Available on line www.jocpr.com

Journal of Chemical and Pharmaceutical Research

ISSN No: 0975-7384
CODEN(USA): JCPRC5

J. Chem. Pharm. Res., 2010, 2(6):344-350

Computational QSAR analysis of some physiochemical and topological descriptors of Curcumin derivatives by using different statistical methods

Raghunath Satpathy, Rajesh Ku. Guru, Rashmiranjan Behera

Department of Biotechnology, MIRC LAB, MITS Engineering College, Rayagada, Odisha, India

ABATRACT

Curcumin is produced from the rhizomes of Curcuma longa plant and having various medicinal and pharmaceutical applications.Here in this work a QSAR study has been performed by taking the 23 analogs of Curcumin.Various structural and physiochemical descriptors were generated. The effect was calculated for each type of descriptors by taking the Andrews coefficient as dependent variable. Multiple regression analysis was performed by Minitab 14 tool. Good correlation R-sq value 0.78 was obtained from the physiochemical descriptors in comparison to structural descriptor calculation. The statistics were also further verified by using SVM (Support vector machines) and ANN (artificial neural networks) based calculation. The results obtained were consistent with MLR statistics and the ANN based method show R-sq value as 0.88 in case of physiological descriptor which is observed to be the highest among above three methods of analysis.

Key words: Descriptors, QSAR, Curcumin, Multiple linear regression, artificial neural network

INTRODUCTION

Curcumin is an alkaloid produced from the turmeric plant *Curcuma longa*, which is a member of the ginger family (Zingiberaceae). Historically the turmeric has been used as a major component of Indian Ayurvedic medicine to treat a wide variety of health problems [1]. Current research has also identified the Curcumin as responsible molecule for most of the biological activity of turmeric. The Curcumin molecules are chemically polyphenols and are responsible for the yellow color of turmeric and can exist in at least two tautomeric forms, keto and enol [2].Curcumin incorporates several functional groups and the aromatic ring systems the carbonyl groups form a diketone [3].Recently numerous clinical trials in humans are going on, investigating the effect of Curcumin on various diseases including multiple myeloma, pancreatic cancer, myelodysplastic syndromes, colon cancer, psoriasis, and Alzheimer's disease, and also deadliest Swine flu [4-5-6-7].

To analyse different potential drug molecules the quantitative structure-activity relationship (QSAR) method is a useful approach.QSAR is basically used to study the biological activities with various properties associated with the structures, which is helpful to explain how structural features in a drug molecule influence the biological activities. The analysis also gathers information that is very much useful for molecular drug design and medicinal Chemistry. Therefore correlating the physicochemical properties or structural features of the important compounds with their biological activity is essential. In addition to this a successful in silico based QSAR analysis also provides the advantages of higher speed and lower costs for bioactivity evaluation of drug as compared to experimental testing [8].

EXPERIMENTAL SECTION

The molecular structure of Curcumin derivatives were collected from Pubchem database available in the NCBI server (http://pubchem.ncbi.nlm.nih.gov/).The structure were drawn by Marvin sketch 5.0 tool (http://www.chemaxon.com/marvin/sketch/index.jsp) and corresponding 3D structure were obtained. The molecules were then energy minimised by PRODRG server [9]. Prodrg is an on line tool where the energy minimization of the molecule was performed by using Gromos 96 force field. Then the energy minimised molecules were fed to Preadmet server (http://preadmet.bmdrc.org/preadmet/index.php) for the calculation of descriptors. Two types of descriptors were chosen physiological and topological types under which the selected descriptors were calculated (Table 1).

Table 1: The various physiochemical and Topological descriptors considered in the study

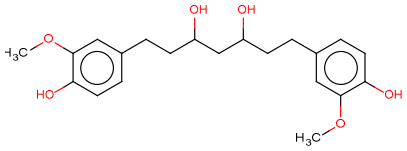
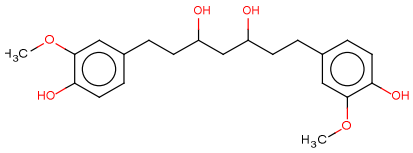
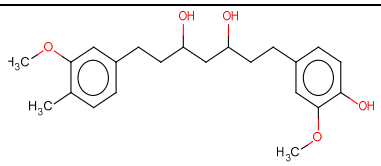
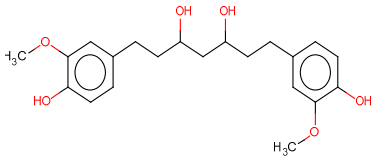
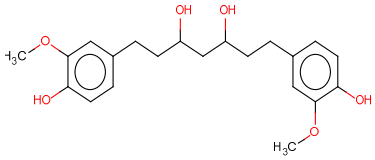
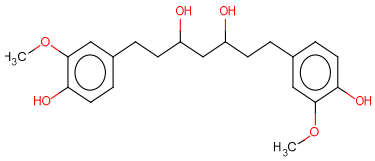
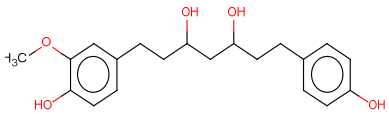
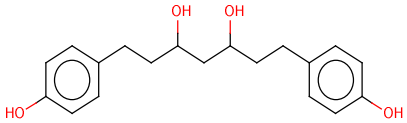
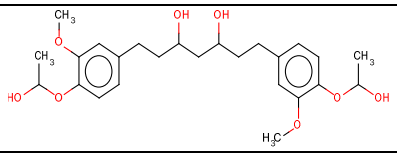
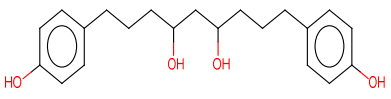
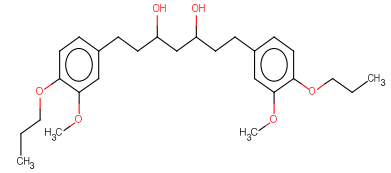
Serial number	Physiochemical	Topological
1	Molecular weight (MW)	Quadratic index (QI)
2	2D Vander walls volume (2DVWV)	Edge based molecular topological index (EMTI)
3	Water solvation free energy (WSE)	Kier symmetry index (KSI)
4	Hydrophobic surface area saturated (HSAS)	Ring degree distance index (RDI)
5	Hydrophobic surface area un-saturated (HSAU)	Eccentric connectivity index (ECI)
6	LogP	Wiener index (WI)

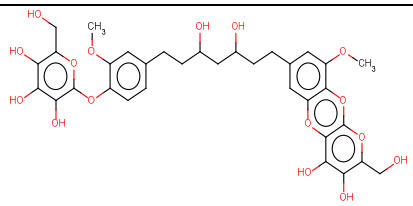
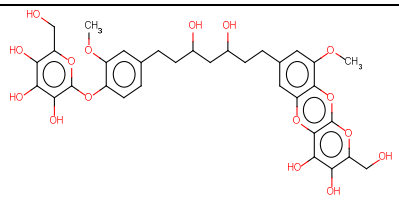
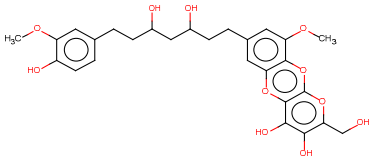
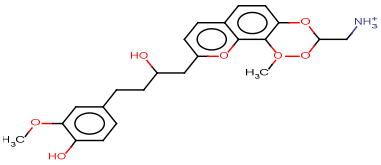
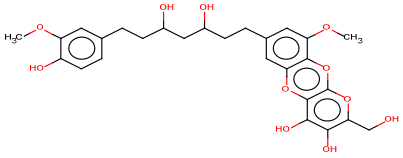
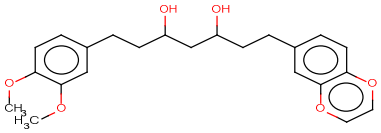
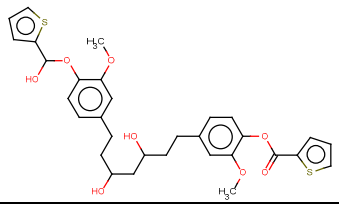
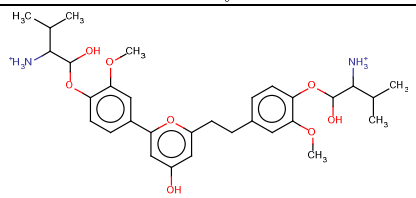
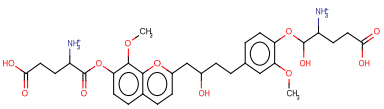
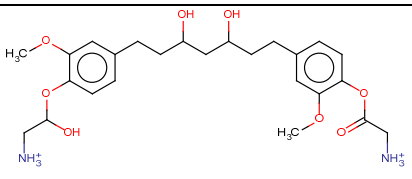
For the two types of descriptor sets, MLR (Multiple linear regressions) analysis was performed by using the MINITAB 14 tool [10].The Andrews affinity was chosen as dependent variable .Andrews affinity is calculated based on the drug receptor binding affinity [11]. For the best model selection the statistical parameters like F value, R-Sq value and mean square deviation etc were considered. The above MLR calculations were also further verified by ANN (Artificial Neural Network) and SVM (Support Vector Machine) based approach by using Molegro tool [12].

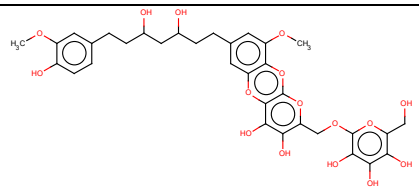
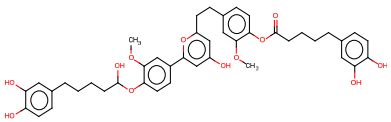
RESULTS AND DISCUSSION

Overall 23 Curcumin analogs were retrieved from Pubchem data base and the same were used for the QSAR analysis (Table 2).

Table 2: Structure of Curcumin analogs considered for the experiment

Curcumin analogs	Structure	Andrews affinity	Molecular weight
1		4.76372	332.182
2		4.83701	368.385
3		3.51782	368.381
4		4.69043	366.413
5		4.76372	370.401
6		4.69043	368.385
7		3.88426	338.359
8		2.85823	308.333
9		6.44934	452.459
10		3.0781	336.387
11		2.85823	448.515

12		22.2063	692.667
13		22.2063	694.683
14		13.485	532.542
15		13.5583	425.437
16		13.485	530.526
17		3.51782	396.439
18		10.1137	588.659
19		25.4309	566.651
20		33.9323	626.615
21		22.4261	482.489

22		22.2063	694.683
23		19.348	744.749

The predicted Andrews affinity on various physiochemical and topological descriptors were calculated by MLR analysis and the regression equations were generated from Mintab 14 tool.
Andrews affinity (physiochemical descriptors) = - 27.0 - 5.11 log p - 0.0291 MW + 0.159 2DVWV + 0.123 WSE - 0.0395 HSAS + 0.156 HSAU
Andrews affinity (topological descriptors) = - 20.5 - 0.18 QI - 0.000022 EMTI + 0.828 KSI - 1.62 RDM + 0.0014 ECI - 0.00108 WI

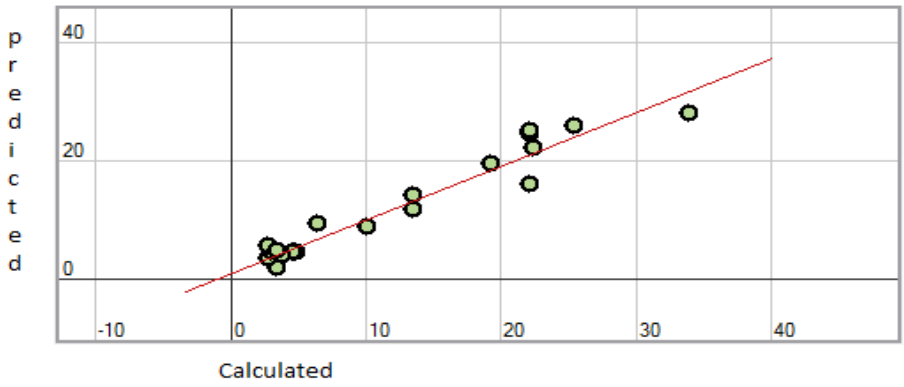


Figure 1: The predicted and calculated affinity relationship in case of physiochemical descriptors by ANN method.

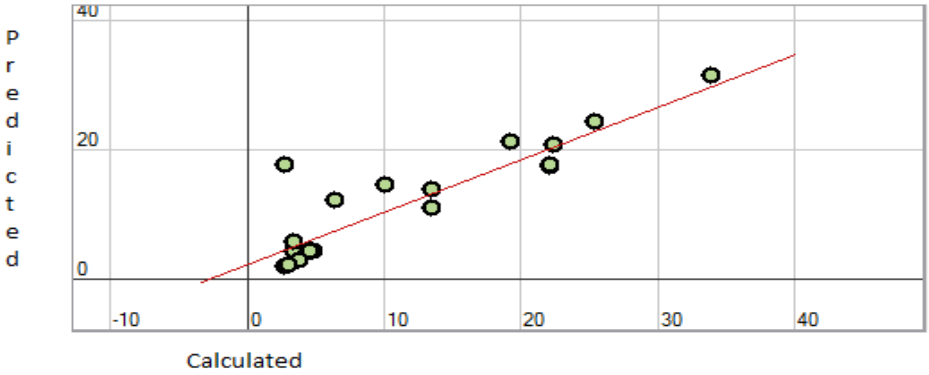


Figure 2: The predicted and calculated affinity relationship in case of topological descriptors by ANN method.

The data set was further verified by Molegro software. In addition to the MLR analysis the SVM and ANN method was used to calculate the statistical variables because the comparative mode of statistical analysis (Multiple linear regression analysis, SVM and ANN based approach) are more reliable to analyse the statistical parameters [13].Default parameter set up was considered in Molegro tool for the ANN and SVM based calculation. In case of ANN based calculation single hidden layer with 3 neurons were chosen. Among all 3 method of analysis

the application of artificial neural networks show the maximum Pearsons coefficient 0.940,Pearsons coefficient square 0.88 and minimum mean square deviation 10.033 in case of physiochemical descriptors that signifies the statistical analysis (Table-3).

Table 3: Comparative statistical parameter calculations by using Molegro tool

Descriptor Type	Methods of analysis	Pearson Correlation (r)	Pearson's coefficient square (r2)	Spearman Rank Correlation (ρ)	Mean Squared Deviation (MSD)	Cross validated squared (q2)
Physiochemical	MLR	0.886	0.7844	0.777	17.081	0.784
	SVM	0.862	0.742	0.805	22.046	0.721
	ANN	**0.940**	**0.884**	**0.77**	**10.033**	**0.873**
Topological	MLR	0.799	0.637	0.656	28.697	0.637
	SVM	0.838	0.703	0.719	23.96	0.697
	ANN	**0.905**	**0.818**	**0.758**	**14.366**	**0.818**

The artificial neural network based calculation provides the improved QSAR model for the effect topological (Figure 1) and physiochemical descriptors (Figure 2) with the Andrews affinity. The successful application of ANN methods to QSAR analysis also has been confirmed for other drug molecules in medicinal chemistry [14]. So the ANN could be used as a promising tool for a good statistical approximation thereby solving complex problems. In general the topological and structural descriptors are very important type of molecular descriptor for bioactivity prediction [15]. Here the results in this work indicate in comparison to topological parameters the physiochemical parameters are more responsible for receptor binding activity of Curcumin.

Acknowledgement:
We are thankful to Chief executive, Majhighariani Institute of Technology & Science, Ryagada for his encouragement and to provide us MIRC lab for computing facility.

REFERENCES

[1]Aggarwal BB; Sundaram C; Malani N; Ichikawa H. *Adv. Exp. Med. Biol.* **2007**, 595: 1–75.
[2]Tsonko M. Kolev; Evelina A.; Velcheva; Bistra A; Stamboliyska; Michael Spiteller. *International Journal of Quantum Chemistry*, **2005**, 102 (6), 1069–79.
[3]Albena T.; Dinkova-Kostova; Paul Talalay *Carcinogenesis*, **1999**, 20 (5), 911-914.
[4]Aggarwal BB; Shishodia S. *Biochemical Pharmacology*, **2006**, 71 (10), 1397–1421.
[5]Hyunsung Choi; Yang-Sook Chun; Seung-Won Kim; Myung-Suk Kim; Jong-Wan Park. *Molecular Pharmacology*, **2006**, 70 (5), 1664–71.
[6]Hatcher H; Planalp R; Cho J; Torti FM; Torti SV. *Cell. Mol. Life Sci.*, **2008**, 65 (11), 1631–1652.
[7]Da-Yuan Chen ; Jui-Hung Shien ; Laurence Tiley ; Shyan-Song Chiou ; Sheng-Yang Wang; Tien-Jye Chang ; Ya-Jane Lee ; Kun-Wei Chan ; Wei-Li Hsu. *FoodChemistry,* **2010**, 119 (4), 1346–1351.
[8] David A. Winkler. *Briefings in Bioinformatics*, **2002**, 3(1), 73-86.
[9]Schuettelkopf AW; D.M.F VanAalten. *Acta Crystallographica*, 2004, D60: 1355-1363.
[10]David J.; Wild J. *Chem. Inf. Model*, **2005**, 459(1), 212
[11]Andrews PR; Craik DJ; Martin JL. *J Med Chem*. **1984**, 27(12), 1648-1657.
[12] www.molegro.com
[13]Shahlaei M; Fassihi A; Saghaie L. *Eur J Med Chem*., **2010**, 45(4), 1572-1582.

[14]Zahra Garkani-Nejad; Fereshteh Saneie. *Bull. Chem. Soc. Ethiop.*, **2010**, 24(3), 317-325.
[15]Du Q, Mezey PG; Chou KC. *J Comput Chem.*, **2005**, 26(5):461-470.