

The relationship between molecular connectivity and partition coefficients

J. C. BOYD, J. S. MILLERSHIP* AND A. D. WOOLFSON

Department of Pharmacy, Medical Biology Centre, The Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL U.K.

Kier & Hall (1976) demonstrated that a relationship existed between molecular connectivity and Log P (the Log of the octanol/water partition coefficient) for several classes of compounds and also for the combined classes. This work utilized a number of Log P values calculated using π values. We have reinvestigated this relationship using experimentally determined Log P values. The results confirm Kier & Hall's findings although different connectivity indices are brought into the equations.

In 1968 Hansch et al demonstrated that a linear relationship existed between the octanol/water partition coefficient and the aqueous solubility for a series of organic liquids. In a recent paper Valvani et al (1981) have investigated the relationship between octanol/water partition coefficients and aqueous solubility for a series of liquid non-electrolytes and demonstrated a similar relationship to that found by Hansch et al (1968). Valvani et al (1981) point out that the use of experimentally determined partition coefficient values is essential for reliable results, and that Hansch et al (1968) had had to use calculated partition values (π values) due to the lack of reliable, experimentally determined partition values.

Kier & Hall (1976) have reported that a linear relationship exists between log P (log of the octanol/water partition coefficient) and $^1X^v$ (a modified first order valence connectivity term) for a series of compounds. An examination of this work showed that, in a number of cases, the log P values used were based on Hansch's calculated log P values (Hansch et al 1968). In connection with our continuing work on molecular connectivity we have been investigating the use of connectivity indices in relation to log P for structure activity studies. During the course of this work we have re-examined the fundamental relationship between log P and molecular connectivity using experimentally derived (albeit from the literature) log P values. In view of the comments by Valvani et al (1981) it seemed propitious to report these findings.

Calculation of molecular connectivity indices

Calculations of molecular and valence connectivity indices (up to and including the sixth order) made throughout this paper follow the general method of

* Correspondence.

Kier & Hall (1976). These values were calculated using the computer program CFUNC. No use was made of the modified $^1X^v$ index as suggested by Kier & Hall (1976).

Log P values

All data used for the log P octanol/water values were taken from Leo et al (1971). The log P values used were all experimentally determined, no calculated log P values were used.

RESULTS AND DISCUSSION

Individual series of compounds were first investigated to find the best single parameter equation relating log P to molecular or valence connectivity indices. The series considered were alcohols, amines, phenols, anilines and acids and the results for these groups of compounds are summarized in Table 1. The results indicate excellent correlation in all cases except the phenols. Superior correlations were obtained in all cases when multi-parameter equations were developed. Equation 6-10 describe the results for the individual series.

Alcohols

$$\text{Log P} = 1.267 {}^1X + 0.612 {}^3X_p - 0.976 {}^3X_p^v - 2.130 \\ r \ 0.999 \quad s = 0.068 \quad n = 14 \quad (6)$$

Table 1. Correlation of connectivity indices with log P.

Compound class	Connectivity index	Slope	Intercept	r	n	s	Eqn
Alcohols	1X	1.099	-1.794	0.996	14	0.135	1
Amines	1X	0.907	-1.432	0.991	12	0.146	2
Phenols	3X_c	2.039	1.657	0.936	16	0.165	3
Anilines	$^6X_{pc}^v$	2.691	0.991	0.976	16	0.167	4
Acids	$^2X^v$	1.475	-0.831	0.998	8	0.099	5

Amines

$$\text{Log P} = 2.750 {}^1X - 1.542 {}^1X^v - 0.852 {}^4X_p^v - 2.433$$

$$r = 0.998 \quad s = 0.082 \quad n = 12 \quad (7)$$

Phenols

$$\text{Log P} = 1.738 {}^3X_c^v - 0.414 {}^4X_{pc} - 159.651 {}^6X_{ch}^v + 6.621$$

$$r = 0.988 \quad s = 0.078 \quad n = 16 \quad (8)$$

Anilines

$$\text{Log P} = 4.957 {}^6X_{pc}^v - 0.497 {}^9X^v + 2.829 \quad (9)$$

$$r = 0.983 \quad s = 0.147 \quad n = 16$$

Acids

$$\text{Log P} = 1.514 {}^2X^v - 0.334 {}^3X_c - 0.732 \quad (10)$$

$$r = 0.999 \quad s = 0.064 \quad n = 8$$

At this point the relationship between log P and ${}^1X^v$ was investigated for the individual compound classes. In this case ${}^1X^v$ was calculated using valence delta values for all heteroatoms and all unsaturated atoms. The results are shown in Table 2 (eqns 11–15). When the combined groups of 66 compounds was investigated the best single parameter equation included ${}^1X^v$, however, the correlation coefficient and the standard error were both poor (eqn 16). A

Table 2. Correlation of ${}^1X^v$ with Log P.

Compound class	Connectivity index	Slope	Intercept	r	n	s	Eqn
Alcohols	${}^1X^v$	1.090	-1.353	0.995	14	0.157	11
Amines	${}^1X^v$	0.887	-1.104	0.987	12	0.175	12
Phenols	${}^1X^v$	0.640	0.467	0.883	16	0.219	13
Anilines	${}^1X^v$	1.148	-1.285	0.958	16	0.220	14
Acids	${}^1X^v$	1.070	-1.148	0.993	8	0.197	15

two parameter equation involving ${}^1X^v$ and ${}^6X_{ch}^v$ gave much improved correlation and lower standard error (eqn 17).

$$\text{Log P} = 1.044 {}^1X^v - 1.063 \quad (16)$$

$$r = 0.943 \quad s = 0.379 \quad n = 66$$

$$\text{Log P} = 0.989 {}^1X^v + 18.739 {}^6X_{ch}^v - 1.193 \quad (17)$$

$$e = 0.970 \quad s = 0.277 \quad n = 66$$

Table 3 compares observed and calculated Log P

Table 3. Comparison of observed and calculated log P values.

Compound	Log P		
	Obs	Calc ^a	Calc ^b
Methanol	-0.82	-0.75	-0.86
Ethanol	-0.32	-0.18	-0.34
n-Propanol	0.34	0.31	0.38
n-Butanol	0.88	0.81	0.86
Isobutanol	0.83	0.67	0.89
s-Butanol	0.61	0.74	0.69
t-Butanol	0.37	0.51	0.40

Table 3—cont.

Compound	Log P		
	Obs	Calc ^a	Calc ^b
n-Pentanol	1.40	1.30	1.40
Isopentanol	1.16	1.16	1.22
2,2-Dimethyl-1-propanol	1.36	0.95	1.30
t-Pentanol	0.89	1.07	0.92
n-Hexanol	2.03	1.80	1.95
n-Octanol	3.15	2.79	3.03
n-Dodecanol	5.13	4.76	5.20
Methylamine	-0.57	-0.62	-0.57
Isopropylamine	-0.03	0.28	0.04
Trimethylamine	0.27	0.13	0.26
n-Butylamine	0.88	0.90	0.77
t-Butylamine	0.40	0.58	0.31
Diethylamine	0.57	0.91	0.72
Methylbutylamine	1.33	1.34	1.33
Dimethylbutylamine	1.70	1.69	1.66
Dipropylamine	1.73	1.89	1.74
Triethylamine	1.44	1.84	1.46
Dibutylamine	2.68	2.88	2.73
Tripropylamine	2.79	3.33	2.75
m-Bromophenol	2.63	2.33	2.58
o-Bromophenol	2.35	2.34	2.33
p-Bromophenol	2.59	2.33	2.55
m-Chlorophenol	2.50	1.95	2.36
o-Chlorophenol	2.15	1.95	2.13
p-Chlorophenol	2.39	1.95	2.33
m-Fluorophenol	1.93	1.24	1.94
o-Fluorophenol	1.71	1.25	1.77
p-Fluorophenol	1.77	1.24	1.91
m-Iodophenol	2.93	3.05	3.00
o-Iodophenol	2.65	3.05	2.69
p-Iodophenol	2.91	3.04	2.97
m-Nitrophenol	2.00	1.88	1.92
o-Nitrophenol	1.79	1.90	1.84
p-Nitrophenol	1.91	1.88	1.89
Phenol	1.46	1.52	1.46
m-Bromoaniline	2.10	2.40	2.14
o-Bromoaniline	2.29	2.40	2.41
p-Bromoaniline	2.26	2.40	2.41
m-Chloroaniline	1.90	2.01	1.78
o-Chloroaniline	1.92	2.02	1.94
p-Chloroaniline	1.83	2.01	1.94
m-Fluoroaniline	1.30	1.30	1.11
o-Fluoroaniline	1.26	1.31	1.08
p-Fluoroaniline	1.15	1.30	1.08
m-Iodoaniline	2.98	3.11	2.82
o-Iodoaniline	3.34	3.12	3.29
p-Iodoaniline	3.34	3.11	3.29
m-Nitroaniline	1.37	1.95	1.56
o-Nitroaniline	1.79	1.95	1.71
p-Nitroaniline	1.39	1.95	1.43
Aniline	0.98	1.59	1.23
Formic acid	-0.54	-0.71	-0.57
Acetic acid	-0.17	-0.28	-0.14
Propionic acid	0.33	0.28	0.26
Butyric acid	0.79	0.77	0.86
Hexanoic acid	1.88	1.76	1.94
m-Toluic acid	2.37	2.29	2.30
p-Toluic acid	2.27	2.29	2.29
Decanoic acid	4.09	3.74	4.08

^a Calculated using equation 17.

^b Calculated for each group of compounds individually using equations 6–10.

values, the first column of calculated values being derived from Equation 17 whilst the second column results from equations 6–10. The results shown in

this Table allow the following conclusions to be drawn:

(1) For individual groups of compounds there is excellent correlation between molecular connectivity and the Log P values.

(2) For the combined group of sixty-six compounds the correlation is less good although still reasonable given the range of Log P values (six Log units) and the diversity of the compounds in the list.

The best single parameter equations (1–5) for the individual groups include a number of different connectivity indices reflecting the different functionalities of the various groups. The multiparameter equations (6–10) give improved correlation and lower standard error indicating the greater degree of structural information encoded by the extended connectivity indices included. Equations (11–15) in which $^1X^v$ is forced into the correlation for each individual group produces results that are only slightly poorer than equations (1–5). This reflects the usefulness of $^1X^v$ as a general descriptor of a molecule taking into account unsaturation and the presence of heteroatoms. This usefulness is further exemplified by the incorporation of $^1X^v$ into equations 16 and 17. The predicted log P values obtained using equation 17 are good but limited when the aromatic compounds are considered. In most cases equation 17 does not produce any discrimination between *ortho*, *meta* and *para* substituted compounds. The calculated log P values using equations 6–10 give excellent results and the discriminatory

power with the phenols and anilines is clearly demonstrated. Thus it would appear that $^3X^v_c$ and $^4X^v_{pc}$ in equation 8 and $^6X^v_{pc}$ in equation 9 are able to encode information relating to the substituent patterns in the ring systems.

These results confirm Kier & Hall's earlier findings (Kier & Hall 1976). However, different connectivity indices are included in the equations generated. The results also add weight to the suggestion of Valvani et al (1981), and other earlier workers, that in structure-activity studies it is always best to use experimentally derived results. It should further be added that the log P values used in this paper, although taken from one compilation, were from a variety of sources which may or may not have resulted in certain inaccuracies.

Acknowledgement

This work was supported by a Postgraduate Studentship awarded to J. C. Boyd by the Department of Education for Northern Ireland.

REFERENCES

- Hansch, C., Quinlan, J. E., Lawrence, G. L. (1968) *J. Org. Chem.* 33: 347–350
- Kier, L. B., Hall, L. H. (1976) *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, London, pp 158–167
- Leo, A., Hansch, C., Elkins, D. (1971) *Chem. Rev.* 71: 525–616
- Valvani, S. C., Yalkowsky, S. H., Roseman, T. J. (1981) *J. Pharm. Sci.* 70: 502–507