Original Research Article

EXPLORING DESCRIPTOR COMBINATION BY CHEMOMETRIC APPROACH TO DEVELOP NEWER MOLECULES ACTIVE THROUGH CORTICOSTEROID BINDING GLOBULIN RECEPTOR

Supriyo Saha*, Vimal Constance, LuvKush[1], Versha Percha

Sardar Bhagwan Singh Post Graduate Institute of Biomedical Sciences and Research, Balawala, Dehradun, India.

ABSTRACT

Quantative Structure Activity Relationship analysis was performed using 30 various steroidal moieties with their IC_{50} value. The QSAR model was pIC_{50} = -0.01499 (+/-0.00582) +0.02256 (+/-0.00794) GATS5e - 0.02574 (+/-0.00344) RDF120v -0.61334 (+/-0.00342) Ds with statistical information: SEE: 0.00581, r^2: 0.99977r^2 adjusted: 0.99973, F: 24389.67031 (DF: 3, 17). This model proposes that by expanding the Sanderson electro negativities and by diminishing the Radial dispersion capacity - 120/ weighted by relative van der Waals volumes and aggregate availability record/ weighted by relative I-state esteem it make a positive reaction. The outcomes from Golbraikh and Tropsha acceptable model was diagrammatized at Table: 6 which represent that Q^2: 0.99928, r^2: 0.99979, |r0^2-r'0^2|:0.0, k: 0.97936, [(r^2-r0^2)/r^2]:0.00004, k': 1.02084, [(r^2-'0^2)/r^2]: 0.00004 which also shown that the predictability of the model is quite high. Applicability domain was identified by Euclidean and Mahalanobis Distance Method. All the points were merely overlapped with observed and predicted IC_{50} value. So this developed QSAR model may work as a navigator for the prediction of newer generation steroidal molecule active through corticosteroid binding globulin receptor pathway.

Keywords: Corticosteroid binding globulin receptor, 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: supriyo9@gmail.com; Tel.: +917895424583

RUNNING TITLE: 2D QSAR study on Corticosteroid Binding Globulin Receptor

INTRODUCTION

Corticosteroid Binding Globulin (CBG) may have been intended to give a cradle store of cortisol. Cortisol is, truth be told, traditionally required in times of anxiety, and there are a large group of extraordinary organic components that initiate the hypothalamic-pituitary-adrenocortical hub. The hypothalamic-pituitary-adrenal pivot is a quickly reacting framework, and the organic entity would increase just a little due to a promptly open pool of CBG-bound cortisol. CBG ties progesterone as hard as it does cortisol, and it is hard to imagine the need for having a quickly access pool of this steroid. CBG-bound cortisol is utilized to serve as a store to allow the free cortisol fixation to stay high at locales of aggravation. Corticosteroid tying globulin (CBG) is the significant transport protein for glucocorticoids in plasma of mammalian species, with more than 90% of flowing corticosteroid atoms being bound by this bearer. Corticosteroid Binding Globulin (CBG) is a 55-kDa monomeric glycoprotein that is discharged primarily by the liver, but on the other hand is delivered in the lung, kidney, and testis [1]. CBG imparts little grouping closeness to other steroid bearers, for example, the vitamin D tying protein, sex hormone tying globulin, or α-fetoprotein. Maybe, the protein displays basic homology to individuals from the super group of serine protease inhibitors (Serpins). Like different serpents, CBG can be cut by proteases, affecting conformational changes in tertiary protein structure and repealing the capacity of the transporter to tie steroids. CBG gives a repository of circling

protein-bound steroids that are organically dormant, and it controls the measure of free hormones that are accessible for passage into the target tissues. In this work we extraordinarily accentuate on the descriptor choice and added to a various straight mathematical statement for the correct improvement of steroid to work through the corticosteroid tying globulin receptor [2]. In this work our main objective is to develop a 2D QSAR model and exploring the descriptor combination to predict the biological activity of a steroidal molecule (which is not restricting to a particular origin as plant or animal). Here corticosteroid binding globulin receptor affinity is bench as a biological activity because for a single steroid this receptor activation is the prima focus of their mechanism of work. So here, we extract 30 different steroids with corticosteroid binding globulin receptor affinity profile from database and by the application of different QSAR techniques such as Dataset deviation, stepwise regression, FA-MLR (Multiple Linear Regression), partial least square techniques and validate the model using various internal and external validation tools. Finally the developed model is cross validated and calculates the predictive value and make a comparison between observed and predicted data to identify the gap between them.

MATERIALS AND METHODS

To build a QSAR model, an arrangement of hypothetical and valuable descriptors was ascertained by utilization of PaDEL-descriptor: an open source programming, ToMoCoMD. QSAR Model was developed by utilization of the MLR Plus Validation Tool. By utilizing PADEL and ToMoCoMD we were ascertained 1875 descriptors incorporate Ghose-Cripen Log Ko/w, Ghose-Crippen molar refractivity, Sum of the nuclear polarizabilities (counting implied hydrogens), Wildman- Crippen LogP and MR, Wildman-Crippen MR, Eccentric Connectivity Index: topological descriptor joining separation and nearness data, H Bond Acceptor Count Number of hydrogen security acceptors, McGowan trademark volume, Wiener Polarity Number, Geary autocorrelation - slack 5/ weighted by Sanderson electronegativities, Radial conveyance capacity - 120/ unweighted, all out openness file/ weighted by relative I-state, worldwide shape list/ weighted by relative polarizabilities, complete availability record/ weighted by relative Sanderson electronegativities, D absolute availability list/ weighted by relative van der Waals volumes, aggregate size file/ weighted by relative van der Waals volumes, Radius of gyration, Gravitational file - hydrogens included. All the clarifications of applicable descriptors were enrolled in Table 1. A descriptor speaks to a quantitative property relies on upon the sub-atomic structure. Hypothetical descriptors are worthwhile because of its free from vulnerability of trial estimation and can be ascertained for mixes before combination.

Dataset and Descriptor Calculation

Dataset of 30 corticosteroid binding globulin receptor agonist were downloaded from http://crdd.osdd.net. All the Molecules SMILES are moved into .Mol arrange by ACDLABS and structures were improved. Figure the 2D and 3D descriptor utilizing PADEL and ToMoCoMD product. Table 2 was demonstrated the point of interest dataset alongside substance, structure, IC_{50} quality and pIC_{50} esteem and Table 1 outcomes some helpful descriptor with clarification.

Descriptor Pretreatment

Removed the interconnected descriptor chose utilizing V-WSP as different cuts of 0.0001and connection coefficient esteems 0.99.

Dataset Division

The dataset was partitioned into Train and Test, utilizing Kennard Stone technique as 21 molecules was in Training set and 9 molecules were in Test set.

Suitable Descriptor Selection

Suitable descriptor determination utilizing Stepwise MLR as F values 3.9 to 4.0. At that point best subset was chosen utilizing 4 descriptor mix and r2 cut off worth 0.6.

The chemometric tool

For the advancement of QSAR mathematical statement two techniques were executed; (1) Stepwise regression (2) multiple linear regressions with component examination as preprocessing variable investigation of variable choice (FA-MLR).

Stepwise Regression

A multi step linear equation, a multistep mathematical statement was fabricated. The fundamental technique included: (i) distinguishing a starting model (ii) rehashing the past venture by adjusting descriptor or a variable mix to attain to better F and r2 esteem. (iii) Calibrate the comparison by legitimizing the qualities in the middle of watching and anticipated qualities. The stepwise MLR was performed utilizing factual programming SPSS and it was judged by parameters as clarified change (r2a), connection coefficient (r), standard slip of assessment (s) and difference proportion (F) at a predefined level of opportunity (DF). All acknowledged MLR comparison had relapse level critical at 95 and 99% levels. The created QSAR comparison was approved by forgetting one or LOO system utilizing Minitab programming and distinctive parameters like cross acceptance r2 (q2), standard deviation taking into account press (SPRESS) and standard deviation of mistake of expecting (SDEP) [3].

FA-MLR

In this situation a last factual apparatus was utilized to build up a QSAR relation, factor examination as an information preprocessing venture to recognize the critical variable to distinguish the essential variables contributing the reaction varies by maintaining a strategic distance from Co straight esteem. The information lattice is initially institutionalized and connection framework and therefore decreased relationship grid. An eigenvalue issue is then tackled and the manufacturing plant example can be acquired from the relating eigenvectors. The principle destinations are to show multidimensional information in space of lower dimensionality with less loss of data (clarifying > 95% of the fluctuation of the information grid) and to concentrate the fundamental highlights behind the information with a definitive objective of translation [4].

QSAR Equation Development

MLR Plus substantial programming was utilized as a part of the advancement of QSAR mathematical statement, where IC_{50} was changed over pIC_{50} esteem [5, 6].

QSAR Equation Validation

Golbraikh and Tropsha acceptable model criteria's 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) [7, 8].

QSAR Equation Cross Validation

The model was cross validated using Leave-One-Out (LOO) process. 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 and Mahalanobis distance method. The distance of a test compound to its nearest neighbor in the training set is compared to the predefined applicability domain threshold [9, 10].

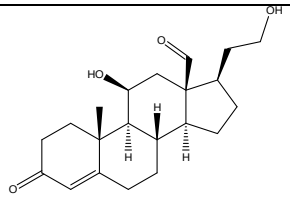
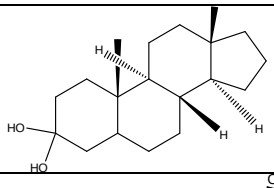
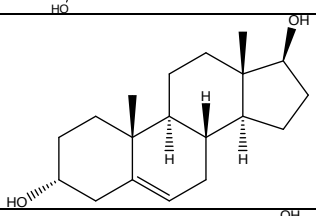
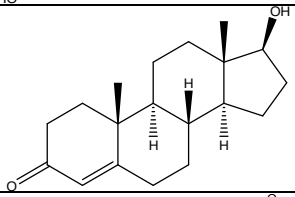
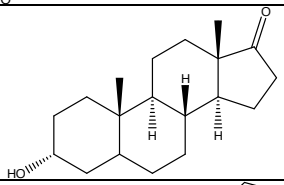
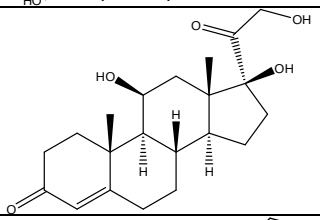
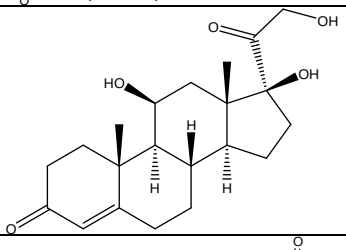
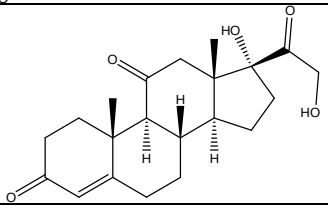
RESULTS AND DISCUSSION

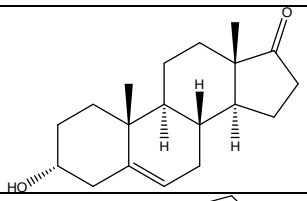
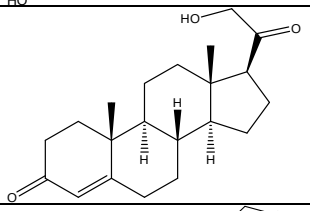
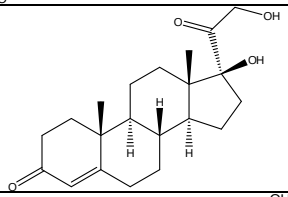
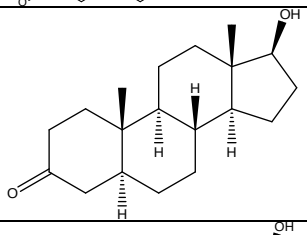
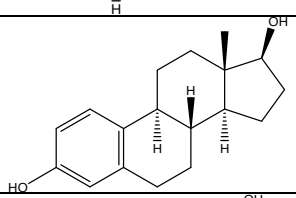
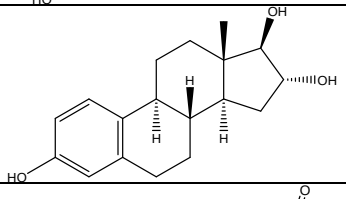
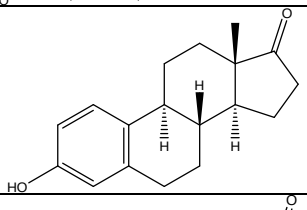
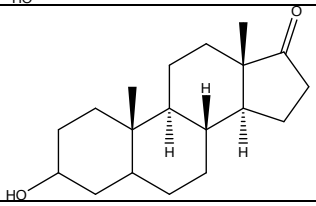
The statistical model for this development is: pIC_{50} = -0.01499 (+/-0.00582) +0.02256 (+/-0.00794) GATS5e -0.02574 (+/-0.00344) RDF120v -0.61334 (+/-0.00342) Ds with statistical information: SEE: 0.00581, r^2: 0.99977r^2 adjusted: 0.99973, F: 24389.67031 (DF: 3, 17). This model proposes that by expanding the Sanderson electro negativities and by diminishing the Radial dispersion capacity - 120/ weighted by relative van der Waals volumes and aggregate availability record/ weighted by relative I-state esteem it make a positive reaction.

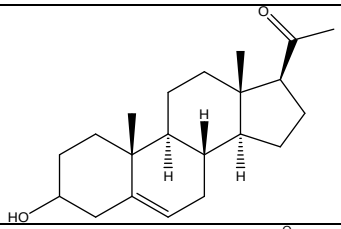
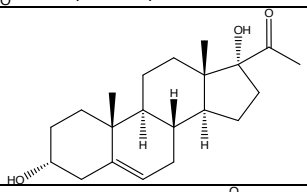
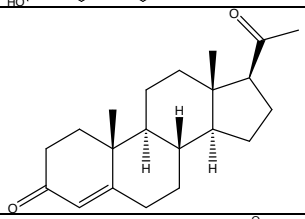
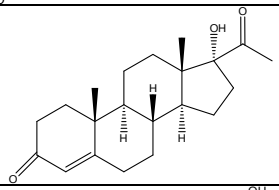
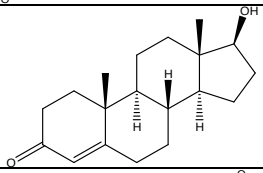
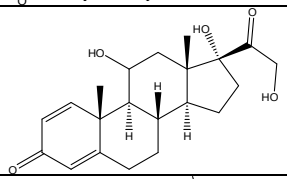
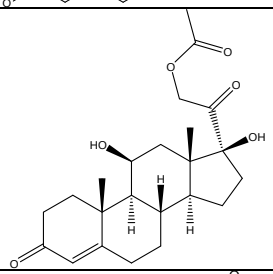
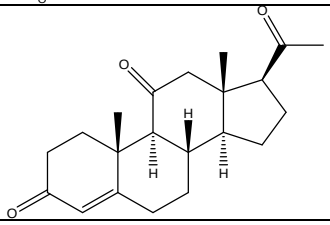
Table 1: Total Descriptor with explanation

Abbreviation descriptors	Explanation of descriptors
AlogP	Ghose-Crippen LogKow
SpMin3_Bhv	Burden modified eigenvalues
ntN	Total number of Nitrogen Atoms
ETA_Beta_ns	A measure of electron-richness of the molecule
Crippen MR	Crippen's molar refractivity
McGowan Volume	McGowan characteristic volume
VABC	Van der Waals volume calculated
nRing	No of Ring
nRotb	No of Rotatable Bonds
Phia Kappa	flexibility Index
Bac	Balaban Centric Index
AlogP	Ghose-Crippen LogKow
Crippen LogP	Crippen's LogP
XLogP	XLogP
AMR	Molar refractivity
TopoPSA	Topological Polar Surface Area
Wpol	Weiner polarity number
MW	Molecular Weight
ETA	Electro Topochemical Descriptor

Table 2: Total Dataset with pIC50 value.

S.No.	CHEMICAL NAME	CHEMICAL STRUCTURE	pIC_{50} VALUES
1.	Aldosterone		-0.883
2.	Androstanediol		-0.896
3.	5-androstenediol		-0.852
4.	4-androstenedione		-0.796
5.	Androsterone		-0.698
6.	Corticosterone		-0.888
7.	Cortisol		-0.788
8.	Cortisone		-0.886

9.	Dehydroepiandrosterone		-0.761
10.	11-deoxycorticosterone		-0.831
11.	11-deoxycortisol		-0.698
12.	Dihydrotestosterone		-0.698
13.	Estradiol		-0.797
14.	Estriol		-0.698
15.	Estrone		-0.797
16.	Etiocholanolone		-0.896

17.	Pregnenolone		-0.896
18.	17a-hydroxypregnenolone		-0.878
19.	Progesterone		-0.838
20.	17a-hydroxyprogesterone		-0.698
21.	Testosterone		-0.772
22.	Prednisolone		-0.857
23.	Cortisolacetat		-0.698
24.	4-pregnene-3,11,20-trione		-0.698

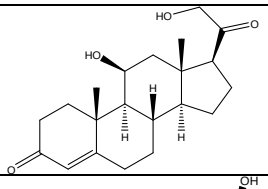
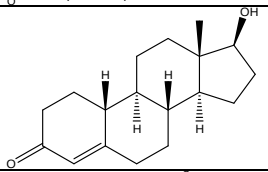
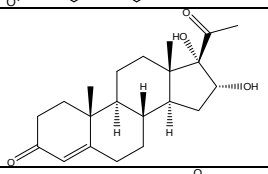
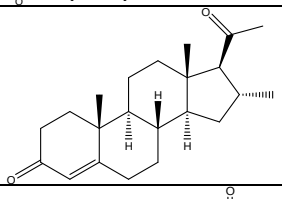
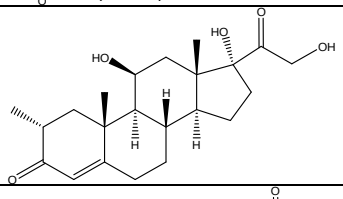
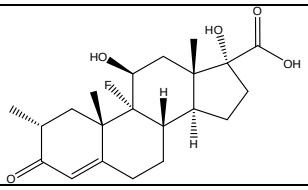
25.	Epicorticosterone		-0.698
26.	19-nortestostrerone		-0.718
27.	16a,17a-dihydroxyprogesterone		-0.875
28.	16a-methylprogesterone		-0.718
29.	2a-methylcortisol		-0.797
30	2a-methyl-19a-fluoro-cortisol		-0.908

Table 3: Leave-One-Out (LOO) Result of the developed Model

Q2	PRESS	SDEP	Parameter	Without scaling	After scaling
0.99928	0.00177	0.00919	rm^2 (Loo)	0.9991	0.99698
			rm^2'(Loo)	0.99916	0.99763
			Average rm^2 (LOO)	0.99913	0.9973
			Delta rm^2 (LOO)	0.00006	0.0065

Table 4: External Validation Parameters (Without Scaling) of the developed Model

r^2	r0^2	Reverse r0^2	rm^2 (test)	Reverse rm^2 (test)	Average rm^2 (test)	Delta rm^2 (test)	Rmsep	rpred^2	Q2f1	Q2f2
0.99979	0.99975	0.99975	0.9935	0.99358	0.99354	0.00008	0.00757	0.99929	0.99929	0.99925

Table 5: Some External Validation Parameters (After Scaling) of the developed Model

rm^2(test)	reverse rm^2(test)	average rm^2(test)	delta rm^2(test)
0.97753	0.97818	0.97785	0.00065

Table 6: Golbraikh and Tropsha acceptable model criteria's of the developed Model

S.N	Parameter	Value	Pass or Not Pass	Threshold Value
1.	Q^2	0.99928	Passed	(Threshold value Q^2>0.5)
2.	r^2	0.99979	Passed	(Threshold value r^2>0.6)
3.	|r0^2-r'0^2|	0	Passed	(Threshold value |r0^2-r'0^2|<0.3)
4.	k OR [(r^2-r0^2)/r^2]	0.97936 OR 0.00004	Passed	(Threshold value: [0.85<k<1.15 and ((r^2-r0^2)/r^2)<0.1]
5.	k' OR [(r^2-r'0^2)/r^2]	1.02084 OR 0.00004	Passed	[0.85<k'<1.15 and ((r^2-r'0^2)/r^2)<0.1])

Table 7: Leave One Out for the Training Set Molecules of the developed Model

Alloted CompNo	Yobs	Ypred	(Residual)^2	(Yobs-Ybar)^2
1	0	0.00275	0.00001	0.02764
2	0	0.00225	0.00001	0.02764
3	0	0.00364	0.00001	0.02764
4	0	-0.00187	0	0.02764
5	0	0.00167	0	0.02764
6	0	-0.00032	0	0.02764
7	0	-0.00705	0.00005	0.02764
8	0	-0.00207	0	0.02764
9	-0.831	-0.86322	0.00104	0.44191
10	0	-0.00336	0.00001	0.02764
11	0	-0.00336	0.00001	0.02764
12	0	-0.00266	0.00001	0.02764
13	0	-0.00222	0	0.02764
14	-0.896	-0.88791	0.00007	0.53255
15	-0.896	-0.89807	0	0.53255
16	0	0.00954	0.00009	0.02764
17	0	-0.00297	0.00001	0.02764
18	0	-0.00106	0	0.02764
19	0	-0.00035	0	0.02764
20	0	0.00825	0.00007	0.02764
21	-0.868	-0.84863	0.00038	0.49247

Table 8: Applicability Domain identified by Euclidean Distance Process for Training Set

Alloted Compd No.	Distance Score	Mean Distance	Normalized Mean Distance
1	10.12467	0.48213	0.00966
2	10.03733	0.47797	0.00737
3	11.29497	0.53786	0.04036
4	10.0688	0.47947	0.00819
5	9.90308	0.47158	0.00385
6	9.75645	0.46459	0
7	15.30833	0.72897	0.14566
8	10.04866	0.47851	0.00767
9	28.98973	1.38046	0.5046
10	10.56881	0.50328	0.02131
11	10.56881	0.50328	0.02131
12	10.24724	0.48796	0.01288
13	10.09095	0.48052	0.00878
14	31.52599	1.50124	0.57114
15	47.87222	2.27963	1
16	12.45309	0.593	0.07075
17	10.37849	0.49421	0.01632
18	9.84207	0.46867	0.00225
19	9.7574	0.46464	0.00002
20	11.99351	0.57112	0.05869
21	27.47289	1.30823	0.46481

Table 9: Applicability Domain identified by Euclidean Distance Process for Test Set

Alloted Compd No.	Distance Score	Mean Distance	Normalized Mean Distance
1	11.29497	0.53786	0.04036
2	13.03805	0.62086	0.0861
3	47.87222	2.27963	1
4	10.21133	0.48625	0.01193
5	13.58967	0.64713	0.10057
6	10.43574	0.49694	0.01782
7	10.09095	0.48052	0.00878
8	11.201	0.53338	0.0379
9	10.04866	0.47851	0.00767

Table 10: Applicability Domain identified by Mahalanobis Distance Process for Training Set

GATS5e	RDF120v	Ds	pIC50	Outlier Info.
0.777812	5.00E-47	0	0	-
0.770761	0.011813	0	0	-
0.906976	0.090141	0	0	-
0.704962	0.104035	0	0	-
0.733933	1.34E-04	0	0	-
0.651121	4.47E-93	0	0	-
0.96654	0.468993	0	0	-
0.580869	8.14E-87	0	0	-

1.262413	0.201172	1.394017	-0.831	-
0.532933	3.13E-14	0	0	-
0.532933	5.59E-17	0	0	-
0.558456	4.64E-70	0	0	-
0.5751	1.11E-60	0	0	-
0.959906	0.892802	1.424518	-0.896	-
0.873019	2.121027	1.380001	-0.896	Outlier
0.998662	0.001727	0	0	-
0.547006	2.42E-71	0	0	-
0.620739	2.73E-35	0	0	-
0.649869	1.16E-121	0	0	-
0.964316	2.03E-06	0	0	-
0.835583	0.079765	1.402142	-0.868	-

Table 11: Applicability Domain identified by Mahalanobis Distance Process for Test Set

GATS5e	RDF120v	Ds	pIC50	AD Info.
0.906976	0.090141	0	0	-
1.039683	0.031504	0	0	-
0.873019	2.121027	1.380001	-0.878	Outside AD
0.562266	9.51E-89	0	0	-
1.069355	2.98E-40	0	0	-
0.812699	6.49E-23	0	0	-
0.5751	7.79E-60	0	0	-
0.791573	0.1989	0	0	-
0.580869	9.63E-83	0	0	-

Table 12: Comparison between pIC50 $_{observed}$ and pIC50 $_{predicted}$ as per developed QSAR model

pIC50 $_{observed}$	pIC50 $_{predicted}$
-0.883	-0.909
-0.896	-0.883
-0.852	-0.887
-0.796	-0.887
-0.698	-0.925
-0.888	-0.914
-0.788	-0.878
-0.886	-0.876
-0.761	-0.903
-0.831	-0.846
-0.698	-0.873
-0.698	-0.873
-0.797	-0.86
-0.698	-0.895
-0.797	-0.912
-0.896	-0.89
-0.896	-0.896

-0.878	-0.896
-0.838	-0.894
-0.698	-0.887
-0.772	-0.931
-0.857	-0.903
-0.698	-0.817
-0.698	-0.845
-0.698	-0.816
-0.718	-0.903
-0.875	-0.86
-0.718	-0.833
-0.868	-0.858
-0.797	-0.908

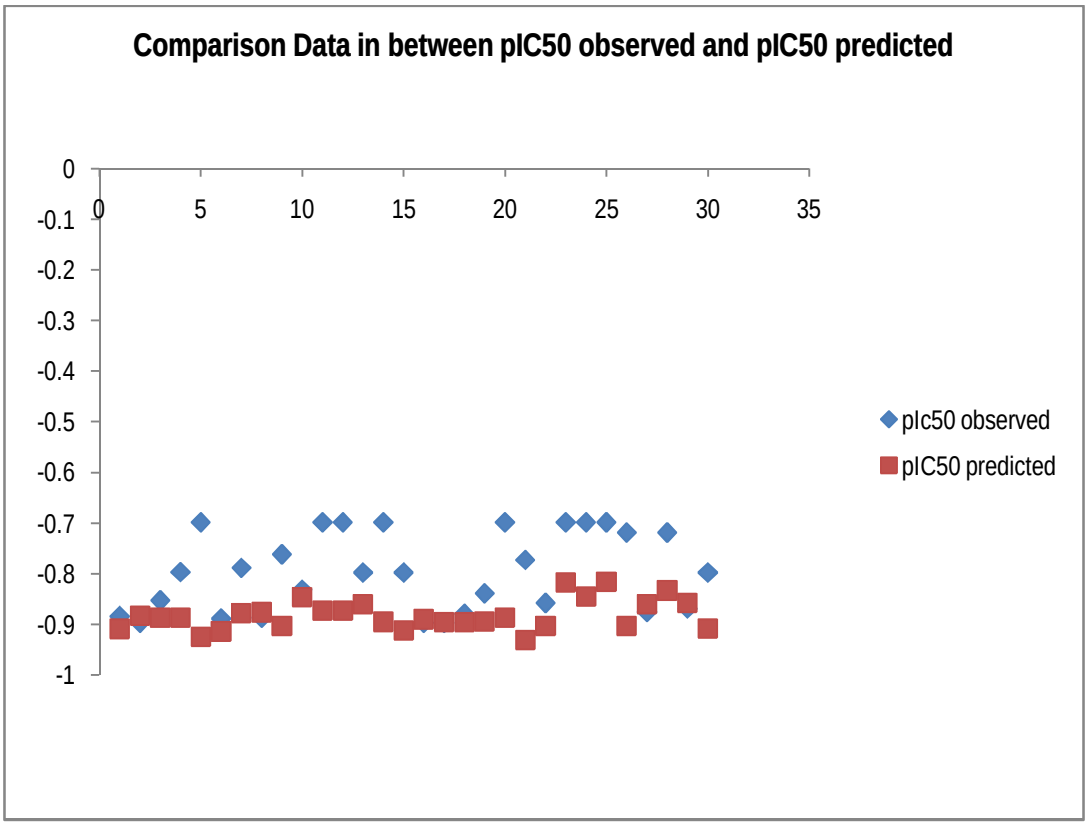


Fig.1: Comparison Data in between $pIC_{50\ observed}$ and $pIC_{50\ predicted}$.

The leave one out and external validation parameter with and without scaling data was reported at Table: 3-5; which represent that the difference between the *r*2 and *r*2adj value is 0.00004 which is less than 0.3 indicates that the number of descriptors involved in the QSAR model is acceptable for prediction the activity profile, F test ratio value 24389.67031 also reflects the greater predictability of the QSAR model. In the LOO result the difference between r2 and q2 (without scaling) is 0.00018 with accessible limit 0.3 which reflect the good predictability of the QSAR model. The rm^2 (test) value for external validation is 0.9935

which must be greater than 0.5 as a good external predictability parameter. For a QSAR model can be considered acceptable if the values of r2m (overall) and r2p are equal to or above 0.5 (or at least near 0.5); in this developed QSAR model r2m is 0.99979 and r2p is 0.99929 [11, 12]. The outcomes from Golbraikh and Tropsha acceptable model was diagrammatized at Table: 6 which represent that Q^2: 0.99928, r^2: 0.99979, |r0^2-r'0^2|:0.0, k: 0.97936, [(r^2-r0^2)/r^2]:0.00004, k': 1.02084, [(r^2-'0^2)/r^2]: 0.00004 which also shown that the predictability of the model is quite high [13]. Applicability domain analyzes data by Euclidean and Mahalanobis distance method was reported at Table: 7-10; which showed that not a single molecule is outside the domain. The comparision data in between observed and predicted biological activity was reported at Table: 11 and graphical representation at Figure 1, so as per the overlapping data, it is quite clear that all the predicted data are in the area of observed data points and also some data are purely overlapped. Though the model comes with very less number of descriptor combinations, there is less chance of error [14, 15].

CONCLUSION

Finally, we conclude that in future when a new steroidal molecule will develop to work through the corticosteroid binding receptor; one only optimize the descriptor combinations to predict the activity profile with ease. This QSAR model will work as navigation and it will a boon for future development.

CONFLICT OF INTEREST

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

ACKNOWLEDGEMENTS

One of the authors is highly thankful to Dr. Veerma Ram for his Unconditional support and well beings.

REFERENCES

1. Zhou A, Wei Z, Stanley PL, Read RJ, Stein PE, Carrell RW. The S-to-R transition of corticosteroid-binding globulin and the mechanism of hormone release. J Mol Biol. 2008; 380 (1): 244–51.
2. Torpy DJ, Lundgren BA, Ho JT, Lewis JG, Scott HS, Mericq V. CBG Santiago: a novel CBG mutation. J Clin Endocrinol Metab. 2012; 97 (1): 151–5.
3. 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.
4. 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.
5. 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.
6. Garg R, Carr JS. Predicting the bioconcentration factor of highly hydrophobic organic chemicals. Food Chem Toxicol. 2014; 69: 252–259.
7. Dearden JC, Hewitt M. QSAR modeling of bioconcentration factor using hydrophobicity, hydrogen bonding and topological descriptor. SAR QSAR Environ Res. 2010; 21: 671–680.
8. Dimitrov SD. Predicting bioconcentration factors of highly hydrophobic chemicals. Effects of molecular size. Pure Appl Chem. 2002; 74: 1823–1830.

9. Krstajic D. Cross-validation pitfalls when selecting and assessing regression and classification models. J Cheminform. 2014; 6(10): 1-15.
10. Zhang S. Jaworska J, Nikolova-Jeliazkova N, Aldenberg T. QSAR Applicability Domain Estimation. J Chem Inf Model. 2006; 46:1984–1995.
11. Ravichandran V. Validation of QSAR Models - Strategies and Importance. Int J Drug Design Discovery. 2011; 3: 511-519.
12. Kim JH. Is it Possible to Predict the ADI of Pesticides using the QSAR Approach?. J Environ Health Sci. 2012; 38(6): 550-560.
13. 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. Res J Pharm Bio Chem Sci. 2014; 5(3): 1884-93.
14. 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. Res in Biotech. 2014; 5(6): 15-23.
15. 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 Chem. 2015; 5(4): 134-148.