Factor Analysis of the Solute–Stationary Phase Interactions in Gas-Liquid Chromatography Using the Oblique Procrastes Transformation

Tatsuya TAKAGI,* Yoko SHINDO, Hideaki FUJIWARA, and Yoshio SASAKI

Faculty of Pharmaceutical Sciences, Osaka University, 1–6 Yamada-Oka, Suita, Osaka 565, Japan. Received September 12, 1988

Factor analyses have been carried out for the gas-chromatographic retention data on 21 stationary phases and 190 solutes using the techniques of the principal factor analysis and the oblique procrastes transformations. The eigenvalues showed the necessity of three factors to explain fully the variations of the retention volumes. Further analyses by using the factor scores after the oblique procrastes transformations have disclosed the meanings of the three factors. Namely, the first factor reflects the dispersion interactions and the formation of cavities in the stationary phase, the second one reflects the polar interactions, and the third one reflects the hydrogen-bonding interactions.

Keywords factor analysis; oblique transformation; procrastes transformation; gas-liquid chromatography; McReynolds' constant; principal factor analysis; scaled particle theory; dipole moment; specific retention volume; dispersion interaction

Recently, gas-chromatographic retention data have been analyzed from many points of view. For example, molecular-scientific analyses such as scaled particle theory[1] can make it possible to predict the retention data. In addition, several studies[2] have been carried out on the factor analyses of the retention indices in gas-liquid chromatography in the past fifteen years. For example, Mc-Closkey and Hawks[2b] have analyzed McReynolds' data on 225 stationary phases and 10 solutes using principal component analysis.[3] They concluded that two factors are needed to reproduce the data within the experimental error, the first component being of a polar nature but the nature of the other component being obscure. Chastrette[2c] suggested that the second factor is related to the hydrogen bonding ability from an analysis of Rohrschneider's data[4] and McReynolds' data.[5]

From these studies[1,2] and another one,[6] we can reach the following conclusions concerning the mechanisms of gas-chromatographic retention phenomena. (1) Dispersion interactions and cavity formation make the main contributions to partition between the gas phase and the liquid phase. (2) Polar interactions make a secondary contribution. (3) Although the remaining contributions are small, they are not negligible.

These contributions might be related to the hydrogen bonding ability.

Despite these outstanding achievements, we consider that the two problems stated below still remain unresolved. (1) It is not easy to identify the physicochemical characters of the extracted factors because of the complicated structures of the stationary phases. So, the physicochemical characters of the factor considered to be the hydrogen bonding ability factor have not been quantitatively cleared up. (2) The polar factor and the hydrogen bonding ability factor may not be orthogonalized. Since an abstract factor does not always have chemical or physical meanings, extracted factors are not necessarily always orthogonalized with each other.

We considered that using data concerning many solutes and several stationary phases would avoid the problem of (1) because correlations between factor scores and the physical properties of the solutes can be investigated. In addition, oblique transformations should be useful to avoid the problem of (2).

In this report, three novel attempts were made to solve these problems as stated below. (1) We dealt with the new data matrix on 21 stationary phases and 190 solutes cited from the McReynolds' data book.[7] The data used are neither retention indices nor retention increments, but the logarithms of the specific retention volumes (V_g). This is because subtracting $\log V_g^\circ$ from $\log V_g^x$ increases the error relatively, where V_g^x means the specific retention volume of a solute and V_g° means that of a corresponding normal alkane. Hence, extracted factors whose eigenvalue is relatively small might be error factors. Also, the physicochemical meaning of the subtracted value of $\log V_g^\circ$ itself may not be unique. (2) We used the oblique procrastes transformation (OPT)[8] after extracting three factors by the principal factor analysis (PFA). The OPT method enables us to fit the factor pattern for a hypothetical factor pattern. It is suitable for our purpose to extract the factors having chemically clear meanings. (3) Correlation analyses between the factor scores and the physical constants of the solutes were carried out. Physicochemical meanings of the extracted factors, especially Chastrette's[2c] second one, can be clarified from this approach.

Calculation

PFA[3] For the initial values of the communalities, we have adopted the squared multiple correlation coefficients (SMC) which indicate the lower bounds for the communalities. Four cycles of computation were needed to converge the factor loadings. The factor scores were estimated by the least-squares method.

OPT[8] We have used the hypothetical factor patterns for the desired matrix to be fitted. The estimation of the factor scores were made in the same way as in the above paragraph.

Computational Details The double precision version of the program package for multivariate statistical analysis, MVA, was used; it had been

TABLE I. List of Stationary Phases Used in This Study

1 APPIEZON L	(API-L)	12 Ethylene glycol adipate	(EGA)
2 CASTOROWAX	(CASTOR)	13 Neopentyl glycol adipate	(NPGA)
3 CARBOWAX 1000	(CW1000)	14 QUADROL	(QUAD)
4 Di-2-ethylhexyl sebacate	(D2EHS)	15 SE30	(SE30)
5 DC 550	(DC550)	16 Sucrose acetate isobutyl-	
6 HALLCOMID	(HCM18)	ate	(SAIB)
7 HYPROSE SP-80	(HYP)	17 Tricresyl phosphate	(TCP)
8 IGEPAL CO 880	(IGE880)	18 TRITON X305	(TX305)
9 Polyphenyl ether 5-ring	(PPE5R)	19 UCON HB-2000	(U2000)
10 Diethylene glycol adipate	(DEGA)	20 VERSILUBE F-50	(VF50)
11 Diisodecyl phthalate	(DIDP)	21 XF 1150	(XF1150)

TABLE II. List of Solutes Used in This Study

1 Methanol	51 Ethylal	101 Butyl ethanoate	151 2,6-Dioxaheptane
2 Ethanol	52 Propanal	102 2-Methylpropyl ethanoate	152 3,7-Dioxanonane
3 1-Propanol	53 Butanal	103 1-Methylpropyl ethanoate	153 4,5-Diethyl-2-ethenyl-1,3-dioxolane
4 2-Methylethanol	54 2-Methylpropanal	104 2-Methyl-2-propyl ethanoate	154 2-Ethyl-2-methyl-1,3-dioxolane
5 1-Butanol	55 Pentanal	105 Pentyl ethanoate	155 Ethyl butyl ether
6 2-Methyl-1-propanol	56 3-Methylbutanal	106 3-Methylbutyl ethanoate	156 Dipropyl ether
7 1-Methyl-1-propane	57 3,3-Dimethylbutanal	107 2-Methylbutyl ethanoate	157 Dibutyl ether
8 2-Methyl-2-propanol	58 Heptanal	108 2-Methyl-2-butyl ethanoate	158 Dipentyl ether
9 1-Pentanol	59 Octanal	109 Hexyl ethanoate	159 Oxolane
10 3-Methyl-1-butanol	60 2-Ethylhexanal	110 4-Methyl-2-pentyl ethanoate	160 2-Methyl-1,2-epoxypropane
11 2-Pentanol	61 2-Propenal	111 2-Ethyl-1-butyl ethanoate	161 2-Methyloxolane
12 3-Pentanol	62 $(CH_2CHCHO)_3$	112 Heptyl ethanoate	162 2-Methyl-2,4-oxole
13 2-Methyl-1-butanol	63 2-Butenal	113 Cyclohexyl ethanoate	163 2,5-Dimethyloxolane
14 2-Methyl-2-butanol	64 2-Ethyl-2-butenal	114 2-Propenyl ethanoate	164 Ethane
15 3-Methyl-2-butanol	65 2-Ethyl-2-hexenal	115 1-Methyl-1-ethenyl ethanoate	165 Butane
16 2,2-Dimethyl-1-propanol	66 2,4-Hexadienal	116 1,1-Ethanediol diethanoate	166 Hexane
17 1-Hexanol	67 2-Propanone	117 1,2-Ethanediol diethanoate	167 Octane
18 2-Hexanol	68 2-Butanone	118 Methyl propyonate	168 Decane
19 3-Hexanol	69 2-Pentanone	119 Ethyl propyonate	169 Dodecane
20 2-Methyl-1-pentanol	70 3-Pentanone	120 1-Propyl propyonate	170 Tetradecane
21 4-Methyl-1-pentanol	71 3-Hexanone	121 1-Butyl propyonate	171 Benzene
22 2-Methyl-2-pentanol	72 3-Methyl-2-pentanone	122 2-Methyl-1-propyl propyonate	172 Methylbenzene
23 3-Methyl-2-pentanol	73 4-Methyl-2-pentanone	123 1-Pentyl-1-propyonate	173 1,2-Dimethylbenzene
24 4-Methyl-2-pentanol	74 3,3-Dimethyl-2-butanone	124 2-Propenyl-propyonate	174 1,3-Dimethylbenzene
25 2-Methyl-3-pentanol	75 2-Heptanone	125 Methyl butylate	175 1,4-Dimethylbenzene
26 3-Methyl-3-pentanol	76 3-Heptanone	126 Ethyl butylate	176 Ethylbenzene
27 2-Ethyl-1-butanol	77 2-Octanone	127 Methylethyl butylate	177 1,2-Diethylbenzene
28 2,2-Dimethyl-1-butanol	78 2-Nonanone	128 Butyl butylate	178 1,3-Diethylbenzene
29 2,3-Dimethyl-2-butanol	79 5-Nonanone	129 Pentyl butylate	179 1,4-Diethylbenzene
30 3,3-Dimethyl-2-butanol	80 Cyclopentanone	130 Ethenyl butylate	180 Chloroethane
31 2-Heptanol	81 Cyclohexanone	131 Butyl methylethanoate	181 Tetrachloromethane
32 3-Heptanol	82 3-Butene-2-one	132 2-Methylpropyl methylethanoate	182 Trichloromethane
33 4-Heptanol	83 5-Hexene-2-one	133 Methyl 3-butenoate	183 2-Chloroethanol
34 2,2-Dimethyl-1-pentanol	84 4-Methyl-3-pentene-2-one	134 Ethyl 3-butenoate	184 3-Hydroxy-2-butanone
35 2,4-Dimethyl-3-pentanol	85 2,3-Butanedione	135 Propyl 3-butenoate	185 2-Hydroxy-2-methyl-3-butanone
36 1-Octanol	86 2,4-Pentanedione	136 Butyl 3-butenoate	186 Dimethoxymethane
37 2-Octanol	87 Ethyl methanoate	137 Methyl 2-methyl-2-propenoate	187 1,4-Dioxixane
38 2-Ethyl-1-hexanol	88 Propyl methanoate	138 3,5-Dioxaheptane	188 1,3,5-Trioxixane
39 2-Ethyl-4-methyl-1-pentanol	89 Methylethyl methanoate	139 2-Methyl-3,5-dioxaheptane	189 1,3,5-Trioxepane
40 Cyclopentanol	90 2-Methylpropyl methanoate	140 3-Methyl-4,6-dioxaoctane	190 Water
41 Cyclohexanol	91 1-Methylpropyl methanoate	141 2,6-Dimethyl-3,5-dioxaheptane	
42 2-Propen-1-ol	92 Pentyl methanoate	142 5,7-Dioxaundecane	
43 2-Propyn-1-ol	93 2-Pentyl methanoate	143 1,3-Dioxolane	
44 2-Buten-1-ol	94 3-Pentyl methanoate	144 4-Methyl-1,3-dioxolane	
45 3-Buten-2-ol	95 Hexyl methanoate	145 1,3-Dioxixane	
46 2-Methyl-2-propen-1-ol	96 2-Propenyl methanoate	146 4-Methyl-1,3-dioxixane	
47 1-Penten-3-ol	97 Methyl ethanoate	147 4,5-Dimethyl-1,3-dioxixane	
48 1-Penten-4-ol	98 Ethyl ethanoate	148 1,3-Dioxepane	
49 2-Methyl-3-butyn-2-ol	99 Propyl ethanoate	149 2-Methyl-1,3-dioxolane	
50 2-Methyl-1-pentyn-3-ol	100 Methylethyl ethanoate	150 2-Methyl-1,4-dioxixane	

coded by three of us and some co-workers.[9] All computations were carried out on personal computers (NEC PC9801 and EPSON PC286V).

Data Matrix Used in the Factor Analyses The stationary phases used in the factor analyses are listed in Table I[7]; various kinds of stationary phases, i.e., nonpolar stationary phases such as SE30, polar stationary phases such as CARBOWAX 1000, and protic stationary phases such as HYPROSE, etc., were used. Hydrocarbons, alcohols, aldehydes, esters, ketones, ethers, water, and some of their chloro derivatives were used for the solutes,[7] as listed in Table II. The abbreviations for the stationary phases listed in Table I and the numbers assigned for the solutes in Table II will be used in the following sections.

Results and Discussion

1. PFA Number of the Extracted Factors: The correlation matrix for the 21 stationary phases is shown in Table III. Although most of the correlation coefficients are close to unity, some of them show deviations from this, such as 0.629 for the correlation coefficient between API-L

and HYP. The correlation coefficients between the nonpolar and the polar stationary phases are small relative to those between the nonpolar stationary phases and those between the polar stationary phases.

We extracted three factors to account for the data satisfactorily (Table IV). The communalities after the convergence are listed in Table V. Even for XF1150 which exhibits the minimum communalities, about 98% of the variations of the retention data are explicable. Considering the errors in the retention data, the variations of the results are sufficiently explicable by the three factors. We also attempted to extract more abstract factors. However, physical or chemical features were not revealed for these higher-order factors. Therefore, we concluded that these three factors are the factors required and sufficient to interpret the data treated here.

Physicochemical Meanings of the Extracted Factors: The

TABLE III. Correlation Coefficients Matrix for 21 Stationary Phases

	API-L	CW1000	DEGA	D2EHS	DIDP	DC550	EGA	IGE880	NPGA	PPE5R	QUAD	SE30	SAIB	TCP	TX305	U2000	VF50	XF1150	CASTOR	HCM18	HYP
API-L	1.000																				
CW1000	0.715	1.000																			
DEGA	0.715	0.994	1.000																		
D2EHS	0.974	0.842	0.837	1.000																	
DIDP	0.963	0.846	0.843	0.990	1.000																
DC550	0.984	0.791	0.797	0.984	0.979	1.000															
EGA	0.736	0.990	0.997	0.852	0.858	0.816	1.000														
IGE880	0.803	0.990	0.984	0.906	0.907	0.866	0.985	1.000													
NPGA	0.852	0.971	0.971	0.940	0.941	0.907	0.976	0.992	1.000												
PPE5R	0.957	0.848	0.860	0.977	0.978	0.988	0.876	0.907	0.943	1.000											
QUAD	0.698	0.984	0.971	0.830	0.827	0.768	0.970	0.973	0.954	0.815	1.000										
SE30	0.992	0.769	0.767	0.988	0.979	0.995	0.787	0.849	0.891	0.973	0.754	1.000									
SAIB	0.894	0.944	0.943	0.967	0.966	0.943	0.953	0.977	0.993	0.968	0.926	0.931	1.000								
TCP	0.900	0.930	0.929	0.966	0.980	0.943	0.937	0.967	0.984	0.966	0.910	0.932	0.991	1.000							
TX305	0.804	0.990	0.984	0.908	0.909	0.866	0.985	0.999	0.992	0.908	0.972	0.850	0.979	0.968	1.000						
U2000	0.864	0.967	0.958	0.949	0.946	0.912	0.963	0.992	0.994	0.939	0.953	0.903	0.991	0.983	0.993	1.000					
VF50	0.994	0.739	0.738	0.977	0.968	0.993	0.760	0.823	0.868	0.965	0.724	0.998	0.912	0.914	0.824	0.881	1.000				
XF1150	0.775	0.968	0.979	0.876	0.886	0.855	0.982	0.974	0.977	0.910	0.947	0.822	0.964	0.951	0.974	0.960	0.799	1.000			
CASTOR	0.937	0.888	0.881	0.985	0.993	0.959	0.892	0.938	0.962	0.965	0.878	0.960	0.978	0.990	0.939	0.969	0.945	0.906	1.000		
HCM18	0.892	0.904	0.883	0.962	0.965	0.909	0.892	0.943	0.956	0.915	0.907	0.920	0.962	0.974	0.944	0.969	0.898	0.892	0.985	1.000	
HYP	0.629	0.984	0.974	0.775	0.777	0.711	0.971	0.958	0.933	0.775	0.990	0.690	0.897	0.878	0.958	0.927	0.656	0.940	0.834	0.863	1.000

TABLE IV. Factor Loadings Calculated by PFA Method

	Factor 1	Factor 2	Factor 3
CCR[a]	0.922	0.990	0.997
API-L	0.898	0.431	−0.001
CASTOR	0.983	0.133	−0.094
CW1000	0.947	−0.315	−0.004
DEGA	0.945	−0.311	0.083
D2EHS	0.968	0.238	−0.044
DIDP	0.969	0.224	−0.040
DC550	0.942	0.321	0.087
EGA	0.953	−0.279	0.089
HCM18	0.971	0.050	−0.227
HYP	0.901	−0.419	−0.052
IGE880	0.981	−0.183	−0.005
NPGA	0.994	−0.095	0.022
PPE5R	0.966	0.209	0.150
QUAD	0.931	−0.326	−0.089
SE30	0.931	0.359	0.000
SAIB	0.998	0.002	0.030
TCP	0.993	0.031	−0.016
TX305	0.982	−0.181	−0.055
U2000	0.994	−0.071	−0.037
VF50	0.912	0.399	0.020
XF1150	0.960	−0.195	0.138

a) Cumulative contribution ratio.

TABLE V. Communalities Calculated by Convergence PFA Method

API-L	0.9913	NPGA	0.9974
CASTOR	0.9926	PPE5R	0.9982
CW1000	0.9965	QUAD	0.9814
DEGA	0.9961	SE30	0.9965
D2EHS	0.9955	SAIB	0.9964
DIDP	0.9900	TCP	0.9876
DC550	0.9983	TX305	0.9960
EGA	0.9946	U2000	0.9953
HCM18	0.9967	VF50	0.9920
HYP	0.9893	XF1150	0.9793
IGE880	0.9955		

first factor, whose loadings amount to 92.2% of the variance, is thought to be a size factor since the PFA has a tendency to extract a factor in common with the variables.

TABLE VI. Correlation Coefficients between Factor Loadings and McReynolds' Constants

	Factor 2	Factor 3
ΔI (benzene)	0.928	0.319
ΔI (butanol)	0.993	0.055
ΔI (2-pentanone)	0.949	0.274
ΔI (nitropropane)	0.950	0.273
ΔI (pyridine)	0.957	0.225
$\sum \Delta I$	0.968	0.219

This size factor reflects dispersion interactions and the energies of cavity formation in the liquid phase. The loading values of the nonpolar stationary phases are positive and large whereas those of the polar stationary phases are negative and small with regard to the second factors. For example, the loading value of API-L is 0.431 and that of SE30 is 0.359, whereas that of CW1000 is −0.315. Thus, the second factor is thought to be a polar one.

Contrary to the above case of the two factors, the physicochemical features of the third factor are somewhat ambiguous because of its small contribution ratio. Nevertheless, the third factor is thought to be a factor relating to hydrogen bondings for the reasons stated below: (1) The loading values of the stationary phases having proton donors such as HCM18 are negative, and those of the stationary phases having proton acceptors such as DEGA, PPE5R, and EGA are positive. (2) The loading values of the nonpolar stationary phases such as API-L and SE30 are approximately zero, and the stationary phases having both proton donors and proton acceptors such as CASTOR, CW1000, and QUAD show slightly negative values.

Table VI shows the correlation coefficients between McReynolds' constants and the factor loading values; good correlations are observed between the loading values of the second factor and the McReynolds' constants. This also indicates that the second factor is a polar one. A plot of the second factor scores against the third ones (Fig. 1) indicates that the third factor plays an important role in separating
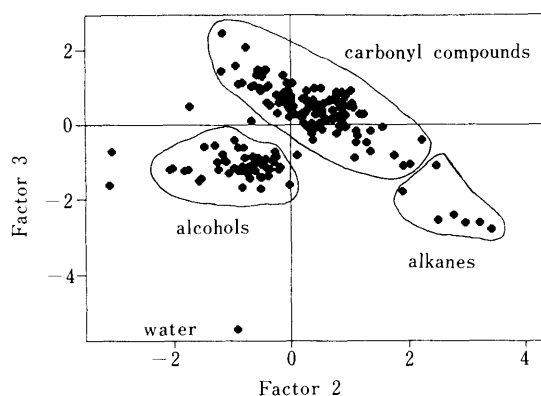
Fig. 1. Two-Dimensional Plot of Gas-Chromatographic Specific Retention Data

Horizontal axis = factor 2 and vertical axis = factor 3 calculated by the PFA method.

TABLE VII. Correlation Coefficients between Factor Scores and Molecular Properties

	MW	bp	μ	H-Index
No. of data	190	99	92	190
Factor 1	0.712[a]	0.959[a]	0.019	0.208
Factor 2	0.587[a]	0.124	0.495[a]	0.716[a]
Factor 3	0.067	0.105	0.413[b]	0.746[a]

a) H_0: $r = 0$ is rejected with $\alpha = 0.0001$. b) H_0: $r = 0$ is rejected with $\alpha = 0.01$.

alcohols from carbonyl compounds.

Then, we calculated the correlation coefficients between the factor scores and the physical constants[10] of the solutes to elucidate the physicochemical features of the three factors (Table VII). The hydrogen-bonding index (H-index) were calculated according to the following rules: (1) + 1 for one proton donating group such as O–H, (2) − 1 for one proton accepting group such as R-O-R(H) or C = O. For example, the H-index of $CH_3COCOCH_3$ is equal to -2 and that of 1-octanol is equal to 0.

According to the above results, we concluded that the second factor is a polar one. Some ambiguities remain concerning the physicochemical meanings of the third factor although the third factor appears to be a factor of hydrogen bonding ability.

2. OPT Procedure for the Transformation: Although the factors obtained in the above section can explain the variations of the retention data, the physicochemical features of the third factor are slightly ambiguous. Therefore, we examined the oblique procrastes transformation[8] to clarify them. The hypothetical factor loadings were prepared as follows. (1) Since preparing a hypothetical factor pattern used for the procrastes transformation is somewhat arbitrary, the principal factor loading matrix was used as the basis pattern. (2) To clarify the second factor to be the polar factor, the order of the loading values of the second factor was changed to fit the order of the McReynolds' constants. Then the remaining loading values of the third factor were calculated from the other two loading values and the contribution ratios of the variables.

The hypothetical factor pattern used in this study is shown in Table VIII.

Physicochemical Features of the Third Factor: The fac-

TABLE VIII. Hypothetical Factor Pattern Used for Procrastes Transformation

	Factor 1	Factor 2	Factor 3
API-L	0.8976	0.4307	−0.0008
CASTOR	0.9829	0.1327	0.0944
CW1000	0.9472	−0.3109	0.0515
DEGA	0.9447	−0.3255	0.0000
D2EHS	0.9681	0.2375	−0.0443
DIDP	0.9687	0.2237	−0.0401
DC550	0.9422	0.3208	0.0875
EGA	0.9533	−0.3151	0.0000
HCM18	0.9708	0.2087	−0.1030
HYP	0.9007	−0.4189	0.0519
IGE880	0.9808	−0.1806	0.0308
NPGA	0.9939	−0.0950	−0.0214
PPE5R	0.9655	0.0497	−0.2520
QUAD	0.9314	−0.1948	0.2756
SE30	0.9312	0.3988	0.0000
SAIB	0.9977	0.0018	0.0297
TCP	0.9932	0.0315	−0.0163
TX305	0.9815	−0.1831	0.0000
U2000	0.9945	−0.0706	0.0373
VF50	0.9125	0.3595	0.1737
XF1150	0.9603	−0.2791	0.0000

TABLE IX. Factor Pattern Calculated by Oblique Procrastes Transformation

	Factor 1	Factor 2	Factor 3
CCR[a]	0.910	0.973	0.997
API-L	0.900	0.380	−0.062
CASTOR	0.987	0.170	0.121
CW1000	0.950	−0.276	0.153
DEGA	0.946	−0.321	0.064
D2EHS	0.971	0.234	0.041
DIDP	0.972	0.220	0.040
DC550	0.943	0.234	−0.117
EGA	0.955	−0.296	0.050
HCM18	0.977	0.171	0.278
HYP	0.904	−0.340	0.227
IGE880	0.984	−0.159	0.120
NPGA	0.997	−0.096	0.069
PPE5R	0.966	0.100	−0.147
QUAD	0.936	−0.237	0.241
SE30	0.934	0.317	−0.041
SAIB	1.000	−0.015	0.034
TCP	0.996	0.037	0.072
TX305	0.985	−0.156	0.119
U2000	0.998	−0.041	0.121
VF50	0.915	0.340	−0.073
XF1150	0.961	−0.250	−0.023

a) Cumulative contribution ratio.

TABLE X. Correlation Coefficients Matrix among Three Factors Calculated by Oblique Procrastes Transformation

	Factor 1	Factor 2	Factor 3
Factor 1	1.00000		
Factor 2	0.01550	1.00000	
Factor 3	−0.07674	−0.29198	1.00000

tor pattern and the correlation coefficient matrix among the three factors are shown in Tables IX and X, respectively. The contribution ratio of the third factor turned out to be larger than that obtained by the PFA method. This has

TABLE XI. Correlation Coefficients between Factor Patterns and McReynolds' Constants

	Factor 2	Factor 3
ΔI (benzene)	0.972	0.372
ΔI (butanol)	0.984	0.612
ΔI (2-pentanol)	0.985	0.420
ΔI (nitropropane)	0.985	0.423
ΔI (pyridine)	0.983	0.461
$\sum \Delta I$	0.992	0.473

TABLE XII. Correlation Coefficients between Molecular Properties and Factor Scores Calculated by Oblique Procrastes Transformation

	MW	bp	μ	H-Index
No. of data	190	99	92	190
Factor 1	0.713[a]	0.961[a]	0.014	0.209
Factor 2	0.458[a]	0.155	0.591[a]	0.512[a]
Factor 3	0.174	0.127	0.253	0.852[a]

a) H_0: $r = 0$ is rejected with $\alpha = 0.0001$.

facilitated clarification of the physicochemical features for the third factor. The correlation coefficients between the factor scores and the McReynolds' constants (Table XI) show that the correlation between the second factor and the constants becomes larger than in the case of the PFA. This indicates that the OPT method has made the physicochemical features clearer.

Correlation coefficients between the factor scores and the physical properties of the solutes are shown in Table XII. The correlation coefficient between the second factor scores and the dipole moments of the solutes turned out to be larger than in the case of the PFA. Considering that dipole moment does not represent the whole of the polar effects, this correlation coefficient is regarded as significant. In fact, the null hypothesis, H_0:$r = 0$, is rejected with the level of significance 0.0001. In addition, the value of correlation coefficient is the maximum one among those between the second factor and the four physicochemical parameters. The correlation coefficient between the H-index and the third factor also turned out to be larger. Considering the coarse definition for the H-index, we think that this value shows a satisfactory correlation. Hence we concluded that the third factor is related to the hydrogen-bonding ability.

The two-dimensional plot of the second factor scores against the third one is shown in Fig. 2. The axis of the third factor is parallel to the line connecting the centers of the two clusters, i.e., the clusters of the carbonyl compounds and alcohols. This result also shows that the physicochemical features of the three factors, especially the third factor, are made clearer than in the PFA method.

Conclusion

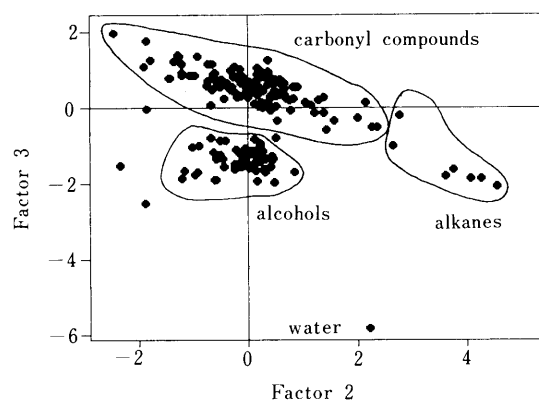(1) At least three factors are needed to explain the variations of gas-chromatographic specific retention volumes.

Fig. 2. Two-Dimensional Plot of Gas-Chromatographic Specific Retention Data

Horizontal axis = factor 2 and vertical axis = factor 3 calculated by the OPT method.

(2) The first factor is considered to be a factor of dispersion interactions and the formation of cavities.

(3) The second factor appears to be a polar factor bacause of the good correlation between this factor and the McReynolds' constants.

(4) The OPT method clarifies that the physicochemical features of the third factor are related to hydrogen-bonding ability.

References

1) a) R. A. Pierotti, Chem. Rev., 76, 717 (1976); b) M. H. Abraham and A. Nasehzadeh, J. Chem. Soc., Faraday Trans. 1, 77, 321 (1981); c) H. Fujiwara, R. Miyagi, I. Ohtaku, and Y. Sasaki, Chem. Lett., 1988, 1389.
2) a) P. H. Weiner and D. G. Howery, Can. J. Chem., 50, 448 (1972); b) D. H. McCloskey and S. J. Hawks, J. Chromatogr. Sci., 13, 1 (1975); c) M. Chastrette, ibid., 14, 13 (1976); d) E. Malinowski, Anal. Chim. Acta, 122, 327 (1980); e) R. F. Hirsch, R. J. Gaydosh, and J. R. Chretien, Anal. Chem., 52, 723 (1980); f) R. Fellous, D. Lafe de Micheaux, L. Lizzani-Cuvelier and R. Luft, J. Chromatogr., 248, 35 (1982); g) D. G. Howery and J. M. Soroka, Anal. Chem., 58, 3091 (1986); h) D. G. Howery, G. D. Williams, and N. Ayala, Anal. Chim. Acta, 189, 339 (1986); i) D. G. Howery and J. M. Soroka, J. Chromatogr. Sci., 25, 149 (1987); j) E. Fernandez-Sanchez, J. A. Garcia-Dominguez, V. Menendez, J. M. Santiuste, and E. Pertierra-Rimada, J. Chromatogr., 402, 318 (1987); k) D. G. Howery and J. M. Soroka, J. Chemom., 1, 91 (1987); l) K. Szymoniak and J. R. Chretien, J. Chromatogr., 404, 11 (1987).
3) K. Takeuchi, H. Yanai, and B. N. Mukherjee, "The Foundations of Multivariate Analysis," Wiley Eastern, New Delhi, 1982.
4) L. Rohrschneider, J. Chromatogr., 22, 6 (1966).
5) W. O. McReynolds, J. Chromatogr. Sci., 8, 685 (1970).
6) R. Kaliszan and H. D. Hoeltje, J. Chromatogr., 234, 303 (1982).
7) W. O. McReynolds, "Gas Chromatographic Retention Data," Preston Pub., Illinois, 1980.
8) H. H. Harman, "Modern Factor Analysis," 3rd ed., The University of Chicago Press, Chicago, 1976.
9) T. Takagi, Y. Shindo, N. Iwata, A. Iwata, K. Tange, T. Katayama, S. Fujii, and Y. Sasaki, "Proceedings of the 4th Software Conference," Osaka, 1988, p. 285.
10) a) A. L. McClellan, "Tables of Experimental Dipole Moments," Freeman, San Francisco, 1963; b) "Kagaku-Benran II," The Chemical Society of Japan, Tokyo, 1984.