International Journal of ChemTech Research
CODEN (USA): IJCRGG ISSN : 0974-4290
Vol.4, No.1, pp 92-98, Jan-Mar 2012

Classical QSAR Modelling of Indole-3-Acetic acids for Cancer Therapy

Love Kumar Soni*

**School of Pharmacy, Devi Ahilya University, Takshashila Campus
Ring Road, Indore - 452017, Madhya Pradesh, India**

Corres.author: lovesoni@hotmail.com

Abstract: A set of 26 compound of the series halogenated Indole-3-Acetic acids as oxidatively activate prodrugs with potential for targeted cancer therapy were subjected to quantitative structure activity relationship (QSAR) analysis using combination of various electronic, thermodynamic and spatial descriptors. Several statistical regression equations were obtained using multiple regression analysis. QSAR analysis showed that the cytotoxin produced from oxidation of IAA with horseradish peroxidase (HRP) have significant correlation with electronic (Energy of Highest molecular orbital), and thermodynamic (Non-1, 4-VDW energy, Henry law constant) properties of the molecule. QSAR analysis suggests that substitution with electron withdrawing and bulkier group is more favorable for inhibitory activity.
Key words: QSAR, halogenated Indole-3-Acetic acids, targeted cancer therapy.

INTRODUCTION

Horseradish peroxidase (HRP) is an important heme-containing enzyme obtained from plant source *Arabidopsis thaliana.* The 3D structure of the enzyme revealed its catalytic intermediates, mechanism of catalysis and the function of specific amino acids residues[1]. Site directed mutagenesis and directed evolution techniques are used to investigate the structure and function of HRP and offers opportunity to develop engineered enzymes for practical applications in natural products and fine chemical synthesis, medical diagnosis bioremediation. Physiological roles traditionally associated with the enzyme includes, IAA metabolism, cross-linking of biological polymers and lignifications are better studied at molecular levels[2]. Progress in this area should result from identification of entire peroxidase gene family of *Arabidopsis thaliana.* Attempts have been done to identify suitable prodrugs that could be used to test the hypothesis that peroxidase activity in cells, either endogenous or enhanced by immunology targeting, can activate prodrugs to cytotoxins. The prototype prodrugs based on derivative of Indole-3-acetic acid (IAA) a plant auxin[3], when activated by peroxidase enzyme (horseradish, HRP) produces peroxy radicals, with deleterious effects. The radical-cations formed on one-electron oxidation of indole-3 acetic acid (IAA) and its ring-substituted derivatives rapidly fragment, eliminating carbon centered free radical (3-indolylmethyl or skatolyl or analogues). The radical is reactive towards DNA and possibly other targets in anoxia, but in oxic and hypoxic cells rapidly adds oxygen to forms a peroxyl radical[4]. Subsequent products like 3-methyl-2-oxindole or analogues are reactive towards cellular nucleophiles such as thiols and DNA. The one electron oxidation of IAA is efficiently achieved by horseradish peroxidase (HRP),

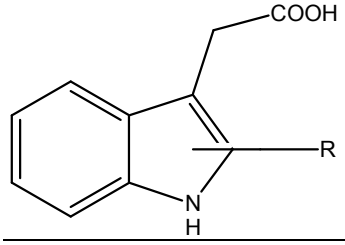
not requiring added hydrogen peroxide cofactor[5]. The cytotoxin so produced by IAA/HRP combination is of IAA are very cytotoxic with HRP even though they are difficult to oxidize. Humans in high doses tolerate IAA and HRP is a robust enzyme that meets the requirements for targeting to tumor by coupling to antibody or polymers, or by gene transfection[8,9]. This reveals that IAA merits further evolution as potential prodrugs for use in cancer therapy based on targeted delivery to tumors.

In the present paper, we describe the QSAR analysis to investigate the relationship between the various physico-chemical parameters and the biological activity data of substituted Indole-acetic acid (IAA)[6], as oxidatively activated prodrugs with potential for targeted cancer therapy that may be helpful in development of potent antibacterial agents.

MATERIAL AND METHODS

Derivatives of substituted Indole-acetic acid (Table 1), as oxidatively activated prodrugs with potential for targeted cancer therapy were taken from the reported work of Rossiter *et al*[6]. The biological activity data *i.e.* rate of oxidation of IAA with HRP (expressed in µm) were converted to negative logarithmic does (-log IC_{50}) in present QSAR analysis. cytotoxic towards mammalian cells, including human tumor cells[6,7]. Some halogen-substituted derivatives The series was subjected to molecular modeling studies using CS Chem-Office Software version 6.0 (Cambridge soft)[10]. The structure of the compounds were drawn in Chem Draw Ultra ver 6.0 and it was copied to Chem 3D Ultra to create the 3D model which was saved as the template model. For every compound, the template model was suitably modified considering its structural features so that every compound maintains same sequence of atoms. These structures were then subjected to energy minimization using molecular mechanics (MM2) until the root mean square (RMS) gradient value became smaller than 0.1kcal/molÅ. Minimized molecules were subjected to re-optimization via Austin model-1 (AM1) method using closed shell (restricted) wave function of MOPAC module until the RMS gradient attained a value smaller than 0.0001 kcal/mol Å. The geometry optimization of the lowest energy structure was carried out using eigenvector (EF) routine. The energy-minimized geometry was used for calculation of various steric, thermodynamic and electronic descriptor as given in table 2. The descriptors value were calculated using "compute properties" module of program.

Table 1: Structure and Activity of Halogenated Indole-3-Acetic acids



Compd. No.	R	IC_{50} (µm)	-log IC_{50}
1	5-CF_3O	0.3	0.523
2	5-F	0.4	0.398
3	1-Me, 5-F	0.4	0.398
4	4-Cl, 6-Cl	0.6	0.222
5	1-Me, 6-Cl	0.9	0.046
6	7-Cl	0.9	0.046
7	4-F	1	0.000
8	4-Cl	1.4	-0.146
9	5-Cl	1.4	-0.146
10	5-Br	1.7	-0.230
11	6-F	1.8	-0.255
12	1-Me	3.7	-0.568
13	H	3.8	-0.580
14	5-I	3.9	-0.591

15	5-Me, 7-F	4	-0.613
16	5-Cl, 7-F	5.3	-0.724
17	4-Me, 7-F	5.8	-0.763
18	6-Cl	7.5	-0.875
19	5-BnO	21	-1.322
20	5-MeO	22	-1.342
21	5-Me	41	-1.613
22	5-Ph	46	-1.663
23	2-Me, 5-F	88	-1.944
24	5-(4-ClC_6H_4)	103	-2.013
25	2-Me	435	-2.638
26	2-Me, 5-MeO	1710	-3.233

Table 2: Descriptors used for QSAR Analysis

S.	Descriptor	Parameter
1	Bend energy (E_b)	Thermodynamic
2	Boiling Point (BP)	Thermodynamic
3	Critical Pressure(P_c)	Thermodynamic
4	Critical Temperature (T_c)	Thermodynamic
5	Critical Volume(P_c)	Thermodynamic
6	Connolly accessible area (SAS)	Steric
7	Connolly molecular area (MS)	Steric
8	Connolly solvent excluded volume (SEV)	Steric
9	Dipole (DPL)	Electronic
10	Dipole-Dipole energy (E_d)	Thermodynamic
11	Electronic energy (ElcE)	Electronic
12	Exact mass (Mass)	Steric
13	Heat of formation(H_f)	Thermodynamic
14	Henry law constant (H)	Thermodynamic
15	HOMO energy (E_{HOMO})	Electronic
16	Ideal gas thermal capacity (C_p)	Thermodynamic
17	LUMO energy (E_{LUMO})	Electronic
18	Partition coefficient (Log P)	Thermodynamic
19	Melting point (MP)	Thermodynamic
20	Molar refractivity (MR)	Thermodynamic
21	Molecular weight (MW)	Steric
22	Non-1,4-vdv energy (E_v)	Thermodynamic
23	Ovality	Steric
24	Partition coefficient (PC)	Thermodynamic
25	Principal moment of inertia-x-axis (PMI-X)	Steric
26	Principal moment of inertia-y-axis (PMI-Y)	Steric
27	Principal moment of inertia-z-axis (PMI-Z)	Steric
28	Repulsion energy(NRE)	Thermodynamic
29	Standard gibb,s free energy (G)	Thermodynamic
30	Stretch energy (E_s)	Thermodynamic
31	Stretch-bend energy (E_{sb})	Thermodynamic
32	Torsion energy (E_t)	Thermodynamic
33	Total Energy (E)	Thermodynamic
34	VDV-1,4-energy(E-1,4)	Thermodynamic

Series was divided into a training set of 18 compounds and a test set of 8 compounds on the basis of structural diversity and cover the complete range of variation in inhibitory activity. Stepwise multiple linear regression analysis method was used to perform QSAR analysis employing in-house VALSTAT[11] program in order to establish a correlation between calculated descriptor as independent variables and biological activity as dependent variable. The best equation was selected on the basis of various statistical parameters such as correlation coefficient (r), standard error of estimation (SE) and sequential Fisher test (F). The equation was further validated on various statistical parameters like leave one out cross validated square correlation coefficient (q^2) using cross validation method[12], boot-strapping square correlation coefficient (r^2_{bs}), boot-strapping standard error (S_{bs}), randomize biological activity data test (Chance) and test for outlier's (Z-score value) which confirm the robustness and applicability of QSAR equation on the structure analogs[13].

Table 3: Observed, Calculated & Predicted Activities of Training Set using Equation 1 and 2.

Compd No.	Observed activity	Equation-1		Equation-2	
		Calculated activity	Predicted Activity	Calculated activity	Predicted Activity
1	0.523	0.708	0.792	0.779	0.903
2	0.398	-0.102	-0.158	-0.137	-0.195
3	0.398	0.506	0.565	0.646	0.767
5	0.046	0.011	-0.005	0.167	0.217
6	0.046	-0.736	-0.785	-0.398	-0.457
8	-0.146	-0.156	-0.161	-0.591	-0.655
9	-0.146	-0.170	-0.161	-0.114	-0.109
10	-0.230	0.095	-0.173	-0.012	0.027
11	-0.255	-0.249	0.154	-0.406	-0.418
12	-0.568	-0.416	-0.248	-0.476	-0.415
14	-0.591	-0.142	-0.314	-0.104	-0.029
15	-0.061	-0.819	-0.072	-0.764	-0.779
17	-0.763	-0.908	-0.837	-1.003	-1.069
19	-1.322	-1.717	-0957	-1.480	-1.509
21	-1.613	-1.307	-1.842	-1.437	-1.406
22	-1.663	-1.971	-1.244	-2.188	-2.519
24	-2.013	-1.578	-2.066	-1.792	-1.628
26	-3.233	-2.797	-1.432	-2.435	-1.871

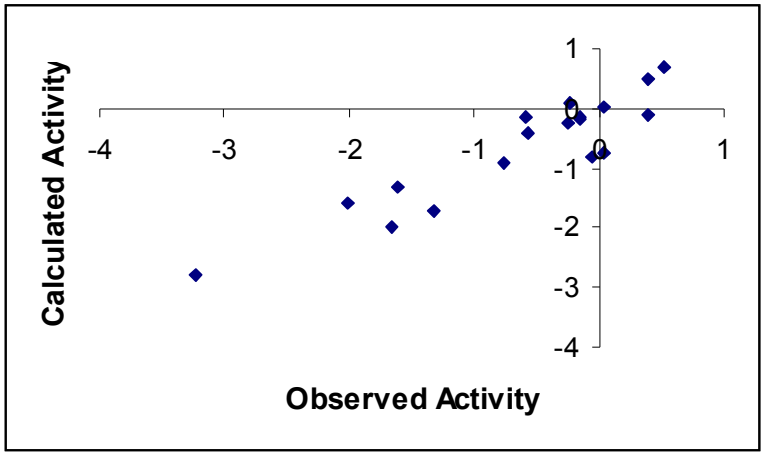


Fig. 1: Graphical representation of observed versus calculated activity using Eqn. 1.

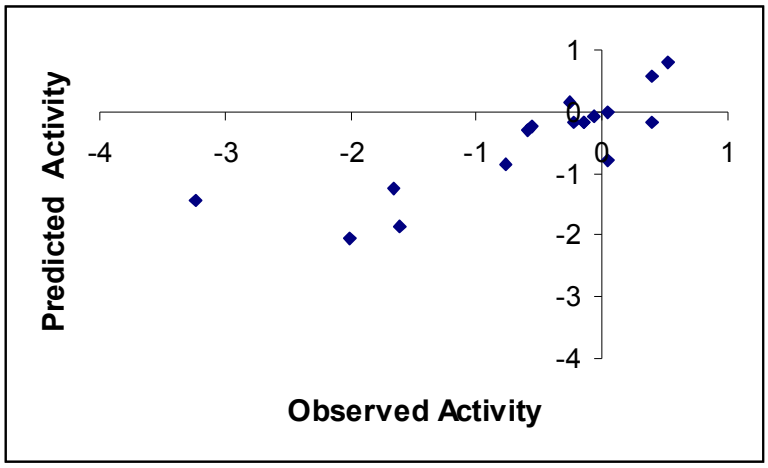


Fig. 2: Graphical representation of observed versus predicted activity using Eqn. 1.

Table 4: Observed & Predicted Activities of Test Set using Equation I and II.

Compd No.	Observed activity	Predicted Activity	
		Equation-1	Equation-2
4	0.222	0.511	-0.004
7	0	-0.422	-0.495
13	-0.579	-1.104	-1.134
16	-0.724	0.149	0.387
18	-0.875	-0.431	-0.413
20	-1.342	-1.913	-1.658
23	-1.944	-0.588	-0.558
25	-2.638	-1.811	-1.659

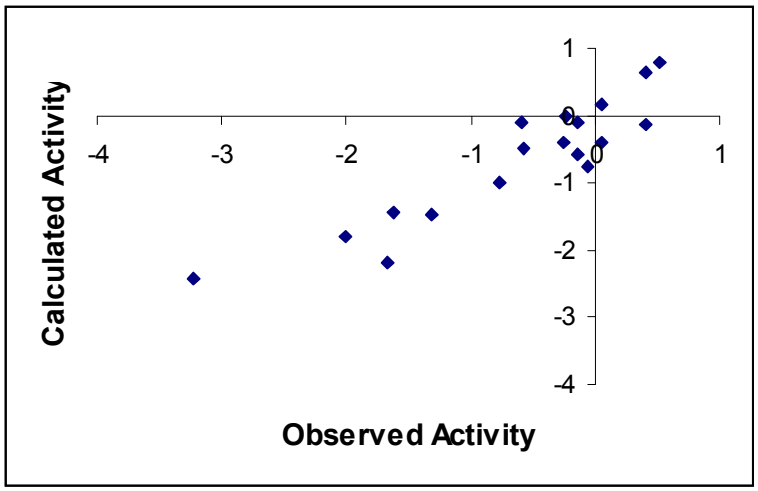


Fig. 3: Graphical representation of observed versus calculated activity using Eqn. 2.

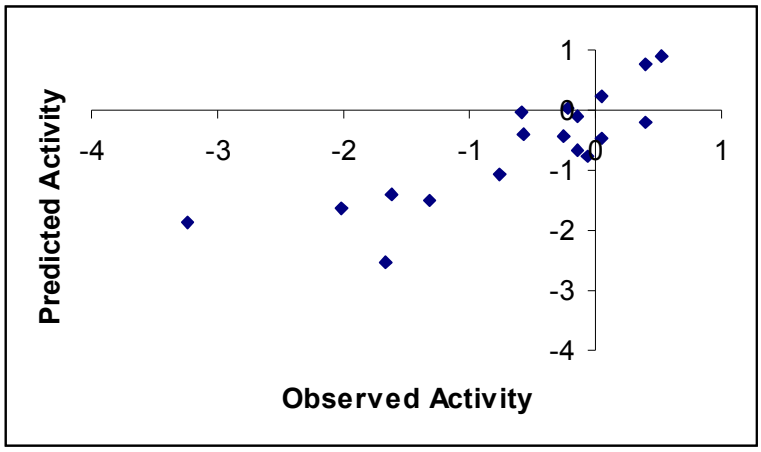


Fig. 4: Graphical representation of observed versus predicted activity using Eqn. 2.

RESULT AND DISCUSSION

When data set were subjected to stepwise multiple linear regression analysis, in order to develop QSAR between as biological activity dependent variables and calculated descriptor as independent variables, several equations (Eqn. 1 and 2) with correlation coefficient (r>0.92) were obtained.

The statistically significant equation 1 with significant coefficient of correlation (r = 0.936) is able to explain wide structural diversity in the substitutions at R as well as to cover the significant explanation for complete range of biological activity data.

- log IC_{50}= - 44.25(± 14.79) - 0.173 (± 0.064) H - 5.39 (± 1.72)] E_{HOMO} + 0.48 (± 0.26) E_v, N= 18, r= 0.876, SE= 0.380, F= 32.9704, r^2_{bs}= 0.832, chance <0.01, q^2= 0.81, S_{press} = 0.471, S_{DEP} = 0.415, r^2_{pred}= 0.327(Eqn. 1)

Equation 1 accounts for 87.6% variance in the inhibitory activity. The value of sequential Fisher test suggested more than 99.9% internal statistical significance as it exceeded the tabulated value $F_{(3.14\ \alpha=0.001)}$= 8.82 (Table 3, fig. 1 and 2). The inter correlation among the parameters (ICAP) was significantly low (ICAP<0.122) suggested that all the three parameters contributed individually and independently to the equation. The selected equation was further statistically evaluated to confirm its robustness. The value of cross validated squared correlation coefficient (q^2 = 0.81) using leave one out method suggested good internal productivity of equation. The value of bootstrapping squared correlation coefficient (r^2_{bs}=0.832), which is at par to the conventional correlation coefficient (r^2 =0.936) suggested that no single compound contributed up to too low or too high extent means equation can be used for wide range of structural analogs. The value of chance statistics suggested less than 1% chance of fortuitous correlation.The observed & predicted activities of test set using equation I has been given in table 4.

Equation 1 shows the negative contribution of HOMO energy (E_{HOMO}) and Henry law constant (H) where as Non-1,4-VDW energy (E_v) contributed positively to the inhibitory activity. HOMO is an electronic descriptor, which indicates the highest energy lavel of the molecule that contains electrons, it governs the molecular properties and reactivities, thus measures the nucleophilicity of the molecule. Negative contribution of HOMO energy in the modal suggested that the molecule may interact on electron rich areas of the receptor and the substitution of electron withdrawing groups in the molecule will impart the positive influence to the inhibitory activity. Non-1,4-vanderwaals energy is thermodynamic descriptor, which represents the sum of pairwise vanderwaals interaction energy terms for atoms, separated by more than three chemical bonds. Its positive coefficient in the equation reflects the necessity of conformation flexibility in the molecule for beneficial activity. Negative contribution of Henry law constant implies that it significantly affects inhibitory activity.

Equation 2 depicts 86.3% variance in cytotoxic inhibitory activity of IAA/HRP combination with standard error of estimation (±0.399).

- log IC_{50}= -37.21(±16.76) - 0.164(±0.068) H - 471(±1.89) E_{HOMO} + 0.183(±0.111) E_t, N=18, r=0.929, r^2 = 0.863, SE = 0.399, F= 29.4184, r^2bs = 0.884, chance< 0.01, q^2= 0.722, Spress= 0.569, S_{DEP}= 0.502, r^2pred=0.22 (Eqn. 2)

Equation 2 depicts 86.3% variance in cytotoxic inhibitory activity of IAA/HRP combination with standard error of estimation (±0.399) (Table 3,

fig. 3 and 4). Sturdiness of the modal is further supported by leave one out cross validation method (q^2=0.722), bootstrapping data (r^2_{bs} =0.884) and scrambled biological activity test (chance <0.01). The observed & predicted activities of test set using equation II has been given in table 4. Positive and negative contribution of torsion energy (E_t) and HOMO energy (E_{HOMO}) and Henry law constant (H) respectively in equation has significant effect. Energy of highest occupied molecular orbital (HOMO), electronic descriptor contributed negatively to the equation, connote the substitution of electron withdrawing groups for maximum activity. Thermodynamic descriptor torsion energy contributed positively to the equation. It indicates the specificity of receptor to blind to the specific conformation of the molecule or it discerns about the better fit of the drug molecule to the receptor site.

In conclusion, the present study provides important structural insights in designing potent halogenated Indole-3-Acetic acids analogues for targeted cancer therapy. The QSAR analysis revealed that electronic parameter (HOMO energy) and thermodynamic parameters (Henry law constant, Non-1,4-vanderwaals energy and torsion energy) play a significant role to explain the variance in biological activity. The QSAR analysis suggested that substitutions of electron withdrawing groups with proper confirmation would enhance the biological activity of the halogenated Indole-3-Acetic acids.

Acknowledgements

Authors are thankful to the Head, School of Pharmacy, Devi Ahilya University, Indore for providing facilities to complete this work.

REFERENCES

1. Veitch NC, HRP: A modern view of classic Enzyme. Phytochemistry 2004;65(3) : 249-259.
2. Wardman P. Indole-3-Acetic Acids and HRP : A New Prodrug / Enzyme Combination for Targeted Cancer Therapy. Curr Pharm Des 2002;8(15): 1363-1374.
3. Folkes LK, Wardman P. Oxidative activation of IAA to cytotoxic Species Potential New Role for Plant Auxins in cancer Therapy. Biochem Pharmacol 2001; 61(2) : 129-136.
4. Greco O, Folkes LK, Wardman P, Tozer GM, Dachs GU, Rossiter S, Kanthou C.: Horseradish Peroxidase-Mediated Gene Therapy: Choice of Prodrug Oxic & Anoxic Tumor conditions. Mol Cancer Ther 2001;1(2):151-160.
5. Folkes L.K., Candeias L.P., Wardman P. Towards Targeted "Oxidation Therapy" of Cancer: Peroxidase- Catalyzed Cytotoxicity of Indole-3-Acetic Acids. Int J Rad Onco Biol Phys 1998; 42(4): 917-920.
6. Rossiter S., Folkes L. K., Wardman, P. Halogenated Indole-3-Acetic Acid As Oxidatively Activated Prodrugs with Potential for Targetted Cancer Therapy. Bioorg Med Chem Lett. 2002; 12: 2523-2526.
7. Greco O, Folkes L K, Torzer G M, Dachs G U. Development of novel enzyme/prodrug combination for gene therapy of cancer; HRP/IAA. Cancer Gene Ther. 2000; 7: 1414-1420.
8. Greco O, Folkes L.K., Wardman P, Stratford M R, Dachs G U. 5-fluroidindole-3-acetic acid : A Prodrug Activated by Peroxidase with Potential for use in Targeted Cancer Therapy. Biochem Pharmcol 2002; 63(2): 265-272.
9. Wardman, P. Free Radicals and Prodrugs for Cancer Therapy. Scientific Year Book, 2001-2002.
10. CS Chem office, Version 8.0, Cambridge soft corporation, Software publishers Association, 1730 M Street, suit 700, Washington D.C. 20036(202) 452-1600, U.S.A.
11. Gupta A K, Arockia B M, Kaskhedikar S G. Indian J Pharm Sci 2004; 66: 396-402.
12. Kubinyi H: Hansch analysis and related approaches. Weinheim: New York; 1993.
13. Hansch C, Leo A. Substituent constants for correlation analysis in chemistry and biology. John Wiley & Sons: New York; 1979.

*****