

to solubility theory at this stage of development. The four partial solubility parameters for solvents, particularly the new acid (δ_a) and base (δ_b) terms, should find use in the pharmaceutical sciences as well as in industrial technology. The multiple regression and triangular plot methods of estimating partial solubility parameters for solutes are promising steps toward characterizing the polar and nonpolar properties of drugs and related biochemicals. The partial solubility parameters of Table I are tentative and no doubt will be changed somewhat as dictated by additional findings.

REFERENCES

- (1) A. Martin, P. L. Wu, A. Adjei, A. Beerbower, and J. M. Prausnitz, *J. Pharm. Sci.*, **70**, 1260 (1981).
- (2) A. Fredenslund, J. Gmehling, and P. Rasmussen, "Vapor-Liquid Equilibrium Using UNIFAC," Elsevier, New York, N.Y., 1977.
- (3) A. Martin, J. Newburger, and A. Adjei, *J. Pharm. Sci.*, **69**, 487 (1980).
- (4) J. H. Hildebrand and R. L. Scott, "The Solubility of Nonelectrolytes," 3rd ed., Dover, New York, N.Y., 1964.
- (5) C. M. Hansen and A. Beerbower, in "Encyclopedia of Chemical Technology," Suppl. Vol., 2nd ed., Wiley, New York, N.Y., 1971, pp. 889-910.
- (6) P. L. Wu, A. Beerbower, and A. Martin, *J. Pharm. Sci.*, **71**, 1285 (1982).
- (7) B. L. Karger, L. R. Snyder, and C. Eon, *J. Chromatogr.*, **125**, 71 (1962).
- (8) B. L. Karger, L. R. Snyder, and C. Eon, *Anal. Chem.*, **50**, 2126 (1978).
- (9) R. A. Keller, B. L. Karger, and L. R. Snyder, in "Gas Chromatography 1970," R. Stock, Ed., Institute of Petroleum, London, 1971, p. 125.
- (10) C. M. Hansen, *J. Paint Technol.*, **39**, 104 (1967).
- (11) C. M. Hansen, *J. Paint Technol.*, **39**, 505 (1967).
- (12) C. M. Hansen and K. Skaarup, *J. Paint Technol.*, **39**, 511

(1967).

- (13) M. J. Kamlet, J. L. M. Abboud, and R. W. Taft, "Progress in Physical Organic Chemistry," Wiley, New York, N.Y., 1981, pp. 485-630.
- (14) T. R. Griffiths and D. C. Pugh, *Coord. Chem. Rev.*, **29**, 129-211 (1979).
- (15) M. J. Chertkoff and A. Martin, *J. Am. Pharm. Assoc., Sci. Ed.*, **49**, 444 (1960).
- (16) J. A. Riddick and W. B. Bunger, "Organic Solvents," Techniques of Chemistry, Vol. II, 3rd ed., Wiley, New York, N.Y., 1970.
- (17) J. H. Hildebrand, J. M. Prausnitz, and R. L. Scott, "Regular and Related Solutions," Van Nostrand Reinhold, New York, N.Y., 1970, p. 22.
- (18) D. C. Ginnings and G. T. Furukawa, *J. Am. Chem. Soc.*, **75**, 522 (1953).
- (19) "United States Pharmacopeia," 20th rev., U.S. Pharmacopeial Convention, Rockville, Md., 1980, p. 929.
- (20) A. Martin, J. Swarbrick, and A. Cammarata, "Physical Pharmacy," 3rd ed., Lea and Febiger, Philadelphia, Pa., 1983, p. 285.
- (21) T. Krishnan, W. C. Duer, S. Goldman, and J. L. Fortier, *Can. J. Chem.*, **57**, 530 (1979).
- (22) G. Allen, J. G. Watkinson, and K. H. Webb, *Spectrochim. Acta*, **22**, 807 (1966).
- (23) IMSL Reference Manual, International Mathematical and Statistical Library, Houston, Texas, 1979.
- (24) J. P. Teas, *J. Paint Technol.*, **40**, 19 (1968).
- (25) A. Martin, P. L. Wu, and A. Beerbower, *J. Pharm. Sci.*, **73**, 188 (1984).
- (26) A. Martin and J. Carstensen, *J. Pharm. Sci.*, **70**, 170 (1981).
- (27) L. R. Snyder, *Chemtech*, **March**, 188 (1980).

ACKNOWLEDGMENTS

Supported in part by a grant from the University of Texas Research Institute. A. Martin is grateful to C. R. Sublett for the endowed professorship which also supported this work.

Expanded Solubility Parameter Approach II: *p*-Hydroxybenzoic Acid and Methyl *p*-Hydroxybenzoate in Individual Solvents

A. MARTIN ^{*x}, P. L. WU ^{*§}, and A. BEERBOWER [‡]

Received September 13, 1982, from the ^{*}Drug Dynamics Institute, College of Pharmacy, University of Texas, Austin, TX 78712 and the [†]Energy Center, University of California at San Diego, La Jolla, CA 92093. Accepted for publication November 17, 1982. [§]Present address: China Chemical & Pharmaceutical Co., Ltd., Taipei, Taiwan.

Abstract □ The recently introduced four-parameter extended Hansen approach was used to study the solubility of *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate in 32 and 35 individual solvents, respectively. The results are compared with those for benzoic acid in 40 solvents. Seventeen of the thirty-two or 53% of the calculated solubilities of *p*-hydroxybenzoic acid were within the established solubility criterion (*i.e.*, <30% error from the experimental value). Twenty-two of thirty-six or 61% of the calculated solubility values for methyl *p*-hydroxybenzoate met the solubility criterion. Experimental excess free energies of solution for *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate were plotted against theoretical values using the expanded four-parameter solubility regression equations. From such results, adjustments may be made in the partial solubility parameters to bring the calculated solubilities in line with experimental results. Multiple regression analyses were used to estimate the total solubility parameter and the four partial solubility parameters of the two benzoic acid derivatives. Satisfactory results were obtained for methyl *p*-hydroxybenzoate, but poor agreement was found for *p*-hydroxybenzoic acid for the total parameter when compared with

the Fedors group contribution method. Both the multiple regression and group contribution methods may yield inaccurate solubility parameters for relatively polar solid solutes. Factor analysis was used to test the adequacy of three- and four-parameter approaches in the evaluation of drug solubility. A principal factor method without iteration and orthogonal factor rotation were used to compare the two expanded solubility parameter approaches. Factor analysis showed that four solubility parameters were significantly more independent and presumably more satisfactory than three parameters.

Keyphrases □ Solubility parameters, expanded—individual solvents, four-parameter extended Hansen approach, *p*-hydroxybenzoic acid, methyl *p*-hydroxybenzoate □ Extended Hansen approach—solubility of *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate in individual solvents, four-parameter system □ Benzoic acid derivatives—*p*-hydroxybenzoic acid, methyl *p*-hydroxybenzoate, use as model drugs, solubility in individual solvents, four-parameter extended Hansen approach

The solubility of a number of drugs in mixtures of two solvents has been analyzed with the extended Hildebrand solubility approach (1, 2). Subsequently, the solubility of

naphthalene and benzoic acid in individual solvents was studied using the Hansen three-parameter solubility approach (3, 4) and a new four-parameter solubility system

Table I—Calculated Solubilities of Methyl *p*-Hydroxybenzoate^a in Individual Solvents at 25°C

No.	Solvent	V ₁	δ _d	δ _p	δ _a	δ _b	Equation 7			
							X ₂ (Obs.)	X ₂ (Calc.)	Residual	Error, %
1	Pentane	116.1	7.1	0.0	0.0	0.0	0.000063	0.000009	0.000054	85.7
2	Hexane	131.6	7.3	0.0	0.0	0.0	0.000077	0.000028	0.000049	63.6
3	Heptane	147.5	7.5	0.0	0.0	0.0	0.000083	0.000074	0.000009	10.8
4	Nonane	179.7	7.7	0.0	0.0	0.0	0.000106	0.000176	-0.000070	-66.0
5	Decane	195.9	7.7	0.0	0.0	0.0	0.000121	0.000176	-0.000055	-45.5
16	Diethyl ether	104.8	7.1	1.4	0.5	6.3	0.0840	0.0749	0.0091	10.8
17	Dipropyl ether	139.4	7.3	1.1	0.4	5.0	0.0314	0.0409	-0.0095	-30.3
18	Dibutyl ether	170.4	7.6	0.8	0.3	3.9	0.0268	0.0326	-0.0058	-21.6
19	Ethyl acetate	98.5	7.4	2.6	5.3	1.9	0.1270	0.0862	0.0408	32.1
20	Propyl acetate	115.7	7.7	2.1	3.6	1.5	0.1366	0.1639	-0.0273	-20.0
21	Butyl acetate	132.6	7.7	1.8	2.8	1.7	0.1326	0.0828	0.0538	39.4
22	Hexyl acetate	164.5	7.8	1.5	1.9	2.2	0.1164	0.0848	0.0316	27.1
31	Methyl alcohol	40.7	7.4	6.0	8.4	7.1	0.1254	0.1155	0.0099	7.9
32	Ethyl alcohol	58.7	7.7	4.3	8.3	5.5	0.1495	0.1745	-0.0250	-16.7
33	Propyl alcohol	75.1	7.8	3.3	7.5	4.8	0.1486	0.1745	-0.0259	-17.4
35	Butyl alcohol	92.0	7.8	2.8	6.4	4.6	0.1484	0.1743	-0.0259	-17.5
39	1-Pentanol	108.6	7.8	2.2	5.4	4.3	0.1528	0.1284	0.0244	20.0
40	1-Hexanol	125.2	8.0	2.1	5.7	3.5	0.1477	0.1694	-0.0217	-14.7
41	1-Heptanol	141.9	8.1	2.0	5.3	3.4	0.1483	0.2013	-0.0530	-35.7
42	1-Octanol	158.4	8.3	1.6	5.2	3.2	0.1381	0.1605	-0.0224	-16.2
43	Ethylene Glycol	55.9	8.3	5.4	17.9	4.4	0.0480	0.0304	0.0176	36.7
44	1,2-Propanediol	73.7	8.2	4.6	14.1	4.6	0.0941	0.0818	0.0123	13.1
45	1,3-Propanediol	72.5	8.1	5.3	10.9	7.4	0.0766	0.0560	0.0206	26.9
46	Glycerol	73.2	8.5	5.9	20.0	5.1	0.0064	0.0021	0.0043	67.2
47	1,4-Butanediol	88.6	8.2	8.1	18.2	3.7	0.1202	0.1438	-0.0236	-19.6
48	Acetic acid	57.6	7.1	3.9	7.0	3.1	0.0532	0.0820	-0.0288	-54.1
49	Propionic acid	75.0	7.2	3.8	6.0	3.0	0.0386	0.1436	-0.1050	-272.0
51	Dimethyl sulfoxide	71.3	9.0	8.0	2.2	5.7	0.5839	0.4965	0.0874	15.0
52	Pyridine	80.9	9.3	4.3	1.4	3.2	0.3243	0.4752	-0.1509	-46.5
53	Formamide	39.9	8.4	12.8	5.7	7.6	0.0765	0.1118	-0.0353	-46.1
54	<i>N</i> -Methylformamide	59.1	8.4	10.1	4.8	3.9	0.2981	0.3770	-0.0789	-26.5
55	<i>N,N</i> -Dimethylformamide	77.4	8.5	6.7	3.4	4.4	0.4605	0.4836	-0.0231	-5.0
56	<i>N,N</i> -Diethylformamide	112.0	8.2	5.6	2.7	3.4	0.4907	0.5346	-0.0439	-8.9
57	<i>N,N</i> -Dimethylacetamide	93.0	8.2	5.6	2.9	4.3	0.5418	0.5020	0.0398	7.3
58	<i>N,N</i> -Diethylacetamide	126.6	8.2	4.1	2.1	3.2	0.5299	0.5223	0.0076	1.4
59	Water	18.1	7.6	7.8	6.7	32.0	0.00027	0.00041	-0.00014	-51.9

^a Solute properties: δ₂ = 12.1, mp 399.65 K, V₂ = 117.5 cm³/mol, X₂¹ = 0.0988, log X₂¹ = -1.005.

(4). The present study is an attempt to use the four-parameter method to reproduce the solubilities of two solid organic compounds, *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate, which serve as models of drug molecules.

The substituted aromatic groups and flexible chains of drug molecules, exhibiting dipolar forces, hydrogen bonding, steric interferences, and ionic charges, make the prediction of solubility in drug research and development a formidable task. The results obtained with benzoic acid in the preceding paper (4) demonstrate the most satisfactory results that have been obtained in this laboratory for drug-like molecules by regression analysis and by the universal functional group activity coefficient (UNIFAC) procedure (5). The correlation found here between experimental data and a four-parameter equation can serve as a benchmark—a goal to be superseded in predicting the solubility of drugs in general. As in the previous paper (4) all solvents are “neat” or individual. The methods used here and the conclusions reached are not generally applicable to the binary solvent systems discussed earlier (1, 2).

EXPERIMENTAL

Methyl *p*-hydroxybenzoate USP¹ was used as obtained. *p*-Hydroxybenzoic acid² was recrystallized from aqueous alcohol (6) and dried at

105°C. Melting points measured by a hot-stage method³ are 485.85°K and 399.65°K for *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate, respectively. Heats of fusions of *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate obtained by differential scanning calorimetry⁴ are 7510 and 5400 cal/mol (7), respectively. The solvents were spectrophotometric grade, ACS grade or redistilled before use. The mole fraction solubilities X₂ of *p*-hydroxybenzoic acid and methyl *p*-hydroxybenzoate at 25°C were determined as described in the preceding paper (4).

RESULTS AND DISCUSSION

The solubilities of methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid at 25°C are found in Tables I and II and in Fig. 1. The properties of the solutes are included in the footnotes of Tables I and II; solvent molar volumes and partial solubility parameters are also listed. Sources and methods of obtaining partial solubility parameters, δ_d, δ_p, δ_a, and δ_b, were described in the preceding paper (4). Briefly, δ_d and δ_p are the dispersion and dipolar solubility parameters of Hansen and Beerbower (8); δ_b, a basic parameter, was obtained using the expression:

$$V_1^{1/2}\delta_b = 60.1\beta + 2.45 \quad (\text{Eq. 1})$$

where V₁ is the molar volume of the solvent and β is a basicity parameter provided by Kamlet *et al.* (9). Since δ_b² = 2δ_aδ_b, it is possible to obtain δ_a, an acid parameter, once δ_b is calculated from Eq. 1. The term, δ_b, is called a hydrogen bonding parameter (4) though it includes all transfer energies.

As in earlier papers of this series, the mole fraction solubilities (X₂) of the two drug prototypes, methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid, are plotted (Fig. 1) as a function of the total solvent solubility parameter, δ₁. Shown in the graphs are the ideal mole fraction solubility line (X₂ⁱ) for the solute and the regular solution curve, calculated using the Hildebrand equation (1). The observed solubilities are

¹ Tenneco Chemical Inc., Piscataway, NJ 08854.

² Matheson, Coleman and Bell, Norwood, OH 45212.

³ Digital Melting Point Analyzer, Model 335, Fisher Scientific Co.

⁴ Perkin-Elmer DSC Model 1B, Norwalk, CT.

Table II—Calculated Solubilities of *p*-Hydroxybenzoic Acid^a in Individual Solvents at 25°C

No.	Solvents	V ₁	δ _d	δ _p	δ _a	δ _b	X ₂ (Obs.)	Equation 8		
								X ₂ (Calc.)	Residual	Error, %
7	Benzene	89.4	9.0	0.5	0.7	0.7	0.000033	0.000012	0.000021	63.6
8	Toluene	106.9	8.8	0.7	0.8	0.6	0.000029	0.000027	0.000002	6.9
10	Ethylene dichloride	79.4	9.3	3.6	2.0	1.0	0.00011	0.00534	-0.00523	-4755.
12	Chloroform	80.8	8.7	1.5	3.0	1.3	0.00015	0.00322	-0.00307	-2047.
16	Diethyl ether	104.8	7.1	1.4	0.5	6.3	0.0521	0.1132	-0.0611	-117.
19	Ethyl acetate	98.5	7.4	2.6	5.3	1.9	0.0737	0.0438	0.0299	40.6
21	Butyl acetate	132.6	7.7	1.8	2.8	1.7	0.0574	0.0048	0.0526	91.6
24	Acetone	74.0	7.6	5.1	2.4	2.4	0.1185	0.1186	-0.0001	-0.1
25	Dioxane	85.7	9.3	0.9	1.0	6.5	0.0844	0.0939	-0.0095	-11.3
28	Acetophenone	117.4	9.6	4.2	1.1	1.5	0.0223	0.0029	0.0194	87.0
29	Benzyl alcohol	103.9	9.0	3.1	5.9	3.8	0.0784	0.1241	-0.0457	-58.3
31	Methyl alcohol	40.7	7.4	6.0	8.4	7.1	0.1142	0.0590	0.0552	48.3
32	Ethyl alcohol	58.7	7.7	4.3	8.3	5.5	0.1213	0.1077	0.0136	11.2
33	Propyl alcohol	75.1	7.8	3.3	7.5	4.8	0.1084	0.1156	-0.0072	-0.6
34	Isopropyl alcohol	76.9	7.7	3.0	7.1	4.5	0.1297	0.1011	0.0286	22.1
35	Butyl alcohol	92.0	7.8	2.8	6.4	4.6	0.1154	0.1094	0.0060	5.2
36	Isobutyl alcohol	92.4	7.4	2.8	6.0	5.1	0.0901	0.0660	0.0241	26.7
39	1-Pentanol	108.6	7.8	2.2	5.4	4.3	0.1145	0.0767	0.0378	33.0
40	1-Hexanol	125.2	8.0	2.1	5.7	3.5	0.1121	0.0815	0.0306	27.3
42	1-Octanol	158.4	8.3	1.6	5.2	3.2	0.1032	0.0445	0.0587	56.9
43	Ethylene glycol	55.9	8.3	5.4	17.9	4.4	0.1132	0.1343	-0.0211	-18.6
44	1,2-Propanediol	73.7	8.2	4.6	14.1	4.6	0.1308	0.1318	-0.0010	-0.8
46	Glycerol	73.2	8.5	5.9	20.0	5.1	0.0301	0.0225	0.0076	25.2
48	Acetic acid	57.6	7.1	3.9	7.0	3.1	0.0444	0.0826	-0.0382	-86.0
49	Propionic acid	75.0	7.2	3.8	6.0	3.0	0.0347	0.0919	-0.0572	-164.8
51	Dimethyl sulfoxide	71.3	9.0	8.0	2.2	5.7	0.3674	0.2769	0.0905	24.6
52	Pyridine	80.9	9.3	4.3	1.4	3.2	0.1044	0.0867	0.0177	17.0
53	Formamide	39.9	8.4	12.8	5.7	7.6	0.0341	0.0331	0.0010	2.9
54	<i>N</i> -Methylformamide	59.1	8.4	10.1	4.8	3.9	0.1025	0.1647	-0.0622	-60.7
55	<i>N,N</i> -Dimethylformamide	77.4	8.5	6.7	3.4	4.4	0.2137	0.2625	-0.0488	-22.8
57	<i>N,N</i> -Dimethylacetamide	93.0	8.2	5.6	2.9	4.3	0.2354	0.2782	-0.0428	-18.2
59	Water, pH 2.0	18.1	7.6	7.8	6.7	32.0	0.00060	0.00104	-0.00044	-73.3

^a Solute properties: δ₂ = 15.3, mp 485.85 K, V₂ = 94.3 cm³/mol, X₂ⁱ = 0.00747, log X₂ⁱ = -2.1267.

plotted as filled circles and the predicted solubilities, obtained using Eq. 2 (below), are plotted as stars.

In the prior study (4), the solubilities of naphthalene and benzoic acid were found to lie on or near the regular solution line. In Fig. 1A most of the experimental points for methyl *p*-hydroxybenzoate fall near X₂ⁱ and the regular solution line, but points 51, 52, and 54-58 (for the solute in strongly dipolar solvents) lie far above X₂ⁱ and the regular solution line. In Fig. 1B, no relationship exists between *p*-hydroxybenzoic acid solubility points and the dashed regular solution line. *p*-Hydroxybenzoic acid is a relatively polar compound (δ₂ = 15.3), and its solubility profile cannot be approximated by regular solution theory. However, it is possible with a four-parameter system to predict the solubilities of such a polar solute

in polar (and nonpolar) solvents within a reasonable degree of accuracy (4). The results obtained here are not highly satisfactory, but adjustment of parameters and consideration of average molecular orientations in specific solute-solvent interactions may lead to improvements.

Estimation of Solubilities Using Four Partial Solubility Parameters—The solubility of a crystalline solid in a number of solvents can be expressed as (4):

$$\frac{\log \alpha_2}{A} = C_1(\delta_{1d} - \delta_{2d})^2 + C_2(\delta_{1p} - \delta_{2p})^2 + 2C_3(\delta_{1a} - \delta_{2a})(\delta_{1b} - \delta_{2b}) + C_0 \quad (\text{Eq. 2})$$

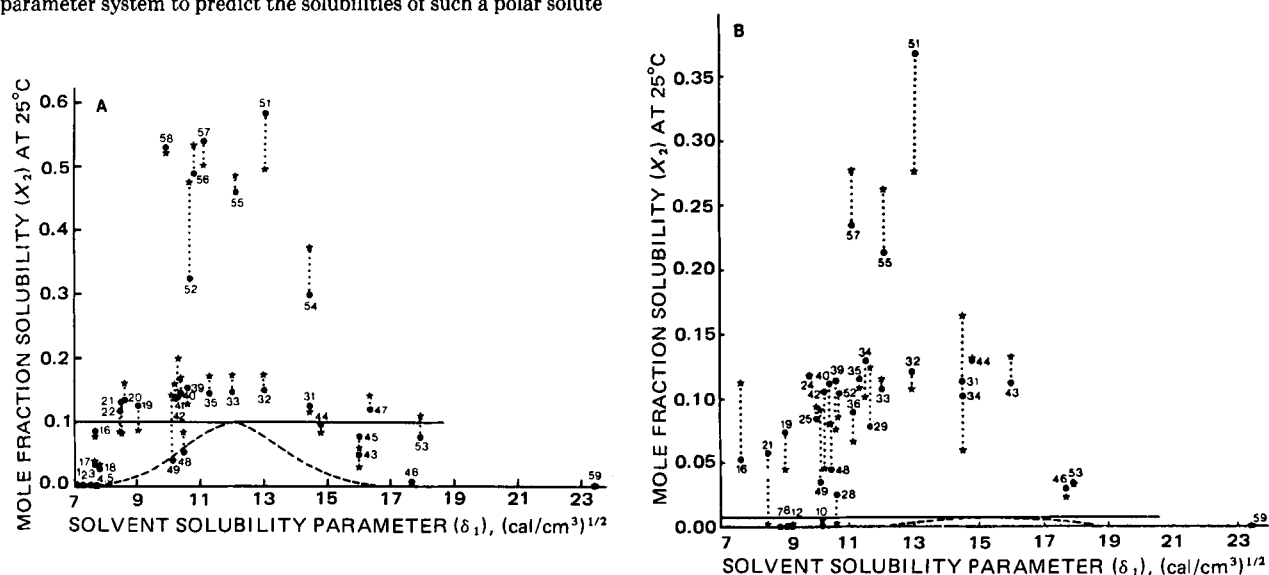


Figure 1—Solubility profile of methyl *p*-hydroxybenzoate (A) and *p*-hydroxybenzoic acid (B) in individual solvents at 25°C. See Tables I-III for solvent numbers. Key: (---) regular solution curve; (—) ideal solubility line, X₂ⁱ = 0.0988 (A) and 0.00747 (B) (●) experimental solubilities; (★) solubilities calculated using Eq. 7(A) or Eq. 8(B).

Table III—Excess Free Energies of Benzoic Acid and Its Derivatives In Individual Solvents at 25°C

No.	Solvent	Benzoic Acid			Methyl <i>p</i> -Hydroxybenzoate			<i>p</i> -Hydroxybenzoic Acid		
		$\overline{\Delta G}_{exp}^E$, cal/mol	$\overline{\Delta G}_{calc}^E$, cal/mol	Residual, cal/mol	$\overline{\Delta G}_{exp}^E$, cal/mol	$\overline{\Delta G}_{calc}^E$, cal/mol	Residual, cal/mol	$\overline{\Delta G}_{exp}^E$, cal/mol	$\overline{\Delta G}_{calc}^E$, cal/mol	Residual, cal/mol
1	Pentane	2157	1992	165	4359	5512	-1153	—	—	—
2	Hexane	1875	1863	12	4240	4839	-599	—	—	—
3	Heptane	1752	1747	5	4196	4264	-68	—	—	—
4	Nonane	1641	1654	-13	4051	3750	301	—	—	—
5	Decane	1589	1654	-65	3972	3750	222	—	—	—
6	Cyclohexane	1833	1461	372	—	—	—	—	—	—
7	Benzene	701	1397	-696	—	—	—	3212	3812	-600
8	Toluene	664	877	-213	—	—	—	3289	3331	-42
9	Chlorobenzene	569	642	-73	—	—	—	—	—	—
10	Ethylene dichloride	696	175	521	—	—	—	2499	199	2300
12	Chloroform	330	226	104	—	—	—	2335	499	1836
13	Carbon tetrachloride	898	1093	-195	—	—	—	—	—	—
16	Diethyl ether	120	-86	206	96	164	-68	-1150	-1610	460
17	Dipropyl ether	—	—	—	679	522	157	—	—	—
18	Dibutyl ether	—	—	—	773	657	116	—	—	—
19	Ethyl acetate	184	171	13	-149	81	-230	-1356	-1048	-308
20	Propyl acetate	—	—	—	-192	-300	108	—	—	—
21	Butyl acetate	167	387	-220	-174	105	-279	-1208	262	-1470
22	Hexyl acetate	—	—	—	-97	91	-188	—	—	—
24	Acetone	114	38	76	—	—	—	-1637	-1638	1
25	Dioxane	-140	-194	54	—	—	—	-1436	-1500	64
27	Nitrobenzene	545	311	234	—	—	—	—	—	—
28	Acetophenone	107	250	-143	—	—	—	-648	561	-1209
29	Benzyl alcohol	264	-45	309	—	—	—	-1393	-1665	272
31	Methyl alcohol	191	630	-439	-141	-93	-48	-1616	-1224	-392
32	Ethyl alcohol	136	276	-140	-245	-337	92	-1651	-1581	-70
33	Propyl alcohol	135	170	-35	-242	-337	95	-1585	-1623	38
34	Isopropyl alcohol	89	168	-79	—	—	—	-1691	-1543	-148
35	Butyl alcohol	65	137	-72	-241	-336	95	-1622	-1590	-32
36	Isobutyl alcohol	231	292	-61	—	—	—	-1475	-1291	-184
39	1-Pentanol	210	142	68	-258	-155	-103	-1617	-1380	-237
40	1-Hexanol	99	91	8	-238	-319	81	-1605	-1416	-189
41	1-Heptanol	—	—	—	-241	-422	181	—	—	—
42	1-Octanol	74	102	-28	-198	-287	89	-1556	-1057	-499
43	Ethylene glycol	554	494	60	428	698	-270	-1610	-1712	102
44	1,2-Propanediol	—	—	—	29	112	-83	-1696	-1700	4
45	1,3-Propanediol	—	—	—	151	336	-185	—	—	—
46	Glycerol	1552	1440	112	1621	2282	-661	-826	-653	-173
47	1,4-Butanediol	—	—	—	-116	-222	106	—	—	—
48	Acetic acid	175	164	11	367	110	257	-1056	-1424	368
49	Propionic acid	104	148	44	557	-222	779	-910	-1487	577
51	Dimethyl sulfoxide	-485	-465	-20	-1053	-956	-97	-2308	-2140	-168
52	Pyridine	-513	-460	-53	-704	-930	226	-1562	-1452	-110
53	Formamide	231	714	-483	152	-73	225	-900	-882	-18
54	<i>N</i> -Methylformamide	-249	-379	130	-162	-793	631	-1552	-1833	281
55	<i>N,N</i> -Dimethylformamide	-462	-468	6	-912	-941	29	-1987	-2109	122
56	<i>N,N</i> -Diethylformamide	—	—	—	-949	-1000	51	—	—	—
57	<i>N,N</i> -Dimethylacetamide	-501	-502	1	-1008	-963	-45	-2044	-2143	99
58	<i>N,N</i> -Diethylacetamide	—	—	—	-995	-986	-9	—	—	—
59	Water	3656	3420	236	3497	3249	248	1494	1168	326

where δ_d , δ_p , δ_a , and δ_b have previously been defined and C_0 , C_1 , C_2 , and C_3 are the coefficients obtained as a result of multiple regression analysis. $\log \alpha_2$ is the logarithm of the solute activity coefficient, where:

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

and

$$A = \frac{V_2 \phi_1^2}{2.303RT} \quad (\text{Eq. 4})$$

where V_2 is the molar volume of the solute, considered as a hypothetical supercooled liquid at the temperature of the solution. The molar volume of methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid are 117.5 and 94.3 cm³/mol, respectively, which are obtained as the average of apparent molar volume in the solutions (10) and by the group contribution method of Fedors (11). The solvent volume fraction, ϕ_1 , and R and T have their usual meanings.

The ideal solubility, X_2^i , of the solute can be approximated by:

$$\log X_2^i = \frac{\Delta H_m^f}{2.303R} \left(\frac{1}{T_m} - \frac{1}{T} \right) \quad (\text{Eq. 5})$$

where ΔH_m^f is the heat of fusion of the solute, R is the gas constant, T_m and T are melting points of the solute and solution temperatures, respectively, in degrees Kelvin.

Once A is calculated, one can use the regression equation (Eq. 2) to

calculate α_2 , and knowing X_2^i for the solute, X_2 is obtained from Eq. 3. To obtain the A value, however, requires a value of the volume fraction of the solvent, ϕ_1 ; but, ϕ_1 depends on the value of X_2 , the quantity sought. This difficulty can be overcome by simple iteration (3), or in a more satisfactory way by use of a root-finding method, as provided in the International Mathematical and Statistical Library (IMSL) (12) and described in the preceding paper (4).

The solubility of a solute, such as methyl *p*-hydroxybenzoate, in a solvent with known parameters can be calculated by expanding Eq. 2, recognizing that the partial solubility parameters for a particular solute are constants. Then:

$$\frac{\log \alpha_2}{A} = D_0 + D_1 \delta_{1d}^2 + D_2 \delta_{1d} + D_3 \delta_{1p}^2 + D_4 \delta_{1p} + D_5 \delta_{1a} \delta_{1b} + D_6 \delta_{1a} + D_7 \delta_{1b} \quad (\text{Eq. 6})$$

where D_0 is a constant of regression and D_1 - D_7 are regression coefficients associated with the partial solubility parameters. The regression equation was obtained by use of a multiple regression program (13) for methyl *p*-hydroxybenzoate in 36 solvents at 25°C and was found to be:

$$\frac{\log \alpha_2}{A} = 7.341 \delta_{1d}^2 - 133.4 \delta_{1d} + 1.138 \delta_{1p}^2 - 16.57 \delta_{1p} + 0.9032 \delta_{1a} \delta_{1b} - 0.6605 \delta_{1a} - 4.231 \delta_{1b} + 623.85$$

$$n = 36, s = 10.61, r^2 = 0.87, F = 27.3, F(7,28,0.01) = 3.36 \quad (\text{Eq. 7})$$

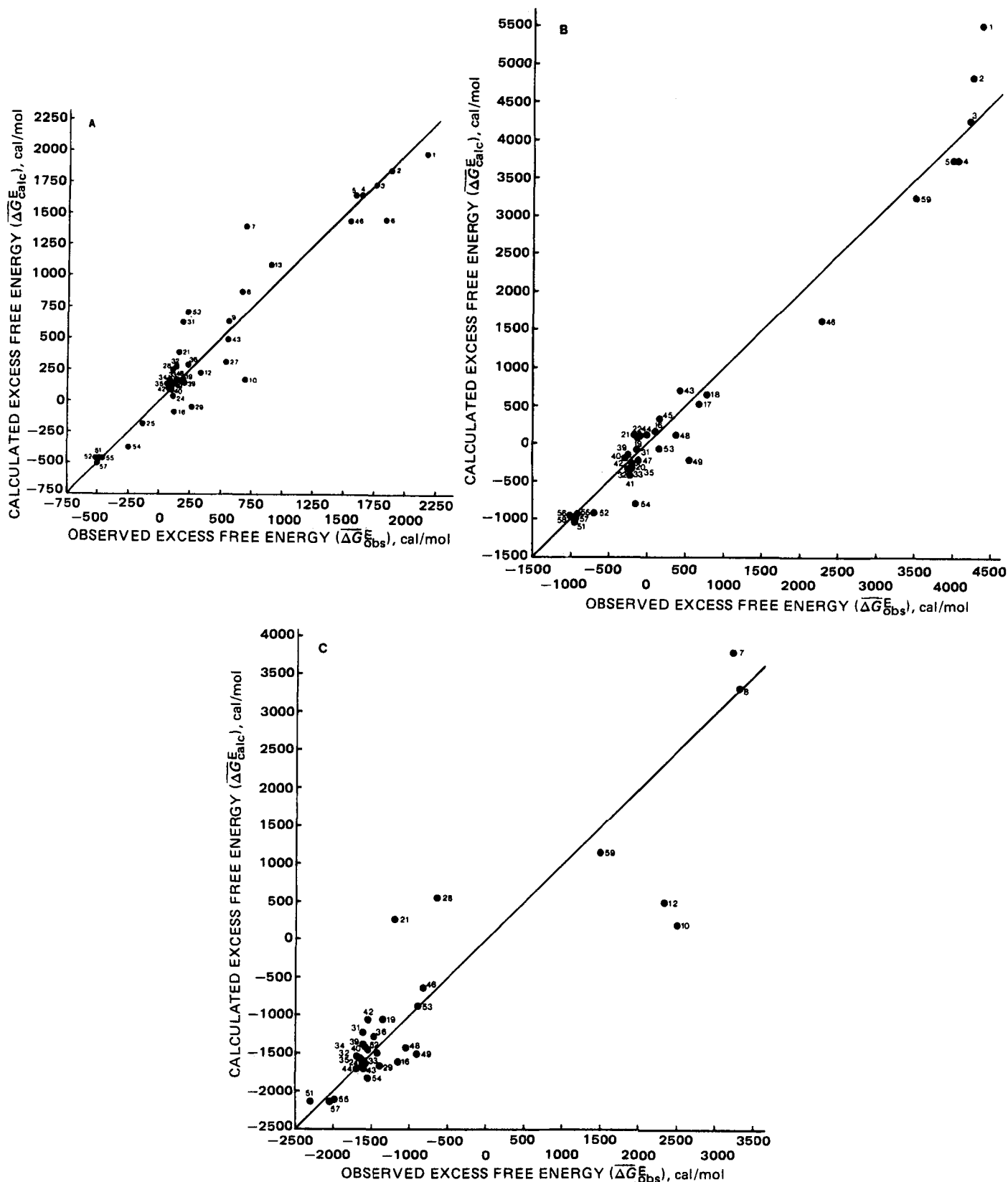


Figure 2—Relationship between excess free energies of mixing of benzoic acid (A), methyl *p*-hydroxybenzoate (B), and *p*-hydroxybenzoic acid (C) derived from measured (Eq. 9) and calculated (Eq. 10 for benzoic acid) solubilities at 25°C. The line is drawn with unit slope and zero intercept. See Tables I–III for solvent numbers.

The estimated solubilities of methyl *p*-hydroxybenzoate obtained from Eq. 7 are listed in Table I. The $(\log \alpha_2)/A$ values fall within a standard deviation of ± 10.61 . The coefficient of determination, r^2 , which gives the fraction of the variance in the dependent variable that is explained by

the regression equation, is 0.87. Therefore 87% of the solubility results (variances) are predicted by Eq. 7. As observed, the calculated F ratio (27.3) is significantly larger than the table value (3.36).

All statistical parameters, therefore, indicate that the relationship is

Table IV—Correlation Coefficients of Solubility Parameters^a

Three-Parameter System				
	δ_d	δ_p	δ_h	
δ_d	1.000			
δ_p	0.076	1.000		
δ_h	-0.183	0.623	1.000	
Four-Parameter System				
	δ_d	δ_p	δ_a	δ_b
δ_d	1.000			
δ_p	0.076	1.000		
δ_a	-0.083	0.439	1.000	
δ_b	-0.143	0.468	0.252	1.000

^a $n = 59$ for each system.

significant. Yet the most meaningful test is how close the calculated solubilities approach observed values. In previous work (3) an error of $\leq 30\%$ from observed solubility was accepted as reasonable. A study of Table I shows that 22 of 36 or 61% of the predicted solubility values meet this criterion. The results are plotted in Fig. 1A where the calculated solubilities for methyl *p*-hydroxybenzoate (stars) are attached by dotted lines to the experimental points (filled circles).

Using the four-parameter system, the regression equation obtained for *p*-hydroxybenzoic acid at 25°C in 32 solvents was:

$$\frac{\log \alpha_2}{A} = 10.65\delta_{1d}^2 - 177.0\delta_{1d} + 1.040\delta_{1p}^2 - 14.73\delta_{1p} + 2.072\delta_{1a}\delta_{1b} - 8.107\delta_{1a} - 12.08\delta_{1b} + 790.82$$

$$n = 32, s = 10.93, r^2 = 0.81, F = 14.6, F(7,24,0.01) = 3.50 \quad (\text{Eq. 8})$$

The calculated solubilities for *p*-hydroxybenzoic acid are shown in Table II and plotted in Fig. 1B. It is observed that 17 of the 32 results or 53% are within 30% of the observed solubility values. These are the best results that can be expected at this stage of the research for these highly irregular solutions. As shown in the preceding paper, both of these solutes show bimodal interactions (*cf.* Figs. 3 and 4 of Ref. 4), and that may be a major cause for perturbation of the single-peak models implied by Eqs. 6–8.

Thermodynamics of Irregular Solutions—Scatchard (14) suggested that nonideal behavior of a solute in solution could be expressed as an "excess free energy" of mixing:

$$\overline{\Delta G}^E = RT \ln (X_2^i/X_2) \quad (\text{Eq. 9})$$

The excess free energies of methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid in solution at 25°C were obtained from Eq. 9 and are listed in Table III as experimental values. The $\overline{\Delta G}^E_{\text{exp}}$ values for benzoic acid, the solubility of which was reported earlier (4), were also included. The calculated results, $\overline{\Delta G}^E_{\text{calc}}$, are obtained from a rearrangement of the four-parameter regression equation, which for benzoic acid (4) is:

$$\overline{\Delta G}^E_{\text{calc}} = V_2\phi_{1,\text{calc}}^2(2.125\delta_{1d}^2 - 36.90\delta_{1d} + 0.2188\delta_{1p}^2 - 2.603\delta_{1p} + 0.6139\delta_{1a}\delta_{1b} - 1.966\delta_{1a} - 2.955\delta_{1b} + 174.25 - 20.271) \quad (\text{Eq. 10})$$

The $\overline{\Delta G}^E_{\text{calc}}$ along with $\overline{\Delta G}^E_{\text{exp}}$ and the residuals for methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid at 25°C are listed in Table III. The previously reported (4) values for benzoic acid are included for comparison. The results for benzoic acid demonstrate that most of the calculated excess free energies, $\overline{\Delta G}^E_{\text{calc}}$, are within 500 cal/mol of the experimental

Table V—Factor Matrix Obtained from Principal Factor Without Iteration

Three-Parameter System				
	Factor 1	Factor 2	Factor 3	
δ_d	-0.158	0.976	0.151	
δ_p	0.880	0.280	-0.383	
δ_h	0.913	-0.101	0.396	
Four-Parameter System				
	Factor 1	Factor 2	Factor 3	Factor 4
δ_d	-0.132	0.972	0.080	0.180
δ_p	0.833	0.273	0.072	-0.476
δ_a	0.718	0.020	-0.650	0.248
δ_b	0.749	-0.152	0.558	0.324

Table VI—Varimax Rotated Factor Matrix for Three and Four Solubility Parameters

Three-Parameter System				
	Factor 1	Factor 2	Factor 3	
δ_d	0.995	0.048	-0.092	
δ_p	0.062	0.941	0.333	
δ_h	-0.114	0.340	0.934	
Four-Parameter System				
	Factor 1	Factor 2	Factor 3	Factor 4
δ_d	0.995	0.048	-0.042	0.072
δ_p	0.058	0.941	0.227	0.243
δ_a	0.046	0.208	0.972	0.103
δ_b	0.080	0.227	0.105	0.965

values, except for two solvents, benzene (-696 cal/mol) and ethylene dichloride (521 cal/mol).

The estimated excess free energies are less satisfactory for the two benzoic acid derivatives. For methyl *p*-hydroxybenzoate, 5 of 36 calculated excess free energies have errors > 500 cal/mol relative to the observed values. For *p*-hydroxybenzoic acid, 6 of 32 calculated excess free energies are > 500 cal/mol in error.

The observed free energies $\overline{\Delta G}^E_{\text{exp}}$ are plotted (horizontal axis) versus $\overline{\Delta G}^E_{\text{calc}}$ (vertical axis) for benzoic acid, methyl *p*-hydroxybenzoate, and *p*-hydroxybenzoic acid in Fig. 2A–C. The straight lines plotted on these three graphs are drawn with slopes of unity and pass through the origin. The vertical distance of each point from the line is a measure of the residuals of Table III and gives an indication of the adjustment that must be made in the four parameters of the predicting equations. From the difference of excess free energies between the observed and estimated values, one can calculate the error in calories per mole caused by one unit difference in a solvent solubility parameter. With methyl *p*-hydroxybenzoate in glycerol, for example, δ_b was changed from 5.1 to 6.1 and δ_a was correspondingly adjusted from 20.0 to 16.8 (to keep δ_h the same) while holding the other values constant. The $\overline{\Delta G}^E_{\text{calc}}$ value obtained by this adjustment became 2070 cal/mol. Comparing this value with the original $\overline{\Delta G}^E_{\text{calc}}$ value (2282 cal/mol) demonstrates that an increase of one unit in δ_b decreased the predicted excess free energy by 212 cal/mol. This example suggests the changes in partial solubility parameters that are required to obtain more exact solubilities.

Estimating the Partial Solubility Parameters of the Solute—As reported earlier (4), the regression equation for a specific solute in a number of solvents can be used to estimate the partial solubility parameters of the solute. For methyl *p*-hydroxybenzoate, Eq. 7 can be rearranged to:

$$\frac{\log \alpha_2}{A} = 7.341(\delta_{1d} - 9.09)^2 + 1.138(\delta_{1p} - 7.28)^2 + 0.9032(\delta_{1a} - 4.68)(\delta_{1b} - 0.73) - 45.59 \quad (\text{Eq. 11})$$

Comparing Eq. 11 with Eq. 2 shows that $\delta_{2d} = 9.09$, $\delta_{2p} = 7.28$, $\delta_{2a} = 4.68$, and $\delta_{2b} = 0.73$ for methyl *p*-hydroxybenzoate. The total solubility parameter δ_{2T} can be calculated by summing the squares of these partial values and taking the square root of the sum:

$$\delta_{2T}^2 = \delta_{2d}^2 + \delta_{2p}^2 + 2\delta_{2a}\delta_{2b} = (9.09)^2 + (7.28)^2 + 2(4.68)(0.73) \delta_{2T} = 11.94 \quad (\text{Eq. 12})$$

This value compares satisfactorily with the value, 12.1, reported elsewhere (7).

The partial solubility parameters of *p*-hydroxybenzoic acid can be obtained by transforming Eq. 8 to the following:

$$\frac{\log \alpha_2}{A} = 10.65(\delta_{1d} - 8.31)^2 + 1.04(\delta_{1p} - 7.08)^2 + 2.072(\delta_{1a} - 5.83)(\delta_{1b} - 3.91) - 44.06 \quad (\text{Eq. 13})$$

It is seen from Eq. 13 that $\delta_{2d} = 8.31$, $\delta_{2p} = 7.08$, $\delta_{2a} = 5.83$, $\delta_{2b} = 3.91$, and $\delta_{2T} = 12.84$ for *p*-hydroxybenzoic acid. The total solubility parameter for *p*-hydroxybenzoic acid calculated by Fedors (11) group contribution (15.3) and the value ($\delta_{2T} = 12.84$) obtained by regression show poor agreement. Undoubtedly, there are faults in both methods. Fedors' procedure, while generally accurate, tends to overestimate δ_T for aromatics.

Factor Analysis for Three- and Four-Parameter Solubility Systems—Factor analysis (15) offers the possibility of analyzing the

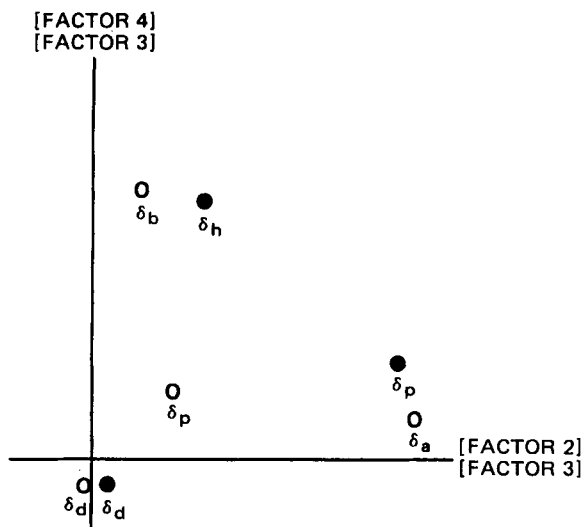


Figure 3—Relative positions of δ_a and δ_b (four-parameter system, \circ) to that of [Factors 3 and 4] in comparison with the relative positions of δ_p and δ_h (three-parameter system, \bullet) to that of [Factors 2 and 3]. Note that the open point δ_a is closer to the [Factor 3] axis than closed point δ_p is to its [Factor 2] axis. Also, the open point δ_b is closer to its axis [Factor 4] than is the closed point δ_h to its [Factor 3] axis.

independencies (noncorrelations) between the independent variables in regression equations. It can be used to test whether the three- or the four-parameter solubility approach forms a better set of variables. In multiple linear regression it is assumed that all the independent variables are mutually independent. In this study, the factor analysis program was obtained from SPSS (13) and was run on the University of Texas Cyber System. The three- and four-parameter systems are listed in Table I of the preceding paper (4); most of the data can also be found in Tables I and II. Table IV lists the correlation coefficients of these solubility parameter systems. As can be seen in Table IV, δ_d is almost uncorrelated with the other partial solubility parameters in both systems. For these solubility parameters, δ_p and δ_h are moderately correlated, with the correlation coefficient equal to 0.623 ($r^2 = 0.388$). This condition is undesirable, for it indicates that the variables are not independent of one another. For the expanded four-parameter solubility system, the correlation coefficient of δ_p and δ_a is 0.439 and that of δ_p and δ_b is 0.468 which are a little high, but are satisfactory in the present study ($r^2 < 0.22$ in both cases). The correlation coefficient between δ_a and δ_b is 0.252, which represents only slight correlation ($r^2 = 0.06$); therefore, these variables are independent of each other, a most desirable characteristic for multiple regression.

Table V lists the factor matrix obtained using the principal factor method without iteration (13). It can be seen from Table V that δ_p and δ_h are heavily loaded on Factor 1 and δ_d is loaded on Factor 2. Factor 3 does not contribute much in the three-parameter system. For the expanded four-parameter solubility system, δ_d is loaded on Factor 2, δ_p is loaded on Factor 1, but also slightly loaded on Factor 4; δ_a and δ_b are loaded on Factor 1 and Factor 3. The unrotated factors in Table V shows that four solubility parameters form a better system of independent variables than three solubility parameters.

Orthogonal factor rotation (13) was also used to compare these two systems of variables. Varimax-type rotation (13) was carried out by assuming the partial solubility parameters to be mutually independent. The procedure attempts to force independent variables to be loaded on different Factors and heavily on one Factor only. The results are found in Table VI. For the three-parameter system, δ_d is completely loaded on Factor 1, but δ_p and δ_h are less successful in loading on Factors 2 and 3, respectively.

For the four-parameter system, δ_d , δ_a , and δ_b are heavily loaded on Factors 1, 3, and 4, respectively; δ_p is less satisfactory in loading on Factor 2. The relative position of δ_a and δ_b to that of Factor 3 and Factor 4 for the four-parameter system was plotted in Fig. 3 in comparison with the

relative position of δ_p and δ_h to that of Factor 2 and Factor 3 for the three-parameter system. It can be concluded that four solubility parameters are more independent than the quantities in the three-parameter solubility system. This is a desirable feature of the four-parameter system.

CONCLUSIONS

The UNIFAC (5) program which had been used earlier in computing solubilities of naphthalene (3) and benzoic acid (4) in individual solvents was not used in this study due to lack of parameters for interaction energies of phenolic —OH/—COOH groups in the data bank of the UNIFAC program.

The regression equation for predicting the solubility of a solute can be cast into the proposed model (Eq. 2) to get partial and total solubility parameters of the solute. The total solubility parameters obtained by this method are close to the values reported elsewhere for benzoic acid and methyl *p*-hydroxybenzoate. For *p*-hydroxybenzoic acid the total solubility parameter, 12.84, obtained from Eq. 13 is different from the value, 15.3, calculated by Fedors' (11) method.

The magnitude of the constant terms in Eqs. 11 and 13, namely -45.59 and -44.06 , suggest that four solubility parameters are still not sufficient to explain all interactions between solute and solvent for the two solutes, methyl *p*-hydroxybenzoate and *p*-hydroxybenzoic acid, as compared with naphthalene and benzoic acid, whose simpler structures result in constant terms closer to zero (-0.24 and 3.13 , respectively) in the transformed regression equations (4). On the other hand, this constant is not merely a catch-all for errors as in most regressions, but contains the entropies of mixing due to disparities in volume and shape between solvent and solute. Therefore, reducing the constant value to zero is not a realistic goal.

REFERENCES

- (1) A. Martin, J. Newburger, and A. Adjei, *J. Pharm. Sci.*, **69**, 487 (1980).
- (2) A. Martin, A. N. Paruta, and A. Adjei, *J. Pharm. Sci.*, **70**, 1115 (1981).
- (3) A. Martin, P. L. Wu, A. Adjei, A. Beerbower, and J. M. Prausnitz, *J. Pharm. Sci.*, **70**, 1260 (1981).
- (4) A. Beerbower, A. Martin, and P. L. Wu, *J. Pharm. Sci.*, **73**, 179 (1984).
- (5) A. Fredenslund, J. Gmehling, and P. Rasmussen, "Vapor-Liquid Equilibrium Using UNIFAC," Elsevier, New York, N.Y., 1977.
- (6) C. K. Hancock, J. N. Pawloski, and J. P. Indoux, *J. Org. Chem.*, **31**, 3801 (1966).
- (7) A. Martin and J. Carstensen, *J. Pharm. Sci.*, **70**, 170 (1981).
- (8) C. M. Hansen and A. Beerbower, "Encyclopedia of Chemical Technology," Suppl. Vol., 2nd ed., Wiley, New York, N.Y., 1971.
- (9) M. J. Kamlet, J. L. M. Abboud, and R. W. Taft, *Prog. Phys. Org. Chem.*, **13**, 485 (1980).
- (10) F. T. Gucker, Jr., W. I. Ford, and C. E. Moser, *J. Phys. Chem.*, **43**, 153 (1939).
- (11) R. F. Fedors, *Polym. Eng. Sci.*, **14**, 147 (1974).
- (12) IMSL Reference Manual, International Mathematical and Statistical Library, Houston, Texas, 1979.
- (13) 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, chap. 20.
- (14) G. Scatchard, *Chem. Rev.*, **8**, 321 (1931).
- (15) R. L. Rummel, "Applied Factor Analysis," Northwestern University Press, Evanston, Ill., 1970.

ACKNOWLEDGMENTS

Supported in part by a grant from the University of Texas Research Institute. A. Martin is grateful to Coulter R. Sublett for the endowed professorship which also supported this work.

The authors acknowledge review of the section on factor analysis by Drs. Mary Whiteside and Francis May, Statistics Division, University of Texas at Austin.