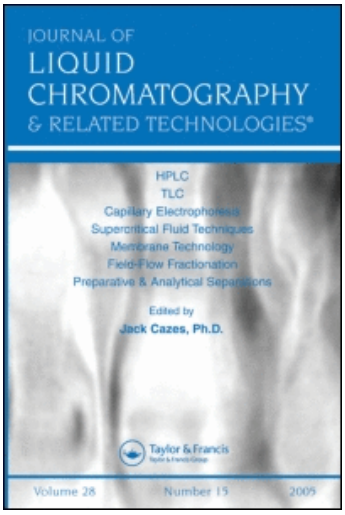


Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

RP-HPLC Retention Data for Measuring Structural Similarity of Compounds for QSAR Studies

Klára Valkó[a]
[a] Central Research Institute for Chemistry Hungarian Academy of Sciences, Budapest, Hungary

To cite this Article Valkó, Klára(1987) 'RP-HPLC Retention Data for Measuring Structural Similarity of Compounds for QSAR Studies', Journal of Liquid Chromatography & Related Technologies, 10: 8, 1663 — 1686
To link to this Article: DOI: 10.1080/01483918708066793
URL: http://dx.doi.org/10.1080/01483918708066793

JOURNAL OF LIQUID CHROMATOGRAPHY, 10(8&9), 1663-1686 (1987)

RP-HPLC RETENTION DATA FOR MEASURING STRUCTURAL SIMILARITY OF COMPOUNDS FOR QSAR STUDIES

Klára Valkó

Central Research Institute for Chemistry
Hungarian Academy of Sciences
H-1525
P. O. Box 17
Budapest, Hungary

ABSTRACT

A quantitative measure has been developed for characterizing the structural similarity of compounds on the basis of their reversed-phase liquid chromatographic retention data. Reversed-phase retention data (expressed by log k' values) of 59 compounds of various types (aniline, phenol, morphine, barbiturate, xantine, sulfonamide, steroid, benzoic acid derivatives) have been investigated as a function of organic modifier concentration (OP%) in the mobile phase. The effect of acetonitrile and methanol concentration in the mobile phase on the log k' values was measured. High correlation was found between the slope and the intercept values of the log k' vs OP% straight lines for only structurally related compounds. In view of these results, the correlation coefficient between the slope and the intercept values of the linear section of the log k' vs OP% straight lines is suggested as a measure of structural similarity of compounds with regard to their partition behaviour in a reversed-phase chromatographic system.

INTRODUCTION

Drug molecules acting at the same receptor sites should be structurally related although their structural formulae are not obviously similar. Structural similarity is important for a good correlation of the biological activity data and physicochemical parameters used in quantitative structure - activity relationship (QSAR) studies according to the Hansch's (1) approach. Structural similarity in QSAR investigations means that each compound in the series shows similar partition behaviour in the biological system and each compound has similar specific activity, they differ from each other only in their affinity towards the receptor.

Retention properties of compounds in the reversed-phase chromatography depend on the partition properties between the polar mobile and the non-polar stationary phases as it has been derived by Snyder and Kirkland (2). This phenomenon was utilized by great number of authors for measuring partition coefficients by reversed-phase liquid chromatography presenting good correlation between the partition coefficients and log k' values. Results were reviewed by Kaliszan (3), Carney (4), and Eadsforth and Moser (7).

Several authors observed outlier compounds from the log k' vs log P (P stands for octanol/water

partition coefficient) relationships, or they found different regression parameters for different classes of compounds (5-9). Others (10,11) published different correlation coefficients for log k' vs log P relationships, when the log k' values were measured in different organic modifier concentrations. Harnisch et. al (12), Snyder et. al. (13) observed the intersections of the log k' vs organic modifier concentration straight lines, for different compounds which suggests that different order of the log k' values can be obtained using different organic modifier concentrations in the mobile phase. All of these findings suggest, that structurally unrelated compounds can behave as an outlier from the relationships.

Although some results published in the literature (14-16) pointed out that the log k' values are not always lineraly related to the organic modifier concentration, we assumed the linear relation of log k' to the OP% values at least within a certain range. According to the results of Hurtubise et. al. (17), the slope and the intercept of the log k' vs. OP% straight lines show high correlation, it also suggests that these straight lines do not cross each other randomly but in one point which is outside of the practically used organic phase concentration range. The same phenomenon was described by Shats et. al. (18), although

they suggest that the intersection points of the log k' vs. OP% straight lines are in different position in the plot for structurally different compound set.

In this paper a number of retention data (log k' values) of various classes of compounds were investigated as a function of organic phase concentration (OP%) in the mobile phase. The correlation coefficients of the slope and the intercept values for the linear section of log k' vs OP% curves were calculated for various subset of compounds.

MATERIALS AND METHODS

The retention time measurements of the model compounds were carried out on a Dimseil C-18 (Chromatronix Inc., Palo Alto, California, USA) column with the dimensions of 250 x 4.6 mm and 10 μm particle size. The mobile phase was pumped by Liquopump Model 312 (LaborMIM, Budapest, Hungary). Injections were made by Rheodyne Model 7010 (Cotati, California, USA) injector with 20 μl sample loop. ISCO Absorbance Monitor Model 226 (ISCO Inc., Lincoln, Nebraska, USA) with 254 nm source screen served as the detector. For detecting the morphine derivatives the 280 nm source screen was used.

Retention time measurements and the data processing were carried out on an Apple II microcomputer (Apple Computer Inc., Cupertino,

California, USA). The model compounds were obtained from the Drug Store of Semmelweis University, Budapest exept of the morphine derivatives which were a kind gift of Dr. Tamás Friedmann (Semmelweis University, Medical School, Department of Pharmacology), and exept for anilines and phenols which were a kind gift of Dr. Tibor Cserháti (Research Institute for Plant Protection, Hungarian Academy of Sci., Budapest). The RP-HPLC conditions for model compounds listed in Table 1 were the following: the mobile phase composition ranged from 5 to 90% (v/v) acetonitrile (HPLC grade, Merck, Darmstatt, West Germany) with 0.05 M phosphate buffer (Reanal, Budapest, Hungary) (pH 4.6) with a flow-rate of 1.00 ml/min. For the benzoic acid and salicylic acid derivatives the pH was adjusted to 2 by 85% phosphoric acid in order to improve the peak symmetry and get the compounds in the non-ionized form. The mobile phase for the morphine derivatives listed in Table 2 contained 0.005 M sodium-butylsulfonate (Aldrich, Milwauke, Wisc., USA) ion pair and the pH was adjusted to 2 in all cases. The acetonitrile concentration was ranged from 20 to 75 %(v/v). The flow-rate was again 1.00 ml/min.

The mobile phase composition for the retention time measurements of the compounds listed in Table 3 ranged from 5 to 80 % methanol and phosphate buffer

Table 1. The slope and the intercept values of the log k' vs OP% straight lines measured in various compositions of acetonitrile and phosphate buffer.

	Compound	Slope	Intercept
1	Resorcine	-0.0150	0.259
2	Sulfadimidine	-0.0280	0.854
3	Sulfamethoxypyridazine	-0.0285	0.892
4	Barbital	-0.0403	1.063
5	Phenobarbital	-0.0319	1.341
6	Chloramphenicol	-0.0414	1.625
7	Salicylamide	-0.0255	0.871
8	Phenacetin	-0.0255	1.002
9	Vanillin	-0.0244	0.866
10	Benzaldehyde	-0.0303	1.575
11	Acetanilide	-0.0270	1.021
12	Nicotinamide	-0.0382	0.250
13	Benzoic acid	-0.0284	1.252
14	Salicylic acid	-0.0301	1.425
15	Acetyl salicylic acid	-0.0272	1.077
16	Coffeine	-0.0299	0.552
17	Hydrochlorothiazide	-0.0456	0.887
18	Cortexolone	-0.0138	0.757
19	Dexamethasone	-0.0139	0.568
20	11-deoxycorticosterone	-0.0147	1.120
21	Sulfaguanidine	-0.0272	0.012
22	Isoniazide	-0.0382	0.060
23	Methylsalycilate	-0.0244	1.727
24	Hydrocortisone	-0.0129	0.435
25	Progesterone	-0.0192	1.831
26	Testosterone	-0.0143	1.085
27	Prednisolon	-0.0097	0.251
28	Triamcinolone	-0.0150	0.727
29	Cortisone	-0.0127	0.432

Table 2. The slope and the intercept values of the log k' vs OP% straight lines for the 12 morfine derivatives obtained by ranging the acetonitrile concentration with buffer (pH 2) and 0.005 M sodium butylsulfonate.

	Compound	Slope	Intercept
1	Azidomorphine	-0.0311	0.975
2	Azidocodeine	-0.0064	0.499
3	N-cyclopropylmethylazido-morphine	-0.0056	0.358
4	Azidoethylmorphine	-0.0088	0.862
5	N-Phenylethylazidoethyl-morphine	-0.0250	2.164
6	N-Phenylethylazido morphine	-0.0150	1.124
7	Acethylazidomorphine	-0.0080	0.594
8	Norazidoethylmorphine	-0.0106	0.880
9	N-cyclopropyl-azidoethyl-morphine	-0.0194	1.644
10	Norazidomorphine	-0.0271	0.851
11	Normorphine	-0.0277	0.344
12	Morphine	-0.0298	0.467

Table 3. The slope and the intercept values of the log k' vs OP% straight lines obtained by various methanol concentrations in the mobile phase.

	Compound	Slope	Intercept
1	Nicotinamide	-0.0379	0.825
2	Coffeine	-0.0334	1.451
3	Resorcine	-0.0255	0.622
4	Hydrochlorothiazide	-0.0342	0.882
5	Barbital	-0.0246	0.978
6	Sulfadimidine	-0.0307	1.245
7	Sulfamethoxypyridazine	-0.0310	1.264
8	Salicylamid	-0.0274	1.280
9	Vanillin	-0.0289	1.376
10	Acetanilide	-0.0244	1.304
11	Chloramphenicol	-0.0275	1.427
12	Phenobarbital	-0.0303	1.610
13	Salicylic acid	-0.0299	1.792
14	Phenacetine	-0.0351	2.065
15	Benzoic acid	-0.0321	1.804
16	Methyl salicylate	-0.0214	1.748
17	O-nitroaniline	-0.0248	1.507
18	M-nitroaniline	-0.0237	1.253
19	P-nitroaniline	-0.0214	0.980
20	2,4-dinitroaniline	-0.0267	1.715
21	2,6-dinitroaniline	-0.0252	1.645
22	2,4,6-trinitroaniline	-0.0264	1.676
23	2-Cl,4-nitroaniline	-0.0280	1.879
24	4-Cl,3-nitroaniline	-0.0289	1.814
25	2,6-diCl,4-nitroaniline	-0.0329	2.568
26	p-nitrophenol	-0.0258	1.486
27	2,4-dinitrophenol	-0.0290	1.789
28	2,5-dinitrophenol	-0.0285	1.803
29	2,6-dinitrophenol	-0.0517	2.984
30	2,4,6-trinitrophenol	-0.0490	2.487
31	3,5-dinitro,4-cianophenol	-0.0421	2.417
32	3-nitro,4-ciano,5-Cl-phenol	-0.0325	2.332
33	3-nitro,4-ciano,5-Br-phenol	-0.0424	2.684
34	3-nitro,4-ciano,5-F-phenol	-0.0285	1.592

0.05 M (pH 4.6). For the phenol derivatives the pH was adjusted to 2 for improving the peak shape and to avoid ionization of these compounds. The flow rate was 1.00 ml/min.

The model compounds were dissolved always in the actual mobile phase to get 0.1% solution and 20 ul was injected. The retention volume of the 1% $NaNO_3$ solution was regarded as the void volume. The retention time measurements were carried out in that mobile phase composition range where the log k' values varied between -0.5 and 2 and the log k' vs organic phase concentration relationship was linear. The slope and the intercept values of the linear section of the log k' vs OP% relationship are listed in the Tables 1-3.

RESULTS AND DISCUSSIONS

The plots of the slope and the intercept values listed in Table 1-3 can be seen in Fig. 1-3. The data obtained by using the same HPLC conditions are shown in the same figure. As it is noticable the position of the data points do not fit to a curve, the distribution of them seems to be random. This finding was really unusual. For example Hurtubise et al. (17) presented the same plot of the slope and the intercept values for 68 compounds and they fit to a straight line. In order to find the explanation of the contradiction several

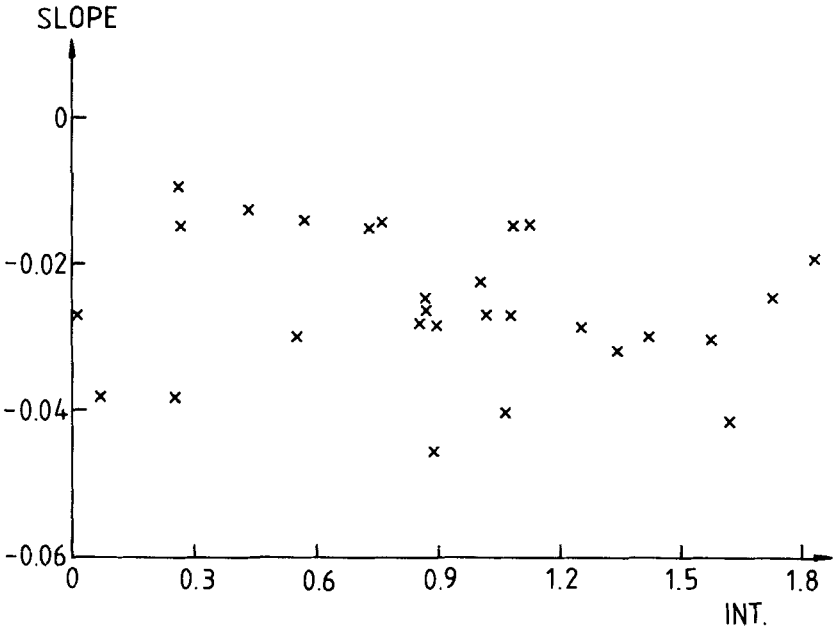


Fig. 1.

The plot of the slope and the intercept values for compounds listed in Table 1.

The number of compounds: 29

Multiple correlation coefficient (R): 0.194

Fisher-test value: 1.06

Significance of the regression (F-ratio): 31.46%

Standard error of the estimate: 0.0098

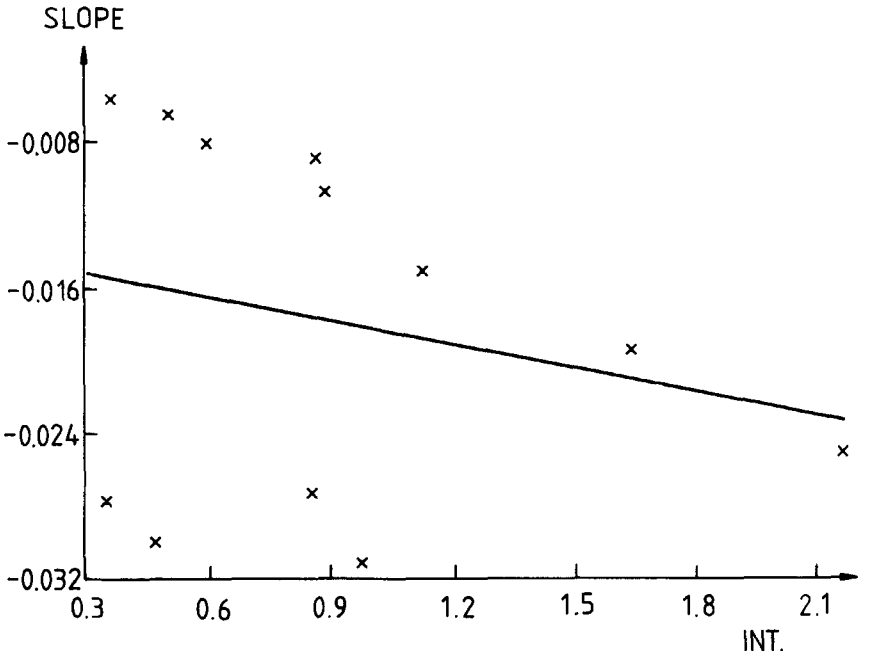


Fig. 2.

The plot of the slope and the intercept values for the morphine derivatives listed in Table 2.

The number of compounds: 12

Multiple correlation coefficient (R): 0.240

Fisher-test value: 0.609

Significance of the regression (F-ratio): 54.15%

Standard error of the estimate: 0.0101

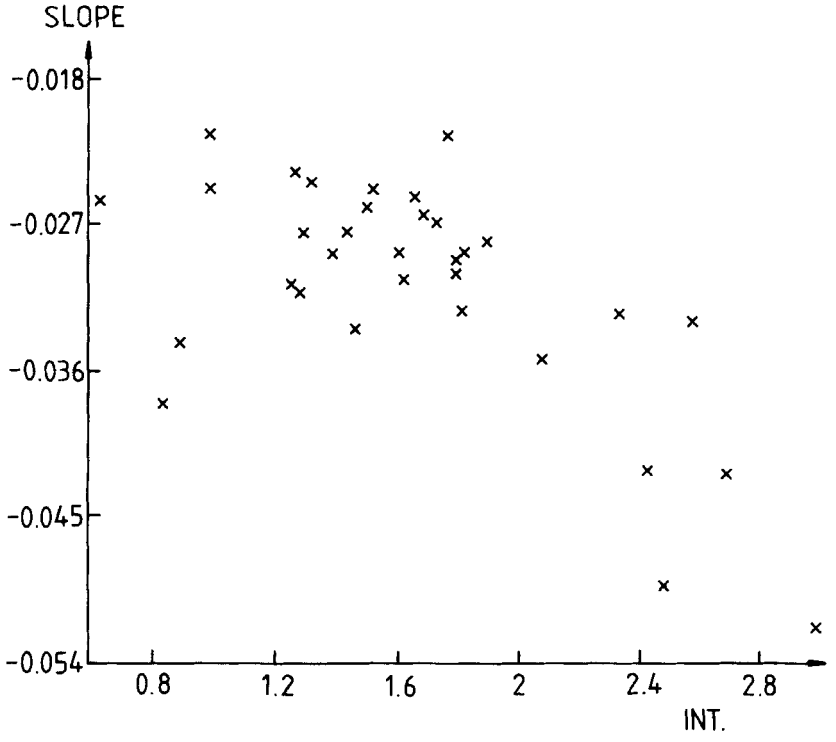


Fig. 3. The plot of the slope and the intercept values obtained by varying the methanol concentration in the mobile phase for compounds listed in Table 3.

The number of compounds: 34

Multiple correlation coefficient (R): 0.642

Fisher-test value: 22.474

Significance of the regression (F-ratio): 0.006%

Standard error of the estimate: 0.005

subsets of the compounds were selected and the correlation of the slope to the intercept values was investigated.

In Fig. 4 the slope and the intercept values for the 9 steroid compounds listed in Table 1 with the serial numbers of 18-20 and 24-29 can be seen. The mathematical statistical parameters of the correlation are listed under the figure. In Fig. 5. the data for 7 substituted phenol and benzoic acid derivatives can be seen (the serial number of the compounds in Table 1 are the following: 1, 7, 9, 11, 13, 14 and 15).

The slope and the intercept values showed again high correlation. The compounds are common in having more or less acidic character and the benzene ring. Fig. 6. shows the good correlation of the slope and the intercept values for the subset of the morphine derivatives containing sterically big substituent on the basic nitrogen or the acidic hydroxyl group on the benzene ring is substituted. The serial numbers of compounds in Table 2 data of which were used in this correlation are the following: 2, 3, 4, 5, 6, 7, 8, 9.

Similarly, structurally related subset of compounds could be found in case of the data obtained by varying the methanol concentration in the mobile phase. Plotting the slope and the intercept values they fit again to a straight line as it is shown in Fig. 7

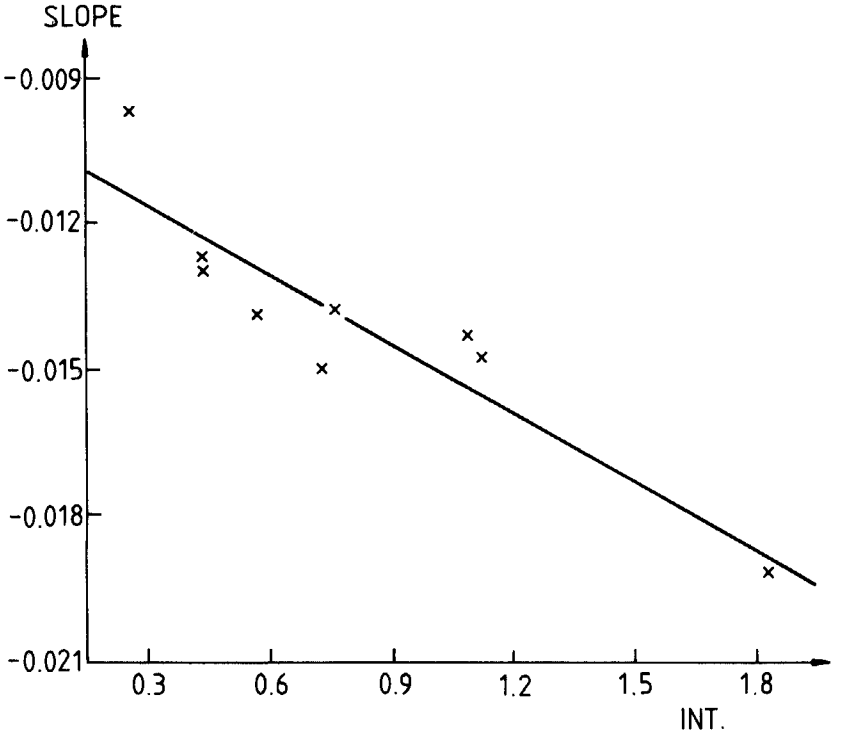


Fig. 4.

The plot of the slope and the intercept values for 9 steroid derivatives listed in Table 1.

The number of compounds: 9

Multiple correlation coefficient (R): 0.915

Fisher-test value: 36.00

Significance of the regression (F-ratio): 0.067%

Standard error of the extimate: 0.001

Regression parameters: -0.00473, -0.01024

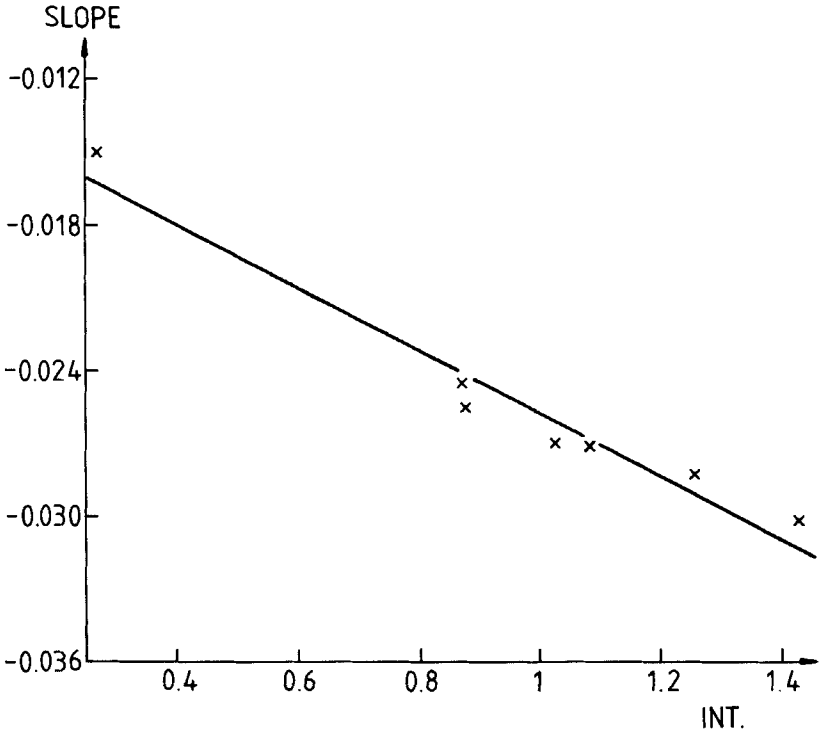


Fig. 5.

The plot of the slope and the intercept values for phenol and benzoic acid derivatives listed in Table 1.

The number of compounds: 7

Multiple correlation coefficient (R): 0.978

Fisher-test value: 109.98

Significance of the regression (F-ratio): 0.033%

Standard error of the estimate: 0.0011

Regression parameters: -0.0130, -0.0128

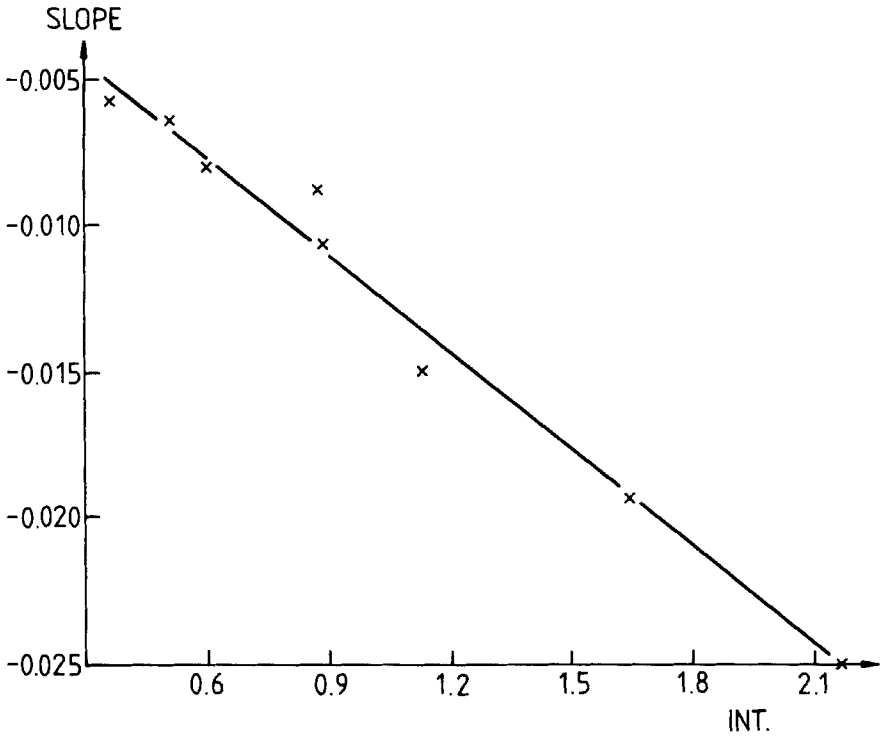


Fig. 6.

The plot of the slope and the intercept values for the structurally related subset of morphine derivatives listed in Table 2.

The number of compounds: 8

Multiple correlation coefficient (R): 0.991

Fisher-test value: 324.69

Sifgnificance of the regression (F-ratio): 0.0016%

Standard error of the estimate: 0.001

Regression parameters: -0.0111, -0.00107

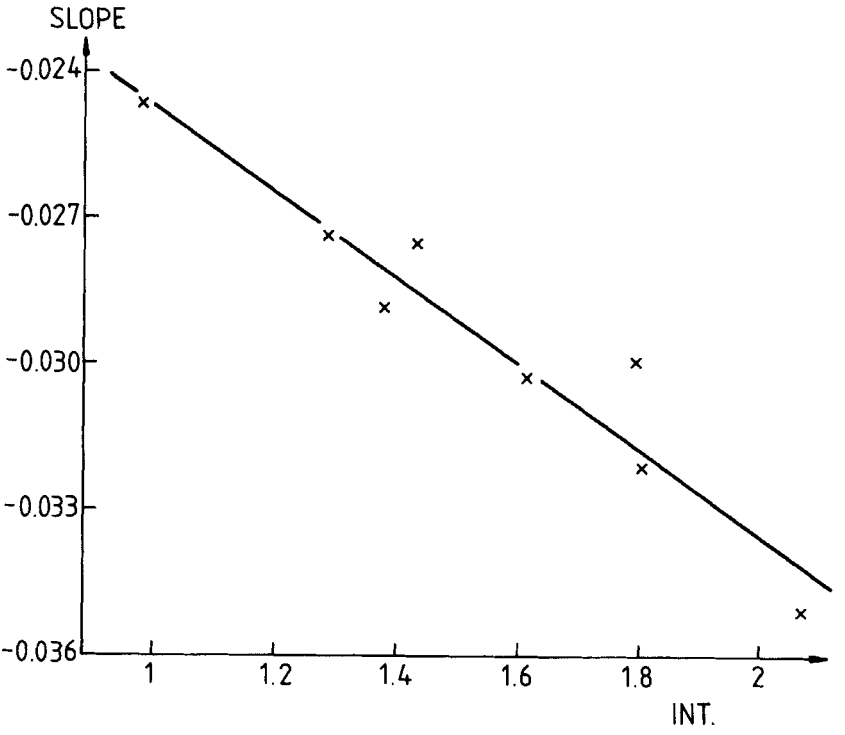


Fig. 7.

The plot of the slope and the intercept values for the slightly acidic subset of compounds listed in Table 3.

The number of compounds: 8

Multiple correlation coefficient (R): 0.956

Fisher-test value: 64.347

Significance of the regression (F-ratio): 0.0326%

Standard error of the estimate: 0.001

Regression parameters: -0.00885, -0.01583

for the subset of slightly acidic drug molecules (serial numbers are in Table 3: 5, 8, 9, 11, 12, 13, 14, 15), and in Fig. 8. for the subset of aniline derivatives. In case of phenols (Fig. 9) the correlation coefficient between the slope and the intercept values is much lower, which can be explained by the big difference in the pK values caused by the substitution. The pH of the mobile phase for the retention data measurements of phenols was 2 which was probably not enough low in some cases for avoiding the dissociation of the highly acidic derivatives.

All of the above mentioned results suggest that always good correlation can be foud between the slope and the intercept values when the data of structurally related compounds are considered. Fig. 10 shows the pictures of the log k' vs OP% straight lines for the structurally related and unrelated compounds. When the correlation coefficient between the slope and the intercept values is low, i. e. the straight lines cross each other, different retention order can be obtained by different mobile phase composition. That can be the explanation of the outlier compounds from the log k' vs log P (the logarithm of octanol/water partition coefficient) relationships (7) and for the different parameters for the log k' vs log P correlations for different classes of compounds (6-9).

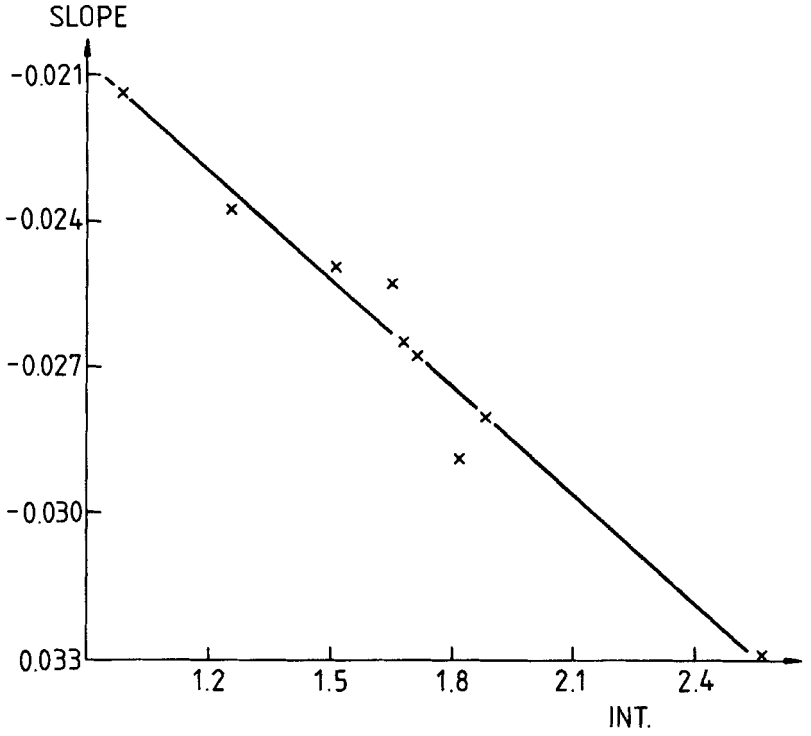


Fig. 8.

The plot of the slope and the intercept values for the aniline derivatives listed in Table 3.

The number of compounds: 9

Multiple correlation coefficient (R): 0.980

Fisher-test value: 172.24

Significance of the regression (F-ratio): 0.0013%

Standard error of the estimate: 0.0007

Regression parameters: -0.00736, -0.01414

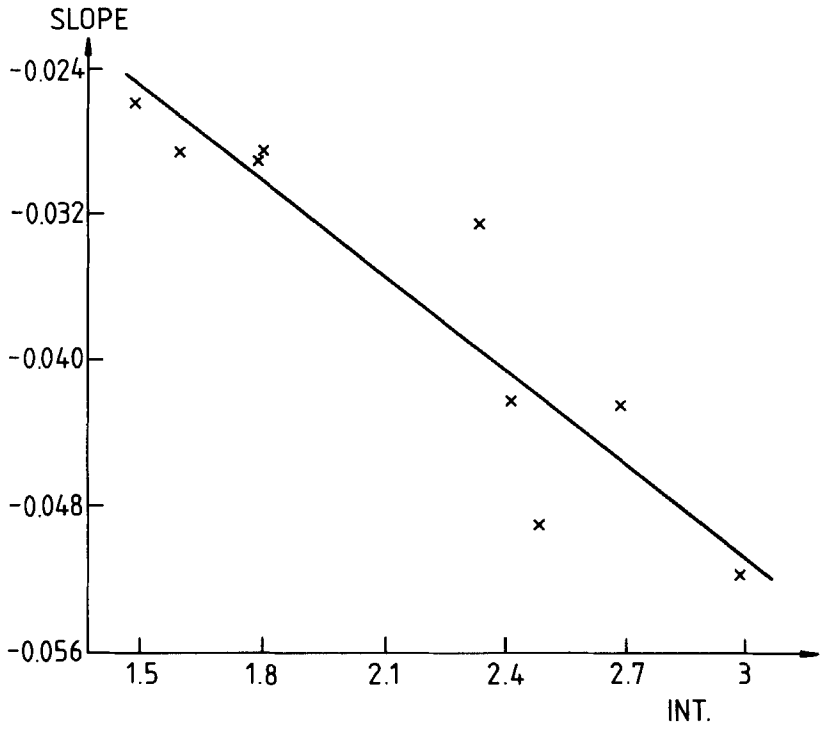


Fig. 9.

The plot of the slope and the intercept values for the phenol derivatives listed in Table 3.

The number of compounds: 9

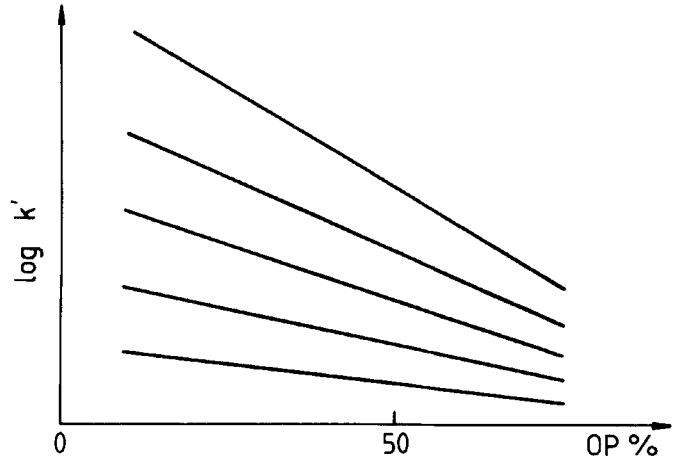
Multiple correlation coefficient (R): 0.921

Fisher-test value: 39.24

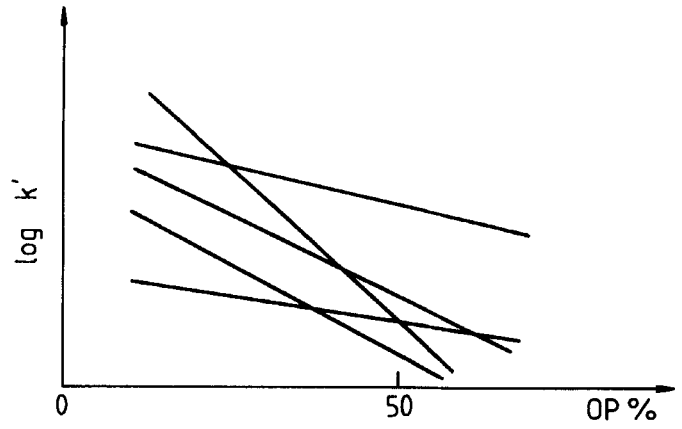
Significance of the regression (F-ratio): 0.053%

Standard error of the estimate: 0.004

Regression parameters: -0.01723, -0.00086



structurally related compounds



structurally unrelated compounds

Fig. 10

The pictures of the log k' vs OP% straight lines for the structurally related and unrelated compounds.

According to the solvophobic theory (19) the slope values are proportional to the hydrophobic contact surface area ofthe compounds, while the intercept values can be regarded as a log k' value referring to the neat water as the mobile phase, and in such a manner they are proportional to the distribution coefficients of the compounds between the mobile and the stationary phase. When the hydrophobic surface area of the compounds are distributed by polar groups in a different way in case of structurally unrelated compounds, the correlation between the hydrophobic surface area and the distribution coefficient, i.e. the slope and the intercept values can be decreased.

In conclusion the structural similarity of a compound set can be revealed by investigating the retention data of the compounds using various mobile phase composition. The correlation coefficient between the slope and the intercept values of the log k' vs OP% straight lines can be suggested as a measure of structural similarity of the compounds regarding their partition behaviour. Such a measure of partition similarity of the compounds is an important parameter in the investigation of quantitative structure - activity relationships for drug design.

REFERENCES

1. Hansch, C., and Fujita, T., ρ-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure, J. Amer. Chem. Soc., 86, 1616, 1964.

2. Snyder, L. R. and Kirkland, J. J., Introduction to Modern Liquid Chromatography, Wiley-Interscience, New York, 1979.

3. Kaliszan, R., Chromatography in Studies of Quantitative Structure - Activity Relationships, J. Chromatogr., 220, 71, 1981.

4. Carney, C. F., Estimation of Molecular Parameters by HPLC, J. Liquid Chromatogr., 8 (15), 2781, 1985.

5. Unger, S. H., Cook, J. R. and Hollenberg, J. S., Simple Procedure for Determining Octanol-Aqueous Partition by Reversed Phase High-Pressure Liquid Chromatography, J. Pharm. Sci., 67 (1), 1364, 1978.

6. Miyake, K. and Terada H., Determination of Partition Coefficients of Very Hydrophobic Compounds by High-Performance Liquid Chromatography on Glyceryl-Coated Controlled-Pore Glass, J. Chromatogr., 240, 9, 1982.

7. Eadsforth, C. V. and Moser, P., Assessment of Reverse-Phase Chromatographic Methods for Determining Partition Coefficients, Chemosphere, 12 (11/12), 1459, 1983.

8. Thus, J. L. G. and Kraak, J. C., Comparison of Phenyl- and Octadecyl-Modified Silica Gel as Stationary Phase for the Prediction of n-Octanol-Water Partition Coefficients by High-Performance Liquid Chromatography, J. Chromatogr., 320, 271, 1985.

9. El Tayar, N., Van der Waterbeend, H. and Testa, B., Lipophylicity Measurements of Protonated Basic Compounds by Reversed-Phase High-Performance Liquid Chromatography, J. Chromatogr., 320, 305, 1985.

10. McDuffie, B., Estimation of Octanol/Water Partition Coefficients for Organic Pollutants Using Reverse-Phase HPLC., Chemosphere, 10, 73, 1981.

11. Hanai, T. and Hubert, J., Retention versus Van Der Waals Volume and Energy in Liquid Chromatography, J. Chromatogr., 290, 197, 1984.

12. Harnisch, M., Möckel, H. J. and Schulze, G., Relationship between log P(ow) Shake-Flask Values and Capacity Factors Derived from Reversed-Phase High-Performance Liquid Chromatography for n-Alkylbenzenes and Some OECD Reference Substances, J. Chromatogr., 282, 315, 1983.

13. Snyder, L. R., Dolan, J. W. and Rigney, M. P., Fast Method Development for Reversed-Phase HPLC. The Use of Computer-Simulation, Submitted to LC/GC Magazine, 1986.

14. Hearn, M. T. W. and Grego, B., Solvent Composition - Capacity Factor Dependencies of Iodiamino Acids. J. Liquid Chromatogr., 7, 1079, 1984.

15. Geng, X. and Regnier, F., Stoichiometric Displacement of Solvent by Non-Polar Solutes in Reversed-Phase Liquid Chromatography, J. Chromatogr., 296, 15, 1984.

16. Junde, W., Jingzhi, L., Rongyu, Y. and Peizhang, L., Influence of the Eluent Composition on Retention in Reversed-Phase High-Performance Liquid Chromatography (HPLC), Scientia Sinica, 28 (7), 680, 1985.

17. Hurtubise, R. J., Cooper, H. A., Scolla, P. L. and Ruckmick, S. C., Prediction of Retention in Liquid Chromatography with Slope-Intercept Relationships, presented in Tenth International Symposium on Column Liquid Chromatography, San Francisco, May 18-23, 1986.

18. Sats, V. D., Sakhartova, O. V., Brivkalve, L. A. and Belikov, V. A., Selection of Eluation Conditions in Reverse-Phase Chromatography. Connectivity Index and Retention of Hydrocarbons and Most Simple Oxygen-Bearing Substances, J. Analiticheskoj Himii, 39 (5), 894, 1984. (in Russ.)

19. Horváth, Cs., Melander, W. and Molnár, I., Solvophobic Interactions in Liquid Chromatography with Nonpolar Stationary Phases, J. Chromatogr., 125, 129, 1976.