Journal of Chromatography, 352 (1986) 141–155
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 770

QUANTITATIVE RETENTION RELATIONSHIPS AS A FUNCTION OF MOBILE AND C_{18} STATIONARY PHASE COMPOSITION FOR NON-COGENERIC SOLUTES

ROMAN KALISZAN* and KRZYSZTOF OSMIALOWSKI
Medical Academy, K. Marksa 107, 80-416 Gdansk (Poland)
and
STERLING A. TOMELLINI, SHIH-HSIEN HSU, STEVEN D. FAZIO and RICHARD A. HARTWICK*
Department of Chemistry, Rutgers University, Piscataway, NJ 08854 (U.S.A.)

SUMMARY

Twelve non-homologous solutes with a wide range of functional groups were chromatographed on a series of well-characterized stationary phases with a range of methanol–water mobile phases. Retention correlations were found for quantum mechanically calculated molecular parameters. These parameters, the total molecular energy, and the maximal excess electronic charge difference, were found to be better descriptors of retention than the dipole moment or fragmental constants. The implication of these results in terms of retention mechanisms on reversed-phase materials are discussed.

INTRODUCTION

At constant temperature, three main variables determine the distribution of a solute between mobile and stationary phases in high-performance liquid chromatography (HPLC): the chemical structure of the solute, the physicochemical properties of the mobile phase, and the physicochemical properties of the stationary phase. The solute distribution in HPLC is easily quantified by means of retention parameters, usually the capacity factor, k'. If one can numerically measure the properties of the solutes, of the mobile phases, and of the stationary phases, one can attempt to derive general relationships linking the four quantities together. Such a relationship, if found, would be of importance for our understanding of the physicochemical interactions involved in the chromatographic separation process.

If the quantitative relationships between the retention data (as dependent variables) and the numerically measured properties of the solutes, of the mobile phases, and of the stationary phases (as independent variables) are of adequate precision, they may be used to predict the retention of a given solute with a given mobile and stationary phase composition. The relationships could also be used for computer simulation of HPLC separations.

0021-9673/86/$03.50

Attempts to relate quantitatively retention data determined for a set of solutes under constant (or normalized) chromatographic conditions to the solute structure have a long history[1,2]. Based on the extrathermodynamic assumption of linear free-energy relationships, correlation analyses have been performed using various retention data in relation to the other experimental free-energy related parameters or substituent contributions[2–6]. The applications of topological indices for the description of gas–liquid chromatographic (GLC) retention indices have been especially successful[7,8].

The majority of reports published to date concerning quantitative structure–retention relationships (QSRR)[4] deal with either homologous or more or less cogeneric groups of compounds, chromatographed under fixed separation conditions. Correlations reported for several sets of solutes between the GLC Kováts retention indices and various molecular descriptors are often very high, especially if the stationary phase is non-polar and the solutes are closely related structurally[7,9].

In the case of liquid chromatography the best relationships have been observed between retention and data determined by the classical shake-flask method[2] or calculated by means of various substituent or fragmental hydrophobic constants[10]. Jinno and Kawasaki[11] described the logarithm of the capacity factors of two separate sets of benzene derivatives, chromatographed on a C_{18} column, in terms of a hydrophobic substituent constant of the solutes and as a function of the volume fraction of the organic modifier in the aqueous eluent mixture. The reported function, derived statistically, is rather complex and its physical meaning is difficult to interpret.

Among members of a homologous series the physicochemical properties that determine chromatographic behavior change in a parallel manner relative to one another. From the existing theories of chromatography[12–14] it is possible to generalize that in reversed-phase HPLC, two types of forces dominate the interactions between solute molecules and the molecules of each phase: the polar forces, arising from permanent or induced electric fields associated with both solute and phase molecules and the non-polar, non-specific forces originating from dispersive interactions. The ability of homologs to participate in polar interactions remains constant to a first approximation. Such changes in the ability of a solute to undergo non-polar, dispersive interactions are most readily parameterized by common measures of solute "bulk" properties, like Van der Waals volumes, surface area, molar refractivity, connectivity indicies or simply the number of carbon atoms.

On the other hand, quantitation of molecular polarity is not as feasible. QSRR concerning more diverse, non-cogeneric sets of solutes are rare and of limited value[15]. The good correlations often reported between reversed-phase liquid chromatographic retention data and hydrophobic substituent constants for various sets of solutes are due to the inclusion of polarity information within the empirically determined substituent constants. Unfortunately, such correlations are of rather limited value for the understanding of chromatographic phenomena, as the interpretation of one empirical quantity in terms of other experimentally derived quantities contributes little fundamental chemical information.

Recently, quantum chemical calculations have been used for the quantitative description of molecular characteristics of solutes. Correlation analyses of routinely calculated CNDO/2 quantum chemical parameters and GLC retention indices have been reported[9,15–17]. As expected, in the case of structurally similar homologues the

correlations obtained are very high[9,16], but significantly lower for a more diverse set of solutes[15]. As with other structural parameters mentioned earlier, the quantum chemical data usually change smoothly over a homologous series, thus allowing a single parameter, *e.g.* total energy, to adequately reflect the overall chemical potential change within the series. However, when structurally diverse sets of solutes are considered, the retention changes can no longer be described by the changes in any single molecular descriptor. To relate retention and molecular structure, more than one independent structural variable must be considered (assuming constant separation conditions and, therefore, that observed retention changes are the result of solute structure differences). Such multiparametric regression analyses have been performed for a diverse set of phenolic derivatives[15], resulting in a two-parameter regression equation relating the Kováts indices to both molecular refractivity and the CNDO/2 calculated dipole moment of the solutes.

In the case of liquid chromatography quite a number of correlations have been reported between retention data and substituent or fragmental empirical constants[2,18,19]. No quantitative relationships have been found however, describing capacity factors in terms of parameters directly related to the more fundamental abilities of solutes to participate in polar and dispersive interactions with both phases. The purposes of the present study was to investigate the relationships existing between quantum mechanically calculated parameters and mobile and stationary phase properties. A limited model involving twelve solutes with a range of functional groups, and a range of well characterized C_{18} stationary phases using methanol–water mobile phases was studied.

THEORY

Over range of compositions for a two-component mobile phase, the logarithm of the capacity factor, $\log k'_{ij}$, for a particular solute i, determined on phase j, depends linearly on the mole fraction, X, of one of the solvent components, as described by the Soczewiński–Wachtmeister relationship[20]

$$\log k'_{ij} = a_{ij}X + b_{ij} \tag{1}$$

where a_{ij} and b_{ij} are constants characteristic for a given solute chromatographed on a specific stationary phase. In the reversed-phase mode, X usually denotes the mole fraction of water in aqueous organic solvent. If the linearity predicted by eqn. 1 is actually observed for the series of solutes analysed with a given stationary phase, then one may assume that the constants a_i and b_i are functions of the solute molecular structure.

Assuming linear free-energy relationships, the molecular properties of the solutes can be expressed as a linear combination of individual structural parameters. Using multiparameter regression analysis, it was found that the constants a_i and b_i can be satisfactorily described by a statistical two-parameter equation involving the quantum chemically calculated total energy of the solute, E_{T_i}, and its polarity parameter, Δ_i, proposed by us (see Results).

Thus, the constants a_{ij} and b_{ij} determined for a particular solute, i, at a specific stationary phase, j, would be:

$$a_{ij} = \alpha_j E_{T_i} + \beta_j \Delta_i + \gamma_i \tag{2}$$

$$b_{ij} = \alpha'_j E_{T_i} + \beta'_j \Delta_i + \gamma'_j \tag{3}$$

where α_j, β_j, γ_j, α'_j, β'_j, γ'_j are regression coefficients, derived using the conventional least-squares method. Having the a_{ij} and b_{ij} data for i compounds determined on j phases from $i \times j$ regression equations in the form of eqn. 1, one can attempt to describe them in terms of E_{T_i} and Δ_i of i-th compound. If eqns. 2 and 3 are statistically significant, then eqn. 1 for the j-th phase may be rewritten as follows:

$$\log k'_i = a_j X + b_j \tag{4}$$

or

$$\log k'_i = (\alpha_j E_{T_i} + \beta_j \Delta_i + \gamma_j) X + (\alpha'_j E_{T_i} + \beta'_j \Delta_i + \gamma'_j) \tag{5}$$

The phases under study differ in C_{18} coverage. For a given solute, i, chromatographed with a fixed mobile-phase composition, X, linearity has been found between the logarithm of the capacity factor, $\log k'_{i(X),j}$ and the C_{18} stationary phase coverage, C_j, for three of the four phases studied:

$$\log k'_{i(X),j} = AC_j + B \tag{6}$$

In such a situation it seems probable that the coefficients a_{ij} and b_{ij} of eqn. 1 or the coefficients α_j, β_j, γ_j, α'_j, β'_j, and γ'_j of eqns. 2 and 3 depend not only on solute structure, as described by the individual E_{T_i} and Δ_i descriptors but also on the stationary phase properties, expressed by the C_{18} stationary phase coverage, C_j. Thus, the most general relation describing capacity factors in terms of solute structure, mobile phase composition and stationary phase surface properties is

$$\log k'_{ij} = (\alpha_j E_{T_i} + \beta_j \Delta_i + \gamma C_j + \delta)X + (\alpha'_j E_{T_i} + \beta'_j \Delta_i + \gamma' C_j + \delta') \tag{7}$$

To derive the regression coefficients α, β, γ, δ, α', β', γ' and δ' the variable matrices of $(i \cdot j) \cdot 4$ dimension are considered:

$$\begin{array}{cccc}
a_{11} & E_{T_1} & \Delta_1 & C_1 \\
a_{21} & E_{T_2} & \Delta_2 & C_1 \\
a_{31} & E_{T_3} & \Delta_3 & C_1 \\
\vdots & \vdots & \vdots & \vdots \\
a_{i1} & E_{T_i} & \Delta_i & C_1 \\
a_{12} & E_{T_1} & \Delta_1 & C_2 \\
a_{22} & E_{T_2} & \Delta_2 & C_2 \\
a_{32} & E_{T_3} & \Delta_3 & C_2 \\
\vdots & \vdots & \vdots & \vdots \\
a_{i2} & E_{T_i} & \Delta_i & C_2 \\
\vdots & \vdots & \vdots & \vdots \\
a_{1j} & E_{T_1} & \Delta_1 & C_j \\
\vdots & \vdots & \vdots & \vdots \\
a_{ij} & E_{T_i} & \Delta_i & C_j
\end{array} \tag{8}$$

and analogously for b_{ij} coefficients.

EXPERIMENTAL

Stationary phase synthesis

All bonded stationary phases were prepared using Whatman Partisil 10 support material (Whatman, Clifton, NJ, U.S.A., batch number 100591). Appropriate quantities of dimethyloctadecylchlorosilane (Petrarch Chemical, Bristol, PA, U.S.A.) were refluxed with *ca.* 5 g of dry silica gel for 12 h in freshly distilled toluene, with 1 ml of pyridine. Bonded phases were extensively washed with toluene, methanol and water. No trimethylmonochlorosilane (TMS) or "capping" agent was added and therefore a monofunctional C_{18} stationary phase was produced.

The parent silica gel support material was analysed by the BET method for surface area and pore volume. BET surface area using nitrogen was 404 m^2 g^{-1}. Nitrogen pore volume was 0.72 cm^3 g^{-1}. The average pore diameter, assuming a cylindrical model, was 68 Å.

The bonded stationary phases were analyzed using a hydrofluoric acid digestion followed by gas chromatographic analysis on an SE 30 capillary column[26]. Analyses were performed in triplicate, and were confirmed using elemental analyses. Stationary phases were packed into 250 × 1 mm ("microbore") glass-lined columns, using previously published procedures[27].

Instrumentation

The chromatographic system consisted of a Brownlee MPLC microbore gradient pump (Brownlee Labs., Santa Clara, CA, U.S.A.), with a nominal 150-μl packed mixing chamber, a Rheodyne (Rheodyne, Cotati, CA, U.S.A.) Model 7413 variable sample loop injection valve set at 0.5 μl, and a Kratos (Kratos Analytical Instruments, Ramsey, NJ, U.S.A.) Model SF 769 UV detector with a 0.5-μl flow cell which was operated at 254 nm. The analog output of the detector was recorded using a strip chart recorder (Kipp en Zonen, BD-40 series), and was simultaneously passed into an Apple II+ computer, equipped with a 12 bit A/D converter, and with a integration/control software package (Dynamic Solutions, Ventura, CA, U.S.A.).

Chromatography of test solutes

A selected group of aromatic test solutes containing substituents of various polarities were chromatographed using a range of methanol–water mobile phases on all stationary phases. HPLC grade methanol (Baker, Phillipsburgh, NJ, U.S.A.) was used to prepare the mobile phases. All mobile phases were prepared volumetrically using double-distilled, deionized water and were sparged with helium before use. Solvent proportioning in all cases was made by the Brownlee MPLC pump. The flow-rate throughout the entire study was kept constant at 50.0 μl min^{-1}.

RESULTS

Structural analysis

Molecular parameters of orbitals were calculated with complete neglect of differential overlap, CNDO/2[21,22]. Standard bond lengths and angles were assumed. The calculations were performed on conformations for which the steric interactions

between parts of the molecule were minimized. Among the variety of quantum chemical indicies, the total energy (E_T), dipole moment (μ), and polarity parameter (Δ) were found to be meaningful for calculating chromatographic capacity factors by correlation analysis. The numerical data are listed in Table II.

Keeping in mind the observations by Karger *et al.*[13] and Scott[12] that compounds such as dioxane or 1,4-dichlorobutane, have total dipole moments near zero (the two dipoles in opposition cancel) but behave as polar solutes, we attempted to study a submolecular measure of polarity reflecting the largest molecular local dipole. The quantum chemical polarity parameter Δ, as proposed by us, is the largest difference in individual atomic electronic excess charge in a given molecule. To determine Δ, first the electron densities on particular atoms in the molecule are calculated and then the atom with the highest electron excess and the atom with the highest electron deficiency are located. The difference between the two is calculated as Δ (see Fig. 1 for illustration).

Analysis of retention data

Two of the fourteen benzene derivatives chromatographed under varying conditions were excluded from correlation analyses after preliminary attempts. These two were toluene and chlorotoluene, highly hydrophobic non-polar solutes, for which the mechanism of chromatographic distribution seemed to be different from that of

TABLE I

CHROMATOGRAPHIC RETENTION DATA OBTAINED ON COLUMN OF VARYING C_{18} COVERAGE

Other conditions as listed in the Experimental section.

No.	*Compound*	*Retention time (min)*							
		Coverage (mol/g)							
		1.54 · 10⁻⁴				*3.18 · 10⁻⁴*			
		Water–methanol				*Water–methanol*			
		65:35	*55:45*	*45:55*	*35:65*	*65:35*	*55:45*	*45:55*	*35:65*
1	Phenol	7.0	5.7	4.8	4.2	8.6	6.2	4.8	4.0
2	Acetophenone	12.7	8.3	6.2	4.9	17.5	10.2	6.7	5.0
3	Nitrobenzene	13.4	9.3	6.8	5.3	18.2	11.4	7.5	5.5
4	Methylbenzoate	20.9	12.0	7.8	5.6	33.4	16.6	9.3	6.1
5	*p*-Cresol	10.6	7.5	5.7	4.6	14.9	9.2	6.2	4.8
6	*p*-Ethylphenol	17.5	10.6	7.1	5.3	28.3	14.2	8.6	5.6
7	*p*-Propylphenol	34.6	17.5	9.7	6.2	63.4	26.6	13.3	7.4
8	4-*sec*-Butylphenol	57.0	25.4	12.4	7.2	113.3	42.5	18.5	9.1
9	Aniline	–	6.9	5.7	4.7	8.6	6.5	5.1	4.3
10	N-Methylaniline	–	9.9	7.2	5.4	16.5	10.4	7.2	5.3
11	4-Chloroacetophenone	–	14.5	8.9	6.0	44.8	20.7	11.1	6.8
12	3,4-Dichloroaceto-phenone	–	27.7	14.1	8.1	112.8	44.2	19.9	10.3
	Uracil	3.8	3.7	3.6	3.6	3.4	3.2	3.1	3.0

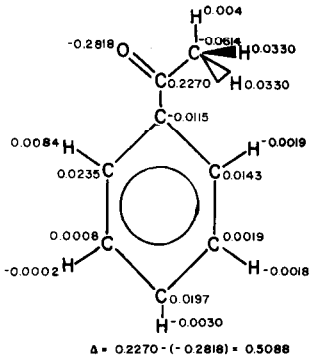

Fig. 1. Excess electronic charge distribution in acetophenone, and calculation of the polarity parameter, Δ.

the remaining polar compounds. The experimental retention data are shown in Table I.

Each of the compounds under study was chromatographed at four methanol–water mobile-phase compositions on four stationary phases with different dimethyloctadecylsilane (C_{18}) coverage. Thus, a total of 191 experimental retention

4.96 · 10⁻⁴					*6.6 · 10⁻⁴*			
Water–methanol					*Water–methanol*			
65:35	*55:45*	*45:55*	*40:60*	*35:65*	*55:45*	*50:50*	*45:55*	*35:65*
13.3	8.9	6.4	5.6	5.0	10.0	8.1	6.8	5.3
30.8	16.3	9.7	7.8	6.5	18.7	14.4	11.1	7.3
32.6	18.9	11.4	9.1	7.5	23.0	17.3	13.4	8.4
65.1	30.3	15.4	11.4	8.9	40.3	27.2	19.4	10.5
–	14.8	9.6	7.2	6.1	17.4	13.1	10.1	6.8
–	26.8	13.6	10.1	8.0	33.9	23.3	16.6	9.4
–	56.4	23.6	15.8	11.5	76.8	48.1	31.0	14.3
–	97.0	35.2	22.3	14.9	139.3	81.6	49.0	19.9
–	9.0	6.6	5.8	5.2	9.5	7.8	7.0	5.5
–	16.8	10.6	8.5	7.2	20.0	14.9	12.0	7.9
–	38.6	18.5	13.2	10.1	51.1	32.9	23.1	11.9
–	91.0	37.0	24.1	16.8	130.0	77.4	49.5	21.3
3.7	3.5	3.4	3.4	3.4	3.5	3.4	3.5	3.3

TABLE II

COEFFICIENTS a AND b OF THE EQUATION $\log k = ax + b$ AND CNDO/2 MO PARAMETERS USED IN CORRELATION ANALYSIS

Solute	*Coverage (mol/g)*								*Total energy (a.u.)*	*Maximum excess charge difference, Δ*	*Dipole moment (debye)*
	$1.54 \cdot 10^{-4}$		$3.18 \cdot 10^{-4}$		$4.96 \cdot 10^{-4}$		$6.6 \cdot 10^{-4}$				
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*			
1	−2.6959	0.4536	−2.5581	0.6639	−2.8606	0.9593	−2.6078	0.9413	−65.5548	0.4328	1.7492
2	−3.0864	0.9465	−3.0568	1.1786	−3.4881	1.5156	−2.5891	1.2605	−81.1756	0.5088	3.0417
3	−2.7960	0.9358	−2.7800	1.1596	−3.5946	1.6078	−2.9937	1.5324	−94.8446	0.7774	5.0589
4	−3.4745	1.3031	−3.5860	1.6051	−3.9130	1.9501	−3.6501	1.9751	−99.6184	0.6836	2.0376
5	−3.0970	0.8482	−2.9067	1.0646	−3.3485	1.4121	−3.0797	1.4014	−74.2375	0.4265	1.7379
6	−3.3854	1.1933	−3.5349	1.5152	−3.7593	1.8154	−3.5988	1.8782	−82.6639	0.4275	2.3324
7	−4.0343	1.6683	−4.1325	2.0059	−4.3633	2.3292	−4.2734	2.4389	−91.3479	0.4270	2.1472
8	−4.4055	1.9693	−4.6025	2.3567	−4.8630	2.7080	−4.7467	2.8292	−100.1758	0.4246	1.8457
9	−2.4448	0.6029	−2.1299	0.5849	−2.5553	0.8762	−2.1672	0.7902	−59.5473	0.3737	1.5206
10	−2.8347	0.9873	−2.6928	1.0879	−2.8933	1.3476	−2.8162	1.4047	−68.2305	0.3564	1.1504
11	−3.4547	1.3868	−3.8131	1.7990	−3.8413	2.0147	−3.8169	2.1259	−96.6241	0.5028	2.2894
12	−3.8477	1.8329	−4.3011	2.2945	−4.3579	2.5474	−4.3741	2.6970	−112.3142	0.4922	1.2809

times were involved in the regression analysis. The capacity factors, k', were calculated from the relationship, $k' = (t_R - t_0)/t_0$, where t_R is the retention time of the solute and t_0 was determined for the unretained tracer uracil under the same conditions.

For each of the compounds chromatographed on all four phases a relation of the form of eqn. 1 was derived by least-squares, yielding 48 a_{ij} and 48 b_{ij} coefficients. In every case, very high correlation coefficients ($r > 0.99$) between log k'_{ij} and mole fraction, X, of water in the mobile phase were observed. These data are presented in Table II.

Studies of the relationship between the log k' for a given solute at a fixed mobile-phase composition and the C_{18} coverage of stationary phase show that linearity is observed between the lowest C_{18} surface concentration of $1.54 \cdot 10^{-4}$ and $4.96 \cdot 10^{-4}$ mol/g. There was a marked deviation from linearity at the highest C_{18} coverage (see Fig. 2 for illustration). The reason for this non-linearity at high cov-

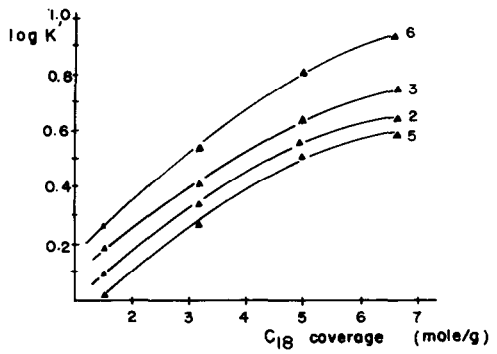

Fig. 2. Plot of log k' *versus* C_{18} coverage for solutes 2, 3, 5 and 6 (see Table I) with a mobile phase of water–methanol (65:35).

erage is not known; however due to this deviation, this stationary phase was excluded from further correlation analysis.

It was interesting to test the validity of the relation described by eqn. 5 for a particular stationary phase. For example, the coefficients a_j and b_j determined for the phase having $3.18 \cdot 10^{-4}$ mol/g C_{18} coverage ($j = 2$) are

$$a_j = 0.0534\ (\pm\ 0.0122)E_{T_i} + 3.3473\ (\pm\ 1.5866)\Delta_i - 0.4034\ (\pm\ 0.9334) \quad (9)$$

$$n = 12,\ s = 0.2481,\ r = 0.9616$$

$$b_j = -0.0415\ (\pm\ 0.0068)E_{T_i} - 2.3986\ (\pm\ 0.8663)\Delta_i - 0.9429\ (\pm\ 0.5083) \quad (10)$$

$$n = 12,\ s = 0.1353,\ r = 0.9805$$

where n is the number of data considered for deriving the regression equations, s is the standard deviation from the regression and, r is the multiple correlation coefficient. In parentheses are the 95% confidence limits as determined by the Student's t-test.

The equation obtained by substitution of a_j and b_j (calculated from eqns. 9 and 10) into eqn. 4 for the stationary phase having a C_{18} coverage of $3.18 \cdot 10^{-4}$ mol/g was then used to calculate $\log k'_i$ for a particular solute at a given mole fraction, X, of water in the mobile phase. The calculated capacity factors, $\log k'_{i,calc.}$, were compared with the previously determined experimental values. Satisfactory agreement was observed, as can be illustrated by the statistical equation:

$$\log k'_{i,obs.} = 0.9766 \log k'_{i,calc.} + 0.0058 \tag{11}$$

$$n = 48,\ s = 0.772,\ r = 0.9868$$

Next, we attempted to derive statistically a general equation relating retention to solute structure, mobile-phase composition, and stationary-phase properties. To derive coefficients a_{ij} and b_{ij} of eqn. 1 as expressed in eqn. 7, the variable matrices of dimension (12 × 3 × 4) (eqn. 8) were considered. Thus the final equation is obtained in the form:

$$\begin{aligned} \log k'_{ij} = {} & [0.0454(\pm 0.0071)E_T + 2.6493(\pm 0.9187)\Delta_i - \\ & 0.1053(\pm 0.0672)C_j - 0.4946(\pm 0.5828)]X + \\ & [-0.0381(\pm 0.0039)E_T - 2.1659(\pm 0.4919)\Delta_i + \\ & 0.1696(\pm 0.0359)C_j - 1.2963(\pm 0.3120)] \end{aligned} \tag{12}$$

The statistics are: for coefficient a_{ij}, $n = 36$, $s = 0.2756$, $r = 0.9251$ and for coefficient b_{ij}, $n = 36$, $s = 0.1476$, $r = 0.9715$. When the two hydrophobic solutes, toluene and 4-chlorotoluene, are included, the statistics deteriorated: For the term a_{ij} the correlation coefficient was 0.9035 and for b_{ij} it was 0.9285.

For the sake of comparison, a similar regression analysis was done replacing the polarity parameter, Δ_i, with the CNDO/2 calculated dipole moment. The statistics obtained were significantly lower: for a_{ij}, calculated in terms of E_{T_i}, μ_i, and C_j the correlation coefficient is equal to 0.8710, and for b_{ij} the corresponding value is 0.9040.

The correlation between the dipole moment, μ_i, and the polarity parameter, Δ_i, for the 12 solutes is not high: $r = 0.7689$. This means that the two quantities provide different structural information. However, their use together in one regression equation is unjustified statistically. For the same reason energies of the highest occupied and lowest empty molecular orbitals were not included in the regression analysis. The intercorrelation between E_{T_i} and Δ_i is only $r = 0.4879$, this means that less than 25% of the structural information may be mutually contained in both parameters.

In deriving eqn. 12 the data obtained on the stationary phase having a C_{18} coverage of $6.6 \cdot 10^{-4}$ mol/g were not included as this phase shows a marked deviation from eqn. 6. For all the remaining data a good correlation was found between retention data observed experimentally and those predicted by eqn. 12:

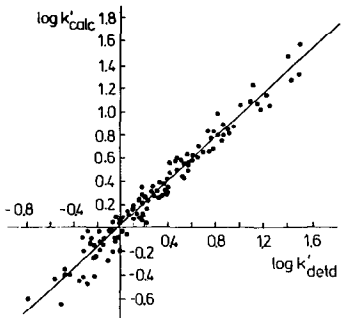

Fig. 3. Correlation of predicted (k'_{calc}) and observed (k'_{detd}) k' values for the twelve test solutes over all mobile and stationary phases examined. Correlation coefficient was 0.9862, $n = 144$.

$$\log k'_{ij,obs} = 0.9524 \log k'_{ij,calc} + 0.0142 \quad (13)$$

$$n = 144, r = 0.9862$$

The relation is presented graphically in Fig. 3 and the numerical data are collected in Table III.

DISCUSSION

As is evident from eqn. 13, the general relationship given in eqn. 12, relating retention to changes in solute structure, mobile phase composition, and stationary phase coverage, yields quantities that are in a good agreement with the experimental data. For the reversed-phase system studied, good linearity was observed between solute retention on a given phase and the composition of the binary methanol–water eluent within the concentrations investigated. Linearity was also found between the logarithm of capacity factor, determined for a given mobile phase composition and the C_{18} surface coverage of stationary phase between $1.54 \cdot 10^{-4}$ and $4.96 \cdot 10^{-4}$ mol/g. Higher C_{18} coverage lead to an observed deviation from linearity. This could be due to the rather narrow pore size of the support material (70 Å), and/or to other causes.

The above mentioned linearities observed in the chromatographic systems under study made the derivation of eqn. 12 simpler than it would be in the case of non-linear dependencies. The derivation of a general relationship, such as eqn. 12, was possible by virture of the ability of a linear combination of the CNDO/2 calculated structural indices to accurately mirror changing solute properties. The best descriptors of solute structure were found to be the total energy, E_T, and the submolecular polarity parameter, Δ. This finding tends to support the common assumption[12,13] that during a chromatographic separation, to a first approximation, two types of intermolecular forces are active, *i.e.* non-specific, non-polar, dispersive forces and structurally specific, polar, Coulombic forces.

Dispersive forces are additive, and their magnitude decreases with the 6th

TABLE III

VALUES OF THE LOGARITHM OF CAPACITY FACTORS, log k'.

For each solute the upper row gives the observed value and the lower row the value calculated by eqn. 12.

Solute	*Coverage (mol/g)*												
	1.54 · 10⁻⁴				*3.18 · 10⁻⁴*				*4.96 · 10⁻⁴*				
	Mole fraction methanol				*Mole fraction methanol*				*Mole fraction methanol*				
	0.1934	*0.2670*	*0.3524*	*0.4526*	*0.1934*	*0.2670*	*0.3524*	*0.4526*	*0.1934*	*0.2670*	*0.3524*	*0.4004*	*0.4526*
1	−0.0746	−0.2672	−0.4771	−0.7782	0.1845	−0.0280	−0.2609	−0.4771	0.4141	0.1883	−0.0544	−0.1891	−0.3274
	0.0443	−0.1387	−0.3510	−0.6002	0.2890	0.0933	−0.1338	−0.4002	0.5546	0.3451	0.1021	−0.0346	−0.1832
2	0.3696	0.0946	−0.1413	−0.4424	0.6177	0.3399	0.0649	−0.1761	0.8648	0.5621	0.2679	0.1120	−0.0401
	0.3766	0.1562	−0.0995	−0.3995	0.6213	0.3883	0.1178	−0.1995	0.8870	0.6401	0.3566	0.1926	0.0175
3	0.4025	0.1799	−0.0512	−0.3259	0.6388	0.4087	0.1521	−0.0729	0.8927	0.6435	0.3716	0.2244	−0.0813
	0.3332	0.1196	−0.1284	−0.4193	0.5780	0.3516	0.0889	−0.2193	0.8436	0.6034	0.3247	0.1681	−0.0022
4	0.6532	0.3509	0.0669	−0.2553	0.9456	0.6220	0.3010	0.0142	1.2200	0.8841	0.5477	0.3716	0.2089
	0.6283	0.3804	0.0927	−0.2448	0.8730	0.6124	0.3100	−0.0448	1.1387	0.8643	0.5458	0.3669	0.1722
5	0.2527	0.0116	−0.2341	−0.5565	0.5292	0.2730	0.0	−0.2218		0.5090	0.2609	0.0483	−0.1001
	0.3093	0.0960	−0.1514	−0.4417	0.5540	0.3281	0.0659	−0.2417		0.5799	0.3017	0.1454	−0.0247
6	0.5569	0.2706	−0.0122	−0.3259	0.8647	0.5362	0.2490	−0.0621		0.8233	0.4771	0.2946	0.1313
	0.5547	0.3135	0.0336	−0.2948	0.7994	0.5455	0.2509	−0.0948		0.7973	0.4867	0.3121	0.1223
7	0.9088	0.5712	0.2290	−0.1413	1.2467	0.8641	0.5172	0.1663		1.1794	0.7739	0.5619	0.3770
	0.8191	0.5398	0.2261	−0.1419	1.0548	0.7718	0.4434	0.0581		1.0237	0.6793	0.4857	0.2752
8	1.1461	0.7683	0.3882	0.0	1.5091	1.0892	0.6962	0.3082		1.4267	0.9709	0.7450	0.5292
	1.0729	0.7726	0.4242	0.0154	1.3176	1.0047	0.6415	0.2154		1.2565	0.8773	0.6642	0.4324
9		−0.0631	−0.2341	−0.5149	0.1845	0.0134	−0.1903	−0.3632		0.1963	−0.0263	−0.1513	−0.2762
		−0.2086	−0.4110	−0.6485	0.2106	0.0234	−0.1937	−0.4485		0.2753	0.0421	−0.0889	−0.2315
10		0.2242	0.0	−0.3010	0.5858	0.3522	0.1214	−0.1154		0.5798	0.3259	0.1761	0.0483
		0.0442	−0.1978	−0.4794	0.4938	0.2742	0.0195	−0.2794		0.5261	0.2553	0.1032	−0.0623
11		0.4652	0.1680	−0.1761	1.0855	0.7709	0.4117	0.1027		1.0012	0.6475	0.4597	0.2946
		0.5663	0.2493	−0.1225	1.0842	0.7983	0.4666	0.0774		1.0502	0.7025	0.5070	0.2945
12		0.8120	0.4649	0.0969	1.5075	1.1076	0.7339	0.3862		1.3979	0.9949	0.7845	0.5956
		0.9894	0.6092	0.1631	1.5618	1.2214	0.8265	0.3631		1.4732	1.0623	0.8313	0.5801

power of distance. This means that if one considers an interaction of a solute molecule with a fixed point on a phase and if the dispersive forces between that point and the closest atom of the solute are important, then the interactions with any other atom of the solute are negligible. During the chromatographic process, individual solute atoms (or groups of atoms) are for approximately equal time exposed to dispersive interaction in both phases. Thus, submolecular contributions to dispersive interactions tend to be additive. As expected, various measures of the molecular ability of a solute to undergo dispersive interactions have been successfully applied for correlations with the retention of homologous series, *e.g.* the sum of bond refractivities and the number of carbon atoms. The CNDO/2 calculated total energy, E_T, has also been applied in the case of homologs for correlation with GLC data[9,16]. As for the calculation of E_T which we have used in our correlation analysis, only the structural formula of the solute is required; the parameter seems to be more reliable than the parameters obtained by summation of fragmental constants.

The Coulombic forces in the case of the HPLC systems considered could result from the dipole–dipole and dipole–induced dipole interactions. These forces are of special importance in the case of correlation analysis of polar solutes and either on polar or non-polar stationary phases. Their quantitation is much more difficult than that of dispersive forces. This explains the fact that there are few published reports concerning QSRR of non-homologous sets of solutes. The success of correlation analysis of GLC data of homologs is due to the fact that their Coulombic interactions with the stationary phase are closely similar and may usually be assumed to be constant for a given series. It is a well known observation that the line describing the relation between retention and dispersive properties of one homologous series is shifted parallel to the analogous line obtained under the same conditions for another homologous series[14,23,24]. The shift is assumed to be the result of differences in the ability of a solute to participate in Coulombic interactions. As a measure of the polarity of a solute, the total dipole moment has thus far been considered. Karger *et al.*[13] determined a relationship between the total dipole moment of a solute, and a GLC retention parameter normalized to the same dispersive properties of the solutes. The authors found a linear dependence for the data considered, but the total dipole moment of the solute determined in solutions differed from the chromatographic "effective" dipole moment. Similar conclusions may be drawn from other reports[12,15].

The total dipole moment determines the net properties of the solute as a whole. Actually, a solute in specific contact with a stationary phase has at least one of its fragments close to the interacting surface, whereas the other submolecular fragments are farther away from it. For example, 1,4-dioxane has a total dipole moment near zero, but its local dipoles makes the solute behave chromatographically like a polar substance. Thus, the polar interactions of a solute would logically be better described by local, rather than total molecular dipoles.

The magnitude of Coulombic forces decreases with the 2nd power of the distance. Thus, not only the dipole formed by the two atoms closest to the interacting surface will count, but also the other, more distant dipoles. It may be extremely difficult, if even possible, to quantify precisely such submolecular increments to the overall polar interactions. In such situations, we have assumed, to a first approximation, that the largest submolecular local dipole will predominate, fully recognizing

the failure of such a simple model to account for the numerous subtleties of conformation and neighboring electronic contributions.

Other sources of error may be deviations from linearity described by eqns. 1 and 6. The independent variable in equations of the type given by eqn. 1 is the mole fraction, X, of water in the binary solvent. Probably the substitution of simple mole fraction X by a precise function of physicochemical properties of the mobile phase (if found) would reduce the error.

One must also realize that the definition of the capacity factor, k', involves some error since the marker substance, uracil, may be either partially retained or excluded from the column[25]. The quantum-chemical calculations are also approximate to some extent, especially as the standard molecular geometry assumed may differ from the actual one. Due to the high correlation between total energy and the energies of lowest empty and highest occupied molecular orbitals it was statistically unjustified to consider these indices together in a correlation analysis. However, there is a change, that charge-transfer complexes between a solute and mobile phase may be playing some role in retention.

Keeping in mind all these approximations and limitations the derived relationship describing the HPLC retention in terms of solute structure and properties of both stationary and mobile phases is encouraging. The predicting power of eqn. 12 is probably not sufficient to discriminate precisely between structurally similar solutes. On the other hand, it supplies information relevant to a general theory of HPLC separations.

ACKNOWLEDGEMENTS

R. A. Hartwick acknowledges the National Science Foundation, Grant 8100224, the donors of the Petroleum Research Fund, grant 16161-AC3, and Biomedical Research Support Grant, for support of this research.

REFERENCES

1 R. Kaliszan, *J. Chromatogr. Sci.*, 22 (1984) 362.
2 E. Tomlinson, *J. Chromatogr.*, 113 (1975) 1.
3 R. Fellous, L. Lizzani-Cuvelier, R. Luft and J. P. Rabine, *J. Chromatogr.*, 110 (1975) 13.
4 B.-K. Chen and Cs. Horváth, *J. Chromatogr.*, 171 (1979) 15.
5 J. K. Haken, A. Nguyen and M. S. Wainwright, *J. Chromatogr.*, 178 (1979) 471.
6 G. Dahlmann, J. J. K. Koser and H. H. Oelert, *Chromatographia*, 12 (1979) 655.
7 R. Kaliszan, *Chromatographia*, 10 (1977) 529.
8 D. Bonchev, O. Mekenjan, G. Protić and N. Trinajstić, *J. Chromatogr.*, 176 (1977) 149.
9 A. García-Raso, F. Saura-Calixto and M. A. Raso, *J. Chromatogr.*, 302 (1984) 107.
10 R. E. Koopmans and R. F. Rekker, *J. Chromatogr.*, 285 (1984) 267.
11 K. Jinno and K. Kawasaki, *Chromatographia*, 18 (1984) 90.
12 R. P. W. Scott, *J. Chromatogr.*, 122 (1976) 35.
13 B. L. Karger, L. R. Snyder and C. Eon, *J. Chromatogr.*, 125 (1976) 71.
14 Cs. Horváth, W. Melander and I. Molnár, *J. Chromatogr.*, 125 (1976) 129.
15 R. Kaliszan and H.-D. Höltje, *J. Chromatogr.*, 234 (1982) 303.
16 F. Sauro-Calixto, A. Garcia-Raso and M. A. Raso, *J. Chromatogr. Sci.*, 22 (1984) 22.
17 L. Buydews, D. L. Massarat and P. Geerlings, *Anal. Chem.*, 55 (1983) 738.
18 R. Kaliszan, *J. Chromatogr.*, 220 (1981) 71.
19 K. Jinno and K. Kawasaki, *Anal. Chim. Acta*, 152 (1983) 25.
20 E. Soczewiński and C. A. Wachtmeister, *J. Chromatogr.*, 7 (1962) 311.

21 J. A. Pople and G. A. Segal, *J. Chem. Phys.*, 44 (1966) 3289.
22 P. A. Dobosh, *QCPE*, 11 (1969) 141.
23 R. Kaliszan, *Chromatographia*, 12 (1979) 171.
24 E. Tomlinson, H. Poppe and J. C. Kraak, *Int. J. Pharmaceut.*, 7 (1981) 225.
25 J. H. Knox, R. Kaliszan and G. J. Kennedy, *Symp. Faraday Soc.*, 15 (1980) 113.
26 S. D. Fazio, S. A. Tomellini, Hsu Shih-Hsien, J. B. Crowther, T. V. Raglione, T. R. Floyd and R. A. Hartwick, *Anal. Chem.*, 57 (1985) 1559.
27 R. M. Meyer and R. A. Hartwick, *Anal. Chem.*, 56 (1984) 2211.