

# Prediction of 1-Octanol/Water Partition Coefficients of Organophosphorus Compounds

Takahiro SUZUKI\*

## Abstract

For predicting the logarithm of the partition coefficient between 1-octanol and water phases for organophosphorus compounds, two types of empirical prediction schemes have been developed. One is a group-contribution method and group contributions of phosphorous containing groups have been determined. The other is a correlation equation between  $\log P$  values and fundamental physicochemical properties of organic compounds. The latter is easy to use because of its employment of calculable molecular descriptors, and it is useful for an alternative prediction scheme for  $\log P$  by the group-contribution approach. The results of this paper will be implemented in an automatic  $\log P$  prediction system CHEMICALC-2 and might be useful for many practical purposes.

**Key words:** group contribution, hydrophobicity,  $\log P$ , partition coefficient, physical property

## 1. Introduction

Legislation in many countries requires a manufacture of chemicals to include physicochemical properties of molecules in its specifications in recent years. Of these properties, the partition coefficient between 1-octanol and water phases is one of the most frequently used properties. Therefore, reliable values of the partition coefficients of chemicals are necessary in the molecular design. An enormous amount of such data has been collected and correlated over the years, however, rapid advance of technology in discovery or synthesis of new compounds maintains a significant gap between demand and availability. The ultimate generalization of the partition coefficient will require a complete understanding of molecular behavior, which we do not have yet. In the mean time, the empirical approaches based on additive-constitutive principles for estimating  $\log P$  (where  $P$  is the 1-octanol/water partition coefficient) may be useful for the many purposes (Leo et al., 1971; Rekker & de Kort, 1979; Broto et al., 1984; Klopmann et al., 1984; Ghose & Crippen, 1986; Sangster, 1997; Mannhold et al., 1998).

In a previous paper (Suzuki & Kudo, 1990), we proposed a convenient method for estimating  $\log P$  based on the group-contribution method and developed a program for its

---

\* Natural Science Laboratory, Toyo University, 5-28-20 Hakusan, Bunkyo-ku, Tokyo 112-8606, Japan

automatic calculation. Besides, we reported an alternative correlation equation that predicts the log  $P$  value in terms of fundamental physicochemical properties of the solute from a set of 167 simple organic compounds (Suzuki et al., 1993). However, the applicability of both methods was limited to working with the organic compounds containing carbon, hydrogen, oxygen, nitrogen, halogens, or sulfur. The purpose of the present paper is the extension of our previous approaches to organophosphorus compounds. Organophosphorus compounds are an important class in pharmacy, in agriculture, in biology, or in environmental chemistry, and the acquisition of their values of the partition coefficients is critical for many practical purposes.

## 2. Principles

### *Group-contribution model*

A group-contribution method for estimating log  $P$  according to our approach can be expressed by the following model without any correction terms (Suzuki & Kudo, 1990),

$$\log P = \sum_i^N n_i \cdot G_i \quad (1)$$

where  $N$  is the total number of groups,  $n_i$  is the number of the  $i$ th group in the molecule, and  $G_i$  is the group-contribution to log  $P$ . One predominant feature of this approach is its combined handling of three sets of group contributions, the Basic Group, the Extended Group and the User Defined Groups Set. The determination of the contribution values of the Basic Group for organophosphorus compounds is performed here.

The definition of the Basic Group is similar to Benson-type groups (Benson, 1976). In most cases a group is taken to contain one polyvalent atom together with its monovalent ligands and its environment (so-called polyvalent atoms and some hetero atoms). The previously determined 415 group values of the Basic Group set were fixed to evaluate contributions for phosphorus containing groups.

### *Regression model*

The logarithm of the partition coefficient for the 1-octanol/water system for simple organic compounds was well correlated by the equation (Suzuki et al., 1993):

$$\log P = 0.00972PA - 0.0223P_c + 5.75q_m^- + 0.969N_F - 0.925N_N - 0.824N_S + 1.72 \quad (2)$$

Regression range:  $-1.51 \leq \log P \leq 6.90$

$n = 172$ ,  $r = 0.973$ ,  $s = 0.370$

where  $PA$  is the parachor,  $P_c$  is the critical pressure,  $q_m^-$  is the atomic charge of most negative atom in the molecule as calculated by the RESCHA method (Baumer et al., 1989), and  $N_F$ ,  $N_N$ , and  $N_S$  are, respectively the numbers of constituent fluorine, nitrogen, and sulfur atoms.

For the application of the above model to organophosphorus compounds, a problem arise. The values of  $P_c$  are not readily available because of the lack of reliable prediction method for the compounds containing phosphorus. Fortunately,  $P_c$  has a minor role in the above model

(Suzuki et al., 1993), therefore, we now find that one obtains almost as good fit for  $\log P$  values by ignoring  $P_c$  term for the original data set:

$$\log P = 0.0116PA + 5.62q_m^- + 1.094N_F - 0.973N_N - 0.935N_S + 0.325 \quad (3)$$

$n = 167, r = 0.969, s = 0.392$

This approximate equation was examined for organophosphorus compounds here and the prediction performance was compared with that by the group-contribution approach. The values of  $PA$  were calculated from sums of atom or group contributions (Reid et al., 1977). The values of  $q_m^-$  were calculated on the basis of chemical formula as in an earlier paper (Suzuki et al., 1993).

### 3. Results and Discussion

The experimental  $\log P$  values for 55 organophosphorus compounds (Hansch & Leo, 1979) in Table 1 were used to evaluate 21 groups contribution values by the least-squares method based on equation (1). The set covers only a part of the all classes of organophosphorus compounds, but it includes complex and important structures for drugs and agricultural chemicals. The group values are given in Table 2.

The group consists of a key group and its environment. A key group (PO, PS, NH<sub>2</sub>, NH, - - -) is taken to contain a central polyvalent atom (C,N,O,S and P) and the attached atoms bonded to it. An attached atom is any atom bonded to only one other atom (e.g., H, halogens, O in CO group, N in CN group, etc.). Certain atomic groups, cyano, nitro, amino, and such, are treated as univalent hetero atoms. The carbon and nitrogen atoms are categorized and the shorthand notation is used as the footnote in Table 2. For the notation of the group, the key group is listed first, followed by the symbol, —, which indicates the group is bonded directly to the nearest neighbors as can be seen in Table 2. Parentheses are placed around the ligands not bonded to one another. A numerical subscript indicates how many of such ligands are attached.

Fourth column of the Table 2 shows the group contributions to  $\log P$  ( $G_i$  in eq.(1)). Some contribution values were assigned arbitrarily to circumvent a colinearity problem in the same way as our previous work (Suzuki & Kudo, 1990). The list of experimental and calculated  $\log P$  values in the original data set is shown in Table 1.

For estimating the  $\log P$  values of phosphorus containing compounds, another fragment or group-contribution method of Broto et al. (1984) is available. Correlations between observed  $\log P$  values and estimated ones by this method and the method of Broto et al. (1984) are shown in the following equations:

This method

$$\log P_{obs} = 0.963 \log P_{est} + 0.103 \quad (4)$$

( $n=55, r=0.982, s=0.26, F=1474$ )

The method of Broto et al.

$$\log P_{obs} = 1.002 \log P_{est} + 0.250 \quad (5)$$

(n=55, r=0.944, s=0.46, F=433)

Although this comparison was made for the available data set (training set), the model developed in this study produced more accurate results from a statistical viewpoint. Thus the usefulness of this approach was confirmed. The results of this study will be included in our automatic log *P* estimation program entitled CHEMICALC-2 (Suzuki, 1991).

The results of calculated log *P* values for 55 compounds in Table 1 by eq.(3) are shown in last column of Table 1 in comparison with the calculated values by the group-contribution method. As can be seen from this table, this regression model gives comparable predictions with the group-contribution model for the majority of the compounds studied. However, the calculated log *P* values for the compounds containing large hydrophilic moiety (such as compounds 3, 16, 18, 21, 22, 23, 25, 28, 33,34,36, 43-49, 51,52, 54, and 55) are rather higher than the measured ones. This shows the limits of the applicability of the present model to the multifunctional compounds. For such complex compounds having large hydrophilic moieties additional descriptors will be required.

**Table 1.** Experimental and calculated log *P* values of organophosphorus compounds.

Compound	Formula	log <i>P</i>		
		Obsd.	Eq.(1)	Eq.(3)
1. Methamidophos	C2H8NO2PS	-0.66	-0.72	-0.78
2. O, S-Dimethyl-N-methylphosphoramidothioate	C3H10NO2PS	-0.07	-0.33	-0.28
3. Dipterex	C4H8Cl3O4P	0.51	0.63	3.02
4. O, S-Diethylphosphoramidothioate	C4H12NO2PS	0.15	0.01	0.15
5. O, S-Dimethyl-N-ethylphosphoramidothioate	C4H12NO2PS	0.07	0.11	0.17
6. O, S-Dimethyl-N-propylphosphoramidothioate	C5H14NO2PS	0.65	0.65	0.64
7. O, S-Dimethyl-N-s-butylphosphoramidothioate	C6H16NO2PS	0.95	0.95	1.10
8. O, S-Dimethyl-N-t-butylphosphoramidothioate	C6H16NO2PS	0.96	0.64	1.10
9. O, S-Dimethyl-N-n-butylphosphoramidothioate	C6H16NO2PS	0.94	1.18	1.10
10. O, S-Dimethyl-N-isobutylphosphoramidothioate	C6H16NO2PS	0.97	1.11	1.10
11. O, S-Dipropylphosphoramidothioate	C6H16NO2PS	1.23	1.38	1.08
12. Hempa	C6H18N3OP	0.28	0.28	0.05
13. O-Methyl-O-(2,4,5-trichlorophenyl)-phosphoramidate	C7H7Cl3NO3P	2.53	2.41	2.83
14. Ronnel	C8H8Cl3O3PS	4.88	4.94	4.52
15. Dicapthon	C8H9ClNO5PS	3.44	3.49	3.30
16. 2-Chloro-dimethylparaoxon	C8H9ClNO6P	1.83	1.80	3.03
17. O-Methyl-O-(2,4,5-trichlorophenyl)-methylphosphoramidate	C8H9Cl3NO3P	3.06	2.79	3.33
18. Dimethyl paraoxon	C8H10NO6P	1.33	0.74	2.57
19. Dimethylparathion	C8H10NO5PS	2.04	2.43	2.84
20. O-Methyl-O-(2,4-dichlorophenyl)methyl-phosphoramidate	C8H10Cl2NO3P	2.41	2.26	2.86
21. O, O-Dimethylphenylphosphonate	C8H11O3P	0.95	0.95	2.56
22. O-Methyl-O-(2-chlorophenyl)-methyl-phosphoramidate	C8H11ClNO3P	1.61	1.62	2.40
23. O, O-Dimethyl-O-phenylphosphate	C8H11O4P	1.22	1.01	2.87
24. Ethoprop	C8H19O2PS2	2.03	2.03	2.27
25. Disulfoton	C8H19O2PS3	1.93	1.92	2.89
26. O-Methyl-O-(2,4,5-trichlorophenyl)ethyl-phosphoramidate	C9H11Cl3NO3P	3.18	3.24	3.79
27. Fenitrothion	C9H12NO5PS	3.38	2.86	3.31
28. 3-Methyl-dimethylparaoxon	C9H12NO6P	1.69	1.15	3.04
29. O-Methyl-O-(2,4,5-trichlorophenyl)propyl-phosphoramidate	C10H13Cl3NO3P	3.70	3.78	4.26
30. Dichlofenthion	C10H13Cl2O3PS	5.14	5.38	5.00
31. Parathion	C10H14NO5PS	3.81	3.40	3.78
32. 3-Ethyl-dimethylparathion	C10H14NO5PS	3.74	3.37	3.77
33. Paraoxon	C10H14NO6P	1.69	1.70	3.51
34. 3-Ethyl-dimethylparaoxon	C10H14NO6P	2.19	1.68	3.51
35. O, O-Diethyl-O-phenylphosphorothioate	C10H15O3PS	3.46	3.67	4.09
36. O, O-Diethyl-O-phenylphosphate	C10H15O4P	1.64	1.98	3.81

**Table 1.** Continued.

Compound	Formula	log <i>P</i>		
		Obsd.	Eq.(1)	Eq.(3)
37. O, O-Diethyl-O-(2,6-dichloro-4-methiophenyl) phosphate	C11H15Cl2O4PS	3.72	4.18	4.15
38. O-Methyl-O-(2,4,5-trichlorophenyl)-isobutylphosphoramidate	C11H15Cl3NO3P	3.86	4.24	4.72
39. O-Methyl-O-(2,4,5-trichlorophenyl)-butylphosphoramidate	C11H15Cl3NO3P	4.18	4.31	4.72
40. O-Methyl-O-(2,4,5-trichlorophenyl)-s-butylphosphoramidate	C11H15Cl3NO3P	3.99	4.07	4.72
41. O-Methyl-O-(2,4,5-trichlorophenyl)-t-butylphosphoramidate	C11H15Cl3NO3P	3.83	3.76	4.72
42. 3-Isopropyl-dimethylparathion	C11H16NO5PS	4.05	4.00	4.23
43. 3-Isopropyl-dimethylparaaxon	C11H16NO6P	2.57	2.31	3.97
44. O, O-Diethyl-O-p-cresylphosphate	C11H17O4P	2.18	2.4	4.26
45. O, O-Diethyl-O-(4-methylthiophenyl)-phosphate	C11H17O4PS	2.24	2.52	3.89
46. O,O-Diethyl-O-(2-Cl-4-methiophenyl)-phosphate	C11H16ClO4PS	3.05	3.58	4.34
47. O-(p-Nitrophenyl)-O-propylethylphosphonate	C11H16NO5P	2.20	2.20	3.65
48. O, O-Diethyl-O-(4-methylsulfonylphenyl)-phosphorothioate	C11H17O5PS2	2.43	2.27	5.08
49. O, O-Diethyl-O-(4-methylsulfonylphenyl)-phosphate	C11H17O6PS	0.00	0.47	4.79
50. O, O-Diethyl-O-(2-methyl-4-methiophenyl)-phosphorothioate	C12H19O3PS2	4.16	3.89	4.64
51. O, O-Diethyl-O-(2-methyl-4-methiophenyl)-phosphate	C12H19O4PS	3.21	2.94	4.35
52. O, O-Diethyl-O-(3-methyl-4-methiophenyl)-phosphate	C12H19O4PS	3.28	2.94	4.35
53. Crufomate	C12H19ClNO3P	3.42	3.24	4.27
54. O-Methyl-O-(3-t-butylphenyl)methyl-phosphoramidate	C12H20NO3P	2.65	2.59	3.80
55. O-Methyl-O-(3-t-butylphenyl)methyl-phosphoramidate	C12H20NO3P	2.71	2.59	3.80

**Table 2.** Group contributions to log *P* of phosphorus containing compounds.

Group*	Frequency of Use	No. of Compounds	Gi	Remarks
PO-(O)2(N)	12	12	-2.686	
PO-(N)(O)(S)	10	10	-2.026	
PO-(C)(O)2	1	1	-1.610	
PO-(Car)(O)2	1	1	-3.030	
PO-(O)3	17	17	-2.516	
PO-(O)(S)2	1	1	-2.940	
PO-(N)3	1	1	-1.931	
PS-(O)3	11	11	-0.821	
PS-(O)2(S)	1	1	-2.000	
NH2-(PO)	3	3	-0.124	≡NH2-(CO), assigned
NH-(C)(PO)	18	18	-0.347	
N-(C)2(PO)	4	2	-0.483	
O-(Car)(PO)	29	29	-0.062	≡O-(C')(CO)
O-(C)(PO)	60	42	-0.062	≡O-(C')(CO)
O-(Car)(PS)	11	11	-0.062	≡O-(C')(CO)
O-(C)(PS)	24	12	-0.062	≡O-(C')(CO)
S-(C)(PO)	12	12	0.275	≡S-(C)(S)
S-(C)(PS)	1	1	0.275	≡S-(C)(S)
CH2-(C)(PO)	1	1	0.226	
CH-(PO)(O)(C)	1	1	0.994	≡CH-(C')(CO)(O)
Car-(Car)2(PO)	1	1	1.044	≡Car-(Car')2(CO)

\* (C)= sp<sup>3</sup> C-atom; (Car)= aromatic sp<sup>2</sup> C-atom; (C')=C, Cd (aliphatic sp<sup>2</sup> C-atom), Ct (sp C-atom), or Car; (Car')=Car or Carf (aromatic fused C-atom); (N)=single bonded N-atom.

## References

- Baumer, I., Sala, G., Sello, G., 1989. Residual Charges on Atoms in Organic Structures: A New Algorithm for Their Calculation. *Tetrahedron Computing Methodology*, **2**, 37-46.
- Benson, S. W., 1976. Thermochemical Kinetics, 2<sup>nd</sup> ed., Wiley, New York.
- Broto, P., Moreau, G., Vanduycke, C., 1984. Molecular Structures, Perception, Autocorrelation Descriptor and SAR Studies; System of Atomic Contributions for the Calculation of the Octanol-Water Partition Coefficient. *Eur. J. Med. Chem.*, **19**, 71-78.
- Ghose, A.K., Crippen, G. M., 1986. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships. I. Partition Coefficients as a Measure of Hydrophobicity. *J. Comput. Chem.*, **7**, 565-577.
- Hansch, C., Leo, A., 1979. Substituent Constants for Correlation Analysis in Chemistry and Biology, Wiley, New York.

- Klopman, G., Namboodiri, K., Schochet, M., 1985. Simple Method of Computing the Partition Coefficient. *J. Comput. Chem.*, **6**, 28-38.
- Leo, A., Hansch, C., Elkins, D., 1971. Partition Coefficients and Their Uses. *Chem. Rev.*, **71**, 525-616.
- Mannhold, R., Rekker, R. F., Dross, K., Bijloo, G., de Vries, G., 1998. The Lipophilic Behaviour of Organic Compounds: 1. An Updating of the Hydrophobic Fragmental Constant Approach. *Quant. Struct.-Act. Relat.*, **17**, 517-536.
- Reid, R. C., Prausnitz, Sherwood, T. K., 1977. The Properties of Gases and liquids, 3rd ed., McGraw-Hill, New York.
- Rekker, R. F., de Kort, H. M., 1979. The Hydrophobic Fragmental Constant; An Extension to a 1000 Data Point Set. *Eur. J. Med. Chem.*, **14**, 479-488.
- Sangster, J., 1997. Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry, John Wiley & Sons, Chichester, England.
- Suzuki, T., Kudo, Y., 1990. Automatic Log P Estimation based on Combined Additive Modeling Methods. *J. Comput.-Aided Mol. Design*, **4**, 155-198.
- Suzuki, T., 1991. Development of an Automatic Estimation System for both the Partition Coefficient and Aqueous Solubility. *J. Comput.-Aided Mol. Design*, **5**, 149-166.
- Suzuki, T., Ohtaguchi, K., Koide, K., 1993. Correlations between Octanol/Water Partition Coefficients and Intrinsic Molecular Properties of Simple Organic Compounds. *J. Chem. Eng. Japan*, **26**, 114-115.