Curcumin analogues and their logPo/w Prediction

Preeti M. Chincholikar[1], Dr. Sulbha Amlathe[2]
[1]Gandhi.P.R. College, Bhopal (M.P) India.
[2]HoD, Applied Chemistry, BUIT, Bhopal (M.P) India.

ABSTRACT

A quantitative structure–activity relationship (QSAR) study was performed to develop models those relate the structures of 41 curcumin compounds to their n-octanol–water partition coefficients (log $P_{o/w}$). The analogues were studied under different quantum-chemical descriptors, electrostatic, constitutional, topological, geometrical that were considered as inputs to the model. The models were constructed using 20 molecules as training set, and predictive ability tested using 11 compounds. Modelling of log $P_{o/w}$ of these compounds as a function of the theoretically derived descriptors was established by multiple linear regression (MLR). The usefulness of the quantum chemical descriptors, calculated at the level of the HF theories using 6-31G basis set for QSAR study of anti-cancer drugs was examined. A multi-parametric equation containing maximum eight descriptors at HF/6-31G* method with good statistical qualities (R2train=0.838, Ftrain=22.93, Q^2_{LOO}=0.843,R^2_{adj}=0.862,Q^2_{LGO}=0.729) was obtained by Multiple Linear Regression using stepwise method. The accuracy of the proposed MLR model was illustrated using the following evaluation techniques: cross-validation, validation through an external test set, and Y randomisation. The predictive ability of the model was found to be satisfactory and could be used for designing a similar group of compounds.*

Keywords: curcumin and its analogues, n-Octanol–water partition coefficients, Quantitative structure–activity relationship (QSAR), Multiple linear regression (MLR).

I 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].

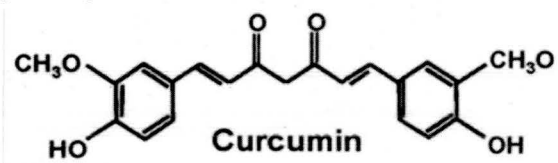


Fig 2 A curcumin molecule

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. Here one of the major aspects of the studies in QSAR is according to Lipinski rule of 5 [8, 9, 10]. 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 [11].

The **partition coefficient** is a ratio of concentrations of un-ionized compound between the two solutions. To measure the **partition coefficient** of ionizable solutes, the pH of the aqueous phase is adjusted such that the predominant form of the compound is un-ionized. The logarithm of the ratio of the concentrations of the un-ionized solute in the solvents is called **log *P***: The log P value is also known as a measure of lipophilicity[12].

$$\log P_{oct/wat} = \log\left(\frac{[solute]_{octanol}}{[solute]^{un-ionized}_{water}}\right)$$

The drug's distribution coefficient strongly affects the ease of any drug that can reach its intended target in the body and how strong an effect it will have once it reaches its target, and how long it will remain in the body in an active form.

LogP is one criterion used in medicinal chemistry to assess the druglikeness of a given molecule, and used to calculate lipophilic efficiency, a function of potency and LogP that evaluate the quality of research compounds.[12][13] For a given compound lipophilic efficiency is defined as the pIC_{50} (or pEC_{50}) of interest minus the LogP of the compound. Here we have used 20 different curcumin analogues to study their logP characters. Experimental determination of log *P*o/w is often complex and timeconsuming and can be done only for already synthesized compounds. For this reason, a number of computational methods for the prediction of this parameter have been proposed. In this work a QSAR study is performed, to develop models that relate the structures of a heterogeneous group of 41 drug compounds to their *n*-octanol–water partition coefficients. However, using *in vivo* methods to measure the logarithmic values of partition coefficient drug concentration ratios (log P) in humans is not possible, and to do so in animal models is expensive and time consuming. Finally, the accuracy of the proposed model was illustrated using the following: leave one out, bootstrapping and external test set, cross-validations and Y-randomisation techniques [14-22].

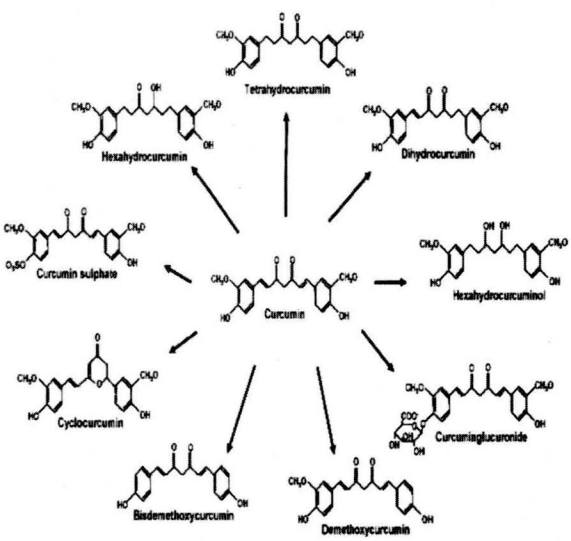


Fig 2 The different curcumin analogues

II 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 ACDchemsketch tool (http://www.acdlabs.com) and corresponding logP of each structure was obtained. DRAGON and GAUSSIAN 05 were used for the descriptor study [23-28,32].

(a) Data Set- The properties data used in this study are the LogPo/w of the set of 20 curcumin derivatives [37-49]. The data set was randomly divided into two subsets: the training set containing 20 compounds (80%) and the test set containing ? compounds (20%). The training set was used to build a regression model, and the test set was used to evaluate the predictive ability of the model obtained. The properties data for the complete set of compounds are presented in Table 2,to derive QSAR models, an appropriate representation of the chemical structure is necessary. For this purpose, descriptors of the structure are commonly used is shown in table-1.

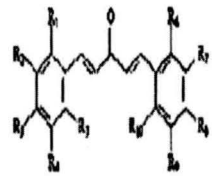
Table 1
The descriptors used in the present study

Descriptors	Symbol	Abbreviation
Quantum chemical descriptors	Molecular Dipole Moment	MDP
	Molecular Polarizability	MP
	Natural Population Analysis	NPA
	Electrostatic Potential	EP
	Highest Occupied Molecular Orbital	HOMO
	Lowest Unoccupied Molecular Orbital	LUMO
	difference between LUMO and HOMO	E GAP
	Electro negativity [χ= -1/2 (HOMO–LUMO)]	X
	El Electro philicity ($\omega=\chi$ 2/2 η)	Ω
	Mulliken E GAPCharge	MC
Chemical properties	Partition Coefficient	Log P
	Mass	M
	Molecule volume	V
	Molecule surface area	MSA
	Hydration Energy	HE
	Refractivity	REF

III RESULTS AND DISCUSSION

Overall 20 Curcumin analogues were retrieved from Pubchem data base and the same were used for the QSAR analysis by the following descriptor studies (Table 2).

Table2
variations considered in the R_1-R_{10} positions of curcumin



Compounds	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
1	-OH	-H	-H	-H	-H	-OH	-H	-H	-H	-H
2	-H	-OH	-H	-H	-H	-H	-OH	-H	-H	-H
3	-H	-H	-OH	-H	-H	-H	-H	-OH	-H	-H
4	-H	-H	-H	-F	-H	-H	-H	-H	-F	-H
5	-H	$-OCH_3$	$-OCH_2OCH_3$	$-OCH_3$	-H	-H	$-OCH_3$	$-OCH_2OCH_3$	$-OCH_3$	-H
6	-H	$-OCH_2OCH_3$	-H	$-OCH_2OCH_3$	-H	-H	$-OCH_2OCH_3$	-H	$-OCH_2OCH_3$	-H
7	-H	-H	-OH	$-OCH_3$	-H	-H	-OH	$-OCH_3$	-H	-H
8	-H	OCH_3		OCH_3	H	H	OCH_3		OCH_3	H
9	-H	$-OCH_3$		-H	-H	-H	$-OCH_3$		-H	-H
10	-H	-H		-H	-H	-H	-H		-H	-H
11	-H	$-CH_3$		$-CH_3$	-H	-H	$-CH_3$		$-CH_3$	-H

(a)Data-All log*P*o/w data for all 41 compounds was taken from the literature.The data set was split into a training set (13compounds) and a prediction set (7 compounds). The log *P*o/w of these compounds are deposited in Journal log as supporting material (see Tables 2). Chemical structure of drugs that illustrated in this study is shown in Table 2.

(b) Molecular descriptor generation-All of the molecules were drawn into the Chemsketch. The Gaussian 03 and DRAGON packages were used for calculating the molecular descriptors(Table 1). Some of the descriptors are obtained directly from the chemical structure, e. g. constitutional, geometrical, and topological descriptors. Other chemical and physicochemical properties were determined by the chemical structure (lipophilicity, hydrophilicity descriptors, electronic descriptors, energies of interaction). In this work, we used Gaussian 03 for ab initio calculations.DFT method at 6-31G* were applied for optimization of anti-cancer drugs and calculation of many of the descriptors. Software hyper Chem and some of the descriptors such as partition coefficient, surface area, hydration energy, and refractivity were calculated through it. The rest of the descriptors were obtained of Gaussian calculations.

A large number of descriptors were calculated by Gaussian package and Chemsketch software. One way to avoid data redundancy is to exclude descriptors that are highly intercorrelated with each other before performing statistical analysis. The molecular structures were saved by the HIN extension and entered on the DRAGON software for the calculation of the 18 different types of theoretical descriptors for each molecule. They included (a) 0D-constitutional (atom and group counts); (b) 1D-functional groups, 1D-atom centered fragments; (c) 2D-topological, 2DBCUTs, 2D-walk and path counts, 2D-autocorrelations, 2D-connectivity indices, 2Dinformation indices, 2D-topological charge indices, and 2D-eigenvalue-based indices; and (d) 3D-Randic molecular profiles from the geometry matrix, 3D-geometrical, 3D-WHIM, and 3DGETAWAY descriptors. A stepwise technique was employed that only one parameter at a time was added to a model and always in the order of most significant to least significant in terms of F-test values. Statistical parameters were calculated subsequently for each step in the process, so the significance of the added parameter could be verified. The goodness of the correlation is tested by the regression coefficient (R^2), the F-test and the standard error of the estimate (SEE). The test and the level of significance, as well as the confidence limits of the regression coefficient, are also reported. The squared correlation coefficient, R^2, is a measure of the fit of the regression model. Correspondingly, it represents the part of the variation in the observed (experimental) data that is explained by the model.

(c) Genetic algorithm for descriptor selection- Genetic algorithm variable selection is a technique that helps identify a subset of the measured variables that are, for a given problem, the most useful for a precise and accurate regression model. The selection of relevant descriptors, which relate the log *P*o/w to the molecular structure, is an important step to construct predictive models. The genetic algorithm was applied to the input set of 53 molecular descriptors for each chemical of the studied data sets and the related response, in order to extract the best set of molecular descriptors, which are, in combination, the most relevant variables in modeling the response of the training set chemicals.

Genetic algorithm (GA), included in the PLS Toolbox version 2.0, was used for variables selection (based on the training set). Using GA-based MLR variable selection procedures, the dependent variables, i.e., the log *P*o/w, were used to find subsets of molecular descriptors that provide a good relationship to the log *P*o/w. Given an X-matrix of descriptors data and a log *P*o/w of values to be predicted, one can choose a random subset of variables from **X** and, through the use of cross-validation and MLR regression method, determine the root-mean-square error ofcross-validation (RMSECV)[31,33] obtained when using only that subset of variables in a regression model. Genetic algorithms use this approach iteratively to

locate the variable subset (or subsets) which gives the lowest RMSECV. The first step of the GA is to generate a large number (e.g., 32, 64, 128) of random selections of the variables and calculate the RMSECV for each of the given subsets. Each subset of variables is called an individual (or chromosome) and the yes/no flags indicating which variables are used by that individual is the gene for that individual. The pool of all tested individuals is the population. The RMSECV values, described as the fitness of the individual, indicate how predictive each individual's selection of variables is for the log *P*o/w [39].

The diversity of the training set and the test set was analyzed using the principal component analysis (PCA) method[40,41]. The PCA was performed with the calculated structure descriptors for the whole data set to detect the homogeneities in the data set[42], and also to show the spatial location of the samples to assist the separation of the data into the training and test sets. The PCA results showed that three principal components (PC1and PC2) described 24.39% of the overall variables, as follows: PC1 = 16.79% and PC2 =7.6.%. Since almost all the variables can be accounted for by the first three PCs, their score plot is a reliable representation of the spatial distribution of the points for the data set. The multi-collinearity between the above seven descriptors were detected by calculating their variation inflation factors (VIF)[43-45], which can be calculated as follows:

$$VIF= \frac{1}{1-r^2}$$

Where r is the correlation coefficient of the multiple regression between the variables in the model. If VIF equals to 1, then no inter-correlation exists for each variable; if VIF falls into the range of 1–5, the related model is acceptable; and if VIF is larger than 10, the related model is unstable and a recheck is necessary [30]. The corresponding VIF values of the seven descriptors are shown in Table 2. As can be seen from this table, most of the variables had VIF values of less than 5, indicating hat the obtained model has statistic significance. To examine the relative importance as well as the contribution of each descriptor in the model, the value of the mean effect (MF)[46-50] was calculated for each descriptor. This calculation was performed with the equation below:

$$MFj= \frac{\beta \sum_{i=1}^{i=n} dij}{\sum_{j}^{m} \beta j \sum_{i}^{n} \beta ij}$$

Where *MFj*represents the mean effect for the considered descriptor *j*, *βj* is the coefficient of the descriptor *j*, *dij*stands for the value of the target descriptors for each molecule and, eventually, *m* is the descriptors number for the model. The MF value indicates the relative importance of a descriptor, compared with the other descriptors in the model. Its sign indicates the variation direction in the values of the activities as a result of the increase (or reduction) of the descriptor values. The mean effect values are shown in Table 3. All descriptors were calculated for the neutral species. The log *P*o/w is assumed to be highly dependent upon the EP26, NPA13, SAPAC22,PW3 ,Mor16m,Mor18m,Mor24m and G2u.

In the present study, the QSAR model was generated using a training set of 33 molecules (Table 2).

Table 3.The linear model based on the eight parameters selected by the GA-MLR method

Descriptor	Chemical meaning	MFa	VIFb
Constant	Intercept	0	0
EP26	Electrostatic potential 26	1.260265	1.148737
NPA13	Natural population analysis 13	-0.15876	1.182888
SAPAC22	Surface area approx atomic charg22	0.00414	1.105926
PW3	Path/walk3-randic shape index	-0.07902	1.284402
Mor16m	3D-MoRSE-signal16/weighted by atomic masses	0.005628	1.321815
Mor18m	3D-MoRSE-signal18/weighted by atomic masses	0.002912	1.105745
Mor24m	3D-MoRSE-signal24/weighted by atomic masses	-0.00102	1.226363
G2u	1st component symmetry directional WHIM index/unweighted	-0.03414	1.099806

a Mean effect
b Variation inflation factors

The test set of 8 molecules (Table 2) with regularly distributed log *P*o/w values was used to assess the predictive ability of the QSARmodels produced in the regression.

(d) MLR analysis-The software package used for conducting MLR analysis was Spss 16. Multiple linear regression (MLR) analysis has been carried out to derive the best QSAR model. The MLR technique was performed on the molecules of the trainingset shown in Table 2.A small number of molecular descriptors (EP26,NPA13,SAPAC22,PW3 ,Mor16m,Mor18m ,Mor24m and G2u) proposed were used to establish a QSAR model. Additional validation was performed on an external data set consisting of 8 drug compounds. Multiple linear regression analysis provided a useful equation that can be used to predict the log *P*o/w of drug based upon these parameters. The best equation obtained for the Lipophilicity of the drug compounds is

LogP=150.269(±37.396)-12.787(±2.570)EP26+3.882(±0.762)NPA13-0.097(±0.025)SAPAC22+30.446(±9.409)PW31.056(±0.236)Mor16m+0.445(±0.168)Mor18m-1.418(±0.258)Mor24m +34.976(±7.513)G2u

N=41 N train=33 N test=8 R^2 train=0.893 F train=24.934 R^2test=0.541 Ftest=-0.045 R^2adj= 0.857 Q^2_{LOO}=0.816 Q^2_{LGO}=0.730.

In this equation, N is the number of compounds; R^2 is the squared correlation coefficient, Q^2_{LOO}, Q^2_{LGO} are the squared cross-validation coefficients for leave one out, bootstrapping and external test set respectively, F is the Fisher F statistic. The figures in parentheses are the standard deviations. The built model was used to predict the test set data and the prediction results are given in Table 1. As can be seen from Table 1, the calculated values for the LogP are in good agreement with those of the experimental values. The predicted values for LogP for the compounds in the training and test sets using equation 1 were plotted against the experimental LogP values in Figure 1. A plot of the residual for the predicted values of LogP for both the training and test sets against the experimental LogP values are shown in Figure 2.

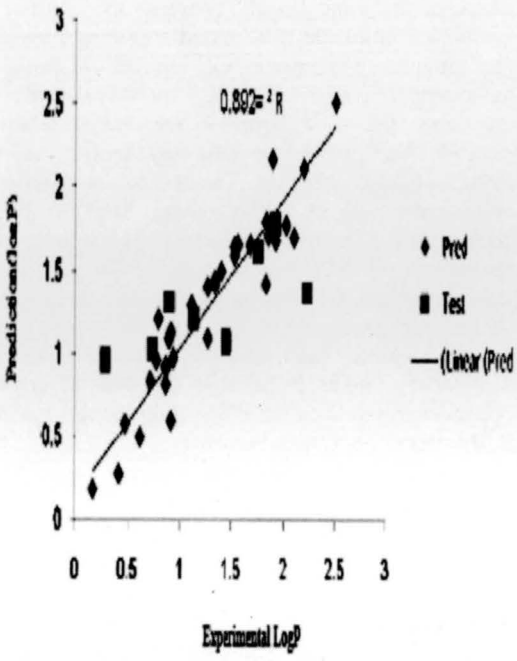


Fig.1 the predicted logP value versus experimental logP by MLR

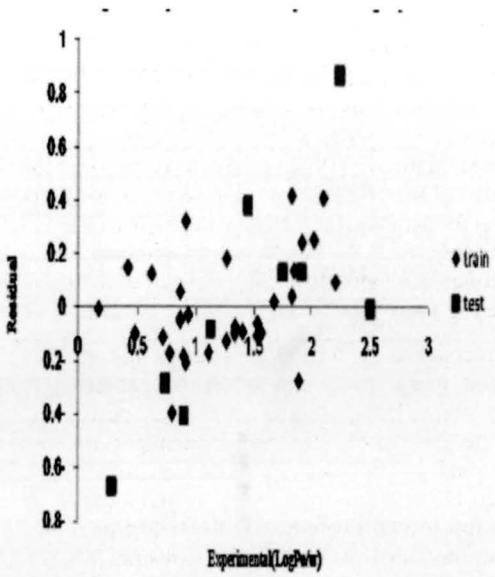


Fig 2 The residual logP versus experimental logP by GA-MLR

Also, in order to assess the robustness of the model, the Y-randomisation test was applied in this study [25–28]. The dependent variable vector (LogP) was randomly shuffled and The new QSAR models (after several repetitions) would be expected to have low R2 and Q2 LOO values (Table 4). If the opposite happens then an acceptable QSAR model cannot be obtained for the specific modeling method and data.

Table 4.
The R^2_{train} and Q^2_{LOO} values after several Y-randomisation tests

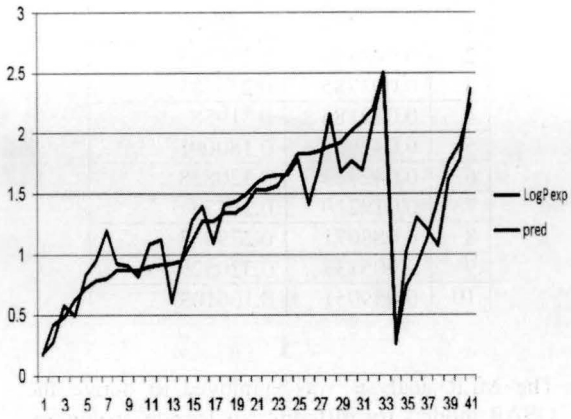
No	Q^2	R^2
1	0.113284	0.472045
2	0.048896	0.230775
3	0.003785	0.234683
4	0.012186	0.31958
5	0.042953	0.180091
6	0.042723	0.320828
7	0.019219	0.21774
8	0.083071	0.279033
9	0.005137	0.320529
10	0.059051	0.166103

The MLR analysis was employed to derive the QSAR models for different Nucleoside analogues. MLR and correlation analyses were carried out by the statistics software SPSS (Table 5). Figure 2 has showed that results were obtained from equation HF/6-31G* to the experimental values.

Table 5. The correlation coefficient existing between the variables used in different MLR and equations with HF/6-31G* method

	EP26	NPA13	SAPAC22	PW3	Mor16m	Mor18m	Mor24m
EP26	1	0	0	0	0	0	0
NPA13	0.054065	1	0	0	0	0	0
SAPAC22	-0.2344	-0.12944	1	0	0	0	0
PW3	0.012593	0.236561	-0.34562	1	0	0	0
Mor16m	-0.33313	-0.24288	-0.00361	-0.23044	1	0	0
Mor18m	0.177918	0.081215	-0.13201	0.061633	-0.19956	1	0
Mor24m	0.157028	0.378326	-0.18368	0.073222	-0.35014	0.252322	1
G2u	.047641	0.049852	-0.17016	0.322559	-0.04335	0.01788	-0.08377

(e) Interpretation of descriptors- The QSAR developed indicated that electrostatic properties (EP), natural population analysis (NPA), surface area approx atomic charge 22 (SAPAC), Path/walk3-randic shape index(PW3) 3D-MoRSE-signal(16,18,24)/weighted by atomic masses (Mor16m,Mor18m, Mor24m), 1st component symmetry directional WHIM index/unweighted (G2u)drug *n*-octanol/water partition coefficients.Positive values in the regression coefficients indicate that the indicated descriptor contributes positively to the value of log *P*o/w, whereas negative values indicate that the greater the value of the descriptor the lower the value of log*P*o/w.In other words, increasing the EP26 and Mor24mwill decrease log *P*o/w and increasing the NPA13,SAPAC22,PW3,Mor16m,G2u and Mor18m increases extent of log *P*o/w of the curcumin. The standardized regression coefficient reveals the significance of an individual descriptor presented in the regression model.



Series 1: the values of log P were obtained by using prediction.
Series 2: the values of log P were obtained by using Experimental methods

Fig 3 The comparison between biological activity (log p) using experimental and prediction

The greater the absolute value of a coefficient, the greater the weight of the variable in the model. Mor16m is the forth descriptor, appearing in the model. It is one of the 3D-molecule representations of structuresbased on electron diffraction (3D-MoRSE) descriptors. The 3DMoRSE descriptors are derived from infrared spectral simulation using a generalised scattering function [31]. Thisdescriptor was proposed as signal (16, 24)/weighted by the atomicmasses which relates to the atomic masses of the molecule.The Mor(16,24)m displays a positive sign, which indicates that theLogPo/wis directly related to this descriptor.The next descriptor is the path/walk 3Randic shape index (PW3), which is one of the topological descriptors. The atomic path/walk indices are defined for each atom as the ratio between the atomic path count and the atomic walk count of the same length. Whereas the number of paths in a molecule is bounded and determined by the molecule's diameter, the number of walks is unbounded. However, being interested only in quotients, the walk count is terminated when it exceeds the maximum allowed length of the corresponding path [31]. The molecular path/walk indices are defined as the average sum of atomic path/walk indices of equal length. As the path/walk count ratio is independent of molecular size, these descriptors can be considered as shape descriptors. As is apparent from Table 2, the PW3 mean effect has a negative sign which indicates that the LogPo/wis inversely related to this descriptor; therefore, increasing the PW3 of molecules leads to a decrease in its LogPo/w values.

IV CONCLUSION

In this article, a QSAR study of 20 curcumin analogues was performed based on the theoretical molecular descriptors calculated by the DRAGON and GAUSSIAN software and selected. The built model was assessed comprehensively (internal and external validation) and all the validations indicated that the QSAR model built was robust and satisfactory, and that the selected descriptors could account for the structural features responsible for the anti-cancer drugs activity of the compounds. The QSAR model developed in this study can provide a useful tool to predict the activity of new compounds and also to design new compounds with high activity.

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.
a. *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
a. Wang; Tien-Jye Chang ; Ya-Jane Lee ; Kun-Wei Chan ; Wei-Li Hsu. *FoodChemistry,*
b. 2010, 119 (4), 1346–1351.

[8] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings". *Adv. Drug Deliv. Rev.* 46 (1-3): 3–26. 2001. doi:10.1016/S0169-409X(00)00129-0. PMID 11259830.

[9] Lipinski CA "Lead- and drug-like compounds: the rule-of-five revolution". *Drug Discovery Today: Technologies* 1(4): 337–341. 2004. doi:10.1016/j.ddtec.2004.11.007.

[10] Oprea TI, Davis AM, Teague SJ, Leeson PD. "Is there a difference between leads and drugs? A historical perspective". *J Chem Inf Comput Sci* 41 (5): 30815.2001) doi:10.1021/ci010366a. PMID 11604031.

[11] Edwards MP, Price DA (2010). "Role of Physicochemical Properties and Ligand Lipophilicity Efficiency in Addressing Drug Safety Risks". *Annual Reports in Medicinal Chemistry.* Annual Reports in Medicinal Chemistry 45: 381–391.doi:10.1016/S0065-7743(10)45023-X. ISBN 978-0-12-380902-5.

[12] Jump up^ Leeson PD, Springthorpe B (November 2007). "The influence of drug-like concepts on decision-making in medicinal chemistry". *Nat Rev Drug Discov* 6 (11): 881–90.doi:10.1038/nrd2445. PMID 17971784.

[13] David A. Winkler. *Briefings in Bioinformatics,* 2002, 3(1), 73-86.

[14] Schuettelkopf AW; D.M.F VanAalten. *Acta Crystallographica,* 2004, D60: 1355-1363.

[15] David J.; Wild J. *Chem. Inf. Model,* 2005, 459(1), 212

[16] Andrews PR; Craik DJ; Martin JL. *J Med Chem.* 1984, 27(12), 1648-1657.

[17] www.pubchem.com

[18] Shahlaei M; Fassihi A; Saghaie L. *Eur J Med Chem.*, 2010, 45(4), 1572-1582.

[19] Zahra Garkani-Nejad; Fereshteh Saneie. *Bull. Chem. Soc. Ethiop.*, 2010, 24(3), 317-325.

[20] Du Q, Mezey PG; Chou KC. *J Comput Chem.*, 2005, 26(5):461-470.

[21] Huuskonen J., Salo M., Taskinen J., *J. Pharm. Sci.*, 86, 450—454 (1997).

[22] P.J. Taylor, C. Hansch, P.G. Sammes, J.B. Taylor, Comprehensive Medicinal Chemistry, Pergamon Press, Oxford, 1990.

[23] I. Moriguchi, S. Hirono, I. Nakagome, H. Hirano, *Chem. Pharm. Bull.* 42 (1994) 976.

[24] W.M. Meylan, P.H. Howard, J. Pharm. Sci. 84 (1995) 83.

[25] V.K. Gombar, K. Enslein, *J. Chem. Inf. Comput. Sci.* 36 (1996) 1127.

[26] S.C. Basak, B.D. Gute, G.D. Grunwald, *J. Chem. Inf. Comput. Sci.* 36 (1996) 1054.

[27] J.J. Huuskonen, D.J. Livingstone, I.V. Tetko, *J. Chem. Inf. Comput. Sci.* 40 (2000) 947.

[28] I.V. Tetko, V.Y. Tanchuk, A.E.P. Villa, *J. Chem. Inf. Comput. Sci.* 41 (2001) 1407.

[29] L. Molnar, G.M. Keseru, A. Papp, Z. Gulyas, F. Darvas, *Bioorg. Med. Chem. Lett.* 14 (2004) 851.

[30] A.F. Duprat, T. Huynh, G. Dreyfus, *J. Chem. Inf. Comput. Sci.* 38 (1998) 586.

[31] J. Ghasemi, S. Shahmirani, E.V. Farahani, *Anal. Chim.*96 (2006) 327.

[32] J. Ghasemi, S. Saaidpour, S.D. Brown, *J. Mol. Struct.* (Theochem.)805 (2007) 27.

[33] J. Ghasemi, S. Saaidpour, *Chem. Pharm. Bull.*55 (2007) 669.

[34] J. Ghasemi, Sh. Ahmadi, *Anal. Chim.*97 (2007) 69.

[35] J. Ghasemi, S. Asadpour, A. Abdolmaleki, *Anal. Chim.Acta* 588 (2007) 200.

[36] S. Qanei Nassab, Z. Bayat, J. Movaffagh, *J. Chem. Pharm. Res.*, 2011, 3(1):64-71

[37] Z. Bayat and S. Vahdani, *J. Chem. Pharm. Res.*, 2011, 3(1):93-102

[38] D.C. Young, Computational Chemistry, John Wiley & SonsInc., 2001.

[39] Pomona college medicinal chemistry project, Claremont, CA91711, LogP database, (C.Hansh and A.Leo), july 1987 edition

[40] E. Frich, P.B. Jones, M.Roed, T.Skovsgaard and N.I.Nissen, *Biochem, pharmacol*, 39(11),

a. 1721-1726(1990) [41] I.Faccheti and A.Vigevani, *pharmacochem.library* 10,138-140(1987)

[41] C.Hansh, A.Leo and D.Hokman, American chemical society,Washington(1995)

[42] E.M.Action, G.L.Tong, D.L.Taylor, D.G.Streeter,J.A.Fillibi and R.L.Wolgemuth, *J.Med.Chem*, 29(10),2074-2079(1986)

[43] Z. Bayat and M. Fakoor Yazdan Abad, *J. Chem. Pharm. Res.*, 2011, *3(1):48-55*

[44] Z. Bayat, S. Qanei Nassab, *J. Chem. Pharm. Res.*, 2010, 2(6):306-315

[45] C.j.Coulson, and V.J. Smith, *J.Pharm.Sci*, 69(7),799-801(1980)

[46] Hoffman, H.G.Berscheid, D.Borttger, H.H.Sedlacek an H.P.Kraemer,*j.med.chem*,33(1),166-171(1990)

[47] E.L.Gabbay, D.Grier, R.E.Fingerle, R.Reimer, R.LevyW.D.Wilson, Biochemistry, 15(10), 2062-2070(1976)

[48] G. Espinosa, D. Yaffe, Y. Cohen, A. Arenas, F. Giralt, *J. Chem. Inf. Comput. Sci.* 2000, 40,859–879.

[49] Todeschini R, Consonni V. Handbook of Molecular Descriptors. Weinheim, Germany: Wiley-VCH, 2000;1–667.