

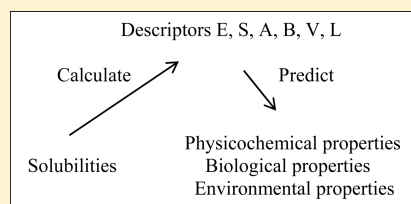
# Deduction of Physicochemical Properties from Solubilities: 2,4-Dihydroxybenzophenone, Biotin, and Caprolactam as Examples

Michael H. Abraham,<sup>\*,†</sup> William E. Acree, Jr.,<sup>‡</sup> Michela Brumfield,<sup>‡</sup> Erin Hart,<sup>‡</sup> Lila Pipersburgh,<sup>‡</sup> Katherine Mateja,<sup>‡</sup> Colleen Dai,<sup>‡</sup> Damini Grover,<sup>‡</sup> and Shoshana Zhang<sup>‡</sup>

<sup>†</sup>Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

<sup>‡</sup>Department of Chemistry, University of North Texas, 1155 Union Circle Drive #305070, Denton, Texas 76203-5017, United States

**ABSTRACT:** Solubilities of 2,4-dihydroxybenzophenone have been measured at 298 K in 19 organic solvents. Combination with a known value of the solubility in water enables corresponding partition coefficients from water to these solvents to be deduced. It is shown that solubilities, via the partition coefficients, can yield the Abraham descriptors for 2,4-dihydroxybenzophenone, and that these, in turn, can be used to estimate a very large number of further solubilities, partition coefficients, and values of biological and environmental properties. This method of extracting information from solubilities is further illustrated using literature data on solubilities of biotin and of caprolactam.



## 1. INTRODUCTION

There are now numerous reports of the determination of solubilities, often as the solubility of a given compound in a series of solvents. Such determinations are very important in the selection of solvents for crystallization, extraction, and purification. However, it is not widely appreciated that solubilities of a given compound in a series of solvents encode a great deal of information on physicochemical properties of the compound and that it is possible to determine these properties (or descriptors) through an analysis of the solubility data. Once these compound descriptors have been determined, they can be used to estimate a large number of further physicochemical, environmental, and biochemical properties of the compound. We illustrate the methods we have used through an analysis of our own data on solubilities of 2,4-dihydroxybenzophenone, CAS Registry No. 131-56-6, and literature data on solubilities of biotin,<sup>1</sup> CAS Registry No. 58-85-5, and caprolactam,<sup>2</sup> CAS Registry No. 105-62-2.

The method used to extract compound descriptors from solubilities has been described several times,<sup>3–8</sup> and reviews are also available.<sup>9,10</sup> In brief, we use two linear equations, eq 1 and eq 2. The dependent variable, SP, is a set of values for a series of solutes in a given system. Thus, SP in eq 1 can be a series of partition coefficients, as  $\log P$ , in a specific water–solvent system, and SP in eq 2 can be a series of gas–solvent partition coefficients, as  $\log K$ , in a specific solvent.

$$SP = c + eE + sS + aA + bB + vV \quad (1)$$

$$SP = c + eE + sS + aA + bB + lL \quad (2)$$

The independent variables, or descriptors, are properties of the solutes as follows:<sup>3–10</sup>  $E$  is the solute excess molar refractivity in units of  $(\text{cm}^3 \text{ mol}^{-1})/10$ ,  $S$  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  $(\text{cm}^3 \text{ mol}^{-1})/100$ , and  $L$  is the logarithm of the gas–

hexadecane partition coefficient at 298 K. The descriptors, in effect, summarize the propensity of a given solute to interact with surrounding solvent molecules. The coefficients in eq 1 and eq 2 are obtained by multiple linear regression analysis and serve to characterize the system under consideration. The coefficients are listed in Table 1 for a number of water to solvent and gas to solvent systems. It is important to note that the coefficients refer to hypothetical partitions between water or the gas phase and a “dry” solvent. More complete lists of systems, including partitions into “wet” solvents<sup>3–8</sup> and various biological systems<sup>7</sup> are available.

Equations 1 and 2 can be applied to solubilities through eqs 3 and 4.

$$P = C/C_w \quad (3)$$

$$K = C/C_g, \quad K_w = C_w/C_g, \quad K = PK_w \quad (4)$$

In these equations  $C$  and  $C_w$  are the solubilities of a solute, in  $\text{mol dm}^{-3}$ , in a solvent and in water respectively, and  $C_g$  is the saturated vapor concentration of the solute, again in  $\text{mol dm}^{-3}$ . Note that  $K$  and  $K_w$  are dimensionless. Equations 3 and 4 require that in the saturated solutions in water and the solvent the same species is in equilibrium with the solid phase, hence that no solvates are present. Solubilities in water and in a series of solvents of a given solute can then be converted into a series of  $\log P$  values between water and the solvents. Once  $\log P$  values are available for partition into a series of solvents for which the coefficients in eq 1 and eq 2 are known, see Table 1, it is possible to set up a series of simultaneous equations that can be solved to yield the unknown solute descriptors. The number of equations can be considerably increased through eq

Received: December 16, 2014

Accepted: April 7, 2015

Published: April 15, 2015

Table 1. Coefficients for Water–Solvent and Gas–Solvent Partitions, Equations 1 and 2, at 298 K

solvent	<i>c</i>	<i>e</i>	<i>s</i>	<i>a</i>	<i>b</i>	<i>v</i>	<i>l</i>
			log <i>P</i> , Equation 1 in <i>V</i>				
dichloromethane	0.319	0.102	−0.187	−3.058	−4.090	4.324	0.000
trichloromethane	0.191	0.105	−0.403	−3.112	−3.514	4.395	0.000
tetrachloromethane	0.199	0.523	−1.159	−3.560	−4.594	4.618	0.000
1,2-dichloroethane	0.183	0.294	−0.134	−2.801	−4.291	4.180	0.000
hexane	0.333	0.560	−1.710	−3.578	−4.939	4.463	0.000
heptane	0.297	0.634	−1.755	−3.571	−4.946	4.488	0.000
octane	0.241	0.690	−1.769	−3.545	−5.011	4.511	0.000
decane	0.172	0.726	−1.750	−3.446	−4.496	4.489	0.000
dodecane	0.114	0.668	−1.644	−3.545	−5.006	4.459	0.000
hexadecane	0.087	0.667	−1.617	−3.587	−4.869	4.433	0.000
cyclohexane	0.159	0.784	−1.678	−3.740	−4.929	4.577	0.000
methylcyclohexane	0.246	0.782	−1.982	−3.517	−4.293	4.528	0.000
isooctane	0.318	0.555	−1.737	−3.677	−4.864	4.417	0.000
benzene	0.142	0.464	−0.588	−3.099	−4.625	4.491	0.000
toluene	0.125	0.431	−0.644	−3.002	−4.748	4.524	0.000
<i>o</i> -xylene	0.083	0.518	−0.813	−2.884	−4.821	4.559	0.000
<i>m</i> -xylene	0.122	0.377	−0.603	−2.981	−4.961	4.535	0.000
<i>p</i> -xylene	0.166	0.477	−0.812	−2.939	−4.874	4.532	0.000
ethylbenzene	0.093	0.467	−0.723	−3.001	−4.844	4.514	0.000
fluorobenzene	0.139	0.152	−0.374	−3.030	−4.601	4.540	0.000
chlorobenzene	0.065	0.381	−0.521	−3.183	−4.700	4.614	0.000
bromobenzene	−0.017	0.436	−0.424	−3.174	−4.558	4.445	0.000
iodobenzene	−0.192	0.298	−0.308	−3.213	−4.653	4.588	0.000
nitrobenzene	−0.152	0.525	0.081	−2.332	−4.494	4.187	0.000
carbon disulfide	0.047	0.686	−0.943	−3.603	−5.818	4.921	0.000
methanol	0.276	0.334	−0.714	0.243	−3.320	3.549	0.000
ethanol	0.222	0.471	−1.035	0.326	−3.596	3.857	0.000
propan-1-ol	0.139	0.405	−1.029	0.247	−3.767	3.986	0.000
butan-1-ol	0.165	0.401	−1.011	0.056	−3.958	4.044	0.000
pentan-1-ol	0.150	0.536	−1.229	0.141	−3.864	4.077	0.000
hexan-1-ol	0.115	0.492	−1.164	0.054	−3.978	4.131	0.000
heptan-1-ol	0.035	0.398	−1.063	0.002	−4.342	4.317	0.000
octan-1-ol	−0.034	0.489	−1.044	−0.024	−4.235	4.218	0.000
decan-1-ol	−0.058	0.616	−1.319	0.026	−4.153	4.279	0.000
propan-2-ol	0.099	0.344	−1.049	0.406	−3.827	4.033	0.000
isobutanol	0.188	0.354	−1.127	0.016	−3.568	3.986	0.000
butan-2-ol	0.127	0.253	−0.976	0.158	−3.882	4.114	0.000
<i>tert</i> -butanol	0.211	0.171	−0.947	0.331	−4.085	4.109	0.000
3-methylbutan-1-ol	0.073	0.360	−1.273	0.090	−3.770	4.273	0.000
pentan-2-ol	0.115	0.455	−1.331	0.206	−3.745	4.201	0.000
2,2,2-trifluoroethanol	0.395	−0.094	−0.594	−1.280	−1.274	3.088	0.000
ethylene glycol	−0.270	0.578	−0.511	0.715	−2.619	2.729	0.000
diethyl ether	0.330	0.401	−0.814	−0.457	−4.959	4.320	0.000
dibutyl ether	0.203	0.369	−0.954	−1.488	−5.426	4.508	0.000
diisopropyl ether	0.472	0.413	−0.745	−0.632	−5.251	4.059	0.000
methyl <i>tert</i> -butyl ether	0.376	0.264	−0.788	−1.078	−5.030	4.410	0.000
tetrahydrofuran	0.207	0.372	−0.392	−0.236	−4.934	4.447	0.000
1,4-dioxane	0.098	0.350	−0.083	−0.556	−4.826	4.172	0.000
methyl acetate	0.351	0.223	−0.150	−1.035	−4.527	3.972	0.000
ethyl acetate	0.328	0.369	−0.446	−0.700	−4.904	4.150	0.000
propyl acetate	0.288	0.363	−0.474	−0.784	−4.938	4.216	0.000
isopropyl acetate	0.000	0.000	0.000	0.000	0.000	0.000	0.000
butyl acetate	0.248	0.356	−0.501	−0.867	−4.973	4.281	0.000
propanone	0.313	0.312	−0.121	−0.608	−4.753	3.942	0.000
butanone	0.246	0.256	−0.080	−0.767	−4.855	4.148	0.000
cyclohexanone	0.038	0.225	0.058	−0.976	−4.842	4.315	0.000
dimethylformamide	−0.305	−0.058	0.343	0.358	−4.865	4.486	0.000
dimethylacetamide	−0.271	0.084	0.209	0.915	−5.003	4.557	0.000
<i>N</i> -methylpyrrolidinone	0.147	0.532	0.225	0.840	−4.794	3.674	0.000
<i>N</i> -formylmorpholine	−0.032	0.696	−0.062	0.014	−4.092	3.405	0.000

Table 1. continued

solvent	<i>c</i>	<i>e</i>	<i>s</i>	<i>a</i>	<i>b</i>	<i>v</i>	<i>l</i>
log <i>P</i> , Equation 1 in <i>V</i>							
<i>N</i> -methylformamide	0.114	0.407	-0.287	0.542	-4.085	3.471	0.000
<i>N</i> -methylacetamide	0.090	0.205	-0.172	1.305	-4.589	3.833	0.000
formamide	-0.171	0.070	0.308	0.589	-3.152	2.432	0.000
acetonitrile	0.413	0.077	0.326	-1.566	-4.391	3.364	0.000
nitromethane	0.023	-0.091	0.793	-1.463	-4.364	3.460	0.000
dimethyl sulfoxide	-0.194	0.327	0.791	1.260	-4.540	3.361	0.000
tributylphosphate	0.327	0.570	-0.837	-1.069	-4.333	3.919	0.000
propylene carbonate	0.004	0.168	0.504	-1.283	-4.407	3.421	0.000
gas-water	-0.994	0.577	2.549	3.813	4.841	-0.869	0.000
log <i>K</i> , Equation 2 in <i>L</i>							
dichloromethane	0.192	-0.572	1.492	0.460	0.847	0.000	0.965
trichloromethane	0.157	-0.560	1.259	0.374	1.333	0.000	0.976
tetrachloromethane	0.217	-0.435	0.554	0.000	0.000	0.000	1.069
1,2-dichloroethane	0.017	-0.337	1.600	0.774	0.637	0.000	0.921
hexane	0.320	0.000	0.000	0.000	0.000	0.000	0.945
heptane	0.284	0.000	0.000	0.000	0.000	0.000	0.950
octane	0.219	0.000	0.000	0.000	0.000	0.000	0.960
decane	0.159	0.000	0.000	0.000	0.000	0.000	0.972
dodecane	0.017	0.000	0.000	0.000	0.000	0.000	0.989
hexadecane	0.000	0.000	0.000	0.000	0.000	0.000	1.000
cyclohexane	0.163	-0.110	0.000	0.000	0.000	0.000	1.013
methylcyclohexane	0.318	-0.215	0.000	0.000	0.000	0.000	1.012
isooctane	0.264	-0.230	0.000	0.000	0.000	0.000	0.975
benzene	0.107	-0.313	1.053	0.457	0.169	0.000	1.020
toluene	0.085	-0.400	1.063	0.501	0.154	0.000	1.011
<i>o</i> -xylene	0.064	-0.296	0.934	0.647	0.000	0.000	1.010
<i>m</i> -xylene	0.071	-0.423	1.068	0.552	0.000	0.000	1.014
<i>p</i> -xylene	0.113	-0.302	0.826	0.651	0.000	0.000	1.011
ethylbenzene	0.059	-0.295	0.924	0.573	0.098	0.000	1.010
fluorobenzene	0.181	-0.621	1.432	0.647	0.000	0.000	0.986
chlorobenzene	0.064	-0.399	1.151	0.313	0.171	0.000	1.032
bromobenzene	-0.064	-0.326	1.261	0.323	0.292	0.000	1.002
iodobenzene	-0.171	-0.192	1.197	0.245	0.245	0.000	1.002
nitrobenzene	-0.296	0.092	1.707	1.147	0.443	0.000	0.912
carbon disulfide,	0.101	0.251	0.177	0.027	0.095	0.000	1.068
methanol	-0.039	-0.338	1.317	3.826	1.396	0.000	0.773
ethanol	0.017	-0.232	0.867	3.894	1.192	0.000	0.846
propan-1-ol	-0.042	-0.246	0.749	3.888	1.076	0.000	0.874
butan-1-ol	-0.004	-0.285	0.768	3.705	0.879	0.000	0.890
pentan-1-ol	-0.002	-0.161	0.535	3.778	0.960	0.000	0.900
hexan-1-ol	-0.014	-0.205	0.583	3.621	0.891	0.000	0.913
heptan-1-ol	-0.056	-0.216	0.554	3.596	0.803	0.000	0.933
octan-1-ol	-0.147	-0.214	0.561	3.507	0.749	0.000	0.943
decan-1-ol	-0.139	-0.090	0.356	3.547	0.727	0.000	0.958
propan-2-ol	-0.048	-0.324	0.713	4.036	1.055	0.000	0.884
isobutanol	-0.003	-0.357	0.699	3.595	1.247	0.000	0.881
butan-2-ol	-0.034	-0.387	0.719	3.736	1.088	0.000	0.905
<i>tert</i> -butanol	0.053	-0.443	0.699	4.026	0.882	0.000	0.907
3-methylbutan-1-ol	-0.052	-0.430	0.628	3.661	0.932	0.000	0.937
pentan-2-ol	-0.031	-0.325	0.496	3.792	1.024	0.000	0.934
2,2,2-trifluoroethanol	-0.092	-0.547	1.339	2.213	3.807	0.000	0.645
ethylene glycol	-0.887	0.132	1.657	4.457	2.355	0.000	0.565
diethyl ether	0.288	-0.347	0.775	2.985	0.000	0.000	0.973
dibutyl ether	0.165	-0.421	0.760	2.102	-0.664	0.000	1.002
diisopropylether	0.114	-0.032	0.685	3.108	0.000	0.000	0.941
methyl <i>tert</i> -butyl ether	0.278	-0.489	0.801	2.495	0.000	0.000	0.993
tetrahydrofuran	0.189	-0.347	1.238	3.289	0.000	0.000	0.982
1,4-dioxane	-0.034	-0.354	1.674	3.021	0.000	0.000	0.919
methyl acetate	0.134	-0.477	1.749	2.678	0.000	0.000	0.876
ethyl acetate	0.182	-0.352	1.316	2.891	0.000	0.000	0.916

Table 1. continued

solvent	<i>c</i>	<i>e</i>	<i>s</i>	<i>a</i>	<i>b</i>	<i>v</i>	<i>l</i>
log <i>K</i> , Equation 2 in <i>L</i>							
propyl acetate	0.165	−0.383	1.264	2.757	0.000	0.000	0.935
butyl acetate	0.147	−0.414	1.212	2.623	0.000	0.000	0.954
propanone	0.127	−0.387	1.733	3.060	0.000	0.000	0.866
butanone	0.112	−0.474	1.671	2.878	0.000	0.000	0.916
cyclohexanone	−0.086	−0.441	1.725	2.786	0.000	0.000	0.957
dimethylformamide	−0.391	−0.869	2.107	3.774	0.000	0.000	1.011
dimethylacetamide	−0.308	−0.736	1.802	4.361	0.000	0.000	1.028
<i>N</i> -methylpyrrolidinone	−0.128	−0.029	2.217	4.429	0.000	0.000	0.777
<i>N</i> -methyl-2-piperidone	−0.264	−0.171	2.086	5.056	0.000	0.000	0.883
<i>N</i> -formylmorpholine	−0.437	0.024	2.631	4.318	0.000	0.000	0.712
<i>N</i> -methylformamide	−0.249	−0.142	1.661	4.147	0.817	0.000	0.739
<i>N</i> -methylacetamide	−0.197	−0.175	1.608	4.867	0.375	0.000	0.837
formamide	−0.800	0.310	2.292	4.130	1.933	0.000	0.442
acetonitrile	−0.007	−0.595	2.461	2.085	0.418	0.000	0.738
nitromethane	−0.340	−0.297	2.689	2.193	0.514	0.000	0.728
dimethyl sulfoxide	−0.556	−0.223	2.903	5.037	0.000	0.000	0.719
tributylphosphate,	0.097	−0.098	1.103	2.411	0.588	0.000	0.844
gas–water	−1.271	0.822	2.743	3.904	4.814	0.000	−0.213

Table 2. Source and Purity for 2,4-Dihydroxybenzophenone and the Organic Solvents Used

chemical	source (manufacturer purity)	purity <sup>a</sup>	method <sup>b</sup>
2,4-dihydroxybenzophenone	Acros Organic (0.99)	0.996	GLC
hexane	Aldrich (0.99+)	0.998	GLC
heptane	Sigma-Aldrich (anhydrous, 0.99)	0.998	GLC
octane	Fluka (0.99)	0.999	GLC
cyclohexane	Sigma-Aldrich (anhydrous, 0.995)	0.999	GLC
methylcyclohexane	Aldrich (anhydrous, 0.99+)	0.998	GLC
isooctane	Aldrich (anhydrous, 0.998)	0.999	GLC
toluene	Aldrich (HPLC grade)	0.999	GLC
ethylbenzene	Aldrich (anhydrous, 0.998)	0.999	GLC
<i>o</i> -xylene	Aldrich (HPLC)	0.998	GLC
<i>m</i> -xylene	Aldrich (anhydrous, 0.99+)	0.998	GLC
<i>p</i> -xylene	Aldrich (HPLC)	0.998	GLC
chlorobenzene	Aldrich (HPLC)	0.998	GLC
butan-1-ol	Sigma-Aldrich (anhydrous, 0.998)	0.999	GLC
pentan-1-ol	Sigma-Aldrich (0.99+)	0.998	GLC
hexan-1-ol	Sigma-Aldrich (anhydrous, 0.99)	0.998	GLC
heptan-1-ol	Avocado Chemicals (0.99)	0.998	GLC
octan-1-ol	Sigma-Aldrich (anhydrous, 0.99+)	0.999	GLC
butan-2-ol	Aldrich (anhydrous, 0.995)	0.999	GLC
diisopropyl ether	Sigma-Aldrich (anhydrous, 0.99)	0.999	GLC
2-ethoxyethanol	Aldrich (0.99)	0.998	GLC
2-isopropoxyethanol	Aldrich (0.99)	0.998	GLC

<sup>a</sup>Mass fraction purity after recrystallization (solute) or distillation (solvents). <sup>b</sup>Analysis method: GLC is gas–liquid chromatography.

4, by taking  $K_w$  as another unknown and converting all of the values of log  $P$  into the corresponding values of log  $K$ . Some simplification is possible because  $E$  can be obtained from known structural fragments of the solute<sup>11</sup> and  $V$  can simply be calculated from structure.<sup>11,12</sup>

## 2. EXPERIMENTAL SECTION

**Materials.** 2,4-Dihydroxybenzophenone was purchased from a commercial source and was recrystallized twice from anhydrous methanol before use. All organic solvents were dried over molecular sieves and distilled prior to use. The sources, mass fraction purities, and analysis methods of the chemicals are given in Table 2.

**Solubility Measurements.** Solubilities were determined by spectrophotometric analysis. An excess amount of 2,4-dihydroxybenzophenone and the solvent were placed in amber glass bottles and allowed to equilibrate in a constant temperature water bath at  $298.2 \pm 0.1$  K for at least 3 days (often longer) with periodic agitation. After equilibration, the samples stood unagitated for several hours in the constant temperature bath to allow any finely dispersed solid particles to settle to the bottom of the container. Attainment of equilibrium was verified both by repetitive measurements performed 2 days later and by approaching equilibrium from supersaturation by pre-equilibrating the samples at a slightly higher temperature for several hours. Aliquots of saturated solutions were transferred into a

tared volumetric flask and then weighed to determine the amount of sample. After weighing the solution was diluted quantitatively with propan-2-ol. Concentrations of the dilute solutions were determined by spectrophotometric analysis at 324 nm (Milton Roy Spectronic 1001 Plus) using a Beer–Lambert law absorbance versus concentration working curve based on the measured absorbances of nine standard solutions having concentrations between  $5.73 \times 10^{-5}$  M and  $1.91 \times 10^{-4}$  M. The calculated molar absorptivity,  $\epsilon = 10,140$  L/(mol cm), was constant over the stated concentration range.

Experimental molar concentrations were converted to (mass/mass) solubility fractions by multiplying by the molar mass of 2,4-dihydroxybenzophenone, volume(s) of volumetric flask(s) used and any dilutions required to place the measured absorbances on the Beer–Lambert law absorbance versus concentration working curve, and then dividing by the mass of the saturated sample analyzed. Mole fraction solubilities were computed from solubility mass fractions using the molar masses of the solute and solvent. Experimental 2,4-dihydroxybenzophenone solubilities,  $x_s$ , in the 21 organic solvents studied are reported in Table 3. Numerical values represent the average of between four and ten independent determinations and were reproducible to within  $\pm 1.5$  % (relative percent error).

**Table 3. Experimental Mole Fraction Solubilities of 2,4-Dihydroxybenzophenone,  $x_s$ , in Select Organic Solvents at 298.2 K<sup>a</sup> and  $p = 0.1$  MPa**

organic solvent	$x_s^b$
hexane	0.0000326
heptane	0.0000382
octane	0.0000443
cyclohexane	0.0000507
methylcyclohexane	0.0000692
isooctane	0.0000331
toluene	0.00364
ethylbenzene	0.00408
<i>o</i> -xylene	0.00429
<i>m</i> -xylene	0.00415
<i>p</i> -xylene	0.00448
chlorobenzene	0.00440
butan-1-ol	0.0109
pentan-1-ol	0.0125
hexan-1-ol	0.0161
heptan-1-ol	0.0180
octan-1-ol	0.0183
butan-2-ol	0.0103
diisopropyl ether	0.00744
2-ethoxyethanol	0.0308
2-isopropoxyethanol	0.0333

<sup>a</sup>Standard uncertainty in temperature is  $u(T) = 0.1$  K. <sup>b</sup>Relative standard uncertainty in the mole fraction solubility is  $u_r(x) = 0.015$ .

Calculation of the solute descriptors requires that experimental mole fraction solubilities be converted to molar solubilities. This is accomplished by dividing  $x_s$  by the ideal molar volume of the saturated solution (i.e.,  $C \approx x_s/[x_s V_{\text{solute}} + (1 - x_s)V_{\text{solvent}}]$ ). A value of  $V_{\text{solute}} = 164.2$  cm<sup>3</sup> mol<sup>-1</sup> was used for the molar volume of the hypothetical subcooled liquid 2,4-dihydroxybenzophenone. Any errors resulting from our estimation of the 2,4-dihydroxybenzophenone's hypothetical subcooled liquid molar volume,  $V_{\text{solute}}$ , or the ideal molar volume approximation should have negligible effect of the

calculated  $C$  values as 2,4-dihydroxybenzophenone exhibits a very low solubility in many of the organic solvents considered in the present investigation. Solubilities of 2,4-dihydroxybenzophenone at 298 K, in mol dm<sup>-3</sup> solution (molarity scale), are in Table 4. The value for the aqueous solubility is that given by Lu and Lu.<sup>13</sup>

**Table 4. Solubilities (mol dm<sup>-3</sup>) of 2,4-Dihydroxybenzophenone at 298 K**

solvent	$C^a$	log $C$	log $C$ (calc)	log $C$ (pred)
water <sup>b</sup>		-3.370		
hexane	0.000248	-3.606	-3.617	
heptane	0.000259	-3.587	-3.578	
octane	0.000271	-3.567	-3.562	
cyclohexane	0.000466	-3.332	-3.229	
methylcyclohexane	0.000539	-3.268	-3.286	
isooctane	0.000199	-3.701	-3.765	
toluene	0.03402	-1.468	-1.371	
<i>o</i> -xylene	0.03531	-1.452	-1.543	
<i>m</i> -xylene	0.03358	-1.474	-1.505	
<i>p</i> -xylene	0.03604	-1.443	-1.636	
ethylbenzene	0.03314	-1.480	-1.584	
chlorobenzene	0.04298	-1.367	-1.179	
butan-1-ol	0.1174	-0.930	-0.844	
pentan-1-ol	0.1141	-0.943	-0.907	
hexan-1-ol	0.1277	-0.894	-0.923	
heptan-1-ol	0.1263	-0.899	-0.943	
octan-1-ol	0.1156	-0.937	-0.912	
butan-2-ol	0.1104	-0.957	-0.852	
diisopropyl ether	0.05211	-1.283	-1.396	
tetrachloromethane				-2.199
diethyl ether				-1.788
propanone				-0.088
ethyl acetate				-0.452
nitrobenzene				-1.171
dimethylformamide				0.860
dimethyl sulfoxide				1.418

<sup>a</sup>From Table 3 after conversion to the molar scale. <sup>b</sup>From ref 13.

### 3. RESULTS AND DISCUSSION

The solubilities of 2,4-dihydroxybenzophenone in Table 4 yield 19 values of the corresponding water–solvent partition coefficients,  $P$ , through eq 3. Then if the value of log  $K_w$  is allowed to float, a further set of 19 gas–solvent partition coefficients can be obtained through eq 4. There are also two extra equations for the gas–water partition, see Table 1, to yield a total of 40 simultaneous equations with coefficients as shown in Table 1. The value of  $E$  was taken<sup>11</sup> as 1.73, and  $V$  was calculated<sup>12</sup> as 1.5982; this leaves  $S$ ,  $A$ ,  $B$ ,  $L$ , and log  $K_w$  as unknowns to be determined from the set of simultaneous equations by a “best-fit” method. With the descriptors given in Table 5, the 40 simultaneous equations were solved with a standard deviation of 0.095 log units. The calculated values of  $P$  can then be converted into calculated solubilities through eq 3, and the latter are given in Table 4. For solubilities in the 19 organic solvents, the average error between the experimental and the calculated values, AE, is 0.002, the absolute average error, AAE, is 0.071, and the standard deviation, SD, is 0.092 log units.

Table 5. Obtained Descriptors

compd	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>L</i>	log $K_w$
2,4-dihydroxybenzophenone	1.73	2.03	0.49	0.70	1.5982	9.062	9.06
biotin	1.35	2.49	0.93	1.31	1.7536	9.563	14.54
caprolactam	0.59	1.72	0.38	0.75	0.9609	4.914	7.98

The descriptors given in Table 5 make general chemical sense. The hydrogen-bond acidity, *A*, seems very low, but the 2-hydroxy group forms a strong intramolecular hydrogen bond with the carbonyl oxygen, so the overall hydrogen bond acidity is entirely due to the 4-hydroxyl group.

Lu and Lu<sup>13</sup> determined the solubility of 2,4-dihydroxybenzophenone not only in water but also in water–ethanol mixtures. We have determined equation coefficients for water–ethanol mixtures across the whole composition range with compositions in volume percent of ethanol.<sup>14</sup> The solubility determinations of Lu and Lu were all at low ethanol concentrations, and the various log *P* values contribute very little with regard to the set of simultaneous equations. We therefore did not use the solubilities in water–ethanol mixtures in the deduction of descriptors for 2,4-dihydroxybenzophenone, but instead we used the obtained descriptors to predict log *P* for water–ethanol mixtures and then to predict log *C* in the mixtures. These are shown in Figure 1 as a function of volume percent of ethanol.

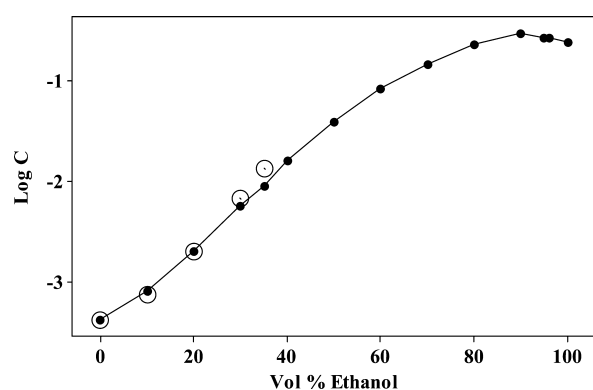


Figure 1. Solubility of 2,4-dihydroxybenzophenone in water–ethanol mixtures: ●, predicted solubility; ○, observed solubility from Lu and Lu.<sup>13</sup>

Biotin, vitamin H, is an important compound. It has a rather unusual structure, and we use it as another illustration of determination of properties, this time from a rather restricted number of solubilities. The procedure is exactly the same as that described for 2,4-dihydroxybenzophenone. Su et al.<sup>1</sup> have determined solubilities of biotin in water and five solvents as shown in Table 6. We have no equation for acetic acid and so are restricted to only four partition coefficients, as found using eq 3. However, we can convert all of the log *P* values to values of log *K* through eq 4; we have two extra equations in log  $K_w$ , see Table 1, and so have a total of ten simultaneous equations. We take *E* as 1.35<sup>11</sup> and *V* as 1.7536,<sup>12</sup> and have easily enough equations to determine the unknowns *S*, *A*, *B*, *L*, and log  $K_w$ , as given in Table 5.

In the case of 2,4-dihydroxybenzophenone and biotin, it was possible to convert all of the solubilities in organic solvents into corresponding partition coefficients through a known value of the solubility in water. There are cases where solubilities of a

Table 6. Solubilities of Biotin, as log *C*, at 298 K

solvent	obs	calc	pred
water	−3.67		
methanol	−2.77	−2.62	
ethanol	−2.97	−3.03	
dimethylformamide	−1.34	−1.37	
dimethyl sulfoxide	−0.31	−0.33	
acetic acid	−1.41		
hexane			−8.81
tetrachloromethane			−6.88
toluene			−5.65
diethyl ether			−4.29
propanone			−3.12
ethyl acetate			−3.75
nitrobenzene			−3.62
dimethylacetamide			−1.02

given compound have been determined in organic solvents, but not in water, so it appears that the partition coefficients required for our analysis cannot be deduced through eq 3. However, it is possible to circumvent this difficulty by using the water solubility, as log  $C_w$ , as another unknown to be determined by the usual trial-and-error method. We use caprolactam as an example of a compound for which only solubilities in organic solvents are available. Guo et al.<sup>15</sup> determined solubilities in five organic solvents. Values of log *C* at 298 K are given in Table 7. With log  $C_w$  as an unknown to be

Table 7. Solubilities of Caprolactam, as log *C*, at 298 K

solvent	obs	calc	pred
water		0.78	
propan-1-ol	0.49	0.48	
propan-2-ol	0.44	0.43	
butan-1-ol	0.39	0.38	
diisopropyl ether	−0.34	−0.32	
methyl <i>tert</i> -butyl ether	0.01	0.01	
hexane			−2.28
tetrachloromethane			−1.07
toluene			−0.31
diethyl ether			0.20
propanone			1.06
ethyl acetate			0.60
nitrobenzene			0.84
dimethylformamide			1.83
dimethyl sulfoxide			2.44

determined, these solubilities yield five water–solvent partition coefficients, through eq 3. Then if log  $K_w$  is taken as another unknown, five gas–solvent partition coefficients are available. In addition, there are two equations, see Table 1, for gas–solvent partitions, making up 12 equations in total. We take *E* as 0.59<sup>11</sup> and *V* as 0.9609,<sup>12</sup> leaving six unknowns to be determined, viz., *S*, *A*, *B*, *L*, log  $C_w$ , and log  $K_w$ . There are easily

enough equations to obtain these unknowns by the method of simultaneous equations. We find with  $\log C_w$  as 0.78, and with the descriptors given in Table 5, calculated solubilities are very close to those observed, see Table 7.

Once descriptors are known for a given compound, they can be combined with coefficients in eq 1 and eq 2 to predict partition coefficients from water and the gas phase to any of the solvents listed in Table 1. Then from either an experimental value of  $\log C_w$ , or from a calculated value, the corresponding solubilities in the solvents can be obtained, trivially. Some predicted solubilities obtained in this way are given in Tables 4, 6, and 7 and illustrate the use of the present method in the selection of solvents for extraction and purification. Partition coefficients from water and the gas phase to wet solvents, and corresponding solubilities in wet solvents can similarly be predicted using the required equation coefficients that we have listed.<sup>3–8</sup> Equation coefficients are available for partition from water and the gas phase into ionic liquids,<sup>16,17</sup> and solubilities in these ionic liquids can then easily be calculated. Coefficients are also available for equations related to various biological and environmental processes,<sup>18–20</sup> and it is a matter of simple arithmetic to obtain values for these processes. All of this information can be obtained from the determination of solubilities in organic solvents and in water. If a water solubility is not available, it can be deduced, as part of the general method, as we have shown for caprolactam. It should be pointed out that the coefficients for partition coefficients in Table 1 refer specifically to 298 K, and hence our predictions of these partition coefficients and the corresponding solubilities also refer to 298 K. However, the obtained descriptors can be used in equations for processes at any temperature. For example, the descriptors in Table 5 can be used to predict solubility in water from 273 K to 573 K.<sup>21,22</sup>

## AUTHOR INFORMATION

### Corresponding Author

\*E-Mail: m.h.abraham@ucl.ac.uk

### Notes

The authors declare no competing financial interest.

## REFERENCES

- (1) Su, J.; Quin, C.; Luo, N.; Xiang, X.; Xu, Y.; Chen, X. Experimental measurements and modelling of the solubility of biotin in six pure solvents at temperatures from 298.15 to 333.85 K. *J. Chem. Eng. Data* **2014**, *59*, 3894–3899.
- (2) Guo, C.; Li, L.; Cheng, J.; Zhang, J.; Li, W. Solubility of caprolactam in different organic solvents. *Fluid Phase Equilib.* **2012**, *319*, 9–15.
- (3) Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. Correlation of the solubility behavior of crystalline 1-nitronaphthalene in organic solvents with the Abraham solvation parameter model. *J. Solution Chem.* **2005**, *34*, 1121–1133.
- (4) Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R.; Acree, W. E., Jr. Prediction of solubility of drugs and other compounds in organic solvents. *J. Pharm. Sci.* **2010**, *99*, 1500–1515.
- (5) Stephens, T. W.; Loera, M.; Calderas, M.; Diaz, R.; Montney, N.; Acree, W. E., Jr.; Abraham, M. H. Determination of Abraham model solute descriptors for benzoin based on measured solubility ratios. *Phys. Chem. Liq.* **2012**, *50*, 254–265.
- (6) Bowen, K. R.; Stephens, T. W.; Lu, J.; Satish, K.; Shan, D.; Acree, W. E., Jr.; Abraham, M. H. Experimental and predicted solubilities of 3,4-dimethoxybenzoic acid in select organic solvents of varying polarity and hydrogen-bonding character. *Eur. Chem. Bull.* **2013**, *2*, 577–583.

(7) Abraham, M. H.; Acree, W. E., Jr. Descriptors for artemisinin and its derivatives; estimation of physicochemical and biochemical data. *Eur. Chem. Bull.* **2013**, *2*, 1027–1037.

(8) Abraham, M. H.; Acree, W. E., Jr. Descriptors for the prediction of partition coefficients for 8-hydroxyquinoline and its derivatives. *Sep. Sci. Technol.* **2014**, *49*, 2135–2141.

(9) Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. The determination of sets of solute descriptors from chromatographic measurements. *J. Chromatogr. A* **2004**, *1037*, 29–47.

(10) Clarke, E. D.; Mallon, L. The Determination of Abraham Descriptors and Their Application to Crop Protection Research. In *Modern Methods in Crop Protection Research*; Jeschke, P., Krämer, W., Schirmer, U., Witschel, M., Eds.; Wiley-VCH: Berlin, 2012.

(11) *Absolv*, version 5.0; Advanced Chemistry Development: Toronto, Ontario, Canada.

(12) Abraham, M. H.; McGowan, J. C. The use of characteristic volumes to measure cavity terms in reversed-phase liquid chromatography. *Chromatographia* **1987**, *23*, 243–246.

(13) Lu, L.-L.; Lu, X.-Y. Solubilities of polyhydroxybenzophenones in an ethanol + water mixture from (293.15 to 343.15) K. *J. Chem. Eng. Data* **2008**, *53*, 1996–1998.

(14) Abraham, M. H.; Acree, W. E., Jr. Partition coefficients and solubilities of compounds in the water-ethanol solvent system. *J. Solution Chem.* **2011**, *40*, 1279–1290.

(15) Guo, C.; Li, L.; Cheng, J.; Zhang, J.; Li, W. Solubility of caprolactam in different organic solvents. *Fluid Phase Equilib.* **2012**, *319*, 9–15.

(16) Stephens, T. W.; Chou, V.; Quay, A. N.; Shen, C.; Dabadge, N.; Tian, A.; Loera, M.; Willis, B.; Wilson, A.; Acree, W. E., Jr.; Twu, P.; Anderson, J. L.; Abraham, M. H. Thermochemical investigations of solute transfer into ionic liquid solvents: Updated Abraham model equation coefficients for solute activity coefficient and partition coefficient predictions. *Phys. Chem. Liq.* **2014**, *52*, 488–518.

(17) Stephens, T. W.; Hart, E.; Kuprasertkul, N.; Mehta, S.; Wadawadigi, A.; Acree, W. E., Jr.; Abraham, M. H. Abraham model correlations for describing solute transfer into ionic liquid solvents: Calculation of ion-specific equation coefficients for the 4,5-dicyano-2-(trifluoromethyl)imidazolide anion. *Phys. Chem. Liq.* **2014**, *52*, 777–791.

(18) Abraham, M. H.; Gola, J. M. R.; Ibrahim, A.; Acree, W. E., Jr.; Liu, X. X. The prediction of blood-tissue partitions, water-skin partitions and skin permeation for agrochemicals. *Pest. Manage. Sci.* **2014**, *70*, 1130–1137.

(19) Abraham, M. H.; Gola, J. M. R.; Ibrahim, A.; Acree, W. E., Jr.; Liu, X. A simple method for estimating *in vitro* air-tissue and *in vivo* blood-tissue partition coefficients. *Chemosphere* **2015**, *120*, 188–191.

(20) Endo, S.; Goss, K.-U. Applications of polyparameter linear free energy relationships in environmental chemistry. *Environ. Sci. Technol.* **2014**, *48*, 12477–12491.

(21) Abraham, M. H.; Acree, W. E., Jr. Prediction of gas to water partition coefficients from 273 to 373 K using predicted enthalpies and heat capacities of hydration. *Fluid Phase Equilib.* **2007**, *262*, 97–110.

(22) Abraham, M. H.; Acree, W. E., Jr. The hydrogen bond properties of water from 273 to 573 K; equations for the prediction of gas-water partition coefficients. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7433–7440.