



ELSEVIER

Journal of Chromatography A, 685 (1994) 203–211

JOURNAL OF
CHROMATOGRAPHY A

Hydrogen bonding XXXV. Relationship between high-performance liquid chromatography capacity factors and water–octanol partition coefficients

Michael H. Abraham^{a,*}, Harpreet S. Chadha^a, Albert J. Leo^b^aDepartment of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK^bMedicinal Chemistry Project, Pomona College, Claremont, CA 91711, USA

First received 31 March 1994; revised manuscript received 19 July 1994

Abstract

The solvation equation $\log SP = c + rR_2 + s\pi_2^H + a\Sigma\alpha_2^H + b\Sigma\beta_2^H + vV_x$ has been applied to reversed-phase HPLC capacity factors, as $\log k'$, for solutes on a C_{18} bonded phase, with various water–methanol mobile phases, using data by Yamagami and Takao. Here, SP is a property for a series of solutes in a fixed solvent system, and the explanatory variables are solute descriptors as follows: R_2 is an excess molar refraction, π_2^H is the solute dipolarity/polarizability, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are the solute overall or effective hydrogen-bond acidity and basicity, and V_x is the McGowan characteristic volume; c , r , s , a , b and v are constants. It is shown that the blend of factors that influence $\log k'$ in any given system is not the same as that which influences $\log P_{\text{oct}}$. In particular, solute hydrogen-bond acidity considerably influences $\log k'$, but has no effect on $\log P_{\text{oct}}$. It follows that when $\log k'$ values are used to estimate $\log P_{\text{oct}}$, great care has to be taken to match the training set of solutes in the correlation equation, with the solutes for which $\log P_{\text{oct}}$ is to be determined.

1. Introduction

The use of reversed-phase high-performance liquid chromatography (RP-HPLC) to determine water–octanol partition coefficients is a well-established and widely practiced method [1–5]. Values of the RP-HPLC capacity factor, k' , are obtained for a training set of solutes with known water–octanol partition coefficients, P_{oct} , using a given stationary phase and a given mobile phase, and a correlation equation of the type

$$\log k' = o \log P_{\text{oct}} + q \quad (1)$$

is constructed. Then further measurements of $\log k'$ in the same system can be used to estimate P_{oct} for other solutes. How good the estimations are, depends not only on the choice of the RP-HPLC system, but also on the choice of the training set of solutes. Xie et al. [2] used a training set of phenols in order to estimate $\log P_{\text{oct}}$ for phenols, and obtained a good correlation between $\log P_{\text{oct}}$ from an RP-HPLC method, and $\log P_{\text{oct}}$ from the traditional shake-flask (sf) method. The RP-HPLC procedure of Xie et al. [2] differs slightly from that summarized as Eq.

* Corresponding author.

1, in that the capacity factors determined at different water–methanol mobile phase compositions were extrapolated to zero methanol composition, and the resulting $\log k'_w$ values used to calculate $\log P_{\text{oct}}$. For 21 phenolic solutes, the relationship between $\log P_{\text{oct}}$ by the sf method, and $\log P_{\text{oct}}$ by the RP-HPLC method is given by

$$\log P_{\text{oct},\text{sf}} = 0.0616 + 0.9861 \log P_{\text{oct},\text{RP-HPLC}}$$

$$n = 21, \rho = 0.9803, \text{S.D.} = 0.12, F = 468 \quad (2)$$

Here and elsewhere, n is the number of solutes, ρ is the correlation coefficient, S.D. is the standard deviation in the dependent variable, and F is the F -statistic. If a wide range of solutes is considered, the regression equation, either in terms of Eq. 1 or Eq. 2, is not so good. For example, Vallat et al. [5] determined $\log k'_w$ both directly using buffered water as the mobile phase, and by extrapolation from buffered water–methanol mobile phases, with a deactivated C_8 stationary phase. For 70 varied solutes they found

$$\log P_{\text{oct}} = 0.30 + 1.09 \log k'_{w,C_8}$$

$$n = 70, \rho = 0.9476, \text{S.D.} = 0.29, F = 599 \quad (3)$$

They attributed the rather poor correlation shown in Eq. 3 to the different hydrogen-bond properties of the RP-HPLC C_8 system and the water–octanol system, on the lines suggested by Kamlet and co-workers [6,7] several years ago. As might be expected, a plot of $\log P_{\text{oct}}$ vs. $\log k'_w$ revealed clear family-dependent lines [5]. Vallat et al. [5] also examined a novel octadecylpolyvinyl alcohol (ODP) stationary phase in a similar way, and obtained

$$\log P_{\text{oct}} = -0.06 + 0.83 \log k'_{w,\text{ODP}}$$

$$n = 40, \rho = 0.9823, \text{S.D.} = 0.23, F = 1048 \quad (4)$$

The ODP system seems to mimic hydrogen-bond and other properties of the water–octanol system rather well [5], and gives a better correla-

tion over a varied set of solutes. However, the RP-HPLC systems mostly used in the determination of $\log P_{\text{oct}}$ values are nearer the C_8 system, and in these cases, the training set of solutes will normally be chosen so as to resemble as closely as possible the solutes for which $\log P_{\text{oct}}$ is unknown. Unless the training set is carefully chosen, Eq. 1 may not hold. An elegant demonstration of this has recently been provided by Yamagami and Takao [4], who obtained $\log k'$ values on a typical C_{18} bonded stationary phase, Capcell Pak, with various water–methanol mixtures as the mobile phase. They used a variety of substituted furans and heterocyclic compounds (see Table 1), as a training set, and showed that Eq. 1 held only for restricted families of solute. Even when taking the two sets of solutes in Table 1 separately, a plot of $\log k'$ against $\log P_{\text{oct}}$ gave rise to families of lines, rather than to one single line, for any given water–methanol mobile phase. Yamagami and Takao [4] also obtained $\log k'_w$ values through extrapolation of $\log k'$ to zero methanol concentration using results from water–methanol mixtures containing 30–70% (v/v) methanol, but found again that a plot of $\log k'_w$ against $\log P_{\text{oct}}$ gave rise to families of lines. Although Yamagami and Takao [4] took the two sets of solutes in Table 1 separately, similar results are found if the two sets are treated together, as shown in Figs. 1–3. Note that solutes 32–35 in set B in Table 1 also occur in set A; we keep both because the $\log k'$ values given by Yamagami and Takao [4] are slightly different from one set to the other.

Yamagami and Takao [4] divided the solutes in set A in Table 1 into three groups of substituents, (i) H, alkyl, halogen, OMe and COMe, (ii) CO_2R and CONMe_2 , and (iii) CONHMe and CONH_2 . They suggested that solutes with hydrogen-bond acidic groups would undergo hydrogen-bonding more effectively with octanol (or rather wet octanol) than with the less basic stationary phase, so that $\log P_{\text{oct}}$ is larger than expected, or $\log k'$ is smaller than expected, by comparison with the other substituted furans. A similar argument was put forward to explain the deviant behaviour of indole and 3-car-

Table 1
Solute and their descriptors used in the calculations

No.	Solute	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x
<i>Set A</i>						
1	Furan	0.369	0.53	0.00	0.13	0.5363
2	2-Methylfuran	0.372	0.50	0.00	0.14	0.6772
3	2-Ethylfuran	0.361	0.50	0.00	0.14	0.8181
4	2-Methoxyfuran	0.408	0.76	0.00	0.25	0.7359
5	2-O-COMe-furan	0.570	1.20	0.00	0.58	0.8925
6	2-CO ₂ Me-furan	0.560	1.00	0.00	0.50	0.8925
7	2-CO ₂ Et-furan	0.560	1.00	0.00	0.50	1.0334
8	2-Br-5-CO ₂ Et-furan	0.785	1.13	0.00	0.45	1.0675
9	2-CONMe ₂ -furan	0.810	1.10	0.00	0.89	1.0745
10	3-CO ₂ Me-furan	0.477	0.84	0.00	0.46	0.8925
11	3-CO ₂ Et-furan	0.429	0.86	0.00	0.46	1.0334
12	3-CONMe ₂ -furan	0.710	0.94	0.00	0.98	1.0745
13	2-CONH ₂ -furan	0.910	1.20	0.10	0.83	0.7927
14	2-CONHMe-furan	0.850	1.15	0.36	0.76	0.9336
15	2-CONHEt-furan	0.800	1.15	0.36	0.81	1.0745
16	3-CONH ₂ -furan	0.810	1.11	0.49	0.67	0.7927
17	3-CONHMe-furan	0.750	0.87	0.36	0.82	0.9336
18	3-CONHEt-furan	0.700	0.89	0.36	0.85	1.0745
<i>Set B</i>						
19	Pyrazine	0.629	0.95	0.00	0.62	0.6342
20	2-CO ₂ Me-pyrazine	0.750	1.28	0.00	0.92	0.9904
21	2-Me-5-CO ₂ Me-pyrazine	0.750	1.28	0.00	0.96	1.1313
22	2-Me-6-CO ₂ Me-pyrazine	0.750	1.28	0.00	0.98	1.1313
23	2-Cl-5-CO ₂ Me-pyrazine	0.850	1.27	0.00	0.85	1.1128
24	2-MeO-6-CO ₂ Me-pyrazine	0.780	1.30	0.00	0.87	1.1900
25	Pyrimidine	0.606	1.00	0.00	0.65	0.6342
26	2-CO ₂ Me-pyrimidine	0.730	1.33	0.00	1.04	0.9904
27	5-CO ₂ Me-pyrimidine	0.730	1.33	0.00	0.82	0.9904
28	Pyridazine	0.670	0.85	0.00	0.81	0.6342
29	4-CO ₂ Me-pyridazine	0.790	1.18	0.00	0.97	0.9904
30	Benzene	0.610	0.52	0.00	0.14	0.7164
31	Methyl benzoate	0.733	0.85	0.00	0.46	1.0726
32	Furan	0.369	0.53	0.00	0.13	0.5363
33	2-CO ₂ Me-furan	0.560	1.00	0.00	0.50	0.8925
34	3-CO ₂ Me-furan	0.477	0.84	0.00	0.46	0.8925
34	2-Br-5-CO ₂ Et-furan	0.785	1.13	0.00	0.45	1.0675
36	N-Methylpyrrole	0.559	0.79	0.00	0.31	0.7180
37	2-CO ₂ Me-pyrrole	0.740	1.06	0.00	0.40	0.9936
38	Thiophene	0.687	0.57	0.00	0.15	0.6411
39	Benzofuran	0.888	0.83	0.00	0.15	0.9053
40	2-CO ₂ Me-benzofuran	1.080	1.31	0.00	0.46	1.2615
41	Indole	1.200	1.12	0.44	0.22	0.9460
42	2-CO ₂ Me-indole	1.320	1.45	0.12	0.52	1.3026
43	3-CO ₂ Me-indole	1.320	1.45	0.44	0.49	1.3026

bomethoxyindole in set B in Table 1. Whether or not such an explanation is quantitatively correct can only be ascertained through an analysis that

includes proper descriptors of solute hydrogen-bond acidity and basicity. This is the aim of the present work.

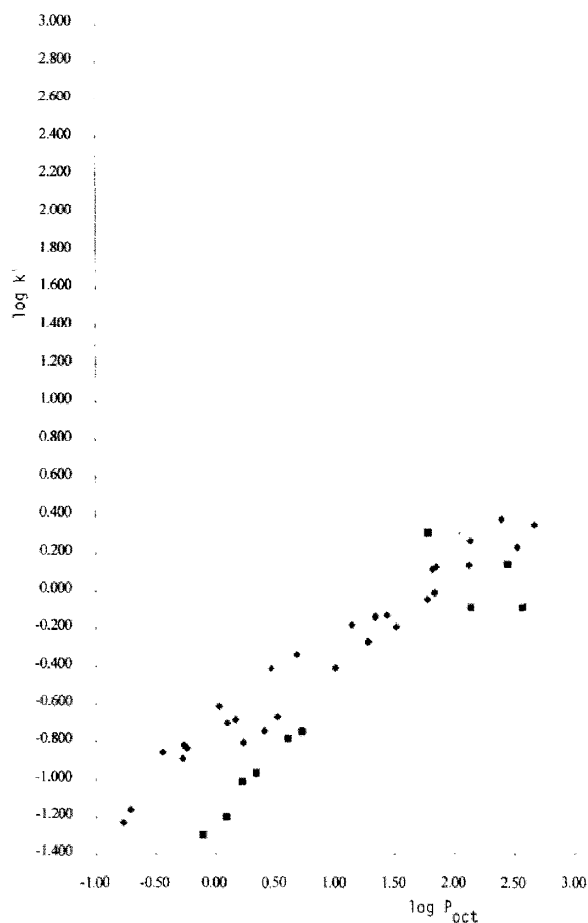


Fig. 1. Plot of $\log k'$ with mobile phase 70% methanol vs. $\log P_{\text{oct}}$ for all the entries in Table 1. Solutes with potential hydrogen bond acidic sites are denoted as ■.

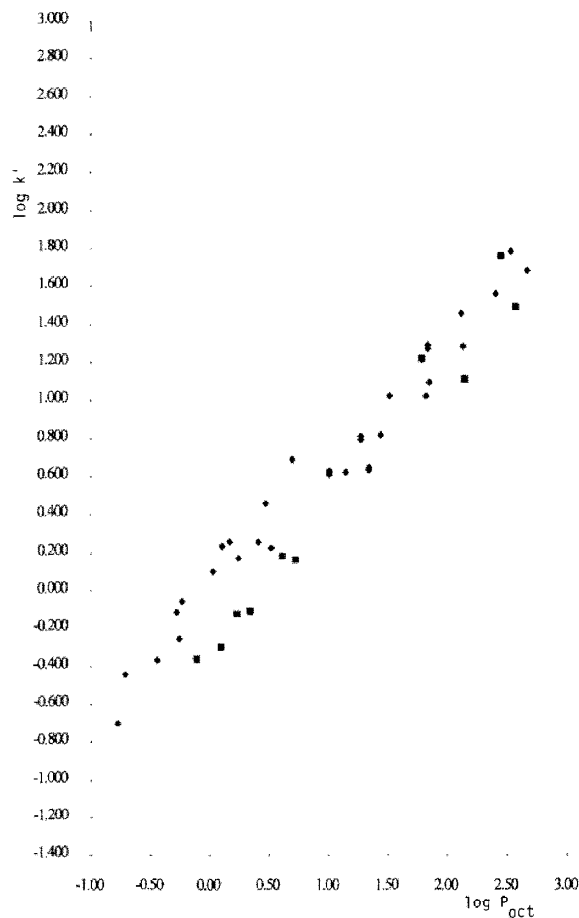


Fig. 2. Plot of $\log k'$ with mobile phase 30% methanol vs. $\log P_{\text{oct}}$ for all the entries in Table 1. Solutes with potential hydrogen bond acidic sites are denoted as ■.

2. Methodology

Our analysis is based [8] on the general solvation equation,

$$\log SP = c + rR_2 + s\pi_2^H + a\Sigma\alpha_2^H + b\Sigma\beta_2^H + vV_x \quad (5)$$

Here, SP is a property for a series of solutes in a fixed solvent system; in this work, SP will either be k' for solutes in a given RP-HPLC stationary phase–mobile phase system, or will be water–octanol partition coefficients for a series of solutes. The explanatory variables in Eq. 5 are

solute descriptors as follows: R_2 is an excess molar refraction that can be obtained for complicated solutes by simple addition of fragments [8]; π_2^H is the solute dipolarity/polarizability, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are the solute overall or effective hydrogen-bond acidity and basicity, and V_x is the McGowan characteristic volume that can be calculated from molecular structure [8]; c , r , s , a , b and v are constants. Hence for any new solute to be examined, three descriptors need to be assigned, viz. π_2^H , $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$. We have shown in some detail [9] exactly how these can be obtained, provided that a number of water–solvent $\log P$ values are available. Yamagami et

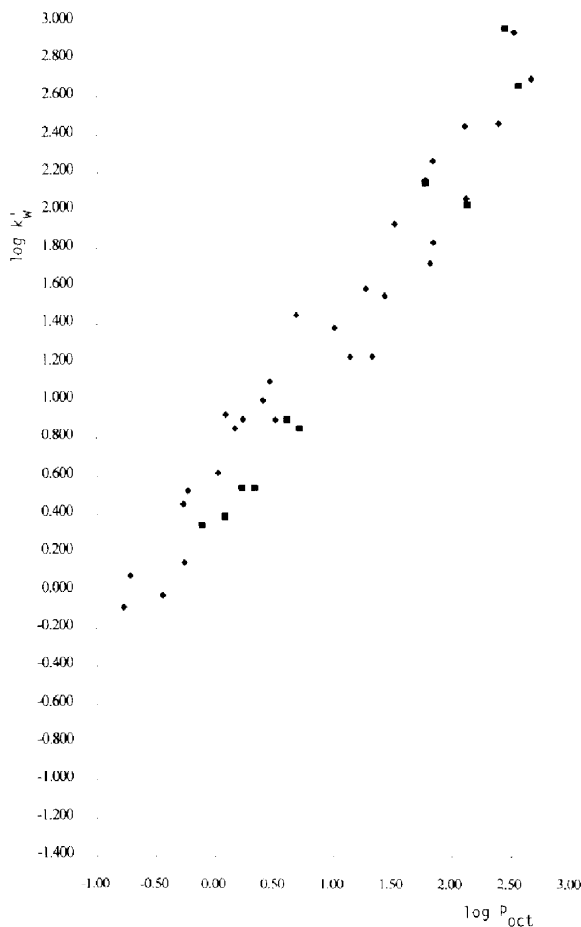


Fig. 3. Plot of $\log k'_w$ vs. $\log P_{\text{oct}}$ for all the entries in Table 1. Solutes with potential hydrogen bond acidic sites are denoted as ■.

al. [10] have determined such $\log P$ values in water–octanol, water–chloroform, water–octane and water–dibutyl ether systems for the pyrazines, pyrimidines, and pyridazines in Table 1, enabling us to calculate the required descriptors. We have already determined those for many of the other compounds [8,9], leaving assignments to be made for solutes 4–18. This was done using structurally related compounds as examples, and the total set of descriptors was tested using the $\log P_{\text{oct}}$ values listed [4]. The assigned descriptors are given in Table 1.

The general Eq. 5 has been applied to a large number of water–solvent partitions [9]. The most relevant is the water–octanol system, for

which the most recent [11] equation is

$$\begin{aligned} \log P_{\text{oct}} = & 0.088 + 0.562R_2 - 1.054\pi_2^{\text{H}} \\ & + 0.034\Sigma\alpha_2^{\text{H}} - 3.460\Sigma\beta_2^{\text{H}} + 3.841V_x \\ n = & 613, \quad \rho = 0.9974, \quad \text{S.D.} = 0.116, \quad F = 23162 \end{aligned} \quad (6)$$

The constants in Eq. 5 can be used to characterize the system in question; thus in Eq. 6 they provide information on the difference in properties of octanol (or wet octanol) and water. The r constant shows that solute polarizability favors octanol, and the s constant that solute dipolarity/polarizability favors water—hence octanol is itself more polarizable than water but water is more dipolar than octanol. The a constant is a measure of the difference in hydrogen-bond basicity between octanol and water (because acidic solutes interact with basic solvents), and shows that the two solvents have almost exactly the same basicity. On the other hand, the b constant indicates that water is much more acidic than octanol. Finally, the large positive v constant shows that octanol will interact with lipophilic solutes to a much greater extent than will water. There is no reason why Eq. 5 should not be applied in the same way to $\log k'$ values; indeed, Miller and Poole [12] have already shown this to be the case for a Baker-bond C_{18} stationary phase with various water–methanol mobile phases. We shall discuss these results later.

3. Results and discussion

The $\log k'$ values of Yamagami and Takao [4] refer to six water–methanol mixtures for the solutes in set A in Table 1, over the range 0–70% methanol, and to three water–methanol mixtures, from 30–70%, for solutes in set B in Table 1. There are not really enough solutes (18) in set A to carry out a complete analysis with Eq. 5, but since such a wide range of mixtures has been used, we thought it of interest to apply the equation, anyway. A summary of the resulting regression equations is in Table 2.

Table 2
Summary of regression Eq. 5 for solutes 1–18^a

Methanol (%)	<i>c</i>	<i>r</i>	<i>s</i>	<i>a</i>	<i>b</i>	<i>v</i>	ρ	S.D.	<i>F</i>
0	-0.84	0.16	-0.10	-1.11	-1.93	3.96	0.9979	0.04	570
5	-0.63	0.36	-0.36	-1.06	-2.11	3.67	0.9969	0.05	384
15	-0.49	0.47	-0.52	-0.96	-2.33	3.34	0.9969	0.06	381
30	-0.39	0.56	-0.68	-0.84	-2.33	2.86	0.9976	0.05	508
50	-0.42	0.51	-0.76	-0.68	-2.09	2.28	0.9989	0.03	1051
70	-0.67	0.32	-0.73	-0.60	-1.74	1.88	0.9975	0.04	482
Log k'_w	-0.17	0.80	-0.66	-1.11	-2.83	3.59	0.9926	0.10	160
30 ^b	-0.32	0.44	-0.58	-0.43	-2.52	3.34	0.9950	0.07	241
50 ^b	-0.41	0.34	-0.74	-0.31	-2.00	2.75	0.9880	0.11	103
70 ^b	-0.71	0.16	-0.44	-0.32	-1.59	2.12	0.9900	0.08	124
100 ^b	-0.98	0.01	-0.19	-0.03	-0.62	1.22	0.9320	0.08	17

^a With a Capcell Pak C₁₈ phase, ref. [4].

^b With a Bakerbond C₁₈ phase, and a different solute set, ref. [12].

Although we do not regard the regression equations as definitive, the constants make general chemical sense, and the interpretation of them follows exactly that of the water–octanol equation, above. The main factors that influence $\log k'$ are solute dipolarity/polarizability, hydrogen-bond acidity and hydrogen-bond basicity that all favor the mobile phase and lead to a reduction in $\log k'$, and solute volume that favors the stationary phase and leads to an increase in $\log k'$. The variation of the constants with mobile phase composition is interesting. As the methanol content increases, the difference between mobile phase and stationary phase basicity decreases and hence the magnitude of the *a* constant decreases; however the *b* constant reaches a maximum in magnitude at around 15–30% methanol. As the difference in hydrophobicity decreases with increasing methanol content, so does the *v* constant decrease. Surprisingly, the *s* constant increases in magnitude as the methanol content increases, so that the difference in dipolarity/polarizability between stationary phase and mobile phase becomes larger. However it must be noted that the stationary phase is not constant over the range of mobile phase composition. At 0% methanol, it will be saturated with water, at 100% methanol it will be saturated with methanol, and at intermediate compositions

it will be saturated with some solvent composition, not necessarily the same as that in the mobile phase. Although the equations of Miller and Poole [12] refer to a different C₁₈ phase, and a different (and also limited) solute set, they resemble quite closely those we find. A selection of the Miller–Poole equations is given in Table 2 for comparison. The trend of the constants in the Miller–Poole equations is practically the same as for the solute 1–18 set, and shows, additionally, that as the mobile phase approaches 100% methanol, the *a* constant becomes negligible.

The equation in $\log k'_w$ is not as good as the others, no doubt because the $\log k'_w$ values themselves are linearly extrapolated from the 30–70% methanol results. Because the linearity in the extrapolation does not hold at low methanol content, the $\log k'_w$ equation is not the same as that at 0% methanol, obtained from the actual experimental $\log k'$ values at 0%.

We can see from a direct comparison of Eq. 6 with the equations summarized in Table 2, why $\log k'$ values will not always be linear with $\log P_{oct}$. Such linearity will hold if the constants in the $\log k'$ equation are similar to those in Eq. 6, or if there is a simple relationship between them. Otherwise, Eq. 1 will not be expected to hold, as a generality, unless the solute set contains only functionally related solutes. Most importantly, if

a solute set contains solutes that differ widely in their hydrogen-bond acidity, then the linear Eq. 1 will break down, simply because $\log P_{\text{oct}}$ is not influenced by solute acidity, whereas $\log k'$ in most water–methanol mixtures is so influenced. The suggestion of Yamagami and Takao [4] that hydrogen-bond acidity of the solute leads to deviations from Eq. 1 is thus confirmed, but in order to quantify the effect we first prefer to analyse results for the total data set in Table 1.

A summary of the regression equations obtained using all 43 data points is in Table 3. There is a satisfactory agreement between the constants in Table 2 (set A) and those in Table 3 (sets A and B), except that the s constant in Table 3 decreases with increase in methanol content, more as expected. All the equations in Table 3 include a substantial a constant, so that our comments on the influence of acidic solutes, above, hold for this larger data set. We can show the effect of our analysis using the various descriptors in Eq. 5, by comparing a plot of observed and calculated $\log k'_w$ values (Fig. 4), with the plot shown in Fig. 3. Another, more quantitative, way is to add $\log P_{\text{oct}}$ as a descriptor to those in Eq. 5, leading to a method of analysis that we refer to as the $\log P_{\text{plus}}$ method.

$$\log SP = c + rR_2 + s\pi_2^H + a\Sigma\alpha_2^H + b\Sigma\beta_2^H + vV_x + o \log P_{\text{oct}} \quad (7)$$

We shall discuss the application of the $\log P_{\text{plus}}$ equation elsewhere, and now just apply it to the results for the 43 data points, as a comparison with the constants in Table 3. The application of

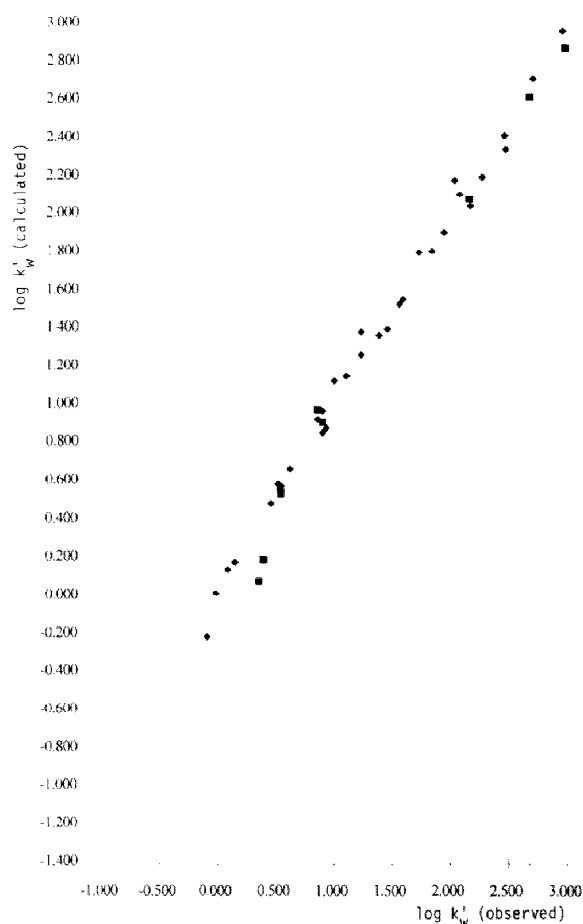


Fig. 4. Plot of observed vs. calculated $\log k'_w$ on Eq. 3 with the constants in Table 3. Solute with potential hydrogen bond acidic sites are denoted as ■.

Eq. 7 is not straightforward with the present data, because of cross-correlations of the original descriptors with $\log P_{\text{oct}}$. The best equations are summarized in Table 4; in all cases, the differ-

Table 3
Summary of regression Eq. 5 for solutes 1–43

Methanol (%)	c	r	s	a	b	v	ρ	S.D.	F
30	-0.43	0.36	-0.61	-0.89	-2.24	2.92	0.9979	0.05	1761
50	-0.45	0.18	-0.47	-0.76	-2.02	2.23	0.9961	0.06	932
70	-0.64	0.05	-0.27	-0.78	-1.72	1.61	0.9838	0.09	223
$\log k'_w$	-0.21	0.70	-0.94	-1.04	-2.60	3.83	0.9948	0.09	711

Capcell Pak C₁₈ phase, ref. [4].

Table 4
Summary of the log P_{plus} regression Eq. 7 for solutes 1–43

Methanol (%)	c	a	v	o	ρ	S.D.	F
30	-0.53	-0.83	0.57	0.65	0.9972	0.05	2283
Log k'_w	-0.31	-0.78	0.98	0.80	0.9968	0.07	1991
30	-0.03	-0.67		0.67	0.9838	0.12	604
50	-0.44	-0.76		0.58	0.9916	0.08	1179
70	-0.80	-0.91		0.46	0.9728	0.11	352
Log k'_w	0.55	-0.50		0.83	0.9714	0.20	335

Capcell Pak C₁₈ phase, ref. [4].

ence between the log k' (or log k'_w) regression with log P_{oct} and the best regression with the log P_{plus} equation, depends on the $a\Sigma\alpha_2^{\text{H}}$ term, together with a vV_x term in two cases. For comparison, the constants are given in all four cases with just a two-term equation. Then it is clear that the $a\Sigma\alpha_2^{\text{H}}$ term is numerically largest for the 70% methanol system, and smallest for the 30% methanol system. This may seem contrary to the results in Table 3, where the $a\Sigma\alpha_2^{\text{H}}$ term is numerically the largest for the 30% system, but the effect of the hydrogen-bond term depends also on the magnitude of the other terms in Eq. 6.

We can now calculate from the two-term equation summarised in Table 4, exactly how the solute hydrogen-bond acidity contributes to the deviation in the log k' against log P_{oct} equation. In the first set of solutes (Table 1), the CONHR-substituted furans have $\Sigma\alpha_2^{\text{H}}$ equal to 0.36 units, which will lead to deviations of 0.24 log units in log k' for 30% methanol, 0.27 log units for 50% methanol, 0.33 units for 70% methanol, and 0.18 units in log k'_w . Because the o constants in Eq. 1 are all less than unity (0.667, 0.573, 0.454 and 0.829 respectively), these deviations in log k' or log k'_w correspond to deviations in any calculated log P_{oct} value of 0.36 (30%), 0.47 (50%), 0.73 (70%) and 0.22 (log k'_w). Yamagami and Takao [4] pointed out the deviant behaviour of indole and 3-carbomethoxyindole in set B, Table 1. Since $\Sigma\alpha_2^{\text{H}}$ is no less than 0.44 units for these two solutes, the calculated deviations will be even larger than those for the CONHR-furans.

We have highlighted the solute hydrogen-bond

acidity as a possible factor in deviations from Eq. 1, but even for non-acidic solutes, such deviations can also arise, but to a less extent. This can be seen by scaling the equations in log k' in order to make the vV_x term exactly the same as that in the log P_{oct} Eq. 6, so that for the 30% methanol system

$$1.32 \log k' = 0.48R_2 - 0.80\pi_2^{\text{H}} - 1.17\Sigma\alpha_2^{\text{H}} - 2.95\Sigma\beta_2^{\text{H}} + 3.84V_x \quad (8)$$

There is no comparison between the a constant in Eq. 8 and that in Eq. 6, as we have discussed above. But also, the s constants differ by 0.20 units, and the b constants by 0.45 units, with the same (adjusted) v constants. Thus if Eq. 1 is applied to solutes with a very wide range of $\Sigma\beta_2^{\text{H}}$ values and similar V_x values, then deviations will also occur.

It might be thought that log k' values determined at 100% methanol could be used to advantage in Eq. 1, because the a constant becomes numerically smaller with increasing methanol content. Indeed the Miller–Poole equation (Table 2) for log k' values on Bakerbond C₁₈ with 100% methanol mobile phase shows that the a constant is not significant. However, this does not help very much in the estimation of log P_{oct} values, because the blend of factors in the given Eq. 5 is still quite different to that in Eq. 6. We can show this in the same way as before, by multiplying Eq. 5 for 100% methanol, in Table 2, by the factor 3.15 through-

$$3.15 \log k' = -3.09 + 0.04R_2 - 0.60\pi_2^H - 0.09\Sigma\alpha_2^H - 1.95\Sigma\beta_2^H + 3.84V_x \quad (9)$$

A comparison of Eq. 9 with Eq. 6 shows that solutes of the same volume but of different R_2 , π_2^H or $\Sigma\beta_2^H$ value cannot lie on the same $\log k'$ vs. $\log P_{\text{oct}}$ plot. This is the reason why 100% methanol is almost never used as the mobile phase in $\log P_{\text{oct}}$ determinations.

We have therefore shown, through the general Eq. 5, how large deviations can arise in any $\log k'$ vs. $\log P_{\text{oct}}$ equation when the solute hydrogen-bond acidity varies through the given solute set if aqueous methanol is used as the mobile phase. Deviations, although smaller in nature, can also occur if other solute properties such as hydrogen-bond basicity vary widely in the solute set. Large deviations can arise when methanol itself is the mobile phase, because the blend of factors that influence $\log k'$ is now quite different to the blend that influences $\log P_{\text{oct}}$. Although our analysis has concentrated on RP-HPLC data for a particular C_{18} stationary phase, the results obtained here are likely to be general for C_{18} and similar phases, and indicate how extreme care must be taken to match the properties of the training set of solutes to those for which $\log P_{\text{oct}}$ values are to be estimated.

Acknowledgement

We are grateful to Dr. Colin F. Poole for a preprint of his paper prior to publication.

References

- [1] A. Nahum and Cs. Horváth, *J. Chromatogr.*, 192 (1980) 315.
- [2] T.M. Xie, B. Hulthe and S. Folestad, *Chemosphere*, 13 (1984) 445.
- [3] C. Yamagami, T. Ogura and N. Takao, *J. Chromatogr.*, 514 (1990) 123.
- [4] C. Yamagami and N. Takao, *Chem. Pharm. Bull.*, 40 (1992) 925.
- [5] P. Vallat, W. Fan, N. El Tayar, P.-A. Carrupt and B. Testa, *J. Liq. Chromatogr.*, 15 (1992) 2133.
- [6] P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft and M.H. Abraham, *Anal. Chem.*, 57 (1985) 2971.
- [7] M.J. Kamlet, M.H. Abraham, P.W. Carr, R.M. Doherty and R.W. Taft, *J. Chem. Soc., Perkin Trans. 2*, (1988) 2087.
- [8] M.H. Abraham, *Chem. Soc. Rev.*, 22 (1993) 73.
- [9] M.H. Abraham, *J. Phys. Org. Chem.*, 6 (1993) 660.
- [10] C. Yamagami, N. Takao and T. Fujita, *J. Pharm. Sci.*, 82 (1993) 155.
- [11] M.H. Abraham, H.S. Chadha, G.S. Whiting and R.C. Mitchell, *J. Pharm. Sci.*, 83 (1994) 1085.
- [12] K.G. Miller and C.F. Poole, *J. High. Resolut. Chromatogr.*, 17 (1994) 125.