

RESEARCH ARTICLES

Extended Hildebrand Solubility Approach and the Log Linear Solubility Equation

A. MARTIN **, P. L. WU *, A. ADJEI *, R. E. LINDSTROM ‡, and P. H. ELWORTHY §

Received July 30, 1981, from the *Drug Dynamics Institute, College of Pharmacy, University of Texas, Austin, TX 78712; †School of Pharmacy, University of Connecticut, Storrs, CT 06268; and the ‡Department of Pharmacy, The University of Manchester, Manchester M139PL, England. Accepted for publication November 10, 1981.

Abstract □ The log linear solubility equation, $\log S = \log S_w + \sigma f$, was studied in relationship to the extended Hildebrand solubility approach. It is shown that the log linear form may be derived beginning with the extended Hildebrand approach. The log linear expression gives a good linear fit for semipolar drugs in a number of water-cosolvent mixtures. It is particularly successful when the solubility parameter, δ_1 , of the cosolvent is 3 or more solubility parameter units larger than the solubility parameter, δ_2 , of the drug. When the cosolvent tends to solvate the drug strongly, the log linear function may even hold where the solubility parameters of the drug and cosolvent are similar. It appears, however, not to be applicable to nonpolar cosolvent systems. An interfacial model for the solubility of drugs in polar mixed solvents is based on σ , a parameter that also figures prominently in the log linear solubility equation. When used to describe mixed solvent systems, the interfacial model applies in the region of the solubility profile (solubility versus solvent composition) where the log linear relationships hold. The extended Hildebrand solubility approach is applicable over a wide range of cosolvent composition in mixed systems from nonpolar organic solvents to water.

Keyphrases □ Solubility—extended Hildebrand approach, log linear equation □ Hildebrand equation—extended solubility approach, log linear solubility □ Log linear solubility equation—relationship to the extended Hildebrand solubility approach

Yalkowsky *et al.* (1), introduced a log linear equation:

$$\log S = \log S_w + \sigma f \quad (\text{Eq. 1})$$

which describes the solubility of some drugs in binary aqueous systems, where S is the solute solubility in moles per liter in a solvent consisting of water and a nonaqueous cosolvent, S_w is the drug's solubility in water, f is the volume fraction of the cosolvent, and σ is a parameter representing the solubilizing power of the cosolvent for the drug and depends on the polarity of the drug and the cosolvent. Equation 1 was found to be applicable to systems where the polarity of the drug was significantly less than either of the solvents in the binary mixture.

A study (2) on the solubility of *p*-aminoacetophenone in propylene glycol-water mixtures, found it necessary to expand Eq. 1 into a fifth degree polynomial of f to account for nonlinearity across the range of cosolvent (propylene glycol) composition. The linear dependence of logarithmic solubility on volume fraction of the cosolvent (Eq. 1) applied when the Hildebrand solubility parameters of both solvent components were much larger than the solubility parameter of the drug.

THEORETICAL

In the present report, it is shown that the linear relationship of Eq. 1 may be considered in terms of the extended Hildebrand solubility approach (3-5); a model in which the solubility parameter of the solute may be larger or smaller than that of either solvent or lie between the solubility parameters of the two solvents. When the range of solubility parameters of the solvent pair approaches the solubility parameter of the solute, the curve may bow sufficiently that a log linear expression of X_2 on f no longer fits the data satisfactorily. A quadratic or higher polynomial of f must then be used as required by the extended Hildebrand method.

The following derivation shows the relationship of the log linear equation to the extended Hildebrand solubility approach. For the solubility of a drug in pure water:

$$\log X_w = \log X^i - \log \alpha_w \quad (\text{Eq. 2})$$

where X_w and X^i are the mole fraction solubility in water and the ideal solubility of the solute, respectively. $\log \alpha_w$ is the logarithm of the solute activity coefficient in water. A general expression for the solubility of a drug in a binary mixture, consisting of water and an organic cosolvent, is:

$$\log X_2 = \log X^i - \log \alpha_2 \quad (\text{Eq. 3})$$

where X_2 and α_2 are the mole fraction solubility and activity coefficient, respectively, for the solute in the solvent mixture. Subtracting Eq. 2 from Eq. 3 results in:

$$\log X_2 = \log X_w + \log \alpha_w - \log \alpha_2 \quad (\text{Eq. 4})$$

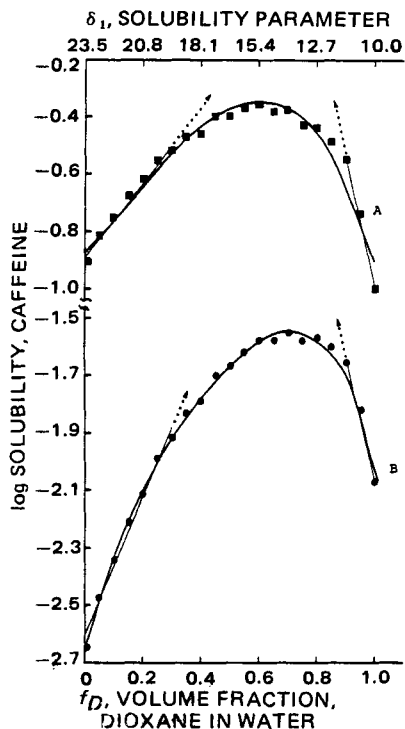


Figure 1—Solubility of caffeine in dioxane–water mixtures at 25°. The curves were calculated using the extended Hildebrand solubility approach (4, 5). Key: (■) log molarity, Curve A; (●) log mole fraction, Curve B.

The logarithmic solubility of a drug in water and the drug's log α_w are constants at a definite temperature. For caffeine in water at 25° (4), log $X_w = -2.64111$ and log $\alpha_w = 1.4764$; therefore, log $X^i = -1.1647$.

According to the extended Hildebrand approach (3–5), the activity coefficient of the drug in a mixed solvent is expressed as:

$$\log \alpha_2 = A(\delta_1^2 + \delta_2^2 - 2W) \quad (\text{Eq. 5})$$

where δ_1^2 and δ_2^2 are the cohesive energy densities of the solvent and solute, respectively, and W is the interaction energy density of solute and solvent. The term A is obtained from regular solution theory (6):

$$A = \frac{V_2\phi_1^2}{2.303RT} \quad (\text{Eq. 6})$$

where V_2 is the solute's liquid molar volume, ϕ_1 is the volume fraction of the pure or binary solvent (total volume fraction of the two solvents in the solution), R is the molar gas constant, and T is the absolute temperature.

Either W or log α_2/A may be regressed in a polynomial on δ_1 to obtain calculated values of log α_2 and mole fraction solubility, X_2 (Eq. 3). Log α_2/A_{calc} may also be obtained by a regression of log α_2/A on f_i , where f_i is the volume fraction of either solvent in a binary mixture (4), for example, water or its cosolvent. For dioxane, D , as the cosolvent:

$$\log \alpha_2/A_{\text{calc}} = C_0 + C_1f_D + C_2f_D^2 + C_3f_D^3 + \dots + C_n f_D^n \quad (\text{Eq. 7})$$

For caffeine in dioxane and water at 25°, the extended Hildebrand equation may be expressed as a fourth degree power series in conformity with Eqs. 4 and 7:

$$\log X_{2\text{calc}} = \log X_w + \log \alpha_w - C_0A - C_1Af_D - C_2Af_D^2 - C_3Af_D^3 - C_4Af_D^4 \quad (\text{Eq. 8})$$

where C_0 is a constant of regression. C_1 , C_2 , C_3 , and C_4 are the coefficients resulting from the regression analysis and remain constant over the range of binary solvent composition for a particular drug at a definite temperature. The value of A varies, although slightly, between 0.093 and 0.103 across the composition of water and dioxane in the caffeine system at 25°.

EXPERIMENTAL

The solubility data in mixed solvents employed in this paper were reported previously (4, 7–9). The drug was brought to equilibrium with the

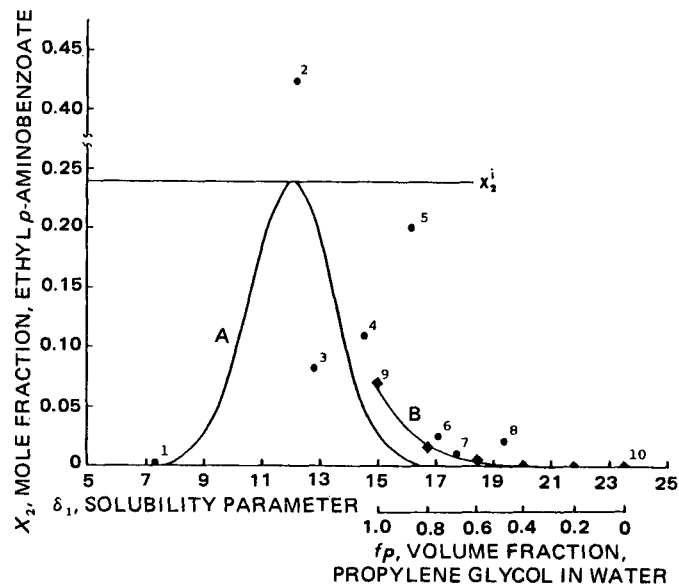


Figure 2—Mole fraction of ethyl *p*-aminobenzoate in pure and binary solvents (9). A, regular solution curve at 25°; B, extended Hildebrand curve for ethyl *p*-aminobenzoate in propylene glycol–water mixture at 37°; $X^i = 0.2404$ at 25°. Key: (●) solubility in individual solvents at 25°; 1, hexane; 2, dimethylformamide; 3, ethanol; 4, methanol; 5, methylformamide; 6, ethylene glycol; 7, glycerin; 8, formamide; 9, propylene glycol (37°); 10, water (37°); (◆) solubility in propylene glycol–water mixtures at 37°, use of Eq. 18b.

mixed solvent in a constant temperature shaker bath and the drug concentration was determined by spectrophotometric or other convenient analytic procedures. Densities were determined in glass pycnometers under controlled conditions. By obtaining the density of each solution, it was possible to express solubilities in units of molar, molal, or mole fraction concentration.

For calculations involving the extended Hildebrand solubility approach, regression analysis was conducted¹ using multiple linear regression programs (10).

RESULTS AND DISCUSSION

Caffeine in Water–Dioxane Mixtures—Introducing the coefficients of regression analysis into Eq. 8 for caffeine in a mixed solvent system of water and dioxane (4), one obtains:

$$\log X_{2\text{calc}} = \log X_w + \log \alpha_w - 14.6274A + 38.4195Af_D - 109.6076Af_D^2 + 114.3535Af_D^3 - 49.3831Af_D^4 \quad (\text{Eq. 9})$$

Equation 9 may be simplified by truncating it after the first term in f_D to yield:

$$\log X_{2\text{calc}} = \log X_w + (\log \alpha_w - 14.6274A) + 38.4195Af_D \quad (\text{Eq. 10})$$

or in general:

$$\log X_{2\text{calc}} = \log X_w + (\log \alpha_w - C_0A) + C_1Af_D \quad (\text{Eq. 11})$$

The quantity in parentheses is sufficiently small to be neglected since log α_w is equal to 1.4764 and $C_0A \cong 14.627 \times 0.1 = 1.4627$. If the regression procedure yielded log α_2/A with perfect accuracy, C_0 would exactly equal (log α_w)/ A . This can be seen by setting f_D equal to zero in Eq. 11, for then the solvent obviously is water, and log $\alpha_2/A = \log \alpha_w/A = C_0$. With the quantity within the parentheses, Eq. 11, equal to zero, one arrives at an equation, expressed in mole fraction:

$$\log X_{2\text{calc}} = \log X_w + \sigma f_D \quad (\text{Eq. 12})$$

analogous to Eq. 1, the log linear expression which, however, has usually been expressed in concentration units of moles per liters or grams per cubic centimeter. The coefficient, $C_1A = 38.4195A$ in Eq. 9, has been obtained by polynomial regression and is equal to σ in Eq. 1 when concentrations are expressed in mole fraction.

One form of the extended Hildebrand expression, Eq. 8, when carried

¹ University of Texas Cyber System.

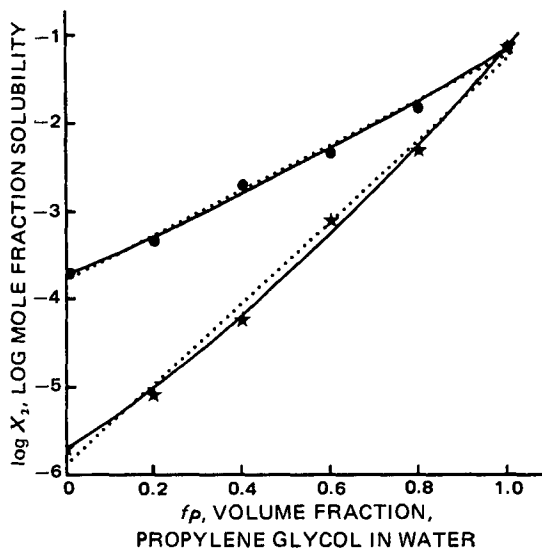


Figure 3—Log mole fraction solubility of ethyl *p*-aminobenzoate (●) and hexyl *p*-aminobenzoate (★) in propylene glycol–water mixtures at 37° (9). Key: (—) from quadratic equations and (.....) from linear regression line.

to the fourth power in f_D , has been shown to reproduce the solubilities of caffeine in dioxane and water within <12% error and for most data points <4% error, a value within the range of experimental accuracy (4)².

In Fig. 1, log solubility of caffeine is plotted both in units of moles per liter and mole fraction versus volume fraction of the cosolvent, dioxane, as abscissa. Also marked along the horizontal axis (top of Fig. 1) are the solubility parameter values, δ_1 , of the mixed solvents, consisting of water and dioxane.

In dealing with caffeine data in earlier reports (4, 5), various polynomial expressions were used. A cubic expression was not as satisfactory as the quartic equation; a quadratic expression yielded poor results; and an equation truncated to the linear term, σf_D , Eq. 12, would be unacceptable as a fit over most of the curve from dioxane ($\delta_1 = 10.01$) to water ($\delta_1 = 23.45$). As observed in Fig. 1, a linear fit is adequate from 0–30% cosolvent composition, although a straight line cannot reproduce exactly the slight curvature of the line. Both the upper and lower curves appear to be parabolic in form, but the curvature is so slight in the 0 to 30% region of dioxane concentration that the data points are adequately fit by a linear function. The intercept is the log solubility in water on a molar (Curve A) or mole fraction (Curve B) scale. In conformity with the criterion for linearity stated earlier, *i.e.*, δ_1 considerably larger than δ_2 , the lowest value of δ_1 is 18.4, well above $\delta_2 = 13.8$ of caffeine. It should be noted that the cosolvent of water in this case was not the pure solvent, dioxane, but rather a mixture of 30% dioxane–70% water. Using a log linear equation beyond 30 or 40% as seen in Fig. 1 would not produce meaningful results. Extrapolation into the region of the dashed lines in the direction of the arrows would produce erroneous values of solubility. Accordingly, the log linear technique, although often useful, should be used with caution over a wide range of solvent compositions. A little more effort is required to apply the extended Hildebrand solubility approach, but it usually can be made to reproduce solubility in mixed solvent systems with considerable fidelity across the entire range of binary solvent composition.

Although parts of the curve (0–30% dioxane in Fig. 1) may appear to be approximately linear, log solubility is ordinarily not a linear function of f (except where a strongly solvating cosolvent is present as seen later), and is better fitted with a power series in f or δ_1 rather than with a straight line.

In a manner similar to that discussed above, a 10% water in dioxane mixture could be taken as the cosolvent of dioxane in the region at the far right side of Fig. 1. At this low water concentration side of the figure, segments that are almost linear are observed in Curves A and B. Here, log solubilities versus volume fraction of water in the mixture produce linear functions only over a range of 10% water ($\delta_1 = 10$ –11.3). The curves begin to bow markedly at 10% water in dioxane because the solvent δ_1 is

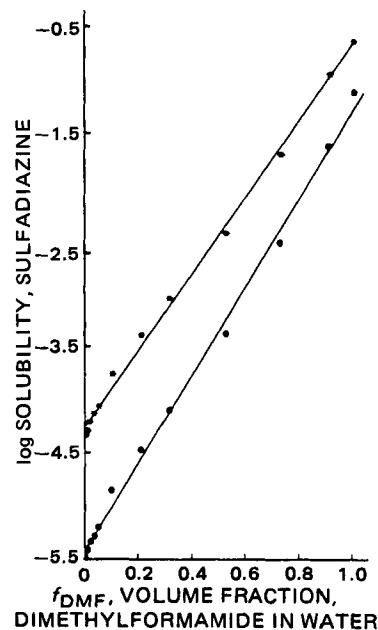


Figure 4—Log solubility of sulfadiazine in dimethylformamide–water mixtures at 20° (7). Key: (●) log mole fraction (lower curve); (★) log *w/w*; linear regression line (upper curve).

approaching the δ_2 value, 13.8, of the solute. According to regular solution theory, the solubility of a compound reaches a peak at a solvent composition where $\delta_1 = \delta_2$. The solubility curve approaches a maximum and then begins to decrease across this region and is greatly curved, as seen in the figure.

Results with theophylline, thebromine, and caffeine in other solvent mixtures have been reported (5). They show similar trends described here for caffeine in mixtures of water and dioxane.

The previous example of a log linear approach in relation to the extended Hildebrand model involved a solute with a solubility parameter, 13.8, between that of water, $\delta_1 = 23.45$ and the cosolvent, dioxane, $\delta_1 = 10.01$. By taking this broad view of the subject, one is able to observe more clearly the regions where a log linear equation is applicable and why it fails when the solubility parameter of the mixed solvent approaches the solubility parameter of the solute.

Alkyl *p*-Aminobenzoates in Water–Propylene Glycol Mixtures—The log linear equation, Eq. 1, was originally introduced by Yalkowsky *et al.* (1), to explain the exponential increase in solubility of poorly water soluble compounds upon the addition of an organic cosolvent to water. They found a number of examples from their own work and the literature in which the solubility data was fit by Eq. 1³. Unlike the case of caffeine in water and dioxane, Eq. 1 applies particularly well where the solubility parameter of the solute is 3 or 4 units below that of the organic component of the aqueous solvent. The solubility of *n*-alkyl *p*-aminobenzoates in propylene glycol–water mixtures and in several pure solvents has been investigated (9). Figure 2 was prepared using data on the solubility of ethyl *p*-aminobenzoate in binary as well as individual solvents. Mole fraction solubility is plotted against the solvent solubility parameter in Fig. 2 as done in studies involving the extended Hildebrand approach (3–5). The regular solution line is plotted in Fig. 2 to show the curve on which the experimental points would fall if these polar solute–solvent systems followed regular solution behavior at 25°. The regular solution line is calculated using the expression:

$$-\log X_2 = -\log X^i + A(\delta_1 - \delta_2)^2 \quad (\text{Eq. 13})$$

where all terms have been defined in earlier equations. The regular solution solubility reaches a peak value at a solubility parameter equal to that of the solute [$\delta_2 = 12.05$ (cal/cm³)^{1/2}]; the mole fraction solubility at this point has the ideal value, $X^i = 0.2404$, for ethyl *p*-aminobenzoate at 25°. At 37° it is 0.3236 ($-\log X_{37}^i = 0.4900$).

It is observed that solutions of ethyl *p*-aminobenzoate in the individual solvents are not regular solutions since they do not lie on the bell-shaped regular solution curve in Fig. 2. The solubility in most of the solvents falls below the ideal solubility line, $X_{25}^i = 0.2404$, and this indicates that ethyl

² In Ref. 4 the volume fraction of water was referred to as ϕ_w . It was used in place of f_D and this led to a polynomial expression different from Eq. 9, but both expressions result in the same calculated solubilities.

³ A recent report (11) states that a number of semipolar drugs exhibit parabolic rather than linear log solubility versus cosolvent composition curves.

Table 1—Regression Equations for Drugs in Binary Solvents Using the Log Linear and The Extended Hildebrand Approaches^a

Mixture	Log Linear and Log Quadratic Approach		Extended Hildebrand Solubility Approach	
	Regression Equation	Log mole Fraction of Solute versus Volume Fraction of Cosolvent	Log α_2/A versus Solubility Parameter of Solvent	
Ethyl <i>p</i> -aminobenzoate in propylene glycol-water $n = 6$, ${}^aF_{(1,4,0.05)} = 7.71$	Linear	$\log X_2 = -3.7861 (\pm 0.0585) + 2.5359 (\pm 0.0965) f$ $R^2 = 0.9942, s = 0.0808, F = 690^a$ (Eq. 14a)	$\frac{\log \alpha_2}{A} = -32.0622 (\pm 1.5389) + 2.7418 (\pm 0.0792) \delta_1$ $R^2 = 0.9967, s = 0.5603, F = 1199^a$ (Eq. 18a)	
	Quadratic	$\log X_2 = -3.7319 (\pm 0.0667) + 2.1296 (\pm 0.3136) f$ $+ 0.4062 (\pm 0.3010) f^2$ $R^2 = 0.9964, s = 0.0736$, overall $F = 417^c$ for f term, $F = 46.1^b$ for f^2 term, $F = 1.8^b$ (Eq. 14b)	$\frac{\log \alpha_2}{A} = -37.7097 (\pm 13.0647) + 3.3430 (\pm 1.3817) \delta_1$ $- 0.0156 (\pm 0.0359) \delta_1^2$ $R^2 = 0.9969, s = 0.6274$, overall $F = 478^c$ for δ_1 term, $F = 5.9^b$ for δ_1^2 term, $F = 0.2^b$ (Eq. 18b)	
Sulfadiazine in dimethylformamide-water $n = 14$, ${}^aF_{(1,12,0.04)} = 4.75$	Linear	$\log X_2 = -5.419 (\pm 0.0234) + 4.2132 (\pm 0.0483) f$ $R^2 = 0.9984, s = 0.0644, F = 7620^d$ (Eq. 15a)	$\frac{\log \alpha_2}{A} = -48.1326 (\pm 1.4470) + 3.0574 (\pm 0.0721) \delta_1$ $R^2 = 0.9934, s = 1.1105, F = 1796^d$ (Eq. 19a)	
	Quadratic	$\log X_2 = -5.4279 (\pm 0.0275) + 4.3429 (\pm 0.2057) f$ $- 0.1438 (\pm 0.2213) f^2$ $R^2 = 0.9985, s = 0.0661$, overall $F = 3627^i$ for f term, $F = 446^e$ for f^2 term, $F = 0.4^e$ (Eq. 15b)	$\frac{\log \alpha_2}{A} = -67.1701 (\pm 7.8444) + 5.2625 (\pm 0.8997) \delta_1$ $- 0.0603 (\pm 0.0245) \delta_1^2$ $R^2 = 0.9957, s = 0.9320$, overall $F = 1278^i$ for δ_1 term, $F = 34.2^e$ for δ_1^2 term, $F = 6.0^e$ (Eq. 19b)	
Methylparaben in formamide-water $n = 10$, ${}^aF_{(1,8,0.05)} = 5.32$	Linear	$\log X_2 = -3.5757 (\pm 0.0128) + 2.7342 (\pm 0.0615) f$ $R^2 = 0.9960, s = 0.0230, F = 1979^g$ (Eq. 16a)	$\frac{\log \alpha_2}{A} = -145.0176 (\pm 3.8840) + 7.4243 (\pm 0.1706) \delta_1$ $R^2 = 0.9958, s = 0.2583, F = 1893^g$ (Eq. 20a)	
	Quadratic	$\log X_2 = -3.6009 (\pm 0.0086) + 3.2824 (\pm 0.1213) f$ $- 1.5847 (\pm 0.3379) f^2$ $R^2 = 0.9990, s = 0.0121$, overall $F = 3599^i$ for f term, $F = 733^h$ for f^2 term, $F = 22^h$ (Eq. 16b)	$\frac{\log \alpha_2}{A} = 439.6364 (\pm 105.2084) - 43.9957 (\pm 9.2520) \delta_1$ $+ 1.1301 (\pm 0.2033) \delta_1^2$ $R^2 = 0.9992, s = 0.1187$, overall $F = 4498^i$ for δ_1 term, $F = 22.6^h$ for δ_1^2 term, $F = 30.9^h$ (Eq. 20b)	
Methylparaben in dimethylformamide-water $n = 10$, ${}^aF_{(1,9,0.05)} = 5.32$	Linear	$\log X_2 = -3.607 (\pm 0.0192) + 5.2419 (\pm 0.0621) f$ $R^2 = 0.9989, s = 0.0326, F = 7123^j$ (Eq. 17a)	$\frac{\log \alpha_2}{A} = -89.4833 (\pm 1.1243) + 5.0712 (\pm 0.0546) \delta_1$ $R^2 = 0.9991, s = 0.3247, F = 8623^j$ (Eq. 21a)	
	Quadratic	$\log X_2 = -3.5917 (\pm 0.0245) + 5.0020 (\pm 0.2457) f$ $+ 0.4927 (\pm 0.4882) f^2$ $R^2 = 0.9995, s = 0.0326$, overall $F = 3570^i$ for f term, $F = 415^k$ for f^2 term, $F = 1.0^k$ (Eq. 17b)	$\frac{\log \alpha_2}{A} = -121.4289 (\pm 12.3283) + 8.1844 (\pm 1.1922) \delta_1$ $- 0.0752 (\pm 0.0290) \delta_1^2$ $R^2 = 0.9995, s = 0.2477$, overall $F = 7412^i$ for δ_1 term, $F = 46.6^k$ for δ_1^2 term, $F = 6.7^k$ (Eq. 21b)	

^a The statistical parameters with each regression equation are n , the number of solvent mixtures used; R^2 , the squared multiple correlation coefficient; s , the standard deviation of the sample; F , the Fisher F ratio, and the tabular value of F with degrees of freedom k and $n-k-1$ at the 95% level. The value, k , is the number of independent variables in the equation. The F values in Column 1 are tabular Fisher ratios. ^b F values in Columns 3 and 4 are to be compared with the tabular aF value; likewise for ^c F , ^d F , ^e F , ^f F , ^g F , ^h F , ⁱ F , ^j F , ^k F .

Table II—Mole Fraction Solubilities of Ethyl *p*-Aminobenzoate at 37° in Propylene Glycol–Water Mixtures ^a

Propylene Glycol, %	δ_1^b	A^c	$\text{Log } \alpha_2$	$\frac{\text{Log } \alpha_2}{A}$	$X_{2\text{obs}} \times 10^4$	$X_{2\text{calc}} \times 10^4$ Linear Eq. 18a	$X_{2\text{calc}} \times 10^4$ Quadratic Eq. 18b
0	23.45	0.1013	3.244	32.020	1.845	1.757	1.810
20	21.74	0.1011	2.839	28.080	4.688	5.278	5.218
40	20.07	0.1002	2.230	22.255	19.055	16.175	15.679
60	18.37	0.0990	1.847	18.657	46.026	49.870	48.389
80	16.68	0.0949	1.338	14.099	148.59	163.16	161.73
100	14.99	0.0775	0.671	8.658	690.24	645.06	661.48

^a Reference 9; solubilities, $X_{2\text{calc}}$, obtained with the extended Hildebrand solubility approach, Eqs. 18a and b; $-\log X_2^i = 0.49000$, $\delta_2 = 12.05$, $V_2 = 144$. ^b Solvent solubility parameter, $(\text{cal}/\text{cm}^3)^{1/2}$. ^c Equation 6.

p-aminobenzoate is not highly solvated. By contrast, ethyl *p*-aminobenzoate associates strongly with *N,N*-dimethylformamide (Point 2 in Fig. 2), producing a solubility ($X_2 = 0.4260$) above the ideal solubility line. The point for ethyl *p*-aminobenzoate in propylene glycol at 37° (Point 9) lies slightly to the right of the regular solution curve which was drawn for solutions at 25°.

The discussion of the solubility parameter of the compound ($\delta_2 = 12.05$) and the regular solution line is also important for the treatment to follow. However, the current report is a study dealing with mixed solvent systems and attention is drawn to the propylene glycol–water mixtures in Fig. 2. One observes that the data points for the solubility of ethyl *p*-aminobenzoate in mixtures of water and propylene glycol (at 37°) lie alongside the 25° regular solution line. On this scale, solubility in water (Point 10) and in 20% propylene glycol (bottom scale) are imperceptible. The points rise to reasonable solubility values for 30–100% propylene glycol solutions. The line passing through the points in Fig. 2 was calculated using the extended Hildebrand solubility approach, Eq. 18a of Table I:

$$\log \alpha_2/A = -37.7097 + 3.3430\delta_1 - 0.0156\delta_1^2$$

Multiplying both sides by A and adding $-\log X^i$ yields:

$$-\log X_{2\text{calc}} = 0.4900 + A(-37.7097 + 3.3430\delta_1 - 0.0156\delta_1^2) \quad (\text{Eq. 22})$$

from which $X_{2\text{calc}}$ is obtained.

Since the solubility parameter of propylene glycol ($\delta_1 = 15$) is sufficiently different from that of ethyl *p*-aminobenzoate ($\delta_2 = 12$), the curve in Fig. 2 may be transformed into a linear form by use of the log linear approach. Figure 3 shows the plots for two esters, ethyl and hexyl *p*-aminobenzoate, in propylene glycol–water mixtures at 37° (9).

The linear least-square line through the data is represented in Table

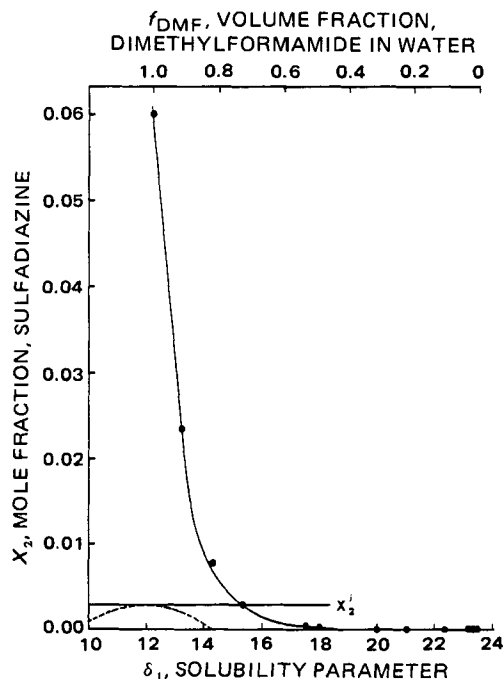


Figure 5—Mole fraction solubility of sulfadiazine ($X_2^i = 0.0030$) in dimethylformamide–water mixtures at 20° (7). Key: (●) experimental solubility; (---) regular solution curve, Eq. 13; (—) extended Hildebrand line, Eq. 23 (cubic expression).

I by Eq. 14a, and the linear form is compared with a quadratic expression, Eq. 14b for the ethyl ester. Although the coefficient of determination is higher for the quadratic, $R^2 = 0.9964$, than for the linear form, $R^2 = 0.9942$, the Fisher F ratio conversely is greater for the linear (690) than for the quadratic (417) expression. Furthermore, the partial F values for introducing f and then f^2 show that f^2 is not significant when its F value, 1.8, is compared with the table value, 10.1, at the 95% level. The log linear fit is, therefore, considered to be satisfactory.

In Table I (Eqs. 18a and b), the extended Hildebrand solubility approach is shown for a regression of $\log \alpha_2/A$ on δ_1 in both linear and quadratic form for the ethyl ester. Again, the F value is significantly larger for the linear than the quadratic, and the partial F values suggest that the quadratic is not as satisfactory as the linear equation. Although the statistical analysis of the hexyl *p*-aminobenzoate in propylene glycol–water mixtures is not included here, the same conclusion was reached in the case of this ester. The comparison of X_2 , observed, with X_2 , calculated, for the ethyl ester using both the linear and quadratic forms, Eqs. 18a and b, is found in Table II. These results show the quadratic equation to yield slightly better solubility predictions. But the curves in Fig. 3 for the ethyl and hexyl esters of *p*-aminobenzoate show that differences between the linear and quadratic approaches are insignificant. Since only six data points are available, further statistical testing would not be fruitful.

Sulfadiazine in Water–Dimethylformamide Mixtures—The solubility of sulfadiazine in mixtures of water ($\delta_1 = 23.45$) and dimethylformamide ($\delta_1 = 12.14$) at 20° has been measured previously (7). The solubility parameter of sulfadiazine is 11.9, as calculated by the Fedors method (12). Yalkowsky *et al.* (1), referred to this study as one which should lend itself to a log linear analysis. The mole fraction solubility and grams of solute per gram of solution at different composition of cosolvent are shown in Fig. 4.

Although the solubility parameter of dimethylformamide is close to that of sulfadiazine, the points do not curve away from linearity as 100% dimethylformamide is approached. Similar results have been found with sulfonamides in dimethylacetamide. Dialkylamides strongly solvate the sulfonamides, and in a series of solvent mixtures, one of which is a dimethylamide, the curve rises to a sharply pointed apex rather than a rounded maximum. This phenomenon may account for the good fit provided by an exponential solubility curve, *i.e.*, a linear fit (Eq. 15a, Table I) on a plot of log solubility *versus* volume fraction. Mole fraction solubility is plotted against the solubility parameter, δ_1 , of dimethylformamide–water mixtures in Fig. 5, in contrast to the log mole fraction plot of sulfadiazine in Fig. 4. In addition to δ_1 , the volume fraction of dimethylformamide (f_{DMF}), in the water–cosolvent mixture is indicated on the horizontal axis (top line of Fig. 5). The extended Hildebrand solubility approach, Eq. 23, was used to plot a curve through the experimental points. The Hildebrand regular solution line, calculated with Eq. 13, is included in the figure to indicate its relationship to the experimental points and the extended Hildebrand line. The regular solution curve is flat rather than peaked as in Fig. 2 because of the scales used in the horizontal and vertical axes. The ideal mole fraction solubility, $X^i = 0.0030$, which corresponds to peak solubility on the regular solution line, is observed to be about one twentieth of the actual solubility of sulfadiazine in dimethylformamide at 20°. The regular solution line does not reproduce the solubility of sulfadiazine in water–dimethylformamide mixtures; whereas the extended Hildebrand solubility equation fits the experimental points well. The regression equation best employed for this purpose is a cubic expression:

$$\log \alpha_2/A_{\text{calc}} = -206.441 + 29.7528\delta_1 - 1.45599\delta_1^2 + 0.0258317\delta_1^3 \quad (\text{Eq. 23})$$

where $n = 14$, $R^2 = 0.9998$, $s = 0.1998$, $F = 18609$, and $F_{(3,10,0.05)} = 3.71$. The linear and quadratic forms, Eq. 19a and b in Table I, are adequate

Table III—Solubility of Sulfadiazine in Water–Dimethylformamide Mixtures at 20° a

f_{DMF}	δ_1 (cal/cm ³) ^{1/2}	V_1 , cm ³ /mole	A , cm ³ /cal	$\log \alpha_2$	($\log \alpha_{2,\text{calc}}$)/ A^b	$X_{2,\text{obs}} \times 10^4$	$X_{2,\text{calc}}^c \times 10^4$	Residuals, %
0	23.45	18.05	0.12448	2.96185	23.71491	0.0327	0.0334	-2.1
0.005	23.39	18.36	0.12448	2.92739	23.47131	0.0354	0.0359	-1.4
0.01	23.33	18.67	0.12448	2.89891	23.23027	0.0378	0.0384	-1.6
0.02	23.21	19.30	0.12448	2.83592	22.75577	0.0437	0.0440	-0.7
0.03	23.09	19.92	0.12448	2.77743	22.29114	0.0500	0.0503	-0.6
0.05	22.86	21.17	0.12448	2.66751	21.42721	0.0644	0.0644	-0.0
0.10	22.26	24.29	0.12447	2.35255	19.32585	0.1330	0.1177	11.5
0.20	21.07	30.54	0.12444	1.95135	15.69879	0.3350	0.3333	0.5
0.30	19.88	36.81	0.12440	1.57822	12.57256	0.7910	0.8173	-3.3
0.50	17.49	49.32	0.12415	0.86574	6.75256	4.080	4.346	-6.5
0.70	15.18	61.44	0.12252	0.00805	0.05759	29.40	29.468	-0.2
0.78	14.32	65.98	0.11971	-0.41290	-3.09509	77.50	70.293	9.3
0.89	13.18	71.93	0.11182	-0.88909	-8.08018	232.0	239.84	-3.4
1.00	12.14	77.40	0.09609	-1.30320	-13.60746	602.0	608.05	-1.0

^a Reference 7; back-calculated solubilities $X_{2,\text{calc}}$ obtained with the extended Hildebrand solubility approach; $\delta_2 = 11.9$ (cal/cm³)^{1/2}; $V_2 = 167$ cm³/mole, $\log X_2^i = -2.5236$.
^b Equation 23, (cubic expression). ^c ($\log \alpha_{2,\text{calc}}$)/ A (above) $\times A$ (Column 4 above) = $\log \alpha_{2,\text{calc}}$, then $\log \alpha_{2,\text{calc}} + 2.5236 = \log X_2$.

for this purpose but the cubic equation (Eq. 23) is better, as indicated by its R^2 , s , and F values. The δ_1 values for the solvent mixtures and the A values are found in Table III together with a comparison of actual solubilities, $X_{2,\text{obs}}$, and calculated values using the cubic expression, Eq. 23, together with Eq. 3 where $\log X^i = -2.5236$ for sulfadiazine at 20°.

The statistical analyses presented in Table I indicate that the log linear equation (Eq. 19a) provides a satisfactory fit of the data across the range of solvent mixtures from water to pure dimethylformamide. Figure 4 also shows that the linear fit is quite adequate. Therefore, the log linear expression is suitable for sulfadiazine in dimethylformamide–water as well as for ethyl and hexyl *p*-aminobenzoate in propylene glycol–water mixtures. However, the curve in Fig. 5, where X_2 is plotted against δ_1 (or f_{DMF}), is best fit by a third degree power series (Eq. 23). The residuals (Table III) are surprisingly small when the data is fit with this cubic equation.

Log Linear Equation and Solubilizing Power, σ —In addition to expressing log solubility versus volume fraction, Eq. 1, in molar terms (1), Yalkowsky *et al.* (13) used mole fraction concentration:

$$\log X_2 = \log X_w + \sigma f \quad (\text{Eq. 24})$$

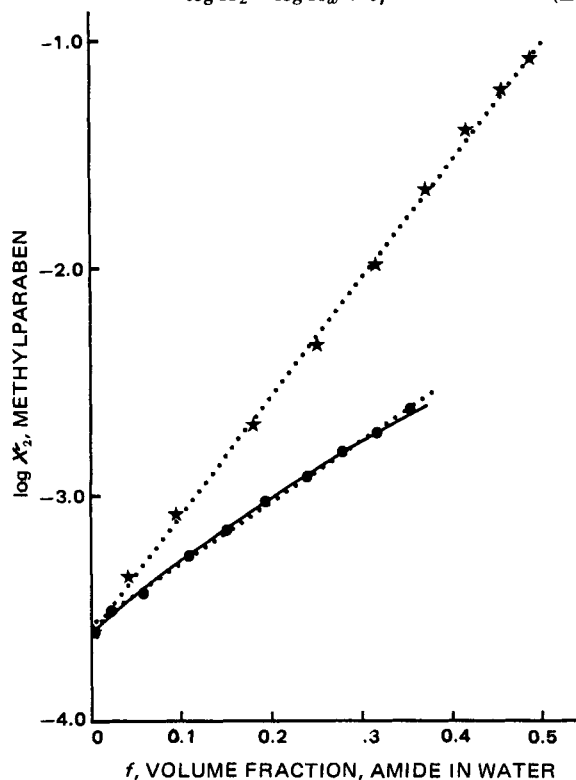


Figure 6—Log mole fraction solubility of methylparaben in formamide–water (●) and dimethylformamide–water (★) at 25° (8). Key: (—) from Eq. 16b (quadratic); (.....) from Eqs. 16a and 17a using linear regression.

X_2 is the mole fraction solubility of the solute at a volume fraction f of the cosolvent, X_w is the mole fraction solubility of the solute in water, and σ is a measure of the solubilizing power of the cosolvent for the drug (13). The term σ , which is the slope of the lines represented by Eq. 24, was defined for mixed solvent systems (12) as:

$$\sigma = \frac{C(\gamma_w - \gamma_c)\text{HYSA}}{2.303kT} \quad (\text{Eq. 25})$$

where γ_w and γ_c are macroscopic interfacial tensions between a model hydrophobic substance, tetradecane, and water, and between tetradecane and the pure cosolvent, respectively; C is a correction term for the macroscopic surface tensions, accounting for the small radius of curvature at the solute–solvent molecular surface; HYSA is the hydrophobic surface area of the semipolar solute as calculated by a computer program originally written by Hermann (14); k is the Boltzmann constant; and T is the absolute temperature. The value for C should remain relatively constant for various solute–mixed solvent systems, and the value was found to vary within a relatively narrow range, 0.52–0.56, for several systems (13). According to Eq. 25, the value of σ should not change for various compositions of water and cosolvent in a binary solvent mixture, σ being the constant slope of a linear expression, Eq. 24.

According to this interfacial model, it should be possible to calculate solubilities of drugs in polar solvent mixtures using either the linear form of the extended Hildebrand solubility approach:

$$\log X_2 = \log X^i - C_0A + \sigma f \quad (\text{Eq. 26})$$

or the log linear expression, Eq. 24, where $\sigma = C_1A$ is defined by Eq. 25.

The interfacial model was tested (8) by determining the solubility of methylparaben ($\delta_2 = 11$) in aqueous systems containing various concentrations of the cosolvents, formamide, methylformamide, and dimethylformamide. Contrary to the requirements of Eq. 25, the σ values were observed to vary in formamide–water from 2.77 to 4.46. For a number of other amide–water cosolvent systems, values for σ varied from 3.40 for acetamide to 6.43 for dimethylpropionamide.

Figure 6 shows the log mole fraction solubility of methylparaben in formamide–water and dimethylformamide–water cosolvent systems. The curve with formamide as the cosolvent is not linear but rather parabolic, and it is well fit by a quadratic expression, Eq. 16b, Table I, of the form:

$$\log X_2 = \log X_w + \sigma f + C_2f^2 \quad (\text{Eq. 27})$$

The regression leading to Eq. 16b gives a log mole fraction solubility of methylparaben in water of -3.6009, whereas the experimental value is -3.6091. This is an error of 1.9% in predicting the experimental solubility, $X_w = 0.000246$, in water, a close estimation using a quadratic expression.

The need for a quadratic equation accounts for a previous inability (8) to obtain a constant value for σ over this range ($f = 0.0$ – 0.35) of formamide compositions. Despite the fact that a linear fit of the points produced a regression line with an R^2 of 0.996, visual observation of the curve indicated that σ (Eq. 24) would not be constant over this composition. One would not predict, however, that σ would vary as much as actually found (2.8–4.5). Despite this variability, it is interesting to calculate an average σ , leaving C as an adjustable parameter as suggested previously (13).

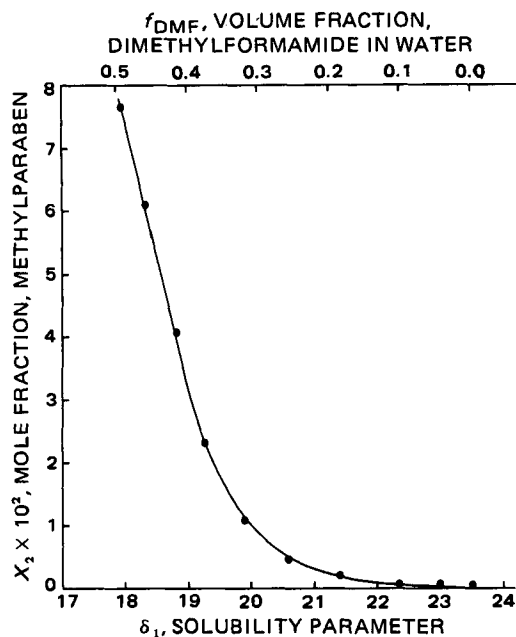


Figure 7—Mole fraction solubility of methylparaben in dimethylformamide–water at 25° (8). Key: (●) experimental solubility; (—) back calculation curve, Eq. 21b, extended Hildebrand approach.

The average value of σ in the formamide–water system, obtained from a least-squares linear fit of the curve in Fig. 6 is 2.734 (Eq. 16a). Referring to Eq. 25, $\gamma_w = 51.9$ for water–tetradecane and $\gamma_c = 31.2$ for formamide–tetradecane. The hydrophobic surface area of methylparaben is 130Å^2 or $130 \times 10^{-16} \text{cm}^2$. At 25°, $T = 298.15^\circ\text{K}$ and the Boltzmann constant equals $1.38 \times 10^{-16} \text{erg molecule}^{-1} \text{degree}^{-1}$. Using Eq. 25, one gets:

$$2.734 = \frac{C(51.9 \text{ erg/cm}^2 - 31.2 \text{ erg/cm}^2) 130 \times 10^{-16} \text{ cm}^2}{(2.303)(1.38 \times 10^{-16} \text{ erg/molecule deg})(298.15 \text{ deg})}$$

$$C = 0.9627 \approx 1.0$$

Yalkowsky *et al.*, accepted C values as large as 1.0, but not ordinarily for this kind of system. Considering the approximations, however, this is an acceptable fit.

In Fig. 6, one observes that for methylparaben in dimethylformamide–water, a plot of $\log X_2$ versus volume fraction, $f = 0.0$ to $f = 0.5$, of the cosolvent appears to be linear. Here, the R^2 for linear and quadratic functions (Table I, Eqs. 17a and b) differ little. The partial F values show that the quadratic term is unnecessary. The volume fraction of dimethylformamide has been carried only to $f = 0.5$ in previous work (8) as observed in Fig. 6. However, the line probably becomes markedly curved at higher volume fractions of cosolvent, and Eqs. 17a and b can no longer be used here. This can be observed by setting $f = 1.0$, *i.e.*, 100% methylformamide, in Eqs. 17a and b, resulting in X_2 values greater than unity, which is not possible. Therefore the solubility of methylparaben in dimethylformamide–water can be considered log linear only to $f = 0.5$.

Both the linear and quadratic equations (Eqs. 17a and b, respectively) have a constant, equivalent to $\log X_w$, of ~ -3.60 (-3.607 for linear and -3.592 for quadratic) or $X_w = 0.000251$. The slope of the line, identified as σ , is ~ 5.2 (5.24 for linear and 5.0 for quadratic). Using Eq. 25 to calculate C , with γ_c for dimethylformamide–tetradecane equal to 4.6, one obtains at 25° in the range of $f = 0.0$ – 0.5 :

$$5.2 = \frac{C(51.9 - 4.6)130 \times 10^{-16}}{(2.303)(1.38 \times 10^{-16})(298.15)}$$

$$C = (5.2/6.49) = 0.80$$

Although C is ordinarily considered to have a value of ~ 0.35 – 0.55 , a value of 0.8 is not unreasonable.

The solubility equation suggested by Yalkowsky *et al.* (13):

$$\log S = \log S_w + \frac{C(\gamma_w - \gamma_c)\text{HYSA}}{2.303kT} \times f \quad (\text{Eq. 28})$$

should therefore provide a satisfactory approach for calculating the solubility of relatively hydrophobic drugs in water–cosolvent systems. The interfacial treatment is predicated on this log linear relationship

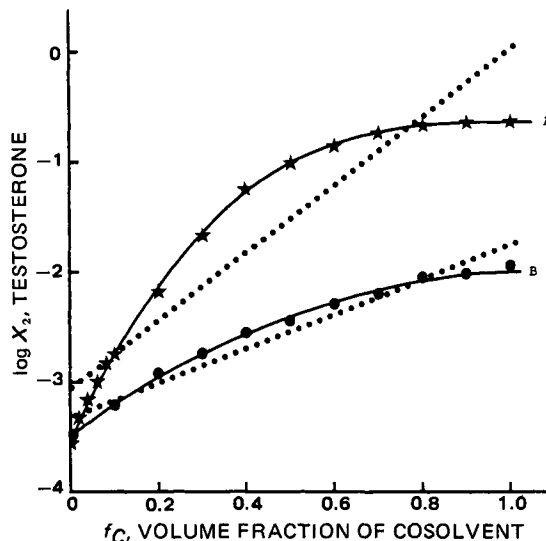


Figure 8—Log mole fraction solubility of testosterone in chloroform–cyclohexane (★); and in isopropyl myristate–cyclohexane (●) at 25°. Key: (---) linear regression line; (—) regression lines, (A) cubic, (B) quadratic.

between drug solubility and cosolvent volume fraction. If, however, the solubility parameter of the solute is similar to that of the cosolvent or is between the δ_1 value of the cosolvent and water, the points on the parabolic curve may be in a region of significant curvature; then the log linear equation and the interfacial model, involving σ , will not apply. On the other hand, the extended Hildebrand solubility approach handles drugs in polar and nonpolar solvent systems, whether the drug's solubility parameter is greater than, less than, or lies between the solubility parameters of a solvent pair, such as water and dimethylformamide. The following example demonstrates the use of the linear form of the extended Hildebrand solubility approach.

The solubility X_2 of methylparaben in dimethylformamide–water mixtures $\delta_1 = 18$ ($f = 0.5$) to $\delta_1 = 23.5$ ($f = 0.0$), from previous works (8) is shown in Fig. 7. The extended Hildebrand equation used to obtain the calculated line for these data may be linear, Eq. 21a, or quadratic, Eq. 21b of Table I. The statistical parameters show these equations to be equivalent. The A values range from 0.058 in pure dimethylformamide to 0.087 in water. The $-\log X^i = 1.0051$. At a volume fraction of dimethylformamide of 0.37, $\delta_1 = 19.28$ and $A = 0.0777$. Equation 21a yields $\log \alpha_2/A = 8.2894$ and Eq. 21b gives $\log \alpha_2/A = 8.4131$ for this mixture. When multiplied by A , the results from Eq. 21a gives $\log \alpha_2 = 0.6441$. This value is added to $-\log X^i$ to obtain 1.6492. Changing the sign and taking the antilog results in $X_{\text{methylparaben}} = 0.0224$. The observed value is 0.0221. The percent error in the back calculation is thus 1.5%.

Nonpolar Cosolvent Systems—In contrast to water–cosolvent systems, it is interesting to consider nonpolar solvent mixtures having solubility parameters lower than the solubility parameters of the solute. An investigation to be published (15) involves a study of the solubility of testosterone ($\delta_2 = 10.9$) and testosterone propionate ($\delta_2 = 9.5$) in cyclohexane ($\delta_1 = 8.2$) with cosolvents such as chloroform ($\delta_1 = 9.1$), ethyl oleate ($\delta_1 = 8.6$), and isopropyl myristate ($\delta_1 = 8.9$). It was found that the log linear relationship did not hold in these systems. Even chloroform, which strongly solvates the steroidal solutes, did not produce a log linear plot. As observed in Fig. 8, the lines are markedly curved when plotted as $\log X_2$ against the volume fraction of the cosolvent. Least-square straight lines are included, but statistical analysis is not needed to show that the experimental data yield nonlinear curves.

A report on benzoic acid in various proportions of hexane and ethyl acetate (16) was also studied for possible log linear characteristics. The solubility parameter of benzoic acid, 11.5, is greater than either hexane, 7.3, or ethyl acetate, 8.9. A plot of $\log X_2$ versus volume fraction of the cosolvent, ethyl acetate, in the solvent mixture showed considerable curvature. The line was well fitted by a quadratic equation of the extended Hildebrand approach but not by the log linear form.

Ternary Solvent Systems—In recent work (17), the solubility of four sulfonamides has been determined in a ternary solvent consisting of dimethylacetamide ($\delta_1 = 11.1$), glycerin ($\delta_1 = 17.7$), and water ($\delta_1 = 23.5$). Two compounds of lower solubility in dimethylacetamide, namely sulfadiazine ($\delta_2 = 13.2$) and sulfisomidine ($\delta_2 = 12.6$), produced nearly

straight lines when $\log X_2$ was plotted against the solubility parameter of the ternary solvent⁴. Two compounds of generally greater solubility, sulfathiazole ($\delta_2 = 13.1$) and sulfamethoxazole ($\delta_2 = 11.6$), produced highly curved lines which require quadratic or higher power series equations to fit the points. The solubility parameters of the drug molecules in these systems fall between the δ_1 values of the highly polar solvents, glycerin and water, and the strongly solvating aprotic solvent, dimethylacetamide. The curves show no peaks, but rather rise to a maximum mole fraction solubility in pure dimethylacetamide.

CONCLUSIONS

It may be concluded that the log linear solubility relationship, *i.e.*, log solubility *versus* volume fraction of cosolvent (Eq. 1) as introduced by Yalkowsky *et al.* (1), gives a good linear fit for semipolar drugs in a number of water-cosolvent mixtures, particularly when the solubility parameter of the solute is ~ 3 solubility parameter units lower than that of the cosolvent of a water-cosolvent mixture. When the cosolvent, such as dimethylformamide, is a strong solvating agent for the solute, sulfadiazine, a straight line function may persist up to 100% cosolvent even though the solubility parameters of drug and cosolvent are quite similar (< 2 or 3δ units apart).

For a system, caffeine, in an ordinary solvent pair, such as water and dioxane, or for theophylline in a moderately solvating mixture, such as water and polyethylene glycol 400 (5), the solubility parameter of the drug lies between the values of water and the cosolvent. Then, the log linear expression may be applied only over a limited range of solvent mixtures on either side of the solubility parameter of the solute. For example, the plot of log solubility *versus* volume fraction in a caffeine solution in water and dioxane was shown to be linear in dioxane from 0 to 30% and nearly linear in water from perhaps 0 to 10%. The log linear expression could not reproduce the solubility data for caffeine over the central range from 30 to 90% dioxane where the curve rises to an apex then falls as the solvent composition approaches pure dioxane.

In some systems, which might be predicted to follow the log linear relationship, a straight line function is not satisfactory but a higher order function, such as a quadratic or cubic in δ_1 or volume fraction provides a fit to the experimental data points. Systems of nonpolar solvents, *e.g.*, cyclohexane, in conjunction with cosolvents such as chloroform, isopropyl myristate, and ethyl acetate have not been found, in limited studies, to

yield log linear relationships. These solutions, like more polar ones, are satisfactorily reproduced by use of quadratic, cubic, or quartic equations employed in the extended Hildebrand solubility approach. Ternary solvent systems, both polar and nonpolar, are under study.

It was found possible to derive the log linear equation from the extended Hildebrand solubility approach, thus providing a quasitheoretical foundation for the log linear equation based on solubility parameters.

REFERENCES

- (1) S. H. Yalkowsky, G. L. Flynn, and G. L. Amidon, *J. Pharm. Sci.*, **61**, 983 (1972).
- (2) S. H. Yalkowsky and G. L. Flynn, *ibid.*, **63**, 1276 (1974).
- (3) A. Martin, J. Newburger, and A. Adjei, *ibid.*, **69**, 487 (1980).
- (4) A. Adjei, J. Newburger, and A. Martin, *ibid.*, **69**, 659 (1980).
- (5) A. Martin, A. N. Paruta, and A. Adjei, *ibid.*, **70**, 1115 (1981).
- (6) J. H. Hildebrand and R. L. Scott, "The Solubility of Nonelectrolytes," 3rd ed., Dover, New York, N.Y., 1964.
- (7) P. H. Elworthy and H. E. C. Worthington, *J. Pharm. Pharmacol.*, **20**, 830 (1968).
- (8) R. E. Lindstrom, *J. Pharm. Sci.*, **68**, 1141 (1979).
- (9) S. H. Yalkowsky, G. L. Amidon, G. Zograf, and G. L. Flynn, *ibid.*, **64**, 48 (1975).
- (10) N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, SPSS, Statistical Package for the Social Sciences, 2nd ed., McGraw-Hill, New York, N.Y., 1975.
- (11) S. H. Yalkowsky and T. J. Rosemann, in "Techniques of Solubilization of Drugs," chap. 3, S. H. Yalkowsky, Ed., Dekker, New York, N.Y., 1981.
- (12) R. F. Fedors, *Polymer. Eng. Sci.*, **14**, 147 (1974).
- (13) S. H. Yalkowsky, S. C. Valvani, and G. L. Amidon, *J. Pharm. Sci.*, **65**, 1488 (1976).
- (14) R. B. Hermann, *J. Phys. Chem.*, **76**, 2754 (1972).
- (15) A. Martin, P. L. Wu, A. Adjei, M. Mehdizadeh, K. C. James, and C. Metzler, *J. Pharm. Sci.*, in press.
- (16) M. J. Chertkoff and A. Martin, *J. Am. Pharm. Assoc., Sci. Ed.*, **49**, 444 (1960).
- (17) S. T. Valesquez-Leon, Thesis, University of Texas, May, 1981.

ACKNOWLEDGMENTS

Support for the study was obtained in part from the professorship provided to A. Martin by Coulter R. Sublett.

The authors thank Alan Beerbower Energy Center, University of California at San Diego, for advice and assistance.

⁴ Volume fraction is not appropriate as the independent variable in a ternary solvent system, but solubility parameter of the solvent mixture may be used instead.