Research Article
CODEN (USA): IJPLCP

[Dubeyet al., 3(8): Aug., 2012]
ISSN: 0976-7126

INTERNATIONAL JOURNAL OF PHARMACY & LIFE SCIENCES

Modeling of lipophilicity of industrial chemicals using topological as well as physico-chemical descriptors

Vinay K. Dubey*, Kamna Singh and Vijay K. Agrawal
QSAR and Computer Chemical Laboratories, A.P.S University, Rewa, (M.P.) - India

Abstract

In the present work we have considered miscellaneous set of 76 industrial chemicals and modeled their logP using topological as well as physicochemical descriptors. The results indicate that the estimation of log P is very much effective when the topological and physicochemical descriptors are used together. The most appropriate model for the estimation (modeling) of log P indicated that by using the combination of topological and physicochemical descriptors. The results are discussed using variety of statistical approaches.

Key-Words: Lipophilicity, log P, Topological descriptors, QSAR

Introduction

The partition coefficient for octanol water (log P_{ow}) has become the preferred measure for lipophilicity in the development of biological active molecules in which transport across biological membranes is often critical[1-4]. Methods for calculating log P were reported in 1997. The fragment based methods are reasonably accurate and very fast but this suffers from a few limitations, such as the need for many parameters and the inability to calculating log P for structures containing complete novel structural fragment.

After the work of Meyer and Overtone [5, 6] lipophilicity has been recognized as a meaningful parameter in structure activity relationship studies. Hansch[7] through his work made log P as a very important parameter in the area of Medicinal Chemistry[8, 9].

The study of log P and its relationship with lipophilicity revealed a wealth of information on molecular structure.

In this work modeling of log P of a diverse set of 76 Industrial chemicals has been carried out. The parameters chosen are physicochemical viz. MW,MR, MV,PC,IR,ST,D,POL and Topological parameters W, J, JhetZ, Jhetm, Jhetv, Jhete, Jhetp, $^{0}\chi$, $^{1}\chi$, $^{2}\chi$, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$.

*** Corresponding Author**

Table 1: Various compounds and their log P values used in the present study

Compd.No.	Compounds	log P (exp)
1	Methanol	-0.77
2	Acetonitrile	-0.34
3	Ethanol	-0.31
4	Acetone	-0.24
5	Ethylamine	-0.13
6	2-Propanol	0.05
7	Propionitrile	0.16
8	Methyl acetate	0.18
9	1-Propanol	0.25
10	2-Butanone	0.29
11	2-Methyl-2-propanol	0.35
12	Tetrahydrofuran	0.46
13	Propylamine	0.48
14	Diethylamine	0.58
15	2-Butanol	0.61
16	Benzamide	0.64
17	Pyridine	0.65
18	Ethyl acetate	0.73
19	2-Methyl-1-propanol	0.76
20	Cyclohexanone	0.81
21	1-Butanol	0.88
22	Diethyl ether	0.89
23	Aniline	0.90
24	2-Pentanone	0.91
25	Butylamine	0.97
26	N,N-Dimethylformamide	1.01
27	4-Fluoroaniline	1.15
28	Ethyl acrylate	1.32

29	Methyl methacrylate	1.38
30	2-Hexanone	1.38
31	4-Toluidine	1.39
32	Benzaldehyde	1.48
33	1,2-Dichloroethane	1.48
34	Amylamine	1.49
35	Isopropyl ether	1.52
36	1-Pentanol	1.56
37	Nitrobenzene	1.85
38	Hexanoic acid	1.92
39	4-Methylphenol	1.94
40	2-Heptanone	1.98
41	1-Hexanol	2.03
42	Hexylamine	2.06
43	Benzene	2.13
44	1,1,2,2-Tetrachloroethane	2.39
45	Trichloroethylene	2.42
46	m-Nitrotoluene	2.45
47	1,1,1-Trichloroethane	2.49
48	n-Heptylamine	2.57
49	Ethyl benzoate	2.64
50	1-Heptanol	2.72
51	Toluene	2.73
52	Tripropylamine	2.79
53	Carbon Tetrachloride	2.83
54	1-Naphthol	2.84
55	1-Octanol	2.97
56	Bromobenzene	2.99
57	o-Xylene	3.12
58	p-Xylene	3.15
59	Ethyl benzene	3.15
60	m-Xylene	3.20
61	Butyl ether	3.21
62	Naphthalene	3.30
63	N,n-Diethylaniline	3.31
64	1,2-Dichlorobenzene	3.38
65	Tetrachloroethylene	3.40
66	Cyclohexane	3.44
67	1,3-Dichlorobenzene	3.60
68	1,2-Dibromobenzene	3.64
69	Isopropylbenzene	3.66
70	1,2,4-Trimethylbenzene	3.78
71	Acenaphthene	3.92
72	1,2,4-Trichlorobenzene	4.02
73	Biphenyl	4.09
74	Butylbenzene	4.26
75	1,2,4,5-Tetrachlorobenzene	4.82
76	Pentachlorobenzene	5.17

Results and Discussion

Several statistically significant models ware obtained using least square method; they are reported in Table 2. The R^2 values for statistically significant model vary from (one-parametric) **0.8106** to **0.9296** (Seven-parametric). The most significant models are obtained as bellow-

One-parametric model:

$\log P = 2.3288+0.8848(\pm0.0497)\,{}^{0}\chi^{v}$ (1)

N = 76, Se = 0.3126, R^2 = 0.8106, R^2A = 0.8080, F = 316.713,Q =2.8800

Two-parametric model:

$\text{LogP}=-2.9673+0.7502(\pm0.0528)\,{}^{0}\chi^{v}+0.5182(\pm0.1120)\text{J hetp}$ (2)

N = 76, Se = 0.2768, R^2 = 0.8535, R^2A = 0.8495, F = 212.724, Q =3.3378

Three-parametric model:

$\log P = -1.4782+0.7956(\pm0.0469)\,{}^{0}\chi^{v}+0.5811(\pm0.0985)\,\text{Jhetp}-0.7953\,(\pm0.1623)\,\text{J}$ (3)

N = 76, Se = 0.2414, R^2 = 0.8902, R^2A = 0.8856, F = 194.489, Q =3.9085

Four-parametric model:

$\log P = -0.7224+0.8674(\pm0.0532)\,{}^{0}\chi^{v}+0.6733(\pm0.1015)\,\text{Jhetp}-1.0476(\pm0.1848)\text{J}-0.0244(\pm0.0095)\text{ST}$ (4)

N =76, Se = 0.2325, R^2 = 0.8995, R^2A = 0.8938, F = 158.814, Q = 4.0791

Five-parametric model:

$\log P = -10.6029+0.8648(\pm0.0572)\,{}^{0}\chi^{v}-1.1472(\pm0.2267)\,\text{J}-0.1005\,(\pm0.0188)\text{ST}+8.9264(\pm2.0196)\text{IR}+0.3242(\pm0.0567)\,\text{Jhetm}$ (5)

N = 76, Se = 0.2116, R^2 = 0.9179, R^2A = 0.9120, F = 156.505, Q = 4.5274

Six-parametric model:

$\text{logP}= -9.8482+0.6593\,(\pm0.1041)\,{}^{0}\chi^{v}-1.2765(\pm0.2270)\,\text{J}-0.0977\,(\pm0.0182)\,\text{ST}+8.1915\,(\pm1.9823)\text{IR}+0.0047\,(\pm0.0020)\,\text{PC}+0.4243\,(\pm0.0680)\text{Jhetz}$ (6)

N = 76, Se = 0.2053, R^2 = 0.9239, R^2A = 0.9172, F = 139.520, Q = 4.46819

Seven- parametric model:

$\text{logP}= -10.6672+0.5896(\pm0.1038)\,{}^{0}\chi^{v}-0.6865(\pm0.1830)\,\text{J}-0.0888\,(\pm0.0189)\,\text{ST}+8.6221(\pm2.1829)\text{IR}+0.0043\,(\pm0.0019)\,\text{PC}-1.7930(\pm0.4279)\text{Jhetv}+1.9287(\pm0.3395)\text{Jhetp}$ (7)

N = 76, Se = 0.1989, R^2 = 0.9296, R^2A = 0.9223, F = 128.1994, Q = 0.8744

All the above models contain $^{0}\chi^{v}$, which has a negative sign, suggesting that the zero order valance connectivity index has a negative effect towards exhibition of log P. Balaban index also shows a negative coefficient which suggests that cyclization is not favorable for exhibition of log P. A negative coefficient of ST suggests that the molecules having high surface tension will have a negative influence on log P. Molecules with high parachor value will support log P. Where as Jhetm and Jhetp will support the exhibition of log P.

It is interesting to note that out of so many topological parameters only $^{0}\chi^{v}$, J, Jhetp, Jhetv, JhetZ, and Jhete are significant. Similarly ST, IR, and PC are the only physicochemical parameters which can be used for modeling log P. On the basis of statistical parameters recorded in Table 2 and 3, it can be referred that the model number 7 is the best model for modeling log P of the compounds used in present study.

Pogliani's quality factor (Q) [4-8] also suggests that the model 7 is the best model. Lower values of PRESS/SSY, PSE, and S_{PRESS} also support these findings. Similarly R^2cv shows 92% variance in case of model 7.

We have estimated log P for all the 76 compounds used in the present study using model 7. Such values are reported in Table 4. A close look of this table shows a very close agreement between observed and estimated log P values, this further confirms our findings. A further confirmation has been obtained by plotting a graph between observed and estimated log P values obtained by model number 7. Such a comparison is depicted in figure 1.

Table 2: Regression parameters and quality of correlation for various models

Model No.	Parameters used	Se	R^2	R^2A	F	R	Q = R/Se
1	$^{0}\chi^{v}$	0.3126	0.8106	0.8080	316.713	0.9003	2.8800
2	$^{0}\chi^{v}$,Jhetp	0.2768	0.8535	0.8495	212.724	0.9239	3.3378
3	$^{0}\chi^{v}$,Jhetp,J	0.2414	0.8902	0.8856	194.489	0.9435	3.9085
4	$^{0}\chi^{v}$,Jhetp,J,ST	0.2325	0.8995	0.8938	158.814	0.9484	4.0791
5	$^{0}\chi^{v}$,J,ST,IR,Jhetm	0.2116	0.9179	0.9120	156.505	0.9580	4.5274
6	$^{0}\chi^{v}$,J,ST,IR,PC,Jhetz	0.2053	0.9239	0.9172	139.520	0.9612	4.6819
7	$^{0}\chi^{v}$,J,ST,IR,PC,Jhetv,Jhetp	0.1989	0.9296	0.9223	128.194	0.9642	4.8477

Table 3: Cross validated parameters for the proposed models.

Model No.	Parameters used	PRESS/ SSY	R^2cv	S_{PRESS}	PSE
1	$^{0}\chi^{v}$	0.2336	0.7664	0.6059	0.5978
2	$^{0}\chi^{v}$,Jhetp	0.1716	0.8284	0.5364	0.5257
3	$^{0}\chi^{v}$, Jhetp, J,	0.1234	0.8766	0.4678	0.4553
4	$^{0}\chi^{v}$,Jhetp, J, ST	0.1118	0.8882	0.4506	0.4356
5	$^{0}\chi^{v}$, J, ST,IR,Jhetm	0.0895	0.9105	0.4102	0.3936
6	$^{0}\chi^{v}$, J, ST,IR,PC,Jhetz	0.0824	0.9176	0.3978	0.3791
7	$^{0}\chi^{v}$, J, ST,IR,PC,Jhetv, Jhetp	0.0758	0.9242	0.3854	0.3645

Table 4: Observed and estimated log P using model 7

Comp.No.	Obs. log P	Est.log P	Residual
1	-0.77	-0.40	-0.37
2	-0.34	-0.53	0.19
3	-0.31	-0.28	-0.03
4	-0.24	0.19	-0.43
5	-0.13	0.10	-0.23
6	0.05	0.14	-0.09
7	0.16	-0.01	0.17
8	0.18	-0.05	0.23
9	0.25	0.17	0.08
10	0.29	0.71	-0.42
11	0.35	0.51	-0.16
12	0.46	0.29	0.17
13	0.48	0.52	-0.04
14	0.58	1.21	-0.63
15	0.61	0.66	-0.05
16	0.64	1.54	-0.90
17	0.65	0.79	-0.14
18	0.73	0.61	0.13
19	0.76	0.66	0.10
20	0.81	1.42	-0.61
21	0.88	0.68	0.20
22	0.89	0.98	-0.09
23	0.90	1.40	-0.50
24	0.91	1.29	-0.38
25	0.97	1.01	-0.04
26	1.01	0.05	0.96
27	1.15	1.22	-0.07
28	1.32	0.82	0.50
29	1.38	0.87	0.52
30	1.38	1.85	-0.47
31	1.39	2.18	-0.79
32	1.48	1.91	-0.43
33	1.48	1.64	-0.16
34	1.49	1.53	-0.04
35	1.52	1.92	-0.40
36	1.56	1.23	0.33
37	1.85	1.04	0.81
38	1.92	1.79	0.13
39	1.94	1.88	0.07
40	1.98	2.42	-0.44
41	2.03	1.78	0.25
42	2.06	2.07	-0.01
43	2.13	1.66	0.47
44	2.39	2.70	-0.31
45	2.42	2.69	-0.27
46	2.45	1.90	0.55
47	2.49	2.45	0.04

48	2.57	2.61	-0.04
49	2.64	2.97	-0.33
50	2.72	2.35	0.37
51	2.73	2.30	0.44
52	2.79	3.46	-0.67
53	2.83	2.67	0.17
54	2.84	3.12	-0.28
55	2.97	2.92	0.05
56	2.99	2.98	0.01
57	3.12	2.93	0.19
58	3.15	2.97	0.18
59	3.15	2.82	0.33
60	3.20	2.95	0.25
61	3.21	3.06	0.15
62	3.30	3.56	-0.26
63	3.31	3.15	0.16
64	3.38	3.36	0.02
65	3.40	2.78	0.62
66	3.44	3.37	0.07
67	3.60	3.08	0.52
68	3.64	4.24	-0.60
69	3.66	3.43	0.23
70	3.78	3.59	0.19
71	3.92	3.92	0.00
72	4.02	3.76	0.26
73	4.09	4.40	-0.31
74	4.26	3.88	0.38
75	4.82	4.43	0.39
76	5.17	5.08	0.09

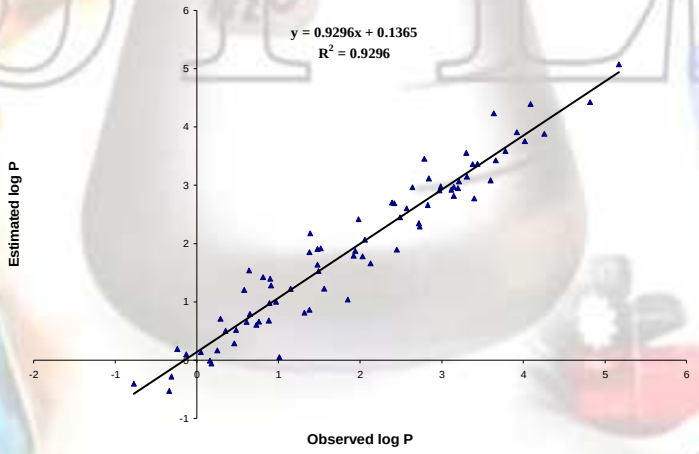

Figure1: Comparison between observed and estimated log P using model 7

References

1. Carrupt, P.; Testa, B.; Gaillard, P.; *In Reviews in Computational Chemistry*, Lipkowitz, K. B.; Boyd, D. B.; Edn., Wiley-VCH, New York, **1997; 1899** Vol.*11*, p.241-345.
2. Hansch, C.; Leo, A. J.; *Substituent Constants for Correlation Analysis in Chemistry and Biology;* Wiley, NewYork, **1979.**
3. Hansch, C.; Leo, A.; *Exploring QSAR Fundamentals and Applications in Chemistry and Biology;* American Chemical Society, Washington, D.D., **1995.**
4. Martin, Y. C.; *Quantitative Drug Design: A Critcal Introduction*, Marcel Dekker, New York, **1978.**
5. Meyer, H.; Welche, I.; *Arch. Exp. Pathol. Pharmakol.*, **1899,***42*, p.109.
6. Overton, E.; *Vierteljahrsschr. Naturforsch. Ges.* Z *rich*, **1899,** *44*, p.87.
7. Hansch, C.; Maloney, P.P.; Fujita, T.; Muir, R. M. *Nature*. **1962**, *194*, p.178.
8. Helmer, F.; Kiehs, K.; Hansch, C.; *Biochemistry*, **1968**, *7*, p.28-58.
9. Rekker, R. F.; *The Hydrophobic Fragmental Constant*, Nauta, W. T.; Rekker, R. F.; Edn., Elsevier, Amsterdam,**1977.**
10. Lybrand, T. P.; *In Reviews in Computational Chemistry*, Lipkowitz, K. B.; Boyd, D. B.; Edn., Wiley-VCH, New York, **1990,** *1*, p.295-320.
11. Straatsma, T. P.; *In Reviews in Computational Chemistry*, Lipkowitz, K. B.; Boyd, D. B.; Edn., Wiley-VCH, New York, **1996,** *9*, p.81-127.
12. Tute, M. S.; *In Advances in Drug Research*, Testa, U.; Meyer, U. A.; Edn., Academic Press, London, **1995,** *26*, p. 45-142.
13. Balbes, L. M.; Mascarella, S. W.; Boyd, D. B.; *In Review in Computational Chemistry*, Lipkowitz, K. B.; Boyd, D. B.; Edn., Wiley-VCH, New York, **1994,** *5*, p. 337-379.
14. Toropov, A. A.; Voropaeva, N. L.; Ruban, I. N.; Rashidova, S. Sh.; *Polymer Sci. Ser. A.*, **1999**, *41*, p.975-985.
15. Toropov, A. A.; Toropov, A. P.; *J. Mol. Struct. (Theochem)*, **2001**, *538*, p.287-193.
16. Toropov, A. A.; Toropov, A. P.; J. *Mol. Struct. (Theochem)*, **2001**, *538*, p.197-199.
17. Peruzzo, P.; Marino, D. M.; Toropov, A. A.; Castro, E. A.; *J. Mol. Struct. (Theochem)*, **2001**,*572*, p.53-60.
18. Mercader, A; *Castro*, E. A.; Toropov, A. A.; *J. Mol. Model.*, **2001**, *7*, p.1-5.
19. Krenkel, G.; Castro, E. A.; Toropov, A. A.; *J. Mol. Struct. (Theochem)*, **2001**, *542*, p.107-113.
20. Krenkel, G.; Castro, E. A.; Toropov, A. A.; *Int. J. Mol.* Sci.,**2001**, *2*, p.57-65.
21. Mercader, A.; Castro, E. A.; Toropov, A. A.; *Chem. Phys. Lett.*, **2000**, *330*, p. 612-623.
22. Krenkel, G.; Castro, E. A.; Toropov, A. A.; *Chem. Phys. Lett.*, **2002**, *355*, p.517-528.
23. Marino,D. J. G.; Peruzzo, P. J.; Castro, E. A.; Toropov, A. A.; *InternetElectron.J.Mol.Des.*,**2002**,*1*,p.115-133.
24. Duehowicz, P.; Castro, E. A.; Toropov, A. A.; *Comp. Chem.*, **2002**, *26*, p.327-332.
25. Balaban, A. T.; Edn., *Chemical Applications of Graph Theory*, Academic Press, New York, **1976**.
26. Randic, M.; Basak, S. C.; *J. Chem. Inf.comput.Sci.*, **1999**, *39*, p.*261-266.*
27. Basak, S. C.; Grunwald, G. D.; *New* J. *Chem.*,**1995,** *19*,p.231 237.
28. Pasyukov, A. V.; Skvortsova, M. I.; Palyulin V. A.; Zefirov, N. S.; *Dokl. Chem.*, **2000,** *3741*, p.224-227.