Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2019, 11(1): 1-11

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

2D-QSAR Study of New 4β-Anilino- 4'-O-demethyl-4-desoxypodophyllotoxin Derivatives as Potential Antitumor Agents

Rajashree Chavan*, Sonal Phule and Kokat Jayashree

Department of Pharmaceutical Chemistry, Pune District Education Association's Seth Govind Raghunath Sable College of Pharmacy, Saswad, Pune- 412 301, India

---

ABSTRACT

In the search of newer and potent antitumor agents, a series of 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives was subjected to 2D Quantitative Structure Activity Relationship (QSAR) analysis. The statistically significant models were generated and the most robust model for 2D QSAR was obtained using random regression method coupled with stepwise forward backward method using V-Life Molecular Design Suite software version 3.5. The physicochemical descriptors, viz., Most +ve &-vePotential distance, vdWSurface Area, SAMostHydrophilic contributed significantly to the biological activity. About 22 QSAR models were generated, among which three significant models were finally selected on the basis of various statistical parameters such as squared correlation co-efficient (r2), and cross-validated square correlation co-efficient (q2). The statistical values of the three significant models viz. model 1, model 2 and model 3 are as r2 to be 0.9320, 0.9576 and 0.9403 respectively and q2 to be 0.8540, 0.8801 and 0.8137 respectively. The descriptors showed by QSAR study can be used further for studying and designing of new compounds. Consequently, this study may prove to be helpful in the development and optimization of existing antitumor activity of this class of compounds.

Keywords: Antitumor; 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives; Random regression; QSAR

INTRODUCTION

Cancer is a major worldwide health problem. Although there has been progress in the development of treatment and prevention for cancer, this disease remains the second major cause of death in the world. Still, the successful treatment of cancer remains a challenge in the 21st century and there is a need to search for newer and safer anticancer agents that have broader spectrum of cytotoxicity to tumor cells [1,2].

Podophyllotoxin derivatives are used clinically against various types of cancers including breast cancer, testicular cancer, small cell lung cancer, lymphoma, Kaposi's sarcoma and childhood leukemia [3-5]. Although being used extensively in clinic, these derivatives exhibited several toxic side effects such as bone marrow depression, increased risk of secondary acute myelogenous leukemia, acquired drug resistance and poor water solubility which

block their further applications [6]. Thus, Wang et al, in their efforts to overcome drug resistance and develop more active, less toxic podophyllotoxin derivatives, synthesized a series of 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives and evaluated them for biological activities. Most of the synthesized compounds exhibited cytotoxic activities against cancer cell lines [7].

The computer aided prediction of biological activity in relation to the chemical structure of a compound is now commonly used technique in drug discovery. Moreover, understanding the QSAR of known compounds helps to facilitate the new drug discovery. Computational chemistry represents molecular structures as a numerical model and simulates their behavior with the equations of quantum and classical physics. The available programs enable scientists to easily generate and present molecular data including geometries, energies and associated properties (electronic, spectroscopic and bulk). The usual paradigm for displaying and manipulating these data is a table in which compounds are defined by individual rows and the molecular properties (or descriptors) are defined by the associated columns [8-10].

QSAR attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity (% activity, IC_{50}, ED_{50}, MIC) for a series of compounds to generate a mathematical expression so that these rules can be used to evaluate new chemical entities. The mathematical expression can then be used to predict the biological response of other chemical structures. 2D-QSAR models are based on descriptors derived from a two-dimensional graph representation of a molecule.

A QSAR generally takes the form of a linear equation:

$$\text{Biological Activity} = \text{Const} + (C_1 \times P_1) + (C_2 \times P_2) + (C_3 \times P_3) + \ldots$$

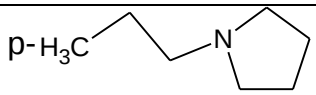
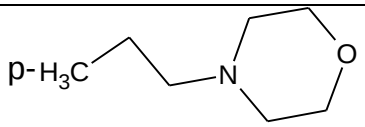
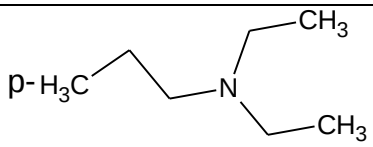
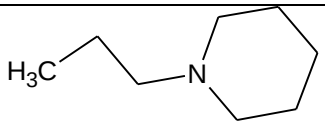
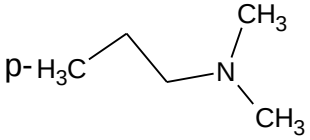
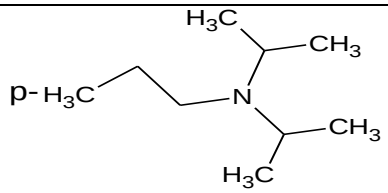
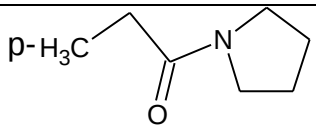
Where the P_1 to Pn are physicochemical parameters value computed for each molecule in the series and C_1 to Cn are the coefficients of parameters. Physicochemical descriptors are based on the physicochemical properties of molecule.

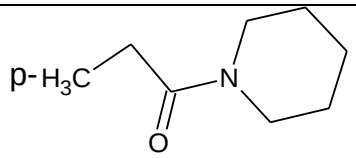
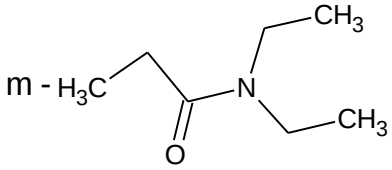
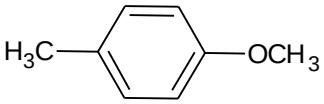
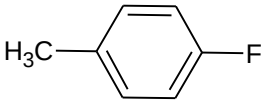
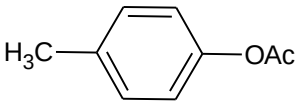
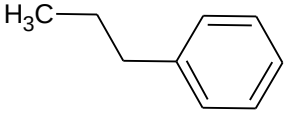
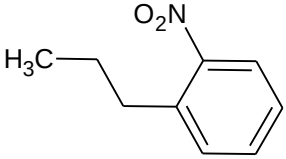
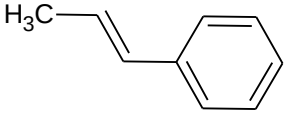
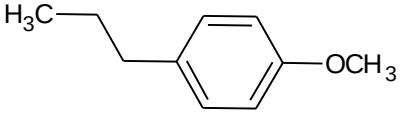
In the present research work, a series of 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives was subjected to 2D quantitative structure activity relationship (QSAR) analysis, for further development of newer antitumor agents. The 2D QSAR study was performed using V-Life Molecular Design Suite software version 3.5.

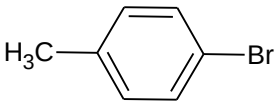
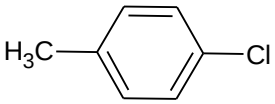
EXPERIMENTAL SECTION

A data set of 19 compounds was taken from the published series as reported. The antitumor activity of the compounds was reported in MIC values. The structure and antitumor activity data of compounds is listed as in Table 1.

Table 1. The chemical structure and biological activity data of compounds

Compound	Structure	MIC	-log MIC
1a		1.57	0.19589
1b		1.27	0.10380
1c		1.53	0.18469
1d		19.1	1.28103
1e		3.95	0.59659
1f		2.20	0.34242
1g		2.20	0.34242

1h		2.33	0.36735
1i		27.2	1.43456
2a		9.28	0.96754
2b		5.31	0.72509
2c		8.48	0.92839
2d		13.7	1.13672
2e		3.70	0.56820
2f		5.30	0.72427
2g		2.80	0.44715
2h		3.08	0.48855

2i		5.72	0.75739
3a		14.21	1.15259

QSAR Study

All the 2D descriptors were calculated for QSAR analysis using Vlife MDS 3.5 software. Thermodynamic parameters describe free energy change during drug receptor complex formation. Spatial parameters are the quantified steric features of drug molecules required for its complimentary fit with receptor. Electronic parameters describe weak non-covalent bonding between drug molecules and receptor. Random regression method is used to generate QSAR equation. For variable selection, stepwise forward-backward method was used.

Criteria for selection of model

n=number of molecules (>15 molecules)

K=number of descriptors in a model (statistically n/5descriptors in a model)

df=degree of freedom (n-k-1) (higher is better)

r^2=coefficient of determination (> 0.7)

q^2=cross-validated r^2 (>0.5)

pred_r^2=r^2 for external test set

F-test=F-test for statistical significance of the model (higher is better, for same set of descriptors and compounds).

Selected models

2D-QSAR investigations of 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives was performed. About 22 QSAR models were generated by using random regression method coupled with stepwise forward-backward method. Among various models, three significant QSAR models were finally selected (Table 2).

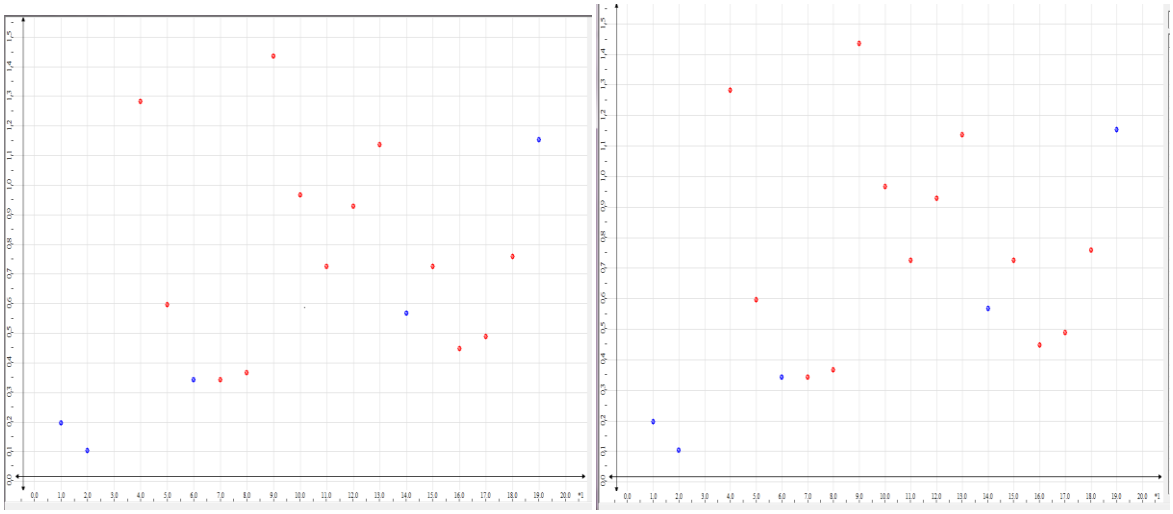
Table 2. Statistics of three models

Model no.	**r^2**	**q^2**	**r^2 sec**	**q^2 sec**	**pred_r^2**	**pred_r^2sec**	**N**	**DF**	**F-test**
1	0.9320	0.8540	0.0881	0.1402	1.0466	0.7726	14	9	30.5831
2	0.9576	0.8801	0.0922	0.1551	1.2409	1.0560	14	9	50.8143

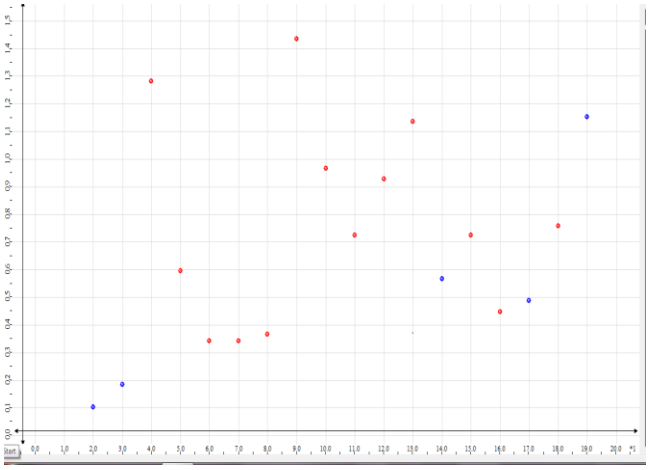
3	0.9403	0.8137	0.1116	0.1971	1.0985	0.7731	14	9	35.4503

RESULTS AND DISCUSSION

For QSAR analysis, regression was performed using MIC values as dependent variables and calculated parameters as independent variables. In any investigation of the effects of molecular properties, it is essential to prove that the results are statistically valid as shown in Figure 1.



(a) (b)

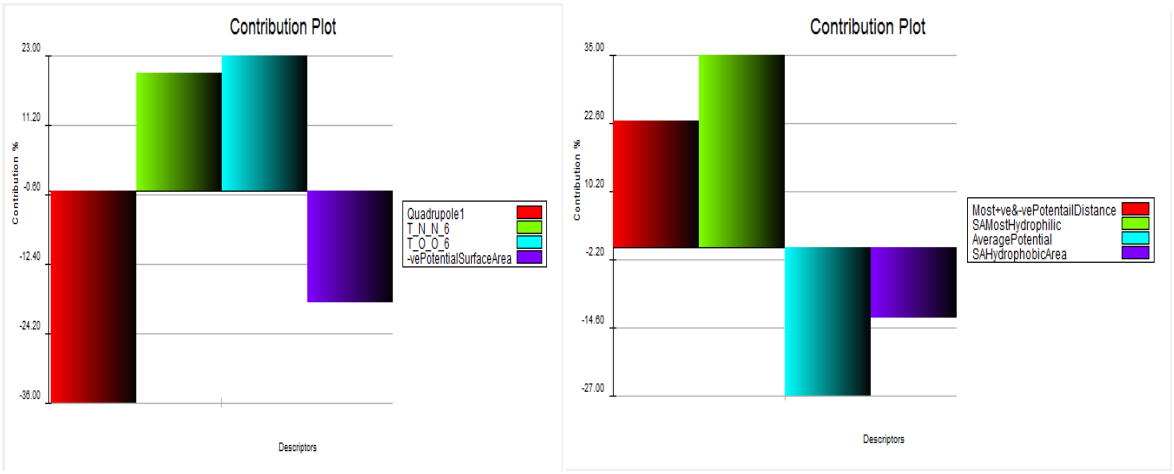


(c)

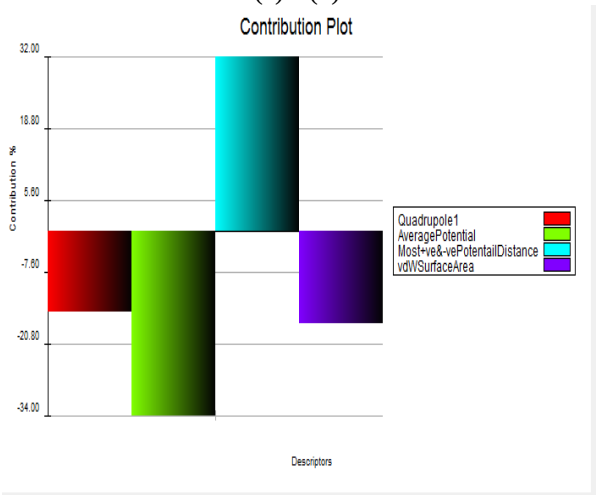
Figure 1. Activity distribution Graph: a) Model 1: all the test set is covered by the training set. b) Model 2: all the test set is covered by the training set. c) Model 3: potent compound is not kept in the test set

Contribution Chart (Figure 2) signifies that the descriptors below the zero line have negative contribution and above the zero line have positive contribution. In model 1, T_O_O_6 is the important descriptor; in model 2,

SAMostHydrophilic is the important descriptor and in the model 3, Most +ve &-vePotential distance is the important descriptor.



(a) (b)



(c)

Figure 2. Contribution charts a) model 1 b) model 2 c) model 3

- Quadrupole1: This descriptor signifies magnitude of first tensor of quadrupole moments.
- -vePotential Surface Area: This descriptor signifies magnitude of first tensor of quadrupole moments.
- Average Potential: This descriptor signifies average of the total electrostatic potential on van der Waals surface area of the molecule.
- Most +ve and -vePotential distance: This descriptor signifies the distance between points having the highest value of +ve and highest value of –ve electrostatic potential on van der Waals surface area of the molecule.
- vdWSurface Area: This descriptor signifies total van der Waals surface area of the molecule.

- SAMostHydrophilic: Most hydrophilic value on the vdW surface. (By Audry Method using Slogp).
- SAHydrophobicArea: vdW surface descriptor showing hydrophobic surface area. (By Audry Method using Slogp).
- T_N_N_6: It determines the distance between two nitrogen atoms by 6 bonds.
- T_O_O_6: It determines the distance between two oxygen atoms by 6 bonds.

Contribution values:

Model 1: Quadrupole1: -36.75%, T_N_N_6: 20.32%, T_O_O_6: 23.07%, -vePotentialSurface Area: -19.86%

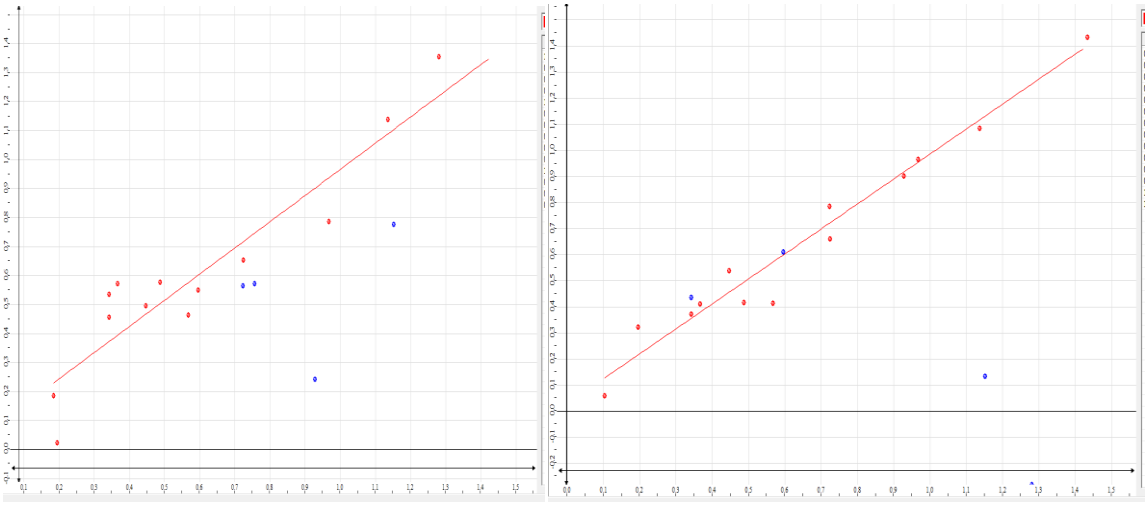
Model 2: Most+ve and -vePotential Distance: 23.98%, SAMostHydrophilic:35.05, Average Potential: -27.68%, SAHydrophobicArea: -13.30%

Model 3: Quadrupole1:-15.32%, AveragePotential: -34.47%, Most+ve&-vePotential Distance: 32.70%, vdWSurfaceArea: -17.51%

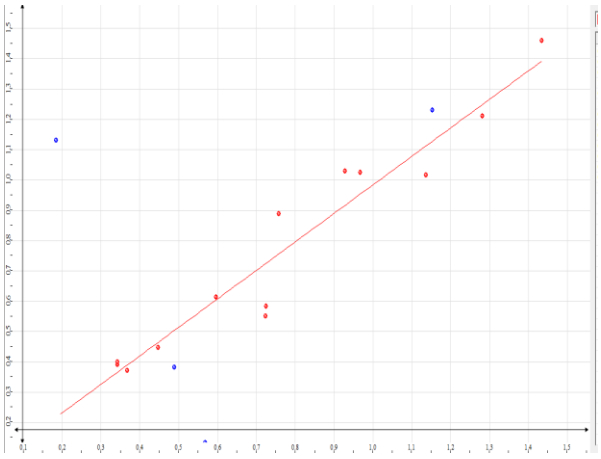
The unicolumn statistics supports the suitability of the selection of test and training set and the average should always be more than the sum (Table 3). Figure 3 shows that the test and training set lie in the plane of fitness plot and these compounds have better predicted activity (Table 4). Most +ve &-vePotential distance is the important descriptor for the model as it is increasing the antitumor activity.

Table 3. Uni-column statistics

Model No.	Column	Average	Max	Min	Standard deviation	Sum
Model 1	Test	0.6484	1.4346	0.1847	0.4044	9.0782
	Training	0.7333	1.1526	0.1038	0.3908	3.6664
Model 2	Test	0.6562	1.4346	0.1038	0.3727	9.1873
	Training	0.7115	1.2810	0.1847	0.4863	3.5573
Model 3	Test	0.7319	1.4346	0.1959	0.3800	10.2468
	Training	0.4996	1.1526	0.2038	0.4144	2.4978



(a) (b)



(c)

Figure 3. Fitness plot a) model 1, b) model 2 c) (Red spot-Training set; Blue spot-Test set)

Equations:

Model 1: sMIC= - 0.0491(±0.0041) Quadrupole1+ 0.6104(± 0.1578) T_N_N_6+ 0.5102(±0.1163) T_O_O_6- 0.0046(±0.0001) -ve Potential Surface Area -0.5597

Model 2: sMIC=+ 0.1322(± 0.0002) Most +ve & -ve Potential Distance+58.7922(±3.4767) SA Most Hydrophilic-149.3640(±15.3396) Average Potential - 0.0025(± 0.0000) SA Hydrophobic Area+5.8658

Model 3: sMIC= - 0.0157(±0.0000) Quadrupole1- 166.9854(±24.2990) Average Potential+0.1043(± 0.0002) Most +ve & -ve Potential Distance- 0.0027(± 0.0000) vdW Surface Area+ 1.8109

From the 19 analogues used for QSAR studies, 2i, 2a and 1e were found to be the most potent compounds with minimal residual activity in model 1, model 2 and model 3, respectively (Table 4).

Table 4. Actual and predicted values of the compounds as predicted by QSAR study

Compoun	Model 1	Model 2	Model 3

d	A*	B*	C*	A	B	C	A	B	C
1a	34560	1.301960	0.132600	57390	42099	)84709	95890	63174	-0.067284
1b	42420	72317	)29894	68200	63725	04475	57390	88782	-0.131392
1c	03800	59001	44799	67540	84422	83118	42420	91470	-0.049050
1d	28390	99641	28749	42420	54714	112294	81030	10090	0.070940
1e	95890	20075	124185	36720	36720	00000	47150	47620	-0.000470
1f	25090	58922	66168	47150	94836	)47686	34560	59730	-0.025170
1g	68200	13393	54807	96590	50061	46529	24270	50321	0.173949
1h	88550	14910	73640	42420	34656	192236	67350	70912	-0.003562
1i	67540	62521	05019	67350	70404	203054	67540	25140	-0.057600
2a	24270	83055	)58785	88550	75350	)86800	28390	29470	-0.101080
2b	47150	38301	)91151	81030	53770	)7274	42420	98773	-0.056353
2c	67350	08950	)41600	95890	21423	74467	25090	82899	0.142191
2d	34560	31750	02810	25090	52353	72737	96590	12933	-0.016343
2e	42420	99668	)57248	70404	67350	03054	67540	62521	0.005019
2f	36720	82390	54330	84690	83806	00884	36720	15500	0.121220
2g	84690	69540	51500	28390	40130	88260	39500	39700	0.099800
2h	96590	09523	)12933	24270	62579	61691	30990	84690	0.946300
2i	42420	34509	)92089	57390	71248	86142	88550	81594	0.169560
3a	52590	32213	20377	52590	75530	77060	52590	30940	0.2165 0

*A: Actual activity, B: Predicted activity, C: Residual activity

CONCLUSION

Among the generated QSAR models; three models were selected on the basis of various statistical parameters such as squared correlation co-efficient (r^2) which is relative measure of quality of fit, Fischer's value (F test) which represents fraction between the variance of calculated and observed activity, standard error (r^2_se) representing absolute measure of quality of fit, cross-validated square correlation coefficient (q^2), standard error of cross-validated square correlation co-efficient (q^2_se), predicted squared regression (pred_r^2) and standard error of predicted squared regression (pred_r^2se) to estimate the predictive potential of the models.

From the derived QSAR model, it can be concluded that antitumor activity of 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin derivatives is strongly influenced by physicochemical descriptors like Quadrupole1, -vePotential Surface Area, Average Potential, Most +ve &-vePotential distance, vdWSurface Area, SAMostHydrophilic and SAHydrophobicArea. Thus on increasing the value of the above stated descriptors, the antitumor activity will increase. T_N_N_6, T_O_O_6 were the descriptors which showed that they have less importance for antitumor activity and on decreasing their values, the activity will improve. The descriptors shown by QSAR study can be used further for studying and designing of new compounds. Consequently, this study may prove to be helpful in development and optimization of existing antitumor activity of this class of compounds.

ACKNOWLEDGMENT

We are greatly thankful to all those persons who gave us their valuable support for carrying out our QSAR study on 4β-anilino- 4'-O-demethyl-4-desoxypodophyllotoxin for developing better analogues with enhanced antitumor activity.

REFERENCES

[1] P Anand; AB Kunnumakkara; C Sundaram; KB Harikumar; ST Tharakan; OS Lai; B Sung; BB Aggarwal. *Pharm Res*. **2008,** 25 (9), 2097-2106.

[2] J Chuanfei; L Yong-Ju; H Hongwu; F Liwu. *Eur J Med Chem*. **2011,** 46, 429-432.

[3] D Guianvarc'h; M Duca; C Boukarim; L Kraus-Berthier; S Leonce; A Pierre; B Pfeiffer; P Renard; PB Arimondo; C Monneret; D Dauzonne. *J Med Chem*. **2004,** 47, 2365-2374.

[4] EL Baldwin; N Osheroff. *Current Med Chem, Anti-Cancer Agents*. **2005,** 5, 363-372.

[5] A Kamal; BA Kumar; M Arifuddin. *Tetrahedron Letters.* **2003,** 44, 8457-8459.

[6] YJ You. *Current Pharm Design*. **2005,** 11, 1695-1717.

[7] L Wang; F Yang; X Yang; X Guan; C Hu; T Liu; Q He; B Yang; Y Hu. *Eur J Med Chem*. **2011,** 46(1), 285-296.

[8] A Katritzky. *Bioorganic Med Chem Letters.* **2010,** 45, 5183-5199.

[9] V Rapatri; T Chitre; K Bothara. *Eur J Med Chem*. **2009,** 44, 3954-3960.

[10] VLife Molecular Design Suite version 3.5; VLife Sciences Technologies Pvt. Ltd., **2010,** Pune, India.