*Journal of Chromatography*, 550 (1990) 213–237
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 1999

# Application of multivariate analysis to the selection of test solutes for characterizing stationary phase selectivity in gas chromatography

THEOPHILUS O. KOLLIE and COLIN F. POOLE*
*Department of Chemistry, Wayne State University, Detroit, MI 48202 (U.S.A.)*

## ABSTRACT

Factors controlling the retention of 28 solutes on 23 stationary phases are identified by multivariate analysis and related to dominant intermolecular interactions. The selectivity of the stationary phases is characterized by the partial molal Gibbs free energy of solution for specific test solutes identified by principal component analysis; nitrobenzene for orientation interactions and *n*-octanol for solvent proton acceptor interactions. No test solute with acceptable certainty was identified for solvent proton donor interactions. The partial molar Gibbs free energy of solution for a methylene group is a convenient parameter for assessing dispersive interactions. For highly cohesive phases, such as OV-275, TCEP, DEGS, QACES and QTAPSO the selectivity parameters were found to be solute-size dependent. This size dependence can be removed by separating the free energy into a cavity term and an interaction term; the interaction term being independent of solute size. Using principal component analysis and cluster analysis the 23 stationary phases were classified into 5 groups based on the similarity of their capacity for specific intermolecular interactions with the phases squalane, QF-1, OV-225, OV-275 and QTAPSO behaving independently. The use of dendrograms is demonstrated to be a useful method for visualizing selectivity differences for chromatographic optimization.

## INTRODUCTION

A very large number of liquid phases have been suggested for use in gas–liquid chromatography (GLC), easily exceeding the capability of any laboratory to stock all, or even a large proportion of these phases [1,2]. Among this collection of phases are many materials which simply duplicate the properties of each other or are industrial materials of poorly defined chemical composition. The bewildering number of phases and the non-standard format used to specify their separation properties has done little to aid the selection of a particular phase for a given separation problem. Even when repeating a literature separation the phase specified may not be in the collection of phases used by a particular laboratory, is not a well characterized phase (so that even if purchased it may not adequately reproduce the original separation), or the material may no longer be commercially available. In these cases, a readily available substitute phase of well defined composition would be preferred, if it could be easily identified. Consequently, there is a general need for a classification method that enables stationary phases to be grouped by their similarity, such that a smaller number of phases could be selected for general use that adequately represent the separation characteristics of all available phases. Phases of a high degree of similarity could be

0021-9673/91/$03.50 

replaced by a single member of the group with the most favorable chromatographic properties (identification of duplicates) and preferred phases from groups with little similartiy (selective phases) chosen for initial exploration in optimizing a separation. This would minimize the number of phases required for general analysis to, probably, a manageable number, and given the general high efficiency of modern GLC, be quite adequate for most situations. This approach is apparent in the literature usage of fused-silica open tubular columns, although in this case selection is based on the limitations of siloxane synthesis, more so than of fundamental considerations of the type and range of intermolecular forces important in GLC [3].

In fact, numerous methods have been discussed for classifying stationary phases, and are reviewed in detail elsewhere [2–7]. Of these methods the system of stationary phase selectivity constants introduced by Rohrschneider [6] and extended by McReynolds [7] has become the most widely used. Rohrschneider published data for 30 test solutes on 23 stationary phases [8] and McReynolds data for 68 test solutes on 25 stationary phases [9] and 10 test solutes on 226 stationary phases [7]. These data sets have been analyzed by multivariate analysis techniques to determine the minimum number of test solutes required to characterize the retention properties of the stationary phases [10–13] and to classify the stationary phases into groups with similar separation properties [11,14–22]. The application of multivariate analysis to chromatography is reviewed in refs. 12 and 23 and in a wider sense to analytical chemistry in refs. 23–27. These methods of factor analysis, principal component analysis, and cluster analysis are ideally suited to deconvoluting large data sets to reveal common, independent variables and for comparing data subsets by a ranking based on similarity. The data of Rohrschneider and McReynolds are generally given in the form of retention index values or as retention index differences with respect to squalane as a reference phase. Rohrschneider suggested that 5 test solutes [6,8] and McReynolds 10 solutes [7] were required to characterize the solvent properties of the liquid phases as judged by the agreement between the predicted and experimental retention indices. Lowry *et al.* [10] suggested that three solutes chosen from the first five test solutes of McReynolds while Fernandez-Sanchez *et al.* [13] found that seven test solutes from the full set of 10 McReynolds test solutes were adequate to reproduce the retention index values with acceptable accuracy. Although the justification for the identity of the test solutes has generally been associated with particular intermolecular interactions, such as dispersion, induction, orientation, and proton donor–acceptor properties, these studies concluded that the identity of the solute was less important than the number of solutes selected for the evaluation. We believe that the two processes of characterizing the importance of specific intermolecular interactions and reproducing retention index values are different processes, since the attempt to increase the precision of the retention index values results in the addition of more test solutes than is logically needed to express the dominant intermolecular interactions.

Two factors were determined to account for about 98% of the total variance in the McReynolds data matrix by principal component analysis [3,11,15,18]. These axes were stated to be general polarity and hydrogen bonding interactions. It seems illogical that so much of the properties of the data matrix could be represented by just two factors given the general complexity of solute–solvent interactions, and therefore, other reasons must exist for the agreement found. Apart from the Rohrschneider/McReynolds data only one other study has used multivariate analysis techniques to

characterize the solvent properties of a group of common liquid phases [28]. A group of 10 chlorophenoxy acid methyl or pentafluorobenzyl esters on 10 stationary phases was used to identify similarities between phases. In this case the limited choice of test solutes provides little information about the selectivity of the phases in terms of intermolecular interactions.

The problem of identifying the relative contribution of intermolecular interactions to retention in GLC, as we see it, is not related to the choice of data analysis technique, but the unreliability of the data sets used for analysis. The data presentation of Rohrschneider/McReynolds is fundamentally flawed when used for characterizing stationary phase selectivity, in a manner which does not preclude its use for predicting retention index values [4,5,29,30]. The following problems with the McReynolds stationary phase selectivity constants for characterizing stationary phase properties have been identified:

(1) Many of the test solutes are too volatile to provide accurate retention values on a wide range of phases having different polarity. Some solutes elute at, or close to, the column dead volume [30–32].

(2) The method used to calculate retention indices ignores the contribution made by interfacial adsorption. Retention index values vary with the ratio of the stationary phase surface area/bulk volume and those parameters which effect this ratio, for example, coating efficiency, support type, etc. The retention index for polar solutes on non-polar phases depends on the degree of support deactivation. On polar phases the *n*-alkane retention index markers are retained almost exclusively by interfacial adsorption and the retention index values are meaningless [30,31,33–38].

(3) The retention index differences used as phase constants are composite terms the magnitude of which depends on both the retention of the index standards as well as that of the test solutes. Using different retention index standards results in different ranking of the phases by selectivity. The magnitude of the phase constants is determined largely by the retention of the *n*-alkanes on most phases and very good correlations exist between the magnitude of the phase constants and the partial molar Gibbs free energy of solution for a methylene group for a wide range of phases [29,30,35,39].

(4) The experimental data compiled by McReynolds are insufficiently accurate for stationary phase characterization. The stationary phase loading is not accurately determined, a wetting agent was added to each phase at the 2% (w/w) level, and several index values were determined retrospectively using an indirect method (bracketing hydrocarbons were not used) [4].

These problems preclude the further use of the McReynolds data set for evaluating stationary phase interactions. Such a data set, like the original, has many uses in GLC and we have started to generate a new data collection. Unlike those of Rohrschneider and McReynolds the data are standardized in the form of the gas–liquid partition coefficient, corrected for interfacial adsorption, and determined under experimental conditions where gas phase imperfections can be safely neglected [5,32]. Stationary phase selectivity is determined by the magnitude of the partial molal Gibbs free energy of solution for a series of test solutes chosen to express the principal intermolecular interactions. The purpose of this paper is to outline the solute selection procedure using multivariate analysis of the partial molal Gibbs free energy of solution for 28 test solutes on 23 stationary phases.

## EXPERIMENTAL

Data for the partial molal Gibbs free energy of solution for the test solutes on the stationary phases identified in Table I were taken from refs. 5, 32, 40 and 41 and are summarized in Table II for the convenience of the reader. All measurements were made at 121.4°C. Multivariate analysis was performed on an Epson Apex 200 computer using Ein*Sight™ version 2.5 (Infometrix, Seattle, WA, U.S.A.) software for data analysis and pattern recognition. The data was entered via a standard spreadsheet program, VP-Planner, version 2.0 (Paperback Software International, Berkeley, CA, U.S.A.). Missing data points were added as follows: for 1,1,2,2-tetrachloroethane on OV-330 the mean, QMES the maximum, QACES the mean and QTAPSO 0.67 of the mean; for dodecafluoroheptanol on QMES the maximum, QACES the mean and QTAPSO 0.5 of the mean; and for 2,4,6-trimethylpyridine on QpTS the maximum (where the mean is the mean of all values for that test solute and the maximum is the

TABLE I

IDENTIFICATION AND ABBREVIATIONS FOR STATIONARY PHASES AND TEST SOLUTES

| Stationary phases | | | Test solutes | |
|---|---|---|---|---|
| No. | Abbreviation | Name | No. | Name |
| 1 | SQ | Squalane | 1 | Partial molar Gibbs free energy of solution for a methylene group |
| 2 | OV-3 | Poly(dimethylmethylphenylsiloxane) 10 mol% phenyl groups | 2 | Benzene |
| 3 | OV-7 | Poly(dimethylmethylphenylsiloxane) 20 mol% phenyl groups | 3 | *n*-Butanol |
| 4 | OV-11 | Poly(dimethylmethylphenylsiloxane) 35 mol% phenyl groups | 4 | 1-Nitropropane |
| 5 | OV-17 | Poly(methylphenylsiloxane) | 5 | Pyridine |
| 6 | OV-22 | Poly(methylphenyldiphenylsiloxane) 65 mol% phenyl groups | 6 | 2-Methyl-2-pentanol |
| 7 | OV-25 | Poly(methylphenyldiphenylsiloxane) 75 mol% phenyl groups | 7 | 2-Octyne |
| 8 | OV-105 | Poly(cyanopropylmethyldimethylsiloxane) | 8 | 1,4-Dioxane |
| 9 | OV-225 | Poly(cyanopropylmethylphenylmethylsiloxane) | 9 | *cis*-Hydrindane |
| 0 | OV-275 | Poly(dicyanoallylsiloxane) | 10 | *n*-Butylbenzene |
| 1 | OV-330 | Poly(dimethylsiloxane)–Carbowax copolymer | 11 | 1-Nitropentane |
| 2 | QF-1 | Poly(trifluoropropylmethylsiloxane) | 12 | Nitrobenzene |
| 3 | CW20M | Poly(ethylene glycol) | 13 | Octanol |
| 4 | DEGS | Poly(diethylene glycol succinate) | 14 | Benzodioxane |
| 5 | TCEP | 1,2,3-Tris(2-cyanoethoxypropane) | 15 | Dihexyl ether |
| 6 | PPE-5 | 1,3-Bis(3-phenoxyphenoxy)benzene | 16 | 1-Dodecyne |
| 7 | QpTS | Tetra-*n*-butylammonium 4-toluenesulphonate | 17 | Dodecane |
| 8 | QPIC | Tetra-*n*-butylammonium 4-picrate | 18 | Benzonitrile |
| 9 | QMES | Tetra-*n*-butylammonium 4-morpholineethanesulphonate | 19 | 1,1,2,2-Tetrachloroethane |
| 0 | QACES | Tetra-*n*-butylammonium 2-(2-acetamido)aminoethane-sulphonate | 20 | 2,4,6-Trimethylpyridine |
| 1 | QTAPSO | Tetra-*n*-butylammonium 3-tris(hydroxymethyl)methyl-amino-2-hydroxy-1-propanesulphonate | 21 | Aniline |
| 2 | DDP | Didecylphthalate | 22 | N-Methylaniline |
| 3 | SE-30 | Poly(dimethylsiloxane) | 23 | N,N-Dimethylaniline |
| | | | 24 | 2,6-Dimethylaniline |
| | | | 25 | Anisole |
| | | | 26 | Nonanal |
| | | | 27 | 2-Octanone |
| | | | 28 | Dodecafluoroheptanol |

TABLE II

PARTIAL MOLAL GIBBS FREE ENERGY OF SOLUTION FOR TEST SOLUTES ON 23 STATIONARY PHASES (cal/mol)

| | $\Delta G_k^0$ ($CH_2$) | Dioxane | 1-Butanol | Nitro-propane | Nitro-pentane | Nitro-benzene |
|---|---|---|---|---|---|---|
| SQ | −521 | −3139 | −2708 | −3137 | −4211 | −4416 |
| OV-3 | −458 | −2976 | −2733 | −3150 | −4078 | −4814 |
| OV-7 | −467 | −3040 | −2733 | −3253 | −4201 | −4986 |
| OV-11 | −475 | −3073 | −2688 | −3329 | −4260 | −5117 |
| OV-17 | −470 | −3104 | −2680 | −3349 | −4275 | −5181 |
| OV-22 | −458 | −3091 | −2586 | −3337 | −4220 | −5183 |
| OV-25 | −431 | −3057 | −2777 | −3301 | −4175 | −5175 |
| OV-105 | −461 | −2919 | −2756 | −3171 | −4097 | −4757 |
| OV-225 | −418 | −3163 | −2967 | −3778 | −4640 | −5497 |
| OV-275 | −279 | −3046 | −2890 | −3697 | −4185 | −5257 |
| OV-330 | −418 | −3284 | −3405 | −3775 | −4613 | −5675 |
| QF-1 | −390 | −2788 | −2389 | −3467 | −4273 | −4838 |
| CW20M | −400 | −3379 | −3533 | −3928 | −4671 | −5914 |
| DEGS | −324 | −3275 | −3143 | −3672 | −4306 | −5552 |
| TCEP | −291 | −3577 | −3386 | −4127 | −4651 | −5808 |
| PPE-5 | −487 | −3420 | −2888 | −3604 | −4578 | −5714 |
| QpTS | −377 | −3300 | −4632 | −4319 | −5062 | −6234 |
| QPIC | −411 | −3439 | −3439 | −4136 | −4965 | −6049 |
| QMES | −398 | −3291 | −4736 | −4293 | −5040 | −6209 |
| QACES | −319 | −3148 | −4351 | −4005 | −4592 | −5909 |
| QTAPSO | −274 | −3007 | −3809 | −3722 | −4261 | −5576 |
| DDP | −511 | −3386 | −3312 | −3802 | −4886 | −5779 |
| SE-30 | −463 | −2899 | −2672 | −2982 | −3914 | −4519 |

| | 1-Octanol | Benzo-dioxane | Dihexyl ether | Benzene | *n*-Butyl-benzene | *cis*-Hydrin-dane |
|---|---|---|---|---|---|---|
| SQ | −4623 | −5497 | −6270 | −3156 | −5201 | −5023 |
| OV-3 | −4607 | −5151 | −5475 | −2803 | −4581 | −4313 |
| OV-7 | −4643 | −5333 | −5472 | −2814 | −4622 | −4317 |
| OV-11 | −4631 | −5484 | −5398 | −2789 | −4621 | −4268 |
| OV-17 | −4604 | −5562 | −5314 | −2690 | −4599 | −4216 |
| OV-22 | −4471 | −5600 | −5122 | −2748 | −4517 | −4121 |
| OV-25 | −4393 | −5630 | −5034 | −2816 | −4497 | −4061 |
| OV-105 | −4609 | −5024 | −5374 | −2739 | −4488 | −4200 |
| OV-225 | −4771 | −5604 | −4668 | −2691 | −4302 | −3668 |
| OV-275 | −3867 | −5305 | −2182 | −2148 | −3107 | −2358 |
| OV-330 | −5198 | −5970 | −4958 | −2897 | −4566 | −3968 |
| QF-1 | −3988 | −4573 | −4317 | −2262 | −3770 | −3350 |
| CW20M | −5233 | −6185 | −4447 | −2917 | −4396 | −3650 |
| DEGS | −4542 | −5838 | −3240 | −2475 | −3708 | −2807 |
| TCEP | −4611 | −5957 | −3215 | −2703 | −3766 | −2751 |
| PPE-5 | −4933 | −6062 | −5395 | −2979 | −4840 | −4329 |
| QpTS | −6317 | −6205 | −4300 | −2870 | −4291 | −3577 |
| QPIC | −5177 | −6009 | −4407 | −2916 | −4426 | −3654 |
| QMES | −6406 | −6172 | −4316 | −2837 | −4267 | −3537 |
| QACES | −5738 | −5878 | −3123 | −2543 | −3652 | −2868 |
| QTAPSO | −5020 | −5558 | −2995 | −2385 | −3358 | −2635 |
| DDP | −5013 | −6043 | −6002 | −3204 | −5177 | −3633 |
| SE-30 | −4554 | −4944 | −5468 | −2772 | −4538 | −4332 |

*(Continued on p. 218)*

TABLE II *(continued)*

| | 2-Octyne | 1-Dodecyne | 2-Methyl-2-pentanol | Dodeca-fluoro-heptanol | Benzo-nitrile | 1,1,2,2-Tetra-chloroethane |
|---|---|---|---|---|---|---|
| SQ | −4154 | −5942 | −3311 | −3194 | −4455 | −4319 |
| OV-3 | −3677 | −5256 | −3023 | −3220 | −4348 | −3968 |
| OV-7 | −3691 | −5265 | −3019 | −3220 | −4450 | −4099 |
| OV-11 | −3647 | −5206 | −2948 | −3131 | −4639 | −4149 |
| OV-17 | −3607 | −5139 | −2895 | −2996 | −4696 | −4169 |
| OV-22 | −3526 | −4950 | −2847 | −2969 | −4711 | −4139 |
| OV-25 | −3493 | −4950 | −2708 | −2745 | −4701 | −4126 |
| OV-105 | −3592 | −5163 | −3042 | −3551 | −4313 | −3932 |
| OV-225 | −3150 | −4644 | −3017 | −3997 | −5063 | −4410 |
| OV-275 | −1940 | −2439 | −2371 | −3391 | −4897 | −4043 |
| OV-330 | −3549 | −5032 | −3343 | −4723 | −5213 | −4331 |
| QF-1 | −2703 | −4164 | −2669 | −3417 | −4468 | −3350 |
| CW20M | −3277 | −4764 | −3327 | −4599 | −5473 | −5046 |
| DEGS | −2560 | −3824 | −2946 | −4077 | −5090 | −4449 |
| TCEP | −2587 | −3534 | −3067 | −4175 | −5415 | −4520 |
| PPE-5 | −3699 | −5277 | −3161 | −3168 | −5117 | −4547 |
| QpTS | −3052 | −4703 | −4198 | −6992 | −5791 | −5606 |
| QPIC | −2991 | −4640 | −3308 | −4498 | −5602 | −4671 |
| QMES | −2992 | −4723 | −4281 | −6992 | −5761 | −5606 |
| QACES | −2369 | −3740 | −3732 | −3767 | −5500 | −4331 |
| QTAPSO | −2009 | −3257 | −3323 | −1883 | −5179 | −2902 |
| DDP | −4008 | −5829 | −3598 | −3767 | −5267 | −4396 |
| SE-30 | −3673 | −5239 | −3009 | −3176 | −4134 | −3826 |

| | Pyridine | 2,4,6-Tri-methyl-pyridine | Aniline | N-Methyl-aniline | N,N-Di-methyl-aniline | 2,6-Dimethyl-aniline |
|---|---|---|---|---|---|---|
| SQ | −3396 | −4784 | −4463 | −5029 | −5286 | −5567 |
| OV-3 | −3222 | −4285 | −4282 | −4673 | −4779 | −5131 |
| OV-7 | −3222 | −4368 | −4429 | −4821 | −4895 | −5285 |
| OV-11 | −3366 | −4416 | −4549 | −4938 | −4984 | −5406 |
| OV-17 | −3395 | −4431 | −4616 | −4999 | −5020 | −5468 |
| OV-22 | −3457 | −4403 | −4679 | −5035 | −5100 | −5491 |
| OV-25 | −3430 | −4397 | −4708 | −5115 | −5044 | −5513 |
| OV-105 | −3159 | −4195 | −4281 | −4635 | −4689 | −5079 |
| OV-225 | −3546 | −4304 | −5145 | −5319 | −5020 | −5768 |
| OV-275 | −3458 | −3566 | −5477 | −5261 | −4506 | −5639 |
| OV-330 | −3686 | −4555 | −5530 | −5625 | −5235 | −6052 |
| QF-1 | −3071 | −3646 | −3977 | −4264 | −4276 | −4665 |
| CW20M | −3858 | −4551 | −5984 | −5871 | −5250 | −6304 |
| DEGS | −3902 | −4641 | −5637 | −5454 | −4822 | −5938 |
| TCEP | −3951 | −4310 | −5971 | −5768 | −5045 | −6244 |
| PPE-5 | −3676 | −4847 | −5168 | −5537 | −5468 | −6021 |
| QpTS | −4038 | −5361 | −6663 | −6558 | −5166 | −6572 |
| QPIC | −3993 | −4523 | −5998 | −6033 | −5436 | −5447 |
| QMES | −4011 | −4270 | −6724 | −6599 | −5148 | −6708 |
| QACES | −3862 | −3896 | −6473 | −6230 | −4794 | −6390 |
| QTAPSO | −3749 | −3810 | −6012 | −5773 | −4604 | −6019 |
| DDP | −3823 | −4329 | −5275 | −5406 | −4979 | −5748 |
| SE-30 | −3137 | −4205 | −4117 | −4502 | −4647 | −4945 |

TABLE II *(continued)*

| | Anisole | Nonanal | 2-Octanone | Dodecane |
|---|---|---|---|---|
| SQ | −4364 | −5178 | −4608 | −6025 |
| OV-3 | −3978 | −4772 | −4250 | −5122 |
| OV-7 | −4062 | −4837 | −4305 | −5060 |
| OV-11 | −4126 | −4836 | −4300 | −4925 |
| OV-17 | −4150 | −4805 | −4257 | −4800 |
| OV-22 | −4158 | −4686 | −4193 | −5016 |
| OV-25 | −4156 | −4610 | −3971 | −4438 |
| OV-105 | −3900 | −4736 | −4465 | −5041 |
| OV-225 | −4149 | −4763 | −3867 | −4006 |
| OV-275 | −3714 | −3512 | −3369 | −1216 |
| OV-330 | −4396 | −4829 | −4359 | −4354 |
| QF-1 | −3490 | −4521 | −4197 | −3913 |
| CW20M | −4484 | −4648 | −4217 | −3816 |
| DEGS | −4103 | −4050 | −3750 | −2168 |
| TCEP | −4284 | −4210 | −4035 | −2188 |
| PPE-5 | −4514 | −5089 | −4555 | −4755 |
| QpTS | −4455 | −4838 | −4371 | −3720 |
| QPIC | −4482 | −4998 | −4640 | −3746 |
| QMES | −4434 | −4792 | −4375 | −3695 |
| QACES | −4125 | −3630 | −3852 | −2409 |
| QTAPSO | −3900 | −3855 | −3588 | −2003 |
| DDP | −4169 | −4554 | −4876 | −5565 |
| SE-30 | −3879 | −4692 | −4184 | −5187 |

largest value found for that test solute on all stationary phases). Missing values were caused by poor chromatographic properties such as adsorption at infinite dilution, irregular peak shape, excessive retention, etc., which precluded accurate determination of the partition coefficient. For similar reasons phenol and 2,4,6-trimethylphenol data from ref. 32 were not included in the data set as these test solutes were not eluted from several phases. Tetraethylammonium 4-toluenesulphonate and tri-*n*-butylammonium 4-toluenesulphonate were not included because they failed to elute, or eluted with poor peak shapes, most of the N-heterocyclic bases [32].

## RESULTS AND DISCUSSION

The principal intermolecular interactions that take place in solution are dispersion, induction, orientation, and various complexation interactions, of which proton donor–acceptor interactions are the most common. Solutes and solvents show a balance of these interactions; the capacity of a solute or solvent to enter into all intermolecular interactions is characterized by its strength, or polarity, and the ratio of one particular interaction to the strength, is its selectivity. Unfortunately, no solute exhibits a single interaction for all phases, since even *n*-alkanes in a polar solvent will experience inductive as well as dispersive interactions. Those solutes that are most useful for characterizing the complementary interactions in a solvent should have a single, dominant type of interaction expressed against a weak background for other interactions. Intuitively, it would seem likely that one test solute to characterize each of

the principal intermolecular interactions should be sufficient. However, neither the number nor the identity of the test solutes required to characterize the selectivity of GLC phases can be predicted in a straightforward manner. An approach which involves the minimum number of *a priori* assumptions is multivariate analysis. In this case an undefined number of factors can be extracted from the combination of a large number of solute–solvent interactions, which hopefully can be described in terms of primary forces. Those solutes which show the strongest correlation with individual factors can be identified as suitable test solutes for defining the relative importance of a particular retention mechanism. The inverse process is also possible, and the similarity of all retention mechanisms can be used to classify stationary phases into groups with similar separation characteristics.

Since it is impractical to commence any study with the full universe of potential candidate test solutes, we used the published opinions of experts in the field of gas chromatography to preselect the 28 solutes identified in Table I [2,4–14,29–33,42]. The strength of intermolecular interactions for these solutes on 23 stationary phases was determined by the partial molal Gibbs free energy of solution [5,32,40,41]. The range of free energy values for different test solutes is quite large, so to avoid inadvertent weighting of the results by extreme values, the data was scaled, using the autoscale routine in the Ein*Sight programming environment [25–27]. This results in each variable being mean centered with a standard deviation of one. A correlation matrix of the scaled data was then produced to evaluate the relationship between individual variables. A correlation of 0.8 to 1.0 was considered to be a reasonable indication that the variables are correlated, that is, displaying the same retention mechanisms. The variables meeting this test are summarized in Table III for the three cases were it seems reasonable to assign a particular interaction as a dominant interaction. Also included are those test solutes which behave independently. Strong correlations were found for the partial molar Gibbs free energy of solution for a methylene group, $\Delta G_k^0$ ($CH_2$) and dihexyl ether, dodecane, 2-octyne, 1-dodecyne, *n*-butylbenzene, *cis*-hydrindane and, to a lesser extent, nonanal. It seems reasonable to class these test solutes as dispersive probes with essentially duplicate interactions. Leaving nonanal aside, there is no additional selectivity observed for the unsaturated hydrocarbons compared to those that are fully saturated. This can be seen from the plot of 1-dodecyne against dodecane ($r = 0.97$) in Fig. 1. These test solutes have nearly identical boiling points (difference *ca.* 1.2°C) and differ insignificantly in molecular size. The two phases that are most discordant in the plot OV-275 (10) and QF-1 (12) can be explained without invoking differences in selectivity. In the case of OV-275 the test solutes have extremely small partition coefficients and are not as accurately determined as are the other values used in calculating the data for Fig. 1. QF-1 behaves anomalously for all test solutes. It is the only fluorocarbon phase in the test set and behaves independently for all probes. Fluorocarbon phases have significantly weaker dispersive interactions than hydrocarbon phases resulting in their different behavior [33,38].

The second group of highly correlated test solutes, with one exception, have fairly large dipole moments and it seems reasonable to associate these test solutes with strong orientation interactions. Dioxane would seem to be anomalous within this group and shows the lowest correlation coefficient ($r = 0.81$). Pyridine is another unexpected member of this group based on its use as a probe for stationary phase proton donor capacity. However, it has a large dipole moment (2.25 D) and shows no

TABLE III

SUMMARY OF CORRELATED VALUES ABSTRACTED FROM THE CORRELATION MATRIX

| | Correlation coefficient | Dipole moment (D) |
|---|---|---|
| *Dispersive interactions* | | |
| $\Delta G_k^0$ ($CH_2$) | 1.00 | |
| Dihexyl ether | 0.98 | |
| Dodecane | 0.98 | |
| 2-Octyne | 0.97 | |
| 1-Dodecyne | 0.96 | |
| *n*-Butylbenzene | 0.95 | |
| *cis*-Hydrindane | 0.93 | |
| Nonanal | 0.83 | |
| *Orientation interactions* | | |
| Nitrobenzene | 1.00 | 3.97 |
| Benzonitrile | 0.99 | 4.08 |
| Pyridine | 0.96 | 2.25 |
| N-Methylaniline | 0.96 | 1.68 |
| Nitropropane | 0.95 | 3.06 |
| Aniline | 0.94 | 1.53 |
| Nitropentane | 0.92 | 3.52 |
| Benzodioxane | 0.91 | 1.43 |
| 2,6-Dimethylaniline | 0.88 | 1.63 |
| Butanol | 0.84 | 1.78 |
| Dioxane | 0.81 | 0.40 |
| *Proton donor–acceptor interactions* | | |
| Octanol | 1.00 | 1.72 |
| 2-Methyl-2-pentanol | 0.98 | |
| Butanol | 0.89 | 1.78 |
| N-Methylaniline | 0.83 | 1.68 |
| *Solutes behaving independently* | | |
| Benzene | | 0.03–0.1 |
| 1,1,2,2-Tetrachloroethane | | 1.67 |
| 2,4,6-Trimethylpyridine | | 2.26 |
| Dodecafluoroheptanol | | |
| N,N-Dimethylaniline | | 1.59 |
| 2-Octanone | | 2.46 |
| Anisole | | 1.25 |

obvious selectivity for proton donor phases over that of nitrobenzene (Fig. 2). Benzodioxan has also been used as a test solute for stationary phase proton donor capacity but again we find that its behavior is more characteristic of that of the orientation probes [5].

The third class of correlated test solutes is comprised of three alcohols and N-methylaniline, the latter is only weakly correlated with 1-octanol ($r = 0.83$). Both N-methylaniline and butanol are also correlated to nitrobenzene, the former strongly ($r = 0.96$) and the latter weakly ($r = 0.84$). We suspect that N-methylaniline is better considered representative of the orientation probes and butanol that of the proton donor probes. In fact, if the data for 1-butanol and 1-octanol are scrutinized (Fig. 3), there are two general trends that can be discerned. Those phases which are strongly

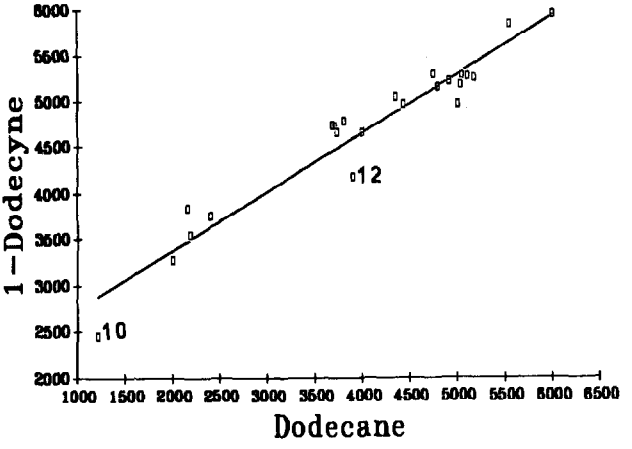



Fig. 1. Plot of the partial molal Gibbs free energy of solution for 1-dodecyne against dodecane for the 23 stationary phases identified in Table I. Phase 10 is OV-275 and 12 is QF-1.

associated (OV-275, DEGS, TCEP, QTAPSO and QACES), form one group with a correlation coefficient of 0.98 ($n$ = 5) displaced from the remaining phases, which have a correlation coefficient of 0.97 ($n$ = 17), QF-1 behaving independently. It seems likely that the reason for the modest correlation between butanol and 1-octanol is due to differences in the cavity term, which is important for solvent with a high degree of cohesion (see later).

The remaining seven solutes in Table III show no strong correlation with other solutes and must be considered as behaving independently. There are several reasons for this. In the case of benzene, weak retention on several phases results in a poor determination of its free energy value causing scatter in the plots against other dispersive probes with which it shows the greatest similarity. 1,1,2,2-Tetrachloro-

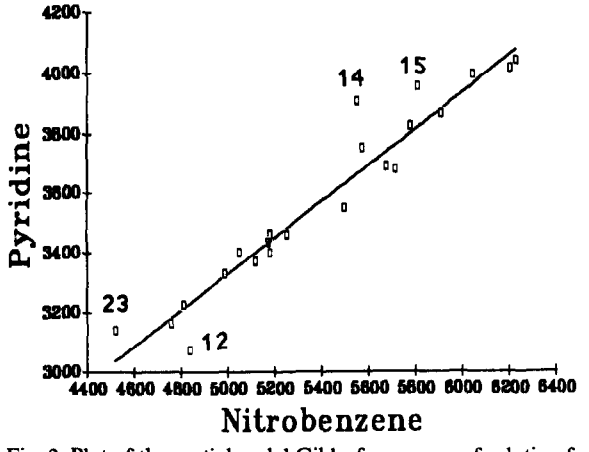



Fig. 2. Plot of the partial molal Gibbs free energy of solution for pyridine against nitrobenzene for the 23 stationary phases identified in Table I. Phase 12 is QF-1, 14 is DEGS, 15 is TCEP and 23 is SE-30.

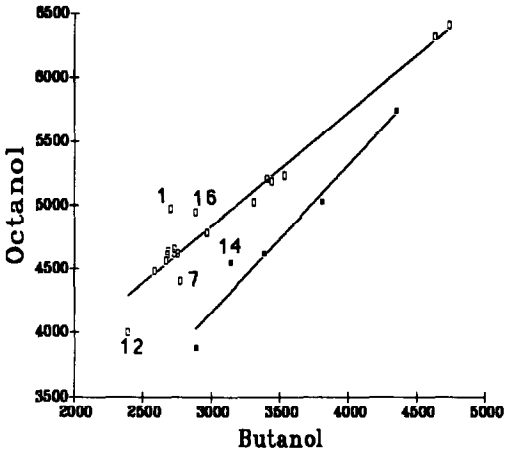



Fig. 3. Plot of the partial molal Gibbs free energy of solution for 1-octanol against 1-butanol for the 23 stationary phases identified in Table I. The strongly associated phases 10, 14, 15, 20 and 21 are displaced and separately correlated from the remaining phases. Phases identified on the figure are: 1 = SQ; 7 = OV-25; 12 = QF-1; 14 = DEGS; 16 = PPE-5.

ethane shows both orientation and proton donor capability, behaving in a mixed mode, and is not useful as a test solute for specific interactions. Dodecafluoroheptanol behaves independently because of the weak dispersive interactions of the fluorocarbon chain compared to hydrocarbon analogs. It is weakly correlated to the other alcohols ($r \approx 0.80$) if only the experimental values are considered (dodecafluoroheptanol was not eluted from all phases and approximate values were added to complete the data matrix, see Experimental). It should be a stronger acid than the alkanols but given the fact that it cannot be eluted successfully from all phases it is not a suitable test solute. 2,4,6-Trimethylpyridine is the most unique test solute and shows no correlation with any of the other solutes, including pyridine. Pyridine and 2,4,6-trimethylpyridine have similar dipole moments, and it was shown earlier that pyridine correlates very strongly with the orientation probes (Fig. 2), making the behavior of 2,4,6-trimethylpyridine the more remarkable. Further studies are required to explain the behavior of 2,4,6-trimethylpyridine. 2-Octanone correlates most strongly with the dispersive probes, with most deviation for the polar phases. 2-Octanone has a reasonable dipole moment (2.46 D), and probably shows mixed behavior, primarily dispersion with some weak orientation capacity. N,N-Dimethylaniline and anisole are most strongly correlated to each other, $r = 0.96$, and then to dioxane and benzodioxane ($r \approx 0.8$). It is likely that these test solutes are retained by a mixed retention mechanism including contributions from orientation and proton acceptor interactions.

To better define the behavior of the test solutes the autoscaled data was subjected to principal component analysis. Eigenvectors were extracted from the data such that the maximum information in the form of variance was preserved with a minimum number of eigenvectors. A summary of the results is presented in Table IV. The first three eigenvectors account for 91.45% of the total variance in the data with additional eigenvectors contributing little further useful information. The results from principal component analysis are usually presented graphically in the form of plots of the

TABLE IV

SUMMARY OF EIGENVECTOR AND PRINCIPAL COMPONENT ANALYSIS OF THE DATA MATRIX

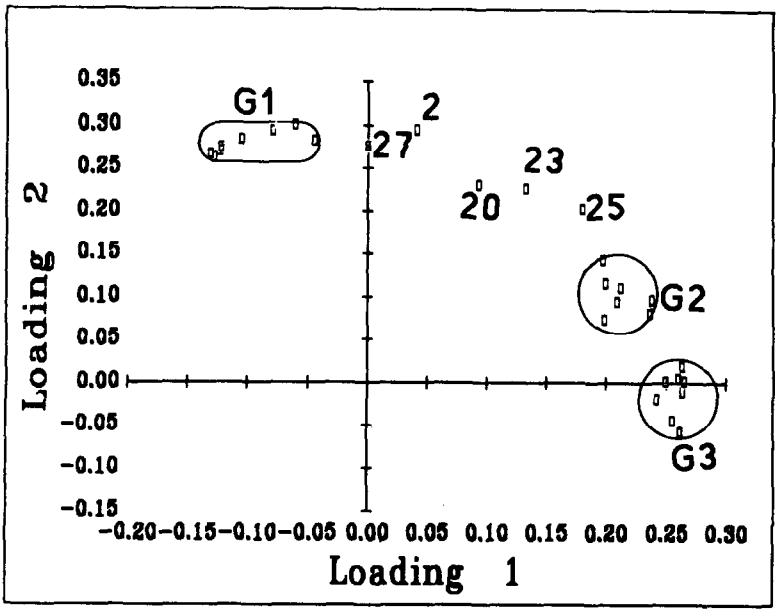
| Eigenvector | Eigenvalue | Percent variance | Percent cumulative variance |
|---|---|---|---|
| *All 23 phases* | | | |
| 1 | 13.7783 | 49.20 | 49.20 |
| 2 | 10.3634 | 37.01 | 86.22 |
| 3 | 1.4652 | 5.23 | 91.45 |
| 4 | 0.7272 | 2.59 | 94.05 |
| 5 | 0.7088 | 2.53 | 96.58 |
| *18 Phases (excluding those phases that are highly cohesive)* | | | |
| 1 | 16.1629 | 57.72 | 57.72 |
| 2 | 7.6740 | 27.40 | 85.13 |
| 3 | 1.5057 | 5.37 | 90.50 |
| 4 | 1.2029 | 4.29 | 94.80 |
| 5 | 0.7788 | 2.78 | 97.58 |

TABLE V

SUMMARY OF THE LOADINGS FOR THE FIRST THREE PRINCIPAL COMPONENTS ($\times 10$)

| Variable | Loading 1 | Variable | Loading 2 | Variable | Loading 3 |
|---|---|---|---|---|---|
| N-Methylaniline | 0.2650 | Butylbenzene | 0.2989 | 2-Methyl-2-pentanol | 0.4202 |
| Benzonitrile | 0.2641 | Benzene | 0.2924 | Dioxane | −0.3891 |
| Nitrobenzene | 0.2634 | 1-Dodecyne | 0.2913 | Octanol | 0.3655 |
| Aniline | 0.2623 | 2-Octanone | 0.2810 | N,N-Dimethylaniline | −0.3397 |
| Pyridine | 0.2600 | Nonanal | 0.2799 | Dodecafluoroheptanol | 0.3188 |
| Nitropropane | 0.2554 | 2-Octyne | 0.2743 | Butanol | 0.3155 |
| 2,6-Dimethylaniline | 0.2493 | Dihexyl ether | 0.2724 | Anisole | −0.2396 |
| Butanol | 0.2425 | $\Delta G_k^0$ ($CH_2$) | 0.2684 | Benzodioxane | −0.2350 |
| Benzodioxane | 0.2385 | Dodecane | 0.2648 | 2,4,6-Trimethylpyridine | −0.1569 |
| Nitropentane | 0.2372 | *cis*-Hydrindane | 0.2611 | Pyridine | −0.1329 |
| Octanol | 0.2121 | 2,4,6-Trimethylpyridine | 0.2287 | 2-Octanone | −0.1181 |
| Dioxane | 0.2093 | N,N-Dimethylaniline | 0.2246 | Dodecane | 0.1171 |
| 2-Methyl-2-pentanol | 0.1996 | Anisole | 0.2007 | 1-Dodecyne | 0.0950 |
| Dodecafluoroheptanol | 0.1987 | 1,1,2,2-Tetrachloroethane | 0.1416 | Dihexyl ether | 0.0718 |
| 1,1,2,2-Tetrachloroethane | 0.1974 | 2,-Methyl-2-pentanol | 0.1142 | Nitropentane | 0.0710 |
| Anisole | 0.1806 | Octanol | 0.1087 | Benzene | −0.0668 |
| N,N-Dimethylaniline | 0.1330 | Benzodioxane | 0.0942 | Nitrobenzene | −0.0633 |
| Dodecane | −0.1314 | Dioxane | 0.0922 | 1,1,2,2-Tetrachloroethane | 0.0420 |
| *cis*-Hydrindane | −0.1288 | Nitropentane | 0.0792 | *cis*-Hydrindane | 0.0394 |
| $\Delta G_k^0$ ($CH_2$) | −0.1231 | Dodecafluoroheptanol | 0.0722 | $\Delta G_k^0$ ($CH_2$) | 0.0394 |
| Dihexyl ether | −0.1229 | Aniline | −0.0586 | Nitropropane | 0.0357 |
| 2-Octyne | −0.1053 | Nitropropane | −0.0459 | Benzonitrile | −0.0340 |
| 2,4,6-Trimethylpyridine | 0.0934 | Butanol | −0.0204 | 2,6-Dimethylaniline | −0.0335 |
| 1-Dodecyne | −0.0793 | Nitrobenzene | −0.0172 | N-Methylaniline | −0.0318 |
| Butylbenzene | −0.0605 | Benzonitrile | −0.0128 | Aniline | 0.0290 |
| Nonanal | −0.0441 | Pyridine | −0.0023 | Butylbenzene | 0.0256 |
| Benzene | 0.0414 | 2,6-Dimethylaniline | −0.0006 | Nonanal | 0.0221 |
| 2-Octanone | 0.0002 | N-Methylaniline | −0.0003 | 2-Octyne | −0.0136 |

loadings and scores. The scores for each sample (phase) result from the projection of the sample data vector onto the principal components (eigenvectors that contain information against eigenvectors which represent noise in the data matrix). The principal component loadings are the elements of the principal components used to



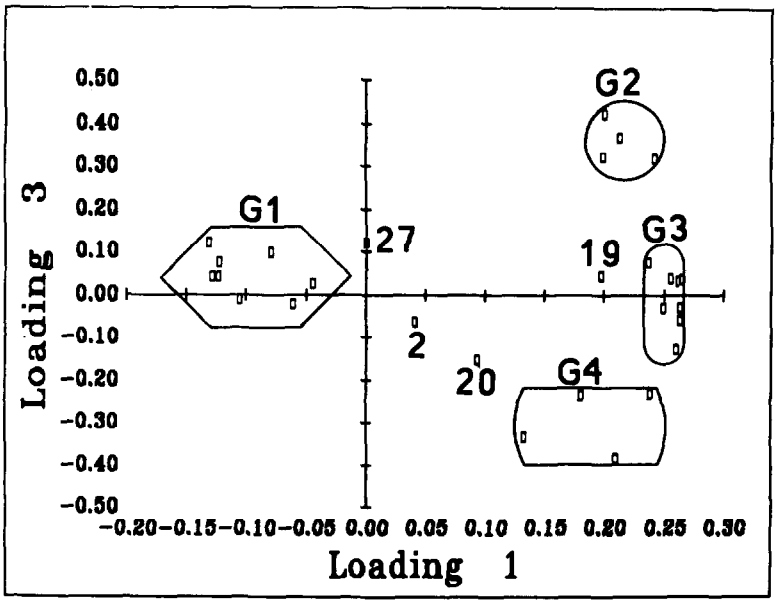




Fig. 4. Principal component plots of loading 1 against loading 2 and loading 3. The compositions of the groups (G) are identified in Tables VI and VII.

produce the scores, equivalent to the coefficients of the linear equation defining the principal components. The loadings describe how much each variable (test solute) contributes to the principal component (retention mechanism). The loadings for the first three principal components are summarized in Table V. The most highly loaded test solutes for loading 1 were previously recognized as orientation interaction probes, those of loading 2 as dispersive interaction probes, and those of loading 3 as proton donor–acceptor probes. The proton donor test solutes are characterized by a positive sign and the proton acceptor probes by a negative sign in the top half of the test solutes showing significant contributions to loading 3.

The plot of loading 1 (orientation) against loading 2 (dispersion) accounts for 86.22% of the cumulative variance (Fig. 4). The test solutes are classified into three reasonably distinct groups (Table VI), with test solutes benzene (2), 2,4,6-trimethylpyridine (20), N,N-dimethylaniline (23), and anisole (25) and 2-octanone (27) behaving independently. The classification of the test solutes into three groups seems intuitively reasonable but some obvious inconsistencies exist. For example, nitropropane and nitropentane are separated into groups 2 and 3, and so are butanol and octanol. Evaluation of the next heavily loaded principal components seem warranted. The plot of loading 1 (orientation) against loading 3 (proton donor–acceptor) is shown in Fig. 4. In this case four distinct groups are obtained (Table VII), with test solutes benzene (2), 1,1,2,2-tetrachloroethane (19), 2,4,6-trimethylpyridine (20) and 2-octanone (27) behaving independently. A more logical classification is obtained than in the previous case with group 1 test solutes containing all the dispersive probes, group 2 the proton donor test solutes, group 3 the orientation test solutes, and group 4 the proton acceptor test solutes. The group 4 test solutes are only diffusely clustered, which we believe is due to the fact that none of these test solutes exhibits a dominant proton acceptor capacity and additional test solutes, not in the current data matrix, will be required to adequately characterize this interaction. This will be the subject of an additional study, still in progress [41].

The plot of loading 1 against loading 4 accounts for almost as much of the cumulative variance (51.79%) as loading 1 against loading 3. However, the three highest weightings in loading 4 are 2,4,6-trimethylpyridine (0.7714), dodecafluoroheptanol (0.3310), and 1,1,2,2-tetrachloroethane (0.3208) which are test solutes showing the most independent behavior. The plot of loading 1 against loading 4 does not provide a good classification of the test solutes. All the dispersive probes are placed

TABLE VI

COMPOSITION OF GROUPS FROM THE CLASSIFICATION OF LOADING 1 *VS.* LOADING 2 (86.22%)

| *Group 1* | *Group 2* | *Group 3* |
|---|---|---|
| 1 $\Delta G_k^0$ ($CH_2$) | 6 2-Methyl-2-pentanol | 3 Butanol |
| 7 2-Octyne | 8 Dioxane | 4 Nitropropane |
| 9 *cis*-Hydridane | 11 Nitropentane | 5 Pyridine |
| 10 Butylbenzene | 13 Octanol | 12 Nitrobenzene |
| 15 Dihexyl ether | 14 Benzodioxane | 18 Benzonitrile |
| 16 1-Dodecyne | 19 1,1,2,2-Tetrachloroethane | 21 Aniline |
| 17 Dodecane | 28 Dodecafluoroheptanol | 22 N-Methylaniline |
| 26 Nonanal | | 24 2,6-Dimethylaniline |

TABLE VII

COMPOSITION OF GROUPS FROM THE CLASSIFICATION OF LOADING 1 *VS.* LOADING 3 (54.43%)

| *Group 1* | *Group 2* | *Group 3* | *Group 4* |
|---|---|---|---|
| 1 $\Delta G_k^0$ ($CH_2$) | 3 Butanol | 4 Nitropropane | 8 Dioxane |
| 7 2-Octyne | 6 2-Methyl-2-pentanol | 5 Pyridine | 14 Benzodioxane |
| 9 *cis*-Hydrindane | 13 Octanol | 11 Nitropentane | 23 N,N-Dimethylaniline |
| 10 Butylbenzene | 28 Dodecafluoroheptanol | 12 Nitrobenzene | 25 Anisole |
| 15 Dihexyl ether | | 18 Benzonitrile | |
| 16 1-Dodecyne | | 21 Aniline | |
| 17 Dodecane | | 22 N-Methylaniline | |
| 26 Nonanal | | 24 2,6-Dimethylaniline | |

into a tight cluster but the other test solutes are diffusely scattered with no distinction between the orientation and proton donor–acceptor probes.

In the same manner that the principal component plots of the loadings can be used to classify the test solutes, the scores plots can be used to classify the stationary phases based on their interactions with the test solutes. The scores for the four most significant principal components are summarized in Table VIII. The plots of score 1

TABLE VIII

PRINCIPAL COMPONENT SCORES FOR THE STATIONARY PHASES

| Stationary phase | Scores | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| SQ | 2.7244 | −5.0862 | 0.2340 | −0.8955 | 0.4778 |
| OV-3 | 4.2911 | −0.9711 | −0.7349 | −0.1927 | 0.3052 |
| OV-7 | 3.3734 | −1.4512 | −0.2678 | −0.1016 | 0.2069 |
| OV-11 | 2.7235 | −1.5169 | 0.2282 | −0.0721 | 0.1537 |
| PV-17 | 2.4245 | −1.2657 | 0.5482 | −0.0810 | 0.2145 |
| OV-22 | 2.3366 | −0.9604 | 0.9207 | −0.2380 | 0.2417 |
| OV-25 | 2.3235 | −0.3772 | 1.1197 | −0.6163 | 0.5262 |
| OV-105 | 4.5357 | −0.6159 | −1.3660 | 0.2446 | 0.0533 |
| OV-225 | −0.1434 | 0.2475 | 0.3363 | 0.5857 | 0.2593 |
| OV-275 | 0.7268 | 8.3215 | 0.8761 | 0.8173 | 0.4481 |
| OV-330 | −1.8062 | −1.9728 | 0.2695 | −0.0987 | 0.0225 |
| QF-1 | 5.5526 | 4.0651 | −1.7563 | 1.9843 | −0.9096 |
| CW20M | −3.6335 | −1.3007 | 0.8783 | 0.0059 | 0.2769 |
| DEGS | −1.6522 | 3.4658 | 1.3332 | 0.2308 | 1.1153 |
| TCEP | −3.8227 | 2.7151 | 1.9583 | 0.7103 | −0.3440 |
| PPE-5 | −0.9774 | −3.7660 | 1.9306 | −0.2435 | −0.0886 |
| QpTS | −7.3026 | −2.1742 | −1.8558 | 0.5764 | 1.5666 |
| QPIC | −3.7661 | −1.7816 | 1.0411 | 1.2617 | −1.0996 |
| QMES | −7.0903 | −1.4666 | −2.5937 | 0.2998 | 0.2683 |
| QACES | −4.2261 | 3.8981 | −1.2826 | −1.5986 | −0.5227 |
| QTAPSO | −1.0408 | 6.2594 | −0.3463 | −2.2583 | −0.8852 |
| DDP | −1.10664 | −3.8031 | −0.2546 | 0.0281 | −2.8139 |
| SE-30 | 5.5558 | −0.4624 | −1.2160 | −0.3481 | 0.5274 |

against score 2, accounting for 86.22% of the total variance, is shown in Fig. 5. Five distinct groups are obtained (Table IX), with squalane (1), OV-225 (9), OV-275 (10), QF-1 (12) and QTAPSO (21) behaving independently. OV-225 is located almost exactly at the cross hairs indicating a balance of interactions while squalane and

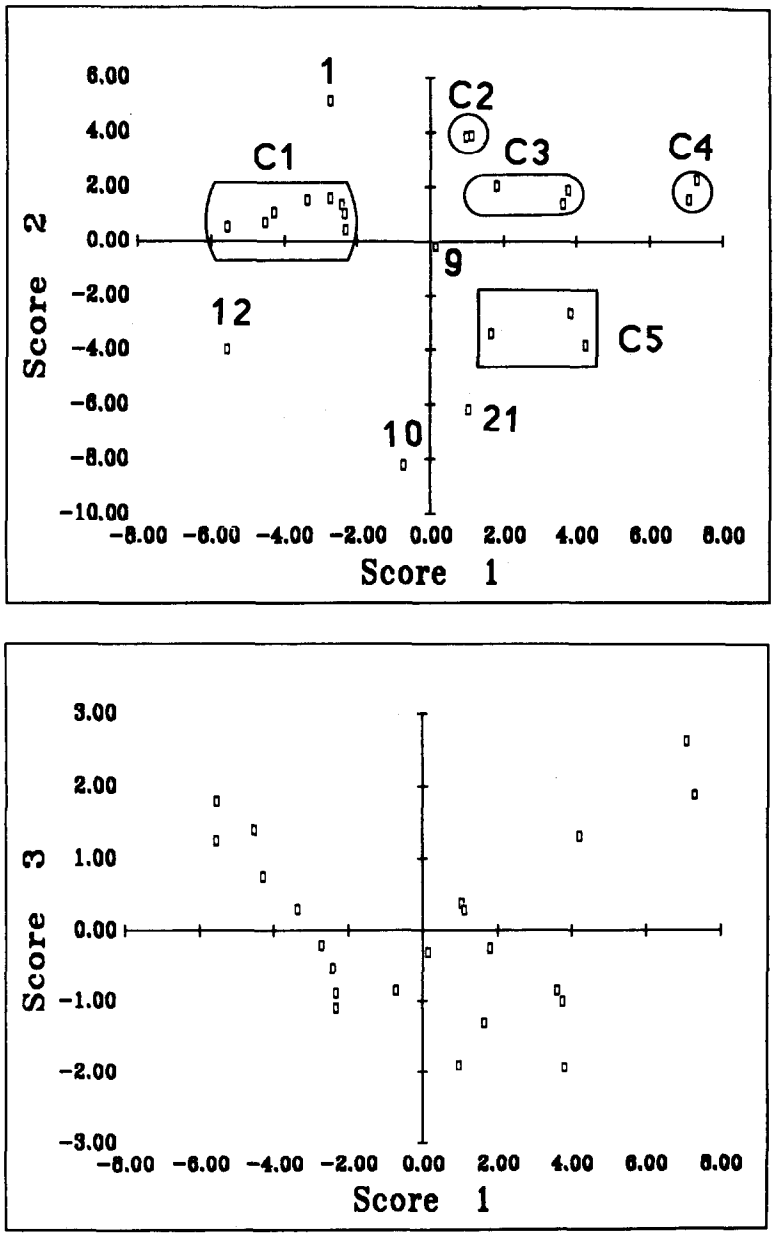



Fig. 5. Principal component plots of score 1 against score 2 and score 3. The composition of the classes (C) is given in Table IX.

TABLE IX

COMPOSITION OF THE GROUPS CLASSIFIED ACCORDING TO SCORE 1 *VS.* SCORE 2 (86.22%)

| *Class 1* | *Class 2* | *Class 3* | *Class 4* | *Class 5* |
|---|---|---|---|---|
| 2 OV-3 | 16 PPE-5 | 11 OV-330 | 17 QpTS | 14 DEGS |
| 3 OV-7 | 22 DDP | 13 CW20M | 19 QMES | 15 TCEP |
| 4 OV-11 | | 18 QPIC | | 20 QACES |
| 5 OV-17 | | | | |
| 6 OV-22 | | | | |
| 7 OV-25 | | | | |
| 8 OV-105 | | | | |
| 23 SE-30 | | | | |

OV-275 are situated towards opposite edges of the figure indicating high selectivity. Few of the groups are compact indicating a transition of properties within the group. The class 1 phases show variation in both the relative contribution of orientation and dispersion interactions with a logical distribution based on increasing polarity as the mole percent of phenyl groups increases for the poly(methylphenylsiloxane) phases. The class 3 and class 4 phases are separated largely by increasing orientation interactions. The division of phases into the different classes seems to be intuitively correct. The plot of score 1 against score 3, representing 54.43% of the total variance, Fig. 5 does not produce significant clustering of the phases. The phases, as a group, vary substantially in their orientation and proton donor–acceptor capacity with little duplication among phases. The desire to have a wide selection of phases expressing different retention mechanisms to adequately test solute properties seems to have been met in the selection procedure. Score 1 against score 4 (51.79% of the total variance) and 1 against 5 (51.73% of the total variance) are somewhat similar to the results of score 1 against score 3 in that they do not lead to a general grouping of the phases into classes.

An alternative to principal component analysis for classifying samples by multivariate analysis is cluster analysis. For cluster analysis a distance matrix is formed from the original scaled (or unscaled) data matrix. The Euclidean distance between any sample (or groups of samples) to another is used as a measure of how similar the two samples are. The output for clustering algorithms are dendrograms. Ein*Sight supports seven different cluster algorithms (group average, centroid, incremental sum of squares, median, Lance and Williams flexible, single-linkage nearest neighbor, and complete-linkage farthest neighbor), all producing slightly different dendrograms for the 23 phases used in this study. The groupings are the same using farthest neighbor, Lance and Williams flexible, incremental sum of squares, centroid, and group average with only the similarity values varying. The nearest neighbor and median produces some significant differences in the position of phases within groups and intuitively are less satisfactory. Fig. 6 illustrates a dendrogram for the complete-linkage farthest neighbor method, which is representative of the other methods excepting the nearest neighbor and median methods. The phases that are most similar are next to each other and are connected. Connections at the extreme left side of the dendrogram have a similarity of 1, representing duplicates, and those at the extreme right a similarity of

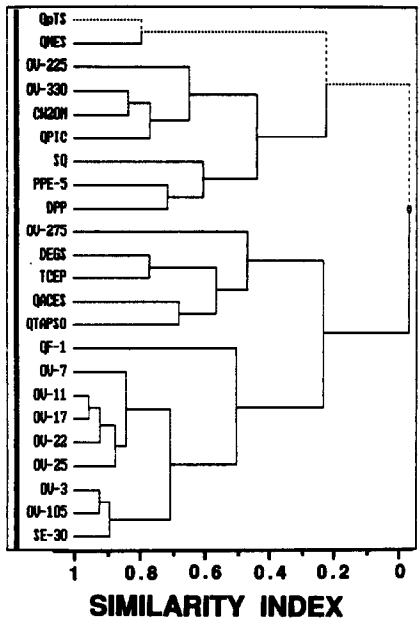



Fig. 6. Similarity of stationary phases using the complete-linkage farthest neighbor dendrogram. Similarity index for connected phases OV-11/OV-17, 0.95; OV-22/OV-25, 0.92; OV-3/OV-7, 0.91; OV-105/SE-30, 0.88; Carbowax 20M/QPIC, 0.80; QpTS/QMES, 0.78; DEGS/TCEP, 0.77; OV-225/OV-330, 0.76; PPE-5/DDP, 0.71; QACES/QTAPSO, 0.68.

zero, and have no features in common. The dendrogram classifies the phases into three basic groups consisting of QpTS, QMES, OV-225, OV-330, CW-20M, QPIC, squalane, PPE-5, and DDP; OV-275, DEGS, TCEP, QACES, QTAPSO; and QF-1, OV-11, OV-17, OV-22, OV-25, OV-3, OV-7, OV-105 and SE-30. Squalane, OV-275 and QF-1 are not connected to any other phase and are behaving independently. The major groups are then divided into subgroups of increasing similarity. A comparison of these subgroups with the classification by principal component analysis using score 1 against score 2 (Table IX) shows very good agreement. The dendrogram is a very useful device for visualizing the relative similarities between phases and for predicting phases likely to show different separation characteristics.

Originally, it was believed that test solutes could be selected by identifying solute types with the desired balance of interactions followed by selecting a homologue of the correct volatility to allow accurate determination of the gas–liquid partition coefficient at the standard measurement temperature. This simple philosophy has one failing. Considering the retention behavior of homologous test solutes in Table I (butanol and octanol, Fig. 3, benzene and *n*-butylbenzene, Fig. 7, and nitropropane and nitropentane, Fig. 8), there are two obvious types of behavior. A small group of phases (OV-275, DEGS, TCEP, QTAPSO and QACES) form a separate correlated group, displaced from the other phases. These phases are highly functionalized and are probably more cohesive than the other phases. They exhibit a marked tendency to repel alkyl groups, such that as the size of the alkyl group increases the gas–liquid partition coefficient increases very little, compared with the other phases. This

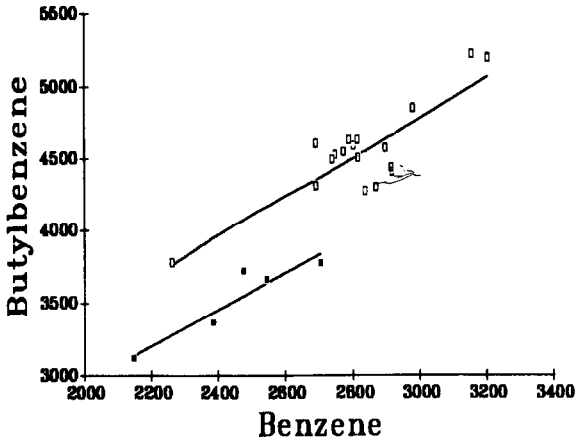



Fig. 7. Plot of the partial molal Gibbs free energy of solution for butylbenzene against benzene for the 23 stationary phases identified in Table I. The strongly associated phases 10, 14, 15, 20 and 21 (■) are displaced and separately correlated ($r$ = 0.95) from the remaining phases (□) ($r$ = 0.85).

difference in behavior is probably accounted for by the difference in the cavity term; the free energy required to separate the solvent molecules to provide a cavity of sufficient size to accommodate the solute molecule. The partial molal Gibbs free energy of solution (soln) for any solute X can be separated into two terms, the cavity term, and the interaction term (int), as shown in eqn. 1

$$(\Delta G_m^0 X)^{soln} = (\Delta G_m^0 X)^{cavity} + (\Delta G_m^0 X)^{int} \quad (1)$$

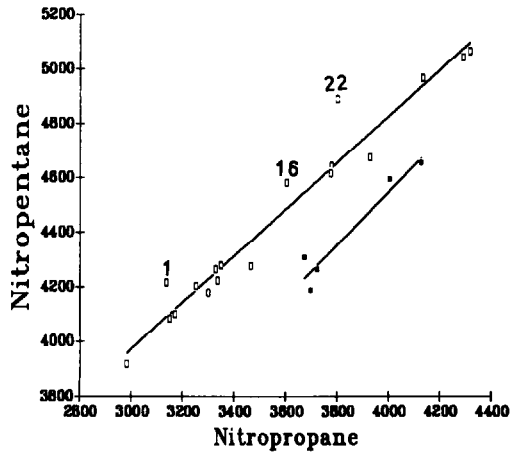



Fig. 8. Plot of the partial molal Gibbs free energy of solution for nitropentane against nitropropane for the 23 stationary phases identified in Table I. The strongly associated phases 10, 14, 15, 20 and 21 (■) are displaced and separately correlated ($r$ = 0.96) from the remaining phases (□) ($r$ = 0.98). Phases identified on the figure: 1 = SQ; 16 = PPE-5; 22 = DDP.

The cavity term is difficult to evaluate by itself, but the cavity term combined with the non-polar interaction term can be expressed by the free energy required to dissolve an *n*-alkane of the same size as the test solute. The interaction term now represents only polar interactions and the additional component of the free energy inadequately accounted for by using the *n*-alkane as a molecular model for the test solute. For example, pentane can be used for butanol and nonane for octanol. If $(\Delta G_m^0 X)^{int}$ for octanol is plotted against that for butanol (Fig. 9), all phases are now found to fall on the same line ($r = 0.99$) with a slope of 1.03 and a near zero intercept ($-6.50$). These results indicate that the polar interactions for butanol and octanol are essentially identical and independent of the molecular weight of the test solute.

In terms of estimating the capacity of a stationary phase for a particular interaction $(\Delta G_m^0 X)^{int}$ would seem to be attractive because of its clear definition. However, although physically elegant, it ignores the practice of chromatography in which solutes of different size will be routinely present in most separations. For those phases that are highly cohesive a lack of solubility of the alkyl group has just as much impact on their chromatographic properties, compared to other phases, as the strength of the polar interactions themselves. To predict the capacity of a solvent for specific intermolecular interactions a scale based on $(\Delta G_m^0 X)^{int}$ seems to be a logical next step, but to predict retention or relative retention a more complex expression that includes the influence of molecular size would be needed.

The highly cohesive phases (OV-275, DEGS, TCEP, QTAPSO and QACES) could be considered as a special case and removed from the data matrix. There is a danger in doing this in that it removes a group of phases that are generally considered among the most selective by chromatographers. Those that remain are fairly representative of orientation interactions, and the liquid organic salts are known to possess strong proton acceptor capacity, but there are probably no strong proton donor phases remaining in the matrix. The data matrix consisting of 28 solutes on 18 phases, excluding those that are highly cohesive, was reanalyzed by multivariate analysis to establish whether the conclusions reached previously are equally applicable

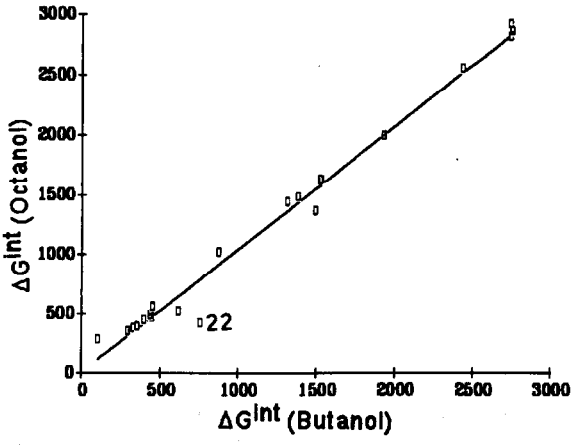



Fig. 9. Plot of the partial molal Gibbs free energy of interaction for octanol against butanol for the 23 stationary phases identified in Table I. Phase 22 is DDP.

in the absence of the highly associated phases. We will discuss this data in summary form only. There are some significant changes in the correlation matrix (Table X), compared to Table III. The dispersive probes remain well correlated except for *cis*-hydrindane, which is weakly correlated to the other test solutes in this group ($r \approx 0.8$), but otherwise could be considered to behave independently. The orientation probes remain highly correlated and likewise the proton donor solutes. However, there is now far more mixing of properties with both groups containing mainly the same test solutes with a different order of the correlation coefficients. Dodecafluoroheptanol is now well correlated to the other alcohols and 1,1,2,2-tetrachloroethane is well correlated to both the orientation and proton donor solutes. Nonanal now behaves independently and seem to have been largely decoupled from the dispersive test solutes by removal of the strongly cohesive stationary phases. Principal component analysis indicates that a greater proportion of the variance is accounted for by component 1 largely at the expense of component 2 (see Table IV). The coefficients for the loading plot remain highly weighted towards loading 1 as an orientation axis and loading 2 as a dispersion axis. Loadings 3 and 5 are now heavily weighted towards solutes that behave independently followed by solutes that are not ranked according to any expectations based on physical interactions. Loading 4 is now weighted towards proton donor–acceptor probes but not as dominantly as was loading 3 for the complete data set (Table V). The plot of loading 1 against loading 2 (Fig. 10) now provides a more obvious grouping of the test solutes than was the case for the full data set (Table VI). The group 1 solutes remain the same except that nonanal is now removed and clustered instead with 2-octanone. The proton acceptor solutes (dioxane, benzodioxane, anisole, N,N-dimethylaniline, and 2,4,6-trimethylpyridine) are now clustered into a single group. The other test solutes, except for benzene (10) are clustered into a third group consisting of the proton donor and orientation probes. The next two significant principal components 1 *vs.* 3 and 1 *vs.* 4 produce scatter plots. With the reduced data set there is a poor ability to differentiate between orientation and proton donor solutes. The scores plot of component 1 *vs.* 2 retains the original classification of stationary phases (Table IX) (Fig. 10). The scores plots of 1 *vs.* 3 and 1 *vs.* 4 are scatter plots. Removing those phases that are highly cohesive from the data matrix has primarily effected the classification of the proton donor and orientation test solutes but not the classification of phases. It is likely that at least in part, this is due to the fact that the highly cohesive phases represented the principal examples of strong proton donor–acceptor phases in the data set.

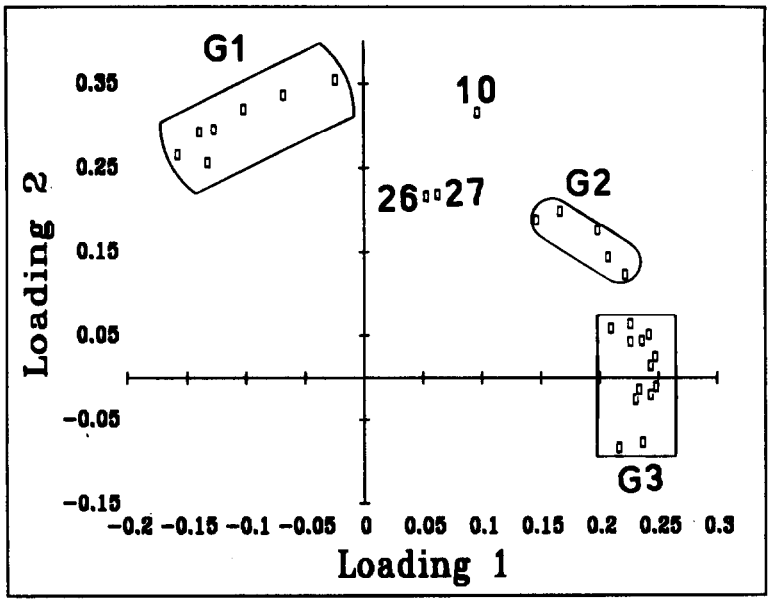
Certain general conclusions can be reached from the data presented in this work. The magnitude of the partial molal Gibbs free energy of solution for specific test solutes can be used to characterize the selectivity of stationary phases that are not highly cohesive. Since many of the test solutes are highly correlated several compounds could be identified as acceptable for this purpose. Based on experience and chromatographic behavior we suggest that the partial molar Gibbs free energy of solution for the methylene group (formally equivalent to the molal free energy since it is calculated by difference and the solvent standard state is self cancelling) be used as a measure of dispersion interactions (which must be augmented by weak induction interactions for polar phases), that nitrobenzene is a suitable test solute for orientation interactions, and that octanol is a suitable test solute for solvent proton acceptor interactions. There is less convincing evidence that any of the test solutes are

TABLE X

SUMMARY OF CORRELATED VALUES ABSTRACTED FROM THE CORRELATION MATRIX FOR 18 PHASES

| | Correlation coefficient |
|---|---|
| *Dispersive interactions* | |
| $\Delta G_k^0$ ($CH_2$) | 1.00 |
| Dihexyl ether | 0.97 |
| Dodecane | 0.94 |
| 2-Octyne | 0.92 |
| 1-Dodecyne | 0.91 |
| *n*-Butylbenzene | 0.87 |
| *Orientation interactions* | |
| Nitrobenzene | 1.00 |
| Benzonitrile | 0.99 |
| Pyridine | 0.98 |
| Nitropentane | 0.97 |
| Aniline | 0.96 |
| N-Methylaniline | 0.96 |
| Nitropropane | 0.95 |
| Dioxane | 0.89 |
| 2,6-Dimethylaniline | 0.88 |
| 1,1,2,2-Tetrachloroethane | 0.88 |
| Butanol | 0.84 |
| Octanol | 0.82 |
| *Proton donor–acceptor interactions* | |
| Octanol | 1.00 |
| Butanol | 0.97 |
| 2-Methyl-2-pentanol | 0.97 |
| Dodecafluoroheptanol | 0.93 |
| N-Methylaniline | 0.92 |
| Aniline | 0.91 |
| 2,6-Dimethylaniline | 0.88 |
| Pyridine | 0.84 |
| Benzonitrile | 0.82 |
| Nitrobenzene | 0.82 |
| Nitropropane | 0.80 |
| Nitropentane | 0.80 |
| *Solutes behaving independently* | |
| Benzene | |
| *cis*-Hydrindane | |
| 2,4,6-Trimethylpyridine | |
| 2-Octanone | |
| Nonanal | |

acceptable for assessing solvent proton donor capacity. For highly cohesive stationary phases, which include OV-275, DEGS, TCEP, QTAPSO and QACES, the partial molal Gibbs free energy of solution for the above test solutes, by itself, is not a good method of classification since its magnitude is dependent on the size of the test solutes in a way that is not correlated with the behavior of less cohesive phases. A more logical classification in terms of solution interactions is obtained by dividing the free energy



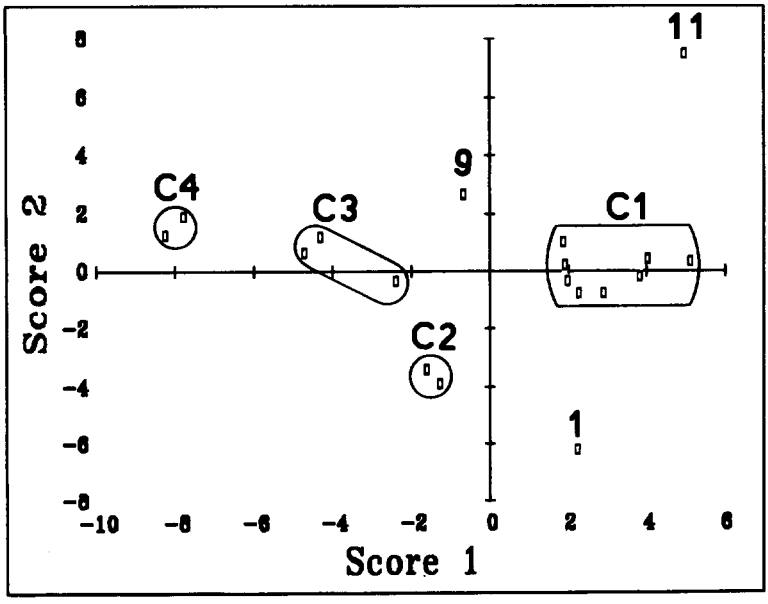




Fig. 10. Plot of loading 1 against 2 and score 1 against 2 for the 18 stationary phases identified in Table I excluding the highly associated phases 10, 14, 15, 20 and 21. See text for the composition of the identified classes (C).

term into a cavity/dispersion term and an interaction term. One method that can be used to determine the cavity/dispersion term is to represent it by the free energy required to accommodate an *n*-alkane of the same size as the test solute. This is not the

only, or necessarily the best solution to the problem [43,44]. Further studies are required to resolve this issue. The collection of test solutes combined with principal component analysis and cluster analysis provides a reasonable method for classifying stationary phases by their similarity of capacity to enter into specific intermolecular interactions. The data matrix is not sufficiently large to warrant the recommendation of a group of preferred phases but does provide useful insight into the selection of phases for initial trial separations based on maximizing the difference in chromatographic selectivity.

ACKNOWLEDGEMENTS

Acknowledgement is made to the donors of the Petroleum Research Fund administered by the American Chemical Society for support of this research. The interest in this project and a gift of OV phases by A. A. Mendicino of Ohio Valley Specialty Chemical are also gratefully acknowledged.

REFERENCES

1 G. E. Baiulescu and V. A. Ilie, *Stationary Phases in Gas Chromatography*, Pergamon Press, New York, 1975.
2 J. A. Yancey, *J. Chromatogr. Sci.*, 24 (1986) 117.
3 T. J. Stark, P. A. Larson and R. D. Dandeneau, *J. Chromatogr.*, 279 (1983) 31.
4 C. F. Poole and S. K. Poole, *Chem. Rev.*, 89 (1989) 377.
5 S. K. Poole and C. F. Poole, *J. Chromatogr.*, 500 (1990) 329.
6 L. Rohrschneider, *Adv. Chromatogr.*, 4 (1967) 333.
7 W. O. McReynolds, *J. Chromatogr. Sci.*, 8 (1970) 685.
8 L. Rohrschneider, *J. Chromatogr.*, 22 (1966) 6.
9 W. O. McReynolds, *Gas Chromatographic Retention Data*, Preston, Evanston, IL, 1966.
10 S. R. Lowry, S. Tsuge, J. J. Leary and T. L. Isenhour, *J. Chromatogr. Sci.*, 12 (1974) 124.
11 M. Chastrette, *J. Chromatogr. Sci.*, 14 (1976) 357.
12 D. L. Massart and H. L. O. De Clercq, *Adv. Chromatogr.*, 16 (1978) 75.
13 E. Fernandez-Sanchez, J. A. Garcia-Dominguez, V. Menendez, J. M. Santiuste and E. Pertierra-Rimada, *J. Chromatogr.*, 402 (1987) 318.
14 J. J. Leary, J. B. Justice, S. Tsuge, S. R. Lowry and T. L. Isenhour, *J. Chromatogr. Sci.*, 11 (1973) 201.
15 S. Wold and K. Anderson, *J. Chromatogr.*, 80 (1973) 43.
16 D. L. Massart, P. Lenders and M. Lauwereys, *J. Chromatogr. Sci.*, 12 (1974) 617.
17 S. Wold, *J. Chromatogr. Sci.*, 13 (1975) 525.
18 D. H. McCloskey and S. J. Hawkes, *J. Chromatogr. Sci.*, 13 (1975) 1.
19 S. R. Lowry, G. L. Ritter, H. B. Woodruff and T. L. Isenhour, *J. Chromatogr. Sci.*, 14 (1976) 126.
20 L. V. Semenchenko and M. S. Vigdergauz, *J. Chromatogr.*, 245 (1982) 177.
21 J. F. K. Huber and G. Reich, *J. Chromatogr.*, 294 (1984) 15.
22 J. A. Garcia-Dominguez, J. Garcia-Munoz, V. Menendez, M. J. Molera and J. M. Santiuste, *J. Chromatogr.*, 393 (1987) 209.
23 E. R. Malinowski and D. G. Howery, *Factor Analysis in Chromatography*, Wiley, New York, 1980.
24 D. L. Massart and L. Kaufman, *The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis*, Wiley, New York, 1983.
25 M. A. Sharaf, D. L. Illman and B. R. Kowalski, *Chemometrics*, Wiley, New York, 1986.
26 O. Strouf, *Chemical Pattern Recognition*, Wiley, New York, 1986.
27 D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte and L. Kaufman, *Chemometrics: A Textbook*, Elsevier, Amsterdam, 1988.
28 J. O. De Beer and A. M. Hyndrickx, *J. Chromatogr.*, 235 (1982) 337.
29 S. K. Poole, B. R. Kersten and C. F. Poole, *J. Chromatogr.*, 471 (1989) 91.
30 B. R. Kersten, C. F. Poole and K. G. Furton, *J. Chromatogr.*, 411 (1987) 43.

31 B. R. Kersten and C. F. Poole, *J. Chromatogr.*, 452 (1988) 191.
32 B. R. Kersten, S. K. Poole and C. F. Poole, *J. Chromatogr.*, 468 (1989) 235.
33 R. M. Pomaville and C. F. Poole, *Anal. Chim. Acta*, 200 (1987) 151.
34 B. R. Kersten and C. F. Poole, *J. Chromatogr.*, 399 (1987) 1.
35 R. M. Pomaville and C. F. Poole, *Anal. Chem.*, 60 (1988) 1103.
36 R. M. Pomaville, S. K. Poole, L. J. Davis and C. F. Poole, *J. Chromatogr.*, 438 (1988) 1.
37 S. K. Poole, K. G. Furton and C. F. Poole, *J. Chromatogr. Sci.*, 26 (1988) 67.
38 R. M. Pomaville and C. F. Poole, *J. Chromatogr.*, 468 (1989) 261.
39 C. F. Poole, S. K. Poole, R. M. Pomaville and B. R. Kersten, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 10 (1987) 670.
40 C. F. Poole, R. M. Pomaville and T. A. Dean, *Anal. Chim. Acta*, 225 (1989) 193.
41 T. O. Kollie and C. F. Poole, unpublished results.
42 A. Hartkopf, S. Grunfeld and R. Delumyea, *J. Chromatogr. Sci.*, 12 (1974) 119.
43 M. H. Abraham and J. C. McGowan, *Chromatographia*, 23 (1987) 243.
44 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.