CHROM. 5976

RETENTION INDEX–STRUCTURE RELATIONSHIPS FOR ALCOHOLS USING FACTOR ANALYSIS

PAUL H. WEINER*, CHARLES J. DACK AND DARRYL G. HOWERY**
The Department of Chemistry, Brooklyn College of the City University of New York, Brooklyn, N.Y. 11210 (U.S.A.)

(Received January 28th, 1972)

SUMMARY

The technique of factor analysis is applied to a data matrix composed of the retention indices of twenty-six saturated and unsaturated alcohols measured on twenty-five solvents. Five factors were found sufficient to span the data space of saturated alcohols, while six factors were required when unsaturated alcohols were included in the data matrix. Successful rotations were made into the following chemically and physically significant solute parameters: (1) unity, (2) molecular weight, (3) molar refraction, (4) carbon number, (5) boiling point (°K), (6) placement of hydroxyl group, (7) ln vapor pressure and (8) molar heat of vaporization. Then, restricting the data matrix to saturated alcohols only, various combinations of five out of the eight identified parameters were used in an attempt to best span the space, the criteria of best fit being the mean error between experimental and predicted retention indices. Using the "best set", equations were developed which permitted the retention index of a new alcohol to be predicted on any of the twenty-five solvents, knowing only the five parameters of that solute. The resulting equations were successfully tested on four alcohol solutes which were not included in the original data matrix but for which experimental retention index data were available.

INTRODUCTION

Recently TAKÁCS *et al.*[1] showed that it is possible to relate the structure of saturated paraffins to their retention indices determined on squalane. This problem was amenable to solution because of the simplicity of the solutes and the use of only one stationary phase. In a more ambitious study, CASTELLO AND D'AMATO[2] have attempted to correlate the retention indices of the alkyl iodides with their molecular structures and physical properties. They correlated retention indices with boiling

---

* Present address: Department of Chemistry, University of Mississippi, University, Miss. 38677, U.S.A.

** To whom all correspondence should be addressed.

point, placement of the iodine atom on the chain, and molecular volume. Again, however, the correlations were obtained on only one stationary phase, tricresyl phosphate. In order to make these studies more general, one would like to correlate data of many solutes taken on a series of stationary phases ranging from nonpolar to highly polar. This more general problem has not been solved because of the lack of an adequate procedure for isolating the controlling factors operative in determining retention indices. For solutes more complex than the paraffins, there exist additional factors, each of which affect the observed retention index to a different extent. The magnitude of each subfactor should depend upon some property of the solute in question, as well as some property of the stationary phase being used. For example, in a series of alcohol solutes, one would expect the number of controlling solute factors to consist of the sum of those present for saturated paraffins of similar carbon chain structure plus additional factors attributable to the presence of the polar hydroxyl group in the solute. These additional factors would include at a minimum dipole–dipole, dipole–induced dipole and hydrogen-bonding interactions. With such a multifactor problem, no method of attack has proven successful in isolating the individual factors.

Recently, the technique of factor analysis has been helpful in unraveling the number and nature of the controlling factors operative in several multi-dimensional problems[3–9]. Factor analysis has been applied to the problems of solvent effects in nuclear magnetic resonance[3–6], activity coefficients determined at infinite dilution by gas chromatography[7] and solute structure–retention index relationships[8,9]. This technique can be utilized whenever the quantity being factor analyzed can be expressed as a linear sum of factors, each factor being a product function of solute and solvent parameters[3]. Recent work indicates that these two criteria are met by the retention index function.

Using the data of ROHRSCHNEIDER[10] who reported retention indices of thirty nonpolar and polar solutes on twenty-three solvents, WEINER AND HOWERY[8,9] applied the technique of factor analysis and found that eight factors were required to span the factor space such that all data were reproduced within a column average error of three retention index units or less. Several physically and chemically significant parameters were successfully associated with the abstract eigenvectors of the space. By a least squares rotational scheme, physically significant parameters such as gas phase fugacity[8] and solute molar polarizations[9] were identified with abstract eigenvectors of the space. Chemically significant parameters identified with abstract eigenvectors were solute hydrogen-bonding ability and electron-donor tendencies[8]. These conclusions may be quite general since a large number of functional groups for both solutes and solvents are represented in the data matrix of ROHRSCHNEIDER.

In the present study, the technique of factor analysis is applied to the problem of identifying the controlling factors which affect the retention indices of saturated and unsaturated alcohols in nonpolar and polar solvents. The data for this study was taken from the compilation of MCREYNOLDS[11] who reported retention indices for many solutes on seventy-six solvents. Retention indices at 120°C of twenty-six saturated and unsaturated alcohols (solutes shown in Table I), taken on twenty-five solvents (solvents shown in Table II), were chosen for factor analysis. The other information shown in Tables I and II will be discussed later under the section entitled *Uniqueness tests.*

RESULTS

Number of fundamental factors

The problem was divided into two parts in order to ascertain if the introduction of unsaturation in the solute molecules affected the number of controlling factors. The data matrix for part one contained retention indices for saturated alcohols on all twenty-five stationary phases; the matrix for part two contained data for both saturated and unsaturated alcohols on all stationary phases. We found that five factors were sufficient to span the saturated-alcohol retention index space while six factors were required for the data matrix containing both saturated and unsaturated alcohols. These conclusions were based on observing the column average errors for the two problems as the number of abstract eigenvectors in the space was systematically increased, and were further substantiated when we tested physically and chemically significant parameters for possible identification with the abstract eigenvectors of both data matrices. The same fit for a given test factor was obtained with five factors in the saturated alcohol problem as with six factors in the problem containing the unsaturated alcohol data. In the remainder of this paper (except for the last section), all results are for the larger problem using six factors in the least squares rotation matrix.

TABLE I

UNIQUENESS TESTS FOR SOLUTES

Solute	*Uniqueness*		*Additional values*
	Number	*Value*	
Methanol	1	0.76	2, 0.31
Ethanol	2	0.20	1, 0.31
Propanol	3	0.10	2, 0.10
Isopropanol	4	0.14	1, 0.13; 2, 0.14; 8, 0.16
Butanol	5	0.12	6, 0.15; 9, 0.12
Isobutanol	6	0.24	5 and 9, 0.15
sec.-Butanol	7	0.11	8, 0.13
tert.-Butanol	8	0.25	4, 0.16; 7, 0.13
Pentanol	9	0.14	5, 0.12; 6 and 13, 0.15; 21, 0.18
2-Pentanol	10	0.05	
3-Pentanol	11	0.09	17, 0.11
2-Methyl-2-butanol	12	0.14	8, 0.14; 16 and 18, 0.12
Hexanol	13	0.19	9, 0.15; 21, 0.28
2-Hexanol	14	0.07	21, 0.12
3-Hexanol	15	0.11	17, 0.12; 18, 0.11; 19, 0.13; 20, 0.14
2-Methyl-2-pentanol	16	0.12	8 and 12, 0.12; 18, 0.10
3-Methyl-3-pentanol	17	0.19	11 and 12 and 15 and 19, 0.12; 18, 0.17
2,3-Dimethyl-2-butanol	18	0.15	12 and 20, 0.12; 17, 0.17
3-Heptanol	19	0.19	15, 0.13; 17, 0.11; 20, 0.18
4-Heptanol	20	0.19	15, 0.14; 19, 0.18; 21, 0.15
Octanol	21	0.46	13, 0.28; 19, 0.19
Cyclopentanol	22	0.43	23, 0.42
Cyclohexanol	23	0.45	22, 0.42
2-Propen-1-ol	24	0.17	22, 0.10; 26, 0.08
2-Propyn-1-ol	25	0.81	26, 0.19
1-Penten-3-ol	26	0.10	24, 0.08; 25, 0.19

Uniqueness tests

Before proceeding to the testing of physically and chemically significant parameters, it is important to determine if any of the solutes or solvents in the scheme exhibit a uniqueness in properties not present in the other compounds in the scheme. This uniqueness test is accomplished by defining a test vector consisting of unity for the solute or solvent suspected of containing the unique factor and of zeros for all other points on the vector[9]. The results of the uniqueness tests for the solutes are shown in Table I, where each row of the table corresponds to a separate uniqueness test vector. The only solutes which seemed to be truly unique were 2-propyn-1-ol (the only solute containing a triple bond) and methanol. Other solutes which seemed to show some slight degree of uniqueness were cyclopentanol and cyclohexanol. For each cycloalcohol uniqueness test vector, the only other high value was for the other cycloalcohol. Another example of the correlation of high values on a given uniqueness test involves the two double bond-containing solutes, 2-propen-1-ol and 1-penten-3-ol. The two double bond-containing alcohols share a common unique property not only with each other, but also with the triple-bonded solute, 2-propyn-1-ol, and interestingly with cyclopentanol.

We can also test for unique properties among the stationary-phase solvents with the uniqueness test vector by using the transposed data matrix[9]. The results are shown in Table II. The sole solvent exhibiting quite unique properties was diglycerol, the only hydroxyl-rich solvent.

TABLE II

UNIQUENESS TESTS FOR SOLVENTS

Solvent	*Uniqueness*		*Additional values*
	Letter	*Value*	
Apiezon L	A	0.32	P, 0.20; R, 0.16; S, 0.21; T, 0.17
Carbowax 600	B	0.26	C, 0.22; M, 0.16; o, 0.15
Carbowax 20M	C	0.29	B and o, 0.22; M, 0.20
Dibutyl tetrachlorophthalate	D	0.10	
Di-2-ethylhexyl sebacate	E	0.12	H, 0.12; N and U, 0.13; X, 0.15
Diglycerol	F	0.93	X, 0.15
Diisodecyl phthalate	G	0.06	
Dioctyl sebacate	H	0.13	N and U, 0.13; X, 0.16
Dow Corning 550 fluid	I	0.13	A, 0.15; J, 0.13; P, 0.20
Dow Corning FS 1265 fluid	J	0.31	L, 0.26; Q, 0.18
Flexol 8N8	K	0.18	P, −0.20; X, 0.13
Hyprose SP-80	L	0.47	J, 0.26; Q, 0.30
IGEPAL CO 880	M	0.15	B, 0.16; C, 0.20; O, 0.17
Isooctyldecyl adipate	N	0.13	E and H and N and U, 0.13; X, 0.16
Pluronic F77	O	0.19	B, 0.15; C, 0.22; M, 0.17
Polyphenyl ether—5 rings	P	0.56	A and I, 0.20; K, −0.20; Q, −0.22
Quadrol	Q	0.44	J, 0.17; L, 0.30; P, −0.22
SE-30	R	0.13	A, 0.16; S, 0.17; T, 0.13
SE-31	S	0.23	A, 0.21; R, 0.17; T, 0.16
SE-52	T	0.13	A, 0.17; R, 0.13; S, 0.16
Sucrose acetate isobutyrate	U	0.17	B and E and H and N, 0.13; X, 0.14
TMP Tripelargonate	V	0.09	E and N and X, 0.10; U, 0.11
Tricresyl phosphate	W	0.05	
Ucon LB-1715	X	0.10	C, 0.10; O, 0.11; Q, 0.12
Zonyl E-7	Y	0.30	E and F, 0.15; H and N, 0.16; U, 0.14

Tests of physically and chemically significant parameters

Before evaluating the fit for a given test factor, one should take into consideration the results of the uniqueness tests, because a high uniqueness value could be caused by errors in the data. If this is the case, then forcing a test vector through this solute or solvent point could cause the entire test vector to be in error. If true uniqueness exists, one must also be cautious in assigning values for that point in any test vector. With the factor analysis scheme in its most sophisticated form, the values for all the points on a suspected vector are not required to test a property by the least square part of the computer program[3]. If all defined points on the test vector rotate into nearly equivalent values on an abstract eigenvector, then it is legitimate to assume that the rotation is successful. If the rotation is acceptable, then values will be generated for those points left "free floating" on the test vector. The capability for "free floating" some test points has proven quite helpful in our previous testing of solute and solvent properties in gas chromatography[8,9]. The results of testing various chemical and physical properties of the alcohols for possible identification with the abstract eigenvectors of the space are given below.

Unity. To test if there is a common factor present in all of the solutes, a test vector composed entirely of ones is created. The results of this test are shown in Table III. The agreement is quite excellent, indicating that for an all-alcohol problem, all solutes share a common factor or factors (as expected for an all-alcohol problem).

Carbon number. The basis of the retention index system is the concept of a carbon number. For the *n*-alkanes, the retention index is defined as one hundred times the number of carbon atoms in the molecule, so that the difference in retention index for each consecutive member of the homologous series is set equal to one hundred retention index units. For any series of compounds containing a common functional group, the retention indices can also be reported relative to the normal straight chain members of that series. One major question amenable to attack by factor analysis is the effects on the carbon number of moving the functional group from the primary position on the solute chain. Since secondary retention indices have been shown to yield consistent results previously, we can define a carbon number test vector such that all *n*-alcohols are assigned a carbon number equal to the number of carbon atoms present in the alcohol, and all other points are "free floated".

The results of this test are shown in Table III. The agreement between the suspected values and the rotated values is excellent. The other predicted values on this vector should then indicate the effects of isomerism on carbon number. These effects seem to depend systematically upon the structure of the alcohols. From a study of the predicted carbon numbers in Table III, the following effects of isomerism on the carbon number can be estimated:

(i) A hydroxyl group in a secondary position causes approximately 0.2 carbon number to be subtracted from the carbon number of the corresponding primary alcohol.

(ii) A hydroxyl group in a tertiary position causes approximately 0.4 carbon number to be subtracted from the carbon number of the corresponding primary alcohol.

(iii) Cyclizing an alcohol causes approximately 0.7 carbon number to be subtracted from the corresponding straight chain alcohol value. We make the following

TABLE III

TESTS FOR CHEMICAL AND PHYSICAL PARAMETERS OF SOLUTES[a]

Values in parenthesis "free floated" in test vectors.

Solute	*Unity*		*Carbon number*		*Substituent effect*		*Molar heat of vaporization*[a]	
	Test	*Predicted*	*Test*	*Predicted*	*Test*	*Predicted*	*Test*	*Predicted*
Methanol	1	0.98	1	1.00	(0)	0.01	9.11	9.03
Ethanol	1	1.01	2	2.00	1	1.00	9.41	9.55
Propanol	1	0.99	3	2.99	1	0.99	9.88	9.71
Isopropanol	1	1.02		2.84	(2)	1.90	9.58	9.53
Butanol	1	0.99	4	4.01	1	1.02	10.47	10.42
Isobutanol	1	0.94		4.06		0.97	10.24	10.14
sec.-Butanol	1	1.00		3.83	(2)	2.03	9.93	9.59
tert.-Butanol	1	1.05		3.70	3	3.00	9.67	9.84
Pentanol	1	1.00	5	4.99	1	1.00	10.60	11.03
2-Pentanol	1	0.99		4.74	(2)	1.96		9.81
3-Pentanol	1	0.97		4.80	(2)	2.03		9.04
2-Methyl-2-butanol	1	1.03		4.64	3	3.00	9.33	9.41
Hexanol	1	1.00	6	6.00	1	0.98		11.44
2-Hexanol	1	1.01		5.80	(2)	2.08		10.52
3-Hexanol	1	0.98		5.82	(2)	2.15		9.62
2-Methyl-2-pentanol	1	1.01		5.53	3	2.99		9.80
3-Methyl-3-pentanol	1	1.00		5.59	3	2.98		9.17
2,3-Dimethyl-2-butanol	1	1.00		5.49	3	3.03		9.30
3-Heptanol	1	1.00		6.89	(2)	2.26		10.39
4-Heptanol	1	0.99		6.81	(2)	2.34		10.11
Octanol	1	1.03	8	8.00	1	1.01	12.70	12.56
Cyclopentanol	1	0.99		4.26		0.67		9.77
Cyclohexanol	1	1.04		5.26		1.15	10.84	10.88
2-Propen-1-ol	1	0.99		2.71		0.16		9.23
2-Propyn-1-ol	1	1.04		2.57		−2.06	9.14	9.19
1-Penten-3-ol	1	0.95		4.42		1.13		8.95

very tentative statements concerning the effects of unsaturation on the carbon number.

(iv) A double bond (based on the predicted value for 2-propen-1-ol) reduces approximately 0.3 carbon number from the *n*-alcohol value.

(v) A triple bond (based on the predicted value for 2-propyn-1-ol) reduces approximately 0.4 carbon number from the *n*-alcohol value. The effects of isomerism and unsaturation may be additive since the carbon number predicted for 1-penten-3-ol (4.42) approximates the carbon number calculated using effects (i) and (iv) (*i.e.*, 5.0 — 0.2 — 0.3).

Substituent effects. A test vector can be defined which relates one of the solute factors of the space to the number and type of substituents on the carbon atom containing the hydroxyl group. As a first approximation to this vector, we specify the number of carbon atoms directly attached to the carbon containing the hydroxyl group. On this substituent test vector, the primary alcohols are assigned a value of one and the tertiary carbons a value of three. As a check of the validity of the test vector, methanol and all secondary alcohols are "free floated". If the test vector is

Vapor pressure[a]		*ln vapor pressure*		*Molar refraction*[a]		*Molecular weight*		*Boiling point*[b]	
Test	*Predicted*	*Test*	*Predicted*	*Test*	*Predicted*	*Test*	*Predicted*	*Test*	*Predicted*
5090	5006	8.535	8.312	8.2	7.9	32.0	31.0	337.9	333.2
3320	3434	8.107	8.365	12.8	13.1	46.1	47.0	351.6	353.5
1730	1794	7.455	7.647	17.4	17.9	60.1	61.6	370.4	368.9
3000	2835	8.006	8.142	17.5	17.5	60.1	60.0	355.6	358.8
853	1142	6.748	6.823	22.1	22.3	74.1	75.3	391.2	391.4
1070	663	6.975	6.651	22.2	22.2	74.1	73.4	381.3	373.8
1660	1387	7.414	7.400	22.2	22.4	74.1	75.3	372.7	374.0
2990	2895	8.003	7.945	22.2	21.9	74.1	72.6	355.8	367.3
400	829	5.991	5.934	26.8	26.7	88.2	88.5	411.3	412.2
760	866	6.633	6.639	26.8	26.6	88.2	87.8	393.1	389.1
	4243		6.885	26.8	27.4	88.2	90.0	389.3	382.0
	1514		7.397	26.8	26.9	88.2	88.5	375.6	381.6
243	523	5.493	5.338	31.5	31.4	102.2	102.3	430.7	433.1
400	644	5.991	6.042	(31.5)	31.4	(102.2)	101.9	413.0	413.3
460	461	6.131	6.101	31.5	31.9	102.2	103.4	408.0	401.7
760	858	6.633	6.527	31.5	30.9	102.2	100.3	396.0	395.4
	935		6.414	31.5	31.6	102.2	102.8	395.7	393.1
	1149		6.435	31.5	31.0	102.2	100.7	394.0	389.7
	317		5.523	36.1	36.7	116.2	117.4	427.0	426.7
	233		5.448	36.1	36.5	116.2	117.1	429.0	321.4
54	−354	3.996	4.216	40.7	40.4	130.2	129.2	467.7	478.2
	−324		0.206	24.7	24.8	86.1	85.6	410.0	409.7
230	35	5.438	5.407	29.4	29.2	100.2	99.8	434.3	440.6
	377		8.931	17.0	16.9	(58.1)	58.5	—	—
	−1341		11.885	(15.4)	16.2	(56.1)	54.7	—	—
	−91		7.338	26.3	25.3	86.1	83.4	—	—

[a] Values of molar heat of vaporization (kcal/mole), vapor pressure (Torr), molar refraction (cc/mole) and boiling point (°K) were taken from ref. 12.

[b] The boiling point test was only tried on the saturated alcohol data.

properly defined, then it is anticipated that methanol will have a value close to zero and the secondary alcohols will have values near two on the scale. The results for this test are shown in Table III. The fit between the test and predicted points is excellent, and the predicted values for methanol and the secondary alcohols are as expected. However, the cyclic and unsaturated alcohols have predicted values smaller than the number of substituent carbons attached to the hydroxyl bearing carbon atom.

The substituent-effect test may be interpreted physically in terms of an induction effect. It is generally believed that methyl groups are electron donors and that unsaturated carbons are electron withdrawing centers. If the substituent effect test factor measures electron density on the carbon atom bearing the hydroxyl group, and, if the inductive effects are additive, then, as a first approximation, the test values one, two and three are justified for the primary, secondary and tertiary alcohols, respectively. The unsaturated alcohols, having electron withdrawing groups, should experience a negative inductive effect which would decrease the electron density at the carbon bearing the hydroxyl group. While the substituent test vector may also be affected to some extent by steric factors, the inductive effects appear to be the major factor influencing the vector. With this explanation, it is harder to understand the small predicted values for the cycloalcohols, which should have values near two. Our results for the two cycloalcohols imply that the cyclo group is behaving as an electron withdrawing group.

Heat of vaporization, vapor pressure and boiling point. The molar heat of vaporization appears to be a true factor of the retention index space (see Table III). Two functional forms of the solute vapor pressure were tested for identification with the abstract eigenvectors of the space: the solute vapor pressure and the natural logarithm (ln) of the vapor pressure. The results of both tests are shown in Table III. Of the two tests, ln vapor pressure gives a much better fit. Although a test of the reciprocal of the normal boiling point gave a moderately good reproduction of test points, the fit for the normal boiling-point test vector was also fair as shown in Table III.

Molar polarization. The fit for solute molar polarization is also shown in Table III. A few points were left unassigned in order to check the validity of the rotation. The agreement is excellent. This factor also tested well[9] on the more general data of ROHRSCHNEIDER[10]. Tests of this type illustrate the power of factor analysis for predicting the mathematical form of unknown interaction terms. KIESLEV *et al.*[13] recently postulated that the molar polarization is directly related to the heat of adsorption of a solute onto a surface.

Dipole moment. Test vectors for both dipole moment and dipole moment squared exhibited the correct trends, but several points in both vectors were in error by over 20%. This implies that the solute dipole moment occurs in an interaction term (which we recall must be written as a product of a solute term and a solvent term) with a power near one.

Dielectric constant. A dielectric constant test vector gave a moderately good fit, whereas a test of the reciprocal of the dielectric constant $1/\varepsilon$ exhibited a very poor reproduction. The results were somewhat unexpected since ANVAER *et al.*[14] found that $1/\varepsilon$ was linear with log (retention volume) for a series of solvents.

Molecular weight. The excellent fit for the solute molecular weight test vector is surprisingly good. The fit for this test vector is shown in Table III. As a check on the vector, several of the solute points were omitted (points shown in parenthesis in Table III). The close agreement between the experimental and predicted values shows that, for this data matrix, molecular weight can be associated with a true factor of the space. The molecular weight test factor did not work well when applied to the retention index data of ROHRSCHNEIDER[10] who measured the retention index of many different types of solutes[9]. These results indicate that one should be cautious in drawing general conclusions from a fit determined on one family of solutes.

Reproduction of the data matrix

If it is possible to identify each of the abstract eigenvectors of the space with a separate physically or chemically significant parameter, then these factors can be used to reproduce the data matrix. This last step of reproduction is quite important since, only if all factors, are accounted for in the test factors, will the data matrix be reproduced within experimental error. If all test factors fit well, but the data matrix cannot be reproduced from them, then at least two of the test factors are being rotated into the same eigenvector and the factor space is not being spanned.

Attempts to find five factors which span the space of the saturated alcohols met with considerable success. The following groups of five test vectors were tested for reproduction of the data matrix:

(A) 1, unity; 2, boiling point (°K); 3, molecular weight; 4, carbon number and 5, molar refraction.

(B) 1, unity; 2, boiling point (°K); 3, substituent effect; 4, carbon number and 5, molar refraction.

(C) 1, unity; 2, ln vapor pressure; 3, substituent effect; 4, carbon number and 5, molar refraction.

(D) 1, unity; 2, heat of vaporization; 3, substituent effect; 4, carbon number and 5, molar refraction.

(E) 1, unity; 2, molecular weight; 3, molar refraction; 4, carbon number and 5, substituent effect.

(F) 1, molecular weight; 2, molar refraction; 3, carbon number; 4, ln vapor pressure and 5, substituent effect.

TABLE IV

CONSTANTS NEEDED TO EVALUATE EQN. 1

Solvent	*Letter*	*Constants*[a]				
		C (1,a)	*C (2,a)*	*C (3,a)*	*C (4,a)*	*C (5,a)*
Apiezon L	A	−1262	4.338	4.273	−76.660	8.008
Carbowax 600	B	−1566	7.379	11.060	19.580	−45.890
Carbowax 20M	C	−1836	8.136	−0.008	−42.130	−2.473
Dibutyl tetrachlorophthalate	D	−1217	4.822	1.765	−50.070	7.693
Di-2-ethylhexyl sebacate	E	−734	3.045	17.700	61.770	−56.960
Diglycerol	F	−1855	10.270	−21.450	−295.100	93.360
Diisodecyl phthalate	G	−1056	4.239	8.261	0.867	−20.160
Dioctyl sebacate	H	−725	3.029	17.410	63.460	−56.460
Dow Corning 550 fluid	I	−1010	3.774	15.210	16.990	−42.640
Dow Corning FS 1265 fluid	J	−969	3.926	15.230	19.740	−43.200
Flexol 8N8	K	−1262	5.327	−0.712	−29.500	9.011
Hyprose SP-80	L	−1557	7.731	13.500	20.280	−55.300
Igepal CO 880	M	−1690	7.463	−1.308	−42.140	3.985
Isooctyldecyl adipate	N	−715	2.999	18.340	67.580	−59.990
Pluronic F77	O	−1743	7.556	−1.648	−39.170	4.205
Polyphenyl ether—5 rings	P	−827	3.138	32.560	83.770	−105.900
Quadrol	Q	−1732	8.133	−1.429	−23.960	−0.984
SE-30	R	−1299	4.816	0.634	−51.770	11.570
SE-31	S	−1373	5.032	−4.416	−84.750	32.870
SE-52	T	−1257	4.613	2.450	−47.610	6.124
Sucrose acetate isobutyrate	U	−985	4.448	16.260	45.940	−54.710
TMP Tripelargonate	V	−848	3.421	17.010	47.500	−53.060
Tricresyl phosphate	W	−1340	5.659	6.596	−13.060	−17.970
Ucon LB-1715	X	−1275	5.384	8.499	35.400	−33.020
Zonyl E-7	Y	−1114	4.975	3.035	−68.740	7.309

[a] Numerals 1–5 represent unity, boiling point (°K), molecular weight, carbon number and molar refraction, respectively.

The average column average error between predicted and experimental retention indices using each set of five test factors were as follows: A, 8.4; B, 11.7; C, 29.0; D, 30.3; E, 18.3; F, 24.6 retention indices. Although the best fit, set A, is still not within experimental error, the major factors appear to be accounted for to a large degree by this choice.

It is now possible, by using set A, to establish a set of equations of the following form:

$$I(i, \alpha) = \sum_{j=1}^{5} C(j, \alpha)\, P(j, i) \tag{1}$$

where $I(i,\alpha)$ is the retention index of solute, i, on stationary phase, α; j corresponds to the five test factors in set A; $C(j,\alpha)$ corresponds to the five multiplicative constants associated with each solvent, α, which are calculated from the reproduction portion of the factor analysis program; and $P(j,i)$ corresponds to the five parameters in set A evaluated for solute, i. The constants $C(j,\alpha)$ for the twenty-five solvents listed in Table II are shown in Table IV.

The validity of Eqn. 1 can be tested by calculating the retention indices on all twenty-five stationary phases of several solutes not included in the original data matrix but for which experimental retention indices are known. In Table V are listed the properties, $P(j,i)$, for four new alcohols not included in the original scheme. Using Eqn. 1, and the data in Tables IV and V, we can predict the retention indices of these four alcohols on all twenty-five stationary phases. The results of this test are shown in Table VI. Considering the simplicity of the five test factors and the complexity of the problem, the agreement is quite remarkable.

TABLE V

PHYSICAL PROPERTIES, $P(j,i)$, FOR ALCOHOLS NOT INCLUDED IN THE ORIGINAL DATA MATRIX[a]

Alcohol	*Boiling point*	*Molecular weight*	*Carbon number*[b]	*Molar refraction*
2-Heptanol	434	116.2	6.8	36.1
2-Octanol	453	130.2	7.8	40.7
1-Nonanol	487	144.2	9.0	45.4
2-Methyl-cyclohexanol	439	114.2	6.3	34.0

[a] All data taken from ref. 12. Boiling point in °K, molecular weight in grams, molar refraction in cc/mole.

[b] Carbon number estimated using the rules given in this paper.

It might be possible with a different set of test factors to produce an even closer fit. Ideally, one should be able to come up with mathematical models for each of the five interaction parameters in terms of the properties of the solutes and stationary phases of the system. Until this stage is reached, however, empirical correlations such as those proposed here can be quite useful in aiding the chromatographer in choosing the best system for a given separation.

TABLE VI

EXPERIMENTAL AND PREDICTED RETENTION INDICES FOR FOUR ALCOHOLS NOT INCLUDED IN THE ORIGINAL DATA MATRIX

	2-Heptanol		*2-Octanol*		*n-Nonanol*		*2-Methylcyclohexanol*	
	Exptl.	*Predicted*	*Exptl.*	*Predicted*	*Exptl.*	*Predicted*	*Exptl.*	*Predicted*
Apiezon L	868	885	966	987	—	1138	946	920
Carbowax 600	1379	1398	1479	1502	—	1712	1477	1499
Carbowax 20M	—	1318	1399	1419	1648	1630	1402	1385
Dibutyl tetrachlorophthalate	999	1018	1101	1119	—	1282	—	1047
Di-2-ethylhexyl sebacate	990	1008	1090	1113	—	1269	1045	1076
Diglycerol	1457	1473	1506	1502	1693	1631	1589	1519
Diisodecyl phthalate	1003	1022	1105	1126	—	1290	1064	1068
Dioctyl sebacate	987	1006	1087	1110	—	1267	1042	1073
Dow Corning 550 fluid	951	972	1053	1077	—	1236	1025	1041
Dow Corning FS 1265 fluid	1061	1079	1164	1188	—	1354	1123	1149
Flexol 8N8	1072	1092	1175	1195	—	1370	1124	1116
Hyprose SP-80	1478	1508	1590	1610	—	1822	1587	1626
Igepal CO 880	1241	1254	1338	1354	—	1554	1323	1307
Isooctyldecyl adipate	994	1011	1094	1117	—	1273	1047	1082
Pluronic F77	1212	1230	1313	1331	—	1534	—	1282
Polyphenyl ether—5 rings	1042	1065	1147	1177	1341	1341	1140	1196
Quadrol	1413	1433	1528	1539	—	1758	1479	1491
SE-30	912	930	1012	1032	—	1195	969	955
SE-31	891	908	997	1008	—	1168	965	915
SE-52	909	927	1008	1029	—	1190	969	956
Sucrose acetate isobutyrate	1154	1172	1258	1278	1467	1453	1222	1252
TMP Tripelargonate	1003	1021	1104	1127	—	1288	1060	1091
Tricresyl phosphate	1125	1145	1229	1249	—	1431	—	1204
Ucon LB-1715	1080	1098	1181	1202	1397	1389	1142	1159
Zonyl E-7	1183	1197	1288	1296	—	1457	—	1232

ACKNOWLEDGEMENTS

The authors would like to thank Dr. J. F. PARCHER of the University of Mississippi for his helpful assistance. We would also like to thank the computer centers at Brooklyn College and the University of Mississippi for the use of their facilities.

This work was supported by a grant from the City University of New York Faculty Research Program.

REFERENCES

1 J. T. TAKÁCS, C. SZITA AND G. TARJAN, *J. Chromatogr.*, 56 (1971) 1.
2 G. CASTELLO AND G. D'AMATO, *J. Chromatogr.*, 58 (1971) 127.
3 P. H. WEINER, E. R. MALINOWSKI AND A. R. LEVINSTONE, *J. Phys. Chem.*, 74 (1970) 4537.
4 P. H. WEINER AND E. R. MALINOWSKI, *J. Phys. Chem.*, 75 (1971) 1207.
5 P. H. WEINER AND E. R. MALINOWSKI, *J. Phys. Chem.*, 75 (1971) 3160.
6 P. H. WEINER, *Ph. D. Thesis*, Stevens Institute of Technology, Hoboken, N. J., 1971; *Diss. Abstr.*, 32 (6) - B3290 No. 72-1634.
7 P. T. FUNKE, E. R. MALINOWSKI, D. E. MARTIRE AND L. Z. POLLARA, *Separ. Sci.*, 1 (1967) 661.
8 P. H. WEINER AND D. G. HOWERY, *Can. J. Chem.*, 50 (1972) 448.
9 P. H. WEINER AND D. G. HOWERY, *Anal. Chem.*, 44 (1972) 1189.

10 L. Rohrschneider, *J. Chromatogr.*, 22 (1966) 6.
11 W. O. McReynolds, *Gas Chromatographic Retention Data*, Preston Technical Abstracts Co., Evanston, Ill., 1966.
12 R. C. Weast (Editor), *Handbook of Chemistry and Physics*, 51st Ed., The Chemical Rubber Co., Cleveland, Ohio, 1970.
13 A. V. Kieslev, N. V. Kovaleva and Y. U. S. Nikitin, *J. Chromatogr.*, 58 (1971) 19.
14 B. I. Anvaer, A. A. Zhuhkouitskii, I. I. Litoutseva, V. M. Sakharor and N. M. Turkel'-taub, *J. Anal. Chem. USSR*, 19 (1964) 162.