Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 6823–6829
Bioorganic & Medicinal Chemistry

Comparative QSAR study of phenol derivatives with the help of density functional theory

F. A. Pasha,[a] H. K. Srivastava[b] and P. P. Singh[b,*]

[a]Department of Chemistry, Bareilly College Bareilly, UP, India
[b]Department of Chemistry, M.L.K. (P.G.) College Balrampur, UP, India

Received 17 May 2005; revised 23 July 2005; accepted 25 July 2005
Available online 19 September 2005

Abstract—Quantum chemical reactivity descriptors based QSAR study of 50 phenol derivatives is presented in this paper. Four different methods have been employed to certify the reliability of QSAR study. The molecular weight, hardness, chemical potential, total energy, and electrophilicity index provide valuable information and have a significant role in the assessment of the toxicity of phenols. The first model has been drawn up with the help of AM1 calculations and in this model the correlation coefficient r^2 is 0.88 and the cross-validation coefficient r^2_{cv} is 0.78. Second and third models have been designed with the PM3 and PM5 calculations, respectively. The values of correlation coefficient r^2 and cross-validation coefficient r^2_{cv} in the second case are 0.85 and .070, while in the third case they are 0.85 and 0.71. Finally, the DFT calculations have been made for the same series of compounds by using a B88-PW91 GGA energy functional with the DZVP basis set. The DFT models have a higher predictive power than AM1, PM3, and PM5 methods, and the reliability of this model is clear from its correlation coefficient r^2 0.91 and cross-validation coefficient r^2_{cv} 0.88. This study is also helpful in determining the effect of any particular phenol derivative of this series over *Tetrahymena pyriformis*.

1. Introduction

In our recent communications,[1–4,29] we have employed semiempirical techniques to develop QSAR/QSPR models in which the quantum chemical descriptors have been successfully correlated with observed biological activity. Success of QSAR not only rests on the development of new drug molecules but also in exploring the toxicological and ecotoxicological characteristics of molecules. Quantitative structure toxicity relationships (QSTR) are predictive tools for a preliminary evaluation of the hazards of chemical compounds by using computer-aided models. Density functional theory (DFT)[5,6] was founded within the two basic theorems provided by Hohenberg and Khon in the 1960s.[7] The performance of the DFT method in the description of structural, energetic, and magnetic molecular properties has been quite substantially reviewed in recent times. DFT methods are, in general, capable of generating a variety of isolated molecular properties, such as ionization energies,[8,9] dipole moment,[10,11] electrostatic potential,[12,13] electron affinities,[14,15] electronegativities,[14,15] and hardness,[14,15] etc., quite accurately.

In the present work, we have taken 50 phenol derivatives and to corroborate the reliability of the present work, we have conducted a comparative QSAR (C-QSAR) study with the help of AM1,[26] PM3,[27] PM5,[28] and DFT techniques. A comparison of all the regression models also indicates that the DFT model provides better results than others.

2. Theory

In DFT, the electronegativity, commonly known to a chemist, is defined as the negative of a partial derivative of energy E of an atomic or molecular system with respect to the number of electrons N with a constant external potential $v(r)$[15]

$$\mu = -\chi = -(\partial E/\partial N)_{v(r)}. \quad (1)$$

In accordance with the earlier work of Iczkowski and Margrave,[16] it should be stated that when assuming a

Keywords: DFT; QSAR; Phenol; Semiempirical; Toxicity.
* Corresponding author. Tel.: +91 5263 232888; fax: +91 5263 232708; e-mail addresses: fpasha@rediffmail.com; hemantkrsri@rediffmail.com

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.07.064

quadratic relationship between E and N and in a finite difference approximation, Eq. 1 may be rewritten as

$$\chi = -\mu = (\mathrm{IE} + \mathrm{EA})/2, \quad (2)$$

where IE and EA are the vertical ionization energy and electron affinity, respectively, thereby recovering the electronegativity definition of Mulliken.[17] Moreover, a theoretical justification was provided for Sanderson's principle of electronegativity equalization, which states that when two or more atoms come together to form a molecule, their electronegativities become adjusted to the same intermediate value.[18–20] The absolute hardness η is defined as[21]

$$\eta = 1/2(\delta\mu/\delta N)v(r) = 1/2(\delta^2 E/\delta N^2)v(r), \quad (3)$$

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

$$\eta = 1/2(\mathrm{IP} - \mathrm{EA}), \quad (4)$$

where IP and EA are the ionization potential and electron affinity, respectively, of the chemical species. According to Koopman's theorem, the IP is simply the eigenvalue of HOMO with change of sign and EA is the eigenvalue of LUMO with change of sign;[15] hence Eqs. 3 and 4 may be written as

$$\eta = 1/2(\varepsilon\mathrm{LUMO} - \varepsilon\mathrm{HOMO}), \quad (5)$$

$$\chi = -\mu = 1/2(\varepsilon\mathrm{LUMO} + \varepsilon\mathrm{HOMO}). \quad (6)$$

With regard to QSAR of a chemical system, the total energy also plays an important role. The 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[22]

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

where P is the density matrix and H is a one-electron matrix.

Parr et al. have introduced the electrophilicity index,[23,24] in terms of chemical potential and hardness. The electrophilicity index is a reliable property of a chemical system and can be used as quantum chemical descriptor, the operational definition of electrophilicity index written as

$$\omega = \mu^2/2\eta. \quad (8)$$

Finally, a more general, but important, property of a molecular system, the molecular weight has also been tested as a descriptor.

3. Materials and method

Fifty substituted phenol derivatives have been used as study material and are given in Table 1 along with their observed (A_{Obsd}) toxicity (pC) against *Tetrahymena pyriformis*.[25] For QSAR prediction, the 3D modeling and geometry optimization of all the phenol derivatives have been performed with the help of CAChe pro software.

Table 1. Phenol derivatives and their toxicity against *Tetrahymena pyriformis*[25]

S.No.	Compound	A_{Obsd}
1	Phenol	−0.431
2	2,6-Difluorophenol	0.396
3	2-Fluorophenol	0.248
4	4-Fluorophenol	0.017
5	3-Fluorophenol	0.473
6	4-Methylphenol	−0.192
7	3-Methylphenol	−0.062
8	2-Chlorophenol	0.277
9	2-Bromophenol	0.504
10	4-Chlorophenol	0.545
11	3-Ethylphenol	0.229
12	2-Ethylphenol	0.176
13	4-Bromophenol	0.681
14	2,3-Dimethylphenol	0.122
15	2,4-Dimethylphenol	0.128
16	2,5-Dimethylphenol	0.009
17	3,4-Dimethylphenol	0.122
18	3,5-Dimethylphenol	0.113
19	3-Chloro-4-fluorophenol	0.842
20	2-Chloro-5-methylphenol	0.64
21	4-Iodophenol	0.854
22	3-Iodophenol	1.118
23	2-Isopropylphenol	0.803
24	3-Isopropylphenol	0.609
25	4-Isopropylphenol	0.473
26	2,5-Dichlorophenol	1.128
27	2,3-Dichlorophenol	1.271
28	4-Chloro-2-methylphenol	0.700
29	4-Chloro-3-methylphenol	0.795
30	2,4–Dichlorophenol	1.036
31	3-*tert*-Butylphenol	0.730
32	4-*tert*-Butylphenol	0.913
33	3,5-Dichlorophenol	1.562
34	2-Phenylphenol	1.094
35	2,4-Dibromophenol	1.403
36	2,3,6-Trimethylphenol	0.418
37	3,4,5-Trimethylphenol	0.930
38	2,4,6-Trimethylphenol	1.695
39	4-Chloro-3,5-dimethylphenol	1.203
40	4-Bromo-2,6-dichlorophenol	1.779
41	2,4,5-Trichlorophenol	2.100
42	4-Bromo-6-chloro-2-methylphenol	1.277
43	4-Bromo-2,6-dimethylphenol	1.278
44	2,4,6-Tribromophenol	2.050
45	2-*tert*-Butyl-4-methylphenol	1.297
46	4-Chloro-2-isopropyl-5-methylphenol	1.862
47	6-*tert*-Butyl-2,4-dimethylphenol	1.245
48	2,6-Diphenylphenol	2.113
49	2,4-Dibromo-6-phenylphenol	2.207
50	2,6–Di-*tert*-butyl-4-methylphenol	1.788

The study is based on semiempirical AM1, PM3, PM5 Hamiltonian, and DFT method. The MOPAC calculations have been performed with MOPAC 2000 software associated with CAChe. The DFT calculations have also been made by optimizing the molecular geometry and by using the B88-PW91 GGA energy functional with the DZVP basis sets on CAChe pro. The values of various descriptors, such as molecular weight (M_W), HOMO energy (eV) (ε_{HOMO}), LUMO energy (eV) (ε_{LUMO}), ionization potential (eV) (IP), electron affinity (eV) (EA), and total energy (Hartree) (T_E) have been directly

Table 2. AM1-based descriptors and predicted biological activities of phenol derivatives by Eq. 9

Compound	M_W	η	μ	T_E	ω	A_{Obsd}	A_{Pred}
1	94.113	4.7565	−2.17925	−49.763	11.295	−0.431	−0.201
2	130.094	4.569	−2.445	−81.555	13.657	0.396	1.108
3	112.103	4.642	−2.3145	−65.657	12.433	0.248	0.419
4	112.103	4.576	−2.2585	−65.659	11.671	0.017	0.295
6	108.14	4.654	−2.114	−56.92	10.399	−0.192	0.004
5[a]	112.103	4.6805	−2.31925	−65.663	12.588	0.473	0.456
7	108.14	4.7075	−2.15875	−56.921	10.969	−0.062	0.02
8	128.558	4.6445	−2.30725	−61.518	12.362	0.277	0.418
9[a]	173.009	4.6265	−2.33775	−59.627	12.642	0.504	0.698
10	128.558	4.61	−2.2575	−61.522	11.747	0.545	0.322
11	122.166	4.71	−2.153	−64.066	10.916	0.229	0.259
12	122.166	4.6915	−2.15275	−64.068	10.871	0.176	0.259
13	173.009	4.6045	−2.29225	−59.633	12.097	0.681	0.601
14	122.166	4.6525	−2.13925	−64.075	10.646	0.122	0.254
15	122.166	4.59	−2.0965	−64.077	10.087	0.128	0.263
16	122.166	4.6225	−2.13675	−64.078	10.552	0.009	0.255
17	122.166	4.6185	−2.09275	−64.074	10.114	0.122	0.257
18	122.166	4.681	−2.1475	−64.079	10.794	0.113	0.256
19	146.548	4.493	−2.379	−77.415	12.714	0.842	0.879
20	142.585	4.5845	−2.28175	−68.676	11.934	0.64	0.583
21	220.009	4.6335	−2.30475	−58.548	12.306	0.854	0.909
22	220.009	4.6495	−2.34875	−58.547	12.825	1.118	1.01
23	136.193	4.691	−2.141	−71.217	10.751	0.803	0.496
24	136.193	4.715	−2.1495	−71.219	10.892	0.609	0.5
25	136.193	4.6815	−2.11775	−71.219	10.498	0.473	0.488
26	163.003	4.5415	−2.43325	−73.279	13.444	1.128	1.064
27	163.003	4.584	−2.423	−73.276	13.456	1.271	1.091
28	142.585	4.5525	−2.23625	−68.68	11.383	0.7	0.52
29	142.585	4.5825	−2.22925	−68.676	11.386	0.795	0.525
30	163.003	4.513	−2.379	−73.277	12.771	1.036	0.911
31	150.22	4.7225	−2.14575	−78.37	10.872	0.73	0.741
32	150.22	4.6805	−2.10875	−78.369	10.407	0.913	0.729
33	163.003	4.626	−2.4555	−73.284	13.946	1.562	1.24
34	170.21	4.4065	−2.26275	−85.952	11.281	1.094	1.022
35	251.905	4.5075	−2.42775	−69.496	13.284	1.403	1.476
36	136.193	4.583	−2.1	−71.233	10.106	0.418	0.507
37	136.193	4.5895	−2.07975	−71.233	9.926	0.93	0.513
38[a]	136.193	4.561	−2.066	−71.235	9.734	1.695	0.531
39[a]	156.612	4.562	−2.2075	−75.834	11.115	1.203	0.746
40	241.899	4.466	−2.49	−83.141	13.845	1.779	1.809
41	197.448	4.4165	−2.48575	−85.036	13.645	2.1	1.487
42	221.481	4.506	−2.354	−78.545	12.485	1.277	1.338
43	201.063	4.5405	−2.22775	−73.948	11.267	1.278	0.988
44	330.801	4.441	−2.531	−79.357	14.224	2.05	2.35
45	164.247	4.7	−2.1695	−85.507	11.061	1.297	1.001
46[a]	184.665	4.5195	−2.19975	−90.133	10.935	1.862	1.221
47	178.274	4.567	−2.057	−92.68	9.662	1.245	1.266
48	246.308	4.2625	−2.23675	−122.138	10.663	2.113	2.226
49	328.003	4.2825	−2.39075	−105.683	12.239	2.207	2.417
50	220.354	4.696	−2.143	−114.106	10.783	1.788	1.957

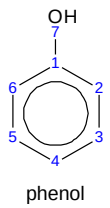
MW = molecular weight, η = hardness, μ = chemical potential, TE = total energy of the system, ω = electrophilicity index, A_{Obsd} = observed toxicity taken from Ref. 25, A_{Pred} = predicted toxicity by Eq. 9.

[a] Data points not included in the deriving equation.

obtained from semiempirical and DFT calculation results. However, the values of hardness (η), softness (S), electronegativity (χ), chemical potential (μ), and electrophilicity index (ω) have been calculated by solving the equations given in theory and the necessary values taken from semiempirical and DFT calculation results. The Project Leader program associated with CAChe pro of Fujitsu, has been used for multiple linear regression (MLR) analysis and various regression equations have been developed for the calculation of toxicity (A_{Pred}).

4. Results

The parent structure of phenol is as follows.



phenol

The assessment of toxicity of a hypothetical compound is of prime interest. The QSAR/QSTR method saves time and cost in determining the toxicity of a series of compounds with the help of toxicity of previously known compounds. The fifty phenol derivatives have been taken in a study out of which 20 are monosubstituted and the remaining 30 are disubstituted and their toxicity values have been taken from the literature[25] and are given in Table 1. A number of quantum chemical descriptors have been identified which are capable of successfully correlating the activity with the structure of a chemical system. In the present study, the following descriptors have been chosen for the QSTR study.

Molecular weight	(M_W)
HOMO energy (eV)	(ε_{HOMO})
LUMO energy (eV)	(ε_{LUMO})
Hardness	(η)
Softness	(S)
Chemical potential	(μ)

Table 3. PM3-based descriptors and predicted biological activities of phenols by Eq. 10

Compound	M_W	η	μ	T_{Energy}	ω	A_{Obsd}	A_{Pred}
1	94.113	4.733	−4.442	−49.819	46.694	−0.431	−0.152
2	130.094	4.6105	−5.0075	−81.643	57.804	0.396	0.879
3	112.103	4.666	−4.73	−65.731	52.196	0.248	0.377
4	112.103	4.608	−4.665	−65.729	50.14	0.017	0.385
5[a]	112.103	4.69	−4.743	−65.733	52.753	0.473	0.370
6	108.14	4.638	−4.315	−57.006	43.178	−0.192	0.055
7	108.14	4.6915	−4.4005	−57.007	45.424	−0.062	0.075
8	128.558	4.5935	−4.6165	−61.585	48.949	0.277	0.387
9[a]	173.009	4.5145	−4.8535	−59.707	53.173	0.504	0.702
10	128.558	4.5285	−4.4805	−61.586	45.455	0.545	0.362
11	122.166	4.694	−4.39	−64.163	45.232	0.229	0.3
12	122.166	4.653	−4.336	−64.158	43.74	0.176	0.29
13	173.009	4.643	−4.67	−59.708	50.629	0.681	0.606
14	122.166	4.6395	−4.3525	−64.189	43.946	0.122	0.3
15	122.166	4.5795	−4.2785	−64.192	41.915	0.128	0.281
16	122.166	4.6135	−4.3495	−64.192	43.639	0.009	0.305
17	122.166	4.604	−4.27	−64.19	41.972	0.122	0.273
18	122.166	4.6685	−4.3735	−64.195	44.648	0.113	0.301
19	146.548	4.49	−4.801	−77.499	51.746	0.842	0.911
20	142.585	4.5355	−4.5575	−68.773	47.103	0.64	0.618
21	220.009	4.216	−4.625	−58.632	45.091	0.854	0.998
22	220.009	4.2735	−4.7735	−58.632	48.689	1.118	1.038
23	136.193	4.671	−4.375	−71.328	44.703	0.803	0.53
24	136.193	4.6985	−4.3985	−71.333	45.45	0.609	0.531
25	136.193	4.674	−4.342	−71.332	44.059	0.473	0.517
26	163.003	4.4475	−4.7625	−73.353	50.438	1.128	0.93
27	163.003	4.5355	−4.7875	−73.354	51.977	1.271	0.899
28	142.585	4.485	−4.448	−68.772	44.367	0.7	0.591
29	142.585	4.507	−4.433	−68.771	44.285	0.795	0.579
30	163.003	4.425	−4.668	−73.351	48.211	1.036	0.908
31	150.22	4.7025	−4.3775	−78.501	45.056	0.73	0.753
32	150.22	4.6725	−4.3165	−78.5	43.529	0.913	0.737
33	163.003	4.576	−4.847	−73.355	53.753	1.562	0.894
34	170.21	4.4695	−4.5915	−86.035	47.113	1.094	1.168
35	251.905	4.4375	−5.0295	−69.596	56.125	1.403	1.442
36	136.193	4.579	−4.28	−71.375	41.94	0.418	0.512
37	136.193	4.58	−4.249	−71.373	41.344	0.93	0.497
38[a]	136.193	4.552	−4.224	−71.378	40.609	1.695	0.490
39[a]	156.612	4.4895	−4.4005	−75.962	43.468	1.203	0.799
40	241.899	4.41	−4.908	−83.24	53.115	1.779	1.652
41	197.448	4.3025	−4.7905	−85.12	49.369	2.1	1.444
42	221.481	4.4535	−4.6965	−78.661	49.116	1.277	1.359
43	201.063	4.561	−4.556	−74.081	47.337	1.278	1.061
44	330.801	4.383	−5.179	−79.491	58.781	2.05	2.163
45	164.247	4.583	−4.219	−85.677	40.789	1.297	0.941
46[a]	184.665	4.4555	−4.3895	−90.283	42.924	1.862	1.261
47	178.274	4.556	−4.213	−92.859	40.433	1.245	1.172
48	246.308	4.1815	−4.4985	−122.252	42.309	2.113	2.408
49	328.003	4.1375	−4.8415	−105.803	48.492	2.207	2.726
50	220.354	4.5495	−4.1475	−114.342	39.13	1.788	1.828

MW = molecular weight, η = hardness, μ = chemical potential, TE = total energy of the system, ω = electrophilicity index, A_{Obsd} = observed toxicity taken from Ref. 25, A_{Pred} = predicted toxicity by Eq. 10.

[a] Data points not included in the deriving equation.

Electrophilicity index (ω)
Total energy (Hartree) (T_E)

The values of these descriptors for all the fifty derivatives have been calculated with the help of AM1, PM3, PM5, and DFT methods. In the formation of first QSTR model based on AM1 Hamiltonian, we have generated various equations by employing all the variables and the best-fitted equation of this class is Eq. 9. The predicted toxicity ($A_{Pred)}$ from Eq. 9 is given in Table 2. The statistical quality of the equation is good as is evident from its correlation coefficient r^2 value = 0.883 and a cross-validation coefficient r^2_{cv} value = 0.787. In this model, the five compounds namely 5, 9, 38, 39, and 46 have been identified as outliers and the descriptor values of these compounds are not included in deriving the regression equation. The predicted toxicity is reliable as is evident from its standard error (SE) value.

Table 4. PM5-based descriptors and predicted biological activities of phenols by Eq. 11

Compound	M_W	η	μ	T_{Energy}	ω	A_{Obsd}	A_{Pred}
1	94.113	4.7135	−4.2395	−49.842	42.359	−0.431	−0.147
2	130.094	4.5385	−5.0695	−81.689	58.319	0.396	1.007
3	112.103	4.6145	−4.6675	−65.761	50.265	0.248	0.424
4	112.103	4.5675	−4.6095	−65.759	48.524	0.017	0.388
5[a]	112.103	4.668	−4.687	−65.765	51.273	0.473	0.447499
6	108.14	4.593	−4.104	−57.017	38.68	−0.192	0.024
7	108.14	4.681	−4.175	−57.018	40.796	−0.062	0.065
8	128.558	4.612	−4.506	−61.596	46.821	0.277	0.394
9[a]	173.009	4.45	−4.697	−59.717	49.088	0.504	0.690383
10	128.558	4.554	−4.432	−61.598	44.726	0.545	0.353
11	122.166	4.681	−4.162	−64.17	40.543	0.229	0.301
12	122.166	4.6525	−4.1725	−64.175	40.499	0.176	0.298
13	173.009	4.5295	−4.5215	−59.72	46.3	0.681	0.65
14	122.166	4.6265	−4.1275	−64.195	39.409	0.122	0.279
15	122.166	4.542	−4.047	−64.194	37.195	0.128	0.238
16	122.166	4.59	−4.115	−64.195	38.862	0.009	0.267
17	122.166	4.5755	−4.0365	−64.194	37.275	0.122	0.241
18	122.166	4.6635	−4.1415	−64.194	39.994	0.113	0.291
19	146.548	4.504	−4.82	−77.521	52.319	0.842	0.923
20	142.585	4.5635	−4.4355	−68.772	44.89	0.64	0.597
21	220.009	3.913	−5.069	−58.612	50.272	0.854	0.947
22	220.009	3.928	−5.181	−58.611	52.719	1.118	0.991
23	136.193	4.656	−4.152	−71.331	40.133	0.803	0.532
24	136.193	4.6855	−4.1565	−71.333	40.475	0.609	0.54
25	136.193	4.63	−4.117	−71.332	39.239	0.473	0.516
26	163.003	4.5255	−4.7595	−73.355	51.258	1.128	0.938
27	163.003	4.563	−4.722	−73.355	50.871	1.271	0.935
28	142.585	4.509	−4.369	−68.775	43.034	0.7	0.561
29	142.585	4.543	−4.352	−68.774	43.022	0.795	0.564
30	163.003	4.4485	−4.6895	−73.354	48.914	1.036	0.888
31	150.22	4.6965	−4.1395	−78.488	40.238	0.73	0.777
32	150.22	4.634	−4.092	−78.489	38.797	0.913	0.749
33	163.003	4.565	−4.852	−73.359	53.734	1.562	0.988
34	170.21	4.458	−4.375	−86.076	42.664	1.094	1.086
35	251.905	4.297	−4.91	−69.597	51.796	1.403	1.467
36	136.193	4.5605	−4.0355	−71.371	37.134	0.418	0.478
37	136.193	4.553	−4.004	−71.367	36.497	0.93	0.468
38[a]	136.193	4.5375	−4.0335	−71.37	36.911	1.695	0.473
39[a]	156.612	4.53	−4.298	−75.95	41.841	1.203	0.783
40	241.899	4.2725	−5.0085	−83.237	53.588	1.779	1.697
41	197.448	4.3835	−4.8505	−85.114	51.566	2.1	1.402
42	221.481	4.3895	−4.6575	−78.658	47.609	1.277	1.372
43	201.063	4.455	−4.385	−74.074	42.831	1.278	1.065
44	330.801	4.1795	−5.1125	−79.48	54.621	2.05	2.248
45	164.247	4.5645	−4.0595	−85.653	37.61	1.297	0.965
46[a]	184.665	4.489	−4.28	−90.264	41.116	1.862	1.247
47	178.274	4.5615	−4.0615	−92.833	37.623	1.245	1.206
48	246.308	4.366	−4.34	−122.311	41.118	2.113	2.303
49	328.003	4.188	−4.815	−105.834	48.548	2.207	2.647
50	220.354	4.555	−4.001	−114.288	36.458	1.788	1.907

MW = molecular weight, η = hardness, μ = chemical potential, TE = total energy of the system, ω = electrophilicity index, A_{Obsd} = observed toxicity taken from Ref. 25, A_{Pred} = predicted toxicity by Eq. 11.
[a] Data points not included in the deriving equation.

$$\begin{aligned} \mathrm{AM1} - A_{\mathrm{Pred}} &= 0.00620256 M_{\mathrm{W}} - 4.61983\eta \\ &\quad + 19.5232\mu - 0.0218789 T_E \\ &\quad + 1.99487\omega + 40.1157 \\ r^2_{\mathrm{cv}} = 0.787087 \quad r^2 &= 0.883616 \quad \mathrm{SE} = .241 \end{aligned} \tag{9}$$

The second QSTR model has been formed with the help of a PM3-based result. In this model, we have generated various equations by employing all the variables and the only best fitted equation is given here. The predicted toxicity (A_{Pred}) from Eq. 10 is reported in Table 3. The statistical quality of the equation is in a better range. The correlation coefficient r^2 value is 0.853, and the value of cross-validation coefficient r^2_{cv} is 0.708. The pre-

Table 5. DFT-based descriptors calculated by the B88-PW91 GGA energy functional with the DZVP basis sets and predicted biological activities of phenols by Eq. 12

Compound	M_{W}	η	μ	T_{Energy}	ω	A_{Obsd}	A_{Pred}
1	94.1130	2.2010	2.911	−307.4590	9.32555	−0.431	−0.185
2	130.094	2.2775	3.0675	−505.9040	10.71513	0.396	0.352
3	112.103	2.2050	3.019	−406.6840	10.04858	0.248	0.115
4	112.103	2.0165	3.0055	−406.6830	9.107553	0.017	0.258
5[a]	112.103	2.2305	3.0245	−406.6840	10.20186	0.473	0.101
6	108.140	2.1210	2.7630	−346.7740	8.096036	−0.192	0.109
7	108.140	2.1685	2.8275	−346.7740	8.668314	−0.062	0.026
8	128.558	2.1545	3.2095	−767.0590	11.09663	0.277	0.487
9[a]	173.009	2.0875	3.4255	−2880.522	12.24741	0.504	0.987
10	128.558	2.0560	3.1440	−767.0560	10.16151	0.545	0.442
11	122.166	2.1720	2.8100	−386.0840	8.575165	0.229	0.210
12	122.166	2.1295	2.7355	−386.0790	7.967482	0.176	0.292
13	173.009	1.9760	3.3770	−2880.519	11.26728	0.681	0.834
14	122.166	2.1470	2.7100	−386.0870	7.883891	0.122	0.267
15	122.166	2.0705	2.6745	−386.0880	7.405092	0.128	0.439
16	122.166	2.1505	2.6755	−386.0880	7.696962	0.009	0.273
17	122.166	2.0895	2.7025	−386.0860	7.630338	0.122	0.383
18	122.166	2.1805	2.6985	−386.0900	7.939094	0.113	0.205
19	146.548	1.9910	3.2910	−866.2770	10.78194	0.842	0.761
20	142.585	2.1570	3.0520	−806.3750	10.04591	0.640	0.531
21	220.009	1.8105	3.4775	−7226.680	10.94719	0.854	0.621
22	220.009	1.8605	3.6315	−7226.680	12.26794	1.118	0.796
23	136.193	2.1230	2.7250	−425.3890	7.882301	0.803	0.499
24	136.193	2.1660	2.8090	−425.3950	8.545391	0.609	0.410
25	136.193	2.1280	2.7380	−425.3950	7.976429	0.473	0.485
26	163.003	2.1080	3.4790	−1226.656	12.75703	1.128	1.245
27	163.003	2.1220	3.4390	−1226.652	12.54815	1.271	1.194
28	142.585	2.0135	3.0535	−806.3700	9.386798	0.700	0.621
29	142.585	2.0390	3.0600	−806.3710	9.546190	0.795	0.607
30	163.003	2.0155	3.4155	−1226.655	11.75605	1.036	1.055
31	150.220	2.1675	2.7995	−464.7050	8.493566	0.730	0.598
32	150.220	2.1345	2.7345	−464.7050	7.980351	0.913	0.664
33	163.003	2.1555	3.5185	−1226.653	13.34237	1.562	1.382
34	170.210	1.6630	3.365	−538.4920	9.415262	1.094	0.973
35	251.905	1.8845	3.6625	−5453.630	12.63925	1.403	1.638
36	136.193	2.1560	2.5850	−425.4040	7.203439	0.418	0.499
37	136.193	2.1160	2.5960	−425.4000	7.130091	0.930	0.586
38[a]	136.193	2.0915	2.5545	−425.4040	6.824011	1.695	0.681
39[a]	156.612	2.0680	2.9200	−845.6870	8.816298	1.203	0.757
40	241.899	1.9190	3.8150	−3799.631	13.96478	1.779	2.089
41	197.448	1.9840	3.6200	−1686.247	12.99956	2.100	1.680
42	221.481	1.9665	3.4505	−3379.371	11.70653	1.277	1.484
43	201.063	2.0235	3.0925	−2959.116	9.675928	1.278	1.067
44	330.801	1.7425	3.9565	−8026.734	13.63845	2.050	2.221
45	164.247	2.0580	2.6020	−504.0130	6.966746	1.297	1.097
46[a]	184.665	2.0725	2.8455	−924.3090	8.390382	1.862	1.141
47	178.274	2.1300	2.5520	−543.3340	6.93603	1.245	1.158
48	246.308	1.7105	3.2315	−769.5510	8.931027	2.113	2.111
49	328.003	1.7250	3.6870	−5684.618	11.72480	2.207	2.340
50	220.354	2.0695	2.5265	−661.2550	6.605019	1.788	1.908

MW = molecular weight, η = hardness, μ = chemical potential, TE = total energy of the system, ω = electrophilicity index, A_{Obsd} = observed toxicity taken from Ref. 25, A_{Pred} = predicted toxicity by Eq. 12. Molecules containing bromine are calculated by using the B88-LYP method.

[a] Data points not included in the deriving equation.

dicted toxicity is reliable as is clear from its standard error (SE) value. The outlier concept is the same as given in Model 1.

$$\begin{aligned} \mathrm{PM3} - A_{\mathrm{Pred}} &= 0.00585765M_{\mathrm{W}} + 0.71477\eta \\ &\quad - 2.40954\mu - 0.0205474T_{\mathrm{E}} \\ &\quad - 0.0998278\omega - 11.1515 \\ r^2_{\mathrm{cv}} = 0.708631 \quad r^2 &= 0.853274 \quad \mathrm{SE} = 0.27. \end{aligned} \tag{10}$$

A third QSAR model has been developed with the help of PM5-based results, and various equations for toxicity prediction have been generated by employing the same variable descriptors. The best-fitted equation is reported here as Eq. 11. The predicted toxicity (A_{Pred}) from Eq. 11 is given in Table 4. The results are reliable as is evident from the correlation coefficient r^2, which is 0.857, and the value of cross-validation coefficient r^2_{cv} is better and is 0.71. The outlier concept is the same as in Model 1.

$$\begin{aligned} \mathrm{PM5} - A_{\mathrm{Pred}} &= 0.00697442M_{\mathrm{W}} - 0.0822995\eta \\ &\quad + 0.36117\mu - 0.0198587T_{\mathrm{E}} \\ &\quad + 0.0348395\omega - 1.34971 \\ r^2_{\mathrm{cv}} = 0.710311 \quad r^2 &= 0.857228 \quad \mathrm{SE} = 0.267. \end{aligned} \tag{11}$$

Finally, the DFT model has been tested against toxicity. The fourth QSAR model has been developed on the basis of descriptor values derived from DFT calculation using the B88-PW91 GGA energy functional with the DZVP basis sets. In this model, we have generated various equations by employing 15 descriptors as variables and the only best fitted equation is reported here. The predicted toxicity (A_{Pred}) from Eq. 12 is given in Table 5. The statistical quality of the equation reveals that a better result could be obtained from the DFT method. The correlation coefficient r^2 value is 0.91, and the value of cross-validation coefficient r^2_{cv} is 0.88. The predicted values of toxicity are closer to observed values and hence are more reliable.

$$\begin{aligned} \mathrm{DFT} - A_{\mathrm{Pred}} &= 0.0140663M_{\mathrm{W}} + 0.00017615\eta \\ &\quad - 6.85443\mu - 8.10552T_{\mathrm{E}} \\ &\quad + 1.33736\omega + 24.7558 \\ r^2_{\mathrm{cv}} = 0.880491 \quad r^2 &= 0.910927 \quad \mathrm{SE} = 0.21 \end{aligned} \tag{12}$$

5. Conclusion

This study results in a framework by which one can calculate the toxicity of any hypothetical compound of the series and their hazards prior to their synthesis. The study is also helpful in the determination of the effect of any particular phenol derivative of this series over *T. pyriformis*. The reliability of this study has been tested by four different methods, AM1, PM3, PM5, and DFT. A comparison of all the methods indicates that the DFT method is more reliable than others and has a high predictive power. The study involves various quantum chemical descriptors [such as molecular weight (M_{W}), hardness (η), chemical potential (μ), total energy (T_{E}), and electrophilicity index (ω)], all descriptors being linearly independent, except for the electrophilicity index (ω), which jointly depends on the hardness (η) and chemical potential (μ).

References and notes

1. Singh, P. P.; Srivastava, H. K.; Pasha, F. A. *Bioorg. Med. Chem.* **2004**, *12*, 171.
2. Singh, P. P.; Pasha, F. A.; Srivastava, H. K. *QSAR & Comb. Sci.* **2003**, *22*, 843.
3. Srivastava, H. K.; Pasha, F. A.; Singh, P. P. *Int. J. Quantum Chem.* **2005**, *103*, 237.
4. Pasha, F. A.; Srivastava, H. K.; Singh, P. P. *Mol. Div.* **2005**, *9*, 215.
5. Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and molecules*; Oxford University Press: New York, 1989.
6. Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.
7. Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864.
8. Chattaraj, P. K.; Cedillo, A.; Parr, R. G. *J. Phys. Chem.* **1991**, *103*, 7645.
9. Ayers, P. W.; Parr, R. G. *J. Am. Chem. Soc.* **2000**, *122*, 2010.
10. De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *250*, 393.
11. Geerlings, P.; De Proft, F.; Martin, J. M. L. In *Theoretical and Computational Chemistry*; Seminario, J., Ed.; Recent Developments in Density Functional Theory; Elsveir: Amsterdam, 1996; Vol. 4, p 773.
12. De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *256*, 400.
13. De Proft, F.; Geerlings, P. *J. Chem. Phys.* **1997**, *106*, 3270.
14. Geerlings, P.; De Proft, F.; Langenaeker, W. *Adv. Quantum Chem.* **1999**, *33*, 303.
15. Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801.
16. Iczkowski, R. P.; Margrave, J. L. *J. Am. Chem. Soc.* **1961**, *83*, 3547.
17. Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782.
18. Sanderson, R. T. *Science* **1955**, *121*, 207.
19. Sanderson, R. T. *Chemical Bonds and Bond Energy*; Academic: New York, 1976.
20. Sanderson, R. T. *Polar Covalence*; Academic Press: New York, 1983.
21. Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.
22. Clare, B. W. *Aust. J. Chem.* **1995**, *48*, 1385.
23. Parr, R. G.; Szentpaly, L. V.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922.
24. Parthasarathi, R.; Subramanian, V.; Roy, D. R.; Chattaraj, P. K. *Bioorg. Med. Chem.* **2004**, *12*, 5533.
25. Hall, L. H.; Vaughn, T. A. *Med. Chem. Res.* **1997**, *7*, 407.
26. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
27. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 209.
28. Stewart J. J. P. MOPAC2002, Fujitsu Limited, Tokyo, Japan, 2002.
29. Pasha, F. A.; Srivastava, H. K.; Singh, P. P. *Int. J. Quantum Chem.* **2005**, *104*, 87.