Journal of Medicinal Chemistry

Volume 18, Number 9 September 1975

Calculation of Hydrophobic Constant (Log P) from π and f Constants†

A. Leo,* P. Y. C. Jow, C. Silipo,‡ and C. Hansch

Department of Chemistry, Pomona College, Claremont, California 91711. *Received February 11, 1975*

The highest level of confidence can be placed in calculated log P values when (1) the log P of a parent solute is known, (2) π constants for the required substituent(s) are available, and (3) the substituents either do not have an effect on groups already present in the parent or else this effect has been previously determined. In some instances there are no values available for any related structures which could serve as a parent; then, rather than substitute groups for hydrogen, it is easier to begin "from scratch", as suggested by Nys and Rekker, and assemble the structure from fragments, each of which has been assigned a hydrophobic value. In the present paper some new log P values for the lower alkanes and the inert gases are analyzed with the view of separating hydrophobic effects according to volume (including branching and flexibility) and polarity. Modified fragment values appear to enable reliable calculations to be made for a wider range of structures than was possible with the originally proposed constants.

Interest in the use of hydrophobic parameters to rationalize interactions of small ligands with various macromolecules in the fields of biochemistry,[2] medicinal chemistry,[3] and environmental science[4] continues its rapid development. Considerable experience in the use of hydrophobic parameters in the study of quantitative structure–activity relationships by regression analysis clearly indicates that *measured* log P's (P = partition coefficient in an octanol–water system) should be used whenever possible. Nevertheless, when the number of compounds is great and the structural variation limited, simple economics make it desirable to measure the log P's for only the key structures and to calculate the remainder where any group interactions not in the measured solutes can be assumed to be negligible. While regression analysis requires the most accurate log P values possible, there is an ever-increasing need for reliable estimates of a host of chemicals for which experimental values may be difficult to obtain. An example would be the thousands of compounds being studied as potential hazards to the environment through bioaccumulation.

Whether the log P calculation is made by adding π values based on eq 1,[5] or by summing f values as shown in

$$\log P_{R\text{-}X} = \log P_{R\text{-}H} + \pi_X \quad (1)$$

$$\log P = \sum_1^n a_n f_n \quad (2)$$

eq 2,[1] there is the implied assumption that the hydrophobicity of any structural fragment is invariant or else its variation is predictable according to rules covering attachment to certain "interacting" fragments.

It is evident that "normal" π and f values must be established with the greatest possible precision before "interaction" or "proximity" effects can be evaluated accurately. For this reason we have carefully measured partition coefficients for hydrogen, the lower alkanes (normal, branched, and cyclic), many of the alkenes and alkynes, and a considerable number of halogen derivatives. Those which have not been published elsewhere[6] appear in Table I. It should be noted that improved gas chromatographic techniques[6] and the analysis of both phases have yielded significantly different and more reliable values for some key solutes, e.g., for pentane and cyclohexane. Taken together with the log P values for the inert gases,[6] these new values provide a basis for establishing a relation between the π value and the substituent volume (including the effects of branching and chain flexibility) and separating volume from polar effects. Measurements on the inert gases are especially valuable in relating log P to volume because their rotational behavior does not affect cavity volume requirements. Their hydrophobic nature appears to correspond with that of the perhalogenated alkanes rather than that of the alkanes. Further work in this area is in progress and will be the subject of a forthcoming paper.

It is noteworthy that f_H, obtained from $\frac{1}{2}$ log P_{H_2} = 0.225, is in satisfactory agreement with that obtained by Nys and Rekker[1] who used a statistical approach. However, the methyl fragment value presents some problems.

When the Δ log P/carbon atom for the n-alkanes is extended beyond C-5 with the data from the alcohols and amines[§] (Figure 1), it is clear that C_1 and C_2 compounds are unique. The methyl fragment value can be calculated as $f_{CH_3} = \frac{1}{2} \log P_{CH_3CH_3} = 0.91$, or $f_{CH_3} = \log P_{CH_4} - f_H = 1.09 - 0.23 = 0.86$, both of which are significantly different from Nys and Rekker's value of 0.705. If the average value of $f_{CH_3} = 0.89$ is used, a constant value for f_{CH_2} cannot be obtained from the higher alkanes, alkanols, and alkylamines.

A reasonable solution to this problem is to assume that every single bond *after the first one* makes a negative contribution to hydrophobicity (perhaps a volume reduction

†This research was supported by Grant No. CA-11110 from the National Institutes of Health and by the Pomona College Medicinal Chemistry Project.

‡NATO Fellow and Visiting Professor of Chemistry from the Institute of Pharmaceutical and Toxicological Chemistry, University of Naples, Italy.

§Using present techniques, it is difficult to accurately measure log P's above 4.0 unless the solute absorbs strongly in the uv.

Table I. New Log P Values (Octanol-Water)

No.	Formula	Name	Log P	SD	No. of analyses
1	O_2	Oxygen	0.65	0.01	5
2	C_4H_{10}	Butane	2.89	0.05	7
3	C_4H_{10}	Isobutane	2.76	0.02	5
4	C_5H_{12}	Neopentane	3.11	0.02	4
5	C_5H_{10}	Cyclopentane	3.00	0.04	10
6	C_6H_{12}	Cyclohexane	3.44	0.04	16
7	CH_3Br	Methyl bromide	1.19	0.04	8
8	$CH_3C{\equiv}CH$	Methylacetylene	0.94	0.04	8
9	CHF_3	Fluoroform	0.64	0.02	9
10	$C_2H_4Cl_2$	1,2-Dichloroethane	1.48	0.02	8
11	$(CH_3)_2C{=}CH_2$	Isobutylene	2.34	0.03	6
12	C_6H_{10}	Cyclohexene	2.86	0.02	4
13	C_4H_6	1,3-Butadiene	1.99	0.02	7
14	C_4H_8	*cis*-2-Butene	2.33	0.02	14
15	C_4H_8	*trans*-2-Butene	2.31	0.03	13
16	C_2H_5OH	Ethanol	−0.31	0.02	12
17	CH_3OCH_3	Dimethyl ether	0.10	0.02	11
18	CH_3OH	Methanol	−0.77	0.02	12
19	C_3H_7OH	Propanol	0.25	0.01	12
20	H_2	Hydrogen	0.45	0.01	4
21	CH_4	Methane	1.09	0.05	7
22	C_2H_6	Ethane	1.81	0.04	11
23	C_3H_8	Propane	2.36	0.07	6
24	C_4H_8	1-Butene	2.40	0.05	9
25	C_5H_{12}	Pentane	3.39	0.09	29
26	N_2	Nitrogen	0.67	0.01	4

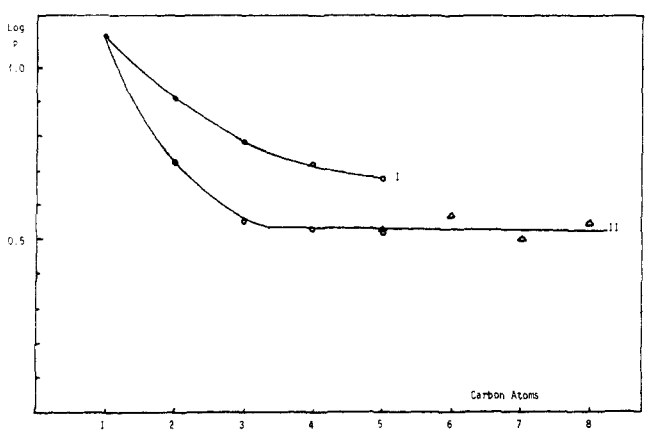

Figure 1. Curve I, log P/C atom; curve II, Δ log P/C atom. (○) from n-alkanes; (△) from 1-alkanols and amines.

through flexibility). Then uniform fragment values can be assigned which are very close to those derived from methane and hydrogen, as shown in Table II, A.

Example

$$\log P_{n\text{-pentane}} = 2f_{CH_3} + 3f_{CH_2} + 3f_b =$$

$$2(0.89) + 3(0.66) + 3(-0.12) = 3.40 \text{ calcd}$$

$$3.39 \text{ obsd}$$

Using f_b for straight chain compounds has the effect of reducing f_{CH_2} to 0.54 which is the Δ log P seen in Figure 1 for C_3 through C_8. Bonds in cyclic alkanes do not lower log P to the same extent, possibly because their reduced flexibility cannot affect volume as greatly. As seen in the following examples, f_b can be assigned a value of −0.09.

Table II. Modified Fragment Constants[a]

A

$f_{CH_3} = 0.89$	$f_b = -0.12$
$f_{CH_2} = f_{CH_3} - f_H = 0.66$	$f_b = -0.09$
$f_{CH} = f_{CH_2} - f_H = 0.43$	$f_{cbr} = -0.13$
$f_C = f_{CH} - f_H = 0.20$	$f_{Rbr} = -0.22$

B

Polar group	Fragment constant f	f_ϕ	$f_{\phi\phi}$	Statistics[b] No. of solutes determining f	Av deviation of f
–Br	0.20	1.09		3	±0.06
–Cl	0.06	0.94		4	±0.04
–F	−0.38	0.37		1	
–I	0.60	1.35		2	±0.02
–N<	−2.16	−1.17	−1.29	1[c]	
$-NO_2$	−1.26	−0.02		3	±0.04
–O–	−1.81[d]	−0.57	0.53	3	±0.03
–S–	−0.79	0.03	0.77	1	
–NH–	−2.11	−1.03	−0.18	11	±0.09
$-NH_2$	−1.54	−1.00[e]		6	±0.02
–OH	−1.64	−0.40[e]		8	±0.06
–CN	−1.28	−0.34		2	±0.01
–C(=O)N<	−3.20	−2.82	−2.09	1	
–C(=O)NH–	−2.71	−1.81	−1.06	1	
$-C(=O)NH_2$	−2.18	−1.26		1	
–C(=O)–	−1.90	−0.32	−0.50	3	±0.00 (3)
–C(=O)O–	−1.49	−0.56	−0.09	3	±0.02
–C(=O)OH	−1.09	−0.03		4	±0.03
$-C_6H_5$	1.90			1[f]	

[a] In addition to those used in ref 1, the following symbols are employed: f_b = single bond between fragments in chains; f_b = single bond between fragments in rings; f_{cbr} = chain branching; f_{Rbr} = group branching; f_ϕ = a fragment attached to one aromatic ring; $f_{\phi\phi}$ = a fragment attached to two aromatic rings. [b] Only for aliphatic fragment constant determinations; f taken from phenyl derivative only. [c] Only Me_3N can be used as higher homologs needed to establish branching factor; see text. [d] For CH_3OCH_3 and c-CH_2CH_2O, use −1.56. [e] Approximately 0.25 higher on α-naphthyl. [f] See text.

	obsd	calcd	difference
 $3f_{CH_2} + 2f_b = 1.72$ (0.66) (−0.09)		1.80	+0.08
 $5f_{CH_2} + 4f_b = 3.00$		2.94	−0.06
 $6f_{CH_2} + 5f_b = 3.44$		3.51	+0.07

If all kinds of branching had the same hydrophobic effect, it could be allowed for by an appropriate reduction in the values for f_{CH} and f_C as derived above but this simplification is not warranted. Adequate treatment of branching on nitrogen atoms as well as chain and group branching on carbon is quite involved and this subject will be treated in detail in a forthcoming paper (however, see below).

Nys and Rekker point out in two recent papers[1] that the use of f values emphasizes the fact that hydrogen itself contributes to hydrophobicity ($f_H \simeq 0.2$) and this approach avoids some of the previous errors in adding π values. It should be clear, however, that f values are subject to the

Chart I

	log P obsd	a. (ref 1) calcd	a. (ref 1) difference	b. (from Table II) calcd	b. (from Table II) difference
A. chain branching					
1. $CH_3CH(CH_3)_2$	2.76				
a. $3f_{CH_3} + f_{CH}$		2.34	−0.42		
(0.702) (0.236)					
b. $3f_{CH_3} + f_{CH} + 2f_b + f_{cbr}$				2.73	−0.03
(0.89) (0.43) (−0.12) (−0.13)					
2. $CH_3C(CH_3)_3$	3.11				
a. $4f_{CH_3} + f_C$		2.95	−0.16		
(0.702) (0.14)					
b. $4f_{CH_3} + f_C + 3f_b + 2f_{cbr}$				3.14	+0.03
(0.89) (0.20) (−0.12) (−0.13)					
B. group branching					
1. $HOCH(CH_3)_2$	0.05				
a. $2f_{CH_3} + f_{CH} + f_{OH}$		0.20	+0.15		
(0.702) (0.236) (−1.44)					
b. $2f_{CH_3} + f_{CH} + f_{OH} + 2f_b + f_{gbr}$				0.11	+0.06
(0.89) (0.43) (−1.64) (−0.12) (−0.22)					
2. $H_2NCH(CH_3)_2$	0.26				
a. $2f_{CH_3} + f_{CH} + f_{NH_2}$		0.26	0		
(0.702) (0.236) (−1.58)					
b. $2f_{CH_3} + f_{CH} + f_{NH_2} + 2f_b + f_{gbr}$				0.21	−0.05
3. $HOC(CH_3)_3$	0.37				
a. $3f_{CH_3} + f_C + f_{OH}$		0.81	+0.44		
b. $3f_{CH_3} + f_C + f_{OH} + 3f_b + 2f_{gbr}$[a]				0.43	+0.06
4. $H_2NC(CH_3)_3$	0.40				
a. $3f_{CH_3} + f_C + f_{NH_2}$		0.87	+0.47		
b. $3f_{CH_3} + f_C + f_{NH_2} + 3f_b + 2f_{gbr}$[a]				0.53	+0.13

[a] When both group and chain branching occur on the same carbon atom, it is treated as two *group* branches.

same proximity effects that π values are and that the fragment concept cannot, as its originators claim, eliminate the need for certain structural corrections such as chain and group branching. This is shown in Chart I in which Nys and Rekker's f values are compared with those of Table II. Adjusting Nys and Rekker's f_{CH} and f_C values to fit the branched alkanes would widen the discrepancy with the alcohols and amines.

Using the new alkane fragment constants in Table II, A, a set of modified polar fragment values was calculated and appears in Table II, B. Note that attachment to an aromatic ring enhances the fragment value and attachment to two rings usually enhances it further. The fragment value for the phenyl ring is dependent upon the degree to which it retains its π electrons after a substituent is attached. The most direct way to obtain f_ϕ for an "undisturbed" phenyl is from $\frac{1}{2}\log P_{\phi-\phi} = (4.04) = 2.02$. However, a log P of over 4 is difficult to obtain with a high degree of accuracy and, since the log P of benzene has been measured with greater care by more laboratories, it could be considered a more suitable standard, especially since a reliable value for hydrogen is now available. We can consider that $f_\phi = \log P_{C_6H_6} - \frac{1}{2}\log P_{H_2} = 2.13 - 0.23 = 1.90$.

Nys and Rekker are correct in stating that partitioning data do not, in themselves, lend support to the conformation where a side chain bearing a polar group is folded over an aromatic ring.[7] This is borne out by the following examples where the polar group fragment constants from the alkyl series (Table II, B) are suitable for use in mixed alkyl–aryl solutes.

	obsd	calcd	difference
$C_6H_5CH_2CH_2OH$			
$f_{C_6H_5} + 2f_{CH_2} + f_{OH} + 2f_b$	= 1.36	1.34	−0.2
$C_6H_5(CH_2)_3OH$			
$f_{C_6H_5} + 3f_{CH_2} + f_{OH} + 3f_b$	= 1.88	1.88	0
$C_6H_5CH_2CH_2NH_2$			
$f_{C_6H_5} + 2f_{CH_2} + f_{NH_2} + 2f_b$	= 1.41	1.44	+0.03
$C_6H_5CH_2CH_2Cl$			
$f_{C_6H_5} + 2f_{CH_2} + f_{Cl} + 2f_b$	= 2.95	3.04	+0.09
$C_6H_5(CH_2)_3Cl$			
$f_{C_6H_5} + 3f_{CH_2} + f_{Cl} + 3f_b$	= 3.55	3.58	+0.03

It is hoped that this continued effort to measure partition coefficients of solutes with structural features of basic importance will enable log P calculations to be made with greater reliability in the future, regardless of whether the substituent or fragment approach is employed.

References and Notes

(1) G. G. Nys and R. F. Rekker, *Chim. Ther.*, 8, 521 (1973); 9, 361 (1974).

(2) (a) K. Bartinek, N. V. Dorovska, and S. D. Varfolomeev, *Biokhimiya*, 37, 1245 (1972); (b) G. M. Cohen and G. J. Manneering, *Mol. Pharmacol.*, 9, 383 (1973); (c) A. Dupaix, J. J. Béchet, and C. Roucous, *Biochemistry*, 12, 2559 (1973); (d) P. B. Hulbert, *Mol. Pharmacol.*, 10, 315 (1974); (e) M. K. Jain and E. H.

Cordes, J. Membr. Biol., 14, 101, 121 (1973); (f) C. Silipo and C. Hansch, Mol. Pharmacol., 10, 954 (1974).
(3) (a) J. L. Barker, Nature (London), 252, 52 (1974); (b) E. J. Lien, J. Kuwahara, and R. T. Koda, Drug Intell. Clin. Pharm., 8, 470 (1974); (c) W. J. Dunn III and C. Hansch, Chem.-Biol. Interact., 9, 75 (1974); (d) J. C. Dearden and J. H. Tubby, J. Pharm. Pharmacol., 26, 73P (1974); (e) S. Inoue, A. Ogino, M. Kise, M. Kitano, S. Tsuchiya, and T. Fujita, Chem. Pharm. Bull., 22, 2064 (1974); (f) T. Fujita, K. Kamoshita, T. Nishioka, and M. Nakajima, Agric. Biol. Chem., 38, 1521 (1974).
(4) W. B. Neely, D. R. Branson, and G. E. Blau, Environ. Sci. Technol., 8, 1113 (1974).
(5) (a) A. Leo, C. Hansch, and D. Elkins, Chem. Rev., 71, 525 (1971); (b) C. Hansch, A. Leo, and D. Nikaitani, J. Org. Chem., 37, 3090 (1972).
(6) C. Hansch, A. Vittoria, C. Silipo, and P. Y. C. Jow, J. Med. Chem., 18, 546 (1975).
(7) C. Hansch and S. M. Anderson, J. Org. Chem., 32, 2583 (1967).

R_m Values of Phenols. Their Relationship with Log P Values and Activity

G. L. Biagi,* O. Gandolfi, M. C. Guerra, A. M. Barbaro, and G. Cantelli-Forti

Istituto di Farmacologia e Farmacognosia, Università di Bologna, Italy. *Received December 31, 1974*

The experimental R_m values for a series of phenols were obtained by a reversed-phase TLC system. The extrapolation from a range of linear relationship between experimental R_m values and acetone concentration provided a set of extrapolated R_m values. This were used for studying the relationship between structure and activity in vitro and in vivo. The possibility to obtain by means of the extrapolation technique the R_m values in a standard system for several series of chemotherapeutic agents is pointed out.

The usefulness of the R_m values in studying the correlation of chemical structure with biological activity was shown with bis(dichloroacetamides) and vitamin K analogs,[1] penicillins and cephalosporins,[2] testosterone esters,[3] and sulfonamides.[4] The chromatographic R_m values, as an expression of the lipophilic character of molecules, were shown to be correlated with the Hansch π values in series of testosterone esters[3] and sulfonamides.[4]

In extending the use of a chromatographic technique for the determination of R_m values, we turned our attention in this paper to a series of phenols. In particular the purpose of the present work was to show that the present chromatographic technique can provide the R_m values of phenols in a standard system, where they can be compared with other series of chemotherapeutic agents. Another aspect of this paper was to further point out the possibility of relationships between partition data in different systems. Finally some experiments were carried out in order to give an additional contribution to the study of structure–activity relationships of phenols.

Experimental Section

Materials and Methods. The phenols reported in Table I were obtained from commercial sources. The experimental data of the present work were the result of at least 4–8 determinations.

R_m Values Determination. The chromatographic technique for the determination of the R_m values as an expression of the lipophilic character of molecules has previously been described.[5,6] The polar mobile phase was represented by veronal acetate buffer at pH 7.4 in various mixtures (v/v) with Me_2CO. The stationary nonpolar phase consisted of a silica gel G layer impregnated with silicone DC 200 (350 cSt) from Applied Sciences Laboratories. The concentration of Me_2CO in the mobile phase ranged from 5 to 55%. The developed plates were dried and sprayed with tetrazotized benzidine.[7] After a few minutes at 105° yellowish spots appeared on a white background. The R_m values were calculated by means of the formula

$$R_m = \log (1/R_f - 1)$$

Antibacterial Activity Determination. The phenolic compounds were assayed against *Staphylococcus aureus* 16 R by means of the turbidimetric method. Test tubes containing 4 ml of brain heart (Difco) liquid medium were added with 0.1 ml of EtOH solutions of phenols in order to obtain various serial dilutions. Finally 0.1 ml of an overnight culture of *Staph. aureus* was added to each tube. Minimal inhibitory concentrations (MIC) were determined after an 18-hr incubation at 37°. In Table IV the MIC's are expressed as log $1/C$ values where C is the average molar concentration which prevents the growth of microorganisms.

Hemolytic Activity Determination. The details of the procedure have already been described.[3] A volume of 3.8 ml of phosphate-buffered saline was added to 0.2 ml of a rat erythrocytes suspension. This was obtained by suspending the erythrocytes, separated from 0.8 ml of rat blood, in phosphate-buffered saline to a volume of 8 ml. EtOH solutions of the test compounds were finally added to the system in 1- to 10-μl amounts. Control and test suspensions were incubated for 3 hr at 37°. After incubation all suspensions were centrifuged and the optical densities of supernatants measured at 540 mμ in the Bausch and Lomb colorimeter. The results were expressed as percent of total hemolysis provoked by distilled H_2O by means of the formula

$$\frac{(\text{OD of sample} - \text{OD of EtOH control}) \times 100}{\text{OD of distilled } H_2O \text{ control}}$$

The linear relationship between the concentration of phenolic compounds and the percent of total hemolysis allowed calculation of the molar concentration of each compound provoking a 50% hemolysis. This is expressed in Table IV as log $1/C$. A "t" test showed the statistical significance of the above linear relationship for each compound.

LD_{50} Determination in Mice. Albino mice, weighing approximately 20–25 g each, were used. Five animals were injected at each dose level. Dimethyl sulfoxide (Me_2SO) solutions of test compounds were administered intraperitoneally. All mice were observed over a period of 24 hr for their death rate. The LD_{50} values were calculated by means of the Spearman and Kärber method.[8] In Table IV are reported the log $1/C$ values where C is the LD_{50} expressed as M/kg. Some preliminary experiments were carried out in order to rule out the possibility of toxic effects of Me_2SO.

Results

R_m and Log P Values. The early results of the chromatographic work had shown that the test compounds did not move from the starting line when the mobile phase was represented only by veronal acetate buffer. The addition of acetone was necessary in order to obtain longer migrations and therefore more reliable R_m values. The plots of the R_m values vs. the composition of the mobile phase showed that for each compound there was a range of linear relationship between R_m values and acetone concentration. The straight lines of Figure 1, where the data of only a few compounds