

Chemo Informatics QSAR Analysis of Nitroaromatic Compounds Toxicity

G.Ayyapan

Associate Professor, Dept. of IT, Bharath University, Chennai -600073, India

ABSTRACT: The use of information technology and management has become a critical part of the drug discovery process. The rational design of new drug molecules involves input from various branches of science. In this context the information and management of bio and chemical information have become the integral part. In addition, it is of utmost importance to enrich potential libraries with those molecules which could be converted to suitable drug candidates or omitted as toxins. During the practice of chemoinformatics, it has been realized that molecular diversity is an essential feature to characterize the reactivity of the molecules. In addition, a paradigm shift in structure- activity relationship has resulted in the integration of various descriptors and quantum chemical descriptors based drug development activities into early stages of lead discovery. In particular, various descriptors are being developed and used to help identify and screen out compounds that are unlikely to become drugs/toxins. This paper highlights the development of recent DFT based chemical reactivity descriptors and the application of these descriptors towards the prediction of chemical reactivity, especially in the prediction of toxicity and biological activities of nitroaromatic compounds.

I.INTRODUCTION

The data base development, management and analysis of biological information are defined as bioinformatics which includes various database management tools, analysis tools and molecular modeling. The term "chemoinformatics" has been introduced in the Annual Reports of Medicinal Chemistry in 1998 by Brown. Chemoinformatics¹⁻⁵ is the amalgamation of those chemical information resources to transform data into vital information and chemical information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and organization. In fact, both bioinformatics and chemoinformatics are generic terms that encompass the design, creation, organization, management, retrieval, analysis, dissemination, visualization and use of chemical and biological information. Chemi-informatics, chemometrics, computational chemistry, chemical informatics, and chemical information management/science are some of the related areas of chemoinformatics. The development of future chemical informatics systems will require a workforce with a solid grounding in chemistry and an expert understanding of the available computer technology. Chemical, agrochemical, pharmaceutical, and biotechnology branches of science require extensive input from both bioinformatics and chemoinformatics.

The quantitative structure- activity relationship (QSAR) and the quantitative structure- property relationship (QSPR) are the important tools of the bio-chemo-informatics which can be built essentially based on the data generated from the molecular modeling and computational chemistry. The QSAR and QSPR attempt to find a mathematical relationship between chemical structure and biological activity or chemical property for a series of homologous compounds. These series of homologous compounds are called the training set. The generated mathematical equation can be used to predict the activity or property of any new compound, which has been built from the chosen training set. Numerous descriptors have been used to develop QSAR and QSPR for different applications. In this regard it is necessary to mention the noteworthy contribution made by Hansch and coworkers⁵ to the development and growth of this area of activity. Nitroaromatic compounds are important materials or intermediates of explosives, pesticides, organic synthesis, and dyestuffs etc. With the development of industry, thousands of these compounds have being introduced into the environment every year and QSAR analysis on the toxicity of these compounds can provide us with valuable information. Hence, in the present investigation experimental toxicity values of 18 nitroaromatics to the algae (*Scenedesmus obliquus*) have been probed with DFT based descriptors.

The DFT offers a strong foundation for various qualitative concepts in the chemical reactivity.^{6, 7} Popular qualitative chemical concepts such as electronegativity and hardness have been widely used in understanding various aspects of chemical reactivity. Recently, Geerlings⁷ and co workers have reviewed the tremendous development in the application of conceptual density functional theory to variety of chemical and biological problems. Although conceptual density functional theory has been used in numerous investigations to probe the chemical reactivity and site selectivity, their applications in the area of structure- activity relationship aspects of biochemoinformatics are limited. Some of the important contributions in this area are highlighted in the following section. Based on the success of these DFT descriptors as revealed in the previous studies^{8,9} and also due to their simple calculation procedure, the usefulness of the DFT descriptors in the QSAR parlance has been probed in detail.

II. THEORETICAL BACKGROUND

According to the density functional theory (DFT),^{10, 11} the chemical potential (μ), and chemical hardness (η) are defined as,

$$\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_{v(\vec{r})} \quad (1)$$

and

$$\eta = \frac{1}{2} \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} = \frac{1}{2} \left(\frac{\partial \mu}{\partial N}\right)_{v(\vec{r})} \quad (2)$$

where E is the total energy of the system, N is the number of electrons in the system and $v(\vec{r})$ is the external potential. μ is identified as the negative of the electronegativity (χ) as defined by Iczkowski and Margrave.¹²

The vertical ionization potential (IP) of a system is the change of energy when an electron is removed from the system and variation of the energy when an electron is added to the system is known as electron affinity (EA). In both the cases of electron withdrawing and electron addition, the external potential has to be kept fixed. Hence, using finite difference approximation, Eqs. (1) and (2) become,

$$\mu = -\frac{(IP+EA)}{2} \quad (3)$$

$$\eta = \frac{IP-EA}{2} \quad (4)$$

However, to obtain the IP and EA and hence μ and η , one needs three energy values (E_N , E_{N+1} and E_{N-1}). To save the computational time, we have calculated chemical potential and chemical hardness by using Koopman's theorem¹¹ as,

and

$$\eta = \frac{E_{LUMO} - E_{HOMO}}{2} \quad (6)$$

where E_{LUMO} is the lowest unoccupied molecular orbital's energy and E_{HOMO} is the highest occupied molecular orbital's

$$\mu = \frac{E_{LUMO} + E_{HOMO}}{2} \quad (5)$$

energy.

Parr et al¹³ introduced the global electrophilicity index (ω) in terms of chemical potential and hardness as

$$\omega = \frac{\mu^2}{2\eta} \quad (7)$$

III. RESULTS AND DISCUSSION

The acute toxicities above the baseline toxicity are elicited by the reaction of the electrophilic nitroaromatic derivative with a nucleophile in cellular nucleic acids of proteins. The electronic parameters including half-wave reduction potential Hammett-Taft constant the energy of the lowest unoccupied molecular orbital, E_{LUMO} and the acceptor delocalizability are used to describe the excess toxicity above the baseline. With the development of the computer technology and quantum chemistry, the quantum chemical treatment of electronic effects is potentially more powerful than the other approaches because it can get the parameters accurately and easily and allow greater flexibility in the construction of the data set.

The experimental toxicity ($-\lg EC_{50}$)¹⁴ values of 18 nitroaromatics to the algae (*Scenedesmus obliquus*) along with that calculated $-\lg EC_{50}$ values are listed in Table 1. The calculated values of electrophilicity for the selected molecules are also presented in Table 1.

Linear regression analysis has been carried out using the experimental toxicity ($-\lg EC_{50}$) values as dependent variable and the calculated electrophilicity values as independent variables. The regression equation is given as,
 $-\lg EC_{50} = 0.983 + 0.791 * \omega$ (8)

$N=18$, $r^2=0.813$, $SD=0.268$

S.NO.	Compounds	Electrophilicity index (eV)	-lg EC ₅₀	
			Experimental	Calculated
1	Nitrobenzene	2.88	3.26	3.26
2	4-Nitrotoluene	3.48	3.74	3.73
3	2-Nitroaniline	2.95	3.33	3.32
4	3-Nitroaniline	3.14	3.48	3.46
5	4-Nitroaniline	2.88	3.4	3.26
6	2-Nitroanisole	3.48	3.44	3.74
7	3-Nitroanisole	3.38	3.71	3.66
8	4-Nitroanisole	3.32	3.65	3.61
9	2-Nitrophenol	3.53	3.51	3.77
10	3-Nitrophenol	3.51	3.75	3.76
11	4-Nitrophenol	3.44	3.57	3.71
12	2,4-Dinitrotoluene	4.51	4.52	4.55
13	2,6-Dinitrotoluene	4.38	4.06	4.44
14	1,2-Dinitrobenzene	4.59	5.04	4.62
15	1,3-Dinitrobenzene	4.68	4.85	4.68
26	1,4-Dinitrobenzene	5.04	4.96	4.97
27	2,4-Dinitroaniline	3.84	4.68	4.02
18	2,4-Dinitrophenol	4.51	4.16	4.55

It can be noted from equation 8 that a high coefficient of determination (r^2) of 0.813 along with a low value of root mean square deviation (SD) of 0.268 has been obtained. Statistical data obtained above clearly shows the importance of DFT based descriptor, electrophilicity in the prediction of toxicity of selected molecules.

A plot between experimental toxicity (-lg EC₅₀) and the calculated toxicity (-lg EC₅₀) values of 18 nitroaromatics compounds are presented in Figure 1.

Table 1. Experimental toxicity values (-lg EC₅₀ (mol l⁻¹)) along with calculated descriptor values for a set of nitroaromatic compounds

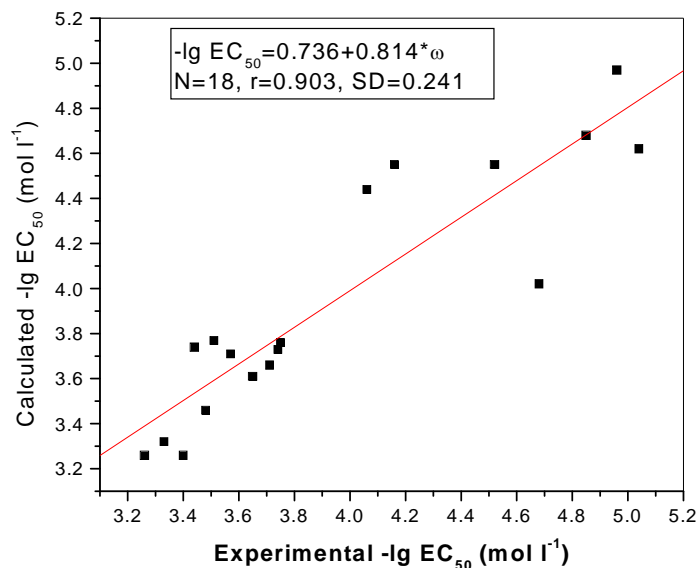


Figure 1. A plot between experimental and calculated toxicity values (-lg EC₅₀ (mol l⁻¹)).

There exists a good correlation coefficient (r) of 0.903 indicating the significance of the developed model in toxicity prediction.

IV. CONCLUSION

The experimental toxicity ($-\lg EC_{50}$) values of 18 nitroaromatics to the algae (*Scenedesmus obliquus*) has been analyzed using QSAR approach. In this study, experimental toxicity values are taken as dependent variables and electrophilicity values is taken as independent variable. A reasonably good coefficient of determination (r^2) of 0.813 is obtained showing the significance of the developed model in toxicity analysis on nitroaromatic compounds.

ACKNOWLEDGEMENTS

The authors thank Dr. V. Subramanian, CLRI, Dr. K. Chitra, QMC, and Dr. V. Kannan, BIHER, for their encouragement.

REFERENCES

1. Hansch, C.; Li, R. I.; Blaney, J. M.; Langridge, R.; *J. Med. Chem.* **1982**, 25, 777.
2. Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. *Chem. Rev.* **2002**, 102, 783.
3. Hansch, C.; Kurup, A.; Garg, R.; Hua Gao. *Chem. Rev.* **2001**, 101, 619.
4. Leach, A.R.; Gillet, V. J. *An Introduction to Chemoinformatics*. Kluwer, Dordrecht, 2003.
5. Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M.; *Nature* (London), **1962**, 194, 178.
6. Chermette, H. *Jour. Comput. Chem.* **1999**, 20, 129.
7. Geerlings, P.; De Profit, F.; Langenaeker, W. *Chem. Rev.* **2003**, 103, 1793.
8. Parthasarathi, R.; Padmanabhan, J.; Subramanian, V.; Maiti, B.; Chattaraj, P. K. *J. Phys. Chem. A.* **2003**, 107, 10346.
9. Parthasarathi, R.; Padmanabhan, J.; Subramanian, V.; Sarkar, U.; Maiti, B.; Chattaraj, P. K. *Internet Electron J. Mol. Des.* **2003**, 2, 798.
10. Pearson, R. G. *Chemical Hardness - Applications from Molecules to Solids*; VCH-Wiley: Weinheim, 1997.
11. Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.
12. Iczkowski, R. P.; Margrave, J. L. *J. Am. Chem. Soc.* **1961**, 83, 3547.
13. Parr, R. G.; Szentpaly, L. V.; Liu, S. *J. Am. Chem. Soc.* **1999**, 121, 1922.
14. Xiu-Fen, Y.; He-Ming, X.; Xue-Dong, G.; Xue-Hai, J. *Chemosphere* **2005**, 59, 467.