IOSR Journal of Applied Chemistry (IOSR-JAC)
e-ISSN: 2278-5736. Volume 18, Issue 4 Ser. 1 (April 2025), PP 12-15
www.iosrjournals.org

Modelling Of IC_{50} Of Phenolic Derivatives By Computational Methods

Sameer Dixit And Arun K Sikarwar
Sr. Lecturer, Associate Professor & Principal
Department Of Chemistry, M. J. P. Govt. Polytechnic College Khandwa, Madhya Paradesh (INDIA)
Department Of Chemistry, Govt, P G. College Harda, Madhya Paradesh (INDIA)

Abstract
Quantitative structure-activity relationship (QSAR) is a statistical modelling approach mostly used in molecular modelling, property and activity prediction of new molecule, drug discovery and prediction of environment. QSAR studies have been performed on 17 molecules of phenol and its derivatives. These quantitative structure-activity relationship models are based on few calculated and experimentally obtained Biological Activities Half Maximal Inhibitory concentration (IC_{50}). The paper deals with structure-activity relationships of phenols and its derivatives for the development of predictive models from several descriptors. To developing the model for IC_{50} of phenol derivatives we used descriptors like Mor04m, Mor23m, FDI, RDF045m, MATS5p, R3e, eHOMO, eLUMO and the best model proposed for Half Maximal Inhibitory concentration (IC_{50}) of Phenol's & its Derivatives. for this we used several statistical parameters like R, PRESS, R_2cv, SSY, SPRESS, PSE, LSE, PE etc. to validate the model.
Keywords: *IC_{50}, QSAR, Molecular descriptors, 3D MoRSE descriptors, FDI descriptors, RDF descriptors, Moreau autocorrelation descriptors, correlation coefficient, PRESS, SPRESS, PSE, LSE, PE*

Date of Submission: 03-04-2025 Date of Acceptance: 13-04-2025

I. Introduction

Half-maximal inhibitory concentration (IC_{50}) is the most widely used and informative measure of a drug's efficacy. It indicates how much drug is needed to inhibit a biological process by half, thus providing a measure of potency of an antagonist drug in pharmacological research. The ability of a compound to penetrate various biological Membranes, tissues and barriers is a primary factor in controlling the interaction of these compounds with biological systems. IC_{50} values were used to measure biological activity, which is defined as median inhibition concentration (concentration that reduces the effect by 50%)[1].

It is a measure of the effectiveness of a compound in inhibiting biological or biochemical function. Often, the compound in question is a drug candidate. This quantitative measure indicates how much of a particular drug or other substance (inhibitor) is needed to inhibit a given biological process (or component of a process, i.e. and enzyme, cell, cell receptor or microorganism) by half. In other words, it is the half maximal (50%) inhibitory concentration (IC) of a substance (50% IC, or IC_{50}). It is commonly used as a measure of antagonist drug potency in pharmacological research. Sometimes, it is also converted to the pIC_{50} scale (-log IC_{50}), in which higher values indicate exponentially greater potency. According to the FDA, IC_{50} represents the concentration of a drug that is required for 50% inhibition in vitro. It is comparable to an EC_{50} for agonist drugs. EC_{50} also represents the plasma concentration required for obtaining 50% of a maximum effect in vivo.

QSAR i.e. Quantitative Structure Activity Relationship provides a way to correlate the effect of structure over activity in terms of mathematical descriptors viz. Topological Indices. Quantitative structure-activity relationships (QSAR) represent an attempt to correlate structural or property descriptors of compounds with activities. These physicochemical descriptors, which include parameters to account for hydrophobicity, topology, electronic properties, and steric effects, are determined empirically or, more recently, by computational methods. Activities used in QSAR include chemical measurements and biological assays. QSAR currently are being applied in many disciplines, with many pertaining to drug design and environmental risk assessment. Generally any attempt to model the relationship between an experimental activity and the chemical structure requires some type of regression analysis. This analysis requires a set of compounds with known activities to make the initial model. This is called the 'training set'. The regression extracts trends between the structures of these training compounds and the activity. The aim is to train an accurate enough model that permits the calculation of the activity of molecules with yet unseen. Making the model requires each chemical structure to be represented in a form that can be easily analyzed. This project will therefore be concerned with molecular descriptor based QSAR, where each molecule is represented by a set of chemical descriptors[2].

II. Material And Method

In Quantitative Structure Activity Relationship **(QSAR)** models in which physicochemical parameters of drugs and the other compounds are correlated with biological activities, lipophilicity (partition coefficient, chromatographic parameters) has a major role. Other important parameters are polarizability, electronic and steric parameters, molecular weight, geometry, conformations entropies etc. Recently many molecular modelling methods based on widely spread Quantitative Structure Property/Activity Relationship (QSPR/QSAR) techniques found their place as an important role for the chemical engineers, chemists and especially for different aims.

In case of modelling Log IC_{50} we found single descriptor is not sufficient to express completely of this activity of given set of compounds. So we use more than one descriptor to achieved goal And this type of analysis known as multiple linear regression analysis 'MLR'. In order to build linear relationship and test model, the 17 compound data sets was used as training to build model. Finally with the selected eight different descriptors, we will build linear models using the training data sets and following equations were obtained.

The ability of a compound to penetrate various biological membranes, tissues and barriers is a primary factor in controlling the interaction of these compounds with biological systems. IC_{50} values were used to measure biological activity, which is defined as median inhibition concentration (concentration that reduces the effect by 50%) (Banarjee et al., 1980)

Modeling of 'IC$_{50}$' of Phenol derivatives

IC_{50} values[3] of phenol derivatives which used in the study are given in **Table-(1.1).** These compounds are interested with respect to the noncreative toxic effects on the microorganism Pseudomonas putida.[4,5]

For developing the model for IC_{50} of phenol derivatives in we used eight descriptors Mor29p, Mor20e, Mor04m, Mor23m, FDI, RDF045m, MATS5p, and R3e. There are 17 observations (molecules) are used to built first model for IC_{50}**.** By regression Statistics we get correlation coefficient is 0.9972, r^2 is 0.9944, Adjusted R Square *is* 0.8789, and Standard Error is 0.3386 for model-I which described by equation 1.

Predicted log IC50 = (39.94343*Mor29p) + (1.256769*Mor20) + (0.495208*Mor04m) + (4.294057*Mor23m) + (-0.08078*FDI) + (0.320863*RDF045m) + (-0.1218*MATS5p) + (0.797296*R3e)…….(1)

Analysis of variance of Model -I for IC_{50}

ANOVA					
	df	***SS***	***MS***	***F***	***Significance F***
Regression	8	183.01	22.876	199.58	2.14E-08
Residual	9	1.0316	0.1146		
Total	17	184.04			

Table (1.1) Observed and Predicted value of Log IC50 using Eq. (1)

S. No	Abbreviations	*log IC50 (mg/L)*	*Predicted log IC50 (mg/L)*	*Residuals*	*Standard Residuals*
1	PH	3.8	3.9699	-0.17	-0.69
2	2MPH	3.39	2.831	0.559	2.2691
3	3MPH	2.36	2.8207	-0.461	-1.87
4	4MPH	3.21	3.2867	-0.077	-0.311
5	2APH	3.44	3.1927	0.2473	1.0038
6	4APH	2.6	2.7467	-0.147	-0.595
7	3NPH	3.05	2.8639	0.1861	0.7553
8	4NPH	3.08	2.9833	0.0967	0.3927
9	24DNPH	2.01	2.1577	-0.148	-0.6
10	2HOPH	3.79	4.2355	-0.445	-1.808
11	3HOPH	4.24	3.9881	0.2519	1.0228
12	4HOPH	3.78	3.7139	0.0661	0.2682
13	123THOB	4.59	4.5109	0.0791	0.3213
14	2CLPH	3.38	3.2183	0.1617	0.6564
15	3CLPH	2.18	2.3345	-0.155	-0.627
16	4CLPH	2.9	2.911	-0.011	-0.045
17	4EtPH	2.92	2.9518	-0.032	-0.129

III. Results And Discussion

Observed value of Log IC_{50} was plotted against and Predicted values Using Eq. (1) shown in Figure below. The figure clearly indicates there is a significant co-relation between Observed and Predicted values of Log IC_{50}. Only 2MPH (2-Methylphenol) shows deviation. Other molecule shows excellent co-relation for Log IC_{50}. (Correlation coefficient is 0.9972, r^2 is 0.9944).

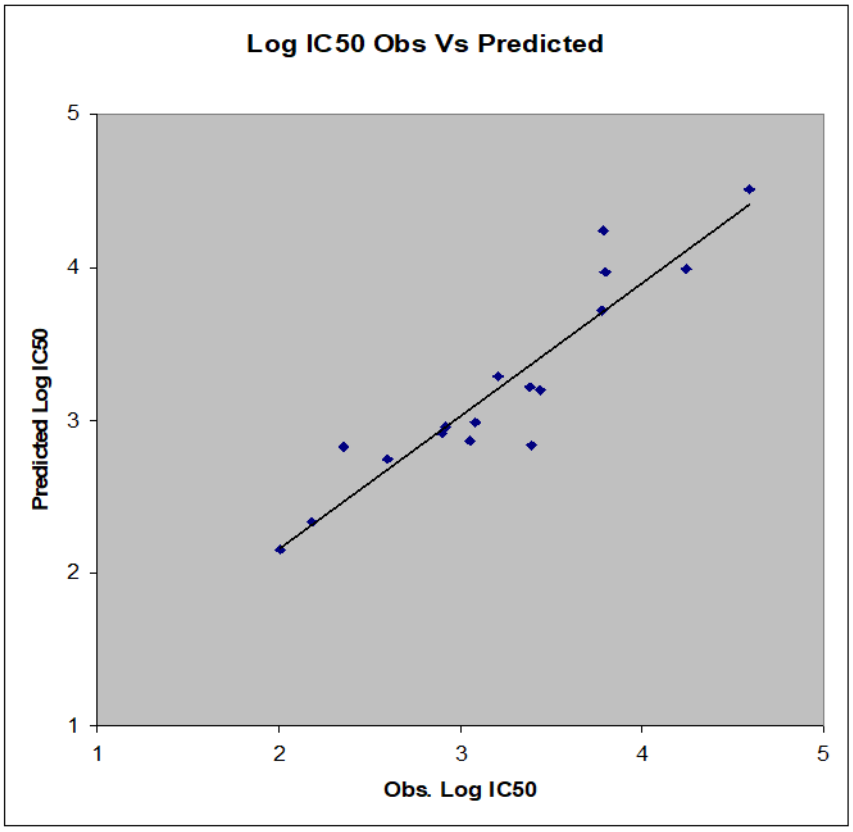

Figure 1.1 Correlation of Observed and Predicted value of Log IC_{50} Using Eq. (1)

Table (1.2) Cross Validation Parameters for modelling of Activity

R	0.997
R^2	0.994
SE or Sd	0.339
N	17
No. of Descriptors	8
PRESS	1.032
SSY	6.896
R^2cv	5.685
SPRESS	0.359
PSE	0.246
R^2A	0.879
LSE	1.032
PE	0.506
Q=r/sd	2.945
PRESS/SSY	0.150

Statistical analysis

In order to validate above model, we used statistical approach. These statistical parameters are support Model-I for LogIC_{50}. We observed that for the models discussed above r value is of order of 0.98 to 0.99. This is better for fitting of values. It is worthy to mention that a model (regression equation) with excellent statistics may not necessary have excellent predictive power. In order to confirm most powerful predictable Model for surface tension we have apply some statistical parameter[6]. Thus, the next step of regression analysis is to examine predictive power of the proposed model this can be easily done by calculating Poglianis quality factor[7] Q. This quality factor is defined as the ratio of correlation coefficient (r) to the standard error 'SE' (standard deviation 'sd'). The Q values are high for Polarizability model-1 (Eq.1) has best predictive powers.

The cross-validated **PRESS and SSY** as recorded in **Table (1.2)** indicates model-I (Eq.1) for Log IC_{50} is a better model and will give excellent result. And according to **SPRESS** and **PSE** values model-I (Eq.1) for Log IC_{50} is a better model and will also give excellent result.

The **PE** values are much greater than correlation coefficients R for Log IC_{50} model-I (Eq.1). So model has best predictive powers. The **LSE** values are low for Log IC_{50} model-I (Eq.1) have best predictive powers.

The PE values are much greater than correlation coefficients R for Biological **Activity LogIC_{50}** model-1 (Eq.1) have better predictive powers.

IV. Conclusion

For QSAR analysis regression was performed using IC_{50} values as dependent variables and calculated parameters as independent variables. In any thorough investigation of the effects of molecular properties, it is essential to prove that the results are both statistically valid and make chemical sense. It would be appropriate to obtain insight into the physical meaning of the correlation obtained as an output of the regression analysis. The magnitude of a descriptor could be used as a guideline to improve the anticancer activity of molecules.

Acknowledgment

Authors are very thankful to Mr. Mukesh Waskel, Principal M. J. P. Govt. Polytechnic College, Khandwa for providing facilities and motivation in the work.

References

[1] Banarjee S., Yalkowsky S. H., Valvani S. C., Water Solubility And Octanol/Water Partition Coefficients Of Organics. Limitations Of The Solubility-Partition Coefficient Correlation, Environ. Sci. Technol, 14, 1227-1229, (1980)
[2] Todeschini R., And Consonni V., Handbook Of Molecular Descriptors. P 688, Weinheim: Wiley – VCH. (2002).
[3] Gul S, Ozturk D, Determination Of Structure-Toxicity Relationship Of Amphiprotic Compounds By Means Of The Inhibition Of The Dehydrogenase Activity Of Pseudomonas Putida, Turk. J. Chem., 22, 341-349, (1998)
[4] Miller M., Wasik S. P., Huang G. L., Shiu W. Y., Mackay D., Relationship Between Water Octanol-Water Coefficient And Aqueous Solubility, Environ. Sci. Technol., 19, 522-529, (1985)
[5] Kemoto I., Motoba K., Suzuki T., Uchida M., Quantitative Structure Activity Relationships Of Nonspecific And Specific Toxicants In Several Organism Species, Environ. Toxicol. Chem., 11, 931-939, (1992)
[6] Dixit S., Sikarwar A. K., 2017, Studies Of Polarizability Of Phenol And Its Derivative Using Computational Methods, Asian Journal Of Orangic & Medicinal Chemistry, 2(4), 134-137
[7] POGLIANI L., 1994, Structure Property Relationships Of Amino Acids And Some Dipeptides, Amino Acids, 6,141-153