ELSEVIER

Journal of Chromatography A, 961 (2002) 171–193

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Column selectivity in reversed-phase liquid chromatography I. A general quantitative relationship

N.S. Wilson[a], M.D. Nelson[a], J.W. Dolan[a], L.R. Snyder[a,*], R.G. Wolcott[b], P.W. Carr[c]

[a]*LC Resources, 2930 Camino Diablo, Suite 110, Walnut Creek, CA 94596, USA*
[b]*Department of Chemistry, Linfield College, McMinnville, OR 97128, USA*
[c]*Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA*

Received 5 December 2001; received in revised form 26 April 2002; accepted 17 May 2002

Abstract

Retention factors *k* have been measured for 67 neutral, acidic and basic solutes of highly diverse molecular structure (size, shape, polarity, hydrogen bonding, pK_a, etc.) on 10 different C_{18} columns (other conditions constant). These data have been combined with *k* values from a previous study (86 solutes, five different C_8 and C_{18} columns) to develop a six-term equation for the correlation of retention as a function of solute and column. Values of *k* can be correlated with an accuracy of ±1–2% (1 standard deviation). This suggests that all significant contributions to column selectivity have been identified (and can be measured) for individual alkyl-silica columns which do not have an embedded polar group. That is, columns of the latter kind can be quantitatively characterized in terms of selectivity for use in the separation of any sample. © 2002 Published by Elsevier Science B.V.

Keywords: Column selectivity; Alkyl-silica columns; Retention prediction; Selectivity; Stationary phases, LC

1. Introduction

The goal of the present, ongoing study is a better quantitative understanding of the physico–chemical factors that determine differences in selectivity for various reversed-phase liquid chromatography (RPLC) columns. This should in turn lead to some practical applications: (a) a more efficient and effective use of column selectivity in the development of RPLC methods, based on a reliable classification of different columns in terms of selectivity, (b) a better characterization and control of column reproducibility for nominally similar columns from different production batches, (c) means for selecting columns of near-equivalent selectivity from different manufacturers—for use as ''second source'' alternatives to a column that may no longer be available and (d) the design and synthesis of new RPLC column packings of unique selectivity. The eventual realization of all of these goals will require a model which describes (i.e., correlates) values of *k* as a function of the reversed-phase column and sample within ±2% (1 standard deviation, SD) or better for most analytes under commonly-used separation conditions of mobile phase composition and temperature. However, we are not proposing the use of such a model for optimizing separation by computer simulation; i.e.,

*Corresponding author. Tel.: +1-510-254-6334; fax: +1-510-254-2386.
E-mail address: lloyd.snyder@lcresources.com (L.R. Snyder).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)00659-3

predictions of retention for ''new'' compounds, based (for example) on molecular structure.

The present paper describes an empirical analysis of RPLC retention for a particular set of experimental conditions and a narrow range of columns, as a first step toward the realization of the above objectives. The following paper (Part II [1]) expands this treatment for changes in temperature or mobile phase composition, and Part III [2] provides additional data for a fundamental interpretation of these combined results.

2. Theory and background

2.1. The ''solvation equation'' for RPLC retention

One approach to the description of RPLC retention (retention factors k) as a function of the column and other conditions is the ''solvation'' relationship introduced by Abraham and co-workers [3,4]:

$$\log k = C_1 + rR_2 + s\pi_2^{\mathrm{H}} + a\sum\alpha_2^{\mathrm{H}} + b\sum\beta_2 + \nu V_x \quad (1)$$

C_1 is a solute-independent constant that includes the phase ratio; it varies with temperature and the stationary and mobile phases. rR_2, $s\pi_2^{\mathrm{H}}$, $a\Sigma\alpha_2^{\mathrm{H}}$, $b\Sigma\beta_2$ and νV_x account for intermolecular interactions of the solute with the mobile and stationary phases; i.e., as a result of dispersion, cavity formation, dipole, polarizability and various hydrogen bonding contributions to retention (see Nomenclature for definitions of symbols in Eq. (1)). Subscripted symbols in Eq. (1) (R_2, π_2^{H}, etc.) represent conditionally invariant solute properties which have been measured or can be estimated for a large number of simple compounds. In the present and following papers [1,2], we will use the term ''column'' to mean a specific stationary phase, apart from any differences in column or particle dimensions.

Given a suitable set of test solutes for which the solute properties of Eq. (1) are known, it is possible to determine the corresponding system parameters (r, s, a, b and ν) for a given column and set of experimental conditions by means of Eq. (1). This in turn has two potential applications. First, a knowledge of the system parameters can provide a semi-quantitative understanding of the physico–chemical basis of retention in a particular system, thereby guiding the choice of optimal conditions for a given separation. Second, when all conditions except the column are fixed, the determination of the system parameters (r, s, a, b and ν) can be used to characterize column selectivity. In principle, values of r, s, a, etc., could be used either to choose columns of different selectivity during RPLC method development, or to verify that selectivity does not change from one batch of (nominally equivalent) column packing to another—although neither application of Eq. (1) has been reported. Eq. (1) and related variants have been widely used in an attempt to understand both isocratic [3,5–19] and gradient [20,21] RPLC separations as a function of the column and other experimental conditions.

The practical utility of Eq. (1) as a means of characterizing RPLC columns is significantly limited by two considerations. First, the predictive accuracy of Eq. (1) is no better than ±10–20% in k. This is totally inadequate for the purpose of characterizing small differences in column selectivity which lead to resolution differences of 0.5–2.0 R_s units (as might occur for different batches of a nominally equivalent packing material). Second, Eq. (1) deliberately excludes certain contributions to reversed-phase retention which are currently outside its scope; e.g., ion-exchange of cationic solutes with ionized silanols [22–24], or ''shape selectivity'' [25,26]. Furthermore, as will be argued in Part III [2], the values of individual solute parameters, π_2^{H}, α_2^{H}, β_2, etc. (especially β_2 [27]) which apply for RPLC separation are likely to differ significantly from commonly assumed values for interactions in solution, thus accounting for some of the error in predicted values of k.

2.2. An alternative to Eq. (1)

An empirical linear-free-energy equation of form similar to Eq. (1) for use in RPLC is easily visualized:

$$\log k = \log k_{\mathrm{ref}} + \eta'\mathbf{H} + \sigma'\mathbf{S} + \beta'\mathbf{A} + \alpha'\mathbf{B} + \kappa'\mathbf{C} + \ldots \quad (2)$$

The quantities η', σ', β', κ' and α' refer to some property of the solute molecule, while **H**, **S**, **A**, **C**

and **B** refer to properties of the system (column, mobile phase, temperature) that are complementary to the solute properties. In the present paper, experimental conditions other than the column are held constant, so **H**, **S**, etc., can be regarded here as properties of the column. The quantity k_{ref} is defined as the value of k for a solute (ethylbenzene) whose retention is determined mainly by ''hydrophobicity'' (i.e., term $\eta'\mathbf{H}$ of Eq. (2)); the inclusion of log k_{ref} in Eq. (2) is intended to correct for differences in the column phase ratio (e.g., surface area, ligand concentration, etc.). However, k for the Ref. solute also depends on its chemical interactions with the column, so each of the interaction terms of Eq. (2) (primarily $\eta'\mathbf{H}$ and $\sigma'\mathbf{S}$, as discussed in Ref. [2]) are relative to corresponding interactions of the reference solute with the column.

Terms $\eta'\mathbf{H}$, $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$ of Eq. (2) will be shown [2] to describe various solute–column interactions which affect retention. The application of Eq. (2) for our interpretation of RPLC retention is described below; during data analysis, we made no a priori assumptions about (a) the physico–chemical origin of terms $\eta'\mathbf{H}$, $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$, (b) values of the various parameters of Eq. (2), or (c) how many terms are required for a pragmatically ''complete'' description of retention. While we initially ignored the possible origins of terms $\eta'\mathbf{H}$, $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$, subsequent reflection [2] suggests that terms $\eta'\mathbf{H}$, $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$ of Eq. (2) are most likely determined, respectively, by hydrophobic interaction ($\eta'\mathbf{H}$), steric selectivity ($\sigma'\mathbf{S}$), hydrogen bonding between acceptor solutes and non-ionized silanols in the stationary phase ($\beta'\mathbf{A}$) or (very tentatively) donor solutes and an unidentified acceptor group in the stationary phase ($\alpha'\mathbf{B}$), and the attraction of protonated bases ($\kappa'\mathbf{C}$) by ionized silanols. Fig. 1 provides a simplified representation of these presumed contributions to RPLC retention and selectivity, corresponding to terms rR_2, $s\pi_2^H$, $a\Sigma\alpha_2^H$, $b\Sigma\beta_2$ and vV_x of Eq. (1). A justification and further description of these retention processes is given in Part III [2], which also suggests that the chemical origin of term v ($\alpha'\mathbf{B}$) is still rather unclear. At a later time, when Eq. (2) is extended to columns of different type (and possibly other solutes), it seems likely that additional terms will be required in order to recognize solute–column interactions (e.g., charge-transfer complexation) that are less significant for the solutes and columns of the present study. Concerning the symbols used in Eq. (2), note that the column parameters (**H**, **S**, etc.) are in bold to distinguish them from other common symbols, and the solute parameters use ''primed'' Greek letters for the same reason. The latter Greek and bold Arabic letters were chosen as abbreviations of *H*ydrophobicity (η' and **H**), *S*teric (σ' and **S**), *A*cidity (α' and **A**), *B*asicity (β' and **B**), and *C*ation-exchange (κ' and **C**).

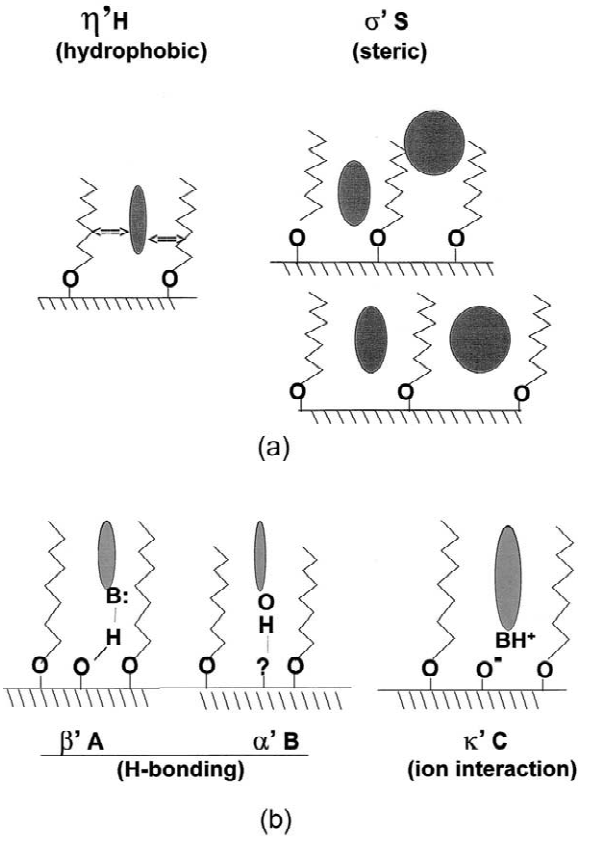

Fig. 1. Tentative representation of retention processes that correspond to various terms of Eq. (1).

As will be seen, Eqs. (1) and (2) exhibit similarities and differences. The potential advantages of Eq. (2) include:

(1) Values of the various parameters of Eq. (2) are derived from RPLC data, rather than assuming similar parameter values as found from other (quite different) chemical systems; much greater predictive

accuracy of Eq. (2) vs. Eq. (1) should therefore be possible. Because of this greater predictive accuracy of Eq. (2), and because Eq. (2) makes no prior assumptions about the interactions responsible for retention, the possibility of overlooked or ignored contributions to retention will be less likely for Eq. (2) than for Eq. (1).

(2) Eq. (2) is intended to be applicable to any kind of solute, not just (as in Eq. (1)) uncharged molecules or molecules of similar shape; resulting column parameters **H**, **S**, etc., should therefore provide a more complete assessment of column selectivity.

The means by which Eq. (2) can be applied for the interpretation of RPLC retention data are described in the Results and discussion section. A similar treatment of data presented here, based on both the solvation equation (Eq. (1)) and principle components analysis (PCA), is described elsewhere [28] with a comparison of results from these three approaches.

3. Experimental

Two separate Shimadzu HPLC systems (Columbia, MD, USA) were used to collect the data reported herein. The 67 solutes used in the present study are listed in Table 1. Some of these compounds were obtained from chemical supply houses, while other samples were the gift of Dennis Hill of the University of Connecticut at Storrs (Storrs, CT, USA). Structures for some of the less obvious solutes of Table 1 are shown in Fig. 1 of Part III [2].

For neutral solutes 1–45 of Table 1, the mobile phase was acetonitrile–water, mixed on-line at 50% (v/v). UV detection was at 205 nm, with a column temperature of 35 °C (see Appendix A for equipment details), and a flow-rate of 1.5 ml/min were employed. For acidic or basic solutes 46–67 of Table 1, the mobile phase was acetonitrile–buffer, where the buffer is 31.2 m*M* potassium phosphate (pH 2.80) prepared by titrating phosphoric acid with KOH; i.e., pH measurements were carried out on the buffer, prior to addition of acetonitrile. Samples were injected individually as 10 μl of 50 μg/ml solutions (500 ng). As discussed in Appendix A, the reproducibility of reported values of log *k* is believed to be ±0.002 log units (±0.5% in *k*, 1 SD).

The 10 columns of the present study are described in Table 2. Five or more columns of each type (from the same production batch) were the generous gift of the manufacturer: GL Sciences (Tokyo, Japan), column 1; Waters (Milford, MA, USA), columns 2 and 7–9; Agilent Technologies (Newport, DE, USA), columns 3–6; and Supelco (Bellefonte, PA, USA), column 10. The stationary phase for column 4 (SB-90%) was prepared in the same way from the same starting materials as for column 3 (SB-100%),

Table 1
Test solutes used in the present study

A. Neutral solutes			B. Basic solutes	C. Acidic solutes (weak acids)
1. Benzene[a]	16. *N*-Benzylformamide	31. Acetophenone[a]	B.1. Strong bases	56. Diclofenate acid
2. Toluene[a]	17. Anisole[a]	32. Benzophenone	46. Amitriptyline	57. Mefenamic acid
3. Ethylbenzene[a]	18. Benzyl alcohol[a]	33. *cis*-Chalcone	47. Diphenhydramine	58. Ketoprofen
4. *p*-Xylene[a]	19. 3-Phenyl propanol[a]	34. *trans*-Chalcone	48. D,L-Propanolol	59. Diflunisal
5. Propylbenzene[a]	20. 5-Phenyl pentanol	35. *cis*-4-Nitrochalcone	49. Nortriptyline	60. 4-*n*-Butylbenzoic acid
6. Butylbenzene[a]	21. Phenol[a]	36. *trans*-4-Nitrochalcone	50. Prolintane	61. 4-*n*-Pentylbenzoic acid
7. Naphthalene[a]	22. *p*-Chlorophenol[a]	37. *cis*-4-Methoxychalcone	B.2. Weak bases	62. 4-*n*-Hexylbenzoic acid
8. *p*-Chlorotoluene[a]	23. 2,3-Dihydroxynaphthalene	38. *trans*-4-Methoxychalcone	51. 4-*n*-Pentylaniline	63. 3-Cyanobenzoic acid
9. Dichlorobenzene[a]	24. 1,3-Dihydroxynaphthalene[a]	39. Prednisone	52. 4-*n*-Hexylaniline	64. 2-Nitrobenzoic acid
10. Benzotrichloride	25. Eugenol[a]	40. Hydrocortisone[a]	53. 4-*n*-Heptylaniline	65. 3-Nitrobenzoic acid
11. Bromobenzene[a]	26. Danthron	41. Mephenytoin[a]	54. *N*-Ethylaniline	66. 2,6-Dimethylbenzoic acid
12. 1-Nitropropane[a]	27. *n*-Propyl formate[a]	42. Oxazepam	55. 2-Phenyl pyridine	67. 2-Fluorobenzoic acid
13. Nitrobenzene[a]	28. Methylbenzoate[a]	43. Flunitrazepam		
14. *p*-Nitrotoluene	29. Benzonitrile[a]	44. 5,5-Diphenylhydantoin		
15. *p*-Nitrobenzyl chloride	30. Coumarin	45. *N*,*N*-Dimethyl acetamide		

[a] ''Ideal'' solutes (see Table 4 and related text).

Table 2
Characteristics of C_{18} columns used in present study; 5-μm particles, 150×4.6 mm column dimensions (see Table 5 for additional details)

Column	Abbreviation[a]	Surface area (m^2/g)	Pore diameter (nm)	% C	μmol/m^2	Metal	
						Fe	Al
1. GL Inertsil ODS-3[b]	Inertsil	436	9.5	14.7	1.74	2.8	<0.5
2. Waters Symmetry C_{18}	Symmetry	343	9	19.7	3.13	<10	<10
3. HP Zorbax SB C_{18}[c]	SB-100%	186	8	10.4	2.08	<1	<1
4. HP Zorbax SB C_{18}^c	SB-90%	188	8	9.20	1.79	<1	<1
5. HP Zorbax SB-300 C_{18}^c	SB-300	52	30	3.25	2.09	<1	<1
6. HP Eclipse XDB-C_{18}	Eclipse	186	8	10.7	3.0	<1	<1
7. YMC Pack Pro C_{18}	YMC 15	322	12.5	15.5	2.51	<10	<10
8. YMC Pack Pro C_{18}	YMC 16	321	12.5	16.3	2.68	<10	<10
9. YMC Pack Pro C_{18}	YMC 17	322	12.5	17.0	2.82	<10	<10
10. Supelco Discovery C_{18}	Discovery	190–220	17–20	12.5	3.12	<20	<1

[a] Shorthand designation of each column.
[b] Values reconfirmed with supplier (a reviewer questioned the value of 1.74 μmol/m^2).
[c] Non-end-capped columns.

except that conditions were deliberately changed to reduce the final bonded phase concentration by 10%. Columns 7–9 (YMC-15, -16 and -17) were also prepared so as to give different final bonded phase concentrations. The stationary phase of column 1 (Inertsil) is synthesized from X_2-methyloctadecylsilane, while columns 2–10 use X-dimethyloctadecylsilane (2, 6–10) or X-di-*i*-butyloctadecylsilane (3–5). Here, ''X-'' refers to a leaving group (usually –Cl). Note that the columns of Table 2 offer a range in pore diameters (8–30 nm) and both fully bonded (1–3, 5, 6, 9, 10) and partially bonded (4, 7, 8) packings. With the exception of columns 3–5, the remaining columns of Table 2 are end-capped. Despite these latter differences, the columns of Table 2 are otherwise (intentionally) similar; i.e., ''monomeric'' C_{18} packings made from type-B silica.

4. Results and discussion

4.1. Application of Eq. (2) to data from the present study

Retention data are summarized in Table 3, and Table 4 summarizes the determination of the solute and column parameters of Eq. (2) from these data. Solute and column ''hydrophobicity'' (herein measured by η' and **H**, respectively), are generally recognized as primarily responsible for RPLC retention, with retention increasing for decreasing polarity of the solute and/or column. That is, more ''hydrophobic'' solutes (typically larger, more hydrocarbon-like molecules) or columns are less polar, leading to increased solute retention. As described below, we have subtracted the contribution of hydrophobicity to retention ($\eta'\mathbf{H}$) from values of log *k* for each solute and column, allowing the (generally smaller) residual retention $\Delta = \log k - \eta'\mathbf{H}$ to be used for the determination of remaining terms $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$ of Eq. (2).

4.1.1. Step 1 of Table 4

This corresponds to a rearrangement of Eq. (2) with definition of ethylbenzene as the reference solute:

$$\log \alpha \equiv \log k - \log k_{ref} = \eta'\mathbf{H} + \sigma'\mathbf{S} + \beta'\mathbf{A} + \alpha'\mathbf{B} + \kappa'\mathbf{C} \quad (3)$$

where k_{ref} is the value of *k* for ethylbenzene. For the definitions of the remaining symbols, see the above text or the Nomenclature at the end of this paper.

4.1.2. Step 2 of Table 4

Fig. 2a shows a plot of log α for the SB-90 column vs. log α for the SB-100 column. These two columns are similar in terms of selectivity, and deviations of the data points from the best-fit line are relatively minor (S.E. = 0.015, or ±3.5% in *k* for all solutes). Fig. 2b shows a similar plot for the more

Table 3
Values of log k for the solutes of Table 1 and columns of Table 2

Solute[a]	Log k for indicated columns[b]									
	1	2	3	4	5	6	7	8	9	10
1	0.675	0.581	0.484	0.480	−0.025	0.510	0.528	0.514	0.530	0.255
2	0.888	0.803	0.702	0.692	0.176	0.734	0.748	0.736	0.752	0.471
3	1.090	1.010	0.910	0.895	0.369	0.947	0.955	0.949	0.957	0.672
4	1.107	1.036	0.924	0.908	0.379	0.964	0.976	0.966	0.979	0.693
5	1.322	1.247	1.143	1.123	0.585	1.188	1.194	1.184	1.196	0.901
6	1.552	1.478	1.373	1.347	0.802	1.424	1.426	1.419	1.429	1.127
7	1.049	0.958	0.860	0.853	0.333	0.890	0.912	0.899	0.913	0.629
8	1.110	1.027	0.920	0.907	0.382	0.957	0.973	0.956	0.955	0.691
9	1.120	1.037	0.928	0.916	0.392	0.965	0.981	0.972	0.986	0.701
10	1.251	1.145	1.057	1.046	0.519	1.093	1.112	1.111	1.111	0.817
11	0.952	0.857	0.757	0.748	0.232	0.787	0.806	0.793	0.809	0.528
12	0.248	0.130	0.066	0.080	−0.396	0.078	0.110	0.086	0.105	−0.160
13	0.521	0.388	0.333	0.346	−0.144	0.343	0.377	0.353	0.370	0.092
14	0.717	0.595	0.538	0.550	0.046	0.551	0.582	0.561	0.576	0.290
15	0.721	0.587	0.538	0.552	0.052	0.554	0.588	0.565	0.580	0.292
16	−0.249	−0.376	−0.380	−0.345	−0.785	−0.413	−0.374	−0.399	−0.380	−0.622
17	0.623	0.516	0.437	0.440	−0.061	0.457	0.483	0.464	0.481	0.202
18	−0.070	−0.186	−0.236	−0.215	−0.666	−0.237	−0.195	−0.220	−0.200	−0.456
19	0.204	0.096	0.051	0.074	−0.390	0.045	0.084	0.062	0.080	−0.178
20	0.567	0.470	0.434	0.449	−0.025	0.424	0.458	0.437	0.453	0.190
21	0.055	−0.066	−0.129	−0.111	−0.564	−0.119	−0.076	−0.100	−0.080	−0.332
22	0.330	0.208	0.138	0.154	−0.308	0.154	0.196	0.174	0.194	−0.063
23	0.158	0.028	−0.043	−0.031	−0.436	−0.039	0.039	0.024	0.032	−0.234
24	0.055	−0.088	−0.142	−0.119	−0.552	−0.131	−0.083	−0.109	−0.092	−0.331
25	0.525	0.417	0.353	0.361	−0.110	0.369	0.405	0.384	0.402	0.129
26	1.069	0.964	0.887	0.893	0.377	0.906	0.948	0.931	0.938	0.645
27	0.220	0.116	0.055	0.067	−0.419	0.059	0.082	0.060	0.079	−0.182
28	0.560	0.446	0.388	0.397	−0.100	0.395	0.423	0.402	0.418	0.124
29	0.378	0.250	0.202	0.219	−0.266	0.204	0.238	0.213	0.230	−0.043
30	0.149	0.069	−0.038	−0.020	−0.446	−0.037	0.029	0.015	0.032	−0.239
31	0.336	0.217	0.175	0.191	−0.294	0.171	0.201	0.178	0.195	−0.076
32	0.910	0.785	0.739	0.745	0.236	0.752	0.780	0.759	0.773	0.479
33	1.034	0.910	0.866	0.872	0.355	0.884	0.911	0.889	0.902	0.601
34	1.125	0.992	0.950	0.956	0.438	0.968	0.999	0.962	0.990	0.684
35	0.988	0.844	0.817	0.829	0.318	0.830	0.866	0.840	0.853	0.545
36	1.109	0.957	0.937	0.950	0.433	0.954	0.996	0.978	0.978	0.656
37	0.992	0.858	0.824	0.834	0.321	0.839	0.870	0.847	0.860	0.555
38	1.097	0.953	0.921	0.932	0.418	0.938	0.976	0.953	0.961	0.650
39	−0.124	−0.274	−0.212	−0.159	−0.621	−0.280	−0.223	−0.253	−0.239	−0.509
40	−0.106	−0.258	−0.205	−0.159	−0.602	−0.260	−0.202	−0.231	−0.218	−0.491
41	0.123	−0.003	−0.045	−0.029	−0.474	−0.041	0.002	−0.023	−0.006	−0.267
42	0.201	0.085	0.042	0.053	−0.353	0.042	0.095	0.074	0.091	−0.162
43	0.448	0.301	0.289	0.311	−0.155	0.284	0.324	0.298	0.312	0.031
44	0.209	0.007	0.053	0.079	−0.400	0.039	0.075	0.047	0.064	−0.210
45	−0.973	−0.994	−0.736	−0.698	−1.240	−1.063	−1.103	−1.139	−1.118	−1.354
46	−0.306	−0.328	−0.113	−0.126	−0.407	−0.165	−0.224	−0.158	−0.141	−0.239
47	−0.689	−0.685	−0.412	−0.419	−0.656	−0.471	−0.563	−0.484	−0.458	−0.524
48	−0.995	−0.959	−0.642	−0.645	−0.871	−0.700	−0.818	−0.721	−0.694	−0.728
49	−0.379	−0.407	−0.190	−0.196	−0.467	−0.240	−0.298	−0.237	−0.214	−0.309

Table 3. Continued

Solute[a]	Log k for indicated columns[b]									
	1	2	3	4	5	6	7	8	9	10
50	−0.757	−0.757	−0.459	−0.467	−0.736	−0.508	−0.637	−0.551	−0.525	−0.596
51	0.543	0.478	0.449	0.433	−0.039	0.425	0.441	0.434	0.447	0.202
52	0.777	0.722	0.685	0.664	0.185	0.664	0.680	0.673	0.688	0.437
53	1.013	0.967	0.923	0.897	0.409	0.904	0.920	0.914	0.930	0.674
54	0.061	−0.040	−0.068	−0.100	−0.579	−0.084	−0.068	−0.083	−0.068	−0.309
55	0.396	0.275	0.248	0.247	−0.238	0.231	0.257	0.241	0.254	−0.018
56	0.884	0.765	0.728	0.716	0.268	0.714	0.764	0.746	0.759	0.484
57	1.104	1.009	0.955	0.936	0.476	0.951	0.995	0.981	0.993	0.720
58	0.480	0.357	0.326	0.326	−0.109	0.315	0.369	0.348	0.362	0.096
59	0.695	0.497	0.471	0.494	0.120	0.349	0.441	0.406	0.414	0.171
60	0.770	0.706	0.645	0.641	0.200	0.632	0.672	0.659	0.677	0.433
61	0.987	0.936	0.867	0.858	0.410	0.856	0.893	0.884	0.901	0.656
62	1.206	1.169	1.093	1.078	0.626	1.083	1.119	1.124	1.130	0.883
63	−0.144	−0.278	−0.293	−0.268	−0.656	−0.342	−0.273	−0.299	−0.286	−0.524
64	−0.268	−0.436	−0.475	−0.448	−0.806	−0.539	−0.441	−0.480	−0.465	−0.709
65	−0.004	−0.136	−0.149	−0.126	−0.510	−0.207	−0.138	−0.165	−0.152	−0.393
66	0.145	0.033	−0.020	−0.016	−0.436	−0.027	0.023	0.001	0.019	−0.230
67	−0.081	−0.208	−0.245	−0.227	−0.626	−0.240	−0.211	−0.237	−0.214	−0.453
t_0	0.915	0.919	0.842	0.864	0.984	0.853	1.022	1.014	0.981	1.132

See the Experimental section for conditions (50% (v/v) ACN–water or buffer, pH 2.80, 35 °C).

[a] Solute numbers as in Table 1.

[b] Column numbers as in Table 2.

Table 4
Steps in the determination of the parameters of Eq. (2)

1.	For each column, calculate log $\alpha = \log (k/k_{ref})$ for each solute
2.	Correlate values of log α for all columns vs. log α for the SB-100 column; identify ''ideal'' neutral[a] solutes (1–9, 11–13, 17–19, 21, 22, 24, 25, 27–29, 31, 41) with SD≤0.010 log units
3.	Correlate values of log α for column i vs. log α for SB-100 column and ''ideal'' neutral solutes: $y = \mathbf{H}_i x$ (assumes $\mathbf{H} = 1$ for SB-100 column)
4.	Calculate $\eta'' = (\log \alpha)/\mathbf{H}$ for each solute and column; calculate average values of η'' ($= \eta''_{avg}$) for each solute and determine SD; select ''nonideal'' solutes with S.E.≥0.017 (equal twice S.E. for ''ideal'' solutes): 20, 23, 30, 32–40, 42–54, 56–65
5.	Determine $\Delta = \log \alpha - (\eta''_{avg}\mathbf{H})$ for each solute and column; cross-correlate value of Δ (10 different columns) for different ''nonideal'' solutes; group solutes which are similar in terms of this correlation; each solute group corresponds to one of the remaining terms of Eq. (3) ($\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\kappa'\mathbf{C}$, or $\alpha'\mathbf{B}$), labeled accordingly ''H'', ''S'', etc.
6.	Determine values of column parameters from value of Δ: $\mathbf{A} = \Delta$ for solute 45 (*N*,*N*-dimethylacetamide) $\mathbf{C}$ = average Δ for solutes 46–50 (strong bases) $\mathbf{B}$ = average Δ for solutes 56–58, 60–65 (acids; 59 excluded) $\mathbf{S}$ = average of: average Δ for solutes 32–38, 39–40, and 43–44
7.	Correlate values of log α for each solute and 10 columns vs. column parameters of 6 (multiple regression via Eq. (3)) to obtain solute parameters (η', σ', β', κ', α') for each solute; calculate log α from Eq. (3) (SD = 0.005)
8.	Correlate (multiple regression via Eq. (3)) values of log α vs. solute parameters from 7 to obtain final column parameters ($\mathbf{H}$, $\mathbf{S}$, $\mathbf{A}$, $\mathbf{C}$, $\mathbf{B}$); calculate log α from Eq. (3) (SD = 0.004)

[a] ''Ideal neutral'' solutes include polar and nonpolar species which have somewhat similar shape and are at most only weak hydrogen bond acceptors; they can be reasonably strong hydrogen bond donors.

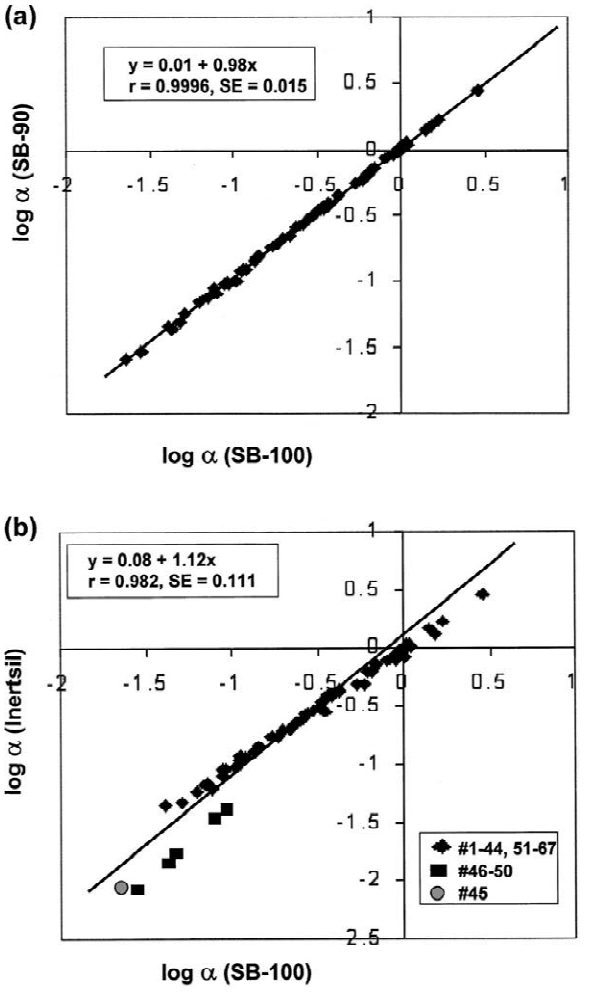

Fig. 2. Retention compared for two columns of Table 2 (SB-90, a, and Inertsil, b) vs. the SB-100 column.

different Inertsil and SB-100 columns. The scatter of data from the best-fit line (S.E.=0.11) is considerably greater in Fig. 2b vs. Fig. 2a, especially for the strong bases (46–50, squares) and the one aliphatic amide (45, circle). Correlations such as those of Fig. 2 allow the average deviation (as measured by the standard error, S.E.) of individual solutes to be calculated for all 10 columns of Table 2. Solutes with sufficiently small deviations (values of S.E.≤ 0.010) can be assumed to be retained almost entirely by hydrophobic interactions, the primary retention mode for RPLC. For these solutes:

$$\log \alpha \approx \eta' \mathbf{H} \tag{4}$$

We have selected 24 of the solutes of Table 4 for which Eq. (4) applies (solutes with S.E.≤0.01 for all 10 columns). The latter solutes (1–9, 11–13, 17–19, 21, 22, 24, 25, 27–29, 31, 41) are defined as ''ideal'' solutes, meaning that their values of σ', β', α' and κ' are approximately zero.

4.1.3. Step 3 of Table 4

If Eq. (4) were to apply exactly for two columns a and b:

$$\log \alpha_b = (\mathbf{H}_b/\mathbf{H}_a) \log \alpha_a \tag{5}$$

where subscripts a and b refer to column a or b, respectively. For the ''ideal'' solutes, log α was correlated for each column (b) vs. the SB-100 column (a) by means of Eq. (5). By defining $\mathbf{H}_a$ equal to 1.000 for the SB-100 column, the slope of these plots becomes equal to $\mathbf{H}_b$ for each column (b).

4.1.4. Step 4 of Table 4

Given experimental values of α, and having determined values of **H** for each of the 10 columns, Eq. (4) was used to calculate tentative (i.e., initial approximation) values of η' (equal to η'') for all 67 solutes and the 10 columns of Table 2. For each solute, values of η'' were averaged over all 10 columns (η''_{avg}), and SDs were determined for each solute and all 10 columns. Larger contributions to log α from terms $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$ of Eq. (3), combined with differences in ratios of the column parameters (**S**/**H**, **A**/**H**, etc.; see discussion of Ref. [29]), result in a greater column-to-column variability in values of η'' for some solutes (and larger values of SD). The latter ''non-ideal'' solutes (20, 23, 30, 32–40, 42–54, 56–65) with values of SD>0.017 were thus identified for the further study of terms $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$

4.1.5. Step 5 of Table 4

For each ''nonideal'' solute i from step 4, deviations Δ from Eq. (4) were determined for each column j:

$$\Delta_{ij} = \log \alpha_{ij} - (\eta''_{avg})_i \mathbf{H}_j \tag{6}$$

Values of $\Delta_{ij} \equiv \Delta$ for various pairs of individual

solutes (x, y) and all 10 columns were correlated in terms of $y=$(constant)x, and solute-pairs were grouped on the basis of high values of the correlation coefficient r and values of S.E./SD$<$0.5; S.E. refers to the standard error of the correlation, while SD refers to the standard deviation of values of Δ for a given solute i and all 10 columns. Fig. 3 illustrates this correlation for various solutes (y) vs. solute 48 (propranolol, x). The resulting solute groupings are summarized in step 6 of Table 4. While it is seen that each solute grouping corresponds to solute molecules of a certain ''type'' (e.g., solutes in group **C** are all strong Bronsted bases), these groupings were in each case determined by whether values of Δ for the members of a group are highly correlated, as in Fig. 3a.

Each solute grouping of step 6 (Table 4) corresponds to one of the terms $\sigma'\mathbf{S}$, $\beta'\mathbf{A}$, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$

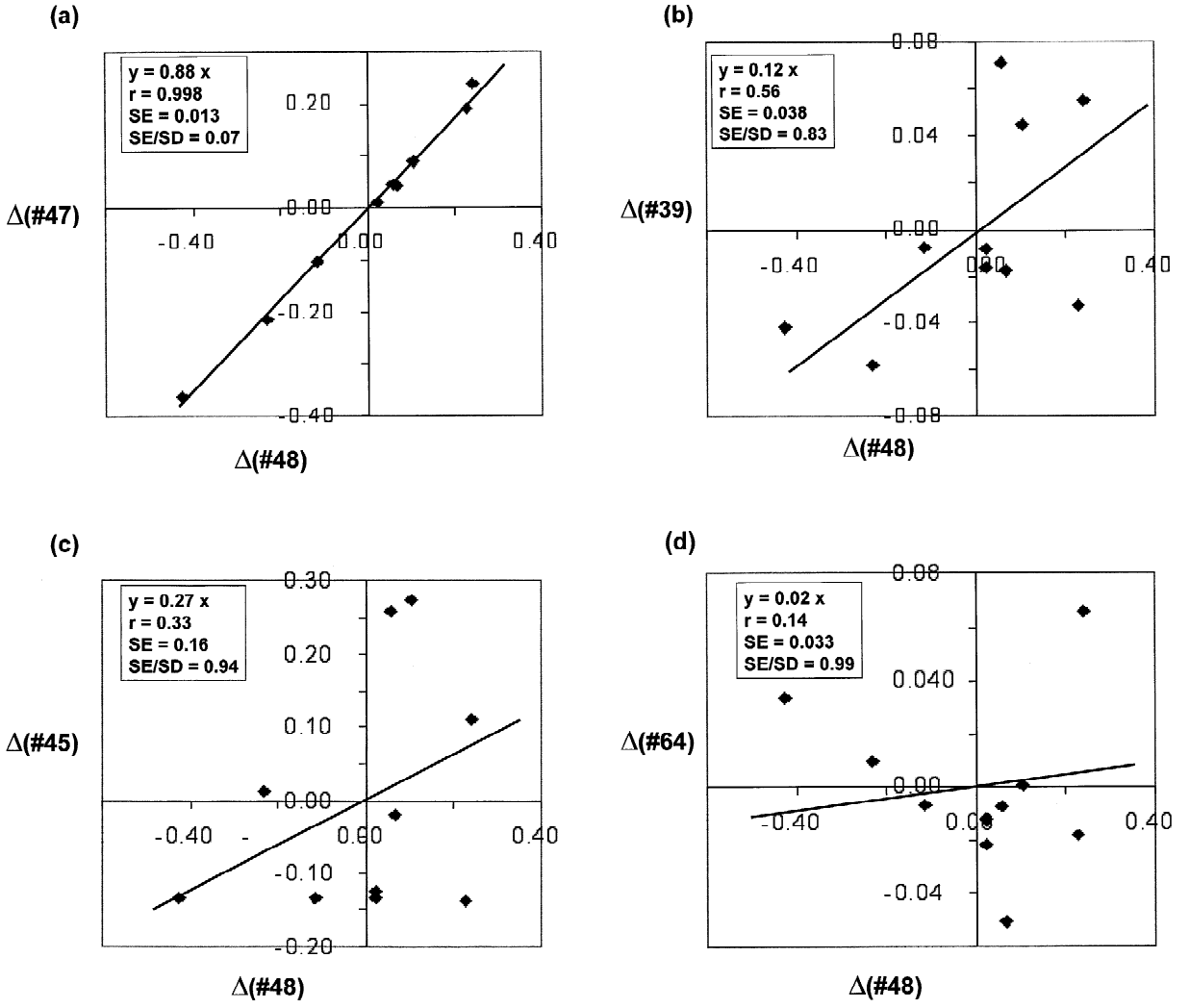

Fig. 3. Correlation ($y=ax$) of values of $\Delta=\log\alpha-\eta'\mathbf{H}$ for various solutes (y) vs. values for solute 48 (propranolol) (x). (a) $y=$solute 47 (diphenhydramine); (b) $y=$solute 39 (prednisone); (c) $y=$solute 45 (*N*,*N*-dimethylacetamide); (d) $y=$solute 64 (2-nitrobenzoic acid). S.E./SD refers to the value of S.E. for the correlation and the value of SD for values of Δ for solute-y (a value of S.E./SD$>$0.5 suggests little correlation). See Table 1 for solute numbering and related text for details.

of Eq. (3), identified in Table 4 by the corresponding column parameter (**S**, **A**, **B** or **C**). The weak bases (51–54) showed modest correlations with solutes from groups **A** and **C**, but average $|\Delta|$ values for these solutes (0.02–0.03) were much smaller than for other solutes assigned to groups **A** and **C** (average $|\Delta|$ equal to 0.15–0.20). Solutes 51–54 were therefore not assigned to any of the latter groups (**S**, **A**, **B** or **C**). For similar reasons, solutes 20, 23, 30, 42 and 59 could not be assigned to any single group. Poor or marginal correlation of Δ values as in Fig. 3b–d is expected for (a) solutes whose retention is substantially affected by more than one of the physico–chemical factors which determine terms σ'**S**, β'**A**, α'**B** and κ'**C** of Eq. (3) or (b) solutes with very small values of Δ—for which experimental error becomes relatively more important.

4.1.6. Step 6 of Table 4

Values of Δ for those solutes which unambiguously belong to one of the four groups defined in step 5 can be used to determine the column parameters **S**, **A**, **B** and **C**. For column parameter **A**, the corresponding solute group contains only one representative solute (45, *N*,*N*-dimethylacetamide), similar to the case of excluded ''nonideal'' solutes 20, 23, 30, 42 and 59 from step 5. However, values of Δ for solute 45 are quite large (-0.13 to 0.28 log units), and in the following section an analysis of retention data from another study [6] further confirms the importance of term (β'**A**) of Eq. (3). Therefore, values of Δ for solute 45 were set equal to the column parameter **A**.

The column parameters **C** and **B** were equated to the average values of Δ for each column and the solutes included in the corresponding group of ''nonideal'' solutes defined in Table 4: 46–50 for **C**, and 56–58, 60–65 for **B**. A similar procedure could have been used for the determination of the column parameter **S**, except that of the 11 solutes included in this group, seven solutes have very similar shapes (32–38). The use of an average value of Δ for the entire group would therefore bias the final value of **S** toward a specific solute shape—which seems unwise for a column parameter that now appears to be related to ''shape'' or ''steric'' selectivity. For this reason, these 11 solutes were first grouped into three sub-groups of related structure (32–38, 39–40, and 43–44), and the average values of Δ for each subgroup were then averaged to obtain a final value of **S**. Values of the column parameters **H**, **S**, etc. (after final adjustment as in step 8) are listed in Table 5.

4.1.7. Step 7 of Table 4

Multiple regression via Eq. (3) of values of log α vs. values of the column parameters **H**, **S**, **A**, **B** and **C** from step 6 yielded final values of the solute parameters η', σ', β', α' and κ' (Table 6). Eq. (3) with these values of the solute and column parameters allows the prediction of values of log α for all 67 solutes and the 10 columns with SD = 0.005. As expected, ''related'' solutes in step 6 of Table 4 exhibit relatively large values of the complementary solute parameter; i.e., $0.66 \leq \sigma' \leq 1.43$ for solutes in group **S**, $\beta' = 1.00$ for the one solute in group **A**, $0.55 \leq \alpha' \leq 1.45$ for the solutes in group **B**, and $0.83 \leq \kappa' \leq 1.23$ for solutes in group **C**.

4.1.8. Step 8 of Table 4

If values of log α are regressed (again) vs. values of η', σ', β', α' and κ' obtained from step 7 via Eq. (3), a small adjustment in the column parameters results (see Table 5 for final values), with SD = 0.004 for the prediction of all values of log α. Since a value of α is the result of two experimental measurements, the implied accuracy of Eq. (2) for log k is $\pm 0.004/2^{1/2} = \pm 0.003$ (1 SD). This can be compared with the experimental repeatability of log k equal to ± 0.002.

Given values of the parameters of Table 5 for columns 1–10, the reliability of Eq. (3) can be further tested by fitting experimental values of k for additional solutes to these column parameters. Part III [2] presents such a test for 23 additional solutes whose structures are in many cases very different from those of compounds 1–67. With the exception of two outliers (dinitro- and trinitrophenols; SDs = 0.016 and 0.025, respectively), the standard deviation was 0.004 log units; i.e., the same as for the 67 solutes of Table 3.

4.2. Application of Eq. (3) to the data of Ref. [5]

Data similar to those of Table 3 have been reported by Tan et al. [5] for five different columns

Table 5
Column parameters **H**, **S**, **A**, **B** and **C** for columns of present study and study of Ref. [5]; final values from step 8 of Table 4

Column	**H**	**S**	**A**	**B**	**C**	Comments[a]
Present study (35 °C)						
1. Inertsil	1.0048	−0.0126	−0.1285	−0.0255	−0.3501	$>$Si(CH$_3$)–C$_{18}$[b]
2. Symmetry	1.0498	−0.0588	0.0104	−0.0289	−0.2071	
3. SB-100	0.9981	0.0211	0.2715	0.0064	0.0854	–Si(*i*-butyl)$_2$–C$_{18}$
4. SB-90	0.9666	0.0418	0.2642	0.0093	0.0505	–Si(*i*-butyl)$_2$–C$_{18}$
5. SB-300	0.8945	0.0426	0.1092	0.0761	0.2204	–Si(*i*-butyl)$_2$–C$_{18}$
6. Eclipse	1.0355	−0.0084	−0.0202	−0.0325	0.0443	
7 YMC 15	1.0022	0.0022	−0.1362	−0.0128	−0.0960	
8 YMC 16	1.0195	−0.0077	−0.1317	−0.0105	0.0088	
9. YMC 17	1.0106	−0.0067	−0.1357	−0.0099	0.0135	
10. Discovery	0.9861	−0.0226	−0.1279	0.0163	0.1899	
Ref. [5] (25 °C)						
1a. Zorbax StableBond C$_{18}$	0.9907	0.0118	0.3429			–Si(*i*-butyl)$_2$–C$_{18}$; same column (different lot) as 3
2a. Zorbax Rx	1.0651	−0.0557	0.3853			
3a. Hypersil C$_{18}$	0.9635	−0.0065	0.0967			
4a. Hypersil C$_8$	0.8536	0.0170	0.0409			
5a. Zorbax C$_8$	0.8267	0.0055	0.0643			

See text for details.

[a] Ligand is –Si(CH$_3$)$_2$–C$_8$ or –Si(CH$_3$)$_2$–C$_{18}$, unless noted otherwise.

[b] Although the ligand is difunctional, the manufacturer claims that this packing is not ''polymeric'', and this claim is confirmed by the data of Table 6 of Ref. [2].

and 86 neutral solutes, 61 of which differ from the compounds of Table 1. We have therefore carried out a similar analysis of the data of Ref. [5] as in Table 4, as a further test of Eq. (3) and in order to obtain parameter values for additional columns and solutes studied by Tan et al. (but not by us). Experimental conditions other than the choice of column or solute were the same for the present study and that of [5], except for temperature (25 °C in Ref. [5] vs. 35 °C in the present study). We have determined the effect of temperature, *T*, on the retention of the solutes of Table 3, as reported in a following paper (Part II [1]) for three of the columns of Table 2 (2–4). This allowed us to compare the retention of several solutes from the present study and Ref. [5] at the same temperature (25 °C). The data of Ref. [5] do not include compounds which are likely to have large values of the solute parameters α' (acids) or κ' (strong bases), thereby precluding measurement of **B** or **C** for the columns of Ref. [5]. We therefore modified Eq. (3) to exclude terms *v* and *vi* for this sample set:

$$\log \alpha \approx \eta'\mathbf{H} + \sigma'\mathbf{S} + \beta'\mathbf{A} \quad (7)$$

In this way (see Appendix B), we were able to obtain values of the solute (η', σ' and β') and column (**H**, **S** and **A**) parameters for several additional columns (1a–5a of Table 5) and solutes (1a–87a of Table 7). Note the special numbering (1a, 2a, etc.) of these latter solutes and columns, both here and in following papers [1,2]. The parameters of Tables 5 and 7 for the data of Ref. [5] allow the calculation of values of log α for these 86 solutes and five columns (at 25 °C) by means of Eq. (7). The overall agreement of experimental and calculated values of log α was ±0.008 (1 SD), corresponding to SD = 0.008/$2^{0.5}$ = 0.006 for log *k*. There was better agreement in log α (±0.002) for the two C$_{18}$ type-B-silica columns (1a, 2a) vs. the three type-A-silica columns (±0.010; 3a, 4a, 5a), which comprise both C$_8$ and C$_{18}$ phases. This twofold greater SD for the data of Ref. [5] vs. the data of Table 3 (SD = 0.008 vs. 0.004) is likely attributable to the greater diversity of columns 1a–5a (C$_8$ and C$_{18}$, type-A and -B silica) compared to columns 1–10 (C$_{18}$ and type-B only).

Small changes in the various solute and column

Table 6
Solute parameters for 67 solutes of present study (35 °C), based on Eq. (3); final values from step 7

Solute	Solute parameter					
	η'	σ'	β'	α'	κ'	SD[a]
1. Benzene	−0.424	−0.203	0.013	−0.041	−0.019	0.002
2. Toluene	−0.206	−0.133	0.004	−0.014	−0.008	0.002
3. Ethylbenzene	0.000	0.000	0.000	0	0	0.000
4. *p*-Xylene	0.018	−0.118	−0.004	0.013	0	0.002
5. Propylbenzene	0.234	0.134	−0.013	0.028	0.001	0.002
6. Butylbenzene	0.464	0.279	−0.028	0.105	0.001	0.002
7. Naphthalene	−0.046	0.057	−0.015	0.154	−0.022	0.003
8. 4-Chlorotoluene	0.012	−0.088	0.006	0.149	−0.024	0.006
9. *p*-Dichlorobenzene	0.024	−0.043	−0.016	0.137	−0.017	0.002
10. Benzotrichloride	0.152	0.412	−0.049	0.126	−0.026	0.003
11. Bromobenzene	−0.149	−0.047	−0.009	0.093	−0.027	0.003
12. 1-Nitropropane	−0.844	−0.036	0.005	−0.112	−0.004	0.002
13. Nitrobenzene	−0.579	0.322	−0.009	0.01	−0.036	0.003
14. 4-Nitrotoluene	−0.376	0.428	−0.007	0.035	−0.027	0.002
15. 4-Nitrobenzyl chloride	−0.373	0.597	−0.031	0.016	−0.026	0.002
16. *N*-Benzylformamide	−1.312	0.040	0.072	0.041	0.032	0.003
17. Anisole	−0.473	0.042	0.001	−0.052	−0.019	0.002
18. Benzyl alcohol	−1.147	−0.143	0.010	−0.102	0.021	0.002
19. 3-Phenyl propanol	−0.865	0.011	0.024	0.12	0.02	0.003
20. 5-Phenyl pentanol	−0.490	0.211	0.035	0.369	0.025	0.004
21. Phenol	−1.031	−0.165	−0.024	−0.035	0.016	0.002
22. *p*-Chlorophenol	−0.760	−0.039	−0.042	0.149	0.001	0.003
23. 2,3-Dihydroxynaphthalene	−0.928	−0.011	−0.113	0.609	−0.031	0.007
24. 1,3-Dihydroxy naphthalene	−1.038	−0.035	−0.056	0.198	0.004	0.005
25. Eugenol	−0.553	0.124	−0.027	0.15	0.01	0.002
26. Danthron	−0.019	0.473	−0.038	0.285	−0.038	0.005
27. *n*-Propyl formate	−0.865	−0.174	0.052	−0.188	0.009	0.002
28. Methylbenzoate	−0.532	0.297	0.027	−0.039	−0.038	0.003
29. Benzonitrile	−0.715	0.245	0.016	−0.02	−0.026	0.002
30. Coumarin	−0.927	−0.554	−0.018	0.648	−0.041	0.011
31. Acetophenone	−0.748	0.186	0.039	−0.047	−0.009	0.001
32. Benzophenone	−0.180	0.660	−0.014	0.089	−0.026	0.001
33. *cis*-Chalcone	−0.052	0.817	−0.024	0.066	−0.021	0.001
34. *trans*-Chalcone	0.032	0.918	−0.030	0.179	−0.042	0.005
35. *cis*-4-Nitrochalcone	−0.102	1.101	−0.044	0.069	−0.035	0.001
36. *trans*-4-Nitrochalcone	0.021	1.434	−0.078	0.013	−0.037	0.005
37. *cis*-4-Methoxychalcone	−0.095	0.965	−0.033	0.057	−0.025	0.001
38. *trans*-4-Methoxychalcone	0.005	1.167	−0.059	0.129	−0.042	0.002
39. Prednisone	−1.167	0.982	0.090	0.023	0.024	0.006
40. Hydrocortisone	−1.151	0.965	0.050	0.096	0.027	0.006
41. Mephenytoin	−0.955	0.112	−0.023	0.047	0.018	0.002
42. Oxazepam	−0.861	0.021	−0.056	0.578	0.03	0.004
43. Flunitrazepam	−0.632	0.752	−0.014	0.158	−0.015	0.002
44. 5,5-Diphenylhydantoin	−0.881	1.284	−0.046	−0.448	0.029	0.008
45. *N,N*-Dimethylacetamide	−1.921	0.000	1.000	0	0	0.003
46. Amitriptyline	−1.096	0.049	−0.030	0.321	0.834	0.003
47. Diphenhydramine	−1.412	−0.057	0.004	0.16	1.022	0.002
48. Propranolol	−1.654	−0.180	0.011	−0.329	1.23	0.002
49. Nortriptyline	−1.169	0.059	−0.036	0.381	0.833	0.002

Table 6. Continued

Solute	Solute parameter					
	η'	σ'	β'	α'	κ'	SD[a]
50. Prolintane	−1.476	0.130	0.050	−0.533	1.08	0.006
51. 4-*n*-Pentylaniline	−0.495	−0.246	0.082	0.257	0.094	0.004
52. 4-*n*-Hexylaniline	−0.258	−0.213	0.076	0.423	0.09	0.005
53. 4-*n*-Heptylaniline	−0.019	−0.175	0.071	0.575	0.086	0.006
54. *N*-Ethylaniline	−1.013	−0.410	0.058	−0.582	0.091	0.010
55. 2-Phenylpyridine	−0.688	0.212	0.052	−0.05	−0.005	0.004
56. Diclofenate acid	−0.192	0.400	−0.036	0.862	−0.031	0.004
57. Mefenamic acid	0.038	0.262	−0.039	0.917	−0.006	0.004
58. Ketoprofen	−0.589	0.296	−0.044	0.546	0.005	0.004
59. Diflunisal	−0.469	0.168	0.152	3.097	−0.428	0.015
60. 4-*n*-Butylbenzoic acid	−0.272	−0.280	0.015	1.024	0.044	0.005
61. 4-*n*-Pentylbenzoic acid	−0.049	−0.307	0.016	1.185	0.047	0.006
62. 4-*n*-Hexylbenzoic acid	0.178	−0.299	0.005	1.35	0.056	0.007
63. 3-Cyanobenzoic acid	−1.215	−0.057	0.031	0.911	−0.042	0.002
64. 2-Nitrobenzoic acid	−1.386	−0.190	0.024	1.454	−0.197	0.003
65. 3-Nitrobenzoic acid	−1.076	−0.016	0.053	1.205	−0.073	0.002
66. 2,6-Dimethylbenzoic acid	−0.929	−0.221	−0.019	0.463	0.008	0.002
67. 2-Fluorobenzoic acid	−1.153	−0.152	−0.004	0.356	0.03	0.008

See Table 4 for basis of calculation.

[a] Standard deviation of fit to Eq. (3) for each solute.

parameters were observed as a result of this 10 °C decrease in temperature and the exclusion of terms *v* and *vi*; i.e., average differences of η', -0.008 ± 0.023 (1 SD); σ', -0.023 ± 0.074; β', -0.001 ± 0.001; **H**, 0.000 ± 0.003; **S**, 0.003 ± 0.009; **A**, 0.000 ± 0.009. The much smaller changes in the column parameters **H**, **S** and **A** (compared to the solute parameters η', σ' and β') are expected; see the discussion in Part II [1].

When values of the solute parameters of Table 6 (at 35 °C) and Table 7 (at 25 °C) are compared for the same compounds, we have the following average differences for values in Table 7 vs. Table 6: η', 0.00 ± 0.02 (1 SD); σ', -0.04 ± 0.09; β', 0.01 ± 0.03. Considering that these solute parameters are likely significant to only ±0.05 units (see the following section), and values of the solute parameters are expected to change with temperature [1], the observed agreement seems satisfactory.

4.3. Significance of the solute parameters of Tables 6 and 7

An interpretation of the solute and column parameters of Eq. (7) in terms of solute molecular structure, stationary phase characteristics, and the nature of the retention process is deferred to Part III [2]. Other comments on the solute parameters follow.

4.3.1. Significance of solute parameter values

By ''significance'' we mean the extent to which values of these parameters reflect an actual physical process, rather than being the result of experimental error and/or ''noise'' in the data reduction process. In order to avoid any over-interpretation of these solute parameters, it is important to have some measure of their repeatability and significance. Some insight is provided by a comparison of solute parameters for (a) molecules of ''similar'' structure or (b) homologous series. Table 8 summarizes some comparisons for solutes of similar structure. The average difference in solute parameters for these compounds is about ±0.1 unit, which defines a maximum uncertainty in each parameter. Keep in mind that slight differences in solute structure can result in real differences in η', σ', etc., as suggested in Ref. [2] for the *cis* vs. *trans* isomers of solutes 33–38.

For a homologous series, regular changes in all molecular properties are expected. Values of each solute parameter (η', σ', etc.) were observed to vary (approximately) linearly with *n*, the number of methylene groups in the molecule (where $n\geq2$). The

Table 7
Solute parameters for 86 solutes of Ref. 5 (25 °C), based on Eq. (3)

Solute	Solute parameter		
	η'	σ'	β'
1a. 1-Butanol	−1.328	−0.626	0.189
2a. 1-Hexanol	−0.820	−0.134	0.138
3a. 1-Octanol	−0.328	0.342	0.097
4a. Isopropanol	−1.702	−1.201	0.238
5a. Cyclohexanol	−1.173	−0.342	0.236
6a. 1-Butanal	−1.017	−0.201	0.247
7a. 1-Hexanal	−0.508	0.360	0.174
8a. 1-Heptanal	−0.270	0.520	0.168
9a. 1-Octanal	−0.045	0.515	0.202
10a. *N*,*N*-Dimethylformamide	−2.062	−0.333	0.892
11a. *N*,*N*-Diethylformamide	−1.528	0.207	0.488
12a. *N*,*N*-Dibutylformamide	−0.559	1.193	0.201
13a. *N*,*N*-Dimethylacetamide	−2.020	0.001	0.992
14a. *N*,*N*-Diethylacetamide	−1.493	0.264	0.525
15a. *n*-Propylformate	−0.876	−0.142	0.055
16a. *n*-Butylacetate	−0.555	0.135	0.071
17a. *n*-Pentylacetate	−0.316	0.364	0.041
18a. *n*-Hexylacetate	−0.076	0.525	0.038
19a. Ethylpropionate	−0.791	−0.112	0.067
20a. Ethylbutyrate	−0.543	0.035	0.035
21a. Diethylether	−0.996	−0.737	0.238
22a. di-*n*-Propylether	−0.355	−0.381	0.122
23a. di-*n*-Butylether	0.175	0.047	0.089
24a. Dioxane	−1.664	−0.714	0.523
25a. Acetone	−1.518	−0.794	0.221
26a. Butane-2-one	−1.227	−0.510	0.142
28a. Heptane-2-one	−0.480	0.262	0.050
29a. Nonane-2-one	0.002	0.589	0.045
30a. Cyclopentanone	−1.227	−0.198	0.239
31a. *n*-Propionitrile	−1.243	−0.378	0.056
32a. *n*-Valeronitrile	−0.770	0.165	−0.005
33a. *n*-Hexanitrile	−0.533	0.326	−0.028
34a. *n*-Hexylcyanide	−0.257	0.517	−0.183
35a. *n*-Heptylcyanide	−0.061	0.669	−0.090
36a. *n*-Octylcyanide	0.176	0.779	−0.100
37a. *n*-Nitropropane	−0.828	0.030	−0.037
38a. *n*-Nitrobutane	−0.594	0.233	−0.057
39a. *n*-Nitropentane	−0.360	0.424	−0.079
40a. Methylene chloride	−0.749	−0.265	−0.021
41a. Chloroform	−0.496	−0.092	−0.040
42a. Dibromomethane	−0.629	−0.159	0.009
43a. Benzyl alcohol	−1.190	−0.367	0.046
44a. 2-Phenyl ethanol	−1.074	−0.237	0.050
45a. 3-Phenyl propanol	−0.912	−0.054	0.050
46a. Benzaldehyde	−0.799	0.097	0.078
47a. *N*-Benzylformamide	−1.359	−0.091	0.102
48a. Methyl benzoate	−0.549	0.099	0.060
49a. Ethyl benzoate	−0.318	0.310	0.001

Table 7. Continued
Solute parameters for 86 solutes of Ref. 5 (25 °C), based on Eq. (3)

Solute	Solute parameter		
	η'	σ'	β'
50a. Anisole	−0.484	−0.060	0.040
51a. Acetophenone	−0.782	0.129	0.095
52a. Propiophenone	−0.511	0.233	0.070
53a. Benzophenone	−0.179	0.772	−0.018
54a. Benzonitrile	−0.712	0.312	0.004
55a. *m*-Toluenitrile	−0.509	0.462	0.014
56a. Benzyl cyanide	−0.704	0.332	−0.033
57a. Nitrobenzene	−0.574	0.325	−0.006
58a. *m*-Nitrotoluene	−0.346	0.483	−0.033
59a. *o*-Nitrotoluene	−0.394	0.475	−0.047
60a. *p*-Nitrotoluene	−0.369	0.431	−0.029
61a. *p*-Nitrobenzyl bromide	−0.304	0.761	−0.049
62a. *p*-Nitrobenzyl chloride	−0.360	0.547	−0.071
63a. Fluorobenzene	−0.417	−0.114	−0.039
64a. Chlorobenzene	−0.213	−0.172	−0.010
65a. Bromobenzene	−0.155	−0.154	0.002
66. Iodobenzene	−0.047	−0.084	0.022
67a. Benzyl bromide	−0.206	0.405	−0.047
68a. *p*-Chlorotoluene	0.015	−0.152	−0.003
69a. *p*-Bromotoluene	0.073	−0.136	0.018
70a. *p*-Dichlorobenzene	0.019	−0.164	0.014
71a. Benzene	−0.434	−0.270	0.010
72a. Toluene	−0.213	−0.204	0.009
73a. Ethylbenzene	0.000	0.000	0.000
74a. *n*-Propylbenzene	0.240	0.145	−0.001
75a. *n*-Butylbenzene	0.480	0.353	−0.010
76a. *tert*.-Butylbenzene	0.338	0.448	−0.028
77a. *p*-Xylene	0.006	−0.250	0.030
78a. Mesitylene	0.231	−0.041	0.031
79a. Biphenyl	0.174	0.219	0.001
80a. Naphthalene	−0.053	−0.078	0.023
81a. Anthracene	0.353	−0.490	0.081
82a. Phenol	−1.024	−0.183	−0.072
83a. *m*-Cresol	−0.860	−0.055	−0.055
84a. *p*-Cresol	−0.860	−0.056	−0.058
85a. *o*-Cresol	−0.800	−0.010	−0.069
86a. *p*-Ethylphenol	−0.660	0.055	−0.062
87a. *p*-Chlorophenol	−0.748	0.035	−0.094

See Table 4 and text for basis of calculation. Data for solute 27 were discarded as inconsistent.

scatter of data for plots of η', σ', etc., vs. n therefore provides a further estimate of the maximum uncertainty in these solute parameters: ±0.01 units (1 SD) for the data of Table 6 and ±0.02–0.07 for the data of Table 7. We conclude that the solute parameters of Tables 6 and 7 are likely reproducible and significant to about ±0.05 units in each parameter. For this reason, differences ≤0.05 units in values of η', σ', etc., for different solutes should not be considered highly significant.

*4.3.2. Relative importance of terms η'**H**, σ'**S**, β'**A**, α'**B** and κ'**C** of Eq. (2)*

The contribution of terms η'**H**, σ'**S**, β'**A**, α'**B** and κ'**C** of Eq. (2) to solute retention can be described in terms of the average change in retention

Table 8
Significance of solute parameters as suggested by some comparisons of solutes with ''similar'' structure

Solute pair	Difference in solute parameters for indicated pairs of solutes				
	η'	σ'	β'	α'	κ'
33, 34	−0.084	−0.101	0.006	−0.113	0.021
35, 36	−0.123	−0.333	0.034	0.056	0.002
37, 38	−0.100	−0.202	0.026	−0.072	0.017
39, 40	−0.016	0.017	0.040	−0.073	−0.003
46, 49	0.073	−0.010	0.006	−0.06	0.001
56, 57	−0.230	0.138	0.003	−0.055	−0.025
Mean	−0.08	−0.08	0.02	−0.05	0.00
SD	0.10	0.17	0.02	0.06	0.02

See text for details.

(δlog k) as a result of a maximum change in the column. As described in Appendix C, it is possible to estimate the relative importance of each term of Eq. (2) in determining values of k (Fig. 4a). The contribution of the hydrophobicity term $\eta'\mathbf{H}$ to changes in retention as a result of change in the column is largest, as expected from the nature of RPLC separation. The relative importance of the remaining terms of Eq. (2) is dependent on the particular solutes represented in Table 6 (and to a lesser extent the columns of Table 5); i.e., the results of Fig. 4a are expected to vary somewhat with the sample.

The contribution of the various terms of Eq. (2) to column selectivity is of greater interest. This can be approximated for solutes 1–67 as described in Appendix C and summarized in Fig. 4b. For adjacent or near-adjacent bands, the hydrophobicity term $\eta'\mathbf{H}$ is now the least important contribution to changes in α as a result of change in the column, because values of η' are highly correlated with k; i.e., adjacent bands tend to have similar values of η', and therefore there is little effect of η' on column selectivity (which is related to differences in each solute parameter for adjacent solutes). Fig. 4c provides a similar plot for the solutes of Table 7 (1a–87a). As in the case of Fig. 4a, the relative values of Fig. 4b and c depend on the solutes and columns chosen, hence explaining the observed differences in Figs. 4b vs. Fig. 4c. However, Fig. 4b and c confirm that, apart from hydrophobicity ($\eta'\mathbf{H}$), each of the remaining terms of Eq. (2) can be important in determining column selectivity.

5. Conclusions

The present study is an initial attempt to achieve a more complete understanding of the basis of column selectivity in RPLC. For the retention of 67 solutes of widely varied structure on 10 different C_{18} columns, we have found that Eq. (2) predicts values of log k with an accuracy of ±0.003 units (±0.7% in k, 1 SD), vs. a repeatability of experimental values of k of ±0.5%. Hence, we believe that Eq. (2) has captured all the chromatographically significant contributions to column selectivity for the present column types and solutes. In Eq. (2), k_{ref}, **H**, **S**, **A**, **B** and **C** are properties of the column (Table 5), and η', σ', β', α' and κ' are properties of the solute (Tables 6 and 7).

The last five terms of Eq. (2) are tentatively believed to correspond to various solute–column interactions illustrated in Fig. 1, with the corresponding column parameters **H**, **S**, etc., providing a reasonably complete characterization of RPLC column selectivity. Eq. (2) represents more than an order of magnitude improvement in predictive accuracy vs. the widely used ''solvation equation'' (Eq. (1); typical SD for predicted values of log k = 0.04–0.06). Similarly, the retention of 86 solutes on five C_8 or C_{18} columns from a previous study [5] exhibited agreement with Eq. (2) (excluding the last two terms, $\alpha'\mathbf{B}$ and $\kappa'\mathbf{C}$) that was only slightly inferior (±0.006 units in log k, 1 SD). These results suggest that Eq. (2) (with the possible inclusion of additional terms which may prove necessary for columns other than C_8 or C_{18}) will prove reliable

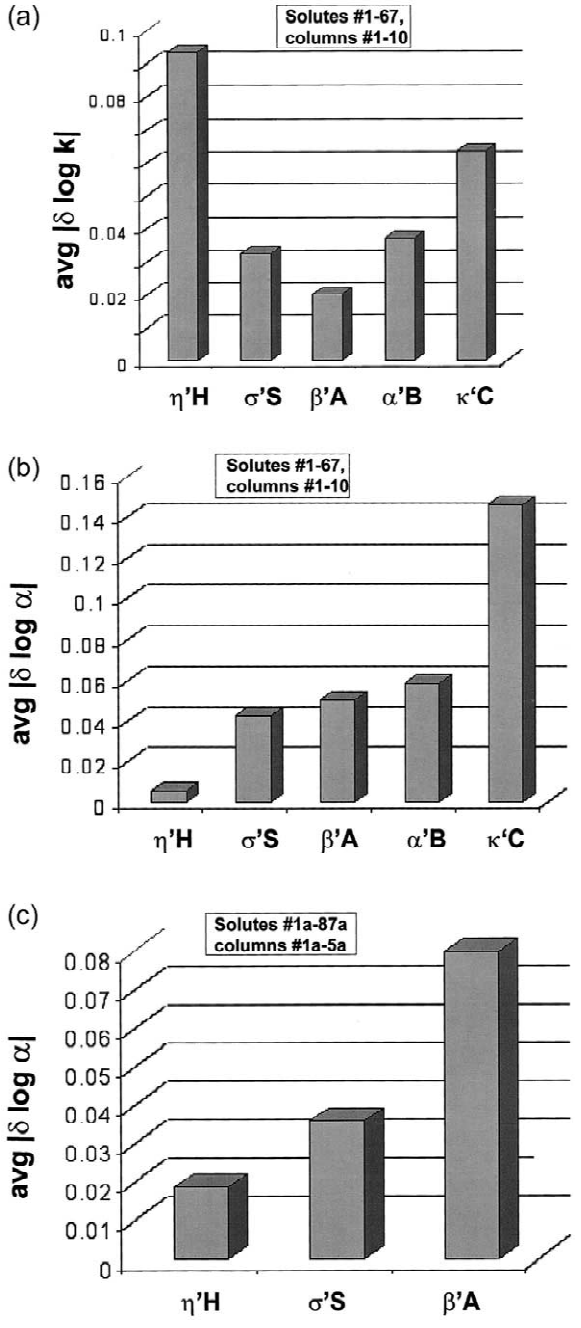

Fig. 4. Contribution of various terms of Eq. (2) to retention (a) and selectivity (b, c). Data for solutes 1–67 are plotted in (a, b); data for solutes 1a–87a are plotted in (c). See text for details.

and precise for the isocratic elution of most solutes and samples; a further test of this conclusion is provided in Part III [2] for an additional 23 solutes, including compounds of quite different structure vs. those in Table 1.

If Eq. (2) is applicable to any analyte or sample (which we believe to be the case), determination of the column parameters **H**, **S**, **A**, **B** and **C** should prove useful for two different goals of RPLC separations. First, columns with very different values of **H**, **S**, **A**, **B** and **C** should exhibit maximal differences in selectivity for different samples, thus facilitating the selection of chemically distinguishable columns for RPLC method development that are more likely to provide maximum changes in selectivity. Second, selectivity can vary from one batch to another of nominally equivalent columns. Measurements of **H**, **S**, **A**, **B** and **C** for columns from different batches should prove useful in determining whether different column batches are sufficiently similar in terms of retention to provide identical separations of any sample (i.e., separation factors α agreeing within $\pm$1–2%). At present, however, we are not suggesting that Eq. (2) be used for the quantitative prediction of separation as an aid for method development, because the determination of required values of the solute parameters η', σ', etc., for a ''new'' sample would require excessive experimental effort.

The present study emphasizes columns that are relatively similar (C_8 and C_{18} ''monomeric'' stationary phases). Thus, data reported here are of limited value for the purpose of choosing columns of very different selectivity, as can be inferred from data reported in Ref. [30] for a wider range of alkyl-silica columns. Our use of these particular columns in this preliminary study was intended for the identification and quantitation of column selectivity effects that are common to all RPLC columns; it appears that this goal has been attained. However, Eq. (2) may prove less accurate for columns with more interactive functionality (embedded polar groups, phenyl or cyano ligands, etc.) than is the case for alkyl ligands; additional terms in Eq. (2) may be needed that recognize solute–column interactions that are unimportant for simple alkyl-silica columns.

The application of Eq. (2) has so far been limited to a single set of experimental conditions, but in the following paper [2] column selectivity is studied as a

function of temperature and mobile phase composition. While no justification has been provided so far for our interpretation in Fig. 1 of the physico–chemical basis of the various terms of Eq. (2), Part III of this series [2] provides an initial attempt in this direction. Finally, it should be noted that the accuracy of Eq. (2) is such as to allow the measurement of the column-selectivity parameters **H**, **S**, etc., for other columns by means of data for only six test solutes (vs. the 67 solutes used in the present study). Thus, the routine characterization of column selectivity in this way should require only 1–2 h per column, as will be described in a later report.

6. Nomenclature

References to a defining equation, table or figure (e.g., Eq. III-4) indicates both the paper (e.g., Part III) and equation number (e.g., 4).

a	Difference in hydrogen bond basicity between the stationary and mobile phases (Eq. (1))
A	Column acidity (Eq. (2))
ACN	Acetonitrile
b	Difference in hydrogen bond acidity between the stationary and mobile phases (Eq. (1))
B	The organic solvent in an organic/buffer mobile phase; also, the isocratic temperature-coefficient of retention (Eq. II-4)
B''	Temperature coefficient of retention in gradient elution (Eq. II-5)
B	A column parameter which appears to measure stationary-phase hydrogen-bond basicity (Eq. (2))
C	A column parameter which measures the attraction of cationic solutes by the negatively charged stationary phase (Eq. (2))
C_1	Constant in Eq. (1)
F	Flow-rate (ml/min)
H	Column hydrophobicity (Eq. (2))
$\mathbf{H}_a$, $\mathbf{H}_b$	Values of **H** for columns a or b (Eq. (5))
IEC	Ion-exchange chromatography
k	Isocratic retention factor (same as capacity factor k') (Eq. (A.2))
k^*	Gradient retention factor (Eq. II-A.1)
k_{ref}	k for ethylbenzene as solute (Eq. (2)); used primarily as a correction for differences in column phase ratio (e.g., surface area)
L	Solute molecular length (see Fig. III-3)
MeOH	Methanol
n	Number of methylene groups (plus methyl) in a homologous alkyl group
N	Column plate number
r	Stationary phase excess molar refraction (Eq. (1)); also, correlation coefficient
RPLC	Reversed-phase liquid chromatography
R_2	Solute excess molar refraction (Eq. (1))
R_s	Resolution of two adjacent bands
s	Dipolarity/polarizability parameter for stationary vs. mobile phase (Eq. (1))
S	Empirical solute parameter from Eq. II-3; equal to $\mathrm{d}(\log k)/\mathrm{d}\phi$
S	Column steric selectivity (Eq. (2))
SD	Standard deviation
S.E.	Standard error
SD/S.E.	See discussion following Eq. (6) (also, Fig. 3)
t_D	System dwell time in gradient elution (min)
t_G	Gradient time (min)
THF	Tetrahydrofuran
t_0	Column dead time (min)
t_R	Retention time (min)
t_R(X)	Value of t_R for X as B-solvent; e.g., X = MeOH, ACN, THF
V_m	Column dead volume (ml)
V_x	Solute molar volume (Eq. (1))
x, y	Variables which define a least-squares fit (e.g., $y = 0.01 + 0.98x$ in Fig. 2a)

α	Separation factor for two adjacent bands; $\alpha = k/k_{ref}$ in Eq. (3)
α_a, α_b	Values of α for a given pair of solutes and column a or b
α'	A tentative measure of the solute hydrogen-bond acidity (Eq. (2))
α_{CH2}	Methylene increment, equal to the separation factor for adjacent homologs; here, the ratio of *k* for *n*-butyl- and *n*-propylbenzene.
$\alpha_{TBN/BaP}$	Ratio of *k*-values for tetrabenzonaphthalene and benzo[*a*]pyrene
α_2^H	Solute hydrogen-bond acidity in solution (Eq. (1))
β'	Solute hydrogen-bond basicity in RPLC (Eq. (2))
β_2	Solute hydrogen-bond basicity in solution (Eq. (1))
δ**H**, δ**S**, etc.	Change in parameters **H**, **S**, etc., as a result of some change in conditions
δ log *k*, δ log α	Change in log *k* or log α
δ log *k* (%B), δ log *k*(*T*), etc.	Change in log *k* as a result of some change in conditions (% B, temperature *T*, etc.)
$\delta\sigma'$	Difference between experimental values of σ' and values from Eq. III-6 (Eq. III-7)
δt_R	A change in gradient retention time t_R as a result of a change in conditions (Eqs. II-8 and 9)
Δ	Difference between experimental and calculated values of log α; see discussion of Eq. (6)
Δ(47), Δ(48), etc.	Value of Δ for solutes 47, 48, etc. (Fig. 3)
$\Delta\phi$	Change in mobile phase composition ϕ during a gradient run
η'	Solute hydrophobicity (Eq. (2))
η''	Approximate value of solute hydrophobicity η' (see Table 4 and related text)
η''_{avg}	Average of η'' values for a given solute and 10 columns (see Table 4 and related text)
ϕ	Volume-fraction of B-solvent in the mobile phase
κ'	A measure of the positive charge on the solute molecule (Eq. (2))
π_2^H	Dipolarity/polarizability parameter for solute (Eq. (1))
σ'	Solute steric selectivity (Eq. (2))
$\Sigma\alpha_2^H$	Solute hydrogen bond acidity (Eq. (1))
$\Sigma\beta_2$	Solute hydrogen bond basicity (Eq. (1))
ν	Free energy to create a cavity in the stationary phase (Eq. (1))

Acknowledgements

The present study (including following papers [1,2]) was supported in part by a Small Business Innovation Research (SBIR) grant from the National Institutes of Health (US Department of Health and Human Services). For all three papers, we are much indebted for the considerable advice, support and critical comments of Drs. J.J. Kirkland (formerly at Agilent Technologies), U.D. Neue (Waters Corp.), S.C. Rutan (Virginia Commonwealth University), D.V. McCalley (University of the West of England), and M.R. Euerby (Astra Zeneca R&D, Loughborough, UK), as well as the companies who provided the columns used in the present study (see Experimental).

Appendix A. Procedures used in present study to minimize experimental error

A1. Reproducibility of reported values of k

Values of log *k* reported in Table 3 and following papers [1,2] were calculated from values of retention time t_R using thiourea as a t_0-marker and taking system extra-column volume into account (see a following section). In each case, we used averages of triplicate measurements of t_R, carried out within a single working day. The average overall reproducibility of these log *k* values is ±0.002 (±0.5% in *k*, 1 SD).

Further comparisons of reproducibility were made by calculating methylene separation factors α_{Me} for

various homologs (*n*-butylbenzene/*n*-propylbenzene/ethylbenzene; 4-*n*-hexyl-/4-*n*-pentyl-/4-*n*-butylbenzoic acids; 4-*n*-heptyl-/4-*n*-hexyl-/4-*n*-pentylanilines). Values of α_{Me} within a given homologous series were constant within ±0.55% (1 SD), implying a reproducibility of the individual values of *k* equal to or better than $0.55/2^{1/2} = \pm 0.4\%$.

A2. Procedures

Our goal in the present study was the measurement of values of *k* with a repeatability of ±0.5% or better, using different operators and HPLC systems, with data collected over a period of several months. Some challenges to the automated collection of highly-reproducible data include (a) errors in online mixing of mobile phase, (b) errors in column temperature and (c) differences in extra-column system volume. For purposes of monitoring retention reproducibility, two system-suitability samples were also run several times each working day throughout the approximately 60 days during which data were collected. For neutral solutes 1–45, the system-suitability sample contained thiourea, phenol, 1-nitropropane, acetophenone, nitrobenzene, toluene and naphthalene. For the ''ionic'' solutes 46–67, the system-suitability sample consisted of thiourea, 2-nitrobenzoic acid, amitriptyline, nitrobenzene and 4-*n*-butylbenzoic acid. These two system-suitability samples were also used to correct raw data for any small fluctuations in experimental conditions as well as differences between different HPLC instruments used in this study (see a following section). Values of *k* reported here were determined by different operators using different equipment over a period of several months, hence requiring specific procedures to achieve the reproducibility reported.

A2.1. Errors in online mixing of mobile phase

For mobile phases (40–60%, ACN–water) prepared accurately by weight, the short-term precision of values of *k* for each HPLC system was equal to 0.1% (RSD). When measurements of *k* were repeated with on-line mixing, it was observed that there was a bias (i.e., error) of 0–0.2% in the apparent % B of the mobile phase, equivalent to errors in *k* as large as 2%. This bias could be corrected by adjusting the nominal % B entered into the system controller, once the bias was known for each system and each desired mobile phase composition (40, 50 or 60% ACN). In this connection, it is desirable to be able to adjust on-line mixing within ±0.01% B, since a rounding error of 0.05% corresponds to an average error in *k* of as much as 0.4%.

A2.2. Errors in column temperature

Peltier, block-heater and hot-air bath heaters were investigated. Typically, the nominal temperature setting was found to deviate from the actual column temperature for reasons discussed previously [31]. This temperature bias could be corrected as follows. First, a suitable length [31] of stainless steel tubing was added between the injector and the column to preheat the solvent before it enters the column. Second, an in-line thermocouple was placed after the column to measure the temperature of the mobile phase leaving the column and determine any temperature bias. The ability to reproduce this temperature was determined by cycling the oven between two temperatures and recording the temperature thermocouple after each equilibration. Short-term resetability was found not to be a problem. The thermocouple was then removed and the temperature setting adjusted to correct the original temperature bias.

A2.3. Differences in extra-column system volume

The extra-column volume V_{ec} of a HPLC system is usually small (40–110 μl in the present instance), and its influence on calculated values of *k* is typically ignored. However, resulting errors in *k* of 5–10% are possible, especially for larger values of V_{ec} that result from the addition of tubing to achieve thermal equilibration of mobile phase prior to the column. Let the extra-column volume and flow-rate be V_{ec} and *F*, respectively. Then a time $t_{ec} = V_{ec}/F$ will be added to both retention time t_R and column dead-time t_0. The retention factor *k* is defined by:

$$t_R = t_0(1 + k) \tag{A.1}$$

for $t_{ec} = 0$. If t'_R and t'_0 refer to apparent values of t_R and t_0, respectively, where $t'_R = t_R + t_{ec}$ and $t'_0 = t_0 + t_{ec}$, then it can be shown that:

$$k = (t'_R - t'_0)/(t'_0 - t_{ec}) \tag{A.2}$$

Values of t_{ec} were determined for each system used in the present study, by removing the column and determining the retention time of an injected solute. Corrected values of k were calculated via Eq. (A.2). Note that values of α (Eq. (3), equal to the ratio of two k values) used in the present data-reduction procedure will be unaffected by the extra-column volume or values of t_{ec}, when all data are collected for a single HPLC instrument (for which the value of t_{ec} does not change). Thus, for use of a single system, values of t_{ec} need not be measured for the accurate determination of values of α (as in Eq. (3)).

A2.4. System-suitability-sample correction

Long-term operation of a HPLC system can result in small, inadvertent changes of both the column temperature and mobile phase composition over time. The effect of these changes on solute retention can be partially corrected by normalizing resulting values of k in terms of k for some standard solute (nitrobenzene in the present case). This correction took the following form:

$$k^0 = k(k^0_{nb}/k_{nb}) \quad \text{(A.3)}$$

Here, k is the observed retention factor on a given day, k^0 is the corrected value of k, k_{nb} is the observed value of k for nitrobenzene on that day, and k^0_{nb} is the average value of k for nitrobenzene determined at the beginning of data collection. The use of Eq. (A.3) or similar procedures cannot be used to correct for large errors in % B or temperature, because changes in k with % B and T are not the same for all solutes [1].

Appendix B. Analysis of retention data of Ref. [5] in terms of Eq. (3)

An earlier study [32] as well as data reported in Ref. [1] suggest that changes in log k with T are similar for a given solute and different columns, at least for columns of related functionality as in Table 2. The temperature-dependence data of [1] were therefore used to adjust the data of Table 3 for a temperature of $T = 25\,°C$, hence allowing a comparison of these adjusted values with data from Ref. [5]. This adjustment for a change in T by 10 °C resulted in a fairly small change in values of log α (average change in log $\alpha = 0.007 \pm 0.021$, 1 SD).

The study of Tan et al. includes a column (StableBond C_{18}) which is nominally equivalent to the SB-100 column of Table 2 (but from a different lot). We next compared values of log α at 25 °C for the two StableBond C_{18} columns (3 and 1a in Table 5) and those solutes which were common to each study. Values of log α for the latter solutes and the two StableBond C_{18} columns agreed with SD = 0.018. The greater SD (0.018) vs. the agreement found with our data and Eq. (3) (SD = 0.004) can be attributed to (a) batch-to-batch differences in column selectivity, (b) errors introduced by our correction for differences in temperature (35 vs. 25 °C), and/or (c) possibly poorer repeatability of the data of Ref. [5].

The data of Ref. [5] do not include compounds which are likely to have large values of the solute parameters α' (acids) or κ' (strong bases), thereby precluding measurement of B or C for the columns of Ref. [5]. We therefore modified Eq. (3) to exclude terms *v* and *vi* for this sample set:

$$\log \alpha \approx \eta'\mathbf{H} + \sigma'\mathbf{S} + \beta'\mathbf{A} \quad \text{(B.1)}$$

Eq. (B.1) was first applied to our data (26 solutes common to both studies, 10 columns, all values of log α adjusted to 25 °C) to obtain values of the corresponding column and solute parameters, in similar fashion as in Table 4.

Solute parameters derived from our adjusted data at 25 °C were then used with the data of Ref. [5] and Eq. (B.1) to obtain values of **H**, **S** and **A** for the five columns of Ref. [5], as summarized in Table 5 (columns 1a–5a). The latter values of **H**, **S** and **A** for these columns permitted calculation of the corresponding solute parameters for the 86 solutes of Ref. [5], as summarized in Table 7.

Appendix C. Derivation of data of Fig. 4

C1. Fig. 4a

For the $\eta'\mathbf{H}$ term, $|\delta \log k|$ is equal to the average absolute value of η' times the maximum difference in **H** among the 10 columns. Thus, for the solutes of

Table 9
Effect of a change in column parameters **H**, **S**, etc., on log α for adjacent or near-adjacent bands

	$\eta'\mathbf{H}$	$\sigma'\mathbf{S}$	$\beta'\mathbf{A}$	$\alpha'\mathbf{B}$	$\kappa'\mathbf{C}$
Average absolute values of $(\eta'_2-\eta'_1)$, $(\sigma'_2-\sigma'_1)$, etc.	0.037	0.423	0.124	0.541	0.258
Maximum change in **H**, **S**, etc.	0.155	0.101	0.403	0.109	0.571
Average change in log α for maximum change in **H**, **S**, etc.	0.006	0.043	0.051	0.059	0.147

See Appendix C for details.

Table 6, the average value of $|\eta'|$ is 0.60, and the maximum difference in **H** is equal to 1.050 (column 2) minus 0.895 (column 5), or 0.155. The average change in $|\log k|$ due to differences in column **H** values is then 0.60×0.155 or 0.093 units. For each term of Eq. (2), Fig. 4a summarizes average changes in $|\log k|$ for a change in column (absolute values calculated in the same way as for $\eta'\mathbf{H}$).

C2. Fig. 4b

A change in the separation factor α for any adjacent band pair 1 and 2 is given as:

$$\begin{aligned}\delta\log\alpha_{12} = {} & (\eta'_2-\eta'_1)\delta\mathbf{H} + (\sigma'_2-\sigma'_1)\delta\mathbf{S} \\ & + (\beta'_2-\beta'_1)\delta\mathbf{A} + (\alpha'_2-\alpha'_1)\delta\mathbf{B} \\ & + (\kappa'_2-\kappa'_1)\delta\mathbf{C} \qquad \text{(C.1)}\end{aligned}$$

where η'_1 and η'_2 refer to values of η' for bands 1 and 2, respectively, and similarly for the remaining solute parameters of Eq. (C.1). The quantities $\delta\mathbf{H}$, $\delta\mathbf{S}$, etc., refer to a change in each column parameter as a result of a change in the column. If the solutes of Table 3 are arranged in order of increasing retention for the SB-100 column, Eq. (C.1) permits the calculation of average values of $(\eta'_2-\eta'_1)$, $(\sigma'_2-\sigma'_1)$, etc., for each adjacent band pair. Similar estimates of the latter quantity can be obtained from the standard deviation of correlations of each solute parameter vs. log *k*. The latter values are listed in the first row of data in Table 9. The second row of Table 9 contains values of the maximum change in each column parameter for the 10 columns of the present study. The third row of Table 9 is the product of values in the first two rows and is therefore equal to the average change in log α as a result of a maximum change in **H**, **S**, etc., for the present 10 columns. The latter values are plotted in Fig. 4b, based on solutes 1–67. The same procedure was used for solutes 1a–87a (Table 7), the results of which are plotted in Fig. 4c.

References

[1] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, P.W. Carr, J. Chromatogr. A (2002) 195.
[2] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, P.W. Carr, L.C. Sander, J. Chromatogr. A (2002) 217.
[3] P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, M.H. Abraham, Anal. Chem. 57 (1985) 2971.
[4] M.H. Abraham, Chem. Soc. Rev. 23 (1993) 660.
[5] L.C. Tan, P.W. Carr, M.H. Abraham, J. Chromatogr. A 752 (1996) 1.
[6] P. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, W. Melander, Cs. Horvath, Anal. Chem. 58 (1986) 2674.
[7] G. Rippel, Gy. Bacsur, A. Bede, A. Sandi, L. Szepesy, J. Liq. Chromatogr. Rel. Technol. 20 (1997) 1667.
[8] D. Bolliet, C.F. Poole, Chromatographia 46 (1997) 381.
[9] M.H. Abraham, M. Roses, C.F. Poole, S.K. Poole, J. Phys. Org. Chem. 10 (1997) 358.
[10] M. Roses, D. Bolliet, C.F. Poole, J. Chromatogr. A 829 (1998) 29.
[11] D. Bolliet, C.F. Poole, M. Roses, Anal. Chim. Acta 368 (1998) 129.
[12] J.A. Blackwell, P.W. Carr, J. High Resolut. Chromatogr. 21 (1998) 427.
[13] A. Sandi, L. Szepesy, J. Chromatogr. A 818 (1998) 1.
[14] A. Sandi, L. Szepesy, J. Chromatogr. A 845 (1999) 113.
[15] M. Reta, P.W. Carr, P.C. Sadek, S.C. Rutan, Anal. Chem. 71 (1999) 3484.
[16] M.A. Al-Haj, R. Kaliszan, A. Nasal, Anal. Chem. 71 (1999) 2976.
[17] M. Reta, P.W. Carr, P.C. Sadek, S.C. Rutan, Anal. Chem. 71 (1999) 3484.
[18] A. Sandi, L. Szepesy, J. Chromatogr. A 893 (2000) 215.
[19] M.A. Al-Haj, R. Kaliszan, B. Buszewski, J. Chromatogr. Sci. 39 (2001) 29.
[20] C.M. Du, K. Valko, C. Bevan, D. Reynolds, M.H. Abraham, J. Chromatogr. Sci. 38 (2000) 503.
[21] J. Li, B. Cai, J. Chromatogr. A 905 (2001) 35.
[22] M.A. Stadalius, J.S. Berus, L.R. Snyder, LC·GC Mag. 6 (1988) 494.
[23] J. Nawrocki, J. Chromatogr. A 779 (1997) 29.

[24] D.V. McCalley, LC·GC 17 (1999) 440.
[25] L.C. Sander, S.A. Wise, J. Chromatogr. 656 (1993) 335.
[26] L.C. Sander, M. Pursch, S.A. Wise, Anal. Chem. 71 (1999) 4821.
[27] L.C. Tan, P.W. Carr, J.M.J. Frechet, V. Smigol, Anal. Chem. 66 (1994) 450.
[28] L.A. Lopez, S.C. Rutan, J. Chromatogr. A, in press.
[29] J. Zhao, P.W. Carr, Anal. Chem. 71 (1999) 2623.
[30] U.D. Neue, B.A. Alden, T.H. Walter, J. Chromatogr. A 849 (1999) 101.
[31] R.G. Wolcott, J.W. Dolan, L.R. Snyder, S.R. Bakalyar, M.A. Arnold, J.A. Nichols, J. Chromatogr. A 869 (2000) 211.
[32] J.W. Dolan, L.R. Snyder, T. Blanc, L. Van Heukelem, J. Chromatogr. A 897 (2000) 37.