Int J Pharm Bio Sci 2012 July; 3(3): (P) 543 - 550

Research Article **Medicinal chemistry**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

A THEORETICAL STUDY OF 1, 1-DIPHENYL-2-PICRYLHYDRAZYL (DPPH) RADICAL SCAVENGING ACTIVITIES OF FLAVONOIDS USING ELECTROTOPOLOGICAL STATE ATOM (E-STATE) PARAMETERS

SUPRATIM RAY

Division of Pharmaceutical Chemistry, Dr. B C Roy College of Pharmacy & Allied Health Sciences, Bidhannagar, Durgapur, 713 206, India

ABSTRACT

This study gives a quantitative structure activity relationship (QSAR) correlation of the thirty five flavonoid derivatives having 1, 1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activities reported by Seyoum et al. The study was performed using electrotopological state atom (E-state) parameter as descriptors. Stepwise regression analysis was used as a chemometric tool. The predictive ability of the model was judged considering internal (Q^2) and external validation (R^2_{pred}) (Q^2= 0.581, R^2_{pred}=0.7284). The developed model indicates the importance of heteroatom oxygen and it is negatively contributed towards activity. Presence of methoxy group in the flavonoid nucleus is detrimental towards activity.

KEY WORDS: QSAR, E-state, Stepwise regression, Flavonoids

SUPRATIM RAY
Division of Pharmaceutical Chemistry, Dr. B C Roy College of Pharmacy & Allied Health Sciences, Bidhannagar, Durgapur, 713 206, India

INTRODUCTION

Chemically, flavonoids are benzo-γ-pyrone derivatives. Common family members of flavonoids include flavones, flavonols, flavanones, isoflavones, biflavanones, catechins and anthocyanidins. Flavonoids are a group of naturally occurring polyphenolic compounds ubiquitously found in fruits and vegetables [1, 2]. Structural diversity of flavonoids allows them to exhibit antioxidant, antineoplastic, antihepatitis, antibacterial, anti-inflammatory, antimutagenic, antiallergic, antithrombic, antiviral and vasodilatory activities [3-5]. 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 [6].The number of flavonoids derivatives is more than 4000 and their antioxidant properties are very different. It is a complex task to select the most effective antioxidants from a large number of flavonoids [7]. Because of their great number and positive biological effects, flavonoids are popular subjects for Quantitative structure-activity relationship (QSAR) studies. In the present paper QSAR study was performed on the 1, 1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activities of flavonoids data set reported by Seyoum et al. [8] using electrotopological state atom (E-state) parameters by stepwise regression analysis.

COMPUTATIONAL

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 [9]. The electrotopological state atom (E-state) index developed by Hall and Kier [10] 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. (1) 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

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

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.

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

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. The present communication will show here the utility of E-state parameters in QSAR studies by exploring QSAR of 1, 1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activities of flavonoids data set reported by Seyoum et al. [8] using electrotopological state atom (E-state) parameters by stepwise regression analysis.

The Data-set and descriptors

The 1, 1-diphenyl-2-picrylhydrazyl (DPPH) radical scavenging activities of flavonoids data set reported by Seyoum et al. [8] were used as the model data-set for the present QSAR analysis (Table 1). The reported IC_{50} values of the compounds were in μM range which was converted to mM range and then to logarithmic scale [log (10^3 / IC_{50})]. The QSAR analysis was performed using electrotopological state atom (E-state) parameter. The whole data set contain thirty five 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 [11]. 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.

Figure 1
Common atom of the molecules

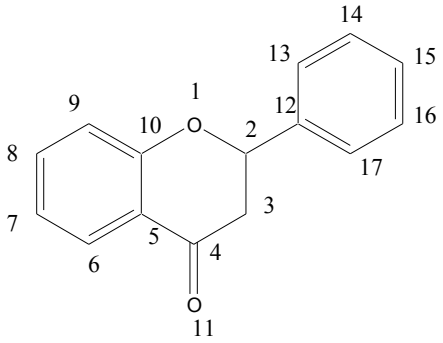
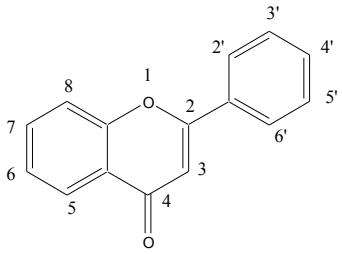
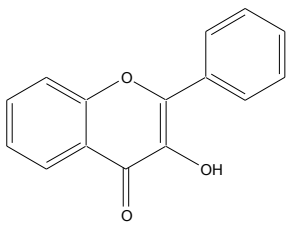
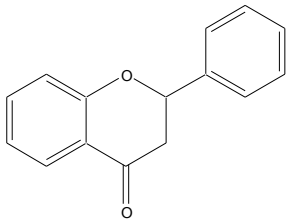


Table 1
Molecular scaffolds of the compounds along with their DPPH radical scavenging activity

Type of Flavonoid	Sl No.	Name of the compound	Substituents	IC_{50}	pC=log(1000/ IC_{50})
Flavone 	1	Luteolin-5-O-glucoside	5-O-glucoside, 7, 3′, 4′-OH	5.73	2.241845
	2	Luteolin	5, 7, 3′, 4′-OH	11.04	1.957031
	3	7, 8-Dihydroxyflavone	7, 8-OH	15.5	1.809668
	4	Cosmosin	7-O-glucoside-5,4′-OH	85.67	1.067171
	5	Luteolin-7-O-glucoside	7-O-glucoside, 5, 3′, 4′-OH	28.17	1.550213
	6	8-Hydroxyacacetin	8-OH	20.28	1.692932
	7	4′-Methoxy-3,6,8-trichloro-5,7-dihydroxyflavone	3,6,8-Cl, 5,7-OH, 4′-OMe	201.52	0.695682
	8	8-Hydroxyflavone	8-OH	166.43	0.778768
	9	5,7-Dihydroxy-3′,4′-dimethoxyflavone	5,7-OH, 3′,4′-OMe	313.18	0.504206
	10	Diosmin	7-O-rutinoside, 5, 3′-OH, 4′-OMe	442.26	0.354322
	11	Apigenin	5,7,4′-OH	463.4	0.334044
	12	Diosmetin	5,7,3′-OH, 4′-OMe	465.13	0.332426
	13	Chrysin	5,7-OH	492.57	0.307532
	14	5-Hydroxy-3′, 4′, 7-trimethoxyflavone	5-OH, 3′, 4′, 7-OMe	539.84	0.267735
	15	7-Hydroxy-5-methyl-4′-methoxyflavone	7-OH, 5-Me, 4′-OMe	808.71	0.092207
	16	Vicenin-2	6,8-C-glucoside, 5,7,4′-OH	171.28	0.766293
	17	Acacetin	5,7-OH, 4′-OMe	529.8	0.275888
Flavon-3-ol 	18	Quercetagenin	3,5,6,7,3′,4′-OH	9.02	2.044793
	19	Rutin	3-rutinoside, 5,7,3′,4′-OH	9.4	2.026872
	20	Hyperoside	3-O-galactoside, 5,7,3′,4′-OH	10.01	1.999566
	21	Robinetin	3,7,3′,4′,5′-OH	11.02	1.957818
	22	Rhamnetin	3,5,3′,4′-OH, 7-OMe	13.5	1.869666
	23	Fisetin	3,7,3′,4′-OH	14.06	1.852015
	24	Quercetin 3,5-di-O-glucoside	3,5-O-glucoside, 7,3′,4′-OH	14.41	1.841336
	25	Morin	3,5,7,2′,4′-OH	17.27	1.762708
	26	Galangin	3,5,7-OH	71.64	1.144844
	27	Quercetin 3,7,3′,4′-tetramethylether	5-OH, 3,7,3′,4′-OMe	261.4	0.582694

	28	3-Hydroxyflavone	3-OH	695.93	0.157434
	29	Quercetin	3,5,7,3′,4′-OH	10.89	1.962972
	30	Kaempferol	3,5,7,4′-OH	28.05	1.552067
	31	Kaempferol 3,5-di-O-glucoside	3,5-O-glucoside, 7,4′-OH	528.37	0.277062
	32	Isoquercitrin	3-O-glucoside, 5,7,3′,4′-OH	9.45	2.024568
Flavanone 	33	Taxifolin	3,5,7,3′,4′-OH	9.27	2.03292
	34	Hesperetin	5,7,3′-OH, 4′-OMe	236.63	0.62593
	35	Hesperidin	7-O-rutinoside, 5,3′-OH, 4′-OMe	281.41	0.55066

Model development

To begin the model development process, the whole data set (n=35) was divided into training (n=27, 75% of the total number of compounds) and test (n=8, 25% of the total number of compounds) sets by *k*-means clustering technique [12] 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. The stepwise regression analysis was performed using statistical software MINITAB [13].

Stepwise Regression

In stepwise regression [14], a multiple term linear equation was built step-by-step. The basic procedures involve (1) identifying an initial model, (2) iteratively "stepping", i.e., repeatedly altering the model of the previous step by adding or removing a predictor variable in accordance with the "stepping criteria", (F = 2 for inclusion; F = 1.9 for exclusion) in our case and (3) terminating the search when stepping is no longer possible given the stepping criteria, or when a specified maximum number steps has been reached. Specifically, at each step all variables are reviewed and evaluated to determine which one will contribute most to the equation. That variable will then be included in the model, and the process started again. A limitation of the stepwise regression search approach is that it presumes that there is a single "best" subset of X variables and seeks to identify it. There is often no unique "best" subset, and all possible regression models with a similar number of X variables as in the stepwise regression solution should be fitted subsequently to study whether some other subsets of X variables might be better.

Statistical qualities

The statistical qualities of the equations were judged by the parameters such as *determination coefficient* (R^2). The generated QSAR equations were validated by leave-one-out *cross-validation* R^2 (Q^2) and *predicted residual sum of squares* (*PRESS*) [15-16] and then were used for the prediction of DPPH radical scavenging activity of the test set compounds. The prediction qualities of the models were judged by statistical parameters like predictive R^2 (R^2_{pred}).

RESULTS AND DISCUSSIONS

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 [17] and the test set members are given in Table 3. The result obtained from stepwise regression analysis is described below and the interpretations of the equations are also depicted.

Table 2
k-Means clustering of compounds using standardized descriptors

Cluster No.	No. of compoun ds in different clusters	Compounds (Sl nos.) in each clusters												
1	13	1	2	4	5	10	12	21	22	23	30	32	34	35
2	5	3	6	7	16	26								
3	8	18	19	20	24	25	29	31	33					
4	9	8	9	11	13	14	15	17	27	28				

Table 3
Observed and calculated DPPH radical scavenging activity

Sl. No.	Obs[a] (pC)	Cal[a]
Training set		
1	2.241845	1.713958
2	1.957031	1.920137
4	1.067171	0.791804
5	1.550213	1.578245
6	1.692932	1.635886
7	0.695682	0.632217
8	0.778768	0.821202
9	0.504206	0.673783
10	0.354322	0.826397
11	0.334044	1.05875
14	0.267735	0.290187
15	0.092207	-0.02882
16	0.766293	1.220604
17	0.275888	0.286005
18	2.044793	1.772775
19	2.026872	1.823137
20	1.999566	1.902772
21	1.957818	2.007349
22	1.869666	1.979552
23	1.852015	2.188278
25	1.762708	1.922458
26	1.144844	0.987933
27	0.582694	0.115672
29	1.962972	2.363135
32	2.024568	1.497443
33	2.03292	2.155344
35	0.55066	0.576607
Test set		
3	1.809668	1.712041
12	0.332426	1.157091
13	0.307532	0.546748
24	1.841336	1.696596
28	0.157434	0.510505
30	1.552067	1.497443
31	0.277062	0.93123
34	0.62593	0.907289

[a] Observed activity (ref. 8); [b] Calculated from eq. (1)

Stepwise regression

Using stepping criteria based on F value (F = 2.0 for inclusion; F = 1.9 for exclusion), the best equation was obtained after successive addition of E-state parameters.

$$\begin{aligned} pC &= 21.21 - 1.49S_1 + 0.50S_7 - 0.58S_9 + 4.3S_{12} - 0.47S_{14} \\ &+ 0.37S_{15} + 1.55S_{16} - 10.0S_{17} \\ n_{Training} &= 27, R^2 = 0.8102, R_a^2 = 0.7518, S = 0.371, \\ Q^2 &= 0.581, PRESS = 7.89, n_{Test} = 8, R_{pred}^2 = 0.7284 \end{aligned} \quad (1)$$

Eq. (1) could explain 75.18% of the variance (adjusted coefficient of variation) and leave – one – out predicted variance was found to be 58.10%. While Eq. (1) was applied for prediction of test set compounds, the predictive R^2 value for the test set was found to be 0.7284. The positive coefficients of S_7, S_{12}, S_{15} and S_{16} indicate that activity increases with increase in E-state value of atom 7, 12, 15 and 16 respectively. Compounds with high values of E-state parameter for atom 7 (S_7) (like 21 and 23) for atom 12 (S_{12}) (like 3) for atom 15 (S_{15}) (like 3) and for atom 16 (S_{16}) (like 26) showed comparatively higher activity. The negative coefficients of S_1, S_9, S_{14} and S_{17} indicate that activity decreases with increase in E-state value of atom 1, 9, 14 and 17 respectively. Compounds with high values of E-state parameter for atom 1 (S_1) (like 14, 15 and 31) for atom 9 (S_9) (like 14, 27 and 28) for atom 14 (S_{14}) (like 13 and 17) and for atom 17 (S_{17}) (like 13, 15 and 28) showed comparatively poor activity. Position 1 indicates the importance of heteroatom oxygen and it is negatively contributed towards activity. Presence of methoxy group in the flavonoid nucleus is detrimental towards activity.

CONCLUSIONS

The whole dataset (n=35) was divided into a training set (27 compounds) and a test set (8 compounds) based on *k*-means clustering of the standardized descriptor matrix and models were developed from the training set. The predictive ability of the models was judged from the prediction of the activity of the test set compounds. The predictive ability of the model was judged considering internal (Q^2) and external validation (R^2_{pred}) (Q^2= 0.581, R^2_{pred}=0.7284). The developed model indicates the importance of heteroatom oxygen and it is negatively contributed towards activity. Presence of methoxy group in the flavonoid nucleus is detrimental towards activity.

REFERENCES

1. Carlo G.D., Mascolo N., Izzo A.A. and Capasso F., Flavonoid: old and new aspects of a class of natural therapeutic drugs, Life Sci., 65: 337-353, (1999).
2. Aherne S.A. and O'Brien N.M., Dietary Flavonols: Chemistry, Food Content, and Metabolism, Nutrition, 18: 75-81, (2002).
3. .Agarwal O.P., The anti-inflammatory action of nepitrin, a flavonoid, Agents Actions, 12: 298-302, (1982).
4. Church M.K., Cromoglycate-like anti-allergic drugs: a review, Drugs Today, 14: 281-341, (1978).
5. Rasulev B.F., Abdullaev N. D., Syrov V.N. and Leszczynski J., A quantitative structure-activity relationship (QSAR) study of the antioxidant activity of flavonoids, QSAR & Comb.Sci, 24: 1056-1065, (2005).
6. Ozawa T., Genetic and functional changes in mitochondria associated with aging, Physiol. Rev., 77: 425-464, (1997).
7. Farkas O., Jakus J. and Heberger K., Quantitative structure-activity relationships of flavonoid compounds, Molecules, 9: 1079-1088, (2004).

8. Seyoum A., Asres K. and El-Fiky F.K., Structure-radical scavenging activity relationships of flavonoids, Phytochemistry, 67: 2058-2070, (2006).
9. Hall L.H., Mohney B. and Kier L.B., The Electrotopological State: An Atom Index for QSAR, Quant. Struct-Act Relat, 10: 43-51, (1991).
10. Kier L.B. and Hall L.H., An Electrotopological State Index for Atoms in Molecules, Pharm. Res., 7: 801-807, (1990).
11. SRETSA is statistical software in Visual Basic, developed by Ray S. and Biswas R. and standardized using known data sets.
12. Leonard J.T. and Roy K., On Selection of Training and Test Sets for the Development of Predictive QSAR models. QSAR Comb Sci, 25: 235-251, (2006).
13. MINITAB is statistical software of Minitab Inc, USA, http://www.minitab.com
14. Darlington R.B., Regression and linear models, McGraw Hill, New York, 1990.
15. Debnath A.K. In: A.K. Ghose and V.N. Viswanadhan (eds.), *Combinatorial library design and evaluation: Principles, software tools, and applications in drug discovery*, Marcel Dekker, New York, 2001, pp. 73–129.
16. Roy K., On some aspects of validation of predictive QSAR models. Expert Opin Drug Discov, 2: 1567-1577, (2007).
17. Roy P.P., Leonard J.K. and Roy K., Exploring the impact of the size of training sets for the development of predictive QSAR models, Chemom. Intell. Lab. Sys. 2008, 90, 31-42.