Comparison of Reliability of log *P* Values for Drugs Calculated by Several Methods

Ikuo MORIGUCHI,[a] Shuichi HIRONO,*,[a] Izumi NAKAGOME,[a] and Hiroyuki HIRANO[b]

School of Pharmaceutical Sciences, Kitasato University,[a] 5–9–1 Shirokane, Minato-ku, Tokyo 108, Japan and Zeria Pharmaceutical Co., Ltd.,[b] Oshikiri, Konan-machi, Saitama 360–01, Japan.

Received October 5, 1993; accepted December 26, 1993

The values of log *P* (partition coefficient in octanol/water) for 22 drugs selected by Rekker *et al.* [*Quant. Struct.-Act. Relat.*, 12, 152 (1993)] were calculated by the simple method of Moriguchi *et al.* [*Chem. Pharm. Bull.*, 40, 127 (1992); *M*] using only 7 parameters, and compared with those obtained from the procedures of Rekker (*R*), Hansch and Leo (*H*), and Suzuki and Kudo (*S*). The descending order of the reliability was *M*, *H*, *R*, and *S* in the standard error of estimate, *R*, *H*, *M*, and *S* in the averaged absolute residual-sums, and *M*, *R*, *H*, and *S* in the regression coefficients between observed and estimated log *P* values. Serious errors of estimate beyond 2.0 in the absolute residual were seen in *R*, *H*, and *S*, but not in *M*. These results showed that the method of Moriguchi *et al.* is not only simple and convenient but also reliable in the application to drugs with complex structures.

Keywords hydrophobicity; log *P*; simple calculation method; octanol/water partition; comparative reliabilty

The importance of log *P* (partition coefficient in octanol/water), which is closely related to the transport properties of drugs and their interaction with receptors, was revealed by Hansch and Fujita[1] in their quantitative structure–activity relationship (QSAR) studies. Since then, there has been an ever-increasing need for prediction of log *P* for various structures, especially those for which experimental values are not available. Recently, the use of non-congeneric QSARs in toxicity for large sets of data has been attempted by regulatory agencies and industry to screen compounds for possible health and environmental hazards. For this purpose, we developed a simple method[2] of calculating approximate values of log *P* applicable to diverse structures of organic molecules using only 13 structural parameters.

Recently, Rekker *et al.*[3] reported a comparative investigation of the results of three typical methods of calculating log *P*: procedures of Rekker,[4] Hansch and Leo,[5] and Suzuki and Kudo.[6] In their report,[3] the log *P* values calculated and observed for 22 selected drugs were overviewed. For comparison, we have calculated the log *P* values of the same 22 drugs without omission using our simple method, and presented its reliability in the application of drugs with complex structures.

Methods

The simple method developed in our laboratory applies an equation with 13 structural parameters,[2] 6 of which were not applicable to any of the 22 drugs. The 7 parameters used for calculation of log *P* in this study were as follows[2]:

CX: summation of weighted numbers of carbon and halogen atoms; the weights are 0.5 for F, 1.0 for C and Cl, 1.5 for Br, and 2.0 for I

NO: total number of N and O atoms

PRX: proximity effect of N/O; 2.0 for X–Y and 1.0 for X–A(–Y) (X, Y: N and/or O; A: C, S, or P; –: saturated or unsaturated bond) with a correction (−1) for –CON< and –SO_2N<

UB: number of unsaturated bonds including semi-polar bonds (such as *N*-oxides and sulfoxides), except those in NO_2

POL: number of aromatic polar substituents (aromatic substituents excluding Ar–C(X)(Y)- and Ar–C(X)=C; X, Y: C and/or H). The upper value was set at 4.0[2]

RNG: indicator variable for the presence of ring structures except benzene and its condensed rings (aromatic, heteroaromatic and hydrocarbon rings)

NO2: number of nitro groups

The remaining parameters appearing in Eq. 1 are as follows[2]:

HB: dummy variable for the presence of intramolecular hydrogen bond as *ortho*-OH and –CO–R, –OH and –NH_2, –NH_2 and –COOH, or 8-OH/NH_2 in quinolines, 5 or 8-OH/NH_2 in quinoxalines, *etc.*

AMP: amphoteric property; α-amino acid, 1.0; aminobenzoic acid, 0.5; pyridinecarboxylic acid, 0.5

ALK: dummy variable for alkane, alkene, cycloalkane, or cycloalkene (hydrocarbons with 0 or 1 double bond)

QN: quaternary nitrogen: >N^+<, 1.0; N-oxide, 0.5

NCS: isothiocyanato (–N=C=S), 1.0; thiocyanato (–S–CN), 0.5

BLM: dummy variable for the presence of β-lactam

Calculations were performed on a Toshiba Sparc LT AS1000/L10 work station using self-written programs.

Results and Discussion

Our simple method of calculating log *P* applies the following equation with 13 parameters and gives good reliability as previously reported.[2]

$$\log P = 1.244(CX)^{0.6} - 1.017(NO)^{0.9} + 0.406PRX - 0.145(UB)^{0.8} + 0.511HB + 0.268POL - 2.215AMP + 0.912ALK - 0.392RNG - 3.684QN + 0.474NO2 + 1.582NCS + 0.773BLM - 1.041 \quad (1)$$

$n=1230$, $r=0.952$, $s=0.411$, $F_0(13, 1216)=900.4$

where n=number of compounds, s=standard error of estimate, and F_0=*F*-statistics for the correlation.

Since 6 parameters in Eq. 1 were not applicable to any of the structures of the 22 drugs to be calculated, a simplified form (Eq. 2) was used for the calculation in this study:

$$\log P = 1.244(CX)^{0.6} - 1.017(NO)^{0.9} + 0.406PRX - 0.145(UB)^{0.8} + 0.268POL - 0.392RNG + 0.474NO2 - 1.041 \quad (2)$$

Values of the 7 parameters included in Eq. 2 for the 22 drugs are listed in Table I. The calculated log *P* values are listed in Table II, along with the observed values and the calculated values by other three methods cited from the report of Rekker *et al.*[3]

The reliability of the log P values calculated by the simple method of Moriguchi *et al.*[2] (M) was compared with those obtained from the methods of Rekker[4] (R), Hansch and Leo[5] (H), and Suzuki and Kudo[6] (S). As shown in Table III, the descending order of accuracy of calculated log P values for the 22 drugs was M, H, R, and S in the standard error of estimate, R, H, M, and S in the averaged absolute residual-sums, and M, R, H, and S in the regression coefficients with the observed log P values.

The regression equations formulated are as follows:

Moriguchi method:

$$\log P = 1.261(\pm 0.283)M - 0.371(\pm 0.740) \quad (3)$$

$n=22$, $r=0.901$, $s=0.764$, $F_0=86.2$

Rekker method:

$$\log P = 0.807(\pm 0.197)R + 0.327(+0.653) \quad (4)$$

$n=22$, $r=0.886$, $s=0.816$, $F_0=73.3$

Hansch–Leo method:

$$\log P = 0.885(\pm 0.237)H + 0.441(\pm 0.688) \quad (5)$$

$n=22$, $r=0.867$, $s=0.877$, $F_0=60.59$

Suzuki–Kudo method:

$$\log P = 0.676(\pm 0.312)S + 1.044(\pm 0.890) \quad (6)$$

$n=22$, $r=0.710$, $s=1.240$, $F_0=20.38$

In these equations, the coefficient for the independent variable (M, R, H, or S) and the intercept must be theoretically 1.000 and 0.000, respectively. The Moriguchi, Rekker, and Hansch–Leo methods meet this requirement within their 95% confidence intervals, but the Suzuki–Kudo method does not.

The Moriguchi method was superior to the Rekker and the Hansch–Leo methods in the standard error of estimate and the regression coefficient, although it was inferior in the averaged absolute residual-sum. The inferiority in the standard error and regression coefficient for the log P values calculated from the latter two methods is relevant to their large values of maximum absolute residual, 2.36 (with verapamil) for the Rekker method and 2.87 (with furosemide) for the Hansch–Leo method, compared to 1.63 (with propafenone) for the Moriguchi method.

In the derivation of Eq. 1, observed values of log P for 10 of the 22 drugs in Table I are included: the observed values used for 3 drugs (atropine, chloramphenicol, and trimethoprim) were the same as those listed in Table II,

TABLE I. Structural Parameters for 22 Drugs

No.	Drug	*CX*	*NO*	*PRX*	*UB*	*POL*	*RNG*	*NO2*
1	Atropine	17	4	2	4	0	1	0
2	Chloramphenicol	13	7	5	4	1	0	1
3	Chlorothiazide	8	7	6	8	2	0	0
4	Chlorpromazine	18	2	0	6	1	0	0
5	Cimetidine	10	6	7	4	0	1	0
6	Diazepam	17	3	1	8	4	1	0
7	Diltiazem	22	6	3	8	3	1	0
8	Diphenhydramine	17	2	0	6	0	0	0
9	Disopyramide	21	4	1	7	0	1	0
10	Flufenamic acid	15.5	3	2	7	4	0	0
11	Furosemide	13	7	4	8	4	1	0
12	Haloperidol	22.5	3	0	7	3	1	0
13	Imipramine	19	2	0	6	2	1	0
14	Lidocaine	14	3	1	4	1	0	0
15	Phenobarbital	12	5	4	6	0	1	0
16	Phenytoin	15	4	3	8	0	1	0
17	Procainamide	13	4	1	4	2	0	0
18	Propafenone	21	4	0	7	2	0	0
19	Propranolol	16	3	0	5	1	0	0
20	Tetracaine	15	4	2	4	2	0	0
21	Trimethoprim	14	7	5	6	4	1	0
22	Verapamil	27	6	0	7	4	0	0

TABLE II. Observed and Calculated Values of log P for 22 Drugs

No.[a]	Obs[b]	M[c]	Δ[d]	R[e]	Δ	H[f]	Δ	S[g]	Δ
1	1.83	2.21	−0.38	1.88	−0.05	1.32	0.51	0.03	1.80
2	1.14	1.23	−0.09	0.32	0.82	0.69	0.45	−0.75	1.89
3	−0.10	−0.36	0.26	−0.68	0.58	−1.24	1.14	−0.44	0.34
4	5.35	3.77	1.58	5.10	0.25	5.20	0.15	3.89	1.46
5	0.40	0.82	−0.42	0.63	−0.23	0.21	0.19	3.33	−2.93
6	2.80	3.36	−0.56	3.18	−0.38	3.32	−0.52	1.23	1.57
7	2.70	2.67	0.03	4.53	−1.83	3.55	−0.85	1.96	0.74
8	3.27	3.26	0.01	3.41	−0.14	2.93	0.34	3.35	−0.08
9	2.58	2.47	0.11	2.57	0.01	1.35	1.23	2.54	0.04
10	5.25	3.86	1.39	5.81	−0.56	5.58	−0.33	5.16	0.09
11	−0.83	0.43	−1.26	1.38	−2.21	2.04	−2.87	1.33	−2.16
12	3.36	4.01	−0.65	3.57	−0.21	3.52	−0.16	3.43	−0.07
13	4.80	3.88	0.92	4.43	0.37	4.41	0.39	3.38	1.42
14	2.26	2.52	−0.26	2.30	−0.04	1.36	0.90	0.91	1.35
15	1.47	0.78	0.69	1.23	0.24	1.37	0.10	1.29	0.18
16	2.47	1.80	0.67	2.76	−0.29	2.09	0.38	2.01	0.46
17	0.88	1.72	−0.84	1.11	−0.23	1.11	−0.23	0.65	0.23
18	4.63	3.00	1.63	4.15	0.48	3.21	1.42	3.86	0.77
19	3.56	2.53	1.03	3.46	0.10	2.75	0.81	2.15	1.41
20	3.73	2.64	1.09	3.55	0.18	3.65	0.08	2.90	0.83
21	0.91	1.26	−0.35	−0.07	0.98	0.66	0.25	0.57	0.34
22	3.79	3.23	0.56	6.15	−2.36	3.53	0.26	6.49	−2.70

a) For drug names, see Table I. *b*) Ref. 3. *c*) Calculated values using the method of Moriguchi *et al.* *d*) $\Delta = \log P$ (obs) − log P (calc). *e*) Calculated values[3] using the method of Rekker. *f*) Calculated values[3] using the method of Hansch and Leo. *g*) Calculated values[3] using the method of Suzuki and Kudo.

TABLE III. Comparison of the Reliabilities of Four Methods

Methods	Standard error[a]	Mean $\lvert\Delta\rvert$[b]	Max $\lvert\Delta\rvert$	Regression coefficient with log P (obs)
Moriguchi (M)	0.83	0.67	1.63	0.901
Rekker (R)	0.88	0.57	2.36	0.886
Hansch–Leo (H)	0.87	0.62	2.87	0.867
Suzuki–Kudo (S)	1.35	1.04	2.93	0.710

a) $\sqrt{\Sigma\Delta^2/n}$. b) $\Delta = \log P$ (obs) $- \log P$ (calc).

while those for 7 drugs (chlorothiazide, chlorpromazine, diazepam, haloperidol, imipramine, phenobarbital, and propranolol) were somewhat different. For reference, the structure–log P model was recalculated excluding the data of the 10 drugs. The regression equation obtained was as follows:

$$\begin{aligned}\log P = {} & 1.246(CX)^{0.6} - 1.017(NO)^{0.9} + 0.408PRX - 0.148(UB)^{0.8} \\ & + 0.511HB + 0.270POL - 2.217AMP + 0.914ALK \\ & - 0.389RNG - 3.687QN + 0.476NO2 + 1.586NCS \\ & + 0.765BLM - 1.048 \qquad (7)\end{aligned}$$

$$n = 1220, \quad r = 0.952, \quad s = 0.411, \quad F_0(13, 1206) = 890.5$$

Equation 7 is substantially the same as Eq. 1: this assures the predicted values of M.

The Moriguchi method was developed to estimate log P values for use in non-congeneric QSAR analyses of large data sets. The intention was that the procedure be simple enough to be applicable to any type of organic molecule except metalloorganic chemicals and polymers, even though the estimation of log P is not as precise.[2] In the present study, this was satisfactorily demonstrated by the results that serious errors of estimate beyond 2.0 in the absolute residual as shown in the other three methods were not seen in the Moriguchi method.

In conclusion, the method of Moriguchi *et al.* is not only simple and convenient but also comparatively reliable in its application to drugs with complex structures.

References

1) C. Hansch, T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).
2) I. Moriguchi, S. Hirono, Q. Liu, I. Nakagome, Y. Matsushita, *Chem. Pharm. Bull.*, **40**, 127 (1992).
3) R. F. Rekker, A. M. ter Laak, R. Mannhold, *Quant. Struct.-Act. Relat.*, **12**, 152 (1993).
4) R. F. Rekker, H. M. de Kort, *Eur. J. Med. Chem.*, **14**, 479 (1979).
5) C. Hansch, A. Leo, "Substituent Constants for Correlation Analysis in Chemistry and Biology," Wiley, New York, 1979, pp. 18—43.
6) T. Suzuki, Y. Kudo, *J. Comput.-Aided Mol. Design*, **4**, 155 (1990).