Talanta 83 (2011) 1118–1125

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Models for liquid–liquid partition in the system formamide–organic solvent and their use for estimating descriptors for organic compounds

Thushara Karunasekara, Colin F. Poole*

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

ARTICLE INFO

Article history:
Available online 4 July 2010

Keywords:
Formamide
Solvation parameter model
Liquid–liquid partition
Descriptor measurements

ABSTRACT

Partition coefficients for varied compounds were determined for the organic solvent–formamide biphasic partition system where the organic solvent is 1,2-dichloroethane, 1-octanol or isopentyl ether. These partition coefficient databases are analyzed using the solvation parameter model facilitating a quantitative comparison of the formamide-based partition systems with other totally organic and water-based partition systems. Formamide is shown to be a less cohesive and hydrogen-bond acidic solvent than water with similar hydrogen-bond basicity and dipolarity/polarizability. Compared with other organic solvents its higher cohesion, dipolarity/polarizability and hydrogen-bonding capability make it a useful base solvent for forming biphasic partition systems with a range of system properties obtained through variation of the counter solvent. Formamide–organic solvent systems offer a complementary approach to water-based partition systems for sample preparation and the determination of descriptors for compounds virtually insoluble in water or unstable in water.

1. Introduction

A resurgence of interest in liquid–liquid partitioning as a sample preparation method stems from a series of developments in different formats that facilitated its use on a small scale (liquid-phase microextraction) thus minimize many of the disadvantages responsible for its replacement by (largely) solid-phase extraction methods over the previous decade [1–3]. Solvent-based methods are generally more tolerant of matrix burden and afford a wider selectivity range than is possible with commonly available sorbents. Solvent properties are more reproducible than those of sorbents and liquid-phase microextraction methods compare favorably in terms of costs and equipment needs compared with sorbent-based methods. The new liquid-phase microextraction methods are viewed as competitive or viable replacements for solid-phase extraction methods, both of which are expected to continue to figure prominently in laboratory practice during the next decade [4–6]. Useful liquid–liquid partition systems require the formation of biphasic systems of low mutual solubility. This tends to dictate that the majority of systems described so far have water as one phase and a low to moderately polar organic solvent as the other [4,7,8]. For many applications this is not a problem, but for compounds and sample matrices of low water solubility, or for compounds that are water unstable, predominantly aqueous biphasic systems are of limited use. Totally organic biphasic systems are an attractive alternative for compounds of this type but limited in choice by the high mutual solubility among organic solvents. Previously described systems include n-heptane–ethylene glycol [9], n-hexane–acetonitrile [10], n-heptane–methanol [11], n-heptane–N,N-dimethylformamide [11], n-heptane–2,2,2-trifluoroethanol [12], n-heptane–1,1,1,3,3,3-hexafluoroisopropanol [12], toluene–perfluoromethylcyclohexane [13], and n-alkane–dimethyl sulfoxide [11,14] systems. These systems are limited by the low solubility of some analytes in n-alkane solvents and by the complete miscibility of the counter solvents listed above with more polar solvents than n-alkanes. They afford systems with complementary properties to aqueous-based partition systems but with limited flexibility. It is desirable to have available alternative totally organic solvent systems that allow a wider range of solubility and selectivity properties to be exploited. This problem we plan to address in this paper.

A common application of water-based biphasic systems is the determination of molecular descriptors for use in the solvation parameter model and other models employed for estimating biopartitioning and environmental distribution properties [8,15]. An attractive feature of predominantly aqueous biphasic systems for descriptor measurements is the availability of complementary systems with large values for the *s*, *a*, and *b* system constants (defined later) which facilitate the calculation of solute descriptors with low uncertainty. Once a full set of descriptors are defined

---

* Corresponding author at: Department of Chemistry, Wayne State University, Rm 181, Detroit, MI 48202, USA. Tel.: +1 313 577 2881; fax: +1 313 577 1277.
E-mail address: cfp@chem.wayne.edu (C.F. Poole).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.06.064

for a compound they can be used to predict the properties of that compound in a wide range of chromatographic [8,15–20], environmental [21–23], and biological [21,24] processes, in addition to liquid-liquid partitioning systems [8]. For the reasons stated in the previous paragraph aqueous biphasic systems are unsuitable for the accurate determination of descriptor values for compounds virtually insoluble or unstable in water. To overcome this problem totally organic biphasic systems with large system constants or distribution properties strongly determined by a limited number of system constants are attractive. The biphasic systems n-heptane–ethylene glycol was shown to be suitable for estimating hydrogen-bonding descriptors for peptides [9]. The biphasic systems n-hexane–acetonitrile, n-heptane–N,N-dimethylformamide and n-heptane–2,2,2-trifluoroethanol were used together with chromatographic retention factors to determine a complete set of descriptors for organosilicon compounds [25,26] and for a variety of other compounds difficult to study using water-based partition systems [15]. These systems provided a working alternative for those compounds that could not be studied using aqueous biphasic systems but do not afford the desired balance between the relative range of system constant values best suited for estimating descriptor values, especially the hydrogen-bond basicity descriptor, *B*, defined later. The second purpose of this paper is to identify additional totally organic biphasic systems suitable for estimating descriptor values for compounds of low water solubility and stability.

2. Formamide, a possible organic water substitute

The search for an organic solvent with the desired qualities to create flexible biphasic systems for extraction and descriptor measurements led us to reflect on the properties of water that have resulted in its widespread use for these applications. These are its high cohesive energy (which allows it to form so many biphasic systems with different solvents) and its overall capacity for polar interactions (which enable it to selectively extract polar compounds). Our goals could be met by a solvent described as water-like but “water light”. This solvent should be sufficiently cohesive to form a reasonable number of biphasic systems with a range of solvents of different selectivities but not so cohesive that compounds of low polarity reside almost totally in the counter solvent. The solvent should also have a sufficient capacity for dipole-type and hydrogen-bonding interactions to provide a reasonable range of selectivity that we would hope to moderate by choice of different counter solvents to enhance selectivity. These considerations led us to evaluate formamide for use as a base solvent and n-heptane, 1,2-dichloroethane, n-octanol and isopentyl ether as counter solvents forming biphasic systems. Relevant solvation properties for water and formamide are summarized in Table 1 [27–30]. Formamide has a high cohesive energy compared with typical organic solvents, roughly two-thirds the value for water. It has an extensive three-dimensional hydrogen-bonded structure similar to water at room temperature [31]. Relatively rare for an organic solvent it has a dielectric permittivity higher than that of water. Spectroscopic measurements of chemical probes indicate that it is almost as dipolar/polarizable as water, a significant hydrogen-bond acid but not as hydrogen-bond acidic as water, and about as hydrogen-bond basic as water. Surfactants are known to form micelles in formamide, a property generally associated with aqueous solvents [30,32,33]. In analytical chemistry formamide has been widely used as a non-aqueous solvent for titration, electrochemistry, and electrophoresis [28,29,34], as a denaturing agent for DNA [35], as an additive in supercritical fluid chromatography to modify the polarity of carbon dioxide [36], and as a stationary phase in high performance liquid–liquid chromatography [37]. Formamide has low toxicity (Table 1) and a low teratogenic effect [38]. Since it has low vapor pressure it is a low risk for inhalation toxicity. It can be absorbed through the skin in quantities sufficient to produce systemic toxicity although it is not very acutely toxic via this route.

Table 1
Characteristic solvent properties of water and formamide.

Property	Water	Formamide
Cohesive energy (J/cm^3)	2302	1575
Dipole moment (D)	1.83	3.37
Dielectric permittivity	78.4	111
Refractive index (20 °C)	1.3325	1.4468
Molecular weight	18.02	45.04
Boiling point (°C)	100	210.5 180 (decomposition)
Vapor pressure (mmHg at 20 °C)	0.08	18
Kamlet–Taft parameters		
π^*	1.09	0.97
α	1.17	0.71
β	0.47	0.48
Reichardt's E_T^N	1.00	0.775
Gutmann's donor number (kJ/mol)	138	151
Gutmann's acceptor number	54.8	39.8
Acute oral toxicity (LD50)		3.15 g/kg [mouse]
Acute dermal toxicity (LD50)		17 g/kg [rabbit]
Inhalation toxicity (LC50)		>3900 ppm/6 H

3. Quantitative structure–partition relationships

Abraham et al. [39] have studied gas–solvent and hypothetical water–solvent partition systems for a number of amides including formamide. They demonstrated that as a solvent it was moderately cohesive, strongly hydrogen-bond basic, and quite dipolar and hydrogen-bond acidic. This combination of solvation properties set it apart from a database of eighteen common organic solvents. In a preliminary study Karunasekara and Poole [40] determined n-heptane–formamide partition coefficients for 84 varied organic compounds and built a suitable model to explain partitioning behavior in this system using the solvation parameter model (also used by Abraham et al. [39] to explain gas–formamide partition coefficients). This paper can be considered an extension of these works applying the solvation parameter model to a wider range of biphasic systems containing formamide to establish the molecular basis of the partition mechanism in these systems.

The solvation parameter model in a form suitable for modeling partition coefficients for neutral compounds, log K_p, in biphasic formamide-containing system is set out below [8,10–12,15]

$$\log K_p = c + eE + sS + aA + bB + vV \qquad (1)$$

The capital letters are solute descriptors defining the capability of the solute to participate in interactions in the two phases and the lower case letters are the system constants defining the difference in the complementary interactions with the solutes in the two immiscible solvent layers. The *E* descriptor defines the solute's capacity for lone pair electron interactions (cm^3/mol/10), the *S* descriptor for interactions of a dipole-type, the *A* and *B* descriptor for hydrogen-bonding interactions with the solute acting as a hydrogen-bond acid or base, and the *V* descriptor is McGowan's characteristic volume (cm^3/mol/100). The system constants are calculated for the biphasic system from experimental partition coefficients for a varied group of compounds with known descriptor values using multiple linear regression analysis. The system constants and their ratios provide a quantitative estimate of selectivity differences for the biphasic systems suitable for comparing the properties of different systems and defining suitable applications for their use. The system constants are also required for the calculation of descriptors for compounds that lack a full set of these values at present.

4. Experimental

4.1. Materials

Formamide was obtained from Acros Organics (Morris Plains, NJ, USA) and dried over molecular sieves before use. 1,2-Dichloroethane, 1-octanol and isopentyl ether were obtained from Sigma–Aldrich (Milwauke, WI, USA). Common chemicals were of the highest purity available and obtained from several sources. The 30 m × 0.32 mm id HP-5 open-tubular column, 0.25 μm film thickness, was obtained from Agilent Technologies (Folsom, CA, USA).

4.2. Instrumentation

Gas chromatographic measurements were made with an Agilent Technologies (Palo Alto, CA, USA) HP 6890 gas chromatograph fitted with a split/splitless injector and flame ionization detector using ChemStation software (rev.B.04.01) for data acquisition. Nitrogen was used as carrier gas at a constant flow rate of 2.5 mL/min (velocity 47 cm/s). The split ratio was set to 30:1, septum purge 1 mL/min, inlet temperature 275 °C, and detector temperature 300 °C. Separations were performed using a temperature program with an initial temperature of 150 °C for 1 min and then raised to 280 °C at 10 °C/min.

4.3. Determination of partition coefficients

The method used to determine partition coefficients is described in detail elsewhere [10–12,39]. The 2.0-mL screw-capped sample vials with PTFE-lined caps (Supelco, Bellefontaine, PA, USA) were charged by syringe with 0.75 mL of formamide, 0.75 mL of counter solvent, 1–10 μL of liquid sample, and 1 μL internal standard. Solid samples were dissolved in either the counter solvent or formamide (depending on solubility) at a concentration of about 0.5–1.5 mg/mL and added to the vial as described for the pure solvent. Smaller sample sizes were used in some cases to avoid saturation in one of the phases. The vials were shaken for 30 s and allowed to stand for 1 h or overnight at room temperature (22 ± 2 °C). Sample volumes of 1 μL from each phase were taken for calculation of the partition coefficients using the relationship

$$K_{\mathrm{p}} = \frac{S_{\mathrm{cs}}}{S_{\mathrm{f}}}\frac{I_{\mathrm{f}}}{I_{\mathrm{cs}}}K_{\mathrm{p}}^{\mathrm{IS}} \qquad (2)$$

where K_p is the partition coefficient for compound S, S_{cs} and S_f the peak area for compound S in the counter solvent and formamide layer, respectively, I_{sc} and I_f the peak area of the internal standard in the counter solvent and formamide layer, respectively, and $K_p{}^{IS}$ the partition coefficient for the internal standard in the counter solvent-formamide system. The internal standard for 1,2-dichloroethane-formamide was 4-chloro-3-methylphenol $K_p = 1.340 \pm 0.004$ ($n = 10$), for 1-octanol–formamide 5-chloro-2-nitroanisole $K_p = 1.309 \pm 0.007$ ($n = 15$), and for isopentyl ether–formamide 5-chloro-2-nitroanisole $K_p = 1.014 \pm 0.006$ ($n = 10$). Formamide has only a weak response to the flame ionization detector and the main peak for the fomamide layer eluting between 1 and 1.2 min in the temperature program described in section 4.2 is a result of the thermal breakdown of formamide [40].

4.4. Calculations

Multiple linear regression analysis and statistical calculations were performed on a Dell Dimension 9200 computer (Austin, TX, USA) using the program PASW v18.0 (PASW, Chicago, IL, USA). The solute descriptors were taken from an in-house database [15,40,41] and are summarized in Table 2 together with the experimental partition coefficients. The Kennard–Stone algorithm programmed in visual basic for use in Excel 2007 (Microsoft Corporation, Redmond, WA) was used to split the data set into a training set and a test set [42].

5. Results and discussion

The solvation parameter model provides a suitable mechanism for studying liquid–liquid partition systems capable of revealing the contribution of intermolecular interactions responsible for differences in individual partition coefficients and for simulating the separation properties (partition coefficients) for compounds with known descriptor values that lack experimental values [8,15,43]. This requires the calculation of the system constants of the solvation parameter models for the partition systems involving the measurement of partition coefficients for a number of varied compounds with known descriptor values. Several methods have been proposed to define a minimum number of compounds to solve Eq. (1) by multiple linear regression analysis [44–46]. Models based on small data sets, even if they meet the minimum number requirement, are often of limited utility for predicting further partition coefficients [47,48]. A contributing factor is that the error in the partition coefficients is not random and tends to be correlated with the size of the partition coefficient. Both large and small partition coefficients have larger errors because of the higher uncertainty in the determination of the low concentration of the compound that exists in one of the phases. The experimental partition coefficients should span a reasonable range of values to facilitate modeling. In practice, the number of solutes should be sufficient to obtain a stable model and to facilitate splitting of the data set into a training set and test set for validation purposes [44,47–50]. The solutes selected to build the model define the descriptor space, which for practical applications should be as wide as possible. The descriptor values for the selected solutes should be somewhat evenly distributed over the descriptor space and each series of descriptors should have a low correlation with each other. The descriptor space for the systems studied here is defined by the minimum and maximum value for each descriptor given in Table 2. These correspond to −0.255–2.292 for E, −0.078–1.942 for S, 0–1.507 for B, 0.775–2.503 for V, and 0–0.927 (or 0–1.312 for isopentyl ether–formamide) for A. Histogram-type plots for each descriptor were scrutinized to visualize how well the descriptor values cover the descriptor space and additional solutes selected where needed [20,43]. Principal component analysis with the descriptor as variables offers an alternative approach to assess how well the descriptors cover the descriptor space [51]. Unintentional correlation between individual or pairwise descriptor series ($r > 0.8$) result in a loss of capability of the multiple linear regression algorithm to distinguish between the complementary system effects [18,47,48]. The cross-correlation matrix for each model was checked to ensure this was not a problem. The three partition systems studied in this report have complementary separation properties and the identity of the solutes used to define each model was optimized for each system to ensure that a useful range of partition coefficients

Table 2
Descriptor values and partition coefficients for varied compounds in the system organic solvent–formamide.

Compound	Descriptors					Partition coefficients[a]		
	E	*S*	*A*	*B*	*V*	Log K_p		
						For/Dce	Oct/For	Ipe/For
Acenaphthene	1.604	1.050	0	0.220	1.2586	−1.906	1.063	1.489
Acenaphthylene	1.557	1.119	0	0.200	1.2156	−1.706	0.821	1.311
Acetanilide	0.960	1.144	0.538	0.708	1.1137	0.288		
Acetophenone	0.806	1.026	0	0.503	1.0138	−0.924	0.297	0.045
3-Aminopropyltriethoxysilane	−0.021	0.487	0.124	1.313	1.898	−1.457	1.119	1.084
Aniline	0.955	1.003	0.249	0.425	0.8162	−0.324	−0.043	
Benzamide	1.258	1.343	0.648	0.664	0.9728	0.817	−0.378	−1.427
Benzaldehyde	0.813	1.025	0.000	0.394	0.8730	−0.839	0.159	
1,4-Benzodioxan	0.884	1.054	0.000	0.354	1.0070	−1.163	0.277	0.455
Benzonitrile	0.742	1.135	0.000	0.331	0.8711	−0.924	0.217	−0.290
Benzophenone	1.224	1.330	0.000	0.576	1.4808	−1.436	0.688	0.735
Benzyl benzoate	1.264	1.280	0.000	0.597	1.6804	−1.970	0.828	1.125
Biphenyl	1.312	0.874	0.000	0.298	1.3242	−1.683	1.101	1.526
1-Bromonaphthalene	1.598	1.005	0.000	0.157	1.2604	−1.853	1.075	1.551
1-Bromoheptane	0.343	0.400	0.000	0.120	1.2699			1.974
1-Bromohexane	0.349	0.400	0.000	0.120	1.1290			1.565
1-Bromooctane	0.339	0.400	0.000	0.120	1.4108	−2.407		2.383
3-Bromophenol	1.081	0.792	0.948	0.201	0.9501		0.445	−0.127
4-Bromophenol	1.080	1.170	0.670	0.200	0.9501		0.103	−0.711
Caffeine	1.518	1.726	0.039	1.232	1.3632	−0.125		
Carbazole	2.025	1.585	0.367	0.231	1.3154	−0.915	0.686	0.519
2-Chloroaniline	1.026	0.965	0.253	0.321	0.9386	−0.570	0.412	0.077
4-Chloroaniline	1.056	1.138	0.325	0.331	0.9386	−0.437	0.149	−0.236
Chlorobenzene	0.718	0.656	0.000	0.056	0.8388	−1.361		
4-Chloro-3-methylphenol	0.920	1.020	0.650	0.230	1.0384	0.127	0.422	−0.028
1-Chloronaphthalene	1.419	0.951	0.000	0.135	1.2078	−1.678	1.186	1.613
4-Chlorophenol	1.016	0.794	0.886	0.205	0.8980			−0.280
Cinnamyl alcohol	1.067	0.959	0.490	0.600	1.1548	−0.123	0.366	
Coumarin	1.269	1.610	0.000	0.524	1.0619	−0.694	−0.034	−0.720
o-Cresol	0.774	0.745	0.621	0.357	0.9160	0.158		−0.104
Decan-1-ol	0.191	0.440	0.344	0.520	1.5763	−1.556		1.598
Dibenzofuran	1.562	1.094	0.000	0.106	1.2087	−1.799	0.942	1.406
Dibenzylamine	1.340	0.985	0.115	1.063	1.7058	−1.443	0.827	1.082
3,4-Dichloroaniline	1.338	1.280	0.545	0.063	1.0610	−0.506	0.336	0.056
1,4-Dichlorobenzene	0.825	0.750	0.000	0.020	0.9612	−1.599	0.927	1.251
2,4-Dichlorophenol	0.960	0.990	0.580	0.140	1.0199		0.425	0.018
Diethyl phthalate	0.729	1.465	0.000	0.869	1.7106	−1.517	0.053	0.386
N,N-Dimethylaniline	0.957	0.840	0.000	0.410	1.0960	−1.436	0.889	1.081
2,6-Dimethylphenol	0.773	0.791	0.408	0.402	1.0569		0.544	0.137
Dimethyl phthalate	0.780	1.410	0.000	0.880	1.4288	−0.996	0.430	−0.341
1,3-Dinitrobenzene	1.027	1.756	0.000	0.399	1.0648	−0.887	−0.334	−0.678
Diphenylamine	1.676	1.204	0.214	0.555	1.4240	−1.267	0.851	0.744
Dodecane	0.000	0.000	0.000	0.000	1.7994			3.660
Ethyl benzoate	0.694	0.886	0.000	0.444	1.2135	−1.187	0.604	0.964
Ethyl 4-hydroxybenzoate	0.860	1.350	0.690	0.450	1.2720	0.164	0.102	
Fluoranthene	2.292	1.486	0.000	0.255	1.5846	−2.284	1.085	1.595
Fluorene	1.664	1.120	0.000	0.252	1.3565	−1.727	1.228	1.553
Glycidoxypropyltrimethoxysilane	0.133	1.090	0.000	0.970	1.8073		0.624	
Heptan-2-one	0.108	0.670	0.000	0.510	1.1106		0.609	0.513
Hexachlorobenzene	1.490	0.990	0.000	0.000	1.4508			2.064
Hexanophenone	0.790	1.026	0.000	0.503	1.5775	−1.842	1.075	1.405
4-Hydroxybenzaldehyde	1.110	1.241	0.927	0.463	0.9317	1.010	−0.324	−1.390
Indole	1.018	1.184	0.390	0.240	0.9464	−0.359	0.155	−0.059
Iodobenzene	1.182	0.784	0.000	0.135	0.9747	−1.557		1.153
Isocyanopropyltriethoxysilane	−0.049	0.642	0.000	0.823	2.0119	−2.306	1.459	1.837
Isopentyl ether	0.000	0.250	0.000	0.450	1.5760	−2.255	1.582	
Methacryloxypropyltrimethoxysilane	0.046	0.869	0.000	1.024	1.9708	−1.853	0.985	1.047
Methyl benzoate	0.738	0.923	0.000	0.439	1.0726	−1.271	0.552	0.514
2-Methoxynaphthalene	1.449	1.140	0.000	0.359	1.2850	−1.695	1.001	1.176
1-Methylnapthalene	1.337	0.915	0.000	0.205	1.2263	−1.855	1.267	1.599
2-Methylnaphthalene	1.304	0.881	0.000	0.244	1.2263	−1.698	1.279	1.619
Naphthalene	1.240	0.906	0.000	0.193	1.0854	−1.624	0.876	1.344
1-Naphthol	1.480	1.157	0.796	0.318	1.1441	−0.064	0.456	−0.012
2-Naphthol	1.457	1.181	0.807	0.345	1.1441	0.070	0.416	−0.237
2-Nitroaniline	1.182	1.441	0.386	0.348	0.9904	−0.315		−0.713
4-Nitroaniline	1.236	1.827	0.597	0.343	0.9904	0.313	−0.406	−1.546
Nitrobenzene	0.846	1.138	0.000	0.269	0.8906	−0.934	0.216	0.050
1-Nitronaphthalene	1.367	1.505	0.000	0.272	1.2569	−1.522	0.571	0.571
2-Nitrophenol	0.962	1.086	0.050	0.371	0.9493	−0.608	0.119	0.132
2-Nitrotoluene	0.866	1.110	0.000	0.270	1.0315	−1.322	0.426	0.427
3-Nitrotoluene	0.874	1.100	0.000	0.250	1.0315	−1.195	0.439	0.436

Table 2 (*Continued*)

Compound	Descriptors					Partition coefficients[a]		
	E	*S*	*A*	*B*	*V*	Log K_p		
						For/Dce	Oct/For	Ipe/For
4-Nitrotoluene	0.918	1.194	0.000	0.264	1.0315	−1.307	0.341	0.343
Nonan-1-ol	0.199	0.440	0.344	0.520	1.4354	−0.945	1.335	1.236
Nonan-2-one	0.113	0.662	0.000	0.496	1.3924	−1.636		1.191
Octanal	0.148	0.633	0.000	0.421	1.2515	−1.805	1.086	1.155
Octan-1-ol	0.199	0.440	0.344	0.520	1.2945	−0.807		0.769
2-Octanone	0.109	0.662	0.000	0.496	1.2515	−1.631	0.883	
Octanophenone	0.779	1.026	0.000	0.503	1.8593	−2.135	1.399	1.722
n-Octyltriethoxysilane	−0.255	−0.078	0.000	0.985	2.5030	−2.985	2.651	3.614
Pentachlorophenol	1.217	0.860	0.610	0.090	1.3870		1.074	0.959
Phthalimide	1.219	1.729	0.214	0.622	1.0208	−0.078	−0.451	−1.354
Phthalonitrile	0.755	1.942	0.000	0.360	1.0256	−0.784	−0.341	−1.418
Phenanthrene	1.997	1.316	0.000	0.279	1.4544	−1.911	1.125	1.304
Phenyl acetate	0.648	1.055	0.000	0.521	1.0726	−1.165	0.427	0.152
2-Phenylacetamide	0.950	1.600	0.520	0.790	1.1140	0.604	−0.761	−1.698
Phenol	0.769	0.759	0.716	0.319	0.7751	0.533	0.205	
Phenyl benzoate	1.330	1.420	0.000	0.470	1.5400	−1.866	0.609	0.849
1-Phenylethanol	0.823	0.819	0.351	0.648	1.0569	0.011		−0.245
2-Phenylethanol	0.787	0.797	0.390	0.636	1.0569	0.032	0.350	
Phenyl ether	1.216	0.912	0.000	0.267	1.3829	−1.942	1.172	1.547
4-Phenylphenol	1.510	1.178	0.853	0.437	1.3829	−0.238	0.632	0.025
Quinoline	1.268	1.090	0.000	0.562	1.0443	−0.778	0.341	0.135
Resorcinol	1.038	0.995	1.312	0.511	0.8338		−0.152	−1.910
Thiophene	0.687	0.560	0.000	0.150	0.6411		0.446	
o-Toluidine	0.966	1.045	0.193	0.491	0.9571	−0.565		−0.036
m-Toluidine	0.946	1.128	0.112	0.516	0.9571	−0.714		−0.020
p-Toluidine	0.923	1.192	0.147	0.396	0.9571	−0.509		−0.509
p-Tolualdehyde	0.862	1.000	0.000	0.420	1.0139	−1.114	0.444	0.128
1,2,4-Trichlorobenzene	1.022	0.748	0.000	0.018	1.0836	−1.736	1.154	1.617
Tri-n-butyrin	0.091	1.230	0	1.507	2.4453	−2.096	0.655	0.655
Valerophenone	0.795	1.026	0.000	0.503	1.4366	−1.599	0.820	1.092

[a] K_p, partition coefficient; For, formamide; Dce, 1,2-dichloroethane; Oct, 1-octanol; Ipe, isopentyl ether.

was maintained. As a consequence, the compounds identified in Table 2 are not the same for each system but cover a similar range of descriptor values.

5.1. Formamide–1,2-dichloroethane partition system

Fitting the partition coefficients ($\log K_p$) in Table 2 to the solvation parameter model gave

$$\log K_p = -0.207(\pm 0.081) - 0.082(\pm 0.046)E + 0.399(\pm 0.056)S + 1.957(\pm 0.060)A + 1.298(\pm 0.079)B - 1.705(\pm 0.058)V \quad (3)$$

$$r = 0.989 \quad r_{adj}^2 = 0.977 \quad SE = 0.122 \quad F = 738 \quad n = 87$$

where r is the multiple correlation coefficient, ${r_{adj}}^2$ the coefficient of determination adjusted for the number of degrees of freedom, SE the standard error of the estimate, F the Fisher statistic, and n the number of compounds with partition coefficients included in the model. The driving force for transfer of solutes to the 1,2-dichloroethane layer is indicated by the system constants with negative coefficients, since the 1,2-dichloroethane-rich layer has a higher density than the formamide-rich layer. This is governed nearly completely by solute size (the v system constant) since the e system constant is small and only just significant at the 95% confidence level (Student's t test). Polar interactions characterized by the s, a, and b system constants favor transfer to the formamide-rich layer (see Section 5.4 for additional details).

To evaluate the predictive ability of the model the data set was split into a training set of 60 compounds and a test set of 27 compounds using the Kennard–Stone algorithm [42]. This approach ensures that the training set and the test set are selected to occupy a similar descriptor space. The model for the training set, Eq. (4), is virtually identical to Eq. (3). Eq. (4) was

$$\log K_p = -0.200(\pm 0.099) - 0.104(\pm 0.045)E + 0.428(\pm 0.058)S + 1.912(\pm 0.065)A + 1.317(\pm 0.081)B - 1.716(\pm 0.067)V \quad (4)$$

$$r = 0.992 \quad r_{adj}^2 = 0.982 \quad SE = 0.120 \quad F = 649 \quad n = 60$$

then used to predict the partition coefficients ($\log K_p$) for the compounds in the test set and the average error, average absolute error, and root mean square error of the difference between the experimental and model predicted values used to assess the ability of Eq. (4) to estimate further values of $\log K_p$ within the same descriptor space. The average error is an indication of bias and at 0.059 indicates that this is not a concern for Eq. (4). The absolute average error (0.112) and root mean square error (0.135) are an indication of the likely error in predicting further partition coefficients based on Eq. (4). Since Eq. (4) is similar to Eq. (3), which is preferred because it is based on a larger number of compounds, it is reasonable to conclude that Eq. (3) should be able to predict partition coefficients to about ±0.13 log units for further compounds with known descriptor values that lie within or close to the descriptor space used to define the model.

5.2. 1-Octanol–formamide partition system

Fitting the partition coefficients ($\log K_p$) in Table 2 to the solvation parameter model gave

$$\log K_p = 0.285(\pm 0.063) + 0.267(\pm 0.034)E - 1.053(\pm 0.043)S - 0.333(\pm 0.038)A - 0.929(\pm 0.066)B + 1.314(\pm 0.046)V \quad (5)$$

Table 3
System constants for biphasic liquid-liquid partition systems.

Partition system	System constants				
	e	*s*	*a*	*b*	*v*
Formamide–1,2-dichloroethane	0.082	−0.399	−1.957	−1.298	1.705
n-Heptane–formamide	0.561	−2.248	−3.250	−1.603	2.384
n-Heptane–N,N-dimethylformamide	0.038	−1.391	−2.160	−0.593	0.486
n-Heptane–2,2,2-trifluoroethanol	0.882	−1.557	−1.312	−2.928	1.301
n-Heptane–methanol	0.186	−0.686	−1.098	−0.951	0.618
n-Heptane–ethylene glycol	0.374	−1.889	−4.072	−1.942	0.618
n-Hexane–acetonitrile	0.349	−1.439	−1.611	−0.874	0.669
Isopentyl ether–formamide	0.564	−1.715	−1.314	−1.407	2.005
1-Octanol–formamide	0.267	−1.053	−0.333	−0.929	1.314
Cyclohexane–water	0.784	−1.678	−3.740	−4.929	4.577
Di-n-butyl ether–water	0.677	−1.506	−0.807	−5.249	4.815
n-Heptane–water	0.670	−2.061	−3.317	−4.733	4.543
1-Octanol–water	0.684	−1.209	−0.185	−3.355	3.846
Toluene–water	0.527	−0.720	−3.010	−4.824	4.545
Water–chloroform	0.183	−0.380	−2.469	−3.426	3.973
Water–1,2-dichloroethane	0.159	0.135	−2.247	−4.776	4.177
Water–methanol–chloroform (3:4:8)	0	−0.333	−1.407	−1.501	1.345

$$r = 0.986 \quad r^2_{\mathrm{adj}} = 0.970 \quad \mathrm{SE} = 0.095 \quad F = 525 \quad n = 82$$

Positive system constant (*v* and *e*) favor transfer to the 1-octanol-rich layer with polar interactions have a negative sign (*s*, *a* and *b*) and favor solubility in the formamide-rich layer. 1-Octanol is more competitive than 1,2-dichloroethane as a reservoir for hydrogen-bonding interactions reducing the value of the *a* and *b* system constants. It is also significantly less competitive for interactions of a dipole-type (*s* system constant) but more cohesive (smaller *v* system constant) than 1,2-dichloroethane. As before, the Kennard–Stone algorithm was used to split the data set into a training set of 59 compounds and a test set of 23 compounds. The model for the training set is given below

$$\log K_{\mathrm{p}} = 0.270(\pm 0.073) + 0.263(\pm 0.036)E - 1.030(\pm 0.044)S - 0.325(\pm 0.043)A - 0.930(\pm 0.069)B + 1.305(\pm 0.050)V \quad (6)$$

$$r = 0.989 \quad r^2_{\mathrm{adj}} = 0.976 \quad \mathrm{SE} = 0.094 \quad F = 467 \quad n = 59$$

and is quite similar to Eq. (5). For the test set the average error was 0.086, the average absolute error 0.114 and the root mean square error 0.101. Thus, Eq. (5) should be able to predict further values of the partition coefficients to about 0.11 log units for compounds with descriptor values that lie within or close to the descriptor space used to define the model.

5.3. Isopentyl ether–formamide partition system

Fitting the partition coefficients ($\log K_{\mathrm{p}}$) in Table 2 to the solvation parameter model gave

$$\log K_{\mathrm{p}} = 0.130(\pm 0.082) + 0.564(\pm 0.041)E - 1.715(\pm 0.055)S - 1.314(\pm 0.047)A - 1.407(\pm 0.074)B + 2.005(\pm 0.056)V \quad (7)$$

$$r = 0.994 \quad r^2_{\mathrm{adj}} = 0.987 \quad \mathrm{SE} = 0.119 \quad F = 1347 \quad n = 88$$

Positive system constant (*v* and *e*) favor transfer to the isopentyl ether-rich layer while polar interactions have a negative sign (*s*, *a* and *b*) and favor solubility in the formamide-rich layer. The relatively low cohesion of isopentyl ether compared with formamide results in a relatively large *v* system constant and the polar characteristics of isopentyl ether are reflected in the intermediate values for the *s* and *a* system constants for the isopentyl ether–formamide partition system. The selectivity of the isopentyl ether–formamide system is closest to the n-heptane–formamide system but with smaller *s* and *a* system constants reflecting the contribution of the ether oxygen to the partition mechanism. The Kennard–Stone algorithm was used to split the data set into a training set of 62 compounds and a test set of 26 compounds. The model for the training set is given below

$$\log K_{\mathrm{p}} = 0.076(\pm 0.101) + 0.568(\pm 0.048)E - 1.713(\pm 0.062)S - 1.308(\pm 0.058)A - 1.403(\pm 0.082)B + 2.027(\pm 0.066)V \quad (8)$$

Table 4
Results from principal component analysis with varimax rotation and Kaiser normalization for the biphasic partition systems indicated in Table 3.

Principal component	Percent variance	Cumulative percent variance
(i) Extraction of principal components		
PC-1	50.50	
PC-2	27.61	78.11
PC-3	18.46	96.57

System constant	PC-1	PC-2	PC-3
(ii) Loading on variables on the principal components			
e	0.499	−0.832	0.094
s	0.135	0.942	0.233
a	−0.156	0.113	0.977
b	−0.977	0.091	0.117
v	0.966	−0.050	−0.120

$$r = 0.995 \quad r^2_{\mathrm{adj}} = 0.988 \quad \mathrm{SE} = 0.127 \quad F = 1033 \quad n = 62$$

and is quite similar to Eq. (7). For the test set the average error was 0.067, the average absolute error 0.087 and the root mean square error 0.108. Thus, Eq. (7) should be able to predict further values of the partition coefficients to about 0.12 log units for compounds with descriptor values that lie within or close to the descriptor space used to define the model.

5.4. Comparison of water and formamide for transfer of neutral solutes to organic solvents

Table 3 summarizes the system constants for the transfer of neutral organic compounds from water to n-heptane [52,53], 1,2-dichloroethane [15,54], and 1-octanol [15,53] and from formamide to n-heptane [40], 1,2-dichloroethane, isopentyl ether, and 1-octanol. System constants are not available for the isopentyl ether–water system and the di-n-butyl ether–water system is used as a surrogate for comparison purposes [15,55,56]. The selectivity of the water–organic solvent and formamide–organic solvent systems is clearly different but certain general trends can be deduced. Formamide is about one-third to one-half as cohesive as water and only about one-quarter to one-third as hydrogen-bond acidic. The relatively high cohesion and hydrogen-bond acidity are the two characteristic properties that tend to set water apart from other common solvents. Formamide and water have similar hydrogen-bond basicity and dipolarity/polarizability being the dominant properties that account for the particular characteristics of formamide. With respect to the above comments it should be kept in mind that the systems being compared refer to the equilibrium solvent compositions in which each phase is saturated with its counter solvent and differences in solvent saturation are not specifically taken into account in these comparisons. Formamide can be seen to possess some of the general characteristic solvation properties of water, but only to an extent, and it should be considered complementary in solvation properties to water rather than a substitute.

5.5. General extraction properties of formamide–organic solvent systems

Principal component analysis with varimax rotation and system constants as variables can be used to compare the extraction properties of the totally organic biphasic systems and the water-based biphasic systems typically used for descriptor measurements, Table 3 [15,26,40,57,58]. The first two principal components describe 78% of the variance but the two-dimensional score plots provide poor a classification of the partition systems, Table 4. The first three principal components explain about 97% of the variance and the three-dimensional plot of these principal components, Fig. 1, is suitable for classification purposes. Principal component 1 (PC-1) mainly expresses information about the *b* and *v* system constants, principal component 2 (PC-2) the *e* and *s* system constants, and principal component 3 (PC-3) the *a* system constant. The water-based biphasic systems (numbered 10–16 in Fig. 1) are grouped at the top of the figure separated in the vertical plane from the totally organic biphasic systems. This highlights the dominant properties of water, its high cohesion and strong hydrogen-bond acidity, which sets the water-based partition systems apart from the other partition systems. The n-heptane–water and cyclohexane–water systems (15 and 11) are indicated as having similar selectivity while the other water-based biphasic systems have complementary properties. The 1-octanol–water and water–chloroform systems are closer to the totally organic partition systems since the water saturated organic counter solvents compete to a greater extent than the other organic solvents as a reservoir of hydrogen-bonding interactions and also reduce the difference in cohesion between the two phases. Just below the water-based biphasic systems are the totally organic biphasic systems with intermediate hydrogen-bond acidity and cohesion. These systems are represented by n-heptane–2,2,2-trifluoroethanol (5), n-heptane–formamide (1), n-heptane–ethylene glycol (9), isopentyl ether–formamide (4), and formamide–1,2-dichloroethane (2). For compounds virtually insoluble or unstable in water these totally organic biphasic systems would be the most useful for estimating the *B* descriptor. The ternary solvent system (water–methanol–chloroform, Folch partition system) and formamide–1,2-dichloroethane are almost selectivity equivalent and for many sample preparation applications one system could be substituted for the other. For determination of the *S* descriptor the n-heptane–formamide (1) and n-heptane–water (15) biphasic systems have the most favorable weighting on PC-2 with 1-octanol–formamide (3) and isopentyl ether–formamide (4) indicated as useful systems because the absolute value of the *s* system constant is intermediate in value and the relative contribution of dipole-type interactions to the partition coefficient is significantly larger than for the other biphasic systems. Of the totally organic biphasic systems n-heptane–formamide (1), n-heptane–N,N-dimethylformamide (6), and n-heptane–ethylene glycol (9) have a favorable loading on PC-3 for determination of the *A* descriptor. The *E* and *V* descriptor can be obtained by calculation and experimental methods are not generally required for their determination [8,15,16].

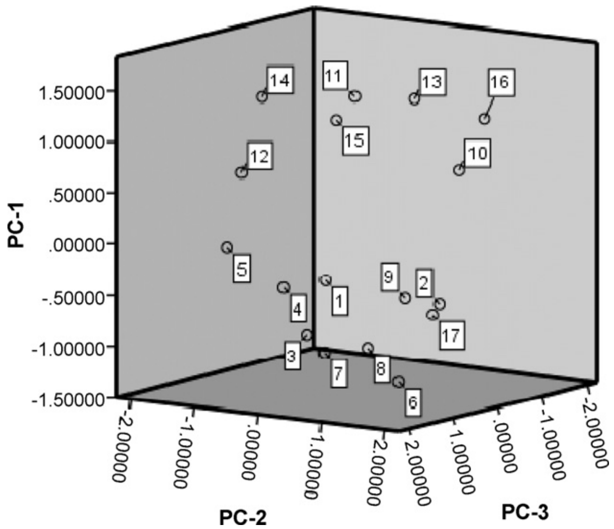

Fig. 1. Score plot of the first three principle components with the system constants as variables for 17 biphasic liquid–liquid partition systems. Identification: 1, n-heptane–formamide; 2, formamide–1,2-dichloroethane; 3, 1-octanol–formamide; 4, isopentyl ether–formamide; 5, n-heptane–2,2,2-trifluoroethanol; 6, n-heptane–N,N-dimethylformamide; 7, n-hexane–acetonitrile; 8, n-heptane–methanol; 9, n-heptane–ethylene glycol; 10, water–chloroform; 11, cyclohexane–water; 12, 1-octanol–water; 13, toluene–water; 14, di-n-butyl ether–water; 15, n-heptane–water; 16, water–1,2-dichloroethane; 17, Folch partition (chloroform–methanol–water).

6. Conclusions

Formamide is demonstrated to be a useful solvent for liquid–liquid partition forming several complementary biphasic systems with organic solvents that should be suitable for sample preparation and the determination of descriptors. Formamide is significantly more cohesive than typical organic solvents but probably about half as cohesive as water. It is moderately hydrogen-bond acidic compared with water and about as hydrogen-bond basic and dipolar/polarizable. For compounds that are virtually insoluble or unstable in water it provides a series of biphasic partition systems that can be used together with n-heptane–2,2,2-trifluoroethanol and n-heptane–N,N-dimethylformamide for the experimental determination of the *S*, *A* and *B* descriptors used in the solvation parameter model. For the calculation of descriptors it is better to use several experimental approaches including chromatographic and solubility methods in conjunction with liquid–liquid partition approaches [8,15].

References

[1] A. Sarafraz-Yazdi, A. Amiri, Trends Anal. Chem. 29 (2010) 1.
[2] M.A. Jeannot, A. Przyjazny, J.M. Kobosa, J. Chromatogr. A 1217 (2010) 2326.
[3] M. Rezaee, Y. Yamini, M. Faraji, J. Chromatogr. A 1217 (2010) 2342.
[4] S.K. Poole, T.A. Dean, J.W. Oudsema, C.F. Poole, Anal. Chim. Acta 236 (1990) 3.
[5] Y. Chen, Z. Guo, X. Wang, C. Qiu, J. Chromatogr. A 1184 (2008) 191.
[6] C.F. Poole, Trends Anal. Chem. 22 (2003) 362.
[7] A. Berthod, S. Carda-Broch, J. Chromatogr. A 1037 (2004) 3.
[8] M.H. Abraham, A. Ibrahim, A.M. Zissimos, J. Chromatogr. A 1037 (2004) 29.
[9] M.H. Abraham, F. Martins, R.C. Mitchell, C.J. Salter, J. Pharm. Sci. 88 (1999) 241.
[10] H. Ahmed, C.F. Poole, J. Chromatogr. A 1104 (2006) 82.
[11] H. Ahmed, C.F. Poole, J. Sep. Sci. 29 (2006) 2158.
[12] J. Qian, C.F. Poole, J. Chromatogr. A 1143 (2007) 276.
[13] F.T.T. Huque, K. Jones, R.A. Saunders, J.A. Platts, J. Fluor. Chem. 115 (2002) 119.
[14] D.F.S. Natusch, B.A. Tomkins, Anal. Chem. 50 (1978) 1429.
[15] C.F. Poole, S.N. Atapattu, S.K. Poole, A.K. Bell, Anal. Chim. Acta 652 (2009) 32.
[16] M.H. Abraham, C.F. Poole, S.K. Poole, J. Chromatogr. A 842 (1999) 79.
[17] C.F. Poole, S.K. Poole, J. Chromatogr. A 965 (2002) 263.
[18] M.A. Vitha, P.W. Carr, J. Chromatogr. A 1126 (2006) 143.
[19] S.K. Poole, C.F. Poole, J. Chromatogr. A 1182 (2008) 1.
[20] C.F. Poole, S.K. Poole, J. Chromatogr. A 1184 (2008) 254.
[21] C.F. Poole, A.D. Gunatilleka, S.K. Poole, Adv. Chromatogr. 40 (2000) 159.
[22] T.H. Nguyen, K.-H. Goss, W.P. Ball, Environ. Sci. Technol. 39 (2005) 913.
[23] S.N. Atapattu, C.F. Poole, J. Environ. Monit. 11 (2009) 815.
[24] M.H. Abraham, G.M.R. Gola, R. Kumarsingh, J.E. Cosmetto-Muniz, W.S. Cain, J. Chromatogr. B 745 (2000) 103.
[25] H. Ahmed, C.F. Poole, G.E. Kozerski, J. Chromatogr. A 1169 (2007) 179.
[26] S.N. Atapattu, C.F. Poole, J. Chromatogr. A 1216 (2009) 7882.
[27] K. Herodes, I. Leito, I. Koppel, M. Roses, J. Phys. Org. Chem. 12 (1999) 109.
[28] L. Geiser, M. Mirgaldi, J.-L. Veuthey, J. Chromatogr. A 1068 (2005) 75.
[29] S.P. Porras, E. Kenndler, Electrophoresis 25 (2004) 2946.
[30] A. Lattes, E. Perez, I. Rico-Lattes, C. R. Chim. 12 (2009) 45.
[31] J.F. Lu, Z.Y. Zhou, Q.Y. Wu, G. Zhao, J. Mol. Struct. Theochem. 724 (2005) 107.
[32] A. Tellez, E. Kenndler, Electrophoresis 30 (2009) 357.
[33] J.J. Silber, R.D. Falcone, N.M. Correra, M.A. Biasutti, E. Abuin, E. Lissi, P. Campodonico, Langmuir 19 (2003) 2067.
[34] A. Hohn, Kirk-Othmer Encyclopedia of Chemical Technology, Wiley, New York, NY, USA, 2001.
[35] J. Fuchs, D. Dell'Atti, A. Buhot, R. Calemezuk, Anal. Biochem. 397 (2010) 132.
[36] J.W. Oudsema, C.F. Poole, J. High Resolut. Chromatogr. 16 (1993) 130.
[37] H. Kazoka, V. Shatz, J. Chromatogr. A 732 (1996) 231.
[38] J.D. George, C.J. Price, M.C. Marr, C.B. Myers, G.D. Jahnke, Toxicol. Sci. 69 (2002) 165.
[39] M.H. Abraham, W.E. Acree, J.E. Cometto-Muniz, New J. Chem. 33 (2009) 2034.
[40] T. Karunasekara, C.F. Poole, J. Sep. Sci. 33 (2010) 1167.
[41] S.N. Atapattu, C.F. Poole, J. Chromatogr. A 1195 (2008) 136.
[42] R.W. Kennard, L.A. Stone, Technometrics 11 (1969) 137.
[43] A.M. Zissimos, M.H. Abraham, M.C. Barker, K.J. Box, K.Y. Tam, J. Chem. Soc. Perkin Trans. 2 (2003) 470.
[44] D. Livingstone, A Practical Guide to Scientific Data Analysis, Wiley, Chichester, 2009.
[45] M. Sergent, D. Mathieu, R. Phan-tan-luu, G. Drava, Chemomet. Intell. Lab. Syst. 27 (1995) 153.
[46] J.P.M. Andries, H.A. Claessens, Y. Vander Heyden, L.M.C. Buydens, Anal. Chim. Acta 652 (2009) 180.
[47] C.F. Poole, S.K. Poole, M.H. Abraham, J. Chromatogr. A 798 (1998) 207.
[48] M.D. Trone, M.G. Khaledi, J. Chromatogr. A 886 (2000) 245.
[49] D.M. Hawkins, S.C. Basak, D. Mills, J. Chem. Inf. Comput. Sci. 43 (2003) 579.
[50] D.M. Hawkins, J. Chem. Inf. Comput. Sci. 44 (2004) 1.
[51] E. Fuguet, C. Rafols, E. Bosch, M.H. Abraham, M. Roses, J. Chromatogr. A 942 (2002) 237.
[52] M.H. Abraham, W.E. Acree, Can. J. Chem. Eng. 83 (2005) 362.
[53] M.H. Abraham, H.S. Chadha, G.S. Whiting, R.C. Mitchell, J. Pharm. Sci. 83 (1994) 1085.
[54] L.M. Sprunger, J. Gibbs, W.E. Acree, M.H. Abraham, Fluid Phase Equilib. 273 (2008) 78.
[55] M.H. Abraham, A.M. Zissimos, W.E. Acree, New J. Chem. 27 (2005) 1041.
[56] M.H. Abraham, A.M. Zissimos, W.E. Acree, Phys. Chem. Chem. Phys. 3 (2001) 3732.
[57] J. Qian, C.F. Poole, J. Sep. Sci. 30 (2007) 2326.
[58] M.H. Abraham, R.E. Smith, R. Luchtefeld, A.J. Boorem, R. Luo, W.E. Acree, J. Pharm. Sci. 99 (2010) 1500.