International Journal of PharmTech Research
CODEN (USA): IJPRIF ISSN : 0974-4304
Vol.4, No.1, pp 56-60, Jan-Mar 2012

QSAR Modeling of inhibitory activity of Flavonoids against Aldose reductase enzyme using Electrotopological state atom (e-state) Parameter

Supratim Ray

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

Corres. Author: supratimray_in@yahoo.co.in
Ph: +91343-2532678, Telefax: +91343-2532679

Abstract: Inhibitory activity against aldose reductase enzyme of nineteen flavonoid derivatives was subjected to classical quantitative structure activity relationship (QSAR) analysis using electrotopological state (E-state) atom parameter. For the development of the QSAR models, statistical techniques like stepwise multiple linear regression and partial least squares (PLS) were used. The best equation is obtained from partial least squares (PLS) considering equation statistics, explained and predicted variance (R^2=0.867, R_a^2=0.834, Q^2=0.692).
Key words: Flavonoids; Aldose reductase; QSAR; E-state; Stepwise regression; PLS.

Introduction

Flavonoids are a group of naturally occurring polyphenolic compounds ubiquitously found in fruits and vegetables (1-3). Common family members of flavonoids include flavones, flavonols, flavanones, isoflavones, biflavanones, catechins and anthocyanidins. Chemically flavonoids are benzo-γ-pyrone derivatives. The structural difference in each flavonoid family results from the variation in the number and substitution pattern of the hydroxyl groups and the extent of glycosylation of these groups (4). Structural diversity of flavonoids allows them to exhibit antineoplastic, antihepatitis, antibacterial, anti-inflammatory, antimutagenic, antiallergic, antithrombic, antiviral and vasodilatory activities (5-7). 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 (8).

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 (9). Because of their great number and positive biological effects, flavonoids are popular subjects for Quantitative structure-activity relationship (QSAR) studies.

In the present paper, we have modeled the inhibitory activity against aldose reductase enzyme of flavonoids data set reported by Okuda et al. (10) using electrotopological state atom (E-state) parameters by stepwise regression.

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 (11).

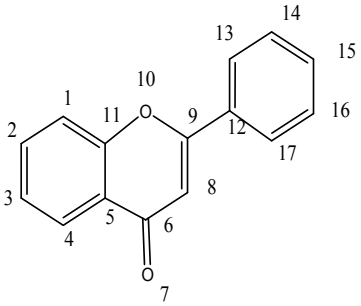


Figure 1: The common atoms of the molecule

The electrotopological state atom (E-state) index developed by Hall and Kier (12) 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

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

In Eq. (1), 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

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

In Eq. (2) 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.

$$S_i = I_i + \sum_{j \neq i} \Delta I_{ij} \qquad (3)$$

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 inhibitory activity against aldose reductase enzyme of flavonoids data set reported by Okuda et al. (10) using electrotopological state atom (E-state) parameters by stepwise regression.

Methods

Stepwise Regression

In stepwise regression (13), 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 = 4 for inclusion; F = 3.9 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.

PLS

PLS is a generalization of regression, which can handle data with strongly correlated and/or noisy or numerous X variables (14-15). 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 nonsignificant. 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.

Data treatment and software

The inhibitory activities of flavonoids against aldose reductase enzyme were used for QSAR analyses as the response variables. The inhibitory activities of flavonoids (10) compounds were used as such reported (log 1/C, pC) for subsequent QSAR analyses as the response variable. For inhibitory activities of flavonoids, 19 compounds (**Table 1) w**ere considered in the present study. 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 Fig.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 (16). 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. The stepwise regression and PLS was performed using statistical software MINITAB (17). The statistical qualities of the equations were judged by parameters like explained variance (R^2_a), correlation coefficient (R). The generated QSAR equations were validated by leave-one-out or LOO method (18-19) using MINITAB software (17) and the calculated parameters are predicted residual sum of squares (PRESS) and cross validation R^2 (Q^2). Q^2 is calculated according to the following formula

$$Q^2 = 1 - \frac{\sum (Y_{obs} - Y_{cal})^2}{\sum (Y_{obs} - \overline{Y})^2} \qquad (4)$$

In Eq. (4), $\overline{Y}$ means average activity value of the entire data set while Y_{obs} and Y_{cal} represent observed and estimated activity values.

Results and Discussion

The results obtained from different statistical methods are described below and the interpretations of the equations are also depicted.

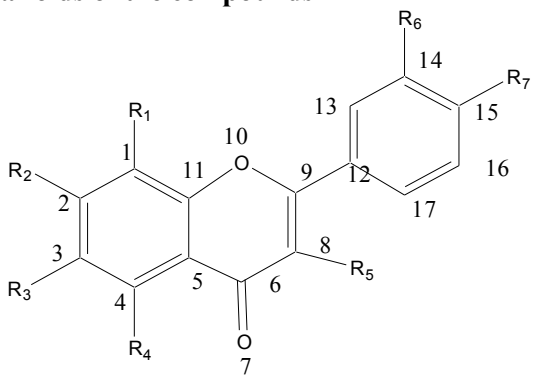
Stepwise regression

Using stepping criteria based on F value (F = 4.0 for inclusion; F = 3.9 for exclusion), the best fit equation was derived.

$$pC = -3.165 - 2.25S_{17} + 1.09S_7$$
$$n = 19, R^2 \quad 0.7658, R^2_a \quad 0.7365, Q^2 = 0.5974, R = 0.875$$
$$PRESS = 0.6677 \qquad (5)$$

Eq. (5) could explain and predict 73.65% and 59.74% respectively of the inhibitory activities. The positive coefficient of S_7 indicate that activity increases with increase in E-state value of atom 7 while the negative coefficient of S_{17} indicate that activity decreases with increase in E-state value of atoms 17.The parameter S_7 implies the impact of the oxygen atom at position 7.

Table 1: Molecular Scaffolds of the compounds



Sl no	R_1	R_2	R_3	R_4	R_5	R_6	R_7	Double bond between C8-C9
1	H	OH	OCH_3	OH	OCH_3	OH	OH	+
2	OCH_3	OCH_3	OCH_3	OCH_3	H	OH	OH	+
3	OCH_3	OCH_3	OH	OCH_3	H	OH	OH	-
4	C_6H_5-CH_2	OH	OCH_3	OH	H	OH	OH	-
5	OCH_3	OCH_3	OCH_3	OH	H	OH	OH	+
6	OCH_3	OCH_3	H	OCH_3	H	OH	OH	+
7	OCH_3	OCH_3	OH	OH	H	OH	OH	+
8	OCH_3	OH	H	OCH_3	H	OH	OH	+
9	OCH_3	OCH_3	H	OH	H	OH	OH	-
10	H	OCH_3	OCH_3	OCH_3	H	OH	OH	-
11	OCH_3	OH	OH	OH	H	OH	OH	-
12	H	OCH_3	OH	OCH_3	H	OH	OH	-
13	OCH_3	OCH_3	OCH_3	OCH_3	H	H	OH	-
14	OH	OCH_3	H	OCH_3	H	OH	OH	-
15	H	OH	OH	OH	H	OH	OH	-
16	H	OCH_3	OCH_3	OH	H	OH	OH	-
17	OCH_3	OH	H	OH	OCH_3	OH	OH	-
18	OCH_3	OCH_3	OH	OCH_3	H	H	OH	-
19	H	OCH_3	OCH_3	OCH_3	H	OH	OH	-

PLS

The number of optimum components was 4 to obtain the final equation (optimized by cross validation). Based on the standardized regression coefficients, the following variables were selected for the final equation:

$$pC = -6.37 + 0.061S_1 + 0.145S_2 - 0.447S_6 + 0.834S_7$$
$$+0.829S_{10} - 0.766S_{11} - 0.49S_{12} - 0.25S_{13} - 0.30S_{16} - 0.664S_{17} \quad (6)$$
$$n = 19, R^2 = 0.867, R_a^2 = 0.834, Q^2 = 0.692, R = 0.930$$
$$PRESS = 0.504$$

Eq. (6) could explain and predict 83.4 and 69.2% respectively of the inhibitory activities. The positive coefficient of the variable S_1, S_2, S_7, S_{10} indicates that the inhibitory activities increases with increase in the E-state value of atom 1, 2, 7 and 10 respectively while the negative coefficients of S_6, S_{11}, S_{12}, S_{13}, S_{16} and S_{17} indicate that the activity decreases with increase in E-state values of atoms 6, 11, 12, 13, 16 and 17. Position 1, 2 indicate that substitution with a group that contain oxygen is beneficial for activity. Position 7 and 10 implies the impact of the oxygen atom at position 7 and 10. Position 6, 11, 12, 13, 16 and 17 indicates the importance of unsaturation in the flavonoid moiety.

Conclusion

For the model, the final equations (5 and 6) obtained from two techniques are of acceptable statistical quality and predictive potential considering the leave-one- out prediction statistics. The models also show the utility of E-state parameters in QSAR study for better understanding about the contribution of atoms or fragments in the molecules towards the biological 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., 1999, 65, 337-353.
2. Haollman P.C.H., and Arts I.C.W., Flavonols, flavones and flavanols-nature, occurrence and dietary burden, J. Sci. Food Agric., 2000, 80, 1081-1093.
3. Aherne S.A., and O'Brien N.M., Dietary Flavonols: Chemistry, Food Content, and Metabolism, Nutrition, 2002, 18, 75-81.
4. Harborne J.B., The Flavonoids, Advances in Research Since 1986, Chapman and Hall, London, 1994.
5. Agarwal O.P., The anti-inflammatory action of nepitrin, a flavonoid, Agents Actions, 1982, 12, 298- 302.
6. Church M.K., Cromoglycate-like anti-allergic drugs: a review, Drugs Today, 1978, 14, 281-341.
7. 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, 2005, 24, 1056-1065.
8. Ozawa T., Genetic and functional changes in mitochondria associated with aging, Physiol. Rev., 1997, 77, 425-464.
9. Farkas O., Jakus J. and Heberger K., Quantitative structure-activity relationships of flavonoid compounds, molecules, 2004, 9, 1079-1088.
10. Okuda J., Miwa K., Inagaki, Horie T. and Nakayama M., Inhibition of aldose reductase from rat and bovine lenses by flavonoids, Biochem. Pharmacol., 1982, 31, 3807-3822.
11. Hall L.H., Mohney B. and Kier L.B., The Electrotopological State: An Atom Index for QSAR, Quant.Struct-Act. Relat., 1991, 10, 43-51.
12. Kier L.B. and Hall L.H. An Electrotopological State Index for Atoms in Molecules, Pharm. Res., 1990, 7, 801-807.
13. Darlington R.B., Regression and linear models, McGraw Hill, New York, 1990.
14. Wold S., PLS for multivariate linear modeling in Van de Waterbeemd H., (Ed.), Chemometric Methods in Molecular Design (Methods and Principles in Medicinal Chemistry), VCH, Weinheim, 1995, 195-218.
15. Fan Y., Shi L.M., Kohn K.W., Pommier Y. and Weinstein J.N., Quantitative structure-antitumor activity relationships of camptothecin analogs: Cluster analysis and genetic algorithm-based studies, J. Med. Chem., 2001, 44, 3254-3263.
16. SRETSA is statistical software in Visual Basic, developed by Ray S. and Biswas R. and standardized on known data sets.
17. MINITAB is statistical software of Minitab Inc, USA, http://www.minitab.com.
18. Word S. and Eriksson L., Statistical validation of QSAR results, in Van de Waterbeemd H., (Ed.) Chemometric Methods in Molecular Design (Methods and Principles in Medicinal Chemistry), VCH, Weinheim, 1995, 312-318.
19. Debnath A.K., Quantitative structure-activity relationship (QSAR): A versatile tool in drug design, in Ghose A.K. and Viswanadhan V.N., (Eds.), Combinatorial Library Design and Evaluation:Principles, Software, Tools and Application in Drug Discovery, Marcel Dekker Inc, NewYork, 2001, 73-129.

*****