A Data Base for Characterizing Reversed-Phase Column Selectivity

L. R. Snyder, N. S. Wilson, and J. W. Dolan, LC Resources Inc.,
2930 Camino Diablo Suite 110, Walnut Creek, CA 94596

Abstract:

Column selectivity can be described by a general equation[1-3]

$$\log (k / k_{ref}) = \eta'H + \sigma'S + \beta'A + \alpha'B + \kappa'C \quad (1)$$

which relates relative retention factors (k / k_{ref}) to properties of the solute (η', σ', β', α', κ') and column (H, S, A, B, C). The selection of columns with sufficiently similar values of H, S, etc. allows the replacement of one column by another for a method procedure (second source). The choice of two or more columns having maximally different values of H, S, etc. provides maximum changes in selectivity for use in HPLC method development.

Values of H, S, etc. have been determined for about 50 reversed-phase columns as the beginning of a column-selectivity data base that will eventually include most commercial columns. The verification and significance of Eq. 1 will be discussed.

1, N. S. Wilson, M. D. Nelson, J. W. Dolan, L. R. Snyder, R. G. Wolcott and P. W. Carr, *J. Chromatogr. A,* in press.
2. N. S. Wilson, M. D. Nelson, J. W. Dolan, L. R. Snyder, and P. W. Carr, *J. Chromatogr. A,* in press.
3. N. S. Wilson, J. W. Dolan, L. R. Snyder, P. W. Carr and L. C. Sander, *J. Chromatogr. A,* in press.

A Column Selectivity Model:

The present study began with measurements of solute retention for 150 compounds of widely varying structure (neutrals, acids, bases, pharmaceuticals); 15 C8 and C18 columns (monomeric phases) were used from different sources. These 1300 values of k can be described (±1-2%) by the following equation:

$$\log(k) = \log(k_{EB}) + \eta'H + \sigma'S + \beta'A + \alpha'B + \kappa'C \qquad (1)$$

k_{EB} is the value of k for ethylbenzene, and the remaining terms have the following meaning:

Solute property	Column property	Significance
η'	H	hydrophobicity
σ'	S	steric interaction
β'	B	H-bond basicity
α'	A	H-bond acidity
κ'	C	electrostatic effect (solute or column charge)

Reversed-phase *retention* is determined primarily by hydrophobic interactions between solute and column (assumes temperature and mobile phase are fixed). However, column *selectivity* is determined mainly by other solute-column interactions; e.g., hydrogen bonding, shape selectivity, etc. Our goal is to characterize column selectivity by identifying and quantifying these various solute-column interactions.

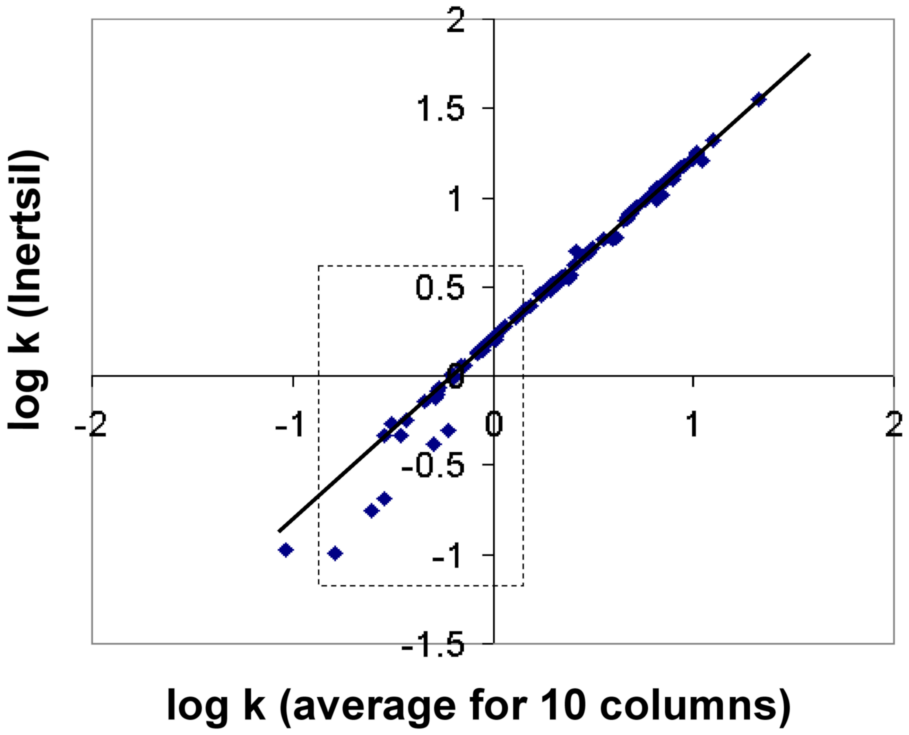
Neutral solutes	Neutral solutes (cont.)	Neutral solutes (cont.)	Basic compounds
1. benzene	24. 1,3 dihydroxy naphthalene	69. 1,3-dinitrobenzene	46. amitriptyline
2. toluene	25. eugenol	70. Nitrocyclohexane	47. diphenhydramine
3. ethylbenzene	26. danthron	71. Biphenyl	48. propranolol
4. p-xylene	27. n-propyl formate	72. 2-Nitrobiphenyl	49. nortriptyline
5. propylbenzene	28. methylbenzoate	73. 3-Nitrobiphenyl	50. prolintane
6. butylbenzene	29. benzonitrile	74. 2-Biphenylmethanol	51. 4-*n*-pentylaniline
7. naphthalene	30. coumarin	75, 2,2'-Biphenol	52. 4-*n*-hexylaniline
8. 4-chlorotoluene	31. acetophenone	76. 4,4'-Biphenol	53. 4-*n*-heptylaniline
9. p-dichlorobenzene	32. benzophenone	77. Diphenylbutyrolactone	54. N-ethylaniline
10. benzotrichloride	33. *cis*-chalcone	78. Fluorescamine	55. 2-phenylpyridine
11. bromobenzene	34. *trans*-chalcone	79. Camphorquinone	90. N-butylaniline
12. 1-nitropropane	35. *cis*-4-nitrochalcone	80. Ferrocene	
13. nitrobenzene	36. *trans*-4-nitrochalcone	81. N,N-diethylacetamide	**Acids**
14. 4-nitrotoluene	37. *cis*-4-methoxychalcone	82. 3-Nitrophenol	
15. 4-nitrobenzyl chloride	38. *trans*-4-methoxychalcone	83. 4-Nitrophenol	56. diclofenate acid
16. N-benzylformamide	39. prednisone	85. 2,5-Nitrophenol	57. mefenamic acid
17. anisole	40. hydrocortisone	87. Fisetin Hydrate	58. ketoprofen
18. benzyl alcohol	41. mephenytoin	88. Biochanin A	59. diflunisal
19. 3-phenyl propanol	42. Oxazepam		60. 4-*n*-butylbenzoic acid
20. 5- phenyl pentanol	43. flunitrazepam		61. 4-*n*-pentylbenzoic acid
21. phenol	44. 5,5 diphenylhydantoin		62. 4-*n*-hexylbenzoic acid
22. p-chlorophenol	45. N,N-dimethylacetamide		63. 3-cyanobenzoic acid
23. 2,3 dihydroxynaphthalene	68. 1,2-dinitrobenzene		64. 2-nitrobenzoic acid
			65. 3-nitrobenzoic acid
			66. 2,6-dimethylbenzoic acid
			67. 2-fluorobenzoic acid

Plus data for 60 additional compounds from Tan, Carr and Abraham, *J. Chromatogr. A,* 752 (1996) 1.

We can isolate these non-hydrophobic effects by plotting values of log k for one column vs. values for another. If only hydrophobic interactions are important, data points should fall close to a straight line. This is illustrated here for a plot of log k for Inertsil as column vs. average values of log k for 10 different columns. The slope of this plot is a measure of column hydrophobicity (= value of the column parameter H).

More important are deviations Δ from this. These are shown better in the expansion of this figure on the following page (left). Solutes for which data deviate from the best-fit solid line by <0.02 units in log k are designated as "non-ideal". Seven such solutes are noted here. Deviations of these "non-ideal" solutes are presumed to arise from non-hydrophobic solute-column interactions.

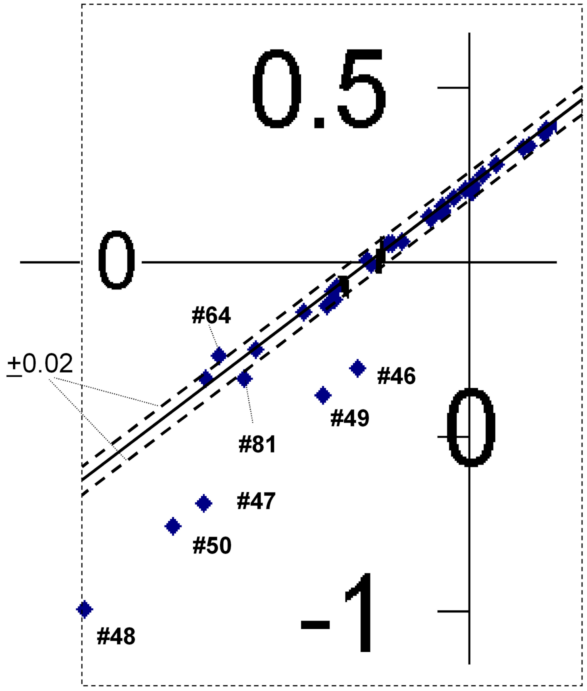
Selection of "non-ideal solutes": deviations Δ of log k values from "ideal solute" correlation line (—)



Slope of plot is the value of the column parameter **H** (equal to 0.989 in this example). Data enclosed in box are expanded and shown below

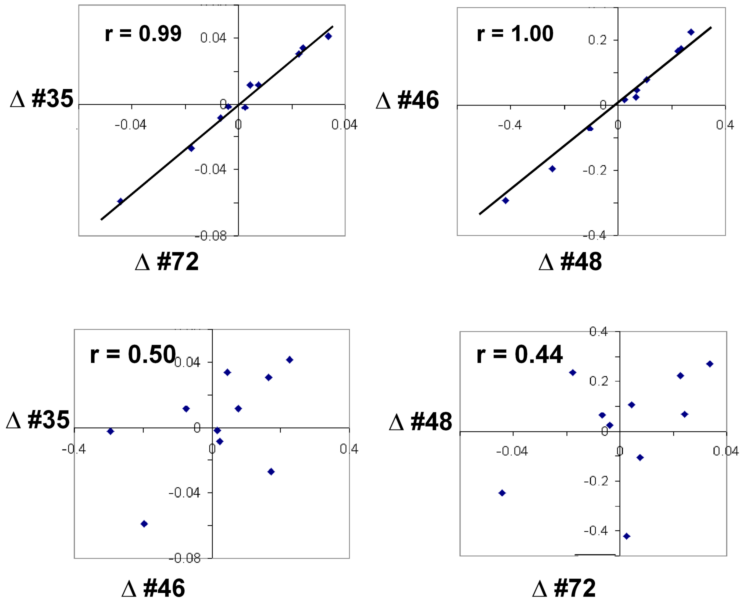
From plots as this for all 10 columns, 45 of the 90 solutes were identified as "non-ideal". The deviations for each solute (and different columns) was determined, and these deviations were cross-correlated to identify solute groups with similar "non-ideal" interactions as shown here (below, right).

"Non-ideal solutes": deviations > 0.02 log units



Deviations > 0.02 units are defined as "non-ideal"; these deviations are the result of non-hydrophobic solute-column interactions. Deviations (Δ) for two solutes (10 columns) which correlate with each other suggest a common non-hydrophobic interaction.

Correlation of deviations Δ for different pairs of solutes (10 columns each, see figures for structures)



It is seen that deviations Δ for solutes #35 and 72 are correlated for the 10 columns, as are Δ-values for solutes #46 and 48. This suggests that similar solute-column interactions are involved for solutes #35 and 72 (or #46 and 48), while different interactions are responsible for solutes #35 and 46 (or #48 and 72).

An Example: Intercorrelation of Solutes Exhibiting Steric Selectivity

	39	40	32	33	34	35	36	37	38	72	73	78	43	77	89	88	44	76
39	1	0.99	0.94	0.94	0.91	0.91	0.87	0.93	0.89	0.88	0.86	0.88	0.94	0.94	0.89	0.8	0.86	0.74
40	0.99	1	0.95	0.96	0.92	0.93	0.9	0.95	0.91	0.91	0.89	0.9	0.97	0.96	0.93	0.8	0.87	0.8
32	0.94	0.95	**1**	**1**	**0.99**	**0.99**	**0.97**	**1**	**0.99**	**0.99**	**0.97**	**0.99**	**0.99**	**0.99**	0.96	0.9	0.94	0.87
33	0.94	0.96	**1**	**1**	**0.98**	**1**	**0.98**	**1**	**0.99**	**0.99**	**0.97**	**0.99**	**0.99**	**1**	0.97	0.9	0.94	0.87
34	0.91	0.92	**0.99**	**0.98**	**1**	**0.98**	**0.94**	**0.98**	**0.97**	**0.99**	**0.97**	**0.99**	**0.97**	**0.99**	0.95	0.9	0.92	0.87
35	0.91	0.93	**0.99**	**1**	**0.98**	**1**	**0.99**	**1**	**1**	**0.99**	**0.98**	**0.99**	**0.98**	**0.99**	0.96	1	0.95	0.88
36	0.87	0.9	**0.97**	**0.98**	**0.94**	**0.99**	**1**	**0.98**	**0.99**	**0.97**	**0.97**	**0.97**	**0.95**	**0.96**	0.95	1	0.93	0.88
37	0.93	0.95	**1**	**1**	**0.98**	**1**	**0.98**	**1**	**0.99**	**0.99**	**0.97**	**0.99**	**0.98**	**0.99**	0.97	0.9	0.94	0.87
38	0.89	0.91	**0.99**	**0.99**	**0.97**	**1**	**0.99**	**0.99**	**1**	**0.99**	**0.99**	**0.99**	**0.97**	**0.98**	0.97	1	0.94	0.9
72	0.88	0.91	**0.99**	**0.99**	**0.99**	**0.99**	**0.97**	**0.99**	**0.99**	**1**	**0.99**	**1**	**0.98**	**0.99**	0.96	1	0.94	0.91
73	0.86	0.89	**0.97**	**0.97**	**0.97**	**0.98**	**0.97**	**0.97**	**0.99**	**0.99**	**1**	**0.99**	**0.96**	**0.97**	0.97	1	0.91	0.93
78	0.88	0.9	**0.99**	**0.99**	**0.99**	**0.99**	**0.97**	**0.99**	**0.99**	**1**	**0.99**	**1**	**0.97**	**0.98**	0.95	1	0.93	0.9
43	0.94	0.97	**0.99**	**0.99**	**0.97**	**0.98**	**0.95**	**0.98**	**0.97**	**0.98**	**0.96**	**0.97**	**1**	**0.99**	0.98	0.9	0.92	0.91
77	0.94	0.96	**0.99**	**1**	**0.99**	**0.99**	**0.96**	**0.99**	**0.98**	**0.99**	**0.97**	**0.98**	**0.99**	**1**	0.98	0.9	0.94	0.89
89	0.89	0.93	0.96	0.97	0.95	0.96	0.95	0.97	0.97	0.96	0.97	0.95	0.98	0.98	1	1	0.89	0.93
88	0.79	0.83	0.93	0.93	0.92	0.95	0.96	0.94	0.97	0.97	0.98	0.96	0.93	0.93	0.95	1	0.88	0.95
44	0.86	0.87	0.94	0.94	0.92	0.95	0.93	0.94	0.94	0.94	0.91	0.93	0.92	0.94	0.89	0.9	1	0.84
76	0.74	0.8	0.87	0.87	0.87	0.88	0.88	0.87	0.9	0.91	0.93	0.9	0.91	0.89	0.93	1	0.84	1

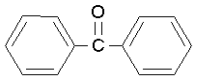
Selected (bolded) solutes: r(avg) = 0.985 All of above solutes: r(avg) = 0.950

All other solute: r(avg) = 0.778

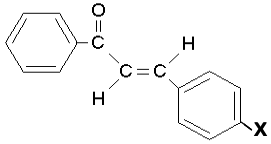
The bolded data correspond to strong correlations between solutes #32-38, 72, 73, 78, 43 and 77. Presumably this reflects a common contribution to column selectivity. The nature of this solute-column interaction can be inferred from solute molecular structure and column properties such as ligand length and concentration, pore diameter and the presence or absence of end-capping.

Solutes Which Exhibit Steric Selectivity (σ'S)

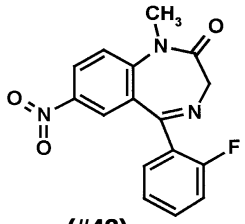
Major (Select)



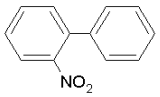
(#32)



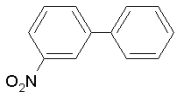
(#33-38)



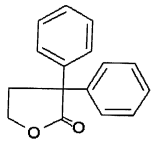
(#43)



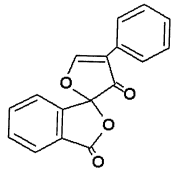
(#72)



(#73)

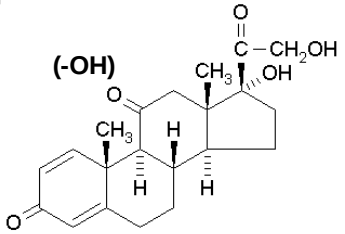


(#77)

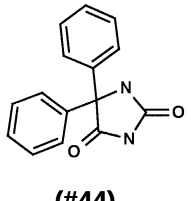


(#78)

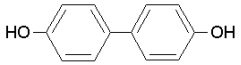
Minor



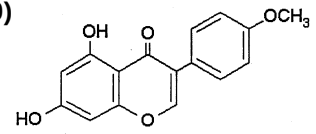
(#39, 40)



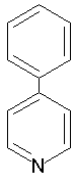
(#44)



(#76)



(#88)



(#89)

Inter-correlation of Other Solutes

Cationic solutes: (κ'**C)**

Selected solutes (#46-50): r(avg) = 0.998

#51-55: r(avg) = 0.93

remaining solutes: r(avg) = 0.80

H-bond Acidic solutes: (α'**B)**

Selected solutes (#56-58, 60-62, 66,67): r(avg) = 0.90

#20, 42, 51-53: r(avg) = 0.86

remaining solutes: r(avg) = 0.71

H-bond basic solutes: (β'**A)**

Selected solutes (#16, 45, 81): r(avg) = 0.94

remaining solutes: r(avg) = 0.77

These solute groupings suggest 4 solute-column interactions plus hydrophobicity (η'**H)**.

Steric Selectivity (σ'S)

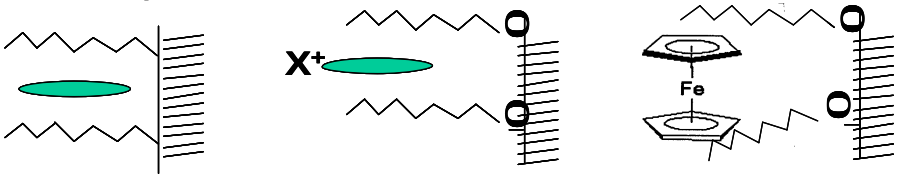
Solute σ' values

- Values of σ' for neutral solutes increase with molecular length: y = -0.90 + 0.155 x, r = 0.89, SE 0.20
- σ' for acids and bases are decreased by 0.6±0.3 units
- "bulky" molecules (#77-80) have larger σ' values (+0.2-0.8 units)

Column S values

- **S** decreases (more difficult solute penetration into the stationary phase) for C_{18} *vs.* C_8, higher ligand concentrations, smaller pore diameter
- **S** does *not* correlate with measures of "shape selectivity"; e.g., values of $\alpha_{TBN/BaP}$

Interpretation



long molecule: SEC-like exclusion

acid or base: partial exclusion

"bulky" molecule: difficult "fit"

H-bonding of Acceptor Solutes (β'A)

Solute β' values

- correlate with solution H-bonding basicity values β_2:

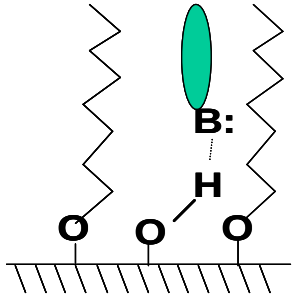
  $\beta' = -0.47 + 1.34\ \beta_2$; r = 0.92 (aliphatic solutes)

  $\beta' = -0.06 + 0.24\ \beta_2$; r = 0.67 (aromatic solutes)

- decrease for increased intramolecular steric hindrance

Column A values

- decrease sharply for end-capped columns
- decrease for increased ligand concentration and narrower pores



Interpretation:

this interaction is clearly due to H-bonding between non-ionized silanols and acceptor groups (**B:**) in the solute molecule

H-bonding of Donor Solutes (α'B)

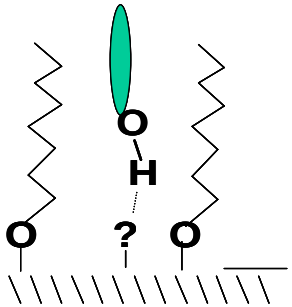
Solute α' values

- increase with Bronsted acid strength ($-pK_a$):

neutrals	α' = 0.0
alcohols (R-OH)	α' = 0.1
phenols	α' = 0.2
R-COOH	
weak	α' = 0.9
strong	α' = 2.3

Column B values

- are unaffected by end-capping
- are unaffected by ligand length or concentration, or pore diameter
- **B** = -0.70 – 0.70 **H**; r = 0.95



Interpretation:

- what these donor solutes are interacting with in the stationary phase is at present uncertain; one possibility is water molecules that are sorbed in the stationary phase

Ion Interaction (κ'C)

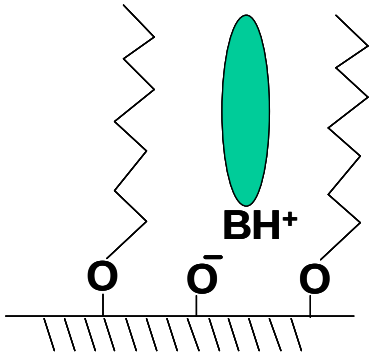
Solute κ' values

- Vary with solute molecule charge; κ' is larger for molecules with a positive charge, smaller for molecules with a negative charge:

neutrals	κ'(avg) = 0.0
weak bases	κ'(avg) = 0.1
strong bases	κ'(avg) = 1.0
weak acids	κ'(avg) = -0.1
strong acids	κ'(avg) = -0.3

Column C values

- decrease sharply with end-capping
- increase with mobile phase pH, especially for non-end-capped columns



Interpretation:

(variable) negative charge on stationary phase due to ionized silanols causes preferential retention of cations; strong (fully ionized) bases primarily affected (e.g., solutes #46-50)

To summarize:

The accuracy of Eq. 1 suggests that all contributions to C8 or C18 column selectivity are accounted for by 5 column parameters: H, S, A, B, C.

Slope values define values of H for each column.

Average deviations (△) for each group of selected solutes provide values of the remaining 4 column parameters (S, A, B, C); values of H, S, A, B, C can also be obtained from just 6 test solutes (1 hr total test-time per column): ethylbenzene, acetophenone (H), nitrochalcone (S), diethylacetamide (A), butylbenzoic acid (B) and amitriptyline (C).

This approach to the characterization of column selectivity has now been extended to about 50 different columns of various types. We are now assembling a data base (i.e., values of H, S, A, B, C) for several hundred reversed-phase columns.

How are values of H, S, A, B, C useful?

- to monitor batch-to-batch column uniformity
- to select alternate columns with "equivalent" selectivity
  - to design separations that can be carried out with either of two columns (same conditions)
- to select columns of very different selectivity
  - for a change in selectivity during method development
  - to develop an orthogonal separation

What about other column types?

We have recently collected retention data for 44 solutes and 8 columns with an embedded polar group (EPG). These columns from different suppliers have amide, urea, or carbamate polar groups in each column. The agreement of these retention data with Eq.1 was poorer (+10%) than observed before for columns without a polar group (±1-2%), presumably because of the diversity of polar groups studied. By recognizing a difference in the interaction of phenolic (-OH) and carboxylic acid (-COOH) groups with these different columns, it was possible to improve the accuracy of predicted values of log k to ±2%. This required an expansion of Eq. 1 to:

$$\log k = \log k_{EB} + \eta'H + \sigma'S + \beta'A + \alpha'_1B_1 + \alpha'_2B_2 + \kappa'C \qquad (2)$$

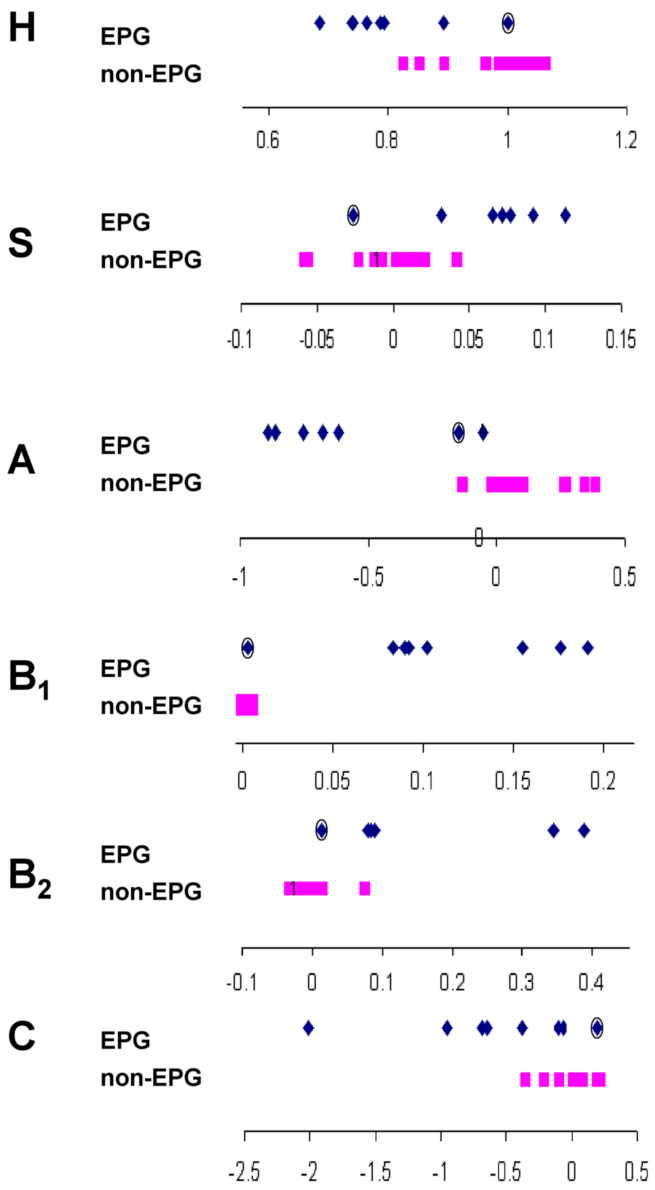
That is, the interaction of H-bond donors with EPG columns requires two terms: α'_1B_1 for phenols, and α'_2B_2 for carboxylic acids. For columns without an embedded polar group, $\alpha'_1B_1 = 0$, and $\alpha'B = \alpha'_2B_2$. Solutes with alcohol groups ($-CH_2-OH$) exhibit negligible hydrogen bonding with EPG columns.

A comparison of values of the column parameters H, S, etc. of some EPG and non-EPG columns is shown here.

Columns with an embedded polar group are generally:

- **more polar and less hydrophobic (smaller H)**
- **more easily penetrated by bulky solute molecules (larger S)**
- **less acidic toward solutes that are H-bond acceptors (smaller A)**
- **more basic toward solutes that are H-bond donors (larger B)**
- **less retentive of protonated bases (smaller C)**

This is largely what we would expect in view of the H-bond basicity of polar embedded groups such as amide, urea, and carbamate.



Summary and Future Plans:

The selectivity of reversed-phase columns can be fully characterized by 6 measurable parameters: H, S, A, B_1, B_2 and C

This allows the selection of columns that are either similar or different in terms of selectivity

- similar columns can be used as a back-up for columns used in a routine test procedure
- different columns allow large changes in selectivity for use in method development

Measurements of H, S, etc. can be carried out for each column with little effort (values of k needed for 6 or 7 test compounds only)

Work is underway to measure values of H, S, etc. for several hundred reversed-phase columns.

We are looking for collaborations with column manufacturers who wish to have their columns included in this column selectivity data base

Acknowledgements:

This work has been supported in part by an SBIR grant from NIGMS

LC RESOURCES inc.

2930 Camino Diablo #110, Walnut Creek, CA 94596 (USA)
http://www.lcresources.com
fax: (925) 930-9136