Comments on “Solvation Parameters. 2. A Simplified Molecular Topology To Generate Easily Optimized Values”

Christina Mintz,† William E. Acree, Jr.,* † and Michael H. Abraham‡

Department of Chemistry, P.O. Box 305070, University of North Texas, Denton, Texas 76203-5070, and Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

Received June 20, 2006


In a recent paper appearing in this Journal Laffort and Héricourt† presented a generalized method to establish the numerical values of the solvation parameters of solutes. The solvation parameters, when combined with the five solvent descriptors, encode meaningful chemical information. For example, the excess molar refraction, E, is derived from the solute refractive index, and hence the e-coefficient gives a measure of general solute−solvent dispersion interactions. The V and L descriptors were set up as measures of the endoergic effect of disrupting solvent−solute bonds. However, solute volume is always well correlated with polarizability, and so the v- and l-coefficients will include not only an endoergic cavity effect but also exoergic solute−solvent effects that arise through solute polarizability. The S descriptor is a measure of dipolarity and polarizability, and hence the s-coefficient will reflect the ability of a solvent to undergo dipole−dipole and dipole−induced dipole interactions with a solute. The A descriptor is a measure of solute hydrogen bond acidity, and hence the a-coefficient will reflect the complimentary solvent hydrogen bond basicity. Similarly the b-coefficient will be a measure of the solvent hydrogen bond acidity. All this is
straightforward for gas-to-solvent partitions, because there are no interactions to consider in the gas phase. For partitions between two condensed phases, the coefficients in eq 7 then refer to differences between the properties of the two phases.

Such information is lost in transforming the Abraham solute descriptors to the solvation parameters suggested in the Laffort and Héricourt paper.\(^1\) To illustrate this point, we write the transformation equations as follows:

\[
\begin{align*}
A & = 0.3540\alpha_2 \\
B & = 1.3736\beta_2 \\
E & = \epsilon_2 \\
S & = 0.6566\alpha_2 + 0.4852\beta_2 + 0.5496\epsilon_2 \\
L & = \delta_2 + 0.3493\alpha_2 + 0.2581\beta_2 + 1.1864\epsilon_2 + 0.115 \\
& (10)
\end{align*}
\]

for the solute’s gas-to-condensed phase transfer property. Similarly, an expression could be developed for converting the equation coefficients for correlations deduced using the Laffort et al. descriptors to equation coefficients based on the Abraham descriptors.

In Table 1 we have converted our existing Abraham equation coefficients\(^5\)\(^-\)\(^7\) for 43 organic solvents to values that would be obtained if one were to use the transformation equations suggested by Laffort and Héricourt.\(^1\) Careful examination of the tabulated numerical entries reveals that all of the alkane solvents have a nonzero \(\beta_2\)-coefficient. The \(\beta_2\) solvation parameter coefficient should provide a meaningful measure of the solvent’s hydrogen-bond acidity. Alkane solvents are incapable of hydrogen-bond formation, and the \(\beta_2\)-coefficient should be zero. As can be seen, the transformed \(\beta_2\)-coefficient provides incorrect information regarding the ability of saturated alkane solvents to function as hydrogen-bond acids.

A more dramatic example of how the suggested transformation leads to incorrect chemical information is found by examining the correlation equations for dry diethyl ether, dry tetrahydrofuran, and dry 1,4-dioxane. The equations obtained by application of eq 6 are as follows, where \(K\) is the gas to solvent partition coefficient, equivalent to the Ostwald solubility coefficient.

\[
\begin{align*}
\log K_{\text{diethyl ether}} & = 0.288 - 0.347E + 0.775S + 2.985A + 0.000B + 0.973L \\
\log K_{\text{THF}} & = 0.189 - 0.347E + 1.238S + 3.289A + 0.000B + 0.982L
\end{align*}
\]

Log \(K_{\text{dioxane}} = -0.034 - 0.347E + 1.674S + 3.021A + 0.000B + 0.919L\) (16)

Eqs 14–16 all make chemical sense. In particular, the \(a\)-coefficients are all quite large, thus reflecting the hydrogen bond basicity of the ethers, and the \(b\)-coefficients are all zero, as expected for solvents that all have zero hydrogen bond acidity.

If we now use the Laffort and Héricourt\(^1\) transformations, the equations are restated as

\[
\begin{align*}
\log K_{\text{diethyl ether}} & = 0.400 + 1.233E + 0.8408S + 1.057A + 0.627B + 0.973\delta_2 \\
\log K_{\text{THF}} & = 0.302 + 1.498E + 1.156O_2 + 1.164A + 0.854B + 0.982\delta_2 \\
\log K_{\text{dioxane}} & = 0.072 + 1.663E + 1.420S + 1.069A + 1.049B + 0.919\delta_2
\end{align*}
\]

The calculated \(\beta_2\)-coefficients are now quite large and positive, so that the transformed equations indicate, quite incorrectly, that ethers are hydrogen bond acids.
There are advantages associated with having an orthogonal set of solute descriptors. However, we note that the transformed descriptors of Laffort and Héricourt\(^1\) will be orthogonal only for their original data set of 369 compounds. For any other data set, there will be loss of orthogonality. For example, in the data set used in deriving eq 19, the largest value of \(R^2\) between descriptors is 0.682 for \(E\) and \(L\) in our descriptors, and \(R^2\) is 0.421 between \(\epsilon\) and \(\delta\) in the Laffort descriptors. The dioxane database contains 105 compounds. For compounds not listed in Table C1 of the Laffort et al.\(^2\) paper, we calculated the Laffort descriptors using the suggested transformation equations (eqs 1–5). PCA on any set of solute descriptors will always yield orthogonal PCs that can be used in linear regressions but has the disadvantage that the equation coefficients cannot be interpreted easily. The procedure of Abraham has the disadvantage that the solute descriptors may be too cross-correlated to yield a stable regression equation but has the advantage that the equation coefficients are easily interpretable. The transformations of Laffort and Héricourt\(^1\) when applied to other data sets than the original 369 compound set will yield descriptors that are not exactly orthogonal and with the disadvantage that the equation coefficients cannot easily be interpreted. We do not believe that the advantage of a possible increase in orthogonality outweighs the chemical information that is lost when the coefficients take on unrealistic values. In several past studies\(^6\)–\(^11\) we have used a number of methods to compare derived correlation equations for biological processes with various correlation equations for gas-to-organic solvent and water-to-organic solvent partitions in order to determine the molecular interactions responsible for the observed biological property. Indeed, such analyses become meaningless if the equation coefficients were to lose their encoded chemical information.

So useful is the comparison of equation coefficients that Ishihama and Asakawa\(^12\) developed a rigorous mathematical procedure for such comparison. A rather different mathematical procedure has been given by Abraham and Martin\(^s\)\(^,\)\(^13\)\(^,\)\(^14\). Given the procedures that have been developed for the comparison of equation coefficients, based on the coefficients encoding specific chemical information, and in view of our own observations, we prefer the existing solute descriptors over the calculated values based on the transformations suggested in the Laffort and Héricourt paper. We are unable at the present time to evaluate the second set of optimized solvation parameters recommended by Laffort and co-workers. There are too few compounds with known solvation parameters for us to determine what the calculated equation coefficients might be. The databases used in deriving the existing Abraham gas-to-organic solvent partitions contain a large number of compounds not found in the 133 compound method 1 solvation parameter database.

REFERENCES AND NOTES