Academic Sciences Asian Journal of Pharmaceutical and Clinical Research
Vol 6, Suppl 1, 2013 ISSN - 0974-2441

Research Article

QSAR MODELING OF ALDOSE REDUCTASE INHIBITORY ACTIVITY OF FLAVONOID COMPOUNDS USING ELECTROTOPOLOGICAL STATE ATOM PARAMETER (E-STATE)

SUPRATIM RAY, ARNAB KUMAR MANNA

Division of Pharmaceutical Chemistry, Dr. B C Roy College of Pharmacy & Allied Health Sciences, Bidhannagar, Durgapur, 713 206, India,Email: supratimray_in@yahoo.co.in

Received:2 December 2012, Revised and Accepted:7 january 2013

ABSTRACT

This study gives a quantitative structure activity relationship (QSAR) correlation of aldose reductase inhibitory activity of seventy flavonoid compounds. The study was performed using electrotopological state atom (E-state) parameter as descriptors. Partial least squares analysis (PLS) is used as chemometric tool. The model indicates the importance of hydroxyl group at various positions of the flavonoid moiety. Presence of methoxy groups attached to the moiety at specific positions is beneficial for aldose reductase inhibitory activity.

Keywords: QSAR, E-state, PLS, Aldose reductase, Flavonoid

INTRODUCTION

Aldose reductase (AR) plays an important role in cataract formation which is produced in diabetic patients due to over expression of aldose reductase. Normally in presence of reduced nicotinamide-adenine-dinuclotide phosphate (NADPH) as cofactor, the enzyme reduces aldose sugar to their alcohol as exemplified by glucose to sorbitol and sorbitol dehydrogenase, oxidizes sorbitol to fructose [1, 2]. But in diabetic condition the balance between sorbitol production and conversion of sorbitol to fructose is disturbed. Excess of sorbitol is produced which accumulated in lens, nerve and retina provokes a hyper-osmotic effect causing lens swelling and opacities that ultimate leads to cataract formation [3].

Flavonoids are a group of naturally occurring polyphenolic compounds ubiquitously found in fruits and vegetables [4-6]. Chemically flavonoids are benzo-γ-pyrone derivatives. Common family members of flavonoids include flavones, flavonols, flavanones, isoflavones, biflavanones, catechins and anthocyanidins. Structural diversity of flavonoids allows them to exhibit antineoplastic, antihepatitis, antibacterial, anti-inflammatory, antimutagenic, antiallergic, antithrombic, antiviral and vasodilatory activities [7-9]. The potent antioxidant activity of flavonoids, their ability to scavenge hydroxyl radicals, superoxide anions and lipid peroxy radicals could be the most important function of flavonoids and underlie many of the above processes in the body [10].

The inhibitory effect of aldose reductase was found in several structurally diverse classes of compounds like tetramethyleneglutaric acid, flavone coumarin, xanthine, naphthalene, flavone, quinazoline derivatives etc [11-14]. But flavonoids derivatives were found to be more potent. A quantitative structure–activity relationship (QSAR) study on a data set of inhibitory activities against AR enzyme of 75 flavonoids was reported using multi linear regression analysis with classical and quantum chemical descriptors was reported [15]. Another study was performed using several types of descriptors using artificial neural network as chemometric tool [16]. In the present work we have modeled the aldose reductase inhibitory activity of 70 flavonoids [17-19] compounds using electrotopological state atom (E-state) parameters by partial least squares.

MATERIALS AND METHODS

Electrotopological state atom (E-state) index

Structural specificity of a drug molecule is exhibited at an atomic or fragmental level instead of the whole molecule. In the drug receptor interaction phenomenon, a portion of the molecule (pharmacophore) may play more important role than the other segments. Though basic information for constitution of topological indices are derived from the atom level (count of atoms, bonds, paths of bonds, etc.), most of the indices are applied to the whole molecule after summing up all components over the whole molecule. Thus QSAR studies at the atomic or fragmental level are justified in the present context [20].

The electrotopological state atom (E-state) index developed by Hall and Kier [21] is an atom level descriptor encoding both the electronic character and topological environment of each skeletal atom in a molecule. The E-state of a skeletal atom is formulated as an intrinsic value I_i plus a perturbation term ΔI_i, arising from the electronic interaction within the molecular topological environment of each atom in the molecule.

The intrinsic value has been defined as the ratio of a measure of electronic state (Kier-Hall valence state electronegativity) to the local connectedness. The count of valence electrons which are the most reactive and involved in chemical reactions and bond formations are considered in the expression of *I* to encode the electronic feature. To reflect differences in electronegativity among the atoms, principal quantum number is employed in the expression of *I*. The topological attribute is included by using adjacency count of atom. The intrinsic value of an atom *i* is defined as

$$I_i = \left[\left(2/N\right)^2 \delta^v + 1\right]/\delta \qquad \text{(i)}$$

In Eq. (i), N stands for principal quantum number and δ^v and δ indicate the count of valence electrons and sigma electrons associated with the atom *i* in the hydrogen suppressed graph. The intrinsic electrotopological state calculated according to Eq. (i) produces different values of an atom in different degrees of substitution (branching). The values are also different for different atoms having differences in electronegativity. The intrinsic values increase with increase in electronegativity or electron-richness and decrease with increase in branching (substitution).

The perturbation factor for the intrinsic state of atom *i* is defined as

$$\Delta I_i = \sum_{j \neq i} \frac{I_i - I_j}{r_{ij}^2} \qquad \text{(ii)}$$

In Eq. (ii) r_{ij} stands for the graph separation factor, i.e., count of skeletal atoms in the shortest path connecting the atoms *i* and *j* including both atoms.

Summation of intrinsic state of an atom and influence of the field is called electrotopological state of the atom.

$$S_i = I_i + \sum_{j \neq i} \Delta I_{ij} \qquad \text{(iii)}$$

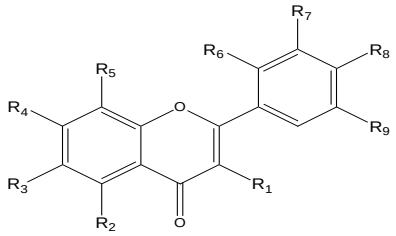
It is a representation of molecular structure information as it varies with changes in structural features including branching, cyclicity, homologation, heteroatom variation, and changes in relative

positions of different groups. The electrotopological state considers both bonded and non-bonded interactions: the bonded component depends simply on differences in electronegativity among the adjacent atoms. The non-bonded interactions may be through inductive effect across the skeleton and is a function of graph separation factor and electronegativity differences. Thus, electrotopological state represents electronic distribution information modified by both local and global topology. The information encoded in the E-state value for an atom is the electronic accessibility at that atom.

Data treatment and software

The inhibitory effects of flavonoid compounds against aldose reductase reported in literatures [17-19] were used as the model data-set for the present QSAR analysis (Table 1). The reported activity [log ($1/IC_{50}$)] was used for QSAR analysis. The QSAR analysis was performed using electrotopological state atom (E-state) parameter. The whole data set seventy compounds and all the compounds contain 17 common atoms (excluding hydrogen). The atoms of the molecules were numbered keeping serial numbers of the common atoms same in all the compounds (as shown in Figure 1). The electrotopological states of the 17 common atoms for all of the compounds were found out using a VISUAL BASIC program SRETSA developed partly by the author [22]. The program uses, as input, only the connection table in a specific format along with intrinsic state values of different atoms. To the output file thus obtained, the biological activity data were introduced to make it ready for subsequent regression analysis.

Table 1: Molecular scaffolds of the compounds along with their activity



Sl No	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	Log $1/IC_{50}$
1	-OCH_3	-OH	-OCH_3	-OH	H	H	-OH	-OH	H	7.52
2	H	-OCH_3	-OCH_3	-OCH_3	-OCH_3	H	-OH	-OH	H	7.49
3	H	-OCH_3	-OH	-OCH_3	-OCH_3	H	-OH	-OH	H	7.47
4	H	-OH	-OCH_3	-OH	-CH_2Ph	H	-OH	-OH	H	7.47
5	H	-OH	-OCH_3	-OCH_3	-OCH_3	H	-OH	-OH	H	7.41
6	H	-OCH_3	H	-OCH_3	-OCH_3	H	-OH	-OH	H	7.35
7	-OCH_3	-OH	-OH	-OH	H	H	-OH	-OH	H	7.24
8	H	-OH	-OH	-OCH_3	-OCH_3	H	-OH	-OH	H	7.19
9	H	-OCH_3	H	-OH	-OCH_3	H	-OH	-OH	H	7.13
10	H	-OH	H	-OCH_3	-OCH_3	H	-OH	-OH	H	7.11
11	H	-OCH_3	-OCH_3	-OCH_3	H	H	-OH	-OH	H	7.04
12	H	-OH	-OH	-OH	-OCH_3	H	-OH	-OH	H	6.92
13	H	-OCH_3	-OH	-OCH_3	H	H	-OH	-OH	H	6.85
14	H	-OCH_3	-OCH_3	-OCH_3	-OCH_3	H	H	-OH	H	6.79
15	H	-OCH_3	H	-OCH_3	OH	H	-OH	-OH	H	6.79
16	-OCH_3	-OCH_3	H	-OCH_3	-OCH_3	H	-OH	-OH	H	6.77
17	H	-OH	-OH	-OH	H	H	-OH	-OH	H	6.69
18	H	-OH	-OCH_3	-OCH_3	H	H	-OH	-OH	H	6.66
19	H	-OH	H	-OCH_3	-OH	H	-OH	-OH	H	6.64
20	-OCH_3	-OH	H	-OH	-OCH_3	H	-OH	-OH	H	6.62
21	H	-OCH_3	-OH	-OCH_3	-OCH_3	H	H	-OH	H	6.6
22	H	-OCH_3	-OCH_3	-OCH_3	H	H	-OH	-OH	H	6.57
23	H	-OH	H	-OH	-OCH_3	H	-OH	-OH	H	6.55
24	-OCH_3	-OCH_3	H	-OH	-OCH_3	H	-OH	-OH	H	6.55
25	H	-$OCOCH_3$	-$OCOCH_3$	--$OCOCH_3$	-OCH_3	H	-$OCOCH_3$	-$OCOCH_3$	H	6.52
26	H	-OH	-OH	-OCH_3	H	H	-OH	-OH	H	6.52
27	-OCH_3	-OCH_3	-OH	-OCH_3	H	H	-OH	-OH	H	6.52
28	-OCH_3	-OH	-OCH_3	-OCH_3	H	H	-OH	-OH	H	6.46
29	H	-OH	-OCH_3	-OH	-OCH_3	H	H	-OH	H	6.39
30	H	-OH	-OCH_3	-OCH_3	-OCH_3	H	H	-OH	H	6.27
31	-OCH_3	-OH	-OH	-OCH_3	H	H	-OH	-OH	H	6.09
32	-OH	-OH	H	-OH	H	H	-OH	-OH	H	6.09
33	H	-OH	-OH	-OCH_3	-OCH_3	H	H	-OH	H	6.07
34	H	-OH	-OH	-OH	-OCH_3	H	H	-OH	H	5.92
35	H	-OH	-OH	-OH	-OCH_3	H	-OCH_3	-OH	H	5.92
36	H	-OH	-OCH_3	-OCH_3	H	H	H	-OH	H	5.85
37	-O-Rh	-OH	H	-OH	H	H	-OH	-OH	H	5.69
38	H	-OH	-OCH_3	-OH	-OCH_3	H	-OCH_3	-OH	H	5.35
39	H	-OCH_3	-OH	-OCH_3	-OCH_3	H	-OCH_3	-OH	H	5.2
40	H	-OH	-OCH_3	-OCH_3	H	H	-OCH_3	-OH	H	5.17
41	H	-OH	-OCH_3	-OH	-OCH_3	H	H	-OCH_3	H	5.14
42	H	-OH	-OH	-OH	-OCH_3	H	H	H	H	5.09
43	H	-OH	-OH	-OCH_3	-OCH_3	H	H	H	H	5.08
44	-$COCH_3$	-OCH_3	-OCH_3	-OCH_3	H	H	-OH	-OH	H	5.05
45	H	-OH	-OCH_3	-OCH_3	H	H	-OH	-O-Glc	H	5.02
46	H	H	H	H	H	H	H	-OH	H	5

47	H	-OH	$-OCH_3$	$-OCH_3$	H	H	$-OCH_3$	-O-Glc	H	4.92
48	H	-OH	$-OCH_3$	$-OCH_3$	H	H	H	-O-Glc	H	4.88
49	-O-Glc	-OH	H	-OH	H	H	-OH	-OH	H	4.79
50	H	-OH	$-OCH_3$	-OH	$-OCH_3$	H	$-OCH_3$	-O-Glc	H	4.78
51	H	$-OCH_3$	$-OCH_3$	$-OCH_3$	$-OCH_3$	H	$-OCH_3$	-OH	H	4.74
52	H	-OH	$-OCH_3$	-O-Glc	$-OCH_3$	H	$-OCH_3$	-OH	H	4.73
53	-Ph	H	H	$-OCH_3$	H	H	H	-OH	H	4.68
54	H	-OH	$-OCH_3$	-OH	$-OCH_3$	H	$-OCH_3$	$-OCH_3$	H	4.67
55	-Ph	H	H	H	H	H	H	-OH	H	4.53
56	-OH	H	$-OCH_3$	H	H	H	H	H	H	4.48
57	-CN	H	H	H	H	H	H	H	H	4.48
58	-COOH	H	H	H	H	H	H	H	H	4.48
59	H	-OH	$-OCH_3$	$-OCH_3$	$-OCH_3$	H	$-OCH_3$	-OH	H	4.42
60	-OH	H	H	H	H	H	H	H	H	4.34
61	-COOH	$-OCH_3$	H	H	-COOH	H	H	H	H	4.34
62	$-OCH_3$	H	H	$-OCH_3$	H	H	H	-OH	H	4.25
63	H	H	H	$-OCH_3$	H	H	H	-OH	H	4.15
64	-OH	-OH	H	-OH	H	H	$-OCH_3$	-OH	H	4
65	H	H	H	H	H	H	H	-OH	H	4
66	-CH3	H	H	H	H	H	H	-OH	H	4
67	H	$-OCH_3$	-OH	$-OCH_3$	$-OCH_3$	H	H	H	H	3.54
68	H	-OH	H	$-OCH_3$	H	$-OCH_3$	H	$-OCH_3$	OH	3.5
69	H	$-OCH_3$	H	-OH	H	H	H	H	H	3
70	H	-OH	H	$-OCH_3$	H	$-OCH_3$	H	-OH	$-OCH_3$	3

Glc=glucose; Rh=rhamnose

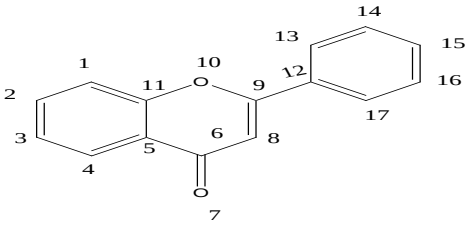


Figure 1: Common atom of the compounds

Model development

To begin the model development process, the whole data set (n=70) was divided into training (n=53, 75% of the total number of compounds) and test (n=17, 25% of the total number of compounds) sets by *k*-means clustering technique [23] applied on standardized descriptor matrix of the E-state parameters. QSAR models were developed using the training set compounds (optimized by Q^2), and then the developed models were validated (externally) using the test set compounds.

PLS

PLS is a generalization of regression, which can handle data with strongly correlated and/or noisy or numerous X variables [24, 25]. It gives a reduced solution, which is statistically more robust than MLR. The linear PLS model finds "new variables" (latent variables or X scores) which are linear combinations of the original variables. To avoid over fitting, a strict test for the significance of each consecutive PLS component is necessary and then stopping when the components are non significant. Application of PLS thus allows the construction of larger QSAR equations while still avoiding over fitting and eliminating most variables. PLS is normally used in combination with cross validation to obtain the optimum number of components. This ensures that the QSAR equations are selected based on their ability to predict the data rather than to fit the data. In case of PLS analysis on the present data set, based on the standardized regression coefficients, the variables with smaller coefficients were removed from the PLS regression until there was no further improvement in Q^2 value irrespective of the components. The PLS analysis was performed using statistical software MINITAB [26].

Statistical parameters

The statistical qualities of various equations were judged by calculating several metrics namely squared correlation variance (R^2), explained variance (R_a^2), standard error of estimate (s) and variance ratio (F) at specified degrees of freedom (df) [27]. Internal validation parameters like Q^2_{int} as well as $r^2_{m(LOO)}$ [28], external validation parameters like $Q^2_{ext(F1)}$, $Q^2_{ext(F2)}$ [29, 30], $r^2_{m(test)}$ [28] and overall validation parameters $r^2_{m(overall)}$ [28] were also reported.

External validation

The statistically internally optimized models were further evaluated for their real predictive power.

$Q^2_{ext(F1)}$ is calculated according to the following formula

$$Q^2_{ext(F1)} = 1 - \frac{\sum (Y_{obs} - Y_{cal})^2}{\sum (Y_{obs} - Y_{training})^2}$$

$Y_{training}$ Means mean activity value of the training set while Y_{obs} and Y_{cal} represent observed and calculated activity values.

$Q^2_{ext(F2)}$ is calculated according to the following formula

$$Q^2_{ext(F2)} = 1 - \frac{\sum (Y_{obs} - Y_{cal})^2}{\sum (Y_{obs} - Y_{test})^2}$$

Y_{test} Means mean activity value of the test set.

An additional parameter which penalizes a model for large differences between observed and predicted values of the prediction set compounds, as well as independent of the mean of training and prediction set, was also calculated for model external predictivity. The expression of r^2_m is defined as:

$$r_m^2 = r^2(1-\sqrt{r^2-r_0^2})$$

Where r^2 and r_0^2 are determination coefficients of linear relations between the observed and predicted values of the compounds with and without intercept respectively. The r_m^2 is applied for test set ($r^2_{m(test)}$), training set ($r^2_{m(LOO)}$) and the overall set ($r^2_{m(overall)}$).

RESULTS AND DISCUSSION

Membership of compounds in different clusters generated using *k*-means clustering technique is shown in Table 2. The test set size was set to approximately 25% to the total data set size [23] and the test set members along with their observed and calculated activity are given in Table 3.

Table 2: *k*-Means clustering of compounds using standardized descriptors

Cluster No.	No. of compounds in different clusters	Compounds (Sl nos.) in each clusters																
1	10	1	6	24	36	37	50	54	57	11	41							
2	11	4	7	12	31	32	52	56	60	8	38	62						
3	7	18	19	63	64	65	55	66										
4	17	16	22	30	33	34	43	44	45	49	51	59	69	70	27	40	47	61
5	13	14	21	28	29	35	42	46	48	58	67	25	39	53				
6	12	2	3	9	10	13	17	20	23	68	5	15	26					

Table 3: Observed and calculated activity from PLS model

Sl. No.	Observed aldose reductase inhibitory activity Log(1/IC_{50})	Calculated activity
	Training Set	
1	7.52	5.28975
2	7.49	6.80282
3	7.47	6.94836
4	7.47	6.49323
6	7.35	6.49937
7	7.24	6.57391
9	7.13	6.56609
10	7.11	6.29316
12	6.92	6.99172
13	6.85	6.54434
14	6.79	5.77346
16	6.77	5.16328
17	6.69	6.57344
18	6.66	6.25748
19	6.64	6.45781
20	6.62	6.42411
21	6.6	5.9688
22	6.57	6.40654
23	6.55	6.2783
24	6.55	7.10528
28	6.46	6.32939
29	6.39	5.82561
30	6.27	5.67988
31	6.09	6.52929
32	6.09	6.135
33	6.07	5.78518
34	5.92	6.118
35	5.92	5.5064
36	5.85	5.22491
37	5.69	6.18206
42	5.09	4.45446
43	5.08	4.25352
44	5.05	6.61885
45	5.02	5.80204
46	5	4.09677
48	4.88	4.55601
49	4.79	4.76661
50	4.78	5.98351
51	4.74	6.57025
52	4.73	6.26274
54	4.67	5.98823
56	4.48	5.08061
57	4.48	4.38442
58	4.48	4.58927
59	4.42	6.29257
60	4.34	4.46052

63	4.15	4.14434
64	4	5.89769
65	4	4.14508
67	3.54	4.79935
68	3.5	3.27716
69	3	4.01685
70	3	4.18694
	Test Set	
5	7.41	6.653537
8	7.19	6.850522
11	7.04	6.478228
15	6.79	6.60105
25	6.52	6.792135
26	6.52	6.432422
27	6.52	6.704153
38	5.35	6.140019
39	5.2	6.580494
40	5.17	5.843587
41	5.14	5.29672
47	4.92	5.220972
53	4.68	4.016358
55	4.53	4.103405
61	4.34	4.595067
62	4.25	4.29963
66	4	4.150476

Observed activity from ref (17, 18 & 19)
Calculated activity from eq. (1)

The number of optimum components was 3 to obtain the final equation (optimized by cross validation). Based on the standardized regression coefficients, the following variables were selected for the final equation:

$$\log(1/IC_{50}) = 9.3098\text{-}0.2586S_1\text{-}0.4081S_3\text{+}0.547S_4 + 0.4325S_5 - 0.2868S_{14} - 0.727C_{15} + 1.3756S_{16} - 3.0245S_{17}$$

$$R^2 = 0.594, R_a^2 = 0.569, PRESS = 40.96, F = 23.9(df = 3,52),$$

$$Q_{\text{int}}^2 = 0.5078, n_{training}=53, r_{m(LOO)}^2 = 0.487, Q_{ext(F1)}^2=0.7626,\ Q_{ext(F2)}^2=0.7620,$$

$$n_{\text{test}} = 17,\ r_{m(test)}^2=0.7269,\ r_{m(\text{overall})}^2 = 0.535$$

(1)

Eq. (1) could explain 56.9% of the variance (adjusted coefficient of variation) and leave – one – out predicted variance was found to be 50.78%.

The positive coefficient of S_4, S_5 and S_{16} indicates that aldose reductase inhibitory activity increases with increase in E-state value of atom 4, 5 and 16 respectively. Compounds with high values of E-state parameter for atom 4 (S_4) (like 65, 63 and 55) showed comparatively higher activity. Position 4 indicates the importance of hydroxyl / methoxy group necessary for activity. The higher value of E-state shows higher activity (like in compounds 65 and 63) than compounds (like 1 and 7) having lower value of E-state for atom 5 (S_5). Higher active compounds (like 63, 65 and 69) possessing higher E state value indicate that no substitution is required for position 16.

The negative coefficients of S_1, S_3, S_{14}, S_{15} and S_{17} indicate that aldose reductase inhibitory activity increases with decrease in E-state value of atoms 1, 3, 14, 15 and 17 respectively. Compounds with low values of E-state parameter for atom 1 (S_1) (like 54, 61 and 67), for atom 3 (S_3) (like 50, 59 and 67), for atom 14 (S_{14}) (like 44 and 64), for atom 15 (S_{15}) (like 64 and 70), for atom 17 (S_{17}) (like 64, 68 and 70) showed comparatively better activity. Substituents containing oxygen like methoxy, carboxylic acid and hydroxyl group at position 1, 3, 14 and 15 is positively contributed towards activity. But for position 17 no substitution is required for activity. The statistical quality of the model showed acceptable internal validation (Q_{int}^2 =0.5078 and $r_{m(LOO)}^2$ =0.487), external validation ($Q_{ext(F1)}^2$ =0.7626, $Q_{ext(F2)}^2$ =0.762, $r_{m(test)}^2$ =0.7269), and overall validation ($r_{m(\text{overall})}^2$ =0.535).

OVERVIEW AND CONCLUSIONS

The whole dataset (n=70) was divided into a training set (53 compounds) and a test set (17 compounds) based on *k*-means clustering of the standardized descriptor matrix and model was developed from the training set (optimized by Q^2). The predictive ability of the models was judged from the prediction of the activity of the test set compounds. The model indicates the importance of hydroxyl group at various positions (like position 4 and 14 etc) of the moiety. Presence of methoxy groups attached to the moiety at positions 1, 4, 3 and 14 is beneficial for aldose reductase inhibitory activity. The model also indicates that no substitution is required for position 16 and 17 respectively.

REFERENCES

1. Verma SD, Kinoshita JH. Inhibition of lens aldose reductase by flavonoids--their possible role in the prevention of diabetic cataracts. Biochem Pharmacol 1976; 25: 2505-2513.
2. Verma SD, Mikuni I, Kinoshita JH. Flavonoids as inhibitors of lens aldose reductase. Science 1975; 188: 1215-1216.
3. Kinoshita JH. Mechanisms initiating cataract formation. Invest Ophthalmol 1974; 13: 713-724.
4. Carlo GD, Mascolo N, Izzo AA, Capasso F. Flavonoid: old and new aspects of a class of natural therapeutic drugs. Life Sci 1999; 65: 337-353.
5. Haollman PCH, Arts ICW. Flavonols, flavones and flavanols-nature, occurrence and dietary burden. J Sci Food Agric 2000; 80: 1081-1093.
6. Aherne SA, O'Brien NM. Dietary Flavonols: Chemistry, Food Content, and Metabolism. Nutrition 2002; 18: 75-81.
7. Agarwal OP. The anti-inflammatory action of nepitrin, a flavonoid. Agents Actions 1982; 12: 298- 302.
8. Church MK. Cromoglycate-like anti-allergic drugs: a review. Drugs Today 1978; 14: 281-341.
9. Rasulev BF, Abdullaev ND, Syrov VN, Leszczynski J. A quantitative structure-activity relationship (QSAR) study of the antioxidant activity of flavonoids. QSAR & Comb Sci 2005; 24: 1056-1065.
10. Ozawa T. Genetic and functional changes in mitochondria associated with aging. Physiol Rev 1997; 77: 425-464.
11. Kador PF, Sharpless NE. Structure-activity studies of aldose reductase inhibitors containing the 4-oxo-4h-chromen ring system. Biophys Chem 1978; 8: 81-85.
12. Kador PF, Kinoshita JH, Sharpless NE. Aldose reductase inhibitors: a potential new class of agents for the

pharmacological control of certain diabetic complications. J Med Chem 1985; 28: 841-849.
13. Pfister JR, Wymann WE, Mahoney JM, Waterbury LD. Synthesis and aldose reductase inhibitory activity of 7-sulfamoylxanthone-2-carboxylic acids. J Med Chem 1980; 23: 1264-1267.
14. Malamas MS, Millen J. Quinazolineacetic acids and related analogues as aldose reductase inhibitors. J Med Chem 1991; 34: 1492-1503.
15. Stefanic-Petek A, Krbavcic A, Solmajer T. QSAR of flavonoids: 4. Differential inhibition of aldose reductase and $p56^{lck}$ protein tyrosine kinase. Croat Chem Acta 2002; 75: 517-529.
16. Fernandez M, Caballero J, Helguera AM, Castroe EA, Gonzalezd MP. Quantitative structure activity relationship study to predict differential inhibition of aldose reductase by flavonoid compounds. Bioorg & Med Chem 2005; 13: 3269–3277.
17. Okuda J, Miwa I, Inagaki K, Horie T, Nakayama M. Inhibition of aldose reductases from rat and bovine lenses by flavonoids. Biochem Pharmacol 1982; 31: 3807-3822.
18. Inagaki K, Miwa I, Yashiro T, Okuda J. Inhibition of aldose reductases from rat and bovine lenses by hydantoin derivatives. Chem Pharm Bull 1982; 30: 3244-3254.
19. Okuda J, Miwa I, Inagaki K, Horie T, Nakayama M. Inhibition of aldose reductase by 3',4'-dihydroxyflavones. Chem Pharm Bull 1984; 32: 767-772.
20. Hall LH, Mohney B, Kier LB. The Electrotopological State: An Atom Index for QSAR. Quant Struct- Act Relat 1991; 10: 43-51.
21. Kier LB, Hall LH. An Electrotopological State Index for Atoms in Molecules. Pharm Res 1990; 7: 801-807.
22. SRETSA is statistical software in Visual Basic, developed by Ray S, Biswas R. and standardized using known data sets.
23. Leonard JT, Roy K. On Selection of Training and Test Sets for the Development of Predictive QSAR models. QSAR Comb Sci 2006; 25: 235-251.
24. Wold S. PLS for multivariate linear modeling. In: Van de Waterbeemd H, editor. Chemometric Methods in Molecular Design (Methods and Principles in Medicinal Chemistry). Weinheim: VCH; 1995. p. 195-218.
25. Fan Y, Shi LM, Kohn KW, Pommier Y, Weinstein JN. Quantitative structure-antitumor activity relationships of camptothecin analogs: Cluster analysis and genetic algorithm-based studies. J Med Chem 2001; 44: 3254-3263.
26. MINITAB is statistical software of Minitab Inc, USA, http://www.minitab.com.
27. Snedecor, G.W.; Cochran, W.G. *Statistical Methods*, Oxford & IBH Publishing Co. Pvt. Ltd.: New Delhi, 1967.
28. Roy PP, Roy K. On some aspects of variable selection for partial least squares regression models. QSAR & Comb Sci 2008; 27: 302-313.
29. Hawkins DM. The problem of overfitting. J Chem Inf Comp Sci 2004; 44, 1-12.
30. Schuurmann G, Ebert RU, Chen J, Wang B, Kuhne R. External validation and prediction employing the predictive squared correlation coefficient-test set activity mean vs training set activity mean. J Chem Inf Model 2008; 48, 2140-2145.