



ELSEVIER

Journal of Chromatography A, 697 (1995) 429–440

JOURNAL OF
CHROMATOGRAPHY A

Application of principal component factor analysis to the cavity model of solvation to identify factors important in characterizing the solvent properties of gas chromatographic stationary phases

Salwa K. Poole, Colin F. Poole*

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

Abstract

Principal component factor analysis was applied to two sets of data consisting of the gas–liquid partition coefficient for 30 solutes on 22 stationary phases or 67 solutes on 10 stationary phases at 121.4°C. Three or four factors were required to characterize the data and were identified as contributions from cavity formation and dispersion interactions, solvent hydrogen-bond base interactions, and orientation interactions typical of aromatic and aliphatic compounds. None of the stationary phases were identified as significant hydrogen-bond acids. The contribution from cavity formation and dispersion were combined into a single factor while orientation interactions were divided into two contributions representing either the additional polarity of polarizable solutes (aromatic compounds) or a family type behavior resulting from a structural contribution to the number of solute–solvent interactions calculated using the cavity model. There is good general agreement between the assignments made using the principal component factor model and results obtained previously using the solvation parameter model.

1. Introduction

A number of attempts have been made to develop a quantitative scale of solvent selectivity to characterize the retention properties of gas chromatographic stationary phases based on the Gibbs free energy of solution for a series of reference compounds [1–5]. Reference compounds were selected assuming that each compound was retained by a single dominant intermolecular interaction and the stationary phases were ranked in accordance with their capacity to retain the reference compounds. This approach had to be abandoned when it was demonstrated

that suitable reference compounds for individual intermolecular interactions did not exist and that the calculated selectivity parameters based on the above approach were dependent on the size of the reference compound as well as its capacity for polar interactions [5–7]. A new approach was required that enabled a separation of molecular size and solute–solvent interactions to be made.

The cavity model of solution is a suitable model for separating the contribution of solute size from specific solute–solvent interactions [8–10]. The size contribution is represented by the cavity term, which accounts for the work that must be done to create a cavity in the solvent of a suitable size to accommodate the solute. This is a solvent property as it depends only on the free

* Corresponding author.

energy required to disrupt solvent–solvent interactions in preparing the cavity. The magnitude of the cavity term will depend on the cohesive character of the solvent and the size of the solute transferred to the solvent. For transfer of the solute from the gas phase to the stationary phase to be favorable the contribution of solute–solvent interactions set up once the solute is transferred to the solvent must exceed the free energy required to disrupt solvent–solvent interactions in preparing the cavity. This difference in free energy is responsible for the variation in gas–liquid partition coefficients observed for different solutes in the same solvent and the same solute in different solvents. Intuitively, this model provides a reasonable qualitative picture of the solvation process but as a quantitative model it is limited by the lack of an exact method to calculate the contribution of solute–solvent and solvent–solvent intermolecular interactions to the solvation process for conditions typical of gas chromatographic experiments. Semi-empirical solutions are available, however.

The first of these solutions is due to Abraham and co-workers, and in a slightly modified version by Carr and co-workers, as outlined in a companion paper [11]. The general equation for the solvation process in gas–liquid chromatography proposed by Abraham and co-workers is represented by Eq. 1.

$$\log K_L = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + l \log L^{16} \quad (1)$$

where K_L is the gas–liquid partition coefficient, R_2 the solute excess molar refraction, π_2^H the effective solute dipolarity/polarizability, α_2^H the effective solute hydrogen-bond acidity, β_2^H the effective solute hydrogen-bond basicity and L^{16} the solute gas–liquid partition coefficient on *n*-hexadecane at 25°C. The explanatory variables listed above are solvation parameters derived from equilibrium constants or calculated from gas chromatographic measurements and are free energy related parameters characteristic of the monomeric solute. The solvent properties r , s , a , b and l are unambiguously defined: the r constant refers to the ability of a solvent to interact

with solute *n*- or π -electron pairs; the s constant to the ability of the solvent to take part in dipole–dipole and dipole–induced dipole interactions; the a constant is a measure of the hydrogen-bond basicity of the solvent; the b constant is a measure of the hydrogen-bond acidity of the solvent and the l constant incorporates contributions from solvent cavity formation and solute–solvent dispersion interactions, and more specifically in gas–liquid chromatography indicates how well the phase will separate members of a homologous series. For an uncharacterized phase the solvent properties r , s , a , b and l are determined from the experimentally derived gas–liquid partition coefficient for a minimum of 15 to 30 varied solutes with known explanatory variables using the statistical analysis technique of multiple linear regression analysis.

An alternative approach is due to Poole and co-workers [6,9] and results in the following general equation for the solvation process

$$\Delta G_S^{\text{Soln}}(\text{X}) = \Delta G_S^{\text{Soln}}(\text{HC})^V + \Delta G_{\text{SQ}}^{\text{P}}(\text{X}) + \Delta G_S^{\text{Int}}(\text{X}) \quad (2)$$

where $\Delta G_S^{\text{Soln}}(\text{X})$ is the partial Gibbs free energy of solution for the transfer of solute X from the gas phase to the stationary phase S, $\Delta G_S^{\text{Soln}}(\text{HC})^V$ is the partial Gibbs free energy of solution for an *n*-alkane with an identical Van der Waals volume to solute X in the stationary phase S, $\Delta G_{\text{SQ}}^{\text{P}}(\text{X})$ is the partial Gibbs free energy of interaction for the polar contribution of solute X in a non-polar reference solvent squalane, SQ, calculated as shown by Eq. 3, and $\Delta G_S^{\text{Int}}(\text{X})$ is the partial Gibbs free energy of interaction for the polar contribution of solute X to solvation in solvent S.

$$\Delta G_{\text{SQ}}^{\text{P}}(\text{X}) = \Delta G_{\text{SQ}}^{\text{Soln}}(\text{X}) - \Delta G_{\text{SQ}}^{\text{Soln}}(\text{HC})^V \quad (3)$$

Experimentally, evaluation of Eq. 2 requires the determination of the gas–liquid partition coefficient for solute X on the stationary phase S and the reference phase squalane as well as the determination of the gas–liquid partition coefficients for an appropriate number of *n*-alkanes on both phases to construct the linear relationship between the Van der Waals volume of the *n*-

alkanes and the logarithm of their gas–liquid partition coefficient. All the terms in Eq. 2 can then be evaluated. The contribution from cavity formation and dispersion interactions to the solvation process is represented by the sum of the first two terms on the right-hand side of Eq. 2. The $\Delta G_s^{Int}(X)$ parameter represents the sum of the polar interactions such as orientation and hydrogen bond formation to the solvation process. For simple molecules $\Delta G_s^{Int}(X)$ depends on the type of functional group present in a solute and can be treated as an incremental constant for prediction purposes. Poole and co-workers have used Eq. 2 to characterize the properties of 28 common stationary phases [6,9], to compare the solvation properties of 12 analogous alkanesulfonate and perfluoroalkanesulfonate liquid organic salts [12], and to study the influence of temperature on the selectivity of 10 stationary phases [13,14].

A comparison of the two models represented by Eqs. 1 and 2 shows acceptable agreement for the contribution of the sum of cavity formation and dispersion interactions (Eq. 4) and the contribution of polar interactions to the solvation process (Eq. 5) at 121.4°C [9,12,13,15,16]. There is, perhaps, a small numerical difference in the magnitude of the sum of cavity formation and dispersion interactions to solvation estimated by both models but this does not affect the agreement in general trends indicated by both models at a constant temperature. A more detailed comparison of the contributions of individual intermolecular interactions to solvation estimated by both models requires that a method be developed to identify these contributions to the free energy terms indicated in Eq. 2. This should be possible using chemometric techniques, such as principal component factor analysis, which are explored in this paper.

$$\Delta G_s^{Soln}(HC)^v + \Delta G_{SO}^P(X) = 1.806 \sum (c + l \log L^{16}) \quad (4)$$

$$\Delta G_s^{Int}(X) = 1.806 \sum (rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H) \quad (5)$$

The application of principal component analy-

sis and factor analysis to gas chromatographic data was pioneered by the research groups headed by Wold, Chastrette, Chretien and Howery and is reviewed in detail elsewhere [17–20]. In most cases the retention index data base of McReynolds was used to identify factors able to explain retention and to identify the similarity of solutes and solvents. The relationship between specific factors and fundamental intermolecular interactions was not made in these studies, perhaps in part due to the unreliable nature of some of the McReynolds data as well as the fact that the retention index is a composite term expressing properties of both the individual solutes and the retention index standards on the different stationary phases [4,9]. We have removed these objections in this paper by using the gas–liquid partition coefficients as the characteristic retention parameter.

2. Experimental

The name, abbreviation and composition of the stationary phases used in this study are summarized in Table 1. The gas–liquid partition coefficients at 121.4°C were taken from previous studies and are corrected for interfacial adsorption [3,15]. For data analysis it is necessary that the gas–liquid partition coefficient is accurately known for all phases in that data set. This resulted in the use of two data sets for evaluation purposes containing the results for 30 solutes on 22 phases and 67 solutes on 10 phases (the 10 phases are also contained in the 22-phase set).

To compare results obtained from the solvation parameter model with those obtained by the application of principal component factor analysis to the cavity model represented by Eq. 2, the appropriate gas–liquid partition coefficients equivalent to the Gibbs free energy terms with a molar standard state were used. For computational purposes the following expressions were used:

$$\log K_L(\text{cavity–dispersion}) = \log ([K_L^{NV}(X)]_s \cdot [K_L(X)]_{SO}) / [K_L^{NV}(X)]_{SO} \quad (6)$$

where $K_L^{NV}(X)$ is the gas–liquid partition coeffi-

Table 1
Identification and abbreviations for stationary phases

No.	Abbreviation	Name ^a
1	SQ	Squalane*
2	SE-30	Poly(dimethylsiloxane)
3	OV-105	Poly(cyanopropylmethyl dimethylsiloxane)*
4	OV-3	Poly(dimethylmethylphenylsiloxane), 10 mol% phenyl groups
5	OV-7	Poly(dimethylmethylphenylsiloxane), 20 mol% phenyl groups
6	OV-11	Poly(dimethylmethylphenylsiloxane), 35 mol% phenyl groups
7	OV-17	Poly(methylphenylsiloxane)*
8	OV-22	Poly(methylphenyldiphenylsiloxane), 65 mol% phenyl groups
9	OV-25	Poly(methylphenyldiphenylsiloxane), 75 mol% phenyl groups
10	OV-330	Poly(dimethylsiloxane)/Carbowax copolymer
11	OV-225	Poly(cyanopropylmethylphenylmethylsiloxane)*
12	QF-1	Poly(trifluoropropylmethylsiloxane)*
13	DDP	Didecylphthalate
14	PPE-5	1,3-bis(3-phenoxyphenoxy)benzene
15	CW-20M	Poly(ethylene glycol)*
16	U50HB	Poly(ethylene glycol) (Ucon 50 HB 660)
17	THPED	N,N,N',N'-Tetrakis(2-hydroxypropyl)ethylenediamine*
18	EGAD	Poly(ethylene glycol adipate)
19	DEGS	Poly(diethylene glycol succinate)*
20	TCEP	1,2,3-Tris(2-cyanoethoxypropane)*
21	QTS	Tetra- <i>n</i> -butylammonium 4-toluenesulfonate*
22	QBES	Tetra- <i>n</i> -butylammonium N,N-(bis-2-hydroxyethyl)-2-aminoethanesulfonate

^a An * indicates membership of the 10-phase data set.

coefficient for an *n*-alkane with an identical Van der Waals volume to solute X and is derived from the linear relationship between $\log K_L^{NV}$ and the Van der Waals volume, V_A , for the *n*-alkanes on each stationary phase.

$$\log K_L^{NV}(X) = m_s V_A + b_s \quad (7)$$

where m_s and b_s are the coefficients obtained by linear regression. $\log K_L(X)$ is the gas–liquid partition coefficient for solute X on stationary phase S or SQ as indicated by the subscript. The contribution of polar interactions is given by

$$\log K_L^{Int}(X) = \log \left(\frac{[K_L(X)]_S}{[K_L(X)]_{SQ}} \right) \cdot \frac{[K_L^{NV}(X)]_{SQ}}{[K_L^{NV}(X)]_S} \quad (8)$$

where $\log K_L^{Int}(X)$ is the gas–liquid partition

coefficient corresponding to the term $\Delta G_S^{Int}(X)$ in Eq. 2.

The Van der Waals volume for the test solutes were calculated with the molecular modeling program MacroModel 2.0 (Department of Chemistry, University of New York, New York, NY, USA) executed on a VAX 11/750 computer (Digital Equipment, Merrimack, NH, USA) [6]. Principal component and cluster analysis for data interpretation was performed using Pirouette V1.1 (Infometrix, Seattle, WA, USA) on a Epson Apex 200 computer (Epson America, Torrance, CA, USA). Raw varimax rotation without additional data preprocessing was used for the principal component factor analysis. The raw rotation method gave consistently better results than any of the weighted or normalized rotation techniques. The characteristic phase constants obtained from application of the solvation parameter model to the data sets discussed in this paper are taken from Ref. [11].

3. Results and discussion

The cavity model of solution provides a reasonable framework for understanding the retention properties of solutes in gas–liquid partition chromatography. It quite reasonably identifies two sets of factors underlying retention attributable to the contributions of solvent–solvent and solute–solvent interactions but does not identify the specific factors involved. These can be fairly confidently identified as dispersion, induction, orientation, hydrogen-bond acid–base, and perhaps electron donor–acceptor interactions (at least for isotropic solvents). The individual importance of these interactions for any system considered is a characteristic property of both the solvent and the solute, and in the absence of explicit information of the solute or solvent contribution, there is no simple approach to establishing the relevant complementary properties of the uncharacterized solvent (or solute) based on the singular observation of retention. One possible solution based on the assignment of characteristic properties to the solute, which are proportional to the specific factors, can be used to isolate the complementary solvent properties by multiple linear regression analysis, as discussed in a companion paper [11]. An alternative approach, based on exploratory data analysis, uses abstract mathematical constraints to reduce the dimensionality of the data to a smaller number of abstract factors retaining the useful information contained in the original data, which are subsequently converted to physically meaningful factors by a factor analytical procedure. There are many approaches to converting the abstract factors to meaningful physical factors as detailed in general texts on factor analysis [20–26], which will not be discussed here. Factor analysis, in general, is performed to acquire a new understanding of a problem and must be judged by logical chemical constraints since, in the end, the answers provided may be true or abstract factors, and the mathematical procedure used in their isolation will not be able to distinguish between the two alternatives.

In this paper we have used the principal component method of data reduction to identify

the abstract factors and the method of varimax rotation to convert the abstract factors to terminal factors for interpretation [21,22,25,26]. A characteristic property of the principal component method is that the first principal component will tend to be a weighted average of all the variables present in the data set and the second and succeeding principal components will tend to contain about equal loadings of opposite sign. In the absence of a single dominant general factor this is unlikely to represent a realistic factor solution for most chemical problems. The mathematical operation of rotation is employed to obtain a simpler structure which it is hoped will more closely resemble the true situation. By rotation the principal component axes are relocated within the factor hyperspace such that as many row points as possible lie close to the final factor axes, with only a small number of points remaining between the rotated axes. Varimax rotation is an example of an orthogonal rotation method that seeks simplicity among the factors by maximizing the total variance of the squared loadings. This is the method employed in these studies as implemented in the Pirouette software environment.

Principal component analysis of either the 10 or 22 stationary phase sets indicates that a single component accounts for >98% of the total variance of the contribution of cavity formation and dispersion interactions to the solvation process. After rotation of the principal components additional factor axis are produced that contain substantial variance (Table 2). All the principal component factors, however, are strongly correlated with each other (both within a rotation set and between different rotation sets), $r^2 > 0.95$. Also, the product of the score and loading coefficients for each solute on all phases are significantly correlated to $\sum(c + l \log L^{16})$ from the solvation parameter model. These and other tests indicate that rotation of the principal components produces phantom factors highly correlated to principal component 1 for the non-rotated factor matrix. It must be concluded that only a single factor can be isolated to describe the contribution of cavity formation and dispersion interactions in the cavity model. This is not

Table 2

Variance extracted by principal component factor analysis applied to the cavity formation and dispersion interaction term of the data set containing 23 solutes and 21 phases

Number of rotations	Principal component	Variance accounted for (%)
0	1	99.83
	2	0.14
	3	0.03
2	1	50.15
	2	49.81
	3	0.03
3	1	20.25
	2	48.15
	3	31.60

so surprising since the free energy involved in cavity formation and the free energy resulting from dispersion interactions in solution are both likely to be dependent on solute size.

The larger number of solutes available for the 10 stationary phase set allowed a greater level of flexibility in assigning the number of factors for the contribution of the polar interaction term to the cavity solution model. Principal component analysis indicated that about 98% of the variance could be accounted for by a single component which after rotation suggested that up to three principal component factors were significant (Table 3). Detailed evaluation of the factor loadings and deletion of certain solute types from the solute set was used to preliminary establish the identity of the factor axes. For the complete data set with three factor rotations the first principal component factor was heavily weighted towards aliphatic solutes capable of strong orientation interactions with dipolar aromatic solutes in an intermediate position and weakly dipolar aromatic and all solutes with a capacity to function as hydrogen-bond acids weakly loaded (Table 4). The second principal component factor was heavily weighted towards solutes with a capacity for hydrogen-bond acid interactions with all other solutes lightly loaded. The third principal component factor was heavily weighted towards aromatic solutes and weakly weighted towards aliphatic and aromatic hydrogen-bond acid solutes. There are two possible

Table 3

Variance extracted by principal component factor analysis applied to the polar interaction term for the data set containing 67 solutes on 10 phases

Principal component	Number of factor rotations		
	0	2	3
<i>Complete data set (67 solutes)</i>			
1	97.83	53.68	42.23
2	1.54	45.69	39.12
3	0.54	0.54	18.56
<i>Aliphatic compounds only (38 solutes)</i>			
1	98.56	75.06	75.40
2	1.24	24.74	20.88
3	0.11	0.11	3.64
<i>Aromatic compounds only (28 solutes)</i>			
1	98.55	55.28	54.17
2	1.18	44.45	23.84
3	0.02	0.21	21.93

assignments for this factor. The characteristic factor used to represent the cavity is the Van der Waals volume. The solvent accessible surface area of the cavity would be a more logical term to account for the number of solute–solvent interactions but this term can not be calculated in a straightforward manner for the types of solvent involved in these studies. The shape of the cavity for aromatic and other cyclic compounds is likely to be different from that for straight-chain aliphatic compounds and the cancellation of interactions proposed in the model for aliphatic compounds by an *n*-alkane of identical Van der Waals volume a better approximation than is obtained for the aromatic and cyclic compounds. Factor three, therefore, could arise in an artificial way to compensate for differences in the capacity of straight chain aliphatic compounds and aromatic (possibly also aliphatic cyclic compounds) to enter into solute–solvent intermolecular interactions. An alternative explanation is to assign principal component factor 3 to a capacity to enter into induction interactions based on the greater polarizability of aromatic compounds compared to aliphatic compounds. Deleting the aromatic compounds from the data set and repeating the principal com-

ponent factor analysis shows that factor three is substantially reduced in the percentage of the variance it contains (3.64% in the 3 factor rotation) and is heavily weighted to the single compound 1,1,2,2-tetrachloroethane. Considering just the aromatic compounds in the original data set then the principal component factor 3 is responsible for a significant fraction of the total variance (21.93% in the 3 factor rotation). In this case the identity of factors 1 and 2 are switched. Of course the variance represented in a data set is not independent of either the identity or the number of solutes in the data set so that the data entries cannot be directly compared in any quantitative sense. The changes are so large in this case that they can safely be used as a qualitative indication of the features involved.

Confirmation of the above indications was sought by comparison to the contributing factors isolated with Abraham's solvation parameter model (see [11]). The comparison was made using the data set for 22 phases and 30 solutes. Principal component factor 1 is well correlated to the product term $s\pi_2^H$ (Fig. 1), with a correlation coefficient $r^2 = 0.92$. The quality of the fit is influenced by the poor agreement of PPE-5 and DDP with the other phases. Principal component factor 2 is correlated to the product term $a\alpha_2^H$

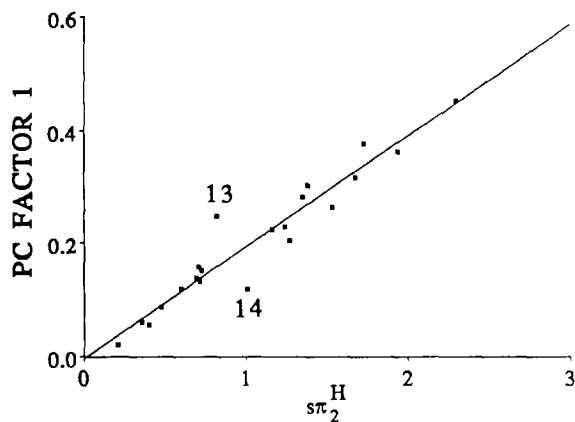


Fig. 1. Plot of principal component (PC) factor 1 extracted from the polar interaction term of the cavity model against the solvation parameter model parameter for dipole–dipole and dipole–induced dipole interactions for nitrobenzene. The numbered stationary phases are identified in Table 1.

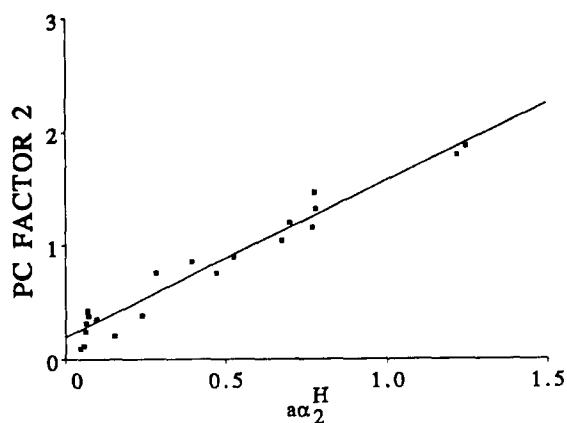


Fig. 2. Plot of principal component factor 2 extracted from the polar interaction term of the cavity model against the solvation parameter model parameter for solvent hydrogen-bond base interactions for octanol.

(Fig. 2), with a correlation coefficient $r^2 = 0.96$. Principal component factor 3 is correlated to the sum $\Sigma(s\pi_2^H + rR_2)$ (Fig. 3), $r^2 = 0.81$ with QF-1 considerably removed from the best line through the other phases (eliminating QF-1, $r^2 = 0.95$). The product term $s\pi_2^H$ is also well correlated to factor 3 but not rR_2 by itself. Inspection of the plots shows that combining $s\pi_2^H$ and rR_2 improves the fit of CW-20M, OV-22, OV-25 and U50HB with the best line through the data but

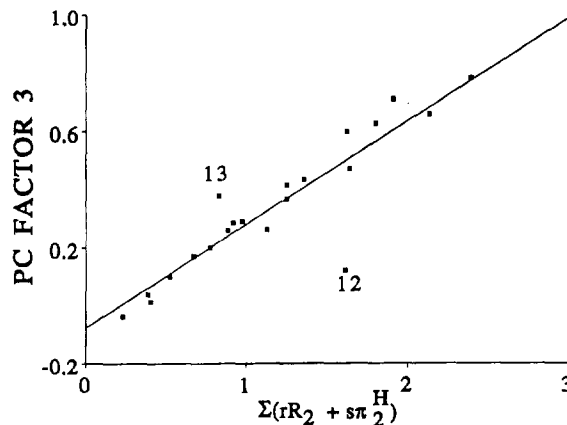


Fig. 3. Plot of principal component factor 3 extracted from the polar interaction term of the cavity model against the sum of the terms representing dipole–dipole, dipole–induced dipole and electron-pair interactions in the solvation parameter model for benzonitrile. The numbered stationary phases are identified in Table 1.

Table 4
Principal component factor loadings for the data set containing 67 solutes on 10 phases

Loading for factor 1		Loading for factor 2		Loading for factor 3	
Component	Loading	Component	Loading	Component	Loading
Dimethyl sulfoxide	0.3165	2-Chlorophenol	-0.4278	Iodobenzene	0.3212
Dimethylacetamide	0.2418	Phenol	-0.3636	2,6-Dimethylaniline	0.2578
Dimethylformamide	0.2348	4-Methylphenol	-0.3408	N,N-Dimethylaniline	0.2550
Dibutylformamide	0.2263	3,5-Dimethylphenol	-0.3363	Bromobenzene	0.2332
Methyl hexadecanoate	0.2030	2,5-Dimethylphenol	-0.3122	Benzodioxane	0.2285
Methyl tetradecanoate	0.1939	2,6-Dimethylphenol	-0.2192	1,2-Dichlorobenzene	0.2197
Dodecan-2-one	0.1904	Octan-1-ol	-0.2042	Anisole	0.2083
Methyl dodecanoate	0.1839	Nonan-1-ol	-0.1979	Benzene	0.2057
Nitrohexane	0.1815	Heptan-1-ol	-0.1957	n-Butylbenzene	0.2017
Methyl undecanoate	0.1774	Butan-1-ol	-0.1901	Chlorobenzene	0.1969
Undecan-2-one	0.1753	Hexan-1-ol	-0.1889	Benzaldehyde	0.1964
Methyl decanoate	0.1735	Pentan-1-ol	-0.1886	Acetophenone	0.1962
Methyl octadecanoate	0.1705	1,1,2,2-Tetrachloroethane	-0.1226	2,4,6-Trimethylpyridine	0.1914
Nonan-2-one	0.1695	Aniline	-0.1211	Ethylbenzene	0.1881
Decan-2-one	0.1693	2-Methyl-2-pentanol	-0.1198	Aniline	0.1838
Methyl nonanoate	0.1686	N-Methylaniline	-0.0918	N-Methylaniline	0.1835
Octan-2-one	0.1637	Dimethyl sulfoxide	-0.0455	Toluene	0.1713
Methyl octanoate	0.1629	2,6-Dimethylaniline	-0.0362	Pyridine	0.1697
Nitropentane	0.1627	Dodec-1-yne	-0.0346	Dioxane	0.1667
Nitropropane	0.1599	Nitrobenzene	-0.0327	cis-Hydrindane	0.1436
Heptan-2-one	0.1581	Benzonitrile	-0.0324	1,1,2,2-Tetrachloroethane	0.1313
Methyl heptanoate	0.1554	Nitrocyclohexane	-0.0289	Nitrobenzene	0.1305
Hexan-2-one	0.1532	Nitrohexane	-0.0273	Oct-2-yne	0.1265
Methyl hexanoate	0.1516	Dimethylformamide	-0.0165	Methyl octanoate	0.1105
Pentan-2-one	0.1460	1,2-Dichlorobenzene	-0.0158	2,6-Dimethylphenol	0.1040
Nonanal	0.1404	Dimethylacetamide	-0.0138	Benzonitrile	0.0940
Butan-2-one	0.1389	Nitropropane	-0.0111	Dihexyl ether	0.0683
Benzonitrile	0.1239	Dibutylformamide	-0.0102	Nitrocyclohexane	0.0560
Nitrocyclohexane	0.1111	Nitropentane	-0.0091	Butan-2-one	0.0559

Nitrobenzene	0.0845	Benzodioxane	-0.0066	2,5-Dimethylphenol	0.0483
Dioxane	0.0810	Methyl octadecanoate	-0.0066	Pentan-2-one	0.0430
Acetophenone	0.0631	Chlorobenzene	-0.0045	Dodec-1-yne	0.0397
2-Methyl-2-pentanol	0.0563	Iodobenzene	0.0006	Nitropropane	0.0342
Butan-1-ol	0.0562	Benzaldehyde	0.0023	Nitropentane	0.0296
Benzaldehyde	0.0560	Bromobenzene	0.0044	Hexan-2-one	0.0278
Pyridine	0.0539	Acetophenone	0.0081	3,5-Dimethylphenol	0.0235
Octan-1-ol	0.0480	Pyridine	0.0104	Methyl heptanoate	0.0222
Nonan-1-ol	0.0454	Ethylbenzene	0.0121	4-Methylphenol	0.0189
Dihexyl ether	0.0369	Toluene	0.0134	Methyl hexanoate	0.0157
2,4,6-Trimethylpyridine	0.0363	Nonanal	0.0137	Heptan-2-one	0.0150
Heptan-1-ol	0.0348	Dodecan-2-one	0.0147	Nonanal	0.0149
Hexan-1-ol	0.0339	Methyl tetradecanoate	0.0156	Dimethylformamide	0.0137
Pentan-1-ol	0.0290	Methyl hexanoate	0.0159	Dimethylacetamide	0.0101
Benzodioxane	0.0252	Undecan-2-one	0.0186	Octan-2-one	0.0065
Dodec-1-yne	0.0241	Methyl dodecanoate	0.0200	Phenol	0.0053
N,N-Dimethylaniline	0.0168	Nonan-2-one	0.0202	Methyl nonanoate	0.0008
Anisole	0.0137	<i>cis</i> -Hydrindane	0.0213	Decan-2-one	-0.0048
Aniline	0.0123	Decan-2-one	0.0214	Nonan-2-one	-0.0052
2-Chlorophenol	0.0091	Methyl undecanoate	0.0224	Methyl decanoate	-0.0086
2,6-Dimethylaniline	0.0078	Oct-2-yne	0.0237	2-Methyl-2-pentanol	-0.0163
N-Methylaniline	0.0006	Octan-2-one	0.0239	Undecan-2-one	-0.0166
Toluene	0.0006	Methyl decanoate	0.0244	Methyl undecanoate	-0.0170
Phenol	-0.0003	Heptan-2-one	0.0266	Dimethyl sulfoxide	-0.0205
Oct-2-yne	-0.0077	Methyl nonanoate	0.0280	Nitrohexane	-0.0209
<i>n</i> -Butylbenzene	-0.0091	2,4,6-Trimethylpyridine	0.0284	Dibutylformamide	-0.0224
4-Methylphenol	-0.0152	Methyl octanoate	0.0295	Methyl dodecanoate	-0.0289
3,5-Dimethylphenol	-0.0209	Hexan-2-one	0.0299	2-Chlorophenol	-0.0318
2,6-Dimethylphenol	-0.0226	Benzene	0.0300	Hexan-1-ol	-0.0336
Benzene	-0.0244	Methyl heptanoate	0.0302	Dodecan-2-one	-0.0383
Ethylbenzene	-0.0252	Anisole	0.0313	Heptan-1-ol	-0.0444
1,1,2,2-Tetrachloroethane	-0.0254	Methyl hexanoate	0.0325	Methyl octadecanoate	-0.0446
Chlorobenzene	-0.0284	Butan-2-one	0.0332	Methyl tetradecanoate	-0.0485
2,5-Dimethylphenol	-0.0315	Pentan-2-one	0.0332	Butan-1-ol	-0.0506
1,2-Dichlorobenzene	-0.0347	Dihexyl ether	0.0332	Nonan-1-ol	-0.0616
Bromobenzene	-0.0516	<i>n</i> -Butylbenzene	0.0395	Methyl hexadecanoate	-0.0679
<i>cis</i> -Hydrindane	-0.0761	N,N-Dimethylaniline	0.0437	Octan-1-ol	-0.0698
Iodobenzene	-0.0761	Dioxane	0.0485	Pentan-1-ol	-0.2852

QF-1 is still in poor agreement with the other phases. Numerically, rR_2 is generally small compared with $s\pi_2^H$ so that its ability to significantly influence the correlation is not very great. Its inclusion with the properties of factor 3 seems reasonable but is not essential in defining the meaning of factor 3. The π_2^H parameter is a dipole/polarizability parameter and its correlation with properties of factor 1 and 3 is not inconsistent with factor 1 representing the properties of dipolar aliphatic compounds and factor 3 the dipolar and polarizability properties of aromatic compounds. For example benzene and cyclohexane are not considered to have significant dipole character but benzene is significantly more polarizable than cyclohexane and this is reflected in their π_2^H parameters of 0.52 for benzene and 0.10 for cyclohexane. The correlation coefficients are independent of the solute's identity. There is good agreement between the product of the scores and loading coefficients for the 22-phase data set and the 10-phase data set for the solutes they have in common that are highly loaded on the individual factor axes. The

score coefficients correlate with the phase constants of Abraham's model for both data sets but the absolute value of the score coefficients depends on the number of data entries and, for this reason, we have made the comparison using the product terms in both models. Thus the original assignments that factor 1 is determined by the capacity for orientation type interactions, factor 2 for solvent hydrogen-bond base interactions, and factor 3 for either induction interactions or aromatic character are reasonable.

Based on these solutions a ranking of stationary phases by their capacity for different intermolecular interactions can be given (Table 5). The relative contributions are scaled within a column but not across rows. The phase QF-1 has significantly different contributions for orientation interactions for aromatic and aliphatic compounds and is omitted to avoid using two scales for this interaction. The score coefficients for principal component factor 1 and principal component factor 3 for the polar interactions obtained from the cavity model are otherwise highly correlated, $r^2 = 0.96$, and differences in

Table 5

A ranking of stationary phases by their capacity for specific interactions based on the score coefficients obtained by principal component factor analysis

Stationary phase	Cavity and dispersion	Stationary phase	Orientation	Stationary phase	Hydrogen-bond basicity
DDP	10.39	TCEP	3.94	QBES	3.10
U50HB	10.03	DEGS	3.26	QTS	2.96
OV-105	9.89	QBES	3.14	TCEP	2.42
OV-3	9.87	QTS	2.74	DEGS	2.18
OV-7	9.79	CW-20M	2.62	CW-20M	1.98
SE-30	9.79	OV-225	2.44	THPED	1.91
OV-11	9.55	EGAD	2.28	EGAD	1.72
OV-17	9.38	DDP	2.15	OV-330	1.48
PPE-5	9.14	THPED	1.98	OV-225	1.40
OV-22	8.99	OV-330	1.94	DDP	1.25
OV-25	8.85	OV-25	1.38	U50HB	1.24
THPED	8.68	OV-22	1.31	OV-25	0.72
OV-330	8.59	U50HB	1.20	OV-22	0.64
OV-225	7.69	OV-17	1.15	PPE-5	0.64
QTS	7.57	OV-11	1.04	OV-17	0.59
EGAD	7.57	PPE-5	1.03	OV-11	0.53
CW-20M	7.31	OV-7	0.76	OV-7	0.42
QBES	5.70	OV-3	0.53	OV-105	0.34
DEGS	5.48	OV-105	0.48	OV-3	0.20
TCEP	5.00	SE-30	0.18	SE-30	0.16

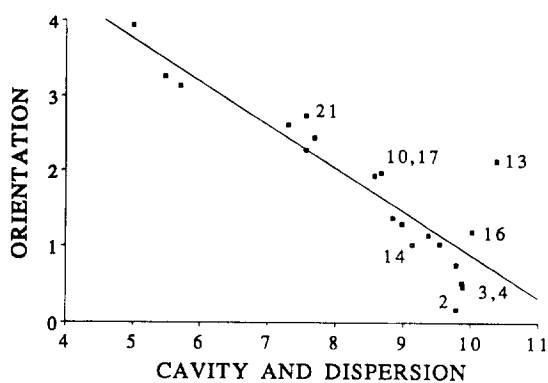


Fig. 4. Plot of the score coefficients representing orientation against the cavity and dispersion contribution from the cavity model. The numbered stationary phases are identified in Table 1.

the ranking are restricted to minor changes in order between coefficients with similar values.

Selectivity differences between phases can be illustrated by plotting the various score coefficients against each other. The plot of orientation against the cavity and dispersion contribution to solution (Fig. 4), indicates the general correlation between the cohesive character of the solvent and its propensity for orientation interactions. TCEP, DEGS and QBES are cohesive solvents with the least favorable contributions for cavity formation. Solutes with a limited capacity for polar interactions will be retained weakly compared to the other phases. DDP is

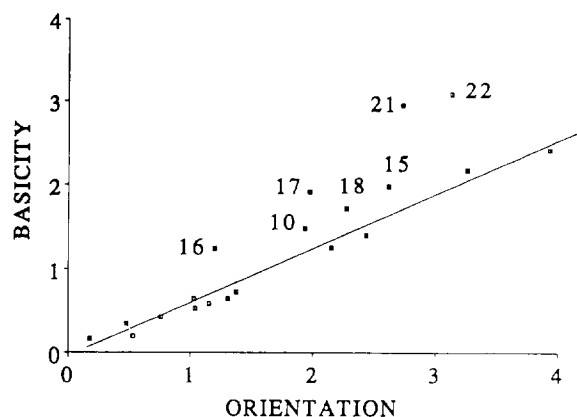


Fig. 5. Plot of the score coefficients representing hydrogen-bond basicity against orientation contributions from the cavity model. The numbered stationary phases are identified in Table 1.

unique in that it has modest demands for cavity formation compared to other solvents of similar polarity. The plot of hydrogen-bond basicity against orientation contributions to solution (Fig. 5) indicates that QBES, QTS, THPED, U50HB, OV-330, CW-20M and EGAD are more basic than other phases of similar orientation capacity. In particular, the two liquid organic salts, QTS and QBES, are strong hydrogen-bond bases. This is expected since they are the only ionic phases present in the data set. As would be anticipated there is no general correspondence between the capacity of a phase to function as a hydrogen-bond base and the cavity and dispersion properties of the same phase. The cohesive character of the solvent will be little influenced by its capacity for hydrogen-bond basicity unless it is simultaneously a hydrogen-bond acid. Since none of the phases were identified as significant hydrogen-bond acids by principal component factor analysis this should not be the general case. The absence of significant hydrogen-bond acidity among the phases studied at the measurement temperature is consistent with other approaches to characterizing the properties of these phases [11].

For the series of poly(methylphenyl)siloxane phases in Table 5 some correspondence between the polar solvent properties of the stationary phases and its composition would be anticipated. This is the general case as illustrated by Eqs. 9–11 below. The individual solvent contributions represented by the score coefficients are well fitted to a second order polynomial function with mol% phenyl groups [P] as the independent variable over the range 0 mol% phenyl groups (SE-30) to 75 mol% phenyl groups (OV-25).

Cavity and dispersion

$$\text{Score} = 9.84 - 1.17 \cdot 10^{-3}[\text{P}] - 1.69 \cdot 10^{-4}[\text{P}]^2$$

$$n = 7, r^2 = 0.984 \quad (9)$$

Orientation

$$\text{Score} = 0.22 + 2.94 \cdot 10^{-2}[\text{P}] - 1.91 \cdot 10^{-4}[\text{P}]^2$$

$$n = 7, r^2 = 0.992 \quad (10)$$

Hydrogen-bond basicity

$$\text{Score} = 0.14 + 1.29 \cdot 10^{-2}[\text{P}] - 7.39 \cdot 10^{-5}[\text{P}]^2$$

$$n = 7, r^2 = 0.966 \quad (11)$$

Principal component factor analysis is a valuable technique for providing insight into the factors affecting the characteristic solvent properties of gas chromatographic stationary phases. Its main limitation is that the solutions provided may not be real in the chemical sense and it is difficult or impossible to distinguish between qualitative factors correlated to a significant extent with real factors. Within these confines the factor solutions extracted from the cavity model seem to be realistic and useful for interpretative purposes. Compared to previous studies where principal component analysis alone was used for interpretation [9,12–16] the factors obtained by rotation are more sensible and are to be preferred.

References

- [1] C.E. Figgins, T.H. Risby and P.C. Jurs, *J. Chromatogr. Sci.*, 14 (1976) 453.
- [2] R.V. Golovnya and T.A. Misharina, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 3 (1980) 51.
- [3] B.R. Kersten, S.K. Poole and C.F. Poole, *J. Chromatogr.*, 468 (1989) 235.
- [4] C.F. Poole and S.K. Poole, *Chem. Rev.*, 89 (1989) 377.
- [5] T.O. Kollie and C.F. Poole, *J. Chromatogr.*, 550 (1991) 213.
- [6] T.O. Kollie and C.F. Poole, *J. Chromatogr.*, 556 (1991) 457.
- [7] M.H. Abraham, G.S. Whiting, R.M. Doherty and W.J. Shuely, *J. Chromatogr.*, 587 (1991) 229.
- [8] M.H. Abraham, P.L. Grellier, I. Hamerton, R.A. McGill, D.V. Prior and G.S. Whiting, *Faraday Discuss. Chem. Soc.*, 85 (1988) 107.
- [9] C.F. Poole, T.O. Kollie and S.K. Poole, *Chromatographia*, 34 (1992) 281.
- [10] P.W. Carr, *Microchem. J.*, 48 (1993) 4.
- [11] S.K. Poole and C.F. Poole, *J. Chromatogr. A*, 697 (1995) 415.
- [12] T.O. Kollie and C.F. Poole, *Chromatographia*, 33 (1992) 551.
- [13] C.F. Poole and T.O. Kollie, *Anal. Chim. Acta*, 282 (1993) 1.
- [14] S.K. Poole, T.O. Kollie and C.F. Poole, *J. Chromatogr. A*, 664 (1994) 229.
- [15] T.O. Kollie, C.F. Poole, M.H. Abraham and G.S. Whiting, *Anal. Chim. Acta*, 259 (1992) 1.
- [16] M.H. Abraham, J. Andonian-Haftvan, I. Hamerton, C.F. Poole and T.O. Kollie, *J. Chromatogr.*, 646 (1993) 351.
- [17] D.L. Massart and H.L.O. De Clercq, *Adv. Chromatogr.*, 16 (1978) 75.
- [18] J.R. Chretien, *Trends Anal. Chem.*, 6 (1987) 275.
- [19] J.R. Chretien, M. Righezza, A. Hassani and B.Y. Meklati, *J. Chromatogr.*, 609 (1992) 261.
- [20] E.R. Malinowski, *Factor Analysis in Chemistry*, Wiley, New York, 1991.
- [21] R.R. Reymont and K.G. Joreskog, *Applied Factor Analysis in the Natural Sciences*, Cambridge Univ. Press, Cambridge, UK, 1993.
- [22] A.L. Comrey and H.B. Lee, *A First Course in Factor Analysis*, Lawrence Erlbaum Associates, Hillsdale, NJ, 2nd ed., 1992.
- [23] J.E. Jackson, *A User's Guide to Principal Components*, Wiley, New York, 1991.
- [24] I.T. Jolliffe, *Principal Component Analysis*, Springer, New York, 1986.
- [25] H.H. Harman, *Modern Factor Analysis*, University of Chicago Press, Chicago, IL, 1976.
- [26] M. Forina, S. Lanteri and R. Leardi, *Trends Anal. Chem.*, 6 (1987) 250.