DFT Based Electrophilicity Index and QSAR study of Phenols as Anti Leukaemia Agent

1, Farhan Ahmad Pasha, 1,2Hemant Kumar Srivastava, 3Yakoob Beg, 1Pashupati Prasad singh
1Department of Chemistry, M. L. K. P. G. College, Balrampur – 271201 India
2Department of Medicinal Chemistry & Natural Products, School of Pharmacy, Faculty of Medicine, Ein Kerem Campus, The Hebrew University of Jerusalem, Jerusalem-91120, ISRAEL
3Department of Chemistry, Bareilly College, Bareilly (U. P.) India

Abstract: Density Functional reactivity indices based QSAR study of 49 phenol derivatives is presented in this paper. Two different models to describe the anti leukaemia activity of phenols have been made. First QSAR model includes molecular properties like molecular weight (Mw), hardness (η), chemical potential (µ), total energy, and electrophilicity index (ω). Various regression models have been made and regression quality indicates that these descriptors provides valuable information and have significant role in assessment of activity of phenols. Klopman gave first quantum chemical treatment to describe the reactivity of a chemical system in terms of acidic softness E_n and basic softness E_m at atomic level. In this paper we have derived the partial electrophilicity by the multiplication of global electrophilicity index (given by Parr etal) and the acidic softness E_n (given by Klopman). This total electrophilicity index has been used as descriptors along with the other atomic properties like highest negative charge (Q_min) etc in second QSAR model. This model also provides good results. The DFT calculations have been performed by using B88-PW91 GGA energy functional with the DZVP basis set on Cache pro software and the regression models have been made on project leader software associated with CAChe. These DFT models have high predictive power and have sufficient reliability to describe the Anti leukaemia activity of phenols which is clear from its correlation coefficient r^2 and cross validation coefficient rcv^2.

Key words: DFT, Electrophilicity Index, QSAR, Phenol, Anti Leukaemia

INTRODUCTION

The synthesis of novel pharmacologically active molecules with reduced toxicity is of prime interest. Recently, QSAR has gained importance in the field of pharmacological sciences\(^1\). Quantitative structure Activity relationships (QSAR) are predictive tools for a preliminary evaluation of the activity of chemical compounds by using computer-aided models. The Hohenberg and Khon theorem based DFT\(^2\-^4\) provide a major boost to the computational chemistry. The performance of DFT method in description of structural, energetic and magnetic molecular properties has been reviewed quite substantially in recent time. DFT methods are in general capable of generating a variety of isolated molecular properties\(^5\-^12\). Quantitative structure–activity relationship (QSAR) techniques increase the probability of success and reduce time and cost involvement in drug discovery process\(^13\-^14\). In this paper a theoretical technique has been discussed by which the biological activity of hypothetical molecule can be measure prior to their synthesis. This technique shall reduce the drug discovery coast, time and efforts.

THEORY

In DFT the electronegativity commonly known by chemist is defined as negative of partial derivative of energy E of an atomic or molecular system with respect to the number of electron N for a constant external potential v(r)\(^15\)

\[ \mu = -\chi = - \left( \frac{\partial E}{\partial N} \right)_{v(r)} \] (1)

In accordance with the earlier work of Iczkowski and Margrave\(^16\), it should be remarked that when assuming a quadratic relationship between E and N and in a finite difference approximation equation-1 may be rewritten as

\[ \chi = -\mu = \frac{(IE+EA)}{2} \] (2)

Where IE and EA are the vertical ionization energy and electron affinity respectively, there by recovering the electronegativity definition of Mulliken\(^17\). More over theoretical justification was provided for Sanderson’s...
principle of electronegativity equalization which state that when two or more atoms come together to form molecule, their electronegativities become adjusted to the same intermediate value\cite{18,20}. The absolute hardness $\eta$ is defined as\cite{21}

$$\eta = \frac{1}{2} \left( \frac{\partial \mu}{\partial N} \right) \nu (r)$$

Where $E$ is the total energy, $N$ the number of electrons of the chemical species and $\nu (r)$ the external potential. The operational definition of absolute hardness and electronegativity is as

$$\eta = \frac{1}{2} (\text{IP} - \text{EA}) \quad (4)$$

Where IP and EA are the ionization potential and electron affinity respectively, of the chemical species. In the matter of QSAR of chemical system the total energy also plays important role. Total energy of a molecular system is the sum of the total electronic energy, $E_{ee}$ and the energy of internuclear repulsion, $E_{nr}$. The total electronic energy of the system is given by\cite{22}

$$E = \frac{1}{2} P (H + F) \quad (7)$$

Where $P$ is density matrix and $H$ is one-electron matrix. Parr et al have introduced the electrophilicity index\cite{23}, in terms of chemical potential and hardness. The electrophilicity index is a reliable property of a chemical system and may be used as quantum chemical descriptor, the operational definition of electrophilicity index may be written as

$$\omega = \frac{\mu^2}{2.\eta} \quad (9)$$

A more general but important property of a molecular system the molecular weight also has been tested as descriptor. The softness of an atom in a molecule was described by Klopman\cite{24} and modified by Singh et al\cite{25}. The Klopman equation is given below.

$$E_n^{1/2} = IP_n - b^2(IP_n - EA_n) - (\chi_r(C_n)^2/R_n)(1 - 1/\varepsilon)$$

$$[\xi_r - 2b^2 \chi_r(C_n)^2] \quad (10)$$

Where

$E_n^{1/2}$ = Softness of Lewis acid
IP = Ionization potential of an atom in a molecule
EA = Electron affinity of an atom in a molecule
$\varepsilon$ = Dielectric constant of the medium in which reaction is carried out.
$R$ and $q$ = Radius and charge of atom s & r
$C$ = Electron density

The ionization potential of an atom in a molecule (IP), electron affinity of an atom in a molecule (EA), charge on atom in a molecule ($q$) and electron density ($C$) of an atom in a molecule are essential requirements for the solution of Klopman equations. The method for calculation of ionization potential of an atom in a molecule (IP) has been described by Dewar and Morita\cite{26}. The charge and electron density of an atom in a molecule are obtained by DFT\cite{2} calculation on CAChe pro software. Water has been chosen for medium hence the value of dielectric constant is taken as 81\cite{27}.

The method for calculation of electron affinity of an atom in a molecule (EA) has been described by us earlier\cite{28}.

Since The Local acidic softness $E_n$ is a measure of electron accepting tendency while the electrophilicity index ($\omega$) of a molecule has been introduced by Parr et al. On the basis of these two important values we may derived a new parameter the partial electrophilicity ($\omega_p$) by multiplying local acidic softness $E_n$ and electrophilicity index ($\omega$) as

$$\omega_p = E_n \times \omega \quad (11)$$

Here this new parameter also has been tested as descriptor in QSAR study.

**MATERIAL AND METHODS**

The current study has been carried out during sept. to Dec. 2005 at cheminformatics laboratory M. L. K. P. G. college Balrampur and Bareilly college Bareilly India. The 49-substituted Phenol derivatives have been used as study material and are reported under table-1 along with their observed activity ($P_{obs}$) against L1210 Leukaemia cells\cite{29}.

For first step of QSAR prediction, we employed similar methodology of our earlier work\cite{35}. In the second step of QSAR study we have made a modification and derived a new parameter Partial electrophilicity ($\omega_p$). Further we have tested this parameter here as a descriptor along with other descriptors. The values of different descriptors for first and second step of study have been calculated by solving the equations given in theory and the necessary values taken from DFT calculation results. The Project Leader program associated with CAChe pro of Fujitsu, have been used for multiple linear regression (MLR) analysis and various regression equations have been developed for the calculation of activity ($A_{pred}$).
RESULTS

The assessment of activity of a hypothetical compound is of prime interest in order to reduce the drug discovery coast. In this paper forty nine phenols.

Table 1: The Phenols derivatives and their Observed Activity against L1210 Leukaemia cells (76)

| No. | Substituents | Obs Act |
|-----|--------------|---------|
| 1   | 4-OCH3       | 4.48    |
| 2   | 4-OC2H5      | 4.64    |
| 3   | 4-OC3H7      | 4.85    |
| 4   | 4-C4H9       | 5.2     |
| 5   | 4-OC6H13     | 5.5     |
| 6   | H            | 3.27    |
| 7   | 4-NO2        | 3.45    |
| 8   | 4-CI         | 4.29    |
| 9   | 4-I          | 3.86    |
| 10  | 4-CHO        | 3.08    |
| 11  | 4-F          | 3.83    |
| 12  | 4-NH2        | 5.09    |
| 13  | 4-OH         | 4.59    |
| 14  | 4-CH3        | 3.85    |
| 15  | 4-C2H5       | 3.86    |
| 16  | 4-NHCOCH3    | 3.73    |
| 17  | 4-CN         | 3.44    |
| 18  | 4-OC6H5      | 4.97    |
| 19  | Bisphenol-A  | 4.07    |
| 20  | 4-Br         | 4.2     |
| 21  | 4-C(CH3)3    | 4.09    |
| 22  | 3-NO2        | 3.48    |
| 23  | 3-NHCOCH3    | 2.65    |
| 24  | 3-CI         | 3.87    |
| 25  | 3-C(CH3)3    | 3.88    |
| 26  | 3-CH3        | 3.54    |
| 27  | 3-OC3H7      | 3.71    |
| 28  | 3-N(CH3)2    | 4.11    |
| 29  | 3-C2H5       | 3.71    |
| 30  | 3-Br         | 3.82    |
| 31  | 3-CN         | 3.11    |
| 32  | 3-F          | 3.46    |
| 33  | 3-OH         | 3.46    |
| 34  | 3-NH2        | 4.11    |
| 35  | 2-CH3        | 3.52    |
| 36  | 2-CI         | 3.22    |
| 37  | 2-F          | 3.2     |
| 38  | 2-OCH3       | 3.78    |
| 39  | 2-C2H5       | 3.75    |
| 40  | 2-OH         | 4.92    |
| 41  | 2-CH3,4CH3   | 5.03    |
| 42  | 2-NH2        | 5.16    |
| 43  | 2-CN         | 3.3     |
| 44  | 2-NO2        | 3.34    |
| 45  | 2-Br         | 3.44    |
| 46  | 2-C(CH3)3    | 4       |
| 47  | 4-C3H7       | 4.04    |
| 48  | 4-C4H9       | 4.33    |
| 49  | 4-C5H11      | 4.47    |

Table 2: The Values of DFT based Global descriptors of Phenols and their Predicted Activity ($A_{pred}$) by equation-12.

| No. | Mw  | TE  | $\eta$ | $\mu$ | $A_{pred}$ |
|-----|-----|-----|--------|-------|------------|
| 1   | 124.139 | -421.974 | 1.895 | -2.563 | 4.571 |
| 2   | 138.166  | -461.291 | 1.889 | -2.529 | 4.772 |
| 3   | 152.193  | -500.603 | 1.887 | -2.521 | 4.934 |
| 4   | 166.219  | -539.914 | 1.889 | -2.513 | 5.149 |
| 5   | 194.273  | -618.538 | 1.887 | -2.507 | 5.419 |
| 6   | 94.113   | -307.459 | 2.201 | -2.911 | 3.339 |
| 7   | 138.11   | -511.97  | 1.568 | -4.465 | 3.337 |
| 8   | 128.558  | -767.056 | 2.056 | -3.144 | 3.689 |
| 9   | 220.009  | -7226.68 | 1.810 | -3.478 | 3.991 |

MW= molecular weight, $\eta$=hardness, $\mu$= chemical potential, $A_{pred}$= predicted toxicity by eqn. 12. a data points not include in deriving equation.
Table 3: The Values of DFT based Electrophilicity Index, Partial electrophilicity and other descriptors of Phenols with their Predicted Activity (APred) by equation-13.

| No. | En    | εHOMO | εLUMO | ω   | ωP   | Mw    | ET    | APred |
|-----|-------|-------|-------|-----|------|-------|-------|-------|
| 1   | 36.8641 | -4.458 | -0.668 | 6.224098 | 229.4458 | 124.139 | -421.974 | 4.39   |
| 2   | 36.71567 | -4.418 | -0.64 | 6.040872 | 221.7952 | 138.166 | -461.291 | 4.639  |
| 3   | 36.74446 | -4.411 | -0.631 | 6.005892 | 220.6832 | 152.193 | -500.603 | 4.856  |
| 4   | 36.70796 | -4.402 | -0.624 | 5.964677 | 218.9511 | 166.219 | -539.914 | 5.069  |
| 5   | 36.73795 | -4.395 | -0.62 | 5.933881 | 217.9986 | 194.273 | -618.538 | 5.491  |
| 6   | 38.05268 | -5.112 | -0.71 | 9.32555 | 354.8622 | 94.113 | -307.459 | 3.183  |
| 7a  | 37.81513 | -6.033 | -2.897 | 15.63 | 591.0505 | 139.11 | -511.97 | 3.16   |
| 8a  | 37.81567 | -5.2 | -1.088 | 10.16151 | 384.2643 | 128.558 | -767.056 | 3.651  |
| 9   | 38.08212 | -5.288 | -1.667 | 10.94719 | 416.8924 | 220.009 | -722.68 | 3.831  |
| 10  | 37.82559 | -5.346 | -1.802 | 13.96372 | 525.4754 | 119.123 | -399.694 | 3.428  |
| 11  | 37.82559 | -5.022 | -0.989 | 9.107553 | 344.4986 | 112.103 | -406.683 | 3.596  |
| 12  | 36.87762 | -3.842 | -0.416 | 3.882207 | 143.1666 | 109.127 | -362.81 | 5.446  |
| 13  | 37.65544 | -4.513 | -0.735 | 6.503238 | 244.882 | 110.112 | -382.674 | 4.197  |
| 14  | 37.50777 | -4.928 | -0.794 | 10.20186 | 375.4281 | 119.123 | -399.694 | 3.428  |
| 15  | 37.62524 | -5.391 | -1.064 | 11.26728 | 429.6436 | 173.009 | -2880.52 | 3.895  |
| 16  | 30.75947 | -4.641 | -0.814 | 7.117508 | 218.9308 | 151.166 | -386.084 | 3.835  |
| 17  | 37.63149 | -5.735 | -1.802 | 13.96372 | 525.4754 | 119.123 | -399.694 | 3.428  |
| 18  | 37.50777 | -5.255 | -0.794 | 10.20186 | 375.4281 | 119.123 | -399.694 | 3.428  |
| 19a | 37.56671 | -4.711 | -0.714 | 7.352138 | 276.1956 | 228.29 | -731.648 | 5.577  |
| 20  | 38.13197 | -5.691 | -2.176 | 11.21001 | 424.0252 | 122.123 | -420.781 | 3.291  |
| 21  | 37.65544 | -4.513 | -0.735 | 6.503238 | 244.882 | 110.112 | -382.674 | 4.197  |
| 22a | 38.4344 | -5.892 | -3.137 | 14.03721 | 539.5119 | 139.11 | -511.968 | 2.578  |
| 23a | 37.48459 | -9.013 | -1.03 | 8.607233 | 322.6386 | 151.166 | -386.084 | 3.835  |
| 24  | 37.62524 | -5.391 | -1.064 | 11.26833 | 423.9735 | 128.558 | -767.057 | 3.542  |
| 25  | 37.74151 | -4.967 | -0.625 | 8.486022 | 302.2753 | 150.22 | -464.705 | 4.101  |
| 26  | 37.62412 | -4.996 | -0.659 | 8.668314 | 326.1377 | 150.22 | -464.705 | 4.101  |

MW= molecular weight, TE = total energy of system, ω= Electrophilicity Index, En is local softness given by Klopman, εHOMO is energy of HOMO, εLUMO is energy of LUMO, ωP is Partial electrophilicity, APred.= predicted toxicity by eqn. 13. a data points not include in deriving equation
derivatives have been taken with their activity from literature\cite{29} and are reported in table-1. The predictive model of QSAR study has been buildup with the help of following important descriptors

- Molecular Weight  \( (M_W) \)
- HOMO Energy  \( (\varepsilon_{\text{HOMO}}) \)
- LUMO Energy  \( (\varepsilon_{\text{LUMO}}) \)
- Hardness  \( (\eta) \)
- Chemical Potential  \( (\mu) \)
- Electrophilicity Index  \( (\omega) \)
- Total Energy (Hartree)  \( (T_E) \)
- Partial electrophilicity  \( (\omega_p) \)

The values of these descriptors for all the forty nine derivatives have been calculated with the help of DFT method. In the formation of first QSAR model we have generated various equations by employing all the variables and the best-fitted equation of this class is equation-12.

\[
P_A=0.0114432*M_W+0.00012912*T_E-1.77179*\eta+1.03748*\mu+9.22152 \tag{12}
\]

\[r^2_{\text{CV}}=0.713606 \text{ and } r^2=0.821269\]

This model includes the molecular weight, total energy, hardness and chemical potential. All these values are molecular property and we already have tested these values as a molecular descriptor in our previous communication\cite{30-35}. The predicted activity \( (A_{\text{Pred}}) \) from equation-12 is reported in table-2. On the basis of statistical quality of result it is clear that one can use this equation to predict the antileukemia activity of a hypothetical compound of similar series. However in search of a more significant model and to recognized the Partial electrophilicity \( (\omega_p) \) as a QSAR descriptor we have performed the study at atomic level and proceed to second step of QSAR study.

The second QSAR model has been formed with the help of newly derived descriptor the Partial electrophilicity \( (\omega_p) \) along with Molecular Weight \( (M_W) \), HOMO Energy \( (\varepsilon_{\text{HOMO}}) \), LUMO Energy \( (\varepsilon_{\text{LUMO}}) \), Electrophilicity Index \( (\omega) \) and Total Energy (Hartree) \( (T_E) \). In this model we have generated various equations by employing all the variables and the only best fitted equation-13 is reported here.

\[
A_{\text{Pred}}=0.272295*E_n+4.65224*\varepsilon_{\text{HOMO}}+0.344202*\varepsilon_{\text{LUMO}}+1.63431*\omega-0.0248194*\omega_p+0.0149739*M_W+0.000174617*E_T+9.05945 \tag{13}
\]

\[r^2_{\text{CV}}=0.713606 \text{ and } r^2=0.821269\]

The predicted activity \( (A_{\text{Pred}}) \) from equation-13 is reported in table-3. On the basis of this model we can also justify the validity of newly derived descriptor Partial electrophilicity \( (\omega_p) \)

**CONCLUSION**

The first model involves all the descriptors which are basically energy related values and they are capable to describe the activity successfully however the second model includes energy values along with the electron accepting tendency of a molecule. Here we have derived a new parameter the Partial electrophilicity \( (\omega_p) \) and tested it as a QSAR descriptor. The good result suggests us to realize the validity of newly derived descriptor electronic exchange in biochemical interaction with in the body. This study results a framework by which one can calculate the activity of any hypothetical compound of the series prior than their synthesis. The study is also helpful in the determination of effect of any particular phenol derivatives of this series over Leukaemia cells.

**ACKNOWLEDGMENT**

We are thankful to Prof. Y. B. Tripathi (Institute of Medical Sciences, BHU, INDIA) for valuable suggestions.

**REFERENCES**

1. Smeyers, Y. G., L. Bouniam, N. J. Smeyers, A. Ezzamarty, A. Hernandez-Laguna, A., and C. I. Sainz-Diaz, 1998 *Eur. J. Med. Chem.*, 33, 103
2. Parr R. G., and W. Yang, 1989, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York
3. Kohn, W., A. D. Becke, and R. G. Parr, 1996 *J. Phys. Chem.*, 100: 12974.
4. Hohenberg, P., and W. Kohn, 1964, *Phys. Rev. B*, 136: 864.
5. Chattaraj, P. K., A. Cedillo, and R. G. Parr, 1991, *J. Phys. Chem.*, 103: 7645.
6. Ayers, P. W., and R. G. Parr, 2000, *J Am Chem Soc*, 122: 2010, 2000.
7. De Proft, F., J. M. L. Martin, and P. Geerlings, 1996, *Chem. Phy Let.*, 250: 393.
8. Geerlings, P., F. De Proft, and J. M. L. Martin, *In Theoretical and Computational Chemistry*; Seminario, J., Ed.; Elsveir; Amsterdam. 1996; Vol-4 (Recent Developments in Density Functional Theory), p773
9. De Proft, F., J. M. L. Martin, and P. Geerlings, 1996, *Chem. Phys Let.*, 256: 400
10. De Proft, F., and P. Geerlings, 1997, *J Chem Phys*, 106: 3270
11. Geerlings, P., F. De Proft, and W. Langenaeker, 1996, *Adv. Quantum Chem.* 33: 303
12. Parr, R. G., R. A. Donnelly, M. Levy, and W. E. Palke, 1978, *J. Chem. Phys.*, 68: 3801
13. Hansch, C., P. G. Sammes, and J. B. Taylor, Computers and the medicinal chemist; in: Comprehensive Medicinal Chemistry, vol. 4, Eds. Pergamon Press, Oxford, 1990, pp. 33–58
14. R. Franke, Theoretical Drug Design Methods, Elsevier, Amsterdam, 1984.
15. Parr, R. G., R. A. Donnelly, M. Levy, and W. E. Palke, 1978, J. Chem. Phys., 68:3801
16. Iczkowski, R. P., and J. L. Margrave, 1961, J Am Chem Soc, 83: 3547
17. Mulliken, R. S., 1934, J. Chem. Phys. 2: 782.
18. Sanderson, R. T., 1955, Science, 121: 207.
19. Sanderson, R. T., 1976, Chemical Bonds and Bond Energy, (Academic, New York), 1976.
20. Sanderson, R. T., 1983, Polar Covalence; Academic Press; New York.
21. Parr, R. G., and R. G. Pearson, 1983, J Am Chem Soc 105: 7512.
22. Clare, B. W., 1995, Structure-Activity Correlations for Psychotomimetics .3. Tryptamines, Aust. J. Chem., 48, 1385.
23. Parr, R. G., L. V. Szentpaly, and S. Liu, 1999, J. Am. Chem. Soc., 121: 1922.
24. Klopman, G., 1968, J. Am. Chem. Soc.; 90: 223.
25. Singh, P. P., S. K. Srivastava, and A. K. Srivastava, 1980, J Inor Nucl Chem, 42: 521.
26. Dewar, M. J. S., and T. F. Morita, 1969, J. Am. Chem. Soc., 91: 796.
27. Daniels, D. J., Surface-penetrating radar-IEE Radar, Sonar, Navigation and Avionics, Series 6: London, The Institute of Electrical Engineers, 1996, p.320.
28. Singh, P. P., F. A. Pasha, and H. K. Srivastava, 2003, QSAR & Combi. Sci., 22: 841.
29. Selassie, C. D., A. J. Shusterman, S. Kapur, R. P. Verma, L. Zhang, and C. Hansch, 1999, J. Chem. Soc. Perkin Trans 2: 2729.
30. Singh, P. P., H. K. Srivastava, and F. A. Pasha, 2004, Bioorg. Med. Chem., 12 (1): 171.
31. Pasha, F. A., H. K. Srivastava, A. Srivastava and P. P. Singh, 2006, QSAR & Comb. Sci., in press.
32. Srivastava, H. K., F. A. Pasha, and P. P. Singh, 2005, Int. J. Quantum. Chem., 103 (3): 237.
33. Pasha, F. A., H. K. Srivastava, and P. P. Singh, 2005, Molecular Diversity, 9(1-3): 215.
34. Pasha, F. A., H. K. Srivastava, and P. P. Singh, 2005, Int. J. Quantum. Chem., 104 (1): 87.
35. Pasha, F. A., H. K. Srivastava, and P. P. Singh, 2005, Bioorg. Med. Chem., 13(24): 6823.