Hindawi Publishing Corporation
Organic Chemistry International
Volume 2010, Article ID 545087, 17 pages
doi:10.1155/2010/545087

Research Article

A Comparative Study of Two Quantum Chemical Descriptors in Predicting Toxicity of Aliphatic Compounds towards *Tetrahymena pyriformis*

Altaf Hussain Pandith,[1] S. Giri,[2] and P. K. Chattaraj[2]

[1] *Department of Chemistry, University of Kashmir, Srinagar, Kashmir 190006, India*
[2] *Department of Chemistry and Center for Theoretical Studies, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721 302, India*

Correspondence should be addressed to P. K. Chattaraj, pkc@chem.iitkgp.ernet.in

Received 9 September 2010; Revised 16 November 2010; Accepted 15 December 2010

Academic Editor: Pierre Esteves

Copyright © 2010 Altaf Hussain Pandith et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Quantum chemical parameters such as LUMO energy, HOMO energy, ionization energy (I), electron affinity (A), chemical potential (μ), hardness (η) electronegativity (χ), philicity (ω^{α}), and electrophilicity (ω) of a series of aliphatic compounds are calculated at the B3LYP/6-31G(d) level of theory. Quantitative structure-activity relationship (QSAR) models are developed for predicting the toxicity ($pIGC_{50}$) of 13 classes of aliphatic compounds, including 171 electron acceptors and 81 electron donors, towards *Tetrahymena pyriformis*. The multiple linear regression modeling of toxicity of these compounds is performed by using the molecular descriptor $\log P$ (1-octanol/water partition coefficient) in conjunction with two other quantum chemical descriptors, electrophilicity (ω) and energy of the lowest unoccupied molecular orbital (E_{LUMO}). A comparison is made towards the toxicity predicting the ability of electrophilicity (ω) versus E_{LUMO} as a global chemical reactivity descriptor in addition to $\log P$. The former works marginally better in most cases. There is a slight improvement in the quality of regression by changing the unit of IGC_{50} from mg/L to molarity and by removing the racemates and the diastereoisomers from the data set.

1. Introduction

The quantitative structure-activity relationship (QSAR) analysis is aimed at deriving empirical models that relate the activity of chemical compounds to their structure [2]. The underlying assumption is that the chemical structure of a compound implicitly determines its behavior towards biological systems. Appropriate structural or functional descriptors are used to represent the chemical structure and the analysis results in a mathematical model describing the relationship between the chemical structure and the biological activity. Different types of descriptors have been employed, which are of constitutional, geometrical, topological, electrotopological, steric, electrostatic, electronic, and quantum chemical origins. The most essential scientific purpose of developing a QSAR model includes: (1) understanding the mechanism of interaction between compounds and biological systems, (2) gaining information about a dose range for the biological effect of a chemical compound which in turn can be useful in the experimental drug design and toxicity testing, and (3) the prediction of the activity of new chemical compounds. Further, QSAR models can save time and experimental resources for synthesizing and biological testing of a large number of compounds and offer possibility of reduction or replacement of animal use in research and toxicity testing. Various statistical methods are used in QSAR analysis. These methods include regression analysis, partial least squares, classification trees, and neural networks [3].

For the development of a useful QSAR model, the foremost important thing is to assess the mode of biochemical action of the toxicant on the biological system, at cellular and molecular levels. There are many approaches to evaluate the mechanistic basis of toxicity. Some of those methods are: *in vitro* tests [4], joint toxicity tests [5], fish acute toxicity

Table 1: Electrophilicity (ω, eV), energy of lowest unoccupied molecular orbital (E_{LUMO}, au), log P, and observed and calculated values of $pIGC_{50}$ for the complete set of aliphatic acceptor compounds with *Tetrahymena pyriformis*.

Molecule	ω	E_{LUMO}	log P*	$pIGC_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
Diols						
1,2-butanediol	0.4500	0.1504	−0.53	−2.048	−1.954	−1.892
1,3-butanediol	0.4643	0.1491	−1.38	−2.301	−2.483	−2.574
1,4-butanediol	0.4458	0.1573	−0.83	−2.236	−2.328	−2.154
1,2-pentanediol	0.4453	0.1506	0.00	−1.627	−1.537	−1.464
1,5-pentanediol	0.4566	0.1548	−0.64	−1.934	−1.926	−1.994
2-methyl-2,4-pentanediol	0.4600	0.1448	−0.68	−1.953	−1.896	−1.996
1,2-hexanediol	0.4443	0.1505	0.53	−1.267	−1.047	−1.036
1,6-hexanediol	0.4514	0.1564	−0.11	−1.495	−1.522	−1.570
1,2-decanediol	0.4320	0.1507	2.64	0.764	0.735	0.667
1,10-decanediol	0.4298	0.1562	2.01	0.224	0.085	0.141
Halogenated alcohol						
2-bromoethanol	0.4711	0.1301	0.18	−0.846	−0.990	−1.071
2-chloroethanol	0.5211	0.1398	−0.06	−1.417	−1.536	−1.541
1-chloro-2-propanol	0.5087	0.1364	0.14	−1.492	−1.320	−1.282
3-chloro-1-propanol	0.5053	0.1403	0.50	−1.399	−1.084	−1.091
4-chloro-1-butanol	0.4787	0.1453	0.85	−0.759	−0.665	−0.939
3-chloro-2,2-dimethyl-1-propanol	0.4923	0.1406	0.81	−0.782	−0.800	−0.843
6-chloro-1-hexanol	0.4711	0.1447	1.59	−0.272	−0.176	−0.309
8-chloro-1-octanol	0.4641	0.1448	2.65	0.488	0.492	0.566
6-bromo-1-hexanol	0.4320	0.1348	1.73	0.007	0.223	0.083
8-bromo-1-octanol	0.4281	0.1353	2.79	1.042	0.867	0.950
2,3-dibromopropanol	0.4953	0.1209	0.63	−0.486	−0.928	−0.440
Saturated alcohol						
methyl alcohol	0.4744	0.1555	−0.77	−2.665	−2.609	−2.497
ethyl alcohol	0.4595	0.1547	−0.31	−1.991	−2.285	−2.164
1-propanol	0.4491	0.1570	0.25	−1.746	−1.864	−1.690
2-propanol	0.4752	0.1468	0.05	−1.882	−1.939	−2.051
1-butanol	0.4482	0.1571	0.88	−1.431	−1.355	−1.209
2-butanol	0.4615	0.1503	0.61	−1.542	−1.530	−1.554
2-methyl-1-propanol	0.4535	0.1550	0.76	−1.372	−1.435	−1.344
2-pentanol	0.4524	0.1510	1.19	−1.159	−1.088	−1.099
3-pentanol	0.4474	0.1476	1.21	−1.244	−1.089	−1.153
3-methyl-2-butanol	0.4469	0.1529	1.28	−0.996	−1.034	−0.991
tert-amyl alcohol	0.4679	0.1453	0.89	−1.173	−1.281	−1.444
2-methyl-1-butanol	0.4519	0.1550	1.22	−0.953	−1.066	−0.993
3-methyl-1-butanol	0.4611	0.1526	1.16	−1.036	−1.084	−1.087
2,2-dimethyl-1-propanol	0.4710	0.1497	1.31	−0.870	−0.928	−1.035
2-methyl-2-propanol	0.4782	0.1442	0.35	−1.791	−1.685	−1.876
1-hexanol	0.4479	0.1569	2.03	−0.379	−0.420	−0.338
3,3-dimethyl-1-butanol	0.4681	0.1511	1.62	−0.737	−0.686	−0.770
4-methyl-1-pentanol	0.4679	0.1519	1.75	−0.637	−0.581	−0.653
1-heptanol	0.4481	0.1568	2.72	0.105	0.142	0.184
2,4-dimethyl-3-pentanol	0.4261	0.1496	1.93	−0.705	−0.575	−0.564
1-octanol	0.4387	0.1568	3.00	0.583	0.338	0.397
2-octanol	0.4391	0.1511	2.90	0.001	0.258	0.203
3-octanol	0.4282	0.1503	2.72	0.031	0.075	0.050
1-nonanol	0.4282	0.1568	3.77	0.855	0.930	0.982
2-nonanol	0.4331	0.1511	3.25	0.618	0.523	0.470
3-ethyl-2,2-dimethyl-3-pentanol	0.4112	0.1458	2.86	−0.169	0.132	0.065

Table 1: Continued.

Molecule	ω	E_{LUMO}	log P*	pIGC$_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
1-decanol	0.4195	0.1568	4.57	1.335	1.551	1.591
4-decanol	0.4093	0.1528	3.78	0.850	0.874	0.909
3,7-dimethyl-3-octanol	0.4331	0.1440	3.52	0.340	0.743	0.530
1-undecanol	0.4126	0.1568	4.53	1.955	1.495	1.560
1-dodecanol	0.4068	0.1568	5.13	2.161	1.964	2.017
1-tridecanol	0.4019	0.1568	5.58	2.450	2.314	2.359
Carboxylic acid						
Propanoic acid	0.4952	0.1524	0.33	−0.512	−0.593	−0.607
Butyric acid	0.5028	0.1500	0.79	−0.572	−0.474	−0.476
Valeric acid	0.4922	0.1517	1.39	−0.267	−0.303	−0.304
Hexanoic acid	0.4868	0.1519	1.92	−0.208	−0.155	−0.153
Heptanoic acid	0.4793	0.1522	2.41	−0.113	−0.016	−0.013
Octanoic acid	0.4700	0.1522	3.05	0.081	0.165	0.170
Nonanoic acid	0.4594	0.1522	3.47	0.351	0.287	0.290
Decanoic acid	0.4495	0.1523	4.09	0.506	0.464	0.468
Undecanoic acid	0.4408	0.1523	4.53	0.898	0.590	0.594
iso-Butyric acid	0.4814	0.1509	0.60	−0.333	−0.509	−0.530
Isovalerianic acid	0.5038	0.1441	1.16	−0.341	−0.375	−0.369
Trimethylacetic acid	0.4789	0.1482	1.47	−0.254	−0.271	−0.281
3-Methylvaleric acid	0.4831	0.1517	1.75	−0.233	−0.198	−0.201
4-Methylvaleric acid	0.4984	0.1499	1.75	−0.272	−0.210	−0.201
2-Ethylbutyric acid	0.4679	0.1478	1.68	−0.152	−0.205	−0.221
2-Propylpentanoic acid	0.4454	0.1547	2.75	0.026	0.103	0.084
2-Ethylhexanoic acid	0.4563	0.1538	2.64	0.075	0.065	0.053
Succinic acid	0.5258	0.1500	−0.59	−0.939	−0.867	−0.870
Glutaric acid	0.5380	0.1425	−0.29	−0.639	−0.795	−0.784
Adipic acid	0.5175	0.1474	0.08	−0.606	−0.679	−0.678
Pimelic acid	0.5170	0.1457	0.42	−0.584	−0.586	−0.581
3,3-Dimethylglutaric acid	0.5309	0.1375	0.16	−0.664	−0.667	−0.654
Suberic acid	0.4998	0.1484	0.95	−0.512	−0.428	−0.430
Sebacic acid	0.4802	0.1496	2.01	−0.268	−0.125	−0.127
1,10-Decanedicarboxylic acid	0.4592	0.1505	3.07	−0.086	0.179	0.176
Crotonic acid	0.5022	0.1131	0.72	−0.545	−0.493	−0.491
trans-2-Pentenoic acid	0.5129	0.1101	1.41	−0.277	−0.314	−0.293
trans-2-Hexenoic acid	0.4982	0.1124	1.94	−0.128	−0.158	−0.142
Halogenated acid						
4-Bromobutyric acid	0.3372	0.1609	0.68	−0.771	−0.623	−0.524
5-Bromovaleric acid	0.3239	0.1634	1.21	−0.693	−0.560	−0.362
4-Chlorobutyric acid	0.3394	0.1844	0.54	−0.677	−0.644	−0.693
3-Chloropropionic acid	0.3668	0.1792	0.41	−0.332	−0.582	−0.709
5-Chlorovaleric acid	0.3211	0.1875	1.07	−0.286	−0.597	−0.533
2-Bromobutyric acid	0.5256	0.1205	1.42	0.122	0.128	−0.070
2-Bromoisobutyric acid	0.3590	0.1550	0.86	−0.584	−0.517	−0.434
2-Bromoisovaleric acid	0.3782	0.1515	1.48	−0.549	−0.332	−0.211
2-Bromovaleric acid	0.5213	0.1208	1.61	−0.042	0.152	−0.008
2-Bromooctanoic acid	0.5175	0.1213	3.19	0.491	0.456	0.512
2-Bromohexanoic acid	0.5193	0.1210	2.14	0.455	0.252	0.165
Monoester						
Ethyl acetate	0.4712	0.1543	0.73	−1.297	−1.400	−1.413
Propyl acetate	0.4783	0.1520	1.24	−1.238	−1.089	−1.104
Isopropyl acetate	0.4834	0.1502	1.02	−1.590	−1.274	−1.338

Table 1: Continued.

Molecule	ω	E_{LUMO}	log P*	pIGC$_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
Butyl acetate	0.4734	0.1525	1.78	−0.486	−0.683	−0.675
Amyl acetate	0.4706	0.1527	2.30	0.163	−0.304	−0.271
Hexyl acetate	0.4666	0.1530	2.83	−0.009	0.090	0.143
Octyl acetate	0.4559	0.1532	3.88	1.057	0.886	0.951
Decyl acetate	0.4403	0.1533	4.94	1.879	1.721	1.762
Ethyl propionate	0.4723	0.1522	1.21	−0.945	−1.073	−1.119
Butyl propionate	0.4692	0.1518	2.30	0.170	−0.295	−0.304
Isobutyl propionate	0.4862	0.1485	2.17	−0.693	−0.491	−0.521
Propyl propionate	0.4753	0.1510	1.77	−0.815	−0.702	−0.738
tert-Butyl propionate	0.4646	0.1468	1.95	−0.409	−0.510	−0.751
Ethyl butyrate	0.4745	0.1511	1.77	−0.490	−0.697	−0.734
Ethyl isobutyrate	0.4705	0.1483	1.55	−1.271	−0.825	−0.999
Ethyl valerate	0.4675	0.1521	2.30	−0.358	−0.285	−0.294
Propyl butyrate	0.4744	0.1506	2.30	−0.414	−0.327	−0.348
Butyl butyrate	0.4689	0.1513	2.83	0.515	0.076	0.081
Propyl valerate	0.4692	0.1511	2.83	0.009	0.073	0.076
Amyl propionate	0.4661	0.1521	2.83	−0.043	0.093	0.111
Ethyl hexanoate	0.4626	0.1524	2.83	0.064	0.115	0.120
Methyl butyrate	0.4761	0.1525	1.29	−1.246	−1.041	−1.048
Methyl valerate	0.4692	0.1535	1.96	−0.845	−0.532	−0.503
Methyl hexanoate	0.4638	0.1538	2.30	−0.561	−0.261	−0.232
Methyl heptanoate	0.4581	0.1541	2.83	0.104	0.143	0.182
Methyl octanoate	0.4515	0.1541	3.36	0.536	0.552	0.585
Methyl nonanoate	0.4436	0.1541	3.88	1.042	0.963	0.983
Methyl decanoate	0.4357	0.1541	4.41	1.378	1.380	1.387
Methyl undecanoate	0.4283	0.1542	4.79	1.425	1.691	1.679
Methyl formate	0.4808	0.1638	0.03	−1.498	−1.946	−1.606
tert-Butyl formate	0.4799	0.1522	0.97	−1.372	−1.287	−1.303
Diester						
Diethyl malonate	0.3493	0.1881	0.96	−0.997	−0.735	−0.674
Diethyl sebacate	0.2748	0.2011	3.90	1.354	1.274	1.192
Diethyl suberate	0.2854	0.2000	2.84	0.702	0.688	0.564
Diethyl succinate	0.3383	0.1896	1.19	−0.851	−0.534	−0.522
Dimethyl malonate	0.3685	0.1856	−0.05	−1.287	−1.371	−1.291
Dibutyl adipate	0.2920	0.1988	3.90	0.792	1.129	1.163
Dimethyl succinate	0.3544	0.1872	0.35	−1.057	−1.064	−1.038
Diethyl adipate	0.2996	0.1979	1.79	−0.126	0.075	−0.071
Dimethyl brassylate	0.2682	0.1998	4.43	1.654	1.579	1.482
Dimethyl sebacate	0.2853	0.1994	2.84	1.011	0.689	0.556
Dimethyl suberate	0.2977	0.1982	1.79	0.296	0.091	−0.066
Diethyl pimelate	0.2880	0.2002	2.31	0.407	0.417	0.259
Dibutyl suberate	0.2788	0.2009	4.96	1.656	1.737	1.803
Diethyl butylmalonate	0.3399	0.1860	3.02	0.557	0.311	0.492
Diethyl ethyl malonate	0.3460	0.1850	1.96	−0.242	−0.238	−0.134
Diethyl 3-oxopimelate	0.3614	0.1624	1.49	−0.378	−0.589	−0.690
Diethyl 4-oxopimelate	0.3731	0.1614	1.54	−0.638	−0.664	−0.673
Diethyl methylmalonate	0.3504	0.1850	1.44	−0.511	−0.519	−0.435
Diethyl propylmalonate	0.3420	0.1857	2.49	0.134	0.044	0.182
Dibutyl succinate	0.3302	0.1907	3.60	0.512	0.664	0.888
Aldehyde						
Propionaldehyde	0.4454	0.1436	0.59	−0.485	−0.536	−0.607

Table 1: Continued.

Molecule	ω	E_{LUMO}	log P*	pIGC$_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
Butyraldehyde	0.4363	0.1447	0.88	−0.380	−0.412	−0.492
Isobutyraldehyde	0.4649	0.1368	0.61	−0.433	−0.615	−0.499
Valeraldehyde	0.4312	0.1452	1.36	−0.022	−0.250	−0.282
2-Methyl-butyraldehyde	0.4249	0.1454	1.14	−0.311	−0.287	−0.385
Hexylaldehyde	0.4194	0.1474	1.78	−0.173	−0.077	−0.123
2-Methylvaleraldehyde	0.4179	0.1464	1.67	−0.474	−0.103	−0.158
2-Ethylbutyraldehyde	0.4216	0.1439	1.67	−0.054	−0.119	−0.122
3,3-Dimethylbutyraldehyde	0.4559	0.1371	1.63	−0.374	−0.279	−0.042
Heptaldehyde	0.4260	0.1456	2.42	−0.002	0.080	0.192
2-Ethylhexanal	0.4136	0.1455	2.73	0.161	0.224	0.334
trans-4-Decen-1-al	0.3360	0.1341	4.05	1.208	0.944	1.097
cis-7-Decen-1-al	0.2795	0.1452	3.52	0.948	1.036	0.695
Ketones						
Acetone	0.4356	0.1425	−0.24	−2.204	−2.250	−2.201
2-Butanone	0.4274	0.1402	0.29	−1.746	−1.828	−1.903
2-Pentanone	0.4089	0.1440	0.91	−1.222	−1.309	−1.285
3-Pentanone	0.4159	0.1430	0.85	−1.456	−1.374	−1.369
4-Methyl-2-pentanone	0.4159	0.1415	1.31	−1.208	−1.028	−1.093
2-Heptanone	0.3989	0.1453	1.98	−0.487	−0.476	−0.437
5-Methyl-2-hexanone	0.4028	0.1444	1.88	−0.646	−0.563	−0.547
4-Heptanone	0.4056	0.1438	1.91	−0.669	−0.548	−0.553
2-Octanone	0.3976	0.1453	2.37	−0.145	−0.179	−0.146
2-Nonanone	0.3965	0.1454	3.14	0.660	0.402	0.427
2-Decanone	0.3957	0.1454	3.73	0.582	0.847	0.866
3-Decanone	0.3997	0.1429	3.49	0.626	0.656	0.580
2-Undecanone	0.3952	0.1455	4.09	1.535	1.119	1.134
2-Dodecanone	0.3948	0.1455	4.55	1.670	1.466	1.476
7-Tridecanone	0.3907	0.1453	5.08	1.521	1.876	1.860

*Taken from [1].

syndromes [6], and the mechanism evaluated on the basis of structural parameters. The mechanism of toxicity ranges from noncovalent effects to electrophilic one involving covalent binding with biological macromolecules. Among varied modes of toxic action, the narcotic mechanism involves the nonspecific non-covalent reversible interactions of the toxicants with cell membranes [7]. Nonpolar narcotics are neutral nonreactive compounds such as aliphatic alcohols, ketones, ethers, and so forth, whose toxic effect is assumed to be determined mainly by the lipid solubility [8]. Polar narcotics are less inert aromatic chemical species, such as phenols and anilines, which usually posses a hydrogen donor group [9].

A large number of QSAR studies of acute toxicity have been reported in the literature [10]. Many authors [11–15] have reported quantitative relationship between toxicity and hydrophobicity, wherein the hydrophobicities are represented by octanol-water partition coefficient (log P_{oct} values) or octanol-water distribution coefficient (log D_{oct} values) as descriptors. These model relationships are assumed to represent a "baseline effect," whereby no completely soluble and nonvolatile chemical compound can exhibit toxicity less than that predicted by such relationships. Schultz et al. [16] have investigated the toxicity of a large data set of 500 aliphatic chemicals towards the protozoan *Tetrahymena pyriformis* in terms of their IGC$_{50}$ values using octanol-water partition coefficient. Some authors [17] have reported that dimyristoyl phosphatidylcholine-water partition coefficients give better statistical fit than octanol-water partition coefficients in QSAR inhibition of *T. pyriformis* population growth for nonpolar narcotics, polar narcotics, and esters. Roberts and Costello [18] have developed QSAR models for the toxicity prediction of 18 nonpolar and polar narcotics to the fish *Poecilia reticulata* using log P_{oct} (octanol-water partition coefficient) and log P_{MW} (membrane-water partition coefficients). Freidig and Hermens [19] have reported QSAR models for the toxicity prediction in the cases of *Poecilia. reticulata* (14 day LC50) and *Pimephales promelas* (4 day LC50). These authors have developed separate one parameter QSAR models for a group of narcotics and reactive compounds, using log P_{oct} as a descriptor for the narcotics and an electronic descriptor for the reactive compounds.

Table 2: Electrophilicity (ω, eV), energy of lowest unoccupied molecular orbital (E_{LUMO}, au), log P, and observed and calculated values of $pIGC_{50}$ for the complete set of aliphatic donor compounds with *Tetrahymena pyriformis*.

Molecule	E_{LUMO}	log P*	ω	$pIGC_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
Aminoalcohols						
2-(methylamino)ethanol	0.1585	−0.94	0.2807	−1.820	−1.764	−1.558
4-amino-1-butanol	0.1539	−1.06	0.3282	−0.975	−0.978	−1.437
2-(ethylamino)ethanol	0.1573	−0.46	0.2830	−1.649	−1.614	−1.343
2-Propylaminoethanol	0.1584	0.07	0.2775	−1.684	−1.588	−1.189
2-amino-1-pentanol	0.1516	0.07	0.3313	−0.672	−0.669	−0.943
3-amino-2,2-dimethyl-1-propanol	0.1491	−0.79	0.3398	−0.925	−0.719	−1.166
6-amino-1-hexanol	0.1563	−0.01	0.3150	−0.958	−0.966	−1.143
2-amino-1-hexanol	0.1515	0.60	0.3312	−0.585	−0.549	−0.747
2-amino-3-methyl-1-butanol	0.1557	−0.06	0.3154	−0.585	−0.969	−1.139
2-amino-3,3-dimethyl-butanol	0.1534	0.34	0.3216	−0.718	−0.772	−0.910
2-amino-3-methyl-1-pentanol	0.1551	0.47	0.3164	−0.659	−0.831	−0.925
2-amino-4-methyl-pentanol	0.1539	0.47	0.3243	−0.619	−0.696	−0.881
2-(tert.butylamino)ethanol	0.1505	0.41	0.2929	−1.673	−1.246	−0.782
diethanolamine	0.1564	−1.43	0.2941	−1.794	−1.646	−1.662
1,3-diamino-2-hydroxy-propane	0.1543	−2.05	0.3205	−1.427	−1.336	−1.809
N-methyl diethanol amine	0.1535	−1.04	0.2656	−1.834	−2.045	−1.415
3-(methylamino)-1,2-propanediol	0.1494	−1.82	0.2969	−1.534	−1.687	−1.552
triethanolamine	0.1496	−1.00	0.2802	−1.749	−1.785	−1.260
Acetylenic alcohols						
3-butyn-2-ol	0.1511	0.14	0.3720	−0.402	−0.721	−0.701
1-pentyn-3-ol	0.1498	0.67	0.3723	−1.178	−0.427	−0.418
2-pentyn-1-ol	0.1421	0.89	0.3370	−0.572	−0.323	−0.346
2-penten-4-yn-1-ol	0.1306	−0.01	0.3022	−0.555	−0.838	−0.912
1-hexyn-3-ol	0.1517	1.20	0.3634	0.657	−0.138	−0.115
1-heptyn-3-ol	0.1520	1.73	0.3615	−0.265	0.154	0.178
4-heptyn-3-ol	0.1458	1.73	0.3353	−0.034	0.141	0.139
2-octyn-1-ol	0.1448	2.48	0.3249	0.194	0.552	0.545
2-nonyn-1-ol	0.1449	3.01	0.3245	0.649	0.845	0.836
2-decyn-1-ol	0.1449	3.54	0.3242	0.985	1.138	1.127
2-tridecyn-1-ol	0.1450	5.13	0.3238	2.366	2.018	2.001
4-methyl-1-pentyn-3-ol	0.1517	1.07	0.3634	−0.027	−0.210	−0.186
4-methyl-1-heptyn-3-ol	0.1546	2.13	0.3510	0.743	0.371	0.414
2-methyl-3-buten-2-ol	0.1494	0.52	0.3114	−1.389	−1.291	−1.305
4-pentyn-1-ol	0.1450	−0.01	0.3782	−1.420	−1.696	−1.663
Unsaturated alcohols						
2-methyl-3-butyn-2-ol	0.1496	0.28	0.3734	−1.311	−1.476	−1.510
trans-3-hexen-1-ol	0.1567	1.40	0.2385	−0.777	−0.621	−0.707
cis-3-hexen-1-ol	0.1494	1.40	0.2525	−0.809	−0.622	−0.565
5-hexyn-1-ol	0.1507	0.52	0.3513	−1.295	−1.293	−1.330
3-methyl-1-pentyn-3-ol	0.1468	1.07	0.3799	−1.323	−0.879	−0.792
4-hexen-1-ol	0.1559	1.40	0.2391	−0.754	−0.621	−0.691
5-hexen-1-ol	0.1564	1.40	0.2748	−0.841	−0.623	−0.701
4-pentyn-2-ol	0.1446	0.12	0.3639	−1.632	−1.597	−1.547
5-hexyn-3-ol	0.1520	0.65	0.3626	−1.404	−1.195	−1.246
3-heptyn-1-ol	0.1515	1.40	0.3024	−0.323	−0.625	−0.606
4-heptyn-2-ol	0.1473	1.18	0.3028	−0.616	−0.791	−0.710
3-octyn-1-ol	0.1522	1.93	0.2993	0.017	−0.223	−0.175

Table 2: Continued.

Molecule	E_{LUMO}	log P*	ω	pIGC$_{50}$		
				Observed*	Cal.(ω, log P)	Cal.(E_{LUMO}, log P)
3-nonyn-1-ol	0.1527	2.46	0.2972	0.340	0.178	0.260
2-propen-1-ol	0.1506	0.17	0.3317	−1.918	−1.557	−1.622
2-buten-1-ol	0.1519	0.34	0.2736	−1.472	−1.425	−1.505
3-buten-2-ol	0.1431	0.12	0.3149	−1.053	−1.594	−1.519
cis-2-buten-1,4-diol	0.1412	−0.81	0.3241	−2.149	−2.298	−2.262
cis-2-penten-1-ol	0.1490	0.87	0.2944	−1.105	−1.025	−1.003
3-penten-2-ol	0.1464	0.65	0.2870	−1.401	−1.191	−1.138
trans-2-hexen-1-ol	0.1576	1.40	0.2322	−0.472	−0.621	−0.726
1-hexen-3-ol	0.1530	1.18	0.3153	−0.811	−0.792	−0.821
cis-2-hexen-1-ol	0.1510	1.40	0.2691	−0.777	−0.623	−0.598
trans-2-octen-1-ol	0.1577	2.45	0.2311	0.365	0.174	0.154
Amines						
Propylamine	0.1586	0.47	0.3178	−0.707	−0.920	−1.002
Butylamine	0.1586	0.97	0.3168	−0.573	−0.682	−0.763
N-Methylpropylamine	0.1577	0.84	0.2729	−0.809	−0.826	−0.809
Amylamine	0.1570	1.49	0.3110	−0.485	−0.444	−0.483
N-Methylbutylamine	0.1586	1.33	0.2709	−0.678	−0.595	−0.589
Hexylamine	0.1568	2.06	0.3108	−0.220	−0.171	−0.207
Isopropylamine	0.1514	0.26	0.3423	−0.863	−0.975	−0.974
Isobutylamine	0.1501	0.73	0.3353	−0.262	−0.763	−0.725
N,N-Dimethylethylamine	0.1543	0.70	0.2383	−0.908	−0.958	−0.815
sec-Butylamine	0.1534	0.74	0.3314	−0.671	−0.765	−0.779
Isoamylamine	0.1561	1.32	0.3254	−0.577	−0.499	−0.550
1-Methylbutylamine	0.1539	1.23	0.3273	−0.685	−0.538	−0.554
1-Ethylpropylamine	0.1561	1.23	0.3153	−0.813	−0.561	−0.593
2-Methylbutylamine	0.1564	1.32	0.3226	−0.477	−0.504	−0.556
N,N-Diethylmethylamine	0.1567	0.95	0.2445	−0.756	−0.826	−0.737
tert-Butylamine	0.1462	0.40	0.3583	−0.897	−0.878	−0.813
tert-Amylamine	0.1474	1.10	0.3499	−0.698	−0.558	−0.500
1,2-Dimethylpropylamine	0.1536	1.10	0.3185	−0.709	−0.617	−0.611
Propargylamine	0.1533	−0.43	0.3450	−0.826	−1.301	−1.339
N-Methylpropargylamine	0.1514	0.08	0.3179	−0.982	−1.107	−1.061
1-Dimethylamino-2-propyne	0.1517	−0.01	0.2876	−1.145	−1.206	−1.110
1,1-Dimethylpropargylamine	0.1517	0.64	0.3342	−0.910	−0.808	−0.797
2-Methoxyethylamine	0.1574	−0.67	0.3294	−1.790	−1.445	−1.528
3-Methoxypropylamine	0.1541	−1.02	0.3305	−1.772	−1.610	−1.638
3-Ethoxypropylamine	0.1542	−0.49	0.3297	−1.703	−1.358	−1.385

*Taken from [1].

Response-surface approach has been widely used for the development of mechanistically comprehensible QSAR models for toxicity. The basic premise of this approach is that the toxic action depends on the biouptake and bioavailability as well as on the electrophilic reactivity of the toxicant at an active site. Researchers have employed log P_{oct} or log D_{oct} as a descriptor encoding biouptake and availability and energy of the lowest unoccupied molecular orbital (E_{LUMO}) or maximum acceptor superdelocalisability (A_{max}) as descriptor encoding the electrophilic reactivity. This approach has been applied to different species, including the bacterium *Vibrio fischeri* [20], the protozoan *Tetrahymena pyriformis* [21, 22], the yeast *Saccharomyces cerevisiae* [23], the mould *Aspergillus nidulans* [24], the algae *Scenedesmus obliquus* [25] and *chlorella vulgaris* [24], the plant *Cucumis sativus* [26, 27], and mice [24]. The response surface approach has been extended by adding additional indicator variables and other parameters to improve the statistical fit of the models [28, 29].

Table 3: Regression models for different groups of aliphatic compounds for estimating their toxicity towards *Tetrahymena pyriformis*.

Molecules	Regression equation	R^2_{adj}	RSS	F
Aliphatic electron acceptors				
Diols ($N = 10$)	$pIGC_{50} = 0.806 \log P - 1.469$	0.9757	0.2091	363.2252
	$pIGC_{50} = -80.678\omega + 34.753$	0.6052	1.9596	21.5274
	$pIGC_{50} = 38.415E_{LUMO} - 7.229$	0.1011	0.2055	0.1729
	$pIGC_{50} = -2.978E_{LUMO} + 0.807 \log P - 1.016$	0.9759	0.2080	365.3740
	$pIGC_{50} = 20.131\omega + 0.962 \log P - 10.503$	0.9852	0.1289	599.4980
Halogenated alcohols ($N = 11$)	$pIGC_{50} = 0.778 \log P - 1.373$	0.9178	0.4210	101.4897
	$pIGC_{50} = -23.659\omega + 10.792$	0.7370	1.1830	29.0317
	$pIGC_{50} = -0.029E_{LUMO} - 0.533$	0.1111	4.4060	6.1491
	$pIGC_{50} = -27.966E_{LUMO} + 0.830 \log P + 2.417$	0.9550	0.2541	213.5748
	$pIGC_{50} = -8.156\omega + 0.577 \log P + 2.748$	0.9228	0.4234	120.5311
Saturated alcohols ($N = 32$)	$pIGC_{50} = 0.774 \log P - 2.003$	0.9800	0.9471	1521.011
	$pIGC_{50} = -49.100\omega + 21.403$	0.7054	10.1754	75.2526
	$pIGC_{50} = 97.442E_{LUMO} - 15.262$	0.0739	4.4639	3.4738
	$pIGC_{50} = 20.518E_{LUMO} + 0.761 \log P - 5.102$	0.9844	0.7404	1962.6124
	$pIGC_{50} = 3.362\omega + 0.814 \log P - 3.578$	0.9808	0.9074	1590.1020
Carboxylic acids ($N = 28$)	$pIGC_{50} = 0.286 \log P - 0.701$	0.9157	0.2974	294.4822
	$pIGC_{50} = -12.760\omega + 5.996$	0.7401	0.7487	77.8974
	$pIGC_{50} = 7.017E_{LUMO} - 1.256$	0.0150	0.1948	1.4109
	$pIGC_{50} = -0.122E_{LUMO} + 0.286 \log P - 0.683$	0.9157	0.2974	294.5405
	$pIGC_{50} = -0.788\omega + 0.272 \log P - 0.292$	0.9163	0.2955	296.8120
Halogenated acids ($N = 11$)	$pIGC_{50} = 0.462 \log P - 0.874$	0.6192	0.4697	17.2608
	$pIGC_{50} = 4.510\omega - 2.109$	0.7594	0.3538	32.5603
	$pIGC_{50} = -12.712E_{LUMO} + 1.664$	0.5035	0.5154	11.1434
	$pIGC_{50} = -5.188E_{LUMO} + 0.331 \log P + 0.085$	0.6627	0.4408	20.6495
	$pIGC_{50} = 3.201\omega + 0.200 \log P - 1.839$	0.8230	0.2792	47.5032
Monoesters ($N = 31$)	$pIGC_{50} = 0.760 \log P - 2.027$	0.9278	1.6643	386.7003
	$pIGC_{50} = -58.295\omega + 3.887$	0.7478	4.7270	89.9960
	$pIGC_{50} = 25.971E_{LUMO} - 4.194$	0.0282	8.1557	0.1782
	$pIGC_{50} = 35.720E_{LUMO} + 0.762 \log P - 7.480$	0.9398	1.4061	469.1512
	$pIGC_{50} = -6.179\omega + 0.696 \log P + 1.003$	0.9298	1.6220	398.4300
Diesters ($N = 20$)	$pIGC_{50} = 0.634 \log P - 1.332$	0.9045	1.2470	181.9855
	$pIGC_{50} = -21.955\omega + 7.200$	0.7249	2.9341	51.0598
	$pIGC_{50} = 46.892E_{LUMO} - 8.767$	0.3304	3.5315	10.3746
	$pIGC_{50} = 12.567E_{LUMO} + 0.579 \log P - 3.593$	0.9255	0.9982	237.1896
	$pIGC_{50} = -8.477\omega + 0.469 \log P + 1.776$	0.9562	0.6042	415.9038
Aldehydes ($N = 13$)	$pIGC_{50} = 0.463 \log P - 0.886$	0.8379	0.4345	63.0319
	$pIGC_{50} = -9.479\omega + 3.887$	0.7835	0.5463	44.4459
	$pIGC_{50} = -34.557E_{LUMO} + 4.916$	0.0096	0.2368	0.8856
	$pIGC_{50} = -14.568E_{LUMO} + 0.453 \log P + 1.218$	0.8519	0.4030	70.0336
	$pIGC_{50} = -4.345\omega + 0.290 \log P + 1.228$	0.8928	0.3044	100.9174
Ketones ($N = 15$)	$pIGC_{50} = 0.772 \log P - 2.031$	0.9726	0.5454	499.1523
	$pIGC_{50} = -85.007\omega + 34.251$	0.7619	3.7968	45.7999
	$pIGC_{50} = -85.007E_{LUMO} + 34.251$	0.4899	5.4965	14.4442
	$pIGC_{50} = 42.290E_{LUMO} + 0.741 \log P - 8.048$	0.9743	0.5124	533.0067
	$pIGC_{50} = -2.870\omega + 0.751 \log P - 0.819$	0.9729	0.5415	502.9870

Table 3: Continued.

Molecules	Regression equation	R^2_{adj}	RSS	F
Aliphatic electron donors				
Aminoalcohols ($N = 13$)	$pIGC_{50} = 0.353 \log P - 1.053$	0.2985	0.9742	8.2357
	$pIGC_{50} = 19.378\omega - 7.151$	0.7322	0.8187	47.4771
	$pIGC_{50} = -28.920E_{LUMO} + 3.235$	0.0308	0.1256	0.4918
	$pIGC_{50} = -35.985E_{LUMO} + 0.362 \log P + 4.486$	0.3473	1.0289	10.0485
	$pIGC_{50} = 17.095\omega + 0.228 \log P - 6.348$	0.8718	0.4609	116.5597
Acetylenic alcohols ($N = 13$)	$pIGC_{50} = 0.551 \log P - 0.807$	0.7620	1.6655	39.4210
	$pIGC_{50} = -13.919\omega + 4.968$	0.0409	1.0372	1.5116
	$pIGC_{50} = 13.169E_{LUMO} - 1.737$	0.0819	0.0799	0.0915
	$pIGC_{50} = 6.275E_{LUMO} + 0.549 \log P - 1.726$	0.7640	1.6552	39.8563
	$pIGC_{50} = 0.494\omega + 0.554 \log P - 0.982$	0.7621	1.6648	39.4504
Unsaturatedalcohols ($N = 25$)	$pIGC_{50} = 0.759 \log P - 1.686$	0.8618	1.0978	150.7421
	$pIGC_{50} = -7.443\omega + 1.290$	0.2703	2.0101	9.8908
	$pIGC_{50} = 86.169E_{LUMO} - 13.940$	0.3432	2.2295	13.5390
	$pIGC_{50} = -19.393E_{LUMO} + 0.839 \log P + 1.156$	0.8712	1.0343	163.327
	$pIGC_{50} = -0.054\omega + 0.757 \log P - 1.668$	0.8619	1.0978	150.7556
Amines	$pIGC_{50} = 0.461 \log P - 1.138$	0.7164	0.7724	61.6300
	$pIGC_{50} = -0.993\omega - 0.524$	0.0376	0.0217	0.1294
	$pIGC_{50} = 7.473E_{LUMO} - 1.990$	0.0394	0.0150	0.0894
	$pIGC_{50} = -17.899E_{LUMO} + 0.480 \log P + 1.611$	0.7383	0.7334	68.7054
	$pIGC_{50} = 1.860\omega + 0.479 \log P - 1.736$	0.7357	0.7381	67.8094

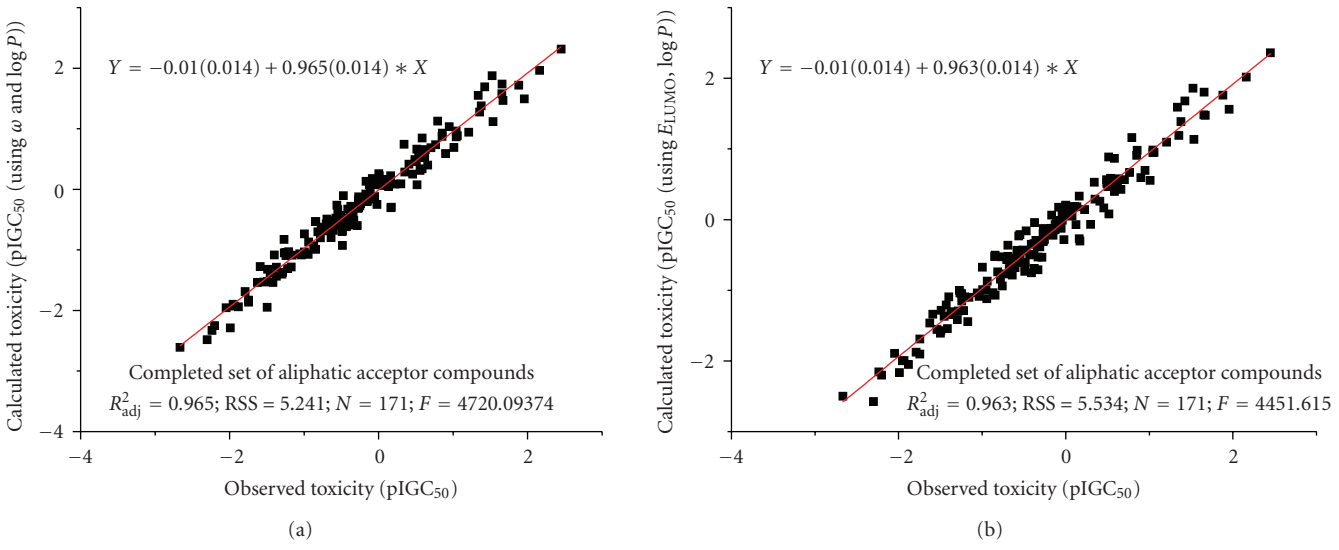

Figure 1: Observed and calculated $pIGC_{50}$ values (a) using electrophilicity index (ω) and log P descriptors and (b) using E_{LUMO} and log P descriptors in a two-parameter regression model, for a complete set of aliphatic acceptors.

Our group has carried out toxicity analysis of a diverse class of systems using conceptual density functional theory-based reactivity/selectivity descriptors like electronegativity, hardness, electrophilicity, and so forth. It has been shown that the toxicity values for a wide variety of polyaromatic hydrocarbons like polychlorinated biphenyls (PCBs), polychlorinated dibenzofurans (PCDFs), polychlorinated dibenzo-p-dioxins (PCDDs) and chlorophenols (CP), as well as arsenic derivatives, and several aliphatic and aromatic toxic molecules, calculated using various conceptual DFT descriptors, especially global and local electrophilicities, correlate well with their corresponding experimental toxicity values [30–39]. In an earlier study, we have reported an atom counting and electrophilicity-based QSTR protocol for predicting the toxicity of aliphatic compounds towards a protozoan, *Tetrahymena pyriformis* [40].

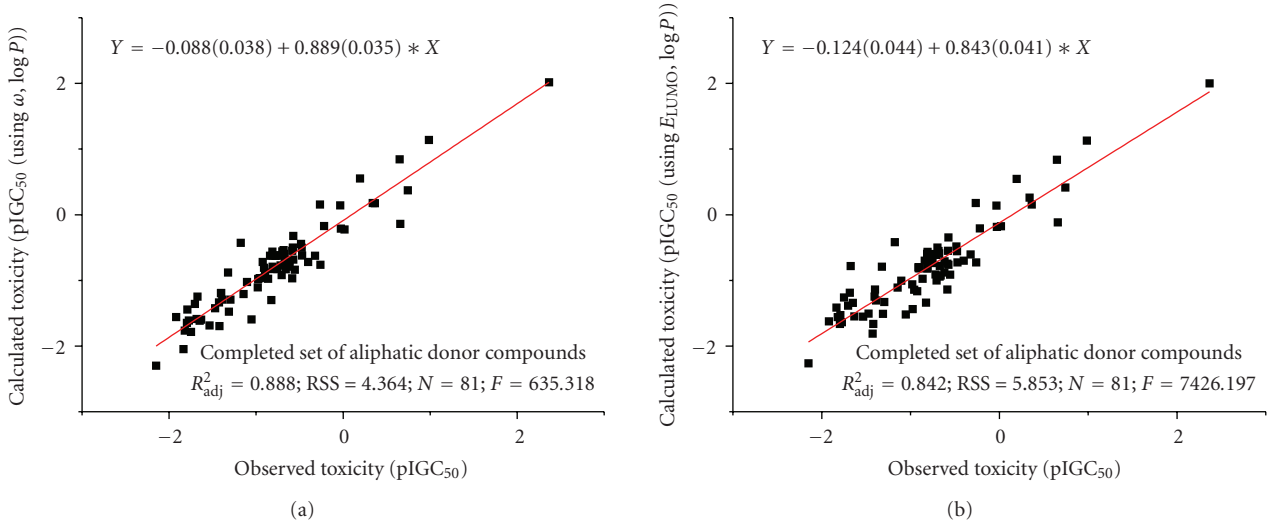

Figure 2: Observed and calculated $pIGC_{50}$ values (a) using electrophilicity index (ω) and log P descriptors and (b) using E_{LUMO} and log P descriptors in a two-parameter regression model, for a complete set of aliphatic donors.

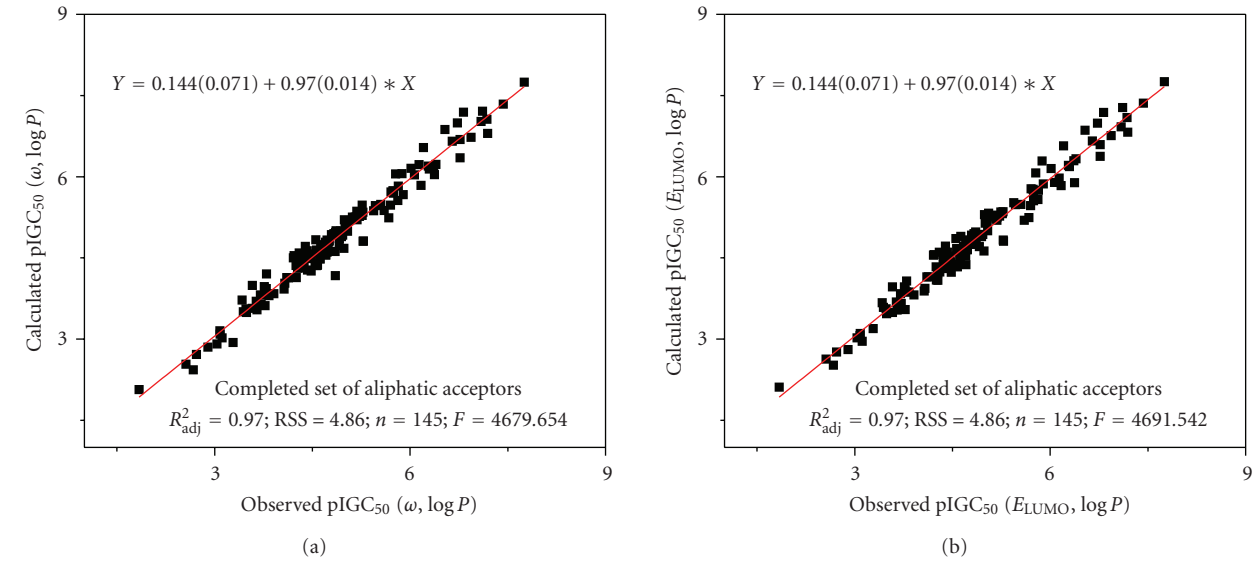

Figure 3: Observed and calculated $pIGC_{50}$ values (a) using electrophilicity index (ω) and log P descriptors and (b) using E_{LUMO} and log P descriptors in a two-parameter regression model, for a complete set of aliphatic acceptors (removing the racemates and the diastereomers from the data set and by changing the unit of IGC_{50} from mg/L to molarity).

In the present work, we develop QSAR models for toxicity of several classes of aliphatic compounds using quantum chemical descriptors, along with the molecular descriptor log P. We attempt to make a comparative evaluation of two quantum chemical parameters namely, electrophilicity index (ω) and energy of the lowest unoccupied molecular orbital (E_{LUMO}), as useful toxicity predicting descriptors towards *Tetrahymena pyriformis*. We intend to check whether the electrophilicity index (ω) is a marginally better toxicity predicting descriptor than LUMO energy when used in addition to log P (a hydrophobicity encoding descriptor).

2. Computational Method

All the geometries are optimized using the GAUSSIAN 03 set of codes [41]. A hybrid density functional theory, using the Becke exchange functional [42] and the correlation functional by Lee et al. [43] and 6–31G(d) basis set are used for the optimization of all the molecules studied in the present work. Frequency analysis is performed on the optimized structures at the same level of theory, and no imaginary frequencies are found. The quantum chemical descriptors such as electron affinity, ionization potential,

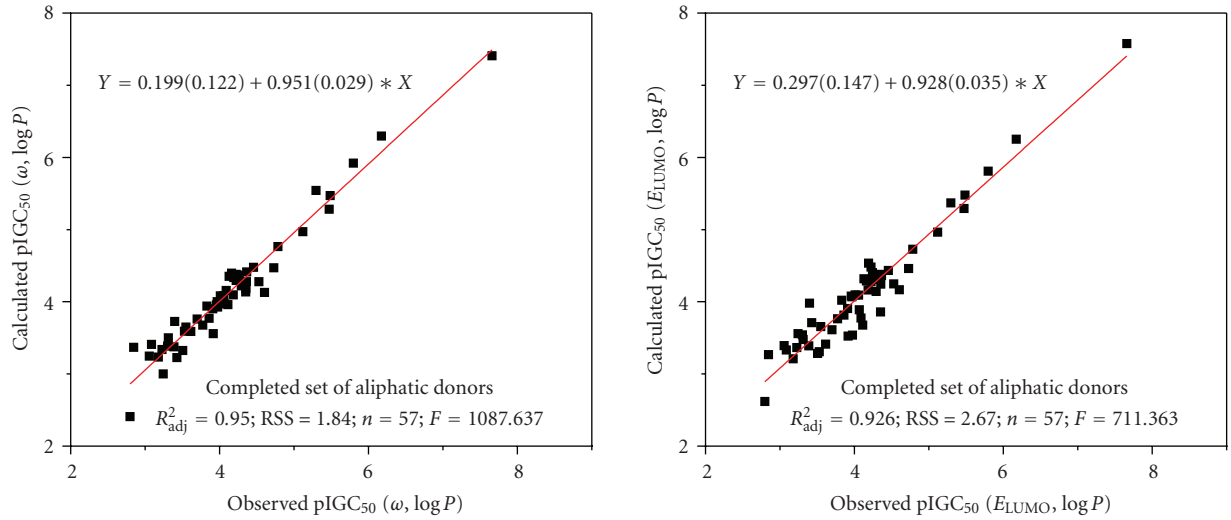

FIGURE 4: Observed and calculated $pIGC_{50}$ values (a) using electrophilicity index (ω) and log P descriptors and (b) using E_{LUMO} and log P descriptors in a two parameter regression model, for a complete set of aliphatic donors (removing the racemates and the diastereomers from the data set and by changing the unit of IGC_{50} from mg/L to molarity).

chemical potential, hardness, and electrophilicity are calculated directly from orbital energies of the optimized geometries.

3. Theoretical Background

3.1. Quantum Chemical Descriptors. Electrophilicity index [44–46] is defined (ω) as a measure of the decrease in energy due to the maximal transfer of electrons from a donor to an acceptor system and is given as

$$\omega = \frac{\mu^2}{2\eta}, \tag{1}$$

where μ and η are the chemical potential [47] and hardness [48], respectively. Chemical potential and hardness can be expressed in terms of ionization energy (I) and electron affinity (A) as given below

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v(r)} \approx -\frac{I+A}{2}, \qquad \eta = \left(\frac{\partial \mu}{\partial N}\right)_{v(r)} \approx I - A. \tag{2}$$

Using Koopmans' approximation, I and A can be expressed in terms of the energies of the highest occupied (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}) as

$$I \approx -E_{HOMO}, \qquad A \approx -E_{LUMO}. \tag{3}$$

The condensed Fukui functions are defined as

$$f_k^{+} = q_k(N+1) - q_k(N), \quad \text{for nucleophilic attack},$$
$$f_k^{-} = q_k(N) - q_k(N-1), \quad \text{for electrophilic attack},$$
$$f_k^{0} = \frac{[q_k(N+1) - q_k(N-1)]}{2}; \quad \text{for nucleophilic attack}, \tag{4}$$

where q_k is the associated electronic population on atom k in a molecule.

The philicity at any atomic site k is defined as [49]

$$\omega_k^{\alpha} = \omega \cdot f_k^{\alpha}, \tag{5}$$

where (α = +, −, and 0) represent local philic quantities describing nucleophilic, electrophilic, and radical attacks, respectively.

3.2. Regression Analysis. The regression analysis is a statistical method wherein a functional dependence of a dependent variable on a set of other independent variables is determined. In linear regression analysis, this dependence has a linear form, which can be expressed as;

$$\acute{Y} = a_1X_1 + a_2X_2 + \cdots + a_pX_p + b, \tag{6}$$

where $a_1, a_2 \cdots a_p$ are regression coefficients, b is the intercept, $X_1, X_2, \cdots X_p$ are independent variables, and $\acute{Y}$ represents expected values of the dependent variable by the regression model.

The above equation represents a hyperplane in the p-dimensional space, where p is the number of independent

Table 4: The general regression equations obtained by using one-parameter and two-parameter models for all the aliphatic acceptors and donors (removing the racemates and the diastereomers from the data set and by changing the unit of IGC_{50} from mg/L to molarity).

(a)

For aliphatic acceptors:
$pIGC_{50} = -7.336\omega + 8.026$
$R^2_{adj} = 0.172$; RSS = 24.81; $n = 145$; $F = 31.08$
$pIGC_{50} = 14.883E_{LUMO} + 2.558$
$R^2_{adj} = 0.052$; RSS = 9.31; $n = 145$; $F = 8.88$
$pIGC_{50} = 0.687 \log P + 3.543$
$R^2_{adj} = 0.815$; RSS = 25.29; $n = 145$; $F = 638.42$
$pIGC_{50} = -1.935\omega + 0.657 \log P + 4.437$
$R^2_{adj} = 0.827$; RSS = 24.11; $n = 145$; $F = 687.52$
$pIGC_{50} = 5.333E_{LUMO} + 0.676 \log P + 2.740$
$R^2_{adj} = 0.823$; RSS = 24.50; $n = 145$; $F = 670.80$

(b)

For aliphatic donors:
$pIGC_{50} = -1.266\omega + 4.521$
$R^2_{adj} = 0.015$; RSS = 0.11; $n = 57$; $F = 0.16$
$pIGC_{50} = -30.337E_{LUMO} + 8.746$
$R^2_{adj} = 0.019$; RSS = 1.41; $n = 57$; $F = 2.08$
$pIGC_{50} = 0.584 \log P + 3.736$
$R^2_{adj} = 0.780$; RSS = 6.80; $n = 57$; $F = 199.29$
$pIGC_{50} = 1.915\omega + 0.592 \log P + 3.145$
$R^2_{adj} = 0.786$; RSS = 6.66; $n = 57$; $F = 206.96$
$pIGC_{50} = -13.231E_{LUMO} + 0.577 \log P + 5.753$
$R^2_{adj} = 0.787$; RSS = 6.65; $n = 57$; $F = 207.58$

variables in the equation. This regression equation can be used for predicting values of the dependent variable from the values of the independent variable.

For determining the quality of the statistical fit, the Pearson correlation coefficient (r) (for regression with single independent variable) or squared coefficient of determination (R^2) is used, which have the following mathematical forms

$$r = \pm\left(\frac{\text{ESS}}{\text{TSS}}\right)^{0.5}, \qquad R^2 = \frac{\text{ESS}}{\text{TSS}} = 1 - \left(\frac{\text{RSS}}{\text{TSS}}\right), \tag{7}$$

where TSS is the total sum of squares, represented as $\sum (Y - Y_{\text{mean}})^2$ and has $N-1$ degrees of freedom, ESS is the explained sum of squares, represented as $\sum (\acute{Y} - Y_{\text{mean}})^2$ and has p degrees of freedom, and RSS is the residual sum of squares, represented as $\sum (Y - \acute{Y})^2$ and has $N - p - 1$ degrees of freedom. Y is the observed value of the dependent variable, $\acute{Y}$ is the predicted value of the dependent variable by the regression model, Y_{mean} is the mean value of the dependent variable, N is the number of observations, and p is the number of independent variables included in the regression model.

If the R^2 value is greater than 0.5, the explained variance by the model (ESS) is larger than the unexplained variance (RSS). The regression equation is considered efficient when the value of R^2 is nearer to 1. The number of independent variables in the equation and the size of the data sample affect the value of R^2. When a new variable is added to the regression equation, the value of R^2 may increase or remain same, even if the added variable does not contribute to reducing of the unexplained variance in the dependent variable. Therefore, another statistical parameter, adjusted R^2 value, is used, which is given by the equation

$$R^2_{\text{adj}} = 1 - \frac{[\text{RSS}/(N-P-1)]}{[\text{TSS}/(N-1)]} = 1 - \left\{(1-R^2) * \left[\frac{(N-1)}{(N-P-1)}\right]\right\}, \tag{8}$$

where, N is the sample size and p is the number of independent variables. The value of R^2_{adj} decreases if an added variable to the equation does not reduce the unexplained variance.

The uncertainty in the model is represented as the standard error of estimate, represented by **s**

$$\mathbf{s} = (\text{RMS})^{0.5} = \left[\frac{\sum \left(Y - \acute{Y}\right)^2}{(N-P-1)}\right]^{0.5}, \tag{9}$$

where RMS is the residual mean square. The standard error of estimate reflects the dispersion of the observed values of the dependent variables about the regression line. Larger values of **s** mean worse statistical fit of the model and less reliability of the prediction.

The statistical significance of a regression equation can be assessed by the means of the Fisher (F) value

$$F = \frac{\text{EMS}}{\text{RMS}} = \frac{R^2(N-p-1)}{p(1-R_2)}, \tag{10}$$

where EMS is the explained mean square given as ESS/p. A regression equation is considered to be statistically significant if the observed F value is greater than a tabulated value for the chosen level of significance and the corresponding degrees of freedom of F. The degrees of freedom of F are equal to p and $N - p - 1$.

A reliable and transparent regression analysis must follow certain basic assumptions, which can be briefly enumerated as follows:

(1) The response variables are not dependent on one another.

(2) The relationship between the dependent and the independent variable(s) is linear.

(3) The residuals (predicted minus observed values of the dependent variable) must follow the normal distribution.

(4) The variance of the residuals is constant for all values of the independent variables.

(5) The independent variables should not show multicollinearity (high level of intercorrelation) and redundancy.

4. Results and Discussion

The quantum chemical descriptors like LUMO energy, HOMO energy, ionization energy, electron affinity, chemical potential, hardness, philicity, and electrophilicity of a series of aliphatic compounds, are calculated from optimized geometries, using (1)–(5) (see Table S1 and Table S2 in Supplementary Materials available online at doi: 10.1155/2010/545087). The main objective of this work is to assess the two quantum descriptors, electrophilicity index (ω), and LUMO energy, which are commonly employed in toxicology studies to represent molecular electrophilicities. We perform a detailed regression analysis using 13 classes of aliphatic compounds, including 171 electron acceptors and 81 electron donors, to develop some model equations, using electrophilicity index (ω), LUMO energy, and log P, to predict toxicity of such chemical compounds towards *Tetrahymena pyriformis*. The general regression equations obtained by using one-parameter and two-parameter models for all the aliphatic acceptors and donors are as follows.

(a) For aliphatic acceptors:

$$\begin{aligned}
&\mathrm{pIGC}_{50} = -5.800\omega + 2.232,\\
&R^2_{\mathrm{adj}} = 0.125; N = 171; F = 25.21; \mathrm{RSS} = 136.66,\\
&\mathrm{pIGC}_{50} = 9.231E_{\mathrm{LUMO}} - 1.708,\\
&R^2_{\mathrm{adj}} = 0.021; N = 171; F = 4.70; \mathrm{RSS} = 152.80,\\
&\mathrm{pIGC}_{50} = 0.618 \log P - 1.444,\\
&R^2_{\mathrm{adj}} = 0.833; N = 171; F = 846.94; \mathrm{RSS} = 26.12,\\
&\mathrm{pIGC}_{50} = -0.725\omega + 0.607 \log P - 1.108,\\
&R^2_{\mathrm{adj}} = 0.834; N = 171; F = 857.92; \mathrm{RSS} = 21.59,\\
&\mathrm{pIGC}_{50} = 0.940E_{\mathrm{LUMO}} + 0.616 \log P - 1.584,\\
&R^2_{\mathrm{adj}} = 0.833; = 171; F = 848.62; \mathrm{RSS} = 21.75,
\end{aligned} \tag{11}$$

(b) For aliphatic donors:

$$\begin{aligned}
&\mathrm{pIGC}_{50} = 1.415\omega - 1.241,\\
&R^2_{\mathrm{adj}} = 0.006; N = 81, F = 1.47, \mathrm{RSS} = 43.56,\\
&\mathrm{pIGC}_{50} = -15.828E_{\mathrm{LUMO}} + 1.606,\\
&R^2_{\mathrm{adj}} = 0.001; N = 81; F = 0.88; \mathrm{RSS} = 43.88,\\
&\mathrm{pIGC}_{50} = 0.537 \log P - 1.164,\\
&R^2_{\mathrm{adj}} = 0.690; N = 81; F = 179.36; \mathrm{RSS} = 13.57,\\
&\mathrm{pIGC}_{50} = 2.426\omega + 0.541 \log P - 1.929,\\
&R^2_{\mathrm{adj}} = 0.704; N = 81; F = 191.42; \mathrm{RSS} = 9.18,\\
&\mathrm{pIGC}_{50} = 4.418E_{\mathrm{LUMO}} + 0.535 \log P - 0.492,\\
&R^2_{\mathrm{adj}} = 0.691; N = 81; F = 180.08; \mathrm{RSS} = 9.40,
\end{aligned} \tag{12}$$

The toxicity values based on these equations, along with the experimentally observed toxicity values are given in Table S3 and Table S4 (see, Supplementary Materials available online at doi: 10.1155/2010/545087). Though, the two parameter equations employing the log P and either of the electronic descriptors (ω or E_{LUMO}) show slightly better correlation as compared to one-parameter model, the overall toxicity predictability of these equations is poor, as is evident from values of the correlation coefficients R^2_{adj} and the calculated toxicity values. It is particularly evident that these generalized equations cannot be used as model equations for accurately predicting the toxicities of the aliphatic compounds.

In order to obtain better predictability and correlation, a stepwise regression analysis is performed by taking each class of chemical compounds separately. The experimentally observed and the calculated toxicity values (pIGC_{50}), along with various descriptors, are presented in Tables 1 and 2 for a set of electron acceptors and a set of electron donors, respectively. The corresponding one-parameter model regression equations (log P, ω, and E_{LUMO}) and two-parameter model regression equations ((log P, ω) and (log P, E_{LUMO})) are shown in Table 3. As is evident from Table 3, the one parameter regression equation based on E_{LUMO} alone does not show any meaningful correlation between the experiment and the calculated toxicity values. The regression equations based on ω show improved correlation coefficients over the equations based on E_{LUMO} for all the electron acceptors and electron donors, except for unsaturated alcohols. However, the adjusted R^2 value is less than 0.70 for diols, acetylenic alcohols, unsaturated alcohols, and amines. For all the electron donor aliphatic compounds, the R^2_{adj} values are negligible, with the exception of amino alcohols. It is remarkable to note that one-parameter regression equations obtained by using log P as an independent variable shows an overall sufficiently improved correlation, compared to that using the electronic descriptors like the electrophilicity (ω) and E_{LUMO}. This result is expected since the hydrophobicity and lipophilicity of the chemical compounds mainly govern their toxic actions at cellular and molecular levels. However as a whole, the stepwise one-parameter model regression analysis based on electronic parameters or log P shows that neither a global electrophilicity descriptor (E_{LUMO} or ω) nor a hydrophobicity descriptor (log P) alone is enough for modeling the toxicity of these compounds with a sufficiently high predictive power.

To improve the predictability of the regression equations and to assess the relative usefulness of the two-quantum descriptors, a two-parameter regression analysis was performed. The results indicate that there is an overall better correlation between the experimental toxicity values and the

Table 5: Regression models for different groups of aliphatic compounds (removing the racemates and the diastereomers from the data set and by changing the unit of IGC_{50} from mg/L to molarity) for estimating their toxicity towards *Tetrahymena pyriformis*.

Molecules	Regression equation	R^2_{adj}	RSS	F
Aliphatic electron acceptors				
Diols ($N = 04$)	$pIGC50 = 0.934 \log P + 3.611$	0.992	0.024	371.304
	$pIGC50 = -88.105\omega + 42.997$	0.552	0.938	4.689
	$pIGC50 = -74.757E_{LUMO} + 15.386$	0.494	0.018	0.008
	$pIGC50 = -66.537E_{LUMO} + 0.933 \log P + 14.002$	0.997	0.010	914.092
	$pIGC50 = 15.860\omega + 1.057 \log P + 8.715$	0.999	0.000	66628.163
Halogenated alcohols ($N = 10$)	$pIGC50 = 0.778 \log P - 1.373$	0.835	0.925	46.700
	$pIGC50 = -23.659\omega + 10.792$	0.694	1.465	21.457
	$pIGC50 = -0.029E_{LUMO} - 0.533$	0.114	0.074	0.082
	$pIGC50 = -27.966E_{LUMO} + 0.830 \log P + 2.417$	0.964	0.231	239.890
	$pIGC50 = -8.156\omega + 0.577 \log P + 2.748$	0.880	0.706	66.997
Saturated alcohols ($N = 22$)	$pIGC50 = 0.774 \log P - 2.003$	0.987	0.659	1624.074
	$pIGC50 = -49.100\omega + 21.403$	0.689	11.448	47.473
	$pIGC50 = 97.442E_{LUMO} - 15.262$	0.142	8.203	4.476
	$pIGC50 = 20.518E_{LUMO} + 0.761 \log P - 5.102$	0.987	0.654	1638.004
	$pIGC50 = 3.362\omega + 0.814 \log P - 3.578$	0.989	0.581	1849.045
Carboxylic acids ($N = 26$)	$pIGC50 = 0.286 \log P - 0.701$	0.904	0.444	237.175
	$pIGC50 = -12.760\omega + 5.996$	0.678	1.136	53.698
	$pIGC50 = 7.017E_{LUMO} - 1.256$	0.027	0.327	1.689
	$pIGC50 = -0.122E_{LUMO} + 0.286 \log P - 0.683$	0.906	0.436	242.920
	$pIGC50 = -0.788\omega + 0.272 \log P - 0.292$	0.906	0.437	241.897
Halogenated acids ($N = 07$)	$pIGC50 = 0.462 \log P - 0.874$	0.024	0.019	0.984
	$pIGC50 = 4.510\omega - 2.109$	0.002	0.196	0.984
	$pIGC50 = -12.712E_{LUMO} + 1.664$	0.180	0.002	0.084
	$pIGC50 = -5.188E_{LUMO} + 0.331 \log P + 0.085$	0.176	0.031	2.278
	$pIGC50 = 3.201\omega + 0.200 \log P - 1.839$	0.024	0.022	1.147
Monoesters ($N = 31$)	$pIGC50 = 0.760 \log P - 2.027$	0.946	1.561	524.241
	$pIGC50 = -58.295\omega + 3.887$	0.749	5.778	90.464
	$pIGC50 = 25.971E_{LUMO} - 4.194$	0.032	0.061	0.056
	$pIGC50 = 35.720E_{LUMO} + 0.762 \log P - 7.480$	0.951	1.412	586.022
	$pIGC50 = -6.179\omega + 0.696 \log P + 1.003$	0.947	1.539	532.620
Diesters ($N = 20$)	$pIGC50 = 0.634 \log P - 1.332$	0.915	1.342	204.539
	$pIGC50 = -21.955\omega + 7.200$	0.710	3.599	47.447
	$pIGC50 = 46.892E_{LUMO} - 8.767$	0.298	4.023	9.082
	$pIGC50 = 12.567E_{LUMO} + 0.579 \log P - 3.593$	0.931	1.165	241.419
	$pIGC50 = -8.477\omega + 0.469 \log P + 1.776$	0.958	0.693	432.461
Aldehydes ($N = 10$)	$pIGC50 = 0.463 \log P - 0.886$	0.936	0.243	132.773
	$pIGC50 = -9.479\omega + 3.887$	0.830	0.583	44.904
	$pIGC50 = -34.557E_{LUMO} + 4.916$	0.073	0.199	0.386
	$pIGC50 = -14.568E_{LUMO} + 0.453 \log P + 1.218$	0.937	0.241	134.476
	$pIGC50 = -4.345\omega + 0.290 \log P + 1.228$	0.964	0.142	238.762
Ketones ($N = 15$)	$pIGC50 = 0.772 \log P - 2.031$	0.978	0.559	616.816
	$pIGC50 = -85.007\omega + 34.251$	0.778	4.522	50.186
	$pIGC50 = -85.007E_{LUMO} + 34.251$	0.050	6.890	14.731
	$pIGC50 = 42.290E_{LUMO} + 0.741 \log P - 8.048$	0.979	0.517	668.897
	$pIGC50 = -2.870\omega + 0.751 \log P - 0.819$	0.979	0.540	639.735

Table 5: Continued.

Molecules	Regression equation	R^2_{adj}	RSS	F
Aliphatic electron donors				
Aminoalcohols ($N = 11$)	pIGC50 = 0.238 log P + 3.783	0.072	0.266	1.770
	pIGC50 = 15.117ω − 0.956	0.647	0.419	19.322
	pIGC50 = −44.539E_{LUMO} + 10.463	0.001	0.176	1.011
	pIGC50 = −37.010E_{LUMO} + 0.216 *log* P + 9.462	0.147	0.345	2.728
	pIGC50 = 14.862ω + 0.221 log P − 0.718	0.803	0.282	41.814
Acetylenic alcohols ($N = 06$)	pIGC50 = 0.655 log P + 3.963	0.931	0.405	68.693
	pIGC50 = 16.352ω + 0.328	0.223	0.166	0.089
	pIGC50 = 131.836E_{LUMO} − 13.121	0.206	1.806	2.298
	pIGC50 = −77.310E_{LUMO} + 0.839 log P − 14.484	0.995	0.030	1017.064
	pIGC50 = −20.562ω + 0.696 log P + 10.499	0.970	0.183	161.877
Unsaturated alcohols ($N = 18$)	pIGC50 = 0.866 log P + 3.198	0.904	0.854	161.154
	pIGC50 = −8.639ω + 6.620	0.215	2.007	5.657
	pIGC50 = 113.888E_{LUMO} − 13.175	0.357	2.485	10.450
	pIGC50 = −35.166E_{LUMO} + 0.998 log P + 8.393	0.922	0.710	200.943
	pIGC50 = 2.723ω + 0.959 log P + 2.305	0.920	0.720	197.680
Amines ($N = 22$)	pIGC50 = 0.489 log P + 3.736	0.809	0.553	89.802
	pIGC50 = −1.665ω + 4.553	0.033	0.059	0.331
	$pIGC_{50}$ = 7.882E_{LUMO} + 2.814	0.045	0.016	0.089
	$pIGC_{50}$ = −17.392E_{LUMO} + 0.508 log P + 6.408	0.830	0.503	103.700
	pIGC50 = 1.613ω + 0.507 log P + 3.219	0.824	0.519	99.074

calculated values upon the addition of an electrophilicity descriptor (E_{LUMO} or ω) to a model, in conjunction with log P. The plots of observed toxicity values ($pIGC_{50}$) versus that predicted on the basis of individual regression equations for a complete set of aliphatic acceptors and donors are presented in Figures 1 and 2. It may be noted that the calculated values of $pIGC_{50}$ in Figures 1 and 2 are obtained from separate regression equations for each individual class of compounds, as reported in Table 3.

These plots reveal that the two-parameter model based on electrophilicity index (ω) and log P (R^2_{adj} = 0.965 for acceptors, 0.888 for donors) is marginally better than that based on E_{LUMO} and log P (R^2_{adj} = 0.963 for acceptors, 0.842 for donors). However, the values of R^2_{adj} for individual groups of molecules while using (ω) and log P as independent variables are only better for electron acceptor compounds, with the exception of halogenated alcohols, saturated alcohols, monoesters, and ketones, where the values are almost the same. In comparison to this, for electron donors the R^2_{adj} values are slightly better when a set of E_{LUMO} and log P values are used in the regression equation as compared to a set of electrophilicity (ω) and log P, except in case of amino alcohols. The calculated toxicity values ($pIGC_{50}$) along with the experimental values, for all the 13 groups of aliphatic compounds studied are reported in Tables 1 and 2.

These results suggest that electrophilicity index (ω) is a marginally better chemical reactivity descriptor in larger cases as compared to E_{LUMO}. We may recommend the toxicity prediction using either of them along with log P. But, a generalized pattern to that effect needs further validation, probably by considering a wide variety of chemical toxicants. Although it is expected that a mechanistic basis of the toxic action may be envisaged from the descriptors used, one should not take the toxicity predictions based on these model relationships without a bit of caution.

As suggested by the Referee, we change the unit of IGC_{50} from mg/L (as used in [1, 16]) to molarity and remove all the racemates and diastereoisomers (also used in those references) from the data set. Respective regression equations are provided in Tables 4 and 5, and the plots of calculated versus observed $pIGC_{50}$ values are presented in Figures 3 and 4. For the individual groups, the correlation improves in most cases. The overall correlation improves in the cases of both electron donors and acceptors, and the overall conclusion remains the same. It may be suggested that log P and ω should be used to predict the toxicity of various aliphatic electron donors and acceptors towards *Tetrahymena pyriformis*.

5. Conclusions

Toxicity of aliphatic compounds considered in this study cannot be completely explained on the basis of the hydrophobicity and the lipophilicity considerations alone. The model QSAR equations with improved toxicity predictability can be developed by taking the electrophilic property of the molecular system into consideration in addition to the hydrophobicity. The "response surface" model proposed by the earlier authors has used mostly E_{LUMO} as the global parameter for the electrophilic reactivity. The results of this

study clearly show that electrophilicity index (ω) and E_{LUMO} are equally capable of describing the contribution of toxicity of aliphatic compounds due to chemical reactivity. The electrophilicity index seems to be a marginally more efficient descriptor for the toxicity prediction as compared to E_{LUMO}. Better QSAR models are obtained by removing the racemates and the diastereoisomers from the data set and by changing the unit of IGC_{50} from mg/L to molarity, as suggested by the Referee.

Supporting Information Available

Quantum chemical parameters such as LUMO energy, HOMO energy, ionization energy (I), electron affinity (A), chemical potential (μ), hardness (η) electronegativity (χ), philicity (ω^{α}) and electrophilicity (ω) of 171 electron acceptors and 81 electron donors, and the experimental and calculated values of $pIGC_{50}$ for electron acceptors and electron donors calculated on the basis of overall regression equations using (ω, log P) and (E_{LUMO}, log P).

In Table S1 contains the calculated HOMO energies, LUMO energies, ionization energies, electron affinities, electronegativities, hardness and chemical potential of 171 aliphatic electron acceptors. Table S2 describes the same as Table S1 for 81 aliphatic electron donors. Electrophilicity (ω), energy of lowest unoccupied molecular orbitals (E_{LUMO}), log P, observed and calculated values of $pIGC_{50}$ for the complete set of aliphatic acceptor and donor compounds with *Tetrahymena pyriformis* are depicted in Tables S3 and S4.

Acknowledgments

We thank Dr. Rana Ashour for the invitation, the Referee for constructive criticism, and Professor A. Basak for helpful discussion. A. H. Pandith thanks Indian Academy of Sciences, Bangalore, for providing him a Summer Research Fellowship to visit IIT Kharagpur. We would like to thank CSIR, New Delhi, for financial support and D. R. Roy, R. Parthasarathi, B. Maiti, V. Subramanian, M. Elango, J. Padmanabhan, P. Bultink, S. Van Damme, and U. Sarkar for their help and assistance in various ways during this study.

References

[1] T. W. Schultz, "Tetratox: *Tetrahymena pyriformis* population growth impairment endpoint—a surrogate for fish lethality," *Toxicology Methods*, vol. 7, no. 4, pp. 289–309, 1997.

[2] Y. C. Martin, *Quantitative Drug Design*, Marcel Dekker, New York, NY, USA, 1978.

[3] I. Lessigiarska, *Development of structure-activity relationship for pharmacotoxicological endpoints relevant to European Union legislation*, Ph.D. thesis, Liverpool John Moores University, Liverpool, UK, 2006.

[4] C. Urani, M. Doldi, S. Crippa, and M. Camatini, "Human-derived cell lines to study xenobiotic metabolism," *Chemosphere*, vol. 37, no. 14-15, pp. 2785–2795, 1998.

[5] S. J. Broderius, M. D. Kahl, and M. D. Hoglund, "Use of joint toxic response to define the primary mode of toxic action for diverse industrial organic chemicals," *Environmental Toxicology and Chemistry*, vol. 14, no. 9, pp. 1591–1605, 1995.

[6] S. P. Bradbury, T. R. Henry, G. J. Niemi, R. W. Carlson, and V. M. Snarski, "Use of respiratory-cardiovascular responses of rainbow trout (*Salmo gairdneri*) in identifying acute toxicity syndromes in fish: Part 3. Polar narcotics," *Environmental Toxicology and Chemistry*, vol. 8, no. 3, pp. 247–261, 1989.

[7] T. W. Schultz, M. T. D. Cronin, J. D. Walker, and A. O. Aptula, "Quantitative structure-activity relationships (QSARS) in toxicology: a historical perspective," *Journal of Molecular Structure (Theochem)*, vol. 622, no. 1-2, pp. 1–22, 2003.

[8] T. W. Schultz, D. T. Lin, T. S. Wilke, and L. M. Arnold, "Quantitative structure-activity relationships for the *Tetrahymena pyriformis* population growth endpoint: a mechanism of action approach," in *Environmental Research and Toxicology*, W. Karcher and J. Devillers, Eds., pp. 61–82, Kluwer Academic Publishers, Boston, Mass, USA, 1990.

[9] G. D. Veith and S. J. Broderius, "Rules for distinguishing toxicants that cause type I and type II narcosis syndromes," *Environmental Health Perspectives*, vol. 87, pp. 207–211, 1990.

[10] I. Lessigiarska, A. P. Worth, and T. I. Netzeva, "Comparative review of QSARs for acute toxicity," EUR Report 21559EN, Joint Research Centre Ispra, Ispra, Italy, 2005.

[11] S. Kapur, A. Shusterman, R. P. Verma, C. Hansch, and C. D. Selassie, "Toxicology of benzyl alcohols: a QSAR analysis," *Chemosphere*, vol. 41, no. 10, pp. 1643–1649, 2000.

[12] J. G. Bundy, A. W. J. Morriss, D. G. Durham, C. D. Campbell, and G. I. Paton, "Development of QSARs to investigate the bacterial toxicity and biotransformation potential of aromatic heterocylic compounds," *Chemosphere*, vol. 42, no. 8, pp. 885–892, 2001.

[13] S. Ren and P. D. Frymier, "Estimating the toxicities of organic chemicals to bioluminescent bacteria and activated sludge," *Water Research*, vol. 36, no. 17, pp. 4406–4414, 2002.

[14] L. E. Sverdrup, T. Nielsen, and P. H. Krogh, "Soil ecotoxicity of polycyclic aromatic hydrocarbons in relation to soil sorption, lipophilicity, and water solubility," *Environmental Science and Technology*, vol. 36, no. 11, pp. 2429–2435, 2002.

[15] A. D. P. Worgan, J. C. Dearden, R. Edwards, T. I. Netzeva, and M. T. D. Cronin, "Evaluation of a novel short-term algal toxicity assay by the development of QSARs and interspecies relationships for narcotic chemicals," *QSAR and Combinatorial Science*, vol. 22, no. 2, pp. 204–209, 2003.

[16] T. W. Schultz, M. T. D. Cronin, T. I. Netzeva, and A. O. Aptula, "Structure-toxicity relationships for aliphatic chemicals evaluated with *Tetrahymena pyriformis*," *Chemical Research in Toxicology*, vol. 15, no. 12, pp. 1602–1609, 2002.

[17] T. W. Schultz and J. R. Seward, "Dimyristoyl phosphatidylcholine/water partitioning-dependent modeling of narcotic toxicity to *Tetrahymena pyriformis*," *Quantitative Structure-Activity Relationships*, vol. 19, no. 4, pp. 339–344, 2000.

[18] D. W. Roberts and J. F. Costello, "Mechanisms of action for general and polar narcosis: a difference in dimension," *QSAR and Combinatorial Science*, vol. 22, no. 2, pp. 226–233, 2003.

[19] A. P. Freidig and J. L. M. Hermens, "Narcosis and chemical reactivity QSARs for acute fish toxicity," *Quantitative Structure-Activity Relationships*, vol. 19, no. 6, pp. 547–553, 2001.

[20] M. T. Croni, G. S. Bowers, G. D. Sinks, and T. W. Schultz, "Structure-toxicity relationships for aliphatic compounds encompassing a variety of mechanisms of toxic action to Vibrio fischeri," *SAR and QSAR in Environmental Research*, vol. 11, no. 3-4, pp. 301–312, 2000.

[21] M. T. D. Cronin and T. W. Schultz, "Development of quantitative structure-activity relationships for the toxicity of aromatic compounds to *Tetrahymena pyriformis*: comparative assessment of the methodologies," *Chemical Research in Toxicology*, vol. 14, no. 9, pp. 1284–1295, 2001.

[22] M. T. D. Cronin, N. Manga, J. R. Seward, G. D. Sinks, and T. W. Schultz, "Parametrization of electrophilicity for the prediction of the toxicity of aromatic compounds," *Chemical Research in Toxicology*, vol. 14, no. 11, pp. 1498–1505, 2001.

[23] X. Wang, C. Yin, and L. Wang, "Structure-activity relationships and response-surface analysis of nitroaromatics toxicity to the yeast (*Saccharomyces cerevisiae*)," *Chemosphere*, vol. 46, no. 7, pp. 1045–1051, 2002.

[24] M. T. Cronin, J. C. Dearden, J. C. Duffy et al., "The importance of hydrophobicity and electrophilicity descriptors in mechanistically-based QSARs for toxicological endpoints," *SAR and QSAR in Environmental Research*, vol. 13, no. 1, pp. 167–176, 2002.

[25] G. H. Lu, X. Yuan, and Y. H. Zhao, "QSAR study on the toxicity of substituted benzenes to the algae (*Scenedesmus obliquus*)," *Chemosphere*, vol. 44, no. 3, pp. 437–440, 2001.

[26] X. Wang, C. Sun, Y. Wang, and L. Wang, "Quantitative structure-activity relationships for the inhibition toxicity to root elongation of *Cucumis sativus* of selected phenols and interspecies correlation with *Tetrahymena pyriformis*," *Chemosphere*, vol. 46, no. 2, pp. 153–161, 2002.

[27] X. Wang, J. Yu, Y. Wang, and L. Wang, "Mechanism-based quantitative structure-activity relationships for the inhibition of substituted phenols on germination rate of *Cucumis sativus*," *Chemosphere*, vol. 46, no. 2, pp. 241–250, 2002.

[28] M. T. D. Cronin, T. I. Netzeva, J. C. Dearden, R. Edwards, and A. D. P. Worgan, "Assessment and modeling of the toxicity of organic chemicals to Chlorella vulgaris: development of a novel database," *Chemical Research in Toxicology*, vol. 17, no. 4, pp. 545–554, 2004.

[29] T. W. Schultz, T. I. Netzeva, D. W. Roberts, and M. T. D. Cronin, "Structure-toxicity relationships for the effects to *Tetrahymena pyriformis* of aliphatic, carbonyl-containing, α,β-unsaturated chemicals," *Chemical Research in Toxicology*, vol. 18, no. 2, pp. 330–341, 2005.

[30] D. R. Roy, R. Parthasarathi, B. Maiti, V. Subramanian, and P. K. Chattaraj, "Electrophilicity as a possible descriptor for toxicity prediction," *Bioorganic and Medicinal Chemistry*, vol. 13, no. 10, pp. 3405–3412, 2005.

[31] R. Parthasarathi, J. Padmanabhan, V. Subramanian, B. Maiti, and P. K. Chattaraj, "Toxicity analysis of 3′4′4′5-pentachloro biphenyl through chemical reactivity and selectivity profiles," *Current Science*, vol. 86, no. 4, pp. 535–542, 2004.

[32] R. Parthasarathi, M. Elango, J. Padmanabhan et al., "Application of quantum chemical descriptors in computational medicinal chemistry and chemoinformatics," *Indian Journal of Chemistry A*, vol. 45, no. 1, pp. 111–125, 2006.

[33] J. Padmanabhan, R. Parthasarathi, V. Subramanian, and P. K. Chattaraj, "Group philicity and electrophilicity as possible descriptors for modeling ecotoxicity applied to chlorophenols," *Chemical Research in Toxicology*, vol. 19, no. 3, pp. 356–364, 2006.

[34] R. Parthasarathi, J. Padmanabhan, V. Subramanian, B. Maiti, and P. K. Chattaraj, "Chemical reactivity profiles of two selected polychlorinated biphenyls," *Journal of Physical Chemistry A*, vol. 107, no. 48, pp. 10346–10352, 2003.

[35] J. Padmanabhan, R. Parthasarathi, V. Subramanian, and P. K. Chattaraj, "Molecular structure, reactivity, and toxicity of the complete series of chlorinated benzenes," *Journal of Physical Chemistry A*, vol. 109, no. 48, pp. 11043–11049, 2005.

[36] J. Padmanabhan, R. Parthasarathi, V. Subramanian, and P. K. Chattaraj, "Chemical reactivity indices for the complete series of chlorinated benzenes: solvent effect," *Journal of Physical Chemistry A*, vol. 110, no. 8, pp. 2739–2745, 2006.

[37] D. R. Roy, U. Sarkar, P. K. Chattaraj et al., "Analyzing toxicity through electrophilicity," *Molecular Diversity*, vol. 10, no. 2, pp. 119–131, 2006.

[38] D. R. Roy, R. Parthasarathi, V. Subramanian, and P. K. Chattaraj, "An electrophilicity based analysis of toxicity of aromatic compounds towards *Tetrahymena pyriformis*," *QSAR and Combinatorial Science*, vol. 25, no. 2, pp. 114–122, 2006.

[39] D. R. Roy, S. Giri, and P. K. Chattaraj, "Arsenic toxicity: an atom counting and electrophilicity-based protocol," *Molecular Diversity*, vol. 13, no. 4, pp. 551–556, 2009.

[40] P. K. Chattaraj, D. R. Roy, S. Giri et al., "An atom counting and electrophilicity based QSTR approach," *Journal of Chemical Sciences*, vol. 119, no. 5, pp. 475–488, 2007.

[41] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., *Gaussian 03, Revision B.01*, Gaussian, Pittsburgh, Pa, USA, 2003.

[42] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.

[43] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, vol. 37, no. 2, pp. 785–789, 1988.

[44] R. G. Parr, L. V. Szentpály, and S. Liu, "Electrophilicity index," *Journal of the American Chemical Society*, vol. 121, no. 9, pp. 1922–1924, 1999.

[45] P. K. Chattaraj, U. Sarkar, and D. R. Roy, "Electrophilicity index," *Chemical Reviews*, vol. 106, no. 6, pp. 2065–2091, 2006.

[46] P. K. Chattaraj and D. R. Roy, "Update 1 of: electrophilicity index," *Chemical Reviews*, vol. 107, no. 9, pp. PR46–PR74, 2007.

[47] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York, NY, USA, 1989.

[48] R. G. Pearson, *Chemical Hardness: Applications from Molecules to Solids*, Wiley-VCH, Weinheim, Germany, 1997.

[49] P. K. Chattaraj, B. Maiti, and U. Sarkar, "Philicity: a unified treatment of chemical reactivity and selectivity," *Journal of Physical Chemistry A*, vol. 107, no. 25, pp. 4973–4975, 2003.