Bull. Environ. Contam. Toxicol. (2003) 70:832–838

DOI: 10.1007/s00128-003-0057-6

Bulletin of Environmental Contamination and Toxicology

Quantitative Structure–Toxicity Relationships for Substituted Aromatic Compounds to *Vibrio fischeri*

G. H. Lu,[1] X. Yuan,[2] C. Wang[1]

[1] College of Environmental Science and Engineering, Hohai University, Nanjing, 210098, People's Republic of China
[2] Department of Environmental Science, Northeast Normal University, Changchun, 130024, People's Republic of China

Received: 24 June 2002/Accepted: 10 December 2002

Substituted aromatic compounds are widely used industrial chemicals, and consequently have a high potential for environmental pollution. They have been reported to be present in the Songhua River of Jilin Province, China (Lang 1993). The presence of many of these chemicals in natural waters is a serious public health problem, so it is beneficial to study their potential hazard to aquatic organisms.

Information on the extent of organic pollutant toxicity is important for risk assessment of chemicals in the environment and for regulating manufacture and use. Quantitative structure-activity relationships (QSARs) are powerful tools in predictive toxicology. The response-surface multiple regression approach to modeling toxicity considers that, in addition to partitioning of the chemical into the biophase, a description of the intrinsic chemical reactivity must be inherent in modeling efforts. The advantage of this approach is that it is toxicokinetic- and toxicodynamic- based, does not require separation of electronic and steric effects, and allows for the use of quantum chemical parameters (Seward et al. 2001).

In this paper, we have measured 15 min EC_{50} values of 40 substituted aromatic hydrocarbons to *Vibrio fischeri.* QSAR models have been developed with molecular weight (M_W), the energy of the highest occupied molecular orbital (E_{HOMO}), the lowest unoccupied molecular orbital (E_{LUMO}) and the logarithm of 1-octanol/water partition coefficient (log P) and used to predict toxicity.

MATERIALS AND METHODS

The freeze-dried powder of *V. fischeri* was supplied by the Nanjing Institute of Soil Science, the Chinese Academy of Science.

The culture was maintained in liquid medium: barm lixiviating extract, 0.5 g; peptone,

---

Correspondence to: G. H. Lu

0.5 g; glycerol, 0.3 g; NaCl, 3 g; Na_2HPO_4, 0.5 g; KH_2PO_4, 0.1 g; distilled water, 100 ml. The pH of the culture medium was adjusted to 7.0±0.5, then the culture was sterilized for 20 min at 121°C. The freeze-dry powder of *V. fischeri* was dissolved using 0.5 ml of 3% NaCl, then was inoculated rapidly into 50 ml of culture medium at 20 °C.

For each compound, 5 concentrations were tested, with the difference in concentration being 0.2 (log Δn). There were three replicates for each concentration and control. 0.5 ml of diluted *V. fischeri* were inoculated into cuvettes containing 2 ml of compound solution or 3% NaCl solution (control). After 15-min of exposure, the toxicity of chemicals to *V. fischeri* was determined using the Microtox Toxicity Analyzer (DXY-2, made by the Nanjing Institute of Soil Science, The Chinese Academy of Science). The concentration values causing 50% reduction of bioluminescene (15min-EC_{50}) were obtained by the one variable regression analyses of the logarithm of compound concentrations and the reduction rates of bioluminescence (Table 1).

The molecular weight (M_W), the energies of the highest occupied molecular orbital (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}) of 40 compounds were calculated by the quantum chemical method MOPAC6.0— AM1 on energy-minimized structures. This method can automatically optimize the bond length, the bond angle and the twist angle, and yield a lot of structural information. The parameter values of studied compounds are presented in Table 1. The linear regression analyses were performed using the SPSS10.0 statistical package.

RESULTS AND DISCUSSION

The experimental results in Table 1 show that the most toxic compound was pentachlorophenol (log $1/EC_{50}$ was 5.11), while the least toxic compound was 2-methylphenol (log $1/EC_{50}$ was 2.44). The range of molecular weights was from 93.13 for aniline to 329.82 for 2,4,6-tribromoaniline; the range of E_{LUMO} from 0.76 for aniline to −2.10 for 2,4-dinitrotoluene; the range of E_{HOMO} from −8.52 for aniline to −10.50 for 2,4-dinitrotoluene, and the range of log P from 1.03 for aniline to 5.04 for pentachlorophenol.

M_W, E_{HOMO}, E_{LUMO} and log P were selected as the structural parameters to establish QSARs. Using bivariate correlation analyses of the toxicity and parameters, a linear correlation matrix was obtained (see Table 2). Moreover, not only E_{HOMO} and E_{LUMO} but also M_W and log P are found to be relatively collinear. M_W can reflect the size of a molecule and appears to be positively correlated with log P and toxicity in this paper. log P is a hydrophobicity parameter. The higher the log P, the stronger the

Table 1. Toxicity data and parameters of studied chemicals

Compounds	Toxicity ($M{\cdot}L^{-1}$) Exp.	Pre.[a]	Er%[a]	Pre.[b]	Er%[b]	Obs.[c]	E_{LUMO} (eV)	$-E_{HOMO}$ (eV)	M_W	log P[d]
Training set										
Aniline	2.64	2.44	7.58	2.25	14.8		0.76	8.52	93.13	1.03
2-Chloroaniline	2.96	2.88	2.70	2.89	2.36		0.19	8.63	127.50	1.76
3-Chloroaniline	2.85	2.87	0.70	2.96	3.86		0.16	8.73	127.50	1.88
2,3-Dichloroaniline	4.51	—	—	—	—		-0.35	9.92	162.02	2.44
2,5-Dichloroaniline	3.96	3.98	0.51	3.71	6.31		-0.46	10.18	162.02	2.80*
2,4,6-Tribromoaniline	4.76	5.08	6.72	4.25	10.7		-0.49	9.10	329.82	4.03
3-Bromoaniline	3.04	3.51	15.5	3.24	6.58		-0.16	9.08	172.03	2.10
4-Bromoaniline	3.47	3.45	0.58	3.28	5.48		-0.11	8.97	172.03	2.26
2-Nitroaniline	3.71	3.17	14.6	3.58	3.50	3.08	-0.94	9.07	138.13	1.85
3-Nitroaniline	3.21	3.26	1.56	3.42	6.54	3.03	-1.03	9.24	138.13	1.37
4-Nitroaniline	3.97	—	—	—	—		-0.88	8.91	138.13	1.39
Phenol	2.63	2.76	4.94	2.64	0.38	2.79	0.40	9.12	94.11	1.46
2-Methylphenol	2.44	2.84	16.4	2.92	19.7	2.70	0.41	9.00	108.14	2.12
3-Methylphenol	2.78	2.85	2.52	2.88	3.60		0.38	9.02	108.14	1.98
2,4-Dichlorophenol	4.02	3.50	12.9	3.74	6.97	4.04	-0.43	9.23	163.00	2.92*
2,4,6-Trichlorophenol	3.85	3.92	1.82	4.29	11.4	4.41	-0.82	9.39	197.45	3.69
Pentachlorophenol	5.11	4.73	7.44	5.22	2.15	5.05	-1.43	9.64	266.34	5.04
2-Nitrophenol	3.53	3.60	1.98	3.71	5.10	3.67	-1.15	9.87	139.11	1.89
3-Nitrophenol	3.34	3.54	5.99	3.83	14.7	3.51	-1.28	9.76	139.11	2.00
4-Nitrophenol	3.70	3.65	1.35	3.81	2.97		-1.21	9.98	139.11	2.04
2,4-Dinitrophenol	4.09	4.23	3.42	3.94	3.67	4.08	-1.81	10.24	184.11	1.54
3- Nitrotoluene	3.74	3.64	2.67	3.87	3.48	3.57	-1.02	9.99	137.14	2.42
4-Nitrotoluene	3.90	3.72	4.62	3.97	1.79	3.80	-1.26	10.14	137.14	2.34
2,4-Dinitrotoluene	4.64	4.35	6.25	4.32	6.90		-2.10	10.50	182.14	2.04
2,6-Dinitrotoluene	4.27	4.32	1.17	4.25	0.47		-2.00	10.44	182.14	2.02
2-Chloronitrobenzene	3.97	3.92	1.26	3.92	1.26	3.68	-1.23	10.15	157.56	2.26
4-Chloronitrobenzene	3.94	4.00	1.52	4.06	3.03	3.56	-1.42	10.29	157.56	2.35
Nitrobenzene	3.26	3.66	12.3	3.75	15.0	3.35	-1.22	10.30	123.11	1.89
3,4-Dichloronitrobenzene	4.11	4.33	5.35	4.52	9.98		-1.52	10.27	192.00	3.29*
2,5-Dichloronitrobenzene	4.22	4.26	0.95	4.25	0.71		-1.34	10.15	192.00	2.90
2-Nitroanisole	3.50	3.89	11.1	3.63	3.71		-1.08	10.17	153.14	1.80
3-Nitroanisole	4.12	3.91	4.76	3.87	6.07	3.67	-1.22	10.21	153.14	2.16
4-Nitroanisole	4.23	3.99	5.67	3.77	10.9	3.54	-1.15	10.37	153.14	2.03
3-Bromonitrobenzene	4.41	4.39	0.45	4.15	5.90	4.22	-1.35	10.21	202.01	2.64
4-Bromonitrobenzene	4.68	4.47	4.49	4.17	10.9		-1.46	10.35	202.01	2.55
Test set										
2,4-Dinitroaniline	4.16	3.77	9.38	3.98	4.33		-1.66	9.37	183.12	1.84
4-Methylphenol	3.16	2.77	12.3	2.83	10.4		0.43	8.88	108.14	1.93
2-Nitrotoluene	3.91	3.65	6.65	3.81	2.56	3.48	-1.01	10.01	137.14	2.30
3-Chloronitrobenzene	3.81	3.92	2.89	4.09	7.35	3.84	-1.36	10.15	157.56	2.49
2,4-Dichloroaniline	4.06	3.97	2.22	3.71	8.62		-0.17	10.15	162.02	2.80*

a. Predicted values and Er% are calculated from Eq. (5); b. Predicted values and Er% are calculated from Eq. (6); c. Observed log $1/IGC_{50}$ to *T. pyriformis* (Schultz et al. 1990); d.The log P values identified with an asterisk are calculated using a fragment constant method (Lyman 1982), the rest are measured values obtained from literature (Deneer et al. 1987, Veith and Mekenyan 1993).

hydrophobicity and the easier the compound is bioconcentrated in an organism. E_{HOMO} is related to ionization potential. The higher the E_{HOMO} values, the stronger the electron donating ability. E_{LUMO} describes how susceptible the molecule is to interactions with a nucleophile and thus is directly related to electron affinity. Both E_{HOMO} and E_{LUMO} show the tendency of chemicals to undergo orbital-controlled reactions.

Table 2. Correlation matrix between toxicity and parameters

R	log 1/EC_{50}	M_W	E_{HOMO}	E_{LUMO}	log P
log 1/EC_{50}	1.000	0.798	-0.591	-0.685	0.625
M_W	0.798	1.000	-0.250	-0.398	0.795
E_{HOMO}	-0.591	-0.250	1.000	0.791	-0.162
E_{LUMO}	-0.685	-0.398	0.791	1.000	-0.217
log P	0.625	0.795	-0.162	-0.217	1.000

There is an extensive literature describing QSARs between toxicity data and log P and molecular orbital interaction descriptors. Veith and Mekenyan (1993) established a set of QSARs for aromatic chemicals by using average superdelocalizability (S^N_{av}), E_{LUMO} respectively, and log P to explain the variation in acute toxicity of substituted benzenes, phenols, and anilines to fish. The square of the correlation coefficient (R^2) is highest (0.81) when either S^N_{av} and log P, or E_{LUMO} and log P are used. The QSAR for acute toxicity using these molecular descriptors defined a toxicity plane, which included several modes of toxic action. Type (I) narcotics are chemicals located in the region of low reactivity where toxicity varies with hydrophobicity alone. Type (II) narcotics are more toxic than Type (I) narcotics at similar values of log P, and the increase can be explained by stronger electronic interactions with cellular soft nucleophiles.

Cronin and Schultz (1996) analyzed data on acute toxicity of phenol derivatives to *Tetrahymena pyriformis*. An excellent correlation was obtained between toxicity and parameters log P and E_{LUMO}. The authors thought that the obtained QSAR model can be explained mechanistically, demonstrating the anticipated importance of hydrophobicity and electronic parameters.

In addition, Lu et al. (2001) developed QSAR models for the toxicity data of substituted benzenes to algae, using log P and E_{LUMO}. A series of equations were obtained for the toxicity values of different subclasses of compounds. Lu et al. thought that the toxicity may be related chiefly to the intracellular reduction of -NO_2, obtaining electron for those compounds containing two -NO_2 groups, while log P contributes most to the QSAR and E_{LUMO} very little for anilines and phenols.

In this study, 40 compounds were randomly divided into two sets. Of these, 35 compounds were included in the training set to develop models and 5 compounds in the test set for prediction. By multiple linear regression analyses, a series of QSAR equations were developed (see Table 3).

Table 3. QSARs for substituted aromatic hydrocarbons

Eq.	$\log 1/EC_{50}=$	n	R^2	SE	F	Sig.
(1)	$0.411(\pm 0.079)\log P-0.523(\pm 0.086)E_{LUMO}+2.391(\pm 0.190)$	35	0.719	0.355	40.94	0.000
(2)	$0.448(\pm 0.090)\log P-0.526(\pm 0.116)E_{HOMO}-2.342(\pm 1.107)$	35	0.630	0.408	27.24	0.000
(3)	$0.009(\pm 0.001)M_W-0.395(\pm 0.077)E_{LUMO}+2.028(\pm 0.188)$	35	0.801	0.299	64.28	0.000
(4)	$0.010(\pm 0.001)M_W-0.447(\pm 0.085)E_{HOMO}-2.150(\pm 0.805)$	35	0.804	0.296	65.79	0.000

n is the number of compounds; R^2 is the square of correlation coefficient; SE is the standard error; F is the mean square ratio and Sig. is the significance level

Equations (1) and (2) in Table 3 show that the toxicity of substituted benzenes to the *V. fischeri* was related chiefly to their ability to penetrate the cell and to the electronic interactions of the chemicals with the active site through a variety of electronic processes. To those compounds not containing $-NO_2$, such as aniline, methylphenols and phenol, E_{LUMO} values were positive, so toxicity to *V. fischeri* was controlled mainly by hydrophobicity instead of electronic factors. For those compounds containing $-NO_2$, E_{LUMO} values were negative, so the toxicity was greater. The enhanced toxicity may be related chiefly to the intracellular reduction of $-NO_2$ obtaining the electron, especially for those compounds whose $\log P$ values were not so high, E_{LUMO} values were lower and toxicity was greater, such as for dinitrobenzenes, dinitroanilines and dinitrophenols.

Since there was obvious correlation between M_W and $\log P$, Eq. (3)-(4) were derived using M_W instead of $\log P$. It is apparent that correlation of Eq. (3) and (4) were better than for (1) and (2). This result was not surprising because of high correlation between M_W and toxicity (R=0.798). For the compounds studied in this paper, the higher the M_W, the stronger the hydrophobicity, and the greater the toxicity, such as 2,4,6-tribromoaniline and pentachlorophenol.

From a predictive standpoint the equations in Table 3 were all not satisfactory. When residuals were studied for Eq. (1)-(4), two common significant outliers could be observed. 2,3-dichloroaniline and 4-nitroaniline were more toxic than predicted by Eq. (1)-(4) and their isomers.

After removal of these two compounds from the training set, the improving QSAR was obtained from stepwise regression:

$$\log 1/EC_{50}=0.010\,M_W(\pm0.001)-0.521\,E_{HOMO}(\pm0.076)-2.926(\pm0.718) \qquad (5)$$

$$n=33,\ R^2=0.861,\ SE=0.257,\ F=92.88,\ Sig.=0.000$$

When using log P anf E_{LUMO} instead of M_W and E_{HOMO}, the following equation was obtained:

$$\log 1/EC_{50}=0.428\log P(\pm0.068)-0.571\,E_{LUMO}(\pm0.072)+2.248(\pm0.164) \qquad (6)$$

$$n=33,\ R^2=0.815,\ SE=0.297,\ F=66.04,\ Sig.=0.000.$$

Eq. (5) and (6) were used to predict the toxicity, and the predicted values and percentage errors are presented in Table 1. Percentage error (Er%) is defined as the absolute difference between the experimental and predicted values for toxicity divided by the experimental values. The average Er% of Eq. (5) was 5.2% for the training set and 6.7% for the test set, while the average Er% of Eq. (6) was 6.4% and 6.7%, respectively.

The correlation of toxocity with log P and E_{LUMO} in *V. fischeri* (i.e., Eq. (6)) was comparable to the correlation with the same parameters to *T. pyriformis* (Cronin and Schultz, 1996):

$$\log 1/IGC_{50}=0.671(\pm0.022)\log P-0.670(\pm0.055)E_{LUMO}-1.123 \qquad (7)$$

$$n=120,\ R^2=0.899,\ SE=0.262,\ F=523$$

Where IGC_{50} is the 48 hr 50% inhibitory growth concentration in m*M*.

Some differences were apparent from the two QSARs. The correlation for *T. pyriformis* had slightly higher slopes for both parameters. Eq. (7) for *T. pyriformis* explained 8% more of the variance of the toxicity data than Eq. (6) for *V. fischeri*. However, the QSAR for *V. fischeri* contained some types of compounds, such as anilines and nitrobenzenes, not included in Eq. (7). The difference of constant mainly resulted from different unit of toxicity data. When comparing the toxicity of 20 same compounds to *V. fischeri* with those from *T. pyriformis* (see Table 1), most phenols were discovered to be more toxic to *T. pyriformis* than to *V. fischeri*, but for nitrobenzenes and anilines, the situation was converse. However, the linear correlation coefficient of toxicity to *V. fischeri* and to *T. pyriformis* was 0.855.

Although the chemicals examined in this study all contain the same parent compound (benzene), different mechanisms of toxic action are represented. For example, phenol itself and the most simple alkyl- and/or halogenated substituted phenols are thought to result in polar narcosis, and penta- and selected nitro-substituted phenols may act as weak acid uncouplers (Cronin and Schultz 1996). Groups of polar narcotics, including phenols and anilines, have shown hydrophobicity-dependent toxicity (Seward et al. 2001). Nitrobenzenes are reactive compounds, may be regarded as pro-electrophiles, yielding the corresponding potentially highly toxic C-nitroso compounds (Dearden et al. 1995).

Acknowledgment. We thank the National Natural Science Foundation of China and the Science Innovation Foundation of Hohai University for funding this research.

REFERENCES

Cronin MTD, Schultz TW (1996) Structure-toxicity relationships for phenols to *Tetrahymena pyriformis*. Chemosphere 32: 1453-1468

Dearden JC, Cronin MTD, Schultz TW, Lin DT (1995) QSAR study of the toxicity of nitrobenzenes to *Tetrehymena pyrifirmis*. Quant Truct -Act Rela 14: 427-432.

Deneer JW, Sinnige TL, Hermens JLM (1987) QSAR in drug design and toxicology. Elsevier Science Publishers BV, Amsterdam

Lang PZ, Long FS, Yuan X, Ding YZ (1993) Study on pollution of toxic organic compounds in the middle reaches of the Songhua River (Shaokou~Songhuajiang Village section). Adv Environ Sci 1: 47-56

Lu GH, Yuan X, Zhao YH (2001) QSAR study on the toxicity of substituted benzenes on the algae (*Scenedesmus obliquus*). Chemosphere 44: 437-440

Lyman WJ (1982) Handbook of chemical property estimation methods. McGraw-Hill Book Co, New York

Schultz TW, Lin DT, Wilke TS, Arnold LM (1990) Quantitative structure-activity relationships for the *Tetrahymena pyriformis* population growth endpoint: a mechanism of action approach. In: Karcher W and Devillers J (ed) Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology, Kluwer Academic Publishers, Dordrecht, p 241-262

Seward JR, Cronin MTD, Schultz TW (2001) Structure-toxicity analyses of *Tetrahymena pyriformis* exposed to pyridines—an examination into extension of surface-response domains. SAR QSAR Environ Res 11: 489-512

Veith GD, Mekenyan OC (1993) A QSAR approach for estimating the aquatic toxicity of soft electrophiles. Quant Struct-Act Relat 12: 349-356