IJPSR (2016), Vol. 7, Issue 7 (Research Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 22 February, 2016; received in revised form, 18 March, 2016; accepted, 04 May, 2016; published 01 July, 2016

2D QSAR STUDY OF POTENT GSK 3β INHIBITOR FOR TREATMENT OF TYPE II DIABETES

Seema Kesar, Pooja Mishra*, Priya Ojha and Sneha Singh

Department of Pharmacy, Banasthali University, Banasthali, 304022, Rajasthan, India

Keywords:

QSAR, 3-Aryl- 4(arylhydrazono)-1H pyrazol-5-ones derivatives, TSAR, MLR, PLS, FFNN

Correspondence to Author:
Pooja Mishra

Research Scholar
Department of Pharmacy,
Banasthali University, Banasthali,
Rajasthan-304022, India

E-mail: pujamshr89@gmail.com

ABSTRACT: The best QSAR model were generated with left of adept and significant descriptors like electronic, lipophilic and topological, using multiple linear regression (MLR) and partial least square (PLS), model further explained by using forward feed neural network analysis (FFNN). QSAR is a kind of technique that directly correlates in between chemical structure to their biological activity. The best MLR statistical expressions were evaluated with good predictive and authenticated ability and the values were S =0.367, F =53.06 r =0.910, r^2 =0.828, $r^2_{(cv)}$=0.780. The r^2 (training and test-set) values of MLR, PLS and FFNN are 0.82, 0.71, 0.82, 0.71 and 0. 81, 0.74 respectively, which predicts the soundness of the model. The model reveals that total dipole moment, bond lipole and kappa 3 are prerequisite descriptors for determining further promising GSK-3β antagonist with high and liable potency against target. In addition to QSAR modelling, Lipinski's rule of five was employed on a series and we found no violation in it, which means 3-aryl- 4-(arylhydrazono) 1H pyrazol-5-ones has enough good pharmacokinetic profile, and it become more accentuated when orally active anti-diabetic agents will formed.

INTRODUCTION: Diabetes is a worldwide problem. It is envisaged that the problem will attain approximately six thousand forty two million patients globally in 2040.[1] It includes a collection of metabolic disorders showing elevated blood sugar levels above an extensive time period.[2] Diabetes mellitus is a metabolic disorder characterized by hyperglycaemia, glycosuria, hyperlipemia, negative nitrogen balance and sometimes ketonemia.[3] In normal physiology, Insulin promotes the conversion of glucose to glycogen in skeletal muscle by stimulating glucose uptake and activating glycogen synthase.[4]

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.7(7).2932-43

Article can be accessed online on: www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.7 (7).2932-43

Defective insulin secretion and insulin resistance causes diabetes and it is of three types-Types I, Type II, Type III (gestational diabetes). Among all these, insulin independent diabetes mellitus (type 2 diabetes) accounts for more than 90% of diabetic cases. Resistance to the biological actions of insulin in tissues like muscle, liver, and adipocytes is a major feature of the pathophysiology in type 2 diabetes.[5] Glycogen synthase kinase (GSK) is a multi-targeted serine/threonine kinase, originally identified as an enzyme, having two identical isoforms namely GSK-3α (51kDa) and GSK-3β (47kDa).[6]

They display 84% overall identity (98% within their catalytic domains) with the main difference being an extra Gly-rich stretch in the N-terminal domain of GSK-3α. [7] The central role of glycogen synthase kinase -3 (GSK-3) in glucose metabolism makes it an exciting target for controlling hyperglycemia.[5] After meals, insulin controls blood

glucose levels by promoting glucose transport into peripheral tissues and enhancing formation of glycogen. At other times, glycogen formation in resting cells is suppressed via phosphorylation and inactivation of the rate-limiting enzyme glycogen synthase (GS), Insulin indirectly relieves GS inhibition through a signalling cascade beginning with phosphorylation of substrates, including insulin receptor substrate 1 (IRS-1), by the tyrosine kinase activity of activated insulin receptor, tyrosine-phosphorylated IRS-1 initiates additional events, including inactivation of glycogen synthase kinase 3 (GSK-3; which is constitutively active in resting cells) and dephosphorylation of GS.[8] Several enzymes have been implicated in the regulation of GS phosphorylation like isoforms of GSK-3. Abnormal over-expression of GSK-3 may contribute to the development of insulin resistance.

Thus, till a date there are several GSK-3β small molecule inhibitors in clinical trials for the treatment of type II diabetes.[5] However, statutory of discovering new lead compounds which having the antagonistic activity against diabetes is essential. The use of quantitative structure-activity relationships (QSAR), since their advent in 1962, has become increasingly helpful in understanding many aspects of chemical-biological interactions in drug and other scientific research.[9] In the present effort, we exaggerate our pursuit of being establishing the relationship between the various physiochemical parameters and anti-diabetic activity of 3-aryl- 4-(arylhydrazono)-1H pyrazol-5-ones derivatives.

MATERIALS AND METHODS:

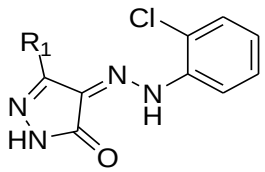
Generation of introductory structure and the art of constructing 3-D optimized structure:

Sketched all the structures of 3-Aryl-4-(Arylhydrazono)-1H-pyrazol-5-ones derivatives **Table 1** and **2** mentioned in literature, [10] on standalone module of accelrys discovery studio (version 2.0) along with their biological activities (k_i values), for in view of sufficient variation in the biological activities, large number of substituent, and all they were should be in negative algorithmic scale, because biological activities are highly prone to become skewed. Thus the mentioned (inhibitory-constant values) were converted into negative log k_i values, and then used for subsequent QSAR analysis as dependent variables. Further studies were performed into the TSAR software (version 3.3Accelrys Inc., Oxford, England). On TSAR worksheet all chemical structures were imported via mol files.

The series had two major substitutions, which was defined using "define substituent" option in the toolbar of TSAR's worksheet. Further by using Charge 2 derive charges option this step is mandatory because alignment of structures according to their molecular weight is necessary with optimizing a 3D model.

In the next step, by "Corina-make3D" option (it is designed by Rudolph, Sadowski and Gasteiger), [11] all loaded 2D molecules and substitutions were converted into their 3Dstructure. Energy optimization of all 3D structures was performed using "Cosmic-Optimize 3D" option of the software which includes valence terms as bond potentials, bond angles, torsional potential, and non-bonded terms as electrostatic interaction and Vander Waals interaction.

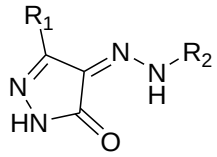
TABLE 1: STRUCTURE AND BIOLOGICAL ACTIVITY DATA OF 3-SUBSITITUTED 4-(2-(2-CHLOROPHENYL) HYDRAZONO)-1H-PYRAZOL-5(4H)-ONE INHIBITORS OF GSK-3β USED IN QSAR ANALYSIS.



Compound Name	R_1	R_2	GSK-3β, K_i values, nM
1	H	1-chloro benzene	1490
2	Me	1-chloro benzene	460
3	Ph	1-chloro benzene	720
4	2-pyridyl	1-chloro benzene	650

5	3-pyridyl	1-chloro benzene	200
6	4-pyridyl	1-chloro benzene	250
7	3-MeO-Ph	1-chloro benzene	850
8	3, 4$(MeO)_2$-Ph	1-chloro benzene	44

TABLE 2: STRUCTURE AND BIOLOGICAL ACTIVITY DATA OF N-, 3-SUBSTITUTED 4-(2-HYDRAZONO)-1H-PYRAZOL-5(4H) -ONE INHIBITORS OF GSK 3 β.



Compound Name	R_1	R_2	GSK-3β, K_i values, nM
9	Ph	Ph	99
10	Ph	3-Cl-Ph	120
11	Ph	4-Cl-Ph	23
12	Ph	3-MeO-Ph	35
13	Ph	3-MeO-Ph	20
14	2-Pyridyl	3-MeO-Ph	640
15	3-Pyridyl	3-MeO-Ph	20
16	4-Pyridyl	3-MeO-Ph	30
17	3-MeO-Ph	3-MeO-Ph	18
18	3-MeO-Ph	3-MeO-Ph	2
19	3, 4, 5-$(MeO)_3$-Ph	3-MeO-Ph	9
20	4-MeO-3-Pyridyl	3-MeO-Ph	2
21	3-MeO-Ph	Ph	39
22	3-MeO-Ph	3-Cl-Ph	8
23	3-MeO-Ph	4-Cl-Ph	9
24	3-MeO-Ph	3-MeO-Ph	5
25	3-MeO-Ph	4-CN-Ph	4
26	3-MeO-Ph	2-Pyridyl	8
27	3-MeO-Ph	3-Pyridyl	30
28	3-MeO-Ph	4-Pyridyl	23
29	3-MeO-Ph	4-CO_2H-Ph’	6.5
30	3-MeO-Ph	4-NMe_2-Ph	9
31	3-MeO-Ph	4-Morpholino-Ph	59
32	4-MeO-Ph	3-Cl-Ph	1.9
33	4-MeO-Ph	4-Cl-Ph	2
34	4-MeO-Ph	3-MeO-Ph	0.8
35	4-MeO-Ph	3-Pyridyl	2
36	4-MeO-Ph	4-Pyridyl	14
37	4-MeO-Ph	4-NMe_2-Ph	110
38	4-MeO-Ph	4-Net_2-Ph	16
39	4-MeO-Ph	4-Morpholino-Ph	8
40	3, 4-$(MeO)_2$-Ph	3-Cl-Ph	0.4
41	3, 4-$(MeO)_2$-Ph	3-MeO-Ph	0.4
42	3, 4-$(MeO)_2$-Ph	4-NMe_2-Ph	87
43	3, 4-$(MeO)_2$-Ph	4-NEt_2-Ph	2
44	3, 4-$(MeO)_2$-Ph	4-Morpholino-Ph	3
45	4-MeO-3-Pyridyl	Ph	4.5
46	4-MeO-3-Pyridyl	4-Cl-Ph	3
47	4-MeO-3-Pyridyl	2-MeO-Ph	270
48	4-MeO-3-Pyridyl	3-MeO-Ph	3.7
49	4-MeO-3-Pyridyl	4-CN-Ph	8.3
50	4-MeO-3-Pyridyl	3-Pyridyl	23
51	4-MeO-3-Pyridyl	4-Pyridyl	16
52	4-MeO-3-Pyridyl	4-CO_2H-Ph	7.7
53	4-MeO-3-Pyridyl	4-Morpholino-Ph	50

Data Set Preparation and Data Reduction:

The main reason behind of calculating molecular descriptors is to explore the structural information about all the chemical structure and respective substituent's and also to acquire a good and predictive QSAR model. More than 200 molecular descriptors were calculated, by using TSAR. TSAR is an integrated analysis package for the interactive investigation of quantitative structure-activity relationship. Since myriad numbers of numerical descriptors of molecular structures were on TSAR's worksheet. The calculated descriptors included molecular attributes, molecular indices, atom counts and VAMP parameters.[12]

Whole data of 53 compounds of 3-aryl- 4-(arylhydrazono) 1H pyrazol-5-ones derivatives were randomly divided into training and test set. Training and test set consists37 and 10 compounds respectively and the predicted values were obtained. External validation of training set compounds was conducted by using MLR, PLS and FFNN for model development. In addition, internal validation of developed models was also validated.

In the process of data reduction, main focus is to check the viability of descriptors. Firstly, pair wise and stepwise correlation analysis was performed on data set. There are large number of descriptors which have high correlative property and leads to low predictivity of the model. If any of two consecutive descriptors contain correlation coefficient greater than 0.5, it shows high correlation with each other and less correlation with biological activity. Then it directly indicates that there is no benefit to keep that type of descriptors, so it was discarded while the other are kept. This process was repeated again and again till highly correlated descriptors with biological activity were attained. Thus, three independent molecular descriptors, total dipole moment (substituent-1), bond lipole (substituent-2), kappa 3 index (substituent-1) were retrieved.

Model Development:

Linear Regression Analysis:

Relationship sets in between statistically analyzed physiochemical descriptors and the biological activity was quantified by MLR (multiple linear regression) and PLS (partial least square). Quantification and predictability of MLR model were based on a y- variable (dependent variable) and x- variables (independent variables). MLR standalone in the field of regression analysis in QSAR methodology, because it describes the relationship between dependent and independent variables. The equation of MLR reveals the correlation behaviour of descriptors with k_i values that help to understand drug receptor binding interaction and designing new chemical entities more precisely. The best model was selected on the basis of statistical parameters such as conventional regression coefficient (r^2), Fischer's ratio (F), and the standard error of estimate (S). PLS (partial least square) has been recommended as an alternative approach to enlarge the information contained in each model and avoids the danger of over fitting.[13] PLS regression can be used with more than one dependent variable [14] to reconfirm or recheck the model and their results are same as that of MLR. In proposed model, cross validation analysis was performed using leave-one-out method.

Non Linear Regression Analysis:

FFNN (Feed Forward Neural Network) is an artificial neural network and it is based on simply three fundamentals namely input, output and hidden nodes. In this network, the information moves in only one track and it goes forward from input nodes through hidden nodes and end into output nodes. In FFNN model, the neural net configuration was modified by changing the percentage of data excluded from prediction of model and also changing the number of nodes in the hidden layer. The best model unveiled the closer and relevant values of test RMS fit and best RMS fit of training set. The graphs reveal the correlation behaviour of descriptors with biological activity (k_i values).

These inhibitors were also possessed suitable pharmacokinetic (ADME) profile. Lipinski's "rule of five" was applied on whole data set. The rule was formulated by Christopher A. Lipinski in 1997, based on the observation that most medication drugs are relatively small and lipophilic molecules.[15, 16] According to this rule, to be drug-like states that the molecules with a molecular weight (>500), log p (>5), hydrogen bond donors (>5), and hydrogen bond acceptors (>10) have poor

adsorption or permeation.[15] This rule describes molecular properties which are important for a drugs pharmacokinetics in the human body. Under consideration, Lipinski's rule of five were calculated and shown in **Table 3**.

TABLE 3: VALUES OF CALCULATED PARAMETERS FOR LIPINSKI'S RULE OF FIVE.

COMP. NAME	ADME (Molecular weight)	ADME (H-bond acceptors)	ADME (H-bond donors)	ADME (Log P)	ADME Rotatable bond	ADME Violations
1	222.65	3	2	2.306	3	0
2	236.68	3	2	1.580	3	0
3	298.75	3	2	3.493	4	0
4	299.74	4	2	2.580	4	0
5	299.74	4	2	2.646	4	0
6	299.74	4	2	2.646	4	0
7	328.78	4	2	3.240	5	0
8	358.81	5	2	2.987	6	0
9	264.31	3	2	2.975	4	0
10	298.75	3	2	3.493	4	0
11	298.75	3	2	3.493	4	0
12	294.34	4	2	2.722	5	0
13	294.34	4	2	2.722	5	0
14	295.33	5	2	2.274	5	0
15	295.33	5	2	1.875	5	0
16	295.33	5	2	1.875	5	0
17	324.37	5	2	2.469	6	0
18	324.37	5	2	2.469	6	0
19	384.43	7	2	1.964	8	0
20	325.36	6	2	2.321	6	0
21	294.34	4	2	2.722	5	0
22	328.78	4	2	3.240	5	0
23	328.78	4	2	3.240	5	0
24	324.37	5	2	2.469	6	0
25	319.35	5	2	2.587	5	0
26	295.33	5	2	2.574	5	0
27	295.33	5	2	1.875	5	0
28	295.33	5	2	1.875	5	0
29	338.35	6	2	2.420	6	0
30	321.37	5	2	2.021	6	0
31	379.49	5	2	2.171	6	0
32	328.78	4	2	3.240	5	0
33	328.78	4	2	3.240	5	0
34	324.37	5	2	2.499	6	0
35	295.33	5	2	1.875	5	0
36	295.33	5	2	1.875	5	0
37	321.37	5	2	2.021	6	0
38	365.48	4	2	3.199	8	0
39	379.46	5	2	2.171	6	0
40	358.81	5	2	2.987	6	0
41	354.4	6	2	2.217	7	0
42	351.4	6	2	1.768	7	0
43	395.51	5	2	2.946	9	1
44	409.49	6	2	1.919	7	0
45	295.33	5	2	2.5741	7	0
46	329.77	5	2	4.1323	6	0
47	325.36	6	2	2.3214	6	0
48	325.36	6	2	2.3214	6	0
49	320.34	6	2	2.3214	6	0
50	296.32	6	2	1.7271	5	0
51	296.32	6	2	1.7271	5	0
52	339.34	7	3	2.2725	6	0
53	380.45	6	2	2.0235	6	0

RESULTS AND DISCUSSION:

Linear Regression Analysis:

The three highly correlated descriptors were left, regression analysis of whole data set of molecular descriptors and the model had the statistical values, which mentioned below in **Table 4**. It shows very poor predictive ability and contemplated that refinement of descriptors can improve the statistical quality of model. An improved model was obtained by deleting outliers. By applying applicability domain (AD) on the compounds of training set for finding of possible outliers. Applicability domain means leverage calculation using the system software. Applicability domain sets an appropriate assumption for whole compounds of model. In William's graph (graph of applicability domain) **Fig.1** all compounds were plotted as a point, and if any one of compound is beyond the zone of domain and shows high leverage value, then it considered as an outlier, because it cannot be associated with a reliable prediction. In this study, taking standard leverage limit is **1.5**, then six compounds 56, 51, 41, 17, 8 and 37 behaved as outliers. They had very low t-values and high leverage value, thus they were deleted.

TABLE 4: STATISTICAL TESTS AND THEIR VALUES OBTAINED AFTER PERFORMING DATA REDUCTION.

Statistical tests	Values
S value	0.52
F value	30.95
F probability	1.27
Regression coefficient, r	0.80
r^2	0.65
Cross validation, $r^2_{(cv)}$	0.56

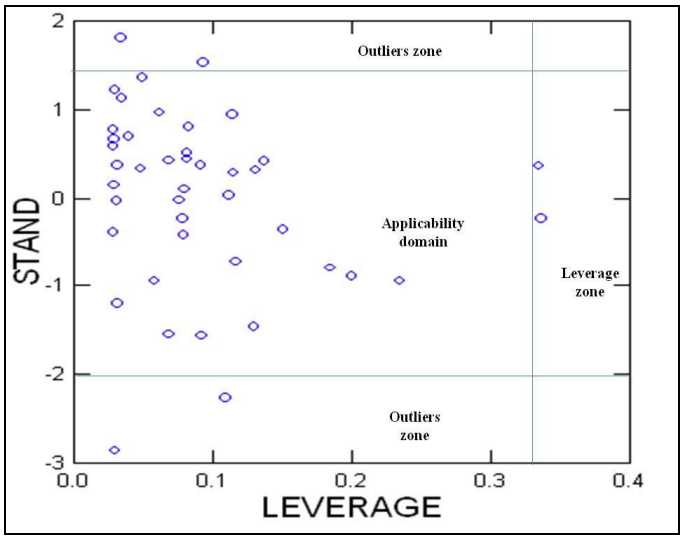


FIG.1: WILLIAM'S PLOT (GRAPH OF APPLICABILITY DOMAIN)

Performing MLR on training set compounds with the three selected descriptors that shows gradually increment in statistical values, were shown in **Table 5** and satisfactory r^2values of (training and test) confirms the robustness of the model as in **Fig. 2** and **3**.

TABLE 5: STATISTICAL TESTS AND THEIR VALUES OBTAINED AFTER PERFORMING MLR ANALYSIS.

Statistical tests	Values
S value	0.36
F value	53.0
Regression coefficient, r	0.91
r^2	0.82
Cross validation, $r^2_{(cv)}$	0.78

Equation 1 represents the MLR equation, after deleting aforesaid outliers:

$$Y = -0.483\times X1 + 0.0706\times X2 + 1.771\times X3 - 3.449 \quad \text{(Equation 1)}$$

To confirm the liability and soundness of the data set, on dimension two, PLS analysis was performed using the same data set. The resulted statistical significance= 0.90, $r^2_{(cv)}$=0.79 and r^2(test and training) values of 0.71 and 0.82 respectively as in **Fig. 4** and **5** clearly explained the authentification and high predictability of the developed PLS model (**Equation 2**).

$$Y = -0.450\times X1 + 0.084\times X2 + 1.760\times X3 - 3.546 \quad \text{(Equation 2)}$$

Where X1= total dipole moment (substituent-1), X2 = bond lipole (substituent-.2), X3 = kappa 3 (substituent-1)

Non Linear Regression Analysis:

Further, in the race of getting the best model, on 3 inputs, 1 hidden node and 1 output, 45% data were excluded, and the feed forward neural network (FFNN) has done on data set of developed model.

TABLE 6: DETAILS OF FFNN

Summary of FFNN	
Net configuration	3-1-1
Test RMS fit	0.113
No. of cycles	814
Best RMS fit	0.085

FFNN also having promising results, the r^2=0.74, 0.81(test and training) values of FFNN as shown in **Fig. 6** and **7** and the plot dependencies were evaluated.

MLR and PLS models were evaluated with comparable r^2 (test and training) values of 0.71, 0.82 and 0.71, 0.82 respectively. Details of the FFNN model, actual and predicted biological activity values of MLR, PLS and FFNN analysis for training and test set are given in **Tables 6, 7** and **8**.

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

Compound Name	Actual activity	Predicted Activity		
		MLR	PLS	FFNN
1	-3.173	-3.035	-3.051	-2.543
4	-2.857	-2.053	-2.063	-2.258
5	-2.812	-3.167	-3.102	-2.596
7	-2.397	-2.698	-2.664	-2.569
9	-1.643	-1.802	-1.793	-2.253
10	-1.995	-2.084	-2.099	-2.268
11	-2.079	-1.910	-1.891	-2.210
15	-1.301	-1.339	-1.373	-1.121
16	-2.806	-2.454	-2.413	-2.572
18	-1.477	-1.978	-1.969	-2.434
19	-1.255	-0.957	-0.962	-1.097
20	-0.301	-0.770	-0.787	-0.458
22	-0.954	-0.493	-0.467	-0.995
23	-0.301	-0.835	-0.848	-0.659
24	1.591	-0.885	-0.877	-1.025
25	-0.903	-0.711	-0.667	-0.864
26	-0.954	-0.639	-0.582	-1.801
27	-0.698	-0.936	-0.936	-1.081
28	-0.602	-0.934	-0.935	-1.071
29	-0.903	-0.910	-0.906	-1.052
30	-1.477	-1.166	-1.213	-1.286
31	-1.361	-1.160	-1.205	-1.278
32	-0.812	-1.053	-1.075	-1.205
33	-0.954	-1.176	-1.225	-1.296
34	-1.770	-1.098	-1.135	-1.178
35	-0.278	-0.528	-0.497	-0.306
36	-0.301	-0.463	-0.419	-0.271
39	-0.301	-0.840	-0.877	-0.467
43	-0.903	-0.913	-0.963	-0.532
45	-0.397	-0.041	-0.012	-0.012
52	-0.301	-0.078	-0.131	-0.122
53	-0.477	-0.481	-0.529	-0.045
54	-0.653	-0.765	-0.765	-0.596
55	-0.477	-0.520	-0.471	-0.412
57	-0.568	-0.814	-0.824	-0.641
58	-0.919	-0.786	-0.798	-0.547
60	-1.204	-1.039	-1.093	-0.832

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

Compound Name	Actual activity	Predicted Activity		
		MLR	PLS	FFNN
2	-2.662	-3.108	-3.110	-2.556
5	-2.301	-2.838	-2.791	-2.582
11	-1.361	-1.022	-0.996	-0.828
12	-1.544	-1.317	-1.347	-1.095

36	-1.146	-0.973	-1.032	-0.612
38	-1.204	-0.406	-0.355	-0.206
41	-0.397	-0.266	-0.280	-0.050
50	-1.361	-1.043	-1.099	-0.835
52	-0.886	-0.923	-0.955	-0.727
53	-1.699	-0.995	-1.040	-1.795

The three highly correlated parameters were left on TSAR sheet, further were used to generate regression equation and analyzed for their relative impacts on the activity of the compounds **Table 9**.

Therefore, it can be concluded that all the *t* –test values, Jacknife SE and Covariance SE values **Table 10** were significant for best model that confirms the importance of each descriptor.

TABLE 9: CORRELATION MATRIX SHOWING CORRELATION BETWEEN BIOLOGICAL ACTIVITY AND PARAMETERS USED.

	Total Dipole (substituent-1)	**Bond Lipole (substituent-2)**	**Kappa 3 (substituent-1)**	**K_i values**
Total Dipole (substituent-1)	1	-0.150	0.320	-0.047
Bond Lipole (substituent-2)	-0.150	1	-0.135	0.108
Kappa 3 (substituent-1)	0.320	-0.135	1	0.827
K_i values	-0.047	0.108	0.827	1

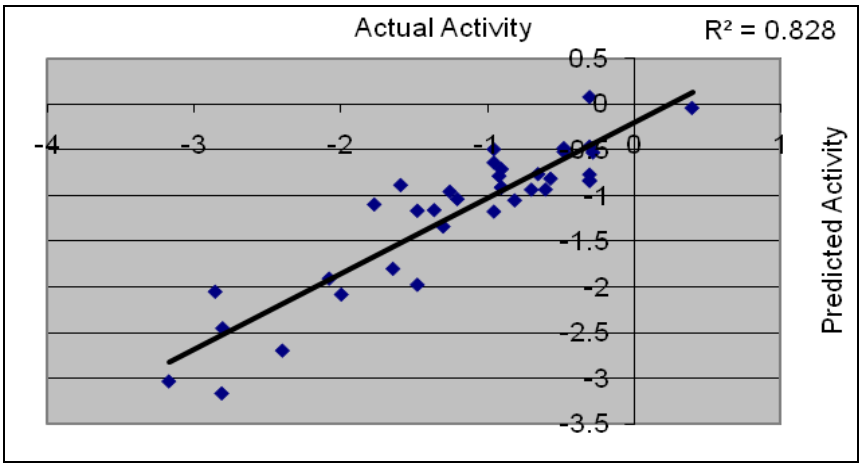


FIG. 2: PLOT OF ACTUAL VERSUS PREDICTED ACTIVITY FOR THE TRAINING SET OF COMPOUNDS DERIVED FROM MLR ANALYSIS.

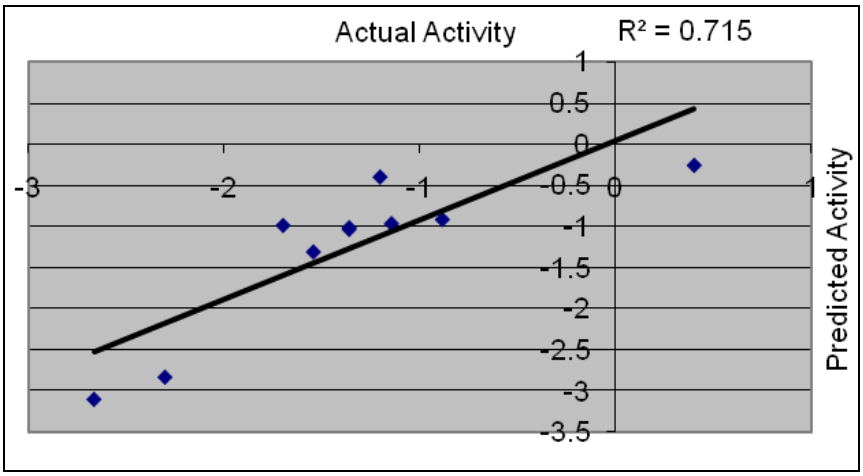


FIG. 3: PLOT OF ACTUAL ACTIVITY VERSUS PREDICTED ACTIVITY FOR THE TEST SET OF COMPOUNDS DERIVED FROM MLR ANALYSIS.

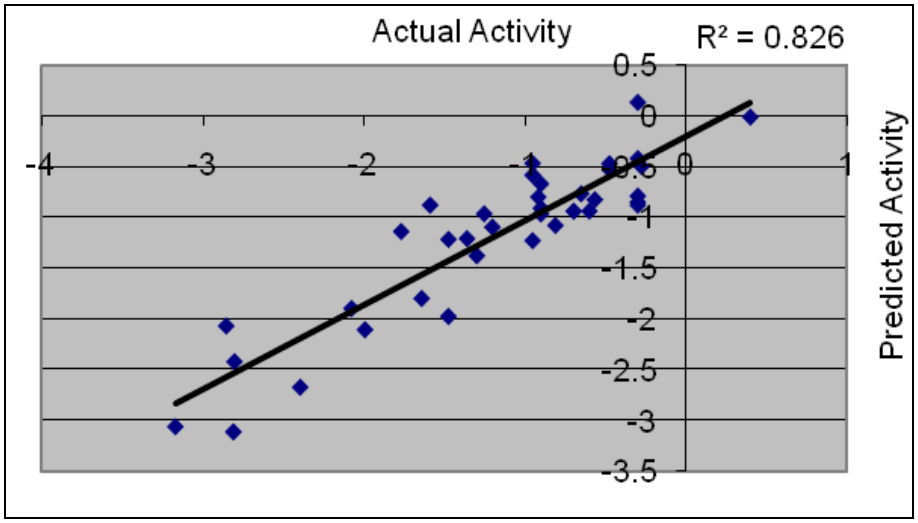


FIG.4: PLOT OF ACTUAL ACTIVITY VERSUS PREDICTED ACTIVITY FOR THE TRAINING SET OF COMPOUNDS DERIVED FROM PLS ANALYSIS.

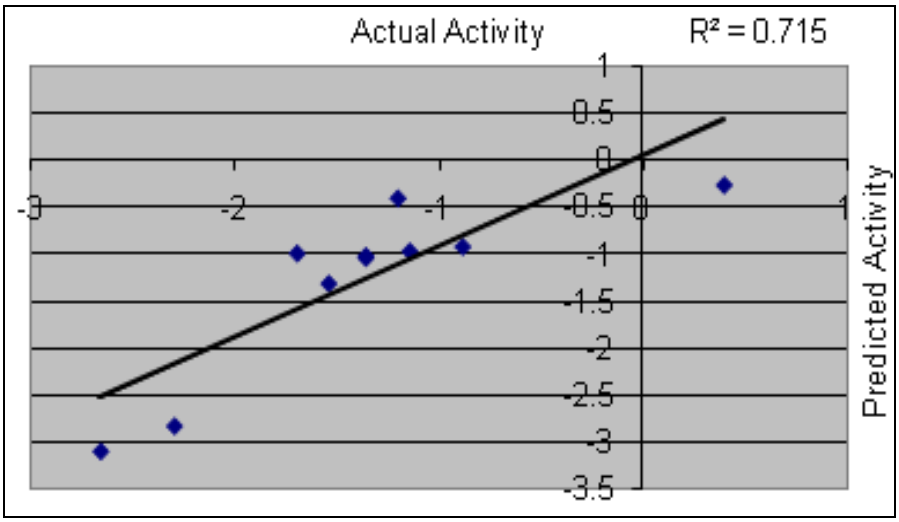


FIG.5: PLOT OF ACTUAL VERSUS PREDICTED ACTIVITY FOR TEST SET COMPOUNDS DERIVED FROM PLS ANALYSIS.

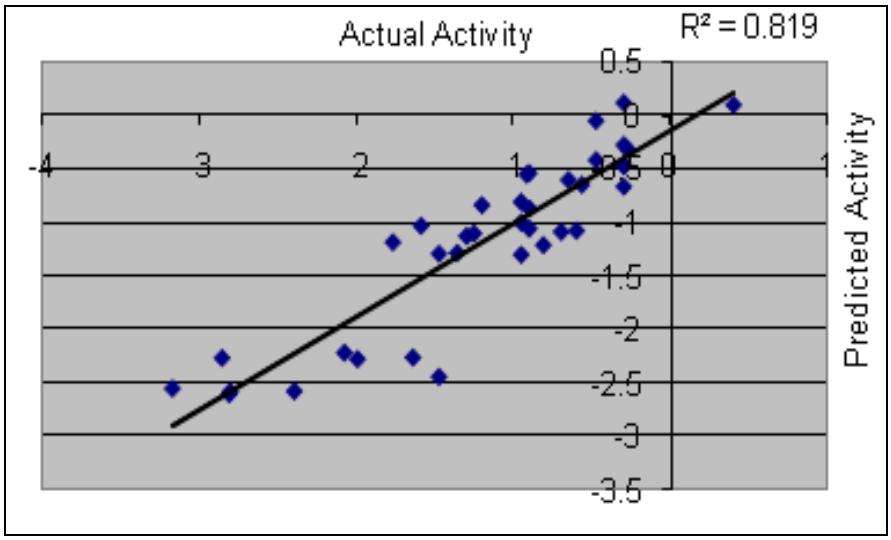


FIG.6: PLOT OF ACTUAL VERSUS PREDICTED ACTIVITY FOR TRAINING SET OF COMPOUNDS DERIVED FROM FFNN ANALYSIS.

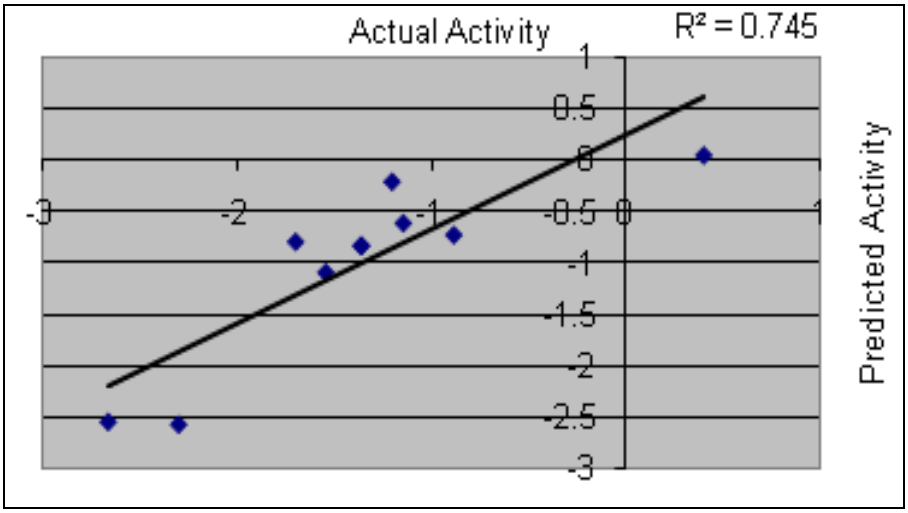


FIG.7: PLOT OF ACTUAL AND PREDICTED ACTIVITY FOR THE TEST SET OF COMPOUNDS OBTAINED FROM FFNN ANALYSIS

Interpretation of Descriptors Entered:

First descriptor i.e. total dipole moment (substituent-1), explains the charge distribution and orientation behaviour in a molecule.[17] In regression equation and FFNN plot dependency **Fig. 8**, total dipole moment descriptor is negatively correlated with the biological activity. It directly indicates that Substitution of such kind of groups in a molecule which decrease the polarity and increase the biological activity simultaneously. It clearly shows that the active site of GSK-3β will definitely have some hydrophobic interactions, and also gives a clue that active site on GSK 3β are lipophilic in nature. The second descriptor, which is bond lipole moment (substituent-2), lipophilicity, means how easily a molecule may be travelled across the biological membrane. Lipole moment descriptor is positively correlates with k_i value of molecule, which is further supported by FFNN Dependency plot **Fig. 9** which means introduction of lipophilic substitution will increase the biological activity.

The third descriptor, kappa 3 index (substituent-1), it is well known and quite elusive topological descriptor that describes the shape or steric configuration of the molecule have a fathomless influence on biological activity. A kappa 3 indices, derived from counts of atoms, bonds and flexibility depict a molecule as being related to the extremes of linear and maximally branched structures, [18] and the parameter positively correlates with the biological activity in the regression equation, which is further supported by FFNN Dependency plot **Fig. 10**. Addition of some linear or branched structures will lead to an increase in biological activity of the lead compound and in the correlation matrix of TSAR, kappa 3 index (substituent-1) were highly correlated with k_i value, that means structural changes should be necessary in further designing of new chemical entities.

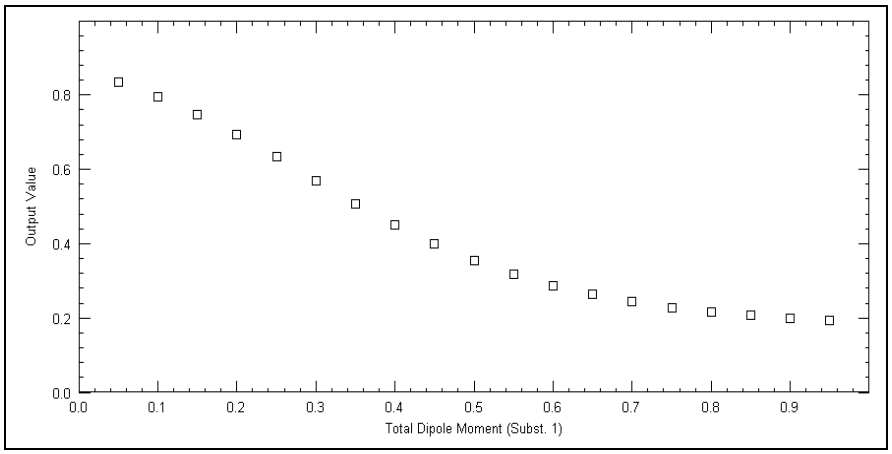


FIG.8: DEPENDENCY PLOT BETWEEN BIOLOGICAL ACTIVITY AND TOTAL DIPOLE MOMENT (SUBSTITUENT-1)

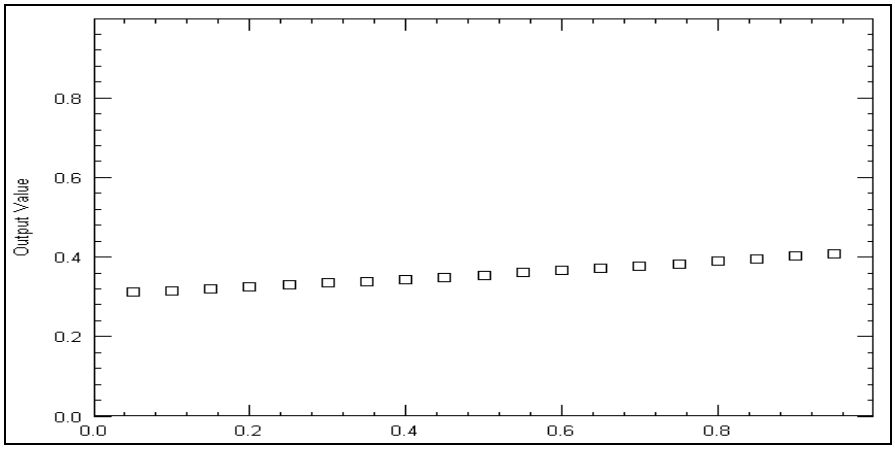


FIG. 9: DEPENDENCY PLOT BETWEEN BIOLOGICAL ACTIVITY AND BOND DIPOLE (SUBSTITUENT-2).

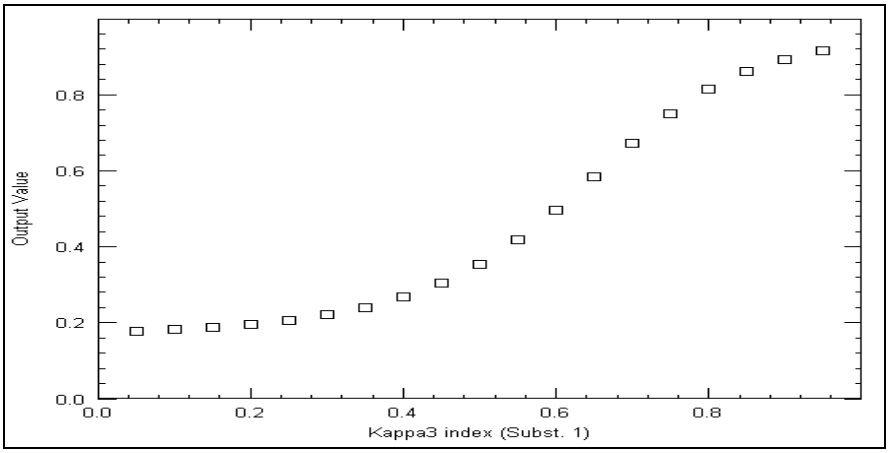


FIG.10: DEPENDENCY PLOT BETWEEN BIOLOGICAL ACTIVITY AND KAPPA3 INDEX (SUBSTITUENT-1).

TABLE 10: t-TEST VALUES, JACKNIFE SE AND COVARIANCE SE FOR THE SELECTED DESCRIPTORS.

Descriptors	t-value	Jacknife SE	Covariance SE
Total dipole moment (substituent-1)	-4.246	0.151	0.113
Bond lipole (substituent-2)	2.587	0.024	0.063
Kappa 3 index (substituent-1)	12.51	0.167	0.141

CONCLUSION: QSAR study was successfully performed on a series of pyrazolones analogues acting against GSK-3β. MLR, PLS and FFNN analysis were performed on model, and wrapped up with possessing very predictive and exhibited comparable results, and also having some useful information about parameters. According to the classical QSAR models presented in the current work, remaining molecular descriptors encoding the shape, lipophilic and polarity architecture of pyrazolones analogs are considered to be important contributors to their biological properties.

ACKNOWLEDGEMENT: Authors are thankful Banasthali University for providing necessary facilities to complete this work.

CONFLICT OF INTEREST: The authors report no conflict of interest. Only the authors are responsible for the contents and writing of paper.

REFERENCES:

1. Franek E, Rutten GEHM, Orsted DD, Baeres FMM, Mota M, Jacob S, Bain SC, Vidal J and Haluzik M: Leader 8: Type 2 Diabetes Patients: A comparison of baseline characteristics of eastern and western european participants with established cardiovascular disease in the leader trial Journal of Diabetes & Metabolism 2016; 7(2):1-6.
2. Gangawane AK, Bhatt B and Matkar S: Skin Infections in Diabetes: A Review 2016; 7(2):1-4**.**
3. Tripathi KD: Essentials of medical pharmacology. Jaypee Brothers Medical Publishers (P) Ltd, seventh edition 2013.
4. Taha MO, Bustanji Y, Al-Ghussein MAS, Mohammad M, Zalloum H, Al-Masri IM and Atallah N: Pharmacophore Modeling, Quantitative Structure Activity Relationship Analysis, and in Silico Screening Reveal Potent Glycogen

Synthase Kinase-3βInhibitory Activities for Cimetidine, Hydroxychloroquine, and Gemifloxacin. Journal of Medicinal Chemistry 2008; 51:2062-2077.

5. Zivkovic JV, Trutic NV, Veselinovic JB, Nikolic GM, Veselinovic AM, Monte Carlo method based QSAR modelling of maleimide derivatives as glycogensynthasekinase-3β inhibitors, Computers in Biology and Medicine 2015; 64:276–282
6. Arnost M, Pierce A, ter Haar E, Lauffer D, Madden J, Tanner K and Green J: 3-Aryl-4-(arylhydrazono)-1H-pyrazol-5-ones: Highly ligand efficient and potent inhibitors of GSK3β.Bioorganic & Medicinal Chemistry Letters 2010; 20(5):1661–1664.
7. Meijer L, Flajolet M and Greengard P: Pharmacological inhibitors of glycogen synthase kinase 3. Trends in Pharmacological Sciences 2004; 25(9):471-480.
8. Ring D B, Johnson KW, Henriksen EJ, Nuss JM, Goff D, Kinnick TR, Ma ST, Reeder JW, Samuels I, Slabiak T, Wagman AS, Hammond M-E W and Harrison SD: Selective Glycogen Synthase Kinase 3 Inhibitors Potentiate Insulin Activation of Glucose Transport and Utilization *In Vitro and In Vivo*. Diabetes 2003; 52:588-595.
9. Paliwal SK, Verma AN and Paliwal S: Structure–activity relationship analysis of cationic 2-phenylbenzofurans as potent anti-trypanosomal agents: a multivariate statistical approach. Monatshefte fur Chemie 2011; 142(10):1069–1086.
10. McManus EJ, Sakamoto K, Armit LJ, Ronaldson L and Shpiro N: Role that phosphorylation of GSK3 plays in insulin and Wnt signalling defined by knocking analysis. The European Molecular Biology Organization Journal 2005;24:1571-1583.
11. Paliwal SK, Verma AN and Paliwal S: Quantitative Structure Activity Relationship analysis of Di-cationic Diphenylisoxazole as potent Anti-Trypanosomal Agents. QSAR & Combinatorial Science 2009; 28(11-12):1367-1375.
12. Paliwal SK, Das S, Yadav D, Saxena M and Paliwal S: Quantitative structure activity relationship (QSAR) of N^6-substituted adenosine receptor agonists as potential antihypertensive agents. Medicinal Chemistry Research 2011; 20:1643-1649.
13. Kubinyi H: Evotionary variable selection in regression and PLS analyses. Journal of Chemometrics1996; 10(2):119-133.
14. Paliwal SK, Seth D, Yadav D, Yadav R, and Paliwal S: Development of a robust QSAR model to predict the affinity of pyrrolidine analogs fordipeptidyl peptidase IV (DPP- IV). Journal of Enzyme Inhibition and Medicinal Chemistry2011; 26(1):129-140.
15. Lipinski CA, Lombardo F, Dominy BW and Feeney PJ: Experimental & Computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews 2001; 46(1-3):3-26.
16. Lipinski CA: Lead- and drug-like compounds: the rule-of-five revolution. Drug Discovery Today: Technologies 2004; 1(4): 337–341.
17. Karelson M: Molecular Descriptors in QSAR/QSPR. John Wiley and Sons Ltd, first edition 2000
18. Devillers J and Balaban AT: In: Topological Indices and Related Descriptors in QSAR and QSPR. Gordan and Breach Science Publishers 1999.

How to cite this article:
Kesar S, Mishra P, Ojha P and Singh S: 2D QSAR Study of Potent GSK3β Inhibitor for Treatment of Type II Diabetes. Int J Pharm Sci Res 2016; 7(7): 2932-43.doi: 10.13040/IJPSR.0975-8232.7(7).2932-43.