185 – 194.
27. 29. Knut Baumann, J. Chem. Inf. Comput. Sci, 42, (**2002**) 26-35.
28. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, PJ. Adv. Drug Delivery Rev., 23, (**1997**)3.
29. Hawkins, P.C.; Skillman, A.G.; Nicholls, A. J Med Chem, 2007, 50, 74–82.
30. Glide, molecular docking tool of Schrodinger Inc., version 5.0, New York, USA.
31. Maestro, Version 8.5 (2008) Schrodinger LLC, New York.
32. Hayes, M. J.; Stein, M.; Weiser, J. J. Phys. Chem. 2004, 108, 3572–3580.
33. Bhabatarak Bhattacharyya,Dulal Panda,Suvroma Gupta,Mithu Banerjee, Medicinal Research Reviews, 2008,Vol. 28, No. 1, 155-183.
34. Tam Luong Nguyen,Connor McGrath, Ann R. Hermone, James C. Burnett, Daniel W. Zaharevitz, Billy W. Day, Peter Wipf, Ernest Hameland Rick Gussio, J. Med. Chem. 2005, 48, 6107-6116.