have the tool necessary to systematically uncover subtle favorable or unfavorable interactions due to molecular features which, in the first approximation, appear to be additive. A new degree of sophistication in QSAR is now possible. We believe that indicator variables will play an increasingly important role in structuring large sets of congeners.

References and Notes

(1) This investigation was supported by U.S. Public Health Service Research Grant No. CA-11110 from the National Cancer Institute.
(2) C. Silipo and C. Hansch, *J. Am. Chem. Soc.*, **97**, 6849 (1975).
(3) M. Yoshimoto and C. Hansch, *J. Med. Chem.*, **19**, 71 (1976).
(4) (a) W. T. Ashton, F. C. Walker III, and J. B. Hynes, *J. Med. Chem.*, **16**, 694 (1973); (b) W. T. Ashton and J. B. Hynes, *ibid.*, **16**, 1233 (1973); (c) J. B. Hynes, W. T. Ashton, D. Bryansmith, and J. H. Freisheim, *ibid.*, **17**, 1023 (1974); (d) J. B. Hynes and W. T. Ashton, *ibid.*, **18**, 263 (1975); (e) J. B. Hynes and C. M. Garrett, *ibid.*, **18**, 632 (1975).
(5) C. Hansch, A. Leo, S. H. Unger, K. H. Kim, and E. J. Lien, *J. Med. Chem.*, **16**, 1207 (1973).
(6) G. M. Furnival and R. W. Wilson, Jr., *Technometrics*, **16**, 499 (1974).
(7) C. Hansch and C. Silipo, *J. Med. Chem.*, **17**, 661 (1974).
(8) C. Hansch, C. Silipo, and E. E. Steller, *J. Pharm. Sci.*, **64**, 1186 (1975).

Dependence of Hydrophobicity of Apolar Molecules on Their Molecular Volume

A. Leo,* Corwin Hansch, and Priscilla Y. C. Jow

Department of Chemistry, Pomona College, Claremont, California 91711. *Received October 9, 1975*

Cavity size is the primary determinant of the partition coefficient (P) of apolar solutes between octanol and water. Although the energy of cavity formation would be expected to be related to cavity *area*, older methods of area calculation give a poorer correlation with log P than does *volume*. Apolar solutes clearly fall into two classes based on their log P/volume relationship, the distinction possibly being whether the solute exposes mostly hydrogen atoms or unbonded electrons.

The great importance of hydrophobic interactions in biological and medicinal chemistry has stimulated much interest in delineating the determinants of this molecular property. Recently, a number of moderately successful efforts have been made to relate hydrophobicity of nonpolar solutes to the surface area or volume of the cavity necessary to contain them.[1-5] We wish to show that, for relatively apolar solutes,[6] the *nature* of the solutes surface and the molecular volume determine hydrophobicity, at least as it is measured by partitioning between octanol and water (coefficient = P).

Harris et al.[2] showed that the surface area of hydrocarbons, as determined by the attachment of small spheres representing water, correlated well with water solubility. Tanford[3] has refined this type of reasoning. Hermann,[4a] in a more sophisticated approach, showed that log S (S = molar solubility in water) of hydrocarbons was linearly related to the surface of a cavity which included the radius of the first layer of water molecules and that aromatic rings were more soluble but displayed nearly the same slope as the aliphatic series. Amidon et al.[4c] simplified Hermann's method of surface area calculation and extended it to include alcohols by inclusion of –OH surface area values. Moriguchi[5] has reported on factoring log P (octanol–water) into a hydrophilic effect of a polar group and a hydrophobic effect due to the free molar volume. He used Quayle's atomic parachor, molar refraction, and Exner's molar volume as parameters relating to molar volume.

We report a study which relates log P (octanol–water) to molecular volume as measured directly from CPK models. Our results show that a variety of apolar solutes[6] clearly separate into two classes (see Figure 1): one class has a surface of covalently bonded hydrogen atoms; the other class has a surface of noncovalently bonded electrons. Alkanes and $Si(CH_3)_4$ constitute the former class, while the rare gases, perhalogenated alkanes, aromatic hydrocarbons, and haloaromatic compounds comprise the latter (Table I).

Equations 1a and 1b show the relationship between log P for the two classes of solutes and surface area as calculated by Bondi.[1] Equations 2a and 2b show that Bondi molar volume is clearly more closely related to log P than is Bondi area, while eq 3a and 3b show the volume taken directly from CPK molecular models is by far the best parameter for class II.

$$\log P_{\mathrm{I}} = 0.513\ (\pm 0.644) + 0.0207\ (\pm 0.005)\ (\text{Bondi area}) \tag{1a}$$

n	r	s
11	0.952	0.206

$$\log P_{\mathrm{II}} = -0.458\ (\pm 0.220) + 0.0263\ (\pm 0.002)\ (\text{Bondi area}) \tag{1b}$$

n	r	s
26	0.985	0.219

$$\log P_{\mathrm{I}} = 0.568\ (\pm 0.282) + 0.0283\ (\pm 0.003)\ (\text{Bondi volume}) \tag{2a}$$

n	r	s
11	0.990	0.096

$$\log P_{\mathrm{II}} = 0.007\ (\pm 0.176) + 0.0288\ (\pm 0.002)\ (\text{Bondi volume}) \tag{2b}$$

n	r	s
26	0.987	0.204

$$\log P_{\mathrm{I}} = 0.728\ (\pm 0.249) + 0.0281\ (\pm 0.003)\ (\text{CPK volume}) \tag{3a}$$

n	r	s
11	0.991	0.090

$$\log P_{\mathrm{II}} = -0.026\ (\pm 0.076) + 0.0279\ (\pm 0.001)\ (\text{CPK volume}) \tag{3b}$$

n	r	s
26	0.998	0.088

The number of data points used is represented by n, r is the correlation coefficient, s the standard deviation, and

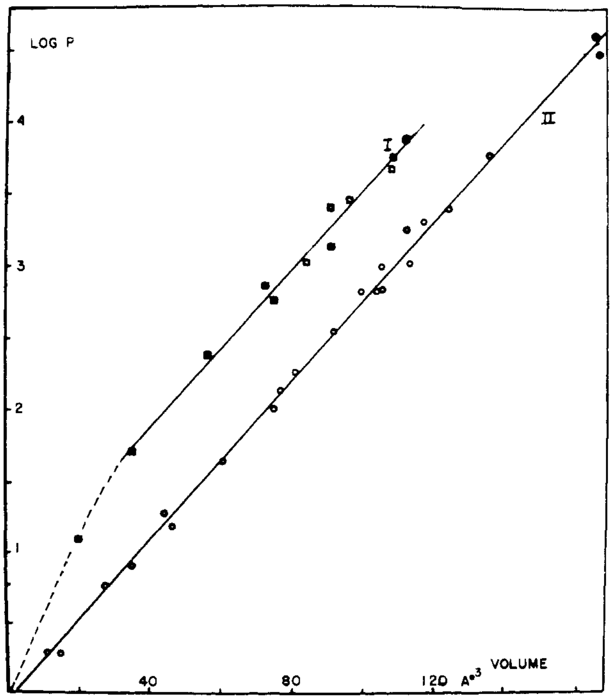

Figure 1. Log *P*-CPK volume relationship for class I and class II solutes.

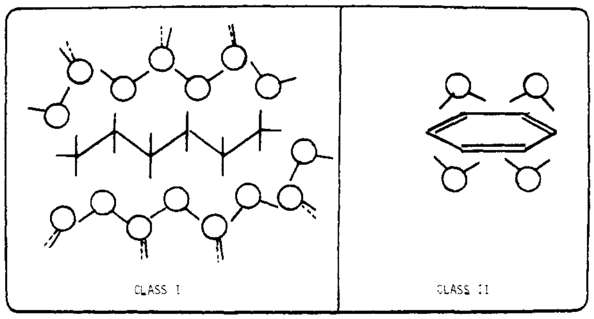

Figure 2. Hypothetical "reversible lining" of cavities for class I and class II solutes.

the numbers in parentheses the 95% confidence intervals. Methane is poorly fit (>2 × standard deviation) by all three parameters and was omitted in deriving eq 1a–3a.

The simplest distinction between class I and class II solutes is that of a surface rich in either protons or electrons. The electronegativity of atoms (e.g., the rare gases vs. the perhaloalkanes) has little if anything to do with hydrophobicity. One possible interpretation is that the aqueous cavities in the two classes of solutes are complementary; i.e., for class I solutes, the linings show an excess of oxygen atoms facing inward, while hydrogen predominates in the cavity surface for class II solutes (see Figure 2). Of course, known aqueous clathrate structures are *not* lined with an excess of either H or O atoms (i.e., those which are not participating in hydrogen bonding) but this objection is not necessarily fatal to the "reversible sweater" hypothesis because the partitioning experiments were not carried out at temperatures or solute concentrations that encourage clathrate formation.[7] The partially H-bonded water, which is thought to exist at 25 °C, can form chains or rings and these might twist to expose either oxygen or hydrogen, depending upon the nearby solute which needed accommodation (see Figure 2). The difference in log *P*/volume between classes I and II suggests that the smallest H-lined "hemiclathrate" can be formed with a lower free-energy expenditure than the corresponding O-lined one but that the enthalpy change (from the very weak H bonds) in any subsequent enlargement of the cavity is about the same for both types and insufficient to offset the entropy of orientation. Until it is supported by additional experimental evidence, the "reversible sweater" hypothesis must be considered highly speculative.

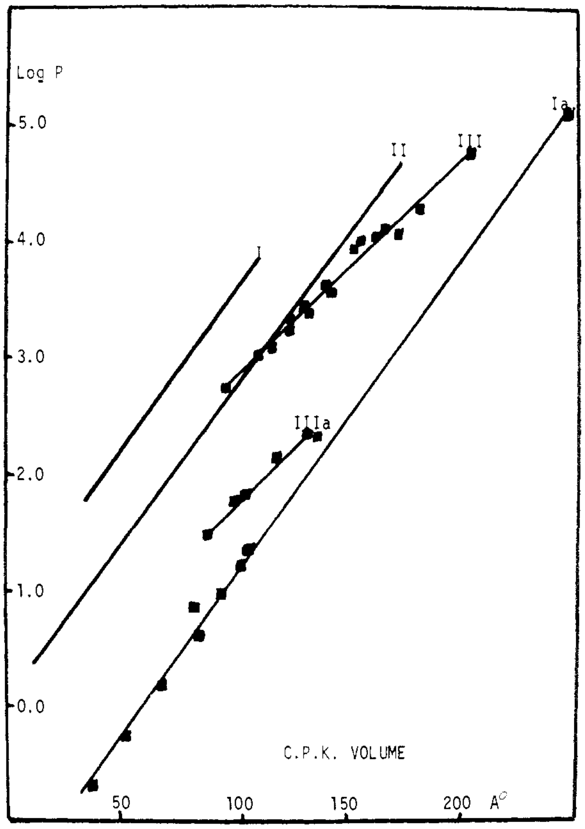

Figure 3. Log *P* of mixed aryl-alkyl solutes less sensitive to volume (classes I and II are the same as in Figure 1).

On the other hand, the lower log *P* of class II solutes could be ascribed to poorer accommodation by the octanol phase if, for instance, the alkyl chains of octanol could accommodate alkyl solutes more readily than aryl rings. However, this appears unlikely in view of the fact that the log *P* of benzene in the system hexane–water is very nearly as great as it is between itself and water (log $P_{C_6H_6}$ in C_6H_6–H_2O = 2.69; log $P_{C_6H_6}$ in C_6H_{14}–H_2O = 2.45). In any event, since the work of cavity formation in water is much greater than in octanol, any difference in cavity surface orientation ought to show a larger effect in that phase.

It is seen in Figure 3 that the log *P* of mixed alkyl-aryl solutes does not increase as rapidly with increased solute volume as does each pure solute type. Class III solutes obey the relationship expressed in eq 4 and are listed in Table I. There is no overlap in slopes of eq 3a or 3b and eq 4 when the 95% confidence intervals is considered. For different reasons, two solutes were omitted in deriving eq 4. Mesitylene (deviation = −0.264) exhibits a sizable dependence of log *P* on concentration which indicates it may "stack" in the aqueous phase. *o*-Chlorotoluene (deviation = 0.196) may contain a polar effect not present

$$\log P_{\mathrm{III}} = 1.033\ (\pm 0.293) + 0.0188\ (\pm 0.002)\ (\mathrm{CPK\ volume}) \qquad (4)$$

n	*r*	*s*
18	0.979	0.116

Table I. Solute Parameters

	Solute	Log *P*	Bondi area[a]	Bondi volume[a]	CPK volume[b]	Deviation	Difference
1	CH_4	1.09	48.15	28.42	20.54	±0.21	−7.88
2	CH_3CH_3	1.81	70.40	45.41	37.34	0.40	−8.07
3	$CH_3CH_2CH_3$	2.36	92.81	62.41	56.80	0.62	−5.61
4	Isobutane	2.76	115.06	79.39	75.68	1.80	−3.71
5	Butane	2.89	115.22	79.39	75.35	0.80	−4.04
6	Cyclopentane	3.00	102.60	83.03	81.68	0.95	−1.35
7	Neopentane	3.11	140.80	96.36	91.67	2.30	−4.69
8	Pentane	3.39	137.60	96.36	90.36	1.00	−6.00
9	Cyclohexane	3.40	125.02	100.00	95.48	1.40	−4.52
10	$(CH_3)_4Si$	3.75	154.00[c]	116.83[c]	107.12[d]	2.50	−9.71
11	2,2-Dimethylbutane	3.82	163.21	113.11	106.96	2.20	−6.15
12	2,3-Dimethylbutane	3.85	159.72	113.11	110.67	2.30	−2.44
			Class II				
13	He	0.28	24.63	11.49	11.49[e]		
14	Ne	0.28	29.80	15.30	15.30[e]		
15	O_2	0.65	33.50	21.58	21.58[e]		
16	N_2	0.67	40.70	26.23	26.23[e]		
17	Ar	0.74	44.41	27.83	27.83[e]		
18	Kr	0.89	51.28	34.52	34.52[e]		
19	CF_4	1.18	76.36	45.37	46.40		1.03
20	Xe	1.28	58.63	42.21	42.21[e]		
21	CF_3Cl	1.65	87.49	55.73	60.15	0.50	4.42
22	CF_3CF_3	2.00	114.54	78.65	75.29	0.80	−3.36
23	CF_2Cl_2	2.16	98.62	66.09	73.54	0.95	7.45
24	Benzene	2.13	99.78	80.33	76.28	0.60	−4.05
25	C_6H_5F	2.27	106.75	85.74	80.75	0.75	−4.99
26	$CFCl_3$	2.53	109.75	76.45	91.93	1.10	15.48
27	$CClF_2CClF_2$	2.82	136.80	91.54	103.52	2.20	11.98
28	CCl_4	2.83	120.88	86.82	105.68	1.90	18.86
29	C_6H_5Cl	2.84	118.55	96.03	98.11	1.05	2.08
30	C_6H_5Br	2.99	123.86	101.21	104.55	1.50	3.34
31	$C_6H_5CF_3$	3.01	145.76	111.52	112.84	2.60	1.32
32	C_6H_5I	3.25	130.17	108.72	111.34	2.40	2.62
33	Naphthalene	3.30	140.10	122.90	115.88	1.05	−7.02
34	*m*-$C_6H_4Cl_2$	3.38	148.44	107.59	124.11	2.60	16.52
35	*p*-$C_6H_4Cl_2$	3.39	148.44	107.59	123.09	2.50	15.50
36	*m*-$C_6H_4Br_2$	3.75	159.07	122.13	134.67	2.70	12.54
37	Phenanthrene	4.60	180.00	165.29	164.28	1.80	−1.01
38	Anthracene	4.45	180.00	165.29	165.31	1.50	0.02
			Class III				
39	Toluene	2.69	123.69	98.81	91.67	1.80	−7.14
40	*o*-Xylene	3.12	147.60	117.33	115.10	2.40	−2.20
41	*p*-Xylene	3.15	147.60	117.33	111.24	2.30	−6.09
42	Ethylbenzene	3.15	146.10	115.79	113.30	2.20	−2.49
43	*m*-Xylene	3.20	147.60	117.33	118.45	2.30	1.15
44	*m*-Chlorotoluene	3.28	142.80	114.55	121.54	2.40	6.99
45	*p*-Chlorotoluene	3.33	142.80	114.55	118.45	2.00	3.90
46	Indan	3.33	132.81	120.11	121.54	2.40	1.43
47	*o*-Chlorotoluene	3.42	142.80	114.55	116.39	2.20	1.84
48	Mesitylene	3.42	171.51	135.85	140.85	3.00	5.00
49	Tetrahydronaphthalene	3.45	155.22	137.09	137.00	2.80	−0.09
50	*n*-Propylbenzene	3.62	168.51	132.77	137.00	2.75	4.23
51	Isopropylbenzene	3.66	168.35	132.77	136.20	2.90	3.43
52	1,2,4,5-Tetramethylbenzene	4.00	195.42	154.37	147.80	2.90	−6.57
53	*tert*-Butylbenzene	4.11	194.09	149.74	149.90	3.10	0.16
54	Diphenylmethane	4.14	199.39	169.20	172.30	3.20	3.10
55	Fluorene	4.18	165.19	158.08	164.30	3.00	6.22
56	9,10-Dihydroanthracene	4.29	187.60	175.06	182.00	3.70	6.95
57	Hexamethylbenzene	4.31	243.30	191.41	168.15	2.90	−23.26
58	Dibenzyl	4.79	221.80	186.18	209.90	4.50	15.72

[a] See ref 1. [b] Corey–Pauling–Koltun models were wrapped with stretchable polyethylene and sealed with tape and the volume was determined by displacement of water. When care was taken not to entrap excess air or to pull the wrap down into the grooves between atoms, the reproducibility of the entire determination (wrapping and displacement) was ±2%. Using the scale of 1.25 cm/Å^3, the conversion factor would be ml(H_2O) × 0.512 = Å. Since room temperature tap water was used, the factor was g(H_2O) × 0.515. [c] Based on neopentane with radius increased as carbon to silicon. [d] Used phosphorus atom with long bonds; CPK red bonds for Si. [e] CPK atoms not available; used van der Waals volumes from Bondi.

in the meta and para derivatives. Adding these two solutes to eq 4 leaves the slope and intercepts essentially unchanged but raises the standard deviation to 0.14. On the basis of the "reversible sweater" hypothesis, one would predict that there would be less order in the water at the "seams" of the two sweater types (see Figure 4) and that the greater entropy could result in a lower effective log *P* for a given volume.

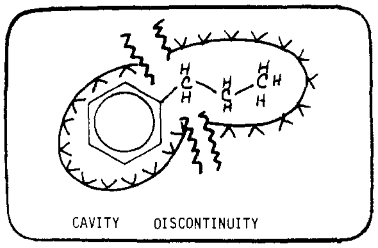

Figure 4. Hypothetical "seam" between unlike linings.

Table II. Solute Volumes

	Solute	Log *P*	CPK volume[a]
	Class Ia		
1	Methanol	−0.77	31.62
2	Ethanol	−0.31	50.68
3	Propanol	0.26	68.96
4	Isobutyl alcohol	0.65	85.39
5	Butanol	0.94	85.28
6	Isopentanol	1.16	99.10
7	Neopentanol	1.34	106.14
8	Pentanol	1.56	107.50
9	Dodecanol	5.12	244.60
	Class Ib		
10	Fluoromethane	0.51	26.34
11	Chloromethane	0.91	41.02
12	Bromomethane	1.19	47.07
13	Chloroethane	1.43	57.37
14	Iodomethane	1.51	58.56
15	Bromoethane	1.61	65.51
16	Chlorobutane	2.39	93.63
	Class IIIa		
17	Phenol	1.48	87.04
18	*p*-Methylphenol	1.94	105.58
19	3-Methylphenol	1.96	108.67
20	3,4-Dimethylphenol	2.28	121.03
21	3,5-Dimethylphenol	2.35	131.30
22	3-Ethylphenol	2.40	129.26

[a] See footnote *b*, Table I.

Since compounds in class Ia fall on a line parallel to those in class I (see Figure 3), it is apparent that a hydroxyl group produces a nearly constant hydrophilic effect regardless of the saturated alkyl fragment to which it is attached. Assuming this also applies to an aromatic –OH, it can be seen that the alkyl-substituted phenols (class IIIa, Figure 3) confirm the special behavior of mixed alkyl–aryl structures (Table II).

The expressions for classes Ia and IIIa are given by eq 5 and 6, respectively.

$$\log P_{\text{Ia}} = -1.612\ (\pm 0.185) + 0.028\ (\pm 0.002)\ (\text{CPK volume}) \qquad (5)$$

n	*r*	*s*
9	0.998	0.119

$$\log P_{\text{IIIa}} = -0.272\ (\pm 0.492) + 0.021\ (\pm 0.004)\ (\text{CPK volume}) \qquad (6)$$

n	*r*	*s*
6	0.989	0.058

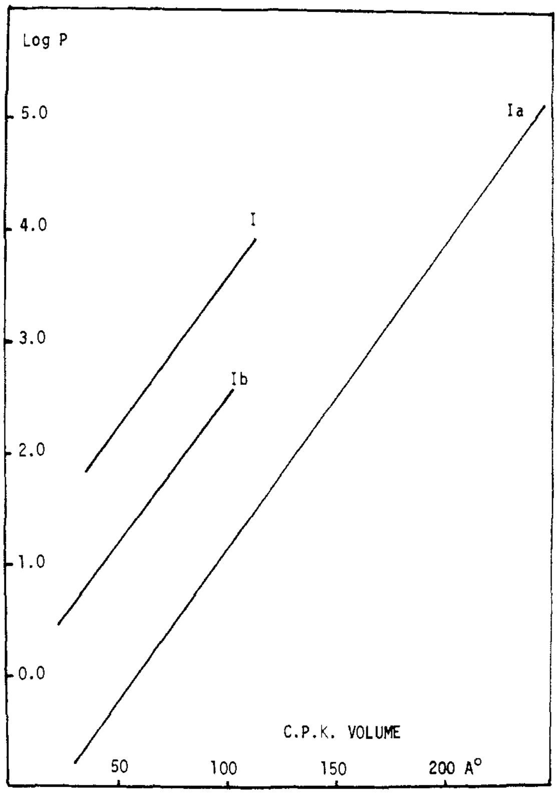

Figure 5. Polar functions on alkyl chains: Ia, alkanols and alkylamines; Ib, alkyl halides.

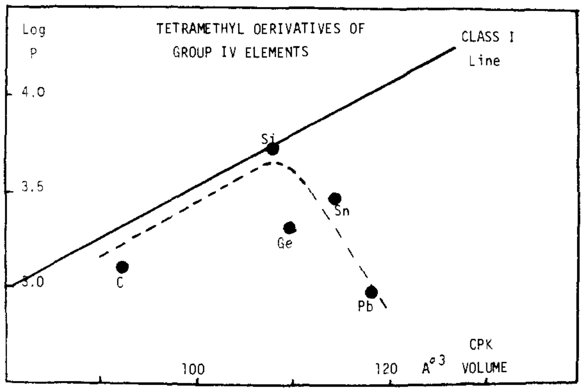

Figure 6. Limits of methyl group shielding.

$$\log P_{\text{Ib}} = -0.185\ (\pm 0.156) + 0.028\ (\pm 0.003)\ (\text{CPK volume}) \qquad (7)$$

n	*r*	*s*
7	0.997	0.053

The introduction of a localized dipole, such as is present in alkyl monohalides, reduces log *P* but not to the same extent as do the H-bonding groups –OH and $-NH_2$ (see Figure 5, class Ib and eq 7). It is assumed in assigning the perhaloalkanes to class II that the bond dipoles are effectively shielded.[8] In this regard, it is interesting to note the trend in the log *P* values of the tetramethyl derivatives of the group IV elements C, Si, Ge, Sn, and Pb. Of course, there is very little dipole in the C–C bonds in neopentane but the electronegativity difference between carbon and silicon (0.76), germanium (0.48), tin (0.78), and lead (0.95) should produce an appreciable bond dipole in these instances.[9]

Neopentane and tetramethylsilane fit the log *P*/volume relation as class I solutes but there is a drastic decline in the following homologues (see Figure 6). It would appear from the CPK models that the methyl groups are no longer able to shield the dipole in the series above germanium and that the aqueous solvent shell can interact with the solute in a manner not possible in the lower members.

It is important to note the high collinearity between surface area and volume in eq 1a and 1b (r^2 = 0.91 and

0.98, respectively). It appears that this also holds for the more recent methods of calculating surface area.[4b,c] It should be emphasized that, regardless of whether it develops that it is a volume or a surface area calculation which can be refined to give the best relation with log *P*, we will still be far from a computerized system of log *P* calculation by merely summing the sizes of assorted fragments. The surface area of fragments can differ appreciably, depending on how they are connected and an even greater log *P* variation can arise by differences in character of the solute surface, depending on the order in which the fragments appear in isomeric solutes or whether or not they are in position to screen a dipole or H-bonding group.

If we are to untangle the effects on hydrophobic bonding caused by dipoles—localized, diffuse, or partly shielded—and by various types of hydrogen bonds, it is evident that we must first have a clear understanding of the very basic role that is played by solute size.

Acknowledgment. This investigation was supported by the Pomona College Medicinal Chemistry Project and by U.S. Public Health Service Research Grant No. CA-11110 from the National Cancer Institute.

References and Notes

(1) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).

(2) M. J. Harris, T. Higuchi, and J. H. Rytting, *J. Phys. Chem.*, **77**, 2694 (1973).

(3) C. Tanford, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 2925 (1974).

(4) (a) R. B. Hermann, *J. Phys. Chem.*, **76**, 2754 (1972); (b) *ibid.*, **79**, 163 (1975); (c) G. L. Amidon, S. H. Yalkowsky, and S. Leung, *J. Pharm. Sci.*, **63**, 3225 (1974).

(5) I. Moriguchi, *Chem. Pharm. Bull.*, **23**, 247 (1975).

(6) This excludes solutes which hydrogen bond or which contain an exposed localized dipole, such as CH_3Cl.

(7) While most solutes with polar functional groups show very little temperature dependence of log *P*, argon and some alkanes show a marked decrease at 0 °C. This is being investigated further.

(8) It should be noted that it is the *localized* bond dipole that is important in determining log *P*, not the measured *overall* dipole; otherwise, *o*-dichlorobenzene would have a much lower log *P* than *p*-dichlorobenzene (ortho, $\mu = 1.58$, log $P = 3.38$; para, $\mu = 0.0$, log $P = 3.39$).

(9) F. A. Cotton and G. Wilkinson, "Advanced Inorganic Chemistry", 2nd ed, Interscience, New York, N.Y., 1966, p 103.

Direct Measurement of Octanol-Water Partition Coefficients by High-Pressure Liquid Chromatography

Margaret S. Mirrlees, Stuart J. Moulton, Christopher T. Murphy, and Peter J. Taylor*

Imperial Chemical Industries Limited, Pharmaceuticals Division, Alderley Park, Macclesfield, Cheshire, SK10 4TG England.

Received June 9, 1975

A technique is presented for the direct measurement of octanol-water partition coefficients by HPLC. The method involves running solutes in octanol-saturated water as the mobile phase against water-saturated octanol entrained on an inert support. Log *P* correlates linearly with log t_c for a number of standards. The measurable range in log *P* (so far) is −0.3 to +3.7. A critical review of chromatographic methods in Hansch analysis is given.

Since the pioneering work of Meyer and Overton,[1] it has gradually become evident that nonspecific distribution plays an important part in determining biological activity. There are two parallel but essentially dissimilar ways in which this understanding may be exploited. The first is to attempt a correlation between biological potency and some thermodynamic quantity that may reasonably be imagined to reflect the distribution process, most commonly partition coefficient. For this purpose, any type of partition coefficient will do; success is the only criterion. The weakness of this approach is its inability to relate one such correlation to another. It was partly to overcome this weakness that Hansch explored the possibility of deriving additive group contributions to the partition coefficient that would allow this to be calculated for unsynthesized compounds (his other main contribution, the introduction of a multiparameter approach to the rationalization of biological activity,[2-4] does not concern us here). Octanol-water was chosen as the standard reference system for reasons connected with the easy availability of pure octanol and its intermediate degree of polarity, arguably much closer in its properties than a hydrocarbon to the biophase.[5] These group contributions or "π values" are

$$\log P_{\mathrm{R}} - \log P_{\mathrm{H}} = \pi_{\mathrm{R}} \qquad (1)$$

defined by eq 1, in which P_{H} and P_{R} are, respectively, the octanol-water partition coefficient for the putative parent compound of a series and for that containing the substituent R. Despite evidence[6] that π values are not always additive, usually for steric reasons but sometimes through electronic and hydrogen bonding effects, there are now many series of compounds in which π appears to provide most of the reason for variation in biological potency.[2-5]

Partition coefficients are generally determined by some variant on the traditional shake-flask method.[7] This method is slow, tedious, often wasteful, and demanding in the standard of purity it requires. Consequently, there have been many attempts to develop chromatographic methods, to which in principle none of the above objections need apply. This is possible because some (not all) forms of chromatography are dominated by partitioning processes. It was shown by Martin and Synge[8] that R_f relates

$$P = \text{constant}\ (1/R_f - 1) \qquad (2)$$

$$R_{\mathrm{M}} = \log\ (1/R_f - 1) \qquad (3)$$

to *P* according to eq 2; it follows that R_{M}, defined by Bate-Smith and Westall[9] according to eq 3, is linearly related to log *P*, and that ΔR_{M} is therefore analogous to π. (Green and Marcinkiewicz[10] had already shown, before Hansch's introduction[2] of π, that ΔR_{M} is additive for a number of benzenoid compounds.) The first exploitation of this relationship in a biological context is due to Boyce and Milbarrow,[11] who showed a relation between the