Research Article [Pandey & Gupta, 7(3): March, 2016:4942-4951]
CODEN (USA): IJPLCP ISSN: 0976-7126

INTERNATIONAL JOURNAL OF PHARMACY & LIFE SCIENCES
(Int. J. of Pharm. Life Sci.)

Molecular Descriptors Analysis of Oxyiminoalkanoic Acid Analogs towards PPARγ Agonist Activity

Jyoti Pandey[1*] and Arun Kumar Gupta[2]
1, Suresh Gyan Vihar University, Jaipur, (RJ) - India
2, RKDF Institute of Pharmaceutical Sciences, Indore, (MP) - India

Abstract

Diabetes mellitus have reached global pandemic proportions with India being designated 'diabetes capital' of the world. Diabetes is a multi-system disorder comprising metabolic and vascular abnormalities resulting from insulin deficiency, with or without insulin resistance. There are a number of groups that endeavor to produce new anti-diabetic agents. In view of above and as a part of our effort to develop newer anti-diabetic agents, molecular modeling analysis was performed to developed QSAR models that show substantial predictive promise for oxyiminoalkanoic acid analogs. The QSAR model explains 81.9 percent variance in activity with standard error of estimation (0.405). Model showed statistical significant internal predictivity (Q^2=0.709) and external predictivity (r^2_{pred}=0.603) values. The detailed structural investigation revealed that the PPARγ agonist activity is predominantly explained by the 2D-autocorrelation descriptors (*MATS8e* & *MATS8v*) and eigenvalue-based descriptors (*BEHp5* & *VRA1*). The structural insights gleaned from the study are helpful in design of agonist with enhanced potency.

Key-Words: QSAR; Oxyiminoalkanoic acid analogs; Anti-diabetic activity; PPARγ agonist; Molecular descriptors

Introduction

According to current estimates, the human population worldwide appears to be in the midst of an epidemic of diabetes. Despite the great strides that have been made in the understanding and management of diabetes, the disease and disease-related complications are increasing unabated. Parallel to this, recent developments in understanding the pathophysiology of the disease process have opened up several new avenues to identify and develop novel therapies to combat the diabetic plague[1].

In diabetes homeostasis of carbohydrate and lipid metabolism is not properly regulated by insulin. Despite the great strides that have been made in understanding and management in this disease, tissue or vascular damage ensues leading to severe diabetic complications such as retinopathy, neuropathy, nephropathy, cardiovascular complications, and ulceration. Thus, diabetes covers a wide range of heterogeneous diseases.[2]

*** Corresponding Author**
E.mail: arunkg73@gmail.com

The graph of diabetes-related mortality is rising unabated. Multiple defects in the pathophysiology of diabetes are mostly imprecisely understood, and therefore warrant not isolating a single drug target to the reversal of all or majority of aspects of the disease.

One novel class of antidiabetic agents that appear to be effective as a treatment for diabetes is peroxisome proliferator-activated receptor (PPAR) agonists such as thiazolidine 2,4-diones (TZDs) represented by ciglitazone, englitazone, pioglitazone and rosiglitazone3. PPARs are the orphan members of the nuclear receptor gene family of ligand activated transcription factors. Three subtypes of PPARs have been cloned from mouse and human: PPAR-α, PPAR-γ, and PPAR-δ. The PPARs are belived to play a physiological role in regulation of glucose and lipid metabolism.

Thiazolidinediones such as pioglitazone, rosiglitazone, and troglitazone have made a great contribution to therapy for type 2 diabetes. However, weight gain has been reported as a side effect of these drugs.[3,4] Thus, improvement of the thiazolidinedione class of antidiabetic agents is still worth pursuing.

Recently novel oxyiminoacetic acid derivatives as potent glucose and lipid lowering agents, (Z)-2-{4-[(5-methyl-2-phenyl-1,3-oxazol-4-yl)methoxy]benzyloxyimino}-2-phenylacetic acid and

(Z)-2-{4-[(5-methyl-2-phenyl-1,3-oxazol-4-yl)methoxy]-benzyloxyimino}-2-(4-phenoxyphenyl)acetic acid were reported.[5] These oxyiminoacetic acids showed transcriptional activity for PPARγ, indicating that the mechanisms of action of oxyiminoacetic acid derivatives involve PPARγ.

Several *in silico* techniques are utilized in the process of drug design and development, one such technique is quantitative structure activity relationship (QSAR), has been traditionally perceived as means of establishing correlation between trends in chemical structure modification and respective changes of biological activity. This quantitative technology can be utilized to improve the structure of the agonist/inhibitor molecule and to interpret the improved structure in terms of favorable biological interactions. Thus, the use of predictive computational (*in silico*) QSAR models allows the biological properties of virtual structures to be predicted, and a more informed choice of target to be selected for synthesis. The use of computational approaches for the estimation of the activity of various molecules as drug candidates prior to their synthesis can save the resources and accelerate the drug discovery procedure. We carried out QSAR analysis and established QSAR models to guide further structural optimization and predict the potency and physiochemical properties of oxyiminoalkanoic acid analogs.

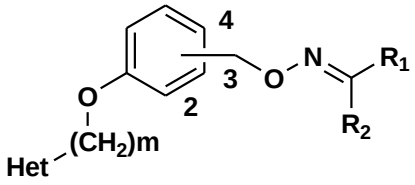
Material and Methods

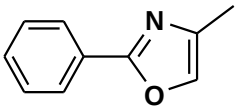
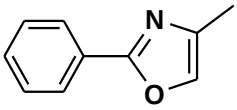
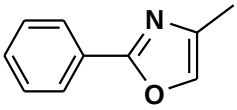
Transactivation of PPARγ activity of oxyiminoalkanoic acids analogs was taken from the reported work of Imoto *et al.*[6] (**Table 1**). In attempting QSAR, these activation data (EC_{50}) were converted to negative logarithmic dose in moles (pEC_{50}) because a QSAR is a linear free energy relationship, and from the van't Hoff isotherm, free energy change during a process is proportional to the logarithm of the rate or equilibrium constant of the process.

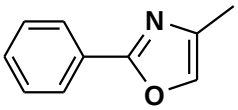
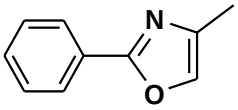
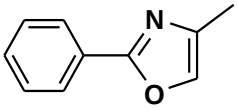
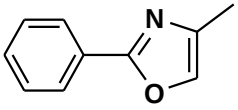
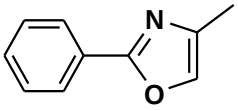
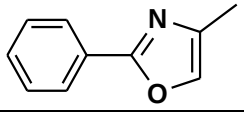
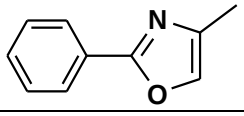
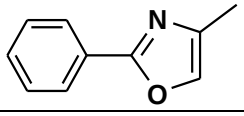
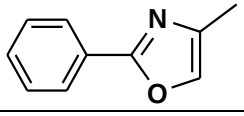
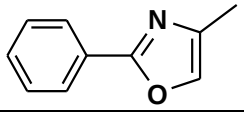
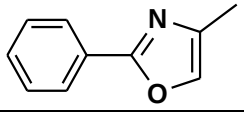
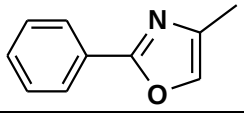
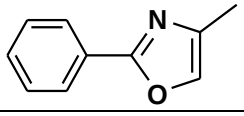
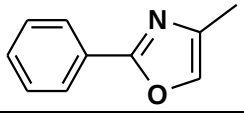
$$\Delta G = -2.303\ RT \log K$$

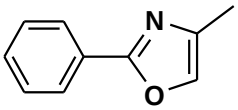
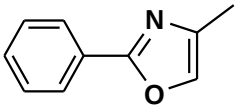
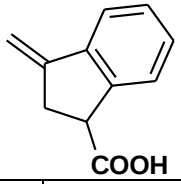
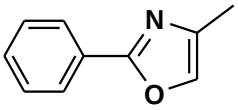
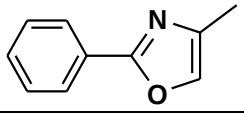
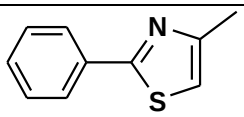
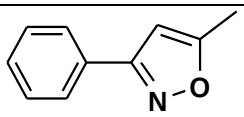
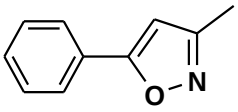
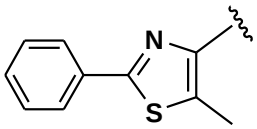
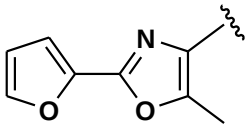
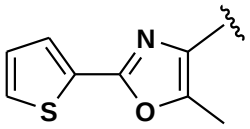
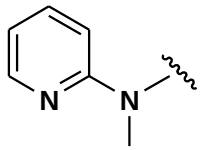
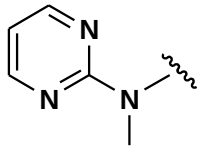
The molecular modeling study was performed using Chemoffice[7] and *DRAGON*[8] program while the regression analysis was carried out on VALSTAT[9]. Structure of all the compounds was sketched using builder module of the program. The sketched structures were subjected to energy minimization using molecular mechanics (MM2) until the root mean square (RMS) gradient value became smaller than 0.1 kcal/mol Å. The energy minimized molecules were subjected to re-optimization via Austin model-1 (AM1) method until the RMS gradient attained a value smaller than 0.0001 kcal/mol Å using MOPAC. The geometry optimization of the lowest energy structure was carried out using EF routine. The minimized molecule was saved as "MOL file". These files were used for calculation of various molecular descriptors with the help of *DRAGON*[8].

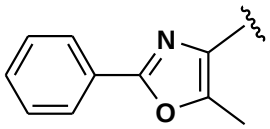
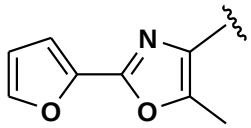
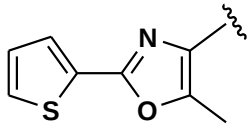
Table 1: Structure and Trnsactivation PPARγ activity (EC_{50} in μM) of Oxyiminoalkanoic Acid Analogs



Comp. No.	R_1	R_2	Sub. position	m	Het	Trnsactivation PPARγ EC50 (μM)	pEC_{50}
1	Phenyl	$-CH_2COOH$	4	1		0.053	7.276
2	Phenyl	$-(CH_2)_2COOH$	4	1		0.024	7.620
3	Phenyl	$-(CH_2)_2COOH$	3	1		0.02	7.699

4	Phenyl	$-(CH_2)_2COOH$	2	1		0.69	6.161
5	Phenyl	$-(CH_2)_3COOH$	4	1		0.062	7.208
6	Phenyl	$-(CH_2)_4COOH$	4	1		0.0025	8.602
7	$-(CH_2)_4COOH$	Phenyl	4	1		0.056	7.252
8	Phenyl	$-(CH_2)_4COOH$	3	1		0.0035	8.456
9	Phenyl	$-(CH_2)_5COOH$	4	1		0.0039	8.409
10	Phenyl	$-(CH_2)_6COOH$	4	1		0.26	6.585
11	Phenyl	$-(CH_2)_7COOH$	4	1		12	4.921
12	4fluorophenyl	$-(CH_2)_2COOH$	4	1		0.047	7.328
13	4phenoxyphenyl	$-(CH_2)_2COOH$	4	1		0.048	7.319
14	2-pyridyl	$-(CH_2)_2COOH$	4	1		0.045	7.347
15	3-pyridyl	$-(CH_2)_2COOH$	4	1		0.25	6.602
16	4-pyridyl	$-(CH_2)_2COOH$	4	1		0.17	6.770
17	Phenyl	$-CONH(CH_2)_2COOH$	4	1		0.11	6.959

18	Phenyl	-$CH_2CONHCH_2COOH$	4	1		0.38	6.420
19	Phenyl	-$(CH_2)_3C(CH_3)_2COOH$	4	1		0.032	7.495
20			4	1		0.83	6.081
21	Phenyl	-$(CH_2)_2COOH$	4	1		0.25	6.602
22	Phenyl	-$(CH_2)_2COOH$	4	1		0.16	6.796
23	Phenyl	-$(CH_2)_2COOH$	4	1		0.98	6.009
24	Phenyl	-$(CH_2)_2COOH$	4	1		0.31	6.509
25	Phenyl	-$(CH_2)_2COOH$	4	1		0.36	6.444
26	Phenyl	-$(CH_2)_2COOH$	4	1		0.042	7.377
27	Phenyl	-$(CH_2)_2COOH$	4	1		0.017	7.770
28	Phenyl	-$(CH_2)_2COOH$	4	2		0.91	6.041
29	Phenyl	-$(CH_2)_2COOH$	4	2		2.1	5.678

30	Phenyl	$-(CH_2)_2COOH$	4	2		0.34	6.469
31	Phenyl	$-(CH_2)_2COOH$	4	1		0.0049	8.310
32	Phenyl	$-(CH_2)_2COOH$	4	1		0.00084	9.076

The data was transferred to the statistical program in order to establish a correlation between physicochemical parameters as independent variables and pEC_{50} as dependent variable employing sequential multiple linear regression (SEQ-MLR) analysis method. In sequential multiple linear regression, the program searches for all the permutation and combination sequentially for the given data set. The statistical quality of the SEQ-MLR equations were assessed by parameters like correlation coefficient (r), standard error of estimate (SEE), sequential Fischer test (F) at specified degree of freedom (df) and explained variance (r^2_{adj}). The internal predictive powers of the equations were validated by (leave one out or loo) method using predicted residual sum of squares (PRESS), cross validation squared correlation coefficient (Q^2), standard deviation based on PRESS (S_{PRESS}), total sum of squares (SSY) and standard deviation of error of prediction (S_{DEP}). Chances of fortuitous correlation were tested with the help of Y-scrambled test. Finally, selected equations have been validated using test set considering predictive squared correlation coefficient r^2_{pred}. The data within the parenthesis is the standard deviation associated with the coefficient of descriptor in regression equation.

Results and Discussion

QSAR analysis in computational research is responsible for the generation of models to correlate biological activity and physicochemical properties of a series of compounds. The quality of a QSAR model, however, depends mainly on the type and quality of the data, and is valid only for the compound structure analogs to those used to build the model. QSAR models can stand alone, augment other computational approaches, or be examined in tandem with equations of a similar mechanistic generate to establish their authenticity and reliability. In the present study, QSAR analysis has been performed to explore structural requirement for PPARγ activation activity of oxyiminoalkanoic acids analogs.

In order to gain an insight to the essential structural and physiochemical requirements for the PPARγ activation activity of this class of molecules, analogs were divided into training set of 21 compounds and test set of 11 compounds. The data was transferred to the statistical program *VALSTAT* in order to establish a correlation between molecular descriptors as independent variables and pEC_{50} as dependent variable employing sequential multiple linear regression (SEQ-MLR) analysis method. The multi-variant expressions were developed on the basis of adjustable correlation coefficient (r^2_{adj}). This parameter explains statistical significance of incorporated physicochemical descriptors in regression. r^2_{adj} takes into account of adjustment of coefficient of determination (r^2). If r^2_{adj} value decline by the addition of a physicochemical descriptor to the equation it is indicated that descriptor was not contributed fairly. Adjustable correlation coefficient is a measure of % explained variation of regression expression. Whereas r^2 value is always increase when an independent variable added to the regression expression.[10]

$$\boldsymbol{pEC_{50}} = -5.403(\pm1.562)GATS5p + 16.478$$
n=21, r=0.622, r^2_{adj}=0.354, SEE =0.766, F=11.971
(eqn. 1)

$$\boldsymbol{pEC_{50}} = -6.811(\pm1.325)GATS5p \ -5.643e\text{-}005(\pm 1.676e\text{-}005)VRA1 + 19.308$$
n=21, r=0.790, r^2_{adj}=0.582, SEE=0.617, F=14.907
(eqn. 2)

$$\boldsymbol{pEC_{50}} = 3.395(\pm0.832)Lop - 6.578(\pm0.971)GATS5p - 12.242\ e\text{-}005(\pm2.030e\text{-}005)VRA1 + 15.990$$
n=21, r=0.890, r^2_{adj}=0.776, SEE=0.451, F=24.116
(eqn. 3)

SEQ-MLR revealed that the r^2_{adj} value is increasing significantly from the uni to the trivariant expressions

i.e. 0.354, 0.582 and 0.776 respectively. Boosting of r^2_{adj} value from uni to trivariant revealed that incorporation of physicochemical descriptors improve the quality of mathematical expression in a comprehensible manner. Significant improvement in r^2_{adj} value emphasizes to explore the higher variant expressions. Therefore several tetravariant expressions were developed through SEQ-MLR method.

pEC_{50} = 0.113(±0.022) Mor02m + 3.559(±0.534) Lop -5.164(±0.681) GATS5p -14.428e-005(± 1.369e-005) VRA1+ 9.785

n=21, r=0.963, r^2_{adj} =0.908, SEE=0.289, F=50.552 (eq^n. 4)

pEC_{50} = 2.653(±0.757) Lop + 4.801(±1.745)BEHp5 -5.286(±0.949)GATS5p -1.302e-004(± 1.747e-005) VRA1-1.813

n=21, r=0.933,r^2_{adj}=0.839, SEE=0.383, F=26.964 (eq^n. 5)

pEC_{50} = 7.676(±1.631)BEHp5 + 8.908(±2.662)MATS8v + 4.807(±1.041)MATS8e - 8.050e-005(±1.397e-005) VRA1 -18.741(±5.464)

n=21, r=0.925, r^2_{adj} =0.819, SEE=0.405, F=23.679 (eq^n. 6)

The robust QSAR models should have to satisfy both statistical quality and predictive power. Therefore, all the expressions were tested for internal and external corroboration. Both the validations put forward decision- making input for selection of QSAR models. Internal corroboration was carried out using leave-one-out cross-validation method and Y-scrambling test while external corroboration confirmed with the help of test set data (**Table 2**). Tetra-variant eqn. 4-6 shows correlation coefficient value in the range of 0.963-0.925, which accounts for more than 81.9% of the explained variance in the activity, calculated as $r^2_{adj} = r^2(1 - 1/F)$ that accounts in percentage when multiplied by 100 . Model revealed that the dependent variable can be predicted from a linear combination of the independent variables. The data show an overall internal statistical significance level better than 99.9% as calculated variance ratio i.e. Fischer value (F) exceeded the tabulated $F_{(4,16\alpha0.001)}$ = 9.084. Fischer value suggested that the equations are applicable for more than 999 out of 1000 times.

Table 2: Statistical data of tetra-variant expressions

Statistical Parameter	Eqn. 4	Eqn. 5	Eqn. 6
r	0.963	0.933	0.925
r^2	0.927	0.871	0.855
SEE	0.289	0.383	0.405
r^2_{adj}	0.908	0.839	0.819
F	50.552	26.964	23.679
PE	0.011	0.019	0.021
QF	3.334	2.435	2.282
FIT	5.465	2.915	2.560
LOF	4.086	7.199	8.053
AIC	0.136	0.239	0.267
ICAP	≅<0.820	≅<0.820	≅<0.650
VIF	≅<3.500	≅<3.500	≅<2.000
Q^2	0.875	0.806	0.709
S_{PRESS}	0.377	0.470	0.575
S_{DEP}	0.329	0.410	0.502
chance	< 0.001	< 0.001	< 0.001
r^2_{BS}	0.915	0.847	0.831
S_{BS}	0.062	0.110	0.111
$r^2_{RANDMEAN}$	0.193	0.199	0.192
S_{RAND}	0.120	0.124	0.118
$r^2_{RANDMAX}$	0.698	0.707	0.664
r^2_{pred}	-0.301	0.344	0.603

The equations were analyzed for the outlier by the Z-score method (Z value), the outlier test helpful in the detection of unexplainable structurally varied analogs.[11] The winning QSAR model should not have any outlier. The Z value for individual compounds lies within the specific range (<2.5), which indicated the absence of outliers. Test revealed that the eq^n. 6 is able to explain the structurally varied analogs and is helpful in the designing of more effective compounds using physiochemical descriptors (**Table 3**).

Table 3: Value of calculated, calculated (leave one out), residual and Z-score of oxyiminoalkanoic acid analogs towards PPARγ agonist activity obtained from model

Comp. No.	Cal[a]	Cal_{res}[b]	Zvalue[c]	Cal(loo)/ Pred[d]	$Cal(loo)_{res}$ / $Pred_{res}$[e]
1	7.112	0.164	0.451	7.054	0.222
2	6.994	0.625	1.725	6.935	0.685
3	7.472	0.227	0.626	7.453	0.246
5	7.452	-0.245	-0.675	7.479	-0.271
7	7.488	-0.236	-0.652	7.599	-0.347
9	7.761	0.648	1.787	7.608	0.801
10	6.443	0.142	0.392	6.402	0.183
11	5.134	-0.213	-0.588	6.047	-1.127
12	6.972	0.356	0.981	6.933	0.395
13	7.779	-0.460	-1.270	7.962	-0.643
15	7.036	-0.434	-1.197	7.078	-0.476
16	6.983	-0.214	-0.589	7.043	-0.274
20	6.650	-0.569	-1.570	6.845	-0.764
22	6.983	-0.187	-0.517	7.016	-0.220
24	6.994	-0.485	-1.337	7.095	-0.586
26	7.524	-0.148	-0.407	7.578	-0.201
28	5.633	0.408	1.125	5.487	0.554
29	5.704	-0.026	-0.072	5.712	-0.035
30	6.070	0.399	1.100	5.947	0.522
31	8.266	0.044	0.121	8.249	0.060
32	8.870	0.206	0.567	8.747	0.329
4	-	-	-	6.409	-0.248
6	-	-	-	8.381	0.221
8	-	-	-	8.155	0.301
14	-	-	-	7.130	0.217
17	-	-	-	6.983	-0.025
18	-	-	-	6.884	-0.464
19	-	-	-	6.364	1.131
21	-	-	-	7.107	-0.505
23	-	-	-	6.977	-0.968
25	-	-	-	6.929	-0.485
27	-	-	-	7.741	0.028

[a] Calculated data of the compounds using model; [b] Residual value of calculated data; [c] Outlier Z-score value obtained from model; [d] Calculated (loo) data of training set compounds using leave-one-out method or predicted data of test set; [e] Residual value of calculated (loo) data / predicted data

The orthogonality of the descriptors in the equations were established through variance inflation factor (VIF) and pair-wise correlations among the descriptors, values are shown in **table 4** and **5** respectively. In case of eq[n]. 6 VIF is less than ≅<2.000 for all the contributing descriptors revealed that the descriptors are fairly independent to each other. The low value of pair-wise correlation (PWC) among the descriptors (< 0.650) also supported comparatively independent contribution.[12,13]

Table 4: t-value and VIF value of the descriptors used in tetra-variant equations

Equations	Intercept/descriptors	t-value	VIF
Eqn. 4	***GATS5p***	7.579	1.341
	VRA1	10.539	3.382
	Mor02m	5.136	1.621
	Lop	6.664	3.004
Eqn. 5	***GATS5p***	5.570	1.476
	VRA1	7.452	3.126
	Lop	3.504	3.428
	BEHp5	2.751	2.526
Eqn. 6	***VRA1***	5.762	1.786
	BEHp5	4.706	1.972
	MATS8v	3.346	1.116
	MATS8e	4.617	1.089

Table 5 Pair wise correlation matrix of physicochemical properties used in QSAR analysis of oxyiminoalkanoic acid analogs towards PPARγ agonist activity

Parameters	*GATS5p*	*VRA1*	*Mor02m*	*Lop*	*BEHp5*	*MATS8v*	*MATS8e*
GATS5p	1.000						
VRA1	0.316	1.000					

Mor02m	0.494	0.507	1.000				
Lop	0.290	0.815	0.398	1.0000			
BEHp5	0.556	0.638	0.818	0.672	1.000		
MATS8v	0.556	0.033	0.182	0.012	0.225	1.000	
MATS8e	0.610	0.168	0.345	0.291	0.285	0.085	1.000

We have also made efforts to investigate predictive power of the proposed model by using quality factor (QF) considering Pogliani's method. The larger value of QF (3.334 - 2.282) signifies better predictive power of tetra-variant equations. For reliability of the QSAR equation, we have calculated regression associated statistical parameter called probable error of correlation (PE), if the value of correlation coefficient (r) is more than six times of PE than the expression is good and reliable. In model the value of correlation coefficient is significantly higher than 6PE supporting reliability and goodness.

The chance of fortuitous correlation is checked with the help of Y-scrambling data test considering ***Chance*** parameter, which is evaluated as ratio of the equivalent regression equations to the total number of randomized sets. ***Chance*** value of 0.001 corresponds to 0.1% chance of fortuitous correlation. ***Chance*** value (less than 0.001) of equations revealed that the result was not based on prospective correlation. Similarly mean randomized r^2 ($r^2_{RANDMEAN}$) values and randomized standard deviation (S_{RAND}) are also supporting that the results are not based on chance correlation.

Internal predictivity of the model was assured with the help of cross-validated constraints like Q^2, S_{PRESS} and S_{DEP} obtained by 'leave one out (LOO)' cross validation method. The value of $Q^2 > 0.5$ is the basic requirement for declaring a model to be a valid one. The internal consistency of the equations fall within satisfactory level i.e. Q^2 value between 0.875-0.709, S_{PRESS} between 0.377 to 0.575 and S_{DEP} lies within 0.329 to 0.502 range (**Table 2**). Although equations show good internal consistency, the high Q^2 does not imply automatically a high predictive ability of the model. Studies indicated that while high Q^2 is the necessary condition for a model to have a high predictive power, it is not a sufficient condition. Such models may not be applicable for the analogs which were never used in the generation of the correlation. Therefore, the external extrapolation power of the equation was further authenticated by a test set of eleven compounds. The value of predictive squared correlation coefficient (r^2_{pred}) for equation 4, 5 and 6 are found to be -0.301, 0.344 and 0.603 respectively (**Table 2**). On the basis of statistical data and predictive power of test set tetra-variant expressions (Eq^n. 6) which fulfill all the corroboration criteria up to significant echelon was considered as QSAR model. The aforementioned discussion indicated that the regression and statistical parameters are good enough to establish model as predictive model.

The mathematically developed and statistically selected QSAR model revealed that *BEHp5*, *MATS8v* and *MATS8e* are contributing positively while *VRA1* contributed negatively to PPARγ agonist activity.

The best QSAR model having coefficient of correlation (r=0.925) which explain 81.9% variance in the activity (**Table 3**). The linear contribution of each physicochemical parameter to the model was significant by more than 99.0% ($p<0.01$). The model showed overall internal statistical significance level more than 99.9% as it exceeded the tabulated $F_{(4,16\ \alpha\ 0.001)}$ =9.084. The value of the bootstrapping squared correlation coefficient (r^2_{BS} = 0.831) and the bootstrapping standard deviation (S_{BS} = 0.111) implies that the equations were proper representatives of the group of analogs. *Chance* value of 0.001 corresponds to 0.1 % chance of fortuitous correlation. *Chance* value (less than 0.001) of model revealed that the result was not based on prospective correlation. Similarly mean randomized squared correlation coefficient ($R^2_{RANDMEAN}$ = 0.192) and randomized standard deviation (S_{RAND} = 0.118) are also supporting that the results are not based on chance correlation. The model showed good internal consistency in leave out test, the cross validated squared correlation coefficient (Q^2) was found to be 0.709 with S_{PRESS} and S_{DEP} 0.575 and 0.502 respectively (**Table 3 and Fig. 1**). The selected mathematical expression (model) are able to predict the activity of test set compound, which supported by r^2_{pred} (0.603) and standard error of prediction (0.537) values. The predicted activity of test set compounds are very close to their actual activity, which indicate the robustness of model (**Table 3 and Fig. 1**). The contributions of descriptors to the model are shown in **Fig. 2**.

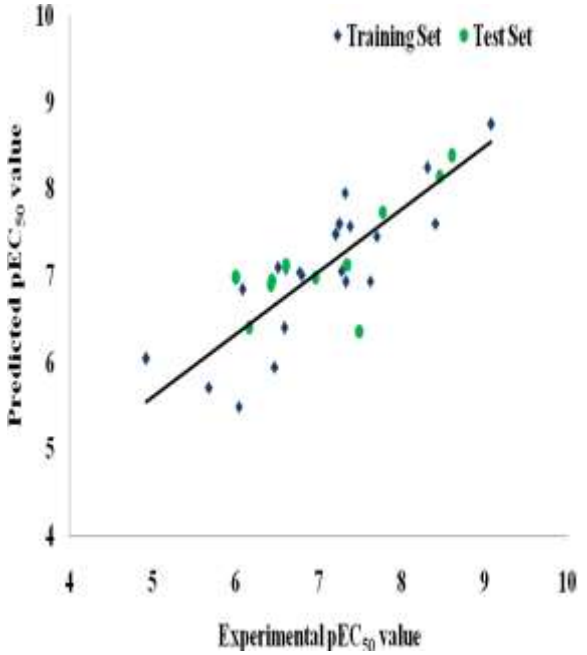


Fig 1: Graphical representation of experimental versus calculated loo pEC_{50} of training set and predicted pEC_{50} of test set

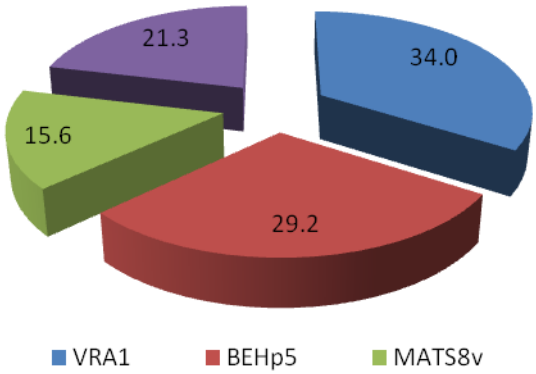


Fig 2: Percentage contribution of descriptors to the model

MATS8e and *MATS8v* belong to 2D-autocorrelation descriptors. It's calculated from the molecular graphs by summing the products of atom weights of the terminal atoms of all the paths of the considered path length (the lag). *MATS8e* Moran autocorrelation - lag 8/ weighted by atomic Sanderson electronegativities while *MATS8v* Moran autocorrelation - lag 8/ weighted by atomic van der Waals volumes.[14]

The Moran autocorrelation descriptors (MATSdw) are given by

$$MATSdw = \frac{\frac{1}{\Delta}\sum_{i=1}^{A}\sum_{j=1}^{A}\delta_{ij}.(w_i-\bar{w}).(w_j-\bar{w})}{\frac{1}{A}\sum_{i=1}^{A}(w_i-\bar{w})^2}$$

Where d is the considered topological distance, A is the atom number, wi and wj are the weights (normalized atomic properties) for atoms i and j respectively. $\bar{w}$ is the average value of property for the molecule δ_{ij}, is a Kronecker delta (δ_{ij},= 1 if δ_{ij} = d, zero otherwise), Δ is the sum of the Kronecker deltas, i.e. the number of vertex pairs at distance equal to d.

MATS8v positively contributed to the agonist activity, reveled that van der Waals volume of a molecule is decisive for explaining the enzyme ligand interaction. Similarly *MATS8e* weighted by atomic Sanderson electronegativities contributed positively, the Sanderson electronegativities allow the estimation of bond energies in a wide range of compounds. The positive contributions suggest that polar substitution is favorable for the activity.

BEHp5 and VRA1 are eigenvalue-based descriptors, in which The VRA indices are defined by applying the Randic operator to the coefficients l_{iA} of the eigenvector associated with the largest negative eigenvalue.

$$VRAl = \sum_b (l_{iA}.\, L_{jA})_b^{-1/2}$$

where the sum runs over all of the bonds in the molecular graph; l_{iA} and L_{jA} are the LOVIs of the two vertices incident to the considered bond. Whereas *BEHp5* belongs to BCUT descriptors, BCUT descriptors (Burden - CAS - University of Texas eigenvalues) are based on a significant extension of the Burden approach, considering three classes of matrices whose diagonal elements correspond to 1) atomic charge-related values, 2) atomic polarizability-related values, and 3) atomic H-bond abilities. Additionally, a variety of definitions were considered for the off-diagonal terms, including functions of interatomic distance, overlaps, computed bond orders, etc. Moreover, for the off-diagonal terms not only was a 2D approach used, but also a 3D approach, to account for geometric interatomic distances.[14] Positive contribution of BEHp5 revealed that 5th heighest eigenvalue of Burden matrix correspond to atomic polarizability is favorable for the PPARγ agonist activity.

Conclusions

In this study, molecular feature based quantification of PPARγ agonist activity have been explored. QSAR results elucidate that the topological distance and eigenvalue-based descriptors affect activities of oxyiminoalkanoic acid analogs towards PPARγ agonist

activity. The values of coefficient of determination and cross validated coefficient of prediction obtained from model are 0.855 and 0.709, respectively. Moreover, test set data showed coefficient of prediction 0.603. The results show that the QSAR model is robust and has good predictive ability. These models are not only able to predict the activity of test compounds but also explained the important structural features of the molecules in a quantitative manner. The study provided useful clues about the structural requirement for effective PPARγ agonist activity and hence for the improvement of the biological activity. In conclusion, the results derived in present study can provide a preliminary valuable guidance for continuing search for potential PPARγ agonist prior to synthesis.

References

1. Tiwari Ashok K. and Rao J. Madhusudana, Diabetes Mellitus and Multiple Therapeutic Approaches of Phytochemicals: Present Status and Future Prospects *Current Science, 2002*, 83, 30-38.
2. Agrawal Y. P., Agrawal M.Y. and Gupta A. K., Design, Synthesis and Evaluation of Rhodanine Derivatives as Aldose Reductase Inhibitors *Chem Biol Drug Des.*, 2015, 85, 172-180.
3. Bastaki Salim, Diabetes mellitus and its treatment, *Int J Diabetes & Metabolism*, 2005, 13, 111-134.
4. O'Moore-Sullivan T. M., Prins J. B., *Med. J. Aust.*, 2002, 176, 381-386.
5. Imoto H., Imamiya E., Momose Y., Sugiyama Y., Kimura H., Sohda T., Chem. Pharm. Bull., 50, 1349—1357 (2002).
6. Imoto H, Sugiyama Y, Kimura H and Momose Y., Studies on Non-Thiazolidinedione Antidiabetic Agents. 2.1) Novel Oxyiminoalkanoic Acid Derivatives as Potent Glucose and Lipid Lowering Agents *Chem. Pharm. Bull.* 2003, 51(2), 138-151.
7. CS Chem Office, Version 11.0,Cambridge Soft Corporation, Software Publishers Association, 1730 MStreet, Suite700, Washington D.C.20036,USA.Tel.:(202) 452-1600.
8. R. Todeschini, V. Consonni, DRAGON-Software for the Calculation of Molecular Descriptors, rel.1.12forWindows(2001).
9. Gupta A. K., B. M. Arockia and S. G. Kaskhedikar, VALSTAT: Validation Program for Quantitative Structure Activity Relationship Studies. Indian J. Pharm.Sci.66 (2004) 396-402.
10. Gupta Arun Kumar, Sabarwal Neetu, Agrawal Yogesh P., Prachand Sumeet, Jain Sanjay, Insights through AM1 calculations into the structural requirement of 3,4,6-substituted-2-quinolone analogs towards FMS kinase inhibitory activity, European Journal of Medicinal Chemistry, 2010, 45, 3472- 3479.
11. Gupta A. K., Singh P. and Sabarwal N, Rationalization of Molecular Descriptors of Aurone Analogs toward Anti-malarial Activity, *Asian J Pharm Clin Res,* 2014, 7, 186-192.
12. Gupta Revathi A, Kaskhedikar S.G., Investigation of the Structural Requirement for 3, 5-disubstituted 4, 5 dihydroisoxazole analogs against *M. tuberculosis*: QSAR Study, International Journal of Drug Design and Discovery, 2010, 1, 81-92.
13. Gupta RA, Gupta AK, Soni LK, Kaskhedikar SG 2-(pyrazin-2-yloxy)acetohydrazide analogs QSAR study: an insight into the structural basis of antimycobacterial activity. Chem Biol Drug Des 2010; 76: 441-450.
14. Todeschini R, Consonni V (2000) Handbook of molecular descriptors, vol 11. Wiley-VCH, Weinheim.

How to cite this article
Pandey J. and Gupta A.K. (2016). Molecular Descriptors Analysis of Oxyiminoalkanoic Acid Analogs towards PPARγ Agonist Activity. *Int. J. Pharm. Life Sci.*, 7(3):4942-4951.
Source of Support: Nil; Conflict of Interest: None declared

Received: 31.01.16; Revised: 15.02.16; Accepted: 10.03.16