International Journal of Pharma and Bio Sciences

QSAR, ADME AND QSTR STUDIES OF SOME SYNTHESIZED ANTI-CANCER 2-INDOLINONE DERIVATIVES

ANKIT K. ROCHANI[1*], B.V. SUMA[1], SURENDAR KUMAR[1], JUDY JAYS[1] AND V. MADHAVAN[2]

*[1]Department of Pharmaceutical Chemistry, M.S. Ramaiah College of Pharmacy, M.S.R.I.T. Post, Bangalore – 560054, India.
[2] Department of Pharmacognosy, M.S. Ramaiah College of Pharmacy, M.S.R.I.T. Post, Bangalore – 560054, India.

Corresponding Author anrochani@gmail.com

ABSTRACT

The present study explores the utility of Quantitative Structure Activity Relationship (QSAR), *in-silico* ADME studies and Quantitative Structure Toxicity Relationship (QSTR) for the established 2-indolinone Lamotrigine Schiff base derivatives. Here, we developed 2D QSAR models for (n=6) 2-indolinone Lamotrigine Schiff base derivatives as cytotoxic agents using the CTC50 values of these compounds obtained by using MTT and SRB bioassay procedure for HEp-2 and DLA cell lines. Multiple regression equations developed using the calculated physicochemical parameters showed that for HEp-2 cell lines, SRB bioassay procedure gave a better correlation between Van der Waals energy, shape flexibility index, surface area and anti-cancer activity (r^2>0.95). Similarly for DLA cell lines, MTT bioassay gave better correlation between HOMO (Highest Occupied Molecular Orbital), LogP, molecular refractivity and anti-cancer activity (r^2>0.99). Also, the *in-silico* ADME and QSTR evaluation showed that structural features of 5a compound had better pharmacokinetic and toxicity profile.

KEYWORDS

QSAR, pharmacokinetic properties, QSTR 4-quinolone, Isatin, Quinolones.

INTRODUCTION

It takes nearly 14 years and approximately 800 million dollars to get a new molecule into the market. Considering that 50% of the compounds fails in preclinical study phases, which leaves unsuitable compounds to progress into clinical testing, great interest has been focused on the determining the pharmacokinetic profile of the new molecules developed prior to its sending for animal or human testing[1]. Important tools for early state studies of drug development are quantitative structure activity relationships (QSAR), *in-silico* ADME studies and quantitative structure toxicological relationship (QSTR) studies, where different types of molecular descriptors have been used to predict the activities and the pharmacokinetic behaviour of untested drugs in the animal body before carrying out the actual experiments[1].

A large number of Isatin derivatives have been reported to have a good anti-bacterial, anti-malarial[4-5], anti-leshminial[6] and anti-cancer[7-8] activities.

The QSAR descriptors used in this study are based on thermodynamic, steric and electronic parameters[2-3]. These parameters

include partition coefficient, molecular volume, surface area, molecular refractivity and others. Also, the structural descriptors which provides information about the various toxicological and pharmacokinetic aspects of the synthesized molecules includes E-state functions, kappa index, Chi index, Lipinski five rules and Wiener index.

ADME screening was carried out in order to understand the pharmacokinetic behaviour of reported derivatives using descriptors like human intestinal absorption (HIA), Lipinski rule of five, blood brain barrier (BBB) penetration, hepato-toxicity, aqueous solubility, CYP450 2D6 inhibition probabilities. This was followed by toxicological evaluation of these derivatives using QSTR (Quantitative Structure Toxicological Relationship) models with an aim of giving rationalised direction for further pharmacological and synthetic chemistry research on these types of molecules.

The purpose of this study is to gain an insight of structural features responsible for anti-cancer activity, to develop *in-silico* toxicity and pharmacokinetic profile for synthesized Isatin derivatives by using QSAR, ADME screening and QSTR studies.

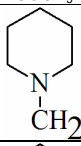
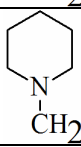
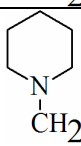
MATERIALS AND METHODS

Softwares used for the above studies are D.S Viewer Pro, TSAR (Oxford soft.), Accord For Excel (v 6.1) and TOPKAT (v 6.2). All these softwares were obtained from Accelrys Inc. CTC50 data for Hep-2 and DLA cell lines for all the (n=6) synthesized derivatives was procured from Asian Journal of Chemistry. QSAR, QSTR and ADME model development and drug design studies were carried out using Hewlett Packard computer systems.

A set of six Schiff bases of 1*H*-Indole-2, 3-dione with amino Lamotrigine was selected from the reported work of G. Nagarajan et al[8]. The biological activity data CTC50 (mentioned in table-1) reported for HEp-2 and DLA cell lines, using MTT and SRB assay procedures were subjected to QSAR, QSTR and ADME screening using Accelrys software modules.

Table 1

Activity data for substituents R and R' of above mentioned structure.

S. No	Compound	R	R'	CTC50 (µg/ml)			
				HEp-2		DLA	
				MTT	SRB	MTT	SRB
1	4a	H	$COCH_3$	479.54	434.85	479.54	447.33
2	4b	Cl	$COCH_3$	346.95	218.02	373.18	427.87
3	4c	Br	$COCH_3$	249.61	438.68	250	428.022
4	5a	H		249.35	438.2	249.35	500
5	5b	Cl		432.78	500	431.87	500
6	5c	Br		105.7	127.56	130.43	217.18

CTC50: Concentration of sample required to kill 50% of the cell.
NTP: National Toxicological Programme.
FDA: Food & Drug Administration

i) QSAR studies

ACD/ChemSketch software was used to draw 2D structures. And these were

transferred to D.S Viewer Pro. in order to convert them to 3D structures. These n=6, 3D structures were converted to .smi file format by using the same software. This .smi data for each derivative was then imported to TSAR program for QSAR studies. The 3D structures imported to TSAR were subjected to energy optimisation. The descriptor values for all the optimized molecules were calculated using "Calculate" module of the program. The data was transferred to the statistical unit of TSAR in order to establish the correlation between physicochemical parameters as independent variables and CTC50 values as dependent variables by employing multiple regression analysis. All the possible combinations of descriptors were considered for QSAR study, only those equations with best correlation coefficient value were selected.

ii) ADME Screening

Here, 2D structures were directly introduced into Accord for Excel for carrying out ADME screening by using "Edit Chemistry" module of software. Now the data for descriptors like BBB (Blood Brain Barrier) penetration and HIA (Human Intestinal Absorption), CYP450 2D6, Lipinski rule of five, plasma protein binding, aqueous solubility and hepatotoxicity were calculated using "Functions" module of software. Descriptor functions like FPSA (Fast Polar Surface Area) and AlogP98 were also used to develop HIA and BBB graphical plots for synthesized derivatives (as shown in figure: 1 and 2)

Figure 1
An ellipse of ADME_ALogP98 vs. ADME_PSA_2D (HIA Plot)

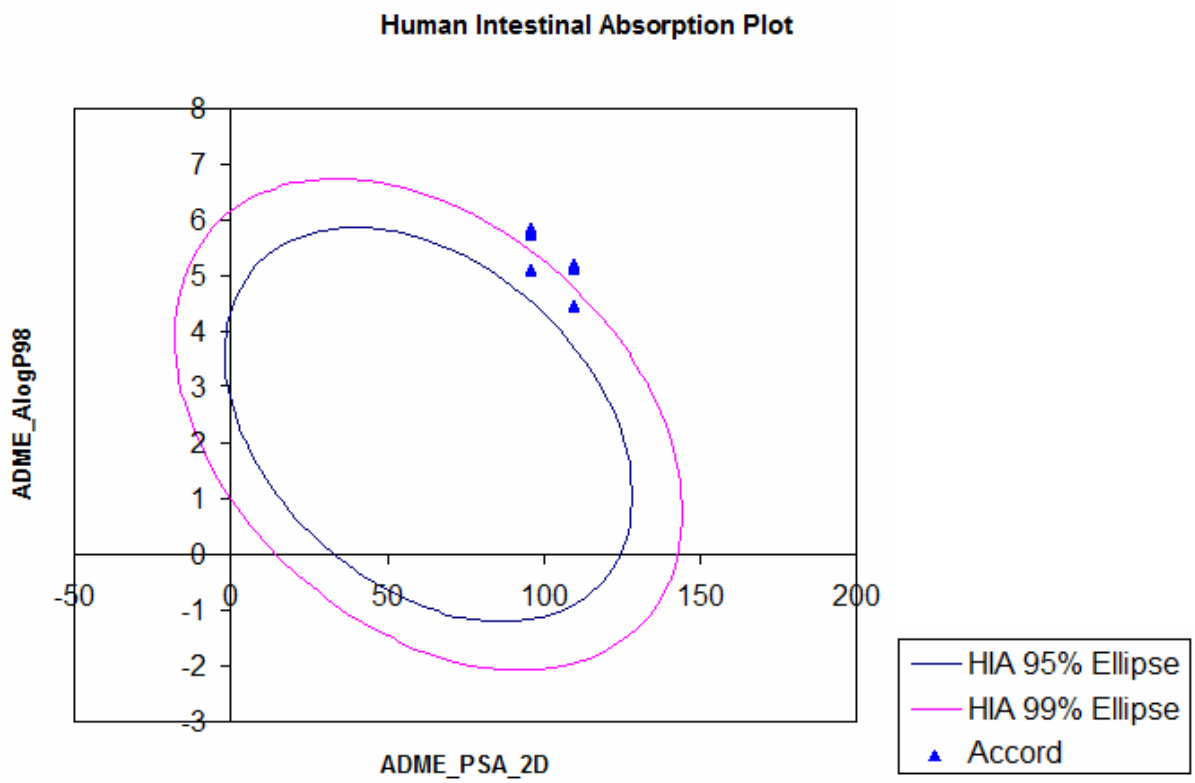
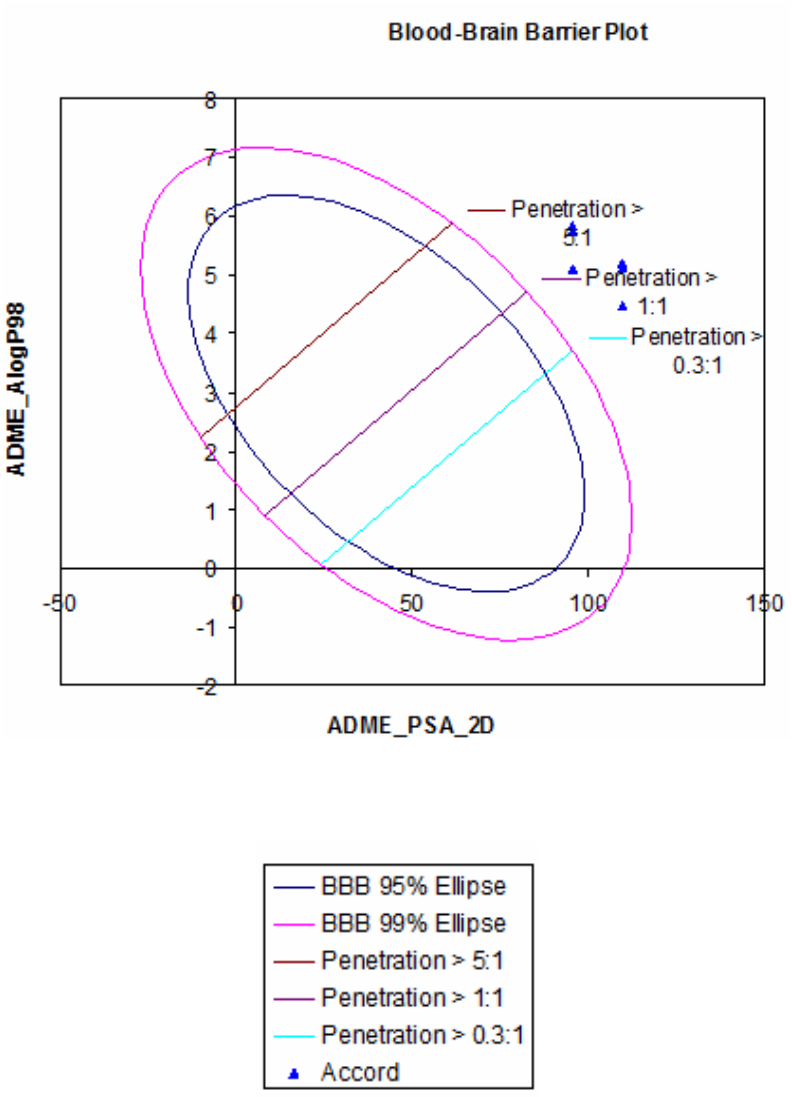


Figure 2

Ellipse of ADME_ALogP98 vs. ADME_PSA_2D (BBB Penetration Plot)



iii) QSTR studies

The smiles notations obtained from D.S Viewer Pro were transferred to TOPKAT for carrying out toxicity studies in various *in-silico* animal models. For the models like Rat LD_{50}, Rat inhalational LC_{50} and LOAEL, values were calculated in the form of dose along with 95% confidence limits. Activity values from Rat LD_{50}, Rat inhalational LC_{50} and LOAEL were subjected to the development of regression equations (same as QSAR studies) using TSAR software. All the descriptor values and results obtained are given in table 4(a-f) and carcinogenicity calls for all the derivatives (as per FDA and NTP norms) were also recorded using TOPKAT (table 5).

Table: 4(a)

Toxicity profile for compounds of 4th and 5th series.

Comp.	Rat LD_{50}		Rat Inhalational LC_{50}		LOAEL	
	Computed values	95% confidence limit	Computed values	95% confidence limit	Computed values	95% confidence limit
4a	572.1 mg/kg	86.0 mg/kg and 3.8g/kg	3.9 $g/m^3/H$	580 $mg/m^3/H$ and > 10$g/m^3/H$	1.4 mg/kg	229.6µg/kg and 8.1 mg/kg
4b	167.7 mg/kg	22.9mg/kg and 1.2g/kg	5.6 $g/m^3/H$	838.8mg/m^3/ H and >10$g/m^3/H$	607.5 µg/kg	89.3µg/kg and 4.1 mg/kg
4c	146.8 mg/kg	20.2mg/kg and 1.1 g/kg	4.7 $mg/m^3/H$	701.3mg/m^3/ H and >10$g/m^3/H$	1.6 mg/kg	265.7µg/kg and 9.5 mg/kg
5a	1.1 g/kg	166.1mg/kg and 7.4 g/kg	16.2 $mg/m^3/H$	1.0$mg/m^3/H$ and 251.2mg/m^3/ H	631.9 µg/kg	103.8µg/kg and 3.8 mg/kg
5b	266.1 mg/kg	39.9mg/kg and 1.8 g/kg	24.7 $mg/m^3/H$	1.6$mg/m^3/H$ and 372.2mg/m^3/ H	250.4 µg/kg	36.7µg/kg and 1.7 mg/kg
5c	222.1 mg/kg	33.5mg/kg and 1.5 g/kg	20.0 $mg/m^3/H$	580 $mg/m^3/H$ and > 10$g/m^3/H$	669.7 µg/kg	229.6µg/kg and 8.1 mg/kg

LD_{50}: Lethal Dose required for killing 50% of animal population
LC_{50}: Lethal Concentration required for killing 50% of animal population.
LOAEL: Lowest-Observed-Adverse-Effect Level.

Table 4(b)
QSTR relationship for Compounds

S. No.	Activity	Equation
1	LD_{50}	B.A = 0.71967334*X_1-0.065252624*X_2-184.15698*X_3 + 5008.6001
2	LOAEL	B.A=600.5401*X_1+16.91551*X_2+586.91998*X_3 -12901.519
3	LC_{50}	B.A= 293.42206*X_1 - 0.44354472*X_2 + 55369.746*X_3 - 563711.19

Table 4(c)
Descriptors for regression equation

S. No.	X_1	X_2	X_3
1.	Total Energy	Heat of Formation	Total Dipole
2.	Cosmic Energy	Molecular Surface Area	Total Lipole
3.	Molecular refractivity	Heat of Formation	Ionisation potential

Table 4(d)

Statistical parameters for QSTR Equations

S. No.	r	r^2	r^2(C.V)	t-probability
1.	0.999941	0.999881	0.987683	< 0.05
2.	0.977649	0.955797	0.426784	< 0.05
3.	0.996359	0.99273	0.627869	< 0.05

Table 4(e)

Toxicity probability results along with its discriminate score for 4th and 5th series of compounds.

Compound	Mutagenicity		Developmental Toxicity Potential		Skin Irritation	
	Prob.	Discriminant Score	Prob.	Discriminant score	Prob.	Discriminant score
4a	0.000	-13.045	0.000	-13.810	0.000	-19.910
4b	0.000	-12.133	0.000	-16.238	0.000	-22.273
4c	0.000	-12.671	0.000	-15.967	0.000	-23.703
5a	0.000	-30.690	0.006	-5.177	0.000	-8.151
5b	0.000	-28.824	0.002	-6.363	0.000	-11.563
5c	0.000	-29.362	0.002	-6.091	0.000	-12.888

Table 4(f)

Toxicity probability results along with its discriminate score for 4th and 5th series of compounds.

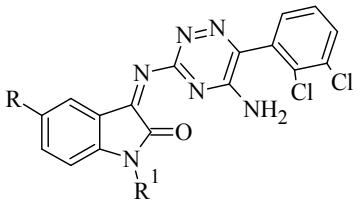
Compound	Occular Irritation		Aerobic Bio Degradability	
	Prob.	Discriminant Score	Prob.	Discriminant Score
4a	1.000	12.098	0.000	-20.354
4b	1.000	13.245	0.000	-21.070
4c	1.000	14.098	0.000	-21.192
5a	1.000	41.091	0.000	-13.659
5b	1.000	31.980	0.000	-41.597
5c	1.000	28.980	0.000	-8.848

Table 5

Carcinogenicity calls for 4th and 5th series compounds.

Series	4a		4b		4c	
Model	Prob.	Discriminant score	Prob.	Discriminant Score	Prob.	Discriminant Score
Male Rat	0.000	-22.992	0.000	-21.538	0.000	-25.403
Female Mouse*	0.783	1.285	0.722	0.955	0.771	1.216
Series	5a		5b		5c	
Model	Prob	Discriminant score	Prob	Discriminant Score	Prob	Discriminant score
Male Rat	0.000	-24.464	0.000	-24.060	0.000	-27.925
Female Mouse*	0.592	0.372	0.509	0.037	0.495	-0.021

** Carcinogenic Vs Non Carcinogenic model.*



RESULTS

i) *QSAR Analysis Results:* Following are the results of multiple regression analysis of CTC50 data from HEp-2 and DLA cell lines.

ii) *ADME Screening Studies*

Following results are obtained after subjecting compounds 4a, 4b, 4c, 5a, 5b and 5c to Accord for Excel.

Table 3(a)
Data table for 4th series compounds.

S. No.	Descriptors	4a	4b	4c
1.	ALOGP98	4.4607	5.1251	5.2091
2.	FPSA	109.6004	109.6004	109.6004
3.	AQ.SOL.LOG	-6.84566	-7.62721	-7.70052
4.	AQ.SOL.LOG.LEV	1	1	1
5.	BBB.LOG.LVL	4	4	4
6.	CYP2D6	0	0	0
7.	CYP2D6.PROB	0.356436	0.277228	0.227723
8.	HEPATOTOX	1	1	1
9.	HEPATOTOX.PROB	0.854305	0.960265	0.953642
10.	HIA.FABS.LEV	1	2	2
11.	HIA.FABS.T2	8.335923	10.86543	11.22068
12.	PROT.BIND.LEV	1	2	2
13.	PROT.BIND.LEV.LOG	1	2	2
14.	HBOND.ACCEPTOR	8	8	8
15.	ALERT	FALSE	FALSE	FALSE
16.	HBOND.DONOR	2	2	2
17.	ALERT	FALSE	FALSE	FALSE
18.	MLOGP.ALERT	FALSE	FALSE	FALSE
19.	WEIGHT.ALERT	FALSE	FALSE	TRUE
20.	RULE.OF.FIVE	0	0	1
21.	ALERT	FALSE	FALSE	FALSE

Table 3(b)

Data table for compounds of 5th series.

S. No.	Descriptors	5a	5b	5c
1.	ALOGP98	5.0948	5.7592	5.8432
2.	FPSA	95.6521	95.6521	95.6521
3.	AQ.SOL.LOG	-7.28539	-8.04877	-8.12209
4.	AQ.SOL.LOG.LEV	1	0	0
5.	BBB.LOG.LVL	4	4	4
6.	CYP2D6	0	0	0
7.	CYP2D6.PROB	0.29703	0.257426	0.19802
8.	HEPATOTOX	0	0	0
9.	HEPATOTOX.PROB	0.298013	0.284768	0.311258
10.	HIA.FABS.LEV	1	2	2
11.	HIA.FABS.T2	8.053926	10.85276	11.24205
12.	PROT.BIND.LEV	1	2	2
13.	PROT.BIND.LEV.LOG	1	2	2
14.	HBOND.ACCEPTOR	8	8	8
15.	ALERT	FALSE	FALSE	FALSE
16.	HBOND.DONOR	2	2	2
17.	ALERT	FALSE	FALSE	FALSE
18.	MLOGP.ALERT	TRUE	TRUE	TRUE
19.	WEIGHT.ALERT	FALSE	TRUE	TRUE
20.	RULE.OF.FIVE	1	2	2
21.	ALERT	FALSE	TRUE	TRUE

ALOGP98: Hydrophobhicity Parameter, FPSA: Fast Polar Surface Area, AQ.SOL.LOG: Log value of Aqueous solubility, AQ.SOL.LOG.LEV: Predicts Aqueous solubility level, BBB.LOG.LVL: Predicts blood-brain-barrier penetration level, CYP2D6: Predicts inhibition or non inhibition of CYP450 2D6 enzyme, CYP2D6.PROB: A scoring function that is a sum of predicted values and CYP2D6 model, HEPATOTOX: Predicts hepatotoxicity or non-hepatotoxicity, HEPATOTOX.PROB: A scoring function that is sum of predicted values of hepatotoxicity model, HIA.FABS.LEV: Predicts passive human intestinal absorption level,HIA.FABS.T2: The Mahalanobis distance for the compound in the FPSA, ALogP98 plane,PROT.BIND.LEV: Predicts Plasma protein binding levels, RULE.OF.FIVE: It's a Lipinski Rule (turns "True" for orally inactive molecules and "False" for orally active molecules in the software).

iii) QSTR studies

Following are the results obtained after subjecting the structures to TOPKAT software and using the Rat LD_{50}, LC_{50} and LOAEL values for building regression equation.

DISCUSSION

It was observed that there was good correlation between the biological activity and the respective physicochemical parameters. But it was observed that Van der Waal's energy, shape flexibility index and surface area of molecules which were recorded for SRB bioassay data gave comparatively better correlation with CTC50 value for HEp-2 cell line (table 2a-2d). Similarly for DLA cell lines the comparative account between MTT and SRB assay procedures showed that correlation of HOMO, LogP, and molecular refractivity of MTT bioassay gave better results with CTC50 (table 2a-2d). From the comparative account of ADME descriptors data for 4th and 5th series of compounds it

was observed that 5c compound had least chances for CYP450 2D6 inhibition, 5th series compounds did not show any dose dependent hepatotoxicity as probability value is near to zero. But 5b and 5c were predicted to have low intestinal absorption. Also, 4a and 5a compounds showed good intestinal absorptions (inside 99% confidence limit) as per HIA.FABS.T2 and Lipinski rule of five values as given table 3(a, b) and shown in figure 1. No compound was found to have BBB penetration as they were found outside the ellipse (figure 2). Protein binding data implies that all the compounds have highest protein binding except in case of 5a compounds (binding ≥ 90%) while all other compounds have protein binding of ≥95%.

Quantitative Structure Toxicity Relationship studies along with Rat LD_{50}, Rat inhalational LC_{50} and LOAEL data is given table 4(a-d). All equations developed gave good regression coefficient (r^2) values. The zero probability values with negative discriminant score shows level of compliance towards not having specific toxic effect. Hence it was observed that all the synthesized compounds showed good safety profiles as can be seen from table 4(a-f). But carcinogenicity studies showed that 5th series compounds are comparatively more safe than 4th series (table 5).

Table 2(a)
QSAR/Multiple Regression Equations

Eq. No.	Cell Lines	Equation
1.	HEp-2 Cell Lines (MTT Assay)	$B.A=3054.9778*X_1-566.17853*X_2-74.267235*X_3$ + 39260.133
2.	HEp-2 Cell lines (SRB Assay)	$B.A=668.92389*X_1+2296.0349*X_2-33.358601*X_3$ - 4153.6558
3.	DLA Cell lines (MTT Assay)	$B.A=8628.2295*X_1+4651.3799*X_2-435.70035*X_3$ + 113934.09
4.	DLA Cell lines (SRB Assay)	$B.A=1550.525*X_1+ 448.40305*X_2- 22.947651*X3$ - 2452.8684

Table 2(b)
Descriptors for Regression Equations

Eq. No.	X_1	X_2	X_3
1.	HOMO	Van der Waal's Energy	Molecular Refractivity
2.	Van der Waal's Energy	Shape flexibility index	molecular surface area
3.	HOMO	LogP	molecular refractivity
4.	Shape flexibility index	Van der Waal's Energy	Molecular surface area

Table 2(c)

Statistical Parameter for Regression Equations

Eq. No.	R	r^2	r^2(C.V)	t-probability*
1.	0.9252	0.855996	0.498544	< 0.05
2.	0.952376	0.90702	0.39776	< 0.05
3.	0.995345	0.990711	0.905728	< 0.05
4.	0.963431	0.9282	0.35094	< 0.05

Table 2(d)
Descriptors values for Regression Equation Calculated Using TSAR software

Eq. No.	Values				
1.	X_1	X_2	X_3	Activity Values	Predicted Values
	-9.3642	3.5253	110.78	479.54	429.542
	-9.4103	2.66121	115.58	346.95	421.265
	-9.4299	2.5374	118.4	249.61	222.044
	-9.0722	2.6695	131.08	249.35	298.497
	-9.0891	1.7967	135.88	432.78	384.175
	-9.1312	1.6872	138.7	105.7	108.407
2.	3.5253	6.0009	347.57	434.85	430.198
	2.66121	6.4477	364.04	218.02	286.835
	2.5374	6.5862	367.33	438.68	412.209
	2.6695	7.3076	417.52	438.2	482.507
	1.7967	7.7612	431.35	500	479.072
	1.6872	7.9049	450.15	127.56	108.489
3.	-9.3642	3.352	110.78	479.54	456.419
	-9.4103	3.87	115.58	373.18	383.947
	-9.4299	4.1438	118.4	250	467.614
	-9.0722	4.6689	131.08	249.35	493.581
	-9.0891	5.1869	135.88	431.87	488.313
	-9.1312	5.4607	138.7	130.43	230.528
4.	6.0009	3.5253	347.57	447.33	456.419
	6.4477	2.66121	364.04	427.87	383.947
	6.5862	2.5374	367.33	428.022	467.614
	7.3076	2.6695	417.52	500	493.581
	7.7612	1.7967	431.35	500	488.313
	7.9049	1.6872	450.15	217.18	230.528

B.A: Biological Activity, HEp-2: Caucasian male larynx epithelium carcinoma, DLA: Dalton's Lymphoma Ascites, MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenly-tetrazolium bromide, SRB: Sulphorhodamine B., r^2(CV): Cross Validation Coefficient.

CONCLUSION

Thus it was found that structural features resembling compound 5a proved to be comparatively better. Hence for further lead modification the features suggested are, there should be no halogen substitution on Isatin ring, small electron donating groups at R position and methylene pyprazinly moiety at R' . Improvement in activity can be done by using the developed QSAR model for derivatives. This study can prove to better for further molecular modelling, synthetic and *in vivo* pharmacological research.

REFERENCE

1) Perez MAC, Gracia AR, Teurel CF, Alvarez IG, Sanz MB, A topological-substructural molecular design (TOPS-MODE) approach to determining pharmacokinetics and pharmacol ogical properties of 6-fluoroquinolone derivatives. European Journal of Pharmaceutics and Biopharmaceutics, 56: 197-206, (2003).
2) F. Choplin. Comprehensive Medicinal Chemistry, Corwin Hansch, Vol 4, Elsevier Pergamon, Oxford: 33-57, (2005).
3) Valentina P, Lango K, Engels M, Rationalization of Physico Chemical Characters of 2-Phenyl-3-hydroxy-4(1H)-quinolinone-7-carboxylic acid Analogs as Topoisomerase Inhibitors: A QSAR Approach. Indian J. Pharm. Educ. Res, 43(3): 284-289, (2009).
4) Joshi KC, Joshi R, Isatin: A VersatileMolecule for the Synthesis of NovelSpiroheterocycles. Indian ChemicalSociety, 76: 643-649, (1999).
5) Bhattacharjee AK, Geyer JA, Woodward CL, Kathcart AK, Nichols DA, Prigge ST, Zhiyu Li, Mott BT, Waters NC, A Three-Dimensional in Silico Pharmacophore Model for Inhibition of Plasmodium falciparum Cylin-Dependent Kinases and Discovery of Different Classes of Novel Pfmrk Specific Inhibitors. Journal of Medicinal Chemistry, 47(22): 5418-26, (2004).
6) Tiwari RK, Rastogi N, Sethi R, Shukla S,Synthesis and Anti-leishmanial Activity of 1-amino methyl-5susbstituted-3-{4'-(3'-chlorobenzloxy)-benzoylhydrazone}-2-indolinones, J. Indian Chem. Soc., 85: 85-88,(2008).
7) Abadi AH, Abou-Seri SM, Abdel-Rahman DE, Klein C, Lozach O, Meijer L, Synthesis of 3-substituted-2-oxoindole Analogues and their Evaluation as Kinase Inhibitors, Anticancer and Antiangiogenic Agents. European Journal of Medicinal Chemistry, 41: 296-305, (2006).
8) Nagarajan G, Balasubramaniam V, Manjunath K.S, Rama Krishna E, Reddy KPK, ReddyEK, Girua K, Synthesis and Invitro Cytotoxicity Activity of Isatin Derivatives. Asian Journal of Chemistry, 20(3): 1665-1669, (2008).