Est. 1984

ORIENTAL JOURNAL OF CHEMISTRY
An International Open Free Access, Peer Reviewed Research Journal

www.orientjchem.org

ISSN: 0970-020 X
CODEN: OJCHEG
2017, Vol. 33, No.(6):
Pg. 2949-2958

QSAR Studies of Lanosterol Synthase Inhibitors as Potential Antihyper Cholesterolemic Compounds

DIVYA G NAIR[1], V.L.PUSHPA[2]*, K. KURIEN THOMAS[1] and K. B. MANOJ[2]

[1]Department of Chemistry, Bishop Moore College, Mavelikara, Kerala-290110, India.
[2]Department of Chemistry, S N College, Kollam, Kerala, India.
*Corresponding author E-mail: drpushpavl@gmail.com

http://dx.doi.org/10.13005/ojc/330631

(Received: August 22, 2017; Accepted: September 20, 2017)

ABSTRACT

The QSAR analysis provides significant structural insight to help the design of novel antihyper cholesterolemic compounds. 2D-QSAR model for a set of 23 LSS inhibitors that have antihyper cholesterolemic activity was derived. The lanosterol synthase inhibitory activity data and various physicochemical descriptors were preferred as dependent and independent variables respectively. The Multiple Linear Regression (MLR) is taken to choose the best model. Here we have derived two best QSAR models to discuss the structural properties significant for the OSC inhibitory activity. In these models, the following parameters 3D- MoRSE parameter (Mor07e), WHIM parameter (G3e), GETAWAY parameter ($R6p_+$), Radial Distribution Function parameters (RDF110u and RDF110e) have provided for the LSS inhibitory prediction. This structure activity relationship obtained for the molecule is comparable with the QSAR results and this agreed with the results developed from QSAR studies.

Keywords: QSAR, Lanosterol synthase inhibitors, Physicochemical descriptors, MLR.

INTRODUCTION

The World Health Organization (WHO) reports that high cholesterol contributes to about 56% of cases of coronary heart disease (CHD) worldwide and causes more than 4 million deaths each year. High cholesterol (hypercholesterolemia) can cause the accumulation of plaque deposits in the arteries it results in atherosclerosis or CHD[1]. Atherosclerosis can lead to plaque ruptures and blockages in the arteries, which increase the risk for heart attacks, stroke, circulation problems and death[2]. Lanosterol synthase (LSS) is (EC 5.4.99.7) is an important enzyme in cholesterol biosynthesis as it functioned as catalyst for the conversion of 2, 3 – oxidosqualene to lanosterol, the first precursor for sterols. This enzyme locates the downstream from essential branching steps in the pathway of cholesterol synthesis and not influence the formation of intermediates needed for other biosynthetic pathways such as synthesis of isoprenoids, coenzyme Q_{10} etc[3-9].

The study of the quantitative structure activity /property relationships (QSAR/QSPR) of compounds is a relevant approach of modern chemistry because of the details obtained is consists of mathematical equations correlating the chemical

structure of the compounds to a large variety of their physical, chemical and biological properties. Once a relation between structure and activity/property is noticed, any number of compounds including those not synthesized yet, can quickly be screened *in silico* for searching of structures with derived properties[10-21]. Hence it is possible for the selection of most favorable structures for synthesis and testing in the laboratory. In this study, twenty three molecules having LSS inhibitory activity were taken from reference[22].

MATERIALS AND METHODS

QSAR study was performed on the data set consists of twenty three compounds having LSS inhibitory activity. The biological activity data IC_{50} (nm) were changed to pIC_{50} for QSAR analysis. The pIC_{50} values of the molecules under study covered a large range from 5 to 8. The chemical structure of all the compounds is shown in table 1, 2 and 3. The structure of the compounds was drawn and 3D structures were generated using Marvin 5.4.1.1 version (Chem Axon Ltd., 2010[23]). Most stable structures of each compound was generated after energy minimization and used for calculating various physicochemical descriptors. The low energy conformers were then allowed for further generation of an extra set of 3,224 molecular descriptors using DRAGON version 5.5[24]. The multi-collinear and redundant descriptors collected from the DRAGON software package were limited from 3,224 to 1497 descriptors by manually. The descriptors were then submitted for stepwise multiple linear regression (MLR) in order to choose suitable descriptors. Regression analysis has been made by the software Build QSAR[25]. The obtaining subset of molecular descriptors is consists of the 3D-MoRSE descriptors such as Mor07e (Signal 07/ weighted by Sanderson electronegativity), Mor10e (Signal 10/weighted by Sanderson electronegativity), Mor20p (Signal 20/weighted by polarizability), WHIM descriptor such as G3e (3rd component symmetry directional whim index/ weighted by Sanderson electronegativity), getaway descriptor such as $R6p_+$ (R- maximal auto-correlation of lag 6/ weighted by polarisability),Connectivity Indices such as X5A (average connectivity index of order 5), Radial Distribution Function descriptors such as RDF 110u (Radial Distribution Function-110/ unweighted) and RDF 110e (Radial Distribution Function-110/ weighted by Sanderson electronegativity)[26-32]. The biological activity and the selected physicochemical parameters of the twenty three LSS inhibitors were shown in Table. 4. Since we have used a cross-validation method there is no requirement to work on training set and test set independently.

RESULTS AND DISCUSSION

The relevance of QSAR analysis mainly relates either or not the molecular descriptors selected are suitable to correlate the biological activity. Regression analysis was carried out by maximum R^2 method[33] choosing stepwise regression. Univariate, bivariate to multivariate regression has been made for understanding the best correlation. The eight selected molecular descriptors were then taken in the formation of QSAR model by MLR. The correlation of the used descriptors and their correlation with the activity are shown in table 5. The results (Table. 5) indicate that the two physicochemical descriptors RDF110u and RDF110e are mutually correlated. When these two are exist together in the regression equation then the model may undergo from the defect due to collinearity[34, 35]. The outcomes of the statistically relevant models are shown in table 6. There is no statistically important mono parametric model is preferred for modeling the LSS inhibitory activity (pIC_{50}). In attaining the statistically significant models, the corresponding 5, 13 and 15 are found to be outliers. Therefore they were removed from the successive regression analysis.

The results shown in Table 6 indicate that there are six bi-parametric regression models out of which the model 5 consists of Mor07e and G3e provide the maximum R^2 result. This model is found as shown below.

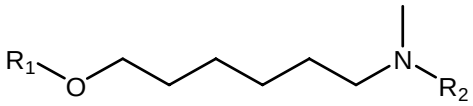
pIC_{50} = –0.7825 (± 0.1251) Mor07e + 24.0892 (± 7.3506) G3e+ 5.7841

n = 20, Se = 0.4012, R = 0.8597, F = 24.0672, Q = 2.1426 ...(1)

From the two better quality tri parametric models, the model 11 consists of Mor07e, G3e and R6p+ provides better results. Mor07e and R6p+ parameters indicate negative correlation and G3e parameter indicates positive correlation with the activity.

Table. 1: The chemical structure of compounds containing the parent structure A

A =



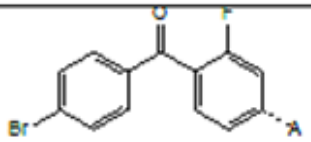
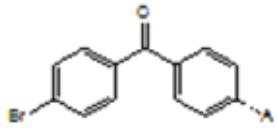
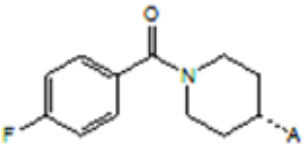
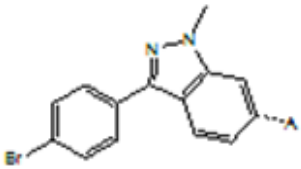
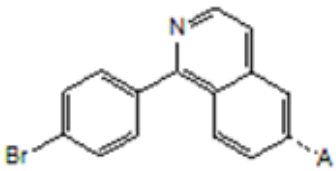
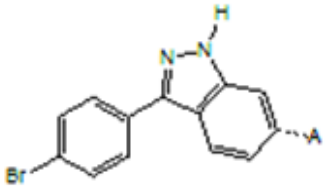
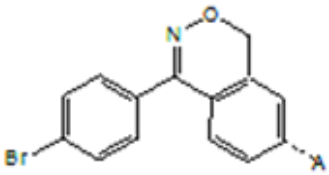
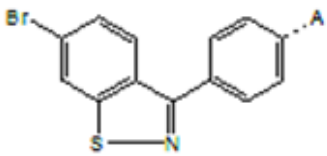
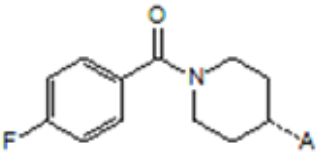
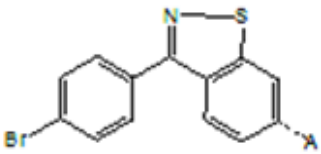
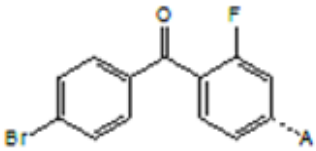
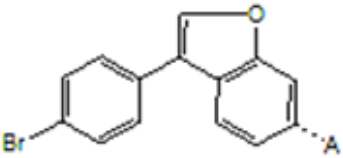
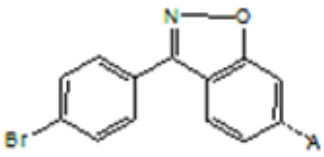
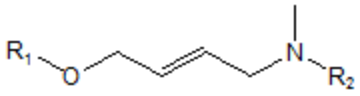
No.	R_1	R_2	No.	R_1	R_2
1			6		
8			9		
11			12		
14			15		
17			18		
20			21		
22					

Table. 2: The chemical structure of compounds containing the parent structure M

M =



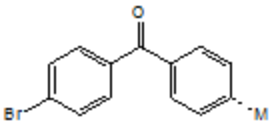
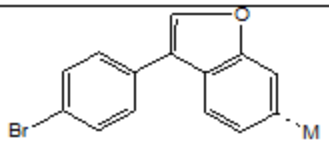
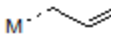
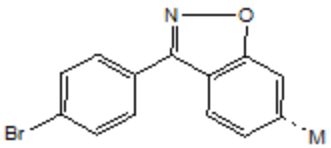
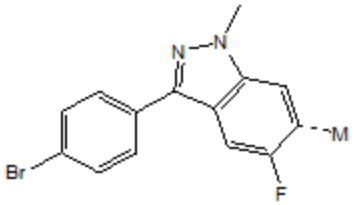
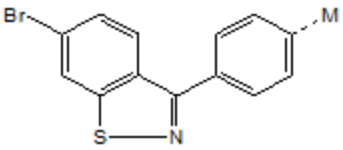
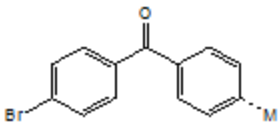
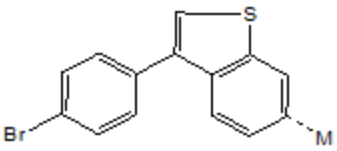
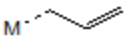
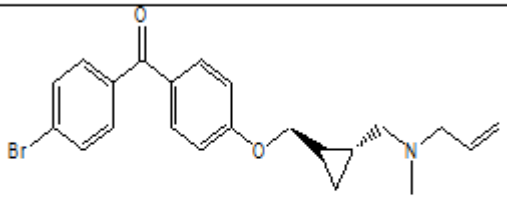
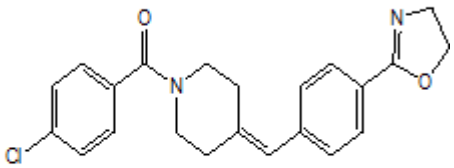
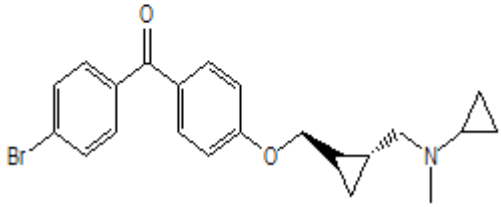
No.	R_1	R_2	No.	R_1	R_2
2			3		
7			10		
13			16		
19					

Table. 3: The chemical structure of compounds 4, 5 and 23

No.	Structure	No.	Structure
4		5	
23			

Tale. 4: The biological activity and the selected physicochemical parameters of the twenty three LSS inhibitors

S.N	1C50 [nm]	pIC50 [nm]	Mor07e	G3e	R6p+	Mor24p	RDF110u	RDF110e	X5A	Mor10e
1	0.098	7.0087	3.392	0.156	0.016	0.206	13.207	13.064	0.082	0.527
2	0.019	7.7212	2.453	0.163	0.016	0.293	13.002	13.237	0.083	0.916
3	0.5	6.301	3.842	0.181	0.024	0.128	12.367	11.947	0.076	1.324
4	0.223	6.6517	3.101	0.157	0.017	0.186	17.386	17.149	0.077	0.893
5	0.036	7.4437	4.258	0.172	0.009	0.03	10.197	9.886	0.079	-0.079
6	0.0054	8.2676	2.221	0.164	0.015	0.034	12.091	12.121	0.087	0.853
7	0.029	7.5376	3.963	0.183	0.019	0.105	6.382	6.607	0.076	0.943
8	0.071	7.1487	3.412	0.153	0.009	-0.195	14.86	14.8	0.087	0.555
9	0.0196	7.7077	3.481	0.183	0.013	0.251	13.021	12.83	0.075	0.63
10	0.029	7.5376	3.819	0.191	0.012	0.142	14.849	14.95	0.072	0.609
11	0.0035	8.4559	2.578	0.185	0.014	0.033	15.832	15.465	0.078	0.689
12	0.039	7.4089	3.102	0.171	0.015	0.052	9.429	9.316	0.077	0.974
13	1.86	5.7305	3.134	0.199	0.012	0.388	7.096	7.378	0.077	1.072
14	0.0041	8.3872	1.844	0.173	0.014	0.231	11.651	11.842	0.078	0.454
15	1.9	5.7212	2.462	0.168	0.011	0.219	8.654	8.677	0.078	0.826
16	0.003	8.5229	2.054	0.168	0.017	0.169	11.548	12.337	0.085	0.323
17	0.048	7.3188	4.291	0.174	0.009	-0.059	11.501	11.188	0.085	0.512
18	0.0029	8.5376	2.395	0.203	0.024	0.378	12.341	12.366	0.077	0.734
19	0.0135	7.8697	2.483	0.172	0.025	0.294	9.959	9.635	0.076	0.788
20	0.0065	8.1871	2.498	0.181	0.014	0.069	13.528	14.194	0.083	1.019
21	0.61	6.2147	3.692	0.163	0.025	0.277	11.267	11.035	0.077	1.147
22	0.38	6.4202	4.078	0.176	0.017	0.118	15.59	15.125	0.077	0.999
23	0.022	7.6576	3.348	0.179	0.013	0.181	14.864	14.618	0.077	0.788

Table 5. Correlation matrix demonstrating correlation of the physicochemical parameters and their correlation with the activity (pIC50)

	pIC50	Mor07e	Mor10e	G3e	R6p+	X5A	RDF110u	RDF110u	$Mor24_P$
pIC50	1								
Mor07e	-0.7577	1							
Mor10e	-0.5351	0.2801	1						
G3e	03722	0.0442	0.0684	1					
R6p+	-0.2039	-0.1045	0.5425	0.1689	1				
X5A	0.2283	-0.2927	-0.306	-0.5516	-0.3911	1			
RDF110u	-0.154	0.0156	-0.105	-0.1096	-0.3492	0.02215	1		
RDF110u	-0.0799	-0.0562	-0.1476	-0.1095	-0.3777	0.07384	0.9911	1	
Mor24P	0.1017	-0.3004	0.1051	02717	0.6481	-04992	-0.1216	-0.1179	1

Table. 6: Regression parameters and quality of the proposed models

Model No	Parameters used	Ai i = 1,2,3,4,5,6	B	Se	R	F- Ratio	Q = R/Se
1	Mor07e	-0.5643 (±0.2238)	9.1449	0.7765	0.4820	6.3569	0.6207
2	Mor10e	-1.5798 (±0.5878)	8.7767	0.6447	0.5351	7.2234	0.8298
3	G3e	9.2697(± 14.5420)	5.7626	0.8778	0.1378	0.4063	0.1570
4	R6p+	-3.4241 (± 38.3007)	7.3272	0.8861	0.0004	0.0080	0.0004
5	Mor07e	-0.7825 (± 0.1251)	5.7841	0.4012	0.8597	24.0672	2.1426
	G3e	24.0892 (±7.3506)					
6	Mor07e	0.6654 (± 0.1425)	10.4151	0.4392	0.8291	18.6789	1.8876
	Mor10e	-1.0344 (± 0.4171)					
7	Mor07e	-0.7946 (± 0.1445)	10.7882	0.4612	0.8094	16.1468	1.7550
	R6p+	-43.9412 (±21.9886)					
8	Mor10e	-1.7759 (± 0.7147)	8.6215	0.6585	0.5449	3.5903	0.8275
	R6p+	18.7997 (± 37.1678)					
9	G3e	24.8066 (±12.9476)	3.9231	0.6973	0.4602	2.2838	0.6601
	R6p+	-42.1597 (±33.5423)					
10	G3e	24.3404 (±10.6436)	4.6110	0.5802	0.6740	7.0753	1.1617
	Mor10e	-1.6627 (± 0.5302)					
11	Mor07e	-0.8242 (± 9.1836x10^{-2})	6.1817	0.2926	0.9324	35.4984	3.1869
	G3e	27.8881 (± 5.4437)					
	R6p+	-56.6141 (±14.1663)					
12	Mor07e	-0.6771 (±9.5088x10^{-2})	6.1071	0.2929	0.9323	35.4101	3.1830
	G3e	25.3406 (± 5.3750)					
	Mor10e	-1.1111 (± 0.2756)					
13	Mor07e	-0.7020 (± 0.1523)	10.6705	0.4448	0.8355	12.3295	1.8782
	Mor10e	-0.7955 (± 0.5275)					
	R6p+	-20.0236 (±26.4826)					
14	G3e	23.9724 (±11.0940)	4.6146	0.5972	0.6751	4.4666	1.1305
	Mor10e	-1.7366 (± 0.6484)					
	R6p+	7.2019 (± 34.1324)					
15	Mor07e	-0.6216 (±8.4701x10-2)	0.8353	0.2522	0.9534	37.4541	3.7800
	G3e	33.5256 (± 5.6228)					
	Mor10e	-0.9482 (± 0.2482)					
16	Mor07e	-0.8407 (± 9.9617)	6.0630	0.2997	0.9336	25.4409	3.1153
	G3e	28.6037 (±11.0940)					
	R6p+	-50.6133 (±18.8095)					
	Mor24p	-0.3665 (± 0.7311)					
17	Mor07e	-0.6698 (±9.0916x10-2)	0.6796	0.2797	0.9424	29.7604	3.3694
	RDF110u	-4.1262 x10-2(± 2.5859 x10-2)					
	G3e	24.4791 (±5.1611)					
	Mor10e	-1.1574 (± 0.2677)					
18	Mor07e	-0.6789 (±9.1827x10-2)	6.7894	0.2828	0.9410	29.0219	3.3274
	RDF110e	-3.9552x10-2 (± 2.6915x10-2)					
	G3e	24.5715 (±5.2166)					
	Mor10e	-1.1639 (± 0.2714)					
19	Mor07e	-0.7609 (±9.4320x10-2)	2.0285	0.2764	0.9438	30.5625	3.4145
	G3e	33.7003 (±6.1635)					
	R6p+	-45.8884 (±14.7787)					
	X5A	-34.9620 (±20.4347)					

20	Mor07e	-0.7589 (±8.8346x10-2)	5.8140	0.2533	0.9531	37.1888	3.7630
	G3e	28.9551 (±5.2166)					
	Mor10e	-0.9876 (±0.2458)					
	Mor24p	-1.2302 (± 0.4863)					
21	Mor07e	-0.8482 (±7.3092x10-2)	7.5748	0.2317	0.9608	45.0928	4.1476
	G3e	27.2749 (±4.3146)					
	R6p+	-71.3508 (±121026)					
	RDF110e	-7.6470 x10-2 (±2.3578x10-2)					
22	Mor07e	-0.8288 (±7.2409x10-2)	7.4821	0.2306	0.9612	45.5277	4.1675
	RDF110u	-7.3823 x10-2(± 2.2520x10-2)					
	G3e	27.1340 (±4.2976)					
	R6p+	-69.7369 (± 11.8635x10-2)					

Table. 7: Various correlation models and their qualities of correlations

No.	Regression expression
1. (5)	pIC50 = -0.7825 (± 0.1251) Mor07e + 24.0892 (±7.3506) G3e + 5.7841
2. (11)	pIC50 = -0.8242 (± 9.1836x10^{-2}) Mor07e + 27.8881 (± 5.4437) G3e -56.6141 (±14.1663) R6p_+ + 6.1817
3. (21)	pIC50 = -0.8482 (±7.3092x10^{-2}) Mor07e + 27.2749 (±4.3146) G3e -71.3508 (±121026)R6p_+ -71.3508 (±121026) R6p_+ - 7.6470 x10^{-2} (±2.3578x10^{-2}) RDF110e + 7.5748
4. (22)	pIC50 = -0.8288 (±7.2409x10^{-2}) Mor07e + 27.1340(±4.2976) G3e -69.7369 (±11.8635x10^{-2}) R6p_+ -7.3823 x10^{-2}(± 2.2520x10^{-2}) RDF110u +7.4821

Table. 8: Observed and predicted pIC_{50} values using model-22

Compd No	Observed pIC50	Predicted pIC50	Residuals
1	7.0087	6.8131	0.1957
2	7.7212	7.7963	-0.0751
3	6.3010	6.6226	-0.3215
4	6.6517	6.7031	-0.0514
5	8.2676	8.1527	0.1149
6	7.5376	7.3671	0.1705
7	7.1487	7.0812	0.0675
8	7.7077	7.6948	0.0129
9	7.5376	7.5666	-0.0289
10	8.4559	8.2202	0.2357
11	7.4089	7.809	-0.4001
12	8.3872	8.8116	-0.4244
13	8.5229	8.3003	0.2226
14	7.3186	7.1705	0.1483
15	8.5376	8.4207	0.1169
16	7.8697	7.6127	0.2569
17	8.1871	8.3481	-0.161
18	6.2147	6.2699	-0.0553
19	6.4202	6.5415	-0.1213
20	7.6576	7.5605	0.0971

pIC_{50} = –0.8242(± 9.1836 x 10^{-2}) Mor07e + 27.8881 (± 5.4437) G3e –
56.6141 (± 14.1663) R6p_+ + 6.1817
n = 20, Se = 0.2926, R = 0.9324, F = 35.4984, Q = 3.1869 ... (2)

The stepwise regression finally obtained two tetra parametric models 21 and 22, possessing better statistics than both the models discussed earlier. The two models differ only due to the presence of RDF 110u and RDF 110e. But the remaining descriptors are same for both the models. These two models were then analyzed for obtaining the relative correlation potential of RDF 110u and RDF 110e, in modeling the activity. RDF 110u and RDF 110e are highly correlated descriptors (r = 0.9911). Therefore the tetra parametric model 22 [equ (4)] containing Mor07e, G3e, R6p_+ and RDF110u is best for modeling the inhibitory activity. The other tetra parametric model 21 equ (3) containing almost equal correlation potential consists of RDF 110e in space of RDF110u. These two models are showed below.

Table. 9: Cross-validation parameters for the proposed models

S.n.	Parameters used	press	ssy	press/ssy	r^2_{cv}
1. (5)	Mor07e, G3e	2.7367	7.7487	0.3532	0.6468
2. (11)	Mor07e, G3e, R6p+	1.3696	9.1158	0.1502	0.8498
3. (21)	Mor07e, G3e, R6p+, RDF110e	0.8050	9.6803	0.0832	0.9168
4. (22)	Mor07e, G3e, R6p+, RDF110u	0.7979	9.6874	0.0824	0.9176

PRESS – Predicted Sum of Squares; SSY – Sum of the Squares of the response value; R^2_{cv} – Cross validation correlation coefficient.

$pIC_{50} = -0.8482(\pm 7.3092 \times 10^{-2})$ Mor07e + 27.2749 (± 4.3146) G3e –
71.3508 (± 12.1026) $R6p_+$ – 7.6470 x 10^{-2} (± 2.3578 x 10^{-2}) RDF110e + 7.5748
n = 20, Se = 0.2317, R = 0.9608, F = 45.0928, Q = 4.1476 ...(3)

$pIC_{50} = -0.8288(\pm 7.2409 \times 10^{-2})$ Mor07e + 27.1340 (± 4.2976) G3e –
69.7369 (± 11.8635) $R6p_+$ – 7.3823 x 10^{-2} (± 2.2520 x 10^{-2}) RDF110u + 7.4821
n = 20, Se = 0.2306, R = 0.9612, F = 45.5277, Q = 4.1675 ...(4)

In order to support our study we have calculated the pIC_{50} activity from the model 22, equ (4) which is explained above. The calculated activities are then compared with observed values (Table. 8).

In order to confirm the inhibitory potential of the selected models, we have again calculated their predictive correlation coefficient (R^2 pred.)[33-35] by plotting graph between observed and predicted pIC_{50} values by equation 4. Such correlations are provided in Fig. 1. from Fig. 1, the R^2 value is obtained as 0.9239 for the model showed by equation 4. This further support that the model obtained from equation 4 has the best predictive power also.

Finally a cross-validation method[33, 35] is applied for deciding the predictive potential of the suggested model. This is required because a model with well-defined statistics may not have good predictive power. The different cross–validation parameters obtained for the recommended models are listed in table 9 and are explained below.

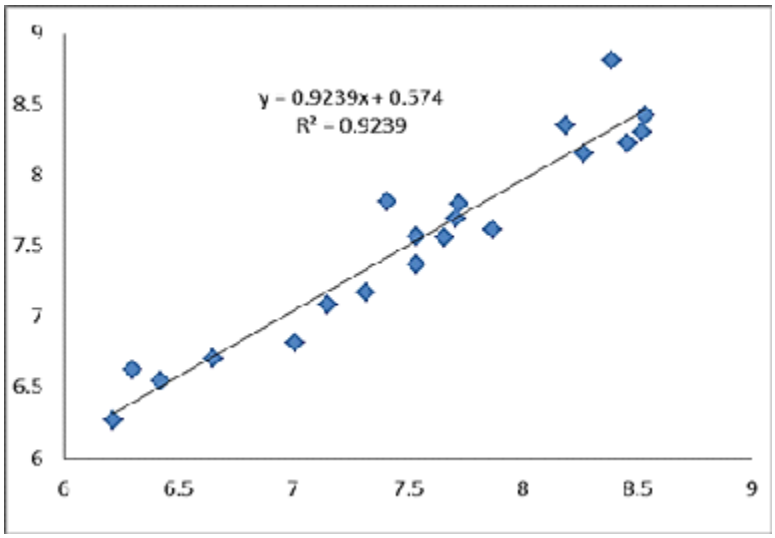


Fig.1. Fitness Plot (Observed vs. Predicted biological activity)

Predictive Residual Sum of Squares value less than the SSY (Sum of the Squares of response value) provide that the model predicts better than chance and can be treated as statistically significant. In this case (Table. 9) PRESS << SSY show that all the models obtained are statistically important and are better than chance. To be a potential QSAR model, the ratio PRESS/SSY should be smaller than 0.4. Here this ratio ranges between 0.08-0.36 showing that all the selected models are reliable QASR models. In this case, the ratio for the models obtained by equation 4 is the smallest. Therefore we summarized that this model is the best among all the models discussed above. The highest value R^2_{cv} (0.9176) provide further strength to our findings.

CONCLUSION

The positive correlation coefficient of G3e indicates that there is a direct relationship with LSS inhibitory activity of these compounds. The negative correlation coefficient of Mor07e, G3e, $R6p_+$ and RDF 110u indicate that there is an inverse relationship with LSS inhibitory activity of these compounds. MoRSE descriptors (here Mor07e) provides a good model performance when the changes in activity correlates with changes in interatomic distances due to variation in bond order and introduction of new atoms. Whim descriptors (here G3e) gave significant 3D molecular information such as molecular size, shape, symmetry and atom distribution with respect to invariant reference frames. Getaway (here $R6p_+$) are geometrical descriptors providing information of the effective position of substituents and fragments in the molecular shape. RDF code (here RDF 110u) gave information regarding interatomic distances in the whole molecule, bond distances, ring types, planar and non-planar systems and atom types.

From these findings, if the count of number of carbon atoms separated from electronegative oxygen atom to charged polar amino group by six bond increases then LSS inhibitory activity of these molecules also increases. The bond angle and bond distances of these carbon atoms, the polar electronegative atoms and aromatic and substituted aromatic rings plays a significant role in LSS inhibitory activity. These findings can be useful in the development and optimization of new LSS inhibitors.

REFERENCES

1. World Health Organization. The World Health Report 2002. Geneva, Switzerland: WHO, **2002**.
2. Gupta,R.; Gupta,S.; Sharma,K.K.; Gupta,A.;Deedwania,P.; *World. J. Cardiol.* **2012***, 4,* 112- 120.
3. Morand,O.H.; Aebi,J.D.; Dehmlow, H.; Ji, Y.H.;Gains,N.;Lengsfeld,H.;Himber,J. *J. Lipid Res.* **1997**, *38*, 373-390.
4. Lenhart.A.; Reinert,D.J.; Aebi,J.D.; Dehmlow,H.; Morand,O.H.; Hegele,R.A.; Huff,M.W. *Cir.Res.* **2003**, *93*, 717-725.
5. Mark,M.; Muller,P.; Maier,R.; Eisele,B. *J.Lipid Res.* **1996**, *37*,148-158.
6. Eisele,B.; Budzinski,R.; Muller,P.; Mark,M. *J.Lipid Res.* **1997**, *38*, 564-575.
7. Dollis,D.; Schuber,F. *Biochem. Pharmacol.* **1994**, *48*, 49-57.
8. Peffley,D.M.; Gayen,A.K.; Morand,O.H. *Biochem. Pharmacol.* **1998***, 56,* 439-449.
9. Gardner,R.G.; Shan,H.; Matsuda, S.P.T.; Hampton,R.Y. *J.Biol.Chem.* **2001**, *276*, 8681-8694.
10. Hansch,C.; Leo,A.; Hoekman,D. *J.Am. Chem. Soc.* **1995**, *117,* 9782.
11. Karelson,M.; Lobonov,V.S.; Katrizky,A.R. *Chem.Rev.* **1996**, *96*, 1027-1043.
12. Devillers,J.; Balaban,A.T. Topological Indices and Related Descriptors in QSAR and QSPR, Gordon & Breach: Williston, VT, **2000**.
13. Todeschini,R.; Cosonni,V. *Handbook of Molecular Descriptors* Wiley-VCH: Weinheim **2000**.
14. Karelson,M. *Molecular Descriptors in QSAR/QSPR*, Wiley: New York, **2000**.
15. Gutman, I. *Graph Theory Notes*, New York, **1994***, 27*, 9.
16. Weiner,H. *J.Am. Chem. Soc.* **1947**, *69*, 17.
17. Diudea,M.V. (Ed) *QSPR/QSAR Studies by Molecular Descriptors*, Babes-Bolyai

University, Cluj, Romania, **2000**.
18. Balaban,A.T. *Chem. Phys. Lett.* **1982**, *89*, 399.
19. Khadikar,P.V.; Deshpande,N.V.; Kale,P.P.; Dobrynin,A.; Gutman,I.; Domotor,G. *J. Chem. Inf. Comput. Sci.* **1995**, *35,* 547.
20. Randic,M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
21. Basak,S.C.; Mills.D. *Commun. Math. Chem (MATCH)* **2001**, *44,* 15.
22. Henrietta,D.; Johannes,D.A.; Synese,J.; YuHua, Ji.; Elisabeth,M.V.M.; Jacques,H.; Oliver,H.M. *J.Med.Chem.* **2003***, 46*, 3354-3370.
23. Chem Axon Ltd. Marvin Sketch, version 5.4.Budapest,Hungary **2010**.
24. Talete srl. DRAGON for windows (Software for Molecular Descriptor Calculations), Version 5.5. Milano, Italy: Talete srl., **2007**.
25. De Oliveira, D.B.; Gaudio, A. C.; BuildQSAR: A New Computer Program for QSAR Analysis. *Quant. Struct.-Activ. Relat.* **2001**, *19*, 599-601.
26. Schuur,J.H.; Selzer,P.; Gasteiger,J. *J. Chem. Inf. Comput. Sci.* **1996**, *36,* 334-344.
27. Gasteiger,J.; Sadowski,J.; Schuur,J.H.; Selzer,P.; Steinhauer,L.; Steinhauer,V. *J. Chem.Inf. Comput. Sci.* **1996**, *36*, 1030-1037.
28. Todeschini,R.; Lasagni,M.; Marengo,E. *J. Chemom.* **1994**, *8*, 263–273.
29. Todeschini,R.; Gramatica,P. 3D QSAR in Drug Design – Vol. 2, H. Kubinyi, G. Folkers, Y. C.Martin (Eds.), Kluwer/ESCOM, Dordrecht (The Netherlands), **1998**, 355–380.
30. Consonni,V.; Todeschini,R.; Pavan,M.; Gramatica,P. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 693–705.
31. Schuur,J.; Gasteiger,J. *Anal. Chem.* **1997**, *69,* 2398–2405.
32. Hemmer,M.C.; Steinhauer,V.; Gasteiger,J. *Vibrat. Spect.* **1999**, *19*, 151–164.
33. Chaterjee,S.; Hadi,A.S.; Price, B. Regression Analysis by Examples, 3rd Ed. Wiley: New York, **2000**.
34. Pogliani,L. *J. Phys. Chem.* **1996**, *100*, 18065.
35. Pogliani,L. *Amino Acids*, **1994***, 6*, 141.