

# Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biochemical Processes

Michael H. Abraham

Chemistry Department, University College London, 20 Gordon Street, London WC1H 0AJ

## 1 Introduction

It is well recognized that there are numerous processes in chemistry and biochemistry in which hydrogen-bonding plays an important rôle. For example, gaseous butanone is more soluble in water at 298 K than is gaseous butane by a factor of around  $1.7 \times 10^4$ ; at least part of this increased solubility must be due to butanone–water hydrogen-bonding. Or consider the inhibition of firefly luciferase activity by aqueous butanone or aqueous butane. Now the increased solvation of butanone over butane in water leads to butane being the more potent by a factor of 74. In order to understand and to interpret these effects quantitatively, it is necessary to separate out the various possible solute–solvent interactions, and to establish numerical scales for solute properties such as hydrogen-bond acidity and hydrogen-bond basicity. It is the purpose of this review to show how these solute property scales can be established, and then to illustrate their application to a number of selected physicochemical and biochemical processes.



To date there have been put forward but few quantitative scales of solute hydrogen-bond acidity or hydrogen-bond basicity that are of any real practical use. Sherry and Purcell,<sup>1</sup> suggested that the enthalpy of the hydrogen-bond complexation reaction (equation 1),  $\Delta H^\circ$ , could be expressed as the product of parameters characteristic of the acid and the base components. Some years later, Raevsky *et al.*, following the general plan of Sherry and Purcell, and also Ioghansen,<sup>2</sup> developed a very comprehensive set of solute parameters based on equation 2.<sup>3</sup> In this equation  $E_A$  is the hydrogen-bond acidity of a given solute,  $E_B$  is the hydrogen-bond basicity of a given solute, and  $\Delta H_{AB}$  is the standard enthalpy of reaction 1 in tetrachloromethane for the particular AB pair, in  $\text{kJ mol}^{-1}$ ,

$$\Delta H_{AB}^\circ = 22.5 E_A E_B \quad (2)$$

To standardize the scale,  $E_A$  is taken as  $-1.00$  for phenol and  $E_B$  is taken as  $+1.00$  for diethylether. Although equation 2 is very useful for the correlation and prediction of  $\Delta H^\circ$  values, it is not so helpful in LSER or QSAR studies for the particular reason that most processes investigated in these studies are Gibbs-energy related and are not enthalpy related. It seems more

relevant to construct solute scales that are derived from  $\log K$  or  $\Delta G^\circ$  values in the examination of Gibbs-energy related processes.

Together with scales based on equation 2, Raevsky *et al.* also derived corresponding Gibbs energy quantities using equation 3,

$$\Delta G_{AB}^\circ = 5.46 C_A C_B \quad (3)$$

Now,  $C_A = -1.00$  for phenol, and  $C_B = +1.00$  for diethylether; the constant 5.46 is taken as  $-\Delta G^\circ$  for the phenol–diethylether complexation in tetrachloromethane. Unfortunately, the methodology based on equation 3 suffers from a considerable disadvantage in the definition of zero acidity or zero basicity. Thus, if the hydrogen-bond acid cyclohexane were to complex with the hydrogen-bond base diethylether, we might expect  $K \rightarrow 0$ , and hence  $\Delta G^\circ \rightarrow \infty$ , leading to a value of  $-\infty$  for  $C_A$  for cyclohexane.\*

A similar difficulty has bedevilled the  $pK_{HB}$  scale of solute hydrogen-bond basicity set up by Taft *et al.*<sup>4</sup> and defined as  $\log K$  for the complexation of bases with 4-fluorophenol in tetrachloromethane at 298 K (equation 1; AH = 4-fluorophenol). Again, it is impossible to define an origin or zero point for the scale; in other words  $pK_{HB}$  for nonbasic compounds such as alkanes cannot be identified.

## 2 Construction of the Hydrogen-bond Scales

Faced with these difficulties, we set out to construct scales of solute hydrogen-bond acidity and hydrogen-bond basicity, using  $\log K$  values for reaction 1 in tetrachloromethane, so that the scales would be Gibbs-energy related. We started with solute hydrogen-bond acidity,<sup>5</sup> and set out  $\log K$  values for a series of acids against a given reference base. We were able to analyse data for no less than 45 reference bases, so that we had 45 such series of  $\log K$  values. We then found that if  $\log K$  values for acids against a given reference base were plotted vs.  $\log K$  values for acids against any other reference base, we obtained a series of straight lines that intersected near a ‘magic point’ of  $-1.1$  log units when  $K$  values were calculated on the molar scale (see Figure 1). This enabled us to construct 45 equations all of the form,

$$\log K (\text{series of acids against reference base B}) = L_B \log K_A^\dagger + D_B \quad (4)$$

where  $L_B$  and  $D_B$  characterize the base, and where the  $\log K_A^\dagger$  values now characterize the series of acids. All 45 equations were constrained to pass through the magic point ( $-1.1$ ,  $-1.1$ ). Examples of equation 4 for a relatively weak base tetrahydrofuran (THF) and a relatively strong base dimethylsulfoxide (DMSO) are,<sup>5</sup>

$$\log K (\text{acids against THF}) = 0.8248 \log K_A^\dagger - 0.1970 \quad (5)$$

$n = 23 \quad \rho = 0.9960 \quad \text{sd} = 0.089$

$$\log K (\text{acids against DMSO}) = 1.2399 \log K_A^\dagger + 0.2656 \quad (6)$$

$n = 51 \quad \rho = 0.9947 \quad \text{sd} = 0.096$

\* Note in proof. Raevsky *et al.* (*Quant. Structure–Activity Relat.*, 1992, **11**, 49) have now introduced a constant term in equation 3. This avoids the zero point difficulty, and puts Raevsky’s method onto the same basis as Abraham’s method.



Michael H. Abraham studied chemistry at the Northern Polytechnic London (B.Sc.), and then at University College London (Ph.D.). He was Reader at the University of Surrey, and is now research fellow at University College London. D.Sc. (London) in 1974. He has published a book, several reviews, and over 200 research papers, and was awarded the 1991 Ebert Prize of the American Pharmaceutical Association.

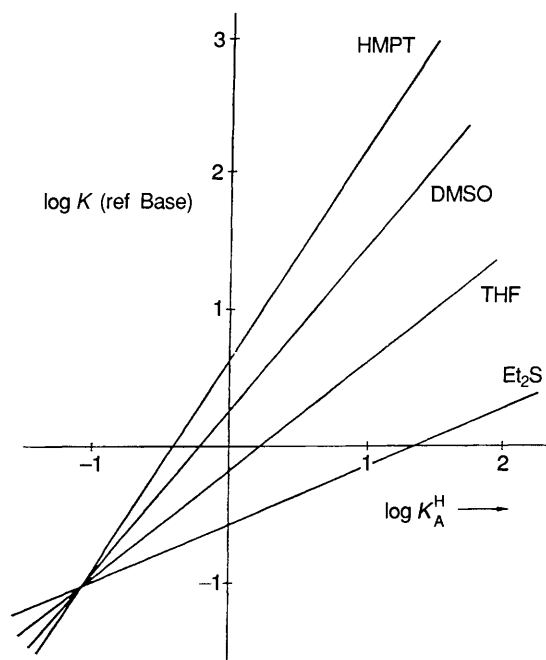


Figure 1 Plots of  $\log K$  (acids against reference base) vs  $\log K_A^H$ , showing the magic point

In the above equations, and elsewhere,  $n$  is the number of data points,  $\rho$  is the correlation coefficient, and  $sd$  the standard deviation. The  $\log K_A^H$  values that characterize the various acids now form a quantitative scale of hydrogen-bond acidity, defined through the 45 equations 4. Not only is the scale Gibbs-energy related, but there is now a natural zero-point, all compounds with zero hydrogen-bond acidity can be assigned  $\log K_A^H = -1.1$  units. It is convenient to shift the origin from  $-1.1$  to zero itself, and at the same time to compress the scale somewhat. This can be done by converting  $\log K_A^H$  to  $\alpha_2^H$ ,

$$\alpha_2^H = (\log K_A^H + 1.1)/4.636 \quad (7)$$

so that equation 7 is the defining equation<sup>5</sup> for the solute hydrogen-bond acidity  $\alpha_2^H$ . Now all compounds with zero acidity, i.e. with  $\log K_A^H = -1.1$ , have  $\alpha_2^H = 0$ .

In an exactly similar way,  $\log K$  values for series of bases against a reference acid A again formed a set of lines through the same magic point, and led to 34 equations of the form,<sup>6</sup>

$$\log K (\text{series of bases against reference acid A}) = L_A \log K_B^H + D_A \quad (8)$$

Now  $\log K_B^H$  characterizes the hydrogen-bond basicity of the series of bases and can again be transformed into a more convenient scale

$$\beta_2^H = (\log K_B^H + 1.1)/4.636 \quad (9)$$

Some values of the hydrogen-bond solute parameters thus obtained are listed in Table 1. Within families, there are connections between hydrogen-bond acidity or basicity and full proton transfer acidity or basicity, but these relationships collapse across families of solutes. For example,  $\alpha_2^H$  is larger for phenol than for simple carboxylic acids, and  $\beta_2^H$  is larger for DMSO than for  $\text{Et}_3\text{N}$ .

A rather simple relationship exists<sup>7</sup> between  $\alpha_2^H$ ,  $\beta_2^H$  and the  $\log K$  value for reaction 1 in tetrachloromethane at 298 K,

$$\log K = 7.354 \alpha_2^H \beta_2^H - 1.094 \quad (10)$$

$n = 1312 \quad \rho = 0.9956 \quad sd = 0.09$

Table 1 Some values<sup>a</sup> of the solute hydrogen-bond acidity and basicity parameters  $\alpha_2^H$  and  $\beta_2^H$ , and the effective parameters<sup>b</sup>  $\sum \alpha_2^H$  and  $\sum \beta_2^H$

Solute	$\alpha_2^H$	$\beta_2^H$	$\sum \alpha_2^H$	$\sum \beta_2^H$
Alkanes	0.00	0.00	0.00	0.00
Alkenes	0.00	0.07	0.00	0.07
$\text{CH}_2\text{Cl}_2$	0.13	0.05	0.10	0.05
$\text{CHCl}_3$	0.20	0.02	0.15	0.02
$\text{CCl}_4$	0.00	0.00	0.00	0.00
$\text{R}_2\text{O}$	0.00	0.45	0.00	0.45
Acetone	0.04	0.50	0.04	0.49
RCOR	0.00	0.48	0.00	0.51
DMSO	0.00	0.78	0.00	0.78
Acetonitrile	0.09	0.44	0.07	0.32
Nitromethane	0.12	0.25	0.06	0.31
$\text{Et}_3\text{N}$	0.00	0.67	0.00	0.79
Benzene	0.00	0.14	0.00	0.14
PhCl	0.00	0.09	0.00	0.07
PhOMe	0.00	0.26	0.00	0.29
PhCOMe	0.00	0.51	0.00	0.48
PhCN	0.00	0.42	0.00	0.33
PhNH <sub>2</sub>	0.26	0.38	0.26	0.41
Water	0.35	0.38	0.82	0.35
Methanol	0.37	0.41	0.43	0.47
Ethanol	0.33	0.44	0.37	0.48
Phenol	0.60	0.22	0.60	0.30
$\text{RCO}_2\text{H}$	0.54	0.42	0.60	0.45

Taken from references 5 and 6. R = n alkyl. <sup>b</sup> Ref 13 and unpublished work.

Although the form of equation 10 is similar to that of Raevsky's equation 3, the constant ( $-1.094$ ) in equation 10 is now related to the magic point and hence to the zero origin. Equation 10 is particularly useful in calculating values of  $\alpha_2^H$  or  $\beta_2^H$  from  $\log K$  values, when either  $\alpha_2^H$  or  $\beta_2^H$  is known.

The usefulness of the entire analysis using the magic point was well illustrated by examination<sup>8</sup> of  $\log K$  values obtained by Hine *et al.*<sup>9</sup> These workers plotted  $\log K$  (bases against a reference acid imide) vs  $\log K$  (bases against 4-fluorophenol) and obtained a set of randomly intersecting lines, Figure 2. Such a plot implies that the relative hydrogen-bond acidity of imides varies according to the actual reference base, especially for weak bases where  $\log K$  values are below around 1.0 log units. However, if Hine's  $\log K$  values are plotted according to equation 8, with the various lines forced through the magic point, Figure 3 is obtained. Now the relative hydrogen-bond acidity of the imines always remains the same, no matter what is the reference base. This, of course, is a condition for the establishment of any general scale.\*\*

The  $\alpha_2^H$  and  $\beta_2^H$  solute scales have been set up using  $\log K$  values for complexation in tetrachloromethane, and it is of some interest to establish whether these scales can be used in other solvents. Abboud *et al.*<sup>10</sup> determined  $\log K$  values for complexation of acids with the reference base pyridine *N*-oxide in cyclohexane at 296.5 K,  $\log K_{\text{PyO}}$ . They found quite a good correlation between  $\alpha_2^H$  and  $\log K_{\text{PyO}}$ , equation 11, which indicates that the  $\alpha_2^H$  (and  $\beta_2^H$ ) scales might be useful with other solvents than tetrachloromethane.

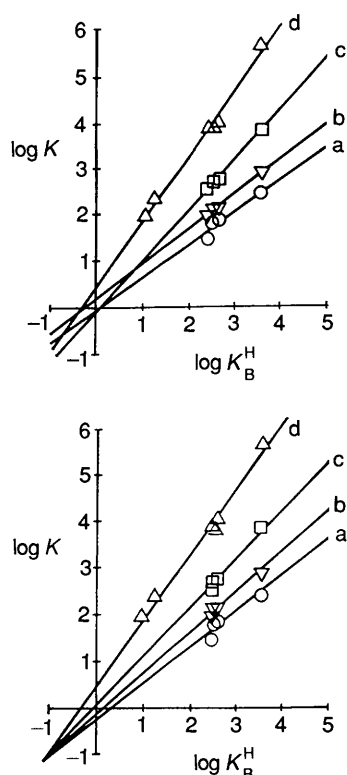
$$\alpha_2^H = 0.185 \log K_{\text{PyO}} + 0.069 \quad (11)$$

$n = 22 \quad \rho = 0.993 \quad sd = 0.03$

However, a comparison of scales set up with solvent 1,1,1-trichloroethane, TCE, and  $\alpha_2^H$  and  $\beta_2^H$  does suggest that some family dependencies may make conversion between these scales rather difficult,<sup>11</sup> within solute families, conversions can again

\* We use the general basicity scale  $\log K_B^H$  in Figure 2 but this makes little difference.

\*\* We note that there are some (known) exceptions with regard to our solute scales. In particular combinations of weak N-H acids with pyridine bases are excluded.<sup>5</sup>



**Figures 2 and 3** Plots of  $\log K$  (bases against reference acids) vs  $\log K_B^H$ . The reference acids are (a) 2-ethyl-2-methylsuccinimide, (b) 2-chloro-3-methylmaleimide, (c) 2,3-dichloro-2-methylsuccinimide, and (d) tetrafluorosuccinimide

be made. Thus for phenols and alcohols (where NMP = *N*-methylpyrrolidinone),<sup>11</sup>

$$\log K (\text{against NMP in TCE}) = 0.870 \log K_A^H + 0.70 \quad (12)$$

$n = 21 \quad \rho = 0.986 \quad \text{sd} = 0.13$

We have therefore established rather comprehensive scales of solute hydrogen-bond acidity and hydrogen-bond basicity, that can be extended using  $\log K$  values in other solvents, as shown above, and also through incorporation of more recent data in tetrachloromethane itself.<sup>12</sup> But all the  $\alpha_2^H$  and  $\beta_2^H$  values we have refer to 1:1 complexation, equation 1. It is by no means obvious that such values are relevant to the solvation situation in which a solute is surrounded by solvent molecules and hence undergoes multiple hydrogen-bonding. In order to test this, we set up a number of multiple linear regression equations, as shown below, and 'back-calculated' the solute hydrogen-bond parameters to check whether or not the 'effective' or 'summation' solute hydrogen-bond acidity and basicity ( $\Sigma\alpha_2^H$  and  $\Sigma\beta_2^H$ ) could be taken as the 1:1  $\alpha_2^H$  and  $\beta_2^H$  scales.<sup>13</sup> In the event, we were fortunate enough to find that in general the 1:1 scales could be used as solute parameters even when the solute was surrounded by a large excess of solvent. There were exceptions, and in Table 1 we compare the 1:1 values with the summation values<sup>13</sup> for a number of representative solutes, all of the monofunctional type. In general, multifunctional solutes cannot be dealt with in terms of 1:1 complexation constants, and at present can only be handled through the back-calculation of summation values.

### 3 Application of Hydrogen-bond Scales

Before these summation scales can be applied to any given solvation process, it is necessary to formulate some model of solvation. We use a simple cavity model, in which the process of dissolution of a gaseous solute in a solvent involves (i) the endoergic creation of a cavity in the solvent, and (ii) incorpora-

tion of the solute in the cavity with consequent setting up of various exoergic solute-solvent interactions. Each of these interactions will require a relevant solute parameter or descriptor.\* After considerable preliminary work (see e.g. ref. 14), the following solute descriptors were selected.  $R_2$  is an excess molar refraction that can be calculated from refractive index or can rather easily be estimated,<sup>14</sup>  $\pi_2^H$  is the solute dipolarity-polarizability obtained to date from gas-liquid chromatography (GLC) of solutes on polar stationary phases,<sup>13</sup>  $\alpha_2^H$  and  $\beta_2^H$  are the solute hydrogen-bond acidity and solute hydrogen-bond basicity (where appropriate  $\Sigma\alpha_2^H$  and  $\Sigma\beta_2^H$  must be used, but we retain the simpler nomenclature),  $L^{16}$  is the solute gas-hexadecane partition coefficient at 298 K,<sup>15</sup> and  $V_x$  is McGowan's characteristic volume.<sup>16</sup> Two general linear solvation energy relationships (LSERs) can be constructed from these parameters,<sup>13</sup>

$$\log SP = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + l\log L^{16} \quad (13)$$

$$\log SP = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + vV_x \quad (14)$$

In these equations, the dependent variable  $\log SP$  refers to some property of a series of solutes in a fixed phase (or phases). Thus  $SP$  could be  $L$ , the gas-liquid partition coefficient for a series of solutes in a given liquid or it could be  $P$ , the partition coefficient for a series of solutes between water and, say, octanol. In the case of biological properties, where  $SP$  can be some biological response as an  $LC_{50}$ , equations 13 and 14 then represent two new families of quantitative structure-activity relationships (QSARs).

Equation 13 is the simpler, and can be applied to processes involving gas  $\rightarrow$  condensed-phase transfer. The terms  $rR_2$ ,  $s\pi_2^H$ ,  $a\alpha_2^H$  and  $b\beta_2^H$  represent specific solute-condensed phase interactions, respectively dispersion, dipole-dipole or dipole-induced-dipole plus some polarizability interaction, solute acid-solvent base, and solute base-solvent acid. The  $\log L^{16}$  term includes both general dispersion interactions and the endoergic cavity term, at the moment it seems not possible to separate out the important cavity term on its own. The constants in equation 13 are found by the method of multiple linear regression analysis and serve to characterize the particular condensed phase.  $r$  is the tendency of the phase to interact through  $\pi$ - and  $n$ -electron pairs,  $s$  is the phase dipolarity-polarizability,  $a$  is the phase hydrogen-bond basicity (because a basic phase will interact with an acidic solute),  $b$  is the phase hydrogen-bond acidity, and  $l$  is a measure of the ability of the phase to distinguish between or to separate homologues in any homologous series.

For processes within condensed phases, equation 14 is used, but now the various constants will reflect differences between the phases. If  $SP$  is  $P_{\text{oct}}$ , the water-octanol partition coefficient, then for example the  $a$ -constant will be a measure of the difference in basicity between water and octanol. The  $vV_x$  term in equation 14 will include both differences in cavity effects and differences in general dispersion interactions, the  $v$ -constant now being a measure of the difference between the hydrophobicity or lipophilicity of the two phases.

#### 3.1 Application to Gas $\rightarrow$ Condensed Phase Processes

Equation 13 has been applied to several large sets of gas-liquid chromatographic (GLC) data.† For the retention of a series of solutes on a given stationary phase at a given temperature, we can take  $SP$  to be  $V_G$ , the specific retention volume, or  $L(K)$ , the gas  $\rightarrow$  liquid partition coefficient, or even  $t$ , the adjusted retention time. Details of the application of equation 13 to data

\* All the processes we shall consider involve a series of solute molecules in a given solvent or solvents. Hence the solvent(s) properties remain constant and it is not necessary to have to deal with solvent parameters. Only the solute is varied and hence only solute parameters are needed.

† For multifunctional solutes we should really refer to  $\Sigma\alpha_2^H$  and  $\beta_2^H$ , the effective or summation hydrogen bond acidity or basicity. In the examples that follow it should be recognized that  $\alpha_2^H$  and  $\beta_2^H$  do indeed refer to  $\Sigma\alpha_2^H$  and  $\Sigma\beta_2^H$  where appropriate.

obtained by Lafford *et al.*<sup>17</sup> and by McReynolds<sup>18</sup> are in Table 2.<sup>13</sup> Poole *et al.*<sup>19</sup> have obtained retention data at 394 K on a variety of phases, including some novel molten salts. Some

**Table 2** Application of equation 13 to GLC data at 393 K<sup>a</sup>

Stationary phase	<i>c</i>	<i>r</i>	<i>s</i>	<i>a</i>	<i>l</i>	<i>n</i>	$\rho$	sd
Carbowax <sup>b</sup>	-2.01	0.25	1.26	2.07	0.429	199	0.997	0.07
DEGS <sup>b</sup>	-1.77	0.35	1.58	1.84	0.383	199	0.997	0.07
PPE <sup>b</sup>	-2.51	0.14	0.89	0.67	0.547	199	0.997	0.06
TCEP <sup>b</sup>	-1.69	0.26	1.93	1.88	0.365	199	0.998	0.06
ZE7 <sup>b</sup>	-1.99	-0.41	1.46	0.77	0.432	199	0.995	0.07
Apiezon J	-0.48	0.24	0.15	0.13	0.596	165	0.999	0.02
Pluronic L 72	-0.54	0.09	0.93	1.42	0.529	163	0.998	0.03

All the phases are non acidic hence  $b = 0$ . <sup>a</sup>Data from ref. 17. Note that  $\log SP = \log L - \log L(\text{decane})$ . The abbreviations are: DEGS diethylenglycol succinate; PPE polyphenylether; TCEP tricyanoethoxypropane; ZE7 zonyl E 7. Data from ref. 18.  $SP = V_c$ .

examples of analyses of Poole's data are in Table 3. Examination of Tables 2 and 3 shows how any GLC stationary phase can now be characterized by the constants in equation 13. Thus squalane and apiezon J are almost nonpolar and nonbasic ( $s$  and  $a$  are small or zero), but are very good at separating homologues ( $l$  is large). Phases such as DEGS and TCEP are quite dipolar and quite basic ( $s$  and  $a$  are large), whereas ZE7 is dipolar but not very basic, and the molten salts are dipolar and very basic ( $a = 3.4$  for the 4-toluene sulfonate and the methane sulfonate).

Not only can GLC stationary phases be characterized through equation 13, but so can any solvent phase. For example, we have applied equation 13 to gas-liquid partition coefficients in a number of amides at 298 K.<sup>20</sup> Using the most recent solvation equation 13, we find for *N*-formylmorpholine (NFM) and also for tri(2-ethylhexylphosphate) (2-EHP) at 298 K the equations,

$$\log L(\text{in NFM}) = -0.53 + 2.57\pi_2^H + 4.32a_2^H + 0.730 \log L^{16} \quad (15)$$

$$n = 45 \quad \rho = 0.9949 \quad \text{sd} = 0.07$$

$$\log L(\text{in 2-EHP}) = -0.07 - 0.26R_2 + 0.91\pi_2^H + 3.74a_2^H + 0.955 \log L^{16} \quad (16)$$

$$n = 22 \quad \rho = 0.9978 \quad \text{sd} = 0.04$$

Since neither of these solvents is acidic, the  $b$ -constant is zero. However, both are quite dipolar and NFM, particularly, is highly basic with  $a = 4.32$  units.\*

The toxicity of gases and vapours is a considerable environmental problem. One procedure for the estimation of such toxicity is to measure the upper respiratory tract irritation.

\* Note that the constants obtained at 298 K in equations 15 and 16 cannot directly be compared with those at 393 K (Tables 2 and 3) because an increase in temperature invariably results in a marked decrease in  $v$ ,  $a$  and  $b$ .

**Table 3** Application of equation 13 to GLC data of Poole *et al.*, at 394.4 K<sup>a</sup>

Stationary phase	<i>c</i>	<i>r</i>	<i>s</i>	<i>a</i>	<i>l</i>	<i>n</i>	$\rho$	sd
Squalane	-0.20	0.12	0.02	-0.10	0.581	39	0.999	0.03
OV 11	-0.30	0.10	0.54	0.17	0.516	39	0.999	0.03
OV 225	-0.51	0.02	1.21	0.96	0.462	39	0.998	0.03
Tetrabutylammonium 4-toluene sulfonate	-0.62	0.01	1.66	3.36	0.440	34	0.989	0.11
Tetrabutylammonium methane sulfonate	-0.63	0.09	1.60	3.41	0.437	34	0.990	0.10
Tetrabutylammonium picrate	-0.54	0.10	1.56	1.42	0.445	36	0.994	0.06

Data from ref. 19.  $SP = L$ . All the phases are non acidic hence  $b = 0$ .

caused by vapours to mice. We have shown<sup>21</sup> that a QSAR for the toxicity of nonreactive compounds can be obtained using our new hydrogen-bond parameters, and a recalculation based on equation 13 yields,

$$-\log \text{FRD}_{50} = 0.96 + 0.81\pi_2^H + 2.55a_2^H + 0.722 \log L^{16} \quad (17)$$

$$n = 39 \quad \rho = 0.987 \quad \text{sd} = 0.12$$

This QSAR is useful for the prediction of toxicity of nonreactive compounds, but also shows that the receptor site is reasonably polar ( $s = 0.81$ ) and has a somewhat lower hydrogen-bond basicity than (2-EHP). Equation 17 can be contrasted with a preliminary equation for the solubility of gaseous solutes in water at 298 K,

$$\log L(\text{in water}) = -1.28 + 0.87R_2 + 2.70\pi_2^H + 4.01a_2^H + 4.80\beta_2^H - 0.210 \log L^{16} \quad (18)$$

$$n = 350 \quad \rho = 0.9952 \quad \text{sd} = 0.19$$

It is clear that the receptor site cannot be any sort of aqueous environment, since equation 18 is completely different to equation 17.

Many other examples of the application of equation 13 can be given, for example to the solubility of gaseous solutes in polymers,<sup>22</sup> or to the solubility of gaseous solutes in blood and other biological systems,<sup>23</sup> but the examples given here show the widespread use of this general solvation equation in the correlation of gas  $\rightarrow$  condensed phase processes.

One other useful feature of equation 13 is that it enables the contribution of specific solute-solvent interactions to the overall  $\log L$  values to be calculated. Thus for a solute with a known  $\pi_2^H$  value, the contribution of the solute-solvent dipolarity-dipolarity term will simply be given by the  $s\pi_2^H$  product. Calculations on these lines are in Tables 4 and 5 for gaseous solubility in NFM.

**Table 4** An analysis of solute-solvent interaction contributions to  $\log L$  in NFM at 298 K

Solute	<i>c</i>	$rR_2$	$s\pi_2^H$	$aa_2^H$	$b\beta_2^H$	$l/\log L^{16}$
Butane	-0.53	0.00	0.00	0.00	0.00	1.18
Hexane	-0.53	0.00	0.00	0.00	0.00	1.91
Benzene	-0.53	0.00	1.34	0.00	0.00	2.03
Butanone	-0.53	0.00	1.80	0.00	0.00	1.67
Ethanol	-0.53	0.00	1.08	1.60	0.00	1.08

**Table 5** An analysis of solute-solvent interaction contributions to  $\log L$  in water at 298 K

Solute	<i>c</i>	$rR_2$	$s\pi_2^H$	$aa_2^H$	$b\beta_2^H$	$l/\log L^{16}$
Butane	-1.28	0.00	0.00	0.00	0.00	-0.34
Hexane	-1.28	0.00	0.00	0.00	0.00	-0.56
Benzene	-1.28	0.53	1.40	0.00	0.67	-0.59
Butanone	-1.28	0.14	1.89	0.00	2.45	-0.48
Ethanol	-1.28	0.21	1.13	1.48	2.30	-0.31

and water. The important dispersion interaction term is not explicitly given in equation 13, because the  $\log L^{16}$  term includes both a positive dispersion interaction effect towards  $\log L$ , and a negative cavity term effect towards  $\log L$ . The analysis of Abraham and Fuchs,<sup>24</sup> and also cavity calculations based on Pierotti's scaled particle theory,<sup>25</sup> both show that dispersion interactions are very large and are nearly always the dominant positive interaction on  $\log L$ .

Possibly the most extraordinary feature of equation 18 is the negative dependence of  $\log L$  on  $\log L^{16}$ , whereas for all nonaqueous solvents we have studied, the  $l$ -constant is quite positive.\* This is a manifestation of the hydrophobic effect. Large gaseous solutes become very soluble in nonaqueous solvents, but in water are no more soluble than small solutes. Both SPT calculations and the analysis of Abraham and Fuchs for hexadecane solvent suggest that the main reason for the peculiar behaviour of water lies in the cavity effect. This becomes much more endoergic as the solute size increases than do cavity effects in nonaqueous solvents, and leads to a large negative contribution to  $\log L$ . Additionally, the dispersion interaction term in nonaqueous solvents becomes more exoergic as solute size increases than does the dispersion term in water. The dispersion interaction effect leads to more positive contributions to  $\log L$  for larger solutes in nonaqueous media than with larger solutes in water.

We can now see from the breakdown in Table 5, exactly what are the factors that lead to increased aqueous solubility of gaseous butanone over gaseous butane at 298 K. The two main terms are  $s\pi_2^H$  and  $b\beta_2^H$  which contribute 1.89 and 2.45 log units to the more favourable solubility of butanone. These correspond to extra exoergic Gibbs energies of 11.0 and 14.0 kJ mol<sup>-1</sup>, due to dipole-dipole and hydrogen-bond (solute base-solvent acid) interactions in the case of butanone. If we consider solvent NFM, Table 4, the increased solubility of gaseous butanone over butane is mainly due to the  $s\pi_2^H$  term, corresponding to an exoergic Gibbs energy of about 10.3 kJ mol<sup>-1</sup> favouring butanone through dipole-dipole interactions. The construction of scales of solute hydrogen-bonding, together with other solute descriptors thus leads, *via* equation 13 and also equation 14, to a quantitative assessment of the rôle of hydrogen-bonding in various processes.

### 3.2 Application to Processes in Condensed Phases

There have been very many applications of equation 14, and so we shall merely give a few examples of recent ones. In principle, any kind of partition between two condensed phases can be examined, and  $SP$  in equation 14 can be  $k'$  the HPLC capacity factor, or  $P$  a liquid-liquid partition coefficient, for example. We have already analysed the very important water-octanol partition coefficient  $P_{\text{OCT}}$ , using an earlier equation<sup>26</sup> and we now recalculate the correlation using equation 14

$$\log P_{\text{OCT}} = 0.08 + 0.58R_2 - 1.09\pi_2^H + 0.03\alpha_2^H - 3.40\beta_2^H + 3.81V_x \quad (19)$$

$$n = 584 \quad \rho = 0.996 \quad \text{sd} = 0.13$$

Equation 19 shows that increase in solute size,  $V_x$ , favours wet octanol, whereas increase in solute dipolarity,  $\pi_2^H$ , or solute hydrogen-bond basicity,  $\beta_2^H$ , favours the aqueous layer. Solute hydrogen-bond acidity has little influence on  $\log P_{\text{OCT}}$ . Since the constants in equation 19 refer to differences in properties of the two phases concerned, we can deduce that the hydrogen-bond basicity of water and wet octanol are almost the same. Similar analyses can be carried out for numerous water-solvent partitions.

There are a large number of biochemical and toxicological processes that involve aqueous solutes interacting with a given

system. In principle, the general solvation equation 14 could be applied to any such process. As an example, we can quote the work of Franks and Lieb<sup>27</sup> on the inhibition of firefly luciferase activity by aqueous nonelectrolytes. Application of equation 14 to the data of Franks and Lieb leads to the regression,

$$-\log EC_{50} = 0.58 + 0.72R_2 - 3.44\beta_2^H + 3.77V_x \quad (20)$$

$$n = 42 \quad \rho = 0.989 \quad \text{sd} = 0.33$$

The crucial factors that determine the potency of aqueous solutes are thus the solute volume that increases inhibition, and solute hydrogen-bond basicity that decreases inhibition.<sup>28</sup> Since the solute hydrogen-bond acidity plays no part, it can be deduced that the target site(s) must be of about the same hydrogen-bond basicity as water itself. On the other hand, the large negative  $b$ -constant in equation 20 indicates that the target site(s) must have a relatively poor hydrogen-bond acidity, *cf* equation 19.

Similar results are obtained for the potency of aqueous nonelectrolytes in inducing general anaesthesia in animals. Again, the general anaesthetic site must be of about the same hydrogen-bond basicity as water, but significantly less acidic.<sup>28</sup>

We can now deal with the other example mentioned in the introduction, namely that the potency of butane as an inhibitor is some 74 times that of butanone, towards firefly luciferase activity. The two main terms governing potency, as  $-\log EC_{50}$ , are the  $b\beta_2^H$  and the  $vV_x$  terms. Since  $V_x$  for butane and butanone is almost the same, the diminished potency of butanone is due to increased hydrogen-bonding with water as compared to the target site(s). With  $\beta_2^H = 0.51$  for butanone, the  $b\beta_2^H$  term contributes  $3.44 \times 0.51 = 1.75$  log units (or a factor of 56) equivalent to a Gibbs energy of 10.0 kJ mol<sup>-1</sup> in favour of butane over butanone as an inhibitor. Thus most of the factor of 74 can be related to hydrogen-bonding of butanone with water.

## 4 Summary

Scales of hydrogen-bond acidity and basicity have been set up using formation constants in tetrachloromethane. These  $\alpha_2^H$  and  $\beta_2^H$  scales are Gibbs-energy related, and are unique for such scales in that they incorporate 'zero points'. It is found that  $\alpha_2^H$  and  $\beta_2^H$  can be used as the basis of general scales that include 'effective' or 'summation' hydrogen-bond acidities and basicities for use as solute parameters or descriptors in LSER and QSAR equations. Such equations that include also various other solute descriptors can be used to correlate and to interpret a wide variety of physicochemical and biochemical processes. The two general LSER and QSAR equations 13 and 14 can in principle be applied to any process involving gas  $\rightarrow$  condensed phase transfer, equation 13, or to any process within condensed phases, equation 14. The main requirement is sufficient values of the dependent variable,  $\log SP$ , that span a variety of solute type. Usually at least five data points are needed for each explanatory variable, and the variety of solute type is necessary so that the explanatory variables or descriptors cover as wide a range as possible and, most importantly, are not subject to significant cross-correlations. In order that these criteria be met, it is necessary to have available the relevant solute descriptors for a rather large number of solutes. In order to aid workers who wish to test equations 13 and 14, we set out the solute descriptors for a reasonably large number and variety of solutes. It is not feasible to list all the available parameters: for example we have values of  $\log L^{16}$  for 100 alkanes, and have recently listed all the parameters for 120 alkylaromatic hydrocarbons,<sup>29</sup> but hopefully those given in Tables 6, 7 and 8 will be sufficient for many purposes. Note that in these tables the parameters are effective or summation values, where appropriate.

Taylor *et al.*<sup>30</sup> have recently discovered an unwelcome complication, in that the hydrogen-bond basicity of certain solutes in water-solvent partitions seems to vary with the particular water-solvent system. We find that for a large number of solutes, including all those given in Tables 6 and 8,  $\Sigma\beta_2^H$  values are

\* Exactly the same results are obtained if solute volume  $V_x$  is used instead of the  $\log L^{16}$  parameter.

**Table 6** Values of  $\log L^{16}$  and  $V_x$  for inorganic gases and alkanes<sup>a</sup>

Solute	$\log L^{16}$	$V_x$
Helium	-1.741	0.068
Neon	-1.575	0.085
Argon	-0.688	0.190
Krypton	-0.211	0.246
Xenon	0.378	0.329
Radon	0.877	0.384
Hydrogen	-1.200	0.109
Oxygen	-0.723	0.183
Nitrogen	-0.978	0.222
Methane	-0.323	0.250
Ethane	0.492	0.390
Propane	1.050	0.531
n-Butane	1.615	0.672
2-Methylpropane	1.409	0.672
n-Pentane	2.162	0.813
2-Methylbutane	2.013	0.813
2,2-Dimethylpropane	1.820	0.813
n-Hexane	2.668	0.954
2-Methylpentane	2.503	0.954
3-Methylpentane	2.581	0.954
2,2-Dimethylbutane	2.352	0.954
2,3-Dimethylbutane	2.495	0.954
n-Heptane	3.173	1.095
2-Methylhexane	3.001	1.095
3-Methylhexane	3.044	1.095
3-Ethylpentane	3.091	1.095
2,2-Dimethylpentane	2.796	1.095
2,3-Dimethylpentane	3.016	1.095
2,4-Dimethylpentane	2.809	1.095
3,3-Dimethylpentane	2.946	1.095
2,2,3-Trimethylbutane	2.918	1.095
n-Octane	3.677	1.236
2,2,4-Trimethylpentane	3.106	1.236
n-Nonane	4.182	1.377
n-Decane	4.686	1.518
n-Undecane	5.191	1.659
n-Dodecane	5.696	1.799
n-Tridecane	6.200	1.940
n-Tetradecane	6.705	2.081
n-Pentadecane	7.209	2.222
n-Hexadecane	7.714	2.363
n-Heptadecane	8.218	2.504
n-Octadecane	8.722	2.645
n-Nonadecane	9.226	2.786
n-Eicosane	9.731	2.927
n-Heneicosane	10.236	3.068
n-Docosane	10.740	3.208
n-Tricosane	11.252	3.349

<sup>a</sup> All these solutes have zero  $R_2$  values. Units of  $V_x$  are ( $\text{cm}^3 \text{mol}^{-1}$ )/100

**Table 7** Descriptors for solutes with varying  $\sum \beta_2$ 

Solute	$R_2$	$\pi_2^H$	$\sum \alpha_2^H$	$\sum \beta_2^H$	$\sum \beta_2^0$	$\log L^{16}$
Aniline	0.955	0.96	0.26	0.41	0.50	3.934
<i>o</i> -Toluidine	0.966	0.92	0.23	0.45	0.59	4.442
<i>m</i> -Toluidine	0.946	0.95	0.23	0.45	0.55	4.463
<i>p</i> -Toluidine	0.923	0.95	0.23	0.45	0.52	4.452
2-Chloroaniline	1.033	0.92	0.25	0.31	0.40	4.674
3-Chloroaniline	1.053	1.10	0.30	0.30	0.36	4.909
4-Chloroaniline	1.060	1.13	0.30	0.31	0.35	4.889
2-Methoxyaniline	0.988	1.00	0.23	0.50	0.67	4.818
3-Methoxyaniline	1.027	1.22	0.25	0.55	0.68	5.023
4-Methoxyaniline	1.050	1.10	0.23	0.65	0.72	4.949
Methyl 4-aminobenzoate	1.078	1.52	0.32	0.59	0.64	
Ethyl 4-aminobenzoate	1.040	1.52	0.32	0.59	0.64	
n-Propyl 4-aminobenzoate	1.030	1.50	0.32	0.59	0.64	
n-Butyl 4-aminobenzoate	1.020	1.47	0.32	0.59	0.64	
Pyridine	0.631	0.84	0.00	0.52	0.47	3.022
2-Methylpyridine	0.598	0.75	0.00	0.58	0.48	3.422
3-Methylpyridine	0.631	0.81	0.00	0.54	0.44	3.631
4-Methylpyridine	0.630	0.82	0.00	0.54	0.43	3.640
2,3-Dimethylpyridine	0.657	0.77	0.00	0.62	0.50	4.045
2,4-Dimethylpyridine	0.634	0.76	0.00	0.63	0.49	4.006
2,5-Dimethylpyridine	0.633	0.74	0.00	0.62	0.49	3.986
2,6-Dimethylpyridine	0.607	0.70	0.00	0.63	0.49	3.760
3,4-Dimethylpyridine	0.676	0.85	0.00	0.62	0.48	4.317
3,5-Dimethylpyridine	0.659	0.79	0.00	0.60	0.44	4.214
2-Ethylpyridine	0.613	0.71	0.00	0.59	0.49	3.844
3-Ethylpyridine	0.640	0.79	0.00	0.57	0.47	4.093
4-Ethylpyridine	0.634	0.80	0.00	0.57	0.47	4.124

constant, and can be used in equations that describe any gas  $\rightarrow$  condensed phase process and any water-solvent partition process. There are, however, solutes for which the general  $\beta_2^H$  (or  $\Sigma \beta_2^H$ ) descriptor has to be modified for certain water-solvent partition processes, specifically those involving solvents that contain a rather high proportion of water when saturated. Thus for the solutes in Table 7 the alternative  $\beta_2^0$  value must be used in LSER equations that describe water-octanol, water-isobutanol, and possibly also water-butyl acetate and water-diethyl ether partitioning. Note that in the water-octanol LSER equation 19, only solutes with an invariant  $\beta_2$  value ( $\beta_2^H$  or  $\Sigma \beta_2^H$ ) have been included.

#### 4.1 A Note on the Calculation of $V_x$

In the various equations we have given,  $V_x$  ( $\text{in cm}^3 \text{mol}^{-1}/100$ ) has been calculated by the procedure of McGowan<sup>16</sup> in which atom constants are simply summed, and  $6.56 \text{ cm}^3 \text{mol}^{-1}$  subtracted for each bond, noting that all bonds (single, double,

**Table 8** Values of solute descriptors used in equations 13 and 14

Solute	$R_2$	$\pi_2^H$	$\sum \alpha_2^H$	$\sum \beta_2^H$	$\log L^{16}$
Cyclopentane	0.263	0.10	0.00	0.00	2.477
Methylcyclopentane	0.225	0.10	0.00	0.00	2.816
Cyclohexane	0.305	0.10	0.00	0.00	2.964
Methylcyclohexane	0.244	0.10	0.00	0.00	3.323
Cycloheptane	0.350	0.10	0.00	0.00	3.704
Methylcycloheptane	0.300	0.10	0.00	0.00	4.034
Cyclooctane	0.413	0.10	0.00	0.00	4.329
<i>trans</i> -Hydrindane	0.439	0.20	0.00	0.00	4.467
<i>cis</i> -Hydrindane	0.439	0.25	0.00	0.00	4.635
Adamantane	0.667	0.66	0.00	0.02	5.095
<i>trans</i> -Decalin	0.467	0.23	0.00	0.00	4.984
<i>cis</i> -Decalin	0.544	0.25	0.00	0.00	5.156
Ethene	0.107	0.10	0.00	0.07	0.289
Propene	0.103	0.08	0.00	0.07	0.946
But-1-ene	0.100	0.08	0.00	0.07	1.491
Pent-1-ene	0.093	0.08	0.00	0.07	2.047

Table 8 Continued

Solute	$R_2$	$\pi_2^H$	$\sum a_2^H$	$\sum \beta_2^H$	$\log L^{16}$
Hex-1-ene	0 078	0 08	0 00	0 07	2 572
Hept-1-ene	0 092	0 08	0 00	0 07	3 063
Oct-1-ene	0 094	0 08	0 00	0 07	3 568
Non-1-ene	0 090	0 08	0 00	0 07	4 073
Dec-1-ene	0 093	0 08	0 00	0 07	4 533
Undec-1-ene	0 091	0 08	0 00	0 07	5 023
Dodec-1-ene	0 089	0 08	0 00	0 07	5 515
Buta-1,3-diene	0 320	0 23	0 00	0 10	1 543
2-Methylbuta-1,3-diene	0 313	0 23	0 00	0 10	2 101
2,3-Dimethylbuta-1,3-diene	0 352	0 23	0 00	0 14	2 690
Cyclopentene	0 335	0 20	0 00	0 10	2 402
1-Methylcyclopentene	0 330	0 20	0 00	0 10	2 864
Cyclohexene	0 395	0 20	0 00	0 10	3 021
1-Methylcyclohexene	0 391	0 20	0 00	0 10	3 483
Cycloheptene	0 414	0 22	0 00	0 10	3 626
1-Methylcycloheptene	0 430	0 22	0 00	0 10	3 957
Ethyne	0 190	0 25	0 21	0 15	0 150
Propyne	0 183	0 25	0 12	0 15	1 025
But-1-yne	0 178	0 23	0 12	0 15	1 520
Pent-1-yne	0 172	0 23	0 12	0 12	2 010
Hex-1-yne	0 166	0 23	0 12	0 10	2 510
Hept-1-yne	0 160	0 23	0 12	0 10	3 000
Oct-1-yne	0 155	0 23	0 12	0 10	3 521
Oct-2-yne	0 225	0 30	0 00	0 10	3 850
Non-1-yne	0 150	0 23	0 12	0 10	4 019
Dec-1-yne	0 143	0 23	0 12	0 10	4 537
Dodec-1-yne	0 133	0 23	0 12	0 10	5 657
Fluoroethane	0 052	0 35	0 00	0 10	0 559
1-Fluoropropane	0 034	0 35	0 00	0 10	1 071
1-Fluorohexane	0 000	0 35	0 00	0 10	2 951
1-Fluorooctane	- 0 020	0 35	0 00	0 10	3 850
Dichloromethane	0 387	0 57	0 10	0 05	2 019
Trichloromethane	0 425	0 49	0 15	0 02	2 480
Tetrachloromethane	0 458	0 38	0 00	0 00	2 823
Chloroethane	0 227	0 40	0 00	0 10	1 678
1,1-Dichloroethane	0 322	0 49	0 10	0 10	2 316
1,2-Dichloroethane	0 416	0 64	0 10	0 11	2 573
1,1,1-Trichloroethane	0 369	0 41	0 00	0 09	2 733
1,1,2-Trichloroethane	0 499	0 68	0 13	0 08	3 290
1-Chloropropane	0 216	0 40	0 00	0 10	2 202
2-Chloropropane	0 177	0 35	0 00	0 12	1 970
1-Chlorobutane	0 210	0 40	0 00	0 10	2 722
1-Chloropentane	0 208	0 40	0 00	0 10	3 223
1-Chlorohexane	0 201	0 40	0 00	0 10	3 777
1-Chloroheptane	0 194	0 40	0 00	0 10	4 282
1-Chlorooctane	0 191	0 40	0 00	0 10	4 772
Chlorocyclohexane	0 448	0 48	0 00	0 10	4 016
1,1-Dichloroethene	0 362	0 34	0 00	0 05	2 110
cis-1,2-Dichloroethene	0 436	0 61	0 11	0 05	2 439
trans-1,2-Dichloroethene	0 425	0 41	0 09	0 05	2 278
Bromomethane	0 399	0 43	0 00	0 10	1 630
Dibromomethane	0 714	0 67	0 10	0 10	2 886
Tribromomethane	0 974	0 68	0 15	0 06	3 784
Bromoethane	0 366	0 40	0 00	0 12	2 120
1-Bromopropane	0 366	0 40	0 00	0 12	2 620
1-Bromobutane	0 360	0 40	0 00	0 12	3 105
1-Bromopentane	0 356	0 40	0 00	0 12	3 611
1-Bromohexane	0 349	0 40	0 00	0 12	4 130
1-Bromoheptane	0 343	0 40	0 00	0 12	4 663
1-Bromooctane	0 339	0 40	0 00	0 12	5 090
1-Bromononane	0 336	0 40	0 00	0 12	5 560
Bromocyclohexane	0 615	0 54	0 00	0 16	4 395
Iodomethane	0 676	0 43	0 00	0 13	2 106
Diiodomethane	1 453	0 69	0 05	0 23	3 857
Iodoethane	0 640	0 40	0 00	0 15	2 573
1-Iodopropane	0 634	0 40	0 00	0 15	3 130
1-Iodobutane	0 628	0 40	0 00	0 15	3 628
1-Iodopentane	0 621	0 40	0 00	0 15	4 130
1-Iodohexane	0 615	0 40	0 00	0 15	4 620
Halothane	0 102	0 38	0 15	0 05	2 177
Teflurane	- 0 070	0 21	0 20	0 02	1 370
Diethylether	0 041	0 25	0 00	0 45	2 015
Di-n-propylether	0 008	0 25	0 00	0 45	2 954
Di-n-butylether	0 000	0 25	0 00	0 45	3 924

Table 8 Continued

Solute	$R_2$	$\pi_2^H$	$\sum a_2^H$	$\sum \beta_2^H$	$\log L^{16}$
Methoxyflurane	0 109	0 67	0 07	0 14	2 864
Isoflurane	- 0 240	0 50	0 10	0 10	1 576
Enflurane	- 0 230	0 40	0 12	0 13	1 750
Fluroxene	0 183	0 30	0 00	0 27	1 600
Acetaldehyde	0 208	0 67	0 00	0 45	1 230
Propionaldehyde	0 196	0 65	0 00	0 45	1 815
Butyraldehyde	0 187	0 65	0 00	0 45	2 270
Iso-butyraldehyde	0 144	0 62	0 00	0 45	2 120
Pentanal	0 163	0 65	0 00	0 45	2 851
Hexanal	0 146	0 65	0 00	0 45	3 357
Heptanal	0 140	0 65	0 00	0 45	3 865
Octanal	0 160	0 65	0 00	0 45	4 361
Nonanal	0 150	0 65	0 00	0 45	4 856
Propenal	0 324	0 72	0 00	0 45	1 656
<i>trans</i> -But-2-ene-1-al	0 387	0 80	0 00	0 50	2 570
2-Methylpropenal	0 400	0 70	0 00	0 50	2 180
Propanone	0 179	0 70	0 04	0 49	1 696
Butanone	0 166	0 70	0 00	0 51	2 287
Pentan-2-one	0 143	0 68	0 00	0 51	2 755
Hexan-2-one	0 136	0 68	0 00	0 51	3 262
Heptan-2-one	0 123	0 68	0 00	0 51	3 760
Octan-2-one	0 108	0 68	0 00	0 51	4 257
Nonan-2-one	0 119	0 68	0 00	0 51	4 735
Decan-2-one	0 108	0 68	0 00	0 51	5 245
Undecan-2-one	0 101	0 68	0 00	0 51	5 732
Dodecan-2-one	0 103	0 68	0 00	0 51	6 167
Cyclopentanone	0 373	0 86	0 00	0 52	3 221
Cyclohexanone	0 403	0 86	0 00	0 56	3 792
Cycloheptanone	0 436	0 86	0 00	0 56	4 376
Methyl formate	0 192	0 68	0 00	0 38	1 285
Ethyl formate	0 146	0 66	0 00	0 38	1 845
<i>n</i> -Propyl formate	0 132	0 63	0 00	0 38	2 433
<i>n</i> -Butyl formate	0 121	0 63	0 00	0 38	2 958
<i>n</i> -Pentyl formate	0 101	0 63	0 00	0 38	3 488
Methyl acetate	0 142	0 64	0 00	0 45	1 911
Ethyl acetate	0 106	0 62	0 00	0 45	2 314
<i>n</i> -Propyl acetate	0 092	0 60	0 00	0 45	2 819
<i>n</i> -Butyl acetate	0 071	0 60	0 00	0 45	3 353
<i>n</i> -Pentyl acetate	0 067	0 60	0 00	0 45	3 844
<i>n</i> -Hexyl acetate	0 056	0 60	0 00	0 45	4 351
<i>n</i> -Heptyl acetate	0 050	0 60	0 00	0 45	4 865
<i>n</i> -Octyl acetate	0 029	0 60	0 00	0 45	5 364
Vinyl acetate	0 223	0 64	0 00	0 43	2 152
Allyl acetate	0 199	0 72	0 00	0 49	2 723
Acetonitrile	0 237	0 90	0 07	0 32	1 739
Propionitrile	0 162	0 90	0 02	0 36	2 082
1-Cyanopropane	0 188	0 90	0 00	0 36	2 548
1-Cyanobutane	0 177	0 90	0 00	0 36	3 108
1-Cyanopentane	0 166	0 90	0 00	0 36	3 608
1-Cyanoheptane	0 159	0 90	0 00	0 36	4 089
1-Cyanoheptane	0 162	0 90	0 00	0 36	4 585
1-Cyanooctane	0 159	0 90	0 00	0 36	4 970
1-Cyanononane	0 156	0 90	0 00	0 36	5 460
1-Cyanodecane	0 154	0 90	0 00	0 36	5 940
Ammonia	0 139	0 35	0 14	0 62	0 680
Methylamine	0 250	0 35	0 16	0 58	1 300
Ethylamine	0 236	0 35	0 16	0 61	1 677
<i>n</i> -Propylamine	0 225	0 35	0 16	0 61	2 141
<i>n</i> -Butylamine	0 224	0 35	0 16	0 61	2 618
<i>n</i> -Pentylamine	0 211	0 35	0 16	0 61	3 139
<i>n</i> -Hexylamine	0 197	0 35	0 16	0 61	3 655
<i>n</i> -Octylamine	0 187	0 35	0 16	0 61	4 520
Dimethylamine	0 189	0 30	0 08	0 66	1 600
Diethylamine	0 154	0 30	0 08	0 69	2 395
Di- <i>n</i> -propylamine	0 124	0 30	0 08	0 69	3 351
Di- <i>n</i> -butylamine	0 107	0 30	0 08	0 69	4 349
Di- <i>n</i> -pentylamine	0 099	0 30	0 08	0 69	4 570
Trimethylamine	0 140	0 20	0 00	0 67	1 620
Triethylamine	0 101	0 15	0 00	0 79	3 040
Nitromethane	0 313	0 95	0 06	0 31	1 892
Nitroethane	0 270	0 95	0 02	0 33	2 414
1-Nitropropane	0 242	0 95	0 00	0 31	2 894
1-Nitrobutane	0 227	0 95	0 00	0 29	3 415
1-Nitropentane	0 212	0 95	0 00	0 29	3 938



Table 8 Continued

Solute	$R_2$	$\pi_2^H$	$\sum a_2^H$	$\sum \beta_2^H$	$\log L^{16}$
1-Nitrohexane	0.203	0.95	0.00	0.29	4.416
Nitrocyclohexane	0.441	0.97	0.00	0.31	4.826
Formamide	0.468	1.30	0.62	0.60	
Acetamide	0.460	1.30	0.54	0.68	
Propionamide	0.440	1.30	0.55	0.68	
Butanamide	0.420	1.30	0.56	0.68	
<i>N</i> -Methylformamide	0.405	1.30	0.40	0.55	
<i>N</i> -Methylacetamide	0.400	1.30	0.40	0.72	
<i>N</i> -Methylpropanamide	0.380	1.30	0.40	0.71	
<i>N,N</i> -Dimethylformamide	0.367	1.31	0.00	0.74	3.173
<i>N,N</i> -Dimethylacetamide	0.363	1.33	0.00	0.78	3.717
Formic acid	0.300	0.60	0.75	0.38	
Acetic acid	0.265	0.65	0.61	0.44	1.750
Propanoic acid	0.233	0.65	0.60	0.45	2.290
Butanoic acid	0.210	0.62	0.60	0.45	2.830
Pentanoic acid	0.205	0.60	0.60	0.45	3.380
Hexanoic acid	0.174	0.60	0.60	0.45	3.920
Heptanoic acid	0.149	0.60	0.60	0.45	4.460
Octanoic acid	0.150	0.60	0.60	0.45	5.000
Nonanoic acid	0.132	0.60	0.60	0.45	5.550
Decanoic acid	0.124	0.60	0.60	0.45	6.090
Undecanoic acid	0.100	0.60	0.60	0.45	6.640
Dodecanoic acid	0.083	0.60	0.60	0.45	7.180
Chloroacetic acid	0.373	1.08	0.74	0.36	
Dichloroacetic acid	0.481	1.20	0.90	0.27	
Trichloroacetic acid	0.589	1.33	0.95	0.28	
Water	0.000	0.45	0.82	0.35	0.260
Methanol	0.278	0.44	0.43	0.47	0.970
Ethanol	0.246	0.42	0.37	0.48	1.485
Propan-1-ol	0.236	0.42	0.37	0.48	2.031
Propan-2-ol	0.212	0.36	0.33	0.56	1.764
Butan-1-ol	0.224	0.42	0.37	0.48	2.601
Pentan-1-ol	0.219	0.42	0.37	0.48	3.106
Hexan-1-ol	0.210	0.42	0.37	0.48	3.610
Heptan-1-ol	0.211	0.42	0.37	0.48	4.115
Octan-1-ol	0.199	0.42	0.37	0.48	4.619
Nonan-1-ol	0.193	0.42	0.37	0.48	5.124
Decan-1-ol	0.191	0.42	0.37	0.48	5.628
Undecan-1-ol	0.181	0.42	0.37	0.48	6.130
Dodecan-1-ol	0.175	0.42	0.37	0.48	6.640
Cyclopentanol	0.427	0.54	0.32	0.56	3.241
Cyclohexanol	0.460	0.54	0.32	0.57	3.758
Cycloheptanol	0.513	0.54	0.32	0.58	4.407
Adamantan-1-ol	0.850	1.20	0.32	0.56	
Prop-2-en-1-ol	0.342	0.46	0.38	0.48	1.951
<i>trans</i> -But-2-en-1-ol	0.350	0.44	0.38	0.48	2.618
2,2,2-Trifluoroethanol	0.015	0.60	0.57	0.25	1.224
Hexafluoropropan-1-ol	-0.240	0.55	0.77	0.10	1.392
Dodecafluoroheptan-1-ol	-0.640	0.50	0.65	0.22	3.089
Ethylthiol	0.392	0.35	0.00	0.24	2.173
<i>n</i> -Propylthiol	0.385	0.35	0.00	0.24	2.685
<i>n</i> -Butylthiol	0.382	0.35	0.00	0.24	3.111
<i>n</i> -Pentylthiol	0.369	0.35	0.00	0.24	3.624
<i>n</i> -Hexylthiol	0.361	0.35	0.00	0.24	4.133
<i>n</i> -Heptylthiol	0.357	0.35	0.00	0.24	4.635
<i>n</i> -Octylthiol	0.353	0.35	0.00	0.24	5.270
<i>n</i> -Nonylthiol	0.347	0.35	0.00	0.24	5.790
<i>n</i> -Decylthiol	0.342	0.35	0.00	0.24	6.318
Allyl thiol	0.542	0.41	0.00	0.20	2.654
Dimethyl sulfide	0.404	0.38	0.00	0.29	2.238
Diethyl sulfide	0.373	0.38	0.00	0.32	3.104
Di- <i>n</i> -propyl sulfide	0.358	0.38	0.00	0.32	4.120
Di- <i>n</i> -butyl-sulfide	0.345	0.38	0.00	0.32	4.950
Sulfur hexafluoride	-0.600	-0.20	0.00	0.00	-0.120
Carbon disulfide	0.877	0.21	0.00	0.07	2.353
Trimethyl phosphate	0.113	1.10	0.00	1.00	
Triethyl phosphate	0.000	1.00	0.00	1.06	4.750
Tri- <i>n</i> -propyl phosphate	-0.050	1.00	0.00	1.15	
Tri- <i>n</i> -butyl phosphate	-0.100	0.90	0.00	1.21	
Benzene	0.610	0.52	0.00	0.14	2.786
Toluene	0.601	0.52	0.00	0.14	3.325
Ethylbenzene	0.613	0.51	0.00	0.15	3.778
<i>o</i> -Xylene	0.663	0.56	0.00	0.16	3.939
<i>m</i> -Xylene	0.623	0.52	0.00	0.16	3.839

Table 8 Continued

Solute	$R_2$	$\pi_2^H$	$\sum a_2^H$	$\sum \beta_2^H$	$\log L^{16}$
<i>p</i> -Xylene	0 613	0 52	0 00	0 16	3 839
<i>n</i> -Propylbenzene	0 604	0 50	0 00	0 15	4 230
<i>n</i> -Butylbenzene	0 600	0 51	0 00	0 15	4 730
1,3,5-Trimethylbenzene	0 649	0 52	0 00	0 19	4 344
Pentamethylbenzene	0 850	0 66	0 00	0 20	5 798
Hexamethylbenzene	0 950	0 72	0 00	0 21	6 557
Styrene	0 849	0 65	0 00	0 16	3 856
Phenylethyne	0 679	0 58	0 12	0 24	3 692
Allyl benzene	0 717	0 60	0 00	0 22	4 136
Diphenylmethane	1 220	1 04	0 00	0 28	6 313
Biphenyl	1 360	0 99	0 00	0 22	6 014
Naphthalene	1 340	0 92	0 00	0 20	5 161
Anthracene	2 290	1 34	0 00	0 26	7 568
Phenanthrene	2 055	1 29	0 00	0 26	7 632
Fluorobenzene	0 477	0 57	0 00	0 10	2 788
Chlorobenzene	0 718	0 65	0 00	0 07	3 657
1,2-Dichlorobenzene	0 872	0 78	0 00	0 04	4 518
1,3-Dichlorobenzene	0 847	0 73	0 00	0 02	4 410
1,4-Dichlorobenzene	0 825	0 75	0 00	0 02	4 435
Benzyl chloride	0 821	0 82	0 00	0 33	4 384
Bromobenzene	0 882	0 73	0 00	0 09	4 041
1,2-Dibromobenzene	1 190	0 96	0 00	0 04	5 456
1,3-Dibromobenzene	1 170	0 88	0 00	0 04	5 327
1,4-Dibromobenzene	1 150	0 86	0 00	0 04	5 324
Benzyl bromide	1 014	0 98	0 00	0 20	4 672
Iodobenzene	1 188	0 82	0 00	0 12	4 502
Methylphenylether	0 708	0 75	0 00	0 29	3 890
Ethylphenylether	0 681	0 70	0 00	0 32	4 242
Benzaldehyde	0 820	1 00	0 00	0 39	4 008
Acetophenone	0 818	1 01	0 00	0 48	4 501
Ethylphenylketone	0 804	0 95	0 00	0 51	4 971
Benzophenone	1 447	1 50	0 00	0 50	
Methyl benzoate	0 733	0 85	0 00	0 46	4 704
Ethyl benzoate	0 689	0 85	0 00	0 46	5 075
<i>n</i> -Propyl benzoate	0 675	0 80	0 00	0 46	5 718
<i>n</i> -Butyl benzoate	0 668	0 80	0 00	0 46	6 210
Phenyl acetate	0 661	1 13	0 00	0 54	4 414
Dimethyl phthalate	0 780	1 41	0 00	0 88	6 051
Benzonitrile	0 742	1 11	0 00	0 33	4 039
Benzylamine	0 829	0 88	0 10	0 72	4 319
Nitrobenzene	0 871	1 11	0 00	0 28	4 557
2-Nitrotoluene	0 866	1 11	0 00	0 27	4 878
3-Nitrotoluene	0 874	1 10	0 00	0 25	5 097
4-Nitrotoluene	0 870	1 11	0 00	0 28	5 154
Benzamide	0 990	1 50	0 49	0 67	5 767
Formanilide	0 970	1 40	0 50	0 50	
Acetanilide	0 870	1 40	0 50	0 67	
Benzoic acid	0 730	0 90	0 59	0 40	
2-Methylbenzoic acid	0 730	0 90	0 60	0 34	
3-Methylbenzoic acid	0 730	0 90	0 59	0 38	
4-Methylbenzoic acid	0 730	0 90	0 60	0 38	
Phenol	0 805	0 89	0 60	0 30	3 766
<i>o</i> -Cresol	0 840	0 86	0 52	0 30	4 218
<i>m</i> -Cresol	0 822	0 88	0 57	0 34	4 310
<i>p</i> -Cresol	0 820	0 87	0 57	0 31	4 312
2-Fluorophenol	0 660	0 69	0 61	0 26	3 453
3-Fluorophenol	0 667	0 98	0 68	0 17	3 842
4-Fluorophenol	0 670	0 97	0 63	0 23	3 844
2-Chlorophenol	0 853	0 88	0 32	0 31	4 178
3-Chlorophenol	0 909	1 06	0 69	0 15	4 773
4-Chlorophenol	0 915	1 08	0 67	0 20	4 775
2-Bromophenol	1 037	0 90	0 35	0 31	4 526
3-Bromophenol	1 060	1 15	0 70	0 16	5 144
4-Bromophenol	1 080	1 17	0 67	0 20	5 135
2-Iodophenol	1 360	1 00	0 40	0 35	4 964
3-Iodophenol	1 370	1 20	0 70	0 18	5 528
4-Iodophenol	1 380	1 22	0 68	0 20	5 492
2-Methoxyphenol	0 837	0 91	0 22	0 52	4 449
3-Methoxyphenol	0 879	1 17	0 59	0 39	4 803
4-Methoxyphenol	0 900	1 17	0 57	0 48	4 773
2-Cyanophenol	0 920	1 33	0 74	0 33	4 531
3-Cyanophenol	0 930	1 55	0 77	0 28	5 181
4-Cyanophenol	0 940	1 63	0 79	0 29	5 420
2-Nitrophenol	1 015	1 05	0 05	0 37	4 760

Table 8 Continued

Solute	$R_2$	$\pi_2^H$	$\sum a_2^H$	$\sum \beta_2^H$	$\log L^{16}$
3-Nitrophenol	1 050	1 57	0 79	0 23	5 692
4-Nitrophenol	1 070	1 72	0 82	0 26	5 876
1-Naphthol	1 520	1 05	0 61	0 37	6 130
2-Naphthol	1 520	1 08	0 61	0 40	6 200
Benzyl alcohol	0 803	0 87	0 33	0 56	4 221
Thiophenol	1 000	0 80	0 09	0 16	4 110
Benzenesulfonamide	1 130	1 55	0 55	0 80	
N-Methylbenzenesulfonamide	1 100	1 50	0 30	0 82	
N,N-Dimethylbenzenesulfonamide	1 100	1 50	0 00	0 86	
Furan	0 369	0 53	0 00	0 13	1 830
Benzofuran	0 888	0 83	0 00	0 15	4 355
Tetrahydrofuran	0 289	0 52	0 00	0 48	2 636
2-Methyltetrahydrofuran	0 241	0 48	0 00	0 53	2 820
1,4-Dioxane	0 329	0 75	0 00	0 64	2 892
Benzodioxan	0 874	1 07	0 00	0 35	4 971
Paraldehyde	0 136	0 68	0 00	0 68	3 169
Piperidine	0 422	0 46	0 10	0 69	3 304
N-Methylpiperidine	0 318	0 39	0 00	0 70	3 330
N-Ethylpiperidine	0 300	0 32	0 00	0 63	3 729
Pyrrrole	0 613	0 73	0 41	0 29	2 865
N-Methylpyrrrole	0 559	0 79	0 00	0 31	2 923
Pyrazine	0 629	0 95	0 00	0 62	2 920
2-Methylpyrazine	0 629	0 90	0 00	0 64	3 254
Pyrimidine	0 606	1 00	0 00	0 65	2 837
Thiophene	0 687	0 56	0 00	0 15	2 819
Benzo(b)thiophene	1 323	0 88	0 00	0 20	5 174
Dibenzothiophen	1 959	1 31	0 00	0 18	7 575
Thiazole	0 800	0 80	0 00	0 45	

or triple) are counted the same. For complicated molecules, it is time consuming to count the number of bonds, but I find that this number is given by the algorithm,

$$B = N - 1 + R$$

where  $B$  is the number of bonds in the molecule,  $N$  is the total number of atoms, and  $R$  is the number of rings. Thus for cyclohexane or benzene  $R = 1$ , and for cyclohexylbenzene, diphenyl, or naphthalene  $R = 2$ . So far, I have found no exception to this rule. Because  $V_x$  is so easily calculated, values are given only in Table 6, partly as examples and partly to list those for the rare gases.

## 5 References

- 1 A D Sherry and K F Purcell, *J Phys Chem*, 1970, **74**, 3535, A D Sherry and K F Purcell, *J Am Chem Soc*, 1972, **94**, 1853
- 2 A B Ioghansen, *Theor Exp Khim (USSR)*, 1971, **7**, 302
- 3 O A Raevsky and V P Novikov, *Khimico-farm Z*, 1982, **16**, 583, O A Raevsky, V V Avidon, and V P Novikov, *Khimico-farm Z*, 1982, **16**, 968, O A Raevsky, V Yu Grigoriev, and V P Soloviov, *Khimico-farm Z*, 1984, **18**, 578, O A Raevsky, in 'QSAR in Drug Design and Toxicology', ed D Hadzi and B Jerman-Balzac. Elsevier, Amsterdam, 1987
- 4 R W Taft, D Gurka, L Joris, P von R Schleyer, and J W Rakshys, *J Am Chem Soc*, 1969, **91**, 4801, J Mitsky, L Joris, and R W Taft, *J Am Chem Soc*, 1972, **94**, 3438
- 5 M H Abraham, P P Duce, P L Grellier, D V Prior, J J Morris, and P J Taylor, *Tetrahedron Lett*, 1988, **29**, 1587, M H Abraham, P L Grellier, D V Prior, P P Duce, J J Morris, and P J Taylor, *J Chem Soc Perkin Trans 2*, 1989, 699
- 6 M H Abraham, P L Grellier, D V Prior, J J Morris, P J Taylor, C Lawrence, and M Berthelot, *Tetrahedron Lett*, 1989, **30**, 2571, M H Abraham, P L Grellier, D V Prior, J J Morris, and P J Taylor, *J Chem Soc Perkin Trans 2*, 1990, 521
- 7 M H Abraham, P L Grellier, D V Prior, R W Taft, J J Morris, P J Taylor, C Lawrence, M Berthelot, R M Doherty, M J Kamlet, J-L M Abboud, K Sraidi, and G Guiheneuf, *J Am Chem Soc*, 1988, **110**, 8534
- 8 M H Abraham, P L Grellier, D V Prior, J J Morris, P J Taylor, and R M Doherty, *J Org Chem*, 1990, **55**, 2227
- 9 J Hine, S Hahn, and J Hwang, *J Org Chem*, 1988, **53**, 884
- 10 J-L M Abboud, K Sraidi, M H Abraham, and R W Taft, *J Org Chem*, 1990, **55**, 2230
- 11 M H Abraham, P P Duce, D V Prior, D G Barrett, J J Morris, and P J Taylor, *J Chem Soc, Perkin Trans 2*, 1989, 1355
- 12 C Laurence, M Berthelot, E Raczynska, J-Y Le Questel, G Duguay, and P Hudhomme, *J Chem Res (s)*, 1990, 250
- 13 M H Abraham, G S Whiting, R M Doherty, and W J Shuely, *J Chromatogr*, 1991, **587**, 213, 229
- 14 M H Abraham, G S Whiting, R M Doherty, and W J Shuely, *J Chem Soc, Perkin Trans 2*, 1990, 1451
- 15 M H Abraham, P L Grellier, and R A McGill, *J Chem Soc, Perkin Trans 2*, 1987, 797
- 16 M H Abraham and J C McGowan, *Chromatographia*, 1987, **23**, 243
- 17 F Patte, M Etcheto, and P Laffort, *Anal Chem*, 1982, **54**, 2239
- 18 W O McReynolds, 'Gas Chromatographic Retention Data', Preston Technical Abstracts, Evanston, Ill, 1966
- 19 B R Kersten, S K Poole, and C F Poole, *J Chromatogr*, 1990, **500**, 329
- 20 M H Abraham, G S Whiting, R M Doherty, and W J Shuely, *J Chem Soc, Perkin Trans 2*, 1990, 1851
- 21 M H Abraham, G S Whiting, Y Alarie, J J Morris, P J Taylor, R M Doherty, R W Taft, and G Nielsen, *Quant Structure-Activity Relat*, 1990, **9**, 6
- 22 M H Abraham, G S Whiting, R M Doherty, W J Shuely, and P Sakellianou, *Polymer*, 1992, **33**, 2162
- 23 M H Abraham and P K Weathersby, unpublished results
- 24 M H Abraham and R Fuchs, *J Chem Soc, Perkin Trans 2*, 1988, 523
- 25 R A Pierotti, *Chem Rev*, 1976, **76**, 717
- 26 M H Abraham, G S Whiting, R Fuchs, and E J Chambers, *J Chem Soc, Perkin Trans 2*, 1990, 291
- 27 N P Franks and W R Lieb, *Nature*, 1984, **310**, 599, 1985, **316**, 349
- 28 M H Abraham, W R Lieb, and N P Franks, *J Pharm Sci*, 1991, **80**, 719
- 29 M H Abraham and G S Whiting, *J Chromatogr*, 1992, **594**, 229
- 30 D E Leahy, J J Morris, P J Taylor, and A R Wait, *J Chem Soc, Perkin Trans 2*, 1992, 705