Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 419–422

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

QSAR Study on Solubility of Alkanes in Water and Their Partition Coefficients in Different Solvent Systems Using PI Index

Padmakar V. Khadikar,[a,*] Dheeraj Mandloi,[b] Amrit V. Bajaj[c] and Sheela Joshi[c]

[a]*Research Division, Laxmi Fumigation and Pest Control, 3, Khatipura, Indore 452007, India*
[b]*Institute of Engineering and Technology, D.A. University, Indore 452017, India*
[c]*School of Chemical Sciences, D.A. University, Indore 452017, India*

Received 11 September 2002; accepted 31 October 2002

Abstract—The aqueous solubility (S_w) of liquids and solids, expressed as log S_w as well as their partition coefficients in different solvent systems viz. P_{oct} (partition coefficient in octanol-water), P_{16} (partition coefficient in water-hexadecane), P_{alk} (partition coefficient in water-alkane), and P_{cyc} (partition coefficient in water-cyclohexane), and aqueous solubility (S_w) have been estimated using the PI (Padmakar–Ivan) index and the results compared with those obtained using the widely used Wiener index (W). Regression analysis of the data using *n*-alkanes show that the PI index gives better results than the W index.

Introduction

The solubility of liquids and solids in water (S_w) as well as partition coefficient of solutes in different solvents viz. partition coefficient in octanol-water (P_{oct}), partition coefficient in water-hexadecane (P_{16}), partition coefficient in water-alkane (P_{alk}), and partition coefficient in water-cyclohexane (P_{cyc}) are very important molecular properties that influence the release, transport, and the extent of absorption of drugs in the body.[1–5] These properties are the key determinants of the environmental fate of agrochemicals and pollutants in the environment.

A number of methods are also reported for estimating the aforementioned parameters using molecular descriptors other than topological indices.[6] However, very little work is done for the estimation of aforementioned parameters using topological indices. This has prompted us to undertake the present investigation in that we have used recently introduced the PI (Padmakar–Ivan) index[7–12] for modeling, monitoring, and estimating P_{oct}, P_{16}, P_{alk}, P_{cyc} and S_w and compared the results with those obtained using widely used Wiener index (W).[13] In doing so we have chosen *n*-alkanes (Table 1) as all these parameters for this set of compounds are easily available in the literature[1,5] which can be adopted.

Results and Discussion

A perusal of Table 1 shows that all the five properties viz. log P_{oct}, log P_{16}, log P_{alk}, log P_{cyc}, and log S_w increase with the size of the alkanes. It means that these properties are the function of size, shape, and branching of the molecules. Hence, it appears that Wiener-type indices will be the most appropriate for modeling, monitoring, and estimating these properties.

The data presented in Table 2 shows that like the aforementioned properties, the magnitude of W and PI indices also increase with the size of the alkanes under present study. This means, both these indices would be quite suitable for modeling the five properties mentioned above.

The data presented in Tables 2 and 3 indicate that simple regressions are enough to model these five properties. In view of this we have carried out simple regressions and the results are discussed below. The regression parameters and quality of correlations needed for this purpose are given in Table 3.

*Corresponding author. Tel./fax: +91-731-531906; e-mail: sachin_g@sancharnet.in

0960-894X/03/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0960-894X(02)00953-8

Table 1. Partition coefficient and solubility of alkanes (refs 1 and 5)

Compd	Alkane	log P_{oct}	log P_{16}	log P_{alk}[a]	log P_{cyc}[a]	−log S_w
1	Methane	1.09	1.14	1.37	1.33	0.90
2	Ethane	1.81	1.83	2.05	2.06	1.36
3	Propane	2.36	2.49	2.73	2.79	1.94
4	*n*-Butane	2.89	3.13	3.38	3.62	2.57
5	*n*-Pentane	3.39	3.87	4.06	4.27	3.18
6	*n*-Hexane	3.90	4.49	4.58	—	3.84
7	*n*-Heptane	4.50	5.14	—	—	4.53
8	*n*-Octane	5.15	5.79	—	—	5.24
9	*n*-Nonane	5.65	6.49	—	—	5.88
10	*n*-Decane	6.34	7.01	—	—	6.98
11	*n*-Undecane	6.54	—	—	—	7.59
12	*n*- Dodecane	6.80	—	—	—	7.67
13	*n*-Tridecane	7.56	—	—	—	—
14	*n*-Tetradecane	8.00	—	—	—	7.96
15	*n*-Hexadecane	9.57	—	—	—	8.40

P_{oct}—Partition coefficient in octanol–water, P_{16}—partition coefficient in water–hexadecane, P_{alk}—partition coefficient in alkanes, P_{cyc}—partition coefficient in water–cyclohexane, and S_w—solubility in water.
[a]Due to a small sample size no regression analyses was carried out for modeling log P_{alk} and log P_{cyc}, as they may result into false models.

Table 2. Wiener (W) and PI indices of alkanes (refer to Table 1)

Compd	Alkane	W	PI	1/W	1/PI	−log S_w
1	Methane	0	0	0.0000	0.0000	0.90
2	Ethane	1	0	1.0000	0.0000	1.36
3	Propane	4	2	0.2500	0.5000	1.94
4	*n*-Butane	10	6	0.1000	0.1667	2.57
5	*n*-Pentane	20	12	0.0500	0.0833	3.18
6	*n*-Hexane	35	20	0.0290	0.0500	3.84
7	*n*-Heptane	56	30	0.0179	0.0333	4.53
8	*n*-Octane	84	42	0.0120	0.0238	5.24
9	*n*-Nonane	120	56	0.0080	0.0179	5.88
10	*n*-Decane	165	72	0.0061	0.0139	6.98
11	*n*-Undecane	220	90	0.0046	0.0111	7.59
12	*n*- Dodecane	286	110	0.0035	0.0091	7.67
13	*n*-Tridecane	364	132	0.0028	0.0076	—
14	*n*-Tetradecane	455	156	0.0022	0.0064	7.96
15	*n*-Hexadecane	680	210	0.0015	0.0048	8.40

Modeling log P_{oct}

The data presented in Table 3 show that the PI index is a better index than the W index for modeling log P_{oct}. The related regression equation is found as follows:

$$\log P_{oct} = 2.6043 + 0.0392\ PI$$
$$n = 15,\ \ S_y = 2.2426,\ \ r = 0.9480 \quad (1)$$

Here, and hereafter, n, is the number of compounds; S_y, is the standard deviation in y (log P_{oct}) or any other property; and r, is the regression coefficient. Based on these statistics we believe that the regression eq 1 is an excellent model for modeling log P_{oct}. This finding is in accordance with our earlier results.[14]

Modeling log P_{16}

Compared to log P_{oct} the data set available for modeling log P_{16} is smaller. Here, we have only 10 log P_{16}, data available for 10 alkanes. The data presented in Table 3 show that in this case also the PI index gives better results than the W index. The corresponding model is found as follows:

$$\log P_{16} = 2.3882 + 0.0681\ PI$$
$$n = 10,\ \ S_y = 1.8719,\ \ r = 0.9244 \quad (2)$$

Modeling log P_{alk} and log P_{cyc}

The data set for modeling log P_{alk} and log P_{cyc} are still smaller than the cases discussed above.

The regression analyses with such a small data will give false models and therefore, their modeling is not attempted.

Table 3. Regression parameters and quality of correlations

Property	Topological index	S_x	S_y	A	B	r	Q
log P_{oct} ($n=15$)	W	64.9020	2.4320	2.7838	0.0365	0.8964	0.3685
	PI	54.1920	2.2426	2.6043	0.0392	0.9480	0.4227
log P_{16} ($n=10$)	W	56.8199	1.8719	2.5615	0.0295	0.8962	0.4787
	PI	25.4034	1.8719	2.3882	0.0681	0.9244	0.4938
log S_w ($n=14$)	W	135.2096	2.5707	3.3914	0.0137	0.7226	0.2811
	PI	64.9020	2.4320	2.7838	0.0365	0.8964	0.3686

S_x—Standard deviation in x (molecular descriptor); S_y—standard deviation in y (property or activity); A—intercept; B—slope; r—correlation coefficient; Q—quality factor.

Table 4. Comparison of estimated parameters with the experimental parameter using the Wiener (W) index

Compd	log P_{oct}			log P_{16}			−log S_w		
	Expt.	Estd.	Res.	Expt.	Estd.	Res.	Expt.	Estd.	Res.
1	1.09	2.78	−1.69	1.14	2.56	−1.42	0.90	3.39	−2.43
2	1.81	2.82	−1.01	1.83	2.59	−0.76	1.36	3.39	−2.49
3	2.36	2.93	−0.57	2.49	2.67	−0.18	1.94	3.44	−1.50
4	2.89	3.14	−0.25	3.13	2.86	0.27	2.57	3.52	−0.95
5	3.39	3.51	−0.12	3.87	3.15	0.71	3.18	3.66	−0.48
6	3.90	4.06	−0.16	4.49	3.59	0.89	3.84	3.87	−0.03
7	4.50	4.82	−0.32	5.14	4.21	0.92	4.53	4.15	0.38
8	5.15	5.84	−0.69	5.79	5.04	0.75	5.24	4.54	0.70
9	5.65	7.16	−1.51	6.49	6.10	0.39	5.88	5.03	0.85
10	6.34	8.80	−2.46	7.01	7.42	−0.41	6.98	5.65	1.33
11	6.54	10.81	−4.27	—			7.59	6.40	1.19
12	6.80	13.22	−6.42	—			7.67	7.30	0.37
13	7.56	16.06	−8.50	—			—		
14	8.00	19.39	−11.39	—			7.96	9.62	−1.66
15	9.57	27.60	−18.03	—			8.40	12.70	−4.30

Expt.—experimental, Estd.—estimated, Res.—residue.

Table 5. Comparison of estimated parameters with the experimental parameter using the PI index

Compd	log P_{oct}			log P_{16}			−logS_w		
	Expt.	Estd.	Res.	Expt.	Estd.	Res.	Expt.	Estd.	Res.
1	1.09	2.60	−1.51	1.14	2.38	−1.24	0.90	2.78	−1.88
2	1.81	2.60	−0.79	1.83	2.38	−0.55	1.36	2.78	−1.42
3	2.36	2.68	−0.32	2.49	2.52	−0.03	1.94	2.85	−0.91
4	2.89	2.83	0.06	3.13	2.80	0.33	2.57	3.00	−0.43
5	3.39	3.07	0.32	3.87	3.21	0.66	3.18	3.22	−0.04
6	3.90	3.38	0.52	4.49	3.75	0.74	3.84	3.51	0.33
7	4.50	3.78	0.72	5.14	4.43	0.71	4.53	3.88	0.65
8	5.15	4.25	0.90	5.79	5.25	0.54	5.24	4.32	0.92
9	5.65	4.80	0.85	6.49	6.20	0.29	5.88	4.83	1.05
10	6.34	5.43	0.91	7.01	7.29	−0.28	6.98	5.41	1.57
11	6.54	6.13	0.41	—			7.59	6.07	1.52
12	6.80	6.91	−0.11	—			7.67	6.80	0.87
13	7.56	7.77	−0.21	—			—		
14	8.00	8.71	−0.71	—			7.96	8.47	−0.51
15	9.57	10.83	−1.26	—			8.40	10.44	−2.04

Expt.—experimental, Estd.—estimated, Res.—residue.

Modeling log S_w

The data set for modeling log S_w is similar to that of log P_{oct}. Hence, we can attempt successful regressions with these sample. Once again the PI index is found better than the W index for modeling log S_w:

$$\log\ S_w = 2.7838 + 0.0365\ \mathrm{PI}$$
$$n = 14,\quad S_y = 2.4320,\quad r = 0.8964 \tag{3}$$

The results discussed so far indicate that the PI index is a better index than W and that it gave best results for modeling log P_{oct}. The data (Tables 4 and 5) show that in modeling the properties, the correlation potential of the PI index follows the sequence:

$$\log\ P_{oct} > \log\ P_{16} > \log\ S_w \tag{4}$$

Conclusion

From the aforementioned results and discussion we conclude that the PI index is a better index than the W index for monitoring, modeling, and estimating log P_{oct}, log P_{16}, and log S_w for the set of alkanes used in the present study. Among these properties the PI index is best suited for modeling log P_{oct}.

Experimental

Properties—P_{oct}, P_{16}, P_{alk}, P_{cyc}, and S_w

All the five properties P_{oct}, P_{16}, P_{alk}, P_{cyc}, and S_w were adopted from the literature[1,5] and used by converting them to their respective log units.

Wiener index (W)

The Wiener index (W)[13] of a graph G is just the sum of distances of all pairs of vertices of G.

$$W = W(G) = {}_{1/2}\sum d(v, u|G) \quad (5)$$

Where $d(v|G)$ is called the distance number (minimum distances) of vertex v and is defined as:

$$d(v|G) = \sum_{u\,V(G)} d(v, u|G) \quad (6)$$

PI index

The PI index[9,10] in its definition relies on the notation of edge types in the chemical graph of the molecule to be characterized and is considered as the modification of Szeged index.[15,16] The PI index is defined as:

$$PI = PI(G) = \sum_{e} [n_{eu}(e|G) + n_{ev}(e|G)] \quad (7)$$

Where $n_{eu}(e|G)$ is the number of edges lying closer to the vertex u than the vertex v. The meaning of n_{ev} is analogous. Edges equidistant from both ends of the edge uv are not counted (taken into account) for the calculation of the PI index.

Regression analysis

All the regression analyses[17] were carried out using Equation Grapher with regression analyzer software.

References and Notes

1. Abraham, M. H.; Chadha, H. C.; Whiting, G. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1085.
2. Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 835. and references therein.
3. Rekker, R. F.; Mannhold, R. *Calculation of Drug Lipopholicity*; VCH: Weinheim, 1992.
4. Abraham, M. H.; Du, C. M.; Platts, J. A. *J. Org. Chem.* **2000**, *65*, 7114.
5. Khadikar, P. V.; Agrawal, V. K.; Karmarkar, S. *Bioorg. Med. Chem.* **2002**, *10*, 3499, and references therein.
6. Khadikar, P. V.; Mathur, K. C.; Shalini, J.; Phadnis, A.; Shrivastava, A.; Mandloi, M. *Bioorg. Med. Chem.* **2002**, *10*, 1761.
7. Khadikar, P. V.; Phadnis, A.; Shrivastava, A. *Bioorg. Med. Chem.* **2002**, *10*, 1181.
8. Khadikar, P.V.; Karmarkar, S.; Singh, S.; Shrivastava, A. *Bioorg. Med. Chem.* **2002**, *10*, 3163.
9. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29*, 134.
10. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
11. Khadikar, P.V.; Bajaj, A.V.; Mandloi, D. *Indian J. Chem.* **2002**, *41A*, 2065.
12. Khadikar, P.V.; Agrawal, V.K.; Karmarkar, S. *Bioorg. Med. Chem.* In press, and references therein.
13. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
14. Khadikar, P. V.; Singh, S.; Shrivastava, A. *Bioorg. Med. Chem. Letters* **2002**, *12*, 1125.
15. Gutman, I. *Graph Theory Notes New York* **1994**, *27*, 9.
16. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
17. Chaterjee, S.; Hadi, A.; Price, B. *Regression Analysis by Examples*, 3rd ed.; John Wiley: New York, 2000.