ELSEVIER

Journal of Chromatography A, 878 (2000) 115–124

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Revised linear solvation energy relationship coefficients for the 77-phase McReynolds data set based on an updated set of solute descriptors

M.H. Abraham[a], D.S. Ballantine[b,*], B.K. Callihan[b]

[a]*Department of Chemistry, University College London, 20 Gordon Street, London WC1H OAJ, UK*
[b]*Department of Chemistry and Biochemistry, Northern Illinois University, DeKalb, IL 60115, USA*

Received 1 September 1999; received in revised form 27 January 2000; accepted 27 January 2000

Abstract

Linear solvation energy relationship (LSER) coefficients for the 77-phase McReynolds data set have been recalculated using updated solute descriptors in the revised solvation equation:

$$\log SP = c_1 + r_1R_2 + s_1\pi_2^H + a_1\sum\alpha_2^H + b_1\sum\beta_2^H + l_1 \log L^{16}$$

These revised LSER coefficients are presented and classified by cluster analysis into groupings of stationary phases which have comparable solubility properties. It was found that the groupings were similar to those proposed by Abraham using the original solvation equation and that any dissimilarities were readily explainable by the grouping methods that were applied. Comparison of the original coefficients with the revised set also shows that several stationary phases which had a statistically insignificant b_1 value with the original equation now have significant b_1 values when utilizing the revised solvation equation. Published by Elsevier Science B.V.

Keywords: Linear solvation energy relationships; Solute descriptors

1. Introduction

Solubility phenomena play a major role in a variety of chemical and biological processes; hence, a thorough understanding of the factors that affect such phenomena is crucial for the utilization and/or optimization of solubility properties in systems under investigation [1]. Ideally, a quantitative measure of solubility processes provides better insight into the fundamental interactions which take place between solute and solvent. Because gas chromatographic retention data can be directly correlated with intrinsic thermodynamic properties, namely the thermodynamic partition coefficient, and because these properties are directly related to the intermolecular interactions between solute and solvent, extensive analysis of several key sets of gas chromatographic retention data [2,3] has been performed in order to develop relationships which predict solubility behavior [3–5].

Of the many sets of gas chromatographic retention data currently available, the McReynolds data set of 77 stationary phases has been utilized to develop several predictive relationships [4,5]. Among the

*Corresponding author.

0021-9673/00/$ – see front matter Published by Elsevier Science B.V.
PII: S0021-9673(00)00143-6

most useful of these studies are the linear solvation energy relationships (LSERs). Linear solvation energy relationships quantify the solubility process by correlating a given solubility property with several additive terms which represent specific solubility interactions. An early example of a linear free energy relationship (LFER) applied to gas chromatographic data was that of Abraham et al. [5]:

$$\log SP = c_1 + r_1 R_2 + s_1 \pi_2^* + a_1 \alpha_2^H + b_1 \beta_2^H + l_1 \log L^{16} \quad (1)$$

In this equation, *SP* refers to a property of probe solutes in a fixed system; thus *SP* can be the adjusted retention time, t_r', the specific retention volume, V_g, or the thermodynamic gas solvent partition coefficient, K or L. The subscript 2 denotes a solute property, and the independent variables are R_2, an excess molar refraction, π_2^*, the dipolarity/polarizability, α_2^H and β_2^H, the hydrogen-bond acidity and basicity, and $\log L^{16}$, where L^{16} is the gas–hexadecane partition coefficient at 25°C. The excess molar refraction was a new solute parameter [5], and π_2^* was taken from Kamlet et al. [6]. The hydrogen-bond descriptors were those developed by Abraham et al. [7,8] from 1:1 hydrogen-bond equilibrium constants, and the L^{16} descriptor was already well known [9].

Eq. (1) was applied to retention data on a number of gas–liquid chromatography (GLC) phases [5], to gas–solvent partition coefficients in amide solvents [10], and to $\log V_g$ values for solutes on all 77 of the McReynolds stationary phases [11]. However, it was recognized that a number of solute descriptors in Eq. (1) could be improved. First, the old dipolarity/polarizability parameter, π_2^*, which had been derived from solvatochromic measurements [6], was replaced by a new descriptor, π_2^H, obtained from GLC measurements [12], and which was then a true free energy property. Second, the hydrogen-bond descriptors derived from 1:1 complexation constants were replaced [12,13] by 'effective' or 'overall' descriptors, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$, that refer to the hydrogen-bond propensity of a solute surrounded by solvent molecules. The revised equation [13] can be stated as:

$$\log SP = c_1 + r_1 R_2 + s_1 \pi_2^H + a_1 \sum \alpha_2^H + b_1 \sum \beta_2^H + l_1 \log L^{16} \quad (2)$$

Because the descriptors used in Eq. (2) are chemically based, the coefficients in the equation encode specific information [14]. The r_1 coefficient indicates the propensity of solutes to interact via π and n-electron pairs (with hexadecane as a standard solvent). The s_1 coefficient is a measure of the solvent phase dipolarity/polarizability; the a_1 coefficient is a measure of the solvent hydrogen-bond basicity (because acidic solutes will interact with basic phases); and the b_1 coefficient is a measure of the solvent hydrogen-bond acidity. The important l_1 coefficient is a measure of the solvent lipophilicity; in other words, how near the solvent lipophilicity is to that of hexadecane, for which $l_1 = 1$ by definition. As regards GLC retention, the l_1 coefficient also represents the ability of a stationary phase to separate members of an homologous series.

Eq. (2) is now one of the best known general LFERs, and has been used to characterize a large number of GLC phases [15], as well as modeling gas–water partition [16], and partition between the gas phase and numerous organic solvents [17–20]. Not only has Eq. (1) been replaced by Eq. (2), but solute descriptors for use in Eq. (2) have now been obtained for a large number of solutes. Table 1 contains details of the present data base of descriptors.

However, one marked omission is that the revised Eq. (2) has never been applied to the McReynolds set [2] of 77 GLC phases. In the present work, we correct this omission by calculating the coefficients (with subscript 1) in Eq. (2) for all 77 phases, using the most up-to-date solute descriptors available. These new coefficients will be presented here as well as a discussion of any significant changes in the terms which characterize these stationary phases. The 77-phase McReynolds set will be divided into groupings by cluster analysis of the revised coefficients. These new groupings will then be com-

Table 1
The available solute descriptors for use in Eq. (2)

Descriptor	Number	Minimum	Maximum
R_2	3430	−1.37	4.62
π_2^H	2950	−0.54	5.60
$\Sigma\alpha_2^H$	3750	0.00	2.10
$\Sigma\beta_2^H$	2600	0.00	4.52
$\text{Log } L^{16}$	1980	−1.74	29.97

pared with previous groupings by both Abraham et al. [11] and Wold [21].

2. Results

In this analysis of the McReynolds 77-phase set, multiple linear regression analysis (MLRA) was performed using retention data for approximately 200 solutes for which all solute descriptors were known. This analysis yielded a complete set of five LSER coefficients (r_1, s_1, a_1, b_1, and l_1) for the 77-phase set. It was found that, in several of these phases, the $b_1\Sigma\beta_2^H$ term was insignificant and could be eliminated. As a result, MLRA was repeated without the $b_1\Sigma\beta_2^H$ term, this time using approximately 240 to 250 solutes, the additional 40 to 50 solutes having no value for $\Sigma\beta_2^H$.This latter analysis gave the r_1, s_1, a_1, and l_1 coefficients designated as the revised coefficients.

Tables 2 and 3 list the revised LSER coefficients based on the new solute descriptors. Because relatively few stationary phases now have a value for b_1 which is significant, Table 2 gives the values for c_1, r_1, s_1, a_1, and l_1 for all 77 phases, which were calculated without the $b_1\Sigma\beta_2^H$ term. Values for b_1 are given in Table 3 for sorbitol, diglycerol, and docosanol, as well as several other phases for which the b_1 coefficient was found to be statistically insignificant when determined with the original solute descriptors but is now statistically significant with the revised solute descriptors. Table 3 also includes the c_1, r_1, s_1, a_1, and l_1 values for these phases when calculated with the $b_1\Sigma\beta_2^H$ included. The R^2 and the standard error for each relationship did not differ significantly from those of Abraham et al.'s original calculation of the coefficients [5].

For a typical GLC phase such as sorbitol, the LSER equation with its appropriate coefficients included would be as follows:

$$\log V_g = -1.98 + 0.47R_2 + 1.25\pi_2^H + 2.09\sum\alpha_2^H + 0.51\sum\beta_2^H + 0.336\log L^{16} \quad (3)$$

To determine the extent of interaction between this phase and any solute for which solute descriptors are known, the corresponding solute descriptor values are inserted into the equation and the resulting V_g (after taking the antilogarithm) should reliably predict the extent of retention on sorbitol at 120°C.

In determining whether a change in any of the coefficients between the original values and the revised values was significant, the amount of the increase/decrease from the original value was compared to the estimated errors associated with these coefficients. The typical standard error for the r_1, s_1, and a_1 coefficients is 0.02–0.03. For the l_1 coefficient, the associated error typically ranged from 0.002 to 0.005. Thus, as a cutoff for determining whether a change in the value of a coefficient was significant or not, two to three times the typical standard error was used as the criterion for listing stationary phases as having a significant change in the value for any particular coefficient. Thus (somewhat arbitrarily), a difference of 0.08 was used as the cutoff for the a_1 and s_1 coefficients, and 0.015 was used for the l_1 coefficient, keeping in mind that differences slightly smaller than this cutoff may or may not be rationalized by chemical explanations set forth below. Many factors are at work that result in differences between the original and revised coefficients: in many cases, three solute descriptors (π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$) have been changed at once; moreover, MLRA on a set of five variables will initiate changes in several coefficients (even if the corresponding solute descriptor remains unchanged). In addition, where trends in the changes in coefficients appear difficult to explain, more than one factor may differ between two stationary phases (two phases which both contain diesters, for example, may also have significantly different carbon skeletons leading to large differences in inductive effects).

Examination of the revised set of LSER coefficients shows that a large number of phases had an insignificant or marginally significant r_1 value. The new coefficients generated were compared with the original coefficients to determine if any trends existed between the original and revised coefficients. For the r_1 coefficient, all values decreased or remained unchanged with the exception of sorbitol, which increased by 0.12. Even though a large number of phases had an insignificant or marginally significant r_1 value, these values were reported to facilitate direct comparison with the original set of coefficients. In cases where the r_1 value was insignificant or marginally significant, elimination of

Table 2
Revised LSER coefficients for the McReynolds 77-phase data set

Stationary phase	c_1	r_1	s_1	a_1	l_1
Squalane	−0.29	0.08	0.08	0.11	0.674
Apiezon M	−0.45	0.23	0.11	0.12	0.600
Apiezon N	−0.48	0.24	0.13	0.11	0.600
Apiezon J	−0.48	0.24	0.14	0.12	0.595
Apiezon L	−0.45	0.25	0.09	0.07	0.600
Versilube P50	−0.38	−0.02	0.27	0.19	0.540
SE-31	−0.36	−0.02	0.25	0.21	0.521
SE-30	−0.34	−0.01	0.31	0.33	0.525
SE-30 NPGA	−0.30	−0.03	0.36	0.45	0.525
SE-52	−0.39	−0.01	0.34	0.26	0.534
Dioctyl sebacate	−0.34	0.09	0.49	0.80	0.589
Di-2-ethylhexyl sebacate	−0.34	0.11	0.49	0.79	0.589
Tripelargonate	−0.40	0.11	0.53	0.82	0.580
Isooctyldecyl adipate	−0.37	0.08	0.55	0.84	0.586
Di-2-ethylhexyl adipate	−0.31	0.04	0.63	1.06	0.643
Diisodecyl phthalate	−0.50	0.07	0.64	0.79	0.588
Dioctyl phthalate	−0.10	0.08	0.68	0.82	0.586
UCON LB-1715	−0.57	0.07	0.78	1.23	0.541
Flexol 8N8	−0.48	0.06	0.74	1.27	0.573
Pluronic L81	−0.49	0.10	0.82	1.32	0.535
Polyphenyl ether, five rings	−0.70	0.15	0.91	0.62	0.561
Polyphenyl ether, six rings	−0.71	0.17	0.91	0.62	0.554
Tricresylphosphate	−0.66	0.14	1.01	1.19	0.550
Sucrose acetate isobutanoate	−0.55	0.01	1.04	1.25	0.510
Hallcomid M18	−0.35	0.13	0.58	1.49	0.592
Hallcomid M18 OL	−0.41	0.14	0.65	1.50	0.584
Pluronic L42	−0.51	0.09	0.92	1.49	0.526
Pluronic L72	−0.54	0.10	0.91	1.44	0.530
Pluronic L61	−0.51	0.07	0.90	1.37	0.527
Pluronic L63	−0.53	0.13	0.95	1.44	0.517
Polytergent J300	−0.54	0.11	0.97	1.55	0.533
Pluronic P84	−0.59	0.14	1.00	1.52	0.515
Pluronic P85	−0.58	0.16	1.00	1.51	0.512
Pluronic L44	−0.56	0.13	1.01	1.57	0.514
Oronite NIW	−0.63	0.13	1.02	1.53	0.523
UCON HB-2000	−0.60	0.16	1.01	1.52	0.515
Ethofat 60-25	−0.58	0.16	1.01	1.59	0.517
Pluronic P65	−0.59	0.15	1.03	1.54	0.511
Pluronic P46	−0.63	0.17	1.09	1.65	0.505
Tergitol NPX	−0.56	0.15	1.03	1.50	0.514
Neopentylglycol adipate, term	−0.65	0.15	1.05	1.42	0.510
Ethylene glycol sebacate	−0.74	0.19	1.08	1.44	0.514
Diethylene glycol sebacate	−0.73	0.23	1.11	1.46	0.496
Neopentylglycol adipate	−0.63	0.17	1.09	1.42	0.489
Neopentylglycol succinate	−0.72	0.16	1.26	1.54	0.468
Pluronic F88	−0.64	0.23	1.16	1.66	0.481
Pluronic F68	−0.66	0.23	1.17	1.69	0.485
Pluronic F77	−0.62	0.20	1.12	1.61	0.493
Igepal CO 880	−0.67	0.23	1.18	1.69	0.486
Triton X 305	−0.82	0.22	1.19	1.68	0.490
Ethylene glycol adipate	−0.91	0.20	1.45	1.79	0.448
Diethylene glycol adipate	−0.92	0.25	1.48	1.83	0.443
XF-1150	−0.70	0.09	1.40	1.41	0.422

Table 2. Continued

Stationary phase	c_1	r_1	s_1	a_1	l_1
Sucrose octaacetate	−0.72	0.06	1.46	1.58	0.429
Carbowax 20M	−0.74	0.27	1.31	1.79	0.465
Carbowax 6000	−0.73	0.26	1.32	1.83	0.466
Carbowax 4000	−0.75	0.25	1.33	1.87	0.465
Carbowax 1540	−0.74	0.25	1.35	1.93	0.454
Carbowax 1000	−0.74	0.25	1.38	1.98	0.452
Carbowax 600	−0.80	0.23	1.45	2.15	0.452
Carbowax 400	−0.74	0.20	1.47	2.19	0.438
Carbowax 300	−0.77	0.19	1.50	2.27	0.433
Quadrol	−0.72	0.01	1.39	2.34	0.471
Hyprose SP80	−0.85	0.09	1.46	2.35	0.419
Triethylene glycol succinate	−0.97	0.26	1.64	1.92	0.410
Diethylene glycol succinate	−0.93	0.28	1.69	1.78	0.373
Dow Corning Fluid 550	−0.46	0.04	0.55	0.33	0.548
Castorwax	−0.47	0.07	0.66	1.08	0.565
Dibutyl tetrachlorophthalate	−0.57	0.17	0.67	0.66	0.593
Citroflex A4	−0.42	0.00	0.87	1.08	0.553
Bis(2-ethoxyethyl) phthalate	−0.59	0.04	1.17	1.26	0.525
Dow Corning Fluid FS 1265	−0.74	−0.22	1.21	0.38	0.449
Kroniflex THFP	−0.72	0.17	1.34	2.31	0.499
Zonyl E-7	−0.77	−0.34	1.50	0.75	0.454
Docosanol	−0.38	0.04	0.37	0.88	0.605
Diglycerol	−1.37	0.48	1.65	2.68	0.236
Sorbitol	−1.98	0.47	1.25	2.09	0.336

Table 3
Revised user coefficients for McReynolds phases having a significant, *b*, value

Stationary phase	c_1 (revised)	r_1 (revised)	s_1 (revised)	a_1 (revised)	b_1 (original)	b_1 (revised)	l_1 (revised)
Polyphenyl ether (six rings)	−0.75	0.20	0.88	0.61	N/A	0.10	0.557
Tricresyl phosphate	−0.70	0.14	1.00	1.21	N/A	0.10	0.553
Sucrose acetate isobutanoate	−0.58	Insignificant	1.02	1.27	N/A	0.07	0.512
Neopentylglycol adipate, term	−0.70	0.18	1.02	1.40	N/A	0.10	0.514
Ethylene glycol sebacate	−0.79	0.22	1.05	1.44	N/A	0.11	0.519
Diethylene glycol sebacate	−0.76	0.24	1.10	1.48	N/A	0.07	0.500
Neopentylglycol adipate	−0.67	0.19	1.08	1.44	N/A	0.07	0.490
Neopentylglycol succinate	−0.75	0.18	1.22	1.52	N/A	0.08	0.471
Ethylene glycol adipate	−0.97	0.29	1.37	1.72	N/A	0.19	0.452
Diethylene glycol adipate	−0.96	0.31	1.41	1.78	N/A	0.14	0.445
Sucrose octaacetate	−0.76	0.10	1.40	1.56	N/A	0.14	0.432
Carbowax 300	−0.80	0.22	1.45	2.23	N/A	0.12	0.434
Quadrol	−0.80	Insignificant	1.30	2.28	N/A	0.24	0.476
Hyprose SP80	−0.95	0.19	1.35	2.24	N/A	0.31	0.424
Triethylene glycol succinate	−1.02	0.34	1.57	1.86	N/A	0.16	0.411
Diethylene glycol succinate	−1.00	0.36	1.64	1.75	N/A	0.17	0.376
Zonyl E-7	−0.88	−0.23	1.37	0.59	N/A	0.37	0.457
Docosanol	−0.38	Insignificant	0.37	0.88	0.34	0.20	0.605
Diglycerol	−1.36	0.48	1.65	2.68	0.52	0.69	0.236
Sorbitol	−1.98	0.47	1.25	2.09	0.34	0.51	0.336

the r_1R_2 term would not severely impact the values for the remaining coefficients.

Stationary phases for which the s_1 value increased significantly ($\geq$0.08) include: SE-30 NPGA, di-2-ethylhexyl adipate, UCON LB-1715, Carbowax 20M, docosanol, and an extremely significant increase with sorbitol (0.44). The only stationary phase which exhibited a significant drop ($\geq$0.08) in the s_1 value was Zonyl E-7.

Examination of the a_1 coefficient showed that several of the 77 stationary phases exhibited a statistically significant increase in the a_1 value: squalane, SE-30 NPGA, di-2-ethylhexyl adipate, polyphenyl ether (five rings), polyphenyl ether (six rings), Pluronic P84, Oronite NIW, Pluronic P46, Pluronic F88, Pluronic F68, Igepal CO 880, ethylene glycol adipate, diethylene glycol adipate, Carbowax 1540, Carbowax 1000, diethylene glycol succinate, Citroflex A-4, Dow Corning Fluid FS 1265, Zonyl E-7, and docosanol (0.08–0.19). Sorbitol showed a much larger increase (0.32). Diglycerol showed a moderate decrease (0.09) in the value for the a_1 coefficient.

Comparing the original l_1 values with the revised numbers, only two stationary phases showed a marked increase in the value of the l_1 coefficient: squalane (0.056) and di-2-ethylhexyl adipate (0.052). One stationary phase, sorbitol, showed a moderate decrease (0.024) in the l_1 value.

In Abraham's original classification of the 77 stationary phase set of McReynolds, the value of the a_1 coefficient for each stationary phase was plotted versus the value of the s_1 coefficient for that same stationary phase. Stationary phases were then classified based on the proximity or clustering of points into distinct groups. In our present analysis, the revised coefficients were analyzed using the Cluster Analysis program contained in Complete Statistical System by StatSoft, ©1988. Table 4 lists the parameters used in the cluster analysis.

A couple of points are worth noting in the analysis of the coefficients by the above approach. First, because the b_1 coefficient was not used in the Cluster Analysis program (less than 30% of the phases have a b_1 value), the r_1, s_1, a_1, and l_1 values used in the analysis for the phases which do have a significant b_1 value are those reported in Table 2; these values for the revised coefficients were calculated with the $b_1\Sigma\beta_2^H$ term removed. These coefficients were used to be consistent with the previous studies; the differences in the coefficients when calculated either with or without the $b_1\Sigma\beta_2^H$ term are small and the effects on the clusters generated should be insignificant.

Table 4
Parameters used for the cluster analysis of the revised LSER coefficients

Number of variables:	Four (r_1, s_1, a_1, and l_1)
Raw data input	
Method:	Joining trees
Cluster cases or variables?	Cases
Distance measure:	Euclidean distances
Amalgamation rule:	Single linkage
Number of cases:	77 (McReynolds phases)

The original stationary-phase groupings of Abraham based on the original solvation equation (Eq. (1)) are compared with the groupings obtained through cluster analysis on the revised coefficients. Both sets of groupings are listed in Table 5. When the cluster analysis was performed on the revised coefficients, all 77 phases were included in the analysis. When the groupings were established, docosanol, sorbitol, and diglycerol were left out (due to a significant $b_1\Sigma\beta_2^H$ term) as well as Dow Corning Fluid 550, Castorwax, dibutyl tetrachlorophthalate, Citroflex A-4, bis(2-ethoxyethyl) phthalate, DCFS 1265, Kroniflex THFP, and Zonyl E-7 (which did not fit with any grouping in Abraham's original clustering).

Comparing Abraham's clusters of the original coefficients with the cluster analysis of the revised coefficients, it is apparent that groups 1, 2, 3, 4, 6, 7, 8, 12, 13, and 16 are identical. Abraham groups 5, 9, 10, and 11 all fall into one group in the current cluster analysis of the revised coefficients, with the exception of neopentylglycol succinate, which clustered with the largest Carbowaxes (group 14). Group 15 (Abraham) seemed to divide into two groups (15A and 15B) in the cluster analysis.

To explain the disparities between Abraham's original groupings and our groupings (with the revised coefficients), a plot of a_1 versus s_1 (Fig. 1) for the revised coefficients was constructed and the data points separated into two groups: Abraham's original groupings and our clusters based on the

Table 5
Proximal groupings (original coefficients) versus cluster analysis (revised coefficients)

Group No.	Original coefficients	Revised coefficients
1	Squalane	
2	Apiezon M Apiezon N Apiezon J Apiezon L	
3	Versilube P50 SE-31 SE-30 SE-30 NPGA SE-52	
4	Dioctyl sebacate Di-2-ethylhexyl sebacate Tripelargonate Isooctyldecyl adipate Di-2-ethylhexyl adipate Diisodecyl phthalate Dioctyl phthalate	
6	Polyphenyl ether (five rings) Polyphenyl ether (six rings)	
7	Tricresyl phosphate Sucrose acetate isobutanoate	
8	Hallcomid M18 Hallcomid M18 OL	
12	Ethylene glycol adipate Diethylene glycol adipate	
13	XF 1150 Sucrose octaacetate	
15	Carbowax 600 Carbowax 400 Carbowax 300 Quadrol Hyprose SP80	15A Carbowax 600 Carbowax 400 Carbowax 300 15B Quadrol Hyprose SP80
16	Triethylene glycol succinate Diethylene glycol succinate	
5	UCON LB 1715 Flexol 8N8 Pluronic L81	Of the 27 stationary phases contained in Abraham's groupings 5, 9, 10, and 11, 26 phases group by cluster analysis into one grouping
9	Pluronic L42 Pluronic L72 Pluronic L61 Pluronic L63 Polytergent J300 Pluronic P84 Pluronic P85 Pluronic L44 Oronite NIW UCON HB 2000 Ethofat 60/25 Pluronic P65 Pluronic P46 Tergitol NPX	which is given the composite designation 5/9/10/11. The 27th phase, neopentylglycol succinate, is grouped with the heavier Carbowaxes (group 14)

Table 5. Continued

Group No.	Original coefficients	Revised coefficients
10	Neopentylglycol adipate, term Ethylene glycol sebacate Diethylene glycol sebacate Neopentylglycol adipate Neopentylglycol succinate	
11	Pluronic F88 Pluronic F68 Pluronic F77 Igepal CO 880 Triton X 305	
14	Carbowax 20M Carbowax 6000 Carbowax 4000 Carbowax 1540 Carbowax 1000	14A Carbowax 20M Carbowax 6000 Carbowax 4000 Carbowax 1540 Carbowax 1000 Neopentylglycol succinate

revised coefficients. To eliminate unnecessary confusion, the groupings which were determined to be identical were not included in the plot. Upon examination of this plot, one finds that, in fact, no significant disparities exist; the original groupings were based on visual proximity, whereas the groupings of the revised coefficients were based on a more statistical treatment with the Cluster Analysis program described earlier. Abraham's original groupings are indicated by solid lines, whereas our clusters, based on the hierarchical clustering, are defined by dashed lines.

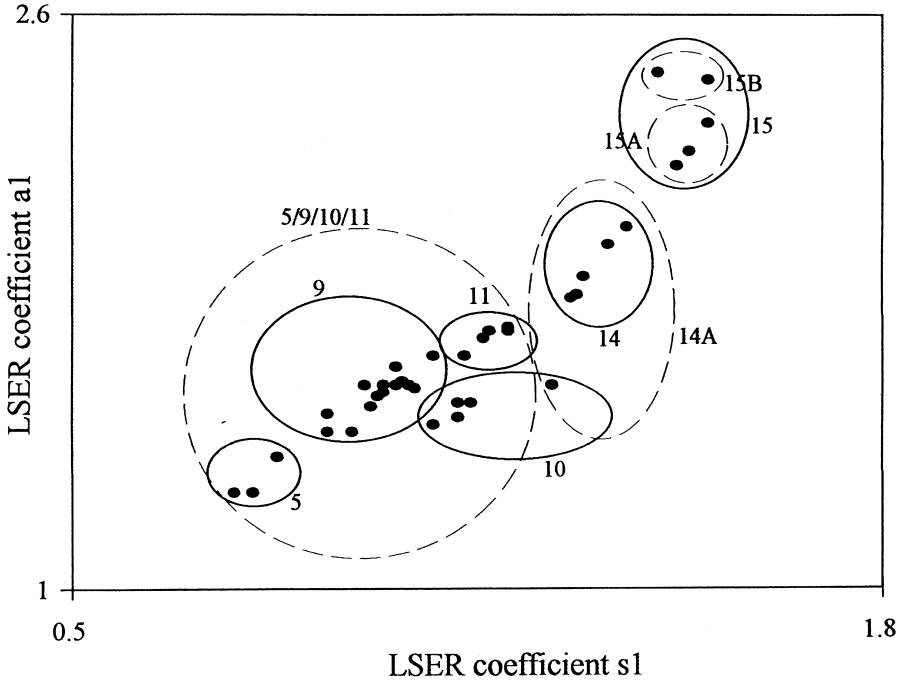

Fig. 1. Comparison of proximal groupings of original coefficients with groupings of revised coefficients by cluster analysis. Original groupings defined by solid ovals; revised groupings defined by dashed ovals.

Neopentylglycol succinate was included in group 10 of Abraham's original groupings; our cluster analysis groups it with the heavier Carbowaxes. This is feasible because the proximity plot shows that this stationary phase lies in a region intermediate between the two groupings; inclusion in either group would be plausible. The phase was included with group 10 (Abraham) due to the similarity in structure to the other phases in the group.

Abraham's group 15 contains three Carbowaxes (Carbowax 600, 400, 300) as well as quadrol and Hyprose SP80. Due to a tighter correlation (because four variables were used) in the groupings with the hierarchical clustering as opposed to the proximity plots (a_1 versus s_1), this group was split into two groups (Carbowaxes in one group and quadrol and Hyprose SP80 in the other group). In addition, groups 5, 9, 10, and 11 (except neopentylglycol succinate) all fell into one group by our statistical clustering.

Of the 226 stationary phases that Wold grouped into clusters by pattern cognition [21], 77 phases were common to the McReynolds data set and, as a result, grouped by Abraham et al. [11]. There was excellent correlation between Abraham et al.'s and Wold's groups; any differences between ours and Abraham's groups would be similar with the Wold groupings.

3. Discussion

The solute descriptors used in Eq. (2) are chemically much more reasonable than those used in the early Eq. (1). As noted above, the new π_2^H descriptor is a free energy parameter, whereas the old π_2^* descriptor is not a thermodynamic property at all. In addition, the new hydrogen-bond acidity and basicity descriptors are designed to relate to a situation in which the solute is surrounded by solvent molecules — as is exactly the case in GLC. We expect, therefore, that if there are differences between the coefficients calculated for Eq. (2) and Eq. (1), that the coefficients calculated for Eq. (2) will be chemically more realistic.

Examination of all stationary phases which now have significant b_1 values shows that each phase contains hydroxyl hydrogens which can participate as weak hydrogen-bond donors. Several phases have negative b_1 values, which is chemically unreasonable. In all of these cases, elimination of the $b_1\Sigma\beta_2^H$ term made the r_1 coefficient statistically more significant; also, the adjusted R^2 term was not negatively impacted, and in some cases actually improved.

Of the many phases which showed a significant change in the value of the a_1 coefficient, all of them have structural features that would facilitate hydrogen-bond acceptor basicity interactions: electron-rich oxygen, nitrogen, and fluorine atoms as well as ester functionalities which readily participate in hydrogen-bond acceptor basicity interactions. As explained previously, several cases were borderline; the changes in the a_1 coefficient could be due to the factors described above which contribute to hydrogen-bond acceptor basicity, or the changes might be a result of the expanded MLRA; several coefficients are being adjusted simultaneously to fit all cases.

The changes in the r_1 and the l_1 coefficients are more than likely a result of the MLRA process; these coefficients change slightly to accommodate the chemically more significant changes in the s_1, a_1, and b_1 coefficients.

Probably, the most difficult changes to explain chemically are those of the s_1 coefficient. Not only were new values of the solute descriptor, π_2^H, used in the updated solvation equation, the utilization of new $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ values decreases, in part, the contribution of hydrogen-bond donor and acceptor behavior to the π_2^H term, which leads to a corresponding change in the s_1 coefficient. Coupled with this is the fact that a large majority of the phases participate in at least some type of dipolarity/polarizability interactions, and, as discussed previously, the MLRA process can account for some of the change in the coefficients.

4. Conclusion

Revised LSER coefficients have been presented for an updated solubility relationship previously presented by Abraham. With this new relationship, several stationary phases now have significant b_1 values.

Comparison of the stationary-phase groupings

made by our cluster analysis with Abraham's proximity groupings and Wold's pattern cognition method shows that, in all three cases, the groupings are essentially the same. This is important because the fundamental chemistry of each group of phases is still the same; the coefficients that characterize each phase now more accurately represent the contribution of each intermolecular interaction to the solvation process.

Acknowledgements

The work reported herein was performed in partial fulfillment of degree requirements for the Ph.D. degree from Northern Illinois University (B.K.C.).

References

[1] M.J. Kamlet, R.W. Taft, Acta Chem. Scand. B 39 (1985) 611, and references cited therein.
[2] W.O. McReynolds, Gas chromatographic retention data, Preston Technical Abstracts, Evanston, IL, 1966.
[3] F. Patte, M. Etcheto, P. Laffort, Anal. Chem. 54 (1982) 2239.
[4] J.F.K. Huber, G. Reich, J. Chromatogr. 294 (1984) 15.
[5] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chem. Soc., Perkin Trans. 2 (1990) 1451.
[6] M.J. Kamlet, R.M. Doherty, J.-L.M. Abboud, M.H. Abraham, R.W. Taft, Chemtech 16 (1986) 566.
[7] M.H. Abraham, P.L. Grellier, D.V. Prior, P.P. Duce, J.J. Morris, P.J. Taylor, J. Chem. Soc., Perkin Trans. 2 (1989) 699.
[8] M.H. Abraham, P.L. Grellier, D.V. Prior, J.J. Morris, P.J. Taylor, J. Chem. Soc., Perkin Trans. 2 (1990) 521.
[9] M.H. Abraham, P.L. Grellier, R.A. McGill, J. Chem. Soc., Perkin Trans. 2 (1987) 797.
[10] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chem. Soc., Perkin Trans. 2 (1990) 1851.
[11] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chromatogr. 518 (1990) 329.
[12] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chromatogr. 587 (1991) 213.
[13] M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chromatogr. 587 (1991) 229.
[14] M.H. Abraham, Chem. Soc. Rev. 22 (1993) 73.
[15] M.H. Abraham, C.F. Poole, S.K. Poole, J. Chromatogr. A 842 (1999) 79.
[16] M.H. Abraham, J. Andonian-Haftvan, G.S. Whiting, A. Leo, R.W. Taft, J. Chem. Soc., Perkin Trans. 2 (1994) 1777.
[17] M.H. Abraham, G.S. Whiting, P.W. Carr, H. Ouyang, J. Chem. Soc., Perkin Trans. 2 (1998) 1385.
[18] M.H. Abraham, G.S. Whiting, W.J. Shuely, R.M. Doherty, Can. J. Chem. 76 (1998) 703.
[19] M.H. Abraham, J. Le, W.E. Acree, Jr., P.W. Carr, J. Phys. Org. Chem. (in press).
[20] M.H. Abraham, J. Le, W.E. Acree, Jr., P.W. Carr, A.J. Dallas (submitted for publication).
[21] S. Wold, J. Chromatogr. Sci. 13 (1975) 525.