Quantum Chemical and Energy Descriptors Based Qsar Studies of Triazines Inhibiting Dihydrofolate Reductase

Mithilesh Tiwari[1], S. K. Singh[2], Lakshmi Gangwar[3]

[1] Sri Ram Murti Smarak Institute of Technology & Management, Lucknow, UP, India

[2] Institute of Engineering & Technology, Sitapur Road, Lucknow, UP, India

[3] SMS Institute of Technology, Lucknow, UP, India

Abstract: ***Among all the 25 QSAR models PA 1 to PA 25, the number of good QSAR models is 10 whose regression coefficient is greater than 0.7. In all the best 10 QSAR models, heat of formation is common. It means the best descriptor to predict the activities are the heat of formation. Also, the predicted activity obtained by taking heat of formation as single descriptor possesses the good value of regression coefficient which is 0.721430.***

Keywords: QSAR**,** Descriptors, triazines, Dihydrofolate reductase, Steric energy, Heat of formation**.**

1. Introduction

On the frontier of chemical structure-activity relationship, especially in bio- and medicinal chemistry, so little solid theory is at hand on which to build that all kinds of purely empirical ideas need to be of great help in sorting out important structure-activity features which can then be used to form more firmly based theory. Techniques such as pattern recognition [1,2], discriminate analysis [3], cluster analysis [2,4], and regression analysis,[5] which has been developed and used heavily out side of chemistry are now beginning to be used by those working with structure-activity relationship. In this paper the Multilinear regression analysis has been applied for QSAR study. The relationship has been worked out between the Log 1/C values of a series of compounds and certain quantum chemical and energy descriptors. The compounds chosen for study are described as below.

Baker and few graduate students [6] synthesized variations of I to achieve II, a drug now in clinical trials against cancer. Baker's group synthesized 256 variations of I and studied their inhibiting effect on dihydrofolate reductase isolated from Walker 256 and L1210 leukemia tumors. He did demonstrate vividly that starting at the enzyme level rather than with whole animals constitutes a powerful technique for drug development. This approach has also been brilliantly exploited by Hitching and his group [7,8] in the development of allopurinol for gout and the new antibacterial agent, trimethoprim. Out of 256 compounds synthesized by Baker the QSAR study of 50 compounds has recently been reported [9]. The remaining compounds leave a wide scope for their QSAR study. It is proposed to make their structural relationship with quantum chemical and energy parameters. The remaining compounds have been divided into three groups as detailed below.

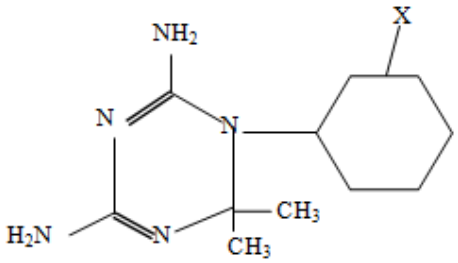
i. The X substituents of compound-1 have SO_2F at position –3 of the phenyl ring.

ii. The X substituents of compounds –1 have SO_2F at position –4 of the phenyl rings.

iii. The X substituents of compounds –1 have no -SO_2F.

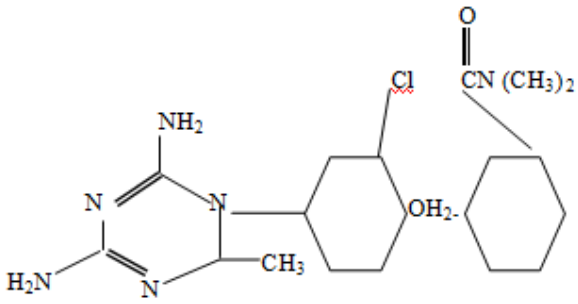
The work of the paper has been limited to compounds of serial –1, that is X substituents having - SO_2F at position-3 of phenyl ring. The compounds of this series are listed in table-1, along with their inhibiting activity in terms of Log 1/C.

Compound-I



[4, 6-diamino-1, 2, dihydro-2, 2,dimethyl-1-1 (x-phenyl)-s-triazines]

Compound-II



[Dimethyl acetyl benzamide derivative of 1]

Table 1: Log 1/C data for reversible inhibition of dihydrofolate reductase by 2, 6-Diamino-1, 2 dihydro-2, 2 dimethyl-1- (X phenyl)-S-triazines

Comp	*I. X*	*Log 1/C*
C1	3-Cl, 4 -OCH2C6H4 -3'-CONHC6H4 -4" -	6.92
C2	3-Cl, 4 -OCH2C6H4 -4'-CONHC6H4 -4" -	6.92
C3	3-OCH2CONHC6H4-4' -SO2F	6.92
C4	3-Cl, 4 -(CH2) 4C6H3-5'-Cl, 2'-SO2F	7.06
C5	3-Cl, 4 -O (CH2) 3 OC6H4-4'-SO2F	7.07
C6	3-Cl, 4 -OCH (CH3)- CONHC6H4-4'-SO2F	7.13
C7	3-Cl, 4 -O (CH2) 2 O (CH2) 2OC6H4-4'-	7.14
C8	3Cl, 4 -O (CH2) 3 CONH-C6H4-4'-SO2F	7.15
C9	3-Cl, 4 -O (CH2) 3 CONHC6H4-3'- SO2F	7.17
C10	3 - (CH2) 2 CONHC6H4-4'-SO2F	7.19
C11	3-Cl, 4 -OCH2C6H4-3"-SO2F	7.24
C12	3-Cl, 4 -OCH2C6H4-2'- CONHC6H4-4"-	7.24
C13	3-Cl, 4 –O (CH2) 4- CONHC6H4-4"-SO2F	7.24
C14	3-Cl, 4 -OCH2C6H3-5'- Cl, 2'-SO2F	7.27
C15	3 -SO2F	7.27
C16	3 -Cl, 4 -O (CH2) 3 NH- CONHC6H4-3'-	7.28
C17	3-Cl, 4 –O (CH2) 3-C6H4-4'-SO2F	7.34
C18	3 -CH2CONHC6H4-4'-SO2F	7.34
C19	3 -Cl, 4 -OCH2C6H4-6-Cl, 3'-SO2F	7.38
C20	3 -Cl, 4 -OCH2C6H3-2'-CH3, 4'-SO2F	7.38
C21	3 -Cl, 4 -S (CH2) 2- CONHC6H4-4'-SO2F	7.39
C22	3-Cl, 4 -O (CH2) C6H4-4'- CONHC6H4-3"-	7.41
C23	3 -Cl, 4 -SCH2CONHC6H4-4'-SO2F	7.42
C24	3 -Cl, 4 -OCH2C6H3-3'-Cl, 2' SO2F	7.42
C25	3 -Cl, 4 -OCH2CONH-C6H4-4'-SO2F	7.43

Quantum chemical [10] and energy descriptors [11, 12] have been prominently used in the last decade for QSAR studies. Following are the descriptors:

(i) Heat of Formation	(ΔHf)
(ii) Steric Energy	(SE)
(iii) Total Energy	(TE)
(iv) HOMO Energy	(∈HOMO)
(v) LUMO Energy	(∈LUMO)
(vi) Absolute Hardness	(η)
(vii) Electronegativity	(χ)

The geometries of all the derivatives mentioned in table-1, have been first optimized, and then the values of the descriptors have been evaluated with the help of software, described in materials and method. MLR analyze the above descriptors to predict value of Log 1/C. The quality of the regression analysis has been adjudged by correlation coefficient [12]. The descriptors or the combination of descriptors, providing the best result, has been recognized. This helps is predicting the Log 1/C value of any new derivative of triazine.

2. Materials and Method

Twenty five derivatives of compound-I have been taken as study material. They are listed in table-1 along with their biological activity in terms of Log 1/C, where C is molar concentration of inhibitor causing 50% reversible inhibitions of enzyme. Enzyme dihydrofolate reductase isolated from Leukemia tumor forms covalent bond with highly reactive derivatives of triazines. The compounds are strongly electrophilic, react through carbonium ion intermediate and form covalent bonds with amino, hydroxyl, carboxyl groups [13]. Consequences are:

(i) bifunctional agents
(ii) DNA-DNA strand and DNA protein cross links.

The values of quantum chemical descriptors such as HOMO energy [14], LUMO energy [15], absolute hardness [16-19], electronegativity [20], total and steric energy [23-25], heat of formation [26] have been evaluated with the help of CAChe software using PM3 hamiltonian. Parr et. Al. [21] define the electronegativity as the negative of chemical potential

$$\chi = -\mu = -(\partial E/\partial N)_v \quad \text{(i)}$$

The absolute hardness η is defined as [22]

$\eta = 1/2(\delta\mu/\delta N)v_{(r)} = 1/2\ (\delta^2E/\delta^2N)v_{(r)}$ (ii) where E is the total Energy, N the number of electrons of the chemical species and $v_{(r)}$ the external potential

The corresponding global softness S, which bears an inverse relationship with the global hardness, is defined as

$S = 1/2\eta = (\partial N/\partial\mu)v_{(r)}$ (iii) The operational definition of absolute hardness, global softness and electronegativity is defined as:

$\eta = 1/2\ (IP-EA)$ (iv)
$S = 1/(IP-EA)$ (v)
$\chi = -\mu = 1/2(IP + EA)$ (vi)

where IP and EA are the Ionization Potential and Electron Affinity respectively, of the chemical species. According to the Koopman's theorem the IP is simply the eigen value of HOMO with change of sign and EA is the eigen value of LUMO with change of sign. [22]; hence, Eqs. 4-6 can be written as

$\eta = 1/2\ (\varepsilon\ LUMO - \varepsilon\ HOMO)$ (vii)
$S = 1/(\varepsilon\ LUMO - \varepsilon\ HOMO)$ (viii)
$\chi = 1/2\ (\varepsilon\ LUMO + \varepsilon\ HOMO)$ (ix)

The energy descriptors [23-26] are useful parameters for describing QSAR of a chemical system. A more useful quantity is the heat of formation of the compound from its elements in their standard state. This is equal to the energy required to ionize the valence electrons of the atoms involved. The heat of formation is defined as

$$\Delta H_f = E_{elect} + E_{nuc} - E_{isol} + E_{atom} \quad \text{(x)}$$

where E_{elect} is the electronic energy, E_{nuc} is the nuclear-nuclear repulsion energy, E_{isol} is the energy required to strip all the valence electrons of all the atoms in the system and E_{atom} is the total heat of atomization of all the atoms in the system. 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-

E = ½.P (H + F) (xi) where P is the density matrix and H is the one electron matrix. These parameters and the charges on atoms has been obtained from PM3 [27] calculations. Multilinear regression (MLR) analysis has been performed by using Project Leader program associated with CAChe Pro software of Fujitsu.

3. Results and Discussion

The values of the descriptors; Heat of Formation (ΔHf), Steric Energy (SE), Total Energy (TE), HOMO Energy (∈HOMO), LUMO Energy (∈LUMO), Absolute Hardness (η) and Electronegativity (χ) of 25 derivatives of triazine have been evaluated and given in the table-2.

Table 2: Values of the quantum chemical and energy descriptors of the derivatives of triazine

Comp	*Heat of Formation*	*Steric Energy*	*Total Energy*	*HOMO Energy*	*LUMO Energy*	*Absolute Hardness*	*Electronegativity*	*Activity*
C1	104.651	-53.108	-295.559	-8.649	-1.100	3.774	4.874	6.920
C2	104.651	-49.867	-295.538	-8.804	-1.170	3.817	4.987	6.920
C5	106.920	-26.048	-258.959	-8.872	-0.853	4.009	4.863	7.070
C6	117.827	-42.847	-266.491	-8.973	-1.085	3.944	5.029	7.130
C7	107.978	-21.483	-278.272	-8.869	-0.842	4.014	4.856	7.140
C8	108.129	-42.078	-273.663	-8.836	-0.954	3.941	4.895	7.150
C9	110.432	-43.999	-273.667	-8.822	-0.865	3.978	4.843	7.170
C10	111.734	-46.895	-242.574	-8.851	-0.912	3.969	4.881	7.190
C11	109.491	-31.693	-232.471	-8.776	-1.032	3.872	4.904	7.240
C12	112.491	-51.374	-295.569	-8.713	-1.097	3.808	4.905	7.240
C13	119.491	-44.016	-280.834	-8.753	-0.971	3.891	4.862	7.240
C14	109.944	-31.410	-244.220	-8.779	-1.225	3.777	5.002	7.270
C16	114.095	-78.672	-283.059	-8.850	-0.833	4.009	4.841	7.280
C17	116.003	-27.797	-246.788	-8.861	-0.947	3.957	4.904	7.340
C18	115.003	-47.702	-235.418	-8.883	-1.055	3.914	4.969	7.340
C19	121.608	-31.713	-244.238	-8.755	-1.267	3.744	5.011	7.380
C20	119.608	-30.705	-239.656	-8.852	-1.043	3.905	4.947	7.380
C21	118.759	-47.005	-263.536	-8.855	-1.063	3.896	4.959	7.390
C22	117.061	-52.844	-295.558	-8.774	-0.869	3.952	4.822	7.410
C23	102.213	-33.088	-256.365	-8.992	-1.251	3.870	5.121	7.420
C24	119.213	-27.095	-244.212	-8.729	-1.156	3.787	4.942	7.420
C25	112.364	-46.087	-259.342	-8.883	-1.042	3.920	4.962	7.430

Outlier Compounds are C3, C4 and C15. Outlier compounds are those compounds which are excluded in multilinear regression (MLR) analysis. Values of regression coefficients (r^2) and cross-validation coefficients (rCV^2) have been calculated for each MLR equation. QSAR model is characterized by the values of regression coefficients (r^2) and cross-validation coefficients (rCV^2). If the value of regression coefficient is greater than 0.5 then the QSAR model is said to have good predictive power besides the value of cross-validation coefficient is greater than 0.2. As the value of regression coefficient increases, the predictive power increases. The maximum value of regression coefficient may be unity. Combination of quantum chemical and energy descriptors in the predicted activities PA1 to PA25 are shown in the table-3.

Table 3: Combination of quantum chemical and energy descriptors in the predicted activities PA1 to PA25

Predicted Activity	*First descriptor*	*Second descriptor*	*Third descriptor*
PA1	Heat of Formation	Steric Energy	
PA2	Heat of Formation	Total Energy	
PA3	Heat of Formation	HOMO Energy	
PA4	Heat of Formation	LUMO Energy	
PA5	Heat of Formation	Absolute Hardness	
PA6	Heat of Formation	Electronegativity	
PA7	Steric Energy	Total Energy	
PA8	Steric Energy	HOMO Energy	
PA9	Steric Energy	LUMO Energy	
PA10	Steric Energy	Absolute Hardness	
PA11	Steric Energy	Electronegativity	
PA12	Total Energy	HOMO Energy	
PA13	Total Energy	LUMO Energy	
PA14	Total Energy	Absolute Hardness	
PA15	Total Energy	Electronegativity	
PA16	HOMO Energy	LUMO Energy	
PA17	HOMO Energy	Absolute Hardness	
PA18	HOMO Energy	Electronegativity	
PA19	LUMO Energy	Absolute Hardness	
PA20	LUMO Energy	Electronegativity	
PA21	Absolute Hardness	Electronegativity	
PA22	Heat of Formation	Steric Energy	Total Energy
PA23	Heat of Formation	Steric Energy	HOMO Energy
PA24	Heat of Formation	Steric Energy	LUMO Energy
PA25	Heat of Formation	Steric Energy	Absolute Hardness

QSAR models PA1 to PA25 are developed and given by the following MLR equations-

1. PA1=0.0579916*ΔHf+0.000449905*SE+0.905356
rCV^2=0.70403
r^2=0.722347

2. PA2=0.0550563*ΔHf+0.00217873*TE+1.78497
rCV^2=0.713752
r^2=0.733793

3. PA3=0.0579875*ΔHf+0.0134586**∈HOMO+1.00593
rCV^2=0.690059
r^2=0.721887

4. PA4=0.0579688*ΔHf-0.00089436**∈LUMO+0.889769
rCV^2=0.653403
r^2=0.72143

5. PA5=0.0579665*ΔHf-0.0811096**η+1.20378
rCV^2=0.686067
r^2=0.722833

6. PA6=0.057986*ΔHf-0.0118822**χ+0.94652
rCV^2=0.584938
r^2=0.721628

7. PA7=-0.00168238*SE+0.0080979*TE+9.64485
rCV^2=0.102934
r^2=0.178305

8. PA8=0.000414605*SE-0.00563828**∈HOMO+7.61439
rCV^2=-1.36326
r^2=0.000496905

9. PA9=0.000696873*SE-0.0555125**∈LUMO+7.61834
rCV^2=-0.0764998
r^2=0.00199318

10. PA10=-0.000254574*SE-0.111437**η+8.06759
rCV^2=-0.0671299
r^2=0.00162192

11. PA11=0.000590646*SE+0.0227784**χ+7.55948
rCV^2=-0.982078
r^2=0.00085062

12. PA12=0.00757178*TE+0.0342852**∈HOMO+9.87485
rCV^2=-0.94715

r^2=0.169602

13. PA13=0.00777357*TE+0.0883477***∈**LUMO+9.71642
rCV^2=-0.00671231
r^2=0.171919

14. PA14=0.00745463*TE-0.0613717***η**+9.78259
rCV^2=0.150068
r^2=0.167469

15. PA15=0.00763948*TE-0.0528724***χ**+9.85034
rCV^2=-1.89273
r^2=0.170502

16. PA16=0.090766***∈**HOMO-0.177534***∈**LUMO+8.26003
rCV^2=-1.40108
r^2=0.00478499

17. PA17=-0.0867681***∈**HOMO-0.355068***η**+8.26003
rCV^2=-1.40108
r^2=0.00478499

18. PA18=0.2683***∈**HOMO+0.355068***χ**+8.26003
rCV^2=-1.40108
r^2=0.00478499

19. PA19=-0.0867681***∈**LUMO-0.181532***η**+8.26003
rCV^2=-1.40108
r^2=0.00478499

20. PA20=-0.2683***∈**LUMO-0.181532***χ**+8.26003
rCV^2=-1.40108
r^2=0.00478499

21. PA21=-0.2683***η**+0.0867681***χ**+8.26003
rCV^2=-1.40108
r^2=0.00478499

22. PA22=0.0549892*ΔHf-0.000106627*SE+0.00222513*TE+1.8005
rCV^2=0.69631
r^2=0.733839

23. PA23=0.0579913*ΔHf+0.000462427*SE-0.000732299***∈**HOMO+0.89949
rCV^2=0.688752
r^2=0.722347

24. PA24=0.0579609*ΔHf+0.000649056*SE-0.0292041***∈**LUMO+0.887091
rCV^2=0.655774
r^2=0.722771

25. PA25=0.0579705*ΔHf+7.51152e-005*SE-0.0729104***η**+1.17459
rCV^2=0.683655
r^2=0.722844

Among all the 25 QSAR models PA 1 to PA 25, the number of good QSAR models is 10 whose regression coefficient is greater than 0.7. Good QSAR models in the decreasing order of predictive power are shown in the table-4 along with the descriptors used in them.

Table 4: Good QSAR models in the decreasing order of predictive power

S. No.	*Predicted Activity*	*rCV^2*	*r^2*	*Descriptors used in QSAR model*
1	PA22	0.696310	0.733839	Heat of Formation, Steric Energy, Total Energy
2	PA2	0.713752	0.733793	Heat of Formation, Total Energy
3	PA25	0.683655	0.722844	Heat of Formation, Steric Energy, Absolute Hardness
4	PA5	0.686067	0.722833	Heat of Formation, Absolute Hardness
5	PA24	0.655774	0.722771	Heat of Formation, Steric Energy, LUMO Energy
6	PA1	0.704030	0.722347	Heat of Formation, Steric Energy
7	PA23	0.688752	0.722347	Heat of Formation, Steric Energy, HOMO Energy
8	PA3	0.690059	0.721887	Heat of Formation, HOMO Energy
9	PA6	0.584938	0.721628	Heat of Formation, Electronegativity
10	PA4	0.653403	0.721430	Heat of Formation, LUMO Energy

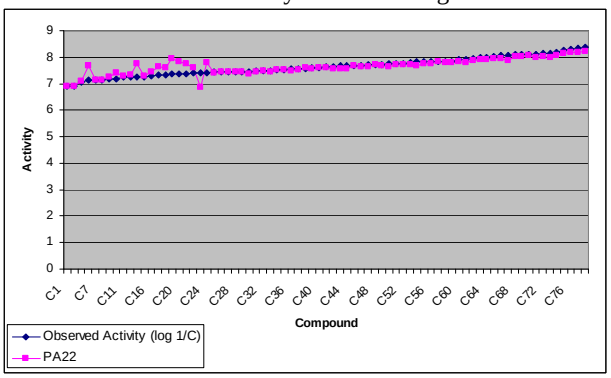
Out of these 10 good QSAR models, best 02 QSAR models are described below:

QSAR Model 1 (PA 22)

This QSAR model contains the descriptors heat of formation, steric energy and total energy. Predicted activity is given by the equation-
PA22=0.0549892*ΔHf-0.000106627*SE+0.00222513*TE+1.8005
rCV^2=0.69631
r^2=0.733839

Value of regression coefficient is 0.733839, which indicates that the predictive power of QSAR model given by PA22 is very good and it can be used to predict the activity of any compound of triazine series. Graph between predicted and observed activities is shown in Graph-1.

Graph 1: Graph between predicted activity PA22 and observed activity in terms of log 1/C



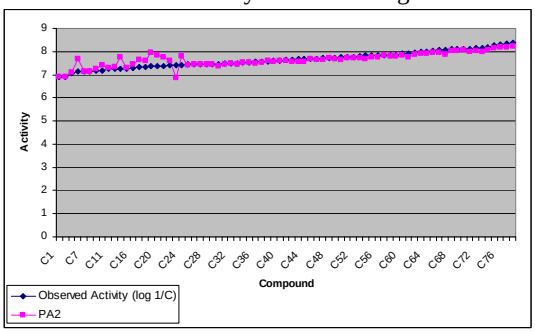
QSAR Model 1 (PA 2)

This QSAR model contains the descriptors heat of formation and total energy. Predicted activity is given by the equation-
PA2=0.0550563*ΔHf+0.00217873*TE+1.78497

rCV^2=0.713752
r^2=0.733793

Value of regression coefficient is 0.733793, which indicates that the predictive power of QSAR model given by PA2 is very good and it can be used to predict the activity of any compound of triazine series. Graph between predicted and observed activities is shown in Graph-2.

Graph 2: Graph between predicted activity PA2 and observed activity in terms of log 1/C



4. Conclusions

Best QSAR model PA 22 contains the descriptors heat of formation, steric energy, total energy and HOMO energy. Value of regression coefficient is 0.73839, which is too greater than 0.5 and hence indicates that the predictive power of QSAR model given by PA22 is very good. In all the best 10 QSAR models, heat of formation is common. It means the best descriptor to predict the activities are the heat of formation. Also, the predicted activity obtained by taking heat of formation as single descriptor possesses the good value of regression coefficient which is 0.721430. Predicted activity using heat of formation as descriptor is given by:

PA4=0.0579688*ΔHf-0.00089436*∈LUMO+0.889769
rCV^2=0.653403
r^2=0.721430

All the combinations of the descriptors give good QSAR model in which the descriptor heat of formation is present.

References

[1] B. R. Kowalski, C. F Bender, J. Am. Chem. Soc., 96, 918,1974.
[2] S. S. Schiffman, Science, 185, 112, 1974.
[3] Y. C. Martin, J. B. Holland, C.H. Jarboe, N. Plotnikoff, J. Med. Chem., 17, 409, 1974.
[4] C. Hansch, S.H.Unger, C.H. Forsythe, J. Med. Chem., 16, 1217, 1973.
[5] C. Hansch, Chem.-Biol. Interact. 9, (1974).
[6] B. R. Baker, G. L. Lourens, J. Med. Chem., 10, 1113,1967.
[7] G. H. Hitchings, J. Infect. Dis., 128, S433, 1973.
[8] J. J. Burchall, J. Infect. Dis., 128, S437, 1973.
[9] P. P. Singh, S. Singh, R. B. Srivastava, Med. Chem. A.I.J., 2(1), 2006.
[10] M. Karelson., V. S. Labanov, A. R. Katritzki, Chem. Rev., 96, 1027, 1996.
[11] F. A. Pasha, H. K. Srivastava, P. P. Singh, Int. J. Quantum Chem., 104, 87, 2005.
[12] F. A. Pasha, H. K. Srivastava, P. P. Singh, Bioorg. Med. Chem., 13(24), 6823-29, 2005.
[13] F. A. Pasha, H. K. Srivastava, P. P. Singh, Bioorg. Med. Chem., 12, 171, 2004, Internet site-anticancer drug.
[14] G. Tripani, A. Carotti, M. Franco, A. Latrofa, G. Genchi, G. Liso, Eur. J. Med. Chem., 28, 13, 1993.
[15] P. G. De Benedetti, M. C. Menziani, M. Cocchi, C. Frassineti, Quant. Struct. Act. Relat., 6, 51, 1987.
[16] Z. Zhou, R. G. Parr, J. Am. Chem. Soc., 112, 5720, 1990.
[17] R. G. Parr, P. K. Chattaraj, J. Am. Chem. Soc., 113, 1854, 1991.
[18] J. R. Sabin, S. B. Trickey, P. Apell, Oddershede, J. Int. J. Quantum Chem., 77, 358, 2000.
[19] P. K. Chattaraj, A. Cedillo, R. G. Parr, J. Chem. Phys., 103, 7645, 1995.
[20] C. J. Umrigar, M. P. Nightingale, K. J. Runge, J. Chem. Phys., 99, 2865, 1993.
[21] R. G. Parr, R. A. Donnelly, M. Levy, W. E. Palke, J. Chem. Phys., 68, 3801, 1978.
[22] R. G. Parr, R. G. Pearson, J. Am. Chem. Soc., 105, 7512, 1983.
[23] B. W. Clare, C. T. Supuran, J. Pharm Sci., 83, 768, 1994.
[24] N. Bodor, Z. Gabanyi, C. K. Wang, J. Am. Chem. Soc., 111, 3783, 1989.
[25] B. W. Clare, Aust J Chem., 48, 1385, 1995.
[26] P. G. De Benedetti, M. C. Menziani, M. Cocchi, C. Frassineti, Quant Struct Act Relat., 6, 51, 1987.
[27] J. J. P. Stewart, J. Comp. Chem., 10, 209, 1989.

Author Profile

S. K. Singh received the M.Sc. degrees in Chemistry from Allahabad University, Allahabad in 1987 and Ph.D. in 1993.Currently he is AP at Deptt of Applies Sciences & Humanities at IET, Lucknow. He has published 45 research papers, 03 Text Books and guiding 08 Ph.D. Students. He has research as well as teaching experience of about 23 years.