Fluid Phase Equilibria 387 (2015) 1–4

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

Analysis of the solubility of betaine: calculation of descriptors and physicochemical properties

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

[a] Department of Chemistry, University College London, 20 Gordon Street, London WC1H OAJ, UK
[b] Department of Chemistry, P.O. Box 305070, University of North Texas, Denton, TX 76203-5070, USA

ARTICLE INFO

Article history:
Received 13 September 2014
Received in revised form 18 November 2014
Accepted 19 November 2014
Available online 2 December 2014

Keywords:
Solubility
Partition coefficients
Abraham descriptors
Linear free energy relationship

ABSTRACT

We have previously constructed equations for the estimation of water–solvent partition coefficients of neutral molecules and of ions, in terms of properties or descriptors of the solute. We convert known solubilities of betaine in water and solvents into water–solvent partition coefficients and use the previous equations to determine descriptors for betaine. We show that equations for neutral compounds do not yield realistic results, but equations that include ionic terms allow descriptors for betaine to be determined. These indicate that betaine, although a very 'polar' molecule is not at all as polar as ion-pairs. From the determined descriptors, solubilities in further solvents and further water–solvent partition coefficients may be predicted.

1. Introduction

There are now numerous methods for the estimation of physicochemical properties, especially water–solvent partition coefficients, *Ps*. These are defined in terms of concentrations of solute as mol dm^{-3} of solution in each phase, that is on the molarity scale. The temperature is normally 298 K, and all measurements we report refer to 298 K. The BioLoom ClogP program [1] calculates water–octanol partition coefficients, as log *Poct*, and these coefficients can also be calculated by a number of well-known programs, the Advanced Chemistry Development ACD program [2], the Environmental EPI program [3] the SPARC method [4], and the ALOGPS program. The ACD and EPI programs are available in the freeware ChemSpider [5], and the ALOGPS program is also free on the web [6]. In addition the EPI software calculates water solubility and Henry's Law constants in water, and the ALOGPS program calculates water solubility. The SPARC method is very general, and water–solvent partition coefficients, solubilities and Henry's Law constants can be calculated for a very wide range of solvents. All these calculations refer to neutral species. To date there are no reliable calculations for ions, and calculations for zwitterions and for compounds such as betaine, $Me_3N^+CH_2CO_2^-$, differ widely, as can be seen from Table 1. SPARC cannot calculate properties of compounds such as betaine. The ALOGPS program [6] gives calculated values of log *Poct* for betaine from eight software packages, with an average of −2.39 ± 1.75, slightly different from the calculated value by ALOGPS itself of −2.70.

Betaine is an important molecule. It occurs naturally, but is also manufactured synthetically for use in the cosmetic and pharmaceutical industries. It is therefore desirable that some reliable method of estimating its physicochemical properties should be put in place. If such a method was realised, then it would be an important step to the estimation of properties of analogs of betaine. Wang et al. [7,8] have determined the solubility of betaine in a number of solvents and we have converted their mol fraction solubilities into solubilities on the molarity scale in mol dm $^{-3}$ at 298 K. The solubility in water has been reported by Huuskonen [9] and listed by Wang et al. [10] as log *Cw*/1 mol dm $^{-3}$) = 0.72. The interest in these solubilities is that we have shown from a series of solubilities in a number of solvents that it is possible to extract the Abraham descriptors of a molecule [11,12]. Once these are available [13,14] then a series of physicochemical properties can be deduced. All this work [10–15] was carried out on neutral molecules. Subsequently, we showed that it was possible to extend the method to ions, both permanent ions such as Na^+ and Br^- and ionic species such as protonated base cations and carboxylate anions, provided that an additional descriptor J^+ was used for cations and an additional descriptor J^- was used for anions [15–19].

Abbreviations: *Cs*, solubility in a solvent; *Cw*, solubility in water; *Ps*, water–solvent partition coefficient; *Poct*, the water–octanol partition coefficient; *E*, the solute excess molar refractivity in units of (cm^3 mol^{-1})/10; *S*, the solute dipolarity/polarizability; *A*, overall or summation hydrogen bond acidity; *B*, overall or summation hydrogen bond basicity; *V*, the McGowan volume in units of (cm^3 mol^{-1})/100; J^+, a descriptor for cations; J^-, a descriptor for cations; SP, solute property.

* Corresponding author. Tel.: +44 2076794639; fax: +44 2076797463.
E-mail address: m.h.abraham@ucl.ac.uk (M.H. Abraham).

http://dx.doi.org/10.1016/j.fluid.2014.11.019

Table 1
Calculated **v**alues of the water–octanol partition coefficient of betaine, with concentrations on the molarity scale, as log *Poct*, at 298 K.

Method	Log *Poct*
ClogP	−8.17 [1]
ACD	−3.25 [2]
EPI	−4.93 [3,5]
SPARC	N/A [4]
ALOGPS	−2.70 [6]
Av of eight	−2.39 [6]

Interestingly, both descriptors were needed to account for properties of ion pairs, which, like betaine, are electrically neutral but which have a total internal charge separation [20].

It seems therefore of some considerable interest to see if our original solubility method for neutral species could be applied to betaine, or if it was necessary to use the more complicated method that we set up for ions [15–19] and ion-pairs [20].

2. Methods

Our method for the analysis of properties of neutral molecules is based on the linear free energy equations, LFERs, Eqs. (1) and (2).

$$\mathrm{SP} = c + eE + sS + aA + bB + vV \tag{1}$$

$$\mathrm{SP} = c + eE + sS + aA + bB + lL \tag{2}$$

In these equations, SP is some solute property, such as log *Poct* for a series of solutes. Eq. (1) is used for processes taking place in solution, and Eq. (2) for gas to solution processes. The independent variables in Eq. (1) are E the solute excess molar refractivity in units of (cm^3 mol^{-1})/10; S the solute dipolarity/polarizability; A and B the overall or summation hydrogen bond acidity and basicity; and V the McGowan volume in units of (cm^3 mol^{-1})/100. For pure liquid solutes, E is obtained from the refractive index of the compound at 293 K or for gases and solids the refractive index can be estimated or E itself can be calculated quite easily [21]. For neutral molecules, the descriptors S,A and B were obtained from various chromatographic measurements and water to solvent partition measurements, as described before [13–17]. The coefficients c, e, s, a, b, v and l are the equation coefficients obtained by multiple linear regression analysis of the dependent variable, SP, against the independent variables in Eqs. (1) and (2).

In the case of ions, we found that Eq. (1) yielded only poor results when applied to water–solvent partitions, and that it was necessary to add an additional term j^+J^+ for cations and an additional term j^-J^- for anions, leading to Eq. (3). It is essential to note that Eq. (3) applies also to neutral molecules for which the solute parameters $J^+ = J^- = 0$, and that the coefficients c, e, s, a, b, and v in Eq. (3) are the same as those in Eq. (1). Thus Eq. (3) with the coefficients $j^+ = j^- = 0$ is the same as Eq. (1) for neutral compounds.

$$\mathrm{SP} = c + eE + sS + aA + bB + vV + j^+J^+ + j^-J^- \tag{3}$$

There were one or two differences in the application of our LFERs to ions as compared to neutral molecules. E was obtained from ionic molar refractions [15], and V was calculated as described [21,22]. Values of descriptors for some neutral molecules, ions and ion-pairs that might be relevant to betaine are shown in Table 2.

3. Results

In order to obtain descriptors for betaine, we use the mol fraction solubility data obtained by Wang et al. [7,8] after conversion to solubilities in mol dm $^{-3}$. The solubility of betaine in water as log *Cw* is 0.72 [9,10], and this can be used to convert solubilities in other solvents, *Cs*, into water–solvent partition coefficients, *Ps*, through Eq. (4).

Table 2
Descriptors for some neutral and ionic species.

Species	*E*	*S*	*A*	*B*	*V*	*J*$^+$	*J*$^-$
Trimethylamine	0.140	0.20	0.00	0.67	0.6311	0.000	0.000
Acetic acid	0.265	0.64	0.62	0.44	0.4648	0.000	0.000
Na^+	−0.020	2.31	1.22	0.00	0.0330	0.316	0.000
Br^-	0.170	2.74	0.00	1.82	0.3066	0.000	1.567
Me_3EtN^+	−0.100	1.37	0.65	0.00	0.7900	1.260	0.000
$MeCO_2{}^-$	0.415	2.19	0.00	2.93	0.4433	0.000	0.275
Me_4NBr	0.070	2.11	0.89	2.18	1.0701	−0.559	0.356
$MeCO_2$ Na	0.395	2.61	1.18	3.01	0.4763	−0.618	1.325

$$Ps = \frac{Cs}{Cw} \tag{4}$$

For most of the solvents studied by Wang et al., including water–ethanol mixtures, we had determined the coefficients in Eq. (3), and these are given in Table 3. The various equations refer to the partition between water and the dry solvent [16–20] or between water and a water–ethanol mixture [23,24]. Although such partitions are hypothetical, the corresponding equations are extremely useful in the analysis of solubilities [11]. Only in the case of solvent wet octanol do the coefficients refer to a practical partition. The % ethanol in Table 3 refers to water–ethanol mixtures with the composition in volume % ethanol.

The molar solubilities from Wang et al. [7,8], are shown in Table 4, and were converted to log *Ps* values through Eq. (4) with log *Cw* taken as 0.72 [9,10]. Then we have 19 equations in log *Ps* for which we have equation coefficients, and can find the unknown solute descriptors that lead to the smallest standard deviation by trial-and-error. We calculated $E = 0.315$, as the sum of values for the trimethylethylammonium ion and the acetate ion, see Table 2, and calculated V by our usual equation [13,25]. We first applied Eq. (1) for neutral solutes to the values of log *Ps* and obtained the descriptors shown in Table 5. Also given in Table 5 are values for the average error (AE), the absolute average error (AAE), and the standard deviation (SD) between the observed and calculated values of log *Ps*; these are the same for the observed and calculated values of log *Cs*. The obtained descriptors make little sense. In particular the value of $S = 0.18$ is far too low to be realistic. The statistics are very poor with SD no less than 0.442 log units. We then used 19 equations on the lines of Eq. (3). Even bearing in mind that there are two more adjustable constants, J^+ and J^- the statistics are now very much better with SD = 0.192 log units. The descriptors are also more realistic. $S = 1.57$ suggests that betaine is very dipolar, and $B = 2.00$ suggests that betaine is a strong hydrogen bond base, compare $B = 2.93$ for the acetate anion. The values of log P (obs), log P (calc), log C (obs) and log C (calc) using Eq. (3) are in Table 4.

Solubilities were also measured in pentan-1-ol and isopentanol [5]. We cannot include these in our analysis because we have no equations on the lines of Eq. (3) for partition from water into these two solvents. However, by comparison to the observed solubilities in the lower alcohols, the solubilities in pentan-1-ol and isopentanol seem to be very much lower than expected.

Our analysis indicates that calculations on properties of betaine using equations that apply to neutral molecules will yield incorrect results. We also found that betaine is much less 'polar' than typical ion-pairs such as Me_4NBr or $MeCO_2$ Na, see Table 2, so that using these as models could also lead to incorrect results. The descriptors using Eq. (3), and the coefficients in Table 3 lead to a value of