PAPER

NJC
www.rsc.org/njc

Correlation and prediction of partition coefficients between the gas phase and water, and the solvents dodecane and undecane

Michael H. Abraham*[a] and William E. Acree Jr.[b]

[a] *The Department of Chemistry, University College London, 20 Gower Street, London UK WC1H OAJ. E-mail: m.h.abraham@ucl.ac.uk*

[b] *Department of Chemistry, P.O. Box 305070, University of North Texas, Denton TX 76203-5070 USA*

Received (in Durham, UK) 23rd July 2004, Accepted 22nd September 2004
First published as an Advance Article on the web 29th October 2004

Literature values of water/dodecane and gas/dodecane partition coefficients, P_{12} and K_{12}, supplemented by some measurements of our own have led to a data base of partition coefficients for 118 solutes. Application of the Abraham solvation equations to the sets of partition coefficients leads to equations that correlate log P_{12} to 0.20 units and log K_{12} to 0.11 units. Consideration of results on independent test sets suggests that log P_{12} could be predicted to 0.2 to 0.3 units, and that log K_{12} could be predicted to 0.1 to 0.2 units. A smaller data set of 66 solutes has been assembled for undecane. Similar equations lead to correlations for log P_{11} and for log K_{11} that indicate these quantities could be predicted to 0.2 to 0.3 log units.

Introduction

Dodecane is an important solvent used in a number of processes as diverse as the extraction of carboxylic acids from water,[1] and the extraction of carotenoids from living cells of *Dunaliella salina*.[2] The Medicinal Chemistry Data Base[3] records numerous partition coefficients for the water/dodecane system, and systematic studies on the water/dodecane partition of drugs,[4] and alcohols and glycol monoethers[5] have been reported. Castells *et al.*[6] have determined activity coefficients for a number of hydrocarbons in dodecane, and Hesse *et al.*[7] have measured gas/dodecane partition coefficients for a range of gases. It seemed timely to collect the available literature data on both water/dodecane and gas/dodecane partition coefficients and to develop correlation equations that could also be used for predictive purposes. Although undecane has not been used as often as dodecane, we thought it useful to collect extraction data in order to compare equations with those for dodecane, and for the lower alkanes.

Gas/dodecane or undecane partition coefficients, K_{12} and K_{11} can be defined through eqn. (1), where C_{ALK} and C_G and are the equilibrium concentrations of a solute in an alkane and the gas phase respectively. If both C_{ALK} and C_G are in mol dm^{-3}, then K_{ALK} is unitless. The water/alkane partition coefficients, P_{12} and P_{11}, and the gas/alkane partition coefficients can be converted into each other through eqn. (2), where K_W is the (unitless) gas/water partition coefficient. This is a particularly useful equation, because it greatly increases the number of values of the partition coefficients that are available.

$$K_{ALK} = C_{ALK}/C_G \quad (1)$$

$$P_{ALK} = K_{ALK}/K_W \quad (2)$$

Methodology

We have already set out[8,9] two general equations that have proved very useful in the correlation and prediction of water/solvent and gas/solvent partition coefficients, as log P and log K, respectively. Eqn. (3) is used for water/solvent partitions, and eqn. (4) is used for gas/solvent partitions.

$$SP = c + e.E + s.S + a.A + b.B + v.V \quad (3)$$

$$SP = c + e.E + s.S + a.A + b.B + l.L \quad (4)$$

In eqn. (3) and eqn. (4), the dependent variable, SP, is a set of solute properties in a given system. Eqn. (3) is used for processes in condensed systems; in the present case SP is log P_{12} or log P_{11} for a series of solutes. Eqn. (4) is used for gas-condensed phase processes; in the present case SP is log K_{12} or log K_{11} for a series of solutes. The independent variables in eqn. (3) and eqn. (4) are solute descriptors as follows.[9] E is the solute excess molar refractivity in units of $(dm^3\ mol^{-1})/10$, S is the solute dipolarity/polarizability, A and B are the overall or summation hydrogen bond acidity and basicity, V is the McGowan characteristic volume in units of $(dm^3\ mol^{-1})/100$, and L is the logarithm of the gas-hexadecane partition coefficient at 298 K. The coefficients in eqn. (3) and eqn. (4) are evaluated through multiple linear regression.

We have already applied eqn. (3) and eqn. (4) to many of the alkane solvents used in extraction processes,[10] and in Table 1 is given a summary of the coefficients. These vary very regularly with the size of the alkane, and so it is of some interest to see if the equations relating to dodecane fit into the pattern shown in Table 1.

Results and discussion

The procedure is very simple. Experimental values of log P_{12} are used as such in eqn. (3), or are converted into corresponding values of log K_{12} *via* eqn. (2), for use in eqn. (4). Conversely, experimental values of log K_{12} are used as such in eqn. (4) or are converted into values of P_{12} for use in eqn. (3). In addition, we use a method that we have developed[11–15] where the solubility of a solute in dodecane, at 298 K in mol dm^{-3}, is the prime experimental datum. If the saturated vapor pressure of the solute at 298 K is available, then log K_{12} can easily be deduced. As before, log K_{12} can then be converted into log P_{12} *via* eqn. (2) if the gas/water partition coefficient is available. Details of these determinations are in Table 2; values of log C_G and log K_w are the same as those used previously for

DOI: 10.1039/b411303a

Table 1 Coefficients in eqns. (3) and (4) for water/alkane and gas/alkane partitions

Alkane	c	e	s	a	b	v
Pentane	0.369	0.386	−1.568	−3.535	−5.215	4.514
Hexane	0.361	0.579	−1.723	−3.599	−4.764	4.344
Heptane	0.325	0.670	−2.061	−3.317	−4.733	4.543
Octane	0.223	0.642	−1.647	−3.480	−5.067	4.526
Nonane	0.240	0.619	−1.713	−3.532	−4.921	4.482
Decane	0.160	0.585	−1.734	−3.435	−5.078	4.582
Undecane	0.056	0.603	−1.661	−3.421	−5.120	4.619
Dodecane	0.114	0.668	−1.644	−3.545	−5.006	4.459
Hexadecane	0.087	0.667	−1.617	−3.587	−4.869	4.433

	c	e	s	a	b	l
Pentane	0.335	−0.276	0.000	0.000	0.000	0.968
Hexane	0.292	−0.169	0.000	0.000	0.000	0.979
Heptane	0.275	−0.162	0.000	0.000	0.000	0.983
Octane	0.215	−0.049	0.000	0.000	0.000	0.967
Nonane	0.200	−0.145	0.000	0.000	0.000	0.980
Decane	0.156	−0.143	0.000	0.000	0.000	0.989
Undecane	0.113	0.000	0.000	0.000	0.000	0.971
Dodecane	0.053	0.000	0.000	0.000	0.000	0.986
Hexadecane	0.000	0.000	0.000	0.000	0.000	1.000

ferrocene,[12] *trans*-stilbene,[11] acenaphthene,[14] anthracene,[13] phenanthrene,[13] pyrene[14] and benzil.[15]

The data that we use are listed in Table 3, where the reference refers to the original experimental value used. There were 119 compounds for which we had all the necessary data to obtain log P_{12} and log K_{12}. One of these, acridine, was an outlier in preliminary calculations, the observed[4] value of 2.32 for log P_{12} being over 0.6 log units smaller than that calculated. Since the observed value was not compatible with log P values in other solvents[3] we omitted acridine, and were left with 118 compounds as shown in Table 3.

In order to ascertain the predictive capability of any equations developed for log P_{12}, we first divided the 118 compounds into a training set of 59 compounds and a test set of 59 compounds, using the Kennard–Stone method[33] of choosing the sets. Application of eqn. (3) to the 59 compound training set resulted in eqn. (5)

$$\text{Log } P_{12} = 0.109 + 0.713\,E - 1.666\,S - 3.531\,A - 5.041\,B + 4.457\,V \quad (5)$$

$$\text{N} = 59,\ \text{SD} = 0.211,\ \text{R}^2 = 0.9983,\ \text{R} = 0.9992,\ \text{F} = 6421$$

Here and elsewhere N is the number of data points, that is the number of solutes, SD is the standard deviation, R is the correlation coefficient and F is the Fischer F-statistic. The statistics of eqn. (5) are excellent, and so we proceeded to predict log P_{12} values for the 59 compounds in the test set, not used to obtain eqn. (5). Results are in Table 4, in terms of the average error, AE, the average absolute error, AAE, and the standard deviation of predicted and observed values. It is clear that eqn. (5) can be used to predict log P_{12} values to about 0.2 log units.

Table 2 Calculation of log K_{12} and log P_{12} from solubilities in dodecane (as log C_{12}) at 298 K

Solute	Log C_{12}	Log C_G	Log K_{12}	Log K_W	Log P_{12}
Ferrocene	−0.83	−6.39	5.57	1.92	3.65
Trans-Stilbene	−1.13	−8.58	7.45	2.78	4.67
Acenaphthene	−0.40	−6.90	6.50	2.36	4.14
Anthracene	−1.91	−9.46	7.55	3.03	4.52
Phenanthrene	−0.55	−7.97	7.42	2.80	4.62
Pyrene	−0.97	−9.65	8.68	3.50	5.18
Benzil	−1.385	−8.89	7.51	4.87	2.64

If all 118 compounds are used, then eqn. (6) results, and we suggest that eqn. (6) be used for prediction of further log P_{12} values. The coefficients in eqn. (6) are the same as those in eqn. (5), within experimental error, as are the associated statistics.

$$\text{Log } P_{12} = 0.114 + 0.668\,E - 1.644\,S - 3.545\,A - 5.006\,B + 4.459\,V \quad (6)$$

$$N = 118,\ \text{SD} = 0.186,\ R^2 = 0.9978,\ R = 0.9989,\ F = 9932$$

Comparison of the coefficients in eqn. (6) with those for partition from water into various alkanes, see Table 1, shows that they are very reasonable. As expected, the factors that influence partition into dodecane are qualitatively and almost quantitatively the same as those influencing partition into the other alkanes.

The 118 values of log K_{12} were also divided into a training set and a test set using the Kennard–Stone method.[33] Application of eqn. (4) to the 59 compound training set resulted in eqn. (7). Once again, this 'training' equation was used to predict values of log K_{12} for the 59 compound test set. Details are in Table 4. The statistics of eqn. (7) are so good that log K_{12} can be predicted to within 0.1 log unit. If the test and training sets are combined, eqn. (8) results. The statistics and coefficients of the two equations, eqns. (7) and (8) are almost identical.

$$\text{Log } K_{12} = 0.075 - 0.029\,E - 0.076\,S - 0.120\,A - 0.005\,B + 0.997\,L \quad (7)$$

$$N = 59,\ \text{SD} = 0.113,\ R^2 = 0.9993,\ R = 0.9997,\ F = 16451$$

$$\text{Log } K_{12} = 0.082 - 0.007\,E - 0.081\,S - 0.160\,A - 0.002\,B + 0.996\,L \quad (8)$$

$$N = 118,\ \text{SD} = 0.100,\ R^2 = 0.9993,\ R = 0.9997,\ F = 33956$$

In both eqn. (7) and eqn. (8), the coefficients of E, S, A and B are hardly significant. Furthermore, for solution of a gaseous

Table 3 Values of log P_{12} and log K_{12} for partition into dodecane, and descriptors for 118 solutes, at 298 K

Solute	Log P_{12}	Log K_{12}	Log K_W	E	S	A	B	V	L	Ref.
Helium	0.45	−1.57	−2.02	0.00	0.00	0.00	0.00	0.0680	−1.741	7
Neon	0.53	−1.43	−1.96	0.00	0.00	0.00	0.00	0.0850	−1.575	7
Argon	0.92	−0.55	−1.47	0.00	0.00	0.00	0.00	0.1900	−0.680	7
Krypton	1.12	−0.09	−1.21	0.00	0.00	0.00	0.00	0.2460	−0.211	7
Xenon	1.50	0.53	−0.97	0.00	0.00	0.00	0.00	0.3290	0.378	16
Hydrogen	0.61	−1.11	−1.72	0.00	0.00	0.00	0.00	0.1086	−1.200	17
Oxygen	0.90	−0.61	−1.51	0.00	0.00	0.00	0.00	0.1100	−0.728	7
Nitrogen	0.95	−0.85	−1.80	0.00	0.00	0.00	0.00	0.2222	−0.978	7
Carbon monoxide	0.84	−0.78	−1.62	0.00	0.00	0.00	0.04	0.2220	−0.836	17
Carbon dioxide	0.27	0.19	−0.08	0.00	0.28	0.05	0.10	0.2809	0.058	18
Methane	1.23	−0.23	−1.46	0.00	0.00	0.00	0.00	0.2495	−0.323	19
Ethane	1.92	0.58	−1.34	0.00	0.00	0.00	0.00	0.3904	0.492	19
Propane	2.55	1.11	−1.44	0.00	0.00	0.00	0.00	0.5313	1.050	17
Butane	3.18	1.66	−1.52	0.00	0.00	0.00	0.00	0.6722	1.615	17
Pentane	3.96	2.26	−1.70	0.00	0.00	0.00	0.00	0.8131	2.162	6
2-Methylbutane	3.86	2.11	−1.75	0.00	0.00	0.00	0.00	0.8131	2.013	20
2,2-Dimethylpropane	3.64	1.80	−1.84	0.00	0.00	0.00	0.00	0.8131	1.820	21
Hexane	4.55	2.73	−1.82	0.00	0.00	0.00	0.00	0.9540	2.668	6
2-Methylpentane	4.41	2.57	−1.84	0.00	0.00	0.00	0.00	0.9540	2.503	6
Heptane	5.20	3.24	−1.96	0.00	0.00	0.00	0.00	1.0949	3.137	6
2,4-Dimethylpentane	4.97	2.89	−2.08	0.00	0.00	0.00	0.00	1.0949	2.809	6
Octane	5.86	3.75	−2.11	0.00	0.00	0.00	0.00	1.2358	3.677	6
2,5-Dimethylhexane	5.41	3.39	−2.02	0.00	0.00	0.00	0.00	1.2358	3.308	6
2,3,4-Trimethylpentane	5.35	3.47	−1.88	0.00	0.00	0.00	0.00	1.2358	3.481	6
Nonane	6.40	4.25	−2.15	0.00	0.00	0.00	0.00	1.3767	4.182	6
Dodecane	8.37	5.84	−2.53	0.00	0.00	0.00	0.00	1.7994	5.696	[a]
Octacosane	18.15	13.81	−4.34	0.00	0.00	0.00	0.00	4.0538	13.780	22
Dotriacontane	20.61	15.73	−4.88	0.00	0.00	0.00	0.00	4.6174	15.790	22
Hexatriacontane	23.21	17.86	−5.35	0.00	0.00	0.00	0.00	5.1810	17.740	22
Cyclopentane	3.35	2.47	−0.88	0.26	0.10	0.00	0.00	0.7045	2.477	23
Cyclohexane	3.85	2.95	−0.90	0.31	0.10	0.00	0.00	0.8454	2.964	6
Ethylcyclohexane	5.39	3.81	−1.58	0.26	0.10	0.00	0.00	1.1272	3.877	6
Ethene	1.31	0.37	−0.94	0.11	0.10	0.00	0.07	0.3474	0.289	17
Pent-1-ene	3.37	2.14	−1.23	0.09	0.08	0.00	0.07	0.7701	2.047	20
3-Methylbut-1-ene	3.31	1.97	−1.34	0.06	0.06	0.00	0.05	0.7701	1.933	20
2-Methylbut-1,3-diene	2.65	2.20	−0.45	0.31	0.23	0.00	0.10	0.7271	2.101	20
cis-Penta-1,3-diene	2.69	2.33	−0.36	0.35	0.23	0.00	0.10	0.7271	2.280	20
trans-Penta-1,3-diene	2.67	2.31	−0.36	0.35	0.23	0.00	0.10	0.7271	2.250	20
Carbon tetrafluoride	1.57	−0.72	−2.29	0.58	−0.26	0.00	0.00	0.3203	−0.819	7
bis(2,2,2-trifluoroethyl)ether	1.34	1.47	0.13	−0.55	−0.10	0.14	0.39	0.8371	1.241	24
Acetone	−0.98	1.81	2.79	0.18	0.70	0.04	0.51	0.5470	1.696	25
Pentane-2,4-dione	−0.16	3.39	3.55	0.41	0.78	0.00	0.63	0.8445	3.335	26
Hexane-2,4-dione	0.38	3.88	3.50	0.41	0.70	0.07	0.63	0.9854	3.728	26
Heptane-2,4-dione	1.07	4.32	3.25	0.41	0.67	0.09	0.61	1.1263	4.256	26
Octane-2,4-dione	1.64	4.74	3.10	0.41	0.67	0.07	0.62	1.2672	4.753	26
Nonane-2,4-dione	2.29	5.29	3.00	0.41	0.66	0.08	0.62	1.4081	5.290	26
Octane-3,5-dione	1.57	4.77	3.20	0.41	0.67	0.09	0.62	1.2672	4.807	26
Nonane-4,6-dione	2.16	5.31	3.15	0.41	0.64	0.11	0.64	1.4081	5.218	26
Diethylcarbonate	0.58	3.23	2.65	0.06	0.58	0.00	0.53	0.9462	3.412	27
Ethanol	−2.16	1.51	3.67	0.25	0.42	0.37	0.48	0.4491	1.485	5
Propan-1-ol	−1.53	2.03	3.56	0.24	0.42	0.37	0.48	0.5900	2.031	5
Butan-1-ol	−0.87	2.59	3.46	0.22	0.42	0.37	0.48	0.7309	2.601	5
Pentan-1-ol	−0.24	3.11	3.35	0.22	0.42	0.37	0.48	0.8718	3.106	5
Hexan-1-ol	0.34	3.57	3.23	0.21	0.42	0.37	0.48	1.0127	3.610	5
Heptan-1-ol	0.83	3.92	3.09	0.21	0.42	0.37	0.48	1.1536	4.115	28
4-Octanol	1.55	4.41	2.86	0.16	0.36	0.33	0.56	1.2945	4.300	29
2-Ethyl-4-methylpentan-1-ol	1.41	4.09	2.68	0.20	0.36	0.37	0.48	1.2945	4.426	29
3-Ethyl-3-hexanol	1.35	4.20	2.85	0.20	0.30	0.31	0.60	1.2945	4.290	29
4-Ethyl-3-hexanol	1.37	4.32	2.95	0.17	0.36	0.33	0.57	1.2945	4.177	29
2-Ethoxyethanol	−2.10	2.81	4.91	0.24	0.52	0.31	0.81	0.7896	2.792	5
2-Propoxyethanol	−1.33	3.37	4.70	0.21	0.50	0.30	0.83	0.9305	3.310	5
2-Butoxyethanol	−0.73	3.86	4.59	0.20	0.50	0.30	0.83	1.0714	3.806	5
2-Pentoxyethanol	−0.13	4.25	4.48	0.19	0.50	0.30	0.82	1.2123	4.171	5
Sulfur Hexafluoride	2.18	−0.05	−2.23	0.60	−0.20	0.00	0.00	0.4643	−0.120	7
Trimethylphosphate	−2.53	3.99	6.52	0.11	1.10	0.00	1.00	0.9707	4.140	30
Triethylphosphate	−1.27	4.55	5.82	0.00	1.00	0.00	1.06	1.3934	4.750	30
Tripropylphosphate	0.61	6.01	5.40	0.05	1.00	0.00	1.15	1.8161	6.180	30
Tributylphosphate	2.40	7.40	5.00	0.10	0.90	0.00	1.21	2.2388	7.370	30
Tetramethyltin	4.26	2.73	−1.53	0.10	0.08	0.00	0.05	1.0431	2.745	21
Ferrocene	3.65	5.57	1.92	1.35	0.85	0.00	0.20	1.1209	5.622	[b]

Table 3 (*continued*)

Solute	Log P_{12}	Log K_{12}	Log K_W	E	S	A	B	V	L	Ref.
trans-Stilbene	4.67	7.45	2.78	1.45	1.05	0.00	0.34	1.5630	7.520	[b]
Acenaphthene	4.14	6.50	2.36	1.60	1.05	0.00	0.22	1.2586	6.469	[b]
Anthracene	4.52	7.55	3.03	2.29	1.34	0.00	0.28	1.4544	7.568	[b]
Phenanthrene	4.62	7.42	2.80	2.06	1.29	0.00	0.29	1.4544	7.632	[b]
Pyrene	5.18	8.68	3.50	2.81	1.71	0.00	0.28	1.5846	8.833	[b]
Anisole	2.06	3.86	1.80	0.71	0.75	0.00	0.29	0.9160	3.890	4
Acetophenone	0.94	4.30	3.36	0.82	1.01	0.00	0.48	1.0139	4.501	4
Acetophenone	1.12	4.48	3.36	0.82	1.01	0.00	0.48	1.0139	4.501	31
3-Methylacetophenone	1.72	5.21	3.49	0.81	1.00	0.00	0.51	1.1548	5.167	31
4-Methylacetophenone	1.63	5.08	3.45	0.84	1.00	0.00	0.52	1.1548	5.081	31
4-Fluoroacetophenone	1.20	4.58	3.38	0.70	1.02	0.00	0.47	1.0316	4.613	31
3-Chloroacetophenone	2.01	5.49	3.48	0.92	1.07	0.00	0.40	1.1363	5.341	31
4-Chloroacetophenone	1.85	5.07	3.22	0.96	1.09	0.00	0.44	1.1363	5.404	31
4-Methoxyacetophenone	0.90	5.94	5.04	0.92	1.58	0.00	0.53	1.2135	5.935	31
3-Nitroacetophenone	0.54	5.97	5.43	1.01	1.50	0.00	0.63	1.1881	5.951	31
4-Nitroacetophenone	0.70	6.08	5.38	1.08	1.54	0.00	0.59	1.1881	6.059	31
3-(CF_3) acetophenone	2.08	4.67	2.59	0.42	0.92	0.00	0.42	1.2079	4.736	31
Benzophenone	2.23	7.03	4.80	1.45	1.50	0.00	0.50	1.4808	6.852	32
Benzil	2.64	7.51	4.87	1.45	1.59	0.00	0.62	1.6374	7.611	[b]
Nitrobenzene	1.43	4.45	3.02	0.87	1.11	0.00	0.28	0.8906	4.557	4
Benzamide	−2.67	5.40	8.07	0.99	1.50	0.49	0.67	0.9728	5.767	4
Anthranilamide	−2.72	6.33	9.05	1.34	1.88	0.47	0.75	1.0726	6.435	4
Acetanilde	−1.76	5.42	7.18	0.90	1.39	0.48	0.67	1.1137	5.567	4
Phenol	−1.04	3.81	4.85	0.81	0.89	0.60	0.30	0.7751	3.766	4
3-Chlorophenol	−0.34	4.51	4.85	0.91	1.06	0.69	0.15	0.8975	4.773	4
Benzyl alcohol	−0.77	4.09	4.86	0.80	0.87	0.39	0.56	0.9160	4.221	4
1-Methylsulfonylacetophenone	−1.56	6.99	8.55	0.83	2.05	0.00	1.05	1.4357	7.071	4
Terfenadine desoxy analog	5.30	18.05	12.75	2.53	1.96	0.35	1.94	3.9545	18.180	3
Terfenadine keto analog	3.25	18.40	15.15	2.60	2.30	0.35	2.25	3.9701	18.640	3
3-Bromoquinoline	2.59	6.59	4.00	1.66	1.21	0.00	0.42	1.2193	6.551	4
Quinoxaline	0.51	5.51	5.00	1.27	1.19	0.00	0.59	1.0032	5.654	4
4-(Dimethylamino)antipyrine	−1.10	9.46	10.56	1.68	1.74	0.00	1.60	1.8662	9.484	4
Caffeine	−2.33	7.13	9.46	1.50	1.72	0.05	1.28	1.3632	7.352	4
Oxindole	−1.95	5.40	7.35	1.15	1.46	0.52	0.59	1.0051	5.563	4
1-Methylisatin	−0.90	6.41	7.31	1.25	1.86	0.00	0.80	1.1617	6.470	4
Diazepam	1.49	10.09	8.60	2.08	1.55	0.00	1.28	2.0739	10.480	4
Chlorpromazine	4.72	11.82	7.10	2.20	1.57	0.00	1.01	2.4056	11.840	3
Butylchlorpromazine	5.07	12.22	7.15	2.20	1.61	0.00	1.03	2.5465	12.380	3
Amitriptyline	4.49	11.24	6.75	1.92	1.40	0.00	1.06	2.3996	11.310	3
Clomipramine	4.81	12.16	7.35	1.98	1.70	0.00	1.04	2.5239	12.110	3
Diethyleneglycol ethyl ether	−2.58	4.02	6.60	0.22	0.72	0.21	1.21	1.1301	4.120	5
Diethyleneglycol propyl ether	−1.83	4.57	6.40	0.20	0.72	0.20	1.19	1.2710	4.640	5
Diethyleneglycol butyl ether	−1.17	5.03	6.20	0.17	0.72	0.20	1.19	1.4119	5.130	5
Diethyleneglycol pentyl ether	−0.60	5.40	6.00	0.17	0.72	0.20	1.17	1.5528	5.600	5
Triethyleneglycol butyl ether	−1.65	6.45	8.10	0.17	0.96	0.28	1.44	1.7524	6.474	5
Triethyleneglycol pentyl ether	−1.07	6.88	7.95	0.17	0.95	0.28	1.44	1.8933	6.930	5
Tetraethyleneglycol butyl ether	−2.13	7.77	9.90	0.18	1.20	0.32	1.72	2.0929	7.800	5
Tetraethyleneglyco lpentyl ether	−1.54	8.31	9.85	0.17	1.20	0.32	1.73	2.2338	8.340	5

[a] Taking $\gamma^{\infty} = 1$. [b] From Table 2.

solute in any solvent, the coefficients of A and B must either be positive or zero. Negative coefficients would imply that the solvent was a poorer hydrogen bond base and poorer hydrogen bond acid than the gas phase. Similarly, for solution of a gaseous solute in a hydrocarbon solvent, the s-coefficient would also be expected to be positive or zero. We therefore set all the statistically non-significant terms in eqn. (7) as zero, and obtained eqn. (9) for the same training set. There is very little difference in the statistics of eqn. (7) and eqn. (9), and very little difference in the predictions of the 59 compound test set, as shown in Table 4. For the predictions of the test set, eqn. (9) shows a small bias, AE = 0.046, as compared to the almost zero bias of the predictions from eqn. (7), AE = 0.019 only. However the goodness of prediction, as judged by the AAE (0.064 and 0.079) and by the SD (0.091 and 0.101), is almost the same. We therefore prefer eqn. (9) to eqn. (7) both on chemical grounds, and on the basis of simplicity. Finally, we examined the entire 118 data set, and obtained eqn. (10). We suggest this latter equation for prediction of further values of log K_{12}, if required.

$$\text{Log } K_{12} = 0.017 + 0.989\,L \quad (9)$$

$$N = 59,\ \text{SD} = 0.127,\ R^2 = 0.9991,\ R = 0.9996,\ F = 64052$$

$$\text{Log } K_{12} = 0.053 + 0.986\,L \quad (10)$$

$$N = 118,\ \text{SD} = 0.112,\ R^2 = 0.9991,\ R = 0.9996,\ F = 135942$$

Table 4 Predictions of the independent test sets

System	N	AE	AAE	SD
Water/dodecane	59	0.027	0.118	0.167
Gas/dodecane (7)	59	0.019	0.064	0.091
Gas/dodecane (9)	59	0.046	0.079	0.101

Although it is true that log P_{12} and log K_{12} are relatively easy to correlate through the general equations, eqn. (3) and eqn. (4), the results can still be regarded as excellent. Eqn. (6) correlates 118 values of log P_{12}, covering a range of no less than 26 log units with an SD of 0.19 log unit, and eqn. (10) correlates 118 values of log K_{12}, covering a range of 20 log units with an SD of 0.11 log unit. We regard these SD values, and the corresponding predictive SD values of 0.17 and 0.10 log unit, as rather too small, and suggest that eqn. (6) could be used to predict log P_{12} values to around 0.2 to 0.3 log unit, and that eqn. (10) could be used to predict log K_{12} values to around 0.1 to 0.2 log unit.

The method of dealing with undecane as an extractant follows exactly the same lines. We assembled data on gas/undecane partition coefficients and on water/undecane partition coefficients, treated these through eqn. (1) and eqn. (2), and thus obtained 66 values of log K_{11} and of log P_{11}, as shown in Table 5. The descriptors are the same as those in Table 3, or have been published.[8–15] Exceptions are those for the 1,3,5-triazine pesticides, studied by Jönsson *et al.*,[34] for which descriptors have only recently been obtained.[35] Although many compounds in Table 5 are the same as, or are close analogs of, those in Table 3,[7,16,36–38] there are a number of sets of compound that are considerably different.[34,39] One test of the generality of an equation is that it holds over a wide range of compound type, and so the data collected in Table 5 will be an interesting test of eqn. (3) and eqn. (4).

With only 66 compounds, there is hardly enough data to separate into reasonably comprehensive training sets and test sets, and so we just analyzed the total set. Application of eqn. (3) yielded,

$$\text{Log } P_{11} = 0.058 + 0.603\,E - 1.661\,S - 3.421\,A - 5.120\,B + 4.619\,V \qquad (11)$$

$$N = 66,\ \text{SD} = 0.214,\ R^2 = 0.9874,\ R = 0.9937,\ F = 942$$

The regression statistics are not quite as good as those for the water/dodecane partitions, but are still very reasonable. The SD values on the coefficients average about 0.16 except for the *v*-coefficient where it is 0.08, somewhat lower. Within these SD values, the coefficients in eqn. (11) are the same as those for the corresponding water/dodecane equation, eqn. (6), except for the *v*-coefficients that differ by 0.16 units.

In the case of the log K_{11} data, we found also that an equation with only the *L* descriptor was just as good as that with all five descriptors:

$$\text{Log } K_{11} = 0.113 + 0.971\,L \qquad (12)$$

$$N = 66,\ \text{SD} = 0.223,\ R^2 = 0.9951,\ R = 0.9975,\ F = 12960$$

The truncated equation is quite expected, because there will be no dipole/dipole or hydrogen bond interactions between solutes and alkane solvents; hence the *s*-, *a*- and *b*-coefficients will be zero. Just as for the corresponding log *P* values, eqn. (12) is not as good as the equation for log K_{12}, see eqn. (10). However, both eqn. (11) and eqn. (12) are good enough to predict further values of log P_{11} and log K_{11} to about 0.3 log unit.

In Fig. 1 are plotted the coefficients in the various water/alkane log *P* equations against C_n, the number of carbon atoms in the alkane. Very smooth lines are obtained, thus showing that the coefficients vary very little with the length of the *n*-alkane, and also that they do not depend on the actual nature of the compounds in the data set. Fig. 2 shows similar plots for the *e*- and *l*-coefficients in the gas/alkane log *K* equations. Again, there are reasonably smooth plots against C_n. Both Figs. 1 and 2 indicate that the coefficients in the log *P* and log *K* equations are not just fitting parameters, but that they encode specific chemical information about solute/solvent interactions.

Table 5 Values of log P_{11} and long K_{11} for partition of 66 compounds into undecane, 298 K

Solute	Log P_{11}	Log K_W	Log K_{11}	Ref.
Helium	0.47	−2.02	−1.55	7
Neon	0.57	−1.96	−1.39	7
Argon	0.94	−1.47	−0.53	7
Krypton	1.15	−1.21	−0.06	7
Xenon	1.52	−0.97	0.55	16
Hydrogen	0.62	−1.72	−1.10	17
Oxygen gas	0.93	−1.51	−0.58	7
Nitrogen gas	0.98	−1.80	−0.82	7
Carbon monoxide	0.87	−1.62	−0.75	17
Methane	1.26	−1.46	−0.20	7
Hexane	4.72	−1.82	2.90	37
Decane	7.28	−2.32	4.96	[a]
Undecane	7.86	−2.38	5.48	[a]
Dodecane	8.50	−2.50	6.00	[a]
Tetrafluoromethane	1.61	−2.29	−0.68	7
Chloroform	1.74	0.79	2.53	36
Carbon tetrachloride	2.74	−0.06	2.68	36
1,1-Dichloroethane	1.48	0.62	2.10	36
1,2-Dichloroethane	1.27	1.31	2.58	36
1,1,1-Trichloroethane	2.49	0.14	2.63	36
1,1,2,2-Tetrachloroethane	1.91	1.81	3.72	36
1,2-Dichloropropane	1.74	0.93	2.67	36
1,2-trans-Dichloroethene	1.95	0.57	2.52	36
Trichloroethene	2.42	0.32	2.74	36
Tetrachloroethene	3.30	−0.07	3.23	36
Bromoform	1.97	1.56	3.53	36
bis(2-Chloroethyl)ether	0.91	3.32	4.23	39
bis(2-Chloroethoxy)ether	0.96	5.35	6.31	39
Acrolein	−0.49	2.51	2.02	39
Acrylonitrile	−0.96	2.39	1.43	39
Sulfur hexafluoride	2.22	−2.23	−0.01	7
Ferrocene	3.66	1.92	5.58	[b]
Benzene	2.20	0.63	2.83	36
Toluene	2.74	0.65	3.39	36
Ethylbenzene	3.27	0.58	3.85	36
trans-Stilbene	4.68	2.78	7.46	[b]
Acenaphthene	4.15	2.36	6.51	[b]
Anthracene	4.52	3.03	7.55	[b]
Phenanthrene	4.63	2.80	7.43	[b]
Pyrene	5.17	3.50	8.67	[b]
Chlorobenzene	2.80	0.82	3.62	36
1,2-Dichlorobenzene	3.35	1.00	4.35	36
Nitrobenzene	1.44	3.02	4.46	39
2-Chlorophenol	0.91	3.34	4.25	36
1-Naphthol	0.56	5.63	6.19	[b]
2-Naphthol	0.36	5.95	6.31	[b]
Furfuryl acetate	0.89	3.45	4.34	38
2-Isobutylthiazole	1.93	2.50	4.43	38
Atrazine	0.56	7.10	7.66	34
Simazine	0.06	7.25	7.31	34
Terbuthylazine	1.60	6.60	8.20	34
Desmetryn	0.60	7.70	8.30	34
Ametryn	1.34	7.10	8.44	34
Prometryn	1.99	6.80	8.79	34
Terbutryn	2.33	6.60	8.93	34
Dimethametryn	2.34	6.40	8.74	34
Dipopetryn	2.60	6.60	9.20	34
Azipotryn	1.84	6.00	7.84	34
Metoprotryn	0.96	8.70	9.66	34
Simetone	−0.56	7.80	7.24	34
Atratone	0.09	7.45	7.54	34
Terbumeton	1.22	6.60	7.82	34
Secbumeton	0.69	7.25	7.94	34
D-Carvone	1.95	2.96	4.91	38
(*Z*) 3-Hexenyl acetate	2.52	2.00	4.52	38
Ethyl 2-methylbutyrate	2.64	1.58	4.22	38

[a] Calculated using an activity coefficient of unity. [b] Values determined by the solubility method.

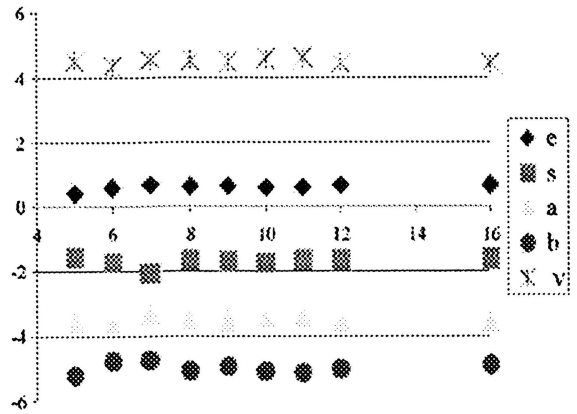

Fig. 1 Plot of the coefficients in eqn. (3) for water/alkane partitions against C_n.

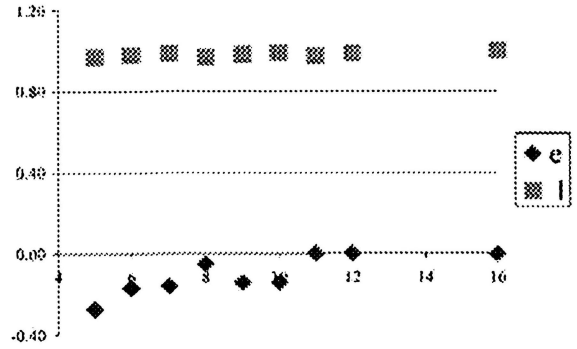

Fig. 2 Plot of the *e*- and *l*-coefficients in eqn. (4) for water/alkane partitions against C_n.

Acknowledgements

We thank Dr Chris Luscombe for help with the Kennard–Stone method, and we are especially grateful to Prof. J. A. Jönsson who kindly provided the original partition data for the triazines.

References

1 A. F. Morales, J. Albet, G. Kyuchoukov, G. Malmary and J. Molinier, *J. Chem. Eng. Data*, 2003, **48**, 874–886.
2 M. A. Hejazi, C. de Lamarliere, J. M. S. Roche, M. Vermue, J. Tramper and R. H. Wijffels, *Biotech. Bioeng.*, 2002, **79**, 29–36.
3 The Med Chem data base 2002, BioByte Corp. and Pomona College, Daylight Chemical Information Systems, 27401 Los Altos, #360 Mission Viejo, CA 92691, USA.
4 D. Bas, D. Dorison-Duval, S. Moreau, P. Bruneau and C. Chipot, *J. Med. Chem.*, 2002, **45**, 151–159.
5 M. Manabe, M. Koda and W. Shirahama, *Bull. Chem. Soc. Jpn.*, 1975, **48**, 3553–3556.
6 C. B. Castells, D. I. Eikens and P. W. Carr, *J. Chem. Eng. Data*, 2000, **45**, 369–375.
7 P. J. Hesse, R. Battino, P. Scharlin and E. Wilhelm, *J. Chem. Eng. Data*, 1996, **41**, 195–201.
8 M. H. Abraham, *Chem. Soc. Rev.*, 1993, **22**, 73–83.
9 M. H. Abraham, A. Ibrahim and A. M. Zissimos, *J. Chromatogr. A*, 2004, **1037**, 29–47.
10 W. E. Acree Jr. and M. H. Abraham, *Can. J. Chem.*, 2001, **79**, 1466–1476.
11 M. H. Abraham, C. E. Green, W. E. Acree Jr., C. E. Hernandez and L. E. Roy, *J. Chem. Soc., Perkin Trans. 2*, 1998, 2677–2681.
12 M. H. Abraham, N. Benjelloun-Dakhama, J. M. R. Gola, W. E. Acree Jr, W. S. Cain and J. E. Cometto-Muniz, *New J. Chem.*, 2000, **24**, 825–829.
13 W. E. Acree Jr and M. H. Abraham, *Can. J. Chem.*, 2001, **79**, 1466–1476.
14 W. E. Acree Jr and M. H. Abraham, *Fluid Phase Eq.*, 2002, **201**, 245–258.
15 W. E. Acree Jr. and M. H. Abraham, *J. Solution Chem.*, 2002, **31**, 293–303.
16 G. L. Pollack and J. F. Himm, *J. Chem. Phys.*, 1982, **77**, 3221–3229.
17 (*a*) *The Solubility Data Project. Hydrogen and deuterium*, ed. C. L. Young, Pergamon Press, Oxford, **Vol. 5/6**, 1981; (*b*) *Carbon monoxide*, ed. R. W. Cargill, Pergamon Press, Oxford, **Vol. 43**, 1990; (*c*) *Propane, butane and 2-methylpropane*, ed. W. Hayduk, Pergamon Press, Oxford, **Vol. 24**, 1986.
18 W. Hayduk, E. B. Walter and P. Simpson, *J. Chem. Eng. Data*, 1972, **17**, 59–61.
19 W. Hayduk and S. C. Cheng, *Can. J. Chem. Eng.*, 1970, **48**, 93–99.
20 A. N. Genkin and B. I. Boguslavskaya, *Neftekhimiya*, 1966, **6**, 626–629.
21 T. M. Letcher, R. C. Baxter, A. Bean and J. D. Sewrey, *J. Chem. Thermodynam.*, 1988, **20**, 581–587.
22 M.-C. Haulait-Pirson, G. Huys and E. Vanstraelen, *Ind. Eng. Chem. Res.*, 1987, **26**, 447–452.
23 T. M. Letcher and W. C. Moollan, *J. Chem. Thermodyn.*, 1995, **27**, 1025–1032.
24 K. Fukuchi, K. Miyoshi, T. Watanabe, S. Yonezawa and Y. Arai, *Fluid Phase Eq.*, 2001, **182**, 257–263.
25 J. R. Loehe and H. C. Van Ness, *Int. DATA Ser., Sel. Data Mix.*, 1979, **78**.
26 H. Watarai, M. Murakami and N. Suzuki, *Bull. Chem. Soc. Jpn.*, 1979, **52**, 937–938.
27 M. H. Cocero, J. A. Gonzalez, I. Garcia, J. C. Cobos and F. Mato, *Int. DATA Ser., Sel. Data Mix*, 1991, **138**.
28 A. E. Beezer and W. H. Hunter, *J. Med. Chem.*, 1983, **26**, 757–759.
29 N. H. Sagert and D. W. P. Lau, *Can. J. Chem.*, 1985, **63**, 789–792.
30 N. H. Sagert and W. Lee, *Can. J. Chem.*, 1980, **58**, 1463–1467.
31 J. Toullec, M. El-Alaoui and P. Kleffert, *J. Org. Chem.*, 1983, **48**, 4808–4816.
32 S. Azizian and A. Haydarpour, *J. Chem. Eng. Data*, 2003, **48**, 1476–1478.
33 R. W. Kennard and L. A. Stone, *Technometrics*, 1969, **11**, 137–148.
34 L. Chimuka, L. Mathiasson and J. A. Jönsson, *Anal. Chim. Acta*, 2000, **416**, 77–86.
35 M. H. Abraham, K. Enomoto, E. Clark, M. Rosés, C. Ràfols and E. Fuguet, unpublished work.
36 T. A. Barbari and C. J. King, *Environ. Sci. Technol.*, 1982, **16**, 624–627.
37 K. N. Marsh and J. B. Ott, *Int. DATA Ser., Sel. Data Mix*, 1983, **222**.
38 S. Rabe, U. Krings, H. Zorn and R. G. Berger, *Lipids*, 2003, **38**, 1075–1084.
39 D. K. Joshi, J. J. Senetar and C. J. King, *Ind. Eng. Chem. Proc. Des. Dev.*, 1984, **23**, 748–754.