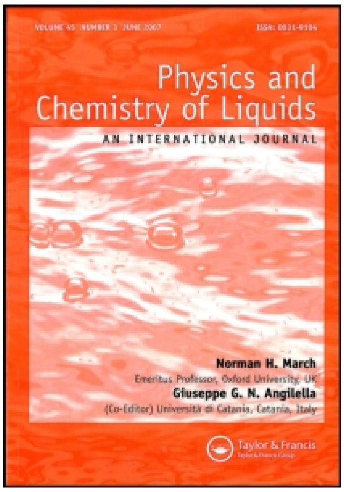

Physics and Chemistry of Liquids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713646857

Computation methodology for determining Abraham solute descriptors from limited experimental data by combining Abraham model and Goss-modified Abraham model correlations

L. M. Sprunger[a]; A. Proctor[a]; W. E. Acree Jr[a]; M. H. Abraham[b]
[a] Department of Chemistry, University of North Texas, Denton, TX 76203-5070, USA [b] Department of Chemistry, University College London, London, WC1H 0AJ, UK

To cite this Article Sprunger, L. M. , Proctor, A. , Acree Jr, W. E. and Abraham, M. H.(2008) 'Computation methodology for determining Abraham solute descriptors from limited experimental data by combining Abraham model and Goss-modified Abraham model correlations', Physics and Chemistry of Liquids, 46: 5, 574 — 585

To link to this Article: DOI: 10.1080/00319100801993407

URL: http://dx.doi.org/10.1080/00319100801993407

Physics and Chemistry of Liquids
Vol. 46, No. 5, October 2008, 574–585

Computation methodology for determining Abraham solute descriptors from limited experimental data by combining Abraham model and Goss-modified Abraham model correlations

L.M. Sprunger[a], A. Proctor[a], W.E. Acree, Jr[a]* and M.H. Abraham[b]

[a]*Department of Chemistry, University of North Texas, Denton, TX 76203-5070, USA;* [b]*Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK*

(*Received 10 January 2008; final version received 16 February 2008*)

A new computation methodology is proposed for calculating the Abraham solute descriptors. The proposed method combines correlations based on both the Abraham model and the Goss-modified version of the Abraham model. The computation methodology is illustrated using published solubility and partition data for benzil in a wide range of organic solvent systems. Calculated solute descriptors back-calculate the observed experimental values to 0.119 log units. The proposed method of using both the Abraham model and the Goss-modified Abraham model correlations significantly increases the number of available equations that can be used in the solute descriptor determination.

Keywords: linear free energy relationships; benzil solubilities; solute descriptors; partition coefficients

1. Introduction

A better understanding of solute–solvent molecular interactions in fluid solution would be of tremendous benefit to individuals working in the pharmaceutical and chemical industries. The design of an effective drug molecule involves delivery of the drug to the desired target site, optimising the release rate from a pharmaceutical formulation, controlling the rate of transport through physiological barriers and favourably altering the drug's solubility through either complexation, cosolvency or structural modifications. Molecular interactions play important roles in each of the aforementioned processes. Molecular interactions are important also in both purification processes and quantitative analyses involving gas–liquid chromatographic and high-performance liquid chromatographic (HPLC) methods. The solute's affinity for the stationary phase *versus* mobile phase is governed by the solute's interactions with the two solubilising media. Naturally, one of the major goals of research in the areas of pharmaceutical chemistry, analytical chemistry and separation science has been the development of solution models that enable *a priori* prediction of solution and partitioning behaviour.

---

*Corresponding author. Email: acree@unt.edu

ISSN 0031–9104 print/ISSN 1029–0451 online

DOI: 10.1080/00319100801993407
http://www.informaworld.com

The general solvation parameter model of Abraham [1–5] is one of the most useful approaches for the analysis and prediction of free energies of partition in chemical and biological systems. The method relies on two linear free energy relationships (LFERS), one for processes within condensed phases

$$\mathrm{SP} = c + e \cdot E + s \cdot S + a \cdot A + b \cdot B + v \cdot V \tag{1}$$

and one for processes involving gas-to-condensed phase transfer

$$\mathrm{SP} = c + e \cdot E + s \cdot S + a \cdot A + b \cdot B + l \cdot L. \tag{2}$$

The dependent variable, SP, is some property of a series of solutes in a fixed phase. The independent variables, or descriptors, are solute properties as follows: E and S refer to the excess molar refraction and dipolarity/polarisability descriptors of the solute, respectively; A and B are measures of the solute hydrogen-bond acidity and basicity; V is the McGowan volume of the solute; and L is the logarithm of the solute gas phase dimensionless Ostwald partition coefficient into hexadecane at 298 K. The first four descriptors can be regarded as measures of the tendency of the given solute to undergo various solute–solvent interactions. The latter two descriptors, V and L, are both measures of solute size, and so will be measures of the solvent cavity term that will accommodate the dissolved solute. General dispersion interactions are also related to solute size, hence, both V and L will also describe the general solute–solvent interactions. The regression coefficients and constants (c, e, s, a, b, v, and l) are obtained by regression analysis of experimental data for a specific process (i.e. a given partitioning process, a given stationary phase and mobile phase combination, etc.). In the case of partition coefficients, where two solvent phases are involved, the c, e, s, a, b, v and l coefficients represent differences in solvent phase properties. For any fully characterised system/process (those with calculated values for the equation coefficients) further values of SP can be estimated for solutes with known values for the solute descriptors. This is the major advantage in using Equations (1) and (2) to correlate solute properties having environmental, pharmaceutical and chemical importance.

The Goss modification of the Abraham equations is shown as Equation (3). The term in $e.E$ is now removed, and the two equations are combined into just one equation. We have shown that regressions based on Equation (3) are statistically almost the same as those based on Equation (1) or Equation (2).

$$\mathrm{SP} = c + s \cdot S + a \cdot A + b \cdot B + v \cdot V + l \cdot L. \tag{3}$$

The purpose of the present communication is to describe a new computation methodology for obtaining solute descriptors from measured solubility and partition coefficient using both the Abraham model and the Goss-modified Abraham model [6–10]. In previous articles [11–16] we have illustrated several computation methods based on the Abraham model. One of the limitations of the previous methodology is that one must have sufficient experimental data for the given solute in a series of solvents/processes and the corresponding Abraham model equation coefficients for the respective solvents/processes must be known. In many instances this can be a problem in that one may not be able to find sufficient experimental data to permit a meaningful computation of solute descriptors. For example, suppose that one was able to find only measured water-to-octanol (wet) partition coefficient, P_{OTOH}, and gas-to-water partition coefficient, K_{w}, gas-to-humic acid partition coefficient, K_{HA}, and gas-to-folvic acid partition coefficient, K_{FA}, data for the solute under consideration. One could calculate the gas-to-octanol (wet) partition coefficient, K_{OTOH}, from the two measured values;

$$\log K_{\mathrm{OTOH}} = \log P_{\mathrm{OTOH}} + \log K_{\mathrm{w}}. \tag{4}$$

However, this would give only four Abraham model equations to use in calculating the six Abraham solute descriptors. Gas-to-water correlations are available for both Equations (1) and (2); however, the correlations for $\log K_{\mathrm{HA}}$ for the gas-to-humic acid partition [7] and $\log K_{\mathrm{FA}}$ for the gas-to-folvic acid partition [10] were reported in terms of Equation (3). If the $\log K_{\mathrm{HA}}$ and $\log K_{\mathrm{FA}}$ values are available one could reanalyse the measured data in terms of the Equation (2). The reanalysis is very simple if one has access to the measured $\log K_{\mathrm{HA}}$ and $\log K_{\mathrm{FA}}$ data. In reporting derived LFER/QSAR (quantitative structure activity relationship) correlations authors may not be able to publish the database used because of proprietary or other concerns. For this example, let us assume that the $\log K_{\mathrm{HA}}$ data are not available, and one is thus not able to reanalyse the experimental data in terms of Equation (2). Reliable estimation methods exist for *E* and *V*. The excess molar refraction, *E*, can be calculated for liquid solutes from the characteristic volume and the measured refractive index at 20°C for the sodium D-line as discussed elsewhere [17,18]. In the case of solid solutes, *E* can be obtained either from the hypothetical refractive index of the liquid solute or through the addition of fragments. The McGowan volume, *V*, can be trivially calculated from the molecular formula and number of rings in the molecule by summation of atom and bond constants [19]. For solutes that are reasonably volatile, *L*, can be determined experimentally by gas chromatography using a hexadecane stationary phase, or, more generally by using a non-polar stationary phase such as squalane. This still leaves the three descriptors *S*, *A* and *B* (and perhaps *L* if the solute is volatile or *E* if the solute is a gas) to be determined with only the four Abraham correlations for $\log P_{\mathrm{OTOH}}$, $\log K_{\mathrm{OTOH}}$ and $\log K_{\mathrm{w}}$. Two additional experimental values and process equations (e.g. $\log K_{\mathrm{HA}}$ and $\log K_{\mathrm{FA}}$ values and the Goss-model correlation for both $\log K_{\mathrm{HA}}$ and $\log K_{\mathrm{FA}}$) would be most useful. Our proposed method of using both the Abraham model and the Goss-modified Abraham model correlations increases the number of available equations, and allows one to use the correlations as they were published.

The proposed computation methodology is illustrated using published solubility data for benzil [20–23] in organic solvents of varying polarity and hydrogen-bonding characteristics. There is more than sufficient experimental data to determine solute descriptors using only the Abraham model. In fact this has already been done [20]. The present study differs from the earlier work in that we compare the calculated solute descriptors based solely on the Abraham model correlations, to those calculated based solely on correlations for the Goss-modified version of the Abraham model, and to those calculated by combining the two sets of correlation equations. As part of the present study we report Goss-modified Abraham model correlations for the solubility of gases in undecane, hexadecane, dimethyl sulfoxide and olive oil (at 310 K) and for the partition of solutes from water to undecane, hexadecane, dimethyl sulfoxide and olive oil (also at 310 K).

2. Computation methodology

Equation (1) actually predicts partition coefficients, and for select solvents both 'dry' and 'wet' equation coefficients have been reported. For solvents that are partially miscible with water, such as 1-butanol and ethyl acetate, partition coefficients calculated as the ratio of the molar solute solubilities in the organic solvent and water are not the same as those obtained from direct partition between water (saturated with the organic solvent) and organic solvent (saturated with water). Care must be taken not to confuse the two sets of partitions. In the case of solvents that are fully miscible with water, such as methanol, no confusion is possible. Only one set of equation coefficients has been reported, and the calculated $\log P$ value must refer to the hypothetical partition between the two pure solvents. For solvents that are 'almost' completely immiscible with water, such as alkanes, cyclohexane, dichloromethane, trichloromethane, tetrachloromethane and most aromatic solvents, there should be no confusion because indirect partition (Equation (5)) will be nearly identical to direct partition.

The determination of solute descriptors is relatively straightforward. We start with the set of equations that we have constructed for the partition of solutes between water and a given solvent. Table 1 gives the coefficients in Equation (1) for the water–solvent partitions we shall consider. The actual numerical values may differ slightly from values reported in earlier publications. Coefficients are periodically revised when additional experimental data becomes available. Note that many of these are 'hypothetical partitions' between pure water and the pure dry solvent; these are shown as 'dry' in Table 1. Although 'hypothetical', these partitions are very useful; as we show later, they can be used to predict solubilities (and activity coefficients) in the pure dry solvent. The partition coefficient of a solid between water and a solvent phase, P, is related to

$$\mathrm{SP} = P = \frac{C_\mathrm{S}}{C_\mathrm{W}} \quad \text{or} \quad \log \mathrm{SP} = \log P = \log C_\mathrm{S} - \log C_\mathrm{W}, \tag{5}$$

the molar solubility of the solid in water, C_W, and in the solvent, C_S, provided that three specific conditions are met. First, the same solid phase must be in equilibrium with the saturation solutions in the organic solvent and in water (i.e. there should be no solvate or hydrate formation). Second, the secondary medium activity coefficient of the solid in the saturated solutions must be unity (or near unity). This condition generally restricts the method to those solutes that are sparingly soluble in water and non-aqueous solvents. Finally, for solutes that are ionised in aqueous solution, C_W refers to the solubility of the neutral form. The molar solubility of benzil in water, $\log C_\mathrm{W} = -4.05$ [25], is used to calculate the experimental solubility ratios, as $\log(C_\mathrm{S}/C_\mathrm{W})$, and to convert the predicted solubility ratios back to predicted solublities.

For partition of solutes between the gas phase and the solvents, Equation (2) is used. (Equation coefficients are given in Table 2 for several organic solvents). Predicted $\log K$ values can also be converted to saturation molar solubilities, provided that the solid saturated vapour pressure at 298.15 K, P°_vap, is available. P°_vap can be transformed into the gas phase concentration, C_G, and the gas–water and gas–solvent partitions, K_W and K_S, can be obtained through

$$\mathrm{SP} = K_\mathrm{W} = \frac{C_\mathrm{W}}{C_\mathrm{G}} \quad \text{or} \quad \log \mathrm{SP} = \log K_\mathrm{W} = \log C_\mathrm{W} - \log C_\mathrm{G} \tag{6}$$

$$\mathrm{SP} = K_\mathrm{S} = \frac{C_\mathrm{S}}{C_\mathrm{G}} \quad \text{or} \quad \log \mathrm{SP} = \log K_\mathrm{S} = \log C_\mathrm{S} - \log C_\mathrm{G} \tag{7}$$

Table 1. Coefficients in Equation (1) of the Abraham model for various processes.[a]

Solvent	*c*	*e*	*s*	*a*	*b*	*v*
1-Octanol (wet)	0.088	0.562	−1.054	0.034	−3.460	3.814
Tetrachloromethane	0.260	0.573	−1.254	−3.558	−4.588	4.589
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
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
Cyclohexane	0.159	0.784	−1.678	−3.740	−4.929	4.577
Methylcyclohexane	0.246	0.782	−1.982	−3.517	−4.293	4.528
Isooctane	0.288	0.382	−1.668	−3.639	−5.000	4.561
Benzene	0.142	0.464	−0.588	−3.099	−4.625	4.491
Toluene	0.143	0.527	−0.720	−3.010	−4.824	4.545
Chlorobenzene	0.040	0.246	−0.462	−3.038	−4.769	4.640
Diethyl ether (dry)	0.308	0.377	−0.813	−0.468	−5.012	4.379
Methyl *tert*-butyl ether (dry)	0.376	0.264	−0.788	−1.078	−5.030	4.410
Dibutyl ether (dry)	0.203	0.369	−0.954	−1.488	−5.426	4.508
1,4-Dioxane (dry)	0.098	0.350	−0.083	−0.556	−4.826	4.172
Tetrahydrofuran (dry)	0.207	0.372	−0.392	−0.236	−4.934	4.447
Methanol (dry)	0.329	0.299	−0.671	0.080	−3.389	3.512
Ethanol (dry)	0.208	0.409	−0.959	0.186	−3.645	3.928
1-Propanol (dry)	0.148	0.436	−1.098	0.389	−3.893	4.036
1-Butanol (dry)	0.152	0.438	−1.177	0.096	−3.919	4.122
1-Pentanol (dry)	0.080	0.521	−1.294	0.208	−3.908	4.208
1-Hexanol (dry)	0.044	0.470	−1.153	0.083	−4.057	4.249
1-Heptanol (dry)	−0.026	0.491	−1.258	0.035	−4.155	4.415
1-Octanol (dry)	−0.034	0.490	−1.048	−0.028	−4.229	4.219
1-Decanol (dry)	−0.062	0.754	−1.461	0.063	−4.053	4.293
2-Propanol (dry)	0.063	0.320	−1.024	0.445	−3.824	4.067
2-Butanol (dry)	0.106	0.272	−0.988	0.196	−3.805	4.110
2-Methyl-1-propanol (dry)	0.177	0.335	−1.099	0.069	−3.570	3.990
2-Methyl-2-propanol (dry)	0.197	0.136	−0.916	0.318	−4.031	4.112
Acetonitrile (dry)	0.413	0.077	0.326	−1.566	−4.391	3.364
Dimethyl sulfoxide	−0.221	0.226	0.878	1.312	−4.604	3.403
Gas phase	−0.994	0.577	2.549	3.813	4.841	−0.869

[a] The solvents denoted as ‘dry’ are those for which partitions refer to transfer to the pure dry solvent. The other partitions are from water (more correctly water saturated with the organic solvent) to the organic solvent saturated with water.

Equations (6) and (7), respectively. As before, the computational method will be valid if conditions discussed above are met. Aihara [25] reported the vapour pressure of benzil as a function of temperature. From his correlation equation

$$\log p^{\circ}_{\text{vap}} = 12.708 - \frac{5148.8}{\text{T}}. \tag{8}$$

A value of $p^{\circ}_{\text{vap}} = 2.92 \times 10^{-5}$ mmHg is calculated, which corresponds to a vapour phase concentration of $\log C_G = -8.80$. A very small adjustment is made to the $\log C_G$ value, $\log C_G = -8.92$, to give better agreement with the published gas–liquid chromatographic

Table 2. Coefficients in Equation (2) of the Abraham model for various processes.[a]

Solvent	c	e	s	a	b	l
1-Octanol (wet)	−0.198	0.002	0.709	3.519	1.429	0.858
Tetrachloromethane	0.282	−0.303	0.460	0.000	0.000	1.047
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
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
Cyclohexane	0.163	−0.110	0.000	0.000	0.000	1.013
Methylcyclohexane	0.318	−0.215	0.000	0.000	0.000	1.012
Isooctane	0.275	−0.244	0.000	0.000	0.000	0.972
Benzene	0.107	−0.313	1.053	0.457	0.169	1.020
Toluene	0.121	−0.222	0.938	0.467	0.099	1.012
Chlorobenzene	0.053	−0.553	1.254	0.364	0.000	1.041
Diethyl ether (dry)	0.288	−0.347	0.775	2.985	0.000	0.973
Methyl *tert*-butyl ether (dry)	0.278	−0.489	0.801	2.495	0.000	0.993
Dioxane (dry)	−0.034	−0.354	1.674	3.021	0.000	0.919
Tetrahydrofuran (dry)	0.189	−0.347	1.238	3.289	0.000	0.982
Methanol (dry)	−0.004	−0.215	1.173	3.701	1.432	0.769
Ethanol (dry)	0.012	−0.206	0.789	3.635	1.311	0.853
1-Propanol (dry)	−0.028	−0.185	0.648	4.022	1.043	0.869
1-Butanol (dry)	−0.039	−0.276	0.539	3.781	0.995	0.934
1-Pentanol (dry)	−0.042	−0.277	0.526	3.779	0.983	0.932
1-Hexanol (dry)	−0.035	−0.298	0.626	3.726	0.729	0.936
1-Heptanol (dry)	−0.062	−0.168	0.429	3.541	1.181	0.927
1-Octanol (dry)	−0.120	−0.203	0.560	3.576	0.702	0.939
1-Decanol (dry)	−0.136	−0.038	0.325	3.674	0.767	0.947
2-Propanol (dry)	−0.060	−0.335	0.702	4.017	1.040	0.893
2-Butanol (dry)	−0.013	−0.456	0.780	3.753	1.064	0.906
2-Methyl-1-propanol (dry)	0.012	−0.407	0.670	3.645	1.283	0.895
2-Methyl-2-propanol (dry)	0.071	−0.538	0.818	3.951	0.823	0.905
Acetonitrile (dry)	−0.007	−0.595	2.461	2.085	0.418	0.738
Dimethyl sulfoxide (dry)	−0.551	−0.319	2.726	4.984	0.316	0.717
SPB-5 (120°C)	−2.620	0.000	0.293	0.212	0.000	0.504
SPB-5 (140°C)	−2.620	0.000	0.278	0.177	0.000	0.453
Water	−1.271	0.822	2.743	3.904	4.814	−0.213

[a] The solvents denoted as 'dry' are those for which partitions refer to transfer to the pure dry solvent.

retention data for benzil on an SPB-5 stationary phase at both 120°C ($\log K = 1.672$) and 140°C ($\log K = 1.272$) (Colin F Poole, unpublished data [26]). The adjustment is discussed in greater detail elsewhere [20].

The number of correlations can be significantly increased by including the Goss–modified Abraham model correlations into the data analysis. The modified correlations [6–10]

$$\log P = c + s \cdot S + a \cdot A + b \cdot B + v \cdot V + l \cdot L \tag{9}$$

$$\log K = c + s \cdot S + a \cdot A + b \cdot B + v \cdot V + l \cdot L \tag{10}$$

Table 3. Coefficients in water-to-organic solvent equation of the Goss-modified Abraham model (Equation (9)) for various processes.[a]

Solvent	*c*	*s*	*a*	*b*	*v*	*l*
Methanol (dry)	0.418	−0.586	0.146	−3.574	3.139	0.122
Ethanol Dry)	0.360	−0.807	0.066	−3.785	3.385	0.148
1-Propanol (dry)	0.373	−0.982	0.141	−4.032	3.271	0.201
2-Propanol (dry)	0.217	−0.959	0.347	−3.986	3.540	0.146
1-Butanol (dry)	0.420	−1.200	−0.030	−4.047	3.182	0.257
2-Butanol (dry)	0.253	−0.950	0.042	−3.891	3.665	0.124
2-Methyl-1-propanol (dry)	0.444	−1.200	−0.123	−3.605	3.171	0.215
2-Methyl-2-propanol (dry)	0.319	−0.980	0.284	−4.079	3.734	0.102
1-Pentanol (dry)	0.606	−1.407	−0.045	−4.074	2.687	0.370
Methylcyclohexane	0.788	−2.085	−3.775	−4.577	2.842	0.457
Diethyl ether (dry)	0.526	−0.818	−0.620	−5.147	3.628	0.208
Dibutyl ether (dry)	0.204	−0.680	−1.673	−5.633	4.291	0.083
Methyl *tert*-butyl ether (dry)	0.490	−0.753	−1.193	−5.102	3.937	0.132
Tetrahydrofuran (dry)	0.446	−0.412	−0.347	−5.104	3.655	0.212
1,4-Dioxane (dry)	0.358	−0.167	−0.714	−4.922	3.309	0.228
Dimethyl sulfoxide (dry)[b]	−0.111	0.896	1.236	−4.613	2.992	0.112
Undecane[c]	0.012	−1.007	−3.265	−5.487	4.443	0.079
Hexadecane[d]	0.502	−1.921	−3.734	−4.942	2.833	0.452
Olive oil (310 K)[e]	0.421	−1.259	−1.779	−4.487	2.718	0.379

[a] The solvents denoted as 'dry' are those for which partitions refer to transfer to the pure dry solvent. The other partitions are from water (more correctly water saturated with the organic solvent) to the organic solvent saturated with water.
[b] Statistical information: 123 data points with a standard deviation of 0.224 log units.
[c] Statistical information: 66 data points with a standard deviation of 0.252 log units.
[d] Statistical information: 357 data points with a standard deviation of 0.130 log units.
[e] Statistical information: 167 data points with a standard deviation of 0.173 log units, a squared correlation coefficient of 0.987 and a Fisher *F*-statistic of 2397.3.

also contain the Abraham solute descriptors, and there is no mathematical reason that this latter set of correlations cannot be utilised in the determination of solute descriptors. The Abraham E solute descriptor has been replaced in Equations (8) and (9) by a second cavity size term. Our past experience [27–29] with Equations (8) and (9), albeit fairly limited, suggests that there is very little difference in descriptive ability of correlations based on the Abraham model and the Goss–modified Abraham model.

Personally we prefer the Abraham model because it is easier to estimate the numerical value of E than it is to estimate the solute's Ostwald coefficient in hexadecane. In many of our recent studies [30–36] the solutes have been either non-volatile drug or prodrug molecules having several functional groups, and the L descriptors were not known. Equation (1) does not require the L descriptors, and we have been able to develop correlations for the partitioning of drugs between blood and various body organs (i.e. blood-to-brain, blood-to-muscle, blood-to-liver, etc.) based on Equation (1). Such correlations would not be possible though with Equations (9) and (10). In Tables 3 and 4 we have assembled the coefficients for the water-to-organic solvent and for gas-to-organic solvent correlations based on the Goss-modified Abraham model. The tabulation includes our newly derived correlations for dimethyl sulfoxide, undecane, hexadecane and olive oil (at 310 K).

Table 4. Coefficients in gas-to-organic solvent equation of the Goss-modified Abraham model (Equation (10)) for various processes.[a]

Solvent	c	s	a	b	v	l
Methanol (dry)	−0.050	1.031	3.885	1.634	0.175	0.703
Ethanol Dry)	−0.060	0.691	3.953	1.334	0.163	0.802
1-Propanol (dry)	−0.183	0.755	4.071	1.075	0.447	0.755
2-Propanol (dry)	−0.269	0.716	4.131	1.184	0.683	0.704
1-Butanol (dry)	−0.085	0.570	3.848	1.041	0.311	0.805
2-Butanol (dry)	−0.235	0.662	3.960	1.243	0.687	0.714
2-Methyl-1-propanol (dry)	−0.009	0.387	3.668	1.674	0.071	0.828
2-Methyl-2-propanol (dry)	−0.211	0.714	4.124	1.028	0.882	0.667
1-Pentanol (dry)	−0.093	0.396	3.857	1.098	0.227	0.855
Methylcyclohexane	0.213	0.000	0.000	0.000	0.407	0.902
Hexadecane	0.000	0.000	0.000	0.000	0.000	1.000
Diethyl ether (dry)	0.031	0.927	3.213	0.000	0.870	0.737
Dibutyl ether (dry)	−0.501	1.559	2.325	−0.849	2.195	0.439
Methyl *tert*-butyl ether (dry)	−0.079	1.187	2.622	0.000	1.640	0.536
Tetrahydrofuran (dry)	−0.118	1.396	3.492	0.000	0.979	0.731
1,4-Dioxane (dry)	−0.200	1.673	3.134	0.000	0.628	0.744
Dimethyl sulfoxide (dry)[b]	−0.533	2.545	5.051	0.415	0.072	0.680
Undecane[c]	−0.044	0.000	0.000	0.000	0.599	0.863
Olive oil (310 K)[d]	−0.266	0.891	1.720	−0.026	0.299	0.793

[a] The solvents denoted as 'dry' are those for which partitions refer to transfer to the pure dry solvent.
[b] Statistical information: 123 data points with a standard deviation of 0.218 log units.
[c] Statistical information: 66 data points with a standard deviation of 0.223 log units.
[d] Statistical information: 215 data points with a standard deviation of 0.177 log units, a squared correlation coefficient of 0.984 and a Fisher *F*-statistic of 2583.8.

3. Results and discussion

Solute descriptors currently exist for benzil ($E = 1.445$, $S = 1.590$, $A = 0.000$, $B = 0.620$, $V = 1.6374$ and $L = 7.6112$) based on an earlier regression analysis [20] involving 51 Abraham model $\log P$ (Equation (1)) and $\log K$ (Equation (2)) correlations. These numerical values will continue to be the solute descriptors for benzil. It is not practical to update continually existing values of solute descriptors every time that new experimental data becomes available, or every time that a new solvent/process correlation is derived. Rather, we are using benzil as an illustrational example to show that it is possible to calculate solute descriptors based solely on the Abraham model (as we currently do), based solely on the Goss-modified Abraham model, and by combining correlations from both correlations. The latter calculation offers several advantages, particularly in cases where one has limited experimental data, or when one has the equation coefficients for only the Goss-modified Abraham model. In presenting derived correlations authors often present the final equation without giving the data used in the regression analysis. The data might not be publishable because of propriety considerations. Determination of Abraham model coefficients would not be possible in such instances. The proposed computation methodology will be considered a success if the three computations yield identical numerical values for the set of six solute descriptors.

To determine the solute descriptors for benzil we must first convert the published benzil solubility data into $\log P$ and $\log K$ values. Most of the experimental data was originally reported as mole fraction solublities, X_S^{exp}. Experimental mole fraction solublities were converted into molar solublities, C_S^{exp}, by dividing X_S^{exp}, by the ideal molar volume of the saturated solution (i.e. $C_S^{exp} \approx X_S^{exp}/[X_S^{exp}\ V_{solute} + (1 - X_S^{exp})\ V_{solvent}]$). A value of $V_{solute} = 183\,cm^3\,mol^{-1}$ was used for the molar volume of the hypothetical subcooled liquid benzil. Any errors resulting from our estimation of the benzil's hyphothetical subcooled liquid molar volume, V_{solute}, or the ideal molar volume approximation, should have negligible effect of the calculated C_S^{exp} values.

Combining the $\log P$ and $\log K$ values calculated from the experimental benzil solubility we have a total of 67 equations for the Abraham model and 33 equations for the Goss-modified Abraham model. The gas–liquid chromatographic retention data for benzil on an SPB-5 stationary phase at 120°C ($\log K = 1.672$) and 140°C ($\log K = 1.272$) [26] provides two additional Abraham model equations. There is also one direct experimental water-to-octanol partition coefficient, $\log P_{OTOH} = 3.38$, in the MedChem database [37]. The aqueous molar solubility of benzil of $\log C_W = -4.05$ [24] can also be included in the regression analyses. The published correlation of Abraham and Le [38]

$$\frac{(\log C_W)}{5} = 0.104 - 0.2011\ E + 0.154\ S + 0.434\ A + 0.848\ B - 0.672\ A \cdot B - 0.797\ V \tag{11}$$

and its updated version (unpublished)

$$\frac{(\log C_W)}{5} = 0.079 - 0.191\ E + 0.064\ S + 0.231\ A + 0.651\ B - 0.157\ A \cdot B - 0.666\ V \tag{12}$$

was used for the aqueous predictions. The cross $A \cdot B$ term was added to the model [38] to account for hydrogen-bond interactions between the acidic and the basic sites in the pure liquid or solid solute. Such interactions are not normally included in solubility ratio and partition coefficient correlations. In practical partitioning studies the solute is generally at very low concentration and is surrounded by solvent molecules. In the case of solubility ratios the same equilibrium solid phase must be present for both C_S and C_W measurements. This allows contributions from breaking of crystal forces to cancel in the calculation of the solubility ratio.

There is sufficient experimental data to calculate the solute descriptors for benzil. The characteristic McGowan volume of benzil of $V = 1.6374$ is calculated from the individual atomic sizes and number of bonds in the molecule and E is estimated as $E = 1.445$. The A descriptor is set equal to zero as benzil has no acidic hydrogens capable of H-bond formation. This leaves the S, B and L solute descriptors to be calculated. Regression analysis using only the Abraham model correlations gave $S = 1.534$, $B = 0.636$ and $L = 7.638$ with a standard deviation between experimental and back-calculated values of $SD = 0.125$ log units, which is in excellent agreement with our existing numerical values for the solute descriptors of benzil. The slight difference in calculated values results from the fact that the current analysis used 73 Abraham model correlations, which is considerably larger than the 51 equations used previously [20]. Regression analysis based solely on the 33 Goss-modified Abraham model equations yielded values of $S = 1.590$, $B = 0.605$ and $L = 7.573$ with $SD = 0.101$ log units, which are in good agreement with values based on Equations (1) and (2). For all practical purposes both calculations were identical. No significance is placed on the slightly smaller deviation of $SD = 0.101$ log units for the

Goss-modified Abraham model. The data set used in the latter analysis was much smaller and had less chemical diversity. Finally, all 106 correlations from both models were combined and analysed to give $S = 1.556$, $B = 0.625$ and $L = 7.622$, with SD = 0.119 log units. All computations were performed using Microsoft Excel Solver. Tables 5 and 6 compare the experimental $\log P$ and $\log K$ values to values back-calculated using $E = 1.445$, $S = 1.556$, $A = 0.000$, $B = 0.625$, $V = 1.6374$ and $L = 7.622$. Solute descriptors calculated by combining the Abraham model and Goss-modified Abraham model correlations are essentially identical to values calculated from analysis of the correlations for each separate model.

Table 5. Comparison between observed $\log P$ and $\log K$ data and back-calculated values based on Equations (1) and (2) of the Abraham Model.[a]

Solvent	$\log P^{\mathrm{exp}}$	$\log P^{\mathrm{calc}}$	$\log K^{\mathrm{exp}}$	$\log K^{\mathrm{calc}}$
1-Octanol (wet)	3.380	3.343	8.250	8.341
Tetrachloromethane	3.940	3.783	8.810	8.540
Hexane	2.690	2.652	7.560	7.510
Heptane	2.700	2.567	7.570	7.533
Octane	2.700	2.832	7.570	7.515
Nonane	2.700	2.732	7.570	7.460
Decane	2.690	2.636	7.560	7.487
Dodecane	2.640	2.694	7.510	7.568
Hexadecane	2.660	2.750	7.530	7.622
Cyclohexane	3.040	3.095	7.910	7.725
Methylcyclohexane	2.960	3.023	7.830	7.721
Isooctane	2.600	2.588	7.470	7.331
Benzene	4.280	4.361	9.150	9.173
Toluene	4.150	4.211	8.970	9.035
Chlorobenzene	4.280	4.294	9.120	9.140
Diethyl ether (dry)	3.810	3.626	8.680	8.409
Methyl *tert*-butyl ether (dry)	3.740	3.609	8.610	8.386
Dibutyl ether (dry)	3.340	3.242		
Dioxane (dry)	4.320	4.290	9.190	9.064
Tetrahydrofuran (dry)	4.390	4.333	9.260	9.099
Methanol (dry)	3.320	3.350	8.190	8.267
Ethanol (dry)	3.290	3.461	8.160	8.263
1-Propanol (dry)	3.240	3.245	8.110	7.988
1-Butanol (dry)	3.200	3.254	8.070	8.145
1-Pentanol (dry)	3.190	3.267	8.060	8.094
1-Hexanol (dry)	3.140	3.351	8.010	8.098
1-Heptanol (dry)	3.110	3.358	7.980	8.166
1-Octanol (dry)	3.080	3.309	7.950	8.054
1-Decanol (dry)	3.090	3.251	7.960	8.012
2-Propanol (dry)	3.080	3.201	7.950	8.005
2-Butanol (dry)	3.124	3.313	7.990	8.112
2-Methyl-1-propanol (dry)	3.065	3.253	7.940	8.090
2-Methyl-2-propanol (dry)	3.060	3.182	7.930	7.979
Acetonitrile (dry)	4.070	3.796	8.940	8.849
Dimethyl sulfoxide (dry)	4.140	4.167	9.010	9.036
SPB-5 (120°C)			1.672	1.677
SPB-5 (140°C)			1.272	1.265

[a] Solute descriptors used in the calculations were: $E = 1.445$, $S = 1.556$, $A = 0.000$, $B = 0.625$, $V = 1.6374$ and $L = 7.622$.

Table 6. Comparison between observed log P and log K data and back-calculated values based on Equations (9) and (10) of the Goss-modified abraham model.[a]

Solvent	log P^{exp}	log P^{calc}	log K^{exp}	log K^{calc}
Methanol (dry)	3.320	3.342	8.190	8.220
Ethanol (dry)	3.290	3.409	8.160	8.229
1-Propanol (dry)	3.240	3.213	8.110	8.150
2-Propanol (dry)	3.080	3.143	7.950	8.069
1-Butanol (dry)	3.200	3.193	8.070	8.097
2-Butanol (dry)	3.124	3.289	7.990	8.137
2-Methyl-1-propanol (dry)	3.065	3.155	7.940	8.067
2-Methyl-2-propanol (dry)	3.060	3.136	7.930	8.071
1-Pentanol (dry)	3.190	3.090	8.060	8.098
Methylcyclohexane	2.960	2.820	7.830	7.754
Hexadecane	2.660	2.510	7.530	7.622
Diethyl ether (dry)	3.810	3.562	8.680	8.515
Dibutyl ether (dry)	3.340	3.284		
Methyl *tert*-butyl ether (dry)	3.740	3.582	8.610	8.539
Tetrahydrofuran (dry)	4.390	4.216	9.260	9.229
1,4-Dioxane (dry)	4.320	4.178	9.190	9.102
Dimethyl sulfoxide (dry)	4.140	4.150	9.010	8.988

[a] Solute descriptors used in the calculations were: $E = 1.445$, $S = 1.556$, $A = 0.000$, $B = 0.625$, $V = 1.6374$ and $L = 7.622$.

Results of our computations clearly show that it is possible to determine solute descriptors based solely on the Abraham model correlations, based solely on the Goss-modified Abraham model correlations, and based on our new computational methodology that combines the derived correlations for both models. The advantages of combining the Abraham model correlations and Goss-modified Abraham model correlations is two-fold. First the number of available equations can be increased without the measurement of new experimental data. For example, one would now be able to write four correlations (Equations (1) and (2) and Equations (9) and (10)) from the solute's measured solubility in ethanol, as opposed to just the two equations based on the Abraham model. The additional equations from the Goss-modified Abraham can be quite useful in cases where experimental data is limited. One would still have to have the required number of linearly independent LFER/QSAR correlations to arrive at a mathematical solution. Naturally one would prefer having more experimental data as this would increase the chemical diversity of the solvents/processes considered; however, from a practical standpoint one may not have the instrumentation or chemicals on hand to perform additional experimental measurements. Second, the new computation methodology allows one to use published correlations based on the Goss-modified Abraham model without having to curve-fit the experimental data in accordance to Equation (1) and/or Equation (2) (and *vice versa*). Reanalysis of experimental data is not always possible as authors may report only the final-derived correlation equation without giving the database used in its development.

Acknowledgements

The authors thank the University of North Texas Research Council for partial support of this work.

References

[1] M.H. Abraham, Chem. Soc. Rev. **23**, 73 (1993).
[2] M.H. Abraham, G.S. Whiting, W.J. Shuely, and R.M. Doherty, Can. J. Chem. **76**, 703 (1998).
[3] M.H. Abraham, J. Le, and W.E. Acree, Jr, Collect. Czech. Chem. Commun. **64**, 1748 (1999).
[4] M.H. Abraham, J. Le, W.E. Acree, Jr, and P.W. Carr, J. Phys. Org. Chem. **12**, 675 (1999).
[5] M.H. Abraham, A. Ibrahim, and A.M. Zissimos, J. Chromatogr. A **1037**, 29 (2004).
[6] K.-U. Goss, Fluid Phase Equilibr. **233**, 19 (2005).
[7] C. Niederer, K.-U. Goss, and R.P. Schwarzenbach, Environ. Sci. Technol. **40**, 5374 (2006).
[8] C.M. Roth, K.-U. Goss, and R.P. Schwarzenbach, Environ. Sci. Technol. **39**, 6638 (2005).
[9] I. Kamprad and K.-U. Goss, Anal. Chem. **79**, 4222 (2007).
[10] C. Niederer, R. P. Schwarzenbach, and K.-U. Goss, Environ. Sci. Technol. **41**, 6711 (2007).
[11] M.H. Abraham, C.E. Green, W.E. Acree, Jr, C.E. Hernández, and L.E. Roy, J. Chem. Soc., Perkin Trans. **2**, 2677 (1998).
[12] C.E. Green, M.H. Abraham, W.E. Acree, Jr, K.M. De Fina, and T.L. Sharp, Pest Manag. Sci. **56**, 1043 (2000).
[13] M.H. Abraham and A.J.M. Al-Hussaini, J. Environ. Monit. **7**, 295 (2005).
[14] M.H. Abraham, J.M.R. Gola, J.E. Cometto-Muñiz, and W.S. Cain, Fluid Phase Equilibr. **180**, 41 (2001).
[15] A.M. Zissimos, M.H. Abraham, M.C. Barker, K.J. Box, and K.Y. Tam, J. Chem. Soc. Perkin Trans. **2**, 470 (2002).
[16] A.M. Zissimos, M.H. Abraham, C.M. Du, K. Valko, C. Bevan, D. Reynolds, J. Wood, and K.Y. Tam, J. Chem. Soc. Perkin Trans. **2**, 2001 (2002).
[17] M.H. Abraham, G.S. Whiting, R.M. Doherty, and W.J. Shuely, J. Chem. Soc. Perkin Trans. **2**, 1451 (1990).
[18] J.D. Weckwerth and P.W. Carr, Anal. Chem. **70**, 4793 (1998).
[19] M.H. Abraham and J.C. McGowan, Chromatographia **23**, 245 (1987).
[20] W.E. Acree, Jr and M.H. Abraham, J. Solution Chem. **31**, 293 (2002).
[21] K.A. Fletcher, S. Pandey, M.E.R. McHale, and W.E. Acree, Jr, Phys. Chem. Liq. **33**, 181 (1995).
[22] W.E. Acree, Jr and J.H. Rytting, J. Pharm. Sci. **71**, 201 (1982).
[23] W.E. Acree, Jr and G.L. Bertrand, J. Solution Chem. **12**, 101 (1983).
[24] M. Shikata, and I. Tachi, J. Electrochem. Assoc. (Jpn), 2, 201 (1934) [Chem. Abstracts, 28, 7173g (19340].
[25] A. Aihara, Bull. Chem. Soc. Jpn. **32**, 1242 (1959).
[26] Colin F. Poole, unpublished data.
[27] K.B. Flanagan, W.E. Acree, Jr, and M.H. Abraham, Fluid Phase Equilibr. **237**, 224 (2005).
[28] C. Mintz, T. Ladlie, and K. Burton, M. Clark, W.E. Acree, Jr, M.H. Abraham, QSAR Comb. Sci. **27,** 627 (2008).
[29] L. Sprunger, W.E. Acree, Jr, and M.H. Abraham, Anal. Chem. **79**, 6891 (2007).
[30] M.H. Abraham and W.E. Acree, Jr, Int. J. Pharm. **294**, 121 (2005).
[31] M.H. Abraham and F. Martins, J. Pharm. Sci. **93**, 1508 (2004).
[32] M.H. Abraham, A. Ibrahim, and W.E. Acree, Jr, Chem. Res. Toxicol. **19**, 801 (2006).
[33] M.H. Abraham, A. Ibrahim, Y. Zhao, and W.E. Acree, Jr, J. Pharm. Sci. **95**, 2091 (2006).
[34] M.H. Abraham, A. Ibrahim, and W.E. Acree, Jr, Eur. J. Med. Chem. **42**, 743 (2007).
[35] M.H. Abraham and A. Ibrahim, Int. J. Pharm. **329**, 129 (2007).
[36] M.H. Abraham and A. Ibrahim, Eur. J. Med. Chem. **41**, 1430 (2006).
[37] A.J. Leo, The Pomona Medicinal Chemistry Project (Pomona College, CA 91711, USA).
[38] M.H. Abraham and J. Le, J. Pharm. Sci. **88**, 868 (1999).