Original Research Article

2D QSAR APPROACH TO DEVELOP NEWER GENERATION SMALL MOLECULES ACTIVE AGAINST SMALL LUNG CANCER CELL LINE DMS 114

Supriyo Saha[1*], Prinsa[2], Mrityunjoy Acharya[3]

1. Sardar Bhagwan Singh Post Graduate Institute of Biomedical Sciences and Research.
1. Balawala, Dehradun, India.
2. Ved Life Savers (P) Limited, Dehradun.
3. Rastriya Bal Suraksha Karyakam, Howrah, West Bengal, India.

ABSTRACT

Quantative Structure Activity Relationship analysis was performed using 38 small molecules without any particular scaffold worked against small lung cancer cell line DMS 114. The QSAR model was pIC_{50} = 32.72228(+/-9.85895) +0.16592(+/-0.11717) ALogP -0.00745(+/-0.00466) AMR -3.74232(+/-1.26299) Mi +0.3363(+/-0.03428) RDF110m. Statistical information for that equation was SEE :0.81811, r^2 :0.8621, r^2 adjusted :0.83584, F :32.82184 (DF :4, 21) which suggested that AlogP (Ghose-Crippen LogKo/w), RDF110m (Radial distribution function - 100 / weighted by relative mass) create positive response and AMR (Molar refractivity), Mi (Mean ionization potentials (scaled on carbon atom)) create negative response towards PIC_{50} value. Then the model was validated through Golbraikh and Tropsha acceptable model criteria's as Q^2:0.77691 Passed (Threshold value Q^2>0.5),r^2: 0.61064 Passed (Threshold value r^2>0.6, |r0^2-r'0^2|: 0.11623 Passed with Threshold value |r0^2-r'0^2|<0.3). As well as the greater q2 value was suggested the model sustainability. Applicabilty domain was identified by Euclidean and Mahalanobis Distance Method. All the points were merely overlapped with observed and predicted IC50 value. So the developed QSAR model will work as a great predictor of its activity with any chemical scaffold.

***Keywords*:** Small Lung Cancer, PADEL, Stepwise regression, FA-MLR, Golbraikh and Tropsha acceptable model, Euclidean and Mahalanobis Distance.

***Correspondence**: Supriyo Saha,** Sardar Bhagwan Singh Post Graduate Institute of Biomedical Sciences and Research, Dehradun-248161. Uttarakhand, INDIA. Email Id: supriyo9@gmail.com. Tel: 07895424583.

RUNNING TITLE: 2D QSAR study against small lung cancer inhibitors.

INTRODUCTION

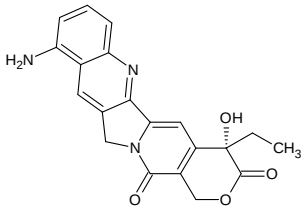
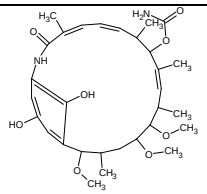
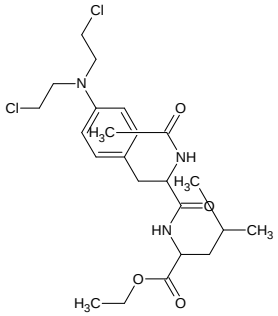
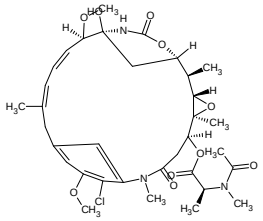
Small cell lung cancer was an ailment in which harmful (disease) cells frame in the tissues of the lung [1, 2]. The most common cause of cancer-related deaths in Europe in 2006 is lung cancer (estimated 334 800 deaths). After prostate cancer, lung cancer is the most frequent type of cancer in men. Age-standardized incidence and mortality rates in 2006 were estimated to be 75.3 and 64.8/100 000/year, respectively, in men, and 18.3 and 15.1/100 000/year in women. Small-cell lung cancer (SCLC) was accounts for 15%–18% of all cases [3]. In recent years the incidence of SCLC has decreased. SCLC is strongly associated with tobacco smoking. Small Cell Lung Cancer was dealt with by chemotherapy and radiation therapy, surgery and Prophylactic Cranial Illumination (PCI) [4]. There were 221,200 new cases and 158,040 deaths occurred in 2015 in the United States [5]. As per Developmental Therapeutics Program NCI,NIH various drug molecules such as Allocolchicine, Amonafide, Busulfan, Camptothecine, Etoposide, ,Melphalan,

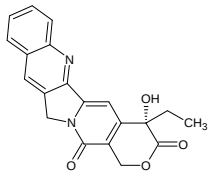
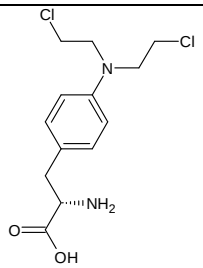
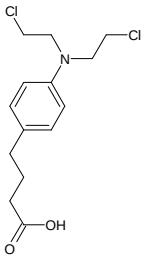
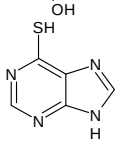
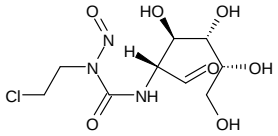
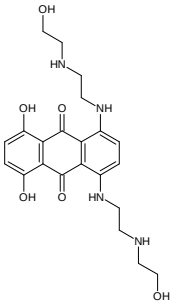
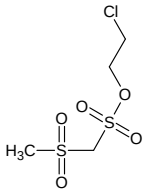
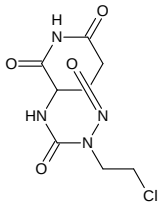
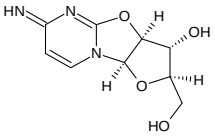
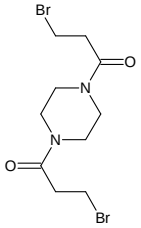
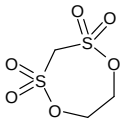
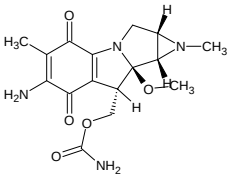
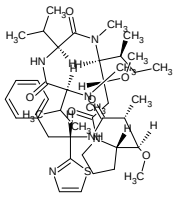
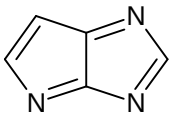
Dolastatin10, Lomustine, Mytansine and lots of other molecules were active against DMS 114 small cancer cell line [6,7,8] as well as all the molecules were not in a particular scaffold of chemical structure. In this recent work, our main intention is to develop a QSAR model to predict the activity spectrum of a new molecule to become active against DMS114 (small lung cancer cell line, established from cells from a mediastinal biopsy of a patient with small cell carcinoma of the lung) as well as till date there was no QSAR model to predict small lung cancer antagonize molecule with diversified scaffold.

MATERIALS AND METHODS

QSAR model was conducted by using a set of theoretical and constructive descriptors, which were calculated by PaDEL-descriptor: an open source software, ToMoCoMD. QSAR Model was constructed by use of MLR Plus Validation Tool. More than 1875 descriptors include Ghose-Cripen Log Ko/w, Ghose-Crippen molar refractivity, Sum of the atomic polarizabilities (including implicit hydrogens), Wildman-Crippen LogP and MR, Wildman-Crippen MR, Eccentric Connectivity Index: topological descriptor combining distance and adjacency information, H Bond Acceptor Count Number of hydrogen bond acceptors, McGowan characteristic volume, Wiener Polarity Number were calculated by PADEL and ToMoCoMD. All the explanations of relevant descriptors were enlisted in Table 1 & 2. A descriptor represents a quantitative property depends on the molecular structure. Theoretical descriptors were advantageous due to its free from uncertainty of experimental measurement and can be calculated for compounds before synthesis. Theoretical descriptors were employed in this QSAR study to model as an inhibitor of small lung cancer cell line DMS 114 as a potential anticancer agent.

Table 1: Molecules in Training Set

SN	Name	Structure	pIC_{50}	SN	Name	Structure	pIC50
1.	Aminocamptothecin		4.8	14.	Macbecin II		9.0
2.	Asalex		5.9	15.	Maytansine		4.1

3.	Camptothecin		4.0	16.	Melphalan		8.0
4.	Chlorambucil		4.1	17.	Mercaptopurine		5.5
5.	Chlorozotocin		3.2	18.	Mitoxantrone		4.0
6.	Clomesone		2.9	19.	PCNU		4.0
7.	Cyclocitidine		3.0	20.	Pipobroman		3.5
8.	Cyclodisone		4.0	21.	Porfiromycin		3.8
9.	Dolastatin 10		8.0	22.	Pyrazoloimidazole		2.0

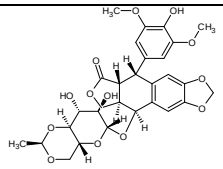
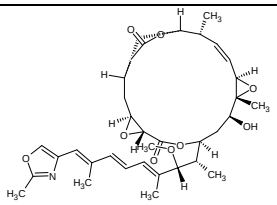
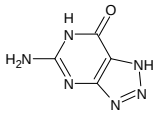
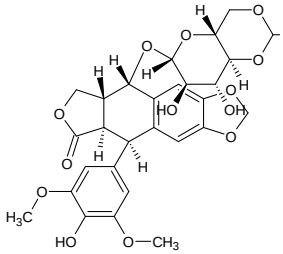
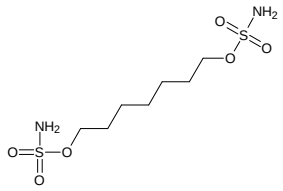
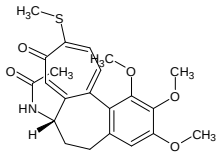
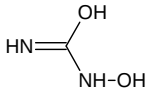
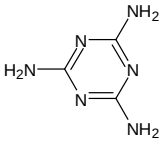
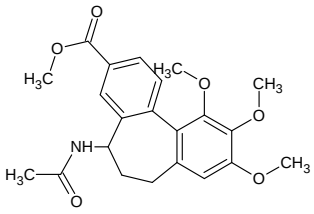
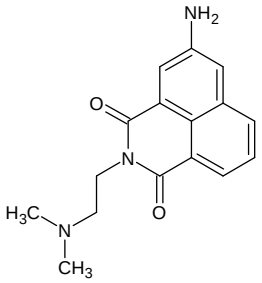
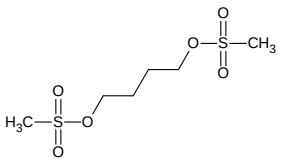
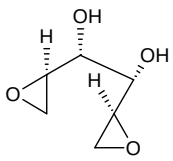
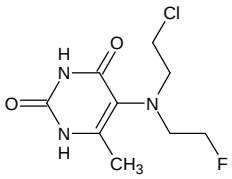
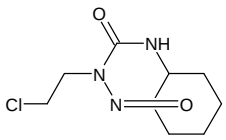
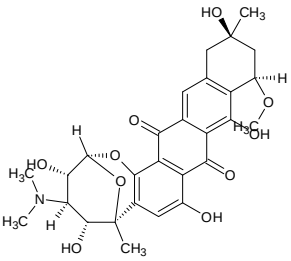
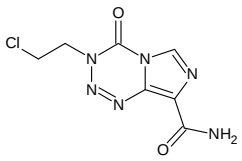
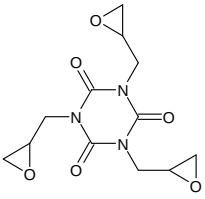
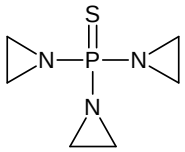
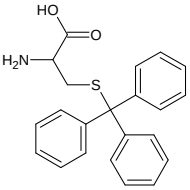
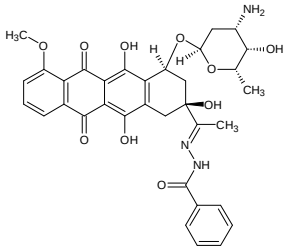
10.	Etoposide		6.0	23.	Rhizoxin		9.3
11.	Guanazole		2.0	24.	Teniposide		4.7
12.	Hepsulfam		2.7	25.	Thiocolchicine		4.0
13.	Hydroxyurea		2.6	26.	Triethylenemelamine		4.2

Table 2: Molecules in Test Set

Name	Structure	pIC_{50}	Name	Structure	pIC50
Allocolchicine		4.0	Amonafide		4.1
Busulafn		3.6	Dianhydrogalacitol		3.8

Flurodpan		3.0	Lomustine		4.2
Menogaril		3.1	Mitozolamide		2.9
Teroxirone		3.8	Thiotepa		4.2
Tritylcysteine		4.2	Zorubicin		5.2

Dataset and Descriptor Calculation

Dataset of 38 small lung cancer cell line inhibitors was downloaded from http://dtp.nci.nih.gov/docs/cancer/searches/standard_mechanism.html. All the molecules SMILE format were transferred into .mol format by ACDLABS and structures were optimized. 2D and 3D descriptors were calculated using PADEL descriptor [9] and ToMoCoMD [10] software. Table 1 & 2 were showed the detail dataset along with chemical structure, LC_{50} value and pLC_{50} value and Table 3 resulted with useful descriptor explanation.

Descriptor Pretreatment

Inter correlated descriptor was cut off using V-WSP as variance cut off 0.0001 and correlation coefficient value 0.99.

Dataset Division

Total dataset of 38 molecules was divided into Training and Test set using Kennard Stone method as 26 molecules were in Training set and 12 molecules in Test set.

Suitable Descriptor Selection

Suitable descriptor selection was performed using Stepwise MLR as F values 3.9 to 4.0. Then best subset was selected using 4 descriptor combination and r2 cut off value 0.6.

The chemometric tool

The development of QSAR equation was implemented two methods (1) Stepwise regression (2) multiple linear regressions with factor analysis as pre processing factor analysis for variable selection (FA-MLR).

Table 3: List of relevant descriptor with explanation

SN	Abbreviation descriptors	Explanation of descriptors
1.	AlogP	Ghose-Crippen LogKow
2.	SpMin3_Bhv	Burden modified eigenvalues
3.	ntN	Total number of Nitrogen Atoms
4.	ETA_Beta_ns	A measure of electron-richness of the molecule
5.	Crippen MR	Crippen's molar refractivity
6.	McGowan Volume	McGowan characteristic volume
7.	VABC	Van der Waals volume calculated
8.	nRing	No of Ring
9.	nRotb	No of Rotatable Bonds
10.	Phia Kappa	flexibility Index
11.	Bac	Balaban Centric Index
12.	AlogP	Ghose-Crippen LogKow
13.	Crippen LogP	Crippen's LogP
14.	XLogP	XLogP
15.	AMR	Molar refractivity
16.	TopoPSA	Topological Polar Surface Area
17.	Wpol	Weiner polarity number
18.	MW	Molecular Weight
19.	ETA	Electro Topochemical Descriptor

Stepwise regression

Multi step linear equation, a multistep equation was built by step by step. The basic procedure involved: (i) Identifying an initial model (ii) Repeating the previous step by altering descriptor or variable combination to achieve better f and r2 value. (iii) Calibrate the equation by justify the values in between observed and predicted values. The stepwise MLR was performed using statistical software SPSS and it was judged by parameters as explained variance (r^2a), correlation coefficient (r), standard error of estimate (s) and variance ratio (F) at a specified degree of freedom (DF). All accepted MLR equation had regression level significant at 95 and 99% levels. The generated QSAR equation was validated by leave one out or LOO method using Minitab software and different parameters like cross validation r2 (q2), standard deviation based on press (S_{PRESS}) and standard deviation of error of prediction (SDEP) [11].

FA-MLR

In this case a final statistical tool was used to develop a QSAR relation, factor analysis as a data pre processing step to identify the important factor to identify the important variables contributing the response variable by avoiding co linear value. The data matrix is first standardized and correlation matrix and subsequently reduced correlation matrix. An eigen value problem is then solved and the factor pattern can be obtained from the corresponding eigen vectors. The main objectives are to display multidimensional

data in space of lower dimensionality with minimum loss of information (explaining > 95% of variance of data matrix) and to extract the basic features behind the data with ultimate goal of interpretation [12].

QSAR Equation Development

MLR Plus valid software was used to developed QSAR equation, where LC50 was converted pLC50 value [13, 14].

QSAR Equation Validation

Golbraikh and Tropsha acceptable model criteria's [15,16,17] to validate a QSAR Equation

1. Q^2 value is Passed (Threshold value Q^2>0.5).
2. r^2 value is Passed (Threshold value r^2>0.6).
3. |r0^2-r'0^2| value is Passed (Threshold value |r0^2-r'0^2|<0.3).

QSAR Equation Validation

The model was cross validated using Leave-One-Out (LOO) process [18]. Applicability domain of the developed QSAR equation was checked based on the response and chemical structure space in which the QSAR model makes predictions with a given reliability. Euclidean distance [19] and Mahalanobis [20] distance method. The distance of a test compound to its nearest neighbor in the training set is compared to the predefined applicability domain threshold.

RESULTS AND DISCUSSION

The statistically suitable QSAR model was pIC_{50} = 32.72228(+/-9.85895) +0.16592(+/-0.11717) ALogP -0.00745(+/-0.00466) AMR -3.74232(+/-1.26299) Mi +0.3363(+/-0.03428) RDF110m. Statistical information for that equation was SEE :0.81811, r^2 :0.8621, r^2 adjusted :0.83584, F :32.82184 (DF :4, 21. The Leave-One-Out (LOO) Result [21, 22] for that derived model was Q2 :0.77691, PRESS :22.73889, SDEP :0.93519 [23, 24]. The model was suggest that AlogP (Ghose-Crippen LogKo/w), RDF110m (Radial distribution function - 100 / weighted by relative mass) create positive response and AMR (Molar refractivity), Mi (Mean ionization potentials (scaled on carbon atom) create negative response towards PIC_{50} value.

Table 4: Applicability Domain data by Euclidean Distance for Training Set

CompNo	Euclidean_Distance	Mean_Distance	Norm_MeanDistance
1	793.996	31.7598	0
2	1218.38	48.7351	0.225446
3	798.808	31.9523	0.002556
4	813.158	32.5263	0.01018
5	803.8	32.152	0.005208
6	913.518	36.5407	0.063494
7	828.975	33.159	0.018582
8	1070.91	42.8365	0.147108
9	2676.41	107.056	1
10	1094.71	43.7883	0.159748
11	1500.24	60.0095	0.375178
12	858.793	34.3517	0.034422

13	1459.27	58.3709	0.353417
14	1722.93	68.9172	0.493481
15	2163.86	86.5546	0.727719
16	890.055	35.6022	0.05103
17	1494.56	59.7825	0.372164
18	853.53	34.1412	0.031627
19	797.16	31.8864	0.001681
20	865.514	34.6205	0.037993
21	833.049	33.322	0.020746
22	1016.19	40.6474	0.118035
23	1186.6	47.4641	0.208566
24	2188.54	87.5418	0.74083
25	802.691	32.1076	0.004619
26	824.799	32.992	0.016364

Table 5: Applicability Domain data by Euclidean Distance for Test Set

CompNo	Euclidean_Distance	Mean_Distance	Norm_MeanDistance
1	242.525	22.0477	0.071257
2	239.774	21.7976	0.059487
3	225.867	20.5334	0
4	302.8	27.5272	0.329095
5	243.046	22.0951	0.073485
6	229.802	20.8911	0.016832
7	434.553	39.5048	0.892699
8	253.757	23.0688	0.119306
9	459.637	41.7852	1
10	453.336	41.2123	0.973044
11	255.02	23.1836	0.124706
12	356.799	32.4363	0.56009

Table 6: Applicability Domain data by Mahalanobis Distance for Training Set

CompNo	MahaDist
1	1.85447
2	1.46744
3	1.95945
4	1.87621
5	1.08202
6	1.32065
7	0.941459
8	1.2401
9	2.9618
10	1.95432
11	3.05896
12	1.85028
13	2.93352
14	2.56264
15	2.27964

16	2.40749
17	2.0962
18	2.41579
19	0.847503
20	1.19229
21	1.4454
22	1.20473
23	2.74802
24	2.40824
25	0.197387
26	0.932951

Table 7: Applicability Domain data by Mahalanobis Distance for Test Set

CompNo	MahaDist
1	1.14769
2	1.89797
3	0.741646
4	1.73997
5	1.95536
6	2.16842
7	1.88626
8	1.41337
9	1.97466
10	2.52221
11	1.35605
12	3.05323

Table 8: Comparison Data in between pIC50 observed and pIC50 predicted

pIC50 observed	**pIC50 predicted**
4.8	3.94
5.9	5.59
4	4.23
4.1	4.3
3.2	3.68
2.9	3.29
3	3.08
4	3.71
8	7.27
6	6.85
2	1.84
2.7	5.04
2.6	2.16
9	8.73
4.1	4.06
8	8.22

5.5	5.72
4	2.38
4	3.46
3.5	3.33
3.8	3.59
2	3.05
9.3	8.1
4.7	5.18
4	4.23
4.2	3.45
4	3.95
4.1	3.81
3.6	3.61
3.8	3.35
3	3.06
4.2	4.67
3.1	3.27
2.9	3.06
3.8	3.41
4.2	3.75
4.2	3.27
5.2	8.87

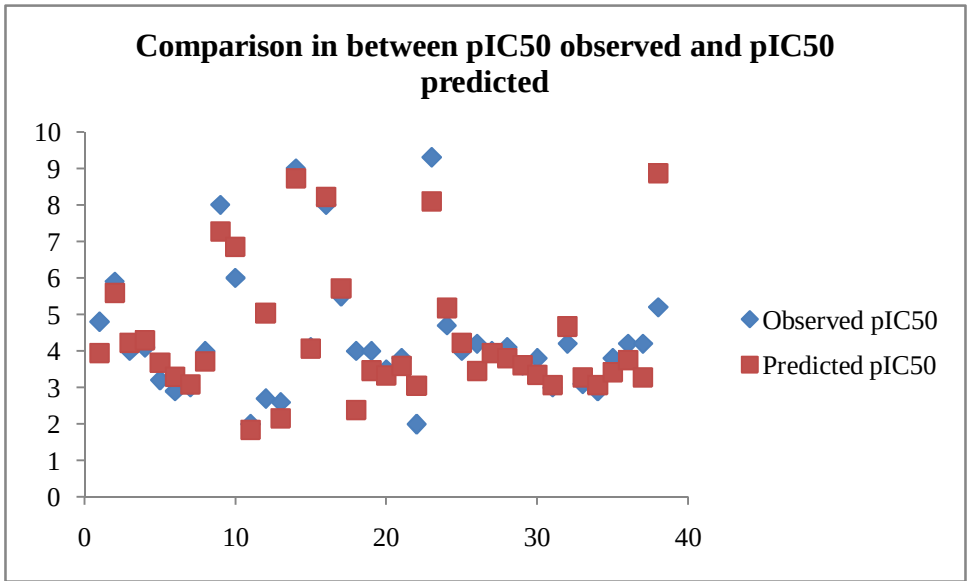


Figure 1: Comparison data in pIC50 observed and pIC50 predicted

The molecular descriptor such as AMR or molar refracrtivity is an important parameter because when a radiation with infinite wavelength, molar refractivity represents the real volume of a molecule but also to the London dispersive forces that act in between drug-receptor interaction, Ghose–Crippen is the SPARTAN default method of calculating log P. This method depends only on the connectivity of the molecule, and it is independent of the wavefunction (i.e., one will get the same results for semi-empirical, HF, and DFT methods but this depends on how the molecule is drawn/connected). The Ghose–Crippen model is

parameterized for 110 atom types, including common bonds of H, C, N, O, S, and the halogens. Avoiding correction factors was obtained evaluating the hydrophobicity on an individual atom basis, accounting for the undeniable intramolecular interactions by employing a large number of atom types, as we know the amount of ionization potential will increase with each removal of electron from the atom; it occurs due to the increasing order in the firmness of the remaining positive charged atom with each other, here in the QSAR model ionization potential must be in decreasing order for the better activity against small lung cancer cell line and in case of Radial distribution function - 100 / weighted by relative mass, which was a descriptors are based on the distance distribution in the molecule, which was interpreted as the probability distribution of finding an atom in a spherical volume of radius *R*. This descriptor encoded with 3D chemical structure weighted by polarizability, electronegativity and molar volume as well as it correlates with the influence of the electronic structure and state in electroluminescence [25, 26]. There were some external validation parameters without scaling and after scaling. The external validation parameters without scaling were r^2 :0.61064, r0^2 :-1.63343, reverse r0^2:0.4828,rm^2(test) :-0.30411,reverse rm^2(test) :0.39231, average rm^2(test) :0.0441, delta rm^2(test) :0.69642, rmsep:1.12933, rpred^2 :-0.36387, Q2f1 :-0.36387, Q2f2 :-2.37916 and after scaling were rm^2(test) :0.13331, reverse rm^2(test) :0.4531, average rm^2(test) :0.2932, delta rm^2(test) :0.31978. Then the model was validated through Golbraikh and Tropsha acceptable model criteria's as Q^2:0.77691 Passed (Threshold value Q^2>0.5),r^2: 0.61064 Passed (Threshold value r^2>0.6, |r0^2-r'0^2|: 0.11623 Passed with Threshold value |r0^2-r'0^2|<0.3). As well as the greater q2 value was suggested the model sustainability. Applicabilty domain was identified by Euclidean and Mahalanobis Distance Method and all the results were diagrammatized at Table 4, 5, 6, 7. The data from Euclidean distance method was confirmed that cyclodisone from training set and mitozolomide from test set were outside the applicability domain. The outcomes from mahalanobis distance was suggested that all the data from training set were normally distributed within 0.197387 to 3.05896 and in case of test set this distribution was occurred in between most of the molecules were inside the 0.741646 to 3.05323. Finally calculate the observed and predicted IC50 value and diagrammatized in Table 8 and Figure 1 was showed that all the points were merely overlapped with each other.

CONCLUSION

It can be easily concluded that if in future we have to develop a small molecule working against small lung cancer cell line, the developed QSAR model will work as a great predictor of its activity with any chemical scaffold and by which we can produce a good molecule with higher activity profile.

CONFLICT OF INTEREST

There is no conflict of interest associated with the authors of this paper.

ACKNOWLEDGEMENTS

One of the authors is highly acknowledged towards Prof. Veerma Ram, Head of the Department SBSPGI for his unconditional support and well beings.

REFERENCES

1. Heighway J, Betticher DC. Lung: Small cell cancer. Atlas Genet Cytogenet Oncol Haematol. 2004; 8(3): 257-259.
2. Prasad US, Naylor AR, Walker WS. Long term survival after pulmonary resection for small cell carcinoma of the lung. Thorax. 1989; 44(10): 784-7.
3. Sorensen M, Johannesma MP, Felip E. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2010; 21(5):120-125.

4. Klasa RJ, Murray N, Coldman AJ. Dose-intensity meta-analysis of chemotherapy regimens in small-cell carcinoma of the lung. J Clin Oncol. 1991; 9 (3): 499-508.
5. Mettler FA Jr, Thomadsen BR, Bhargavan M, et al.: Medical radiation exposure in the U.S. in 2006: preliminary results. Health Phys 95 (5): 502-7, 2008.
6. Grever MR, Schepartz SA, Chabner BA. The National Cancer Institute: Cancer Drug Discovery and Development Program. Seminars in Oncology. 1992; 19(6): 622-638.
7. Boyd MR, Paull KD. Some Practical Considerations and Applications of the National Cancer Institute In Vitro Anticancer Drug Discovery Screen. Drug Development Research. 1995; 34: 91-109.
8. Shoemaker RH. The NCI60 Human Tumour Cell line Anticancer Drug Screen. Nature Reviews. 2006; 6: 813-823.
9. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011; 32(7): 1466-74.
10. Garcia J, Marrero-Ponce CR, Acevedo-Martínez Y, Barigye L, Valdes-Martini, SJ, Contreras-Torres JR. QuBiLS-MIDAS: A Parallel Free-Software for Molecular Descriptors Computation based on Multi-Linear Algebraic Maps. J Comput Chem. 2014; 35: 1395–1409.
11. Roy K, Kar S. The rm2 metrics and regression through origin approach: reliable and useful validation tools for predictive QSAR models (Commentary on 'Is regression through origin useful in external validation of QSAR models?'). Eur J Pharm Sci. 2014; 62: 111– 114.
12. Roy K, Mitra I. On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. Comb Chem High Throughput Screen. 2011; 14(6): 450-74.
13. Roy K, Das RN, Paul LA. Quantitative structure–activity relationship for toxicity of ionic liquids to Daphnia magna Aromaticity vs. lipophilicity. Chemosphere. 2014; 112: 120-127.
14. Garg R, Carr JS. Predicting the bioconcentration factor of highly hydrophobic organic chemicals. Food Chem Toxicol. 2014; 69: 252–259.
15. Dearden JC, Hewitt M. QSAR modeling of bioconcentration factor using hydrophobicity, hydrogen bonding and topological descriptor. SAR QSAR Environ Res. 2010; 21: 671–680.
16. Dimitrov SD, Dimitrova NC, Walker JD, Veith GD, Mekenyan OG. Predicting bioconcentration factors of highly hydrophobic chemicals. Effects of molecular size. Pure Appl Chem. 2002; 74: 1823–1830.
17. Ravichandran V, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK. Validation of QSAR Models - Strategies and Importance. Int J Drug Design Discovery. 2011; 3: 511-519.

18. Krstajic D, Buturovic LJ, Leahy DE, Thomas S. Cross-validation pitfalls when selecting and assessing regression and classification models. J Cheminform. 2014; 6(10): 1-15.

19. Zhang S, Golbraikh A, Oloff S, Kohn H, Tropsha A. J Chem Inf Model. 2006; 46:1984–1995.

20. Jaworska J, Nikolova-Jeliazkova N, Aldenberg T. QSAR Applicability Domain Estimation by Projection of the Training Set in Descriptor Space: A Review. ATLA. 2005; 33: 445–459.

21. In silico ADME-Toxicity Profiling, Prediction of Bioactivity and CNS Penetrating Properties of some Newer Resveratrol Analogues. J Pharma Sci Tech. 2014; 3(2): 98-105.

22. Saha S, Acharya M. Discovery of Hydrazinecarboxamide or Hydrazinecarbothioamide Bearing Small Molecules as Dual Inhibitor of Ras Protein and Carbonic Anhydrase Enzyme as Potential Anticancer Agent Using Validated Docking Study and *In-silico* ADMET Profile. Research J Pharmaceutical Biological Chemical Sciences. 2014; 5(3): 1884-93.

23. Saha S, Acharya M. Hydrazinecarboxamide or hydrazinecarbothioamide Bearing small molecules as Dual inhibitor of Ras protein and Carbonic anhydrase enzyme as potential anticancer agent – A MLR approach based on docking energy. Research in Biotechnology. 2014; 5(6): 15-23.

24. Saha S, Banerjee A, Rudra A. 2D QSAR approach to develop newer generation moleculesactive aganist ERBB2 receptor kinase as potential Anticancer Agent. Int J Pharma Chemistry. 2015; 5(4): 134-148.

25. Laura G, Salvatore G. Comparing Log *p* calculations by the ghose- crippen and villar methods. www.iupac.org/publications/cd/medicinal_chemistry. 2006; 1-13.

26. Fernandes AG, Stefani R. Quantitative structure-property relationships of electroluminescent materials: Artificial neural networks and support vector machines to predict electroluminescence of organic molecules. Bull Mater Sci. 2013; 36(7): 1307–1313.