Partition Coefficients and Solubilities of Compounds in the Water–Ethanol Solvent System

Michael H. Abraham · William E. Acree Jr.

Received: 25 November 2010 / Accepted: 29 December 2010 / Published online: 30 June 2011
© Springer Science+Business Media, LLC 2011

Abstract A new method for the prediction of solubility in the water–ethanol system has been devised. The method is based on the construction of equations for the prediction of partition coefficients from water to water–ethanol mixtures, based on the Abraham descriptors for compounds. Once partition coefficients have been predicted for a particular compound, only an experimental value for the solubility in ethanol is required for the prediction of solubilities in water–ethanol mixtures. Results of predictions are comparable to the Jouyban-Acree model that requires experimental solubilities in both water and ethanol.

Keywords Water–ethanol · Solubility · Partition coefficients · Abraham descriptors · Linear free energy relationship · Multiple linear regression

1 Introduction

The solubility of compounds, especially drugs, in water–cosolvent systems is of ongoing interest, and several methods have been put forward for the correlation and prediction of solubilities. The most important cosolvents are ethanol, propylene glycol and glycerin, and in this paper we deal with the water–ethanol system. A number of equations have been put forward by Yalkowsky et al. [1–5], the most recent being a bilinear model. When applied specifically to water–ethanol mixtures, the model was formulated as:

$$\log_{10} \left( \frac{S_{\text{mix}}}{S_w} \right) = \sigma_A f + (\sigma_B - \sigma_A) f/(1 + 10^{-3.6(f-1)})$$

(1)
In Eq. 1, \( S_{\text{mix}} \) and \( S_w \) are the solubility of a given compound in a water–ethanol mixture and in pure water, \( f \) is the volume fraction of ethanol in the mixture, \( \sigma_A \) is the initial slope of a plot of \( S \) against \( f \) as \( f \to 0 \), and \( \sigma_B \) is the final slope of the plot of \( S \) against \( f \) as \( f \to 1 \). Thus, in order to apply Eq. 1 to a given compound, the solubility in water, ethanol, and a number of water–ethanol mixtures must be known. For 53 compounds where there were between 5 and 14 solubility determinations of each compound in water–ethanol mixtures, Eq. 1 fitted the experimental solubilities with a root mean square error (RMSE) of 0.067 log units. However, no actual predictions of solubilities were given. The Wilson method as previously modified [6] has been used to correlate solubilities of salicylic acid in water–ethanol mixtures [7]. The method uses two adjustable constants but requires knowledge of the molar enthalpy of fusion and melting point of the solute. These quantities are also used in a solubility parameter model [8] that uses three adjustable constants. The NRTL method again requires knowledge of the molar enthalpy of fusion and melting point of the solute, but now uses six adjustable constants to correlate the solubility data [7].

A different approach combines a model due to Acree [9] with Abraham solvation parameters [10, 11]. Unlike previous models, the Jouyban-Acree model [12] was trained on experimental solubilities in the water–ethanol system to yield Eq. 2:

\[
\log_{10} S_{m,T} = f_1 \log_{10} S_{1,T} + f_2 \log_{10} S_{2,T} + (f_1 f_2 / T)(558.45 + 358.60E + 22.01S \\
- 352.97A + 130.48B - 297.10V) + (f_1 f_2 [f_1 - f_2] / T)(45.67 - 165.77E \\
- 321.55S + 479.48A - 409.51B + 827.63V) \\
+ (f_1 f_2 [f_1 - f_2]^2 / T)(-493.81 - 341.32E + 866.22S - 36.17A + 173.41B \\
- 555.48V)
\]  

(2)

In Eq. 2, \( S_m \), \( S_1 \), and \( S_2 \) are the solubilities of the solute in a water–ethanol mixture of composition \((f_1, f_2)\), water and ethanol, respectively, at a temperature \( T \), \( f_1 \) is the volume fraction of water in the mixed solvent, \( f_2 \) is the volume fraction of ethanol in the mixed solvent, and \( E, S, A, B, \) and \( V \) are Abraham solute descriptors as set out below. Since Eq. 2 is already trained on experimental data, there are no adjustable parameters, and the only input solubilities are those in water and ethanol. Predicted values could be used for \( \log_{10} S_{1,T} \) and \( \log_{10} S_{2,T} \) in Eq. 2; however, this would introduce additional error and result in poorer predicted solubilities for the mixed solvents [13].

The Abraham parameters are descriptors of solute molecules as follows [10, 11]. \( E \) is the solute excess molar refractivity in units of \((\text{cm}^3 \cdot \text{mol}^{-1})/10\), \( S \) is the solute dipolarity/polarizability, \( A \) and \( B \) are the overall or summation hydrogen bond acidity and basicity, and \( V \) is the McGowan characteristic volume in units of \((\text{cm}^3 \cdot \text{mol}^{-1})/100\). They have mostly been used to correlate and to predict water–solvent partition coefficients, \( P \), through Eq. 3, where \( P \) is a set of partition coefficients for a series of solutes in the same water–solvent system. Note that Eq. 3 has been applied to ‘practical’ partition coefficients, for example those for partition between water and wet octanol, and to ‘hypothetical’ partition coefficients in systems such as water–dry solvents [14]. The coefficients in Eq. 3 are obtained by multiple linear regression, and serve to characterize the system under consideration. The determination of the independent variables, or descriptors, in Eq. 3 has been set out previously [10, 11], and in addition it has been shown that descriptors can be obtained just from solubilities in organic solvents [15, 16].

\[
\log_{10} P = c + eE + sS + aA + bB + vV
\]  

(3)