

# On the Effect of Temperature on Aqueous Solubility of Organic Solids

Simon Black\* and Frans Muller

AstraZeneca, Pharmaceutical Development, Macclesfield, SK10 2NA, England

## Abstract:

The variation of solubility with temperature in water is an important consideration in the design of crystallizations and other processes. Solubility was found to vary exponentially with temperature (in °C) for a literature data set of 50 organic molecules in water. This complements a similar recent finding for the solubility of organic molecules in organic solvents. The ‘solubility doubling temperature’ is a convenient way to express this variation and provides a numerical link between enthalpy of dissolution and yield. The ‘solubility doubling temperature’ in water varies with solute, and the median value from this data set is 21.4 °C, consistent with ‘Black’s rule’.

## 1. Introduction

The variation of solubility with temperature is an important consideration in the design of crystallizations and other processes, and is intimately linked to the yield from cooling crystallizations.<sup>1–3</sup> Knowledge of the average variation of solubility with temperature sets general expectations about such yields, and assists the selection of methods for solubility measurement and crystallisation process development.<sup>3</sup>

A recent review of 110 data sets for molecular solids in organic solvents indicated that in 85% of cases the variation of solubility with temperature (°C) was exponential.<sup>3</sup> The ‘solubility doubling temperature’ was introduced as a convenient way of expressing this relationship for a given solvent and solute. The median value of this ‘solubility doubling temperature’ ( $\Delta T_{sd}$  hereafter) was found to be 20 °C. This process heuristic, denoted ‘Black’s Rule’, was used to set a general expectation for yields from cooling crystallizations and to justify a simple approach for designing cooling crystallisation processes.

Of the 110 data sets examined,<sup>3</sup> only six were in pure water. Aqueous solubilities are relevant to ‘water washes’, in vitro dissolution testing and bioavailability. The discovery of a compilation of 50 sets of aqueous solubility data in the literature<sup>4</sup> triggered this investigation to further test the hypotheses that:

- Aqueous solubilities of organic solids vary exponentially with temperature.
- The median value of  $\Delta T_{sd}$  is ~20 °C (Black’s Rule).

\* Author to whom correspondence may be sent. E-mail: simon.black@astrazeneca.com.

- (1) Parsons, A. R.; Black, S. N.; Colling, R. Automated Measurement of Metastable Zones for Pharmaceutical Compounds. *Trans. IChemE* **2003**, *81*, 700–704.
- (2) Barrett, P.; Glennon, B. Characterising the metastable zone width and solubility curve using the Lasentec FBRM and PVM. *Chem. Eng Res. Des.* **2002**, *80*, 799–805.
- (3) Muller, F. L.; Fielding, M.; Black, S. A Practical Approach for Using Solubility to Design Cooling Crystallisation. *Org. Process Res. Dev.* **2009**, *13*, 1315–1321. DOI: 10.1021/op9001438.
- (4) Mullin, J. W. *Crystallization*, 4th ed.; Butterworth-Heinemann: Woburn, MA, 2001; ISBN 0-7506-4833-3.

During the search for suitable literature data sets, a different method of analyzing and fitting temperature/solubility data was noted.<sup>5,6</sup> It is common practice to fit data to three-parameter equations such as eq 1:

$$\ln x = A + B/T + C(\ln T) \quad (1)$$

where  $x$  is the solubility expressed as mole fraction and  $T$  is the temperature in Kelvin.

Values of  $A$ ,  $B$ , and  $C$  (for example in Table 2 of refs 5 and 6) show changes in sign and magnitude, even for the same solute, and no physical meanings are assigned to these parameters. Moreover, the link between this expression and solubility data (in grams per litre) and yields is not intuitive. In contrast, the hypothesis above, that solubility varies exponentially with temperature, implies a two-parameter fit in which both parameters have physical meanings and are constrained to be positive. Therefore, an additional aim of this work was to compare these two approaches from a theoretical viewpoint. This comparison is presented in section 2, justifying a simple method for analysing temperature/solubility data which is illustrated in section 3 and used to analyse the literature data set<sup>4</sup> in section 4.

## 2. Theory

A mathematical representation of ‘solubility varies exponentially with temperature’ is:

$$S = c \cdot e^{(dT)} \quad (2)$$

where  $S$  is the solubility,  $T$  is the temperature in °C,  $c$  is the solubility at 0 °C and  $d$  is a constant. In this paper the units of solubility are grams per litre of solvent. Where solubility, expressed as mole fraction, is <0.1, it is directly proportional to the solubility expressed as grams per litre of solvent, so this will affect the constant  $c$ , but not the constant  $d$ .

The ‘solubility doubling temperature’ is defined as:

$$\Delta T_{sd} = (\ln(2)/d) = 0.693/d \quad (3)$$

$\Delta T_{sd}$  describes the variation of solubility with temperature using a single number. A large  $\Delta T_{sd}$  corresponds to a shallow solubility/temperature curve and a low yield from a cooling crystallization. A small  $\Delta T_{sd}$  corresponds to a steep solubility/temperature curve and a high yield from a cooling crystallisation.

- (5) Dang, L.; Du, W.; Black, S.; Wei, H. Solubility of Fumaric Acid in Propan-2-ol, Ethanol, Acetone, Propan-1-ol and Water. *J. Chem. Eng. Data* **2009**, *54*, 3112–3113. DOI: 10.1021/je9001637.
- (6) Yu, Q.; Black, S.; Wei, H. Solubility of Butanedioic Acid in Different Solvents at Temperatures between 283 and 333 K. *J. Chem. Eng. Data* **2009**, *54*, 2123–2125. DOI: 10.1021/je900021g.

As discussed previously,<sup>3</sup> the variation of solubility with temperature can be described by considering the change in Gibbs free energy on dissolution, and then simplifying for a two-parameter fit, as shown in eq 4.

$$\ln x_1 = (\Delta S_d/R) - (\Delta H_d/RT_1) \quad (4)$$

where  $x_1$  = solubility in mole fraction at temperature  $T_1$ ;  $\Delta S_d$  = entropy of dissolution, in J/mol·K;  $\Delta H_d$  = enthalpy of dissolution, in J/mol;  $R$  = molar gas constant = 8.314 J/mol·K; and  $T_1$  = absolute temperature in Kelvin.

This can be simplified to eq 5:

$$\ln x_1 = a - (b/T_1) \quad (5)$$

where  $a = \Delta S_d/R$  and  $b = \Delta H_d/R$ , provided that  $a$  and  $b$  are independent of temperature. These two expressions appear independently in Mullin<sup>4</sup> (eqs 3.6 and 3.18b). Equation 5 resembles eq 1 with  $C = 0$ , clarifying the physical meanings of the constants  $A$  and  $B$ .

It was noted previously<sup>3</sup> that eq 5 gave a good data fit for 85% of solute/solvent data sets, justifying the assumptions that  $\Delta S_d$  and  $\Delta H_d$  are constant for a given solute/solvent pair over the temperature range of interest. One reason for poor data fits could be the appearance of solvates, hydrates, or enantiotropic polymorphs over part of this temperature range.

Comparison of eq 2 with eq 4 shows that, although both are two-parameter fits, solubility and temperature are treated differently. In most systems of interest to the pharmaceutical industry,  $x < 0.1$ , and thus, the different units for solubility will not be significant. In eq 2 temperature is quoted in °C, whereas eq 4 uses the reciprocal of the absolute temperature. Clearly both equations cannot be exactly correct. The two equations can be compared by considering how eq 4 treats the ratio of solubilities  $x_1$  and  $x_2$  at two different temperatures,  $T_1$  and  $(T_1 + \Delta T)$ :

$$\ln(x_2/x_1) = (\Delta H_d/R) \cdot (1/T_1 - 1/(T_1 + \Delta T)) \quad (6)$$

$\Delta T_{sd}$  is the temperature difference for which the solubility ratio  $x_2/x_1 = 2$ :

$$\Delta T_{sd} \cdot H_d = RT_1 \cdot \ln(2) \cdot (T_1 + \Delta T_{sd}) \quad (7)$$

In general,  $T_1 \gg \Delta T_{sd}$

$$\Delta T_{sd} \cdot \Delta H_d \approx RT_1^2 \cdot \ln(2) \quad (8)$$

Higher values of  $\Delta H_d$  will lead to lower values of  $\Delta T_{sd}$  and, hence, steeper solubility curves and higher yields from cooling crystallizations. Equation 8 confirms that it is not possible for both  $\Delta T_{sd}$  and  $\Delta H_d$  to be exactly constant over a wider temperature range such as from 0 to 100 °C. However, over narrower temperature ranges, if either  $\Delta H_d$  or  $\Delta T_{sd}$  is constant, the other will be approximately constant. For example, close to 20 °C and substituting  $R = 8.314$  J/mol·K,  $\ln(2) = 0.693$ ,  $T_1 = 20$  °C = 293.15 K:

$$\Delta T_{sd} \cdot H_d \approx 500 \text{ K} \cdot \text{kJ/mol} \quad (9)$$

Hence, the median value of  $\Delta T_{sd} \approx 20$  °C ('Black's rule') arises from the median value of  $\sim 25$  kJ/mol for  $\Delta H_d$ .<sup>3</sup>

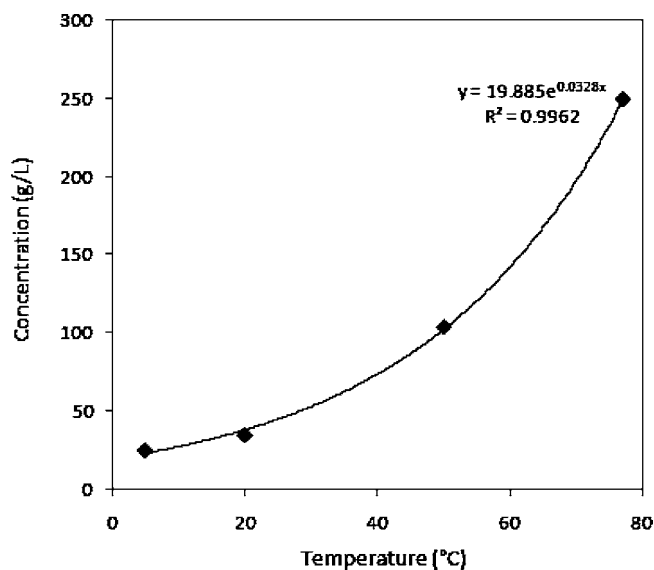


Figure 1. Example of fitting solubility data using equation 2.

### 3. Demonstration for a Pharmaceutical Compound

In this section the use of  $\Delta T_{sd}$  is illustrated with an example from the literature.<sup>7</sup> The published solubility data for this compound were presented as a graph which included a 'line to guide the eye'. Figure 1 shows the same data together with an exponential fit in the form of eq 2. The fit for these four data points is  $R^2 = 0.996$ , and the 'solubility doubling temperature' is given by eq 3 as  $0.693/0.0328 = 21$  °C.

It follows that a yield of 87.5% is predicted for a cooling crystallization over a temperature range of 63° (=  $3 \times \Delta T_{sd}$ ). In practice, a cooling crystallization gave a yield of 86% over four batches at pilot-plant scale.<sup>7</sup>

Figure 2 shows the same data plotted as  $\ln(\text{mole fraction})$  against the reciprocal of the absolute temperature. From the gradient of the black straight line ( $R^2 = 0.989$ ), the enthalpy of dissolution is 26 kJ/mol. The product of the enthalpy of dissolution and the doubling temperature is 546 K·kJ/mol, in reasonable agreement with eq 9.

The green line in Figure 2 shows the exponential fit to the data from Figure 1. This illustrates the impossibility of fitting the same data exactly using both eqs 2 and 4. However, Figure 2 shows that, for this data set, either approach is acceptable.

The dotted line in Figure 2 shows the ideal solubility, calculated using the known melting temperature and enthalpy of melting (195 °C, 36 kJ/mol).<sup>7</sup> The gradient of the 'ideal solubility' line is too steep, because the enthalpy of melting is much larger than the enthalpy of dissolution. This illustrates the difficulty in using ideal solubility as a basis for fitting (as opposed to estimating) the variation of solubility with temperature. This is a general phenomenon that is discussed more fully elsewhere.<sup>3,8,9</sup>

(7) Black, S. N.; Quigley, K.; Parker, A. A Well-Behaved Crystallisation of a Pharmaceutical Compound. *Org. Process Res. Dev.* **2006**, *10*, 241–244.

(8) Granberg, R. A.; Rasmuson, A. C. Solubility of paracetamol in pure solvents. *J. Chem. Eng. Data* **2000**, *44*, 1391–1395.

(9) Nordström, F. L.; Rasmuson, A. C. Prediction of solubility curves and melting properties of organic and pharmaceutical compounds. *Eur. J. Pharm. Sci.* **2009**, *36*, 330–344.

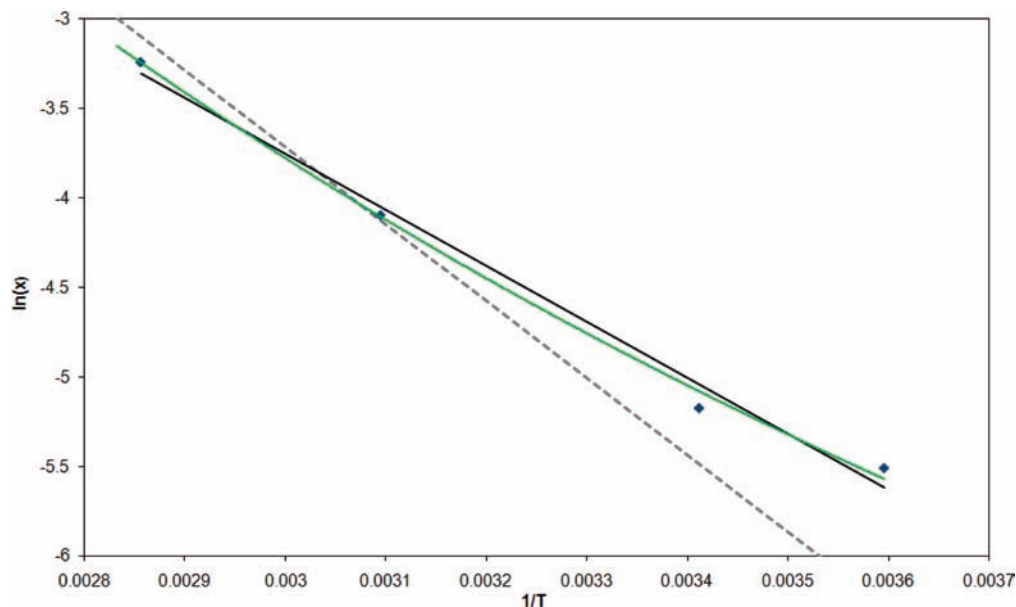


Figure 2. Fitting solubility data from Figure 1 using eq 2 - green line; eq 4 - black line. Ideal solubility - dotted grey line.

#### 4. Organic Solids in Water

Table A.5 in J. W. Mullin's book on *Crystallisation*<sup>4</sup> contains data on aqueous solubility as a function of temperature for several simple organic solids. These data were analysed to answer the following questions:

- Do aqueous solubilities vary exponentially with temperature?
- Is eq 2 as good as eq 4 for data fitting?
- What is the median doubling temperature?
- Is the proposed relationship between  $\Delta T_{sd}$  and  $\Delta H_d$  in eq 9 valid?

**4.1. Comments on the Data.** The table contains 51 data sets, each consisting of solubility data in g/100 g water at up to eight temperatures between 0 and 100 °C. There are no references, and the method by which the solubilities were determined is not given. Five monohydrates, one dihydrate, and one pentahydrate are identified, but there is no mention of polymorphism. The chemical formulae for *m*-aminophenol and *p*-aminophenol give nitro groups in place of amino groups; the names have been used here. There are separate entries for *o*-hydroxybenzoic acid and salicylic acid, although the molecular formula and the solubility data are identical—one of these data sets was removed before the start of the analysis, leaving a total of 50.

Thirty of these sets contain data for all eight temperatures, and for 11 sets only one temperature is missing. There are three sets containing 6 data, two with 5 data, three with 4 data, and one with 2 data. This last set could not be used to assess the exponential fit. By inspection, all 50 data sets showed solubility increasing with temperature.

**4.2. Data Fitting.** The 50 data sets were fitted using eq 2 and eq 4 in an Excel spreadsheet which is available as Supporting Information. Figure 3 shows the distribution of the  $R^2$  values for eq 2 (solid diamonds) when sorted in order of decreasing size. The corresponding  $R^2$  values for eq 4 are shown as hollow squares.

As Figure 3 shows, in general the data fits are good for both eqs 2 and 4. It is possible to extract meaningful values for  $\Delta T_{sd}$

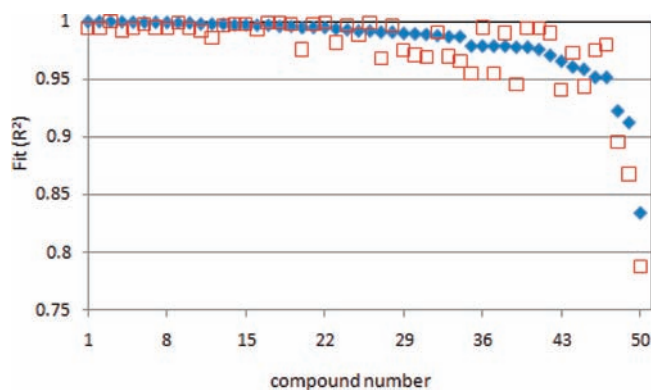


Figure 3. Distribution of  $R^2$  values:  $\blacklozenge$  - eq 2;  $\square$  - eq 4.

and  $\Delta H_d$  for all 50 data sets. The three compounds which give the poorest fits ( $R^2 < 0.94$  by both methods) are *o*-aminophenol (0.835), *p*-aminophenol (0.913), and benzamide (0.922). Excluding these three compounds, the average values of  $R^2$  are 0.988 (eq 1) and 0.985 (eq 3). In some cases eq 2 gives a better fit, and in some cases eq 4 gives a better fit, but the differences are small—either equation can be used to fit each data set. The fits for the seven hydrates (numbers 11, 15, 21, 26, 33, 40, and 46 in Figure 3) are not significantly different from those of the other compounds.

**4.3. Solubility Doubling Temperature.**  $\Delta T_{sd}$  was calculated for each compound using eq 3, and the spread of solubility doubling temperatures is shown in Figure 4. In this Figure, compound 1 is fructose with  $\Delta T_{sd} = 231$  °C. Compound 50, with the lowest  $\Delta T_{sd}$  of 7.1 °C, is benzamide, which gave a poor data fit.  $\Delta T_{sd}$  is not constant, and the median is 21.4 °C, consistent with the hypothesis that 'on average  $\Delta T_{sd} \approx 20$  C°' (Black's Rule).

**4.4. The Link to Enthalpy of Dissolution.** The enthalpy of dissolution was calculated for each compound using eq 4. The median value for  $\Delta H_d$  is 23.4 kJ/mol. Figure 5 shows a comparison between the enthalpies of dissolution calculated from eq 4 and the solubility doubling temperatures calculated using eq 2 for 49 of the 50 compounds analysed. The curve is

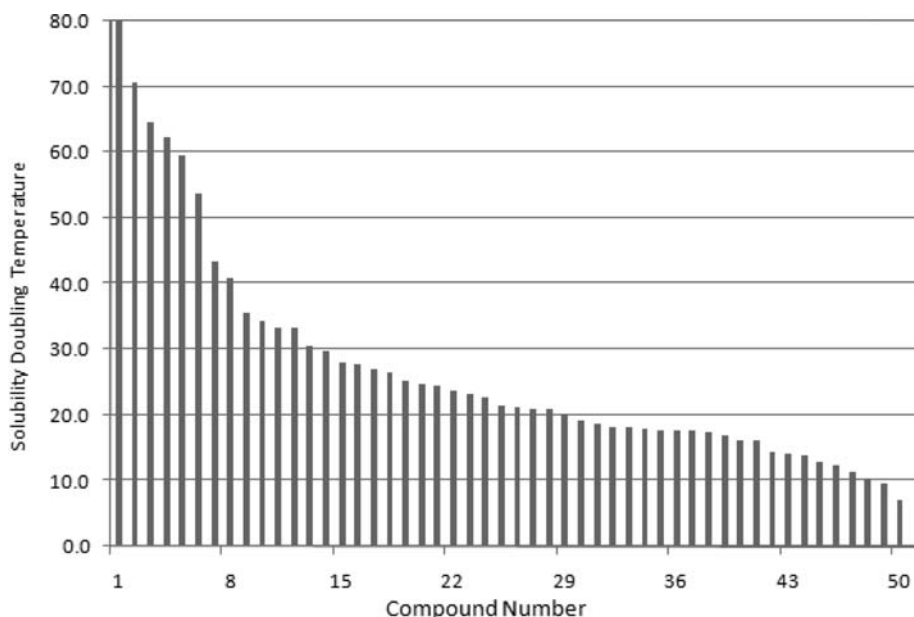


Figure 4. Solubility Doubling Temperatures

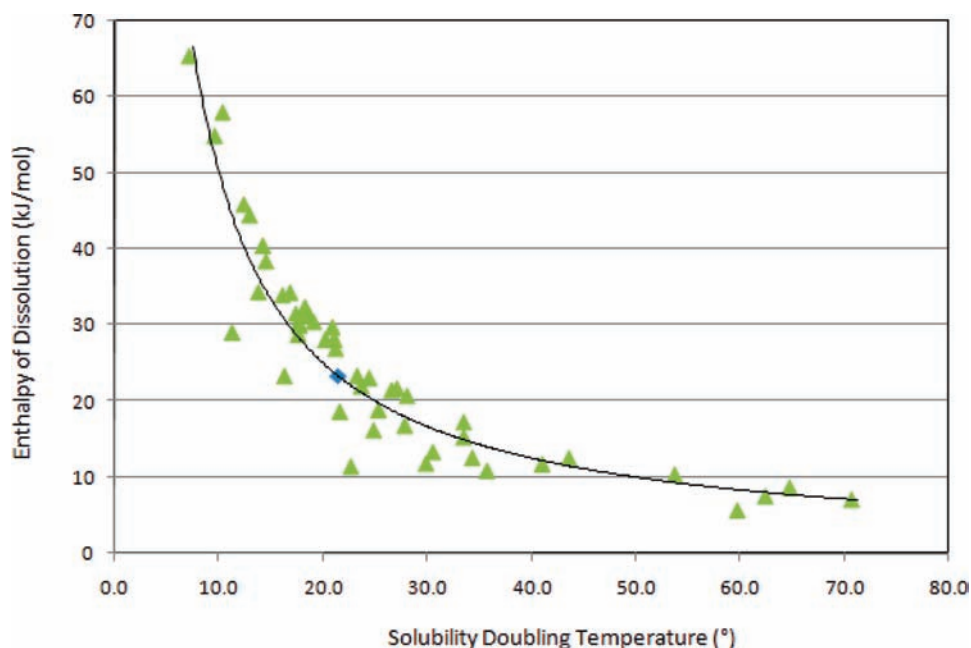


Figure 5. Solubility doubling temperatures and enthalpies of dissolution. The blue diamond indicates the medians. The solid line corresponds to eq 9.

the relationship between doubling temperature and enthalpy of dissolution predicted by eq 9. Fructose ( $\Delta T_{sd} = 231$  °C,  $\Delta H_d = 2.1$  kJ/mol;  $\Delta T_{sd} \times \Delta H_d = 485$  K·kJ/mol) is omitted from the Figure, although it would lie close to the curve. The average value of  $\Delta T_{sd} \times \Delta H_d$  for this complete data set is 511 K·kJ/mol, in good agreement with eq 9.

## 5. Discussion

Figure 3 demonstrates that the exponential fit according to eq 2 is as good as a more conventional two-parameter fit based on eq 4. Equation 2 is easier to use and is linked directly to the ‘solubility doubling temperature’ which can be used for process design. Out of 50 compounds, 47 gave good fits ( $R^2 > 0.95$ ) with this model. On this evidence, the general use of three-parameter fits such as eq 1 is hard to justify. One benefit from

fitting the data using two parameters is that potentially erroneous data are identified. Although many of the fits could have been improved by omitting one of the data points, further analysis was outside the scope of this study.

The goodness of fit and the spread and median of  $\Delta T_{sd}$  are consistent with the previous analysis of solubility of organic solids in predominantly organic solvents.<sup>3</sup> There is no evidence here for specific interactions between organic molecules and water that affect the variation of solubility with temperature. Taken together, these two studies cover 160 solubility data sets for organic compounds in aqueous and nonaqueous solvents. This supports the general use of eq 2 and  $\Delta T_{sd}$  when designing experiments and analysing data on the solubility of organic compounds.

There is a spread in the distribution of enthalpies of dissolution, as shown in Figure 5. The median is similar to that reported previously for other solvents.<sup>3</sup> The correlation between  $\Delta T_{sd}$  and  $\Delta H_d$  follows directly from the similar data fits to eq 2 and 4 which is demonstrated in Figure 3. The good agreement with the curve predicted by eq 9 supports the theoretical basis for the concept of 'solubility doubling temperatures', which are in general close to  $500/\Delta H_d$  as predicted. 'Black's Rule' can now be understood as arising from the median value for the enthalpy change in the relevant process (here dissolution). This is analogous to the justification for 'reaction kinetics double every 10 °C' in terms of activation energies.

There is no evidence in this analysis that hydrates of organic compounds behave differently. The behaviour of organic salts has not been addressed here, primarily due to lack of data. It is anticipated that the extensive literature on the solubility of inorganic salts in water will be a useful guide.

## 5. Conclusions

This study of the effect of temperature on solubility of 50 organic compounds in water showed no significant differences with a previous study<sup>3</sup> of the effect of temperature on the solubility of organic compounds in organic solvents. Taken together, 138 out of 160 data sets show an exponential increase of solubility with temperature. The median 'solubility doubling temperature' is close to 20 °C.

## Supporting Information Available

An Excel spreadsheet of the 50 data sets fitted using eqs 2 and 4. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review January 11, 2010.

OP100006Y