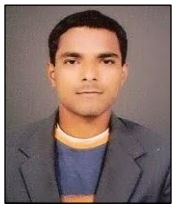
PHARMATUTOR
PHARMACY INFOPEDIA
PRINT ISSN: 2394-6679 | E-ISSN: 2347-7881

2D QSAR study on Saponins of Pulsatilla koreana as an Anticancer agent



Sagar Alone*, Sharda Deore, Bhushan Bawiskar
Department of Pharmacognosy & Phytochemistry
Govt. College of pharmacy, Amravati, Maharashtra, India
sagaralone123@gmail.com

ABSTRACT

Total seventeen saponins previously isolated from roots of *Pulsatilla koreana* having cytotoxic activity against 4 different cancer cell line (A-549, SK-OV-3, SK-MEL-2, HCT15) were used for 2D QSAR using V-life Molecular design suit. Using multiple linear regression method against 4 different cell lines develops QSAR model. QSAR model was generated by using training set of 11 and test set of 6 molecules having correlation coefficient (r^2), significant cross validated correlation coefficient (q^2) and F-test (For statistical significance) is as given below (A-549: r^2- 0.9281, q^2- 0.8691, F-test- 51.6079), (SK-OV-3: r^2- 0.9554, q^2- 0.9184, F-test- 85.7357), (SK-MEL-2: r^2- 0.9160, q^2- 0.8285, F-test- 43.6084), (HCT15: r^2- 0.9203, q^2- 0.8357, F-test- 46.1887). In this QSAR study Alignment independent descriptors such as T_2_C_7, T_O_O_5 and physicochemical descriptors like Chain path count such as 6 chain count and Chi chain such as Chi 6 chain were most responsible descriptors for anticancer activity.

Keywords: QSAR; Pulsatilla koreana; cytotoxic activity; multiple linear regressions

INTRODUCTION

Cancer is a disease of cell characterized by progressive, persistent, abnormal, purposeless, and uncontrolled proliferation of tissues. Currently, cancer is most dominating cause of death in world [1, 2]. *Pulsatilla koreana* belongs to the family *Ranunculaceae* and is an endemic species in Korea containing 17 saponins in which eight lupane-type (1, 3, 5, 7, 9, 11, 13, 15) and nine oleanane-type (2, 4, 6, 8, 10, 12, 14, 16, 17), isolated from *Pulsatilla koreana* roots. The roots of this plant have been widely used in traditional medicine for the treatment of several diseases, particular malaria and amoebic dysentery [3]. This plant was evaluated for their cytotoxic activity against four human solid tumor cell lines (A-549, SK-OV-3, SK-MEL-2, and HCT-15). ED50 values was observed that the saponins 5—17, possessed free carboxylic group at C-28, exhibited moderate to considerable cytotoxic activity (ED_{50}; 1.57—174.34m M) against tumor lines, whereas their sugar-bonded esters, that is, disdesmoside saponins 1, 2, 3, and 4 were inactive [1] (ED_{50} >300m M). Saponins molecules isolated from the *Pulsatilla koreana* as anticancer agent are taken from literature for QSAR analysis [4].

ED_{50} value was defined as the concentration of compound needed to reduce a 50% of absorbance relative to the Quantitative structure activity relationship (QSAR) gives information relating chemical to Biological activities by developing a QSAR model. Different molecular descriptors are used to determine the structural feature of lead molecule. The purpose of using QSAR descriptors is to calculate the properties of molecules that serve as numerical descriptions, using such an approach one could predict the activities of newly designed compounds before a decision is being made whether these compounds should be really synthesized and tested.

With respect to the above subjects and scope QSAR study is performed on *Pulsatilla koreana* saponins analogues in order to get a better understanding of their structural features and anticancer activity.

How to cite this article: S Alone, S Deore, B Bawiskar; 2D QSAR study on Saponins of Pulsatilla koreana as an Anticancer agent; PharmaTutor; 2015; 3(6); 24-28

PHARMATUTOR
PHARMACY INFOPEDIA
PRINT ISSN: 2394-6679 | E-ISSN: 2347-7881

MATERIAL AND METHOD

2D QSAR Methodology

Data set:

17 saponins molecules isolated from the *Pulsatilla koreana* as anticancer agent are taken from literature for QSAR analysis. [4] 2D Structure of all above 17 saponins analogues molecules are sketched by using 2D structure drawing function of VLife MDS system and all 2D structures are converted to 3D by using VLife MDS (Molecular Design Suite) ™ 3.5 software supplied by VLife Sciences Technologies Pvt. Ltd., Pune, India. Batch optimization of 3D structure was done by using MMFF. To get more minimization we input 100000 cycles with the converse criteria (RMS gradient) of 0.01. The distance dependent function was kept at 1.0 with analytical gradient type. The 2D-QSAR models were generated for this series using multiple linear regression (MLR) against 4 different cancer cell lines.

Selection of molecules in the training set and test is a key and important feature of any QSAR model. Therefore, the care was taken in such a way that biological activity of all compounds in test lie within the maximum and minimum value range of biological activities of training set of compounds. The Uni-Column Statistics of test and training sets further reflected the correct selection of test and training sets. A Uni-Column statistics for training set and test set were generated to check correctness of selection criteria for trainings and test set molecules those best models which come out with promising results are discussed here. QSAR models were generated by a training set of 11 molecules for each model. And a test set of 6 molecules with uniformly distributed biological activities. The structures of all the compounds along with their actual and predicted biological activities are presented in Table 1.

Table 1: Experimental and Predicted activity of *Pulsatilla koreana* saponins analogues

Sr.No.	Saponins	ED50 (µM)							
		A-549		Sk-OV-3		SK-MEL-2		HCT15	
		Expt.	Pred.	Expt.	Pred.	Expt.	Pred.	Expt.	Pred.
1	PK01	>300.0	341.798	>300.0	333.826	>300.0	340.972	>300.0	340.723
2	PK02	>300.0	310.389	>300.0	313.994	>300.0	294.399	>300.0	294.808
3	PK03	>300.0	317.875	>300.0	308.737	>300.0	311.381	>300.0	311.289
4	PK04	>300.0	286.467	>300.0	288.905	>300.0	372.206	>300.0	272.733
5	PK05	38.14	80.1154	36.27	73.2989	37.44	119.877	40.06	122.038
6	PK06	11.25	17.2984	13.27	20.1392	3.04	11.9355	11.86	15.4923
7	PK07	145.62	104.038	120.38	93.2087	167.82	127.275	174.34	129.396
8	PK08	13.27	17.2984	11.41	14.9598	12.16	19.3334	13.58	22.8506
9	PK09	135.28	104.038	114.32	93.2087	165.54	112.479	171.29	114.68
10	PK10	2.56	17.2984	2.31	14.9598	1.57	4.53761	8.36	8.13401
11	PK11	43.64	80.1154	38.90	68.1195	39.67	90.2853	49.04	92.6047
12	PK12	13.49	-6.624	13.71	-10.1294	14.12	-17.6561	14.17	-13.9409
13	PK13	155.32	104.038	124.12	93.2087	145.68	127.275	138.73	129.396
14	PK14	4.24	17.2984	3.95	14.9598	3.47	19.3334	5.50	22.8506
15	PK15	41.35	80.1154	38.91	68.1195	40.02	75.4896	40.86	77.8882
16	PK16	9.58	-6.624	11.39	-10.1294	10.37	-32.4518	12.74	-28.6574
17	PK17	10.73	41.2208	10.66	40.049	9.81	48.925	14.09	52.2837

Expt. - Experimental activity
Pred. - Predicted activity

PHARMATUTOR
PHARMACY INFOPEDIA
PRINT ISSN: 2394-6679 | E-ISSN: 2347-7881

QSAR Study:

All types of 2D Physicochemical descriptors including Individual, Chi chain, Path count, Chiv, Element count and Kappa categories and Alignment independent including topological structure descriptors were calculated for QSAR analysis using Vlife MDS software. Multiple linear regression method is used to generate QSAR equation. For variable selection, stepwise forward-backward method was used. For validating the quality of the models. Selection of molecules in the training set and test is a key and important feature of any QSAR model. Therefore, the care was taken in such a way those biological activities of all compounds in test lie within the maximum and minimum value range of biological activities of training set of compounds. The Uni-Column Statistics of test and training sets further reflected the correct selection of test and training sets. A Uni-Column statistics for training set and test set were generated to check correctness of selection criteria for trainings and test set molecules.

Table 2: Uni-column statistics of the training and test sets for QSAR models:

Cell line	Training/Test set	Average	Max	Min	Std. Dev	Sum
A-549	Training set	94.1109	300.000	4.2400	114.9137	1035.2200
	Test set	131.5417	300.000	2.5600	138.64.97	789.2500
SK-OV-3	Training set	88.6536	300.0000	3.9500	112.7344	975.1900
	Test set	127.3850	300.0000	2.3100	139.3976	764.3100
SK-MEL-2	Training set	94.0000	300.0000	3.0400	116.5289	1034.0000
	Test set	136.1183	300.0000	1.5700	139.8473	816.7100
HCT-15	Training set	96.4845	300.0000	5.5000	114.9111	1061.3300
	Test set	138.8817	300.0000	8.3600	138.1815	833.2900

Table 3: Statistics of the Models:

Statistics	A-549	SK-OV-3	SK-MEL-2	HCT-15
N	**11**	**11**	**11**	**11**
Degree of freedom	8	8	8	8
r2	0.9281	0.9554	0.9160	0.9203
q2	0.8691	0.9187	0.8285	0.8357
F test	51.6079	85.7357	43.6084	46.1887
r2 se	34.4579	26.6109	37.7639	36.2697
q2 se	46.4909	35.9315	53.9485	52.0702
pred_r2	0.9538	0.9686	0.8950	0.8899
pred_r2se	31.0692	25.8051	47.7110	48.3659

Here:

N - Number of molecules,
K - Number of descriptors in a model,
DOF - Degree of freedom (higher is better),
r2 - Coefficient of determination (> 0.7),
q2 - Cross-validated r (>0.5),
pred_r2 - r for external test set (>0.5),
F-test - F-test for statistical significance of the model (higher is better, for same set of descriptors and compounds),
r2_Se, q2_se, pred_r2_se = error for r2, q2, pred_r2 respectively.

Equations:

For (A-549):

ED_{50}= + 55.3309(± 6.5096) T_2_C_7 + 23.9224(± 8.9012) 6ChainCount - 616.7280

PHARMATUTOR
PHARMACY INFOPEDIA
PRINT ISSN: 2394-6679 | E-ISSN: 2347-7881

For (SK-OV-3):
ED_{50}= + 53.2371(± 0.5713) T_2_C_7 + 553.1007(± 130.2520) chi6chain -612.2515

For (SK-MEL-2):
ED_{50} = + 53.9707(± 7.2274) T_2_C_7 + 7.3979(± 2.5587) T_O_O_5 - 493.8088

For (HCT15):
ED_{50} = + 53.2728(± 6.9414) T_2_C_7 + 7.3583(± 2.4574) T_O_O_5 - 484.2732

VALIDATION OF QSAR MODEL

Validation of QSAR study is important to test the internal stability and predictive ability of the QSAR models & was validated by the following procedure as given below. There are two types of validation

1) Internal validation

2) External validation

Internal validation:

It was carried out using leave-one-out (q2, LOO) method. For calculating q2, each molecule in the training set was eliminated once and the activity of the eliminated molecule was predicted by using the model developed by the remaining molecules. The q2 was calculated using the equation (Eq. 2), which describes the internal stability of a model.

$$q^2 = 1 - \frac{\sum [yi(Act) - yi(Pred)]^2}{\sum [yi - y\,(mean)]^2}$$

Where *yi (Act)* and *yi (Pred) are* the actual and predicted activity of the *i*th molecule in the training set, respectively, and y mean is the average activity of all molecules in the training set.

External validation:

The predictive ability of the selected model was also confirmed by external validation of test set compounds, which is also denoted with pred_r2. The pred_r2 value is calculated as follows

$$pred\_r2 = 1 - \frac{\sum [yi(Act) - yi(Pred)]^2}{\sum [yi - y\,(mean)]^2}$$

Where *yi* and *yi* are the actual and predicted activity of the *I* th molecule in the training set, respectively, and y mean is the average activity of all molecules in the training set. We have done both Internal & External validation with these formula and following values are obtained.

Table 4: Internal & External validation data:

No	Validation	A-549	SK-OV-3	SK-MEL-2	HCT-15
1	Internal (q2)	0.9281	0.9512	0.8798	0.9209
2	External (pred_r2)	0.9281	0.9512	0.8841	0.9209

RESULT & DISCUSSION

The aim of our study was to evaluate a series of analogs of Pulsatilla koreana saponins by doing 2D QSAR with the help of different descriptors, from the QSAR studies it was found that

- The Alignment Independent category in which T_2_C_7 descriptor is most powerful descriptor responsible for anticancer activity to all 4 types of cell lines at range of 70% - 80% (A-549, SK-OV-3, SK-MEL-2, and HCT15).
- T_O_O_5 descriptor is responsible for anticancer activity in SK-MEL-2 and HCT15 cancer cell line at range 20%- - 30%.
- 6 chain count and chi6chain responsible for anticancer activity for A-549, SK-OV-3 respectively.
- Existence of free carboxylic group is at C-28 position is most responsible for cytotoxic activity and hydroxyl group at C-23 had a negative effect on the cytotoxic activity.
- It is due to electron donating effect of two loan pair towards C-3 of aglycon. The results obtained from this 2D-QSAR study are in agreement with the observed SAR of Pulsatilla koreana saponins studied.
- There is difference in cytotoxic between oleanane and lupane saponins generally cytotoxicity of lupane type saponins were much weaker than those of oleanane type saponins that means oleanane type Pulsatilla koreana saponins analogues having more anticancer activity than that of lupane type.

PHARMATUTOR
PHARMACY INFOPEDIA
PRINT ISSN: 2394-6679 | E-ISSN: 2347-7881

Table 5: Descriptors responsible for anticancer activity

Sr NO	Cancer cell line	Correlation between activities and chemical descriptors	Predicted accuracy of QSAR
1	A-549	92.81%	86.91%
2	SK-OV-3	95.54%	91.84%
3	SK-MEL-2	91.60%	82.85%
4	HCT-15	92.81%	83.57%

Hence, the model proposed in this work is useful and can be employed to design new saponins derivative of *Pulsatilla koreana* as most active anticancer agent.

CONCLUSION

I have evaluated a 17 series of *Pulsatilla koreana* analogues by using 2D QSAR with the help of different descriptors in order to determine the better structural characteristics and to study different descriptors responsible for anticancer activity.

From the QSAR study it was mainly concluded that the alignment independent descriptor category is most responsible for all four types of cancer cell lines which includes

T_2_C_7: This is the count of Number of double bounded atoms (i.e. - Any double bonded atoms, T-2) separated from carbon atom by 7 bonds in a molecule.

T_O_O_5: This is the count of No. of oxygen atoms (Single double or triple bounded) separated from other oxygen by 5 bond distance in a molecule.

6-chain count: These descriptors signify total numbers of six membered rings in a compound.

Chi6chain: These descriptors signify a retention index for six membered rings.

It was concluded that all descriptors have positive correlation with the anticancer activity. Hence increase in all descriptors values will help to increase the anticancer activity of *Pulsatilla koreana* saponin analogues.

Hence we concluded that the model proposed in this 2D QSAR study is very useful. It can be employed to design new saponin derivative of *Pulsatilla koreana* as most active anticancer agent.

FUTURE SCOPE

1) To determine better structural characteristic of Pulsatilla koreanna saponin analogues.
2) To study the different descriptors (3D Descriptors) responsible anticancer activity (3D QSAR Study)
3) To identify all other cancer cell line to which Pulsatilla koreanna saponin analogues exhibits its affinity/inhibition potential.
4) To design new saponin derivative of Pulsatilla koreanna as most active anticancer agent.

Acknowledgment: The authors thanks to entire team of Vlife MDS for their constant support.

REFERENCES

1. Varmus H. The new era in cancer research. Science, 2006 312:1162–11650; doi: 10.1126/science.1126758.
2. Kingston DG, Newman DJ. Taxoids: cancer-fighting compounds from Nature, Drug discovery development, 2007; 10(2); 130-144
3. Bae K. H. The Medicinal Plants of Korea. Kyo-Hak Press, 1999, Seoul, 139
4. Seong-Cheol Bang et al, Antitumor Activity of Pulsatilla koreana Saponins and Their Structure–Activity Relationship; Chem. Pharm. Bull, 2005; 53(11); 1451-1454