Quantum chemical QSAR study of 1-phenyl-X-benzimidazoles as inhibitors of PDGFR tyrosin kinase

Article *in* International Journal of PharmTech Research · October 2011

3 authors:

Rita Elias
University of Basrah
22 PUBLICATIONS **52** CITATIONS
SEE PROFILE

Sadiq M-H Ismael
University of Basrah
11 PUBLICATIONS **14** CITATIONS
SEE PROFILE

Bahjat Saeed
University of Basrah
68 PUBLICATIONS **250** CITATIONS
SEE PROFILE

International Journal of PharmTech Research
CODEN (USA): IJPRIF ISSN : 0974-4304
Vol.3, No.4, pp 2183-2189, Oct-Dec 2011

Quantum Chemical QSAR Study of 1-phenyl-X- benzimidazoles as Inhibitors of PDGFR Tyrosin Kinase

Rita S. Elias[1], Sadiq M-H. Ismael[2] and Bahjat A. Saeed*[2]

[1]Department of Pharmaceutical Chemistry, College of Pharmacy,
[2]Department of Chemistry, College of Education,
University of Basrah, Iraq.

**Corres. Author: bahjat.saeed@uobasrah.edu.iq*
Phone No: 07802410050

Abstract: A series of 1-phenyl-X-benzimidazole that already tested for their inhibitory action against the PDGF-β-receptor tyrosin kinase were subjected to QSAR study using the quantum chemical descriptors; ε_{HOMO}, ε_{LUMO}, ΔE, η, χ, S and Mullikan atomic charges. In this study several QSAR equations were formulated which was able to explain 86 -93% of the variance in the data. R^2 values were in the range 0.86-0.93 while s values were in the range 0.06-0.27. Atomic charges on the nitrogen atoms was found to have important role in determining the biological activity of the studied benzimidazoles, since the activity increases with increasing charge on nitrogen atom number 1, and it could be concluded that biological activity may be improved if the phenyl ring in these molecules is substituted by electron donating substituents.
Keywords: benzimidazoles. QSAR, tyrosin kinase, RM1, PDGFR.

Introduction

Quantitative structure–activity relationships (QSAR) are widely used in modeling a variety of physico-chemical parameters as well as biological activity of chemically active compounds[1]. It is a central concept in chemistry that molecular properties are intimately related to molecular structure. With the large variety of quantum chemical programs available today, it is possible to draw correct picture about the molecular structure of most molecules. To establish a QSAR model it is necessary to define the molecular structure and the influences governing measured values of the physical properties or biological activity. Quantum chemical descriptors have been extensively used in Quantitatine Structure Activity Relationship studies in biochemistry[2]. The use of quantum chemical descriptors in the development of QSAR has received attention due to reliability and versatility of prediction by these descriptors. For the calculation of the quantum chemical molecular descriptors used in QSAR studies semi empirical methods such as AM1, PM3, RM1 and PM6 mainly have been used[3]. However, DFT method has been used recently for the prediction of physiological and biological properties of organic molecules[4]. HOMO and LUMO energies[5], HOMO-LUMO energy gap (ΔE), dipole moment (μ), atomic charges[6], softness (S)[7], and electronegativity (χ)[8] are of descriptors that usually used in quantum based QSAR studies.

Softness (S) [9], is a property of molecules that measures the extent of chemical reactivity,

$$S = 1/2\eta \qquad (1)$$

Using Koopmans' theorem for closed-shell molecules, η and χ can be defined as follows[10],

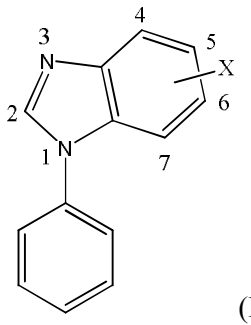
$\eta = (I - A) = ½ (\varepsilon LUMO - \varepsilon HOMO)$ (2)

$\chi = ½(I + A)$ (3)

$I = -\varepsilon HOMO$ and $A = -\varepsilon LUMO$ (4)

Where, I and A are ionization potential and electron affinity of the molecule respectively.

The receptors having integral PTK are known as receptor tyrosine kinases (RTKs). RTKs form a large and important class of cell-surface receptors whose ligands are soluble or membrane-bound peptide/protein, including insulin and epidermal growth factor. Binding of a ligand to this type of receptor stimulates the receptor's protein tyrosine kinase activity, which subsequently stimulates a signal-transduction cascade leading to changes in cellular physiology and/or patterns of gene expression[11]. This work represents a Quantitative Structure Activity Study of the inhibitory action of 1-phenyl-X-benzimidazoles (I) against the phosphorylation of a model glutamate tyrosine copolymer substrate by isolated mouse PDGF-β-receptor tyrosin kinase.



(I)

Experimental

The studied 1-phenyl-X-benzimidazoles have been taken with their reactivity from literature[12].Chemical structures and experimental biological activities are gathered in Table 1. Biological activities against the phosphorylation of a model glutamate tyrosine copolymer substrate by isolated mouse PDGF-β receptor tyrosine kinase enzyme are represented as log 1/C. The final geometries were obtained with the semi-empirical RM1 method in the Hyperchem program. The resulted optimized geometries were subjected to further single point calculations using B3LYP method at the 6-31G basis set using Firefly program[12]. Firefly program was employed for calculation of different quantum chemical descriptors including, dipole moment (μ), local Mullikan charges (Q), HOMO and LOMO energies. Softness (s) and electronegetivity (χ) were calculated using Eqs. 1 and 3 respectively. A multiple regression analysis performed on the data had revealed the correlations in Equations 5-11. In these Equations, n is the number of data points, R^2 is the coefficient of multiple determination, s is standard error of estimate, F is the F-ratio between the variances of calculated and observed activities.

Results and discussion

To build successful QSAR models, the studied benzimidazoles are best categorized into three groups. This may be due to the fact that these compounds act by different mechanisms or interact with the receptor in different binding modes.

Group A (10 compounds), that includes compounds 1-7, 17, 31, and 37 and are mainly mono-substituted at position 5 with phenolic and alkoxy groups.

Group B (8 compounds), that includes compounds 12, 13, 16, 22, 23, 39, 40 and 41 and are mainly di-substituted at 4 and 5 or 4 and 6 positions with phenolic or alkoxy groups as well as halogens.

Group C (23 compounds), that includes compounds 4-11, 14, 15, 18-21, 24-30, 32-36, and 38 and are mainly include hetero atomic nitrogen and sulpher substituents.

The biological activity of group A compounds (as represented by log 1/C) are best predicted by the hepta-parametric regression equations 5 and 6.

$\log 1/C = 4320.17(\pm 423220.46)\ HOMO + 43193.92(\pm 423314.64)\ LUMO + 6.48\ (\pm 25.28)\ \Delta E + 86387.55(\pm 846634.13)\ \chi + 41.60(\pm 86.43)\ Q_{N1} - 4.36(\pm 214.82)\ Q_{N3} - 14.08(\pm 139.14)$ …….. (5)

$n = 10,\ R^2 = 0.80,\ s = 0.70,\ F = 2.0$

Eq. 5, which seems with low statistical validation, could be better if compound 7 is excluded (although it is not an outlier) and give Eq. 6 with R^2 value 0.98,

$\log 1/C = 35081.63\ (\pm 284656.65)\ HOMO + 3574.39\ (\pm 284652.75)\ LUMO + 3.91\ (\pm 18.04)\ \Delta E + 70149.10\ (\pm 569308.65)\ \chi - 45.77\ (\pm 58.83)\ Q_{N1} + 4.52\ (\pm 145.67)\ Q_{N3} - 2.27\ (\pm 97.52)$ …….. (6)

$n = 9,\ R2 = 0.98,\ s = 0.27,\ F = 16.0.$

Table 1. Chemical structures of benzimidazoles studied and their observed activities aginst Platlet-Drived Growth Factor Receptor (PDGFR) Tyrosine Kinase.

no.	substituents	obsd. log 1/C
1	H	5.03
2	4-OH	4.85
3	5-OH	6.36
4	5-OMe	6.37
5	5-OC_2H_5	6.62
6	5-OC_3H_7	6.6
7	5-$OCHMe_2$	5.51
8	5-OC_4H_9	5.89
9	5-OCH_2(oxyranyl)	6.5
10	SH	5.48
11	5-SMe	6.13
12	4,5-di-OH	4.6
13	4-OH,5-OMe	5.15
14	4-OMe,5-OH	4.3
15	4,5-di-OMe	4.3
16	4-Br,5-OH	4.3
17	4-$CH_2CH=CH_2$,5-OH	4.3
18	5,6-di-OMe	5.92
19	5,6-di-OH	5.64
20	5-OMe,6-Me	6
21	5-OH,6-Me	5.6
22	5-OMe,6-COOH	4.68
23	5-OH,6-COOH	5.37
24	5-OMe,6-COOMe	6.06
25	5-OMe,6-CH_2OH	6.43
26	5-OMe,6-CHO	6
27	5-$OCH_2CH=CH_2$	6.22
28	5-$O(CH_2)_4OH$	6.35
29	5-$OCH_2CH(OH)CH_2OH$	6.51
30	5-$O(CH_2)_2NH_2$	6.19
31	5-$O(CH_2)_2NMe_2$	5.82
32	5-$O(CH_2)_3NMe_2$	6.82
33	5-$O(CH_2)_4NMe_2$	6.8
34	5-$O(CH_2)_2$N-moropholinyl	6.14
35	5-$O(CH_2)_3$N-morpholinyl	6.77
36	5-$O(CH_2)_4$N-morpholinyl	6.57
37	5-$S(CH_2)_3$N-morpholinyl	4.3
38	5-$OSCNMe_2$	5.34
39	4-Br,5-$OCH_2CH=CH_2$	4.3
40	4,5-$CH_2CH(Me)O$-	4.54
41	5,6-OCH_2O-	5.66

Table 2. The values of DFT-based descriptors and their predicted activities (represented as log 1/C) by Eq. 5.

no.	ε_{HUMO}	ε_{LUMO}	ΔE	χ	Q_{N9}	Q_{N7}	A_{obsd}	A_{pred}	Resd.
1	-5.90546	-0.85526	5.0502	3.38036	-0.688739	-0.507394	5.03	5.03	-0.00
2	-5.74568	-0.87049	4.70406	3.30808	-0.698787	-0.541156	4.85	4.77	0.08
3	-5.50409	-0.86179	4.64229	3.18294	-0.690822	-0.513638	6.36	6.35	0.013
4	-5.46953	-0.83131	4.63821	3.15042	-0.691124	-0.51343	6.37	6.58	-0.21
5	-5.4483	-0.17164	4.63114	2.80997	-0.691902	-0.513734	6.62	6.645	-0.02
6	-5.43551	-0.81008	4.62542	3.12279	-0.692024	-0.513807	6.6	6.46	0.14
17	-5.38925	-0.69416	4.69508	3.0417	-0.637985	-0.52508	4.3	4.52	-0.22
31	-5.2687	-0.58586	4.68284	2.92728	-0.636946	-0.517036	5.82	5.65	0.17
37	-5.44966	-0.77852	4.67114	3.11409	-0.63361	-0.506387	4.3	4.25	0.05

χ: electronegetivity; ΔE: $\varepsilon_{LUMO} - \varepsilon_{HOMO}$; ε_{LUMO}: energy of LUMO (eV); ε_{HOMO}: energy of HOMO (eV); Q_{N9}: Mullikan atomic charge on N atom number 9; A_{obsd}: observed biological activity; A_{pred}: predicted biological activity; Resd.: residual.

The data concerning Eq. 6 are gathered in Table 2, and shown as observed log 1/C versus predicted log 1/C in Fig. 1. The agreement between the observed and the calculated activities is excellent. In this model the positive values of HOMO, LUMO, ΔE, χ, and N3 atomic charge (Q_{N3}) suggest that the activity increases with increase values of these descriptors. On the other hand the negative value of N1 atomic charges (Q_{N1}) suggests the opposite.

Concerning group B compounds a hepta-parametric model was defined with R^2 nearly reached unity,

log 1/C= 1.632 (±10.44) χ + 39121.2 (±50030.2) ΔE – 39124.68 (±500312.60) LUMO + 39130.39 (±500330.82) HOMO – 7.60 (±100.47) Q_{N3} + 52.34 (±270.50) Q_{N1} + 68.12 (±263.57) …….. (7)

n= 8, R^2= 0.998, s= 0.06, F= 81

From Eq. 3 the positive values of χ, ΔE, HOMO, and N1 atomic charge suggests that the activity of the studied compounds is increased with increasing values of these descriptors, while it decreases with increasing values of both LUMO and N3 atomic charge since they have a negative values in this equation. The predicted log 1/C values of the studied compounds as calculated by Eq. 3 are gathered in Table 3, and also shown schematically in Fig. 2. The agreement between the observed and predicted values is excellent.

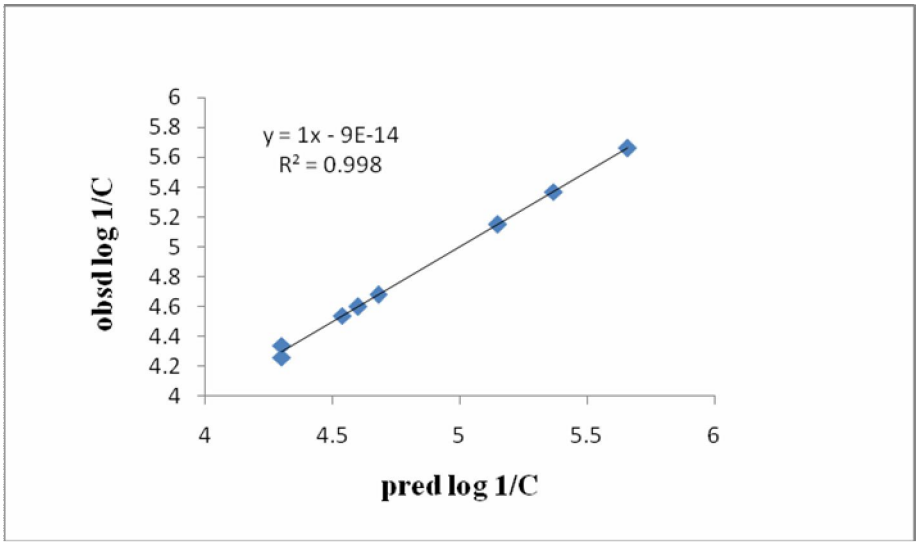


Fig. 1. Observed vs predicted log 1/C of benzimidazoles calculated by Eq. 1

Table 3. The values of DFT-based descriptors and their predicted activities (represented as log 1/C) by Eq. 7.

no.	χ	ΔE	ε_{LUMO}	ε_{HOMO}	Q_{N9}	Q_{N7}	A_{obsd}	A_{pred}	Resd.
12	3.01504	4.33861	-0.84573	-5.18435	-0.701901	-0.545611	4.6	4.60	0.00
13	2.98238	4.33535	-0.81471	-5.15006	-0.703477	-0.54583	5.15	5.15	0.00
16	3.37641	4.60474	-1.07404	-5.67878	-0.695908	-0.498316	4.3	4.26	0.04
22	2.89862	4.09888	-1.54398	-5.64286	-0.698738	-0.513621	4.68	4.68	0.00
23	3.68091	4.04309	-1.65936	-5.70246	-0.693515	-0.513846	5.37	5.37	0.00
39	3.3594	4.60012	-1.05934	-5.65946	-0.697204	-0.498751	4.3	4.34	0.04
40	2.93122	4.54869	-0.65688	-5.20557	-0.637374	-0.52581	4.54	4.54	0.00
41	2.89639	4.448	-0.67239	-5.1204	-0.641693	-0.512616	5.66	5.66	0.00

χ: electronegetivity; ΔE: ε_{LUMO} - ε_{HOMO}; ε_{LUMO}: energy of LUMO (eV); ε_{HOMO}: energy of HOMO (eV); Q_{N9}: Mullikan atomic charge on N atom number 9; A_{obsd}: observed biological activity; A_{pred}: predicted biological activity; Resd.: residual.

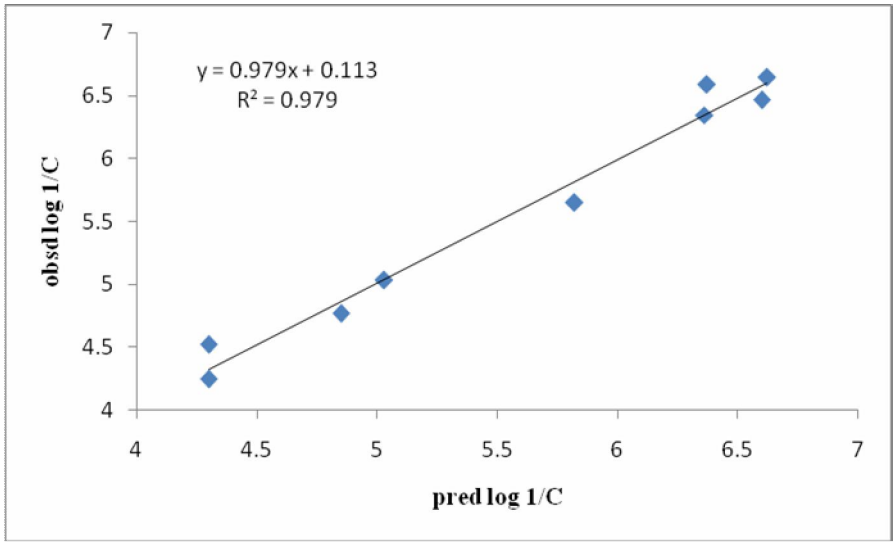


Fig. 2. Observed vs predicted log 1/C of group B benzimidazoles calculated by Eq. 7.

It could be seen from Eqs. 2 and 3 that both HOMO and LUMO have a significant role in the biological activity of the studied benzimidazoles.

For group C, four equations with two, three, five and six descriptors were modeled in which R^2 value increases with increasing numbers of the descriptors. From Eqs. 2 and 3 it is clear that the Mullikan charges of the nitrogen atoms N1 and N3 have some role in the biological activity of the benzimidazoles, accordingly a tri-parametric model including only the atomic charges of N3 and N1 atoms has R^2= 0.86.

log 1/C= 9.70 (±3.71) Q_{N1} – 24.55 (±3.71) Q_{N3} + 1.66 (±1.22) …….. (8)

n= 23, R^2= 0.86, s= 0.27, F= 59.7

The tetra-parametric model in which ΔE is included in addition to N3 and N1 atomic charges has R^2= 0.88,

log 1/C= 10.97 (±1.79) Q_{N1} – 26.86 (±3.27) Q_{N3} - 0.117 (±0.06) ΔE + 0.20 (±0.55) … (9)

n= 23, R^2= 0.88, s= 0.25, F= 46.7

Better R^2 was obtained with the hexa-parametric equation (Eq. 10) which includes the dipole moment (μ) and softness (S) in addition to the descriptors that define Eq. 5,

log 1/C= 12.64 (±7.38) Q_{N1} – 31.20 (±14.05) Q_{N3} – 0.24 (±0.42) S – 0.11 (±0.16) μ – 0.69 ΔE + 2.28 (±3.64) …….. (10)

n= 23, R^2= 0.91, s= 0.23, F= 32.8

Table 4. The values of DFT-based descriptors and their predicted activities (represented as log 1/C) by Eq. 11.

no.	ε_{HOMO}	ε_{LUMO}	χ	ΔE	Q_{N9}	Q_{N7}	A_{obsd}	A_{pred}	Resd.
8	-5.42953	-0.80655	3.11804	4.62297	-0.69118	-0.51359	5.89	5.99	-0.10
9	-5.71797	-0.75321	3.23559	4.96475	-0.63556	-0.51344	6.5	6.61	-0.11
10	-5.50735	-0.97254	3.23994	4.53481	-0.68942	-0.50857	5.48	5.67	-0.19
11	-5.26898	-0.89362	3.0813	4.37535	-0.69062	-0.50887	6.13	5.93	0.20
14	-2.77721	1.62562	0.57579	4.40283	-0.20257	-0.25464	4.3	4.30	0.00
15	-2.77286	1.64358	0.56466	4.41644	-0.20273	-0.25451	4.3	4.29	0.01
18	-5.21782	-0.69471	2.95626	4.52311	-0.69672	-0.51439	5.92	5.67	0.24
19	-5.3585	-0.73117	3.04483	4.62733	-0.69495	-0.51361	5.64	5.66	-0.02
20	-5.36014	-0.76546	3.0628	4.59467	-0.69614	-0.51489	6	5.96	0.04
21	-5.43415	-0.77961	3.10688	4.65454	-0.69339	-0.51469	5.6	5.72	-0.12
24	-6.01349	-1.12084	3.51278	4.89264	-0.63933	-0.51342	6.06	6.06	0.00
25	-5.6328	-0.70069	3.16674	4.9321	-0.63831	-0.51775	6.43	6.69	-0.26
26	-5.72776	-1.73419	3.73097	3.99357	-0.69986	-0.51148	6	5.88	0.12
27	-5.45837	-0.95158	3.20497	4.50678	-0.69162	-0.51186	6.22	6.09	0.13
28	-0.19974	-0.02972	0.11473	0.17002	-0.69157	-0.51370	6.35	6.36	-0.01
29	-5.50218	-0.54831	3.02524	4.95387	-0.63673	-0.51646	6.51	6.37	0.16
30	-5.51497	-0.64192	3.07844	4.87305	-0.63566	-0.51410	6.19	6.36	-0.17
32	-5.39606	-0.60409	3.00052	4.79196	-0.59739	-0.50542	6.82	6.87	-0.05
33	-5.58626	-0.69498	3.14062	4.89128	-0.63587	-0.51405	6.8	6.36	0.44
34	-5.40068	-0.63838	3.01953	4.7623	-0.63637	-0.51408	6.14	6.37	-0.23
35	-5.34245	-0.67321	3.00783	4.66923	-0.63552	-0.51406	6.77	6.66	0.11
36	-5.31578	-0.67457	2.99517	4.64121	-0.63595	-0.51574	6.57	6.45	0.12
38	-5.15768	-1.08193	3.1198	4.07575	-0.69534	-0.50580	5.34	5.62	-0.28

χ: electronegetivity; ΔE: ε_{LUMO} - ε_{HOMO}; ε_{LUMO}: energy of LUMO (eV); ε_{HOMO}: energy of HOMO (eV); Q_{N9}: Mullikan atomic charge on N atom number 9; A_{obsd}: observed biological activity; A_{pred}: predicted biological activity; Resd.: residual.

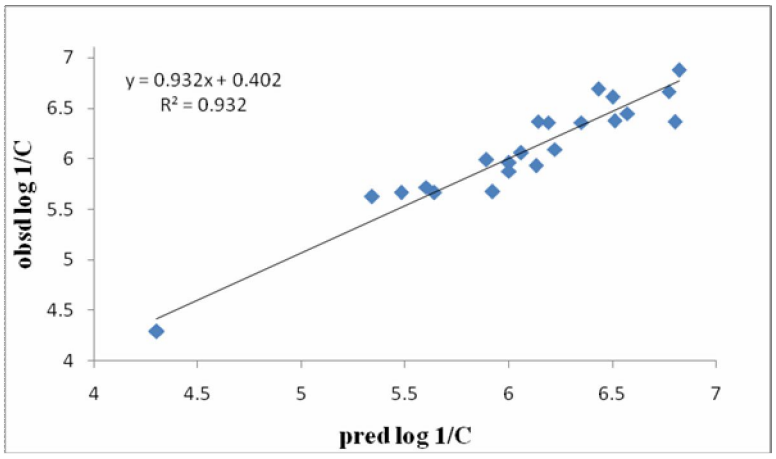


Fig. 3. Observed vs calculated log 1/C of group C benzimidazoles calculated by Eq. 11.

In this model both softness, dipole moment and energy gap have negative signs which suggests that the biological activity decreases with increasing values of these parameters. The atomic charges on N3 and N1 have the significant rule suggesting the importance of the substituents on the studied compounds on the biological activity.

The best model predicted in this study is the hepta-parametric regression equation (Eq. 11) in which the descriptors HOMO, LUMO, and χ as well as N3, N1 charges and ΔE are included,

log 1/C= 26081.62 (±26212.89) LUMO - 26070.0 (±26210.29) HOMO + 10.92 (±12.57) χ – 26076.55

(±26211.63) ΔE + 12.52 (±5.50) Q_{N1} – 27.25 (±7.97) Q_{N3} – 1.04 (±2.35) …….. (11)

n=23, R^2= 0.93, s= 0.20, F= 37.0

In this model LUMO, χ, and Q_{N9} have positive effect on the biological activity, while ΔE, HOMO, and Q_{N7} have negative effect with significant role for LUMO, HOMO, and ΔE. The predicted biological activities (log 1/C) of the studied benzimidazoles as calculated by Eq. 11 are gathered in Table 4, in addition a comparison between observed and predicted values of log 1/C for benzimidazoles used in development of Eq. 11 is shown in Fig. 3.

The general feature in the previously discussed models is that the biological activity increases with increasing Mullikan atomic charge on N1 atom, this implies that the biological activity of such compounds could be enhanced by adding electron donating substituents on the phenyl ring.

The values of R^2 for the QSAR models (Eqs. 5-11) range from 0.86-0.93 which suggests that these models explain 86 -93% of the variance in the data. In addition the smaller the value of s and the larger the value of F, the better the QSAR model. The values of s in the Eq. 5-11 range from 0.06-0.27, while the values of F range from 16.0-81.0 which are statistically significant at the 99% level. The values of R^2, s and F suggest that the QSAR models (Eqs. 5-11) are predictive and validate.

References

1. Singh, J., Shaik, B., Singh, S., Agrawal, V.K., Chadikar, P.V., Deeb, O., and Supuran, C.T., Comparative QSAR Study on Para-substituted Aromatic Sulphonamides as CAII Inhibitors: Information versus Topological (Distance-Based and Connectivity) Indices, Chem Biol Drug Des 2008,71, 244–259.
2. Karelson, M. and Lobanov, V. S., Quantum-Chemical Descriptors in SAR/QSPR Studies, Chem. Rev. 1996, 96, 1027-1043.
3. Puzyn, T., Suzuki, N., Haranczyk, M., and Rak, J., Calculation of Quantum-Mechanical Descriptors for QSPR at the DFT Level: Is It Necessary? J. Chem. Inf. Model. 2008, 48, 1174–1180.
4. Kumar, P., Chattaraj, Chakraborty, A., and Giri, S., Net Electrophilicity, J. Phys. Chem. A 2009, 113, 10068–10074.
5. Afantitis, A., Melagraki, G., Sarimveis, H., Koutentis, P., Markopoulosd, J. and Igglessi-Markopoulou, O., A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes, Bioorganic and Medicinal Chemistry, 2006, 14, 6686–6694.
6. Kiralj, R. and Ferreira, M. M. C., Comparative Chemometric and QSAR/SAR Study of Structurally Unrelated Substrates of a MATE Efflux Pump VmrA from V. parahaemolyticus: Prediction of Multidrug Resistance, QSAR Comb. Sci., 2008, 27, 314 – 329.
7. Chattaraj, P. K., Pe´rez, P., Zevallos, J. and Toro-Labbe, A., Ab Initio SCF and DFT Studies on Solvent Effects on Intramolecular Rearrangement Reactions, J. Phys. Chem. A, 2001, 105, 4272-4283.
8. Jenssen, H., Fjell, C., Cherkasov, A. and Hanckok, R. W., QSAR modeling and computer-aided design of antimicrobial peptides, J. Pept. Sci., 2008, 14, 110–114.
9. De Proft, F., Liu, S. and Parr, R. G., Chemical potential, hardness, hardness and softness kernel and local hardness in the isomorphic ensemble of density functional theory, J. Chem. Phys., 1997, 107, 1887- 1893.
10. Xia, Y., Yin, D., Rong, C., Xu, Q., Yin, D. and Liu, S., Impact of Lewis Acids on Diels-Alder Reaction Reactivity: A Conceptual Density Functional Theory Study, J. Phys. Chem. A, 2008, 112, 9970–9977.
11. Kurup, K., Garg, R. and Hansch, C., Comparative QSAR Study of Tyrosine Kinase Inhibitors, Chem. Rev., 2001, 101, 2573-2600.
12. Alex A. Granovsky, Firefly version 7.1.G, www.http://classic.chem.msu.su/gran/ firefly /index.html.
13. Polanski, J., Gieleciak, A. R. and Magdziarz, T., Modeling robust QSAR, J. Chem. Inf. Model, 2006, 46, 2310-2318.

*****